# THÈSE

## En vue de l'obtention du
## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

**Délivré par l'Université Toulouse 3 - Paul Sabatier**

**Présentée et soutenue par**

**Mégane BOUDINEAU**

Le 1 février 2019

# Vers la résolution « optimale » de problèmes inverses non linéaires parcimonieux grâce à l'exploitation de variables binaires sur dictionnaires continus : applications en astrophysique.

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse**

Spécialité : **Informatique et Télécommunications**

Unité de recherche :
**IRAP - Institut de Recherche en Astrophysique et Planetologie**

Thèse dirigée par
**Hervé CARFANTAN**

Jury

**Mme Laure BLANC-FÉRAUD,** Rapporteur
**M. Jean-François GIOVANNELLI,** Rapporteur
**M. David MARY,** Examinateur
**M. Michaël BAZOT,** Examinateur
**Mme Sonia CAFIERI,** Examinateur
**M. Hervé CARFANTAN,** Directeur de thèse

# Remerciements

Il est l'heure de terminer ce projet par le chapitre le moins scientifique et avouons-le, c'est avec le cœur lourd que j'écris ces « quelques » lignes symbolisant la fin de cette aventure, cinq ans presque jour pour jour après avoir mis un premier pied à l'IRAP. À chaque personne que j'ai pu rencontrer pendant cette thèse : voici un joyeux merci d'avoir croisé votre chemin, et je m'excuse par avance si j'oublie de vous citer dans les gratitudes plus particulières qui suivent... C'est parti !

Deux personnes ouvrent nécessairement le bal, tout simplement parce que sans elles, j'aurais pu prendre une direction différente. Un infini merci à Sofiane Karoui et Fatima Zohra.

Je souhaite grandement remercier les membres de mon jury de thèse, Laure Blanc-Féraud, Jean-François Giovannelli, David Mary, Sonia Cafieri et Michaël Bazot, pour avoir accepté d'examiner mon manuscrit, un cadeau de Noël jugé assez volumineux ! Merci notamment pour l'intérêt que vous avez porté à mes travaux ; les deux heures de questions sont pour moi passées bien vite. Vos encouragements sur mes recherches et sur ma carrière me seront précieux.

Évidemment, viens ensuite mon illustre directeur de thèse, Hervé Carfantan. Merci pour tout ce que tu m'as appris et apporté durant ces cinq années de collaboration ; pour ta patience lors de la rédaction (il en fallait beaucoup !), ta compréhension des nombreux « Boudinisme », pour tes conseils et ta disponibilité. C'est un réel plaisir de travailler avec toi et j'espère vraiment que l'on pourra continuer, même si nos bureaux ne sont plus adjacents. En définitive, je suis fière d'avoir été assez folle pour répondre à l'offre effrayante de stage sur le Bernoulli-Gaussien !

À mon encadrant officieux, Michaël Bazot, je dois aussi bien des remerciements. En effet, il paraît que sans toi, le projet n'aurait jamais vu le jour ! Merci pour ton implication dans ce projet, pour ton soutien lors de la rédaction, et pour avoir su garder ton calme quand la pression me faisait perdre le mien. Et puis pour le chocolat, les pizzas, les framboises, et le retour de Puff !

Comme cela apparaît tout au long de ce manuscrit, cette thèse est dans la continuité directe de celle de Sébastien Bourguignon, et même de ces travaux plus récents. Alors merci pour avoir si brillamment ouvert la voix et d'avoir su la faire prospérer ensuite. Merci de m'avoir invitée aux réunions de l'ANR MIMOSA ; et donc en élargissant le spectre, je tiens aussi à remercier ses membres, dont Marcel Mongeau, Jordan Ninin, Jérôme Idier et Ramzi Ben Mhenni.

Il en est des collaborations imprévues qui peuvent s'avérer ô combien enrichissantes ; je pense ici à Baptiste Klein, sans qui la partie astrophysique de cette thèse n'aurait pas la même saveur. Merci pour m'avoir aidée à comprendre la physique derrière les données. Et surtout, pour ton intarissable enthousiasme ; tu es un véritable exemple de soleil actif !

Durant cette thèse, j'ai aussi eu la chance d'enseigner à l'IUT GE2I. J'ai particulièrement apprécié faire partie de l'équipe des cours d'algorithmique, et leur motivation pédagogique ; merci notamment à Jérémie, Philippe, Claude, Samir, Stéphanie et Isabelle. Merci aussi à Gaëlle, Gilles et Bruno pour leur accueil et leur encadrement.

Enfin, cette thèse fut aussi riche en collaboration grâce à l'encadrement de stages ou de projets étudiant. Je tiens donc à remercier Alexis, Julien, Kévin, Charlotte, Marine et Antoine.

Un doctorat, c'est aussi une étape de vie au sein d'une équipe de recherche et plus largement d'un laboratoire. Faire partie de l'équipe SISU pendant toutes ces années a été particulièrement enrichissant - et donc je tiens à remercier l'ensemble de ses membres ; Yannick, Éric, Laurent, Hervé, Shahram, Jean-François et Sylvie, mais aussi Ibrahim, Tony, Rémi, Ahmed, Wilhem, Andréa, Mehdi, Axel, Victor, Simon, Sakr, Yohann et Aurélie. D'ailleurs, merci de m'avoir laissé utiliser Shannon de façon si outrageuse ; prenez-en soin maintenant ! Les rituelles pauses café m'ont aussi permis de rencontrer des individus en dehors de l'équipe, et j'ai suivi avec un grand intérêt les conversations scientifiques et les débats animés autour de ce doux breuvage. Donc merci à tous pour ce rendez-vous régulier, dont Marc, Fred, François et Thierry. Ilane, Hayley, Mathieu, Babak, Louise et Geoffroy, merci aussi à vous. Merci à Audrey pour son soutien régulier et ces déjeuners partagés. Et aussi, que serait le labo sans les regards bienveillants d'Ella et Guess. Merci à elles et merci à Frédéric de nous permettre de les côtoyer un peu.

Il est maintenant temps de parler du fameux Bureau B141, aux trois ingénieux surnoms. Le Bon et le Truand ont accueilli sans ciller la Brute pour former un trio dantesque ; merci Ahmed et Wilhem pour ce joyeux accueil dans le monde du doctorat ! Au départ du Bon, l'horizon s'est assombri avec l'arrivée de Darth Andréa ; mais avouons-le, jamais « Neurasthéniques » n'auront été si heureux ensemble ! Merci pour tous ces fous rires, les 6h16, les mots fléchés, la Taupe, les peluches par dizaine... bref pour tous ces moments de franche rigolade. Surtout, merci d'avoir été aussi présents et à l'écoute dans les autres moments, ceux qui donnaient à notre bureau son surnom et qui ont aussi forgé notre amitié. L'atmosphère s'est un peu détendue quand Mehdi nous a rejoint, because its suit and its self-confidence ont fait de lui a true Barney. Merci de reprendre mes travaux avec autant de sérieux et de classe ; pour le coup, le BG sonne bien ! Merci à vous deux de m'accueillir encore un peu au MacLaren's. N'oublions pas un visiteur régulier du B141 ; Axel, un grand merci pour les entraînements de course à pied, et ce en te pliant à mes horaires tordus - même à 7h du matin par -5 ° C ! Merci aussi pour le tiramisu à la fraise, Emelie, deux œufs dans une poêle, Totoro géant, et pour être un bel exemple de « vivre ses rêves ».

Petite liste non exhaustive de gratitudes qui ont rendu la vie plus légère durant ces trois années, merci : les chocopépites, les 120 biscuits gaufrés enrobés de chocolat, Rancho (les deux), Bulbi&Imer, la vieille bécane, le héron, les levers de soleil sur la terrasse, Bertie (rip), Mike (les deux), Yakulu, les tortues ninja, les baleines, KG records & Imaginary families, l'herbe gelée qui craque sous le pied, Opiniâtre Robinson et Salamandre, Sirius, et trois petits points !

Aussi riche la vie fut-elle au labo, elle le fut aussi en dehors. Merci notamment aux Bougainvillains, Déborah, Justine, Oyer, Manu et Thibo, un premier accueil fantastique à Toulouse et un soutien indomptable durant toutes ces années, entre dîners de fin d'année et soirées GoT. Merci aux amis d'ici et d'ailleurs ; clin d'œil à ceux que j'ai eu le plaisir de voir en ville rose, notamment Diane et Héloise - MDH power ! - et les amis manceaux, Justine, Paul et Clothilde.

Merci aussi aux « Firbaux » ; Manu, Dominique, Laurine, Yannick, Juliette et la petite Emy, pour votre accueil et tous les bons moments en terre volcanique, oléronnaise et même normande ! Normandie en effet si loin à pied mais si proche de cœur. À PaPa et Maman, merci pour votre soutien inébranlable toutes ces années d'études et pour la confiance accordée tout au long du chemin, même si celui-ci m'emmène souvent loin. David, tu as allumé la curiosité scientifique chez moi (si si !) il y a bien longtemps, et tu as toujours su m'inspirer pour tant d'aspects de la vie alors merci pour tout grand frère. Et puis, les remerciements s'étendent à Laëtitia et au petit Elyo qui nous apporte tant de joie. Et pour terminer, un grand merci à mon compagnon de vie, qui a su m'accompagner tout au long de cette thèse, et surtout m'a permis de me réaliser en dehors du travail. Alors merci Manu, pour le froid de l'Islande et le soleil de l'Écosse, le vent d'Oléron, pour me motiver à me dépasser en rando et en trail et pour accepter ma folie au quotidien. Merci de m'avoir fait tant rire - vive le dénivelé du Sancy - et merci pour avoir accepté ce ralentissement de vie ces derniers mois. Une pause ravito, et c'est reparti !

# Résumé

Cette thèse s'intéresse à la résolution de problèmes inverses non linéaires exploitant un a priori de parcimonie ; plus particulièrement, des problèmes où les données se modélisent comme la combinaison linéaire d'un faible nombre de fonctions non linéaires en un paramètre dit de « localisation » (par exemple la fréquence en analyse spectrale ou le décalage temporel en déconvolution impulsionnelle). Ces problèmes se reformulent classiquement en un problème d'approximation parcimonieuse linéaire (APL) en évaluant les fonctions non linéaires sur une grille de discrétisation arbitrairement fine du paramètre de localisation, formant ainsi un « dictionnaire discret ». Cependant, une telle approche se heurte à deux difficultés majeures. D'une part, le dictionnaire provenant d'une telle discrétisation est fortement corrélé et met en échec les méthodes de résolution sous-optimales classiques comme la pénalisation L1 ou les algorithmes gloutons. D'autre part, l'estimation du paramètre de localisation, appartenant nécessairement à la grille de discrétisation, se fait de manière discrète, ce qui entraîne une erreur de modélisation. Dans ce travail nous proposons des solutions pour faire face à ces deux enjeux, d'une part via la prise en compte de la parcimonie de façon exacte en introduisant un ensemble de variables binaires, et d'autre part via la résolution « optimale » de tels problèmes sur « dictionnaire continu » permettant l'estimation continue du paramètre de localisation. Deux axes de recherches ont été suivis, et l'utilisation des algorithmes proposés est illustrée sur des problèmes de type déconvolution impulsionnelle et analyse spectrale de signaux irrégulièrement échantillonnés. Le premier axe de ce travail exploite le principe « d'interpolation de dictionnaire », consistant en une linéarisation du dictionnaire continu pour obtenir un problème d'APL sous contraintes. L'introduction des variables binaires nous permet de reformuler ce problème sous forme de « programmation mixte en nombres entiers » (Mixed Integer Programming – MIP) et ainsi de modéliser de façon exacte la parcimonie sous la forme de la « pseudo-norme L0 ». Différents types d'interpolation de dictionnaires et de relaxation des contraintes sont étudiés afin de résoudre de façon optimale le problème grâce à des algorithmes classiques de type MIP. Le second axe se place dans le cadre probabiliste Bayésien, où les variables binaires nous permettent de modéliser la parcimonie en exploitant un modèle de type Bernoulli-Gaussien. Ce modèle est étendu (modèle BGE) pour la prise en compte de la variable de localisation continue. L'estimation des paramètres est alors effectuée à partir d'échantillons tirés avec des algorithmes de type Monte Carlo par Chaîne de Markov. Plus précisément, nous montrons que la marginalisation des amplitudes permet une accélération de l'algorithme de Gibbs dans le cas supervisé (hyperparamètres du modèle connu). De plus, nous proposons de bénéficier d'une telle marginalisation dans le cas non supervisé via une approche de type « Partially Collapsed Gibbs Sampler. » Enfin, nous avons adapté le modèle BGE et les algorithmes associés à un problème d'actualité en astrophysique : la détection d'exoplanètes par la méthode des vitesses radiales. Son efficacité sera illustrée sur des données simulées ainsi que sur des données réelles.

**Mots clés** : Problèmes inverses, Approximations parcimonieuses, Variables binaires, Dictionnaire continu, Modèle Bernoulli-Gaussien, Estimation

# Abstract

This thesis deals with solutions of nonlinear inverse problems using a sparsity prior ; more specifically when the data can be modelled as a linear combination of a few functions, which depend non-linearly on a "location" parameter, i.e. frequencies for spectral analysis or time-delay for spike train deconvolution. These problems are generally reformulated as linear sparse approximation problems, thanks to an evaluation of the nonlinear functions at location parameters discretised on a thin grid, building a "discrete dictionary". However, such an approach has two major drawbacks. On the one hand, the discrete dictionary is highly correlated ; classical sub-optimal methods such as L1-penalisation or greedy algorithms can then fail. On the other hand, the estimated location parameter, which belongs to the discretisation grid, is necessarily discrete and that leads to model errors. To deal with these issues, we propose in this work an exact sparsity model, thanks to the introduction of binary variables, and an optimal solution of the problem with a "continuous dictionary" allowing a continuous estimation of the location parameter. We focus on two research axes, which we illustrate with problems such as spike train deconvolution and spectral analysis of unevenly sampled data. The first axis focusses on the "dictionary interpolation" principle, which consists in a linearisation of the continuous dictionary in order to get a constrained linear sparse approximation problem. The introduction of binary variables allows us to reformulate this problem as a "Mixed Integer Program" (MIP) and to exactly model the sparsity thanks to the "pseudo-norm L0". We study different kinds of dictionary interpolation and constraints relaxation, in order to solve the problem optimally thanks to MIP classical algorithms. For the second axis, in a Bayesian framework, the binary variables are supposed random with a Bernoulli distribution and we model the sparsity through a Bernoulli-Gaussian prior. This model is extended to take into account continuous location parameters (BGE model). We then estimate the parameters from samples drawn using Markov chain Monte Carlo algorithms. In particular, we show that marginalising the amplitudes allows us to improve the sampling of a Gibbs algorithm in a supervised case (when the model's hyperparameters are known). In an unsupervised case, we propose to take advantage of such a marginalisation through a "Partially Collapsed Gibbs Sampler." Finally, we adapt the BGE model and associated samplers to a topical science case in Astrophysics : the detection of exoplanets from radial velocity measurements. The efficiency of our method will be illustrated with simulated data, as well as actual astrophysical data.

**Keywords** : Inverse problems, Sparse Approximations, Binary variables, Continuous dictionary, Bernoulli-Gaussian model, Estimation

# Table des matières

**Remerciements** i

**Résumé** v

**Abstract** vii

**Table des matières** ix

**Notations et acronymes** xiii

**1 Introduction** 1
- 1.1 Problèmes inverses non linéaires exploitant un *a priori* de parcimonie . . . 2
  - 1.1.1 Problèmes inverses non linéaires considérés . . . 2
  - 1.1.2 Un premier exemple : déconvolution de trains d'impulsions . . . 4
  - 1.1.3 Un second exemple : analyse spectrale de signaux irrégulièrement échantillonnés . . . 5
- 1.2 Approximations parcimonieuses linéaires et différentes approches pour leur résolution . . . 8
  - 1.2.1 Une reformulation linéaire . . . 8
  - 1.2.2 Approches déterministes . . . 9
  - 1.2.3 Approches stochastiques . . . 15
  - 1.2.4 Bilan sur les deux approches en APL . . . 20
- 1.3 Approximations parcimonieuses non linéaires sur dictionnaires infinis et continus 21
  - 1.3.1 Norme atomique et relaxation convexe sur dictionnaire infini . . . 21
  - 1.3.2 Approximation parcimonieuse non linéaire sur dictionnaire continu . . . 23
  - 1.3.3 Bilan sur les approximations parcimonieuses non linéaires . . . 26
- 1.4 Conclusion sur les objectifs et les pistes suivies . . . 26
  - 1.4.1 Pistes choisies pour répondre aux objectifs . . . 26
  - 1.4.2 Plan du manuscrit et contributions . . . 28
- 1.5 Données tests . . . 29
  - 1.5.1 Données tests d'analyse spectrale en échantillonnage irrégulier . . . 29
  - 1.5.2 Données tests de déconvolution impulsionnelle . . . 31

**2 Interpolation de dictionnaire et reformulation MIP** 33
- 2.1 Principe de l'interpolation de dictionnaire et travaux existants . . . 34
  - 2.1.1 Cadre général de l'interpolation de dictionnaire pour l'APNL . . . 35
  - 2.1.2 Résolution du problème d'APLsc avec $\mathcal{D}_x = \mathbb{R}_+$ . . . 38
  - 2.1.3 Différentes interpolations de dictionnaire avec $\mathcal{D}_x = \mathbb{R}_+$ . . . 40
  - 2.1.4 Une extension aux amplitudes complexes . . . 44
  - 2.1.5 Bilan . . . 45

2.2 Étude de l'espace des contraintes . . . 47
2.2.1 Cadre général . . . 47
2.2.2 Cas des amplitudes à valeurs réelles . . . 48
2.2.3 Cas des amplitudes à valeurs complexes . . . 51
2.2.4 Bilan . . . 52
2.3 Vers une résolution optimale des problèmes d'APLsc . . . 52
2.3.1 Une résolution optimale avec la pseudo-norme $\ell_0$ . . . 53
2.3.2 Prise en compte des contraintes pour la résolution avec pénalisation $\ell_1$ . . . 56
2.3.3 Bilan . . . 58
2.4 Contributions pour l'interpolation de dictionnaire . . . 59
2.4.1 Interpolation polaire . . . 59
2.4.2 Approximation par K-SVD . . . 65
2.5 Application sur les données de déconvolution impulsionnelle . . . 72
2.5.1 Réglages des paramètres des méthodes pour les tests . . . 73
2.5.2 Comparaison des méthodes de résolution $\ell_1$ et $\ell_0$ avec interpolation polaire 75
2.5.3 Comparaison des interpolations de dictionnaire avec résolution $\ell_0$ . . . 77
2.5.4 Bilan . . . 79
2.6 Extension de l'interpolation de dictionnaire pour l'application à l'analyse spectrale 80
2.6.1 Modèle *cosinus et sinus* et parcimonie structurée . . . 81
2.6.2 Modèle *cosinus+phase* et paramètres de forme . . . 86
2.6.3 Comparaison des différentes formulations proposées sur le signal test . . . 90
2.7 Conclusion . . . 92

**3 Modèle Bernoulli-Gaussien étendu et échantillonnage stochastique 95**
3.1 Modèle Bernoulli-Gaussien pour la résolution d'APL . . . 97
3.1.1 Expression de la loi *a posteriori* . . . 97
3.1.2 Estimation à partir d'échantillons de la loi *a posteriori* . . . 98
3.1.3 Échantillonnage stochastique de la loi *a posteriori* . . . 101
3.2 Modèle BGE : premiers algorithmes d'échantillonnage de la loi jointe . . . 106
3.2.1 Modèle BG étendu et algorithme Gibbs-hybride . . . 106
3.2.2 Modèle BG à dimension variable et algorithme RJGS . . . 110
3.2.3 Choix du modèle BGE avec échantillonnage Gibbs-hybride . . . 111
3.3 Échantillonnage de la loi *a posteriori* marginalisée dans le cadre supervisé . . . 113
3.3.1 Algorithme d'échantillonnage de la loi marginalisée . . . 115
3.3.2 Astuces calculatoires pour l'échantillonnage de $q_j$ . . . 117
3.3.3 Astuces calculatoires pour l'échantillonnage des $\boldsymbol{\nu}_j$ . . . 123
3.3.4 Bilan sur l'échantillonnage de la loi *a posteriori* marginalisée dans le cadre supervisé . . . 124
3.4 Échantillonnage exploitant la marginalisation dans le cadre non supervisé . . . 126
3.4.1 Échantillonnage de type *Partially Collapsed Gibbs Sampler* pour le modèle BGE . . . 127
3.4.2 Astuces calculatoires pour l'implémentation du PCGS . . . 129
3.4.3 Conseils pour le réglage des lois *a priori* pour les hyperparamètres . . . 131
3.5 Application du modèle BGE et de l'échantillonneur associé aux données tests . . . 133
3.5.1 Modèle BGE pour un problème d'analyse spectrale . . . 134
3.5.2 Étude d'une réalisation typique des échantillonneurs et proposition d'une nouvelle stratégie d'estimation . . . 136
3.5.3 Comparaison des échantillonneurs sur 50 graines . . . 140
3.5.4 Étude du réglage des lois *a priori* des hyperparamètres . . . 141

3.5.5 Application sur les données de déconvolution impulsionnelle . . . . . 142
3.5.6 Bilan sur les tests du modèle BGE et des échantillonneurs associés . . . . 144
3.6 Conclusion . . . . . 144

**4 Application à la détection d'exoplanètes** **147**
4.1 Détection d'exoplanètes par l'analyse des vitesses radiales . . . . . 148
4.1.1 Principe de la détection d'exoplanètes par l'analyse des vitesses radiales : formulation d'un problème inverse . . . . . 150
4.1.2 Méthodes pour l'estimation des paramètres orbitaux . . . . . 155
4.1.3 L'activité stellaire, un véritable enjeu . . . . . 158
4.1.4 Le nombre d'exoplanètes . . . . . 161
4.1.5 Bilan : la parcimonie, une alternative . . . . . 162
4.2 Modélisation BGE et échantillonneur associé pour la détection d'exoplanètes . . 163
4.2.1 Reparamétrisation du modèle . . . . . 164
4.2.2 Reformulation en APNL via le découpage du domaine des localisations . . 165
4.2.3 Modèle BGE et échantillonnage associé . . . . . 166
4.2.4 Détection des exoplanètes via l'estimation des paramètres . . . . . 169
4.2.5 Bilan . . . . . 170
4.3 Tests sur des données simulées sans activité stellaire . . . . . 171
4.3.1 Les données . . . . . 171
4.3.2 Un premier essai : algorithmes gloutons avec descente de gradient . . . . . 172
4.3.3 Exploitation du modèle BGE et de l'échantillonneur associé . . . . . 174
4.3.4 Bilan . . . . . 182
4.4 Tests sur des données avec activité stellaire . . . . . 182
4.4.1 Les données . . . . . 182
4.4.2 Réglages du modèle BGE et de l'échantillonneur associé . . . . . 183
4.4.3 Résultats . . . . . 184
4.5 Premiers tests sur des données réelles . . . . . 189
4.5.1 Système HD40307 . . . . . 189
4.5.2 Système CoRoT-7 . . . . . 191
4.6 Conclusion . . . . . 194

**Conclusion générale** **197**

**Annexes** **201**

**A Annexe du chapitre 2** **203**
A.1 Démonstration du théorème 1 p. 61 sur l'expression des paramètres $r$ et $\theta$ pour l'interpolation polaire . . . . . 203
A.2 Démonstration du théorème 2 p. 68 sur la K-SVD continue . . . . . 204
A.3 Expressions analytiques de $\mathbf{B}_\infty$ pour l'approximation par K-SVD continue . . . . 206
A.3.1 Analyse spectrale - modèle *cosinus et sinus* . . . . . 206
A.3.2 Analyse spectrale - modèle *cosinus+phase* . . . . . 206

**B Annexe du chapitre 3** **207**
B.1 Rappel de quelques théorèmes sur les matrices . . . . . 207
B.2 Calcul de la loi marginale . . . . . 207
B.3 Lois *a posteriori* conditionnelles en $q_j$ . . . . . 210
B.4 Simplification de $\Psi_{q_j}$ grâce aux matrices $\mathbf{T}$ et $\boldsymbol{\gamma}$ . . . . . 210

B.5 Simplification des matrices $\mathbf{T}$ et $\boldsymbol{\gamma}$ avec la factorisation de Cholesky $\mathbf{F}$ . . . . . 212
B.6 Mise à jour du facteur de Cholesky $\mathbf{F}$ . . . . . 212
B.7 Cadre non supervisé : échantillonnage des variances avec une étape de Métropolis-Hastings . . . . . 215

**Table des figures** **217**

**Liste des tableaux** **221**

**Bibliographie** **232**

# Notations et acronymes

Afin d'accompagner le lecteur dans la compréhension du manuscrit, ce chapitre détaille les acronymes et les notations les plus utilisés.

## Acronymes

| | |
|---|---|
| **APL** | Approximation Parcimonieuse Linéaire |
| **APLsc** | Approximation Parcimonieuse Linéaire sous contraintes |
| **APNL** | Approximation Parcimonieuse Non Linéaire |
| **ASD** | Average Spike Distance |
| **BF** | Facteur de Bayes |
| **BG** | Bernoulli Gaussien |
| **BGE** | Bernoulli Gaussien étendu |
| **(C)CBP** | (Complex) Continuous Basis Pursuit |
| **CE** | Correct Estimate |
| **EAP** | Espérance *a posteriori* |
| **ELR** | Exact Location Recovery |
| **MAP** | Maximum *a posteriori* |
| **MCMC** | Monte-Carlo par Chaînes de Markov |
| **MH** | Métropolis-Hastings |
| **MIP** | Mixed Integer Program |
| **MIQCP** | Mixed Integer Quadratically Constrained Program |
| **MIQP** | Mixed Integer Quadratic Program |
| **MP** | Matching Pursuit |
| **MPSRF** | Multivariate Potential Scale Reduction Factor |
| **MSE** | Erreur quadratique moyenne |
| **OLS** | Orthogonal Least Square |
| **OMP** | Orthogonal Matching Pursuit |
| **PCGS** | Partially Collapsed Gibbs Sampler |
| **POR** | Posterior Odds Ratio |
| **QP** | Quadratic Program |
| **RSB** | Rapport Signal sur Bruit |
| **SISU** | Signal Image en Sciences de l'Univers |
| **SMLR** | Single Most Likely Replacement |
| **SOCP** | Second Order Cone Program |
| **SVD** | Singular Value Decomposition |

## Notations mathématiques

| | | |
|---|---|---|
| $\mathbb{F}$ | Espace $\mathbb{R}_+$, $\mathbb{R}$ ou $\mathbb{C}$ | |
| $[\![1; N]\!]$ | Ensemble des entiers entre 1 et $N$ | |
| $\text{Card}(\cdot)$ | Nombre d'élements dans l'ensemble | |
| $\odot$ | Produit de Hadamart | |
| $.^T$ | Transposée | |
| $.^\dagger$ | Transposée conjuguée | |
| $\|\boldsymbol{u}\|^2_{\mathbf{A}}$ | Norme pondérée : $\|\boldsymbol{u}\|^\dagger_{\mathbf{A}} = \boldsymbol{u}^T\mathbf{A}\boldsymbol{u}$ avec $\mathbf{A}$ matrice semi-définie positive | |

## Notations : problèmes inverses non linéaires

| | | |
|---|---|---|
| $\boldsymbol{t}$ | Instants d'échantillonnage | $N \times 1$ |
| $\boldsymbol{y}$ | Signal, données | $N \times 1$ |
| $\boldsymbol{\epsilon}$ | Bruit additif | $N \times 1$ |
| $\mathbf{h}(\cdot)$ | Atome continu | $N \times m$ |
| $\boldsymbol{x}$ | Amplitude d'un atome continu | $m \times 1$ |
| $\boldsymbol{\nu}$ | Paramètres non linéaires | $p$ |
| $\nu$ | Paramètre de localisation | 1 |
| $\boldsymbol{\upsilon}$ | Paramètres de forme | $p-1$ |
| $\mathcal{I}_\nu$ | Domaine de localisation $\mathcal{I}_\nu = [\nu_{\min}; \nu_{\max}]$ | |
| $\mathcal{I}_{\boldsymbol{\upsilon}}$ | Espace de définition des paramètres de forme | |
| $L$ | Ordre du modèle | |

## Notations : APL

| | | |
|---|---|---|
| $\mathcal{G}$ | Grille de discrétisation des paramètres de localisation | $J$ |
| $\ddot{\nu}_j$ | Paramètre de localisation discrétisé | |
| $\mathbf{h}_j$ | Atome discret $\mathbf{h}_j = \mathbf{h}(\ddot{\nu}_j)$ | $N \times m$ |
| $\boldsymbol{x}_j$ | Amplitude de l'atome discret | $m \times 1$ |
| $\mathbf{H}$ | Dictionnaire discret $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_J]$ | $N \times mJ$ |
| $\boldsymbol{x}$ | Amplitudes | $mJ \times 1$ |
| $\Omega$ | Support $\Omega = \mathsf{Supp}\,(\boldsymbol{x}) = \{j \in [\![1; J]\!],\ \boldsymbol{x}_j \neq \mathbf{0}\}$ | |
| $\mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H})$ | Terme d'attache aux données | |
| $\|\boldsymbol{x}\|_0$ | Pseudo-norme $\ell_0$ : $\|\boldsymbol{x}\|_0 = \text{Card}(\mathsf{Supp}\,(\boldsymbol{x}))$ | |
| $\widehat{\boldsymbol{x}}^{\text{MC}}$ | Minimiseur des moindres carrés | |
| $\overline{\mathbf{H}}$ | Dictionnaire réduit (atomes d'amplitudes non nulles) | $N \times mL$ |
| $\overline{\boldsymbol{x}}$ | Amplitudes réduites (non nulles) | $mL \times 1$ |

## Notations : APNL

| | | |
|---|---|---|
| $\mathcal{I}_j$ | Partition du domaine de localisation : $\mathcal{I}_\nu = \cup_{j=1}^J \mathcal{I}_j$ | |
| $\Delta$ | Largeur des intervalles $\mathcal{I}_j$ : $\Delta = (\nu_{\max} - \nu_{\min})/J$ | |
| $\boldsymbol{\nu}$ | Vecteur regroupant l'ensemble des paramètres non linéaires<br>$\boldsymbol{\nu} = [\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_J]$, $\boldsymbol{\nu}_j = (\nu_j, \boldsymbol{\upsilon}_j)$ | $pJ \times 1$ |
| $\mathfrak{I}, \mathfrak{I}_j$ | Domaine de définition de $\boldsymbol{\nu}$<br>$\boldsymbol{\nu} \in \mathfrak{I} \Leftrightarrow \forall j \in [\![1; J]\!],\ \boldsymbol{\nu}_j \in \mathfrak{I}_j\ :\ \nu_j \in \mathcal{I}_j$ et $\boldsymbol{\upsilon}_j \in \mathcal{I}_{\boldsymbol{\upsilon}}$ | |
| $\mathbf{H}(\boldsymbol{\nu})$ | Dictionnaire continu $\mathbf{H}(\boldsymbol{\nu}) = [\mathbf{h}(\boldsymbol{\nu}_1), \ldots, \mathbf{h}(\boldsymbol{\nu}_J)]$ | $N \times mJ$ |

## Notations : interpolation de dictionnaire

| | | |
|---|---|---|
| $\delta_j$ | Décalages : $\nu_j \in \mathcal{I}_j \Leftrightarrow \delta_j = \nu_j - \ddot{\nu}_j \in \mathcal{I}_\Delta$ | |
| $\mathcal{I}_\Delta$ | Intervalle des décalages : $\delta_j \in \mathcal{I}_\Delta = [-\frac{\Delta}{2}; \frac{\Delta}{2}]$ | |
| $\boldsymbol{h}_j(\delta_j)$ | Autre écriture de l'atome continu $\boldsymbol{h}_j(\delta_j) = \boldsymbol{h}(\nu_j)$ | $N \times 1$ |
| $\mathcal{D}_x$ | Domaine de définition des amplitudes scalaires $\mathcal{D}_x = \mathbb{R}_+$, $\mathbb{R}$ ou $\mathbb{C}$ | |
| | | |
| $\widetilde{\mathbf{h}}_j$ | Atome interpolé | $N \times K$ |
| $\widetilde{\boldsymbol{x}}_j$ | Amplitude de l'atome interpolé | $K \times 1$ |
| $\Psi(\cdot)$ | Interpolateur | $K \times 1$ |
| $g$ | Changement de variables de $\mathcal{D}_x \times \mathcal{I}_\Delta$ vers $\mathcal{S}_j$ | |
| $\mathcal{S}_j$ | Espace image du changement de variables | |
| $\mathcal{H}_j$ | Relaxation convexe de $\mathcal{S}_j$ | |
| $\mathcal{S}_j^+, \mathcal{H}_j^+$ | Espace image et relaxation pour $\mathcal{D}_x = \mathbb{R}_+$ | |
| $\mathcal{S}_j^{\mathbb{R}}, \mathcal{H}_j^{\mathbb{R}}$ | Espace image et relaxation pour $\mathcal{D}_x = \mathbb{R}$ | |
| $\mathcal{S}_j^{\mathbb{C}}, \mathcal{H}_j^{\mathbb{C}}$ | Espace image et relaxation pour $\mathcal{D}_x = \mathbb{C}$ | |
| $\mathcal{P}(\cdot)$ | Projection de $\mathcal{H}_j^+$ vers $\mathcal{S}_j^+$ | |
| $\mathcal{C}$ | Espace des contraintes | |
| $\mathcal{I}_\Delta^P$ | Intervalle $\mathcal{I}_\Delta$ discrétisé en $P$ points | $P$ |

$\mathcal{H}_j^+$ respecte l'*implication de nullité* : pour $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+$, $\alpha_j = \widetilde{x}_{j,1} = 0 \Rightarrow \widetilde{\boldsymbol{x}}_j = \mathbf{0}$
*Contrainte polyhédrale* : contrainte linéaire inégalité sur deux composantes.
*Contrainte de disjonction* : $(\widetilde{\boldsymbol{x}}_j^+ = \mathbf{0})$ ou $(\widetilde{\boldsymbol{x}}_j^- = \mathbf{0})$

## Notations : modèle BGE

| | | |
|---|---|---|
| $\mathbb{E}\{\cdot\}$ | Espérance d'une variable aléatoire | |
| $t$ | Indices des échantillons des algorithmes MCMC | $T$ |
| $\widehat{\cdot}$ | Estimateur de l'EAP par moyenne des échantillons MCMC | |
| $\widehat{\boldsymbol{q}}_\alpha$ | Séquence binaire provenant de $\widehat{\boldsymbol{q}}$ seuillée à $\alpha$ | |

## Lois de probabilité usuelles

| | |
|---|---|
| $u \sim \mathcal{B}(\lambda)$ | $\Pr(u = 1) = \lambda$, $\Pr(u = 0) = 1 - \lambda$ |
| $\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ | $p(\boldsymbol{u}) = (2\pi)^{-N/2} \lvert\boldsymbol{\Gamma}\rvert^{-1/2} \exp(-\frac{1}{2}\lVert\boldsymbol{u} - \boldsymbol{\mu}\rVert^2_{\boldsymbol{\Gamma}^{-1}})$ $\quad$ ($\boldsymbol{u}$ (dim. $N \times 1$)) |
| $u \sim \mathcal{U}([a; b[)$ | $p(u) = \dfrac{1}{b-a}\ 1_{[a;b[}(u)$ |
| $u \sim \mathcal{B}e(\alpha, \beta)$ | $p(u) = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u^{\alpha-1}(1-u)^{\beta-1}\ 1_{[0;1]}(u)$ |
| $u \sim \mathcal{IG}(\alpha, \beta)$ | $p(u) = \dfrac{\beta^\alpha}{\Gamma(\alpha)} \left(\dfrac{1}{u}\right)^{1+\alpha} \exp(-\beta/u)$ |

# Chapitre 1

# Introduction

## Sommaire


**1.1 Problèmes inverses non linéaires exploitant un *a priori* de parcimonie . . . . 2**
1.1.1 Problèmes inverses non linéaires considérés . . . . . 2
1.1.2 Un premier exemple : déconvolution de trains d'impulsions . . . . . 4
1.1.3 Un second exemple : analyse spectrale de signaux irrégulièrement échantillonnés . 5
**1.2 Approximations parcimonieuses linéaires et différentes approches pour leur résolution . . . . . 8**
1.2.1 Une reformulation linéaire . . . . . 8
1.2.2 Approches déterministes . . . . . 9
1.2.3 Approches stochastiques . . . . . 15
1.2.4 Bilan sur les deux approches en APL . . . . . 20
**1.3 Approximations parcimonieuses non linéaires sur dictionnaires infinis et continus . . . . . 21**
1.3.1 Norme atomique et relaxation convexe sur dictionnaire infini . . . . . 21
1.3.2 Approximation parcimonieuse non linéaire sur dictionnaire continu . . . . . 23
1.3.3 Bilan sur les approximations parcimonieuses non linéaires . . . . . 26
**1.4 Conclusion sur les objectifs et les pistes suivies . . . . . 26**
1.4.1 Pistes choisies pour répondre aux objectifs . . . . . 26
1.4.2 Plan du manuscrit et contributions . . . . . 28
**1.5 Données tests . . . . . 29**
1.5.1 Données tests d'analyse spectrale en échantillonnage irrégulier . . . . . 29
1.5.2 Données tests de déconvolution impulsionnelle . . . . . 31


Ce manuscrit regroupe les travaux de la thèse intitulée : *Vers la résolution « optimale » de problèmes inverses non linéaires parcimonieux grâce à l'exploitation de variables binaires sur dictionnaires continus : applications en astrophysique.* Cette thèse a été réalisée au sein de l'équipe *Signal Images en Sciences de l'Univers* (SISU) de l'*Institut de Recherche en Astrophysique et Planétologie* (IRAP), à Toulouse.

L'objectif principal de cette thèse est la résolution de problèmes inverses particuliers à l'aide d'une reformulation en problèmes d'approximation parcimonieuse non linéaire (APNL). Des méthodes de résolution dites « optimales » de problèmes d'approximation parcimonieuse linéaire (APL) seront étendues et/ou adaptées pour permettre l'estimation continue des paramètres suivant lesquels le modèle n'est pas linéaire. Par résolution « optimale », nous sous-entendons deux critères à respecter. D'une part, la parcimonie (le fait n'avoir qu'un faible nombre de composantes d'amplitudes non nulles) est prise en compte de façon exacte via l'introduction de

variables binaires, permettant de définir clairement le support de la solution. La détection de ce support est en effet le point crucial de tels problèmes. D'autre part, les méthodes de résolution employées visent à garantir de trouver la solution optimale des problèmes posés dans ce cadre. Des travaux ont proposé de telles solutions « optimales » pour des problèmes d'approximation parcimonieuse linéaire, et notre contribution se place dans ce cadre pour des problèmes non linéaires. Nos travaux se placent dans la lignée directe de ceux menés dans l'équipe SISU sur les problèmes inverses en astrophysique. Le but ultime sera alors leur application à des problèmes d'actualité en astrophysique : l'analyse spectrale de signaux irrégulièrement échantillonnés et plus précisément la détection d'exoplanètes via la méthode des vitesses radiales.

Dans ce premier chapitre, nous introduisons les différentes notions, définitions et notations qui permettront de comprendre le contexte général de cette thèse et de justifier les pistes de recherche choisies. Nous commençons par présenter le type de problèmes inverses auquel nous nous intéressons dans l'ensemble de ce travail : les données se modélisent comme la combinaison linéaire d'un faible nombre de fonctions non linéaires en un paramètre dit de « localisation », et possiblement en des paramètres supplémentaires dits de « forme ». Deux exemples classiques nous permettent d'illustrer ce type de problèmes : l'analyse spectrale de signaux irrégulièrement échantillonnés et la déconvolution impulsionnelle. La reformulation classique en problème d'approximation parcimonieuse linéaire (APL) est ensuite abordée, ainsi que les approches de type déterministe et stochastique pour sa résolution. Nous mettons en évidence deux approches qui, via l'introduction de variables binaires, permettent une résolution dite « optimale » du problème d'APL. Cependant, cette reformulation entraîne une estimation discrète du paramètre de localisation ; nous donnons ensuite quelques pistes déjà explorées pour obtenir une estimation continue du paramètre de localisation mais aussi des paramètres de formes. Nous revenons enfin sur l'objectif de ce travail et justifions les pistes de recherche envisagées. Les trois chapitres suivants, qui regroupent nos travaux, sont introduits et nous récapitulons les contributions que nous avons apportées, ainsi que nos articles scientifiques associés. Des données tests en analyse spectrale en échantillonnage irrégulier ainsi qu'en déconvolution impulsionnelle seront décrites en fin de chapitre ; elles permettrons de tester les deux approches proposées sur des données similaires.

## 1.1 Problèmes inverses non linéaires exploitant un *a priori* de parcimonie

Nous formalisons dans cette section le type particulier de problèmes inverses non linéaires auxquels nous nous intéressons, et nous l'illustrons avec deux exemples concrets en sciences de l'Univers.

### 1.1.1 Problèmes inverses non linéaires considérés

Commençons d'abord par définir un problème inverse pour des séries temporelles dans un cadre général. On considère une grandeur physique dépendant du temps $t$. Le problème *direct* consiste à modéliser cette grandeur physique en fonction du temps et de paramètres $\boldsymbol{\Theta}$ pour donner un modèle $f(t;\boldsymbol{\Theta})$. Cette grandeur physique peut être observée ; on obtient ainsi le signal $\boldsymbol{y} = [y_1, \ldots, y_N]^T$, provenant de la mesure de cette grandeur physique aux instants d'échantillonnage $\boldsymbol{t} = [t_1, \ldots, t_N]^T$, tel que $y_n = y(t_n)$. Les observations $\boldsymbol{y}$ peuvent alors se modéliser comme la somme du modèle $f(\boldsymbol{t};\boldsymbol{\Theta})$ et d'un terme de bruit additif regroupant de possibles erreurs de modèle, de mesures, ou encore des signaux parasites. Le problème *inverse* consiste alors à retrouver les paramètres physiques $\boldsymbol{\Theta}$ à partir de l'observation $\boldsymbol{y}$.

Dans ce manuscrit, nous nous focalisons sur un type de problèmes inverses non linéaires avec *a priori* de parcimonie, où la grandeur physique se modélise comme la combinaison linéaire d'un faible nombre de fonctions non linéaires en certains paramètres. Plus précisément, les signaux $\boldsymbol{y}$ considérés sont des séries temporelles, à partir desquelles il s'agit de résoudre le problème suivant :

$$\text{Estimer } L \text{ faible et les } \{\boldsymbol{\nu}_\ell, \boldsymbol{x}_\ell\}_\ell, \text{ tels que } \boldsymbol{y} = \sum_{\ell=1}^{L} \mathbf{h}(\boldsymbol{\nu}_\ell)\boldsymbol{x}_\ell + \boldsymbol{\epsilon} \tag{1.1}$$

avec, en notant $\mathbb{F} = \mathbb{R}$ ou $\mathbb{C}$ (la nature des espaces, complexes ou réels, dépendra du problème et du modèle choisi) :

- $\boldsymbol{y} \in \mathbb{F}^N$ : le *signal* / les *données*, un vecteur colonne de dimension $N$, provenant de l'échantillonnage du signal continu $y(t)$ aux $N$ instants $t_n$ ;
- $\mathbf{h}(\cdot) : \mathbb{F}^p \to \mathbb{F}^{N\times m}$ : l'*atome continu*[1], une fonction dépendant non linéairement des paramètres $\boldsymbol{\nu}_\ell$. La matrice $\mathbf{h}(\boldsymbol{\nu}_\ell)$ est constituée de $m \geq 1$ colonnes et $N$ lignes, puisqu'elle vient de l'échantillonnage temporel d'un modèle continu $h(t; \boldsymbol{\nu}_\ell)$ : $\mathbf{h}(\boldsymbol{\nu}_\ell) = h(\boldsymbol{t}; \boldsymbol{\nu}_\ell)$, où $\boldsymbol{t} = [t_1, \ldots, t_N]^T$ sont les $N$ instants d'échantillonnage ;
- $\boldsymbol{\nu}_\ell \in \Lambda \subset \mathbb{R}^p$ : par abus de langage, ces paramètres seront nommés *paramètres non linéaires*, pour refléter la dépendance non linéaire de l'atome continu $\mathbf{h}(\cdot)$ en ces paramètres. Chacune des $p$ composantes de $\boldsymbol{\nu}_\ell$ a un sens physique différent, qui sera le même pour tout $\ell \in [\![1; L]\!]$. En écrivant $\boldsymbol{\nu}_\ell = [\nu_\ell, \boldsymbol{\upsilon}_\ell]$, on distingue deux natures différentes de paramètres non linéaires :
  * le *paramètre de localisation* : $\nu_\ell \in \mathcal{I}_\nu = [\nu_{\min}; \nu_{\max}] \subset \mathbb{R}$. On appellera *domaine de localisation* l'intervalle $\mathcal{I}_\nu = [\nu_{\min}; \nu_{\max}]$.
  * les *paramètres de forme* : $\boldsymbol{\upsilon}_\ell \in \mathcal{I}_{\boldsymbol{\upsilon}} \subset \mathbb{R}^{p-1}$, qui seront parfois absents de la modélisation ($p = 1 \Leftrightarrow \boldsymbol{\upsilon}_\ell = \varnothing$).

  Ces deux types de paramètres seront illustrés davantage par les exemples présentés dans cette section, et prendront tout leur sens dans le cadre de l'*approximation parcimonieuse* en § 1.2.1 et § 1.3.2 ;
- $\boldsymbol{x}_\ell \in \mathbb{F}^m$ : les *paramètres linéaires* ou *amplitudes*, un vecteur colonne de dimension $m$ qui regroupe les amplitudes des $m$ colonnes de $\mathbf{h}(\boldsymbol{\nu}_\ell)$ ;
- $L$ : *l'ordre du modèle*, le nombre inconnu de composantes $\mathbf{h}(\boldsymbol{\nu}_\ell)\boldsymbol{x}_\ell$ pour approcher $\boldsymbol{y}$ ;
- $\boldsymbol{\epsilon}$ : un *bruit additif*, contenant les erreurs de mesures ou de modèle, ou encore des signaux parasites.

Un tel problème est particulièrement difficile à résoudre : l'ordre du modèle $L$ est inconnu, et la non-linéarité de l'atome $\mathbf{h}(\cdot)$ en $\boldsymbol{\nu}_\ell$ complique systématiquement des tentatives d'estimation par optimisation ou par échantillonnage stochastique. Par ailleurs, la présence du bruit complique d'autant plus le problème, puisque le signal $\boldsymbol{y}$ ne peut pas s'exprimer exactement comme les combinaisons linéaires des atomes continus $\mathbf{h}(\boldsymbol{\nu}_\ell)$. Leur résolution nécessite alors la prise en compte de connaissances *a priori* sur le problème : l'ordre du modèle $L$ est faible. Il s'agit d'un *a priori* de *parcimonie*. En sciences de l'Univers, de nombreux problèmes peuvent s'exprimer sous la forme d'un tel problème inverse non linéaire ; les plus courants étant la déconvolution impulsionnelle et la recherche de périodicités dans des signaux irrégulièrement échantillonnés. Dans le cadre de cette thèse, nous nous sommes intéressés à ces deux problèmes, que nous décrivons plus en détails dans les paragraphes suivants. Notons qu'ils peuvent tous deux être présentés comme des problèmes de déconvolution, le premier dans le domaine temporel et le second dans le domaine fréquentiel, mais ils diffèrent par les spécificités liées à l'application scientifique. Nous verrons que dans ces deux problèmes, l'intérêt principal est de retrouver le nombre $L$ d'atomes $\mathbf{h}(\cdot)$ (les échos pour la déconvolution impulsionnelle ou les sinusoïdes pour l'analyse spectrale)

[1] La dénomination *atome continu* est en lien avec la reformulation en problème d'*approximation parcimonieuse non linéaire* qui sera présentée en § 1.3.2.

ainsi que leur localisation caractérisée par le paramètre de localisation $\nu_\ell$ (le décalage ou la fréquence). On parlera donc de problèmes de *détection* d'évenements caractérisés par leur nombre et leur localisation. Une autre approche possible est la modélisation de type Processus ponctuels marqués (*Marked point process*) utilisée pour la détection d'objets notamment en traitement d'images (voir par exemple [DZ02] ou plus récemment [MCM+16]) associée à des algorithmes stochastiques de type MCMC à saut réversibles. Nous n'aborderons pas davantage une telle approche dans ce manuscrit, mais nous verrons au Chapitre 4 qu'elle est parfois employée pour la détection d'exoplanètes.

### 1.1.2 Un premier exemple : déconvolution de trains d'impulsions

Le problème de déconvolution de trains d'impulsions est présent dans de nombreux domaines, tel que l'exploration sismique pour l'étude de la structure géophysique de la Terre (voir par exemple [CCL96]) ou le contrôle non destructif de matériaux (voir par exemple [CBIS13]). Dans les deux cas, une onde (sismique ou ultrasonore) est envoyée vers le milieu à étudier, et l'onde réfléchie donne un signal $y(t)$ (en temps continu). Ce signal peut alors se modéliser comme la convolution entre une onde (connue ou non) $g(t)$ et un train d'impulsion $x(t)$ plus un terme de bruit additif $\epsilon(t)$. Le modèle convolutif s'écrit donc de la façon suivante :

$$y(t) = (x * g)(t) + \epsilon(t) = \int_{-\infty}^{+\infty} g(t-u)x(u)\,\mathrm{d}u + \epsilon(t) \tag{1.2}$$

L'objectif est alors d'estimer le train d'impulsion $x(t)$, permettant de repérer (par exemple), les défauts dans le matériau testé ou encore les différentes couches de la surface terrestre. En écrivant le train d'impulsion $x(t)$ comme une somme de $L$ impulsions de Dirac aux instants $\nu_\ell$ et de poids $x_\ell$, soit $x(t) = \sum_{\ell=1}^{L} x_\ell\,\delta(t-\nu_\ell)$, on obtient le modèle suivant :

$$y(t) = \sum_{\ell=1}^{L} x_\ell\, g(t-\nu_\ell) + \epsilon(t)$$

Après échantillonnage des signaux avec en général un pas régulier (soit $y_n = y(nT_e)$ où $T_e$ est la période d'échantillonnage), ce problème de déconvolution peut alors s'écrire sous la forme du problème inverse non linéaire que nous avons défini précédemment :

$$\text{Estimer } L \text{ et les } \{\nu_\ell, x_\ell\}_\ell, \text{ tels que } \boldsymbol{y} = \sum_{\ell=1}^{L} x_\ell\, \boldsymbol{h}(\nu_\ell) + \boldsymbol{\epsilon}$$
$$\text{avec } \boldsymbol{h}(\nu_\ell) = h(\boldsymbol{t};\nu_\ell) = g(\boldsymbol{t}-\nu_\ell)$$

où le paramètre non linéaire est seulement composé du paramètre de localisation $\nu_\ell$ (soit $p=1$), correspondant aux décalages temporels indiquant la localisation temporelle de l'écho, et les amplitudes $x_\ell$ ($m=1$).

Résoudre ce problème peut être particulièrement difficile, notamment quand des échos voisins $\{\boldsymbol{h}(\nu_\ell)\}_\ell$ se recouvrent et les rendent difficilement discernables. Notons aussi que la forme $g(t)$ des ondes peut être inconnue ; on parlera alors de déconvolution *aveugle*. Dans un cadre *semi-aveugle*, la forme de l'onde $g(t)$ peut être connue mais paramétrée par un ou plusieurs *paramètres de forme* inconnus ; par exemple, la forme de l'onde que nous utiliserons dans nos tests[2] (voir § 1.5.2) peut être paramétrée par trois paramètres de formes $\boldsymbol{\upsilon} = (\alpha, f_c, \phi)$ :

$$g(t;\boldsymbol{\upsilon}) = e^{-\alpha t^2}\cos(2\pi f_c t + \phi)$$

[2]Pour nos tests sur la déconvolution impulsionnelle, ces paramètres de forme seront supposés connus.

Le problème inverse non linéaire peut alors se réécrire de la façon suivante :

$$\text{Estimer } L \text{ et les } \{\boldsymbol{\nu}_\ell, x_\ell\}_\ell, \text{ tels que } \boldsymbol{y} = \sum_{\ell=1}^{L} x_\ell \, \boldsymbol{h}(\boldsymbol{\nu}_\ell) + \boldsymbol{\epsilon}$$
$$\text{avec } \boldsymbol{h}(\boldsymbol{\nu}_\ell) = h(\boldsymbol{t}; \boldsymbol{\nu}_\ell) = g(\boldsymbol{t} - \nu_\ell; \boldsymbol{\upsilon}_\ell)$$

où les paramètres non linéaires $\boldsymbol{\nu}_\ell$ regroupent le *paramètre de localisation* $\nu_\ell$, indiquant la position de l'écho, et les *paramètres de forme* $\boldsymbol{\upsilon}_\ell = (\alpha, f_c, \phi)_\ell$ paramétrant l'allure de cet écho.

De nombreuses techniques ont été développées afin de traiter directement ce problème de façon *paramétrique*. Une autre approche plus classique consiste en la discrétisation de l'intégrale (1.2) en une convolution discrète et la recherche sur un ensemble fini d'échos en imposant un *a priori* de parcimonie. Cette approche sera l'objet de la section suivante dans un cadre plus général.

### 1.1.3 Un second exemple : analyse spectrale de signaux irrégulièrement échantillonnés

En astrophysique, la recherche de périodicités dans des signaux est d'une importance capitale dans de nombreuses applications. Par exemple, l'observation d'une étoile pour la recherche de pulsations périodiques peut permettre d'analyser sa structure interne et son évolution (il s'agit de l'astérosismologie, voir par exemple [BCBC04]), mais aussi de mettre en exergue la présence d'exoplanètes en rotation autour de l'étoile exerçant une influence gravitationnelle (par spectroscopie, e.g. [MQ95]).

Ces problèmes reviennent souvent à la recherche d'un nombre inconnu $L$ (et souvent faible) de variations sinusoïdales de fréquence $\nu_\ell$ (souvent exprimée en jours$^{-1}$ en astrophysique) dans une série temporelle à valeurs réelles $\boldsymbol{y} \in \mathbb{R}^N$ et peuvent être écrits sous la forme d'un problème inverse non linéaire :

$$\text{Estimer } L \text{ et } \{\nu_\ell, \boldsymbol{x}_\ell\}_\ell, \text{ tels que } \boldsymbol{y} = x_0 + \sum_{\ell=1}^{L} \mathbf{h}(\nu_\ell)\boldsymbol{x}_\ell + \boldsymbol{\epsilon} \tag{1.3}$$

où l'atome continu est composé de deux colonnes : $\mathbf{h}(\nu_\ell) = [\cos(2\pi\nu_\ell \boldsymbol{t}) \ \sin(2\pi\nu_\ell \boldsymbol{t})]$ et les amplitudes $\boldsymbol{x}_\ell$ de deux composantes associées $\boldsymbol{x}_\ell = [a_\ell, b_\ell]^T$ (soit $m = 2$). Les fréquences $\nu_\ell$ jouent le rôle de *paramètre de localisation*, puisqu'elles permettent de distinguer différents modes de pulsations (ou encore différentes exoplanètes - voir Chapitre 4), et sont donc l'unique composante des paramètres non linéaires ($p = 1$). En fait, deux autres modèles, équivalents à ce dernier que nous nommons modèle en *cosinus et sinus*, peuvent être employés (on omet les indices $\ell$) :

- le modèle *exponentielles complexes* : $\mathbf{h}(\boldsymbol{\nu})\boldsymbol{x} = \mathbf{h}(\nu)\boldsymbol{x} = (Ae^{i2\pi\nu\boldsymbol{t}} + Be^{-i2\pi\nu\boldsymbol{t}})$ où $\boldsymbol{x} = [A, B]^T \in \mathbb{C}^2$ et $B = A^*$ est le conjugué de $A$.
- le modèle *cosinus+phase* : $\mathbf{h}(\boldsymbol{\nu})\boldsymbol{x} = x \, \boldsymbol{h}([\nu, \phi]) = x \, \cos(2\pi\nu\boldsymbol{t} + \phi)$, avec $x \in \mathbb{R}_+$ et $|\phi| \leq \pi$, ou $x \in \mathbb{R}$ et $|\phi| \leq \frac{\pi}{2}$.

Ces modèles se différencient par la dimension des modèles/amplitudes associées, à valeurs réelles ou complexes, et par la dimension des paramètres non linéaires ; ces caractéristiques sont rappelées en Tableau 1.1. Notamment, les paramètres non linéaires du modèle *cosinus+phase*, contrairement aux deux autres modèles, comprend aussi un *paramètre de forme* via la phase $\phi_\ell$, en plus du *paramètre de localisation*, la fréquence $\nu_\ell$. Ces différents modèles auront des avantages ou des inconvénients suivant les méthodes d'estimation envisagées, comme nous le verrons notamment au Chapitre 2.

| Modèles | Paramètres non linéaires $\boldsymbol{\nu}$ | Atomes continus $\mathbf{h}(\boldsymbol{\nu})$ | Amplitudes $\boldsymbol{x}$ |
|---|---|---|---|
| *cosinus et sinus* | $\nu \in \mathbb{R}$ | $[\cos(2\pi\nu\boldsymbol{t})\ \sin(2\pi\nu\boldsymbol{t})] \in \mathbb{R}^{N\times 2}$ | $[a,b]^T \in \mathbb{R}^2$ |
| | $p=1$ | $m=2$ | |
| *cosinus+phase* | $[\nu,\phi] \in \mathbb{R}^2$ | $[\cos(2\pi\nu\boldsymbol{t}+\phi)] \in \mathbb{R}^{N\times 1}$ | $x \in \mathbb{R}$ |
| | $p=2$ | $m=1$ | |
| *exponentielles complexes* | $\nu \in \mathbb{R}$ | $[\exp^{-i2\pi\nu\boldsymbol{t}}\ \exp^{+i2\pi\nu\boldsymbol{t}}] \in \mathbb{C}^{N\times 2}$ | $[A, B=A^*]^T \in \mathbb{C}^2$ |
| | $p=1$ | $m=2$ | |

TAB. 1.1 – Récapitulatif des trois modèles de sinusoïdes pour de l'analyse spectrale, dans le cadre des problèmes inverses non linéaires.

Dans le cas de l'observation d'objets célestes depuis le sol, le problème d'analyse spectrale est particulier, notamment à cause du caractère irrégulier de l'échantillonnage temporel des mesures. En effet, les conditions d'observation sont entre autres influencées par la disponibilité des télescopes, les conditions météorologiques, les possibles éclipses de l'objet d'intérêt par un autre corps céleste ou encore les alternances jour/nuit.

Pour rappel, l'échantillonnage d'un signal continu consiste à le multiplier par une somme de Dirac centrés aux instants d'échantillonnage $\boldsymbol{t} = [t_1, \ldots, t_N]^T$ ; le spectre $Y_e(f)$ d'un signal échantillonné $\boldsymbol{y}$ s'écrit alors comme la convolution du spectre du signal continu $Y(f)$ par la fenêtre spectrale $\mathcal{W}(f)$ :

$$y_e(t) = y(t)\sum_{n=1}^{N}\delta(t-t_n) \Rightarrow Y_e(f) = Y * \mathcal{W}(f) \quad \text{avec} \quad \mathcal{W}(f) = \sum_{n=1}^{N} e^{2j\pi f t_n}.$$

La Figure 1.1 permet de constater l'influence de l'échantillonnage régulier ou irrégulier sur la fenêtre spectrale d'un signal de $N = 100$ échantillons sur une durée de 5 jours. Pour un échantillonnage régulier de fréquence d'échantillonnage $f_e$ (ou de période d'échantillonnage $T_e$, soit $t_n = (n-1)T_e$), la fenêtre spectrale est un sinus cardinal périodisé de période $f_e$ dont le lobe principal est de largeur $2f_e/N$ et dont les lobes secondaires sont d'amplitude décroissante (voir Figure 1.1). Cela entraîne le phénomène de repliement spectral énoncé par le théorème de Nyquist-Shannon : les fréquences supérieures à $f_e/2$ d'un signal échantillonné régulièrement se retrouvent « repliées » dans l'intervalle $[-f_e/2; f_e/2]$, si bien que le spectre $Y_e(f)$ du signal régulièrement échantillonné contient des alias des hautes fréquences dans cet intervalle, et que l'on peut uniquement rechercher des fréquences inférieures à la moitié de la fréquence d'échantillonnage : $f_{max} < f_e/2$.

Pour un échantillonnage irrégulier, la forme de la fenêtre spectrale ne s'exprime plus comme un sinus cardinal périodisé ; on peut cependant l'observer dans les deux cas représentés en Figure 1.1. Au milieu, les N points ont été distribués aléatoirement sur l'intervalle. On constate la disparition de la périodisation de la fenêtre spectrale. En bas, l'échantillonnage est irrégulier mais présente des « trous » périodiques dûs à l'alternance jours/nuits. Si on perd aussi la périodisation de la fenêtre spectrale, on peut cependant constater l'apparition de forts lobes secondaires (dont l'amplitude dépasse 60% de celle du lobe principal) aux alentours de 1 jours$^{-1}$.

La non-régularité de l'échantillonnage implique donc un certain avantage, dans le sens où la

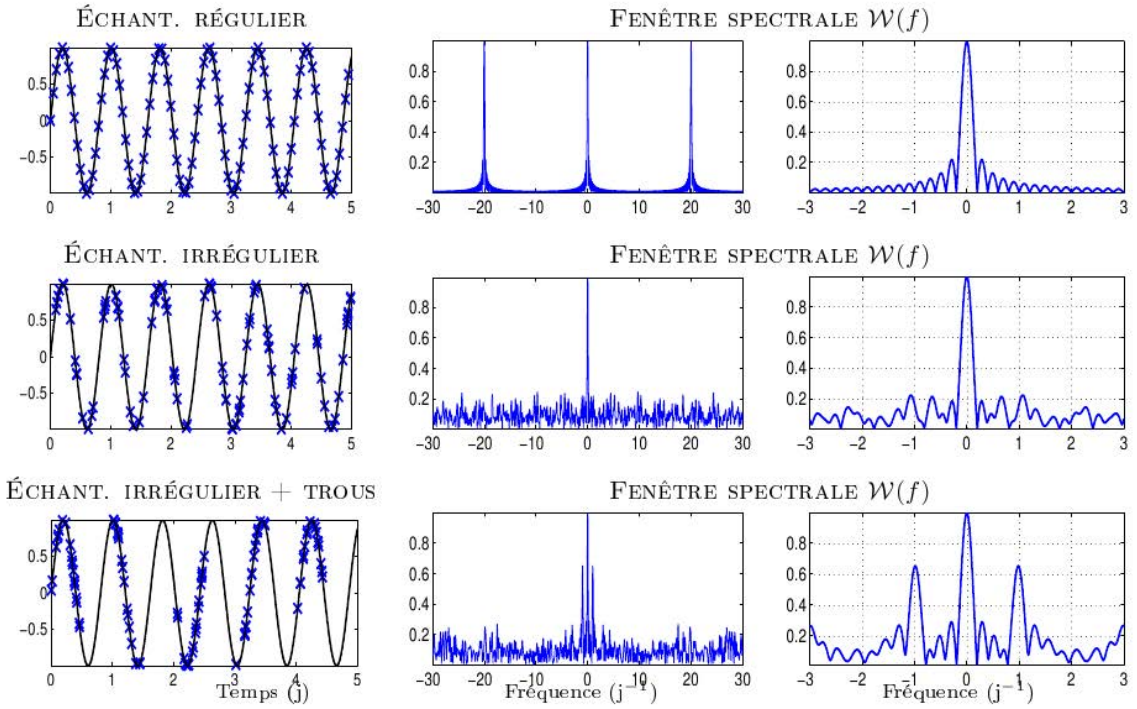



Fig. 1.1 – Influence de l'échantillonnage régulier ou irrégulier sur la fenêtre spectrale. A gauche les schémas d'échantillonnage. Au milieu et à droite le module des fenêtres spectrales associées (les figures de droite correspondent à un zoom autour de $f = 0$). Source : [BC13].

limite de Shannon n'existe plus - d'après [Bou05], la fréquence limite est reléguée à la fréquence critique $f_c = 1/2\tau$, où $\tau$ est le plus grand diviseur de tous les instants d'échantillonnage. On peut donc théoriquement rechercher des fréquences bien plus élevées que dans le cas de l'échantillonnage régulier. Cependant, le caractère irrégulier de l'échantillonnage produit de nombreux lobes secondaires dont l'amplitude ne décroît pas avec la fréquence comme dans le cas du sinus cardinal. De plus, la présence de trous périodiques dans les observations, provoqués par exemple par l'alternance jour/nuit, peut entraîner des lobes secondaires d'amplitudes très élevées rendant ces problèmes d'analyse spectrale très difficiles. En effet, la convolution d'un spectre de raies par la fenêtre spectrale va faire apparaître de nombreux pics correspondant à ces lobes secondaires et il sera délicat d'analyser le spectre du signal ainsi échantillonné.

Cet échantillonnage irrégulier représente donc la principale difficulté pour ces problèmes d'analyse spectrale. Notamment, il met en échec les méthodes de type *périodogrammes de Lomb-Scargle* et dérivés [Lom76, Sca82, FM81] (adaptations au cas irrégulier du périodogramme de Schuster du cas régulier), dans le cas où l'on recherche plusieurs fréquences ; dû à la convolution d'un spectre de raies par la fenêtre spectrale, les maxima de tels périodogrammes ne s'identifient plus avec les vraies fréquences présentes dans le signal [BC13]. Des méthodes de type déconvolution itérative peuvent alors être envisagées, comme les méthodes de type « prewhitening » [GD73] ou CLEAN [RLD87] proposées en astrophysique il y a plus de 30 ans, et qui s'interprètent comme des algorithmes gloutons. Plus de détails sur les méthodes classiquement employées pour résoudre ces problèmes peuvent être trouvés dans [Bab12] ou [Eng13].

Enfin, notons que la détection d'exoplanètes par la méthode des vitesses radiales peut être vue comme un cas particulier d'analyse spectrale en échantillonnage irrégulier (avec cette fois-

ci plusieurs paramètres de forme). Nous ne rentrons pas ici dans les détails de ce problème, puisqu'une étude plus détaillée lui sera consacrée au chapitre 4.

### Vers une reformulation en approximation parcimonieuse linéaire

Nous avons vu que ces problèmes inverses non linéaires ont des difficultés bien spécifiques liées à leur application, rendant particulièrement complexe leur résolution. Une direction a cependant été communément exploitée pour les résoudre : il s'agit de la reformulation en problème d'approximation parcimonieuse linéaire, qui est abordée dans un cadre général dans la section suivante.

## 1.2 Approximations parcimonieuses linéaires et différentes approches pour leur résolution

Dans cette section, nous présentons la reformulation des problèmes inverses non linéaires présentés précédemment en problèmes d'*approximation parcimonieuse linéaire* (APL), ainsi que les différentes approches utilisées pour les résoudre.

### 1.2.1 Une reformulation linéaire

Afin de résoudre les problèmes inverses non linéaires présentés précédemment, la reformulation du problème (1.1) comme un problème d'*approximation parcimonieuse linéaire* (APL) peut être employée. Pour les problèmes d'analyse spectrale vus précédemment, cette reformulation a par exemple été utilisée dans les travaux de [BC05]. Pour les problèmes de déconvolution impulsionnelle, il s'agit de leur reformulation naturelle (voir par exemple [OSK94]) : l'intégrale (1.2) est approchée par une convolution discrète et les décalages sont alors recherchés parmi les instants d'échantillonnage (voir § 1.5.2).

On considère ici que les paramètres non linéaires du modèle se réduisent au *paramètre de localisation* $\nu \in \mathcal{I}_\nu$ (la fréquence en analyse spectrale ou la position de l'écho en déconvolution impulsionnelle). La reformulation en APL consiste à discrétiser le *domaine des localisations* $\mathcal{I}_\nu = [\nu_{\min}; \nu_{\max}]$ pour former une grille régulière $\mathcal{G} = \{\ddot{\nu}_j\}_{j=1:J}$ arbitrairement fine de dimension $J \gg L$. Les paramètres de localisation $\nu$ seront alors recherchés au sein de cette grille et leur nombre $L$ sera supposé faible. Cette discrétisation permet de calculer le modèle $\mathbf{h}(\cdot)$ pour chaque $\ddot{\nu}_j$ de la grille et ainsi de regrouper les *atomes* $\mathbf{h}_j = \mathbf{h}(\ddot{\nu}_j) \in \mathbb{F}^{N\times m}$ en un *dictionnaire* $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_J] \in \mathbb{F}^{N\times mJ}$. Les $J$ amplitudes $\boldsymbol{x}_j$ de chaque atome $\mathbf{h}_j$ sont regroupées au sein d'un vecteur colonne $\boldsymbol{x} = [\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_J^T]^T$ de dimension $mJ$. Le problème d'APL s'écrit alors :

$$\text{Estimer } \boldsymbol{x} \textit{ parcimonieux} \text{ tel que } \boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{\epsilon} = \sum_{j=1}^{J} \mathbf{h}_j \boldsymbol{x}_j + \boldsymbol{\epsilon}. \tag{1.4}$$

Le modèle non linéaire $\sum_\ell \mathbf{h}(\nu_\ell)\boldsymbol{x}_\ell$ est donc approché par la combinaison linéaire $\sum_j \mathbf{h}_j \boldsymbol{x}_j$ où seulement un faible nombre d'amplitudes $\boldsymbol{x}_j$ sont non nulles ; il s'agit donc d'un *a priori* de *parcimonie* sur les paramètres $\boldsymbol{x}$. Dans notre modèle, notons que nous autorisons une *parcimonie structurée* ($m > 1$), dans le sens où $\boldsymbol{x}_j = \mathbf{0}$ signifie que les $m$ composantes de $\boldsymbol{x}_j$ sont nulles simultanément. On définit alors le *support* $\Omega = \mathsf{Supp}\,(\boldsymbol{x})$ comme l'ensemble des indices $j$ tels que $\boldsymbol{x}_j \neq \mathbf{0}$ :

$$\Omega = \mathsf{Supp}\,(\boldsymbol{x}) = \{j \in [\![1; J]\!],\ \boldsymbol{x}_j \neq \mathbf{0}\} \tag{1.5}$$

Cette formulation en APL permet de simplifier, en théorie, l'estimation de l'ordre du modèle $L$, puisqu'il s'agira du nombre de $\boldsymbol{x}_j$ estimés non nuls ; soit de la dimension du support d'une solution $\widehat{\boldsymbol{x}}$ :

$$\widehat{L} = \mathrm{Card}\left(\mathsf{Supp}\left(\widehat{\boldsymbol{x}}\right)\right) \tag{1.6}$$

De plus, elle a aussi pour avantage de faire disparaître la non linéarité du modèle : les paramètres de localisation sont directement estimés comme les $\ddot{\nu}_j$ associés aux $\boldsymbol{x}_j \neq \mathbf{0}$. Notons qu'une telle linéarisation du problème n'est plus valable quand l'atome continu dépend également de paramètres de forme dont la non-linéarité ne peut pas être supprimée par discrétisation[3]. Le problème de détection - estimer les $\{\nu_\ell\}_\ell$ et leur nombre $L$ - est donc équivalent à l'estimation du support $\Omega$, soit de trouver les indices pour lesquels les amplitudes sont non nulles. L'estimation de ces dernières est en soi un problème facile une fois le support $\Omega$ estimé ; elle peut être réalisée par moindres carrés par exemple. Dans le cadre de la résolution des problèmes inverses non linéaires considérés (1.1), la résolution du problème d'APL doit donc être vue au sens de la détection du support $\Omega$, et non comme la recherche d'une bonne approximation $\mathbf{H}\boldsymbol{x}$ de $\boldsymbol{y}$.

Notons d'ores et déjà que la discrétisation du paramètre de localisation induit nécessairement une *erreur de précision* qui se rajoute au bruit, puisque l'estimation des paramètres non linéaires se fait de façon discrète. Nous reviendrons sur cette complication en § 1.3.

Dans la suite de cette section, nous allons rapidement aborder les principales approches utilisées pour résoudre le problème d'APL (1.4).

### 1.2.2 Approches déterministes

Une première approche pour résoudre le problème d'APL (1.4) est de le formuler comme un problème d'optimisation bi-objectif. En effet, on cherche à estimer les amplitudes $\boldsymbol{x}$ qui permettent à la fois d'approcher au mieux le signal $\boldsymbol{y}$ comme une combinaison linéaire des *atomes* $\mathbf{h}_j$ du *dictionnaire* $\mathbf{H}$, mais aussi de minimiser le nombre $L$ d'atomes utilisés pour cette approximation. Dans un cadre déterministe, le problème (1.4) s'écrit alors comme le problème d'optimisation suivant :

$$\min_{\boldsymbol{x}} \left(\mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H}), \|\boldsymbol{x}\|_0\right). \tag{1.7}$$

où on définit les deux critères à minimiser simultanément suivant $\boldsymbol{x}$ :

- $\mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H})$ est le terme d'*attache aux données*, et caractérise la fidélité entre de l'approximation $\mathbf{H}\boldsymbol{x}$ vis-à-vis du signal $\boldsymbol{y}$. Il est généralement exprimé comme la norme $\ell_p$ du résidu à la puissance $p$ : $\mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H}) = \|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|_p^p$, où $p$ désigne ici le degré de la norme[4] (1, 2, $\infty$...). En général, on omettra le degré de la norme $\ell_2$ : $\|\cdot\| = \|\cdot\|_2$.
- $\|\boldsymbol{x}\|_0 = L$ compte le nombre d'amplitudes $\boldsymbol{x}_j$ non nulles. Bien qu'il ne s'agisse pas d'une norme, ce critère est couramment appelé *pseudo-norme* $\ell_0$, dénomination que nous utiliserons dans la suite de ce manuscrit. Elle sera utilisée dans nos travaux au sens de la *parcimonie structurée* : $\|\boldsymbol{x}\|_0 = \mathrm{Card}(\mathsf{Supp}\left(\boldsymbol{x}\right))$.

Trois formulations permettent d'écrire ce problème bi-objectif comme un problème d'optimisation avec ou sans contraintes, que nous donnons dans les équations suivantes :

$$\begin{aligned}
\textit{Contrainte sur la parcimonie}\ (L_0) \quad &: \quad \min_{\boldsymbol{x}} \mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H}) \quad \text{tel que } \|\boldsymbol{x}\|_0 \leq L_0 \\
\textit{Contrainte sur l'erreur}\ (\epsilon_0) \quad &: \quad \min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0 \quad \text{tel que } \mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H}) \leq \epsilon_0 \\
\textit{Pénalisation}\ (\lambda_0) \quad &: \quad \min_{\boldsymbol{x}} \mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H}) + \lambda_0 \|\boldsymbol{x}\|_0
\end{aligned} \tag{1.8}$$

[3] Il est bien sûr possible de discrétiser les paramètres de forme de la même façon que les paramètres de localisation, mais la taille de la grille résultante augmenterait de façon exponentielle. Un des objectifs de cette thèse étant l'estimation continue des paramètres non linéaires, nous n'insistons pas sur cette piste.

[4] Ne pas confondre ici le degré $p$ de la norme avec la dimension des paramètres non linéaires $\boldsymbol{\nu}$.

Ces trois formulations dépendent chacune d'un paramètre auxiliaire appelés par la suite *hyper-paramètre*, qui permet de régler le compromis entre l'attache aux données et la parcimonie : $L_0$ est l'ordre maximal du modèle, $\epsilon_0$ est une erreur d'approximation à ne pas dépasser, et $\lambda_0$ règle le poids de la parcimonie vis-à-vis de l'attache aux données. Dans la suite de ce manuscrit, on utilisera les termes *problèmes* $\ell_0$ et *solutions* $\ell_0$ pour faire référence respectivement aux formulations (1.8) et à leur solution. Quelle que soit la formulation choisie, le problème d'optimisation est dit *NP-difficile*, dans le sens où il revient à un problème d'*exploration combinatoire* à cause de la présence de la pseudo-norme $\ell_0$ [Nat95]. L'exploration de l'ensemble des configurations possibles est en effet impossible en pratique. Nous revenons dans les paragraphes suivants sur les différentes directions qui ont été suivies pour tâcher d'approcher au mieux les solutions des problèmes (1.8). Dans la suite de ce chapitre, et pour l'ensemble du manuscrit, le terme d'attache aux données sera choisi comme l'erreur quadratique :

$$\mathcal{Q}(\boldsymbol{x};\boldsymbol{y},\mathbf{H}) = \|\boldsymbol{y}-\mathbf{H}\boldsymbol{x}\|^2$$

L'erreur quadratique pourra si besoin être pondérée par la matrice de covariance du bruit $\boldsymbol{\Sigma}$ (on reviendra sur cette pondération en § 1.2.3) :

$$\mathcal{Q}(\boldsymbol{x};\boldsymbol{y},\mathbf{H},\boldsymbol{\Sigma}) = \|\boldsymbol{y}-\mathbf{H}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}^{-1}} = (\boldsymbol{y}-\mathbf{H}\boldsymbol{x})^\dagger\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\mathbf{H}\boldsymbol{x})$$

Par souci de clarté, nous omettons $\boldsymbol{\Sigma}$ dans la suite de ce chapitre. Pour clarifier les explications suivantes, nous définissons ici le *minimiseur des moindres carrés* (dans le cas où la dimension de $\boldsymbol{x}$ est inférieure à celle de $\boldsymbol{y}$) :

$$\begin{aligned}
\widehat{\boldsymbol{x}}^{\mathrm{MC}}(\boldsymbol{y},\mathbf{h}) &= \arg\min_{\boldsymbol{x}}\|\boldsymbol{y}-\mathbf{h}\boldsymbol{x}\|^2 \\
&= (\mathbf{h}^T\mathbf{h})^{-1}\mathbf{h}^T\boldsymbol{y} && \text{(signaux à valeurs réelles)} \\
&= (\mathbf{h}^\dagger\mathbf{h})^{-1}\mathbf{h}^\dagger\boldsymbol{y} && \text{(signaux à valeurs complexes)}
\end{aligned}$$

La notation $\widehat{\boldsymbol{x}}^{\mathrm{MC}}(\boldsymbol{y},\mathbf{h})$ correspondra donc, par la suite, au minimiseur d'un tel critère en fonction d'un signal $\boldsymbol{y}$ et d'une matrice $\mathbf{h}$, dont la dépendance sera omise quand aucune confusion n'est possible ; on notera alors uniquement $\widehat{\boldsymbol{x}}^{\mathrm{MC}}$.

**Algorithmes gloutons et heuristiques d'exploration.** Les algorithmes gloutons (ou *greedy algorithms*, voir par exemple [Tro04]) sélectionnent, itérativement, les « meilleurs » atomes du dictionnaire, au sens du terme d'attache aux données $\|\boldsymbol{y}-\mathbf{H}\boldsymbol{x}\|^2$. Nous pouvons décrire leur principe via le plus connu d'entre eux, le *Matching Pursuit* (MP) [MZ93]. Cependant nous généralisons ici cet algorithme au cas d'atomes à $m > 1$ colonnes car cela nous sera utile par la suite. À partir de l'initialisation du résidu $\boldsymbol{r}^0 = \boldsymbol{y}$ et de l'ensemble des indices des atomes sélectionnés $\Omega^0 = \varnothing$ (ce qui se ramène au support de la solution courante), on répète les étapes de *sélection* et de *mise à jour* décrites ci-dessous, pour l'étape $\ell$ :

1. *la sélection* : parmi les atomes $(\mathbf{h}_j)_{j\notin\Omega}$ du dictionnaire, on sélectionne celui qui permet de minimiser l'erreur quadratique vis-à-vis du résidu $\boldsymbol{r}^{\ell-1}$ :

   $$\Omega^\ell = \Omega^{\ell-1}\cup\{j_\ell\} \text{ avec } j_\ell = \arg\min_j \|\boldsymbol{r}^{\ell-1}-\mathbf{h}_j\widehat{\boldsymbol{x}}^{\mathrm{MC}}_j\|^2$$

   Dans le cas à une seule colonne ($m = 1$), cette sélection se ramène classiquement à $j_\ell = \arg\min_j|\boldsymbol{h}_j^\dagger\boldsymbol{r}^{\ell-1}|$ dans le cas de dictionnaires normalisés ($\|\boldsymbol{h}_j\|^2 = 1$).

2. *la mise à jour* : le résidu et les amplitudes sont mis à jour en retirant la contribution de l'atome sélectionné :

   $$\boldsymbol{r}^\ell = \boldsymbol{r}^{\ell-1}-\mathbf{h}_{j_\ell}\widehat{\boldsymbol{x}}^{\mathrm{MC}}(\boldsymbol{r}^{\ell-1},\mathbf{h}_{j_\ell}) \quad\text{et}\quad \boldsymbol{x}_{j_\ell} = \boldsymbol{x}_{j_\ell}+\widehat{\boldsymbol{x}}^{\mathrm{MC}}(\boldsymbol{r}^{\ell-1},\mathbf{h}_{j_\ell})$$

De tels algorithmes nécessitent la définition d'un critère d'arrêt, afin de décider au bout de combien d'itérations $L_0$ les sélections s'arrêtent ; ils peuvent être vus comme des algorithmes sous-optimaux tentant d'approcher la solution des problèmes définies en (1.8), de type *contrainte par la parcimonie* (minimiser le terme d'attache aux données avec un ordre $L < L_0$) ou *bornée par l'erreur* (ajouter itérativement des atomes le terme d'attache aux données passe sous un seuil $\epsilon_0$). Le caractère itératif de ces algorithmes est à la fois leur principal atout - le coût de calcul est très réduit par rapport à une exploration combinatoire - mais aussi leur principal défaut. Notamment, le *Matching Pursuit* ne remet jamais en cause les atomes sélectionnés ni les amplitudes associées et n'explore qu'une partition réduite de l'ensemble des configurations possibles. Notons tout de même que des formes plus avancées d'algorithmes gloutons permettent une meilleure exploration ; par exemple, l'*Orthogonal Matching Pursuit* (OMP) [PRK93] améliore l'étape de mise à jour en réévaluant l'ensemble des amplitudes des atomes sélectionnés par moindres carrés :

$$\boldsymbol{r}^{\ell} = \boldsymbol{y} - \overline{\mathbf{H}}_{\ell}\, \widehat{\boldsymbol{x}}^{\mathrm{MC}}(\boldsymbol{y}, \overline{\mathbf{H}}_{\ell})$$

où $\overline{\mathbf{H}}_{\ell}$ regroupe l'ensemble des atomes $(\mathbf{h}_j)_{j\in\Omega_{\ell}}$ sélectionnés à l'étape $(\ell)$. L'*Orthogonal Least Square* (OLS) [CBL89] optimise quant à lui l'étape de sélection sur le même principe ; le nouvel atome sélectionné sera celui qui, associé aux atomes déjà sélectionnés, minimise l'erreur quadratique avec le signal $\boldsymbol{y}$ :

$$\Omega^{\ell} = \Omega^{\ell-1} \cup \{j_{\ell}\} \text{ avec } j_{\ell} = \arg\min_{j\notin\Omega^{\ell-1}} \|\boldsymbol{y} - [\overline{\mathbf{H}}_{\ell-1},\ \mathbf{h}_j]\, \widehat{\boldsymbol{x}}^{\mathrm{MC}}\|^2$$

Malgré ces modifications, les atomes sélectionnés ne sont toujours pas remis en cause. Cependant, des types d'algorithmes dit *forward-backward* permettent l'ajout mais aussi le retrait des atomes déjà sélectionnés, tentant ainsi de corriger une fausse détection. Parmi ces approches on peut citer l'algorithme FoBa de [Zha11] basé sur l'OMP mais permettant le retrait d'atomes. L'algorithme *OMP with replacement* proposé par [JTD11] propose quant à lui de remplacer un atome déjà détecté par un nouveau.

Par ailleurs, l'algorithme *Single Best Replacement* (SBR), proposé par [SIBD11], qui s'inspire de l'algorithme SMLR que nous verrons en § 1.2.3, peut être vu comme une approche *forward-backward* bien qu'il ne s'agisse pas d'un algorithme glouton par définition. Cet algorithme vise en effet à résoudre le problème *pénalisé par la pseudo-norme* $\ell_0$ (1.8). Des approches similaires ont été envisagées par [HD10] pour adapter des algorithmes gloutons comme l'OMP ou l'OLS à la résolution bayésienne du problème *pénalisé par la pseudo-norme* $\ell_0$.

Même si ces heuristiques permettent une exploration plus avancée de l'ensemble des configurations possibles, ce sont tout de même des méthodes sous-optimales ne donnant aucune garantie d'obtenir le support optimum du problème dans un cadre général.

**Relaxation $\ell_1$ de la pseudo-norme $\ell_0$.** Une des solutions les plus populaires est de relaxer la pseudo-norme $\ell_0$ par la norme $\ell_1$, dont la non dérivabilité en 0 favorise les solutions parcimonieuses [Ela10, TW10]. Le problème d'optimisation devient continu et convexe pour $\mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H})$ convexe, et donc notamment pour $\mathcal{Q}(\boldsymbol{x}; \boldsymbol{y}, \mathbf{H}) = \|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2$. Dans le cas d'amplitudes monodimensionnelles ($m = 1$), la norme $\ell_1$ s'exprime comme la somme de la valeur absolue (ou module) des amplitudes, $\|\boldsymbol{x}\|_1 = \sum_j |x_j|$ et la *formulation pénalisée* des équations (1.8) prend la forme classique suivante :

$$\min_{\boldsymbol{x}} \|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2 + \lambda \|\boldsymbol{x}\|_1 \tag{1.9}$$

Dans le cas d'amplitudes multi-dimensionnelles ($m > 1$), la norme $\ell_1$ du vecteur des amplitudes $\boldsymbol{x}$ favorise la parcimonie sur la globalité de ses $mJ$ composantes, mais de façon indépendantes.

Il est alors proposé d'utiliser une « norme mixte » (voir par exemple [KT09]) pour imposer la *parcimonie structurée*, soit pour favoriser la mise à zéro simultanée des composantes des $\boldsymbol{x}_j$. Dans le cadre des APL, la norme mixte $\ell_{p,q}$ est la norme $\ell_q$ du vecteur composé des $J$ normes $\ell_p$ des amplitudes $\boldsymbol{x}_j$. La norme mixte la plus utilisée dans ce contexte est la norme $\ell_{1,2}$, qui s'écrit :

$$\|\boldsymbol{x}\|_{1,2} = \sum_{j=1}^{J} \|\boldsymbol{x}_j\|_2 \tag{1.10}$$

Elle permet en effet de conserver un terme de pénalisation convexe et favorise bien la parcimonie structurée.

L'utilisation de ces relaxations permet d'obtenir un problème d'optimisation convexe pour lequel il existe un unique minimum. La parcimonie est favorisée par la non-dérivabilité de la norme $\ell_1$ (ou $\ell_{1,2}$) en 0, mais qui rend sa résolution plus compliquée. On parle alors de problème d'optimisation non-lisse. Le problème d'optimisation (1.9) peut être formulé comme un problème de programmation quadratique (QP) pour lequel de nombreux algorithmes génériques existent. Des algorithmes plus spécifiques à la parcimonie ont été proposés, par exemple, des algorithmes de seuillage itératif [DDDM04], d'*homotopie* [OPT00], de gradients proximaux [BT09] et de *Block Coordinate Descent* [BCI07].

L'hyperparamètre $\lambda$ règle le compromis entre la parcimonie de la solution et l'attache aux données. Plus il sera élevé, plus la pénalisation $\ell_1$ aura d'importance dans le critère et plus la solution sera donc parcimonieuse ; mais plus le terme d'attache aux données sera élevé. Il nécessite donc un réglage approprié, pour lequel nous donnons deux méthodes couramment utilisées :

- *sélection de $\lambda$ dans le « coude » de la « courbe en L » :* pour plusieurs $\lambda$ compris dans un intervalle $[\lambda_{\min}; \lambda_{\max}]$, on calcule les solutions $\widehat{\boldsymbol{x}}_\lambda$ du problème (1.9). On trace alors le terme de pénalisation $\|\widehat{\boldsymbol{x}}_\lambda\|_1$ en fonction de l'erreur quadratique $\mathcal{Q}(\widehat{\boldsymbol{x}}_\lambda, ; \boldsymbol{y}, \mathbf{H}) = \|\boldsymbol{y} - \mathbf{H}\widehat{\boldsymbol{x}}_\lambda\|^2$. Il a été montré dans [Han92] que cette courbe avait une forme de « L » et que les solutions dans le « coude » du L donnait un bon compromis, dans le cas d'une pénalisation quadratique. Cela reste une bonne indication dans le cas $\ell_1$.
- *diminution de $\lambda$ à partir d'un $\lambda_{max}$* : ici, on calcule les solutions $\widehat{\boldsymbol{x}}_\lambda$ de (1.9) à partir d'un $\lambda = \lambda_{\max}$ que l'on fait décroître jusqu'à ce que le résidu $\boldsymbol{r} = \boldsymbol{y} - \mathbf{H}\boldsymbol{x}$ puisse s'identifier à du bruit. Une telle procédure est particulièrement intéressante associée à des algorithmes d'*homotopy*, puisque ces derniers calculent de fait toutes les solutions du problème pour les valeurs de $\lambda$ inférieures à $\lambda_{max}$ jusqu'à une valeur minimale souhaitée.

Un des inconvénients majeurs du problème (1.9) réside dans la *sous-estimation* des amplitudes non nulles. En effet, à support et attache aux données équivalents, une solution est meilleure si sa pénalisation $\ell_1$ et donc ses amplitudes $\boldsymbol{x}_j$ non nulles sont plus faibles. Pour résoudre ce problème, il est nécessaire de réévaluer les amplitudes après résolution de (1.9). À partir d'une solution $\widehat{\boldsymbol{x}}_\lambda$ et du support associé $\Omega = \mathsf{Supp}\,(\widehat{\boldsymbol{x}}_\lambda)$, on réévalue les amplitudes non nulles par moindres carrés sur le dictionnaire réduit au support $\overline{\mathbf{H}}_\Omega = [\mathbf{h}_j]_{j\in\Omega}$, soit $\overline{\boldsymbol{x}}_\Omega = \widehat{\boldsymbol{x}}^{\mathrm{MC}}(\boldsymbol{y}, \overline{\mathbf{H}}_\Omega)$. Cette réévaluation permet ainsi d'obtenir les meilleurs amplitudes $\overline{\boldsymbol{x}}_\Omega$ en terme d'attache aux données, pour le support $\Omega$ estimé à partir de la solution de (1.9). Cependant, ce support fait souvent apparaître de fausses détections à cause de la pénalisation $\ell_1$ des amplitudes, qu'une telle réévaluation ne peut corriger.

**D'autres relaxations de la pseudo-norme $\ell_0$.** La norme $\ell_1$ n'est pas la seule à être employée pour relaxer la pseudo-norme $\ell_0$. Notamment, des relaxations non convexes ont également été

étudiées ; elles permettent une meilleure prise en compte de la parcimonie au prix de devoir traiter des problèmes d'optimisation non convexe. Par exemple, on peut citer la relaxation « lisse » non convexe envisagée par [CJPT13] ou la relaxation continue CEL0 proposée par [SBFA15]. Une synthèse de ces approches est donnée par [SBFA17], où des conditions sont données pour garantir que la relaxation non convexe conserve l'optimum global de la solution initiale parmi l'ensemble des minimiseurs locaux. Cependant, aucun algorithme à ce jour n'est garanti d'obtenir le minimum global de ces critères relaxés.

**Équivalence avec le problème $\ell_0$.** Les algorithmes gloutons, les relaxations de la pseudo-norme $\ell_0$, et les autres méthodes vues précédemment sont des méthodes sous-optimales cherchant à approcher la solution d'une des formulations $\ell_0$ (1.8). Dans le cas des algorithmes gloutons et de la relaxation $\ell_1$, des conditions suffisantes d'équivalence entre les solutions de ces problèmes existent et se basent sur des indicateurs tels que la cohérence mutuelle (e.g. [Tro04, Fuc05]), la constante d'isométrie restreinte (e.g. [CRT06]) ou le coefficient d'*exact recovery* (e.g. [Tro04, Tro06]). Ces indicateurs peuvent être vus comme des formes de « mesures de corrélation » des atomes $\mathbf{h}_j$ du dictionnaire $\mathbf{H}$. Pour résumer, moins le dictionnaire est *corrélé*, et plus les solutions des méthodes de type gloutonnes ou relaxation $\ell_1$ s'approcheront de celles du problème $\ell_0$. Ces conditions peuvent être couramment respectées dans le cadre de l'*acquisition compressée* où la matrice $\mathbf{H}$ s'exprime généralement comme la multiplication d'une matrice de mesure et d'une matrice orthogonale. Au contraire, ce sera rarement le cas dans le cadre de nos problèmes inverses non linéaires, où les atomes $\mathbf{h}_j$ correspondent à la même fonction $\mathbf{h}(\cdot)$ exprimée en différents paramètres de localisation $\ddot{\nu}_j$ discrétisés sur une grille : $\mathbf{h}_j$ et $\mathbf{h}_{j+1}$ sont par exemple très similaires. Il a par exemple été montré que les indicateurs nommés ci-dessus ne respectaient les conditions suffisantes dans le cas d'un problème d'analyse spectrale en échantillonnage irrégulier [BC13]. En pratique, pour ce problème, les algorithmes gloutons les plus classiques (type MP ou OMP) donnent de nombreuses fausses détections dans le support. L'utilisation de la relaxation $\ell_1$ ou d'algorithmes de type OLS donnent généralement des solutions plus satisfaisantes, sans pour autant garantir d'obtenir le bon support (voir [BC13]). Une conclusion similaire est donnée dans [BSCI11] pour le problème de déconvolution impulsionnelle, où des fausses détections sont constatées, tant pour des algorithmes gloutons que pour la relaxation $\ell_1$. Cette non-équivalence avec le problème $\ell_0$ et les fausses détections au sein du support estimé sont préjudiciables dans le cadre des problèmes inverses où l'estimation correcte du support est primordiale pour répondre au problème de détection.

**Reformulation des problèmes $\ell_0$ en MIP.** Afin de dépasser les problèmes rencontrés par les méthodes sous-optimales, il a été proposé récemment par [BNCM16] de s'attaquer à la résolution *exacte* des problèmes $\ell_0$ (1.8), dans le sens où il est garanti de trouver l'optimum global des problèmes $\ell_0$. Plus précisément, des *variables binaires* $\boldsymbol{b} \in \{0,1\}^J$ sont introduites pour caractériser la nullité des $\boldsymbol{x}_j$ :

$$b_j = 0 \Leftrightarrow \boldsymbol{x}_j = \mathbf{0} \tag{1.11}$$

À partir de cette équivalence, la pseudo-norme $\ell_0$ des amplitudes $\boldsymbol{x}$ peut directement s'exprimer comme la somme des variables binaires :

$$L = \|\boldsymbol{x}\|_0 = \|\boldsymbol{b}\|_0 = \sum_{j=1}^{J} b_j \tag{1.12}$$

La relation (1.11) entre les $b_j$ et $\boldsymbol{x}_j$ ne peut pas être directement prise en compte dans un problème d'optimisation ; c'est pourquoi [BNCM16] a choisi de la reformuler en *contraintes*

*linéaires d'inégalités* en faisant l'hypothèse dite du *big-M* : l'ensemble des composantes de $\boldsymbol{x}$ sont bornées par une constante $M$. Cette hypothèse n'est pas un problème dans le cadre de nos problèmes inverses, puisque les amplitudes sont finies et donc bornées par nature. En pratique, on peut donc généralement définir *a priori* une telle borne $M$. L'équivalence (1.11) se réécrit[5] alors :

$$-Mb_j\mathbf{1}_m \leq \boldsymbol{x}_j \leq Mb_j\mathbf{1}_m \tag{1.13}$$

ce qui correspond à $2m$ contraintes d'inégalités, 2 pour chaque composante de $\boldsymbol{x}_j$. L'introduction des variables binaires $\boldsymbol{b}$ et l'hypothèse de *big-M* permettent alors de reformuler les problèmes d'optimisation faisant intervenir la pseudo-norme $\ell_0$ (1.8) en Problème d'optimisation mixte en nombres entiers (*Mixed Integer Programming*, MIP). Par exemple, la reformulation en MIP de la formulation *contrainte par la parcimonie*, que nous utiliserons au Chapitre 2, s'écrit de la façon suivante :

$$\min_{\boldsymbol{x},\boldsymbol{b}} \|\boldsymbol{y}-\mathbf{H}\boldsymbol{x}\|^2 \text{ s.t. } \begin{cases} \boldsymbol{b} \in \{0,1\}^J \\ -M\boldsymbol{b} \leq \boldsymbol{x} \leq M\boldsymbol{b} \\ \sum_{j=1}^J b_j \leq L_0 \end{cases} \tag{1.14}$$

Ici, la somme des $b_j$ inférieure à $L_0$, associée aux contraintes de *big-M*, remplacent la contrainte de parcimonie $\|\boldsymbol{x}\|_0 \leq L_0$ du problème (1.8). L'inégalité $-M\boldsymbol{b} \leq \boldsymbol{x} \leq M\boldsymbol{b}$ se comprend comme l'inégalité (1.13) pour tout $j \in [\![1;J]\!]$. Le problème (1.14) est un *Mixed-Integer Quadratic Program* (MIQP), dans le sens où il s'agit de minimiser un critère quadratique sous contraintes linéaires, avec simultanément des variables réelle et des variables entières. Notons que [BNCM16] propose les reformulations en MIP pour chaque formulation de (1.8), et ce pour différents termes d'attaches aux données (erreur quadratique, mais aussi normes $\ell_1$ et $\ell_\infty$ du résidu).

La reformulation des problèmes (1.8) en MIP ne diminue pas la complexité calculatoire de leur résolution, puisqu'ils sont toujours NP-difficile et requièrent *a priori* une exploration combinatoire pour les variables binaires $\boldsymbol{b}$. Cependant, elle permet de bénéficier des récentes avancées en optimisation discrète et linéaire, qui proposent des algorithmes de résolution garantissant de trouver l'optimum global tout en évitant l'exploration exhaustive de toutes les configurations [Bix12]. Des logiciels de résolution efficace des MIP sont disponibles ; notamment, le logiciel CPLEX connu pour être l'un des plus performants, est disponible gratuitement sous licence académique. La plupart sont basés sur des algorithmes dits de *Branch&Cut*, dont le principe est d'utiliser une stratégie de *Branch&Band* - exploration des configurations sur un *arbre de décision* - grâce à des relaxations continues successives des variables entières permettant de « couper » certaines branches de l'arbre sans risque de manquer la configuration optimale. La résolution de tels problèmes ne rentrant pas dans le cadre de cette thèse, nous nous référons à [BNCM16] et aux références qui y figurent pour plus de détails.

Ainsi, [BNCM16] a pu calculer les solutions des problèmes $\ell_0$ (1.8) grâce aux reformulations MIP en utilisant le logiciel CPLEX dans des temps raisonnables, pour des problèmes de déconvolution impulsionnelle linéaires (problème d'APL). Par exemple, pour des données de taille $N = 120$ bruitées avec un RSB $= 20$ dB, une grille de taille $J = 100$ et un ordre du modèle $L = 7$ (supposé connu, soit en posant $L_0 = L = 7$), la solution du problème MIP (1.14) issu de la formulation *contrainte par la parcimonie* est obtenue (en moyenne sur 50 signaux différents) en 0.6s (temps CPU), et son optimalité est prouvée en moyenne en 1.3s (alors qu'il faudrait plus

[5] On aura bien une équivalence $b_j = 0 \Leftrightarrow \boldsymbol{x}_j = \mathbf{0}$ quand on utilisera la contrainte du *big-M* $|\boldsymbol{x}_j| \leq Mb_j$ dans la reformulation en MIP. Par exemple, dans le cas de la reformulation (1.14), si $\boldsymbol{x}_j = \mathbf{0}$, avoir $b_j = 0$ est nécessairement plus avantageux puisque cela autorise un autre $b_{k\neq j}$ et donc $\boldsymbol{x}_{k\neq j}$ à être non nuls, et donc à minimiser davantage le terme d'attache aux données.

d'un an pour un parcours exhaustifs de toutes les configurations possibles). Le temps de calcul augmente bien sûr fortement avec l'ordre du modèle $L$, et le coût de calcul reste globalement beaucoup plus élevé que pour les méthodes sous-optimales vues précédemment. Cependant, les reformulations en MIP pourront bénéficier à la fois des améliorations des performances de calculs des ordinateurs, mais aussi des adaptations des MIP aux spécificités des problèmes parcimonieux (voir [MBN19]) ce que ne peuvent pas faire des logiciels de résolution générique comme CPLEX. En particulier, cette adaptation spécifique fait actuellement l'objet d'une thèse[6] au LS2N à Nantes.

**Bilan sur les approches déterministes.**

Nous avons vu que le problème d'APL pouvait s'exprimer comme un problème d'optimisation bi-objectif faisant intervenir la pseudo-norme $\ell_0$, où l'on souhaite minimiser à la fois l'attache aux données et l'ordre du modèle. La solution de ce problème d'optimisation peut être approchée à l'aide de méthodes sous-optimales comme les algorithmes gloutons, la relaxation $\ell_1$, ou d'autres relaxations et heuristiques, ou alors peut-être calculée de façon exacte grâce à une reformulation en MIP, au prix d'un coût de calcul plus élevé.

### 1.2.3 Approches stochastiques

Nous considérons maintenant le problème d'APL (1.4) de façon probabiliste : les paramètres d'amplitudes $\boldsymbol{x}$ ainsi que le bruit $\boldsymbol{\epsilon}$ sont supposés être les réalisations de variables aléatoires. Notons que dans l'ensemble de ce manuscrit, on ne distinguera pas dans nos notations la variable aléatoire de sa réalisation. Le signal $\boldsymbol{y}$ est donc un processus aléatoire résultant de la réalisation du modèle linéaire suivant :

$$\underbrace{\boldsymbol{y}}_{\text{Aléa.}} = \underbrace{\mathbf{H}}_{\text{Déter.}} \underbrace{\boldsymbol{x}}_{\text{Aléa.}} + \underbrace{\boldsymbol{\epsilon}}_{\text{Aléa.}}$$

où le dictionnaire $\mathbf{H}$ est déterministe. Le problème d'APL revient alors à estimer, à partir d'une réalisation du signal $\boldsymbol{y}$, la réalisation des amplitudes $\boldsymbol{x}$ qui est « responsable » de l'obtention d'un tel signal. Une première approche est d'utiliser les connaissances que nous avons sur le bruit ; en effet, la loi probabilité du signal $\boldsymbol{y}$ connaissant les amplitudes $\boldsymbol{x}$ s'écrit suivant la loi *a priori* sur le bruit :

$$p(\boldsymbol{y} \,|\, \boldsymbol{x}) = p(\mathbf{H}\boldsymbol{x} + \boldsymbol{\epsilon} \,|\, \boldsymbol{x}) = p_{\boldsymbol{\epsilon}}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x} \,|\, \boldsymbol{x})$$

Prise comme une fonction des amplitudes $\boldsymbol{x}$, cette fonction est appelée vraisemblance (*likelihood* en angais). C'est une mesure de l'adéquation entre la réalisation $\boldsymbol{y}$ des données et la réalisation $\boldsymbol{x}$, sans *a priori* sur les amplitudes $\boldsymbol{x}$. Une deuxième approche consiste à étudier la loi de probabilité *a posteriori* des amplitudes $\boldsymbol{x}$ connaissant $\boldsymbol{y}$, qui prend en compte la vraisemblance, mais aussi une information *a priori* sur les amplitudes $\boldsymbol{x}$. En effet, d'après la règle de Bayes, la probabilité *a posteriori* s'écrit :

$$\underbrace{p(\boldsymbol{x} \,|\, \boldsymbol{y})}_{\substack{\text{loi } a \text{ posteriori}}} = \underbrace{\frac{1}{p(\boldsymbol{y})}}_{\substack{\text{constante de} \\ \text{normalisation}}} \underbrace{p(\boldsymbol{y} \,|\, \boldsymbol{x})}_{\text{vraisemblance}} \underbrace{p(\boldsymbol{x})}_{\substack{\text{loi} \\ a \text{ priori}}}$$

La prise en compte d'une loi *a priori* pour les amplitudes $\boldsymbol{x}$ associée à la vraisemblance plonge le problème d'APL dans le paradigme bayésien. C'est bien la prise en compte d'une connaissance

[6] Thèse en cours (2016-2019) menée par Ramzi Ben Mhenni et sous la direction de Sébastien Bourguignon.

*a priori* sur les amplitudes qui différencie la loi *a posteriori* de la vraisemblance.

La règle de Bayes permet d'écrire la loi *a posteriori* des paramètres d'intérêts $\boldsymbol{x}$, que l'on peut alors exploiter pour estimer ces paramètres. Les estimateurs les plus utilisés sont l'estimateur du *maximum a posteriori* (MAP) et l'estimateur de l'*espérance a posteriori* (EAP) qui s'écrivent :

$$\begin{array}{lrcl} \textit{Maximum a posteriori :} & \widehat{\boldsymbol{x}}_{\mathrm{MAP}}(\boldsymbol{y}) & = & \arg\max_{\boldsymbol{x}} p(\boldsymbol{x} \,|\, \boldsymbol{y}) \\ \textit{Espérance a posteriori :} & \widehat{\boldsymbol{x}}_{\mathrm{EAP}}(\boldsymbol{y}) & = & \mathbb{E}\{\boldsymbol{x} \,|\, \boldsymbol{y}\} = \int \boldsymbol{x}\, p(\boldsymbol{x} \,|\, \boldsymbol{y})\, \mathrm{d}\boldsymbol{x} \end{array}$$

Le choix de l'estimateur n'est pas une question évidente. L'estimateur du MAP donne les amplitudes $\boldsymbol{x}$ maximisant la loi *a posteriori*, soit les plus probables au regard des données et des connaissances *a priori*. Il peut être vu comme la minimisation en moyenne d'une erreur de type tout ou rien. L'estimateur de l'EAP correspond au *minimiseur de l'erreur quadratique moyenne* (MMSE), peut cependant donner des valeurs peu probables dans le cas de lois *a posteriori* multimodales.

Dans le cas de l'APL, la loi *a priori* sur les amplitudes $\boldsymbol{x}$ doit modéliser du point de vue statistique la parcimonie de la solution. Nous présentons ici le lien entre le modèle probabiliste et les approches déterministes décrites précédemment ; nous commençons par relier l'estimateur du MAP avec le problème d'optimisation bi-objectif $\ell_0$, puis nous détaillons un modèle probabiliste structurel de la parcimonie : le modèle Bernoulli-Gaussien.

**Interprétation probabiliste du problème bi-objectif au sens du MAP.** Le calcul de l'estimateur du *maximum a posteriori* (MAP) est un problème d'optimisation, puisque l'on cherche à maximiser la loi *a posteriori*. La constante de normalisation ne dépendant pas de $\boldsymbol{x}$, l'estimateur du MAP peut se réécrire sous la forme suivante :

$$\begin{array}{rcl} \widehat{\boldsymbol{x}}_{\mathrm{MAP}}(\boldsymbol{y}) & = & \arg\max\limits_{\boldsymbol{x}} \quad p(\boldsymbol{y} \,|\, \boldsymbol{x})\, p(\boldsymbol{x}) \\ & = & \arg\max\limits_{\boldsymbol{x}} \quad \underbrace{\log(p(\boldsymbol{y} \,|\, \boldsymbol{x}))}_{\text{log-vraisemblance}} \quad + \quad \underbrace{\log(p(\boldsymbol{x}))}_{\text{log-prior}} \end{array}$$

On voit alors apparaître un critère à deux termes : la log-vraisemblance et le logarithme de la loi *a priori*, ce qui rappelle la forme des problèmes d'optimisation bi-objectif (1.7) vue précédemment. Ce n'est pas étonnant, puisqu'en voulant maximiser la loi *a posteriori* selon $\boldsymbol{x}$, on cherche à la fois à maximiser la vraisemblance, une mesure de l'adéquation de $\boldsymbol{x}$ aux données $\boldsymbol{y}$, et la probabilité *a priori* d'obtenir $\boldsymbol{x}$. Dans le problème bi-objectif, l'utilisation de la pseudo-norme $\ell_0$ indique en effet que la solution est *a priori* parcimonieuse. Nous donnons ici quelques exemples de lien entre les problèmes déterministes et leur interprétation bayésienne.

Comme nous l'avons vu précédemment, la vraisemblance s'écrit suivant la loi sur le bruit. Ainsi, si l'on considère un bruit gaussien centré *indépendent et identiquement distribué* (*i.i.d.*) de variance $\sigma^2$, alors la log-vraisemblance est proportionnelle à la norme du résidu au carré, qui est le terme classique d'attache aux données vu précédemment :

$$\log p(\boldsymbol{y} \,|\, \boldsymbol{x}) \propto -\frac{\|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|_2^2}{2\sigma^2}$$

Cela justifie notamment la prise en compte de la variance du bruit (voir d'une matrice de covariance) dans le terme d'attache aux données dans les approches déterministes. De la même façon, la log-vraisemblance sera proportionnelle à $\|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|_1$ si le bruit est *i.i.d.* suivant une loi de Laplace.

En ce qui concerne la loi *a priori*, une loi de Laplace donne un terme de pénalisation dont le logarithme est proportionnel à la norme $\ell_1$ des amplitudes, ce qui est strictement équivalent à la classique relaxation $\ell_1$. Il n'existe pas de loi $p(\boldsymbol{x})$ qui s'exprime directement en fonction de la pseudo-norme $\ell_0$ des amplitudes. Cependant, l'introduction de variables binaires $\boldsymbol{q}$ suivant une loi de Bernoulli et conditionnant la nullité des amplitudes $\boldsymbol{x}$ permet d'introduire un tel *a priori* en considérant la loi jointe :

$$\log p(\boldsymbol{q}, \boldsymbol{x}) \propto \|\boldsymbol{x}\|_0 = \sum_j q_j$$

avec un *a priori* conditionnel particulier pour $\boldsymbol{x}$ (voir [SIBD11]). Nous revenons par la suite sur un tel modèle probabiliste.

Finalement, une telle interprétation probabiliste au sens du MAP des problèmes d'optimisation déterministes étudiés précédemment permet de justifier la forme des termes d'attache aux données en fonction des connaissances *a priori* sur le bruit. Elle donne aussi une meilleure interprétation des hyperparamètres réglant le compromis entre l'attache aux données et la parcimonie, notamment dans le cas de la formulation *pénalisée* (1.8), puisqu'ils peuvent s'écrire directement[7] en fonction des lois *a priori* sur le bruit et sur les amplitudes. Enfin, le contexte bayésien permet de considérer d'autres estimateurs que celui du MAP, comme celui de l'espérance *a posteriori* EAP, voire d'étudier la loi *a posteriori* dans son ensemble, puisqu'elle contient bien plus d'informations que simplement ces estimateurs. Le paragraphe suivant s'intéresse donc à un modèle bayésien modélisant structurellement la parcimonie, et permettant d'aller au-delà du cadre classique de l'estimateur du MAP : le modèle Bernoulli-Gaussien.

**Le modèle Bernoulli-Gaussien, modèle probabiliste pour la parcimonie.** Afin de décrire de manière structurelle un *a priori* de parcimonie sur les amplitudes, un modèle *Bernoulli-Gaussien* (BG) a été introduit pour la première fois par [KM82] pour la déconvolution de train d'impulsions.

Ce modèle repose sur l'introduction d'une séquence binaire $\boldsymbol{q} = [q_1, \ldots, q_J] \in \{0,1\}^J$ de la même dimension que la grille de discrétisation $\mathcal{G}$, telle que pour tout indice $j$, la variable $q_j$ indique si l'amplitude associée $\boldsymbol{x}_j$, est nulle via l'implication suivante :

$$q_j = 0 \Leftrightarrow \boldsymbol{x}_j = \boldsymbol{0} \tag{1.15}$$

Si l'introduction de variables binaires conditionnant la nullité des amplitudes associées rappelle la reformulation du problème $\ell_0$ en MIP vue dans le cadre déterministe, ce n'est pas un hasard ; l'idée de la reformulation en MIP présentée dans [BNCM16] est le fruit d'une collaboration avec des chercheurs en *Recherche Opérationnelle*, initiée après la présentation de travaux exploitant le modèle Bernoulli-Gaussien.

Les lois *a priori* sur la séquence binaire $\boldsymbol{q}$ et sur les amplitudes $\boldsymbol{x}$ permettent de traduire cette implication. La séquence binaire $\boldsymbol{q}$ est supposée *i.i.d.* et suit une loi *a priori* de Bernoulli :

$$q_j \sim \mathcal{B}(\lambda) \quad \Leftrightarrow \quad \left\{ \begin{array}{rcl} \Pr(q_j = 1) & = & \lambda \\ \Pr(q_j = 0) & = & 1 - \lambda \end{array} \right. \tag{1.16}$$

où l'hyperparamètre $\lambda \in [0;1]$ est la probabilité que chaque $q_j$ soit égal à 1.

[7] Notons tout de même que ces lois *a priori* dépendent elle-même d'hyperparamètres parfois inconnus, comme la variance du bruit ou encore la probabilité de Bernoulli dans le cas du modèle Bernoulli-Gaussien (voir paragraphe suivant). Ils ont cependant un « sens » en terme probabiliste qui peut être exploité pour les régler, contrairement aux hyperparamètres des formulations (1.8).

Conditionnellement à $q_j$, la loi *a priori* sur les amplitudes $\boldsymbol{x}_j$ est la suivante :

$$\boldsymbol{x}_j \,|\, q_j \sim \mathcal{N}(\mathbf{0}, q_j \sigma_x^2 \mathbf{I}_m) \quad \Leftrightarrow \quad \left\{ \begin{array}{lcl} \boldsymbol{x}_j \,|\, q_j = 1 & \sim & \mathcal{N}(\mathbf{0}, \sigma_x^2 \mathbf{I}_m) \\ \boldsymbol{x}_j \,|\, q_j = 0 & \sim & \text{loi Dirac en zéro} \end{array} \right. \tag{1.17}$$

Les amplitudes $\boldsymbol{x}_j$, supposées mutuellement indépendantes, sont gaussiennes centrées de variances $\sigma_x^2$ quand $q_j = 1$ et la loi Dirac en zéro force les amplitudes à être nulles dans le cas contraire. Cette définition *conditionnelle* de la loi *a priori* sur les paramètres $(\boldsymbol{q}, \boldsymbol{x})$ permet de respecter l'équivalence[8] de l'équation (1.15).

Les variables de Bernoulli $q_j$ permettent alors de caractériser la présence ($q_j = 1$) ou l'absence ($q_j = 0$) de l'atome $\mathbf{h}_j$ (et donc du $\ddot{\nu}_j$ associé) dans le signal $\boldsymbol{y}$. Notons que l'on peut directement exprimer l'ordre du modèle $L$ (soit la pseudo-norme $\ell_0$ des amplitudes) comme la somme des $q_j$ :

$$L = \|\boldsymbol{x}\|_0 = \|\boldsymbol{q}\|_0 = \sum_{j=1}^{J} q_j$$

La loi *a priori* de Bernoulli pour les variables binaires est donc bien un *a priori* de parcimonie sur les amplitudes $\boldsymbol{x}$, à condition que l'hyperparamètre $\lambda \in [0; 1]$, la probabilité *a priori* que $q_j = 1$, soit faible.

Notons que dans la littérature sur le modèle BG, un autre type de loi *a priori* est aussi considéré pour les amplitudes $\boldsymbol{x}_j$. Plutôt que d'utiliser une forme conditionnelle de la loi *a priori* comme [CCL96] par exemple, l'article fondateur du modèle BG [KM82] ou plus récemment [SIBD11, DHD12] utilise une forme produit, où chaque amplitude $\boldsymbol{x}_j$ s'écrit comme le produit de $q_j$ avec une nouvelle amplitude $\boldsymbol{a}_j$ ; ces dernières suivent toutes une loi *a priori* gaussienne, indépendamment de la valeur du $q_j$ associé. D'un point de vue théorique, ce deux formes sont strictement équivalentes puisque la loi *a priori* sur $\boldsymbol{x}_j$ (marginalisée par rapport à $q_j$) est strictement identique : $p(\boldsymbol{x}_j) = \lambda\, g(\boldsymbol{x}_j; \mathbf{0}, \sigma_x^2 \mathbf{I}_m) + (1 - \lambda)\delta_{\mathbf{0}}(\boldsymbol{x})$ (où $g(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Gamma})$ est la densité de probabilité d'une variable $\boldsymbol{x}$ de moyenne $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Gamma}$). Mais en pratique, cela amène à la manipulation de lois *a posteriori* différentes. Dans le cadre de ce manuscrit, nous utilisons la forme conditionnelle de la loi *a priori* dans la continuité des travaux de [Bou05].

Un tel modèle BG est généralement associé à un bruit gaussien centré de matrice de covariance $\boldsymbol{\Sigma}$. Le modèle est alors entièrement décrit par les paramètres $(\boldsymbol{q}, \boldsymbol{x})$ et les hyperparamètres $\boldsymbol{\theta} = (\lambda, \sigma_x^2, \boldsymbol{\Sigma})$ décrivant les lois *a priori* et la loi du bruit. On distingue le cadre *supervisé* où les hyperparamètres $\boldsymbol{\theta}$ sont supposés connus et où l'on étudie la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$, et le cadre *non supervisé* où les $\boldsymbol{\theta}$ sont inconnus et où l'on étudie $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$. L'estimation des paramètres peut se faire au sens du MAP ou de l'EAP.

Le problème d'optimisation résultant du calcul de l'estimateur du MAP des paramètres du modèle BG est très similaire aux problèmes d'optimisation faisant intervenir la pseudo-norme $\ell_0$. En effet, d'après [SIBD11], le calcul dans le cadre supervisé du maximiseur de la loi $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$ revient, pour un bruit *i.i.d.* de variance $\sigma_\epsilon^2$, à résoudre le problème de minimisation suivant[9] :

$$\arg\min_{\boldsymbol{x}, \boldsymbol{q}} \quad \|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2 + \lambda_0 \|\boldsymbol{x}\|_0 + \frac{\sigma_\epsilon^2}{\sigma_x^2} \|\boldsymbol{x}\|^2$$

[8] L'implication directe est évidente ($q_j = 0 \Rightarrow \boldsymbol{x}_j = \mathbf{0}$), et on a l'implication réciproque presque partout. En effet, si $\boldsymbol{x}_j \sim \mathcal{N}(\mathbf{0}, \sigma_x^2 \mathbf{I}_m,)$, la probabilité que $\boldsymbol{x}_j = \mathbf{0}$ est un ensemble de mesure nulle.

[9] Cette équivalence entre le calcul de l'estimateur du MAP et le problème pénalisé par la pseudo-norme $\ell_0$ est en fait démontrée par [SIBD11] pour le modèle BG *produit*, où $\boldsymbol{x}_j = q_j \boldsymbol{a}_j$. On peut cependant sans difficulté obtenir le même résultat en utilisant le modèle BG conditionnel.

avec $\lambda_0 = 2\sigma_\epsilon^2 \log(1/\lambda - 1)$. Cela revient, quand on fait tendre $\sigma_x^2$ vers l'infini[10], à la formulation *pénalisée* du problème $\ell_0$ (1.8) que nous avons vue dans le cas des approches déterministes.

Dans le cadre supervisé (et pour $\sigma_x^2 < \infty$), des algorithmes ont été proposés pour le calcul du MAP, comme l'algorithme SMLR (*Single Most Likely Replacement*), proposé par [KM82], qui permet une exploration locale de l'espace des configurations. Le critère à maximiser suivant $\boldsymbol{q}$ vient soit de la loi *a posteriori* jointe $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$, soit de la loi *a posteriori* marginale $\Pr(\boldsymbol{q} \,|\, \boldsymbol{y})$. Notons tout de même qu'il a été montré que l'estimation suivant la loi jointe donnait lieu à davantage de fausses détections [CGI96]. Le principe du SMLR est de partir d'une séquence $\boldsymbol{q}^{(0)}$ initiale, et d'améliorer le critère itérativement en ne modifiant qu'une seule composante de la séquence courante $\boldsymbol{q}^{(t)}$, plus particulièrement celle dont la modification augmente le plus la loi *a posteriori*. L'algorithme SMLR n'a aucune garantie de converger vers le maximum global de la loi *a posteriori* mais seulement vers un maximum local. Dans le cadre non supervisé, des solutions ont été proposées pour l'estimation des hyperparamètres au sens du MAP, basées pour la plupart sur des algorithmes d'optimisation alternée ; la loi *a posteriori* est d'abord maximisée en $\boldsymbol{q}$ avec le SMLR ($\boldsymbol{\theta}$ fixé), puis elle est maximisée suivant $\boldsymbol{\theta}$ à $\boldsymbol{q}$ fixé (voir par exemple [CGI96, KHN15]). Dans tous les cas, la méthode envisagée est toujours sous-optimale à cause de l'emploi du SMLR.

À la différence du calcul du MAP, le calcul de l'estimateur de l'EAP revient par définition à un problème de calcul d'intégrale qui n'est pas résolvable analytiquement. Plus précisément, l'intégration sur les amplitudes $\boldsymbol{x}$ se fait aisément à $\boldsymbol{q}$ fixé. Le calcul de l'EAP revient alors à faire une moyenne pondérée sur toutes les configurations possibles pour $\boldsymbol{q}$, ce qui revient à une exploration combinatoire. La méthode la plus utilisée est alors de générer des échantillons des paramètres d'intérêt suivant la loi *a posteriori* afin de pouvoir calculer leur moyenne empirique. Par exemple, si l'on possède $T$ échantillons $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}) \sim p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$, alors la *moyenne a posteriori* marginale de la séquence de Bernoulli s'estime de la façon suivante :

$$\widehat{\boldsymbol{q}}_{\mathrm{EAP}} \approx \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{q}^{(t)}$$

Même si la loi *a posteriori* ne revêt pas d'expression simple permettant de simuler directement des échantillons, elle se prête bien à l'échantillonneur de Gibbs, un type d'algorithme *Monte-Carlo par Chaînes de Markov* (MCMC), solution proposée par [CCL96]. Le principe est de tirer des échantillons de chaque paramètre suivant les lois *a posteriori* conditionnelles en les autres paramètres ; en effet, dans le cas du modèle BG, la loi *a posteriori* conditionnelle $p(q_j, \boldsymbol{x}_j \,|\, \boldsymbol{y}, \boldsymbol{q}_{-j}, \boldsymbol{x}_{-j}, \boldsymbol{\theta}\,|\,)$ (où $\boldsymbol{u}_{-j}$ correspond au vecteur $\boldsymbol{u}$ sans la $j^{\mathrm{e}}$ composante) est une loi Bernoulli-Gaussienne et se simule aisément. Nous reviendrons sur cet échantillonnage au Chapitre 3. Un tel échantillonneur permet d'obtenir *asymptotiquement* des échantillons tirés suivant une loi cible, ici la loi *a posteriori*. Notons que les échantillons de la loi *a posteriori* nous permettent non seulement de calculer l'estimateur de l'EAP, mais aussi de pouvoir approcher l'estimateur du MAP (voir [KTHD12] par exemple) et enfin de mieux comprendre la loi *a posteriori* ; les échantillons nous permettent par exemple d'estimer leur loi marginale sur la base de leur histogramme, et d'estimer leur variance et leurs corrélations.

### Bilan sur les approches stochastiques

Nous avons vu que le problème d'APL peut se formaliser de façon probabiliste, en introduisant une loi *a priori* favorisant la parcimonie pour les amplitudes $\boldsymbol{x}$. Cette modélisation mène à la définition de deux estimateurs et donc à deux problèmes différents :

[10] On peut alors parler de modèle Bernoulli-Uniforme (BU).

1. Estimateur du MAP : nous avons vu que le calcul du MAP se ramène à un problème d'optimisation très proche (voir identique) à ceux décrits dans le cadre déterministe en § 1.2.2. Le modèle bayésien apporte cependant une interprétation statistique des hyperparamètres.
2. Estimateur de l'EAP : cet estimateur nécessite l'utilisation d'algorithme d'échantillonnage MCMC afin d'obtenir des échantillons de la loi *a posteriori*. Bien que le coût de calcul d'une telle démarche soit important, notamment quand le nombre de paramètres à simuler augmente, l'échantillonnage MCMC permet notamment une estimation des hyperparamètres dans un cadre non supervisé, mais aussi d'obtenir une meilleure compréhension de la loi *a posteriori*.

Par la suite, nous différencierons donc les approches déterministes, qui visent idéalement à minimiser les problèmes $\ell_0$ (1.8) (et se ramènent donc à l'estimateur du MAP), et les approches stochastiques, qui consistent en une étude globale de la loi *a posteriori* des paramètres d'intérêt, grâce à des échantillonnages de type MCMC.

### 1.2.4 Bilan sur les deux approches en APL

Nous avons vu dans cette section que les problèmes inverses non linéaires considérés pouvaient se reformuler comme des problèmes d'*approximation parcimonieuse linéaire* (APL), en discrétisant le paramètre de localisation $\nu$ sur une grille $\mathcal{G}$ pour construire un dictionnaire $\mathbf{H}$ d'atomes $\mathbf{h}_j = \mathbf{h}(\ddot{\nu}_j)$, $\ddot{\nu}_j \in \mathcal{G}$, qui permet d'approcher le signal $\boldsymbol{y}$ par $\mathbf{H}\boldsymbol{x}$, avec $\boldsymbol{x}$ parcimonieux.

Des méthodes classiques existent pour résoudre les problèmes d'APL, et se classent en deux catégories : les approches déterministes et les approches stochastiques. Dans ces approches, nous avons mis en exergue deux méthodes permettant une résolution du problème d'APL que nous qualifions d'« optimale », dans le sens où elles permettent une prise en compte exacte de la parcimonie (dans les deux cas grâce à l'introduction de variables binaires conditionnant la nullité des amplitudes), et où la résolution du problème associé se fait de façon optimale. Nous précisons ces méthodes dans chaque approche :

- Approche déterministe : la reformulation des problèmes $\ell_0$ en MIP, grâce à l'introduction de variables binaires (équivalentes aux variables de Bernoulli dans le cadre stochastique) est la seule méthode, à notre connaissance, permettant une résolution « optimale » respectant les critères définis ci-dessus. Les relaxations de la pseudo-norme $\ell_0$ ne prennent pas en compte de façon exacte la parcimonie, tandis que les algorithmes de type gloutons (et heuristiques de type SMLR) ne permettent pas une résolution optimale du problème d'optimisation mis en jeu. Dans les deux cas, ces approches sont mises en difficulté par la forte corrélation des atomes du dictionnaire.
- Approche stochastique : le modèle Bernoulli-Gaussien permet une prise en compte exacte de la parcimonie, grâce au lien probabiliste reliant les variables (binaires) de Bernoulli aux amplitudes $\boldsymbol{x}$. L'étude de la loi *a posteriori* des paramètres est faite de façon optimale à partir d'échantillons de cette loi, simulés grâce à un échantillonnage de type MCMC.

L'utilisation de telles méthodes « optimales » est nécessaire dans le cadre des problèmes inverses considérés, afin de garantir une estimation correcte du support de la solution pour répondre au problème de détection. Le prix à payer pour l'utilisation des variables binaires est cependant élevé : les deux méthodes sont particulièrement lourdes en terme de coût de calcul, d'autant plus que la taille des données $N$, l'ordre du modèle $L$ ou la taille de la grille $J$ augmente.

Enfin, notons que même pour ces méthodes optimales de résolution du problème d'APL, la perte de *précision* induite par la discrétisation du paramètre de localisation pour la reformulation du problème inverse non linéaire en APL reste problématique. En effet, cette reformulation

en APL est une *approximation* du modèle non linéaire : le paramètre de localisation est une variable continue, et il est fort peu probable qu'il appartienne exactement à la grille de discrétisation. Outre l'erreur de précision que cela induit, le support de la solution optimale du problème d'APL peut être différent de celui de la solution du problème initial et on échoue alors à résoudre le problème de détection. L'augmentation de la taille $J$ de la grille n'est pas une solution ; en conservant une estimation discrète, elle augmente davantage la corrélation du dictionnaire compliquant d'autant plus la résolution par les méthodes sous-optimales, mais apportant aussi des instabilités numériques associées à un coût de calcul d'autant plus élevé pour les méthodes optimales utilisant des variables binaires. Dans tous les cas, les paramètres de forme ne peuvent pas être pris en compte dans une reformulation en APL.

La section suivante expose des réponses pouvant être envisagées pour une *estimation continue des paramètres non linéaires* pour une reformulation équivalente des problèmes inverses non linéaires avec *a priori* de parcimonie considérés.

## 1.3 Approximations parcimonieuses non linéaires sur dictionnaires infinis et continus

Revenons maintenant aux problèmes inverses parcimonieux initiaux (1.1), où l'on cherche à estimer les paramètres non linéaires $\{\boldsymbol{\nu}_\ell\}_\ell$ de façon continue, en plus du faible ordre du modèle $L$ et des amplitudes $\{\boldsymbol{x}_\ell\}_\ell$ des atomes continus. On peut trouver dans la littérature deux approches différentes permettant notamment l'estimation continue des paramètres non linéaires. La première vise à résoudre directement ce problème en exploitant la *norme atomique* sur un *dictionnaire infini*. La seconde se place dans le cadre des approximations parcimonieuses, mais utilise un *dictionnaire continu* de dimension finie dont les atomes dépendent continument des paramètres non linéaires, plutôt que le dictionnaire discret des APL.

Nous distinguerons dans ce manuscrit les *dictionnaires infinis* et les *dictionnaires continus*. Dans la littérature, le terme de *dictionnaire continu* est employé dans les deux approches.

### 1.3.1 Norme atomique et relaxation convexe sur dictionnaire infini

La *norme atomique* a été introduite par [CRPW12] et est notamment utilisée dans le cadre de l'acquisition compressée (*Compress Sensing*) pour l'estimation de spectres de raies [TBSR13]. Le principe est de s'exempter de la discrétisation du paramètre continu en considérant l'ensemble suivant :

$$\mathcal{A} = \left\{ \boldsymbol{h}(\boldsymbol{\nu}) \in \mathbb{F}^N,\ \boldsymbol{\nu} \in \Lambda \subset \mathbb{R}^p \right\}$$

Nous appelons cet ensemble $\mathcal{A}$ le *dictionnaire infini*, dans le sens où il est de dimension infinie et où il s'agit du cas limite du dictionnaire discret $\mathbf{H}$ quand le nombre d'atomes $J$ tends vers l'infini. Notons que les paramètres non linéaires $\boldsymbol{\nu}$ peuvent ici contenir à la fois le paramètre de localisation mais aussi les paramètres de forme ; pour rendre les explications plus simples, nous considérons ici que les atomes $\boldsymbol{h}(\cdot)$ sont composés d'une seule colonne ($m = 1$). La *norme atomique* d'un vecteur $\boldsymbol{z} \in \mathbb{F}^N$ est alors définie pour le dictionnaire infini $\mathcal{A}$ :

$$\|\boldsymbol{z}\|_{\mathcal{A}} = \inf \left\{ \sum_j |x_j|,\ \boldsymbol{z} = \sum_j x_j\, \boldsymbol{h}(\boldsymbol{\nu}_j) \text{ avec } \boldsymbol{h}(\boldsymbol{\nu}_j) \in \mathcal{A} \right\}$$

Cette norme correspond à la généralisation au dictionnaire infini de la norme $\ell_1$ sur les amplitudes du dictionnaire discret : $\|\boldsymbol{z}\|_{\mathcal{A}}$ est la plus petite norme $\ell_1$ des amplitudes $\boldsymbol{x}$ permettant d'écrire

le signal $\boldsymbol{z}$ comme une combinaison linéaire d'un nombre arbitraire d'éléments de $\mathcal{A}$. On peut d'ailleurs considérer que la norme atomique est la relaxation convexe d'une « pseudo-norme 0-atomique » que nous notons par abus de langage $\|\cdot\|_{\mathcal{A},0}$ et définie comme :

$$\|\boldsymbol{z}\|_{\mathcal{A},0} = \inf \left\{ \|\boldsymbol{x}\|_0,\ \boldsymbol{z} = \sum_j x_j\ \boldsymbol{h}(\boldsymbol{\nu}_j) \text{ avec } \boldsymbol{h}(\boldsymbol{\nu}_j) \in \mathcal{A} \right\}$$

Ce terme permettrait une prise en compte exacte de la parcimonie, mais est bien évidemment difficile à prendre en compte dans un problème d'optimisation, à l'image de la pénalisation $\ell_0$ pour les APL. On utilise donc en pratique la norme atomique.

Dans le cas de signaux $\boldsymbol{y}$ bruités[11], une reformulation du problème (1.1) similaire à celle de la pénalisation $\ell_1$ est alors envisagée [BTR13] et nommée *Atomic soft thresholding* (AST) par analogie :

$$\min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{z}\|^2 + \lambda \|\boldsymbol{z}\|_{\mathcal{A}} \tag{1.18}$$

La norme atomique est une régularisation parcimonieuse [CRPW12], au même titre que la norme $\ell_1$ des amplitudes, puisque que ce problème revient à trouver $\boldsymbol{z}$ approchant suffisamment bien le signal $\boldsymbol{y}$ et s'exprimant comme une combinaison linéaire d'éléments de $\mathcal{A}$ dont les amplitudes sont de norme $\ell_1$ minimale. La résolution du problème d'AST (1.18) n'est pas une tâche facile ; cependant, dans des applications en acquisition compressée comme l'estimation de spectres de raies [BTR13] ou encore celle de décalages temporels en déconvolution [PSC17], le problème d'AST peut être reformulé en *Semi Definite Program* (SDP) permettant ainsi sa résolution à l'aide de routine d'optimisation[12].

La résolution du problème d'AST a gagné en popularité grâce à de nombreux théorèmes prouvant son optimalité, notamment pour des problèmes d'estimation de spectres de raies en acquisition compressée, sous des hypothèses d'absence (ou de faible) bruit et de fréquences suffisamment séparées [TBSR13, BTR13]. Cependant, elle présente quelques inconvénients dans le cadre des problèmes inverses non linéaires considérés dans nos travaux. D'abord, les garanties de bonnes estimations ne sont plus toujours valables quand l'amplitude du bruit augmente. Ensuite, la reformulation en SDP du problème d'AST dépendra de l'application en jeu, et donc des fonctions $\boldsymbol{h}(\cdot)$ du dictionnaire infini. Par exemple, dans le cadre d'analyse spectrale en échantillonnage irrégulier, [HBLC16] explique que l'échantillonnage irrégulier typique des problèmes en astrophysique rend impossible la reformulation en SDP[13]. Enfin, [TBR13] démontre que la solution du problème d'AST peut être approchée par celle de la formulation de l'APL avec pénalisation $\ell_1$ quand la grille de discrétisation devient de plus en plus fine. Bien que le problème de la corrélation du dictionnaire semble être dépassé, le problème d'AST est une relaxation convexe du problème « optimal » qui utiliserait la pénalisation « atomique-0 » ; il souffrira donc des mêmes problèmes que les méthodes sous-optimales dans le cas de l'APL, notamment pour nos applications.

Nous nous intéressons donc à un deuxième axe de recherche pour la résolution du problème inverse non linéaire.

[11] Le norme atomique a d'abord été envisagée dans un cadre non bruité où le signal $\boldsymbol{y}$ peut exactement s'écrire comme une combinaison linéaire du dictionnaire infini [TBSR13].

[12] Les paramètres non linéaires $\boldsymbol{\nu}$ peuvent être retrouvés à partir du problème dual de (1.18) [BTR13].

[13] En effet, la reforfumation en SDP nécessite de pouvoir construire un vecteur $\tilde{\boldsymbol{t}}$ d'instants d'échantillonnage régulièrement espacés à partir des temps $\boldsymbol{t}$ initiaux. Cela est possible dans le cas de l'acquisition compressée, où l'échantillonnage est régulier avec des données manquantes, mais c'est impossible dans le cas de l'échantillonnage spécifique aux problèmes d'astrophysique ; même si on arrive à trouver un dénominateur commun permettant de construire le vecteur $\tilde{\boldsymbol{t}}$, ce dernier serait de dimension trop élevée pour une résolution par SDP.

### 1.3.2 Approximation parcimonieuse non linéaire sur dictionnaire continu

La deuxième approche reste dans le cadre des approximations parcimonieuses à dictionnaire de dimension finie, dans le sens où le problème inverse non linéaire (1.1) est abordé en cherchant à approcher le signal $\boldsymbol{y}$ comme une combinaison linéaire parcimonieuse d'éléments d'un dictionnaire, à la différence près que les éléments du dictionnaire ne sont plus constants et dépendent des paramètres non linéaires.

Plaçons nous d'abord dans le cas où les paramètres non linéaires sont composés uniquement d'un paramètre de localisation, comme pour les APL (*i.e.* $\boldsymbol{\upsilon} = \varnothing$). Pour obtenir cette reformulation en APL, le paramètre de localisation $\nu$ est discrétisé en $J$ valeurs discrètes $\ddot{\nu}_j$ du domaine de localisation $\mathcal{I}_\nu = [\nu_{\min}; \nu_{\max}]$. Au contraire, l'intervalle $\mathcal{I}_\nu$ est ici découpé en $J$ sous-intervalles disjoints $\mathcal{I}_j$ de longueurs égales[14] $\Delta$ :

$$\mathcal{I}_j = [\nu_{\min} + (j-1)\Delta; \nu_{\min} + j\Delta], \text{ avec } \Delta = \frac{\nu_{\max} - \nu_{\min}}{J} \tag{1.19}$$

Une telle partition du domaine des localisations, schématisée en Figure 1.2, permet donc de considérer $J$ paramètres $\nu_j$ évoluant continument dans leur intervalle respectif $\mathcal{I}_j$, au lieu des $J$ paramètres $\ddot{\nu}_j$ discrets appartenant à la grille $\mathcal{G}$. Pour simplifier, on parlera encore de grille pour parler de cette partition de $\mathcal{I}_\nu$.

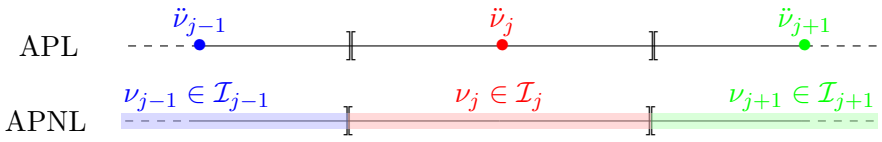



FIG. 1.2 – Représentation de la discrétisation du domaine des localisations pour la reformulation en APL, et de sa partition pour la reformulation en APNL.

Nous pouvons maintenant associer des paramètres de formes $\boldsymbol{\upsilon}_j \in \mathcal{I}_{\boldsymbol{\upsilon}}$ à chaque paramètre de localisation $\nu_j$ pour former les paramètres non linéaires $\boldsymbol{\nu}_j = [\nu_j, \boldsymbol{\upsilon}_j]$ ; ils permettent de préciser la forme de l'atome $\mathbf{h}(\cdot)$ localisé dans l'intervalle $\mathcal{I}_j$. En regroupant l'ensemble des $J$ paramètres non linéaires $\boldsymbol{\nu}_j$ dans le vecteur $\boldsymbol{\nu} = [\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_J]$, on introduit la notation suivante :

$$\boldsymbol{\nu} \in \mathfrak{I} \Leftrightarrow \forall j \in [\![1; J]\!],\ \boldsymbol{\nu}_j \in \mathfrak{I}_j : \nu_j \in \mathcal{I}_j \text{ et } \boldsymbol{\upsilon}_j \in \mathcal{I}_{\boldsymbol{\upsilon}}$$

Le dictionnaire discret $\mathbf{H}$ est alors remplacé par un dictionnaire de même dimension que nous notons $\mathbf{H}(\boldsymbol{\nu})$, et que nous appelons *dictionnaire continu* dans le sens où chacun de ses $J$ éléments $\mathbf{h}(\boldsymbol{\nu}_j)$, que nous appelons *atome continu*, dépend continument des paramètres non linéaires :

$$\mathbf{H}(\boldsymbol{\nu}) = [\mathbf{h}(\boldsymbol{\nu}_1), \ldots, \mathbf{h}(\boldsymbol{\nu}_J)] \in \mathbb{F}^{N \times mJ} \text{ défini pour } \boldsymbol{\nu} \in \mathfrak{I}$$

Le problème inverse non linéaire (1.1) peut alors se reformuler de la façon suivante :

$$\text{Estimer } \boldsymbol{x} \textit{ parcimonieux} \text{ et } \boldsymbol{\nu} \in \mathfrak{I} \text{ tels que } \boldsymbol{y} = \mathbf{H}(\boldsymbol{\nu})\boldsymbol{x} + \boldsymbol{\epsilon} = \sum_{j=1}^{J} \mathbf{h}(\boldsymbol{\nu}_j)\boldsymbol{x}_j + \boldsymbol{\epsilon}. \tag{1.20}$$

Le problème (1.20) sera appelé par la suite problème d'*approximation parcimonieuse non linéaire* (APNL), par analogie avec les problèmes d'APL. Le problème d'APNL permet de formaliser *exactement* le problème inverse non linéaire (1.1) dans le sens où il autorise l'estimation

[14] Ce découpage régulier de $\mathcal{I}_\nu$ n'est pas obligatoire en théorie : on peut au contraire considérer des intervalles de tailles différentes $\Delta_j$. Cependant, on simplifie ici en supposant un découpage régulier, que l'on adoptera dans l'ensemble de nos travaux.

continue des paramètres non linéaires, contrairement au problème d'APL. Ceci se fait au prix de la réintroduction de la non-linéarité dans le modèle puisque chaque atome continu dépend de façon non linéaire des $\boldsymbol{\nu}_j$. Cependant, la partition en $J$ intervalles du domaine des localisations $\mathcal{I}_\nu$ apporte un double avantage. D'une part, elle permet l'estimation directe de l'ordre $L$ du modèle comme la dimension du support $\Omega = \mathsf{Supp}\,(\boldsymbol{x})$, de façon similaire aux problèmes d'APL (voir (1.6)). D'autre part, elle restreint l'espace de recherche des paramètres de localisations $\nu_j$ dans des intervalles de taille $\Delta$, $J$ fois plus petits que la taille de $\mathcal{I}_\nu$. Pour les problèmes inverses considérés, la détection correcte du support $\Omega$ et des paramètres non linéaires $\{\boldsymbol{\nu}_j\}_{j\in\Omega}$ correspondant sera donc primordiale. Notons par ailleurs que le nombre $J$ d'atomes du dictionnaire continu $\mathbf{H}(\boldsymbol{\nu})$ pour les APNL peut être bien plus faible que le nombre d'atome du dictionnaire discret $\mathbf{H}$ pour les APL, puisque que ce dernier nécessite une fine discrétisation du paramètre de localisation pour réduire les erreurs de précision. Pour les APNL, le choix de la partition du domaine des localisations se basera par exemple sur des critères de résolution pour l'analyse spectrale (au sens de la distance minimale entre deux fréquences pour pouvoir être distinguées), comme nous le verrons en § 1.5.1. Trois types d'approches ont été proposées pour bénéficier de ces avantages et essayer de résoudre ce problème d'APNL.

**Interpolation de dictionnaire : approximation linéaire des atomes continus.** Dans le cadre de signaux à valeurs réelles et amplitudes positives, d'atomes composé d'une seule colonne et de l'absence de paramètres de forme, [ETS11] a proposé de bénéficier du fait que les paramètres de localisation $\nu_j$ appartiennent à un petit intervalle pour effectuer une approximation linéaire des atomes continus, permettant de bénéficier d'une « interpolation de dictionnaire ». Le problème d'APNL (1.20) est alors approché par le suivant :

$$\text{Estimer } \widetilde{\boldsymbol{x}} \in \mathcal{C} \textit{ parcimonieux} \text{ tels que } \boldsymbol{y} = \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}} + \widetilde{\boldsymbol{\epsilon}} = \sum_{j=1}^{J} \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j + \widetilde{\boldsymbol{\epsilon}}. \tag{1.21}$$

On retrouve un problème d'APL classique (1.4), à la différence de la présence des contraintes $\widetilde{\boldsymbol{x}} \in \mathcal{C}$ : il sera donc nommé problème d'*approximation parcimonieuse linéaire sous contraintes* (APLsc). Ces contraintes sont issues de la nécessité de pouvoir identifier l'amplitude $x_j$ et le paramètre de localisation $\nu_j$ à partir de la nouvelle amplitude $\widetilde{\boldsymbol{x}}_j$. Notons que $\widetilde{\boldsymbol{x}}$ parcimonieux s'entend toujours au sens de la parcimonie structurée : en effet, la linéarisation de la colonne $\boldsymbol{h}(\nu_j)$ donne un nouvel atome $\widetilde{\mathbf{h}}_j$ constitué de $k > 1$ colonnes et appelé « atome interpolé » ; $\widetilde{\mathbf{H}}$ est appelé le « dictionnaire interpolé ». Le principe d'une telle reformulation en APLsc sera nommé « interpolation de dictionnaire » dans ce manuscrit. Le problème d'APLsc est alors résolu à l'aide des méthodes classiques d'APL, adaptées à la présence des contraintes ; [ETS11] a proposé d'utiliser la relaxation $\ell_1$ des amplitudes tandis que [KYHP14, FDJ15] adaptent des algorithmes gloutons. Par ailleurs, différents types d'interpolation de dictionnaire ont été proposés, des plus basiques comme l'approximation de Taylor au premier ordre à des plus complexes comme l'interpolation polaire proposée par [ETS11]. Ils se différencient majoritairement par l'erreur d'approximation qu'elles induisent. Notons que [FDJ15] a généralisé l'interpolation polaire et la résolution avec pénalisation $\ell_1$ aux cas de signaux à valeurs complexes et amplitudes quelconques.

Les approches ci-dessus sont sous-optimales pour la résolution de ce problème d'APLsc, puisqu'elles s'appuient sur de la relaxation $\ell_1$ ou sur des algorithmes gloutons. Leur performance peut donc être entravée à la fois par l'erreur d'approximation induite par l'approximation linéaire des atomes continus et par les méthodes sous-optimales utilisées.

**Modèle Bernoulli-Gaussien étendu.** Nous définissons le modèle Bernoulli-Gaussien étendu comme l'extension au problème d'APNL du modèle BG classique vu en § 1.2.3, en prenant en compte les paramètres non linéaires $\boldsymbol{\nu}$. Dans le cadre bayésien, on étudiera donc la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$.

Un premier modèle Bernoulli Gaussien étendu aux non linéarités a d'abord été proposé par [Dub96]. Ce dernier proposait de faire une approximation linéaire de son atome continu pour obtenir un problème similaire à celui d'APLsc (1.21) - bien qu'antérieur à ces travaux. Il proposait alors l'estimation au sens du MAP des paramètres, en maximisant $p(\boldsymbol{q}, \widetilde{\boldsymbol{x}} \,|\, \boldsymbol{y})$ avec l'algorithme SMLR mais sans prendre en compte les contraintes reliant les nouvelles amplitudes $\widetilde{\boldsymbol{x}}$ aux amplitudes et fréquences initiales.

Une autre extension a ensuite été proposée par [BC06] pour l'analyse spectrale de signaux irrégulièrement échantillonnés. La loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ est échantillonnée suivant un algorithme dit de « Gibbs-hybride » : l'échantillonnage est similaire au cas des APL pour les variables $(q_j, \boldsymbol{x}_j)$ mais les paramètres $\nu_j$ (ici les fréquences des raies recherchées) sont échantillonnés à l'aide d'une étape de *Metropolis Hastings*. La probabilité $\lambda = \Pr(q_j = 1)$ représente alors la probabilité d'avoir une fréquence dans l'intervalle $\mathcal{I}_j$. Les auteurs ont cependant noté des ralentissements dans l'échantillonnage par rapport à celui du BG classique et quelques problèmes théoriques, qui mériteraient d'être étudiés.

Plus récemment, des méthodes s'appuyant sur des algorithmes d'optimisation locale ont été proposés pour l'estimation des paramètres du modèle Bernoulli Gaussien étendu[15] pour des problèmes d'estimation de spectres de raies en acquisition compressée [BHF17, HFR18]. Il est proposé dans [HFR18] d'effectuer une optimisation de type *Block Coordinate Descent* exploitant les propriétés du SMLR. [BHF17] se place quant à lui dans le cadre « variationnel » consistant à approcher la loi *a posteriori*, pour la rendre plus simple à maximiser par des algorithmes d'optimisation locale de type *Expectation-Maximisation* mêlés à des algorithmes gloutons.

Les travaux de Kail [KHN15, KHN15] utilisent un autre type de modèle Bernoulli Gaussien étendu : en restant dans le cadre du problème d'APL, les paramètres de localisation sont discrétisés très finement sur la grille, permettant une grande réduction de l'erreur de précision. Le modèle BG classique est alors adapté pour prendre en compte une *contrainte de distance minimum* entre deux $(q_{j_1}, q_{j_2})$ égaux à un. Cela permet notamment de diminuer le coût de calcul lié à une grille de grande taille $J$ et aussi d'éviter des instabilités numériques. La résolution du problème d'APL sous ce modèle BG adapté est alors effectuée, soit au sens du MAP en adaptant l'algorithme SMLR, soit via un échantillonnage stochastique de la loi *a posteriori*.

**APL et optimisation locale.** Une autre solution envisagée pour résoudre le problème d'APNL (1.20) est d'associer aux techniques classiques de résolutions d'APL des étapes d'optimisation locale en les paramètres non linéaires $\boldsymbol{\nu}_j$ pour améliorer localement le critère d'attache aux données. C'est par exemple le principe de l'approche CLEANEST [Fos95] et du logiciel `Period04` [LB04], utilisés dans la communauté astrophysique, et qui consistent en une optimisation locale à partir du support trouvé, à chaque itération d'un algorithme glouton. Dans [Bou14], des étapes de descente de gradient ont été introduites au sein d'algorithmes gloutons pour la détection d'exoplanètes.

[15] Le modèle BGE est ici légèrement différent ; les paramètres de localisation $\nu_j$ associés aux variables de Bernoulli $q_j$ appartiennent tous au même intervalle, alors que pour le problème d'APNL, nous les contraignons à appartenir à des intervalles distincts, provenant du découpage du domaine des localisations.

### 1.3.3 Bilan sur les approximations parcimonieuses non linéaires

Nous avons vu dans cette section deux façons d'appréhender l'estimation continue des paramètres non linéaires pour la résolution des problème inverses considérés dans nos travaux (1.1). La première se base sur les *dictionnaires infinis* $\mathcal{A}$, des espaces de dimension infinie qui regroupent l'ensemble des atomes continus $\mathbf{h}(\boldsymbol{\nu})$ pour $\boldsymbol{\nu} \in \Lambda \subset \mathbb{R}^p$. Elle permet de définir le problème d'*Atomic Soft Thresholding* (1.18), équivalent sur dictionnaire infini de la pénalisation $\ell_1$ sur dictionnaire discret. La seconde se place dans le cadre des approximations parcimonieuses, où au lieu de fixer les paramètres de localisation sur une grille de $J$ paramètres discret $\ddot{\nu}_j$, on les relâche dans $J$ intervalles $\mathcal{I}_j$ disjoints. Cela permet de définir le problème d'*approximation parcimonieuse non linéaire* (APNL, (1.20)) sur le *dictionnaire continu* $\mathbf{H}(\boldsymbol{\nu})$ de dimension $(N \times mJ)$, dont les atomes dits *atomes continus* dépendent de façon non linéaire des $\boldsymbol{\nu}_j \in \mathfrak{I}_j$. Des reformulations du problème d'APNL et des adaptations des méthodes classiques de résolution d'APL ont été proposées pour résoudre ce problème de façon sous-optimale.

Les travaux présentés dans cette section nous ont amenés à nous orienter vers deux axes de recherche présentés en section suivante.

## 1.4 Conclusion sur les objectifs et les pistes suivies

Nous revenons dans cette section sur l'objectif principal de cette thèse et nous justifions les pistes de recherches qui ont été poursuivies. Le plan général du manuscrit est ensuite décrit, accompagné des principales contributions scientifiques et des publications associées.

### 1.4.1 Pistes choisies pour répondre aux objectifs

Comme annoncé dès le début de ce chapitre, cette thèse vise à résoudre les problèmes inverses non linéaires avec *a priori* de parcimonie présentés en § 1.1 dans un cadre général. Pour répondre à la difficulté de tels problèmes, les résolutions envisagées, qualifiées d'« optimales », doivent répondre aux critères suivants :

1. la parcimonie doit être prise en compte de façon exacte ;
2. les méthodes de résolution employées visent à garantir de trouver la solution optimale des problèmes posés.

Nos recherches ont aussi pour objectif de bénéficier de ces travaux pour l'application à des problèmes d'actualité dans la communauté astrophysique : l'analyse spectrale de signaux irrégulièrement échantillonnés et plus particulièrement la détection d'exoplanètes par la méthode des vitesses radiales.

*Remarque : À l'opposé de nombreux travaux, nous favorisons donc la recherche de solutions optimales au prix d'un coût de calcul plus élevé. Néanmoins, ce n'est pas un problème pour les applications d'astrophysique qui nous intéressent ; l'acquisition des données se fait généralement sur plusieurs jours, et leur traitement ne nécessite pas une mise en œuvre en temps réel. Bien sûr, nous chercherons à améliorer l'efficacité de nos méthodes, mais l'optimalité des solutions reste la priorité.*

Au regard des méthodes existantes pour la résolution des problèmes d'APL vues en § 1.2, et des récentes propositions pour l'estimation continue des paramètres non linéaires résumées en § 1.3, nous avons décidé de focaliser nos recherches sur la reformulation en problème d'*approximation parcimonieuse non linéaire* (APNL) utilisant un *dictionnaire continu*, plutôt que l'exploita-

tion d'un *dictionnaire infini* pour la résolution de l'*Atomic Soft Thresholding* (AST). En effet, comme nous l'avons vu précédemment, le problème d'AST est l'équivalent pour le dictionnaire infini de la relaxation de la pseudo-norme $\ell_0$ par la norme $\ell_1$ sur le dictionnaire discret. Cette formulation peut être considérée comme « sous-optimale » car elle ne prend pas en compte de façon exacte de la parcimonie. Elle souffrira des mêmes problèmes de sous-estimation des amplitudes et des fausses détections, préjudiciables pour les problèmes inverses considérés. À cela s'ajoute la difficulté de reformuler le problème d'AST en *Semi-Definite Programming* (SDP) pour certaines applications, notamment pour celles qui nous intéressent : le cas de l'échantillonnage irrégulier complique la mise en œuvre des algorithmes et l'application à la détection d'exoplanètes semble impossible. Au contraire, les problèmes d'APNL nous semblent plus prometteurs pour ces applications. Nous avons donc décidé d'étudier les possibilités d'adapter aux problèmes d'APNL les méthodes de résolution d'APL répondant aux critères d'optimalité dans les cadres déterministe et stochastique, formant deux axes de recherche que nous décrivons dans les paragraphes suivant.

Le premier axe de recherche est donc l'estimation, par une approche déterministe, des amplitudes $\boldsymbol{x}$ et des paramètres non linéaires $\boldsymbol{\nu}$ pour résoudre le problème d'APNL. Dans le cadre des APL, seule la reformulation en MIP du problème $\ell_0$ permet une résolution « optimale » du problème, tant sur la prise en compte de la parcimonie que sur la méthode de résolution. Dans le cadre des APNL, la non-linéarité induite par les paramètres $\boldsymbol{\nu}_j$ rend cependant les problèmes d'optimisation non convexes et nécessairement sensibles aux minima locaux. Afin de garder un problème linéaire, on utilisera la reformulation du problème d'APNL en problème d'APL sous contraintes par le principe d'*interpolation de dictionnaire* proposé initialement par [ETS11]. Le premier verrou sera donc d'adapter la reformulation en MIP du problème $\ell_0$ proposé par [BNCM16] à la présence des contraintes afin d'obtenir une résolution optimale, puisque les méthodes de résolution actuellement proposées pour ces problèmes d'APLsc sont des méthodes sous-optimales se basant sur les algorithmes gloutons ou la relaxation $\ell_1$. Par ailleurs, le problème d'APLsc est bien une approximation du problème d'APNL initial, puisqu'il nécessite la linéarisation des atomes $\mathbf{h}(\cdot)$. Un deuxième verrou sera donc de minimiser l'erreur d'approximation en étudiant les différentes propositions d'approximation. Enfin, des extensions au cas de la présence de paramètres de forme, ou d'atomes composés de plusieurs colonnes seront envisagées.

Le deuxième axe s'appuie sur une approche stochastique et consiste en la prise en compte des paramètres non linéaires au sein d'un modèle Bernoulli-Gaussien étendu, associé à un échantillonnage stochastique de la loi *a posteriori* pour le calcul de l'estimateur de l'EAP. Le modèle BGE proposé par [BC06] nous semble le plus approprié aux objectifs de cette thèse. L'approche envisagée par [KTHD12], d'un modèle BG classique sur grille très fine associé à une contrainte de distance minimale et à un échantillonnage stochastique, ne permet pas, quant à elle, une véritable estimation continue des paramètres non linéaires, et ne peut pas se généraliser à la présence de paramètres de formes. Nos travaux viseront donc à résoudre les problèmes d'échantillonnage relevés par [BC06], notamment au niveau de la convergence de l'échantillonneur, dont les échantillons peuvent rester bloqués dans des modes locaux, mais aussi des incohérences sur le schéma d'échantillonnage.

Notons qu'une étude des algorithmes gloutons avec optimisation locale pour la détection d'exoplanètes a en parallèle été menée au cours de cette thèse. Ce type de méthodes, considérées comme sous-optimales, ne rentre pas dans le cadre choisi pour ce manuscrit. Des premières pistes de recherche peuvent cependant être trouvées dans le rapport [Bou14], et nous avons choisi d'illustrer rapidement ces travaux dans le Chapitre 4 sur l'application à la détection des exoplanètes.

Nous détaillons maintenant le plan de la suite de ce manuscrit, en résumant les contributions apportées dans chaque axe de recherche et en y associant les publications scientifiques relatives.

### 1.4.2 Plan du manuscrit et contributions

Ce manuscrit sera composé, après ce chapitre introductif, de trois chapitres. Les Chapitres 2 et 3 exposent les résultats des recherches menées dans les deux axes de recherche décrits précédemment pour la résolution des problèmes d'APNL. Le Chapitre 4 s'attèle à une application concrète de ces travaux : la détection d'exoplanètes par la méthode des vitesses radiales. Nous décrivons plus en détails le contenu de ces chapitres dans les paragraphes suivants.

Le Chapitre 2, intitulé « Interpolation de dictionnaire et reformulation MIP », s'attaque à la reformulation en problème d'APL sous contraintes du problème d'APNL grâce au principe de l'*interpolation de dictionnaire*. Nous commençons par étudier les travaux existant, tant sur le plan des approximations permettant d'obtenir le dictionnaire interpolé, que sur les méthodes de résolution du problème d'APL sous contraintes. Cette étude nous a révélé que le principe d'interpolation de dictionnaire n'avait pas de cadre « clair » ; nous avons donc commencé par le clarifier comme un problème d'optimisation *convexe* sous contraintes. Nos deux contributions majeures consistent en (i) une étude plus aboutie des différentes approximations de la littérature ayant permis à des corrections pour l'interpolation polaire proposée par [ETS11] et étendue par [FDJ15] et à une revalorisation de l'approximation par K-SVD proposée par [KYHP14] grâce à une étude théorique plus poussée ; (ii) la résolution du problème d'APL sous contraintes en prenant en compte de façon exacte la parcimonie, grâce à l'introduction des variables binaires permettant une reformulation en *Mixed Integer Program*. Cette dernière contribution a fait l'objet de l'article de conférence [BC18], comparant la résolution d'APL sous contraintes avec pénalisation $\ell_1$ et avec contrainte $\ell_0$, pour des problèmes de déconvolution impulsionnelle.

Le Chapitre 3, intitulé « Modèle Bernoulli-Gaussien étendu et échantillonnage stochastique », rapporte nos travaux sur l'échantillonnage stochastique de la loi *a posteriori* du modèle Bernoulli-Gaussien étendu proposé par [BC06]. Après une étude rapide du modèle Bernoulli-Gaussien classique dans un cadre d'échantillonnage stochastique, nous revenons d'abord sur des travaux antérieurs à cette thèse permettant de valider en pratique la démarche d'échantillonnage hybride proposée par [BC06]. Les contributions de nos travaux se trouvent à la fois dans le cadre supervisé et non supervisé. Dans le cadre supervisé, nous proposons un échantillonnage de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y})$ marginalisée suivant les amplitudes $\boldsymbol{x}$, accélérant la convergence de l'échantillonnage et assurant un meilleur « mélange » des échantillons. Dans le cadre non supervisé, l'échantillonnage direct des hyperparamètres de variances suivant leur loi *a posteriori* conditionnelle marginalisée en $\boldsymbol{x}$ n'est plus possible, mais nous montrons qu'un échantillonneur de type *Partially Collapsed Gibbs Sampler* (PCGS) [vDP08] de la loi jointe $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ permet d'exploiter l'échantillonnage suivant la loi marginale des paramètres d'intérêt : la séquence de Bernoulli $\boldsymbol{q}$ et les paramètres non linéaires $\boldsymbol{\nu}$. Ces contributions ont été publiées dans l'article de conférence [BCBB16], où l'échantillonnage du modèle BGE est appliqué à un problème d'analyse spectrale en échantillonnage irrégulier.

Dans un souci d'uniformisation, les méthodes développées dans ces deux chapitres sont testées sur les mêmes données : ce sont des signaux simulés pour des problèmes de déconvolution impulsionnelle (données de [BC18]) et d'analyse spectrale en échantillonnage irrégulier (données de [BCBB16]). Ces données sont décrites dans la section suivante, à laquelle on se référera au cours de ces deux chapitres. Notons que nous avons publié, au début de nos recherches sur

l'interpolation de dictionnaire, une première comparaison des deux approches dans l'article de conférence [BC17], sur les données d'analyse spectrale en échantillonnage irrégulier. Plus précisément, les résultats obtenus dans [BCBB16] sont comparés à ceux obtenus en appliquant directement la méthode de relaxation $\ell_1$ pour la résolution du problème d'APL sous contraintes proposées par [FDJ15]. Le Chapitre 2 complète donc cette étude préliminaire, en exploitant nos contributions pour ce problème d'analyse spectrale. Cependant, malgré les améliorations que nous avons apportées au concept de l'interpolation de dictionnaire, le modèle BGE semble plus adapté au problème de la détection des exoplanètes.

Le Chapitre 4, intitulé « Application à la détection d'exoplanètes », exploite nos travaux sur l'estimation des paramètres du modèle BGE pour un problème d'actualité en astrophysique : la détection d'exoplanètes par la méthode des vitesses radiales, qui est un cas particulier d'analyse spectrale en échantillonnage irrégulier. Une étude de ce problème nous permettra de mettre en exergue les principaux enjeux visés par les astrophysiciens : l'estimation du nombre d'exoplanètes et la prise en compte de l'activité stellaire. Nous pourrons alors montrer que notre modèle BGE et l'algorithme d'échantillonnage associé s'adaptent particulièrement bien à ce problème de détection et sont une excellente alternative aux méthodes actuellement utilisées par les astrophysiciens pour répondre à ces enjeux. Nous testerons d'abord notre méthode sur des données simulées sans activité stellaire afin de montrer son efficacité vis-à-vis d'autres méthodes. Elle sera ensuite testée sur des données simulées avec activité stellaire issues de l'article [NFB+18] et enfin sur des données réelles.

Avant d'entrer dans le vif du sujet, nous présentons dans la section suivante les données tests utilisées dans les Chapitres 2 et 3.

## 1.5 Données tests

Nous décrivons dans cette section les signaux utilisés pour tester les méthodes développées aux Chapitres 2 et 3. Nous regroupons aussi des définitions utiles, spécifiques à ces applications, qui seront utilisées dans ces deux chapitres. Notons que le *Rapport Signal sur Bruit* (RSB) qui mesure la puissance moyenne du signal $P_{\text{moy}}$ par rapport à celle du bruit sera défini comme :

$$\text{RSB} = 10 \log_{10} \left( \frac{P_{\text{moy}}}{\sigma_\epsilon^2} \right)$$

où correspond $\sigma_\epsilon^2$ à la variance du bruit.

### 1.5.1 Données tests d'analyse spectrale en échantillonnage irrégulier

Les données présentées ici s'inspirent de problèmes classiques d'analyse spectrale en échantillonnage irrégulier en astrophysique. Il s'agit des mêmes données que celles traitées par [BC06] et que nous avons réutilisées dans [BCBB16]. Le signal est représenté en Figure 1.3 et est simulé en calculant la somme de cinq sinusoïdes dont les fréquences sont comprises entre 0.1 et 0.16 cycles par jour, avec un bruit blanc gaussien additif. Le RSB est égal à 5dB, où la puissance moyenne du signal est calculée comme $P_{\text{moy}} = \sum_\ell (x_\ell^2(1) + x_\ell^2(2))/2$. La simulation s'effectue suivant le modèle suivant :

$$\boldsymbol{y} = \sum_{\ell=1}^{5} (x_\ell(1) \cos(2\pi\nu_\ell \boldsymbol{t}) + x_\ell(2) \sin(2\pi\nu_\ell \boldsymbol{t})) + \boldsymbol{\epsilon}$$

| Indices | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Fréquences (jours$^{-1}$) | 0.0997 | 0.1120 | 0.1275 | 0.1433 | 0.1571 |
| Amplitudes | 0.4<br>0.3 | -0.72<br>-0.2 | -0.6<br>-0.8 | 0.45<br>0.6 | -0.46<br>0.2 |

TAB. 1.2 – Valeurs de simulation des fréquences et des amplitudes pour le signal $\boldsymbol{y}$.

où $\boldsymbol{t}$ regroupe les $N$ instants non réguliers d'échantillonnage et où les fréquences $\nu_\ell$ ainsi que les amplitudes $(x_\ell(1), x_\ell(2))$ sont récapitulées dans le Tableau 1.2.

Les $N = 250$ données sont échantillonnées irrégulièrement sur une durée de 500 jours ; la présence de trous périodiques dans l'échantillonnage rend ce problème particulièrement difficile. Pour des raisons liées au problème physique, les fréquences sont recherchées entre $\nu_{\min} = 0$ et $\nu_{\max} = 0.25$ jours$^{-1}$.

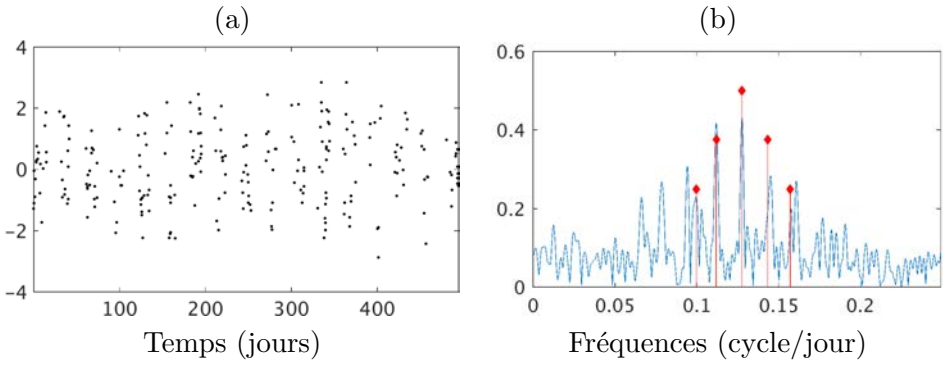



FIG. 1.3 – (a) Signal irrégulièrement échantillonné contenant 5 sinusoïdes. (b) Spectre de Fourier (-) et vraies fréquences ($\diamond$).

**Choix de la grille.** Dans le cas des problèmes d'APNL, on rappelle que le nombre d'intervalles $J$ de la grille s'écrit $J = (\nu_{\max} - \nu_{\min})/\Delta$, où $\Delta$ est la longueur des intervalles $\mathcal{I}_j$ dans lesquels peuvent évoluer les fréquences. Pour définir le nombre d'intervalles/leur largeur, on se base sur la résolution spectrale induite par la fenêtre spectrale relative à l'échantillonnage. En effet, on peut considérer qu'on ne pourra pas trouver deux fréquences éloignées de moins de la largeur du lobe principal, noté $\Delta_f$. En analogie avec le cas de l'échantillonnage régulier, la largeur du lobe principal de la fenêtre spectrale en échantillonnage irrégulier est de $\Delta_f = 2/T$, où $T = t_N - t_1$ correspond à la durée de l'observation. On peut donc choisir une largeur d'intervalle $\Delta$ qui satisfait la résolution spectrale, *i.e.* qui est suffisamment petite pour ne contenir qu'une fréquence.

$$\Delta \leq \Delta_{\max} = \Delta_f = \frac{2}{T}$$

Soit, formulée en nombre d'intervalles :

$$J \geq J_{\min} = \frac{\nu_{\max} - \nu_{\min}}{\Delta_{\max}} = \frac{T}{2}\,(\nu_{\max} - \nu_{\min})$$

Pour le signal défini dans cette section, cela donne $J_{\min} = 63$ et $\Delta_{\max} = 0.0040$ j$^{-1}$. Nous avons choisi les valeurs suivantes :

$$J = 100 \Leftrightarrow \Delta = 0.0025\text{j}^{-1}.$$

### 1.5.2 Données tests de déconvolution impulsionnelle

Les données présentées ici s'inspirent des travaux menés sur la déconvolution impulsionnelle, en reprenant notamment l'ondelette utilisée par [CBIS13], qui est de la forme suivante :

$$g(t) = \beta e^{-\alpha t^2} \cos(2\pi f_c t + \phi)$$

Elle est représentée en Figure 1.4(a), échantillonnée avec une période $T_e = 1s$ sur une durée de $37T_e$ comportant la quasi totalité de son énergie.

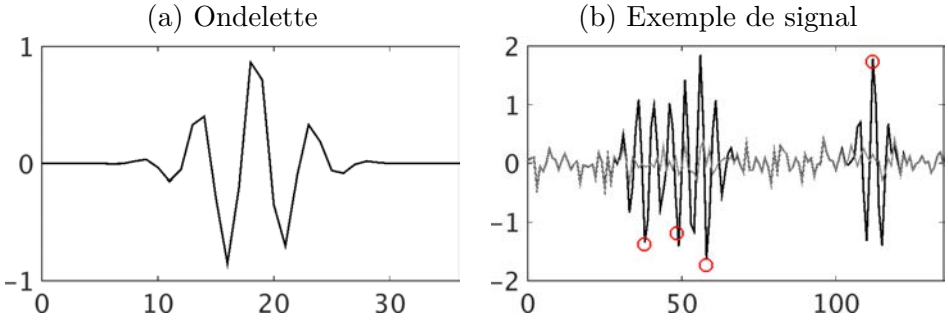



FIG. 1.4 – (a) Ondelette utilisée. (b) Exemple de signal simulé avec RSB=10dB ; signal en noir, bruit en gris, positions réelles du maximum des échos signalées avec les ronds rouges.

Les signaux sont simulés suivant le modèle suivant :

$$\boldsymbol{y} = \sum_{\ell=1}^{L} x_\ell \, \boldsymbol{h}(\nu_\ell) + \boldsymbol{\epsilon}, \quad \text{avec } \boldsymbol{h}(\nu_\ell) = h(\boldsymbol{t}; \nu_\ell) = g(\boldsymbol{t} - \nu_\ell)$$

À partir de ce modèle, nous simulons une bibliothèque de 200 signaux non bruités suivant les choix suivants :

- $L = 4$ : 4 échos décalés sont utilisées pour simuler chaque signal.
- $\nu_\ell \sim \mathcal{U}([0; J-1[)$ : les positions $\nu_\ell$ des échos (correspondant à la position de l'écho) sont tirées aléatoirement dans l'intervalle $[0; J-1[$ avec $J = 100$.
- $x_\ell \sim \mathcal{U}([-M_0; -0.5] \cup [0.5; M_0])$ : les amplitudes sont tirées avec un signe aléatoire et une valeur absolue comprise entre 0.5 et $M_0 = 2$. La borne inférieure de 0.5 a été choisie afin d'éviter les écueils liés à des amplitudes trop petites.
- $N = 136$ : les données $\boldsymbol{y}$ sont simulées dans $\mathbb{R}^N$, avec les instants d'échantillonnage $\boldsymbol{t}$ tels que $t_n = (n-1)$.

Enfin, le bruit $\boldsymbol{\epsilon}$ sera généré suivant une loi gaussienne centrée $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ avec trois RSB différents : 10, 20 et 30 dB, pour lesquels la puissance moyenne du signal est calculée comme $P_{\text{moy}} = \|\boldsymbol{y}_{\text{mod}}\|^2/N$, où $\boldsymbol{y}_{\text{mod}}$ correspond au signal non bruité. On obtient ainsi une bibliothèque de 600 signaux différents, dont un exemple est donné en Figure 1.4(b).

**Grille de recherche.** La recherche des décalages se fait entre $\nu_{\min} = 0$ et $\nu_{\max} = 99$. La grille la plus naturelle suit l'échantillonnage du signal : $\mathcal{G} = \{0, \ldots, J-1\}$, donnant une taille d'intervalle $\Delta = T_e = 1$. En effet, le modèle de convolution continu en équation (1.2) est en général approximé par le modèle de convolution discret suivant :

$$\begin{aligned} y_n &= \int x(u) g(nT_e - u) \, \mathrm{d}u \\ &\approx \sum_j x(jT_e) h((n-j)T_e), \end{aligned}$$

dans le cas de la recherche de décalage discret (appartenant à la grille). Prendre cette grille revient donc à chercher des décalages entre deux instants d'échantillonnage du signal. La grille de recherche est représentée en Figure 1.5.

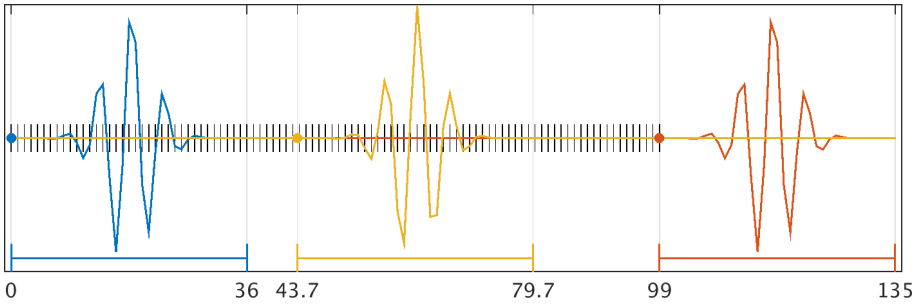



FIG. 1.5 – Schéma représentant l'échantillonnage des données et la grille. En noir, la grille de recherche entre 0 et $J-1=99$. En bleu, l'écho $\boldsymbol{h}(0)$ et en rouge $\boldsymbol{h}(99)$, où $\nu \in \mathcal{G}$. En jaune, l'écho $\boldsymbol{h}(43.7)$ où $\nu \notin \mathcal{G}$. En-dessous de chaque écho est représentée sa durée énergétique.

**Indices de qualité pour ce problème de déconvolution.** Deux indices de qualité seront utilisés pour comparer les résultats des différentes méthodes :

- ELR : l'*Exact Location Recovery* (« Estimation exacte des localisations ») est une valeur binaire égale à 1 uniquement si $L=4$ pics sont retrouvés et si les erreurs de localisations entre les vrai pics et ceux estimés ne dépassent pas $\frac{\Delta}{2}$.
- ASD : l'*Average Spike Distance* (« Distance moyenne des pics »)(inspiré du critère utilisé en neuro-science [VR01]) compare les trains d'impulsions réel et estimé en calculant l'erreur quadratique moyenne entre leur convolution avec un noyau Laplacien (avec un écart-type de $\frac{\Delta}{2}$). Ce critère prend à la fois en compte les estimations de la localisation et de l'amplitude ; il est moins strict que l'ELR puisque l'impact de fausses détections de basses amplitudes est réduit.

# Chapitre 2

# Interpolation de dictionnaire et reformulation MIP

**Sommaire**


**2.1 Principe de l'interpolation de dictionnaire et travaux existants . . . . . 34**
2.1.1 Cadre général de l'interpolation de dictionnaire pour l'APNL . . . . . 35
2.1.2 Résolution du problème d'APLsc avec $\mathcal{D}_x = \mathbb{R}_+$ . . . . . 38
2.1.3 Différentes interpolations de dictionnaire avec $\mathcal{D}_x = \mathbb{R}_+$ . . . . . 40
2.1.4 Une extension aux amplitudes complexes . . . . . 44
2.1.5 Bilan . . . . . 45
**2.2 Étude de l'espace des contraintes . . . . . 47**
2.2.1 Cadre général . . . . . 47
2.2.2 Cas des amplitudes à valeurs réelles . . . . . 48
2.2.3 Cas des amplitudes à valeurs complexes . . . . . 51
2.2.4 Bilan . . . . . 52
**2.3 Vers une résolution optimale des problèmes d'APLsc . . . . . 52**
2.3.1 Une résolution optimale avec la pseudo-norme $\ell_0$ . . . . . 53
2.3.2 Prise en compte des contraintes pour la résolution avec pénalisation $\ell_1$ . . . . . 56
2.3.3 Bilan . . . . . 58
**2.4 Contributions pour l'interpolation de dictionnaire . . . . . 59**
2.4.1 Interpolation polaire . . . . . 59
2.4.2 Approximation par K-SVD . . . . . 65
**2.5 Application sur les données de déconvolution impulsionnelle . . . . . 72**
2.5.1 Réglages des paramètres des méthodes pour les tests . . . . . 73
2.5.2 Comparaison des méthodes de résolution $\ell_1$ et $\ell_0$ avec interpolation polaire . . . . 75
2.5.3 Comparaison des interpolations de dictionnaire avec résolution $\ell_0$ . . . . . 77
2.5.4 Bilan . . . . . 79
**2.6 Extension de l'interpolation de dictionnaire pour l'application à l'analyse spectrale . . . . . 80**
2.6.1 Modèle *cosinus et sinus* et parcimonie structurée . . . . . 81
2.6.2 Modèle *cosinus+phase* et paramètres de forme . . . . . 86
2.6.3 Comparaison des différentes formulations proposées sur le signal test . . . . . 90
**2.7 Conclusion . . . . . 92**

Ce chapitre présente nos travaux sur la reformulation du problème d'approximation parcimonieuse non linéaire (APNL) en problème d'approximation parcimonieuse linéaire sous contraintes (APLsc) grâce au principe de l'« interpolation de dictionnaire » introduit initialement par [ETS11], et à leur reformulation sous la forme de *Problème d'optimisation mixte en nombres entiers* (« Mixed integer program »/MIP).

Les travaux entourant l'interpolation de dictionnaire sont assez disparates, tant sur l'approximation linéaire permettant de linéariser l'atome continu que sur les méthodes de résolution du problème d'APLsc associé. La première section de ce chapitre § 2.1 consiste donc en la définition d'un *cadre général* pour l'interpolation de dictionnaire pour des problèmes d'APNL ; nous y définissons les notations et les termes techniques que nous utiliserons dans la suite du chapitre, en nous basant sur les travaux existants. *Notons qu'en l'absence de théorie bien établie, les termes que nous employons sont souvent propres à ce manuscrit. Notamment, l'appellation « interpolation de dictionnaire » n'est pas employée couramment à notre connaissance, mais elle illustre bien la méthodologie présentée.*

Nous développons ensuite notre étude des différentes étapes composant l'interpolation de dictionnaire. En § 2.2, nous commençons par formuler de façon correcte l'espace des contraintes de façon générale quand les amplitudes des atomes continus sont à valeurs réelles ou complexes. Ensuite, nous abordons en § 2.3 la résolution des problèmes d'APLsc, et présentons notre principale contribution : la reformulation des problèmes d'APLsc en MIP, permettant la prise en compte exacte de la parcimonie grâce à l'introduction de variables binaires, et ainsi de répondre à un des objectifs de cette thèse. Nous revenons par ailleurs sur les résolutions de type pénalisation $\ell_1$ déjà proposée pour les problèmes d'APLsc pour y apporter quelques améliorations. En § 2.4, nous étudions plus en détails deux interpolations de dictionnaire proposées dans la littérature pour y apporter des nouveautés : l'interpolation polaire et l'approximation par K-SVD. Les contributions apportées dans ces trois sections sont testées en § 2.5 sur les signaux tests de déconvolution impulsionnelle.

Enfin, nous proposerons en § 2.6 une extension du cadre général de l'interpolation de dictionnaire présenté en § 2.1 pour une application aux données tests d'analyse spectrale en échantillonnage irrégulier. Cela permettra notamment de révéler le potentiel de l'approximation par K-SVD grâce aux améliorations que nous lui avons apportées.

*Remarque : Les travaux que nous avons réalisés sur l'interpolation de dictionnaire sont présentés de façon intuitive et très graphique pour une bonne compréhension du lecteur. Un formalisme mathématique plus rigoureux pourrait sans doute être pris en compte pour une présentation plus synthétique, mais peut-être au prix d'une difficulté de compréhension accrue.*

## 2.1 Principe de l'interpolation de dictionnaire et travaux existants

Le principe de l'« interpolation de dictionnaire » a été introduit dans le cadre des approximations parcimonieuses par [ETS11]. Des extensions ont ensuite été proposées, notamment par [KYHP14, FDJ15, PPZ17]. Nous proposons dans cette section de présenter ce principe dans le cadre général des problèmes inverses non linéaires et de leur reformulation en problèmes d'APNL. Nous pourrons ainsi définir des notations et des termes génériques qui pourront être utilisés tout au long de ce chapitre, et pour commencer dans cette section, pour décrire les différents travaux existants.

### 2.1.1 Cadre général de l'interpolation de dictionnaire pour l'APNL

Nous commençons par définir un *cadre général* de l'interpolation de dictionnaire pour les problèmes inverses non linéaires parcimonieux (1.1), qui repose sur les hypothèses suivantes :

1. Les paramètres non linéaires se réduisent au paramètre de localisation : $\boldsymbol{\nu}_\ell = \nu_\ell \in \mathbb{R}$.
2. Les amplitudes des atomes continus sont scalaires : $\boldsymbol{x}_\ell = x_\ell \in \mathcal{D}_x$ $(m = 1)$, où $\mathcal{D}_x$ peut être $\mathbb{R}_+$, $\mathbb{R}$ ou $\mathbb{C}$. Les atomes continus sont donc constitués d'une seule colonne.

C'est dans ce cadre que la plupart des travaux sur l'interpolation de dictionnaire ont été abordés.

Le problème inverse (1.1) peut être reformulé en problème d'APNL, en découpant le domaine de localisation $\mathcal{I}_\nu = [\nu_{\min}; \nu_{\max}]$ en $J$ intervalles $\mathcal{I}_j$ de largeur $\Delta$. Nous introduisons ici des notations et des appellations spécifiques à l'interpolation de dictionnaire. En considérant la grille $\mathcal{G} = \{\ddot{\nu}_j\}_j$ telle que $\ddot{\nu}_j$ est le milieu de l'intervalle $\mathcal{I}_j$, nous introduisons les *décalages* $\delta_j \in \mathcal{I}_\Delta = [-\Delta/2; \Delta/2]$, caractérisant l'éloignement des paramètres de localisation $\nu_j$ du centre de leur intervalle : $\nu_j = \ddot{\nu}_j + \delta_j$, tels qu'illustrés en Figure 2.1. L'atome continu peut alors se réécrire uniquement suivant $\delta_j$ puisque $\ddot{\nu}_j$ est fixé : $\boldsymbol{h}(\nu_j) = \boldsymbol{h}(\ddot{\nu}_j + \delta_j) = \boldsymbol{h}_j(\delta_j)$ ; soit, en récapitulatif de ces notations :

$$\begin{aligned} \textit{décalages} \ : \quad & \delta_j & = \quad & \nu_j - \ddot{\nu}_j \in \mathcal{I}_\Delta \\ \textit{intervalle des décalages} \ : \quad & \mathcal{I}_\Delta & = \quad & [-\Delta/2; \Delta/2] \\ \textit{atome continu} \ : \quad & \boldsymbol{h}_j(\delta_j) & = \quad & \boldsymbol{h}(\ddot{\nu}_j + \delta_j) = \boldsymbol{h}(\nu_j) \end{aligned}$$

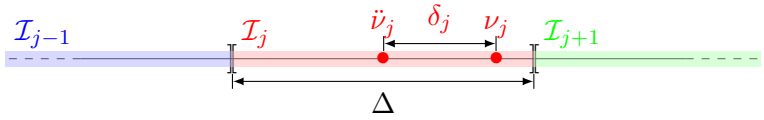



FIG. 2.1 – Représentation de la partition du domaine de localisation pour la reformulation en APNL spécifique au cadre de l'interpolation de dictionnaire.

Dans le cadre général et avec ces notations, le problème d'APNL s'écrit de la façon suivante :

$$\text{Estimer } \boldsymbol{x} \textit{ parcimonieux et } \boldsymbol{\delta} \in \mathcal{I}_\Delta^J \text{ tels que } \boldsymbol{y} = \sum_{j=1}^{J} x_j \, \boldsymbol{h}_j(\delta_j) + \boldsymbol{\epsilon}. \tag{2.1}$$

Résoudre ce problème dans le cadre déterministe revient à résoudre un problème d'optimisation bi-objectif, comme nous l'avons vu pour l'APL en § 1.2.2 ; on veut minimiser à la fois l'ordre du modèle $L = \|\boldsymbol{x}\|_0$ et un terme d'attache aux données, qui ici sera supposé quadratique $\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\delta}; \boldsymbol{y}) = \|\boldsymbol{y} - \sum_j x_j \boldsymbol{h}_j(\delta_j)\|^2$. Cependant, la non-linéarité des atomes continus en les décalages $\delta_j$ rend ce problème particulièrement difficile : on parle d'*optimisation non linéaire* dans le sens où le critère à minimiser n'est pas linéaire et n'est certainement pas convexe. D'après [BV04], il n'existe pas de méthodes efficaces pour résoudre un problème d'optimisation non linéaire dans le cadre général.

Le principe de l'interpolation de dictionnaire est alors de se ramener à un problème d'APL sous contraintes (APLsc) en approchant les « signaux élémentaires » non linéaires $x_j \, \boldsymbol{h}_j(\delta_j)$ par des combinaisons linéaires $\widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j$, dont seules les amplitudes $\widetilde{\boldsymbol{x}}_j$ sont à estimer. Cette « approximation linéaire » s'effectue en deux étapes :

1. *Approximation.* On exploite le fait que les décalages $\delta_j$ sont faibles (inférieurs à $\Delta/2$ en valeur absolue) pour approcher l'atome continu comme une combinaison de $K$ vecteurs

colonnes $\{\widetilde{\boldsymbol{h}}_{j,k}\}_k$ indépendants de $\delta_j$ :

$$\underbrace{\mathbf{h}_j(\delta_j)}_{N\times 1} \approx \sum_{k=1}^{K} \Psi_k(\delta_j)\, \widetilde{\boldsymbol{h}}_{j,k} = \underbrace{\widetilde{\mathbf{h}}_j}_{N\times K}\, \underbrace{\Psi(\delta_j)}_{K\times 1}$$

La matrice $\widetilde{\mathbf{h}}_j$, que nous nommons « atome interpolé », regroupe les $K$ vecteurs colonne $\boldsymbol{h}_{j,k}$. Nous nommons la fonction $\Psi(\delta_j)$ l'« interpolateur » ; il modélise la dépendance en les décalages $\delta_j$. Par exemple, l'approximation la plus naturelle est le développement de Taylor à l'ordre 1 : on a alors $\boldsymbol{h}_j(\delta_j) \approx \boldsymbol{h}_j(0) + \delta_j \boldsymbol{h}'_j(0)$, et donc $\widetilde{\mathbf{h}}_j = [\boldsymbol{h}_j(0), \boldsymbol{h}'_j(0)]$ et $\Psi(\delta_j) = [1, \delta_j]^T$ $(K = 2)$.

2. *Linéarisation par changement de variables.* Le signal élémentaire peut donc s'approcher via l'approximation précédente : $x_j\, \boldsymbol{h}_j(\delta_j) \approx \widetilde{\mathbf{h}}_j\, x_j \Psi(\delta_j)$. La linéarisation s'effectue en considérant le changement de variables $\widetilde{\boldsymbol{x}}_j = g_j(x_j, \delta_j)$ défini de la façon suivante[1] :

$$g \;:\; \left\{ \begin{array}{ccccccc} ( & \mathcal{D}_x & \times & \mathcal{I}_\Delta & ) & \subset \mathbb{F}^2 & \longrightarrow \quad \mathcal{S}_j \subset \mathbb{F}^K \\ ( & x & , & \delta & ) & & \longmapsto \quad \widetilde{\boldsymbol{x}} = x\Psi(\delta) \end{array} \right. \tag{2.2}$$

L'espace image du changement de variables est noté $\mathcal{S}_j$ et est inclus dans $\mathbb{F}^K$. *Notons que cet espace image est défini par les points de $\mathbb{F}^K$ qui sont multiples de la « courbe » $\Psi(\mathcal{I}_\Delta) = \{\Psi(\delta)\}_{\delta\in\mathcal{I}_\Delta}$. On dira que l'espace $\mathcal{S}_j$ est engendré par la « dilatation » de la courbe $\Psi(\mathcal{I}_\Delta)$.*

On obtient donc bien une approximation linéaire $x_j\, \boldsymbol{h}_j(\delta_j) \approx \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j$, dont l'erreur dépendra de l'approximation choisie. Le problème d'APNL (2.1) peut alors être remplacé par le problème d'APL sous contraintes (APLsc) suivant :

$$\text{Estimer } \widetilde{\boldsymbol{x}} \in \mathcal{C} \text{ } \textit{parcimonieux} \text{ tels que } \boldsymbol{y} = \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}} + \widetilde{\boldsymbol{\epsilon}} = \sum_{j=1}^{J} \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j + \widetilde{\boldsymbol{\epsilon}}. \tag{2.3}$$

où l'espace $\mathcal{C}$ prend en compte l'ensemble des contraintes à respecter par $\widetilde{\boldsymbol{x}}_j$ pour que $\widetilde{\boldsymbol{x}}_j = g(x_j, \delta_j)$ satisfasse $x_j \in \mathcal{D}_x$ et $\delta_j \in \mathcal{I}_\Delta$ (cela sera précisé ultérieurement). On appelle $\widetilde{\mathbf{H}}$ le « dictionnaire interpolé » pour le distinguer du *dictionnaire discret* $\mathbf{H}$ et du *dictionnaire continu* $\mathbf{H}(\boldsymbol{\nu})$. Le bruit $\widetilde{\boldsymbol{\epsilon}}$ prend maintenant en compte l'erreur d'approximation induite par l'approximation linéaire des signaux élémentaires. Ce problème d'APLsc est plus simple à résoudre que le problème d'APNL, grâce à la linéarité du modèle ; mais la difficulté est transférée au niveau de la prise en compte des contraintes. Pour les méthodes de résolution actuellement proposées (pénalisation $\ell_1$, approches gloutonnes, et pour $\mathcal{D}_x = \mathbb{R}_+$ ; nous y revenons en § 2.1.2), il peut être formulé comme un *problème d'optimisation convexe*, dans le sens où le critère et les contraintes sont convexes [BV04]. Plus précisément, ces problèmes peuvent se reformuler sous la forme suivante :

$$\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T\mathbf{R}\boldsymbol{x} + \boldsymbol{f}^T\boldsymbol{x} \text{ avec } \left\{ \begin{array}{rcll} \mathbf{A}_i\boldsymbol{x} & \leq & \boldsymbol{d}_i & \text{contraintes linéaires inégalités} \\ \mathbf{A}_e\boldsymbol{x} & = & \boldsymbol{d}_e & \text{contraintes linéaires égalités} \\ \boldsymbol{l}^T\boldsymbol{x} + \boldsymbol{x}^T\mathbf{Q}\boldsymbol{x} & \leq & r & \text{contraintes quadratiques inégalités} \end{array} \right. \tag{2.4}$$

où le critère à minimiser est quadratique (avec une matrice $\mathbf{R}$ semi-définie positive), et où les contraintes définissant un « espace de contraintes convexe », peuvent être de type inégalités ou

[1] Bien que pour certaines approximations linéaires, il est possible que l'interpolateur $\Psi$, et donc le changement de variables $g$, dépendent de l'intervalle pour lequel on considère l'approximation, nous n'indiquerons pas l'indice $j$ de l'intervalle pour des raisons de clarté.

égalités linéaires, et inégalités quadratiques. Pour ces dernières, soit la matrice $\boldsymbol{Q}$ doit être semi-définie positive, soit elles peuvent se reformuler en une contrainte de type $a_1^2x_1^2+a_2^2x_2^2-a_3^2x_3^2 \leq 0$ ou $a_1^2x_1^2+a_2^2x_2^2-a_{3,4}^2x_3x_4 \leq 0$ (contraintes de type cône du second degré («Second Order Cone»/ SOC) ; nous ne rentrons pas dans les détails ici, voir [BV04]). Cette forme générale d'un problème d'optimisation convexe peut être résolue par des logiciels, comme le logiciel CPLEX que nous utiliserons. Dans la suite de ce chapitre, nous dirons pour simplifier que, par abus de langage, le problème est un problème d'optimisation convexe s'il peut se mettre sous la forme (2.4).

Pour revenir à notre problème d'APLsc (2.3), les contraintes $\widetilde{\boldsymbol{x}} \in \mathcal{C}$ compliquent l'utilisation des méthodes classiques de résolution d'APL, mais elles sont nécessaires pour garantir une bonne « identification » des paramètres du problème initial $(x_j, \delta_j)$ à partir des $\widetilde{\boldsymbol{x}}_j$. En effet, si la fonction $g$ est bijective[2], alors contraindre $\widetilde{\boldsymbol{x}}_j \in \mathcal{S}_j$ permet de retrouver un unique $(x_j, \delta_j)$ par inversion du changement de variables $(x_j, \delta_j) = g^{-1}(\widetilde{\boldsymbol{x}}_j)$. Dans ce cas idéal, on a :

$$\textit{Cas idéal :} \quad \widetilde{\boldsymbol{x}} \in \mathcal{C} \Leftrightarrow \forall j \in [\![1; J]\!],\ \widetilde{\boldsymbol{x}}_j \in \mathcal{S}_j.$$

Cependant, l'espace image $\mathcal{S}_j$ peut rarement être considéré comme un espace de contraintes convexe comme défini en (2.4) ; les auteurs sont généralement amenés à proposer une relaxation convexe de cet espace, que nous notons $\mathcal{H}_j$. L'espace des contraintes s'écrit alors :

$$\textit{Cas relâché :} \quad \widetilde{\boldsymbol{x}} \in \mathcal{C} \Leftrightarrow \forall j \in [\![1; J]\!],\ \widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j.$$

Dans ce cas, l'*identification* des paramètres initiaux $(x_j, \delta_j)$ ne peut plus se faire par inversion du changement de variables, puisqu'un $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j$ estimé peut ne pas appartenir à $\mathcal{S}_j$. Différentes solutions sont alors proposées.

Le problème, tel que nous l'avons posé, peut se séparer en deux parties distinctes et indépendantes :

1. la définition d'une approximation linéaire de faible erreur $\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j$ et de son espace image $\mathcal{S}_j$ ; si besoin, calculer son enveloppe convexe $\mathcal{H}_j \supset \mathcal{S}_j$ de type (2.4) pour définir un espace de contraintes $\mathcal{C}$ et donner une méthode d'identification des paramètres initiaux. L'ensemble de ces étapes seront regroupées par la suite sous le terme *interpolation de dictionnaire*.
2. la résolution du problème d'APLsc : adapter les différentes méthodes de résolution d'APL (notamment la prise en compte de la parcimonie) à la présence des contraintes convexes.

Nous allons maintenant détailler, pour chacune de ces parties, les différentes variantes qui ont été proposées dans la littérature. Une grande majorité des travaux, basés sur ceux de [ETS11], ont été réalisés dans le cas $\mathcal{D}_x = \mathbb{R}_+$, soit pour des amplitudes du modèle initiale $x_j$ positives. Une extension au cas $\mathcal{D}_x = \mathbb{C}$ (amplitudes $x_j$ complexes) a été proposée par [FDJ15]. Nous commençons donc par détailler les deux parties *pour des amplitudes positives*, avant de résumer succinctement l'extension aux amplitudes complexes.

Pour le cas où $\mathcal{D}_x = \mathbb{R}_+$, nous définissons les notations spécifiques suivantes : $\mathcal{S}_j^+$ est l'espace image de $g$ sur $\mathbb{R}_+ \times \mathcal{I}_\Delta$ et $\mathcal{H}_j^+$ une relaxation convexe. Nous définissons de plus quelques notations supplémentaires :

– Dans un cadre général, on écrira $\widetilde{\boldsymbol{x}}_j \in \mathcal{C}_j$ où $\mathcal{C}_j$ représente l'espace des contraintes pour les amplitudes $\widetilde{\boldsymbol{x}}_j$, qui pourra être l'espace image $\mathcal{S}_j^+$ ou sa relaxation convexe $\mathcal{H}_j^+$.

[2] La bijectivité de la fonction $g$ est à comprendre de son espace de définition vers son espace image (elle est donc par définition surjective), et en omettant le cas où $x_j = 0$. En effet, quel que soit l'interpolateur considéré, $\widetilde{\boldsymbol{x}}_j = 0$ correspond à une infinité de couples $(x_j, \delta_j) = (0, \delta_j)$. Cela n'a pas d'importance dans la cadre de nos travaux, puisque $\delta_j$ n'a pas d'influence sur le modèle quand $x_j = 0$.

- À partir de l'intervalle $\mathcal{I}_\Delta$, nous définissons l'ensemble discret $\mathcal{I}_\Delta^P$ constitué de $P$ valeurs discrètes de $\delta$ réparties uniformément dans $\mathcal{I}_\Delta$ : $\delta_p = -\frac{\Delta}{2} + \frac{p-1}{P-1}\Delta \in \mathcal{I}_\Delta^P \subset \mathcal{I}_\Delta$, pour $p \in [\![1; P]\!]$.
- L'opérateur $\mathcal{P}(\cdot)$ est un opérateur de projection permettant de passer de l'espace relâché $\mathcal{H}_j^+$ à l'espace image $\mathcal{S}_j^+$.
- Pour un vecteur $\widetilde{\boldsymbol{x}}_j$, on pourra être amené à écrire ses trois premières composantes de la façon suivante : $[\widetilde{x}_{j,1}, \widetilde{x}_{j,2}, \widetilde{x}_{j,3}] = [\alpha_j, \beta_j, \gamma_j]$.

### 2.1.2 Résolution du problème d'APLsc avec $\mathcal{D}_x = \mathbb{R}_+$

Le problème d'APLsc (2.3) diffère des problèmes d'APL étudiés en § 1.2 par la présence des contraintes $\widetilde{\boldsymbol{x}} \in \mathcal{C}$, spécifiques à l'interpolation de dictionnaire employée. Dans le cadre déterministe, le problème (2.3) s'écrit toujours comme un problème d'optimisation bi-objectif similaire à (1.7), mais où les amplitudes sont contraintes :

$$\min_{\widetilde{\boldsymbol{x}} \in \mathcal{C}} \left( \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2, \|\widetilde{\boldsymbol{x}}\|_0 \right)$$

Notons que contrairement au problème d'APNL initial où les amplitudes $x_j$ sont supposées scalaires, les amplitudes du problème d'APLsc sont vectorielles : $\widetilde{\boldsymbol{x}}_j \in \mathbb{R}^K$. *La parcimonie doit donc être prise en compte de façon structurée ;* on rappelle l'expression de la pseudo-norme $\ell_0$ dans ce cas :

$$\|\widetilde{\boldsymbol{x}}\|_0 = \mathrm{Card}(\mathsf{Supp}\,(\widetilde{\boldsymbol{x}})) \quad \text{avec} \quad \mathsf{Supp}\,(\widetilde{\boldsymbol{x}}) = \{j \in [\![1; J]\!],\ \widetilde{\boldsymbol{x}}_j \neq \mathbf{0}\}$$

Dans certains cas particulier d'interpolation de dictionnaire, on sera amené à voir qu'il est possible de considérer la parcimonie *uniquement sur le première composante des* $\widetilde{\boldsymbol{x}}_j$, car l'espace des contraintes $\mathcal{C}$ respecte l'implication suivante :

$$\widetilde{\boldsymbol{x}} \in \mathcal{C} \Rightarrow (\forall j \in [\![1; J]\!],\ \alpha_j = \widetilde{x}_{j,1} = 0 \Rightarrow \forall k > 1,\ \widetilde{x}_{j,k} = 0) \tag{2.5}$$

que l'on appellera par la suite « implication de nullité ». Quand elle est respectée, il suffit donc de minimiser $\|\boldsymbol{\alpha}\|_0$ où $\boldsymbol{\alpha}$ contient les premières composantes $\alpha_{j,1} = \widetilde{x}_j$ de $\widetilde{\boldsymbol{x}}$.

La résolution des problèmes d'APLsc a jusqu'à présent été abordée uniquement à l'aide de méthodes sous-optimales telles que la pénalisation $\ell_1$ ou les approches gloutonnes.

#### Pénalisation $\ell_1$ : Continuous Basis Pursuit (CBP)

En introduisant le principe de l'interpolation de dictionnaire, [ETS11] propose dans un cadre général[3] de prendre en compte la parcimonie structurée en relaxant la pseudo-norme $\ell_0$ par la norme mixte $\ell_{1,2}$ (définie en (1.10)), afin d'obtenir la formulation suivante :

$$\min_{\widetilde{\boldsymbol{x}} \in \mathcal{C}} \quad \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 + \lambda \|\widetilde{\boldsymbol{x}}\|_{1,2}$$

Ce problème peut être résolu en utilisant des méthodes classiques d'optimisation convexe (les auteurs citent des méthodes de « point intérieur »), *à condition que l'espace des contraintes* $\mathcal{C}$ *soit convexe.*

En pratique, les auteurs n'utilisent pas la norme mixte $\ell_{1,2}$, car l'espace des contraintes qu'ils définissent relatif à l'interpolation de dictionnaire qu'ils emploient (interpolation polaire, voir

[3] Les auteurs proposent une formalisation de la résolution du problème dans un cadre général (voir § IV.C, « General Interpolation »), où l'interpolation de dictionnaire n'est pas précisée.

§ 2.1.3) respecte l'implication de nullité. Ils utilisent donc la norme $\ell_1$ du vecteur $\boldsymbol{\alpha}$ regroupant les premières composantes $\alpha_j$ des amplitudes $\widetilde{\boldsymbol{x}}_j$ :

$$\min_{\widetilde{\boldsymbol{x}} \in \mathcal{C}} \quad \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 + \lambda \|\boldsymbol{\alpha}\|_1 \tag{2.6}$$

Par analogie avec la dénomination *Basis Pursuit* employée pour la résolution du problème pénalisé par la norme $\ell_1$ pour les problèmes d'APL, les auteurs nomment leur méthode *Continuous Basis Pursuit*, dans le sens où la solution du problème (2.6) leur permet de retrouver de façon continue les décalages $\delta_j$.

### Approches gloutonnes

À notre connaissance, deux articles ont proposé d'approcher la solution de (2.3) via l'utilisation d'algorithmes gloutons de type OMP : ainsi, [KYHP14] propose le « Continuous-OMP » (COMP) et [FDJ15] le « Interpolating Band-excluded OMP » (IBOMP). Nous commençons par décrire l'algorithme COMP qui est l'adaptation naturelle de l'OMP pour le problème d'APLsc, puis le IBOMP présentant d'importantes variantes.

En partant de l'initialisation $\boldsymbol{r}^0 = \boldsymbol{y}$, $\Omega^0 = \varnothing$, $\overline{\mathcal{C}}^0 = \varnothing$ (espace des contraintes pour les atomes sélectionnés) et $\overline{\mathbf{H}}^0 = [\,]$, l'algorithme COMP de [KYHP14] consiste à répéter les étapes de sélection et de mise à jour que nous décrivons pour l'itération $\ell$ :

1. *sélection* : parmi les atomes $(\widetilde{\mathbf{h}}_j)_{j \notin \Omega^{\ell-1}}$ du dictionnaire interpolé, on sélectionne celui qui permet de minimiser l'erreur quadratique vis-à-vis du résidu $\boldsymbol{r}^{\ell-1}$ :

$$\left.\begin{array}{rcl} \Omega^\ell & = & \Omega^{\ell-1} \cup \{j_\ell\} \\ \overline{\mathcal{C}}^\ell & = & \overline{\mathcal{C}}^{\ell-1} \times \mathcal{C}_{j_\ell} \\ \overline{\mathbf{H}}^\ell & = & [\overline{\mathbf{H}}^{\ell-1}, \widetilde{\mathbf{h}}_{j_\ell}] \end{array}\right\} \text{ avec } j_\ell = \arg\min_{j \notin \Omega^{\ell-1}} \left( \min_{\widetilde{\boldsymbol{x}}_j \in \mathcal{C}_j} \|\boldsymbol{r}^{\ell-1} - \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j\|^2 \right)$$

2. *mise à jour* : on réévalue l'ensemble des amplitudes des atomes sélectionnés en minimisant le critère des moindres carrés *contraints* pour mettre à jour le résidu :

$$\boldsymbol{r}^\ell = \boldsymbol{y} - \overline{\mathbf{H}}^\ell \widetilde{\boldsymbol{x}}^\ell \quad \text{où} \quad \widetilde{\boldsymbol{x}}^\ell = \arg\min_{\widetilde{\boldsymbol{x}} \in \overline{\mathcal{C}}^\ell} \|\boldsymbol{y} - \overline{\mathbf{H}}^\ell \widetilde{\boldsymbol{x}}\|^2$$

Les contraintes associées aux amplitudes $\widetilde{\boldsymbol{x}}_j$ empêchent la minimisation analytique du critère des moindres carrés comme décrite en § 1.2.2. La réalisation des étapes de sélection et de mise à jour nécessite donc la résolution de problèmes d'optimisation convexe. Notons qu'il est également proposé une mise à jour minimisant un critère non convexe basé sur le modèle non linéaire $x\boldsymbol{h}(\delta)$, en passant dans le domaine de Fourier (sans rapport avec l'interpolation utilisée à l'étape de sélection).

L'algorithme IBOMP proposé par [FDJ15] exploite le dictionnaire interpolé, mais est très différent du COMP. En gardant les mêmes notations que précédemment, nous décrivons l'itération $\ell$ de l'algorithme :

1. *sélection* : l'indice du support $j_\ell$ est sélectionné sur le dictionnaire discret $\mathbf{H}$ constitué des atomes $\{\boldsymbol{h}_j(0)\}_j$ ; ensuite, le décalage $\delta_{j_\ell}$ est estimé en exploitant l'interpolation de dictionnaire :

(a) $\Omega^\ell = \Omega^{\ell-1} \cup \{j_\ell\}$ avec $j_\ell = \arg\min_{j \notin B_\eta(\Omega^{\ell-1})} \|\boldsymbol{r}^{\ell-1} - \widehat{x}_j^{\mathrm{MC}} \boldsymbol{h}_j(0)\|^2$

(b) $\overline{\boldsymbol{\nu}}^\ell = [\overline{\boldsymbol{\nu}}^{\ell-1}, \ddot{\nu}_{j_\ell} + \delta_{j_\ell}]$ avec $(x_{j_\ell}, \delta_{j_\ell}) = g^{-1}\left(\mathcal{P}\left(\arg\min_{\widetilde{\boldsymbol{x}} \in \mathcal{C}_{j_\ell}} \|\boldsymbol{r}^{\ell-1} - \widetilde{\mathbf{h}}_{j_\ell} \widetilde{\boldsymbol{x}}\|^2\right)\right)$

2. *mise à jour* : l'ensemble des amplitudes des atomes continus (mais de décalages fixés) sélectionnés sont réévaluées par moindres carrés pour mettre à jour le résidu :

$$\boldsymbol{r}^\ell = \boldsymbol{y} - \mathbf{H}(\overline{\boldsymbol{\nu}}^\ell)\widehat{\boldsymbol{x}}^{\mathrm{MC}} \quad \text{avec } \widehat{\boldsymbol{x}}^{\mathrm{MC}} = (\mathbf{H}(\overline{\boldsymbol{\nu}}^\ell)^T\mathbf{H}(\overline{\boldsymbol{\nu}}^\ell))^{-1}\mathbf{H}(\overline{\boldsymbol{\nu}}^\ell)^T\boldsymbol{r}^\ell$$

La sélection de l'indice $j_\ell$ à l'étape 1.(a) ci-dessus s'effectue dans le *complémentaire*[4] de l'espace $B_\eta(\Omega^{\ell-1})$ nommé « bande $\eta$-cohérence » (« $\eta$-coherence band »), regroupant les indices des atomes discrets dont la cohérence avec ceux déjà sélectionnés est supérieure à $\eta$ ; il s'agit du caractère « Band-excluded » de IBOMP, inspiré de [FL12]. Notons que les auteurs proposent d'affiner le résultat obtenu avec IBOMP en utilisant leur version[5] du *Continuous Basis Pursuit* sur un support réduit constitué des indices estimés $\Omega^L$ ainsi que leur voisinage.

Ces deux algorithmes de type OMP diffèrent donc dans leur exploitation de l'interpolation de dictionnaire. Que ce soit pour l'étape de sélection ou de mise à jour, le coût de calcul de l'algorithme IBOMP est bien inférieur à celui du COMP ; IBOMP ne requiert qu'une seule optimisation convexe sous contraintes à chaque itération, contre $J$ optimisations similaires (sélection) et une optimisation convexe sous contraintes de plus grande dimension (mise à jour) pour le COMP.

D'un point de vue « qualité de la solution », la sélection du COMP est plus rigoureuse puisqu'il est possible qu'IBOMP ne fournisse pas l'indice correspondant au meilleur atome interpolé (et donc au meilleur atome continu, à condition que l'erreur d'approximation soit faible). Pour l'étape de mise à jour, les deux algorithmes présentent des avantages. Le COMP permet, via la réévaluation de l'ensemble des amplitudes des atomes interpolés, celle de l'ensemble des décalages par identification (mais sans l'exploiter). L'IBOMP permet quant à lui de corriger l'erreur d'approximation en calculant le résidu avec le modèle initial, mais en fixant les décalages estimés. *Une mise à jour hybride entre les deux algorithmes sembleraient donc plus optimale : la réévaluation de l'ensemble des amplitudes des atomes interpolés sélectionnés, puis l'identification des décalages permettant la mise à jour du résidu sur le modèle initial.* Une telle mise à jour hybride est nécessairement plus coûteuse que celle de l'IBOMP, et la qualité de la solution dépendra de l'erreur d'approximation induite par l'interpolation.

Pour les deux algorithmes, les critères d'arrêt sont relativement simples ; IBOMP s'arrête quand l'ordre du modèle attendu $L$ a été atteint et COMP quand le résidu devient suffisamment faible.

*Remarque : les optimisations lors de l'étape de sélection s'effectuent sur des espaces de faibles dimensions ($\mathcal{C}_j$ est de dimension $K$ avec $K$ faible généralement). En pratique, il est possible de trouver une solution analytique à de tels problèmes, et donc d'accélérer nettement cette étape. Cependant, nous n'approfondirons pas les approches gloutonnes dans la suite de ce chapitre.*

### 2.1.3 Différentes interpolations de dictionnaire avec $\mathcal{D}_x = \mathbb{R}_+$

*Comme dans l'ensemble de cette partie, nous nous intéressons à l'approximation linéaire d'un signal élémentaire $x_j\,\boldsymbol{h}_j(\delta_j)$ pour un indice $j$ fixé, nous nous permettons d'omettre cet indice afin d'alléger les notations.*

[4] Notons qu'une erreur s'est glissée dans l'algorithme décrivant le IBOMP (voir « Algorithm 1 », [FDJ15]) : il y est en effet indiqué que la sélection de l'indice doit se faire *dans* $B_\eta(\Omega^{\ell-1})$, et non dans son complémentaire. Cependant, il s'agit bien d'*exclure* les indices $B_\eta(\Omega^{\ell-1})$, comme cela est correctement expliqué dans le corps de l'article.

[5] Les auteurs étendent en effet le cadre de l'interpolation de dictionnaire aux atomes à amplitudes complexes ; nous revenons plus en détail sur leur version du CBP en § 2.1.4.

Dans la littérature, différentes approximations linéaires $\widetilde{\mathbf{h}}\widetilde{\boldsymbol{x}}$ de signaux élémentaires $x\,\boldsymbol{h}(\delta)$ avec $x \in \mathcal{D}_x = \mathbb{R}_+$ ont été proposées pour l'interpolation de dictionnaire : [ETS11], l'article fondateur, a ainsi proposé l'approximation de Taylor à l'ordre 1 (que nous notons (T1)) et l'interpolation polaire (Po). [KYHP14] a formalisé l'approximation de Taylor à l'ordre 2 (T2) (déjà évoquée par le précédent) et proposé l'approximation par K-SVD (K-SVD). Plus récémment, [PPZ17] a proposé l'approximation « Minimax » (Minimax).

Dans le Tableau 2.1, nous récapitulons les différentes étapes de ces interpolations de dictionnaire : l'approximation de l'atome continu et les hypothèses à respecter, la linéarisation par changement de variables, la définition de l'espace image $\mathcal{S}^+$ et de sa relaxation convexe $\mathcal{H}^+$, et enfin l'identification des paramètres initiaux à partir de $\widetilde{\boldsymbol{x}} \in \mathcal{H}^+$. Nous détaillons ensuite davantage ces étapes.

| | T1 [ETS11] | T2 [KYHP14] | Minimax [PPZ17] | Polaire [ETS11] | K-SVD [KYHP14] |
|---|---|---|---|---|---|
| Hypothèses | Dérivabilité de $\boldsymbol{h}(\delta)$ | Idem T1 + de $\boldsymbol{h}'(\delta)$ | $\varnothing$ | Invariance par translation | $\delta = \delta_p \in \mathcal{I}_\Delta^P$ |
| Atome interpolé $\widetilde{\mathbf{h}}$ | $[\boldsymbol{h}(0), \boldsymbol{h}'(0)]$ | $[\boldsymbol{h}(0), \boldsymbol{h}'(0), \frac{\boldsymbol{h}''(0)}{2}]$ | Optim. (2.7) | $[\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}]$ | $[\boldsymbol{u}_k]_k$ |
| Dimension $K$ | 2 | 3 | $K$ (au choix) | 3 | $K$ (au choix) |
| Interpolateur $\Psi(\delta) \in \mathbb{R}^K$ | $[1, \delta, \dots, \delta^{K-1}]^T$ | | | $\begin{bmatrix} 1 \\ r\cos(\frac{2\theta}{\Delta}\delta) \\ r\sin(\frac{2\theta}{\Delta}\delta) \end{bmatrix}$ | $\boldsymbol{z}_p = \Psi(\delta_p)$ |
| Changement de variables $\widetilde{\boldsymbol{x}} = g(x, \delta)$ | $\widetilde{x}_k = x\delta^{k-1}$ | | | $\widetilde{\boldsymbol{x}} = \begin{bmatrix} x \\ xr\cos(\frac{2\theta}{\Delta}\delta) \\ xr\sin(\frac{2\theta}{\Delta}\delta) \end{bmatrix}$ | $\widetilde{\boldsymbol{x}} = x\boldsymbol{z}_p$ |
| $\widetilde{x}_1 \geq 0$ ? | oui | | | | non |
| Inversion $(x, \delta) = g^{-1}(\widetilde{\boldsymbol{x}})$ | $x = \widetilde{x}_1$ $\delta = \widetilde{x}_2/\widetilde{x}_1$ | | | $x = \widetilde{x}_1$ $\delta = \frac{2\theta}{\Delta}\arctan(\frac{\widetilde{x}_3}{\widetilde{x}_2})$ | $\widetilde{\boldsymbol{x}} = x\boldsymbol{z}_p$ $\delta = \delta_p$ |
| $\mathcal{S}^+$ | $\widetilde{x}_1 \geq 0$ $-\frac{\Delta}{2}\widetilde{x}_1 \leq \widetilde{x}_2 \leq \frac{\Delta}{2}\widetilde{x}_1$ $\forall k > 2, \quad \widetilde{x}_k = \widetilde{x}_2^{k-1}/\widetilde{x}_1^{k-2}$ | | | $\widetilde{x}_1 \geq 0$ $\widetilde{x}_2 \geq r\cos\theta\widetilde{x}_1$ $\widetilde{x}_2^2 + \widetilde{x}_3^2 = r^2\widetilde{x}_1^2$ | $\exists p \in [\![1; P]\!]$, $\exists x \in \mathbb{R}_+$, $\widetilde{\boldsymbol{x}} = x\boldsymbol{z}_p$ |
| $\mathcal{S}^+$ convexe ? | oui | non | oui (K=2) non (K>2) | non | non |
| $\mathcal{H}^+$ | $\widetilde{x}_1 \geq 0$ $\forall k > 1, \begin{matrix} \min(0, (-\frac{\Delta}{2})^{k-1})\,\widetilde{x}_1 - \widetilde{x}_k \leq 0 \\ -(\frac{\Delta}{2})^{k-1}\ \widetilde{x}_1 + \widetilde{x}_k \leq 0 \end{matrix}$ | | | $\widetilde{x}_1 \geq 0$ $\widetilde{x}_2 \geq r\cos\theta\widetilde{x}_1$ $\widetilde{x}_2^2 + \widetilde{x}_3^2 \leq r^2\widetilde{x}_1^2$ | Non spécifié |
| Implication de nullité | oui | | | | |
| Identification $(x, \delta) \leftarrow \widetilde{\boldsymbol{x}} \in \mathcal{H}^+$ | $(x, \delta) = g^{-1}(\widetilde{\boldsymbol{x}})$ | Non spécifié | Proj. sur $\boldsymbol{h}(\delta)$ (2.8) | $\widetilde{\boldsymbol{x}} = \mathcal{P}(\widetilde{\boldsymbol{x}})$ $(x, \delta) = g^{-1}(\widetilde{\boldsymbol{x}})$ | $\widetilde{\boldsymbol{x}} = \mathcal{P}(\widetilde{\boldsymbol{x}})$ $(x, \delta) = g^{-1}(\widetilde{\boldsymbol{x}})$ |

TAB. 2.1 – Récapitulatif des propriétés des différentes interpolations de dictionnaire existantes.

**Interpolation de dictionnaire par approximations de Taylor et Minimax**

Les approximations de Taylor [ETS11, KYHP14] et l'approximation Minimax [PPZ17] sont très similaires, car la définition de l'interpolateur $\Psi(\delta)$ est identique[6] : $[\delta^{k-1}]_k$. L'espace image $\mathcal{S}^+$, sa relaxation $\mathcal{H}^+$ et l'identification des paramètres initiaux sont aussi donc identiques. La différence réside dans la définition de l'approximation de l'atome continu :

- l'approximation de Taylor se base sur le développement de Taylor à l'ordre $K$ autour de $\delta = 0$ : à l'ordre 2 par exemple, on a $\boldsymbol{h}(\delta) \approx \boldsymbol{h}(0) + \delta\boldsymbol{h}'(0) + \frac{1}{2}\delta^2\boldsymbol{h}''(0)$. L'erreur d'approximation commise sera donc plus faible pour les $\delta$ plus proches de 0. L'interpolateur $\Psi(\delta)$ et l'atome interpolé $\widetilde{\mathbf{h}}$ découlent donc directement de cette approximation.
- l'approximation Minimax quant à elle, fixe d'abord l'interpolateur (de façon identique à celui des approximations de Taylor), et définit l'atome interpolé $\widetilde{\mathbf{h}}$ comme celui qui *minimise l'erreur maximale d'approximation*. Pour résoudre le problème d'optimisation, l'auteur discrétise le décalage $\delta$ sur une grille $\mathcal{I}_\Delta^P$ très fine et obtient le problème suivant :

$$\widetilde{\mathbf{h}} = \arg\min_{\widetilde{\mathbf{h}}} \max_p \|\boldsymbol{h}(\delta_p) - \widetilde{\mathbf{h}}\Psi(\delta_p)\|_2, \quad \text{avec } \delta_p \in \mathcal{I}_\Delta^P \tag{2.7}$$

Cette optimisation est résolue en la reformulant en « Second order cone program » (SOCP). La définition de l'espace image $\mathcal{S}^+$ se fait à partir de l'écriture du changement de variables. Notons que pour $K > 2$, les composantes $\widetilde{x}_k$ sont fixées en fonction des deux premières, puisque les paramètres initiaux s'expriment directement en fonction de $\widetilde{x}_1$ et $\widetilde{x}_2$ ; par exemple, pour $K = 2$, $x\boldsymbol{h}(\delta) \approx x\widetilde{\mathbf{h}}_1 + x\delta\widetilde{\mathbf{h}}_2 + x\delta^2\widetilde{\mathbf{h}}_3 = \widetilde{\mathbf{h}}\widetilde{\boldsymbol{x}}$. Ainsi, la convexité de l'espace $\mathcal{S}^+$ va dépendre de la dimension $K$ ; si $K = 2$, l'espace est convexe (contraintes d'inégalités linéaires), mais si $K > 2$, l'espace n'est pas convexe (contraintes d'égalité quadratiques). La solution employée par [KYHP14] pour (T2) et par [PPZ17] pour (Minimax) est alors de définir une relaxation convexe $\mathcal{H}^+$ de $\mathcal{S}^+$ à l'aide de contraintes linéaires entre la première composante $\widetilde{x}_1$ et chacune des autres composantes $\widetilde{x}_k$, à partir de la constatation suivante, pour $\widetilde{\boldsymbol{x}} \in \mathcal{S}^+$ :

$$\frac{\widetilde{x}_k}{\widetilde{x}_1} = \left(\frac{\widetilde{x}_2}{\widetilde{x}_1}\right)^{k-1} = \delta^{k-1} \quad \Rightarrow \quad \left|\frac{\widetilde{x}_k}{\widetilde{x}_1}\right| \leq \left(\frac{\Delta}{2}\right)^{k-1}$$

Les contraintes données en Tableau 2.1 s'obtiennent grâce au fait que la première composante $\widetilde{x}_1$, qui s'identifie avec l'amplitude initiale $x$, est positive. Nous nommons ce type de contraintes « contraintes polyhédrales », dans le sens où à $\widetilde{x}_1$ fixé, l'espace contraint dans les plans $(\widetilde{x}_{k_1}, \widetilde{x}_{k_2})$ est un rectangle. Comme les composantes $\widetilde{x}_k$ sont encadrées linéairement par $\widetilde{x}_1$, l'implication de nullité est ici respectée. Les espaces images $\mathcal{S}^+$ ainsi que la relaxation convexe $\mathcal{H}^+$ sont représentés en Figure 2.2, pour les approximations de Taylor à l'ordre 1 et 2 (et donc équivalent à l'approximation Minimax pour $K = 2$ et $K = 3$). *Notons que pour l'approximation (T2), la contrainte quadratique d'égalité $\widetilde{x}_3\widetilde{x}_1 = \widetilde{x}_2^2$ pourrait être relâchée avec une contrainte quadratique d'inégalité $\widetilde{x}_3\widetilde{x}_1 \geq \widetilde{x}_2^2$ respectant les conditions de contraintes convexes ; cela reviendrait à considérer le volume à l'intérieur du cône parabolique décrit par $\mathcal{S}^+$.*

Enfin, l'identification des paramètres initiaux $(x, \delta)$ à partir d'un $\widetilde{\boldsymbol{x}} \in \mathcal{H}^+$ dépend de la dimension $K$ ; si $K = 1$, ils peuvent être directement retrouvés par inversion du changement de variables car $\widetilde{\boldsymbol{x}} \in \mathcal{S}^+$. Si $K \geq 2$, $\widetilde{\boldsymbol{x}}$ n'appartient pas nécessairement à $\mathcal{S}^+$. [PPZ17] propose alors soit de projeter $\widetilde{\boldsymbol{x}}$ sur $\mathcal{S}^+$ (par exemple, en fixant les deux premières composantes et en exprimant toutes les autres en fonction d'elles) puis d'inverser le changement de variables, soit de résoudre le problème suivant en discrétisant $\delta_p \in \mathcal{I}_\Delta^P$ :

$$\delta = \arg\max_{\delta_p \in \mathcal{I}_\Delta^P} \left(\widetilde{\mathbf{h}}\,\widetilde{\boldsymbol{x}}\right)^T \boldsymbol{h}(\delta_p) \tag{2.8}$$

[6] Comme défini par [KYHP14], l'interpolateur pour l'approximation (T2) s'écrit légèrement différemment : $\Psi(\delta) = [1, \delta, \delta^2/2]^T$. Ici, nous avons reporté le coefficient $\frac{1}{2}$ sur la colonne concernée de l'atome interpolé.

ce qui revient (pour $P$ élevé) à projeter l'approximation linéaire $\widetilde{\mathbf{h}}\,\widetilde{\boldsymbol{x}}$ estimée sur l'espace des atomes continu.

Les deux autres approximations linéaires (interpolation polaire et K-SVD) seront décrites plus en détails en § 2.4. Nous décrivons brièvement leurs principales caractéristiques.

### Interpolation de dictionnaire par interpolation polaire

Pour obtenir une erreur d'approximation plus faible que celle induite par les approximations de Taylor, les auteurs d'[ETS11] propose de bénéficier d'une propriété géométrique des atomes continus qu'ils considèrent : leur invariance par translation, soit $\forall \delta \in \mathcal{I}_\Delta$, $\|\boldsymbol{h}(\delta)\|_2$ est constant. Nous reviendrons plus en détails sur la définition de cette approximation en § 2.4.1. Cette approximation permet d'obtenir un interpolateur $\Psi(\delta)$ de dimension $K = 3$, tel que l'espace $\mathcal{S}^+$ décrive une portion de surface d'un cône (voir Figure 2.2). Cet espace n'est pas convexe à cause de la contrainte quadratique d'égalité $\widetilde{x}_2^2 + \widetilde{x}_3^2 = r^2\widetilde{x}_1^2$. La définition de l'espace convexe $\mathcal{H}^+$ consiste à relâcher cette contrainte en une contrainte inégalité (on obtient une contrainte inégalité quadratique de type SOCP). Notons que comme les approximations de Taylor et Minimax, la première composante $\widetilde{x}_1$ correspond à l'amplitude initiale $x$ et est donc positive, et l'espace $\mathcal{H}^+$ respecte l'implication de nullité (grâce à la contrainte quadratique d'inégalité).

Enfin, à partir de $\widetilde{\boldsymbol{x}} \in \mathcal{H}^+$, l'auteur propose d'identifier les paramètres initiaux $(x, \delta)$ en projetant $\widetilde{\boldsymbol{x}}$ sur $\mathcal{S}^+$ de la façon suivante :

$$\mathcal{P}(\widetilde{\boldsymbol{x}}) = \begin{bmatrix} \widetilde{x}_1 & r\widetilde{x}_1 \dfrac{\widetilde{x}_2}{\sqrt{\widetilde{x}_2^2 + \widetilde{x}_3^2}} & r\widetilde{x}_1 \dfrac{\widetilde{x}_3}{\sqrt{\widetilde{x}_2^2 + \widetilde{x}_3^2}} \end{bmatrix}^T \in \mathcal{S}^+ \tag{2.9}$$

puis en inversant le changement de variables : $(x, \delta) = g^{-1}(\mathcal{P}(\widetilde{\boldsymbol{x}}))$. *Notons qu'appliquer $g^{-1}$ comme défini dans le Tableau 2.1 directement à $\widetilde{\boldsymbol{x}} \in \mathcal{H}^+$ donne le même résultat.*

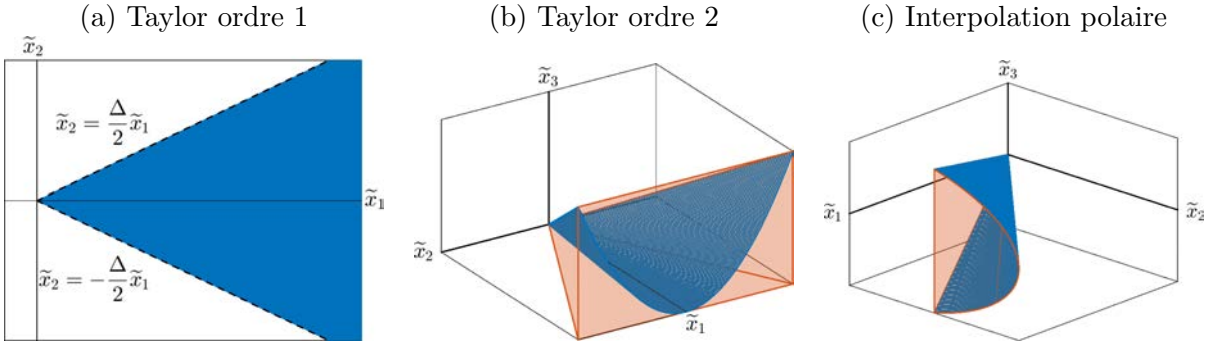



FIG. 2.2 – Illustration de l'espace image $\mathcal{S}^+$ (en bleu) et de sa relaxation convexe $\mathcal{H}^+$ (en rouge) pour les différentes interpolations de dictionnaire.

### Interpolation de dictionnaire par approximation K-SVD

Les auteurs de [KYHP14] ont proposé une approximation linéaire se basant sur la *Décomposition en Valeurs Singulières* (SVD) de la matrice $\mathbf{F}_P$, où la matrice $\mathbf{F}_P$ est constituée des $P$ colonnes $\boldsymbol{h}(\delta_p)$ avec $\delta_p \in \mathcal{I}_\Delta^P$ pour $P$ grand. Cette décomposition s'écrit $\mathbf{F}_P = \mathbf{U}_P\mathbf{S}_P\mathbf{V}_P^T$, où $\mathbf{S}_P$ est la matrice diagonale des valeurs singulières. Comme un atome $\boldsymbol{h}(\delta_p)$, colonne de $\mathbf{F}_P$, peut s'écrire en fonction de sa SVD, l'approximation par K-SVD consiste alors à ne considérer que les $K$ plus grandes valeurs singulières pour approcher $\boldsymbol{h}(\delta_p)$. Elle s'écrit alors $\boldsymbol{h}(\delta_p) \approx \sum_{k=1}^{K} z_{p,k}\boldsymbol{u}_k$,

où les $\boldsymbol{u}_k$ correspondent aux $k$ premières[7] colonnes de $\mathbf{U}$, et les $z_{p,k}$ s'expriment en fonction des valeurs singulières et des lignes $p$ de $\mathbf{V}$. Nous reviendrons plus en détails sur la définition de cette approximation en § 2.4.2.

La grande particularité de cette approximation est qu'elle s'exprime sur l'espace discret $\mathcal{I}_\Delta^P$ et non sur l'intervalle continu $\mathcal{I}_\Delta$. Par conséquent, l'espace image $\mathcal{S}^+$, construit par *dilatation* des points $\Psi(\delta_p)$, ne peut pas être convexe, puisque constitué d'hyperdroites distinctes passant par l'origine. Bien que la relaxation $\mathcal{H}^+$ de l'espace image proposées par les auteurs ressemblent aux contraintes *polyhédrales* proposées par [PPZ17] pour l'approximation Minimax, elle n'est cependant pas clairement explicitée et comporte quelques erreurs. Nous y reviendrons en § 2.4.2.

L'identification des paramètres initiaux se fait par projection de $\mathcal{H}^+$ vers $\mathcal{S}^+$ en résolvant le problème suivant :

$$(x, \delta_p) = g^{-1}(x\boldsymbol{z}_p) \quad \text{avec} \quad (x, \boldsymbol{z}_p) = \arg\min_{x,\boldsymbol{z}_p} \|\widetilde{\boldsymbol{x}} - x\boldsymbol{z}_p\|^2$$

Notons que les auteurs proposent cette approximation, car l'approximation par K-SVD est celle qui donne l'erreur d'approximation moyenne la plus faible pour une approximation de rang $K$ ; soit en considérant la matrice $\mathbf{F}_P^K (\text{dim. } N \times P)$ l'approximation par K-SVD de $\mathbf{F}$ : $\|\mathbf{F}_P - \mathbf{F}_P^K\| \leq \|\mathbf{F} - \mathbf{A}\|$, pour toute matrice $\mathbf{A}$ de rang $K$ et pour la norme de Frobenius.

*Remarque : Parmi toutes les interpolations de dictionnaire présentées, deux méthodes ressortent pour l'identification des paramètres initiaux à partir de $\widetilde{\boldsymbol{x}} \in \mathcal{H}^+$ : soit la projection sur l'espace image $\mathcal{S}^+$ pour pouvoir inverser le changement de variables, soit la projection sur l'espace des atomes continus en maximisant selon $\delta$ la corrélation entre l'approximation estimée et $\boldsymbol{h}(\delta)$. Notons que les deux approches sont acceptables, et nous ne pouvons dire laquelle est préférable. Dans la suite de ce chapitre, nous n'avons pas abordé la question, et nous supposerons que nous utilisons la première option quand cela est possible.*

### 2.1.4 Une extension aux amplitudes complexes

Jusqu'à maintenant, nous avons vu les méthodes de résolution des problèmes d'APLsc et les approximations linéaires proposées (associées à leur espace de contraintes convexes) uniquement pour des signaux à valeurs réelles et à amplitudes positives. À notre connnaissance, seul [FDJ15] a proposé une extension aux signaux à valeurs et amplitudes complexes, dans le cadre de l'optimisation convexe ; les articles travaillant dans ce cadre se basent généralement sur cette extension (par exemple [HBFR, FMC17]). *Notons d'ores et déjà que l'approche proposée par [FDJ15] est très peu justifiée et souffre de nombreuses inexactitudes ; nous la décrivons donc succinctement dans ce paragraphe et nous y reviendrons en § 2.2.*

Les auteurs proposent son extension associée à l'interpolation polaire. Sans justification théorique, ils supposent que pour $\boldsymbol{h}_j(\delta_j) \in \mathbb{C}^N$, l'interpolation polaire définie en Tableau 2.1 est toujours valable : $\boldsymbol{h}_j(\delta_j) \approx \widetilde{\mathbf{h}}_j \Psi(\delta_j)$ où $\widetilde{\mathbf{h}}_j \in \mathbb{C}^{N\times 3}$ et $\Psi(\delta_j) \in \mathbb{R}^3$. Pour pouvoir résoudre le problème d'APLsc avec des contraintes convexes, les auteurs réécrivent $\widetilde{\boldsymbol{x}}_j = x_j \Psi(\delta_j) \in \mathbb{C}^3$ comme la somme des parties réelle et imaginaire, positive et négative :

$$\widetilde{\boldsymbol{x}}_j = \widetilde{\boldsymbol{x}}_j^{r+} - \widetilde{\boldsymbol{x}}_j^{r-} + i\,\widetilde{\boldsymbol{x}}_j^{i+} - i\,\widetilde{\boldsymbol{x}}_j^{i-} \quad \text{avec} \quad \widetilde{\boldsymbol{x}}_j^{\cdot\cdot} \in \mathbb{R}_+^3$$

Le problème revient donc à estimer l'ensemble des amplitudes $\{\widetilde{\boldsymbol{x}}^{\cdot\cdot}\}_{j,\cdot\cdot}$ que les auteurs contraignent aux espaces convexes $\mathcal{H}_j^+ \cap \mathbb{R}_+^3$, avec $\mathcal{H}_j^+$ défini en Tableau 2.1. Les auteurs proposent alors de résoudre le problème d'APLsc en utilisant la formulation pénalisée par la norme $\ell_1$ :

$$\min_{\widetilde{\boldsymbol{x}}^{ri+-} \in \mathcal{C}} \left\| \boldsymbol{y} - \widetilde{\mathbf{H}}^{ri+-} \widetilde{\boldsymbol{x}}^{ri+-} \right\|^2 + \lambda \|\boldsymbol{\alpha}^{ri+-}\|_{1,2} \tag{2.10}$$

[7] Si les valeurs singulières sont rangées dans l'ordre décroissant.

$$\begin{array}{rrcl} \text{où} & \widetilde{\mathbf{H}}^{ri+-} & = & [\widetilde{\mathbf{H}}, -\widetilde{\mathbf{H}}, i\widetilde{\mathbf{H}}, -i\widetilde{\mathbf{H}}] \\ & \widetilde{\boldsymbol{x}}^{ri+-} & = & [\widetilde{\boldsymbol{x}}^{r+T}, \widetilde{\boldsymbol{x}}^{r-T}, \widetilde{\boldsymbol{x}}^{i+T}, \widetilde{\boldsymbol{x}}^{i-T}]^T \\ & \widetilde{\boldsymbol{x}}^{ri+-} \in \mathcal{C} & \Leftrightarrow & \forall (j, \cdot\cdot),\ \widetilde{\boldsymbol{x}}_j^{\cdot\cdot} \in \mathcal{H}_j^+ \cap \mathbb{R}_+^3 \end{array}$$

Une norme mixte $\ell_{1,2}$ est utilisée sur les premières composantes pour favoriser leur mise à zéro simultanée. Cette formulation est nommée *Complex Continuous Basis Pursuit* (CCBP), par extension du CBP proposé par [ETS11].

Une fois qu'une estimation des $\widetilde{\boldsymbol{x}}^{ri+-} \in \mathcal{C}$ est obtenue par la résolution du problème ci-dessus, l'auteur ne précise pas comment reconstruire les paramètres initiaux $(x_j, \delta_j)$, mais on peut déduire la méthode employée à partir des codes mis à disposition de la communauté[8]. Nous y reviendrons en § 2.2.3.

### 2.1.5 Bilan

Dans cette section, nous avons donc pu exposer le principe de l'interpolation de dictionnaire pour la reformulation en APLsc des problèmes inverses non linéaires parcimonieux, et poser un ensemble de définitions et de notations qui nous serviront dans la suite de ce chapitre.

Notons déjà que nous n'approfondirons pas les approches gloutonnes pour la résolution du problème d'APLsc, ni les méthodes d'identification des paramètres initiaux à partir des $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+$ estimés. Pour cela, nous nous référons à la remarque effectuée en fin de § 2.1.3.

Pour rappel, l'objectif de nos travaux est de résoudre les problèmes inverses parcimonieux non linéaires présentés en § 1.1 de façon « optimale », notamment dans le sens où la parcimonie doit être prise de façon exacte. Les travaux sur l'interpolation de dictionnaire, tels que présentés dans cette section, présentent donc pour le moment des verrous importants pour atteindre notre objectif.

Un premier verrou est l'exploitation de l'interpolation de dictionnaire de façon correcte et justifiée pour des atomes continus à amplitudes réelles ou complexes ($\mathcal{D}_x = \mathbb{F}$). En effet, les travaux ont d'abord été proposés dans le cas où $\mathcal{D}_x = \mathbb{R}_+$ par [ETS11] et ont ainsi pu être utilisés dans ce cadre (voir par exemple [MMSC]). Cependant, l'extension aux amplitudes réelles ou complexes proposée par [FDJ15] est très peu justifiée, et nous lui avons trouvé certaines limites lors de notre étude. Par ailleurs, les auteurs de [DP17], dans leur étude des propriétés du CBP de [ETS11] avec approximation de Taylor dans le cas $\mathcal{D}_x = \mathbb{R}_+$, précisent bien que le passage à $\mathcal{D}_x = \mathbb{F}$ demande une étude plus approfondie. Ces derniers proposent l'extension de [FCPC13] au cas $\mathcal{D}_x = \mathbb{C}$ ; cependant, cette dernière nécessite l'utilisation d'optimisation non convexe et non lisse, et ne rentre donc pas dans le cadre de nos travaux. Par conséquent, nous étudierons en § 2.2, de façon plus approfondie, l'extension au cas $\mathcal{D}_x = \mathbb{F}$ dans le cadre de l'optimisation convexe.

Ensuite, la principale limite réside dans les méthodes de résolution du problème d'APLsc, de type pénalisation $\ell_1$ ou approches gloutonnes. Ces approches sous-optimales ne prennent pas en compte de façon exacte la parcimonie. [FDJ15] relevait d'ailleurs des limites à ces méthodes dans le cas de l'application à des problèmes d'estimation de fréquences ; bien que l'approximation linéaire permettrait en théorie de choisir un dictionnaire interpolé moins dense ($J$ plus faible que pour un problème d'APL) et donc moins corrélé, il est tout de même nécessaire de limiter la taille des intervalles $\Delta$ afin que l'erreur d'approximation ne soit pas trop élevée. Le dictionnaire interpolé peut donc montrer une corrélation suffisamment élevée pour mettre en échec ces méthodes sous-optimales. Afin de prendre en compte de façon exacte la parcimonie, nous proposerons donc une adaptation de la reformulation en MIP du problème exploitant la pseudo-norme

[8] Voir à l'adresse suivante `www.sparsesampling.com/cpe`, la fonction `ccbp.m`.

$\ell_0$, en l'adaptant à la présence des contraintes, et aux extensions à $\mathcal{D}_x = \mathbb{F}$ effectuées en § 2.2. Il s'agira de la principale contribution de ce chapitre. Notons que nous reviendrons aussi sur la résolution par pénalisation $\ell_1$, étendue à $\mathcal{D}_x = \mathbb{F}$. L'ensemble de ces travaux seront présentés en § 2.3.

Enfin, nous avons vu de nombreuses propositions d'approximations linéaires, mais la caractérisation de celle qui permet la meilleure détection du support n'est pas évidente. En Figure 2.3, nous avons récupéré des illustrations des erreurs d'approximation[9] à partir de [ETS11, PPZ17]. L'interpolation polaire, qui nous semble être la plus utilisée (par exemple [MMSC, FMC17, MD16]), donne généralement une erreur d'approximation moyenne faible (davantage que les approximations de Taylor), mais l'erreur maximale est plus élevée que celle des approximations Minimax ou par K-SVD. Nous verrons tout au long de ce chapitre que l'erreur d'approximation n'est pas le seul critère de choix de l'approximation linéaire. Nous reviendrons davantage sur l'interpolation polaire et l'approximation par K-SVD en § 2.4. Pour la première, nous reviendrons sur la définition de l'approximation, notamment pour $\boldsymbol{h}_j(\delta_j) \in \mathbb{C}^N$ et nous proposerons un espace de contraintes convexe défini à partir de contraintes linéaires. Pour la seconde, nous reviendrons sur la définition de l'espace des contraintes et sur son caractère discret, afin de définir une variante continue. L'approximation Minimax semble intéressante mais sa proposition étant récente, nous n'avons pas pu l'étudier plus en détails. Notons cependant qu'elle minimise bien l'erreur maximale d'approximation, *mais pour un interpolateur* $\Psi(\cdot)$ *donné*, ce qui est un caractère limitant.

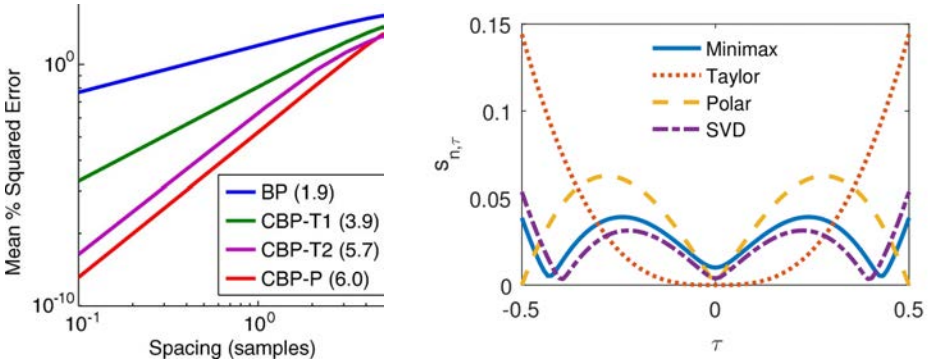



FIG. 2.3 – Illustration des erreurs d'approximation pour différentes approximations. À gauche : erreur d'approximation moyenne en fonction de la taille $\Delta$ des intervalles, pour l'approximation discrète/APL (BP), les approximations de Taylor à l'ordre 1 (T1) et 2 (T2) et pour l'interpolation polaire (P). Crédit : [ETS11]. À droite : erreur d'approximation en fonction de $\tau = \delta \in \mathcal{I}_\Delta$ pour un intervalle donné, et pour les approximations Minimax, Taylor à l'ordre 1, l'interpolation polaire et l'approximation par K-SVD. Crédit : [PPZ17].

Après avoir testé nos différentes contributions sur les données tests de déconvolution impulsionnelle en § 2.5, nous essaierons d'étendre le cadre général de l'interpolation de dictionnaire au cas de la présence de paramètres de forme ou dans le cas d'amplitudes $\boldsymbol{x}_j$ vectorielles pour l'appliquer au problème d'analyse spectrale en échantillonnage irrégulier.

[9] Notons que ces erreurs d'approximation sont à considérer pour un modèle $\boldsymbol{h}_j(\delta_j)$ donné.

## 2.2 Étude de l'espace des contraintes

Dans cette section, nous allons étudier l'espace des contraintes pour une approximation linéaire quelconque $x_j \boldsymbol{h}_j(\delta_j) = \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j$ avec $\widetilde{\boldsymbol{x}}_j = g(x_j, \delta_j) \in \mathcal{S}_j$. Notre principal objectif est d'étendre *de façon correcte et justifiée* la définition de l'espace des contraintes aux cas où les atomes continus sont à amplitudes réelles ($x_j \in \mathcal{D}_x = \mathbb{R}$) puis à valeurs complexes ($x_j \in \mathcal{D}_x = \mathbb{C}$). En effet, nous verrons que la solution proposée par [FDJ15], non justifiée, peut facilement être mise en échec. De plus, nous insisterons sur le fait que l'espace des contraintes convexe doit être le plus proche possible de l'espace image.

### 2.2.1 Cadre général

Dans un cadre général, on considère, pour un indice $j$, que l'atome continu $\boldsymbol{h}_j(\delta_j)$ ainsi que son amplitude $x_j$ peuvent être à valeurs réelles ou complexes : $x_j \in \mathcal{D}_x = \mathbb{F}$ et $\mathbf{h}_j(\delta_j) \in \mathbb{F}^N$. On considère donc l'approximation linéaire suivante :

$$x_j \, \boldsymbol{h}_j(\delta_j) \approx \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j \quad \text{avec} \quad \widetilde{\boldsymbol{x}}_j = x_j \Psi(\delta_j) = g(x_j, \delta_j) \quad \text{et} \quad g \; : \; \begin{cases} (\mathbb{F} \times \mathcal{I}_\Delta) & \longrightarrow & \mathcal{S}_j \subset \mathbb{F}^K \\ (x_j, \delta_j) & \longmapsto & \widetilde{\boldsymbol{x}}_j = x_j \Psi(\delta_j) \end{cases}$$

Dans le cas où l'espace image $\mathcal{S}_j$ n'est pas convexe, il est nécessaire d'en définir une relaxation convexe[10] $\mathcal{H}_j \supset \mathcal{S}_j$ pour définir l'espace des contraintes convexes $\mathcal{C}_j$. Une telle relaxation ne contraint plus le signal $\widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j$ avec $\widetilde{\boldsymbol{x}}_j \in \mathcal{C}_j$ à être une approximation du signal élémentaire $x_j \, \boldsymbol{h}_j(\delta_j)$. Comme ces contraintes compliquent la résolution du problème (puisqu'il faut adapter les méthodes classiques d'APL), une question se pose : quitte à choisir une relaxation convexe de $\mathcal{S}_j$, pourquoi ne pas choisir l'espace $\mathbb{R}^K$ tout entier et retourner à un classique problème d'APL ? Pour répondre à cette question, observons l'erreur quadratique minimale entre un signal composé uniquement d'un signal élémentaire $\boldsymbol{y} = x_j \, \boldsymbol{h}_j(\delta_j)$ fixé et une approximation linéaire $\widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j$ pour $\widetilde{\boldsymbol{x}}_j$ appartenant à différents espaces de contraintes ; comme $x_j \Psi(\delta_j) \in \mathcal{S}_j \subseteq \mathcal{H}_j \subseteq \mathbb{R}^K$, on obtient la relation suivante :

$$\min_{\widetilde{\boldsymbol{x}}_j \, \in \, \mathbb{R}^K} \|\boldsymbol{y} - \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j\|^2 \leq \min_{\widetilde{\boldsymbol{x}}_j \, \in \, \mathcal{H}_j} \|\boldsymbol{y} - \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j\|^2 \leq \min_{\widetilde{\boldsymbol{x}}_j \, \in \, \mathcal{S}_j} \|\boldsymbol{y} - \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j\|^2 \leq \|\boldsymbol{y} - x_j \, \widetilde{\mathbf{h}}_j \Psi(\delta_j))\|^2$$

Ainsi, plus on relâche l'espace des contraintes pour $\widetilde{\boldsymbol{x}}_j$, meilleure sera l'approximation, mais plus le $\widetilde{\boldsymbol{x}}_j$ optimal a de chance de s'éloigner de $x_j \Psi(\delta_j)$, et donc les $(x_j, \delta_j)$ identifiés de s'éloigner des $(x_j, \delta_j)$ initiaux. Pour le type de problèmes auxquels on s'intéresse, c'est bien la *détection et localisation* des événements que l'on recherche, et non une bonne approximation des signaux. De ces observations, nous concluons la propriété suivante :

*La relaxation convexe $\mathcal{H}_j$ de l'espace image $\mathcal{S}_j$ doit être la plus proche possible de $\mathcal{S}_j$.*

Dans l'idéal, on souhaiterait donc que l'espace image $\mathcal{S}_j$ soit convexe. Nous avons vu en § 2.1 que parmi les approximations proposées dans la littérature, seule l'approximation de Taylor à l'ordre 1 et l'approximation Minimax[11] pour $K = 2$ ne nécessitaient pas de relaxation convexe[12]. Cela n'est pas un hasard. Si l'on note $K_0$ la dimension du domaine de définition $\mathcal{D}_x \times \mathcal{I}_\Delta$ ($K_0 = 2$ pour $\mathcal{D}_x = \mathbb{R}$ et $K_0 = 3$ pour $\mathcal{D}_x = \mathbb{C}$), alors l'espace image $\mathcal{S}_j$ sera au maximum de même

[10] Dans le cas où $\mathcal{D}_x = \mathbb{C}$ et donc où $\mathcal{S}_j \subset \mathbb{C}^K$, on considère $\mathbb{C}$ comme un $\mathbb{R}$-espace vectoriel de dimension 2 (paramétrage parties réelle/partie imaginaire, ou module/phase par exemple), sur lequel la convexité a un sens.

[11] Pour $K = 2$, l'espace des contraintes défini par l'auteur correspond à l'espace image.

[12] L'approximation par K-SVD étant discrète, une relaxation est nécessaire pour obtenir un espace des contraintes continu, même pour $K = 2$.

dimension ; nous avons d'ailleurs vu que l'espace image $\mathcal{S}_j$ correspondait à une surface (dimension $2 = K_0$) dans un volume (dimension 3) pour les approximations de Taylor 2 et polaire (voir Figure 2.2). Nous émettons alors les deux postulats suivants :

1. *Si $K_0 > K$, la fonction $g$ ne peut pas être bijective.*
2. *Si $K_0 < K$, l'espace image $\mathcal{S}_j$ a peu de chance d'être convexe, sauf s'il appartient à un « hyperplan » passant par l'origine.*

Pour obtenir un espace image convexe, il est donc préférable de choisir un interpolateur $\Psi(\cdot)$ appartenant à un espace de même dimension que le domaine de définition de $g$, soit dans notre cas, $K = 2$.

Forts de ces remarques sur la définition de l'espace des contraintes convexes dans un cadre général, nous allons maintenant étudier l'extension aux cas des amplitudes à valeurs réelles et complexes.

### 2.2.2 Cas des amplitudes à valeurs réelles

Dans ce paragraphe, nous considérons maintenant l'espace image de $\mathbb{R}\times\mathcal{I}_\Delta$ par le changement de variables $g$, que nous notons $\mathcal{S}_j^{\mathbb{R}}$. Nous définissons les espaces $\mathcal{S}_j^+ = g(\mathbb{R}^+\times\mathcal{I}_\Delta)$ et $\mathcal{S}_j^- = g(\mathbb{R}^-\times\mathcal{I}_\Delta)$, respectivement les espaces images pour les amplitudes à valeurs positives et négatives. Nous commençons par montrer en quoi obtenir un espace convexe à partir de $\mathcal{S}_j^{\mathbb{R}}$ est difficile, puis nous décrivons la stratégie choisie.

#### Enveloppe convexe de l'espace $\mathcal{S}_j$ trop étendue

De par la définition de l'espace $\mathcal{S}_j$, nous avons les propriétés suivantes :

$$
\begin{array}{cccl}
 & \mathcal{S}_j^{\mathbb{R}} & = & \left\{\widetilde{\boldsymbol{x}}_j = x_j\Psi(\delta_j) \in \mathbb{R}^K,\ x_j \in \mathbb{R},\ \delta_j \in \mathcal{I}_\Delta\right\} \\
 & & \Downarrow & \\
(1) & \mathcal{S}_j^{\mathbb{R}} & = & \mathcal{S}_j^+ \cup \mathcal{S}_j^- \\
(2) & \mathcal{S}_j^- & = & \left\{\widetilde{\boldsymbol{x}}_j = x_j\Psi(\delta_j) \in \mathbb{R}^K,\ x_j \in \mathbb{R}^-,\ \delta_j \in \mathcal{I}_\Delta\right\} \\
 & & = & \left\{\widetilde{\boldsymbol{x}}_j = -x_j\Psi(\delta_j) \in \mathbb{R}^K,\ x_j \in \mathbb{R}^+,\ \delta_j \in \mathcal{I}_\Delta\right\}
\end{array}
$$

L'espace $\mathcal{S}_j^-$ se déduit donc de $\mathcal{S}_j^+$ par symétrie centrale. En section § 2.1, nous avons vu pour $\mathcal{D}_x = \mathbb{R}_+$ des exemples d'approximations linéaires où l'espace $\mathcal{S}_j^+$ est convexe (Taylor 1) et où ce dernier n'est pas convexe mais où une relaxation convexe $\mathcal{H}_j^+$ peut être définie sans difficulté (Taylor 2, Polaire). En Figure 2.2, nous avons observé l'espace $\mathcal{S}_j^+$ des différentes approximations linéaires ; nous observons maintenant l'espace $\mathcal{S}_j^{\mathbb{R}}$ en Figure 2.4. Dans tous les cas, nous faisons les remarques suivantes :

1. l'intersection de $\mathcal{S}_j^+$ et de $\mathcal{S}_j^-$ ne comprend que $\mathbf{0}$ : $\mathcal{S}_j^+ \cap \mathcal{S}_j^- = \{\mathbf{0}\}$ ;
2. l'espace $\mathcal{S}_j^{\mathbb{R}}$ n'est pas convexe.

Pour ces exemples d'approximation, il serait donc nécessaire de définir une relaxation convexe $\mathcal{H}_j^{\mathbb{R}}$ de $\mathcal{S}_j^{\mathbb{R}}$, même pour l'approximation de Taylor à l'ordre 1. Cependant, comme on peut le constater sur la Figure 2.4, l'enveloppe convexe de $\mathcal{S}_j^{\mathbb{R}}$ correspond à l'espace $\mathbb{R}^K$ tout entier ; une telle relaxation est en contradiction avec le conseil que nous donnons en § 2.2. Une autre stratégie doit donc être adoptée.

(a) Taylor ordre 1 (b) Taylor ordre 2 (c) Interpolation polaire

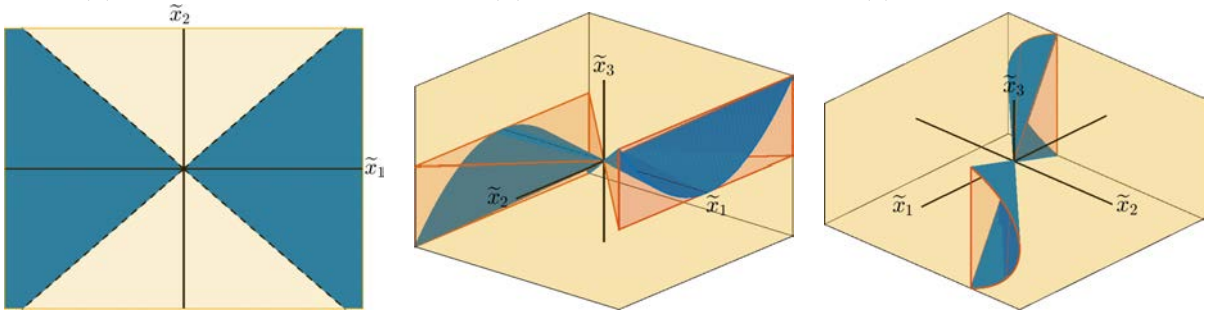



FIG. 2.4 – Illustration de l'espace image $\mathcal{S}_j$ (en bleu) et des relaxations convexes $\mathcal{H}_j^+$ et $\mathcal{H}_j^-$ (en rouge) pour les différentes approximations linéaires. La relaxation convexe $\mathcal{H}_j$ de $\mathcal{S}_j$ est représentée en jaune.

Nous faisons les postulats suivants pour généraliser nos observations à un cas quelconque d'approximation linéaire :

$$\begin{array}{lrcl} (1) & \mathcal{S}_j^+ \cap \mathcal{S}_j^- = \{\mathbf{0}\} & \Rightarrow & \mathcal{H}_j^{\mathbb{R}} = \mathbb{R}^K \quad \text{(ou un hyperplan de } \mathbb{R}^K) \\ (2) & g \text{ bijective sur } \mathbb{R} \times \mathcal{I}_\Delta & \Rightarrow & \mathcal{S}_j^+ \cap \mathcal{S}_j^- = \{\mathbf{0}\} \\ (3) & \exists k, \forall \delta_j \in \mathcal{I}_\Delta,\ \Psi_k(\delta_j) = \text{constante} & \Rightarrow & \mathcal{S}_j^+ \cap \mathcal{S}_j^- = \{\mathbf{0}\} \end{array}$$

Le premier postulat affirme que la disjonction des espaces $\mathcal{S}_j^+$ et $\mathcal{S}_j^-$ entraîne toujours que l'enveloppe convexe de $\mathcal{S}_j^{\mathbb{R}}$ soit l'espace $\mathbb{R}^K$ tout entier (sauf si $\mathcal{S}_j^{\mathbb{R}}$ git dans un hyperplan de $\mathbb{R}^K$). Les deuxième et troisième postulats donnent des conditions suffisantes pour obtenir cette disjonction.

Le postulat (2) est important ; il est souhaitable que la fonction $g$ soit bijective sur son domaine de définition, afin que $\widetilde{\boldsymbol{x}}_j \in \mathcal{S}_j^{\mathbb{R}}$ identifie un unique $(x_j, \delta_j)$. Cela signifie que pour toute approximation linéaire « idéale » ($g$ bijective), la définition d'un espace de contraintes convexe proche de $\mathcal{S}_j^{\mathbb{R}}$ est problématique.

La plupart des approximations linéaires présentées en § 2.1.3 sont dans le cas du postulat (3) : la première composante de l'interpolateur est en effet égale à 1 pour les approximations de Taylor, Minimax et Polaire. Il est intéressant de remarquer que cette propriété est suffisante pour induire que les espaces $\mathcal{S}_j^+$ et $\mathcal{S}_j^-$ sont disjoints (le signe de $\widetilde{x}_{j,k} = x_j \Psi_k(\delta_j)$ est du signe de $\Psi_k(\cdot)$ si $x_j > 0$ et du signe inverse si $x_j < 0$).

*Remarque : Au regard de ces postulat, nous avons l'intuition que quelle que soit l'approximation linéaire employée, l'enveloppe convexe $\mathcal{H}_j^{\mathbb{R}}$ de $\mathcal{S}_j^{\mathbb{R}}$ s'étendra dans tout l'espace auquel il appartient (soit $\mathbb{R}^K$ quand la courbe $\Psi(\mathcal{I}_\Delta)$ n'appartient pas à un hyperplan passant par l'origine).*

### Solution : Dédoublement des variables $\widetilde{\boldsymbol{x}}_j$

Pour définir un espace de contraintes convexe n'englobant pas tout l'espace $\mathbb{R}^K$, nous suivons l'idée de [FDJ15] qui a proposé de remplacer la recherche d'un $\widetilde{\boldsymbol{x}}_j$ dans $\mathcal{H}_j$ (probablement égal à $\mathbb{R}^K$) par celle d'un $\widetilde{\boldsymbol{x}}_j$ dans $\mathcal{H}_j^+$ *ou* dans $\mathcal{H}_j^-$ (deux espaces convexes n'englobant pas tout $\mathbb{R}^K$). Pour définir l'espace des contraintes associés à cette recherche, on est alors amené à dédoubler les variables $\widetilde{\boldsymbol{x}}_j$ et les atomes $\widetilde{\mathbf{h}}_j$ de la façon suivante :

$$\widetilde{\mathbf{h}}_j^{+-} = \begin{bmatrix} \widetilde{\mathbf{h}}_j & -\widetilde{\mathbf{h}}_j \end{bmatrix} \quad \text{et} \quad \widetilde{\boldsymbol{x}}_j^{+-} = \begin{bmatrix} \widetilde{\boldsymbol{x}}_j^+ \\ \widetilde{\boldsymbol{x}}_j^- \end{bmatrix} \quad \text{où} \quad (\widetilde{\boldsymbol{x}}_j^+, \widetilde{\boldsymbol{x}}_j^-) \in {\mathcal{H}_j^+}^2$$

Le modèle s'écrit alors :

$$\widetilde{\mathbf{h}}_j^{+-}\widetilde{\boldsymbol{x}}_j^{+-} = \widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j^{+} + (-\widetilde{\mathbf{h}}_j)\widetilde{\boldsymbol{x}}_j^{-} = \widetilde{\mathbf{h}}_j \underbrace{\widetilde{\boldsymbol{x}}_j^{+}}_{\in\mathcal{H}_j^+} + \widetilde{\mathbf{h}}_j \underbrace{(-\widetilde{\boldsymbol{x}}_j^{-})}_{\in\mathcal{H}_j^-}$$

Ce modèle nous permet bien de rechercher un $\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j$ avec $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+$ et $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^-$. Cependant, il est nécessaire de contraindre au moins l'un des deux $\widetilde{\boldsymbol{x}}_j$ *à être nul*, et ce pour deux raisons :

1. On cherche une unique approximation $\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j$ avec $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+ \cup \mathcal{H}_j^-$.
2. Le modèle peut se réécrire $\widetilde{\mathbf{h}}_j^{+-}\widetilde{\boldsymbol{x}}_j^{+-} = \widetilde{\mathbf{h}}_j(\widetilde{\boldsymbol{x}}_j^{+} - \widetilde{\boldsymbol{x}}_j^{-})$ et est donc équivalent à une seule approximation $\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j$ avec $\widetilde{\boldsymbol{x}}_j = \widetilde{\boldsymbol{x}}_j^{+} - \widetilde{\boldsymbol{x}}_j^{-}$. Cependant, rien ne garantit que $(\widetilde{\boldsymbol{x}}_j^{+} - \widetilde{\boldsymbol{x}}_j^{-})$ appartiennent à $\mathcal{H}_j^+ \cup \mathcal{H}_j^-$, notamment car cet espace est non convexe[13]. Contraindre au moins l'une des deux amplitudes à 0 garantit que l'amplitude équivalente appartienne à l'un des deux espaces convexes.

Par conséquent, l'espace des contraintes associés au modèle $\widetilde{\mathbf{h}}_j^{+-}\widetilde{\boldsymbol{x}}_j^{+-}$ s'écrit de la façon suivante :

$$\mathcal{C}_j = \left\{ \widetilde{\boldsymbol{x}}_j^{+-} = \begin{bmatrix} \widetilde{\boldsymbol{x}}_j^{+} \\ \widetilde{\boldsymbol{x}}_j^{-} \end{bmatrix},\ (\widetilde{\boldsymbol{x}}_j^{+}, \widetilde{\boldsymbol{x}}_j^{-}) \in {\mathcal{H}_j^+}^2 \text{ et } (\widetilde{\boldsymbol{x}}_j^{+} \text{ ou } \widetilde{\boldsymbol{x}}_j^{-}) = \mathbf{0} \right\} \tag{2.11}$$

À cause de la contrainte ($\widetilde{\boldsymbol{x}}_j^{+}$ ou $\widetilde{\boldsymbol{x}}_j^{-} = \mathbf{0}$), que nous nommons « contrainte de disjonction », l'espace $\mathcal{C}_j$ n'est pas un espace de contraintes convexe comme nous l'avons défini en (2.4). Nous verrons en § 2.3 qu'une formulation du problème d'APLsc avec la pénalisation $\ell_1$ rend impossible la prise en compte d'une telle contrainte dans le cadre d'un problème d'optimisation convexe. Cependant, la méthode de résolution faisant appel au MIP (programmation mixte en nombres entiers) que nous proposerons pourra prendre en compte une telle contrainte de façon linéaire.

*Remarque : Dans la formulation du problème d'APLsc sur cet espace de contrainte, la parcimonie devra être imposée de façon structurée sur l'ensemble des composantes de* $\widetilde{\boldsymbol{x}}_j^{+-}$.

Enfin, ce dédoublement de variables ressemble à celui proposé par [FDJ15] dans le cas de signaux à valeurs complexes si l'on se restreint aux parties réelles. Cependant, il diffère en deux points particulièrement importants :

1. Bien que la contrainte de disjonction soit difficile à prendre en compte dans la résolution du problème d'optimisation convexe avec pénalisation $\ell_1$ que l'auteur veut résoudre, ce dernier n'en parle pas alors qu'elle nous semble cependant capitale ;
2. L'auteur contraint l'ensemble des composantes de $(\widetilde{\boldsymbol{x}}_j^{+}, \widetilde{\boldsymbol{x}}_j^{-})$ à être positives, soit

   $$(\widetilde{\boldsymbol{x}}_j^{+}, \widetilde{\boldsymbol{x}}_j^{-}) \in (\mathcal{H}_j^+ \cap \mathbb{R}_+^K)^2$$

   Cette contrainte tronque une partie de l'espace $\mathcal{H}_j^+$, ce qui rend impossible la reconstruction de certaines valeurs de $\delta_j$. En effet, le fait que $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+$ contraint généralement sa première composante à être positive, mais les autres peuvent être de signe quelconque. Par exemple, dans le cas de l'interpolation polaire qu'il emploie, la troisième composante de $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+$ peut s'écrire $\widetilde{x}_{j,3} = x_j r_j \sin(\frac{\Delta}{2\theta_j}\delta_j)$ ; la contraindre à être positive revient à exclure l'ensemble des $\delta_j$ négatifs.

Notre approche nous semble donc plus cohérente.

[13] En effet, on peut facilement démontrer que pour tous les espaces relâchés $\mathcal{H}_j^+$ présentés en § 2.1.3, l'union $\mathcal{H}^+ \cup \mathcal{H}^- = \mathcal{H}^+ \cup (-\mathcal{H}^+)$ n'est pas convexe, tout comme $\mathcal{S}^+ \cup \mathcal{S}^-$.

### 2.2.3 Cas des amplitudes à valeurs complexes

Dans ce paragraphe, nous considérons maintenant l'espace image de $\mathbb{C}\times\mathcal{I}_\Delta$ par le changement de variables $g$ que nous notons $\mathcal{S}_j^{\mathbb{C}}$. Nous conservons les notations $\mathcal{S}_j^{\mathbb{R}} = g(\mathbb{R}\times\mathcal{I}_\Delta)$ et $\mathcal{S}_j^+$ et $\mathcal{S}_j^-$.

Dans le cadre des problèmes d'optimisation, nous sommes amenés à considérer des variables réelles uniquement ; l'espace des complexes $\mathbb{C}$ doit donc être considéré comme un $\mathbb{R}$-espace vectoriel de dimension 2. Afin de conserver la linéarité de modèle $\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j$, il est nécessaire de considérer la paramétrisation en parties réelle/partie imaginaire plutôt que celle en module/phase ; ainsi, si on écrit $\widetilde{\boldsymbol{x}}_j = \widetilde{\boldsymbol{x}}_j^r + i\,\widetilde{\boldsymbol{x}}_j^i$ (avec $(\widetilde{\boldsymbol{x}}_j^r, \widetilde{\boldsymbol{x}}_j^i) \in \mathbb{R}^{2K}$)), le modèle reste linéaire : $\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j = \widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j^r + i\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j^i$.

#### Parties réelles et imaginaires liées

Nous avons vu en § 2.1.3 que les interpolateurs des approximations présentées dépendent soit uniquement de $\delta$ (approximations de Taylor), soit aussi de paramètres à valeurs réelles ($(r_j, \theta_j)$ pour l'interpolation polaire pour des atomes complexes proposées par [FDJ15]). Dans tous les cas, l'interpolateur est à valeurs réelles : $\Psi(\delta) \in \mathbb{R}^K$, même si l'atome continu approché est à valeurs complexes Par conséquent, les parties réelle et imaginaire d'un $\widetilde{\boldsymbol{x}}_j \in \mathcal{S}_j$ appartiennent toutes deux à $\mathcal{S}_j^{\mathbb{R}}$ :

$$\text{Si } \widetilde{\boldsymbol{x}}_j = x_j\,\Psi(\delta_j) \in \mathcal{S}_j \quad \text{alors} \quad \begin{aligned} \widetilde{\boldsymbol{x}}_j^r &= \mathfrak{Re}(\widetilde{\boldsymbol{x}}_j) = \mathfrak{Re}(x_j)\Psi(\delta_j) \in \mathcal{S}_j^{\mathbb{R}} \\ \widetilde{\boldsymbol{x}}_j^i &= \mathfrak{Im}(\widetilde{\boldsymbol{x}}_j) = \mathfrak{Im}(x_j)\Psi(\delta_j) \in \mathcal{S}_j^{\mathbb{R}} \end{aligned}$$

Cependant, la réciproque est fausse : $(\widetilde{\boldsymbol{x}}_j^r, \widetilde{\boldsymbol{x}}_j^i) \in \mathcal{S}_j^{\mathbb{R}^2}$ n'impliquent pas que $\widetilde{\boldsymbol{x}}_j \in \mathcal{S}_j^{\mathbb{C}}$ : en effet, *les parties réelle et imaginaire d'un $\widetilde{\boldsymbol{x}}_j \in \mathcal{S}_j^{\mathbb{C}}$ sont nécessairement liées, puisqu'elles sont multiples d'un même vecteur* $\Psi(\delta_j)$, ce qui n'est pas nécessairement le cas de $(\widetilde{\boldsymbol{x}}_j^r, \widetilde{\boldsymbol{x}}_j^i)$ pris au hasard dans $\mathcal{S}_j^{\mathbb{R}^2}$. Ainsi, on peut en déduire que le modèle $\widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j$ avec $\widetilde{\boldsymbol{x}}_j \in \mathcal{S}_j^{\mathbb{C}}$ est équivalent au modèle suivant :

$$\widetilde{\mathbf{h}}_j^{ri}\widetilde{\boldsymbol{x}}_j^{ri}, \text{ avec } \widetilde{\mathbf{h}}_j^{ri} = \begin{bmatrix}\widetilde{\mathbf{h}}_j & i\,\widetilde{\mathbf{h}}_j\end{bmatrix} \text{ et } \widetilde{\boldsymbol{x}}_j^{ri} \in \left\{ \widetilde{\boldsymbol{x}}_j^{ri} = \begin{bmatrix}\widetilde{\boldsymbol{x}}_j^r \\ \widetilde{\boldsymbol{x}}_j^i\end{bmatrix},\ (\widetilde{\boldsymbol{x}}_j^r, \widetilde{\boldsymbol{x}}_j^i) \in \mathcal{S}_j^{\mathbb{R}^2},\ \widetilde{\boldsymbol{x}}_j^r \propto \widetilde{\boldsymbol{x}}_j^i \right\} \tag{2.12}$$

où la contrainte $\widetilde{\boldsymbol{x}}_j^r \propto \widetilde{\boldsymbol{x}}_j^i$, que nous nommons contrainte de liaison, indique que parties réelle et imaginaire doivent être proportionnelles au même vecteur.

#### Définition de l'espace de contraintes

Cependant, l'espace associé au modèle $\widetilde{\mathbf{h}}_j^{ri}\widetilde{\boldsymbol{x}}_j^{ri}$ défini en équation (2.12) ne peut être choisi comme espace des contraintes. Nous avons en effet vu précédemment que l'espace $\mathcal{S}_j^{\mathbb{R}}$ n'est généralement pas convexe et que son enveloppe convexe revient souvent à $\mathbb{R}^K$. Nous proposons donc comme [FDJ15] de faire appel au même dédoublement en parties positives et négatives des parties réelle et imaginaire, afin de pouvoir considérer l'espace $\mathcal{S}_j^+$ et sa relaxation convexe $\mathcal{H}_j^+$ si nécessaire. On considère donc le modèle suivant :

$$\widetilde{\mathbf{h}}_j^{ri+-} = \begin{bmatrix}\widetilde{\mathbf{h}}_j & -\widetilde{\mathbf{h}}_j & i\,\widetilde{\mathbf{h}}_j & -i\,\widetilde{\mathbf{h}}_j\end{bmatrix} \quad \text{et} \quad \widetilde{\boldsymbol{x}}_j^{ri+-} = \begin{bmatrix}\widetilde{\boldsymbol{x}}_j^{r+} \\ \widetilde{\boldsymbol{x}}_j^{r-} \\ \widetilde{\boldsymbol{x}}_j^{i+} \\ \widetilde{\boldsymbol{x}}_j^{i-}\end{bmatrix} \quad \text{où} \quad (\widetilde{\boldsymbol{x}}_j^{r+}, \widetilde{\boldsymbol{x}}_j^{r-}, \widetilde{\boldsymbol{x}}_j^{i+}, \widetilde{\boldsymbol{x}}_j^{i-}) \in \mathcal{H}_j^{+4}$$

auquel on associe l'espace des contraintes suivant :

$$\mathcal{C}_j = \left\{ \widetilde{\boldsymbol{x}}_j^{ri+-} = \begin{bmatrix} \widetilde{\boldsymbol{x}}_j^{r+} \\ \widetilde{\boldsymbol{x}}_j^{r-} \\ \widetilde{\boldsymbol{x}}_j^{i+} \\ \widetilde{\boldsymbol{x}}_j^{i-} \end{bmatrix}, \begin{array}{l} (\widetilde{\boldsymbol{x}}_j^{r+}, \widetilde{\boldsymbol{x}}_j^{r-}, \widetilde{\boldsymbol{x}}_j^{i+}, \widetilde{\boldsymbol{x}}_j^{i-}) \in \mathcal{H}_j^{+4} \\ (\widetilde{\boldsymbol{x}}_j^{r+} \text{ ou } \widetilde{\boldsymbol{x}}_j^{r-}) = \mathbf{0} \\ (\widetilde{\boldsymbol{x}}_j^{i+} \text{ ou } \widetilde{\boldsymbol{x}}_j^{i-}) = \mathbf{0} \\ (\widetilde{\boldsymbol{x}}_j^{r+} - \widetilde{\boldsymbol{x}}_j^{r-}) \propto (\widetilde{\boldsymbol{x}}_j^{i-} - \widetilde{\boldsymbol{x}}_j^{i-}) \end{array} \right\} \tag{2.13}$$

Encore une fois, nous rappelons que cette formalisation de l'espace des contraintes semble très similaire à celle proposée par [FDJ15], mais en diffère pour les mêmes raisons que celles explicitées précédemment. De plus, l'auteur ne prend pas en compte la contrainte de liaison entre parties réelle et imaginaire, sans expliciter qu'il la relâche. Nous verrons en § 2.3 qu'en pratique, nous ne pourrons pas nous non plus imposer cette contrainte.

### 2.2.4 Bilan

Nous avons donc vu que pour décrire des espaces de contraintes les plus proches possibles des espaces images $\mathcal{S}_j^{\mathbb{R}}$ et $\mathcal{S}_j^{\mathbb{C}}$, il était nécessaire de dédoubler (voir de quadrupler dans le cas complexe) les atomes interpolés et les amplitudes associées.

Dans un cadre général, le problème d'APLsc s'écrira par la suite de la façon suivante :

$$\text{Estimer } \widetilde{\boldsymbol{x}} \in \mathcal{C} \textit{ parcimonieux} \text{ tels que } \boldsymbol{y} = \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}} + \widetilde{\boldsymbol{\epsilon}} = \sum_{j=1}^{J} \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j + \widetilde{\boldsymbol{\epsilon}} \tag{2.14}$$

où $\widetilde{\boldsymbol{x}} \in \mathcal{C} \Leftrightarrow \forall j,\ \widetilde{\boldsymbol{x}}_j \in \mathcal{C}_j$, et où on distingue les cas suivants :

- amplitudes positives : $\mathcal{C}_j = \mathcal{H}_j^+ \subset \mathbb{R}^K$,
- amplitudes réelles : $\widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j = \widetilde{\mathbf{h}}_j^{+-} \widetilde{\boldsymbol{x}}_j^{+-}$ et espace des contraintes $\mathcal{C}_j \subset \mathbb{R}^{2K}$ défini en défini en équation (2.11),
- amplitudes complexes : $\widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j = \widetilde{\mathbf{h}}_j^{ri+-} \widetilde{\boldsymbol{x}}_j^{ri+-}$ et espace des contraintes $\mathcal{C}_j \subset \mathbb{R}^{4K}$ défini en défini en équation (2.13).

La parcimonie devra s'imposer de façon structurée à l'ensemble des composantes de $\widetilde{\boldsymbol{x}}_j$, soit simultanément à ces parties positive et négative (et réel et imaginaire dans le cas complexe).

Dans les cas à amplitudes réelles et complexes, nous avons vu qu'il était nécessaire d'imposer des contraintes de disjonction et de liaison (pour les amplitudes complexes) pour que l'espace des contraintes ne s'éloignent pas trop de l'espace image. Cependant, ces contraintes ne rentrent pas dans le cadre des contraintes convexes que nous avons définies en (2.4). Nous verrons dans la section suivante comment les méthodes de résolution de tels problèmes peuvent (ou ne peuvent pas) prendre en compte ces contraintes.

## 2.3 Vers une résolution optimale des problèmes d'APLsc

Nous avons vu en § 2.1.2 que les méthodes actuellement utilisées pour résoudre les problèmes d'APLsc (approches gloutonnes, pénalisation $\ell_1$) sont sous-optimales, dans le sens où la parcimonie n'est pas prise en compte de façon exacte. La contribution majeure de cette section est la proposition d'une résolution optimale via la reformulation du problème d'APLsc avec exploitation de la pseudo-norme $\ell_0$ en MIP (programmation mixte en nombre entier) permettant une prise en compte exacte de la parcimonie via l'introduction de variables binaires. Cette reformulation sera étudiée pour les trois cas possibles d'espaces des contraintes (amplitudes positives, réelles et complexes). Par ailleurs, la méthode par pénalisation $\ell_1$ n'a été correctement

définie par [ETS11] que pour les signaux à amplitudes positives. Nous commencerons donc par étudier sa formulation pour les espaces de contraintes liés aux problèmes à amplitudes réelles et complexes que nous avons proposé précédemment. Nous allons donc voir, pour chacune des résolutions de type $\ell_1$ et $\ell_0$, comment peut être imposée la parcimonie structurée sur l'ensemble des composantes des $\widetilde{\boldsymbol{x}}_j$ et les contraintes de disjonction et de liaison.

Dans cette section, nous nous plaçons dans le cas où l'espace relaxé $\mathcal{H}_j^+$ respecte l'implication de nullité en équation (2.5), qui pour rappel s'écrit :

$$\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+ \subset \mathbb{R}^K \ : \ \alpha_j = \widetilde{x}_{j,1} = 0 \ \Rightarrow \ \widetilde{\boldsymbol{x}}_j = \mathbf{0}$$

Pour rappel, les espaces $\mathcal{H}_j^+$ définis en Tableau 2.1 pour les interpolations de dictionnaire avec approximations de Taylor et interpolation polaire respectent cette contrainte. Ce sera également le cas pour l'espace $\mathcal{H}_j^+$ que nous définirons pour l'approximation par K-SVD en § 2.4.2, dans le cas où les premières composantes $\widetilde{x}_{j,1}$ sont sous contraintes de positivité. Le respect de l'implication de nullité par l'espace des contraintes permet d'appliquer la parcimonie uniquement aux premières composantes de $\widetilde{\boldsymbol{x}}$ (notées $\boldsymbol{\alpha}_j$, composées des premières composantes des parties positive/négative et réelle/imaginaire le cas échéant). Notons que dans ce cas, la contrainte de disjonction peut s'écrire comme une contrainte égalité quadratique :

$$\text{Avec implication de nullité,} \quad \widetilde{\boldsymbol{x}}_j^+ = \mathbf{0} \text{ ou } \widetilde{\boldsymbol{x}}_j^- = \mathbf{0} \quad \Leftrightarrow \quad \alpha_j^+ \cdot \alpha_j^- = 0.$$

Cependant, cette contrainte égalité quadratique ne correspond pas à une contrainte convexe comme nous l'avons définie en § 2.1 (équation (2.4)) et elle ne peut donc pas être prise en compte facilement (voir [BV04]). Il sera donc nécessaire d'imposer la contrainte de disjonction par un autre moyen.

### 2.3.1 Une résolution optimale avec la pseudo-norme $\ell_0$

Nous avons vu en § 1.2.2 qu'il est possible de formuler de façon exacte la parcimonie pour les problèmes d'APL grâce à une reformulation en *Programmation mixte en nombres entiers* (MIP) proposée par [BNCM16] permettant l'exploitation de la pseudo-norme $\ell_0$ via l'introduction de variables binaires.

Dans cette section, nous proposons l'exploitation des MIP pour la résolution des problèmes d'APLsc. La différence principale avec les problèmes d'APL est bien sûr la présence des contraintes. Elles ne représentent pas une grande difficulté, puisque la plupart des logiciels permettant la résolution des MIP (dont le logiciel CPLEX que nous utiliserons) permettent la prise en compte de contraintes du même type qu'en programmation convexe (voir (2.4)). En effet, la résolution du problème se fait via une exploration combinatoire (heureusement non exhaustive) des configurations possibles des nombres entiers ; pour chaque configuration (nombres entiers fixés), l'algorithme MIP résout alors un problème d'optimisation convexe. La principale difficulté de l'adaptation des reformulations MIP aux problèmes d'APLsc est la prise en compte des contraintes de disjonction pour les cas où $\mathcal{D}_x = \mathbb{R}$ ou $\mathbb{C}$. Notons d'ores et déjà que nous n'avons pas réussi à prendre en compte la contrainte de liaison dans le cas $\mathcal{D}_x = \mathbb{C}$.

L'adaptation des MIP aux problèmes d'APLsc pourrait se faire en théorie pour toutes les formulations données par [BNCM16] (différents types d'attache aux données et différentes façons de prendre en compte la parcimonie - voir (1.8) p. 9 pour ces dernières). Nous avons cependant choisi pour l'APLsc la formulation *erreur quadratique avec contraintes sur la parcimonie* (donnée pour le problème d'APL en équation (1.14)). Le choix d'un terme d'attache aux données quadratique est naturel vis-à-vis des travaux menés précédemment. En ce qui concerne la prise en compte de la parcimonie (en « contraintes », en « critère », ou en « pénalisation »), notre

choix se justifie par des raisons théoriques et pratiques : en effet, de par la nature discrète du problème, ces problèmes d'optimisation ne sont pas convexes et n'admettent donc pas nécessairement une unique solution. Cependant, d'après [BNCM16], pour l'erreur quadratique, les formulations « contraintes » et « pénalisation » peuvent donner une unique solution, contrairement à la formulation « critère ». De plus, cette dernière représente un coût de calcul bien plus élevé, à cause de la contrainte quadratique sur l'erreur (MIQCP[14] contre MIQP[15] pour les autres formulations). Enfin, le réglage de l'hyperparamètre est bien plus simple dans le cas « contraintes » puisqu'il est entier et donc discret ($L_0$, borne maximale pour l'ordre du modèle). Il suffit alors de calculer la solution pour certaines valeurs de $L_0$ afin de trouver le plus optimal. Nous reviendrons sur cette stratégie en § 2.5.1.

Dans un cadre général, le problème d'APLsc (2.14) se formule avec contrainte de parcimonie via la pseudo-norme $\ell_0$ de la façon suivante :

$$\min_{\widetilde{\boldsymbol{x}}\in\mathcal{C}} \quad \|\boldsymbol{y}-\widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 \quad \text{tel que} \quad \|\boldsymbol{\alpha}\|_0 \leq L_0 \tag{2.15}$$

où la parcimonie peut porter uniquement sur les premières composantes $\boldsymbol{\alpha}_j$ grâce à l'hypothèse de respect de l'implication de nullité. Nous proposons donc la reformulation en MIP de ce problème dans les cas $\mathcal{D}_x = \mathbb{R}_+$, $\mathbb{R}$ et $\mathbb{C}$.

### Amplitudes positives

Dans le cas où $\mathcal{D}_x = \mathbb{R}_+$, nous introduisons $J$ variables binaires $b_j$ qui permettent de conditionner la nullité des $\widetilde{\boldsymbol{x}}_j$ via la contrainte de *big-M*. Le problème se formalise alors de la façon suivante :

$$\min_{\substack{\widetilde{\boldsymbol{x}}\in\mathbb{R}^{KJ}\\ \boldsymbol{b}\in\{0,1\}^J}} \|\boldsymbol{y}-\widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 \text{ s.t. } \begin{cases} -M\boldsymbol{b} \leq \boldsymbol{\alpha} \leq M\boldsymbol{b} \\ \|\boldsymbol{b}\|_0 \leq L_0 \\ \widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+ \end{cases} \tag{2.16}$$

où pour rappel, $\|\boldsymbol{b}\|_0$ donne bien la somme des variables binaires $\mathbf{1}_J^T\boldsymbol{b} = \sum_j b_j$. La nature de ce problème d'optimisation dépendra de la nature des contraintes $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+$ : un MIQCP dans le cas de contraintes quadratiques (comme pour l'interpolation polaire) et un MIQP pour des contraintes linéaires (approximations de Taylor par exemple). Remarquons de plus que la contrainte de *big-M* porte sur les amplitudes $\alpha_j$ et non sur les amplitudes d'origine $x_j$ ; il sera donc important de définir un $M$ adapté aux valeurs que peuvent prendre les premières composantes ; ce sera le cas pour tous les $\mathcal{D}_x$ et nous y reviendrons en § 2.5.1.

### Amplitudes à valeurs réelles

Dans le cas où $\mathcal{D}_x = \mathbb{R}$, nous rappelons que les variables $\widetilde{\boldsymbol{x}}_j \in \mathbb{R}^{2K}$ s'écrivent $[\widetilde{\boldsymbol{x}}_j^{+T}\ ,\ \widetilde{\boldsymbol{x}}_j^{-T}]^T$. La formalisation en MIP doit alors imposer de façon structurée la parcimonie sur l'ensemble des premières composantes de $\widetilde{\boldsymbol{x}}_j$, mais aussi garantir la contrainte de disjonction.

Une première solution serait de garder $J$ variables binaires $b_j$, et de les associer à la première composante de $\widetilde{\boldsymbol{x}}_j^+$ et de $\widetilde{\boldsymbol{x}}_j^-$ via le *big-M* : $-Mb_j \leq \alpha_j^+, \alpha_j^- \leq Mb_j$. La parcimonie serait alors imposée de façon structurée. Cependant, le problème ainsi posé ne nous permet pas de contraindre la disjonction entre parties positive et négative. Pour rappel, la contrainte $\alpha_j^+ \cdot \alpha_j^- = 0$, satisfaisante en théorie, est une contrainte quadratique égalité qui ne peut pas être prise en compte en pratique. Il est nécessaire d'imposer la disjonction à l'aide de contraintes

[14] Mixed-Integer Quadratically Constrained Program.
[15] Mixed-Integer Quadratic Program.

additionnelles ; sans elles, il pourrait exister une infinité de $\widetilde{\boldsymbol{x}}_j^+$ et $\widetilde{\boldsymbol{x}}_j^-$ appartenant à $\mathcal{H}_j^+$ et donnant le même $\widetilde{\boldsymbol{x}}_j = \widetilde{\boldsymbol{x}}_j^+ - \widetilde{\boldsymbol{x}}_j^-$ et donc la même valeur du critère quadratique ; mais la variable $\widetilde{\boldsymbol{x}}_j$ ainsi reconstruite n'appartiendrait pas nécessairement à $\mathcal{H}_j^+ \cup \mathcal{H}_j^-$.

La solution que nous avons trouvée pour imposer la contrainte de disjonction est le dédoublement des variables binaires $\boldsymbol{b} = [\boldsymbol{b}^+ \ , \ \boldsymbol{b}^-] \in \{0,1\}^{2J}$ où les $\boldsymbol{b}^+$ conditionnent la nullité des $\boldsymbol{\alpha}^+$ et $\boldsymbol{b}^-$ conditionne la nullité des $\boldsymbol{\alpha}^-$ via des contraintes de *big-M*. La contrainte de disjonction peut alors être imposée par la contrainte linéaire $b_j^+ + b_j^- \leq 1$ (au moins l'un des deux sera nul). La pseudo-norme $\ell_0$ de $\widetilde{\boldsymbol{x}}$ se calcule toujours comme la somme des variables binaires ($\|\boldsymbol{b}\|_0$), puisque à $j$ fixé, $b_j^+ + b_j^- \leq 1$. Le problème s'écrit alors de la façon suivante :

$$\min_{\substack{\widetilde{\boldsymbol{x}} \in \mathbb{R}^{2KJ} \\ \boldsymbol{b} \in \{0,1\}^{2J}}} \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 \text{ s.t. } \begin{cases} -M\boldsymbol{b}^+ \leq \boldsymbol{\alpha}^+ \leq M\boldsymbol{b}^+ \\ -M\boldsymbol{b}^- \leq \boldsymbol{\alpha}^- \leq M\boldsymbol{b}^- \\ b_j^+ + b_j^- \leq 1 \end{cases} \text{ et } \begin{array}{l} \|\boldsymbol{b}\|_0 \leq L_0 \\ \widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+ \end{array} \tag{2.17}$$

Le dédoublement des variables binaires permet donc de prendre en compte la contrainte de disjonction via une contrainte linéaire ; elle génère cependant une augmentation du coût de calcul considérable puisqu'on double le nombre de variables entières.

### Amplitudes à valeurs complexes

Dans le cas où $\mathcal{D}_x = \mathbb{C}$, nous rappelons que les variables $\widetilde{\boldsymbol{x}}_j \in \mathbb{R}^{4K}$ sont composées de $(\widetilde{\boldsymbol{x}}_j^{r+}, \widetilde{\boldsymbol{x}}_j^{r-}, \widetilde{\boldsymbol{x}}_j^{i-}, \widetilde{\boldsymbol{x}}_j^{i-})$. À partir de la formulation du problème réel, deux difficultés apparaissent :

1. L'amplitude reconstruite $\alpha_j$ s'écrit $(\alpha_j^{r+} - \alpha_j^{r-}) + i\ (\alpha_j^{i+} - \alpha_j^{i-})$. Les parties positive et négative ne doivent pas être non nulles en même temps, mais indépendamment entre parties réelle et imaginaire. Par exemple, on peut avoir $\alpha_j = \alpha_j^{r+} - i\ \alpha_j^{i-}$. Afin d'imposer la contrainte de disjonction de façon indépendante entre parties réelle et imaginaire, on doit (encore) dédoubler les variables binaires : $\boldsymbol{b} = [\boldsymbol{b}^{r+}, \boldsymbol{b}^{r-}, \boldsymbol{b}^{i+}, \boldsymbol{b}^{i-}] \in \{0,1\}^{4J}$, où les paires $(\boldsymbol{b}^{r+}, \boldsymbol{b}^{r-})$ et $(\boldsymbol{b}^{i+}, \boldsymbol{b}^{i-})$ permettront d'appliquer les contraintes de disjonction. Cela donne le problème suivant :

$$\min_{\substack{\widetilde{\boldsymbol{x}} \in \mathbb{R}^{4KJ} \\ \boldsymbol{b} \in \{0,1\}^{4J}}} \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 \text{ s.t. } \begin{cases} -M\boldsymbol{b}^{r+} \leq \boldsymbol{\alpha}^{r+} \leq M\boldsymbol{b}^{r+} \\ -M\boldsymbol{b}^{r-} \leq \boldsymbol{\alpha}^{r-} \leq M\boldsymbol{b}^{r-} \\ b_j^{r+} + b_j^{r-} \leq 1 \\ \|\boldsymbol{b}\|_0 \leq 2L_0 \end{cases} \text{ et } \begin{array}{l} -M\boldsymbol{b}^{i+} \leq \boldsymbol{\alpha}^{i+} \leq M\boldsymbol{b}^{i+} \\ -M\boldsymbol{b}^{i-} \leq \boldsymbol{\alpha}^{i-} \leq M\boldsymbol{b}^{i-} \\ b_j^{i+} + b_j^{i-} \leq 1 \\ \widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+ \end{array}$$

   La contrainte $\|\boldsymbol{b}\|_0 \leq 2L_0$ donne une borne maximale à la somme des variables binaires, car la somme (à $j$ fixé) $b_j^{R+} + b_j^{R-} + b_j^{I+} + b_j^{I-}$ est comprise entre 0 et 2 (grâce aux contraintes de disjonction).

2. Malheureusement, cette contrainte n'est pas équivalente à la contrainte de parcimonie originale $\|\widetilde{\boldsymbol{x}}\|_0 \leq L_0$ (où la norme $\ell_0$ est à prendre de façon structurée). En effet, la somme des variables binaires correspond ici aux nombres de parties réelles et imaginaires non nulles, et peut donc être inférieure au double du nombre de $\widetilde{\boldsymbol{x}}_j$ non nuls. Par exemple, prenons le cas suivant ($J = 2, L_0 = 1$) :

$$\begin{array}{rcccccccccc} \boldsymbol{b} & = & [ & b_1^{r+} & b_1^{r-} & b_1^{i+} & b_1^{i-} & , & b_2^{r+} & b_2^{r-} & b_2^{i+} & b_2^{i-} & ] \\ \boldsymbol{b}_a & = & [ & 1 & 0 & 0 & 1 & , & 0 & 0 & 0 & 0 & ] \\ \boldsymbol{b}_b & = & [ & 0 & 0 & 0 & 1 & , & 1 & 0 & 0 & 0 & ] \end{array}$$

Dans le premier cas, on a $\|\boldsymbol{b}_a\|_0 = 2 \leq 2L_0 = 2$ et $\|\widetilde{\boldsymbol{x}}\|_0 = 1 = L_0$ : la contrainte de parcimonie est bien respectée. Dans le second cas, on a aussi $\|\boldsymbol{b}_b\|_0 = 2$ mais les deux $\widetilde{\boldsymbol{x}}_j$ sont non nuls si bien que $\|\widetilde{\boldsymbol{x}}\|_0 = 2 > L_0 = 1$ ; la contrainte de parcimonie n'est plus respectée. De façon plus générale, sous cette contrainte, un atome avec parties réelle et imaginaire non nulles ($\|\boldsymbol{b}_j\|_0 = 2$) équivaut alors à deux atomes avec l'une ou l'autre annulée $\|\boldsymbol{b}_{j_1}\|_0 = \|\boldsymbol{b}_{j_2}\|_0 = 1$.

Nous avons donc proposé de rajouter une contrainte égalité linéaire sur les variables binaires pour que $\|\boldsymbol{b}_j\|_0 \in \{0, 2\}$ : $b_j^{r+} + b_j^{r-} = b_j^{i+} + b_j^{i-}$. La norme $\|\widetilde{\boldsymbol{x}}\|_0$ est alors égale à deux fois la somme des variables binaires et on obtient la formulation suivante :

$$\min_{\substack{\widetilde{\boldsymbol{x}} \in \mathbb{R}^{4KJ} \\ \boldsymbol{b} \in \{0,1\}^{4J}}} \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 \text{ s.t. } \begin{cases} -M\boldsymbol{b}^{r+} \leq \boldsymbol{\alpha}^{r+} \leq M\boldsymbol{b}^{r+} \\ -M\boldsymbol{b}^{r-} \leq \boldsymbol{\alpha}^{r-} \leq M\boldsymbol{b}^{r-} \\ b_j^{r+} + b_j^{r-} \leq 1 \\ \|\boldsymbol{b}\|_0 \leq 2L_0 \end{cases} \text{ et } \begin{cases} -\boldsymbol{b}^{i+} \leq \boldsymbol{\alpha}^{i+} \leq M\boldsymbol{b}^{i+} \\ -\boldsymbol{b}^{i-} \leq \boldsymbol{\alpha}^{i-} \leq M\boldsymbol{b}^{i-} \\ b_j^{i+} + b_j^{i-} \leq 1 \\ \widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+ \end{cases} \quad (2.18)$$
$$b_j^{r+} + b_j^{r-} = b_j^{i+} + b_j^{i-}$$

Cette formulation induit une augmentation du nombre de variables binaires et donc du coût de calcul. Notons cependant que même si elle s'écrit ici avec $4J$ variables binaires, elle est en fait équivalente à un problème avec $3J$ variables binaires, grâce à la contrainte égalité introduite ci-dessus. Elle sera testée en § 2.6 sur les données tests d'analyse spectrale.

### 2.3.2 Prise en compte des contraintes pour la résolution avec pénalisation $\ell_1$

Au regard des travaux précédents, il nous semble nécessaire de revenir sur la résolution du problème d'APLsc avec pénalisation $\ell_1$ proposée par [FDJ15] pour $\mathcal{D}_x = \mathbb{C}$ (et $\mathcal{D}_x = \mathbb{R}$ par déduction) déjà évoquée en § 2.1.4. En effet, la prise en compte des contraintes de disjonction avec la reformulation MIP vue précédemment est rendue possible grâce aux variables binaires, absentes de la formulation du problème avec pénalisation $\ell_1$. Sans ces variables binaires, imposer ces contraintes dans le cadre de l'optimisation convexe avec pénalisation $\ell_1$ semble donc impossible. Par conséquent, nous allons voir dans quelles mesures ces contraintes seront néanmoins respectées, sans les imposer à l'aide de contraintes fermes.

Dans un premier temps, revenons au cas $\mathcal{D}_x = \mathbb{R}_+$, ne nécessitant pas le respect de ces contraintes. La formulation proposée par [ETS11] s'écrit de la façon suivante, dans le cas où l'espace des contraintes $\mathcal{C}_j = \mathcal{H}_j^+$ respecte l'implication de nullité :

$$\min_{\widetilde{\boldsymbol{x}} \in \mathcal{C}} \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 + \lambda\|\boldsymbol{\alpha}\|_1 \quad (2.19)$$

Dans le cas de l'interpolation polaire et des approximations de Taylor utilisées par [ETS11], mais aussi pour l'approximation par K-SVD dans le cas où l'espace des contraintes respecte l'implication de nullité, l'espace $\mathcal{H}_j^+$ contient une contrainte de positivité sur la première composante (voir Tableau 2.1 et § 2.4.2) : $\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+ \Rightarrow \alpha_j = \widetilde{x}_{j,1} \geq 0$. Cette contrainte de positivité sur les $\alpha_j$ permet d'écrire la norme $\ell_1$ de $\boldsymbol{\alpha}$ comme la somme (linéaire) de ces composantes, si bien que le problème (2.19) est bien un problème d'optimisation convexe comme présenté en (2.4).

Nous allons maintenant voir comment cette formulation se généralise quand les amplitudes $x_j$ sont réelles ou complexes.

### Amplitudes à valeurs réelles

Dans le cas où $\mathcal{D}_x = \mathbb{R}$, il est nécessaire d'imposer la parcimonie de façon structurée à l'ensemble des premières composantes de $\widetilde{\boldsymbol{x}}_j^+$ et $\widetilde{\boldsymbol{x}}_j^-$, soit $\alpha_j^+$ et $\alpha_j^-$. En les regroupant dans le vecteur $\boldsymbol{\alpha}_j$, la parcimonie structurée peut donc être imposée via la pénalisation $\ell_{1,2}$ de la façon suivante :

$$\|\boldsymbol{\alpha}\|_{1,2} = \sum_{j=1}^{J} \|\boldsymbol{\alpha}_j\|_2 = \sum_{j=1}^{J} \sqrt{{\alpha_j^+}^2 + {\alpha_j^-}^2}$$

Cette pénalisation favorise la mise à zéro simultanée de $\alpha_j^+$ et $\alpha_j^-$ ; cependant, elle ne semble pas favoriser la contrainte de disjonction. Remarquons cependant que si cette dernière est respectée, la norme $\ell_{1,2}$ se réécrit comme la norme $\ell_1$ de $\boldsymbol{\alpha}$ :

$$\text{Si } \alpha_j^+ \cdot \alpha_j^- = 0, \text{ alors } \|\boldsymbol{\alpha}\|_{1,2} = \sum_{j=1}^{J} \sqrt{{\alpha_j^+}^2 + {\alpha_j^-}^2} = \sum_{j=1}^{J} |\alpha_j^+| + |\alpha_j^-| = \|\boldsymbol{\alpha}\|_1$$

Comme nous l'avons vu précédemment, quelle que soit l'interpolation de dictionnaire envisagée respectant l'implication de nullité, l'espace $\mathcal{H}_j^+$ contient une contrainte de positivité sur la première composante : $(\widetilde{\boldsymbol{x}}_j^+, \widetilde{\boldsymbol{x}}_j^-) \in {\mathcal{H}_j^+}^2 \Rightarrow \alpha_j^+, \alpha_j^- \geq 0$. La norme $\ell_1$ de $\boldsymbol{\alpha}$ peut être écrite encore plus simplement :

$$\text{Si } \alpha_j^+ \cdot \alpha_j^- = 0 \text{ et } \boldsymbol{\alpha}_j \in \mathbb{R}_+^2, \text{ alors } \|\boldsymbol{\alpha}\|_1 = \sum_{j=1}^{J} |\alpha_j^+| + |\alpha_j^-| = \sum_{j=1}^{J} \alpha_j^+ + \alpha_j^- \tag{2.20}$$

Cette écriture de la norme $\ell_1$ rappelle la reformulation en problème de programmation quadratique (QP) pour un problème classique d'APL avec pénalisation $\ell_1$ sur des amplitudes réelles [Fuc04]. Pour un problème d'APL, on peut facilement montrer par l'absurde que la formulation satisfait la contrainte de disjonction. On peut alors suivre la même logique, pour essayer de montrer que la pénalisation par la norme $\ell_1$ de $\boldsymbol{\alpha}$ entraîne nécessairement un respect de la contrainte de disjonction.

On suppose donc par l'absurde que $\widetilde{\boldsymbol{x}}_0$ minimise le critère $\mathcal{J}(\widetilde{\boldsymbol{x}}) = \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 + \lambda \sum_j (\alpha_j^+ + \alpha_j^-)$ sur $\mathcal{H}^{+2J}$ *sans respecter la contrainte de disjonction.* Il existe alors un indice $j$ tel que $\alpha_{0,j}^+$ et $\alpha_{0,j}^-$ sont strictement supérieurs à 0. On peut alors construire un $\widetilde{\boldsymbol{x}}_1$ identique à $\widetilde{\boldsymbol{x}}_0$ sauf à l'indice $j$. À cet indice, on écrit $\alpha_{1,j}^+ = \max(0, \alpha_{0,j}^+ - \alpha_{0,j}^-)$ et $\alpha_{1,j}^- = -\min(0, \alpha_{0,j}^+ - \alpha_{0,j}^-)$. Le vecteur $\boldsymbol{\alpha}_{1,j}$ respecte alors la contrainte de disjonction et par inégalité triangulaire donne une norme $\ell_1$ plus faible : $\|\boldsymbol{\alpha}_1\|_1 < \|\boldsymbol{\alpha}_0\|_1$. En l'absence de contraintes (comme en APL), on peut construire le reste des composantes de $\widetilde{\boldsymbol{x}}_j$ pour que l'erreur quadratique en $\widetilde{\boldsymbol{x}}_0$ soit égale à celle en $\widetilde{\boldsymbol{x}}_1$ (par exemple, si $\alpha_{1,j}^+ = \alpha_{0,j}^+ - \alpha_{0,j}^-$ alors $\widetilde{\boldsymbol{x}}_{1,j}^+ = \widetilde{\boldsymbol{x}}_{0,j}^+ - \widetilde{\boldsymbol{x}}_{0,j}^-$ et $\widetilde{\boldsymbol{x}}_{1,j}^- = \mathbf{0}$), si bien qu'on met en évidence l'inégalité $\mathcal{J}(\widetilde{\boldsymbol{x}}_1) < \mathcal{J}(\widetilde{\boldsymbol{x}}_0)$, contraire à l'hypothèse de départ. Cependant, la présence des contraintes pour notre problème nous empêche de conclure cette contradiction. En effet, on a bien $\mathcal{J}(\widetilde{\boldsymbol{x}}_1) < \mathcal{J}(\widetilde{\boldsymbol{x}}_0)$, *mais rien ne prouve que* $\widetilde{\boldsymbol{x}}_{1,j} \in {\mathcal{H}_j^+}^2$.

Dans le cas des problèmes d'APLsc, nous n'avons donc pas pu démontrer que la pénalisation par la norme $\ell_1$ des $\boldsymbol{\alpha}$ suffisait à satisfaire la contrainte de disjonction. Cependant, sur nos divers tests avec cette pénalisation, cette contrainte a toujours été respectée, sans avoir à l'imposer dans l'algorithme d'optimisation, comme nous le verrons en § 2.5.

Nous conseillons donc la formulation suivante du problème (2.14) :

$$\min_{\widetilde{\boldsymbol{x}} \in \mathcal{H}^{+2J}} \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 + \lambda \|\boldsymbol{\alpha}\|_1 \tag{2.21}$$

#### Amplitudes à valeurs complexes

Dans le cas où $\mathcal{D}_x = \mathbb{C}$, la pénalisation $\ell_{1,2}$ s'écrit de la façon suivante :

$$\|\boldsymbol{\alpha}\|_{1,2} = \sum_{j=1}^{J} \|\boldsymbol{\alpha}_j\|_2 = \sum_{j=1}^{J} \sqrt{\alpha_j^{r+2} + \alpha_j^{r-2} + \alpha_j^{i+2} + \alpha_j^{i-2}}$$

Il s'agit de la même pénalisation que celle proposée par [FDJ15], mais pour un espace de contraintes différent puisque pour rappel, ce dernier considère $\widetilde{\boldsymbol{x}}_j = \widetilde{\boldsymbol{x}}_j^{ri+-} \in (\mathcal{H}_j^+ \cap \mathbb{R}_+^K)^4$. Nous n'avons pas trouvé de moyen pour garantir les contraintes de disjonction et de liaison dans ce contexte. Sans ces contraintes, le problème se reformule donc de la façon suivante :

$$\min_{\widetilde{\boldsymbol{x}} \in \mathcal{H}^{+4J}} \|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2 + \lambda \|\boldsymbol{\alpha}\|_{1,2} \tag{2.22}$$

En pratique, nous avons testé cette formulation sur le problème d'analyse spectrale en § 2.6 ; les solutions de ce problème ne respectaient ni les contraintes de liaison, ni les contraintes de disjonction.

#### Biais introduit par la norme $\ell_1$ et importance de la réévaluation des amplitudes

Tout comme dans le cas des problèmes d'APL classique, la pénalisation par des normes de types $\ell_1$ introduit un biais sur les amplitudes $\boldsymbol{\alpha}$, qui sont souvent sous-estimées. Ce biais se répand sur l'ensemble des composantes de $\widetilde{\boldsymbol{x}}$ à cause de l'implication de nullité (souvent dues au fait que les autres composantes sont bornées par les premières). Dans le cadre des problèmes d'APLsc provenant de l'interpolation de dictionnaire, cette sous-estimation est davantage préjudiciable quand on veut identifier les paramètres initiaux $(x_j, \delta_j)$ à partir des $\widetilde{\boldsymbol{x}}_j$ sous-estimés (projection sur $\mathcal{S}_j$ puis inversion du changement de variables).

À partir d'une solution $\widetilde{\boldsymbol{x}}_0$, nous proposons donc la réévaluation des amplitudes sur le support $\mathsf{Supp}\,(\widetilde{\boldsymbol{x}}_0)$ de $\widetilde{\boldsymbol{x}}_0$. Dans le cas où $\mathcal{D}_x = \mathbb{R}_+$, elle s'écrit :

$$\min_{\widetilde{\boldsymbol{x}}_j \in \mathcal{H}_j^+} \|\boldsymbol{y} - \sum_{j \in \mathsf{Supp}(\widetilde{\boldsymbol{x}}_0)} \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j\|^2 \tag{2.23}$$

Dans le cas où $\mathcal{D}_x = \mathbb{R}$ ou $\mathbb{C}$, on ne réévalue uniquement que les parties non nulles pour garantir la contrainte de disjonction (en supposant que cette dernière ait été respectée lors de la résolution du problème pénalisé). Ainsi, les amplitudes cette fois-ci estimées minimisent bien l'erreur quadratique pour ce support, contrairement à la solution via la pénalisation $\ell_1$. Nous montrerons le gain de cette réévaluation en § 2.5.

Maintenant que nous avons reformulé la pénalisation $\ell_1$ sous-optimale pour notre problème d'APLsc, nous allons proposer une reformulation prenant en compte de façon exacte la parcimonie via la pseudo-norme $\ell_0$.

### 2.3.3 Bilan

Dans cette section, nous avons donc pu formuler les problèmes d'APLsc (2.14) sous contraintes de parcimonie exploitant la pseudo-norme $\ell_0$ pour les 3 cas d'espace de définition $\mathcal{D}_x = \mathbb{R}_+$, $\mathbb{R}$ et $\mathbb{C}$, sous l'hypothèse que l'espace $\mathcal{H}_j^+$ respecte l'implication de nullité. Nous sommes aussi revenus sur la formulation avec pénalisation $\ell_1$ pour les cas $\mathcal{D}_x = \mathbb{R}$ et $\mathbb{C}$.

Les reformulations en MIP du problème contraint avec la pseudo-norme $\ell_0$, que nous donnons pour chaque espace de définition, permettent de respecter les contraintes de disjonction. Cela

s'effectue cependant au prix d'une augmentation significative du coût de calcul, passant par le doublement ($\mathcal{D}_x = \mathbb{R}$) voir le quadruplement ($\mathcal{D}_x = \mathbb{C}$) du nombre de variables binaires. Néanmoins, une telle formulation permettra de prendre en compte de façon exacte la parcimonie. Nous n'avons pas trouvé de solution pour prendre en compte la contrainte de liaison dans le cas $\mathcal{D}_x = \mathbb{C}$. Les formulations sont récapitulées en équations (2.16) ($\mathcal{D}_x = \mathbb{R}_+$), (2.17) ($\mathcal{D}_x = \mathbb{R}$) et (2.18) ($\mathcal{D}_x = \mathbb{C}$).

Pour la pénalisation $\ell_1$, le problème (2.14) peut être formulé dans le cas $\mathcal{D}_x = \mathbb{R}_+$ selon l'équation (2.19) comme cela avait été proposé par [ETS11]. Dans le cas où $\mathcal{D}_x = \mathbb{R}$, on ne peut pas imposer la contrainte de disjonction à l'aide de contraintes convexes, mais on peut intuiter qu'elle sera néanmoins respectée dans la formulation (2.21). Pour $\mathcal{D}_x = \mathbb{C}$, les contraintes de disjonction et de liaison ne sont pas imposées et ont peu de chance d'être respectées pour la formulation (2.22).

Notons que la reformulation en MIP s'adapterait très bien au cas où l'espace des contraintes ne respecteraient pas l'implication de nullité, en faisant porter les contraintes de big-M sur l'ensemble des composantes de $\widetilde{\boldsymbol{x}}$[16]. Nous verrons un exemple en § 2.6. Au contraire, la formulation avec pénalisation $\ell_1$ se compliquerait davantage, car il serait alors nécessaire d'utiliser une norme mixte pour imposer la parcimonie structurée ; les contraintes de disjonction ne seraient sûrement pas respectées dans ce cas.

Maintenant que nous avons étudié la définition de l'espace des contraintes et les résolutions des problèmes d'APLsc dans un cadre général, nous allons revenir plus en détails sur deux types d'approximations linéaires qui ont attirés notre attention : l'interpolation polaire et l'approximation par K-SVD.

## 2.4 Contributions pour l'interpolation de dictionnaire

Dans cette section, nous revenons plus en détails sur deux approximations qui ont attirées notre attention lors de l'étude de l'interpolation de dictionnaire : l'interpolation polaire et l'approximation par K-SVD. La première est la plus utilisée dans la littérature (avec l'approximation de Taylor à l'ordre 1) ; nous revenons sur la définition de l'approximation pour corriger quelques inexactitudes présentes dans la littérature, et nous proposons un espace de contraintes linéaires, qui est moins coûteux à prendre en compte que les contraintes proposées par [ETS11]. La seconde est très peu utilisée, bien qu'elle donne une erreur d'approximation moyenne minimale. Nous proposons donc à la fois une définition correcte de l'espace de contraintes convexes et une extension en approximation par K-SVD continue, pour révéler le potentiel de cette interpolation de dictionnaire. Les exemples et illustrations seront calculés à partir du modèle utilisé pour les tests numériques en § 2.5, pour un problème de déconvolution de train d'impulsions.

*Comme dans l'ensemble de cette partie, nous nous intéressons à l'approximation linéaire d'un signal élémentaire $x_j \, \boldsymbol{h}_j(\delta_j)$ pour un indice $j$ fixé, nous nous permettons d'omettre cet indice afin d'alléger les notations.*

### 2.4.1 Interpolation polaire

Dans cette section, nous allons d'une part revenir sur la définition de l'arc paramétré permettant l'interpolation polaire d'un atome continu, afin de justifier l'utilisation de la définition de [Eka12] plutôt que celle de [FDJ15], tout en démontrant l'expression de cette approximation pour des signaux à valeurs complexes. D'autre part, nous allons revenir sur l'établissement de

[16] Il faudrait néanmoins adapter la borne $M$ selon l'ordre de grandeur de la composante.

l'espace convexe $\mathcal{H}^+$ donné en Tableau 2.1 pour en montrer une limite, et pour proposer un nouvel espace convexe permettant une réduction drastique du coût de calcul.

### Une définition correcte de l'arc paramétré

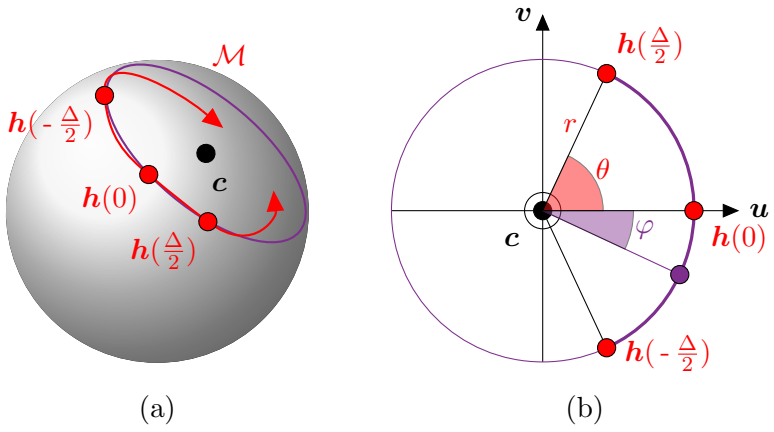
Nous commençons par revenir plus en détails sur la définition de l'arc paramétré qui permet d'approcher un atome continu dans le cadre de l'interpolation polaire. L'approximation par interpolation polaire est définie par [ETS11] pour des atomes respectant des hypothèses de conservation de la norme sous translation et de courbure constante par symétrie. D'après [FDJ15], ces hypothèses s'écrivent de la façon suivante :

$$\forall \delta \in \mathcal{I}_\Delta, \quad \underbrace{\|\boldsymbol{h}(\delta)\|_2 = R}_{\text{invariance par translation}} \quad \text{et} \quad \underbrace{\|\boldsymbol{h}_0 - \boldsymbol{h}(\delta)\|_2 = \|\boldsymbol{h}_0 - \boldsymbol{h}(-\delta)\|_2}_{\text{symétrie}}$$

où $R$ est une constante et $\boldsymbol{h}_0 = \boldsymbol{h}(0)$. De cette hypothèse d'invariance par translation, on déduit que la courbe $\mathcal{M} = \boldsymbol{h}(\mathcal{I}_\Delta)$, représentée en rouge en Figure 2.5(a), appartient à la surface d'une hypersphère (sphère grise sur l'illustration). L'unique arc de cercle passant par les trois points $[\boldsymbol{h}(-\frac{\Delta}{2})), \boldsymbol{h}(0), \boldsymbol{h}(+\frac{\Delta}{2}))] = [\boldsymbol{h}_m, \boldsymbol{h}_0, \boldsymbol{h}_p]$ (appartenant au cercle violet sur l'illustration), de centre $\boldsymbol{c}$, est schématisé en Figure 2.5(b) et peut être paramétré par l'angle $\varphi \in [-\theta; \theta]$ de la façon suivante :

$$\boldsymbol{f}(\varphi) = \boldsymbol{c} + r\cos(\varphi)\boldsymbol{u} + r\sin(\varphi)\boldsymbol{v} \quad \text{avec} \quad |\varphi| \leq \theta \tag{2.24}$$

où $(r, \theta)$ sont respectivement le rayon et le demi angle de l'arc.




(a) (b)

FIG. 2.5 – Illustration du principe de l'interpolation polaire. (a) Représentation de l'hypersphère, de l'espace $\mathcal{M}$ en rouge et de l'unique cercle (en violet) appartenant à l'hypersphère et passant par le trois points. (b) Représentation de l'arc de cercle (en violet) paramétrisé par l'angle $\varphi$ dans le plan du cercle. Figure inspirée de [ETS11].

Les vecteurs $[\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}]$ se calculent en inversant le système suivant :

$$\begin{array}{rcccccccc}
\boldsymbol{h}_m & = & \boldsymbol{f}(-\frac{\Delta}{2}) & = & \boldsymbol{c} & + & r\cos\theta\ \boldsymbol{u} & - & r\sin\theta\ \boldsymbol{v} \\
\boldsymbol{h}_0 & = & \boldsymbol{f}(0) & = & \boldsymbol{c} & + & r\ \boldsymbol{u} & & \\
\boldsymbol{h}_p & = & \boldsymbol{f}(\frac{\Delta}{2}) & = & \boldsymbol{c} & + & r\cos\theta\ \boldsymbol{u} & + & r\sin\theta\ \boldsymbol{v}
\end{array}$$

$$\Updownarrow$$

$$\begin{aligned}
[\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}] &= [\boldsymbol{h}_m, \boldsymbol{h}_0, \boldsymbol{h}_p] \begin{bmatrix} 1 & 1 & 1 \\ r\cos\theta & r & r\cos\theta \\ -r\sin\theta & 0 & r\sin\theta \end{bmatrix}^{-1} \\
&= [\boldsymbol{h}_m, \boldsymbol{h}_0, \boldsymbol{h}_p]\ \mathbf{A}(r, \theta)^{-1}
\end{aligned} \tag{2.25}$$

Ainsi, l'arc $\boldsymbol{f}(\varphi)$ dépend uniquement des points $(\boldsymbol{h}_m, \boldsymbol{h}_0, \boldsymbol{h}_p)$ et des paramètres $(r, \theta)$. Il peut être défini pour n'importe quelles valeurs de $(r, \theta)$, tant que la matrice $\mathbf{A}(r, \theta)$ est inversible.

Le principe de l'interpolation polaire est alors d'approcher l'atome $\boldsymbol{h}(\delta)$ par cet arc $\boldsymbol{f}(\varphi)$ appartenant aussi à l'hypersphère, et en supposant que le décalage $\delta$ est proportionnel à l'angle $\varphi$ par la relation $\frac{\delta}{\Delta/2} = \frac{\varphi}{\theta}$. L'approximation par interpolation polaire s'écrit donc de la façon suivante :

$$\boldsymbol{h}(\delta) \approx \boldsymbol{f}(\varphi) = \widetilde{\mathbf{h}}\Psi(\varphi) \quad \text{avec} \quad \varphi = \frac{2\theta}{\Delta}\delta \quad \Rightarrow \quad \left\{ \begin{array}{rcl} \widetilde{\mathbf{h}} & = & [\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}] \\ \Psi(\varphi) & = & [1, r\cos\varphi, r\sin\varphi]^T \in \mathbb{R}^3 \end{array} \right.$$

où nous retrouvons les expressions données en Tableau 2.1. L'approximation est exacte aux points d'interpolation $\delta \in \{-\frac{\Delta}{2}, 0, \frac{\Delta}{2}\}$.

Dans la littérature, deux expressions différentes sont évoquées pour $(r, \theta)$. Si leurs expressions n'est pas explicitées dans [ETS11], le premier article traitant de l'interpolation polaire, elles sont cependant données dans le thèse du même auteur [Eka12] :

$$\begin{array}{rcl} \theta_E & = & 2\,\text{Angle}(\boldsymbol{h}_0 - \boldsymbol{h}_p, \boldsymbol{h}_m - \boldsymbol{h}_p) = 2\arccos\left(\dfrac{(\boldsymbol{h}_0 - \boldsymbol{h}_p)^T(\boldsymbol{h}_m - \boldsymbol{h}_p)}{\|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2\|\boldsymbol{h}_m - \boldsymbol{h}_p\|_2}\right) \\ r_E & = & \sqrt{\dfrac{\|\boldsymbol{h}_p - \boldsymbol{h}_0\|_2^2}{2(1-\cos\theta)}} \end{array} \tag{2.26}$$

Dans [FDJ15], qui étend l'interpolation polaire aux atomes complexes sans justifications, les expressions sont différentes :

$$\begin{array}{rcl} r_F & = & \|\boldsymbol{h}_p\|_2 = \|\boldsymbol{h}_m\|_2 = \|\boldsymbol{h}_0\|_2 \\ \theta_F & = & \text{Angle}(\boldsymbol{h}_m, \boldsymbol{h}_0) = \arccos\left(\dfrac{\mathfrak{Re}(\boldsymbol{h}_0^\dagger \boldsymbol{h}_m)}{\|\boldsymbol{h}_0\|_2\|\boldsymbol{h}_m\|_2}\right) \end{array} \tag{2.27}$$

Pour un triplet $(\boldsymbol{h}_m, \boldsymbol{h}_0, \boldsymbol{h}_p)$ donné, ces deux expressions différentes de $(r, \theta)$ donnent lieu à des arcs paramétrés différents par l'équation (2.25). En Figure 2.6, nous avons représenté les deux arcs différents dans un cas particulier ($N = 3$) permettant de les visualiser en 3D. On observe bien que les deux arcs sont différents ; notamment, le point $\boldsymbol{c}$ est au centre de l'arc de cercle pour $(r_E, \theta_E)$, tandis que celui associé à $(r_F, \theta_F)$ en est très éloigné ; l'arc en question ressemble plutôt à une ellipse. Par ailleurs, si l'on observe le zoom autour des différents arcs (figure de droite), on remarque que l'arc associé à $(r_E, \theta_E)$ appartient bien à l'hypersphère, tandis que celui associé à $(r_F, \theta_F)$ s'en éloigne parfois (il disparait à l'intérieur de la sphère par intermittence).

À partir de ces observations, nous avons démontré le théorème suivant pour des atomes à valeurs réelles ou complexes (démonstration en Annexe A.1) :

**Théorème 1** *Soit* $(\boldsymbol{h}_m, \boldsymbol{h}_0, \boldsymbol{h}_p) \in \mathbb{F}^N$ *tels que :*
*(H1)* $\|\boldsymbol{h}_m\|_2 = \|\boldsymbol{h}_0\|_2 = \|\boldsymbol{h}_p\|_2$
*(H2)* $\|\boldsymbol{h}_0 - \boldsymbol{h}_m\|_2 = \|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2$
*Alors l'arc paramétré* $\boldsymbol{f}(\varphi)$ *appartient à l'hypersphère pour tout* $\varphi \in [-\theta; \theta]$, *soit* $\|\boldsymbol{f}(\varphi)\|_2 = \|\boldsymbol{h}_0\|_2$, *si et seulement si les paramètres* $(r, \theta)$ *s'écrivent :*

$$\begin{array}{rcl} \theta & = & 2\arccos\left(\dfrac{\mathfrak{Re}((\boldsymbol{h}_0 - \boldsymbol{h}_p)^\dagger(\boldsymbol{h}_m - \boldsymbol{h}_p))}{\|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2\|\boldsymbol{h}_m - \boldsymbol{h}_p\|_2}\right) \\ r & = & \dfrac{\|\boldsymbol{h}_p - \boldsymbol{h}_0\|_2}{\rho\sqrt{2(1-\cos\theta)}} \end{array} \tag{2.28}$$

*extension directe de l'expression de [Eka12] valables pour les complexes, et les vecteurs* $[\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}]$ *respectent les propriétés suivantes :*

$$\begin{gathered}\mathfrak{Re}(\boldsymbol{c}^\dagger \boldsymbol{u}) = \mathfrak{Re}(\boldsymbol{c}^\dagger \boldsymbol{v}) = \mathfrak{Re}(\boldsymbol{u}^\dagger \boldsymbol{v}) = 0 \\ \|\boldsymbol{u}\|_2 = \|\boldsymbol{v}\|_2 = \rho\end{gathered}$$

Le scalaire $\rho$ règle la norme des vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ ; le fixer à 1 permet d'obtenir des vecteurs normés. On obtient l'expression donnée dans [Eka12] dans le cas d'atomes à valeur réelles, pour lesquels les vecteurs $(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v})$ sont alors orthogonaux. Notre propriété la généralise aux atomes à valeurs complexes, où l'orthogonalité se conserve uniquement en partie réelle.

Par conséquent, nous démontrons par cette propriété que cette définition des paramètres $(r, \theta)$ est la seule qui permet à l'arc $\boldsymbol{f}(\varphi)$ d'appartenir à la même hypersphère que la fonction $\boldsymbol{h}(\delta)$ qu'il approche. Notons d'ailleurs que pour $\boldsymbol{h}(\delta) \in \mathbb{R}^3$, elle correspond à l'illustration donnée par [ETS11] qui a inspiré la Figure 2.5.

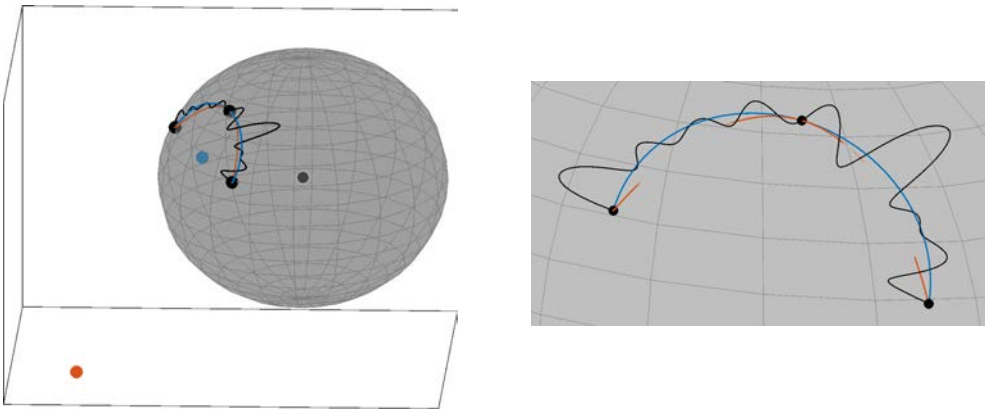


FIG. 2.6 – Illustration en 3D ($\boldsymbol{h}(\delta) \in \mathbb{R}^3$) des arcs interpolateurs proposés par [Eka12] (en bleu) et [FDJ15] (en rouge). Les points de la même couleur représentent les positions des vecteurs $\boldsymbol{c}$ relatifs à chaque arc. En noir sont représentés les trois points appartenant à la sphère, ainsi qu'une courbe $\boldsymbol{h}(\delta)$ appartenant aussi à la sphère et passant par ces trois points. À droite est montré un zoom au niveau des arcs interpolateurs, où l'on a masqué l'intérieur de la sphère.

Comme nous l'avons constaté sur la Figure 2.6, l'arc proposé par [FDJ15] n'appartient généralement pas à l'hypersphère ; le choix d'un tel arc pour approcher l'atome continu $\boldsymbol{h}(\delta)$ est donc incohérent avec les hypothèses du problème. De plus, les vecteurs $(\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v})$ ne seront généralement pas orthogonaux[17]. Les expressions $(r_F, \theta_F)$ données par l'auteur proviennent certainement d'une incompréhension des explications données par [ETS11], puisque ce dernier ne donne l'expression correcte que dans son manuscrit de thèse. C'est cependant cette expression $(r_F, \theta_F)$ qui est la plus couramment utilisée et rappelée dans la littérature (voir [FMC17] par exemple). Bien qu'en pratique, les deux arcs correspondants à ces deux expressions différentes soient très similaires, nous insistons sur le fait que le choix de l'expression en (2.28) devrait donner de meilleurs résultats.

Dans la suite de nos travaux, c'est donc cette écriture qui sera utilisée.

[17] La propriété d'orthogonalité est particulièrement intéressante dans le cas d'une minimisation de l'erreur quadratique entre un signal et un unique atome interpolé $\|\boldsymbol{y} - \widetilde{\mathbf{h}}\widetilde{\boldsymbol{x}}\|^2$, comme par exemple lors de la sélection d'un algorithme glouton type COMP (voir § 2.1.2). Elle permet notamment de trouver une solution analytique à la minimisation du critère sur l'espace $\mathcal{S}^{\mathbb{R}}$ (nous ne détaillons pas les calculs dans ce manuscrit). Nous émettons l'hypothèse qu'une telle propriété doit aussi faciliter la résolution de problèmes de type $\ell_1$ ou $\ell_0$.

### Une nouvelle relaxation convexe de l'espace image

À partir de l'arc interpolateur, on peut définir le changement de variables permettant d'obtenir une approximation linéaire de l'atome continu. Comme nous l'avons vu en § 2.2, il suffit de s'intéresser à l'espace image tel que $\mathcal{D}_x = \mathbb{R}_+$ pour définir l'espace des contraintes, les cas $\mathcal{D}_x = \mathbb{R}$ et $\mathbb{C}$ pouvant en être déduits. Pour rappel, le changement de variables et l'espace image $\mathcal{S}^+ = g(\mathbb{R}_+ \times \mathcal{I}_\Delta)$ s'écrivent, en suivant [ETS11] :

$$\tilde{\boldsymbol{x}} = g(x, \delta) = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} x \\ x\, r\, \cos\varphi \\ x\, r\, \sin\varphi \end{bmatrix}, \; \varphi = \frac{2\theta}{\Delta}\delta \quad \text{et} \quad \mathcal{S}^+ = \left\{ \tilde{\boldsymbol{x}} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \in \mathbb{R}^3, \; \begin{array}{c} \alpha \geq 0 \\ \beta \geq r\alpha\cos\theta \\ \beta^2 + \gamma^2 = r^2\alpha^2 \end{array} \right\}$$

De plus, l'angle $\varphi$ est identifié par $\varphi = \arctan(\gamma/\beta)$. Nous commençons par faire une rapide remarque sur la dépendance en $\theta$ de la relaxation convexe $\mathcal{H}^+$ de $\mathcal{S}^+$ proposée par [ETS11], qui consiste à relâcher la contrainte égalité en inégalité : $\beta^2 + \gamma^2 \leq r^2\alpha^2$. La Figure 2.7 représente les espaces $\mathcal{S}^+$ et $\mathcal{H}^+$ pour un $\theta$ inférieur et supérieur à $\pi/2$. Dans le cas où $\theta \geq \pi/2$, deux particularités ressortent sur la reconstruction de l'angle $\varphi$ :

1. Dans le cas où $\beta < 0$, l'angle $\varphi$ à identifier est supérieur à $\pi/2$ en valeur absolue ; l'identification $\varphi = \arctan(\gamma/\beta)$ proposée par [ETS11] est donc inadaptée puisque cette dernière donne un angle appartenant à $]-\pi/2; \pi/2[$.
   *Dans le cas où $\theta \geq \pi/2$, l'identification de $\varphi$ doit donc se faire via la fonction atan2 : atan2$(\gamma, \beta)$ donne l'unique angle $\varphi$ dans $]-\pi; \pi[$ tel que $\beta = R\cos\varphi$ et $\gamma = R\sin\varphi$ ($R$ correspond au rayon du point $(\beta, \gamma)$.*
2. L'espace $\mathcal{H}^+$ contient des points tel que l'angle associé est supérieur à $\theta$ en valeur absolue (les points correspondant au quadrillage non bleuté sur la figure).

Il est donc préférable d'éviter les valeurs $\theta$ supérieures à $\pi/2$, puisque les $\delta$ identifiés pourraient ne pas appartenir à l'espace de départ $\mathcal{I}_\Delta$ par la relation de proportionnalité avec $\varphi$.

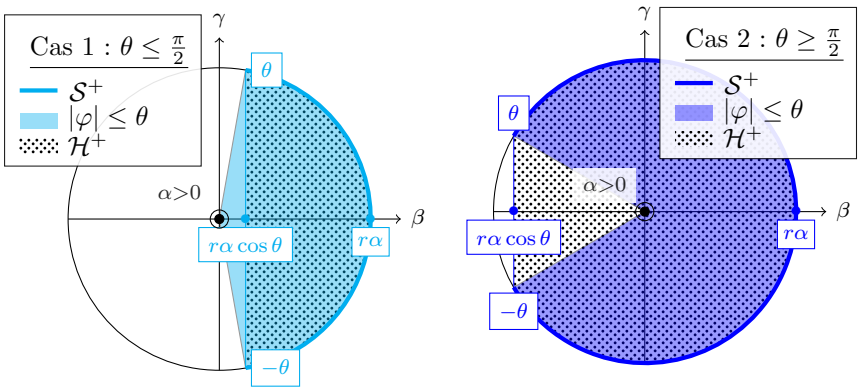



FIG. 2.7 – Comparaison des espaces $\mathcal{S}^+$ et $\mathcal{H}^+$ de l'interpolation polaire pour deux $\theta$ différents.

De par la définition de l'angle $\theta$ en équation (2.28), il est évident qu'il dépend fortement de la largeur $\Delta$ de $\mathcal{I}_\Delta$. Notamment, plus $\Delta$ est grand et plus l'angle $\theta$ devrait l'être aussi. En Figure 2.8, nous avons représenté l'évolution de $\theta$ en fonction de $\Delta$ pour un atome continu $\boldsymbol{h}(\delta)$ du même type que celui que nous utiliserons en § 2.5. Cette courbe confirme notre hypothèse.

*Conclusion : en pratique pour une application donnée (soit un atome $\boldsymbol{h}(\delta)$ donné), il sera nécessaire de choisir un $\Delta$ suffisamment petit pour assurer un angle $\theta$ inférieur à $\pi/2$.*

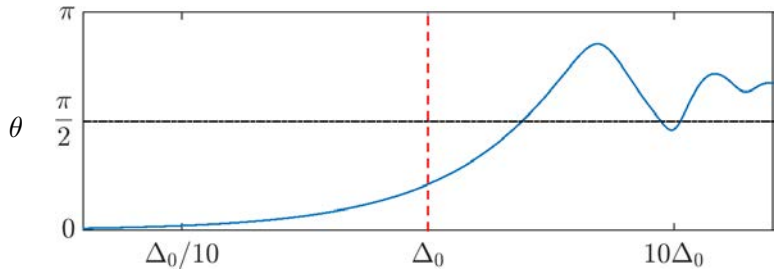



FIG. 2.8 – Évolution de $\theta$ en fonction de $\Delta$ en échelle logarithmique. La valeur $\Delta_0$ est celle que nous utiliserons pour les tests en § 2.5.

Comme nous l'avons rappelé précédemment, l'espace relâché $\mathcal{H}^+$ comporte une contrainte quadratique : $\beta^2 + \gamma^2 \leq r^2\alpha^2$. La prise en compte de cette contrainte entraîne un coût calculatoire bien plus important, puisqu'elle nécessite l'utilisation de programmation de type QCQP (résolution de type $\ell_1$) ou MIQCP (résolution de type $\ell_0$). Nous avons donc proposé de remplacer cette contrainte quadratique par des contraintes linéaires : les espaces correspondant, que nous notons $\mathcal{H}^+_{\text{Lin}}$, englobent l'enveloppe convexe $\mathcal{H}^+$ et sont définis à l'aide de tangentes en $\mathcal{S}^+$. Nous les représentons en Figure 2.9.

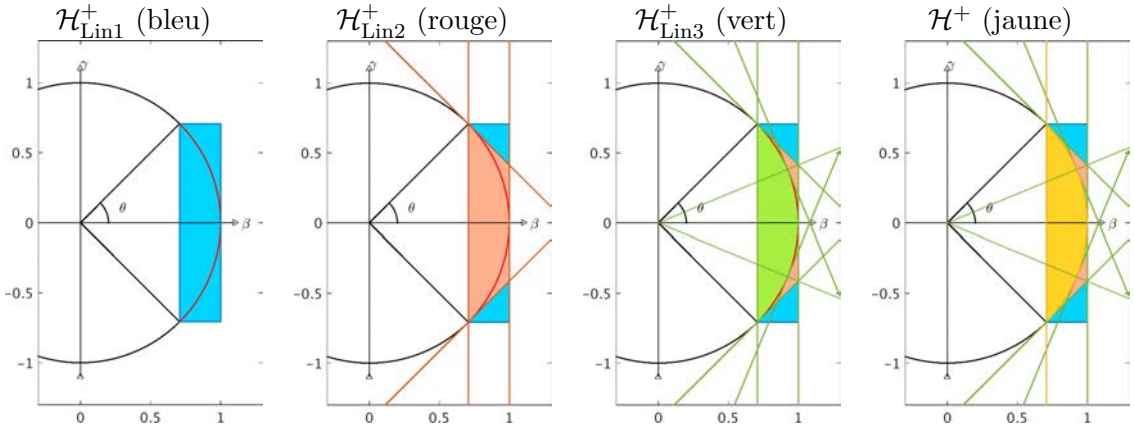



FIG. 2.9 – Représentation des différentes relaxations convexes de l'espace $\mathcal{S}^+$ à $\alpha > 0$ fixé dans le cas de l'interpolation polaire. De gauche à droite : de la relaxation la plus grande à l'enveloppe convexe.

Nous proposons ici 3 relaxations convexes de l'espace $\mathcal{S}^+$. La première fait partie des contraintes polyhédrales, comme définies en § 2.1.3, dans le sens où à $\alpha$ fixé, l'espace associé est un rectangle.

$$\mathcal{H}^+_{\text{Lin1}} = \left\{ \tilde{\boldsymbol{x}} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \in \mathbb{R}^3, \quad \begin{array}{ccccc} & & \alpha \geq 0 & & \\ r\alpha\cos\theta & \leq & \beta & \leq & r\alpha \\ -r\alpha\sin\theta & \leq & \gamma & \leq & r\alpha\sin\theta \end{array} \right\}$$

Pour le deuxième espace $\mathcal{H}^+_{\text{Lin2}}$, on contraint maintenant la composante $\gamma$ avec les tangentes en $\mathcal{S}^+$ aux points $(r\alpha\cos\theta, \pm r\alpha\sin\theta)$ :

$$\mathcal{H}^+_{\text{Lin2}} = \left\{ \tilde{\boldsymbol{x}} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \in \mathbb{R}^3, \quad \begin{array}{ccccc} & & \alpha \geq 0 & & \\ r\alpha\cos\theta & \leq & \beta & \leq & r\alpha \\ \dfrac{-r}{\sin\theta}\alpha + \dfrac{1}{\tan\theta}\beta & \leq & \gamma & \leq & \dfrac{r}{\sin\theta}\alpha - \dfrac{1}{\tan\theta}\beta \end{array} \right\}$$

Pour le troisième espace $\mathcal{H}^+_{\text{Lin3}}$, on ajoute deux nouvelles contraintes à l'aide des tangentes à l'arc en $(r\alpha\cos\frac{\theta}{2}, \pm r\alpha\sin\frac{\theta}{2})$ :

$$\mathcal{H}^+_{\text{Lin3}} = \left\{ \widetilde{\boldsymbol{x}} = \begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} \in \mathbb{R}^3, \quad \begin{array}{rcccl} & & \alpha \geq 0 & & \\ & r\alpha\cos\theta \leq & \beta & \leq & r\alpha \\ \frac{-r}{\sin\theta}\alpha + \frac{1}{\tan\theta}\beta \leq & \gamma & \leq & \frac{r}{\sin\theta}\alpha - \frac{1}{\tan\theta}\beta \\ \frac{-r}{\sin\frac{\theta}{2}}\alpha + \frac{1}{\tan\frac{\theta}{2}}\beta \leq & \gamma & \leq & \frac{r}{\sin\frac{\theta}{2}}\alpha - \frac{1}{\tan\frac{\theta}{2}}\beta \end{array} \right\} \tag{2.29}$$

Ce dernier espace est très proche de l'enveloppe convexe $\mathcal{H}^+$, comme on le constate en Figure 2.9. $\mathcal{H}^+$ est en fait le cas limite de $\mathcal{H}^+_{\text{Lin}\infty}$ où on aurait ajouté une infinité de contraintes via des tangentes en l'arc $\mathcal{S}^+$. Nous faisons donc la supposition que cette relaxation n'aura que très peu d'influence sur la performance de l'estimation. En contre-partie, un gain important en terme de coût de calcul est prévisible, comme la contrainte quadratique initiale, $\beta^2 + \gamma^2 \leq r^2\alpha^2$ est remplacée par 3 (ou 5) contraintes linéaires. Nous confirmerons ces suppositions en § 2.5.

### 2.4.2 Approximation par K-SVD

L'approximation par K-SVD a été proposée par [KYHP14], et a été très peu utilisée par la suite. Elle se distingue des autres approximations linéaires existantes (interpolation polaire, approximation de Taylor) notamment par un espace image $\mathcal{S}^+$ discret. Dans cette section, nous allons d'abord revenir plus en détails sur la définition de l'approximation linéaire, pour en déduire un espace de contraintes convexe. Nous proposerons ensuite une extension à cette approximation, que nous nommerons *Approximation par K-SVD continue*, qui permet d'obtenir un espace image $\mathcal{S}^+$ continu. Nous étudierons les propriétés de cette approximation. Cette approximation ne sera étudiée que pour des atomes $\boldsymbol{h}(\delta)$ à valeurs réelles. Le cas des signaux à valeurs complexes n'a pas été étudié. Comme nous l'avons vu en § 2.2, il est suffisant d'étudier l'espace image $\mathcal{S}^+$ et sa relaxation convexe $\mathcal{H}^+$, puisque l'espace des contraintes s'écrit à partir de ces espaces pour $\mathcal{D}_x = \mathbb{R}$.

#### Définition de l'approximation linéaire par K-SVD

Le principe proposé par [KYHP14] est de construire un sous-dictionnaire $\mathbf{F}_P = [\boldsymbol{h}(\delta_p)]_p$ de dimension $N \times P$, avec $\delta_p \in \mathcal{I}^P_\Delta$ (ensemble regroupant les $P$ discrétisations de $\delta$ sur $\mathcal{I}_\Delta$). La décomposition en valeurs singulières (SVD) de $\mathbf{F}_P$ s'écrit alors :

$$\mathbf{F}_P = \mathbf{U}\mathbf{S}\mathbf{V}^T \text{ où } \left\{ \begin{array}{ll} \mathbf{U} & (\text{dim. } N \times N) \\ \mathbf{V} & (\text{dim. } P \times P) \\ \mathbf{S} & (\text{dim. } N \times P) \end{array} \right. \begin{array}{l} \\ \\ \text{matrices orthonormales des vecteurs singuliers} \\ \text{matrice diagonale des valeurs singulières} \end{array}$$

La matrice $\mathbf{S}$ est diagonale, dans le sens où seuls ses coefficients diagonaux, appelés *valeurs singulières*, peuvent être non nuls ; ces dernières sont positives ou nulles et *rangées dans l'ordre décroissant* ; nous les notons $\{\lambda_p\}_p \in \mathbb{R}_+^{\min(P,N)}$. La matrice $\mathbf{F}_P$ peut alors être approchée en ne prenant en compte que les $K < \min(P, N)$ plus grandes valeurs singulières de la décomposition. En notant $\boldsymbol{u}_k$ les colonnes de $\mathbf{U}_P$ et $\boldsymbol{v}_k$ celles de $\mathbf{V}$, on obtient :

$$\mathbf{F}_P = \sum_{k=1}^{\min(P,N)} \lambda_k \boldsymbol{u}_k \boldsymbol{v}_k^T \approx \mathbf{F}_K = \sum_{k=1}^{K} \lambda_k \boldsymbol{u}_k \boldsymbol{v}_k^T \quad (\text{dim. } N \times P)$$

La matrice $\mathbf{F}_K$ est la meilleure approximation de rang $K$ de la matrice $\mathbf{F}_P$ au sens de la norme de Frobenius. L'atome continu $\boldsymbol{h}(\delta_p)$, qui correspond à la colonne $p$ de $\mathbf{F}_P$, peut alors être approché

par la colonne $p$ de $\mathbf{F}_K$, qui s'exprime comme la combinaison des $K$ premiers vecteurs singuliers $\boldsymbol{u}_k$ de coefficients $z_{p,k} = \lambda_k v_{p,k}$, où $v_{p,k}$ correspond à la ligne $p$ et à la colonne $k$ de $\mathbf{V}$ :

$$\boldsymbol{h}(\delta_p) \approx \sum_{k=1}^{K} \lambda_k v_{p,k} \boldsymbol{u}_k = \widetilde{\mathbf{h}} \Psi(\delta_p) \quad \Rightarrow \quad \left\{ \begin{array}{rcl} \widetilde{\mathbf{h}} & = & [\boldsymbol{u}_1, \dots, \boldsymbol{u}_K] \\ \boldsymbol{z}_p = \Psi(\delta_p) & = & [\lambda_1 v_{p,1}, \dots, \lambda_K v_{p,K}]^T \in \mathbb{R}^K \end{array} \right.$$

L'interpolateur $\Psi(\delta_p)$ ne s'exprime pas analytiquement en fonction de $\delta_p$ ; nous le noterons par la suite $\boldsymbol{z}_p$, où $z_{p,k}$ représentera sa $k^{\text{e}}$ composante.

*Remarque : Une propriété intéressante de l'atome interpolé* $\widetilde{\mathbf{h}}$ *est que ses colonnes sont orthonormées, puisque ce sont des vecteurs singuliers de la matrice* $\mathbf{F}_P$.

Le changement de variables définissant l'approximation linéaire s'écrit $\widetilde{\boldsymbol{x}} = g(x, \delta_p) = x\boldsymbol{z}_p$ et l'espace image $\mathcal{S}^+$ est la collection discrète des $P$ hyperdroites issues de l'origine et passant par les points $\boldsymbol{z}_p$ de $\mathbb{R}^K$ :

$$\mathcal{S}^+ = \left\{ \widetilde{\boldsymbol{x}} \in \mathbb{R}^K, \ \exists\, p \in [\![1; P]\!], x \geq 0, \ \widetilde{\boldsymbol{x}} = x\boldsymbol{z}_p \right\}$$

Cet espace étant discret, il ne peut pas être convexe.

### Étude de la relaxation convexe $\mathcal{H}^+$ de $\mathcal{S}^+$

On cherche à définir un espace $\mathcal{H}^+ \supset \mathcal{S}^+$ convexe pouvant servir d'espace de contraintes. Pour chaque $\widetilde{\boldsymbol{x}} \in \mathcal{S}^+$, il existe $(x, p) \in (\mathbb{R}_+ \times [\![1; P]\!])$ tel que $\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{x}}_p = x\boldsymbol{z}_p$, c'est à dire que chaque composante s'écrit $\widetilde{x}_{p,k} = xz_{p,k}$. Contrairement aux approximations de Taylor ou à l'interpolation polaire, il n'existe pas de composante $\widetilde{x}_k$ qui contient uniquement l'amplitude de départ. Pour écrire des contraintes qui ne dépendent pas de l'amplitude $x$, on peut alors choisir, comme le propose [KYHP14], une composante libre (généralement la première composante $\widetilde{x}_{p,1}$, puisque c'est celle qui contient le plus d'énergie) puis considérer le rapport entre $\widetilde{x}_{p,k}$ et $\widetilde{x}_{p,1}$ qui ne dépend pas de $x$ :

$$\forall k \geq 2, \ \frac{\widetilde{x}_{p,k}}{\widetilde{x}_{p,1}} = \frac{xz_{p,k}}{xz_{p,1}} = \frac{z_{p,k}}{z_{p,1}} = \zeta_{p,k} \quad \text{pour } \widetilde{\boldsymbol{x}}_p = x\boldsymbol{z}_p \in \mathcal{S}^+ \text{ et } x > 0$$

Le coefficient $\zeta_{p,k}$ correspond au coefficient directeur de la droite passant par l'origine et le point $(z_{p,1}, z_{p,k})$ dans le plan $(\widetilde{x}_1, \widetilde{x}_k)$. À partir de cette observation, [KYHP14] propose de considérer l'espace suivant[18] comme espace des contraintes :

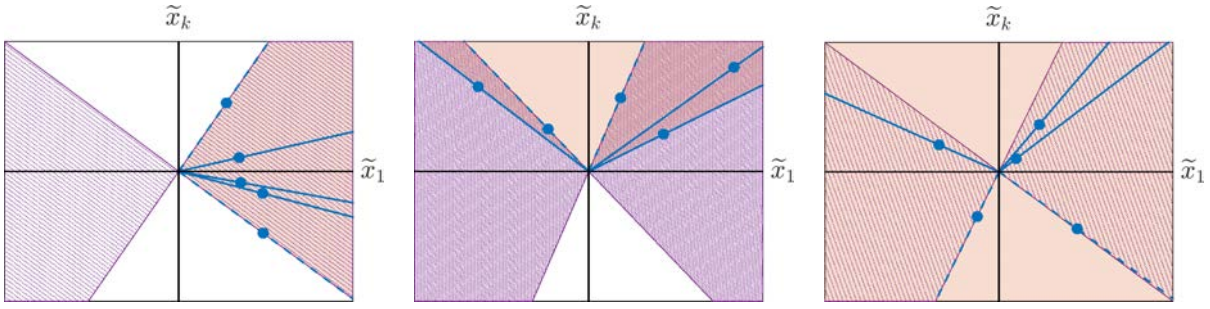
$$\mathcal{C}_0^+ = \left\{ \widetilde{\boldsymbol{x}} \in \mathbb{R}^K, \ \forall k > 1, \ \min_p \zeta_{p,k} = \zeta_{\min,k} \leq \frac{\widetilde{x}_k}{\widetilde{x}_1} \leq \zeta_{\max,k} = \max_p \zeta_{p,k} \right\}$$

Notons que ces contraintes sont polyhédrales, comme celles proposées pour l'approximation Minimax par [PPZ17]. Pour ce dernier, elles permettaient de définir un espace de contraintes convexes, *car la première composante* $\widetilde{x}_1$ *peut être contrainte à être positive.* Cependant, ce n'est pas le cas pour l'approximation par K-SVD. En effet, la première composante de l'interpolateur $\Psi(\delta_p)$ s'écrit $z_{p,1} = \lambda_1 v_{p,1}$ et son signe est donc le même que celui de l'élément $v_{p,1}$ de la matrice $\mathbf{V}$ de la K-SVD, qui peut être différent suivant le point $p$ considéré. En Figure 2.10, nous

[18] En fait, [KYHP14] propose d'encadrer le rapport $\widetilde{x}_k/\widetilde{x}_1$ par les valeurs minimale et maximale (en $p$) de $z_{p,k}$, plutôt que $z_{p,k}/z_{p,1}$. Cependant, cela n'est pas cohérent puisque cette contrainte n'est pas forcément vérifiée pour les points $\boldsymbol{z}_p$. Il s'agit sûrement d'une erreur de notation ou d'une hypothèse sur le fait que $z_{p,1}$ pourrait être égal à 1 ; nous avons donc décidé de considérer une contrainte plus cohérente.

représentons 3 exemples différents pour illustrer quelques cas typiques. Dans chaque cas, nous avons projeté dans le plan $(\widetilde{x}_1, \widetilde{x}_k)$, les $P = 5$ points $\Psi(\mathcal{I}_\Delta^5)$ de coordonnées $(z_{p,1}, z_{p,k})$ ainsi que les espaces $\mathcal{S}^+$, $\mathcal{C}_0^+$ et $\mathcal{H}_0^+$, l'enveloppe convexe de la projection de $\mathcal{S}^+$ dans le plan $(\widetilde{x}_1, \widetilde{x}_k)$. Une première remarque s'impose : l'espace $\mathcal{C}_0^+$ englobe toujours $\mathcal{S}^+$ mais n'est convexe dans aucun cas.

- Dans l'exemple (a), il suffirait de contraindre la première composante $\widetilde{x}_1$ à être positive, puisque $z_{p,1} > 0$ pour tout $p$ ; on aurait alors $\mathcal{S}^+ \subset \mathcal{C}_0^+ = \mathcal{H}_0^+$.
- Dans l'exemple (b), le signe de $z_{p,1}$ dépend du point considéré, on ne peut donc pas imposer $\widetilde{x}_1$ à être de signe constant. Notons cependant que l'espace $\mathcal{H}_0^+$ peut se décrire avec des contraintes linéaires : $\widetilde{x}_k \geq \zeta_i \widetilde{x}_1$, pour $\zeta_i = \min_{p \,|\, z_{p,1}>0} \zeta_{p,k}$ et $\zeta_i = \max_{p \,|\, z_{p,1}<0} \zeta_{p,k}$.
- Dans l'exemple (c), les points $(z_{p,1}, z_{p,k})$ ne sont pas contenus dans un demi plan ; l'enveloppe convexe de $\mathcal{S}^+$ est donc l'ensemble du plan 2D, ce qui revient à n'imposer aucune contrainte sur $\widetilde{x}_k$.




Exemple (a) Exemple (b) Exemple (c)

FIG. 2.10 – Illustration de l'espace $\mathcal{C}_0^+$ dans le plan $(\widetilde{x}_1, \widetilde{x}_k)$ pour 3 cas différents, chacun construit à partir $P = 5$ points $\boldsymbol{z}_p$. Pour chaque cas : les points bleus représentent les points $(z_{p,1}, z_{p,k})$, les droites bleues représentent l'espace engendré par ces points ($\mathcal{S}^+$ projeté dans le plan $(\widetilde{x}_1, \widetilde{x}_k)$), les droites en pointillées sont celles ayant les coefficients directeurs le plus faible $(\zeta_{\min,k})$ et le plus élevé $(\zeta_{\max,k})$ parmi les 5 droites, l'espace hachuré en violet correspond à $\mathcal{C}_0^+$ et celui coloré en rouge représente $\mathcal{H}_0^+$.

Pour définir un espace de contraintes convexe à partir des points $\boldsymbol{z}_p$ et des coefficients directeurs associés $\zeta_{p,k}$, nous proposons de définir des contraintes polyhédrales entre chaque $(\widetilde{x}_1, \widetilde{x}_k)$ avec la procédure suivante :

(a) Si l'ensemble des $z_{p,1}$ sont de même signe (exemple (a)), on impose[19] $\widetilde{x}_1 \geq 0$ et la contrainte prend la forme proposée par [KYHP14] :

$$\zeta_{\min,k}\, \widetilde{x}_1 \leq \widetilde{x}_k \leq \zeta_{\max,k}\, \widetilde{x}_1$$

où $\zeta_{\min,k}$ et $\zeta_{\max,k}$ sont les coefficients directeurs minimal et maximal.

Si les $z_{p,1}$ sont de signes différents (exemples (b-c)), on calcule alors le plus grand angle entre deux points de $\{\boldsymbol{z}_p\}_p$.

(b) Si cet angle est inférieur à $\pi$ (exemple (b)), on calcule :

$$\zeta^+_{\min,k} = \min_{p \,|\, z_{p,1}>0} \zeta_{p,k} \quad \text{et} \quad \zeta^-_{\max,k} = \max_{p \,|\, z_{p,1}<0} \zeta_{p,k}$$

[19] Si les $z_{p,1}$ sont négatifs, il suffit de les multiplier par (-1) ainsi que la colonne de l'atome interpolé associé puisque $z_{p,1}\boldsymbol{u}_1 = (-z_{p,1})(-\boldsymbol{u}_1)$.

$$\text{Soit} \quad \forall p,\ z_{p,k} \geq \zeta^{+}_{\min,k} z_{p,1} \quad \text{alors on impose} \quad \begin{cases} \widetilde{x}_k & \geq & \zeta^{+}_{\min,k}\, \widetilde{x}_1 \\ \widetilde{x}_k & \geq & \zeta^{-}_{\max,k}\, \widetilde{x}_1 \end{cases} ;$$

$$\text{Soit} \quad \forall p,\ z_{p,k} \leq \zeta^{+}_{\min,k} z_{p,1} \quad \text{alors on impose} \quad \begin{cases} \widetilde{x}_k & \leq & \zeta^{+}_{\min,k}\, \widetilde{x}_1 \\ \widetilde{x}_k & \leq & \zeta^{-}_{\max,k}\, \widetilde{x}_1 \end{cases} .$$

(c) Enfin, si cet angle est supérieur à $\pi$, alors nous sommes dans le cas de l'exemple (c) et aucune contrainte n'est imposée sur $\widetilde{x}_k$. On dira dans ce cas que les points $(z_{p,1}, z_{p,k})$ « recouvrent plus d'un demi-espace ».

*Il est important de noter que l'espace des contraintes ainsi défini respecte l'implication de nullité* uniquement *dans le cas (a) : quand on peut contraindre les premières composantes $\widetilde{x}_1$ à être positives. En effet, les contraintes définies dans le cas (b) n'impliquent pas que $\widetilde{x}_k = 0$ si $\widetilde{x}_1 = 0$ (dans ce cas, $\widetilde{x}_k$ est juste contraint à être positif ou négatif). Par conséquent, les formulations des problèmes avec pénalisation $\ell_1$ et contraintes $\ell_0$, proposées en § 2.3 sous hypothèse d'implication de nullité, ne sont applicables que dans le cas (a). Pour les autres cas, il est nécessaire d'imposer la parcimonie de façon structurée sur l'ensemble des composantes de $\widetilde{\boldsymbol{x}}$. Comme nous l'avons dit précédemment, cela peut se faire sans grande difficulté avec la reformulation MIP (résolution de type $\ell_0$), mais cette prise en compte est difficilement envisageable pour la résolution avec pénalisation $\ell_1$ dans le cas où les amplitudes sont à valeurs réelles, car la contrainte de disjonction ne sera sûrement pas respectée.*

Bien que nous ayons maintenant correctement défini un espace de contraintes convexe, nous remarquons que cet espace est bien plus « grand » que l'espace $\mathcal{S}^+$, puisque l'on passe d'un nombre fini d'hyperdroites à un volume de dimension $K$, ce qui est contraire aux recommandations que nous avons formulées en § 2.2 pour obtenir une bonne identification des paramètres initiaux. Le paragraphe suivant, via un passage à la limite $P \to \infty$, propose une solution à ce problème.

**Approximation par K-SVD continue : passage d'hyperdroites à l'hyperespace.**

Afin d'obtenir une meilleure approximation des $\boldsymbol{h}(\delta)$ via K-SVD, il semble logique de prendre le plus de points $P$ possible dans $\mathcal{I}_\Delta$ afin de prendre en compte l'ensemble des $\{\boldsymbol{h}(\delta)\}_{\delta\in\mathcal{I}_\Delta}$. La Figure 2.11 permet d'appréhender l'évolution des points $\boldsymbol{z}_p = \Psi(\delta_p)$ quand le nombre $P$ de $\delta_p$ pris pour calculer la SVD augmente, dans le cas $K = 3$. Plus exactement, à partir du modèle $\boldsymbol{h}(\delta)$ et de l'espacement $\Delta = \Delta_0$ que nous utiliserons pour les tests en § 2.5, on crée la sous-grille $\mathcal{I}^P_\Delta$ en discrétisant $\mathcal{I}_\Delta$ de la façon suivante : pour $P = 3$, on choisit $\mathcal{I}^3_\Delta = [-\frac{\Delta}{2}, 0, \frac{\Delta}{2}]$, puis on augmente $P$ en ajoutant un point au milieu de chaque intervalle. À partir de la sous-grille $\mathcal{I}^P_\Delta$, on peut alors calculer le sous-dictionnaire $\mathbf{F}_P$, sa SVD, et extraire les $P$ points $\boldsymbol{z}_p$. Dans notre exemple, les points se dissocient dans l'espace, mais semblent tendre vers une courbe continue, notamment quand on regarde les points pour $P = 65$. Par ailleurs, on peut constater que les points $\boldsymbol{z}_p$ semblent évoluer continument en fonction du $\delta_p$ auquel ils correspondent.

Au regard de ces observations, on peut penser qu'en prenant $P$ suffisamment grand, on aura une approximation par K-SVD de l'ensemble des atomes $\{\boldsymbol{h}(\delta)\}_{\delta\in\mathcal{I}_\Delta}$, ce que démontre le théorème suivant.

**Théorème 2 (K-SVD continue)** *Soit $\mathbf{F}_P$ la matrice de dimension $N \times P$ dont les colonnes sont constituées des vecteurs $[\boldsymbol{h}(\delta_p)]_{p=1..P}$, avec $\delta_p \in \mathcal{I}^P_\Delta$ et $P > N$. Soit $\mathbf{U}_P\mathbf{S}_P\mathbf{V}_P^T = \mathbf{F}_P$ sa décomposition en valeurs singulières compacte (SVDc), où les matrices $\mathbf{U}_P$, $\mathbf{S}_P$ et $\mathbf{V}_P$ sont respectivement de dimension $N \times N$, $N \times R$ et $R \times P$, où $R$ est le nombre de valeurs singulières non nulles de $\mathbf{F}_P$ et où les valeurs singulières sont rangées par ordre décroissant dans $\mathbf{S}_P$.*
*Alors, on a les résultats suivants :*

1. $\lim_{P\to\infty} \mathbf{U}_P = \mathbf{U}_\infty$, *où* $\mathbf{U}_\infty$ *vient de la SVD de la matrice* $\mathbf{B}_\infty = \mathbf{U}_\infty \mathbf{S}_\infty \mathbf{U}_\infty^T$, *telle que* $\mathbf{B}_\infty = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \boldsymbol{h}(\delta)\,\boldsymbol{h}(\delta)^T d\delta$.
2. $\boldsymbol{h}(\delta) \approx \mathbf{U}_{K\infty}\Psi(\delta)$ *est l'approximation limite quand* $P \to \infty$ *de* $\boldsymbol{h}(\delta)$ *via K-SVD, où* $\mathbf{U}_{K\infty}$ *contient les* $K$ *premières colonnes de* $\mathbf{U}_\infty$ *et* $\Psi(\delta) = \mathbf{U}_{K\infty}^T \boldsymbol{h}(\delta)$.

La démonstration de ce théorème est décrite en Annexe A.2. Elle s'inspire principalement des travaux réalisés dans [KYO18] pour l'évaluation de caractéristiques de réseaux de microphones via une « SVD de dimension infinie ». Elle repose sur le fait que l'on peut calculer la matrice $\mathbf{U}_P$ à partir de la diagonalisation de la matrice $\mathbf{A}_P = \mathbf{F}_P\mathbf{F}_P^T = \mathbf{U}_P\mathbf{S}_P^2\mathbf{U}_P^T$, de dimension $N \times N$. Ce théorème permet de retourner dans le paradigme des approximations usuelles (Taylor, polaire) : on passe d'un interpolateur $\Psi_P(\delta_p)$ discret à un interpolateur continu $\Psi(\delta)$ ; d'un espace $\mathcal{S}^+$ discret composé des hyperdroites $x\Psi(\delta_p)$ à un hyperespace $\mathcal{S}^+$ continu. Pour cette raison, nous appelons cette approximation *K-SVD continue*, pour la différencier de l'approximation par K-SVD proposée par [KYHP14].

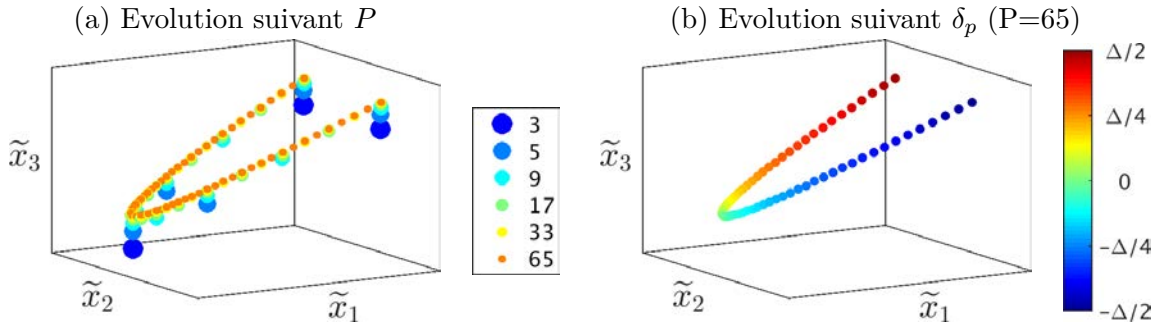



FIG. 2.11 – Représentations des points $\boldsymbol{z}_p$ en discrétisant de plus en plus finement l'intervalle $\mathcal{I}_\Delta$. (a) On augmente progressivement $P$. (b) Pour $P = 65$, on regarde la répartition des points suivant les $\delta_p$ auxquels ils correspondent. Modèle utilisé pour les tests en § 2.5.

L'intérêt de considérer cette approximation par K-SVD continue est double :
- Aucune relaxation entre les hyperdroites n'est plus nécessaire.
- Le gain en mémoire de calcul est important : en K-SVD discrète, l'atome interpolé $\widetilde{\mathbf{h}} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_K]$ se calcule via la K-SVD de la matrice $\mathbf{F}_P$ de dimension $N \times P$ avec $P$ arbitrairement grand ; en K-SVD continue, $\widetilde{\mathbf{h}} = \mathbf{U}_{K,\infty}$ se calcule en diagonalisant la matrice $\mathbf{B}_\infty$ de dimension $N \times N$, elle-même calculée par intégration (numérique ou analytique, si on peut calculer une primitive de la fonction à intégrer).

Nous verrons en § 2.6 que ce dernier avantage est particulièrement intéressant dans le cas où l'atome continu $\boldsymbol{h}(\cdot)$ dépendrait de deux paramètres non linéaires (voire davantage), et où le nombre $P$ de points nécessaires pour obtenir une approximation satisfaisante devient exponentiel.

L'approximation linéaire d'un atome $\boldsymbol{h}(\delta)$ par K-SVD continue s'écrit donc de la façon suivante :

$$x\,\boldsymbol{h}(\delta) \approx \widetilde{\mathbf{h}}\,\widetilde{\boldsymbol{x}} \text{ avec } \widetilde{\boldsymbol{x}} = g(x,\delta) = x\Psi(\delta) \text{ et } \begin{cases} \widetilde{\mathbf{h}} &= \mathbf{U}_{K,\infty} = [\boldsymbol{u}_{1,\infty}, \ldots, \boldsymbol{u}_{K,\infty}] \\ \Psi(\delta) &= \mathbf{U}_{K,\infty}^T \boldsymbol{h}(\delta) \end{cases}$$

Tout comme les approximations continues classiques telles que les approximations de Taylor ou l'interpolation polaire, il est nécessaire d'étudier le caractère bijectif du changement de variables $g(x,\delta) = x\Psi(\delta)$ sur $\mathbb{R}_+ \times \mathcal{I}_\Delta$ ainsi que la convexité de l'espace image $\mathcal{S}^+$ afin de pouvoir définir

un espace de contraintes convexes. L'extension aux atomes à amplitudes réelles/complexes se fera alors comme décrit en § 2.2.

Néanmoins, nous n'avons pas poussé cette étude de façon théorique, bien qu'un certain nombre de postulats, que nous énoncerons à la fin de cette section, mériteraient d'être démontrés. Pour définir l'espace des contraintes convexe, nous utiliserons la méthodologie que nous avons décrite au paragraphe précédent pour l'approximation par K-SVD discrète, en discrétisant l'interpolateur continu $\Psi(\delta)$ afin d'obtenir $P$ points $\{\boldsymbol{z}_p\}_p = \{\Psi(\delta_p)\}_p$. Une étude plus poussée permettrait certainement de trouver des contraintes permettant de définir une relaxation $\mathcal{H}^+$ plus proche de l'espace image $\mathcal{S}^+$.

Nous terminons par observer en Figure 2.12, dans le cas pratique de l'atome continu utilisé pour nos tests en § 2.5, l'allure de l'interpolateur $\Psi(\delta)$ et de ses *dilatations*, permettant de caractériser l'espace image $\mathcal{S}^+$. Cela nous permet de mieux appréhender cette approximation par K-SVD continue et d'émettre des postulats sur la bijectivité du changement de variables $g$ et sur la convexité de l'espace image. Notons que sur ces observations, on peut déduire la non-bijectivité du changement de variables quand des dilatations de l'interpolateur se coupent (alors un point $\widetilde{\boldsymbol{x}}$ peut s'identifier à deux couples $(x, \delta)$ différents).

Pour $\Delta$ faible (égal à $\Delta_0$, la valeur utilisée en § 2.5), la première composante $\Psi_1(\delta)$ de l'interpolateur est toujours positive, et dans ce cas l'espace $\mathcal{S}^+$ est convexe pour $K = 2$. Les différentes dilatations ne se coupent pas quel que soit $K$ : le changement de variables semble bijectif. Pour $\Delta = 4\Delta_0$ la première composante de $\Psi(\delta)$ change de signe ; le changement de variables semble bijectif pour $K = 3$, mais pas pour $K = 2$. Enfin, pour $\Delta = 8\Delta_0$, il n'est bijectif ni pour $K = 3$ ni pour $K = 2$. Ces observations nous ont mené aux postulats suivant :

1. Pour $\Delta$ suffisamment faible, le changement de variables $g$ est bijectif sur $\mathbb{R}_+ \times \mathcal{I}_\Delta$, la première composante de l'interpolateur est de signe constant et l'espace $\mathcal{S}^+$ est convexe pour $K = 2$.

2. Il existe un $\Delta$ limite, au delà duquel les trois propriétés précédentes ne sont plus vérifiées.

Des études plus approfondies pourraient mener à la démonstration de ces postulats.

*Remarque : En pratique, on peut définir une heuristique pour déterminer si le changement de variables est bijectif sur $\mathbb{R} \times \mathcal{I}_\Delta$ pour $K = 2$. Il suffit de regarder si l'angle $\varphi(\delta) = atan2(\Psi_2(\delta), \Psi_1(\delta))$ est de progression monotone relativement*[20] *à $\varphi(-\Delta/2)$ (bijectivité sur $\mathbb{R}_+ \times \mathcal{I}_\Delta$) et si $|\varphi(\Delta/2) - \varphi(-\Delta/2)|$ ne dépasse pas $\pi$ (bijectivité sur $\mathbb{R} \times \mathcal{I}_\Delta$).*

*Remarque 2 : Pour les tests effectués en section suivante, l'approximation par K-SVD continue donne un espace image similaire à celui illustré en Figure 2.12 pour $\Delta = \Delta_0$ pour chaque indice $j$ de la grille : le changement de variables est bijectif (au moins pour $K = 2$) et les premières composantes $\Psi_1(\delta)$ sont de signes constants, si bien que l'espace des contraintes peut être défini comme dans le cas (a) vu précédemment, respectant l'implication de nullité. Cela signifie que les formulations (2.21) (pénalisation $\ell_1$) et (2.17) (formulation MIP/$\ell_0$) peuvent être utilisées directement pour résoudre le problème d'APLsc.*

[20] Comme la fonction atan2$(\cdot,\cdot)$ renvoie un angle compris entre $-\pi$ et $\pi$, il est nécessaire de considérer des angles relatifs pour éviter les effets de bord autour de $\pm\pi$.

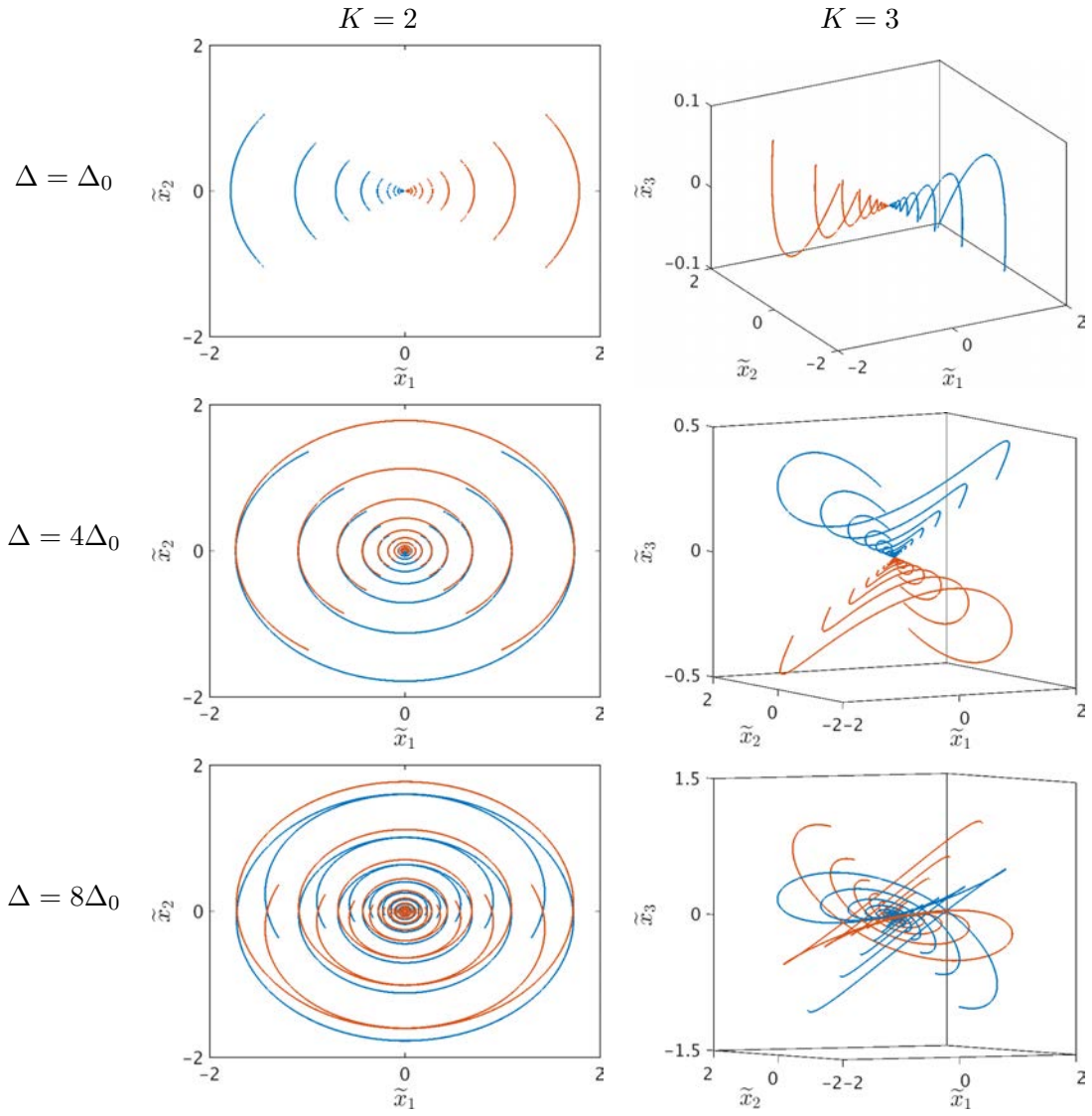



FIG. 2.12 – Illustration de l'espace $\mathcal{S}^{\mathbb{R}}$ pour $K = 2$ (gauche) et $K = 3$ (droite) et pour $\Delta = \Delta_0$, $4\Delta_0$ et $8\Delta_0$. Les lignes bleues (resp. rouges) correspondent à des dilatations positives (resp. négatives) de $\Psi(\delta)$.

**Bilan**

L'approximation par K-SVD est une approximation linéaire intéressante dans le cadre de l'interpolation de dictionnaire. Notamment, comme la K-SVD d'une matrice $\mathbf{F}$ est sa meilleure approximation de rang $K$ au sens de la norme de Frobenius, on peut considérer qu'à $K$ fixé, il s'agit de l'approximation qui donnera la plus faible erreur d'approximation en moyenne sur $\mathcal{I}_\Delta$, soit plus faible que l'approximation de Taylor à l'ordre 1 pour $K = 2$ et que l'interpolation polaire pour $K = 3$. Pourtant, cette approximation n'a été que très peu utilisée ; sûrement à cause de la difficulté de proposer une relaxation convexe, du coût calculatoire nécessaire pour calculer la K-SVD discrète de la matrice $\mathbf{F}_P$ avec $P$ grand, ou encore car *a priori*, elle ne permet de reconstruire qu'un décalage $\delta_p$ discret. À travers les contributions présentées dans ce chapitre, soit notamment la définition correcte d'un espace de contraintes convexe et le passage à l'approximation par K-SVD continue, cette approximation semble donc prometteuse. Par ailleurs, elle ne nécessite pas d'hypothèses particulières sur le modèle contrairement à l'interpolation polaire qui nécessite l'invariance par translation.

## 2.5 Application sur les données de déconvolution impulsionnelle

Les données tests de déconvolution de train d'impulsion présentées en § 1.5.2 se prêtent bien au test de nos propositions pour l'interpolation de dictionnaire. En effet, le modèle $\mathbf{h}_j(\delta_j)\boldsymbol{x}_j$ associé aux données respecte les conditions dans lesquelles l'interpolation de dictionnaire peut être (pour l'instant) appliquée : les amplitudes sont scalaires $\boldsymbol{x}_j = x_j(\text{dim. } 1)$ et les paramètres non linéaires ne sont constitués que du seul paramètre de localisation, la position de l'écho. Les atomes continus $\boldsymbol{h}_j(\delta_j)$ respectent l'invariance par translation, permettant l'utilisation de l'interpolation polaire. De plus, les amplitudes étant à valeurs réelles, nous allons pouvoir tester l'approche que nous avons proposée en § 2.2.2. Nous allons donc comparer les deux méthodes de résolution du problème d'APLsc pour $\mathcal{D}_x = \mathbb{R}$ : la formulation (2.21) pour la pénalisation $\ell_1$ associée à la réévaluation des amplitudes que nous proposons en (2.23), ainsi que la formulation MIP (2.17) de la résolution exploitant la pseudo-norme $\ell_0$. Par abus de langage, nous parlerons dans cette section de « résolution $\ell_1$ » et de « résolution $\ell_0$ ». De même, nous testerons les diverses interpolations de dictionnaire que nous avons étudiées (Taylor à l'ordre 1 (T1), polaire (Po), 2-SVD (K2) et 3-SVD (K3)). Notons qu'une partie des résultats présentés dans cette section sont issus de l'article de conférence [BC18], qui comparent les méthodes de résolution $\ell_1$ et $\ell_0$ avec l'interpolation polaire.

Pour rappel, les données sont composées de 200 modèles tests de $L = 4$ échos aléatoirement positionnés, avec 3 niveaux de bruits différents (RSB = 10, 20, 30 dB). La grille de recherche est constituée de $J = 100$ intervalles, tels que $\Delta = \Delta_0 = T_e = 1s$.

**Une première comparaison des erreurs d'approximation.** L'erreur d'approximation d'une interpolation donnée peut avoir une grande influence sur les résultats. En Figure 2.13 sont représentées les erreurs d'approximations théoriques moyennes pour chaque interpolation testée, pour $\Delta$ variant de $\Delta_0/10$ à $10\Delta_0$, à un indice $j$ quelconque que nous ne précisons pas par la suite. Plus précisément, pour un $\Delta$ fixé, on calcule l'erreur quadratique moyenne entre le modèle $\boldsymbol{h}(\delta)$ et son approximation $\widetilde{\mathbf{h}}\Psi(\delta)$ par intégration numérique selon $\delta$ sur $\mathcal{I}_\Delta$ :

$$\xi^{\text{moy}}(\Delta) = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \|\boldsymbol{h}(\delta) - \widetilde{\mathbf{h}}\Psi(\delta)\|^2 \,\mathrm{d}\delta$$

Sur la Figure 2.13, on voit l'augmentation de l'erreur moyenne suivant l'augmentation de $\Delta$. Pour $\Delta = \Delta_0$, on a logiquement :

$$\xi_{BP}^{\text{moy}} > \xi_{T1}^{\text{moy}} > \xi_{K2}^{\text{moy}} > \xi_{Po}^{\text{moy}} > \xi_{K3}^{\text{moy}}$$

Notamment, $\xi_{T1}^{\text{moy}} > \xi_{K2}^{\text{moy}}$ et $\xi_{Po}^{\text{moy}} > \xi_{K3}^{\text{moy}}$ illustre l'optimalité des approximations par K-SVD quand on cherche à approcher un signal avec une combinaison de $K$ vecteurs. De plus, l'erreur d'approximation est logiquement plus faible pour les approximations à 3 atomes (Po, K3) que pour les approximations à 2 atomes (T1, K2). Nous verrons que cette hiérarchie ne se transmettra pas nécessairement aux résultats des problèmes d'APLsc.

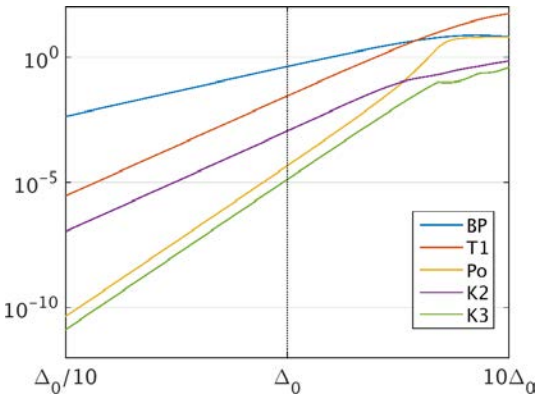



FIG. 2.13 – Erreur d'approximation théorique moyenne pour chaque interpolation testée, pour $\Delta$ variant de $\Delta_0/10$ à $10\Delta_0$. En bleu (BP) : sur grille ($\delta = 0$). En rouge (T1) : approximation de Taylor. En jaune (Po) : interpolation polaire. En violet (K2) : approximation par 2-SVD continue. En vert (K3) : approximation par 3-SVD continue.

Nous commençons par détailler davantage la mise en œuvre des méthodes de résolution et des approximations linéaires, avant de présenter les résultats.

### 2.5.1 Réglages des paramètres des méthodes pour les tests

#### Critère d'arrêt pour la comparaison des résolutions de type $\ell_1$ et $\ell_0$

Afin de pouvoir comparer les méthodes de résolution de type $\ell_1$ et $\ell_0$, il est nécessaire de définir un « critère d'arrêt » similaire dans les deux cas, permettant de régler le degré de parcimonie de la solution. En effet, nous avons vu en § 2.3 que chacune des deux méthodes de résolution dépend d'un hyperparamètre ($\lambda$ pour la résolution $\ell_1$ et $L_0$ pour la résolution $\ell_0$) réglant le compromis entre la parcimonie de la solution et l'erreur de reconstruction. Plus $L_0$ augmente - respectivement plus $\lambda$ diminue, plus le terme d'attache aux données ($\|\boldsymbol{y} - \widetilde{\mathbf{H}}\widetilde{\boldsymbol{x}}\|^2$ dans les deux cas) diminue.

Comme nous l'avons proposé dans l'article [BC18], on peut comparer les méthodes de la façon suivante : en partant du degrés de parcimonie le plus élevé soit $L_0 = 0$ et $\lambda = \lambda_{\max}$ tel que la solution $\widetilde{\boldsymbol{x}} = \mathbf{0}$, on peut diminuer la parcimonie (*i.e.* augmenter $L_0$ et diminuer $\lambda$) jusqu'à ce que l'erreur de reconstruction passe en dessous d'un seuil prédéfini.

Plus précisément, nous choisissons comme erreur de reconstruction l'erreur quadratique entre le signal $\boldsymbol{y}$ et le modèle original du problème d'APNL, évaluée en la solution $(\boldsymbol{x}, \boldsymbol{\nu})$ identifiée à partir de la solution $\widetilde{\boldsymbol{x}}$ du problème d'APLsc, et réduite au support de $\boldsymbol{x}$. Elle s'écrit donc $\|\boldsymbol{r}\|^2$ avec[21] $\boldsymbol{r} = \boldsymbol{y} - \overline{\mathbf{H}}(\overline{\boldsymbol{\nu}})\widehat{\boldsymbol{x}}^{\text{MC}}$. On peut alors définir un seuil à partir duquel le résidu $\boldsymbol{r}$ peut

[21] Les amplitudes sont réévaluées sur le dictionnaire $\mathbf{H}(\widehat{\boldsymbol{\delta}})$ par moindres carrés.

être assimilé à du bruit. Par exemple, en supposant que le bruit est gaussien centré de variance $\sigma^2$, on sait calculer $T$ tel que Proba $\left(\frac{\|\boldsymbol{r}\|^2}{\sigma^2} < T\right)$ =95% (quantile de la loi du $\chi^2$ à $N$ degrés de liberté). On s'arrête alors d'augmenter $L_0$ ou de diminuer $\lambda$ quand $\frac{\|\boldsymbol{r}\|^2}{\sigma^2} < T$. On dira par la suite qu'on atteint le critère d'arrêt quand la résidu atteint le « seuil du bruit ».

Notons que le choix de $\lambda_{\max}$ et le pas de décroissance associé n'est pas aussi évident que l'augmentation de $L_0$ à partir de 0. Nous avons choisi $\lambda_{\max}$ suivant sa définition pour les problèmes d'APL (voir par exemple [BC06]) :

$$\forall \lambda \geq \lambda_{\max} = \max_j \mathbf{H}^T \boldsymbol{y},\ \arg\min_{\boldsymbol{x}}(\|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2 + \lambda\|\boldsymbol{x}\|_1) = \mathbf{0}$$

où $\mathbf{H}$ est le dictionnaire discret, soit le dictionnaire continu exprimé en $\boldsymbol{\nu} = \ddot{\boldsymbol{\nu}}$. En pratique, la résolution $\ell_1$ du problème d'APLsc donnait toujours une solution à support nulle ($\widetilde{\boldsymbol{x}} = \mathbf{0}$) pour $\lambda \geq \lambda_{\max}$. À partir de ce $\lambda_{\max}$, on diminue progressivement $\lambda$ par pas de $0.01\lambda_{\max}$[22].

### Normalisation des atomes et conséquences sur les contraintes

Il est souvent nécessaire de normaliser les colonnes des atomes interpolés afin de minimiser les effets de changement d'échelle (sauf dans le cas de l'interpolation K-SVD, où les colonnes des atomes sont orthonormées par définition). Par conséquent, les colonnes des atomes s'écrivent $\widetilde{\widetilde{\mathbf{h}}}_{j,k} = \widetilde{\mathbf{h}}_{j,k}/\|\widetilde{\mathbf{h}}_{j,k}\|_2$. Cette normalisation a une influence directe sur la définition des espaces de contraintes, mais est facile à prendre en compte. En effet, l'approximation linéaire s'écrit alors :

$$x_j\, \boldsymbol{h}(\delta_j) \approx \widetilde{\widetilde{\mathbf{h}}}_j \widetilde{\widetilde{\boldsymbol{x}}}_j = \sum_k \widetilde{\widetilde{x}}_{j,k}\, \widetilde{\widetilde{\mathbf{h}}}_{j,k} \text{ où } \widetilde{\widetilde{x}}_{j,k} = \widetilde{x}_{j,k}\|\widetilde{\mathbf{h}}_{j,k}\|_2$$

Les contraintes doivent donc porter sur les amplitudes $\widetilde{\widetilde{x}}_{j,k}$ ; l'adaptation se fait simplement en remplaçant $\widetilde{x}j,k$ par $\widetilde{\widetilde{x}}_{j,k}/\|\widetilde{\mathbf{h}}_{j,k}\|_2$ dans les contraintes inégalités décrites pour chaque interpolation de dictionnaire (voir Tableau 2.1 pour les approximations de Taylor et polaire).

### Réglage du big-M

Dans le cas des résolutions de type $\ell_0$ via la reformulation en MIP, il est aussi nécessaire de régler le paramètre $M$, qui pour rappel permet d'imposer la nullité des amplitudes via la contrainte de bornes de type $-Mb_j \leq x_j \leq Mb_j$ (APL) ou $-Mb_j \leq \alpha_j \leq Mb_j$ (APLsc). Pour rappel de § 1.5.2, les amplitudes des atomes continus sont bornées puisqu'elles sont tirées aléatoirement avec $|x_j| \leq M_0 = 2$. Dans le cas de la résolution sur dictionnaire discret, il est donc naturel de choisir $M = M_0 = 2$.

Cependant, dans le cas des formulations pour les problèmes d'APLsc, la contrainte de big-M porte sur la première composante $\alpha_j = \widetilde{x}_{j,1}$ des amplitudes des atomes interpolés. Il est donc nécessaire de prendre en compte la relation entre les $x_j$ et les $\alpha_j$ via le changement de variables $g$.

Le Tableau 2.2 récapitule le réglage du paramètre $M$ pour différents types d'interpolation de dictionnaire : pour les approximations classiques (Taylor et polaire), la première composante $\widetilde{x}_{j,1}$ correspond à l'amplitude $x_j$ de l'atome continue ; il est cependant nécessaire de prendre en compte la normalisation des atomes. Pour l'approximation par K-SVD (pour qui l'espace des contraintes que nous avons défini en § 2.4.2 respecte l'implication de nullité pour l'atome continu de cette application), ce n'est généralement plus le cas ; mais on sait que la première composante $\widetilde{x}_1$ s'écrit $x\boldsymbol{u}_{\infty,1}^T \boldsymbol{h}(\delta)$. Le choix que nous avons fait pour $M$ provoque un certain relâchement sur les amplitudes $x$ reconstituées qui pourront être légèrement plus élevées que $M_0$.

[22] Si le critère d'arrêt n'est toujours pas atteint pour $\lambda = 0.01\lambda_{\max}$, on continue de le diminuer de façon géométrique en le multipliant par 0.99.

| Interpolation | Relation $\alpha \leftrightarrow x$ | $\lvert x\rvert \leq M_0 \Rightarrow \lvert\alpha\rvert \leq M$ |
|---|---|---|
| K-SVD | $\alpha = x(\boldsymbol{u}_{\infty,1}^T \boldsymbol{h}(\delta))$ | $M = M_0 \max_\delta \lvert\boldsymbol{u}_{\infty,1}^T \boldsymbol{h}(\delta)\rvert$ |
| Autres | $\alpha = x\lVert\widetilde{\mathbf{h}}_1\rVert_2$ | $M = M_0\lVert\widetilde{\mathbf{h}}_1\rVert_2$ |

TAB. 2.2 – Réglage du *big-M* suivant les différents cas d'interpolation.

**Logiciel de calcul et toolbox.**

Les diverses optimisations sont effectuées avec le logiciel IBM ILOG CPLEX V12.6.0, dont une version gratuite est disponible pour les étudiants et académiques. Elles sont lancées via une interface Matlab sur un ordinateur avec un processeur Intel Xeon *E5-2680 processors* (40 threads) avec des CPU tournant à 2.8 GHz. Dans le cas des résolutions $\ell_0$, les calculs peuvent être parallélisés sur 4 cœurs.

### 2.5.2 Comparaison des méthodes de résolution $\ell_1$ et $\ell_0$ avec interpolation polaire

Les résultats présentés ici sont issus de l'article [BC18]. Nous y comparons les résolutions $\ell_1$ et $\ell_0$ sur dictionnaire interpolé avec *interpolation polaire (avec contraintes quadratiques)* et sur dictionnaire discret (APL). Nous définissons les notations suivantes :

- $\mathcal{P}_{2+1}^{\mathcal{D}}$ : résolution $\ell_1$ sur dictionnaire discret.
- $\mathcal{P}_{2/0}^{\mathcal{D}}$ : résolution $\ell_0$ sur dictionnaire discret.
- $\mathcal{P}_{2+1}^{\mathcal{C}*}$ : résolution $\ell_1$ sur dictionnaire continu sans réévaluation.
- $\mathcal{P}_{2+1}^{\mathcal{C}}$ : résolution $\ell_1$ sur dictionnaire continu avec réévaluation.
- $\mathcal{P}_{2/0}^{\mathcal{C}}$ : résolution $\ell_0$ sur dictionnaire continu.

La réévaluation évoquée ici correspond à celle proposée en (2.23)(p. 58) permettant de corriger le biais induit par la pénalisation $\ell_1$. Ces méthodes ont été testées sur les 200 modèles décrits en § 1.5.2 pour des RSB de 10, 20 et 30 dB. Pour chaque résultat, nous avons calculé les indices ELR (pour rappel, ELR=1 si le support théorique est retrouvé avec une erreur inférieure à $\Delta/2$ pour l'estimation de chaque localisation) et ASD (qui indique une bonne estimation quand il est faible) décrits en § 1.5.2 (p.31), ainsi que la norme au carré du résidu $\zeta^2 = \|\boldsymbol{y} - \mathbf{D}\widehat{\boldsymbol{x}}^{\mathrm{MC}}\|^2$. Pour chaque résolution, nous utilisons le critère d'arrêt décrit précédemment.

Notons d'abord que l'ensemble des solutions des problèmes d'APLsc par la résolution $\ell_1$ respectait bien la *contrainte de disjonction*, décrite en § 2.2.2 et non imposée à l'aide de contraintes fermes dans la formulation (2.21).

La Figure 2.14 donne les résultats des diverses méthodes pour le signal présenté en Figure 1.4(b) (p. 31) avec RSB=10dB. Dans le cas de ce signal, trois échos de l'ondelette se recouvrent, ce qui rend le problème plus difficile.

Les solutions des méthodes discrètes $\mathcal{P}^{\mathcal{D}}$ présentent des fausses détections, qui compensent l'erreur de discrétisation afin d'obtenir un résidu suffisamment bas pour satisfaire le critère d'arrêt[23]. Toujours dans le cas discret, la solution des problèmes $\ell_0$ est meilleure que celle du problème $\ell_1$, en terme d'ASD et de résidu.

Une fausse détection est aussi présente dans le résultat de la méthode continue $\mathcal{P}_{2+1}^{\mathcal{C}}$, mais elle donne un meilleur ASD que les méthodes discrètes. Seule la solution de $\mathcal{P}_{2/0}^{\mathcal{C}}$ retrouve exactement le support (ELR=1) ; elle donne aussi un meilleur ASD que la méthode $\ell_1$ continue.

[23]Notons que dans le cas de RSB=10dB, le critère d'arrêt autorise un résidu plus élevé que pour un RSB plus grand. Dans ce cas, les fausses détections sont moins présentes que pour des niveaux de bruits moins élevés.

La figure en bas à gauche montre les résultats des quatre méthodes autour du 4e pic : l'estimation discrètes ($\ell_1$ et $\ell_0$) des localisations donne la même position sur la grille, tandis que les estimations continues proposent des positions similaires plus proches de la vraie localisation.

Notons que la définition de l'espace des contraintes déduit pour $\mathcal{D}_x = \mathbb{R}$ de l'extension du [FDJ15] au cas $\mathcal{D}_x = \mathbb{C}$ donne de moins bons résultats, puisqu'il serait impossible dans ce cas d'identifier des décalages $\delta_j$ négatifs.

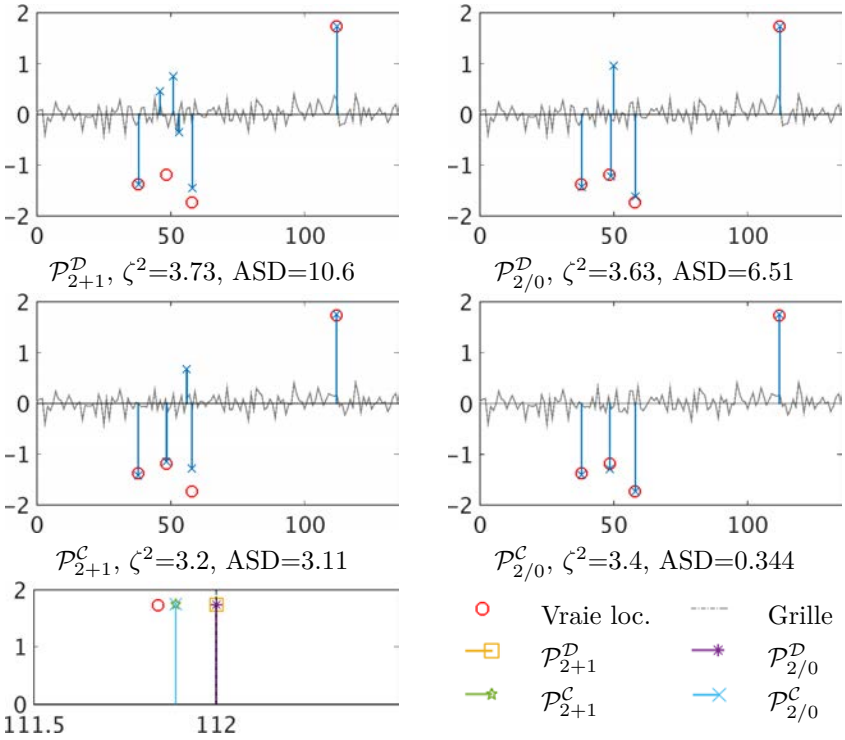



FIG. 2.14 – Résultats d'estimation sur le signal en Figure 1.4(b) avec RSB=10dB pour les méthodes d'approximation parcimonieuse : $\mathcal{P}^{\mathcal{D}}_{2+1}$, $\mathcal{P}^{\mathcal{D}}_{2/0}$, $\mathcal{P}^{\mathcal{C}}_{2+1}$ et $\mathcal{P}^{\mathcal{C}}_{2/0}$. Les cercles rouges montrent les vraies localisations des pics et les croix bleues leur positions estimées. Pour chaque méthode, le résidu est tracé en gris, et sa norme $\zeta^2$ et l'ASD associé sont donnés. La figure du bas montre les résultats des 4 estimations en zoomant autour du 4e pic.

Après cette observation d'un cas particulier, nous avons généralisé ces résultats en calculant le ratio d'ELR et l'ASD moyen sur l'ensemble des 200 modèles et pour les 3 niveaux de bruits, pour avoir un meilleur aperçu global des résultats sur l'ensemble des tests effectués. Ils sont récapitulés dans le Tableau 2.3 pour les méthodes continues $\mathcal{P}^{\mathcal{C}}_{2/0}$, $\mathcal{P}^{\mathcal{C}}_{2+1}$ mais aussi $\mathcal{P}^{\mathcal{C}*}_{2+1}$, la méthode $\ell_1$ sans réévaluation. Nous n'y avons pas inclus les résultats des méthodes discrètes (APL), puisque ces dernières ont un taux de 0% d'ELR.

Une claire amélioration des résultats est apportée grâce à la réévaluation des amplitudes pour les méthodes $\ell_1$ ($\mathcal{P}^{\mathcal{C}*}_{2+1}$), pour tout niveau de bruit, tant en terme d'ELR que d'ASD. Cela confirme la nécessité de procéder à cette réévaluation.

Toutes les méthodes rencontrent des difficultés pour des RSB bas ; les meilleurs résultats sont obtenus pour RSB=20dB et deviennent légèrement moins bons pour RSB=30dB. Cette baisse de qualité des résultats peut être expliquée par l'erreur d'approximation due à l'interpolation polaire, qui n'est plus négligeable par rapport au bruit quand celui-ci est faible. Pour que le résidu atteigne le seuil du bruit (soit une erreur de reconstruction suffisamment basse), les différentes méthodes doivent sélectionner plus de $L = 4$ atomes continus, comme nous pouvons le constater dans le tableau.

| RSB (dB) | % ELR | | | ASD | | |
|---|---|---|---|---|---|---|
| | $\mathcal{P}^{\mathcal{C}*}_{2+1}$ | $\mathcal{P}^{\mathcal{C}}_{2+1}$ | $\mathcal{P}^{\mathcal{C}}_{2/0}$ | $\mathcal{P}^{\mathcal{C}*}_{2+1}$ | $\mathcal{P}^{\mathcal{C}}_{2+1}$ | $\mathcal{P}^{\mathcal{C}}_{2/0}$ |
| 10 | 24.5 (10.5 / 18 / 47) | 31 (12 / 22.5 / 34.5) | 67.5 (24 / 5.5 / 3) | 8.1 (17) | 6.2 (9.2) | 3.6 (7.9) |
| 20 | 31 (0.5 / 1.5 / 67) | 39.5 (0.5 / 3 / 57) | 90.5 (3 / 2.5 / 4) | 2.6 (5.2) | 1.6 (4.1) | 0.78 (4.3) |
| 30 | 30.5 (0 / 0.5 / 69) | 34.5 (0 / 0 / 65.5) | 88 (0 / 0.5 / 11.5) | 1.3 (5.1) | 0.68 (2.3) | 0.045 (0.31) |

TAB. 2.3 – Pourcentage d'ELR et ASD moyen (et écart-type) calculés sur l'ensemble des 200 tests pour $\mathcal{P}^{\mathcal{C}}$. À droite des pourcentages d'ELR sont indiqués les pourcentages des solutions telles que, de haut en bas, $\widehat{L} < 4$, $\widehat{L} = 4$ mais sans ELR et $\widehat{L} > 4$.

Dans tous les cas, la solution $\ell_0$ donne le meilleur pourcentage d'ELR (jusqu'à 90.5% pour un RSB=20dB) et le meilleur ASD ; la solution $\ell_1$ souffre d'un fort taux de fausse-détection ($\widehat{L} > 4$) quel que soit le niveau de bruit. Au contraire, pour un fort niveau de bruit (RSB=10dB), une majorité des solutions sans ELR obtenues avec la résolution $\ell_0$ ont un support de taille $\widehat{L} < 4$ : dans ces cas, la sélection de 3 atomes interpolés à suffit à atteindre le critère d'arrêt. Cela montre bien que la résolution $\ell_1$ échoue à trouver la solution optimale.

Ces résultats ont permis de montrer l'atout de la solution $\ell_0$ par rapport à la solution $\ell_1$. Nous allons maintenant étudier la performance de la solution $\ell_0$ pour les différentes approximations étudiées dans ce manuscrit.

### 2.5.3 Comparaison des interpolations de dictionnaire avec résolution $\ell_0$

Nous avons donc testé notre résolution $\ell_0$ associée à différentes interpolations de dictionnaire : approximation de Taylor à l'ordre 1 (T1), interpolation polaire (Po), approximation par 2-SVD (K2) et 3-SVD (K3). Nous testons aussi l'interpolation polaire avec l'espace des contraintes linéaires $\mathcal{H}^{+}_{\text{Lin3}}$ que nous avons proposé en (2.29) (voir § 2.4.1). Pour le différencier de (Po) (contraintes quadratiques), nous l'appelons (PoLin). Nous avons donc calculé les solutions de $\mathcal{P}^{\mathcal{C}}_{2/0}$ pour les 200 modèles et les 3 niveaux de bruits, pour chacune de ces approximations. Les ratios d'ELR et les ASD moyens sont répertoriés en Tableau 2.4. Nous avons aussi lancé les résolutions $\ell_0$ en recherchant exactement $L_0 = 4$ atomes et donc sans utiliser le critère d'arrêt se basant sur l'erreur de reconstruction. Le pourcentage d'ELR associé (que nous appellerons ELR pour $L_0 = 4$ pour le différencier de l'ELR au seuil du bruit) est aussi récapitulé dans ce tableau.

En terme d'ELR, l'ensemble des approximations linéaires donnent des résultats assez similaires, sauf (T1) pour un bruit faible (RSB=30dB) ; pour ce niveau de bruit, le pourcentage d'ELR est de 29.5% pour (T1) tandis qu'il est supérieur à 80% pour les autres approximations. En effet, comme nous l'avons vu en Figure 2.13, l'approximation (T1) est celle qui donne l'erreur d'approximation moyenne la plus élevée (en dehors du dictionnaire discret). Cette erreur est alors compensée par un nombre $\widehat{L}$ d'atomes sélectionnés plus élevés (70% des solutions donnent un support de taille supérieur à 4). Néanmoins, le pourcentage d'ELR de (T1) pour RSB=10dB est supérieur ou égal aux autres approximations. De plus, le pourcentage d'ELR pour $L_0 = 4$ (sans critère d'arrêt sur l'erreur de reconstruction) est supérieur à ceux de l'interpolation polaire, pour

chaque niveau de bruit. Par conséquent, quand l'ordre du modèle est connu, l'approximation de Taylor est dans cet exemple tout aussi compétitive que les autres approximations, bien qu'elle commette une erreur d'approximation plus élevée.

| RSB | % ELR | | | | | ASD | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| (dB) | T1 | Po | PoL | K2 | K3 | T1 | Po | PoL | K2 | K3 |
| 10 | 23.5<br>68 6<br>85 2.5 | 24<br>67.5 5.5<br>82 3 | 24<br>67.5 5.5<br>82 3 | 24<br>68 5.5<br>84.5 2.5 | 24<br>67.5 6<br>82 2.5 | 3.5<br>(7.7) | 3.6<br>(7.9) | 3.6<br>(7.9) | 3.6<br>(7.9) | 3.5<br>(7.6) |
| 20 | 2<br>84.5 0.5<br>97.5 13 | 3<br>90.5 2.5<br>95 4 | 3<br>90.5 2.5<br>95 4 | 2.5<br>94 1<br>97.5 2.5 | 2.5<br>92 1.5<br>96.5 4 | 0.35<br>(1) | 0.78<br>(4.3) | 0.78<br>(4.3) | 0.3<br>(1) | 0.41<br>(1.2) |
| 30 | 0<br>29.5 0.5<br>98.5 70 | 0<br>88 0.5<br>95 11.5 | 0<br>88 0.5<br>95 11.5 | 0.5<br>85.5 0.5<br>99 13.5 | 0<br>85 0<br>97.5 15 | 0.17<br>(0.43) | 0.045<br>(0.31) | 0.042<br>(0.3) | 0.12<br>(0.49) | 0.056<br>(0.33) |

TAB. 2.4 – Pourcentage d'ELR et ASD moyen (et écart-type) calculés sur l'ensemble des 200 tests pour les différentes approximations linéaires avec résolution $\ell_0$. À droite des pourcentages d'ELR sont indiqués les pourcentages des solutions telles que, de haut en bas, $\widehat{L} < 4$, $\widehat{L} = 4$ mais sans ELR et $\widehat{L} > 4$. En dessous des pourcentages d'ELR est indiqué le pourcentage d'ELR pour les solutions $\ell_0$ *où l'on a remplacé le critère d'arrêt par* $\widehat{L} = 4$.

Les pourcentages d'ELR pour les approximations par K-SVD sont assez similaires, mais celui de (K2) est néanmoins toujours plus élevé que (K3). Cette tendance s'observe aussi sur les pourcentages d'ELR pour $L_0 = 4$. Bien que (K3) donne une meilleure erreur d'approximation que (K2), elle souffre nécessairement d'un espace de contraintes trop relâché par rapport à l'espace image. Par ailleurs, (K2) donne des pourcentages d'ELR supérieurs à ceux de (Po) pour RSB=10dB et 20dB : bien que l'erreur d'approximation soit meilleure pour (Po), l'utilisation de (K2) permet d'utiliser un espace de contraintes identiques à l'espace image, et ainsi d'obtenir une meilleure estimation du support. Pour RSB=30dB, (K2) souffre du même problème que (T1) : son erreur d'approximation empêche le résidu d'atteindre le seuil du bruit pour $\widehat{L} = 4$ et le pourcentage d'ELR est donc inférieur à celui de (Po). Cependant, on constate bien que le pourcentage d'ELR pour $L_0 = 4$ est plus élevé que pour (Po).

Enfin, notons que les résultats de (Po) et (PoLin) sont presque identiques ; la relaxation des contraintes quadratiques par des contraintes linéaires ne modifient donc pas les résultats, pour le problème illustré ici.

Pour terminer cette analyse des résultats, nous avons représenté en Figure 2.15 le pourcentage d'ELR pour $L_0 = 4$ en fonction du temps de calcul moyen de la solution pour $L_0 = 4$, pour chaque interpolation de dictionnaire et chaque niveau de bruit, pour la résolution $\ell_0$[24].

On peut déjà constater que le remplacement des contraintes quadratiques par des contraintes linéaires pour l'interpolation polaire accélèrent les calculs, notamment pour RSB=10dB où le temps de calcul moyen est presque divisé par 4.

Ensuite, cette figure illustre bien l'atout des méthodes telles que $K = 2$ (T1 et K2), quand le nombre $L$ d'atomes est connu, puisqu'elles présentent les taux d'ELR les plus élevés et le temps de calcul moyen les plus faibles, pour chaque niveau de bruit.

[24]Nous ne comparons pas les temps de calcul avec la résolution $\ell_1$ puisqu'ils sont bien plus faibles que ceux de la résolution $\ell_0$.

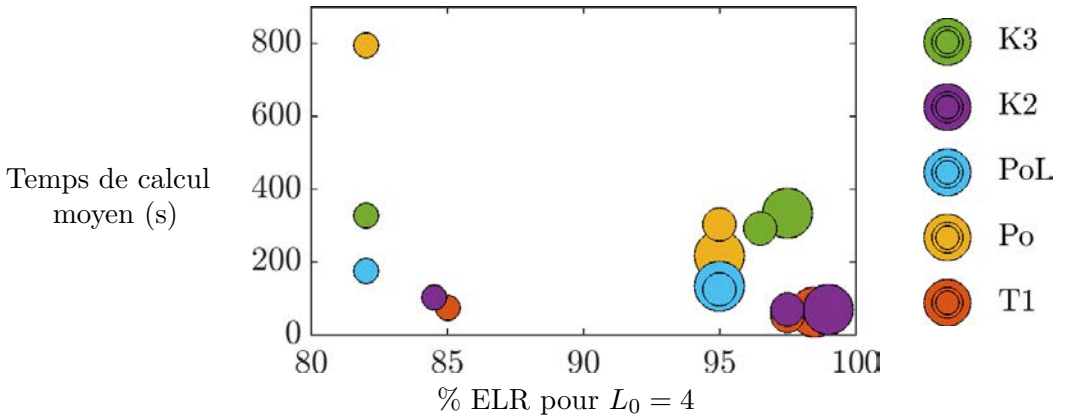



FIG. 2.15 – Représentation du temps moyen (en secondes) de calcul pour la résolution $\ell_0$ en fonction du pourcentage d'ELR, pour différentes approximations linéaires et différents niveaux de bruits. La taille des ronds est associée avec le RSB (le plus grand rond correspond à RSB=30dB).

### 2.5.4 Bilan

Les résultats que nous avons présentés nous ont permis de montrer la bien meilleure efficacité des méthodes de résolution de type $\ell_0$ que nous proposons par rapport aux méthodes de type $\ell_1$, même avec la réévaluation des amplitudes proposée. De plus, le remplacement des contraintes quadratiques par des contraintes linéaires pour l'interpolation polaire donne des résultats similaires mais avec un vrai gain au niveau du temps de calcul.

Dans le cas où l'ordre du modèle est connu (et où l'on peut fixer $L_0$ à cette valeur pour la résolution $\ell_0$), nous avons montré l'efficacité des approximations telles que l'atome interpolé est constitué de $K = 2$ colonnes, soit l'approximation de Taylor à l'ordre 1 et l'approximation par 2-SVD. Pour un coût de calcul réduit, la résolution $\ell_0$ trouve en effet plus souvent le bon support que pour les approximations avec $K = 3$. Même si ces dernières donnent une erreur d'approximation moyenne plus faible, la relaxation convexe nécessaire de l'espace image pour définir l'espace des contraintes entraîne une moins bonne estimation du support.

Dans le cas où l'ordre du modèle est inconnu, le constat est plus mitigé, notamment dans le cas de signaux peu bruités (ici pour RSB=30dB). L'erreur d'approximation induite par les approximations linéaires avec $K = 2$, plus élevée que pour $K = 3$, empêche l'erreur de reconstruction de passer sous le seuil de bruit pour $L = 4$, même si le bon support est détecté ; la taille du support estimé sera alors souvent supérieur à l'ordre du modèle théorique. On peut alors envisager deux pistes de recherche différentes. D'une part, on peut chercher à prendre en compte l'erreur d'approximation dans le critère d'arrêt, afin de définir un seuil moins restrictif pour les signaux peu bruités. D'autre part, on peut chercher à minimiser l'erreur d'approximation induite par l'interpolation du dictionnaire. Chercher une nouvelle approximation linéaire avec $K = 2$ donnant une erreur d'approximation inférieure à celle de la 2-SVD nous semble difficile, puisque pour rappel, la 2-SVD est celle qui minimise l'erreur moyenne pour $K = 2$ par définition. Une solution serait alors de proposer pour l'approximation par 3-SVD un espace des contraintes plus proche de l'espace image. Cela pourrait s'envisager via une étude plus approfondie des propriétés de cette approximation, ou plus spécifique aux atomes continus qu'on cherche à approcher.

Dans tous les cas, on pourrait aussi envisager de diminuer la largeur des intervalles $\Delta$ (et donc d'augmenter leur nombre $J$). L'erreur d'approximation diminuerait, mais au prix d'une augmentation du coût de calcul.

Nous avons donc pu proposer une résolution du problème d'APNL via l'interpolation de dictionnaire pour des atomes continus à valeurs réelles (ce qui n'avait été proposé que par [FDJ15], et de façon moins rigoureuse et entrainant de moins bons résultats même dans le cas d'une pénalisation $\ell_1$). Nous allons maintenant essayer d'appliquer ces travaux au problème d'analyse spectrale présenté en § 1.5.1 ; ce problème nécessitera de proposer des extensions pour des problèmes à parcimonie structurée ou présentant des paramètres de formes, ce que ne permet pas, pour l'instant, le cadre général de l'interpolation de dictionnaire défini en § 2.1.

## 2.6 Extension de l'interpolation de dictionnaire pour l'application à l'analyse spectrale

Dans cette section, nous souhaitons exploiter l'interpolation de dictionnaire pour la formulation en APNL du problème d'analyse spectrale présenté en § 1.5.1. Pour rappel, il s'agit de retrouver les 5 fréquences présentes dans un signal bruité irrégulièrement échantillonné.

Nous avons vu en § 1.1.3 les 3 différentes façons de modéliser une sinusoïde : les modèles *cosinus et sinus* ($\boldsymbol{x}_j \in \mathbb{R}^2$, $\boldsymbol{v}_j = \varnothing$), *cosinus+phase* ($x_j \in \mathbb{R}$, $\boldsymbol{v}_j = \phi_j$) et *exponentielles complexes* ($\boldsymbol{x}_j \in \mathbb{C}^2$, $\boldsymbol{v}_j = \varnothing$ et symétrie hermitienne), récapitulées dans le Tableau 1.1. Quelle que soit la modélisation envisagée, aucune ne rentre dans le cadre de l'interpolation de dictionnaire comme décrite dans les sections précédentes ; soit à cause de la présence d'un paramètre de forme, soit à cause des amplitudes vectorielles (avec contrainte de symétrie hermitienne pour le modèle *exponentielles complexes*). Il sera donc nécessaire de proposer des extensions au cadre général de l'interpolation de dictionnaire.

Dans le cadre de la formalisation du problème d'analyse spectrale en problème d'APL (fréquences fixées sur une grille), le modèle *exponentielle complexe*, qui s'écrit pour un indice $j$ fixé $\mathbf{h}_j\boldsymbol{x}_j = [e^{-2\pi\ddot{\nu}_j\boldsymbol{t}} \quad e^{2\pi\ddot{\nu}_j\boldsymbol{t}}][A_j \quad A_j^*]^T$, est généralement choisi pour la résolution avec pénalisation $\ell_1$ comme le justifie [Bou05]. D'une part, la pénalisation porte sur le module des amplitudes spectrales : $|A| = \frac{1}{2}\sqrt{a^2+b^2}$. D'autre part, le problème peut être résolu en considérant les deux colonnes de $\mathbf{h}_j$ de façon distincte : $\boldsymbol{h}_j = e^{i2\pi\ddot{\nu}_j\boldsymbol{t}}$ et $\boldsymbol{h}_{-j} = e^{i2\pi(-\ddot{\nu}_j)\boldsymbol{t}}$, pour constituer un dictionnaire discret $\mathbf{H}$ de taille $2J$ sur lequel le problème d'APL avec pénalisation $\ell_1$ est posé. Bien que les amplitudes $(A_j, A_{-j})$ soient alors estimées en omettant la contrainte de symétrie hermitienne $A_{-j} = A_j^*$, une solution minimisant le critère et respectant la symétrie hermitienne peut cependant être construite à partir d'une solution initiale, grâce à la convexité non stricte du critère[25].

Ces observations nous ont d'abord menés à suivre la même stratégie pour la formulation du problème d'APLsc via interpolation du dictionnaire : choisir le modèle *exponentielles complexes* en relâchant la contrainte d'hermicité, si bien que les atomes du dictionnaire continu s'écrivent $\boldsymbol{h}_j(\delta_j) = e^{j2\pi(\ddot{\nu}_j+\delta_j)\boldsymbol{t}}$ et ne sont constitués que d'une seule colonne. Le problème ainsi posé rentre donc bien dans le cadre de l'interpolation de dictionnaire général, pour des atomes d'une seule colonne à valeurs et amplitudes complexes. Au début de notre étude sur l'interpolation de dictionnaire, nous avons donc testé, dans l'article de conférence [BC17], cette formulation avec pénalisation $\ell_1$ et interpolation polaire, en adaptant directement[26] le CCBP de [FDJ15] rappelé en équation (2.10). Cependant, nous nous sommes vite rendu compte des limites d'une telle formulation. En effet, si la relaxation de la contrainte de symétrie hermitienne ne pose pas de problème dans le cas des problèmes d'APL, il en va différemment pour le problème d'APLsc :

[25] Si $\widehat{\boldsymbol{x}}_0$ est un minimiseur de $\|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2 + \lambda\|\boldsymbol{x}\|_1$ avec $\mathbf{H}$ le dictionnaire des exponentielles complexes, alors $\widehat{\boldsymbol{x}} = (\widehat{\boldsymbol{x}}_0 + \widehat{\boldsymbol{x}}_0^H)/2$ (où $\cdot^H$ est l'opérateur symétrie hermitienne : $\widehat{x}_j^H = \widehat{x}_{-j}^*$) minimise aussi le critère tout en étant à symétrie hermitienne.

[26] Nous avons cependant corrigé les erreurs de formulation dont nous avons parlé en § 2.2.

les deux amplitudes $\widetilde{\boldsymbol{x}}_j$ et $\widetilde{\boldsymbol{x}}_{-j}$ n'identifient généralement pas la même fréquence par inversion du changement de variables. Cela s'ajoute aussi au fait qu'on doive relâcher la contrainte de liaison entre partie réelle et partie imaginaire (que ce soit pour les résolutions $\ell_1$ ou $\ell_0$). Dans tous les cas, la solution du problème d'APNL identifiée à partir de celle du problème d'APLsc avec cette formulation a peu de chance de donner le bon support, car l'approximation linéaire s'éloigne trop du modèle initial.

Par conséquent, nous allons donc nous intéresser plus particulièrement aux modèles *cosinus et sinus* et *cosinus+phase* pour la résolution de ce problème d'analyse spectrale. Pour le premier, nous allons montrer que la formulation des problèmes d'APLsc est en fait équivalente au cas de l'interpolation d'atomes à valeurs et amplitudes complexes, et ne nécessite donc pas d'adaptation particulière. Pour le second, nous étendons le principe de l'interpolation de dictionnaire à la présence de paramètres de forme grâce à l'approximation par K-SVD. Nous présenterons ensuite les résultats des différentes méthodes que nous proposons pour ce problème.

*Remarque : nous serons parfois amenés à ne pas noter l'indice $j$ des approximations linéaires ou des atomes continus, pour alléger les notations.*

### 2.6.1 Modèle *cosinus et sinus* et parcimonie structurée

Nous proposons donc de modéliser les sinusoïdes avec une combinaison linéaire de sinus et de cosinus. En discrétisant les fréquences sur l'axe positif pour obtenir la grille $\mathcal{G} = [\ddot{\nu}_1, \dots, \ddot{\nu}_J]$, le problème d'APNL s'écrit de la façon suivante :

$$\text{Estimer } \delta_j \in \mathcal{I}_\Delta \text{ et } \boldsymbol{x} \text{ parcimonieux tel que } \boldsymbol{y} \approx \sum_{j=1}^{J} \mathbf{h}_j(\delta_j)\boldsymbol{x}_j, \tag{2.30}$$

avec $\mathbf{h}_j(\delta_j) = [\boldsymbol{h}_j^c(\delta_j), \boldsymbol{h}_j^s(\delta_j)]$, $\boldsymbol{h}_j^c(\delta_j) = \cos(2\pi(\ddot{\nu}_j + \delta_j)\boldsymbol{t})$, $\boldsymbol{h}_j^s(\delta_j) = \sin(2\pi(\ddot{\nu}_j + \delta_j)\boldsymbol{t}))$ et $\boldsymbol{x}_j = [a_j, b_j]^T$ qu'on notera aussi $\boldsymbol{x}_j = [x_j^c, x_j^s]^T$ Les atomes continus $\mathbf{h}_j(\delta_j)$ sont composés de deux colonnes et ce modèle ne rentre donc plus dans le cadre de l'interpolation de dictionnaire classique décrite en § 2.1.1. Nous allons donc étendre la notion d'interpolation de dictionnaire au cas d'atomes continus à deux colonnes, en montrant en quoi ce problème est similaire à la formulation des problèmes d'APLsc pour les signaux à amplitudes complexes. Nous montrerons ensuite comment les problèmes se formalisent pour l'approximation de Taylor à l'ordre 1, l'interpolation polaire et l'approximation par K-SVD.

#### Équivalence des formulations des problèmes d'APLsc pour des atomes à deux colonnes d'amplitudes réelles et des atomes à une colonne d'amplitude complexe

Pour étendre l'interpolation de dictionnaire au cas où l'atome continu est constitué de deux colonnes, on peut considérer l'approximation de chaque colonne de façon indépendante, soit dans notre cas :

$$\begin{array}{rcl} x_j^c\, \boldsymbol{h}_j^c(\delta_j) & \approx & \widetilde{\mathbf{h}}_j^c \widetilde{\boldsymbol{x}}_j^c \quad \text{avec} \quad \widetilde{\boldsymbol{x}}_j^c = x_j^c \Psi^c(\delta_j) \in \mathcal{S}_j^c \\ x_j^s\, \boldsymbol{h}_j^s(\delta_j) & \approx & \widetilde{\mathbf{h}}_j^s \widetilde{\boldsymbol{x}}_j^s \quad \text{avec} \quad \widetilde{\boldsymbol{x}}_j^s = x_j^s \Psi^s(\delta_j) \in \mathcal{S}_j^s \end{array}$$

L'approximation linéaire s'écrit alors :

$$\mathbf{h}_j(\delta_j)\boldsymbol{x}_j \approx \widetilde{\mathbf{h}}_j^{cs} \widetilde{\boldsymbol{x}}_j^{cs} \quad \text{avec} \quad \widetilde{\mathbf{h}}_j^{cs} = \begin{bmatrix} \widetilde{\mathbf{h}}_j^c & \widetilde{\mathbf{h}}_j^s \end{bmatrix} \text{ et } \widetilde{\boldsymbol{x}}^{cs} = \begin{bmatrix} \widetilde{\boldsymbol{x}}_j^c \\ \widetilde{\boldsymbol{x}}_j^s \end{bmatrix} = \begin{bmatrix} x_j^c \Psi^c(\delta_j) \\ x_j^s \Psi^s(\delta_j) \end{bmatrix} \in \mathcal{S}_j^{sc}$$

Cette formulation de l'approximation linéaire est donc très similaire à celle utilisée pour l'approximation linéaire d'un atome continu à valeurs et amplitudes complexes, comme nous l'avons

vu en § 2.2.3 : les amplitudes de l'atome interpolé se séparent en deux parties (ici cosinus/sinus, précédemment réelle/imaginaire) qui sont liées par leur dépendance en $\delta_j$. La définition de l'espace des contraintes peut donc suivre la même logique ; on considère l'approximation $\mathbf{h}_j(\delta_j)\boldsymbol{x}_j \approx \widetilde{\mathbf{h}}_j\widetilde{\boldsymbol{x}}_j$ avec $\widetilde{\boldsymbol{x}}_j \in \mathcal{C}_j$ telle que :

$$\widetilde{\mathbf{h}}_j = \begin{bmatrix} \widetilde{\mathbf{h}}_j^c & -\widetilde{\mathbf{h}}_j^c & \widetilde{\mathbf{h}}_j^s & -\widetilde{\mathbf{h}}_j^s \end{bmatrix} \quad \text{et} \quad \mathcal{C}_j = \left\{ \widetilde{\boldsymbol{x}}_j = \begin{bmatrix} \widetilde{\boldsymbol{x}}_j^{c+} \\ \widetilde{\boldsymbol{x}}_j^{c-} \\ \widetilde{\boldsymbol{x}}_j^{s+} \\ \widetilde{\boldsymbol{x}}_j^{s-} \end{bmatrix}, \begin{array}{l} (\widetilde{\boldsymbol{x}}_j^{c+}, \widetilde{\boldsymbol{x}}_j^{c-}) \in \mathcal{H}_j^{c+2} \\ (\widetilde{\boldsymbol{x}}_j^{c+} \text{ ou } \widetilde{\boldsymbol{x}}_j^{c-}) = \mathbf{0} \\ (\widetilde{\boldsymbol{x}}_j^{s+}, \widetilde{\boldsymbol{x}}_j^{s-}) \in \mathcal{H}_j^{s+2} \\ (\widetilde{\boldsymbol{x}}_j^{s+} \text{ ou } \widetilde{\boldsymbol{x}}_j^{s-}) = \mathbf{0} \end{array} \right\}$$

Tout comme dans le cas de la séparation en parties réelle et imaginaire, il est nécessaire de dédoubler les parties positive et négative puisque les amplitudes $x_j^c$ et $x_j^s$ sont à valeurs réelles. Nous n'intégrons pas ici de contrainte de liaison car les formulations de résolutions de type $\ell_1$ et $\ell_0$ proposées en § 2.3 pour le cas $\mathcal{D}_x = \mathbb{C}$ ne peuvent pas la prendre en compte.

Ainsi, le problème d'APLsc peut être abordée via la formulation du type (2.14) avec $\mathcal{D}_x = \mathbb{C}$, en remplaçant le modèle et l'espace de contraintes par ceux présentés précédemment. Il pourra être résolu à l'aide des formulations (2.22) (résolution $\ell_1$ dans le cas $\mathcal{D}_x = \mathbb{C}$) et (2.18) (résolution $\ell_0$ dans le cas $\mathcal{D}_x = \mathbb{C}$, que nous n'avons pas encore testée pour le moment).

Parmi les diverses approximations étudiées dans ce manuscrit, nous choisissons d'utiliser l'approximation de Taylor à l'ordre 1, l'interpolation polaire et l'approximation par K-SVD. Nous vérifions ici que les hypothèses pour chacune des approximations sont vérifiées.

### Approximation de Taylor à l'ordre 1

Les deux colonnes de l'atome non linéaire, $\boldsymbol{h}_j^c(\delta)$ et $\boldsymbol{h}_j^s(\delta)$, étant toutes les deux continues et dérivables, elles peuvent être approchées par une approximation de Taylor à l'ordre 1. Alors on obtient les approximations suivantes :

$$\begin{array}{rcccl} \boldsymbol{h}_j^c(\delta_j) & \approx & \boldsymbol{h}_j^c(0) + \delta_j \boldsymbol{h}_j^{c'}(0) & = & \boldsymbol{h}_j^c(0) + \delta_j(-2\pi \boldsymbol{t} \odot \boldsymbol{h}_j^s(0)) \\ \boldsymbol{h}_j^s(\delta_j) & \approx & \boldsymbol{h}_j^s(0) + \delta_j \boldsymbol{h}_j^{s'}(0) & = & \boldsymbol{h}_j^s(0) + \delta_j(2\pi \boldsymbol{t} \odot \boldsymbol{h}_j^c(0)) \end{array}$$

où $\odot$ est le produit de Hadamard (multiplication terme à terme de vecteurs). On peut alors écrire les atomes interpolés et les interpolateurs :

$$\begin{array}{rcl} \widetilde{\mathbf{h}}_j^c & = & [\boldsymbol{h}_j^c(0), -2\pi \boldsymbol{t} \odot \boldsymbol{h}_j^s(0)] \\ \widetilde{\mathbf{h}}_j^s & = & [\boldsymbol{h}_j^s(0), \quad 2\pi \boldsymbol{t} \odot \boldsymbol{h}_j^c(0)] \end{array} \quad \text{et} \quad \Psi^c(\delta_j) = \Psi^s(\delta_j) = [1, \delta_j]^T$$

Comme les interpolateurs $\Psi^c(\delta_j)$ et $\Psi^s(\delta_j)$ sont identiques, les espaces images $\mathcal{S}_j^{c+}$ et $\mathcal{S}_j^{s+}$ le sont aussi. Comme vu précédemment, l'espace $\mathcal{S}_j^+$ pour l'approximation de Taylor est convexe et s'exprime à l'aide de contraintes d'inégalités (voir Tableau 2.1).

### Interpolation polaire

Le cas de l'interpolation polaire est plus difficile, car l'hypothèse d'invariance par translation n'est pas respectée par les atomes continus $\boldsymbol{h}_j^c(\delta_j)$ et $\boldsymbol{h}_j^s(\delta_j)$ : leur norme n'est pas constante pour $\delta_j \in \mathcal{I}_\Delta$. Cependant, cette hypothèse est respectée par le modèle *exponentielle complexe* que nous notons $\boldsymbol{h}_j^e(\delta_j)$ : $\forall \delta_j \in \mathcal{I}_\Delta$, $\|\boldsymbol{h}_j^e(\delta_j)\|_2 = \|e^{i2\pi(\ddot{\nu}_j+\delta_j)\boldsymbol{t}}\|_2 = 1$. L'interpolation polaire peut alors être calculée dans le cas de l'exponentielle complexe, pour obtenir l'approximation suivante :

$$\boldsymbol{h}_j^e(\delta_j) \approx \boldsymbol{c}_j + r\cos(\varphi_j)\boldsymbol{u}_j + r\sin(\varphi_j)\boldsymbol{v}_j, \quad \varphi_j = \frac{2\theta_j}{\Delta}\delta_j$$

où les vecteurs $(\boldsymbol{c}_j, \boldsymbol{u}_j, \boldsymbol{v}_j)_j \in \mathbb{C}^N$ et les paramètres $(r_j, \theta_j) \in \mathbb{R}^2$ sont calculés comme expliqué en § 2.4.1. Puis, en remarquant que $\boldsymbol{h}^c_j(\delta_j)$ et $\boldsymbol{h}^s_j(\delta_j)$ sont les parties réelle et imaginaire de $\boldsymbol{h}^e_j(\delta_j)$, on peut alors déduire les approximations suivantes :

$$\begin{array}{rcl} \boldsymbol{h}^c_j(\delta_j) = \mathfrak{Re}(\boldsymbol{h}^e(\delta_j)) & \approx & \mathfrak{Re}(\boldsymbol{c}_j + r\cos(\varphi_j)\boldsymbol{u}_j + r\sin(\varphi_j)\boldsymbol{v}_j)) \\ \boldsymbol{h}^s_j(\delta_j) = \mathfrak{Im}(\boldsymbol{h}^e(\delta_j)) & \approx & \mathfrak{Im}(\boldsymbol{c}_j + r\cos(\varphi_j)\boldsymbol{u}_j + r\sin(\varphi_j)\boldsymbol{v}_j)) \\ \boldsymbol{h}^c_j(\delta_j) & \approx & \boldsymbol{c}^r_j + r\cos(\varphi_j)\boldsymbol{u}^r_j + r\sin(\varphi_j)\boldsymbol{v}^r_j \\ \boldsymbol{h}^s_j(\delta_j) & \approx & \boldsymbol{c}^i_j + r\cos(\varphi_j)\boldsymbol{u}^i_j + r\sin(\varphi_j)\boldsymbol{v}^i_j \end{array}$$

où nous notons $\cdot^r$ et $\cdot^i$ les parties réelle et imaginaire des vecteurs $(\boldsymbol{c}_j, \boldsymbol{u}_j, \boldsymbol{v}_j)$. Cela permet donc de définir les atomes interpolés et les interpolateurs :

$$\begin{array}{rcl} \widetilde{\mathbf{h}}^c_j & = & [\boldsymbol{c}^r_j, \boldsymbol{u}^r_j, \boldsymbol{v}^r_j] \\ \widetilde{\mathbf{h}}^s_j & = & [\boldsymbol{c}^i_j, \boldsymbol{u}^i_j, \boldsymbol{v}^i_j] \end{array} \quad \text{et} \quad \Psi^c(\delta_j) = \Psi^s(\delta_j) = \Psi(\delta_j) = [1, r\cos(\varphi_j), r\sin(\varphi_j)]^T$$

L'interpolateur étant identique au cas classique pour l'interpolation polaire, l'espace des contraintes peut être défini comme vu précédemment (voir Tableau 2.1). On pourra notamment utiliser le remplacement des contraintes quadratiques par des contraintes linéaires que nous avons proposé en § 2.4.1 (p. 64).

Cependant, comme nous l'avons signalé en § 2.4.1, il est nécessaire de vérifier que l'angle $\theta_j$ est bien inférieur à $\frac{\pi}{2}$ afin d'éviter une relaxation supplémentaire. Notons tout d'abord que quel que soit l'indice $j$ de la grille, les angles $\theta_j$ sont identiques ; nous parlerons donc par le suite de l'angle $\theta$. Pour une grille de taille $J = 100$, soit pour un $\Delta = \Delta_0$ défini en § 1.5.1, l'angle $\theta$ vaut $2.22 > \frac{\pi}{2}$. Nous avons donc calculé les angles $\theta$ pour différentes valeurs de $J$ (et donc différentes valeurs de $\Delta$), et nous les avons tracés en Figure 2.16(b) afin de trouver la limite au delà de laquelle $\theta$ devient supérieur à $\pi/2$. Il faut alors au moins $J = 178$ points pour obtenir un angle $\theta \leq \frac{\pi}{2}$. En pratique, nous testerons les méthodes pour $J = 100$ et $J = 200$.

Notons que pour $J = 100$ où $\theta > \pi/2$, les espaces $\mathcal{H}^+_{\text{Lin}}$ définis à partir des contraintes linéaires s'éloignent d'avantage de l'espace image $\mathcal{S}^+$. Comme illustré en Figure 2.16(a), il peut être raisonnable d'ajouter des contraintes linéaires supplémentaires, ce que nous ferons en pratique.

### Approximation par K-SVD

Une dernière approche pour obtenir une approximation linéaire des atomes continus $\boldsymbol{h}^c_j(\delta_j)$ et $\boldsymbol{h}^s_j(\delta_j)$ est leur approximation par K-SVD. Rappelons la méthode pour construire les atomes interpolés $\widetilde{\mathbf{h}}^{c/s}_j$ et les interpolateurs $\Psi^{c/s}(\delta_j)$ pour obtenir l'approximation $\boldsymbol{h}^{c/s}_j(\delta_j) \approx \widetilde{\mathbf{h}}^{c/s}_j \Psi^{c/s}(\delta_j)$ des deux atomes non linéaires (où $^{c/s}$ vaut pour $c$ et $s$ et où on omet l'indice $j$) :

1. Calculer les matrices $\mathbf{B}^{c/s}_\infty$ par intégration : $\mathbf{B}^{c/s}_\infty(m,n) = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h^{c/s}_m(\delta) h^{c/s}_n(\delta) \mathrm{d}\delta$. Ces intégrations ont une expression analytique donnée en Annexe A.3.1.
2. Calculer les atomes interpolés $\widetilde{\mathbf{h}}^{c/s}$ par K-SVD des matrices $\mathbf{B}^{c/s}_\infty$.
3. En déduire l'interpolateur : $\Psi^{c/s}(\delta) = (\widetilde{\mathbf{h}}^{c/s})^T \boldsymbol{h}^{c/s}(\delta)$.

(a) Espace des contraintes pour $\theta \geq \frac{\pi}{2}$ (b) Évolution de $\theta$ en fonction de $J$

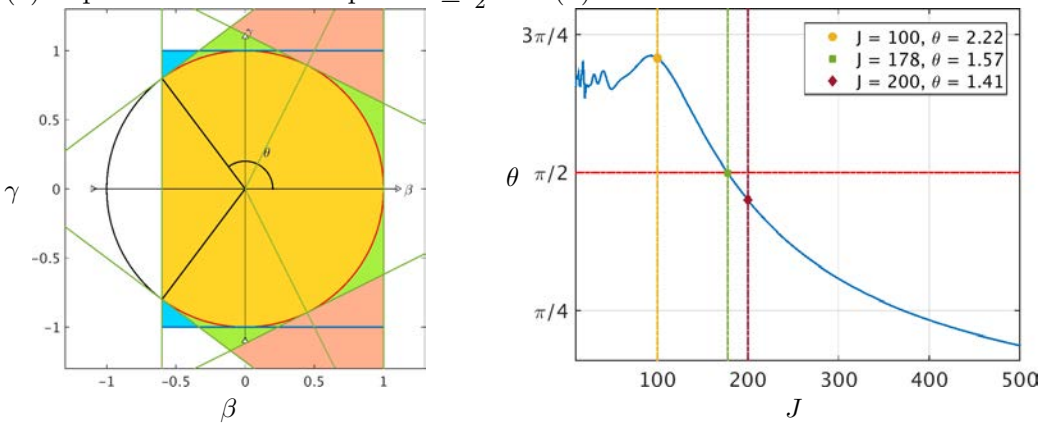



FIG. 2.16 – Illustrations pour l'interpolation polaire en analyse spectrale. (a) Espaces des contraintes pour $\theta = 2.22 \geq \frac{\pi}{2} : \mathcal{H}^+$ en jaune, $\mathcal{H}^+_{\text{Lin2}}$ en jaune+vert+rouge, $\mathcal{H}^+_{\text{Lin3}}$ en jaune+vert. Les droites en bleues représentent des contraintes supplémentaires permettant de mieux approcher $\mathcal{H}^+$. (b) Évolution de l'angle $\theta$ en fonction du nombre d'intervalles $J$, inversement proportionnel à la taille $\Delta$ des intervalles.

La Figure 2.17 illustre l'allure des interpolateurs $\Psi^c(\delta_j)$ ($\Psi^s(\delta_j)$ ayant généralement la même forme) ainsi que l'espace $\mathcal{S}^{\mathbb{R}}$ autour de la même fréquence $\ddot{\nu} = 0.0165$ pour différentes valeurs de $\Delta$ (soit différents nombres de points $J$ sur la grille : $J = 100$ et $J = 200$) et pour $K = 3$. Notons que la projection de $\mathcal{S}^{\mathbb{R}}$ dans l'espace $(\widetilde{x}_1, \widetilde{x}_2)$ revient en fait à observer l'interpolateur et l'espace $\mathcal{S}^{\mathbb{R}}$ pour $K = 2$. Pour $J = 100$, la projection de $\mathcal{S}^{\mathbb{R}}$ dans l'espace $(\widetilde{x}_1, \widetilde{x}_2)$ montre ainsi que le changement de variables n'est pas bijectif pour $K = 2$, contrairement au cas $J = 200$. La projection dans $(\widetilde{x}_1, \widetilde{x}_3)$ montre l'impossibilité de définir une contrainte polyhédrale sur $(\widetilde{x}_1, \widetilde{x}_3)$ dans le cas $J = 100$, car l'espace $\mathcal{S}^{\mathbb{R}}$ recouvre plus du demi-espace.

De façon plus générale, nous avons évalué, pour différents $J$, d'une part le caractère bijectif des changements de variables dans le cas $K = 2$ (nous avons vu en § 2.4.2 que cela pouvait se vérifier grâce à une heuristique) et d'autre part si l'ensemble des composantes $\{\widetilde{\boldsymbol{x}}_k\}_{k>1}$ de $\widetilde{\boldsymbol{x}}$ sont contraintes à l'aide des contraintes polyhédrales décrites en § 2.4.2, ou si certaines sont laissées libres. Ainsi, pour différents $J \in [\![100; 220]\!]$, nous avons calculé :

- pour $K = 2$, le pourcentage sur les $J$ intervalles pour lesquels le changement de variables n'est pas bijectif ;
- pour $K = 3$, le pourcentage sur les $J$ intervalles pour lesquels on a dû relâcher une des contraintes.

La Figure 2.18 récapitule ces résultats. Nous constatons que c'est seulement à partir de $J \geq 180$ que l'approximation par K-SVD devient satisfaisante, tant en terme de bijectivité que d'espace des contraintes.

Finalement, quelle que soit l'approximation linéaire choisie pour cette formulation du problème d'APLsc, les résultats souffriront sûrement de la relaxation de la contrainte de liaison entre les parties cosinus et sinus. Nous allons maintenant voir que l'exploitation du modèle *cosinus+phase* va nous permettre d'éviter ce problème.

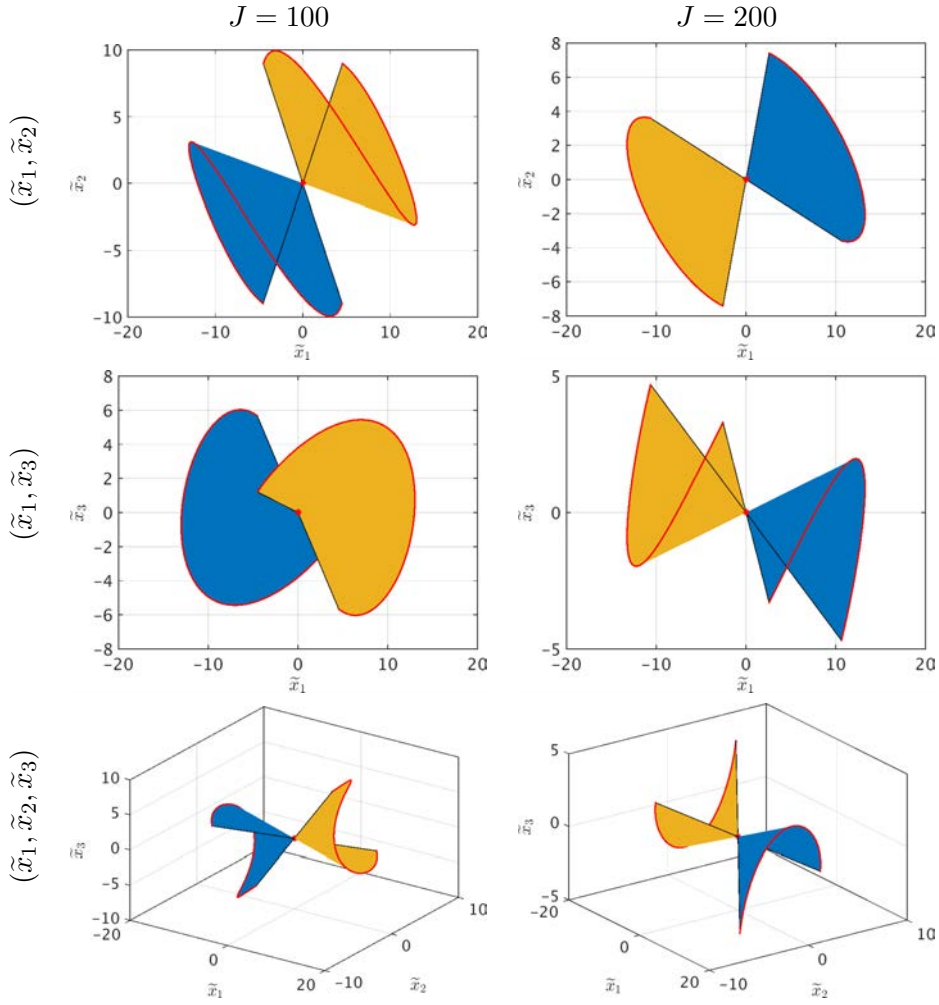



FIG. 2.17 – Approximation K-SVD. Allure des interpolateurs $\Psi^c(\delta_j)$ (rouge) et des espaces $\mathcal{S}^+$ (bleu) et $\mathcal{S}^-$ (jaune) autour d'une fréquence $f_0$, pour $J = 100$ et $J = 200$ (et donc deux $\Delta$ différents). Première et deuxième ligne : projection dans les espaces $(\widetilde{x}_1, \widetilde{x}_2)$ et $(\widetilde{x}_1, \widetilde{x}_3)$.

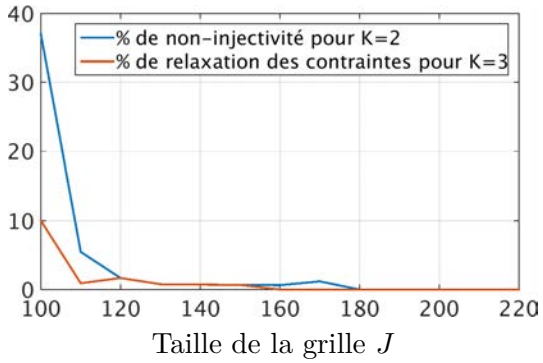



FIG. 2.18 – Pourcentage de non-injectivité (K=2) et de relaxation des contraintes (K=3) pour l'espace des contraintes avec approximation par K-SVD suivant le nombre d'intervalles $J$.

### 2.6.2 Modèle *cosinus+phase* et paramètres de forme

Dans ce paragraphe, nous allons considérer le modèle *cosinus+phase*, soit :

$$\boldsymbol{h}_j(\delta_j, \phi_j) = \cos(2\pi(\ddot{\nu}_j + \delta_j)\boldsymbol{t} + \phi_j)$$

où $\delta_j \in \mathcal{I}_\Delta$ et $\phi \in [-\pi; \pi]$. Ici, la fréquence $\nu_j = \ddot{\nu}_j + \delta_j$ joue le rôle de paramètre de localisation, tandis que la phase $\phi_j$ joue celui du paramètre de forme. Il est intéressant de remarquer que le modèle *cosinus et sinus* vient de la linéarisation du modèle *cosinus+phase*. Cela peut être considéré comme une première linéarisation de dictionnaire, exacte, dans le sens où l'on a remplacé un atome continu en la phase par une combinaison linéaire de deux nouveaux atomes : $\boldsymbol{h}^c$ et $\boldsymbol{h}^s$.

À cause de la présence du paramètre de forme, il est nécessaire de proposer une adaptation des approximations linéaires présentées dans ce chapitre. Nous commençons par faire quelques remarques sur le modèle.

#### Différentes valeurs pour les amplitudes et la phase

L'espace de départ des paramètres $(x, \phi)$ peut être vu de deux façons différentes pour le modèle *cosinus+phase*, par la relation trigonométrique $\cos(\theta + \pi) = -\cos(\theta)$ :

1. $x \in \mathbb{R}^+$ et $\phi \in \mathcal{I}_{2\pi}$, un intervalle de largeur $2\pi$.
2. $x \in \mathbb{R}$ et $\phi \in \mathcal{I}_\pi$, un intervalle de largeur $\pi$.

Ainsi, soit on considère que les amplitudes sont positives et que la phase parcourt l'ensemble du cercle trigonométrique (intervalle de largeur $2\pi$), soit les amplitudes sont réelles et il suffit de chercher la phase dans un intervalle de largeur $\pi$. On écrira $\mathcal{S}^+_{2\pi,j}$ l'espace image du changement de variables $g$ sur $(\mathbb{R}_+ \times \mathcal{I}_\Delta \times \mathcal{I}_{2\pi})$ et $\mathcal{S}_{\pi,j}$ l'espace image de $g$ sur $(\mathbb{R} \times \mathcal{I}_\Delta \times \mathcal{I}_\pi)$ (et $\mathcal{S}^+_{\pi,j}$ quand $\mathcal{D}_x = \mathbb{R}_+$).

#### Une approximation linéaire avec $K$ nécessairement supérieur ou égal à 3

On cherche une approximation linéaire d'un atome continu s'écrivant de la façon suivante :

$$x_j\, \boldsymbol{h}_j(\delta_j, \phi_j) \approx \widetilde{\mathbf{h}}_j \widetilde{\boldsymbol{x}}_j \text{ avec } \widetilde{\boldsymbol{x}}_j = g(x_j, \delta_j, \phi_j) = x_j \Psi(\delta_j, \phi_j)$$

Nous avons vu en § 2.2.1 qu'une condition nécessaire pour que le changement de variables $g$ soit bijectif est que l'espace d'arrivée soit de dimension au moins égale à l'espace de départ. Dans le cas des sections précédentes, où on considérait un seul paramètre non linéaire, il était donc nécessaire de choisir un $K \geq 2$ pour pouvoir retrouver $(x, \delta)$ à partir de $\widetilde{\boldsymbol{x}}$. Ici, nous cherchons aussi à retrouver la phase, soit $(x, \delta, \phi)$ à partir de $\widetilde{\boldsymbol{x}}$ ; ce dernier doit donc au minimum être de dimension $K = 3$.

En théorie, l'approximation de Taylor à l'ordre 1 permet d'obtenir une approximation linéaire avec $K = 3$ ; l'atome interpolé $\widetilde{\mathbf{h}}_j$ serait alors composé de $\boldsymbol{h}_j(0, 0)$ et des dérivées partielles en $\delta_j$ et $\phi_j$ exprimée en $(0, 0)$. Cependant, cette approximation n'est pas satisfaisante en $\phi_j$, puisque ce dernier n'est pas restreint à un petit intervalle comme $\delta_j \in \mathcal{I}_\Delta$. Même en prenant $\phi \in \mathcal{I}_\pi$, l'erreur d'approximation sera donc nécessairement trop élevée.

Notons qu'il serait sûrement possible d'étendre l'interpolation polaire à une double non-linéarité comme le suggérait [ETS11] ; cependant, nous n'avons pas exploré cette piste - l'atome continu du modèle *cosinus+phase* $\boldsymbol{h}_j(\delta_j, \phi_j)$ ne satisfait pas l'invariance par translation (la norme ne reste pas constante pour tous $(\delta_j, \phi_j)$).

Nous proposons donc une interpolation de dictionnaire par K-SVD continue, qui va s'avérer particulièrement bien adaptée à notre problème.

### Interpolation de dictionnaire par K-SVD continue

L'approximation par K-SVD peut facilement s'étendre au cas des atomes continus à deux paramètres non linéaires, de la façon suivante :

1. Calculer les matrices $\mathbf{B}_{\infty,j}$ par double intégration :

$$\mathbf{B}_{\infty,j} = \int_{\mathcal{I}_\theta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \boldsymbol{h}_j(\delta,\phi)\boldsymbol{h}_j(\delta,\phi)^T \mathrm{d}\delta\mathrm{d}\phi$$

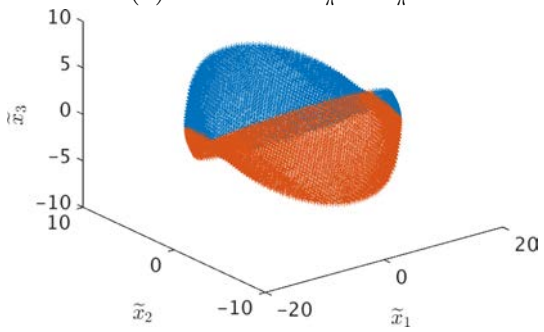
   où $\mathcal{I}_\theta$ vaut pour $\mathcal{I}_\pi$ ou $\mathcal{I}_{2\pi}$ suivant l'espace de départ choisi pour $(x,\phi)$. Ces intégrations ont une expression analytique donnée en Annexe A.3.2.
2. Calculer les atomes interpolés $\widetilde{\mathbf{h}}_j$ par K-SVD des matrices $\mathbf{B}_{\infty,j}$.
3. En déduire l'interpolateur : $\Psi(\delta_j,\phi_j) = (\widetilde{\mathbf{h}}_j)^T \boldsymbol{h}_j(\delta_j,\phi_j)$.

Le gain en mémoire dû au passage de la K-SVD discrète à la K-SVD continue est ici considérable. En effet, pour calculer $\widetilde{\mathbf{h}}_j$ par K-SVD discrète, il faudrait calculer la SVD de la matrice $\mathbf{F}_{P\times Q}$, cette dernière étant constituée de $P \times Q$ atomes $\boldsymbol{h}_j(\delta_{j,p},\phi_{j,q})$ où les paramètres non linéaires ont chacun été discrétisés : $\delta_{j,p} \in \mathcal{I}_\Delta^P$ et $\phi_{j,q} \in \mathcal{I}_\theta^Q$. Pour obtenir une approximation correcte, le nombre de points $P$ et $Q$ doit être suffisamment grand ; cependant, pour une discrétisation avec $P = Q = 1000$ points, le calcul de la K-SVD de la matrice $\mathbf{F}_{P\times Q}$ n'est pas réalisable (au moins sur le matériel utilisé pour nos travaux) car trop coûteuse en mémoire.

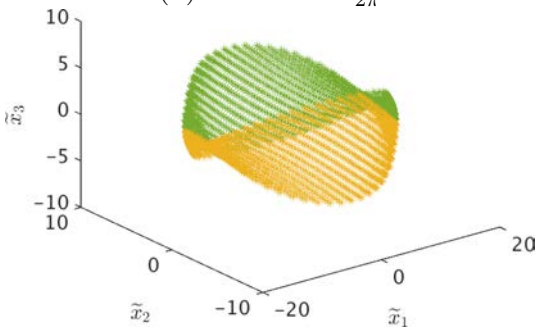
Nous nous intéressons dans un premier temps à la 3-SVD. Afin de mieux appréhender les espaces $\mathcal{S}^+_{\pi,j}$ et $\mathcal{S}^+_{2\pi,j}$, générés par $x_j\Psi(\delta_j,\phi_j)$, nous les avons illustrés en Figure 2.19 autour de la fréquence $\ddot{\nu} = 0.0165$ (nous omettons l'indice $j$ par la suite), en discrétisant $(\delta_p,\phi_q) \in \mathcal{I}_\Delta^P \times \mathcal{I}_\theta^Q$ pour tracer les points $\Psi(\delta_p,\phi_q)$ associés (les espaces $\mathcal{S}$ peuvent se voir comme la dilatation de ces courbes) :

- (a-b) Pour un $\Delta$ correspondant à une grille de $J = 300$ points, nous avons représenté :
  (a) $\Psi_\pi(\delta,\phi)$ (bleu) et $-\Psi_\pi(\delta,\phi)$ (rouge) en discrétisant $\delta \in \mathcal{I}_\Delta$ et $\phi \in \mathcal{I}_\pi = [0;\pi]$.
  (b) $\Psi_{2\pi}(\delta,\phi)$ en discrétisant $\delta \in \mathcal{I}_\Delta$ et $\phi \in \mathcal{I}_{2\pi}[0;2\pi]$ - les points en vert pour $\phi \leq \pi$ et ceux en jaune pour $\phi \geq \pi$.
  Les espaces décrits dans les deux figures sont très similaires. Quel que soit le cas considéré, les espaces $\mathcal{S}^+$ ne semblent pas convexes. Il est donc nécessaire de pouvoir définir une relaxation convexe. Nous proposons d'étudier uniquement $\mathcal{S}^+_{2\pi}$.
- (c) À partir des points $\Psi_{2\pi}(\delta,\phi)$ décrit en (b), nous montrons des courbes à $\delta = \delta_0$ fixé en faisant varier la phase $0 \leq \phi \leq 2\pi$ ; pour $\delta_0 \in \{-\frac{\Delta}{2}, 0, \frac{\Delta}{2}\}$, nous obtenons alors les cercles jaune, rouge et violet. La majeure partie des points $\Psi_{2\pi}(\delta,\phi)$ semblent en fait être compris entre les deux plans décrits par les cercles $\Psi_{2\pi}(\pm\frac{\Delta}{2},\phi)$.
- (d) Ces deux plans sont représentés sur cette figure. Ils passent par l'origine et permettent de décrire deux parties convexes de l'espace $\mathbb{R}^3$. Les dilatations positives de $\Psi(\delta,\phi)$, soit l'espace $\mathcal{S}^+_{2\pi}$, semblent occuper tout l'espace défini par ces plans.
- (e-f) Ces deux figures sont les équivalentes des figures (c-d), mais pour un $\Delta$ correspondant à $J = 100$ points. Les deux plans décrits précédemment permettent toujours de diviser en deux parties convexes l'espace, mais certains points $\Psi_{2\pi}(\delta,\phi)$ semblent être exclus.
- (g) On observe pour différentes tailles de grille $J$, le pourcentage des points $\boldsymbol{z} = \Psi(\delta_p,\phi_q)$ appartenant bien à l'espace décrit par les deux plans, pour l'ensemble des intervalles de la grille. On remarque que plus la taille de la grille augmente, mieux les points sont décrits par les deux plans. Cependant, même pour une grille fine ($J = 300$), un faible pourcentage de points s'échappent de cet espace.

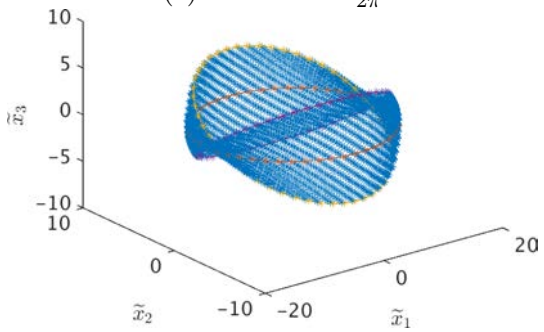
(a) $J = 300$ - $\mathcal{S}_{\pi}^{+} \cup \mathcal{S}_{\pi}^{-}$



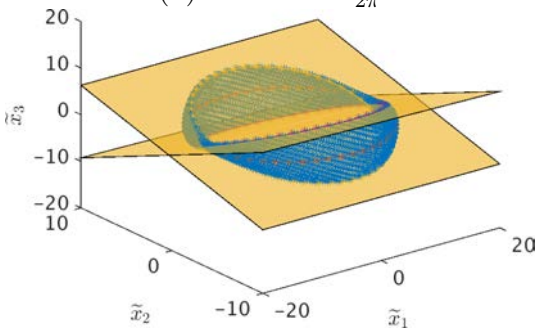

(b) $J = 300$ - $\mathcal{S}_{2\pi}^{+}$



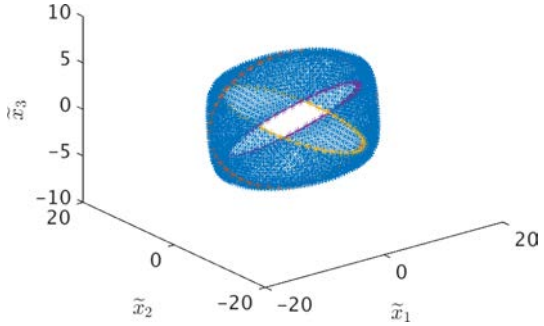

(c) $J = 300$ - $\mathcal{S}_{2\pi}^{+}$



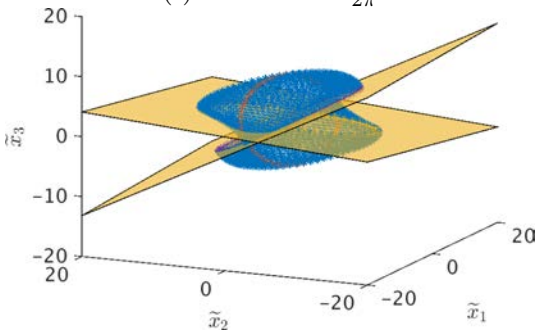

(d) $J = 300$ - $\mathcal{S}_{2\pi}^{+}$




(e) $J = 100$ - $\mathcal{S}_{2\pi}^{+}$




(f) $J = 100$ - $\mathcal{S}_{2\pi}^{+}$




(g) Pourcentage de points appartenant à l'espace décrit par les deux plans selon $J$.

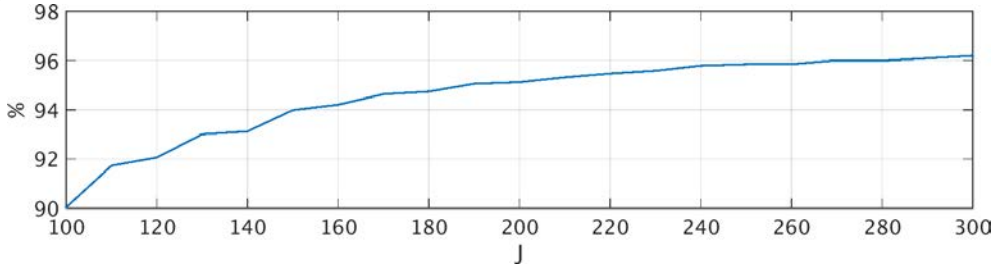



FIG. 2.19 – Représentation des espaces $\mathcal{S}_{\pi}^{+} \cup \mathcal{S}_{\pi}^{-}$ et $\mathcal{S}_{2\pi}^{+}$ pour plusieurs tailles de grille $J$, autour de la fréquence $f_0$. (a) En bleu $\mathcal{S}_{\pi}^{+}$ et en rouge $\mathcal{S}_{\pi}^{-}$. (b) $\mathcal{S}_{2\pi}^{+}$ : En vert, points correspondant à $\phi \leq \pi$ et en jaune points correspondants à $\phi \geq \pi$. (c) Les cercles en jaune, rouge et violet représentent respectivement $\Psi(-\frac{\Delta}{2}, \phi)$, $\Psi(0, \phi)$ et $\Psi(\frac{\Delta}{2}, \phi)$ pour $\phi$ variant de 0 à $2\pi$. (d) Proposition de plan séparateur reposant sur les cercles $\Psi(\pm\frac{\Delta}{2}, \phi)$ pour contraindre l'espace $\mathcal{S}_{2\pi}^{+}$. (e-f) Espace $\mathcal{S}_{2\pi}^{+}$ pour $J = 100$ - cercles et plans séparateurs. (g) Pourcentage de points appartenant à l'espace décrit par les deux plans selon $J$.

Au regard de ces observations, nous prenons donc les décisions suivantes pour utiliser l'approximation par 3-SVD continue sur le modèle *cosinus+phase* :

1. Nous cherchons les $(x, \delta, \phi)_j$ dans $\mathbb{R}^+ \times \mathcal{I}_\Delta \times \mathcal{I}_{2\pi}$ ; l'espace image sera donc $\mathcal{S}^+_{2\pi}$. En effet, alors que dans les sections précédentes, la séparation en $\mathcal{S}^+$ et $\mathcal{S}^-$ permettait de considérer des espaces convexes ne s'éloignant pas trop de l'espace image, cela ne semble pas être le cas ici : $\mathcal{S}^+_\pi$ et $\mathcal{S}^-_\pi$ semblent tous les deux non convexes et une relaxation convexe semble difficile à trouver.
2. L'espace image $\mathcal{S}^+_{2\pi}$ n'est pas convexe ; on peut cependant le séparer en deux espaces convexes à l'aide des plans tangents en $\Psi(\pm\frac{\Delta}{2}, \phi)$. De la même façon que l'on séparait $\mathcal{S}^+$ et $\mathcal{S}^-$ dans le cas des atomes non linéaires en un seul paramètre, nous séparons ici les deux espaces compris entre les plans. Ainsi, les deux plans sont décrits par les équations suivantes :
$$f_i(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3) = a_i\widetilde{x}_1 + b_i\widetilde{x}_2 + c_i\widetilde{x}_3 = 0,\ i \in \{1, 2\}$$
telles que les plans soient orientés face à face (ou dos à dos suivant l'espace). Les espaces seront alors décrits par les équations suivantes :
$$\begin{aligned}(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3) \in \mathcal{C}_a \quad &\Leftrightarrow \quad \forall i \in \{1, 2\},\ f_i(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3) \geq 0 \\ (\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3) \in \mathcal{C}_b \quad &\Leftrightarrow \quad \forall i \in \{1, 2\},\ f_i(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3) \leq 0\end{aligned}$$

Les équations des plans seront déterminées en calculant leur normale à l'aide de trois points décrits par $\Psi(\pm\frac{\Delta}{2}, \phi_0)$ avec $\phi_0 \in \{0, 2\frac{\pi}{3}, 4\frac{\pi}{3}\}$ (par exemple).

Comme nous l'avons vu précédemment, certains points de $\mathcal{S}^+_j$ échapperont à l'union $\mathcal{C}_a \cup \mathcal{C}_b$ : ainsi, certains $\widetilde{\boldsymbol{x}} = g(x, \delta, \phi)$ n'appartiendront pas à l'espace des contraintes. Nous supposons que cela aura peu d'influence sur les résultats de l'estimation. En effet, puisque l'identification des paramètres initiaux $(x, \delta, \phi)_j$ ne peut pas se faire analytiquement, elle sera donc effectuée en minimisant la distance entre l'atome interpolé estimé et ceux appartenant à l'espace image en discrétisant $(\delta, \phi)$, comme le propose [PPZ17] (voir § 2.1.5). De plus, nous avons remarqué que les points $(x, \delta, \phi)$ appartenant à $\mathcal{S}^+_{2\pi}$ mais pas à $\mathcal{C}_a \cup \mathcal{C}_b$ ont pour la majorité un $\delta$ proche de $\pm\frac{\Delta}{2}$ ; c'est aussi pour ces valeurs que l'erreur d'approximation est la plus élevée et où les méthodes d'estimation par approximation parcimonieuse éprouveront des difficultés.

Enfin, la formalisation du problème d'APLsc se fait de façon similaire à celle pour les signaux à amplitudes réelles en § 2.2.2 : le modèle s'écrira :

$$x\,\boldsymbol{h}(\delta, \phi) \approx \widetilde{\mathbf{h}}^{ab}\,\widetilde{\boldsymbol{x}}^{ab} \quad \text{avec} \quad \widetilde{\mathbf{h}}^{ab} = [\widetilde{\mathbf{h}} \quad \widetilde{\mathbf{h}}] \text{ et } \widetilde{\boldsymbol{x}}^{ab} = \begin{bmatrix}\widetilde{\boldsymbol{x}}^a \\ \widetilde{\boldsymbol{x}}^b\end{bmatrix} \quad \text{avec} \quad \begin{array}{c}(\widetilde{\boldsymbol{x}}^a, \widetilde{\boldsymbol{x}}^b) \in (\mathcal{C}_a, \mathcal{C}_b) \\ \widetilde{\boldsymbol{x}}^a \text{ ou } \widetilde{\boldsymbol{x}}^b = \mathbf{0}\end{array}$$

On peut donc utiliser les formulations dans le cas $\mathcal{D}_x = \mathbb{R}$ des résolutions $\ell_1$ (§ 2.3.2) et $\ell_0$ (§ 2.3.1), mais à une différence près : nous ne pouvons plus imposer la parcimonie structurée via la première composante de $\widetilde{\boldsymbol{x}}_j$ uniquement, comme dans le cas usuel, car l'espace des contraintes que nous avons décrit ne respecte plus l'implication de nullité. La parcimonie devra donc être imposée sur l'ensemble des composantes des $\widetilde{\boldsymbol{x}}_j$ ; nous avons vu que cela ne posait pas de problème particulier pour les formulations de type $\ell_0$ puisque la parcimonie peut être imposée à l'ensemble des composantes grâce à la contrainte de big-M des formulations MIP. Cependant, cela sera beaucoup plus problématique pour la pénalisation $\ell_1$ pour laquelle la parcimonie structurée requiert l'utilisation de norme mixte.

Enfin, notons que nous avons généralisé les constations faites pour l'approximation par 3-SVD à celle par 4-SVD sur l'espace image, mais ce cas reste plus délicat. En effet, nous avons constaté que pour la 4-SVD, où les amplitudes $\widetilde{\boldsymbol{x}}$ appartiennent à $\mathbb{R}^4$, l'espace image $\mathcal{S}^+_{2\pi}$ projeté dans les plans $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3)$ et $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_4)$ avaient la même forme que ceux décrits par dans le cas 3-SVD.

Sans détailler davantage, notons qu'il est alors possible de définir un espace de contraintes en considérant les projections dans les espaces $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3)$ et $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_4)$, comme nous le faisions pour la définition de contraintes polyhédrales en l'absence de paramètre de forme.

### 2.6.3 Comparaison des différentes formulations proposées sur le signal test

Le Tableau 2.5 récapitule les diverses dimensions pour la résolution $\ell_0$ des problèmes d'APLsc, selon l'approximation linéaire et le modèle choisi. Nous notons les diverses approximations de la façon suivante, pour alléger les notations :

- pour le modèle *cosinus et sinus* : approximation de Taylor à l'ordre 1 (T1), interpolation polaire avec contraintes linéaires (PoL), approximation 2-SVD (K2) et 3-SVD (K3).
- pour le modèle *cosinus+phase* : approximation 3-SVD (K3p) et 4-SVD (K4p).

Comme on peut le voir, la résolution via le modèle *cosinus et sinus* sera plus coûteuse que celle via le modèle en *cosinus+phase*, puisqu'elle nécessite 2 fois plus de variables binaires, pour le même nombre d'intervalles sur la grille.

| Modèle | *cosinus et sinus* | | | | *cosinus+phase* | |
|---|---|---|---|---|---|---|
| Approximation | T1 | K2 | PoL | K3 | K3p | K4p |
| Nb. variables binaires | $4J$ | $4J$ | $4J$ | $4J$ | $2J$ | $2J$ |
| $K$ | 2 | 2 | 3 | 3 | 3 | 4 |
| Taille $\widetilde{\boldsymbol{x}}$ | $8J$ | $8J$ | $12J$ | $12J$ | $6J$ | $8J$ |

TAB. 2.5 – Récapitulatif des diverses dimensions pour la résolution $\ell_0$ des problèmes d'approximation parcimonieuse sur dictionnaires interpolés.

#### Une première observation des erreurs d'approximation

En Figure 2.20, nous avons calculé l'erreur d'approximation moyenne $\xi_J$ pour différentes valeurs de $\Delta_J$, en moyenne sur l'ensemble des $J$ intervalles :

$$\xi_J = \frac{1}{J}\sum_{j=1}^{J} \frac{1}{\Delta_J\, 2\pi} \int_{\mathcal{I}_\Delta} \int_{\mathcal{I}_{2\pi}} \|\boldsymbol{h}_j(\delta,\phi) - \widetilde{\mathbf{h}}_j \Psi_j(\delta,\phi)\|^2 \mathrm{d}\delta \mathrm{d}\phi$$

où pour le modèle en *cosinus et sinus*, $\widetilde{\mathbf{h}}_j \Psi_j(\delta,\phi)$ s'écrit :

$$\widetilde{\mathbf{h}}_j \Psi_j(\delta,\phi) = \widetilde{\mathbf{h}}_j^c(\cos\phi \Psi_j^c(\delta)) + \widetilde{\mathbf{h}}_j^s(-\sin\phi \Psi_j^s(\delta))$$

On remarque déjà que pour la grille proposée en § 1.5.1 (soit $J = 100$), l'approximation de Taylor donne un erreur d'approximation plus mauvaise que sans approximation. Les autres approximations donnent un meilleur résultat, mais les erreurs restent assez élevées.

Les approximations Polaire et par K-SVD sur le modèle en *cosinus et sinus* donnent les erreurs d'approximation les plus faibles. Cependant, la relaxation de la contrainte de liaison amènera nécessairement des problèmes de performance pour les méthodes d'approximation parcimonieuse. De plus, rappelons que pour résoudre ce problème via la reformulation en MIP, il est nécessaire de considérer $5J$ variables binaires, là où le modèle en *cosinus+phase* n'en nécessite que $2J$.

Enfin, l'approximation par 3-SVD sur le modèle *cosinus+phase* ne donne pas la meilleure approximation ; on observe d'ailleurs que l'erreur d'approximation ne décroit que très légèrement pour $J \geq 200$. Cela est dû au fait qu'en augmentant $J$, on diminue la taille $\Delta_J$ de l'intervalle $\mathcal{I}_\Delta$, mais pas celui des phases $\phi$ qui reste toujours un intervalle de largeur $2\pi$. Pour ce modèle, il faut aller jusqu'à l'ordre 4 en K-SVD pour arriver à réduire cette erreur d'approximation.

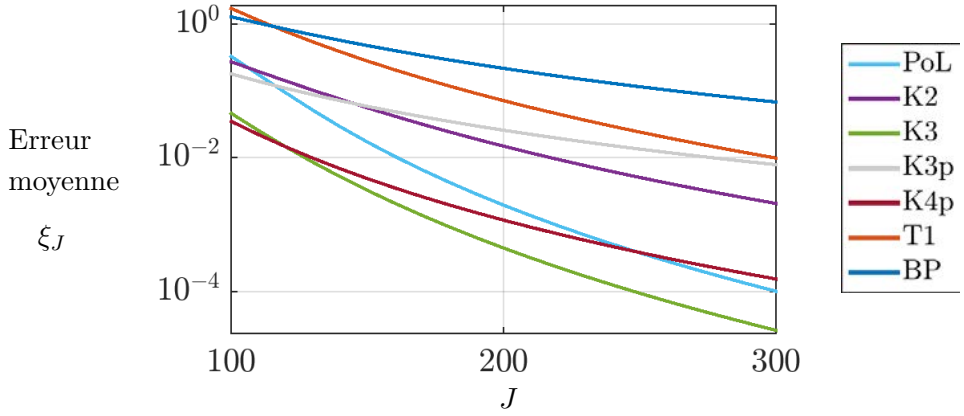



FIG. 2.20 – Erreur d'approximation moyenne $\xi_J$ pour différentes valeurs de taille de grille $J$ et pour les différentes approximations linéaires proposées. BP signifie sur grille ($\delta_j = 0$).

### Résultats des tests

Nous avons commencé par tester la formulation $\ell_1$ pour le modèle *cosinus et sinus*, équivalente à celle que nous proposons dans le cas $\mathcal{D}_x = \mathbb{C}$ en § 2.3.2. Nous avons ainsi pu constater que la solution du problème d'APLsc ne respectait jamais la contrainte de disjonction. Par ailleurs, les résultats n'étaient pas satisfaisants en terme de détection de support.

Nous avons donc testé les approximations linéaires (T1), (PoL), (K2) et (K3) avec le modèle *cosinus et sinus* et (K3p) et (K4p) avec le modèle *cosinus+phase*, avec la résolution $\ell_0$ en fixant $L_0 = 5$, le nombre de fréquences à retrouver, et pour $J = 100$ et $J = 200$. Pour chaque test, le respect de l'ELR ainsi que le passage de l'erreur de reconstruction en dessous du seuil du bruit sont évalués.

En Figure 2.21 est représenté le temps de calcul en fonction de l'erreur moyenne sur l'estimation des fréquences pour chaque test effectué.

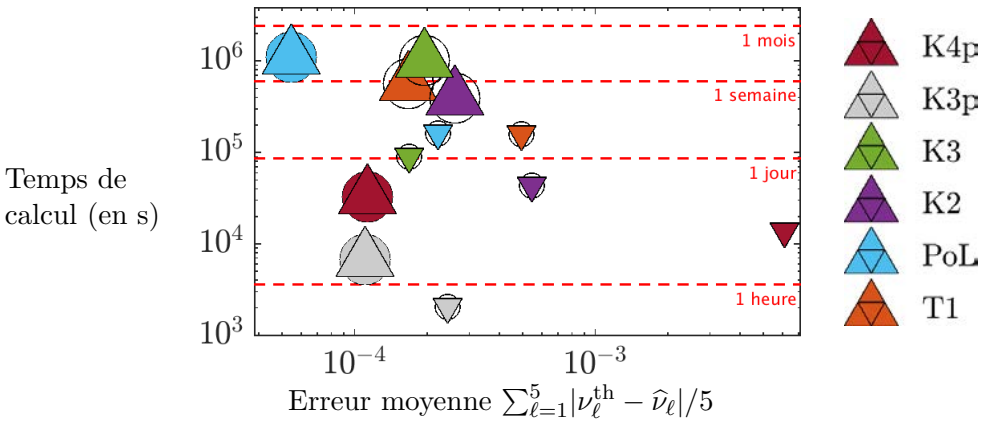



FIG. 2.21 – Représentation du temps de calcul pour la résolution $\ell_0$ en fonction de l'erreur moyenne sur l'estimation des fréquences, pour différentes interpolations de dictionnaire avec $J = 100$ et $J = 200$. La taille des triangles est relative à $J$ : le plus grand correspond à $J = 200$. Quand le triangle est entouré d'un cercle noir, la solution satisfait l'ELR. Quand le cercle est coloré, le résidu de la solution est en dessous du seuil d'arrêt.

Une première constatation est que le seul résultat ne respectant pas l'ELR est celui pour (K4p, $J = 100$). L'espace des contraintes défini par les plans tangents doit trop s'éloigner de

l'espace image. Comme attendu, les méthodes se basant sur le modèle *cosinus et sinus* sont très lourdes en temps de calcul : de 12 heures (K2, $J = 100$) à 13 jours (PoL, $J = 200$). En effet, le nombre de variables binaires nécessaires pour ce modèle est de $4J$ tandis qu'il n'est que de $2J$ pour le modèle *cosinus+phase*.

Enfin, seules les approximations linéaires (PoL), (K3p) et (K4p) pour $J = 200$ permettent d'obtenir une erreur de reconstruction inférieure au seuil du bruit. L'approximation (K3p, $J = 100$) atteint l'ELR et l'erreur moyenne sur l'estimation des fréquences est relativement faible.

Nous pouvons en conclure que pour ce problème d'analyse spectrale, le modèle *cosinus+phase* associé aux approximations linéaires se basant sur la K-SVD continue semble être la meilleure option.

## 2.7 Conclusion

Dans ce chapitre, nous avons généralisé dans le cadre des problèmes d'APNL le principe d'interpolation de dictionnaire, introduit par [ETS11]. À partir des travaux existants, nous avons pu déterminer un cadre général sur lequel nous nous sommes basés pour proposer de nombreuses améliorations, permettant d'atteindre certains objectifs de cette thèse. Notamment, nous avons pu proposer une prise en compte exacte de la parcimonie grâce à la reformulation du problème d'APLsc exploitant la pseudo-norme $\ell_0$ en problème d'optimisation mixte en nombre entier (MIP), inspiré des travaux de [BNCM16]. Cette formulation dans le cadre de l'interpolation de dictionnaire n'est rendue possible que grâce à une étude préliminaire de l'espace image du changement de variables, qui a permis de mettre en exergue la nécessité de dissocier l'espace image $\mathcal{S}^{\mathbb{R}}$ quand les atomes continus à approcher ont des amplitudes à valeurs réelles. Enfin, l'étude plus détaillée des interpolations de dictionnaire a notamment permis de proposer un espace de contraintes linéaires pour l'interpolation polaire, permettant un important gain en coût de calcul ; de plus, la définition d'un espace de contraintes correct pour l'approximation par K-SVD, et son extension à une approximation continue, la rendent maintenant exploitable. Les tests menés sur les données de déconvolution impulsionnelle ont permis d'illustrer nos contributions, notamment en montrant l'amélioration amenée par la résolution de type $\ell_0$ vis-à-vis de celle de type $\ell_1$ en terme de détection de support.

Pour tenter d'appliquer le principe d'interpolation de dictionnaire à un problème d'analyse spectrale, quelques extensions au cadre général présenté dans l'ensemble du chapitre ont été proposées, notamment pour la prise en compte de paramètres de forme ou pour l'adaptation à des atomes continus multi-colonnes (parcimonie structurée). Ces tests ont permis de révéler le potentiel de l'approximation par K-SVD pour ce type de problème.

Une limite de l'interpolation de dictionnaire pour la résolution des problèmes d'APNL est sa difficile formulation dans le cas où les atomes continus sont à valeurs et amplitudes complexes, ou composés de plusieurs colonnes, comme c'est le cas pour le modèle *cosinus et sinus* en analyse spectrale. En effet, l'impossibilité à ce jour de prendre en compte la contrainte de liaison pour définir un espace de contraintes convexe entraîne une relaxation de l'espace image préjudiciable à la bonne détection du support.

Par ailleurs, l'exploitation de la résolution de type $\ell_0$ permet une meilleure compréhension des interpolations de dictionnaire. Elle a ainsi permis de mettre en évidence le compromis entre l'erreur d'approximation et la convexité de l'espace image : les approximations linéaires telles que $K = 2$ (Taylor 1, 2-SVD), permettent de considérer l'espace image du changement de variables comme espace de contraintes convexe, mais l'erreur d'approximation les caractérisant

peut mettre en échec la détermination de l'ordre du modèle. Au contraire, les approximations avec $K$ supérieur (interpolation polaire, 3-SVD) présentent une erreur d'approximation plus faible, mais la relaxation nécessaire de l'espace image vers un espace convexe peut faire échouer la bonne détection du support. Une perspective intéressante serait alors une étude plus approfondie et plus rigoureuse de l'approximation par K-SVD, qui permettrait de définir un espace de contraintes plus proche de l'espace image pour $K > 2$, et ainsi de bénéficier d'une erreur d'approximation plus faible que pour $K = 2$. De même, une extension de l'interpolation de dictionnaire de type Minimax proposée par [PPZ17] pourrait être envisagée ; comme elle se base sur la définition d'un atome interpolé minimisant l'erreur maximale d'approximation vis-à-vis d'un interpolateur prédéfini $\Psi(\delta)$, ce dernier pourrait être choisi tel que l'espace image $\mathcal{S}^+ = \{x\Psi(\delta)\}$ soit convexe, même pour $K > 2$. Il pourrait aussi être envisageable de prendre en compte l'erreur d'approximation dans le critère pour le choix de l'ordre de modèle.

Une autre possibilité pour diminuer l'erreur d'approximation est la diminution de la taille $\Delta$ des intervalles auxquelles appartiennent les décalages, ce qui entraînerait une augmentation du coût de calcul, lequel est déjà particulièrement élevé pour les méthodes de résolution de type $\ell_0$. Cette piste pourra néanmoins être envisagée en adaptant les algorithmes de résolution des MIP aux spécificités liées aux problèmes d'APL (et d'APLsc).

# Chapitre 3

# Modèle Bernoulli-Gaussien étendu et échantillonnage stochastique

**Sommaire**


**3.1 Modèle Bernoulli-Gaussien pour la résolution d'APL . . . . . 97**
3.1.1 Expression de la loi *a posteriori* . . . . . 97
3.1.2 Estimation à partir d'échantillons de la loi *a posteriori* . . . . . 98
3.1.3 Échantillonnage stochastique de la loi *a posteriori* . . . . . 101
**3.2 Modèle BGE : premiers algorithmes d'échantillonnage de la loi jointe . . . . . 106**
3.2.1 Modèle BG étendu et algorithme Gibbs-hybride . . . . . 106
3.2.2 Modèle BG à dimension variable et algorithme RJGS . . . . . 110
3.2.3 Choix du modèle BGE avec échantillonnage Gibbs-hybride . . . . . 111
**3.3 Échantillonnage de la loi *a posteriori* marginalisée dans le cadre supervisé . . 113**
3.3.1 Algorithme d'échantillonnage de la loi marginalisée . . . . . 115
3.3.2 Astuces calculatoires pour l'échantillonnage de $q_j$ . . . . . 117
3.3.3 Astuces calculatoires pour l'échantillonnage des $\boldsymbol{\nu}_j$ . . . . . 123
3.3.4 Bilan sur l'échantillonnage de la loi *a posteriori* marginalisée dans le cadre supervisé 124
**3.4 Échantillonnage exploitant la marginalisation dans le cadre non supervisé . . 126**
3.4.1 Échantillonnage de type *Partially Collapsed Gibbs Sampler* pour le modèle BGE . 127
3.4.2 Astuces calculatoires pour l'implémentation du PCGS . . . . . 129
3.4.3 Conseils pour le réglage des lois *a priori* pour les hyperparamètres . . . . . 131
**3.5 Application du modèle BGE et de l'échantillonneur associé aux données tests 133**
3.5.1 Modèle BGE pour un problème d'analyse spectrale . . . . . 134
3.5.2 Étude d'une réalisation typique des échantillonneurs et proposition d'une nouvelle stratégie d'estimation . . . . . 136
3.5.3 Comparaison des échantillonneurs sur 50 graines . . . . . 140
3.5.4 Étude du réglage des lois *a priori* des hyperparamètres . . . . . 141
3.5.5 Application sur les données de déconvolution impulsionnelle . . . . . 142
3.5.6 Bilan sur les tests du modèle BGE et des échantillonneurs associés . . . . . 144
**3.6 Conclusion . . . . . 144**

Dans ce chapitre, nous allons étudier et améliorer une deuxième méthode pour résoudre les problèmes d'approximation parcimonieuse non linéaire (APNL) présentés en § 1.3.2 : la modélisation Bernoulli-Gaussienne étendue aux non-linéarités associée à une estimation au sens de l'EAP grâce à l'échantillonnage stochastique de la loi *a posteriori*. Cette approche a été proposée par [BC06] et consiste en (i) la prise en compte d'un modèle BG classique via l'introduction de variables binaires $q_j$ suivant une loi *a priori* de Bernoulli conditionnant la nullité de l'amplitude associée $\boldsymbol{x}_j$ et (ii) la considération des paramètres non linéaires $\boldsymbol{\nu}_j$ dans le problème d'estimation bayésien. Elle repose sur la simulation d'échantillons des paramètres d'intérêt suivant la loi *a posteriori* grâce à un échantillonneur de type Gibbs-hybride.

Les travaux présentés dans ce chapitre consistent en l'amélioration de l'estimation des paramètres du modèle BGE, tant sur le point de vue échantillonnage que calcul d'estimateur, en se basant sur divers travaux transverses. Les différentes sections s'organisent de la façon suivante :

1. Les principes du modèle Bernoulli-Gaussien classique dans le cadre de l'échantillonnage stochastique sont d'abord rappelés.

2. Nous revenons sur une première tentative d'amélioration effectuée en stage de Master [Bou14] visant à palier les limites de l'échantillonneur de Gibbs-hybride proposé par [BC06] en incluant une étape de Reversible-Jump [Gre95], dans le cadre supervisé. Nous montrons que l'algorithme proposé revient à un algorithme de Gibbs-hybride, où les paramètres $(q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j)_j$ sont proposés simultanément via une étape de Métropolis-Hastings, et que l'algorithme de Gibbs-hybride proposé par [BC06] est plus efficace.

3. Nous proposons ensuite de modifier l'algorithme de Gibbs-hybride *dans le cadre supervisé*, en considérant l'échantillonnage de la loi *a posteriori* marginalisée en les amplitudes plutôt que la loi jointe. Bien que non recommandé par [BC06] notamment pour des raisons calculatoires, l'échantillonnage de la loi marginalisée peut être effectué pour un coût de calcul plus raisonnable grâce aux astuces calculatoires proposées par [GILC11] dans le cas d'un modèle BG classique. L'apport amené ici consiste en (i) l'adaptation de ces astuces calculatoires à l'échantillonnage des paramètres non linéaires $\boldsymbol{\nu}_j$ via une étape MH, (ii) l'extension des calculs de la marginalisation au cas où le bruit est non *i.i.d.* (matrice de covariance $\boldsymbol{\Sigma}$ dans l'optique d'une application à la détection d'exoplanètes) et où les amplitudes sont non-scalaires $\boldsymbol{x}_j \in \mathbb{R}^m$ et (iii) une modification améliorant le conditionnement de matrices utilisées dans les calculs.

4. Le *cadre non supervisé* sera ensuite abordé : nous proposons l'utilisation d'un algorithme de type *Partially Collapsed Gibbs sampler* (PCGS) échantillonnant la loi *a posteriori* jointe mais bénéficiant d'un échantillonnage des paramètres d'intérêt $(\boldsymbol{q}, \boldsymbol{\nu})$ suivant la loi marginalisée. Un PCGS étant déjà utilisé par [GILC11] pour un modèle BG classique, nous l'adaptons à l'échantillonnage des paramètres non linéaires et nous prouvons sa validité suivant les principes donnés par [vDP08]. Des recommandations quant à l'échantillonnage des hyperparamètres via le réglage de leur loi *a priori* seront données.

5. Nous comparerons les différents échantillonneurs sur les données d'analyse spectrale (§ 1.5.1) où nous montrerons les avantages des échantillonneurs que nous proposons et pour lequel nous proposons une nouvelle stratégie d'estimation de la séquence de Bernoulli, basée sur une étape de fusion des échantillons. Des tests statistiques seront aussi réalisés sur les données de déconvolution impulsionnelle (§ 1.5.2).

La plupart des résultats de ce chapitre (théories et tests) sont synthétisés dans l'article de conférence [BCBB16]. Des tests plus approfondis seront réalisés au chapitre suivant sur des données de vitesses radiales d'étoile pour la détection d'exoplanètes.

## 3.1 Modèle Bernoulli-Gaussien pour la résolution d'APL

Dans cette première section, on se replace temporairement dans le cadre du problème d'APL (1.4), où l'on cherche à estimer les amplitudes $\boldsymbol{x}$ parcimonieuses pour approcher au mieux le signal $\boldsymbol{y}$ par $\mathbf{H}\boldsymbol{x}$, une combinaison linéaire d'atomes du dictionnaire. Nous rappelons les concepts importants du modèle BG classique, afin de pouvoir les étendre dans le cas des problèmes d'APNL dans les sections suivantes.

Afin de décrire de manière structurelle un *a priori* de parcimonie sur les amplitudes, le modèle *Bernoulli-Gaussien* (BG) a été introduit pour la première fois par [KM82] pour la déconvolution de train d'impulsions. Depuis, ce modèle a été utilisé dans de nombreux domaines tels que l'échographie [Cha93], l'astrophysique [Bou05], ou encore les télécommunications [FJA05], et pour résoudre des problèmes différents : la déconvolution de trains d'impulsion [CCL96, CGI96, GILC11, KTHD12], l'estimation de spectres de raies [BC05, Dub96, BHF17, HFR18], la reconstruction de phase [DK15] ou encore l'apprentissage de dictionnaire [DT10] et la séparation de sources [DMTC09].

L'estimation des paramètres du modèle BG a d'abord été réalisée au sens du MAP [KM82], notamment grâce à l'algorithme SMLR dont nous avons parlé en § 1.2.3. Une estimation au sens de l'EAP a quant à elle été envisagée par [CCL96] à partir d'échantillons de la loi *a posteriori* tirés grâce à un algorithme de Gibbs, un échantillonneur de type MCMC. Nous revenons dans cette section sur le modèle Bernoulli-Gaussien classique pour l'échantillonnage de la loi jointe *a posteriori*, dans les cadres supervisé et non supervisé.

### 3.1.1 Expression de la loi *a posteriori*

Nous revenons sur le modèle Bernoulli-Gaussien (BG) décrit en § 1.2.3, pour donner les expressions de la loi *a posteriori*. Par souci de clarté, nous pourrons être amené à rappeler des notions déjà vues précédemment. Pour rappel, le modèle BG, que nous illustrons en Figure 3.1, repose sur l'introduction de variables binaires $\boldsymbol{q} = [q_1, \ldots, q_J]$ indépendantes qui permettent une prise en compte exacte de la parcimonie en conditionnant la nullité des amplitudes $\boldsymbol{x}_j$. Les lois *a priori* pour les variables binaires $q_j$ et pour les amplitudes $\boldsymbol{x}_j$ sont rappelées ci-dessous :

$$\begin{array}{rcll} q_j & \sim & \mathcal{B}(\lambda) & \text{(voir équation (1.16))} \\ \boldsymbol{x}_j \,|\, q_j & \sim & \mathcal{N}(\mathbf{0}, q_j \sigma_x^2 \mathbf{I}_m) & \text{(voir équation (1.17))} \end{array}$$

L'hyperparamètre $\lambda \in [0;1]$ correspond à la probabilité *a priori* de la présence d'un atome $\mathbf{h}_j$ dans le signal ; plus il est proche de 0 et plus l'*a priori* de parcimonie est important. Notons que l'estimation du support $\Omega$ (soit les indices tels que $\boldsymbol{x}_j \neq \mathbf{0}$ - voir (1.5)), le point crucial des problèmes d'APL dans le cadre des problèmes inverses considérés, revient à l'estimation de la séquence de Bernoulli $\boldsymbol{q}$.

Nous avons vu précédemment que le modèle BG est généralement associé à un bruit additif gaussien centré de matrice de covariance $\boldsymbol{\Sigma}$, si bien que le modèle est entièrement décrit par les paramètres $(\boldsymbol{q}, \boldsymbol{x})$ et les hyperparamètres $\boldsymbol{\theta} = (\lambda, \sigma_x^2, \boldsymbol{\Sigma})$. La loi *a posteriori* peut alors se décrire dans le cadre *supervisé* - $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$ - où les hyperparamètres sont considérés fixés et connus, ou dans le cadre *non supervisé* - $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ - où ces derniers sont inconnus. Par la règle de Bayes, on obtient alors :

$$\begin{array}{lrcl} \text{Cadre supervisé} & p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y}, \boldsymbol{\theta}) & \propto & p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta})\, p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{\theta}) \\ \text{Cadre non supervisé} & p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y}) & \propto & p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta})\, p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta}) \\ & & \propto & p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\Sigma})\, p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{\theta})\, p(\boldsymbol{\theta}) \end{array} \tag{3.1}$$

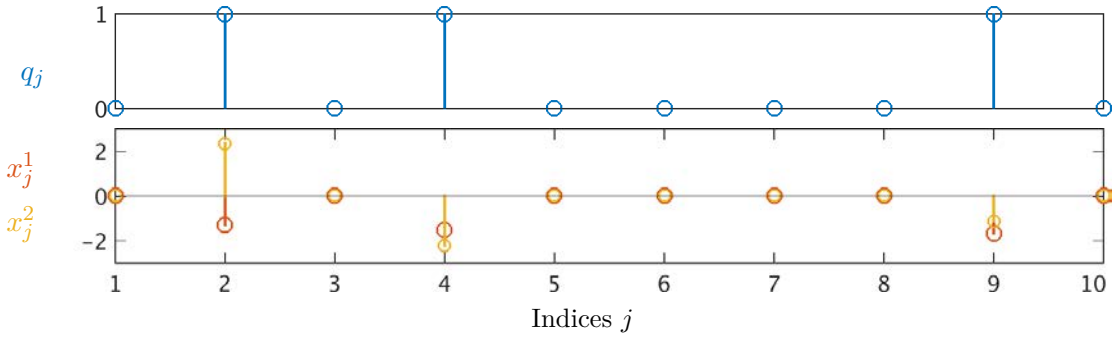



FIG. 3.1 – Illustration du modèle Bernoulli-Gaussien pour des amplitudes bidimensionnelles.

Les expressions de la vraisemblance et des densités *a priori* des variables $(\boldsymbol{x}, \boldsymbol{q})$, qui permettent d'obtenir la loi *a posteriori* dans le cadre supervisé, sont données ci-dessous.

**Expression de la loi *a priori* jointe pour $(\boldsymbol{q}, \boldsymbol{x})$.** Similairement au dictionnaire réduit $\overline{\mathbf{H}}$ et aux amplitudes réduites $\overline{\boldsymbol{x}}$ définies au chapitre 1, nous définissons $\overline{\boldsymbol{q}}$ regroupant les $q_j = 1$ et $\overline{\boldsymbol{q}}^C$ les $q_j = 0$. Par indépendance mutuelle des variables $(q_j, \boldsymbol{x}_j)$, la loi jointe *a priori* sur $(\boldsymbol{x}, \boldsymbol{q})$ s'écrit, conditionnellement aux hyperparamètres :

$$\begin{aligned} p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{\theta}) &= p(\boldsymbol{q}, \boldsymbol{x} \,|\, \lambda, \sigma_x^2) \\ &= \underbrace{\lambda^L (1-\lambda)^{J-L}}_{p(\boldsymbol{q} \,|\, \lambda)} \underbrace{(2\pi)^{-mL/2} \sigma_x^{-mL} \exp\left(-\frac{1}{2\sigma_x^2} \|\overline{\boldsymbol{x}}\|^2\right)}_{p(\overline{\boldsymbol{x}} \,|\, \overline{\boldsymbol{q}}, \sigma_x^2)} \underbrace{\delta_{\mathbf{0}}(\overline{\boldsymbol{x}}^C)}_{p(\overline{\boldsymbol{x}}^C \,|\, \overline{\boldsymbol{q}}^C)} \end{aligned} \tag{3.2}$$

On voit apparaître l'ordre du modèle $L = \|\boldsymbol{q}\| = \sum_j q_j$. La notation $L$ cache la dépendance en les paramètres $q_j$, mais permet de clarifier les expressions. Notons aussi la dépendance de la loi *a priori* en les deux hyperparamètres $\lambda$ et $\sigma_x^2$.

**Expression de la vraisemblance.** Comme le bruit est additif et suit une loi normale centrée de matrice de covariance $\boldsymbol{\Sigma}$, on peut directement en déduire l'expression de la vraisemblance :

$$\begin{aligned} p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta}) &= p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\Sigma}) \\ &= (2\pi)^{-N/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2} \|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}^{-1}}\right) \\ &\propto |\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2} \|\boldsymbol{y} - \overline{\mathbf{H}}\overline{\boldsymbol{x}}\|^2_{\boldsymbol{\Sigma}^{-1}}\right) \end{aligned} \tag{3.3}$$

Pour rappel, la notation $\|\boldsymbol{u}\|^2_{\mathbf{A}}$ désigne le produit $\boldsymbol{u}^T \mathbf{A} \boldsymbol{u}$. La dépendance de la vraisemblance en les paramètres de Bernoulli $\boldsymbol{q}$ se fait au travers des amplitudes $\boldsymbol{x}$ et le produit $\mathbf{H}\boldsymbol{x}$ peut se réduire à $\overline{\mathbf{H}}\overline{\boldsymbol{x}}$.

### 3.1.2 Estimation à partir d'échantillons de la loi *a posteriori*

Supposons que nous disposons de $T$ échantillons des paramètres de la loi *a posteriori* dans le cadre non supervisé :

$$\forall t \in [\![1; T]\!],\ (\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta})^{(t)} \sim (\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$$

La méthode d'échantillonnage stochastique pour obtenir ces échantillons sera exposée en § 3.1.3. Nous donnons ici les procédures d'estimation des paramètres à partir de ces échantillons dans le cadre non supervisé. Notons que le cadre supervisé se déduit simplement en supposant les hyperparamètres connus.

**Estimation au sens de l'EAP.** Rappelons dans un premier temps quelques propriétés de l'estimateur de l'EAP, en commençant par sa définition, pour un paramètre $\boldsymbol{u}$ à valeur continue ou discrète, conditionnellement au signal $\boldsymbol{y}$ :

$$\begin{aligned} \boldsymbol{u} \text{ à valeur continue :} \quad \widehat{\boldsymbol{u}}_{\text{EAP}}(\boldsymbol{y}) &= \mathbb{E}\{\boldsymbol{u} \,|\, \boldsymbol{y}\} = \int p(\boldsymbol{u} \,|\, \boldsymbol{y})\ \boldsymbol{u}\ \mathrm{d}\boldsymbol{u} \\ \boldsymbol{u} \text{ à valeur discrète :} \quad \widehat{\boldsymbol{u}}_{\text{EAP}}(\boldsymbol{y}) &= \mathbb{E}\{\boldsymbol{u} \,|\, \boldsymbol{y}\} = \sum_{\boldsymbol{u}} \Pr(\boldsymbol{u} \,|\, \boldsymbol{y})\ \boldsymbol{u} \end{aligned}$$

Pour alléger les écritures, on notera cet estimateur $\widehat{\boldsymbol{u}}$ dans le suite de ce paragraphe. Par la loi des grands nombres, cet estimateur peut être approché par la moyenne empirique d'échantillons suivant cette loi $\boldsymbol{u}^{(t)} \sim \boldsymbol{u} \,|\, \boldsymbol{y}$ :

$$\widehat{\boldsymbol{u}} = \mathbb{E}\{\boldsymbol{u} \,|\, \boldsymbol{y}\} \approx \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{u}^{(t)} \quad \Leftrightarrow \quad \forall i,\ \widehat{u}_i = \mathbb{E}\{u_i \,|\, \boldsymbol{y}\} \approx \frac{1}{T} \sum_{t=1}^{T} u_i^{(t)} \tag{3.4}$$

On en déduit que l'estimation par moyenne empirique des composantes $\widehat{u}_i$ de $\widehat{\boldsymbol{u}}$ se fait marginalement vis-à-vis des autres composantes.

Revenons au modèle Bernoulli-Gaussien. Pour définir des estimateurs de $(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta})$ au sens de l'EAP, une première façon a été proposée par [CCL96] : il s'agit simplement d'estimer l'EAP *simultanément* sur l'ensemble des paramètres en suivant (3.4). Plus précisément :

$$(\widehat{\boldsymbol{q}}, \widehat{\boldsymbol{x}}, \widehat{\boldsymbol{\theta}}) \approx \frac{1}{T} \sum_{t=1}^{T} \left( \boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{\theta}^{(t)} \right)$$

Deux résultats très intéressants peuvent être déduits de ces estimateurs :

1. $\widehat{q}_j = \mathbb{E}\{q_j \,|\, \boldsymbol{y}\} = \sum_{q_j} q_j \Pr(q_j \,|\, \boldsymbol{y}) = 1 \Pr(q_j = 1 \,|\, \boldsymbol{y}) + 0 \Pr(q_j = 0 \,|\, \boldsymbol{y})$

$$\widehat{q}_j = \Pr(q_j = 1 \,|\, \boldsymbol{y}) \approx \frac{1}{T} \sum_{t=1}^{T} q_j^{(t)}$$

   On en déduit que l'estimateur EAP de la séquence de Bernoulli est équivalent à la probabilité *a posteriori* d'obtenir $q_j = 1$ (ou de façon équivalence celle de la présence de l'atome $\mathbf{h}_j$ dans le signal $\boldsymbol{y}$), que la moyenne empirique des échantillons $q_j^{(t)}$ permet d'approcher.

2. $\widehat{\boldsymbol{x}}_j = \mathbb{E}\{\boldsymbol{x}_j \,|\, \boldsymbol{y}\} = \Pr(q_j = 1 \,|\, \boldsymbol{y}) \mathbb{E}\{\boldsymbol{x}_j \,|\, \boldsymbol{y}, q_j = 1\} + \Pr(q_j = 0 \,|\, \boldsymbol{y}) \mathbb{E}\{\boldsymbol{x}_j \,|\, \boldsymbol{y}, q_j = 0\}$

$$\widehat{\boldsymbol{x}}_j = \widehat{q}_j\ \mathbb{E}\{\boldsymbol{x}_j \,|\, \boldsymbol{y}, q_j = 1\} \approx \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{x}_j^{(t)}$$

   Ainsi, l'estimateur $\widehat{\boldsymbol{x}}_j$ s'exprime comme l'espérance *a posteriori* de $\boldsymbol{x}_j$ conditionnellement à $q_j = 1$ pondérée par la probabilité *a posteriori* que $q_j = 1$. Cet estimateur entraîne donc une sous-estimation des amplitudes $\boldsymbol{x}_j$ (son approximation par la moyenne empirique prend d'ailleurs en compte les amplitudes nulles). Par ailleurs, il empêche un résultat *exactement parcimonieux*, puisque qu'en théorie on n'aura presque sûrement jamais $\Pr(q_j = 1 \,|\, \boldsymbol{y}) = 0$ (bien que cela puisse être vérifié en pratique par la moyenne empirique des échantillons).

Ces résultats montrent qu'une telle estimation « en moyenne » sur les amplitudes ne convient pas au problème d'APL (1.4), notamment dans un cadre de *détection* où l'objectif principal est d'estimer le support de la solution, soit ici la séquence de Bernoulli $\boldsymbol{q}$. En effet, l'estimateur $\widehat{\boldsymbol{q}}$ n'est plus binaire et ne permet plus de prendre en compte de façon exacte la parcimonie.

La procédure envisagée par [Bou05], basée sur une première étape de *détection* de la séquence de Bernoulli puis sur une deuxième étape d'*estimation* des amplitudes, est plus appropriée au problème d'APL. Le schéma d'estimation est le suivant :

1. Estimer $\widehat{\boldsymbol{q}}$ et $\widehat{\boldsymbol{\theta}}$ au sens de l'EAP par moyennage des échantillons, comme vu précédemment.
2. Seuiller la séquence $\widehat{\boldsymbol{q}}$ avec un seuil $\alpha \in [0;1]$ pour obtenir une séquence binaire $\widehat{\boldsymbol{q}}_\alpha$ :

$$\widehat{\boldsymbol{q}}_\alpha = [\widehat{q}_{1,\alpha}, \ldots, \widehat{q}_{J,\alpha}]^T \text{ tel que } \widehat{q}_{j,\alpha} = \begin{cases} 1 & \text{si } \widehat{q}_j \geq \alpha \\ 0 & \text{si } \widehat{q}_j < \alpha \end{cases}$$

   La séquence $\widehat{\boldsymbol{q}}_\alpha$, binaire, modélise exactement la parcimonie.
3. Estimer les amplitudes conditionnellement à la séquence binaire $\widehat{\boldsymbol{q}}_\alpha$ et à $\widehat{\boldsymbol{\theta}}$. On peut démontrer (voir en Annexe B.2) que la loi *a posteriori* conditionnelle $(\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{\theta})$ est une loi gaussienne de moyenne $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Gamma}$, qui dépendent de $\boldsymbol{q}$ et $\boldsymbol{\theta}$ :

$$(\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{\theta}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma}) \Leftrightarrow \begin{cases} \text{avec} & \boldsymbol{\Gamma}^{-1} &= \sigma_x^{-2}\mathbf{I}_{mL} + \overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}} \\ & \boldsymbol{\mu} &= \boldsymbol{\Gamma}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \end{cases} \tag{3.5}$$

   et que $(\overline{\boldsymbol{x}}^C \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{\theta}) \sim \delta_{\mathbf{0}}(\overline{\boldsymbol{x}}^C)$. L'estimation au sens du MAP et au sens de l'EAP conditionnellement à $(\widehat{\boldsymbol{q}}, \widehat{\boldsymbol{\theta}})$ donnent donc le même résultat, à savoir $\mathbf{0}$ pour $\overline{\boldsymbol{x}}^C$ et la moyenne $\boldsymbol{\mu}_\alpha$ de la loi $(\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, \widehat{\boldsymbol{q}}_\alpha, \widehat{\boldsymbol{\theta}})$ comme définie en (3.5) pour $\overline{\boldsymbol{x}}$ :

$$\overline{\boldsymbol{x}}_\alpha = \arg\max_{\overline{\boldsymbol{x}}} p(\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, \widehat{\boldsymbol{q}}_\alpha, \widehat{\boldsymbol{\theta}}) = \mathbb{E}\{\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, \widehat{\boldsymbol{q}}_\alpha, \widehat{\boldsymbol{\theta}}\} = \boldsymbol{\mu}_\alpha$$

Ce schéma d'estimation permet d'obtenir une estimation binaire de la séquence de Bernoulli $\widehat{\boldsymbol{q}}_\alpha$, permettant de prendre en compte exactement la parcimonie dans l'estimation conditionnelle des amplitudes. Contrairement au schéma d'estimation précédent, ces dernières ne seront pas sous-estimées. Notons que la séquence $\widehat{\boldsymbol{q}}$ à valeurs réelles n'est pas sans intérêt, puisque, comme nous l'avons vu précédemment, ses composantes $\widehat{q}_j$ permettent d'estimer la probabilité *a posteriori* de la présence d'un atome $\mathbf{h}_j$ dans le signal $\boldsymbol{y}$. Le seuil $\alpha$ correspond donc à un *degré de confiance* pour la sélection des $\widehat{q}_{j,\alpha} = 1$. Ces différentes séquences estimées sont illustrées en Figure 3.2.

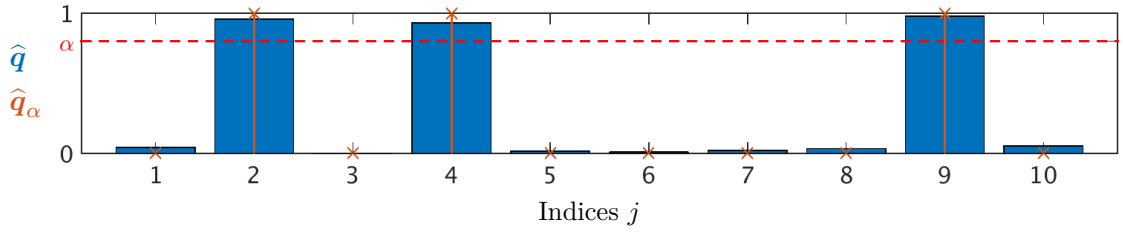



FIG. 3.2 – Illustration de l'estimation au sens de l'EAP de la séquence de Bernoulli $\widehat{\boldsymbol{q}}$ et de la séquence détectée $\widehat{\boldsymbol{q}}_\alpha$ avec un seuil $\alpha = 0.8$.

**Estimation au sens du MAP.** En parallèle, une estimation des paramètres au sens du MAP peut être approchée à partir des $T$ échantillons. Pour rappel, nous avons vu en § 1.2.3 que l'estimation au sens du MAP en maximisant la loi *a posteriori* marginale $\Pr(\boldsymbol{q} \,|\, \boldsymbol{y})$ plutôt que la loi jointe $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$ donne moins de fausses détections. Par ailleurs, la maximisation de $\Pr(\boldsymbol{q} \,|\, \boldsymbol{y})$ est un problème d'optimisation combinatoire pour lequel l'algorithme SMLR donne un optimum local. Une alternative est alors d'approcher l'estimateur du MAP marginal avec celui qui maximise $\Pr(\boldsymbol{q} \,|\, \boldsymbol{y})$ parmi les $T$ échantillons :

$$\widehat{\boldsymbol{q}}_{\text{MAP}}(\boldsymbol{y}) = \arg\max_{\boldsymbol{q}} \Pr(\boldsymbol{q} \,|\, \boldsymbol{y}) \approx \arg\max_{(t)} \Pr(\boldsymbol{q}^{(t)} \,|\, \boldsymbol{y})$$

Cette estimation de $\widehat{\boldsymbol{q}}_{\mathrm{MAP}}(\boldsymbol{y})$ est bien sûr sous-optimale, puisque les $T$ échantillons ne suffiront généralement pas à explorer l'ensemble des $2^J$ configurations possibles, dont notamment $\widehat{\boldsymbol{q}}_{\mathrm{MAP}}(\boldsymbol{y})$. Cependant, la probabilité pour que les échantillons $\boldsymbol{q}^{(t)}$ visitent le vrai maximiseur $\widehat{\boldsymbol{q}}_{\mathrm{MAP}}(\boldsymbol{y})$ est élevée, puisqu'ils suivent la loi *a posteriori*, ce qui justifie donc le recours à cette approche.

Maintenant que nous avons défini les estimateurs des paramètres à partir des échantillons $(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta})^{(t)}$ distribués suivant la loi *a posteriori*, nous expliquons comment obtenir ces derniers.

### 3.1.3 Échantillonnage stochastique de la loi *a posteriori*

La loi *a posteriori* jointe $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ est proportionnelle au produit de la vraisemblance (3.3) avec les lois *a priori*, sur les paramètres (3.2) et sur les hyperparamètres dans le cas non supervisé. La loi *a posteriori* marginale $p(\boldsymbol{q}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ peut aussi être calculée en intégrant la loi jointe vis à vis des amplitudes (nous y reviendrons en § 3.3). Dans les deux cas, l'expression de ces lois ne permet pas un échantillonnage direct de la loi *a posteriori*. Pour obtenir des échantillons, il est donc nécessaire d'avoir recours à des méthodes d'échantillonnage stochastique de type *Monte-Carlo par Chaînes de Markov* (MCMC). Pour résumer, de telles méthodes permettent de simuler des échantillons suivant des lois plus simples que la loi cible (pour nous, la loi *a posteriori*), mais telles qu'il existe des garanties de convergence *asymptotique* vers des échantillons de la loi cible, quelles que soient les valeurs initiales. Les algorithmes de Gibbs et de Métropolis-Hastings (MH), que nous utiliserons dans ce manuscrit, sont tous les deux des exemples d'échantillonneur MCMC.

L'échantillonnage des paramètres a d'abord été envisagé en prenant pour loi cible la loi jointe $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ [CCL96], pour laquelle un algorithme de Gibbs se formule aisément, à la fois dans le cadre supervisé et non supervisé. Nous reviendrons sur l'échantillonnage de la loi marginale $p(\boldsymbol{q}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ en § 3.3.

Nous commençons par quelques rappels et notations pour les échantillonneurs de Gibbs et de Métropolis-Hastings, puis nous présentons les algorithmes de Gibbs pour l'échantillonnage de la loi jointe *a posteriori* du modèle BG, dans les cadres supervisé et non supervisé.

**Quelques rappels et notations pour les échantillonneurs MCMC.** L'échantillonnage de Métropolis-Hastings forme avec l'échantillonnage de Gibbs les deux grandes familles classiques d'échantillonneur MCMC. Dans cet exemple, on veut simuler des échantillons d'une variable $\boldsymbol{u}$ de dimension $N_u$ suivant une loi $p(\boldsymbol{u} \,|\, \boldsymbol{y})$ que nous appelons la *loi cible*.

Le principe de l'échantillonneur de Gibbs est de tirer des échantillons selon les lois conditionnelles *a posteriori*, dans le cas où celles-ci peuvent se simuler directement. L'échantillonnage de la $n^{\mathrm{e}}$ composante à l'itération $(t+1)$ se fait conditionnellement aux $(n-1)$ premières composantes déjà échantillonnées à l'itération $(t+1)$ et aux $(N_u - n)$ composantes suivantes dont on ne connait que les échantillons à l'itération $(t)$. L'algorithme de Gibbs classique est décrit dans le Tableau 3.1.

| |
|---|
| Itération $(t+1)$, à partir de l'échantillon courant $(\boldsymbol{u}^{(t)})$ : |
| Générer, pour tout $n \in [\![1; N_u]\!]$, $u_n^{(t+1)} \sim p(u_n \,|\, \boldsymbol{u}_{-n}^{(t+1,t)}, \boldsymbol{y})$ |

Tab. 3.1 – Algorithme d'échantillonnage de Gibbs classique.

On définit les notations :
– $\boldsymbol{u}_{-n}^{(t+1,t)} = [u_1^{(t+1)}, \ldots, u_{n-1}^{(t+1)}, u_{n+1}^{(t)}, \ldots, u_{N_u}^{(t)}]$ ;
– $\boldsymbol{u}_{n,n+1}^{(t+1,t)} = [u_1^{(t+1)}, \ldots, u_n^{(t+1)}, u_{n+1}^{(t)}, \ldots, u_{N_u}^{(t)}]$.
Cette dernière sera notamment utilisée pour l'échantillonnage de lois hiérarchiques comme la loi Bernoulli-Gaussienne. Ces notations seront uniquement utilisées dans cette section, afin de bien illustrer le caractère itératif de l'échantillonnage. Dans la suite du manuscrit, les indications d'échantillons $(t)$ ne seront indiquées que quand cela semblera nécessaire ; la notation $\boldsymbol{u}_{-n}$ sera quant à elle toujours utilisée pour désigner le vecteur sans la $n^{\text{e}}$ composante (et ce au sens de la parcimonie structurée : par exemple, $\boldsymbol{x}_{-j}$ réunit l'ensemble des $\{\boldsymbol{x}_i\}_{i\neq j}$).

L'échantillonnage de Métropolis-Hastings s'utilise quand on ne peut pas simuler directement suivant la loi cible ou suivant ces lois conditionnelles. À partir de l'échantillon courant $\boldsymbol{u}^{(t)}$, une itération de l'échantillonneur consiste en deux étapes :

1. proposer $\boldsymbol{u}'$ suivant une *loi de proposition* que nous notons $Q(\boldsymbol{u}' \,|\, \boldsymbol{u}^{(t)})$ ;
2. accepter la proposition $\boldsymbol{u}^{(t+1)} \leftarrow \boldsymbol{u}'$ suivant la probabilité d'acceptation $\alpha$ :

$$\alpha = \min(1, \alpha_0) \text{ avec } \alpha_0 = \frac{p(\boldsymbol{u}' \,|\, \boldsymbol{y})}{p(\boldsymbol{u}^{(t)} \,|\, \boldsymbol{y})} \frac{Q(\boldsymbol{u}^{(t)} \,|\, \boldsymbol{u}')}{Q(\boldsymbol{u}' \,|\, \boldsymbol{u}^{(t)})}$$

Pour simplifier les notations par la suite, nous ne ferons pas la différence entre la probabilité d'acceptation $\alpha$ et la valeur $\alpha_0$.

**Échantillonnage de la loi jointe dans le cadre supervisé.** Dans le cadre supervisé, la loi cible pour l'estimation des paramètres du modèle BG est $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$, que l'on écrira ici $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$. On peut montrer que les loi *a posteriori* conditionnelles $p(q_j, \boldsymbol{x}_j \,|\, \boldsymbol{q}_{-j}, \boldsymbol{x}_{-j}, \boldsymbol{y})$ sont des lois Bernoulli-Gaussiennes grâce au caractère conjugué des lois *a priori* sur les amplitudes et sur le bruit (gaussiennes toutes les deux), et se simulent donc facilement. Plus précisément, elles s'écrivent de la façon suivante (voir [Bou05]) :

$$\begin{array}{rcll} q_j & | & \boldsymbol{q}_{-j}, \boldsymbol{x}_{-j}, \boldsymbol{y} & \sim & \mathcal{B}(\lambda_j) \\ \boldsymbol{x}_j & | & \boldsymbol{q} \quad, \boldsymbol{x}_{-j}, \boldsymbol{y} & \sim & \mathcal{N}(\boldsymbol{\mu}_j, q_j \boldsymbol{\Gamma}_j) \end{array} \tag{3.6}$$

où les paramètres $\lambda_j, \boldsymbol{\mu}_j, \Gamma_j$ se calculent de la façon suivante :

$$\begin{aligned} \lambda_j &= \left(1 + \frac{1-\lambda}{\lambda} \frac{\sigma_x^2}{|\boldsymbol{\Gamma}_j|^{1/2}} \exp\left(-\frac{1}{2} \|\boldsymbol{\mu}_j\|^2_{\boldsymbol{\Gamma}_j^{-1}}\right)\right)^{-1} \\ \boldsymbol{\Gamma}_j^{-1} &= \mathbf{h}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j + \sigma_s^{-2} \mathbf{I}_m \\ \boldsymbol{\mu}_j &= \boldsymbol{\Gamma}_j \mathbf{h}_j^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\zeta}_j, \text{ avec } \boldsymbol{\zeta}_j = \boldsymbol{y} - \sum_{i \neq j} \mathbf{h}_i \boldsymbol{x}_i \end{aligned}$$

L'échantillonneur de Gibbs associé est donné en Tableau 3.2. Comme tout échantillonneur de type MCMC, cet échantillonneur de Gibbs est assuré de converger et donc de donner des échantillons de la loi *a posteriori* jointe au bout d'un grand nombre d'itérations.

**Échantillonnage de la loi jointe dans le cas non supervisé.** Dans le cadre non supervisé, la loi cible $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ prend en compte les hyperparamètres à échantillonner, avec pour rappel $\boldsymbol{\theta} = [\lambda, \sigma_x^2, \boldsymbol{\Sigma}]$ : l'*a priori* de parcimonie $\lambda$, la variance des amplitudes $\sigma_x^2$ et la matrice de covariance du bruit $\boldsymbol{\Sigma}$. Nous présentons ici l'échantillonnage de l'ensemble des hyperparamètres, mais notons que l'on peut choisir d'en fixer certains à des valeurs cohérentes ($\lambda$ en lien avec un *a priori* sur l'ordre de parcimonie, $\sigma_x^2$ avec l'ordre de grandeur des amplitudes et $\boldsymbol{\Sigma}$ suivant celui du bruit).

Itération $(t+1)$, à partir des échantillons courants $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)})$ :
Générer pour tout $j \in [\![1; J]\!]$,

$$\begin{array}{rcll} q_j^{(t+1)} & \sim & p(q_j \,|\, \boldsymbol{q}_{-j}^{(t+1,t)}, \boldsymbol{x}_{-j}^{(t+1,t)}, \boldsymbol{y}) & : \ \mathcal{B}(\lambda_j) \\ \boldsymbol{x}_j^{(t+1)} & \sim & p(\boldsymbol{x}_j \,|\, \boldsymbol{q}_{j,j+1}^{(t+1,t)}, \boldsymbol{x}_{-j}^{(t+1,t)}, \boldsymbol{y}) & : \ \mathcal{N}(\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_j) \end{array}$$

TAB. 3.2 – Algorithme d'échantillonnage de Gibbs pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$.

La loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ s'écrit de la même façon que dans le cadre supervisé, à la différence près qu'il faut prendre en compte les lois *a priori* sur les hyperparamètres $p(\boldsymbol{\theta})$ d'après (3.1). L'échantillonnage des paramètres $(\boldsymbol{q}, \boldsymbol{x})$ ne sera donc pas modifié. En supposant l'indépendance entre $\lambda$, $\sigma_x^2$ et $\boldsymbol{\Sigma}$, les lois *a posteriori* conditionnelles en chacun des hyperparamètres s'écrivent de la façon suivante :

$$\begin{array}{rcl} p(\lambda \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta}_{-\lambda}) & \propto & \lambda^L (1-\lambda)^{J-L} \, p(\lambda) \\ p(\sigma_x^2 \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta}_{-\sigma_x^2}) & \propto & (\sigma_x^2)^{-mL/2} \exp\left(-\dfrac{1}{2\sigma_x^2} \|\overline{\boldsymbol{x}}\|^2\right) \, p(\sigma_x^2) \\ p(\boldsymbol{\Sigma} \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta}_{-\boldsymbol{\Sigma}}) & \propto & |\boldsymbol{\Sigma}|^{-1/2} \exp(-\dfrac{1}{2} \|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}^{-1}}) \, p(\boldsymbol{\Sigma}) \end{array} \tag{3.7}$$

Pour rappel, le paramètre $L$ dépend des $q_j$ ($L = \|\boldsymbol{q}\|$). En l'absence d'information complémentaire, on choisira des lois *a priori* $p(\lambda)$, $p(\sigma_x^2)$ et $p(\boldsymbol{\Sigma})$ non-informatives (afin de ne pas influencer l'estimation) et conjuguées à leur loi *a posteriori* conditionnelle, afin que cette dernière puisse être simulée aisément dans le cadre d'un échantillonneur de Gibbs, comme proposé par [Bou05, GILC11] :

- Une loi *a priori* beta pour $\lambda \sim \mathcal{B}e(\alpha, \beta)$, si bien que sa loi *a posteriori* conditionnelle est toujours une loi beta. Le cas $\alpha = \beta = 1$ correspond à une loi *a priori* uniforme sur $[0; 1]$.
- Une loi *a priori* inverse-gamma pour $\sigma_x^2 \sim \mathcal{IG}(\alpha_x, \beta_x)$. Sa loi *a posteriori* conditionnelle sera aussi une loi inverse-gamma. Le cas où $\alpha_x = \beta_x = 0$ correspond à une loi *a priori* de Jeffreys.
- Le cas de la matrice de covariance pour le bruit est plus délicat, puisqu'il s'agit d'une matrice de $N^2$ composantes. Cependant, dans le cas *i.i.d.* elle s'écrit $\boldsymbol{\Sigma} = \sigma_\epsilon^2 \mathbf{I}_N$ et dépend seulement de la variance $\sigma_\epsilon^2$. De même que pour la variance des amplitudes, on choisit une loi *a priori* inverse-gamma pour $\sigma_\epsilon^2 \sim \mathcal{IG}(\alpha_\epsilon, \beta_\epsilon)$ et sa loi *a posteriori* le sera aussi.

Les correspondances entre loi *a priori* et loi *a posteriori* conditionnelle des hyperparamètres sont données dans le Tableau 3.3. Notons que si l'on prend un *a priori* de Jeffreys pour $\sigma_x^2$, *i.e.* $\alpha_x = \beta_x = 0$, la loi *a posteriori* conditionnelle est impropre dans le cas où $L = \sum_j q_j = 0$, puisqu'il s'agit aussi d'une loi de Jeffreys.

Finalement, l'échantillonneur de Gibbs (Tab. 3.2) dans le cas supervisé peut aisément être adapté au cadre non supervisé (Tab. 3.4), en ajoutant les étapes d'échantillonnage des hyperparamètres à chaque itération. Notons que l'ordre d'échantillonnage entre $(\lambda, \sigma_x^2, \sigma_\epsilon^2)$ n'a pas d'importance, puisque leurs lois *a posteriori* conditionnelles sont indépendantes.

*Remarque : Notons que [AD99] proposait un modèle différent pour les variances du bruit et des amplitudes. Notamment, il introduisait un terme en lien avec le rapport signal sur bruit (typiquement $\sigma_x^2/\sigma_\epsilon^2$) lui permettant de se ramener à une dépendance plus simple en les hyperparamètres. Nous n'avons pas utilisé cette possibilité dans nos travaux, notamment parce que le cas où le bruit n'est pas indépendant complique sérieusement ce type de stratégie.*

| Hyperparamètre $\theta$ | Loi *a priori* $p(\theta)$ | Loi *a posteriori* conditionnelle $p(\theta \mid \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta}_{-\theta})$ |
|---|---|---|
| $\lambda$ | $\mathcal{B}e(\alpha, \beta)$<br>$\mathcal{U}([0;1[) \leftrightarrow \mathcal{B}e(1,1)$ | $\mathcal{B}e(L+\alpha, J-L+\beta)$<br>$\mathcal{B}e(L+1, J-L+1)$ |
| $\sigma_x^2$ | $\mathcal{IG}(\alpha_x, \beta_x)$ | $\mathcal{IG}(\alpha_x + mL/2, \beta_x + \|\boldsymbol{x}\|^2/2)$ |
| $\sigma_\epsilon^2$ | $\mathcal{IG}(\alpha_\epsilon, \beta_\epsilon)$ | $\mathcal{IG}(\alpha_\epsilon + N/2, \beta_\epsilon + \|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2/2)$ |

TAB. 3.3 – Récapitulatif des correspondances entre les loi *a priori* sur les hyperparamètres et les lois *a posteriori* conditionnelles qui en résultent dans le cadre du modèle Bernoulli-Gaussien.

Itération $(t+1)$, à partir des échantillons courants $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{\theta}^{(t)})$ :

Générer pour tout $j \in [\![1; J]\!]$,

$$\begin{array}{rcll} q_j^{(t+1)} & \sim & p(q_j \mid \boldsymbol{q}_{-j}^{(t+1,t)}, \boldsymbol{x}_{-j}^{(t+1,t)}, \boldsymbol{\theta}^{(t)}, \boldsymbol{y}) & : \ \mathcal{B}(\lambda_j) \\ \boldsymbol{x}_j^{(t+1)} & \sim & p(\boldsymbol{x}_j \mid \boldsymbol{q}_{j,j+1}^{(t+1,t)}, \boldsymbol{x}_{-j}^{(t+1,t)}, \boldsymbol{\theta}^{(t)}, \boldsymbol{y}) & : \ \mathcal{N}(\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_j) \end{array}$$

Puis générer les hyperparamètres :

$$\begin{array}{rcll} \lambda^{(t+1)} & \sim & p(\lambda \mid \boldsymbol{q}^{(t+1)}, \boldsymbol{x}^{(t+1)}, \boldsymbol{\theta}_{-\lambda}^{(t)}, \boldsymbol{y}) & : \ \mathcal{B}e(L+\alpha, J-L+\beta) \\ \sigma_x^{2(t+1)} & \sim & p(\sigma_x^2 \mid \boldsymbol{q}^{(t+1)}, \boldsymbol{x}^{(t+1)}, \lambda^{(t+1)}, \sigma_\epsilon^{2(t)}, \boldsymbol{y}) & : \ \mathcal{IG}(\alpha_x + \dfrac{Lm}{2}, \beta_x + \dfrac{\|\boldsymbol{x}\|^2}{2}) \\ \sigma_\epsilon^{2(t+1)} & \sim & p(\sigma_\epsilon^2 \mid \boldsymbol{q}^{(t+1)}, \boldsymbol{x}^{(t+1)}, \boldsymbol{\theta}_{-\sigma_\epsilon^2}^{(t+1)}, \boldsymbol{y}) & : \ \mathcal{IG}(\alpha_\epsilon + \dfrac{N}{2}, \beta_\epsilon + \dfrac{1}{2}\|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2) \end{array}$$

TAB. 3.4 – Algorithme d'échantillonnage de Gibbs pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \mid \boldsymbol{y})$ dans le cadre non supervisé.

**Contrôle de convergence des échantillonneurs.** Les échantillonneurs de Gibbs décrits précédemment sont garantis de donner, au bout d'un certain nombre d'itérations, des échantillons suivant la loi *a posteriori* ; $p(\boldsymbol{q}, \boldsymbol{x} \mid \boldsymbol{y})$ dans le cadre supervisé (Tab. 3.2) et $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \mid \boldsymbol{y})$ dans le cadre non supervisé (Tab. 3.4). Cependant, ce nombre d'itérations est inconnu et peut-être *a priori* particulièrement élevé.

Différentes méthodes peuvent être utilisées pour définir un critère d'arrêt ; par exemple, l'échantillonnage peut être arrêté quand l'estimation des paramètres au sens de l'EAP n'évolue plus ; voir par exemple [BC05]. Une méthode plus justifiable pour déterminer si un échantillonnage a convergé est par exemple l'utilisation du diagnostic de [BG98], le *Multivariate Potential Scale Reduction Factor* (MPSRF), permettant de comparer les variances inter-chaînes et intra-chaînes. Plus précisément, on considère une variable $\boldsymbol{u}$ pour laquelle on dispose de $N_c$ chaînes de Markov indépendantes, constituées chacune de $T$ échantillons. L'échantillon $t$ de la chaîne $n$ sera noté $\boldsymbol{u}^{t,n}$, et on remplacera l'indice ($t$ ou $n$) par un point pour désigner la moyenne sur l'ensemble des échantillons et/ou chaines ; par exemple, $\boldsymbol{u}^{\cdot,n}$ est la moyenne des $T$ échantillons de la $n^{\text{e}}$ chaine. On peut alors définir les matrices de covariances intra- et inter-classes suivantes :

$$\mathbf{V}_{\text{intra}} = \frac{1}{N_c(T-1)} \sum_{n=1}^{N_c} \sum_{t=1}^{T} (\boldsymbol{u}^{t,n} - \boldsymbol{u}^{\cdot,n})(\boldsymbol{u}^{t,n} - \boldsymbol{u}^{\cdot,n})^T$$
$$\mathbf{V}_{\text{inter}} = \frac{1}{N_c - 1} \sum_{n=1}^{N_c} (\boldsymbol{u}^{\cdot,n} - \boldsymbol{u}^{\cdot,\cdot})(\boldsymbol{u}^{\cdot,n} - \boldsymbol{u}^{\cdot,\cdot})^T$$

Le MPSRF s'écrit alors :

$$\text{MPSRF} = \frac{T-1}{T} + \frac{N_c+1}{N_c} \text{eig}(\mathbf{V}_{\text{intra}}^{-1} \mathbf{V}_{\text{inter}}) \tag{3.8}$$

où eig$(\cdot)$ renvoie la plus grande valeur propre de la matrice en argument. Un MPSRF proche de 1 permet de diagnostiquer la convergence des échantillonnages ; [BG98] préconise de prendre 1.2 pour valeur seuil. Notons que ce diagnostic est utilisé pour l'échantillonnage d'un modèle BG pour de la déconvolution impulsionnelle linéaire dans [GILC11], sur lequel nous nous sommes basés pour ces explications.

Ce diagnostic par MPSRF est plus juste théoriquement que le premier critère d'arrêt proposé. En effet, après convergence, les échantillons sont tirés suivant la loi *a posteriori*, pouvant par exemple être bimodale et entraînant l'estimation EAP à évoluer au cours des échantillons. Le MPSRF évalue en plus si les échantillons tirés se ressemblent entre plusieurs chaînes indépendantes ; au sein d'une même chaîne, les échantillons peuvent évoluer de façon significative, mais si c'est le cas pour toutes les chaînes et de façon assez similaire, on peut alors supposer qu'il y a convergence. Le coût de calcul du MPSRF est à prendre en compte, puisqu'il nécessite la simulation de plusieurs chaînes en parallèle, et les calculs coûteux de valeurs propres et d'inversions matricielles.

**Défauts de l'algorithme de Gibbs et solutions pour accélérer la convergence.** Bien qu'en théorie, les algorithmes de Gibbs soient assurés de converger (*i.e.* les échantillons finissent par suivre la loi jointe *a posteriori*), cette convergence peut s'avérer particulièrement lente. L'algorithme de Gibbs est en effet connu pour rester bloqué dans des modes locaux. Par exemple, la Figure 3.3 illustre deux cas particuliers de configuration (en gris) où l'échantillonneur peut rester bloqué. La configuration à atteindre est représentée en couleurs.

- indices $(2, 3, 4)$ : la configuration optimale est difficile à atteindre, puisqu'elle nécessite de mettre à zéro $q_2$ et $q_4$, et de mettre à un $q_3$.
- indices $(8, 9)$ : la configuration optimale est difficile à atteindre, puisqu'elle nécessite de mettre à zéro $q_9$ et de mettre à un $q_8$.

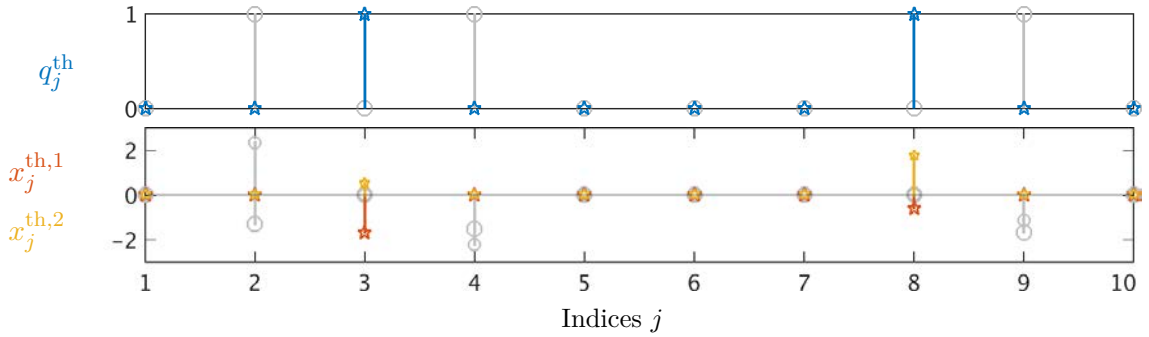



FIG. 3.3 – Exemples typiques de configurations favorisant les modes locaux pour l'échantillonneur de Gibbs. En gris est représentée la configuration courante des échantillons (identique à celle du schéma 3.1). En couleurs sont représentées les valeurs théoriques vers lesquelles l'échantillonneur devrait converger.

Des solutions ont été proposées afin d'accélérer la convergence de l'algorithme, et d'échapper à de tels modes locaux, dans le cadre du modèle Bernoulli-Gaussien :

- Une étape d'échantillonnage additionnelle adaptée au problème. Dans le cadre de l'analyse spectrale en échantillonnage irrégulier, [Bou05] a proposé d'ajouter des étapes d'échantillonnage supplémentaires de la séquence de Bernoulli, pour proposer des « mouvements » permettant, en une itération, de sortir des modes locaux identifiés en Figure 3.3.
- Un échantillonnage par *K-uplets*. Plutôt que d'échantillonner les paramètres $(q_j, \boldsymbol{x}_j)$ les uns après les autres, [GILC11] a proposé de les échantillonner par K-uplets, afin de permettre des transitions plus rapides.
- Un échantillonnage de la loi *a posteriori* marginalisée suivant les amplitudes $p(\boldsymbol{q}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ plutôt que la loi jointe. Cette perspective a longtemps été ignorée pour des raisons de calcul, comme le justifie [Bou05]. Cependant, suite aux travaux de [GILC11] permettant une accélération des calculs, l'échantillonnage de la loi *a posteriori* marginalisée peut s'avérer plus efficace. Nous reviendrons plus en détail sur cette méthode en § 3.3.

Nous avons ainsi vu comment le modèle Bernoulli-Gaussien peut être exploité pour résoudre des problèmes classiques d'APL. Nous revenons maintenant au problème d'APNL qui nous intéresse, où les atomes dépendent des paramètres non linéaires $\boldsymbol{\nu}_j$ à estimer. Le modèle BG va donc être étendu à la prise en compte des $\boldsymbol{\nu}_j$.

## 3.2 Modèle BGE : premiers algorithmes d'échantillonnage de la loi jointe

Nous revenons maintenant aux problèmes d'approximation parcimonieuse non linéaire que nous avons présentés en § 1.3.2. Pour rappel, ces problèmes s'écrivent :

$$\text{Estimer } \boldsymbol{x} \textit{ parcimonieux} \text{ et } \boldsymbol{\nu} \in \mathfrak{I} \text{ tels que } \underbrace{\boldsymbol{y}}_{\text{Aléa.}} = \sum_{j=1}^{J} \underbrace{\mathbf{h}}_{\text{Déter.}} \underbrace{(\underbrace{\boldsymbol{\nu}_j}_{\text{Aléa.}})}_{} \underbrace{\boldsymbol{x}_j}_{\text{Aléa.}} + \underbrace{\boldsymbol{\epsilon}}_{\text{Aléa.}} \tag{3.9}$$

où nous avons indiqué quels paramètres proviennent de la réalisation d'une variable aléatoire. Pour estimer les paramètres non linéaires avec une approche stochastique grâce à une extension du modèle Bernoulli-Gaussien vu précédemment, il est donc nécessaire de définir leur loi *a priori* pour les introduire dans le modèle bayésien et en déduire une procédure d'estimation. Dans cette section, nous rappelons d'abord le modèle BG étendu proposé par [BC06] dans le cas de l'analyse spectrale, puis nous revenons sur une modélisation BG à dimension variable reposant sur des travaux antérieurs à cette thèse [Bou14]. Cette modélisation nous permet de justifier la validité de l'échantillonnage des paramètres du modèle BGE.

### 3.2.1 Modèle BG étendu et algorithme Gibbs-hybride

Le modèle Bernoulli-Gaussien étendu (BGE), introduit par [BC06] et plus en détails dans [Bou05], a été utilisé pour des problèmes d'analyse spectrale ; de tels problèmes ne présentent qu'un paramètre non linéaire, la fréquence, qui joue le rôle de paramètre de localisation. Cependant, le modèle, la procédure d'estimation ainsi que l'algorithme d'échantillonnage s'étendent très facilement à la présence de paramètres de localisation. Le modèle BGE présenté ici est donc celui proposé par [BC06], avec la légère adaptation que nous amenons à la présence des paramètres de forme, soit $\boldsymbol{\nu}_j = (\nu_j, \upsilon_j)$.

**Modèle BGE.** Le modèle BGE ajoute l'ensemble des paramètres $\{\boldsymbol{\nu}_j\}_j$ au modèle BG classique, si bien que les paramètres à estimer dans le cadre supervisé sont les $\{q_j, \boldsymbol{\nu}_j, \boldsymbol{x}_j\}_j$. Il repose sur un *a priori* d'indépendance statistique entre les paramètres non linéaires $\boldsymbol{\nu}_j$ et les paramètres BG $(q_j, \boldsymbol{x}_j)$ ; [BC06] propose ainsi de donner un *a priori* uniforme sur l'intervalle $\mathcal{I}_j$ défini en équation (1.19) pour chaque paramètre de localisation $\nu_j$, quelle que soit la variable $q_j$ correspondante. En supposant les paramètres de forme $\boldsymbol{\upsilon}_j$ indépendant des $\nu_j$ et en leur donnant une loi *a priori* uniforme dans leur intervalle de définition, on obtient :

$$\forall j \in [\![1; J]\!],\ \boldsymbol{\nu}_j \sim \mathcal{U}(\mathfrak{I}_j) \ \Leftrightarrow\ \nu_j \sim \mathcal{U}(\mathcal{I}_j) \text{ et } \boldsymbol{\upsilon}_j \sim \mathcal{U}(\mathcal{I}_{\boldsymbol{\upsilon}})$$

Notons que cette loi *a priori* n'introduit pas d'hyperparamètre supplémentaire, et que les paramètres $\{\boldsymbol{\nu}_j\}_j$ sont *a priori* indépendants entre eux. Le modèle BGE est illustré en Figure 3.4, dans le cas où $\boldsymbol{\upsilon}_j = \varnothing$. Contrairement aux amplitudes $\boldsymbol{x}_j$ dont la nullité est induite par celle de $q_j$, les paramètres de localisation $\nu_j$ prennent bien une valeur indépendante de $q_j$.

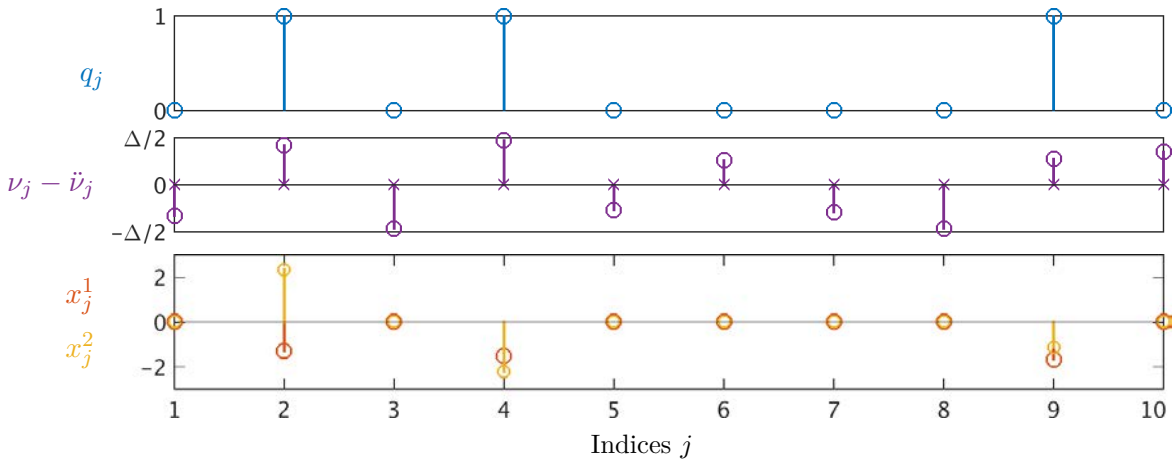



FIG. 3.4 – Illustration du modèle Bernoulli-Gaussien étendu. Pour les paramètres de localisation (2e figure - en violet), la valeur représentée est la différence entre le paramètre $\nu_j$ et le milieu $\ddot{\nu}_j$ de l'intervalle $\mathcal{I}_j$.

La loi *a posteriori* du modèle BGE s'écrit de la façon suivante, dans les cadres supervisé et non supervisé :

$$\begin{aligned} \text{Cadre supervisé} \quad & p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x} \,|\, \boldsymbol{y}, \boldsymbol{\theta}) \ \propto\ p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta})\ p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{\theta})\ p(\boldsymbol{\nu}) \\ \text{Cadre non supervisé} \quad & p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y}) \ \propto\ p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta})\ p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\theta}) \\ & \qquad\qquad \propto\ p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\Sigma})\ p(\boldsymbol{\nu})\ p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{\theta})\ p(\boldsymbol{\theta}) \end{aligned} \tag{3.10}$$

Les seules différences avec les lois *a posteriori* dans le cas du modèle BG classique (3.1) sont (i) la présence de la loi *a priori* sur les paramètres non linéaires, qui n'interviendra que dans les lois *a posteriori* conditionnelles $p(\boldsymbol{\nu}_j \,|\, \boldsymbol{y}, -)$ et (ii) la dépendance de la vraisemblance en les paramètres non linéaires. Cette dernière a la même expression analytique que celle de l'équation (3.3), à la différence près que le dictionnaire réduit $\overline{\mathbf{H}}$ dépend maintenant des paramètres non linéaires :

$$p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta}) \propto |\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}\|\boldsymbol{y} - \overline{\mathbf{H}}(\overline{\boldsymbol{\nu}})\overline{\boldsymbol{x}}\|^2_{\boldsymbol{\Sigma}^{-1}}\right) \tag{3.11}$$

où on définit la notation $\overline{\boldsymbol{\nu}}$ qui contient les paramètres $\boldsymbol{\nu}_j$ associés à $q_j = 1$. Par la suite, on notera $\overline{\mathbf{H}}(\overline{\boldsymbol{\nu}})$ simplement $\overline{\mathbf{H}}$ pour alléger les notations. Grâce à l'indépendance statistique

entre les $\boldsymbol{\nu}_j$ et $(q_j, \boldsymbol{x}_j)$, les lois *a posteriori* conditionnelles en les paramètres $\{q_j, \boldsymbol{x}_j\}_j$ (3.6) et $\boldsymbol{\theta}$ (3.7), nécessaires pour un échantillonnage de Gibbs, sont donc identiques à celles du modèle BG classique ; la dépendance du dictionnaire $\mathbf{H}$ et des atomes $\mathbf{h}(\cdot)$ en les paramètres $\boldsymbol{\nu}_j$ doit apparaître, mais les formules analytiques restent inchangées. Notons qu'il en va de même pour la loi *a posteriori* conditionnelle en les amplitudes réduites :

$$(\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma}) \tag{3.12}$$

où $(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ ont la même expression qu'en équation (3.5), mais dépendent de $\overline{\boldsymbol{\nu}}$. L'estimation se fait toujours grâce à des échantillons $(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta})^{(t)}$ de la loi cible à partir desquels les estimateurs peuvent être décrits.

**Procédure d'estimation.** L'estimation des paramètres au sens de l'EAP à partir des échantillons $(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta})^{(t)}$ suit la même démarche séquentielle que dans le cas du modèle BG classique : on commence par calculer $\widehat{\boldsymbol{\theta}}$ et $\widehat{\boldsymbol{q}}$ en effectuant la moyenne empirique des échantillons, puis la séquence de Bernoulli binaire $\widehat{\boldsymbol{q}}_\alpha$ par seuillage de $\widehat{\boldsymbol{q}}$. Les amplitudes $\overline{\boldsymbol{x}}$ peuvent toujours être estimées en prenant le maximum/la moyenne $\boldsymbol{\mu}$ de sa loi *a posteriori* conditionnelle $p(\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, \widehat{\boldsymbol{q}}_\alpha, \widehat{\overline{\boldsymbol{\nu}}}_\alpha, \widehat{\boldsymbol{\theta}})$ (3.12), mais cela nécessite alors une estimation préalable des paramètres non linéaires $\widehat{\overline{\boldsymbol{\nu}}}_\alpha$. Ces derniers sont estimés via les échantillons $\overline{\boldsymbol{\nu}}_\alpha^{(t)}$ (la réduction de $\boldsymbol{\nu}$ à $\overline{\boldsymbol{\nu}}_\alpha$ se fait via les indices $j$ tels que $\widehat{q}_{j,\alpha} = 1$) conditionnellement à $\widehat{\boldsymbol{q}}_\alpha$, soit :

$$\widehat{\overline{\boldsymbol{\nu}}}_\alpha = \mathbb{E}\{\overline{\boldsymbol{\nu}}_\alpha \,|\, \boldsymbol{y}, \widehat{\boldsymbol{q}}_\alpha\} \approx \frac{1}{T_\alpha} \sum_{t \,|\, \boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha} \overline{\boldsymbol{\nu}}_\alpha^{(t)}, \text{ avec } T_\alpha = \text{Card}\left\{t \in [\![1; T]\!] \,|\, (\boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha)\right\} \tag{3.13}$$

Par ailleurs, les échantillons $\overline{\boldsymbol{\nu}}_\alpha^{(t)}$ donnent aussi accès à un estimateur de leur variance. Par exemple, pour le paramètre de localisation, on a :

$$\text{Pour } j \text{ tel que } \widehat{q}_{j,\alpha} = 1, \quad \widehat{\sigma}^2_{\nu_{j,\alpha}} \approx \frac{1}{T_\alpha - 1} \sum_{t \,|\, \boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha} (\nu_j^{(t)} - \widehat{\nu}_{j,\alpha})^2 \tag{3.14}$$

Cette définition s'étend facilement aux paramètres de forme grâce à leur indépendance statistique avec les paramètres de localisation. Cette estimation est illustrée en Figure 3.5, uniquement pour les paramètres de localisation.

**Algorithme d'échantillonnage.** Afin d'obtenir des échantillons suivant la loi cible $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$, l'échantillonneur de Gibbs classique (Tab. 3.4) doit être étendu à la présence des non-linéarités. Les paramètres $(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta})$ peuvent s'échantillonner directement suivant leur loi conditionnelle, comme dans le cas classique ; celles-ci dépendront néanmoins des $\boldsymbol{\nu}_j^{(t)}$ courants. Au contraire, la loi *a posteriori* conditionnelle $p(\boldsymbol{\nu}_j \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{\nu}_{-j}, \boldsymbol{x}, \boldsymbol{\theta})$ ne se simule pas directement, puisque le modèle dépend de façon non linéaire des $\boldsymbol{\nu}_j$ :

$$p(\boldsymbol{\nu}_j \,|\, \boldsymbol{y}, \boldsymbol{q}, \boldsymbol{\nu}_{-j}, \boldsymbol{x}, \boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{\zeta}_j - \mathbf{h}(\boldsymbol{\nu}_j)\boldsymbol{x}_j\|^2_{\boldsymbol{\Sigma}^{-1}}\right) 1_{\mathfrak{I}_j}(\boldsymbol{\nu}_j) \tag{3.15}$$

où $\boldsymbol{\zeta}_j = \boldsymbol{y} - \sum_{i \neq j} \mathbf{h}(\boldsymbol{\nu}_i)\boldsymbol{x}_i$. [BC06] propose donc d'échantillonner $\nu_j$ grâce à une étape de Métropolis-Hastings, que nous étendons à la présence des paramètres de forme $\boldsymbol{\upsilon}_j$. La proposition d'un candidat $\boldsymbol{\nu}'_j$ se fait suivant la loi de mélange suivante :

$$\begin{array}{rrcll} & Q(\boldsymbol{\nu}'_j \,|\, \boldsymbol{\nu}_j^{(t)}) & = & p_{\text{rw}}\, Q_1(\boldsymbol{\nu}'_j \,|\, \boldsymbol{\nu}_j^{(t)}) + (1 - p_{\text{rw}})\, Q_2(\boldsymbol{\nu}'_j \,|\, \boldsymbol{\nu}_j^{(t)}) & \\ \text{avec} & Q_1(\boldsymbol{\nu}'_j \,|\, \boldsymbol{\nu}_j^{(t)}) \;:\; \boldsymbol{\nu}'_j & \sim & \mathcal{N}(\boldsymbol{\nu}_j^{(t)}, \sigma^2_{\text{rw}} \mathbf{I}_p) & \text{(marche aléatoire autour de } \boldsymbol{\nu}_j^{(t)}) \\ & Q_2(\boldsymbol{\nu}'_j \,|\, \boldsymbol{\nu}_j^{(t)}) \;:\; \boldsymbol{\nu}'_j & \sim & \mathcal{U}(\mathfrak{I}_j) & \text{(loi } a \text{ priori)} \end{array} \tag{3.16}$$

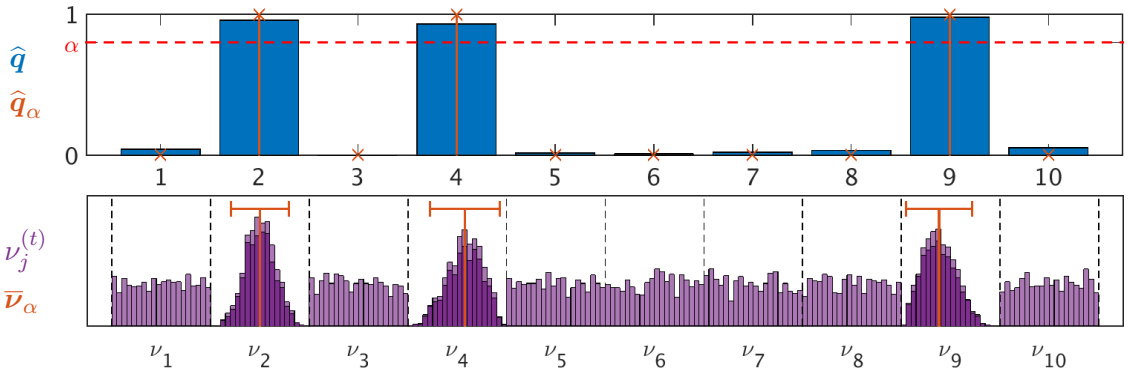



FIG. 3.5 – Illustration de l'estimation séquentielle de la séquence de Bernoulli et des paramètres de localisation. En haut : estimation au sens de l'EAP de la séquence de Bernoulli $\widehat{\boldsymbol{q}}$ et de la séquence détectée $\widehat{\boldsymbol{q}}_\alpha$ avec un seuil $\alpha = 0.8$. En bas : en violet clair, histogrammes des échantillons $\nu_j^{(t)}$ ; en violet foncé, histogrammes des échantillons $\overline{\boldsymbol{\nu}}^{(t)}$ tels que $\boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha$ ; en rouge, localisations estimées $\widehat{\overline{\boldsymbol{\nu}}}_\alpha$ associées à leur barre d'erreur à $\pm 2\,\widehat{\sigma}_{\nu_j,\alpha}$.

Cette loi de mélange permet d'assurer un taux d'acceptation suffisamment élevé (grâce à un $\sigma^2_{\text{rw}}$ suffisamment petit - ce dernier peut être adapté à chaque composante de $\boldsymbol{\nu}_j$) et une exploration de l'espace efficace grâce aux propositions suivant la loi *a priori*. La probabilité d'acceptation faisant intervenir le rapport des lois *a priori* $p(\boldsymbol{\nu}'_j)/p(\boldsymbol{\nu}_j)$, elle sera nécessairement nulle si $\boldsymbol{\nu}'_j$ est proposé en dehors de l'intervalle $\mathfrak{I}_j$ à cause de la marche aléatoire. La proposition est alors répétée jusqu'à ce que $\boldsymbol{\nu}'_j \in \mathcal{I}_j$ ; on peut considérer que le rapport des lois *a priori* est égal à 1. Comme $Q(\boldsymbol{\nu}'_j \,|\, \boldsymbol{\nu}_j^{(t)})$ est symétrique (grâce à l'*a priori* uniforme sur $\boldsymbol{\nu}_j$), la probabilité d'acceptation $\alpha$ s'écrit uniquement comme le rapport de la loi cible (3.15), ce qui donne :

$$\alpha = \exp\left(-\frac{1}{2}(\|\boldsymbol{\zeta}_j - \mathbf{h}(\boldsymbol{\nu}'_j)\boldsymbol{x}_j\|^2_{\boldsymbol{\Sigma}^{-1}} - \|\boldsymbol{\zeta}_j - \mathbf{h}(\boldsymbol{\nu}_j^{(t)})\boldsymbol{x}_j\|^2_{\boldsymbol{\Sigma}^{-1}})\right) \tag{3.17}$$

Remarquons que dans le cas où $q_j = 0$ (et donc $\boldsymbol{x}_j = \mathbf{0}$), cette probabilité d'acceptation est égale à 1 et la proposition $\boldsymbol{\nu}'_j$ est toujours acceptée. Tant que $q_j = 0$, le paramètre $\boldsymbol{\nu}_j$ évolue donc indépendamment de $\boldsymbol{y}$ en suivant la loi de proposition.

L'échantillonneur proposé est donc une version hybride de l'algorithme de Gibbs classique, puisqu'il inclut une étape de Métropolis-Hastings. Sa convergence vers la loi cible est garantie [Rob96]. Il est récapitulé dans le cas non supervisé en Tableau 3.5.

**Limites de l'échantillonneur de Gibbs-hybride.** La procédure d'estimation basée sur les échantillons de la loi jointe $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ obtenus grâce à l'échantillonneur Gibbs-hybride (Tab. 3.5) permet donc en théorie l'estimation de l'ensemble des paramètres, dont notamment la séquence de Bernoulli et les paramètres de localisation (les fréquences dans le cadre de l'analyse spectrale). [BC06] souligne néanmoins que la lenteur de convergence de l'échantillonneur de Gibbs et notamment la possibilité de se retrouver bloqué dans un mode local, peut s'aggraver. En effet, l'indépendance statistique supposée entre les variables BG $(q_j, \boldsymbol{x}_j)$ et les fréquences $\nu_j$ est nécessaire pour faciliter l'échantillonnage, mais entraîne une incohérence. La proposition d'un $\nu'_j$ lors de l'étape MH sera acceptée conditionnellement à la valeur courante de l'amplitude $\boldsymbol{x}_j$. Cela peut entraîner un fort taux de rejet, puisque l'amplitude adaptée pour la fréquence $\nu_j$ courante ne l'est pas nécessairement pour la proposition $\nu'_j$. Au problème d'incohérence dans l'échantillonnage s'ajoute la *non-existence statistique* des paramètres non linéaires $\boldsymbol{\nu}_j$ quand $\boldsymbol{q}_j = 0$, puisque l'amplitude nulle associée annule la participation de $\mathbf{h}(\boldsymbol{\nu}_j)$ dans la vraisemblance. C'est dans

Itération $(t+1)$, à partir des échantillons courants $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{\nu}^{(t)}, \boldsymbol{\theta}^{(t)})$ :

Générer pour tout $j \in [\![1; J]\!]$,

$$\begin{array}{rcll} q_j^{(t+1)} & \sim & (q_j \,|\, \boldsymbol{q}_{-j}, \boldsymbol{x}_{-j}, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y}) & : \quad \mathcal{B}(\lambda_j) \\ \boldsymbol{x}_j^{(t+1)} & \sim & (\boldsymbol{x}_j \,|\, \boldsymbol{q}, \boldsymbol{x}_{-j}, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y}) & : \quad \mathcal{N}(\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_j) \\ \boldsymbol{\nu}_j^{(t+1)} & \sim & (\boldsymbol{\nu}_j \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}_{-j}, \boldsymbol{\theta}, \boldsymbol{y}) & : \quad \text{étape Métropolis-Hastings (3.16)-(3.17)} \end{array}$$

Puis, générer les hyperparamètres :

$$\begin{array}{rcll} \lambda^{(t+1)} & \sim & (\lambda \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\theta}_{-\lambda}, \boldsymbol{y}) & : \quad \mathcal{B}e(L+\alpha, J-L+\beta) \\ \sigma_x^{2\,(t+1)} & \sim & (\sigma_x^2 \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \lambda, \sigma_\epsilon^2, \boldsymbol{y}) & : \quad \mathcal{IG}(\alpha_x + \dfrac{Lm}{2}, \beta_x + \dfrac{\|\boldsymbol{x}\|^2}{2}) \\ \sigma_\epsilon^{2\,(t+1)} & \sim & (\sigma_\epsilon^2 \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\theta}_{-\sigma_\epsilon^2}, \boldsymbol{y}) & : \quad \mathcal{IG}(\alpha_\epsilon + \dfrac{N}{2}, \beta_\epsilon + \dfrac{1}{2}\|\boldsymbol{y} - \mathbf{H}(\boldsymbol{\nu})\boldsymbol{x}\|^2) \end{array}$$

TAB. 3.5 – Algorithme d'échantillonnage de Gibbs-hybride pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ dans le cadre non supervisé.

l'optique de résoudre ces problèmes que nous avons proposé, lors du projet de master [Bou14], un algorithme d'échantillonnage faisant appel à une étape de Reversible-Jump (RJ) [Gre95]. Nous revenons donc rapidement sur cet algorithme.

### 3.2.2 Modèle BG à dimension variable et algorithme RJGS

Dans le projet de Master [Bou14], nous avions proposé de modéliser la *non-existence* statistique des paramètres non linéaires $\boldsymbol{\nu}_j$ (et donc des atomes non linéaires associés) quand $q_j = 0$ par leur retrait du paradigme d'estimation, accompagné de celui des amplitudes. Ainsi, la variable $\boldsymbol{\kappa}_j = (\boldsymbol{x}_j, \boldsymbol{\nu}_j)$ varie de dimension suivant la valeur de $q_j$ :

$$\begin{array}{lll} \boldsymbol{\kappa}_j \,|\, q_j = 0 & \text{est de dimension} & 0 \\ \boldsymbol{\kappa}_j \,|\, q_j = 1 & \text{est de dimension} & m+p \quad \text{(dimensions de } \boldsymbol{x}_j \text{ et } \boldsymbol{\nu}_j\text{)} \end{array}$$

Cet *a priori* de *non-existence statistique* imposé par le changement de dimension est plus rigoureux que le modèle BGE proposé par [Bou05], puisqu'il permet d'inférer la non-influence des paramètres non linéaires sur la vraisemblance quand $q_j = 0$. Au contraire du modèle BGE où les paramètres $\boldsymbol{\nu}_j$ évoluent aléatoirement et indépendamment de $\boldsymbol{y}$ quand $q_j = 0$, ici ils n'existent tout simplement pas.

#### Échantillonneur de Gibbs avec étape de Reversible-Jump

Pour échantillonner les paramètres $\boldsymbol{\kappa}_j$ de dimension variable, nous avons fait appel à une étape de Reversible-Jump [Gre95], insérée au sein d'un algorithme de Gibbs. L'algorithme associé, appelé *Reversible-Jump Gibbs Sampler* (RJGS), est donné en Tableau 3.6. Le principe est de proposer 3 événements différents suivant la valeur courante de $q_j$ :

**Naissance** quand $q_j = 0$, on propose de passer à $q_j = 1$ en faisant « naître » les paramètres $\boldsymbol{\kappa}'_j$ suivant leur loi *a priori* $p(\boldsymbol{\kappa}'_j)$ ; un atome est ainsi ajouté à l'approximation parcimonieuse.

**Mort** quand $q_j = 1$, on propose avec une probabilité $p_{10}$ de passer à $q_j = 0$ et de faire « mourir » les paramètres $\boldsymbol{\kappa}_j$ en les supprimant ; un atome est ainsi retiré à l'approximation parcimonieuse.

**Vie** quand $q_j = 1$, on propose avec une probabilité $p_{11} = 1 - p_{10}$ de rester à $q_j = 1$ et de laisser « vivre » les paramètres $\boldsymbol{\kappa}_j$ en les faisant évoluer vers $\boldsymbol{\kappa}'_j$ grâce à une marche aléatoire par exemple (tout autre loi de proposition symétrique $Q(\boldsymbol{\kappa}'_j \,|\, \boldsymbol{\kappa}_j) = Q(\boldsymbol{\kappa}_j \,|\, \boldsymbol{\kappa}'_j)$ donnant la même probabilité d'acceptation $\alpha_{11}(\boldsymbol{\kappa}_j, \boldsymbol{\kappa}'_j)$) ; l'atome déjà présent dans l'approximation parcimonieuse est ainsi modifié suivant l'évolution de $\boldsymbol{\kappa}_j$ vers $\boldsymbol{\kappa}'_j$.

La validité de ces évenements a été vérifiée en suivant l'article fondateur du Reversible-Jump [Gre95] et le tutoriel [WS01]. Nous montrons dans le paragraphe suivant que le formalisme amené par le *Reversible-Jump* permet en fait de montrer la validité de l'échantillonnage du modèle BGE.

Itération $(t+1)$. Suivant un état $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{\nu}^{(t)})$ :
– pour $j = 1...J$, on tire $(q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j) \,|\, (\boldsymbol{y}, \boldsymbol{q}_{-j}, \boldsymbol{x}_{-j}, \boldsymbol{\nu}_{-j})$ de la façon suivante :
  * À partir de l'état courant $X = (\pi, \boldsymbol{\kappa}_j)$, où $\pi = q_j$ (0 ou 1) et :

$$\begin{aligned} \boldsymbol{\kappa}_j \,|\, (\pi = 0) &= \varnothing \\ \boldsymbol{\kappa}_j \,|\, (\pi = 1) &= (\boldsymbol{\nu}_j, \boldsymbol{x}_j) \end{aligned}$$

  * on propose d'abord $\pi' \sim \mathcal{B}(p_{\pi\pi'})$, puis 3 cas sont possibles pour la proposition de $\boldsymbol{\kappa}'_j$ :
    1. $\pi = 0, \pi' = 1$ : *Naissance*, $q_j$ passe de 0 à 1, on fait « naître » les paramètres $\boldsymbol{\kappa}'_j$ suivant leur loi *a priori* : $Q(\boldsymbol{\kappa}'_j \,|\, \pi, \pi') = p(\boldsymbol{\kappa}'_j)$.
    2. $\pi = 1, \pi' = 0$ : *Mort*, $q_j$ passe de 1 à 0, on fait « mourir » les paramètres $\boldsymbol{\kappa}_j$.
    3. $\pi = 1, \pi' = 1$ : *Vie*, $q_j$ reste à 1, on fait évoluer les paramètres $\boldsymbol{\kappa}_j$ vers $\boldsymbol{\kappa}'_j$ à l'aide d'une loi de proposition symétrique.
  * on accepte la proposition $X' = (\pi', \boldsymbol{\kappa}'_j)$ avec la probabilité $\alpha_{\pi\pi'}$ définie par :

$$\alpha_{01}(\boldsymbol{\kappa}'_j) = \frac{\lambda}{1-\lambda} \cdot \frac{\exp\left(-\frac{1}{2}\|\boldsymbol{\zeta}_j - \mathbf{h}(\boldsymbol{\nu}'_j)\boldsymbol{x}'_j\|^2_{\boldsymbol{\Sigma}^{-1}}\right)}{\exp\left(-\frac{1}{2}\|\boldsymbol{\zeta}_j\|^2_{\boldsymbol{\Sigma}^{-1}}\right)} \cdot p_{10} \qquad \text{(Naissance)}$$

$$\alpha_{10}(\boldsymbol{\kappa}_j) = \frac{1-\lambda}{\lambda} \cdot \frac{\exp\left(-\frac{1}{2}\|\boldsymbol{\zeta}_j\|^2_{\boldsymbol{\Sigma}^{-1}}\right)}{\exp\left(-\frac{1}{2}\|\boldsymbol{\zeta}_j - \mathbf{h}(\boldsymbol{\nu}_j)\boldsymbol{x}_j\|^2_{\boldsymbol{\Sigma}^{-1}}\right)} \cdot \frac{1}{p_{10}} \qquad \text{(Mort)}$$

$$\alpha_{11}(\boldsymbol{\kappa}_j, \boldsymbol{\kappa}'_j) = \frac{\exp\left(-\frac{1}{2}\|\boldsymbol{\zeta}_j - \mathbf{h}(\boldsymbol{\nu}'_j)\boldsymbol{x}'_j\|^2_{\boldsymbol{\Sigma}^{-1}}\right)}{\exp\left(-\frac{1}{2}\|\boldsymbol{\zeta}_j - \mathbf{h}(\boldsymbol{\nu}_j)\boldsymbol{x}_j\|^2_{\boldsymbol{\Sigma}^{-1}}\right)} \frac{\exp\left(-\frac{1}{2\sigma_x^2}\|\boldsymbol{x}'_j\|^2\right)\ p(\boldsymbol{\nu}'_j)}{\exp\left(-\frac{1}{2\sigma_x^2}\|\boldsymbol{x}_j\|^2\right)\ p(\boldsymbol{\nu}_j)} \qquad \text{(Vie)}$$

L'algorithme ainsi que les équations sont adaptés du Chapitre 3 de [Bou14]. Notons que pour les étapes de naissance et de mort, le rapport des lois de proposition se simplifie avec celui des lois *a priori*.

TAB. 3.6 – Algorithme d'échantillonnage de Gibbs avec étapes de Reversible-Jump (RJGS) pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu} \,|\, \boldsymbol{y})$ .

### 3.2.3 Choix du modèle BGE avec échantillonnage Gibbs-hybride

Nous comparons maintenant le modèle BG étendu et et le modèle BG à dimension variable, ainsi que les échantillonneurs associés.

**Un autre échantillonneur du modèle BGE, équivalent au RJGS sur modèle à dimension variable.** Nous revenons au modèle BGE, où les paramètres $(\boldsymbol{\nu}_j, \boldsymbol{x}_j)$ existent indépendamment de la valeur de $q_j$. Nous proposons ici un autre algorithme d'échantillonnage, nommé *Metropolis-Hastings Gibbs sampler* (MHGS), donné en Tableau 3.7 : pour tout $j$, $(q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j)$ sont échantillonnés simultanément avec une étape MH, contrairement au Gibbs-hybride où l'étape MH concerne uniquement les paramètres non linéaires $\boldsymbol{\nu}_j$ puisque $(q_j, \boldsymbol{x}_j)$ sont directement simulés suivant leur loi conditionnelle *a posteriori*.

Itération $(t+1)$. Suivant un état $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{\nu}^{(t)})$ :
- pour $j = 1...J$, on tire $(q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j) \,|\, (\boldsymbol{y}, \boldsymbol{q}_{-j}, \boldsymbol{x}_{-j}, \boldsymbol{\nu}_{-j})$ suivant la méthode suivante :
  * À partir de l'état courant $(q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j)$, on fait la proposition $(q'_j, \boldsymbol{x}'_j, \boldsymbol{\nu}'_j)$ à l'aide de la loi de proposition $Q(q'_j, \boldsymbol{x}'_j, \boldsymbol{\nu}'_j \,|\, q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j) = Q(\boldsymbol{x}'_j, \boldsymbol{\nu}'_j \,|\, q'_j, q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j) Q(q'_j \,|\, q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j)$ :
    - si $q_j = 0$, $Q(q'_j = 1 \,|\, q_j = 0, \boldsymbol{x}_j, \boldsymbol{\nu}_j) = 1$ et $Q(\boldsymbol{x}'_j, \boldsymbol{\nu}'_j \,|\, q'_j, q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j) = p(\boldsymbol{x}'_j, \boldsymbol{\nu}'_j)$.
    - si $q_j = 1$, $Q(q'_j \,|\, q_j = 1, \boldsymbol{x}_j, \boldsymbol{\nu}_j) = p_{11} q'_j + p_{10}(1 - q'_j)$ avec $p_{11} + p_{10} = 1$, et

$$\begin{array}{ll} Q(\boldsymbol{x}'_j, \boldsymbol{\nu}'_j \,|\, q'_j = 1, q_j = 1, \boldsymbol{x}_j, \boldsymbol{\nu}_j) & \text{est une loi de proposition symétrique} \\ Q(\boldsymbol{x}'_j, \boldsymbol{\nu}'_j \,|\, q'_j = 0, q_j = 1, \boldsymbol{x}_j, \boldsymbol{\nu}_j) & = \quad p(\boldsymbol{x}'_j, \boldsymbol{\nu}'_j) \end{array}$$

  * les probabilités d'acceptation sont les mêmes que celles du RJGS (voir Tableau 3.6) suivant « l'évènement » proposé (naissance, mort ou vie).

TAB. 3.7 – Algorithme d'échantillonnage de Gibbs avec étapes de Métropolis-Hastings (MHGS) pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu} \,|\, \boldsymbol{y})$ .

De par le caractère *hiérarchique* du modèle BGE, la loi de proposition fait naturellement apparaître les événements de Naissance ($q_j : 0 \to 1$), Vie ($q_j : 1 \to 1$) et Mort ($q_j : 1 \to 0$), rappelant le RJGS. En prenant pour $(\boldsymbol{\nu}_j, \boldsymbol{x}_j)$ les mêmes lois de proposition que le RJGS lors des Naissances et des Vies, et les lois *a priori* pour les proposer lors des Morts (ici, $(\boldsymbol{\nu}_j, \boldsymbol{x}_j)$ existent si $q_j = 0$), les probabilités d'acceptation du RJGS et du MHGS sont les mêmes : *les deux algorithmes sont donc équivalents.*

**Bilan : le RJGS permet de valider en pratique l'échantillonnage des paramètres non linéaires du modèle BGE, malgré leur non-existence statistique.** Le formalisme « Reversible-Jump » semble donc inutile, puisque que l'algorithme MHGS est équivalent ; les échantillons sont tirés suivant les mêmes lois (sauf les amplitudes $\boldsymbol{x}_j$ : quand $q_j = 0$, elles n'existent pas avec le RJGS mais elles sont nulles avec le MHGS) et donneront donc les mêmes estimations en pratique. Cependant, cette formalisation a le mérite de montrer que l'échantillonnage des paramètres non linéaires du modèle BGE, malgré leur non-existence statistique, est en pratique correct, puisque l'échantillonnage équivalent avec étape de Reversible-Jump est théoriquement juste. On peut étendre cela à l'échantillonneur de Gibbs-hybride et conclure que l'échantillonnage des paramètres non linéaires du modèle BGE ne pose pas de problème théorique.

Nous avons maintenant deux algorithmes pour échantillonner le modèle BGE : (i) le Gibbs-hybride de [Bou05] qui échantillonne *directement* les paramètres $(q_j, \boldsymbol{x}_j)$ suivant leur loi *a posteriori* conditionnelle (à $\boldsymbol{\nu}_j$) puis $\boldsymbol{\nu}_j$ conditionnellement à $(q_j, \boldsymbol{x}_j)$ via une étape MH et (ii) le MHGS présenté dans ce paragraphe qui échantillonne *simultanément* les paramètres $(q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j)$ avec une étape MH. Nous expliquons au paragraphe suivant pourquoi le Gibbs-hybride nous semble préférable.

**Choix du Gibbs-hybride plutôt que le MHGS.** La motivation pour l'algorithme RJGS/MHGS était de pouvoir simuler simultanément les amplitudes $\boldsymbol{x}_j$ et les paramètres non linéaires $\boldsymbol{\nu}_j$ afin de résoudre le problème d'incohérence de l'échantillonnage de l'algorithme Gibbs-hybride. En effet, ce dernier effectue une étape MH conditionnellement aux amplitudes $\boldsymbol{x}_j$ pour l'échantillonnage de $\boldsymbol{\nu}_j$, risquant un fort taux de rejet.

Même si *a priori* l'algorithme MHGS semble résoudre cette incohérence, nous trouvons l'algorithme Gibbs-hybride préférable pour les raisons suivantes :

- L'échantillonnage simultané de $(q_j, \boldsymbol{x}_j, \boldsymbol{\nu}_j)$ doit lui aussi présenter un fort taux de rejet, puisqu'il est fort improbable de proposer simultanément suivant leur loi *a priori* des $\boldsymbol{x}_j$ et $\boldsymbol{\nu}_j$ cohérents avec les données.
- L'algorithme de Gibbs-hybride permet de simuler $(q_j, \boldsymbol{x}_j)$ directement suivant leur loi *a posteriori* conditionnelle (loi BG). Cela a notamment une grande importance pour la séquence de Bernoulli $q_j$, quand celle-ci vaut 1. En effet, l'algorithme de MHGS/RJGS proposera une « Vie » ou une « Mort » avec une certaine probabilité indépendante des données, qui sera acceptée ou non selon la proposition $(\boldsymbol{x}'_j, \boldsymbol{\nu}'_j)$, tandis que le Gibbs-hybride effectue l'événement de « Mort » en lien direct avec les données, et seulement la « Vie » est réalisée ou non suivant la proposition $(\boldsymbol{\nu}'_j)$.

Dans le rapport [Bou14], nous avons d'ailleurs noté que le RJGS n'apportait en pratique pas d'amélioration notable vis à vis du Gibbs-hybride. Nous concluons que l'échantillonnage proposé par l'algorithme Gibbs-hybride est plus satisfaisant.

*Remarque : Dans la suite de ce manuscrit, nous continuerons d'utiliser les appellations de « Naissance » et de « Mort » pour évoquer les transitions du paramètre Bernoulli $q_j$ (resp. $0 \to 1$ et $1 \to 0$), bien que les paramètres $(\boldsymbol{x}_j, \boldsymbol{\nu}_j)$ existent toujours quelle que soit la valeur de $q_j$. De même, on parlera de « Vie » pour qualifier l'échantillonnage des paramètres $\boldsymbol{\nu}_j$ quand $q_j = 1$. Par abus de langage, on parlera de « naissance », « vie » ou « mort » de l'atome $\mathbf{h}_j = \mathbf{h}(\boldsymbol{\nu}_j)$, puisque les évolutions de $q_j$ et $\boldsymbol{\nu}_j$ influencent directement celles de $\mathbf{h}_j$. Enfin, notons que parler de l'étape de « naissance » (et des autres événements) regroupe 2 sous-étapes : le calcul de la probabilité de transition/modification, et sa réalisation le cas échéant.*

## 3.3 Échantillonnage de la loi *a posteriori* marginalisée dans le cadre supervisé

Bien que satisfaisant en théorie, l'échantillonnage de la loi jointe $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ avec l'échantillonneur Gibbs-hybride proposé par [BC06] rencontre les problèmes classiques des échantillonneurs de Gibbs : le blocage dans des modes locaux ralentissant la convergence de l'échantillonneur. Dans un cadre général, il a été montré que la convergence d'un algorithme de Gibbs était accélérée en marginalisant certains paramètres de la loi cible [LWK94]. Par exemple, l'échantillonneur de Gibbs est sensé converger en moins d'itérations si la loi cible est $p(\boldsymbol{u} \,|\, \boldsymbol{y}) = \int_{\boldsymbol{v}} p(\boldsymbol{u}, \boldsymbol{v} \,|\, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{v}$ plutôt que la loi jointe $p(\boldsymbol{u}, \boldsymbol{v} \,|\, \boldsymbol{y})$, d'autant plus si les paramètres $\boldsymbol{u}$ et $\boldsymbol{v}$ sont fortement corrélés.

Dans le cas du modèle BG classique (pour des problèmes d'APL), la marginalisation de la loi jointe $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ suivant les amplitudes $\boldsymbol{x}$ se fait analytiquement, pour obtenir la *loi marginalisée* (ou *loi marginale*) $p(\boldsymbol{q}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$. C'est d'ailleurs cette loi marginale (ou plutôt $\Pr(\boldsymbol{q} \,|\, \boldsymbol{y})$ dans le cadre supervisé) que l'algorithme SMLR cherche à maximiser dans l'article fondateur du modèle BG [KM82]. Cependant, l'échantillonnage selon la loi marginalisée a longtemps été abandonné au profit de celui de la loi jointe, pour deux principales raisons que donne [Bou05] : (i) dans le cadre supervisé, le coût de calcul pour simuler des échantillons $q_j^{(t)}$ suivant la loi marginale conditionnelle $\Pr(q_j \,|\, \boldsymbol{y}, \boldsymbol{q}_{-j})$ est bien plus élevé que la simulation suivant la loi jointe

conditionnelle $\Pr(q_j \,|\, \boldsymbol{q}_{-j}, \boldsymbol{x}_{-j}, \boldsymbol{y})$ et rend l'échantillonnage moins efficace en temps de calcul et (ii) dans le cadre non supervisé, la dépendance non linéaire de la loi *a posteriori* marginalisée en les hyperparamètres de variances $\sigma_x^2$ et $\sigma_\epsilon^2$ ne permet plus un échantillonnage direct de leur loi *a posteriori* conditionnelle.

Ces deux problèmes ont cependant trouvé des solutions proposées par [GILC11] dans le cadre de problèmes de déconvolution impulsionnelle linéaire (modèle BG classique). Ainsi, des astuces calculatoires en partie inspirées de [CGI96] permettent de simuler $q_j$ suivant sa loi marginalisée conditionnelle de façon efficace. De plus, l'utilisation d'un échantillonneur de type *Partially Collapsed Gibbs Sampler* (PCGS), introduit par [vDP08], permet de bénéficier de l'échantillonnage des $q_j$ suivant leur loi marginalisée conditionnelle et d'échantillonner les hyperparamètres suivant leur loi jointe conditionnelle, ce qui permet d'aborder le problème dans un cadre non supervisé. L'auteur montre alors que la convergence de leur échantillonneur est plus rapide que l'algorithme de Gibbs classique, autant en nombre d'itérations qu'en temps de calcul.

Dans le cas du modèle BG étendu aux non-linéarités (pour des problèmes d'APNL), la simulation des paramètres suivant la loi *a posteriori* marginalisée $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ devrait donc accélérer la convergence de l'échantillonneur. Elle permettrait aussi de résoudre le problème de l'échantillonnage incohérent du Gibbs-hybride remarqué par [Bou05] : la simulation de $\boldsymbol{\nu}_j$ conditionnellement aux amplitudes $\boldsymbol{x}_j$ présentant un faible taux d'acceptation.

Nous proposons donc dans cette section d'aborder l'échantillonnage de la loi *a posteriori* marginale dans le *cadre supervisé* : $p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$. Après avoir décrit les lois conditionnelles en $q_j$ et $\boldsymbol{\nu}_j$ nécessaires pour l'algorithme de Gibbs-hybride, nous adaptons les astuces calculatoires proposées par [GILC11] au modèle BGE, et notamment à l'échantillonnage des paramètres non linéaires $\boldsymbol{\nu}_j$ suivant leur loi marginalisée conditionnelle. Un tel échantillonnage n'a pas encore été proposé à notre connaissance, et il s'agit de la contribution majeure apportée par ce chapitre. Une autre nouveauté vis-à-vis des travaux existants (échantillonnage de $q_j$) est l'adaptation des calculs à notre cas plus général d'APNL, dans l'optique d'utiliser cet échantillonneur pour la détection d'exoplanètes au chapitre suivant. Les différences résident dans le bruit non *i.i.d.* (présence d'une matrice de covariance $\boldsymbol{\Sigma}$) et les amplitudes $\boldsymbol{x}_j$ à valeurs vectorielles (dim. $m$) et dont la variance *a priori* $\sigma_x^2$ est inconnue. Ces adaptations nécessitent la démonstration de résultats très similaires à ceux présentés dans [GILC11] ; nous essaierons donc d'insister sur les différences et sur nos apports, en reportant la majeure partie des démonstrations en Annexe B. De plus, nous recommandons une modification permettant d'améliorer la stabilité numérique des matrices en jeu, notamment quand la variance $\sigma_x^2$ est inconnue.

Notons que le *cadre non supervisé* sera abordé dans la section suivante.

*Remarque 1 : Que l'on dispose d'échantillons $(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta})^{(t)}$ suivant la loi jointe ou marginale ne change pas la stratégie de détection/estimation au sens de l'EAP proposée par [BC06] pour le modèle BGE, puisque les estimations se font marginalement aux autres paramètres. Dans ces deux sections, nous aborderons donc uniquement les algorithmes permettant d'obtenir de tels échantillons.*

*Remarque 2 : l'ensemble des lois* a posteriori *conditionnelles pour le modèle BGE en les paramètres $(q_j, \boldsymbol{x}_j)$, $\overline{\boldsymbol{x}}$ et $\boldsymbol{\theta}$ gardent la même expression analytique que pour le modèle BG (voir équations (3.6),(3.5) et (3.7)), mais où la dépendance des matrices $\mathbf{h}$, $\mathbf{H}$ ou $\overline{\mathbf{H}}$ en les paramètres $\boldsymbol{\nu}_j$ est cachée. Pour alléger les notations, nous gardons cette notation « cachée » de l'influence des paramètres non linéaires, notamment en écrivant $\mathbf{h}_j = \mathbf{h}(\boldsymbol{\nu}_j)$ ou $\mathbf{H} = \mathbf{H}(\boldsymbol{\nu})$. Nous spécifions la dépendance en les paramètres $\boldsymbol{\nu}_j$ uniquement quand cela nous semble nécessaire.*

### 3.3.1 Algorithme d'échantillonnage de la loi marginalisée

Dans le cadre supervisé, la loi *a posteriori* marginale est $p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$. L'algorithme de Gibbs que nous associons à cette loi cible est donné en Tableau 3.8, et consiste en l'échantillonnage successif de $q_j$ et $\boldsymbol{\nu}_j$ suivant leur loi *a posteriori* conditionnelle, soit $\Pr(q_j \,|\, \boldsymbol{q}_{-j}, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y})$ et $p(\boldsymbol{\nu}_j \,|\, \boldsymbol{q}, \boldsymbol{\nu}_{-j}, \boldsymbol{\theta}, \boldsymbol{y})$. Nous montrons que la première est une loi de Bernoulli dont on donne l'expression analytique, puis que la deuxième ne s'échantillonne pas directement et nécessite une étape de Métropolis-Hastings, comme pour l'échantillonnage de la loi jointe conditionnelle vu en § 3.2.1. Les calculs sont donnés en Annexes B.2 et B.3.

Itération $(t+1)$, à partir des échantillons courants $(\boldsymbol{q}^{(t)}, \boldsymbol{\nu}^{(t)})$ :
Générer pour tout $j \in [\![1; J]\!]$,

$$\begin{array}{rcll} q_j^{(t+1)} & \sim & (q_j \,|\, \boldsymbol{q}_{-j}, \boldsymbol{\nu}, \boldsymbol{y}) & : \text{ loi de Bernoulli (voir (3.19))} \\ \boldsymbol{\nu}_j^{(t+1)} & \sim & (\boldsymbol{\nu}_j \,|\, \boldsymbol{q}, \boldsymbol{\nu}_{-j}, \boldsymbol{y}) & : \text{ étape Métropolis-Hastings (voir (3.20))} \end{array}$$

TAB. 3.8 – Échantillonneur de la loi *a posteriori* marginalisée $p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$.

Nous commençons par calculer la loi *a posteriori* marginalisée $p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$ en intégrant la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$ par rapport aux amplitudes $\boldsymbol{x}$. Cela se fait aisément puisque l'on sait d'après (3.12) que la loi $p(\overline{\boldsymbol{x}} \,|\, \boldsymbol{y}, -)$ est gaussienne et $p(\overline{\boldsymbol{x}}^C \,|\, \boldsymbol{y}, -)$ une loi Dirac en zéro :

$$\begin{array}{rcl} p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y}, \boldsymbol{\theta}) & \propto & |\mathbf{B}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \mathbf{B}^{-1} \boldsymbol{y}\right) \lambda^L (1-\lambda)^{J-L} p(\boldsymbol{\nu}) p(\lambda) p(\sigma_x^2) p(\boldsymbol{\Sigma}) \\ \text{avec} \quad \mathbf{B} & = & \boldsymbol{\Sigma} + \sigma_x^2 \overline{\mathbf{H}} \overline{\mathbf{H}}^T \quad (\text{dim. } N \times N) \end{array} \tag{3.18}$$

On voit apparaître dans cette expression la matrice $\mathbf{B}$, de dimension (dim. $N \times N$). Cette matrice dépend de presque tous les paramètres du modèle BGE : $\sigma_x^2$ et $\boldsymbol{\Sigma}$, mais aussi $\boldsymbol{q}$ et $\boldsymbol{\nu}$ via le dictionnaire réduit $\overline{\mathbf{H}}$. De plus, le paramètre $L$ s'écrit $L = \|\boldsymbol{q}\|$ et compte le nombre d'atomes présents dans le dictionnaire réduit $\overline{\mathbf{H}}$.

À partir de cette expression, les lois *a posteriori* conditionnelles sont calculées en ne gardant que les dépendances en $q_j$ d'une part et $\boldsymbol{\nu}_j$ d'autre part. Nous définissons ici la fonction $\Psi(\cdot, \cdot; \boldsymbol{y})$ qui apparaîtra dans les calculs qui suivent :

$$\Psi(\mathbf{A}_1, \mathbf{A}_2; \boldsymbol{y}) = \log \frac{|\mathbf{A}_1|}{|\mathbf{A}_2|} + \left(\boldsymbol{y}^T (\mathbf{A}_1 - \mathbf{A}_2) \boldsymbol{y}\right)$$

#### Échantillonnage des paramètres de Bernoulli

À partir de la loi *a posteriori* marginalisée (3.18), on peut calculer la loi *a posteriori* conditionnelle en $q_j$ en ne conservant que les composantes dépendant de $q_j$ et en calculant la constante de normalisation associée. Cette loi est toujours une loi de Bernoulli, et s'exprime de la façon suivante :

$$\begin{array}{rcl} \Pr(q_j \,|\, -) & = & \left(1 + \exp(-(\Psi_{q_j} + \eta_j)/2)\right)^{-1} \quad \text{avec} \quad \eta_j = 2(-1)^{q_j} \log\left(\frac{1-\lambda}{\lambda}\right) \\ & & \text{et} \quad \Psi_{q_j} = \Psi(\mathbf{B}_{(1-q_j)}^{-1}, \mathbf{B}_{(q_j)}^{-1}; \boldsymbol{y}) \end{array} \tag{3.19}$$

où $\mathbf{B}_{(0)}$ (resp. $\mathbf{B}_{(1)}$) correspond à la matrice $\mathbf{B}$ avec $q_j = 0$ (resp. $q_j = 1$). Nous ne notons pas la dépendance en $\boldsymbol{\nu}_j$ puisque ce dernier est fixé lors de l'échantillonnage de $q_j$.

### Échantillonnage des paramètres non linéaires

Dans le cas de l'échantillonnage de la loi jointe du modèle BGE vue en § 3.2.1, nous avons vu qu'il était impossible d'échantillonner directement la loi conditionnelle *a posteriori* en $\boldsymbol{\nu}_j$ du fait des non-linéarités du modèle. C'est d'autant plus vrai pour la loi marginalisée, où la dépendance en $\boldsymbol{\nu}_j$ se retrouve dans la matrice $\mathbf{B}$ via $\overline{\mathbf{H}}$ quand $q_j = 1$ (voir équation (3.18)). Il est donc toujours nécessaire de passer par une étape de Métropolis-Hastings pour échantillonner les paramètres non linéaires $\boldsymbol{\nu}_j$ dans un échantillonneur de Gibbs. On suppose ici que la proposition $\boldsymbol{\nu}'_j$ est faite suivant la loi de mélange symétrique définie en (3.16), si bien que la probabilité d'acceptation $\alpha$ s'écrit comme le rapport des lois cibles où le rapport des lois *a priori* se simplifie[1]. Cette dernière se déduit simplement à partir de l'expression de la loi marginalisée (3.18), en ne faisant apparaître que les termes dépendant de $\boldsymbol{\nu}_j^{(t)}$ et $\boldsymbol{\nu}'_j$ :

$$\begin{aligned} \alpha &= \frac{p(\boldsymbol{\nu}'_j \,|\, -)}{p(\boldsymbol{\nu}_j^{(t)} \,|\, -)} = \frac{|\mathbf{B}_{(q_j,\boldsymbol{\nu}'_j)}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \mathbf{B}^{-1}_{(q_j,\boldsymbol{\nu}'_j)} \boldsymbol{y}\right)}{|\mathbf{B}_{(q_j,\boldsymbol{\nu}_j^{(t)})}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \mathbf{B}^{-1}_{(q_j,\boldsymbol{\nu}_j^{(t)})} \boldsymbol{y}\right)} \\ \text{si } q_j = 1, \quad \alpha &= \exp(-\Psi_{\boldsymbol{\nu}_j}/2) \text{ avec } \Psi_{\boldsymbol{\nu}_j} = \Psi(\mathbf{B}^{-1}_{(1,\boldsymbol{\nu}'_j)}, \mathbf{B}^{-1}_{(1,\boldsymbol{\nu}_j^{(t)})}; \boldsymbol{y}) \\ \text{si } q_j = 0, \quad \alpha &= 1 \end{aligned} \tag{3.20}$$

où l'écriture $\mathbf{B}_{(q_j,\boldsymbol{\nu}_j)}$ traduit la dépendance de la matrice $\mathbf{B}$ en $q_j$ et $\boldsymbol{\nu}_j$. Dans le cas où $q_j = 0$, la matrice $\mathbf{B}_{(0,\boldsymbol{\nu}_j)}$ ne dépend pas de $\boldsymbol{\nu}_j$ de par son expression en équation (3.18) : $\mathbf{B}_{(0,\boldsymbol{\nu}_j)} = \mathbf{B}_{(0)}$. Tout comme dans l'échantillonneur de la loi jointe, la proposition est alors toujours acceptée.

### Implémentation inefficace de l'échantillonneur

L'algorithme de Gibbs-hybride (Tab. 3.8) peut être implémenté en suivant les expressions de la loi *a posteriori* en $q_j$ conditionnelle $\Pr(q_j \,|\, -)$ (3.19) et de la probabilité d'acceptation $\alpha$ de l'échantillonnage MH de $\boldsymbol{\nu}_j$ (3.20). Ces deux expressions sont très similaires : elles font intervenir toutes les deux la fonction $\Psi(\mathbf{B}'^{-1}, \mathbf{B}^{-1}; \boldsymbol{y})$, où $\mathbf{B}'$ et $\mathbf{B}$ diffèrent uniquement selon la valeur du paramètre à échantillonner ($q_j$ ou $\boldsymbol{\nu}_j$). Dans les deux cas, le calcul de cette fonction représente la majorité du coût de calcul nécessaire à l'échantillonnage. L'implémentation d'une étape de l'échantillonneur se fait récursivement de la façon suivante, à partir de la matrice $\mathbf{B}^{-1}$ courante :

| | | | |
|---|---|---|---|
| Matrice courante $\mathbf{B}^{-1}$ $\rightarrow$ | Calcul de $\mathbf{B}'^{-1}$ $\rightarrow$ | Calcul de $\Psi$ $\rightarrow$ | Calcul des probabilités |
| ($\mathbf{B}^{-1}_{(q_j)}$ ou $\mathbf{B}^{-1}_{(1,\boldsymbol{\nu}_j^{(t)})}$) | ($\mathbf{B}^{-1}_{(1-q_j)}$ ou $\mathbf{B}^{-1}_{(1,\boldsymbol{\nu}'_j)}$) | ($\Psi_{q_j}$ ou $\Psi_{\boldsymbol{\nu}_j}$) | ($\Pr(q_j \,|\, \boldsymbol{y})$ ou $\alpha$) |
| | | | $\downarrow$ |
| | Si modification ($q_j \leftarrow 1 - q_j$ ou $\boldsymbol{\nu}_j \leftarrow \boldsymbol{\nu}'_j$) $\rightarrow$ | | Échantillonnage |
| | Mise à jour : $\mathbf{B}^{-1} \leftarrow \mathbf{B}'^{-1}$ | | |

L'implémentation de cet algorithme nécessite donc pour chaque étape le calcul d'une matrice $\mathbf{B}'^{-1}$ de dimension $N \times N$ et de son déterminant pour le calcul de $\Psi$. De telles opérations sont très coûteuses, si bien que même si l'algorithme converge en moins d'itérations, le temps de calcul associé sera bien plus élevé que pour l'échantillonneur de la loi jointe ; cela a longtemps dissuadé l'échantillonnage de la loi marginalisée au profit de celui de la loi jointe.

[1] Nous avons en effet vu en § 3.2.1 que la proposition était répétée jusqu'à ce que $\boldsymbol{\nu}'_j \in \mathfrak{I}_j$ ; le rapport des lois *a priori*, uniformes sur le même intervalle, est donc égal à 1.

Dans la suite de cette section, nous proposons donc d'adapter à notre échantillonneur du modèle BGE les astuces calculatoires proposées par [GILC11] dans le cadre du modèle BG classique rendant l'échantillonneur de la loi marginale plus efficace.

D'une part, nous adaptons ces calculs pour l'échantillonnage des variables de Bernoulli $q_j$ aux spécificités de notre problème, *i.e.* la présence d'un bruit non-*i.i.d.* (matrice de covariance $\mathbf{\Sigma}$), et d'amplitudes $\boldsymbol{x}_j$ multidimensionnelles dont la variance $\sigma_x^2$ est inconnue. Les calculs suivent donc la même logique mais présentent quelques différences. Ils exploitent des matrices de dimension plus faibles que $\mathbf{B}$ pour le calcul de la loi conditionnelle en (3.19), notamment un facteur de Cholesky $\mathbf{F}$(dim. $mL \times mL$), qui nécessite d'être mis à jour après chaque modification des $q_j$.

D'autre part, nous exploitons ces astuces calculatoires pour l'échantillonnage des paramètres non linéaires $\boldsymbol{\nu}_j$, notamment pour un calcul efficace de la probabilité d'acceptation (3.20) et pour la mise à jour nécessaire du facteur de Cholesky $\mathbf{F}$.

### 3.3.2 Astuces calculatoires pour l'échantillonnage de $q_j$

L'implémentation efficace de l'échantillonnage de $q_j$ via sa loi conditionnelle $\Pr(q_j \,|\, -)$ (3.19) repose les résultats suivants, issus de [GILC11] :

1. La fonction $\Psi_{q_j} = \Psi(\mathbf{B}_{(1-q_j)}^{-1}, \mathbf{B}_{(q_j)}^{-1}; \boldsymbol{y})$ peut se réécrire en fonction de matrices de faible dimension par rapport à $\mathbf{B}$ : $\mathbf{T}_{(q_j)}$(dim. $m \times m$) et $\boldsymbol{\gamma}_{(q_j)}$(dim. $m \times 1$). Ces matrices s'expriment en fonction de la matrice $\mathbf{B}^{-1}$ courante, soit $\mathbf{B}_{(q_j)}^{-1}$ ; il n'est donc plus nécessaire de calculer la matrice $\mathbf{B}_{(1-q_j)}^{-1}$ (ni son déterminant) à chaque itération, mais uniquement quand une Naissance ou une Mort sera réalisée pour mettre à jour la matrice $\mathbf{B}^{-1}$ courante.
2. La matrice $\mathbf{B}^{-1}$, et par transition $\mathbf{T}$, $\boldsymbol{\gamma}$, et $\Psi_{q_j}$, peuvent s'écrire en fonction d'un facteur de Cholesky $\mathbf{F}$(dim. $mL \times mL$), rendant le calcul de la matrice $\mathbf{B}^{-1}$ inutile : à partir de la matrice $\mathbf{F}$ courante $\mathbf{F}_{(q_j)}$, on peut calculer $\mathbf{T}_{(q_j)}$, $\boldsymbol{\gamma}_{(q_j)}$, et $\Psi_{q_j}$, et mettre à jour le facteur de Cholesky $\mathbf{F}$ si une Naissance ou une Mort est réalisée.

Le schéma de l'implémentation est donc le suivant :

| | | | | |
|---|---|---|---|---|
| Matrice courante $\mathbf{F} = \mathbf{F}_{(q_j)}$ | $\rightarrow$ | Calcul de $\mathbf{T}_{(q_j)}, \boldsymbol{\gamma}_{(q_j)}$ en fonction de $\mathbf{F}_{(q_j)}$ | $\rightarrow$ Calcul de $\Psi_{q_j}$ en fonction de $\mathbf{T}_{(q_j)}, \boldsymbol{\gamma}_{(q_j)}$ | $\rightarrow$ Calcul de $\Pr(q_j \,|\, \boldsymbol{y})$ $\downarrow$ Échantillonnage |
| | | Si modification (Naissance ou Mort) Mise à jour du facteur de Cholesky $\mathbf{F}$ | $\leftarrow$ | |

Dans les paragraphes suivants, nous rentrons plus en détails sur ces simplifications, pour aboutir à un pseudo-code de l'implémentation résultante. *Notons que nous proposons une modification dans le schéma d'implémentation donné ci-dessous, permettant d'améliorer la stabilité numérique des matrices* $\mathbf{T}$ *utilisées.* Comme lors de l'échantillonnage de $q_j$, les paramètres $\boldsymbol{\nu}_j$ sont fixés, la dépendance des atomes continus en $\boldsymbol{\nu}_j$ est cachée afin d'alléger les notations : $\mathbf{h}_j = \mathbf{h}(\boldsymbol{\nu}_j)$ . L'ensemble de ces résultats sont démontrés en Annexe B.4, B.5 et B.6.

#### Introduction des matrices $\mathbf{T}_{(q_j)}$ et $\boldsymbol{\gamma}_{(q_j)}$ pour la réécriture de $\Psi_{q_j}$

La fonction $\Psi_{q_j} = \Psi(\mathbf{B}_{(1-q_j)}^{-1}, \mathbf{B}_{(q_j)}^{-1}; \boldsymbol{y})$ peut se simplifier car les matrices $\mathbf{B}_{(q_j)}$ et $\mathbf{B}_{(1-q_j)}$ sont très similaires. On peut en effet montrer que :

$$\mathbf{B}_{(1-q_j)} - \mathbf{B}_{(q_j)} = (-1)^{q_j} \sigma_x^2 \, \mathbf{h}_j \mathbf{h}_j^T$$

À l'aide des lemmes d'inversion et de déterminants matriciels (voir Annexe B.1), on peut simplifier l'expression de $\Psi_{q_j}$ (et donc de la loi $\Pr(q_j \,|\, -)$) en faisant ressortir la matrice $\mathbf{T}_{(q_j)}$ (dim. $m \times m$) et le vecteur $\boldsymbol{\gamma}_{(q_j)}$ (dim. $m \times 1$), dont la dimension $m$ est faible (typiquement $m = 1$ ou 2 dans nos applications) comparée à la dimension $N$ de la matrice $\mathbf{B}$, le nombre de données. Nous obtenons les simplifications suivantes :

$$\begin{aligned} \boldsymbol{y}^T(\mathbf{B}_{(1-q_j)}^{-1} - \mathbf{B}_{(q_j)}^{-1})\boldsymbol{y} &= -\boldsymbol{\gamma}_{(q_j)}^T \mathbf{T}_{(q_j)}^{-1} \boldsymbol{\gamma}_{(q_j)} \\ \frac{|\mathbf{B}_{(1-q_j)}|}{|\mathbf{B}_{(q_j)}|} &= (-1)^{mq_j} \sigma_x^{2m} |\mathbf{T}_{(q_j)}| \end{aligned} \tag{3.21}$$

avec les définitions suivantes pour $\mathbf{T}_{(q_j)}$ et $\boldsymbol{\gamma}_{(q_j)}$ :

$$\begin{aligned} \mathbf{T}_{(q_j)} &= (-1)^{q_j} \sigma_x^{-2} \mathbf{I}_m + \mathbf{h}_j^T \mathbf{B}_{(q_j)}^{-1} \mathbf{h}_j \quad (\text{dim. } m \times m) \\ \boldsymbol{\gamma}_{(q_j)} &= \mathbf{h}_j^T \mathbf{B}_{(q_j)}^{-1} \boldsymbol{y} \quad (\text{dim. } m \times 1) \end{aligned} \tag{3.22}$$

Par conséquent, la fonction $\Psi_{q_j}$ peut s'écrire en fonction des matrices $\mathbf{T}_{(q_j)}$ et vecteurs $\boldsymbol{\gamma}_{(q_j)}$ :

$$\Psi_{q_j} = \Psi(\mathbf{B}_{(1-q_j)}^{-1}, \mathbf{B}_{(q_j)}^{-1}; \boldsymbol{y}) = \log\left((-1)^{mq_j} \sigma_x^{2m} |\mathbf{T}_{(q_j)}|\right) - \boldsymbol{\gamma}_{(q_j)}^T \mathbf{T}_{(q_j)}^{-1} \boldsymbol{\gamma}_{(q_j)} \tag{3.23}$$

Ainsi, pour échantillonner le paramètre $q_j$, il suffira de calculer les matrices $\mathbf{T}_{(q_j)}$ et $\boldsymbol{\gamma}_{(q_j)}$ en fonction de la matrice $\mathbf{B}_{(q_j)}^{-1}$ courante ; la matrice $\mathbf{B}_{(1-q_j)}^{-1}$ n'aura besoin d'être calculée que dans le cas où $\mathbf{B}^{-1}$ doit être mise à jour : passage de $q_j$ à $1 - q_j$ (soit une Naissance ou une Mort).

### Introduction du facteur de Cholesky F pour le calcul de T et $\gamma$

En posant $\mathbf{G} = \sigma_x \overline{\mathbf{H}}$, matrice de taille $(N, mL)$, on peut écrire :

$$\begin{aligned} \mathbf{B} &= \boldsymbol{\Sigma} + \mathbf{G}\mathbf{G}^T \\ \mathbf{B}^{-1} &= \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{G}\mathbf{C}^{-1}\mathbf{G}^T\boldsymbol{\Sigma}^{-1} \\ \text{où} \quad \mathbf{C} &= \mathbf{I}_{D^*} + \mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G} \end{aligned}$$

La matrice $\mathbf{C}$ est de taille $mL$, ce qui est bien inférieur à $N$ en pratique, puisque qu'on cherche à obtenir un faible $L = \|\boldsymbol{q}\|$ en accord avec l'*a priori* de parcimonie. La matrice $\mathbf{C}^{-1}$ étant définie-positive, on peut utiliser sa factorisation de Cholesky $\mathbf{C}^{-1} = \mathbf{F}^T\mathbf{F}$, où son facteur de Cholesky $\mathbf{F}$ est une matrice triangulaire supérieure de taille $mL$. Les évaluations de $\mathbf{T}_{(q_j)}$ et $\boldsymbol{\gamma}_{(q_j)}$ reviennent alors à :

$$\begin{aligned} \mathbf{T}_{(q_j)} &= (-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j - (\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j)^T(\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j) \\ \boldsymbol{\gamma}_{(q_j)} &= \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - (\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j)^T(\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}) \end{aligned} \tag{3.24}$$

Le calcul des matrices $\mathbf{T}_{(q_j)}$ et $\boldsymbol{\gamma}_{(q_j)}$ se fait donc à partir des matrices courantes $\mathbf{F}_{(q_j)}$ et $\mathbf{G}_{(q_j)}$, et ne fait plus intervenir les matrices $\mathbf{B}$. Si la valeur de $q_j$ change durant l'échantillonnage de Gibbs, il est nécessaire de mettre à jour les matrices $\mathbf{G}$ et $\mathbf{F}$ pour prendre en compte ce changement d'état. L'intérêt de l'introduction du facteur de Cholesky $\mathbf{F}$ est double : il permet d'une part de simplifier le calcul de $\Psi_{q_j}$ et donc de $\Pr(q_j \,|\, \boldsymbol{y})$, et d'autre part sa mise à jour peut se faire en un temps réduit, grâce aux méthodes de *mise à jour de Cholesky de rang 1*. Ces dernières permettent de réaliser, en $\mathcal{O}(a^2)$, la mise à jour d'un facteur de Cholesky $\mathbf{A}(\text{dim. } a \times a)$ vers $\mathbf{A}'$ tel que $\mathbf{A}'^T\mathbf{A}' = \mathbf{A}^T\mathbf{A} \pm \boldsymbol{d}\boldsymbol{d}^T$, où $\boldsymbol{d}$ est un vecteur colonne. Nous généralisons cette mise à jour dans le cas où $\boldsymbol{d} = \mathbf{d}$, une matrice de $R$ colonnes : nous notons $\mathbf{A}' = \text{cholupdate}(\mathbf{A}, \mathbf{d}, \pm)$

la mise à jour de Cholesky réalisant $R$ mises à jour de rang 1 successives[2].

Pour la mise à jour des matrices $\mathbf{F}$ et $\mathbf{G}$, on différencie la naissance $q_j : 0 \to 1$ (les matrices courantes sont $\mathbf{F}_{(0)}$ et $\mathbf{G}_{(0)}$ et on veut les mettre à jour vers $\mathbf{F}_{(1)}$ et $\mathbf{G}_{(1)}$) et la mort $q_j : 1 \to 0$ (les matrices courantes sont $\mathbf{F}_{(1)}$ et $\mathbf{G}_{(1)}$ et on veut les mettre à jour vers $\mathbf{F}_{(0)}$ et $\mathbf{G}_{(0)}$).

### Ordre et mise à jour de G

La matrice $\mathbf{G}$ s'écrivant $\sigma_x\overline{\mathbf{H}}$, une naissance ou une mort de l'atome $\mathbf{h}_j$ consiste à insérer ou retirer $\sigma_x\mathbf{h}_j$ de la matrice $\mathbf{G}$. Jusqu'à présent, nous n'avons pas précisé d'ordre de rangement des atomes $\{\sigma_x\mathbf{h}_i\}_i$ composant $\mathbf{G}$ ; cet ordre n'a en effet aucune influence sur la valeur des matrices $\mathbf{B}$ et $\mathbf{T}$. Cependant, il a une influence directe sur la matrice $\mathbf{C}$ et donc sur les facteurs de Cholesky $\mathbf{F}$. Pour une naissance ou une mort à l'indice $j$, on définit la matrice $\widetilde{\mathbf{G}}_j$ comme la matrice $\mathbf{G}_{(1)}$ où l'atome $\sigma_x\mathbf{h}_j$ se situe dans les dernières colonnes :

$$\widetilde{\mathbf{G}}_j = \begin{bmatrix}\mathbf{G}_{(0)} & \sigma_x\mathbf{h}_j\end{bmatrix} \tag{3.25}$$

Pour la mise à jour de la matrice $\mathbf{G}$, nous choisissons la convention suivante pour la naissance :

*Lors de la naissance de l'atome $\mathbf{h}_j$, on insère $\sigma_x\mathbf{h}_j$ en fin de matrice $\mathbf{G}_{(0)}$ :*
$$\mathbf{G}_{(1)} = \widetilde{\mathbf{G}}_j = \begin{bmatrix}\mathbf{G}_{(0)} & \sigma_x\mathbf{h}_j\end{bmatrix}$$

On définit le vecteur *d'ordre* $\boldsymbol{o} = [o_1, \ldots, o_J]$, tel que $o_j$ indique en quelle position se trouve $\sigma_x\mathbf{h}_j$ dans $\mathbf{G}$ ; si $q_j = 0$ alors $o_j = 0$. Par exemple, si $\mathbf{G} = \sigma_x[\mathbf{h}_3, \mathbf{h}_2]$ avec $J = 4$, alors $\boldsymbol{o} = [0\ 2\ 1\ 0]$ ; après naissance de l'atome $\mathbf{h}_1$, il devient $\boldsymbol{o} = [3\ 2\ 1\ 0]$ ($o_1 \leftarrow L = 3$). Ce vecteur d'ordre permet donc de localiser la contribution $\sigma_x\mathbf{h}_j$ dans la matrice courante $\mathbf{G}$. Ainsi, lors de la mort de l'atome $\sigma_x\mathbf{h}_j$, il suffit de le localiser et de le retirer de la matrice $\mathbf{G}_{(1)}$ :

*Lors de la mort de l'atome $\mathbf{h}_j$, on retire $\sigma_x\mathbf{h}_j$ de la matrice $\mathbf{G}_{(1)}$ grâce à la position[3] $o_j$ :*
$$\mathbf{G}_{(1)} = [\mathbf{G}_a \quad \underset{\substack{\uparrow\\ \text{position } o_j}}{\sigma_x\mathbf{h}_j} \quad \mathbf{G}_b] \quad \Rightarrow \quad \mathbf{G}_{(0)} = [\mathbf{G}_a\ \mathbf{G}_b]$$

Après la mort de l'atome $\mathbf{h}_j$, il est nécessaire de mettre à jour le vecteur d'ordre ; $o_j$ devient nul, et tous les atomes qui étaient placés après lui dans $\overline{\mathbf{H}}$ gagnent une place (voir le pseudo-code 1).

### Relations de base pour les mises à jour de Cholesky

On note $\widetilde{\mathbf{C}}_j$ et $\widetilde{\mathbf{F}}_j$ les matrices $\mathbf{C}_{(1)}$ et $\mathbf{F}_{(1)}$ quand $\mathbf{G}_{(1)} = \widetilde{\mathbf{G}}_j$. On démontre alors la relation suivante, en posant $\widetilde{\mathbf{F}}_j = \begin{bmatrix}\tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{12} \\ \mathbf{0} & \tilde{\mathbf{f}}_{22}\end{bmatrix}$, où $\tilde{\mathbf{f}}_{11}$ et $\tilde{\mathbf{f}}_{22}$ sont des matrices triangulaires supérieures :

$$\begin{aligned}\widetilde{\mathbf{F}}_j^T\widetilde{\mathbf{F}}_j &= \begin{bmatrix}\mathbf{F}_{(0)}^T\mathbf{F}_{(0)} + \mathbf{b}_0\mathbf{T}_{(0)}\mathbf{b}_0^T & \sigma_x^{-1}\mathbf{b}_0 \\ \sigma_x^{-1}\mathbf{b}_0^T & \sigma_x^{-2}\mathbf{T}_{(0)}^{-1}\end{bmatrix} = \begin{bmatrix}\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12} \\ (\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12})^T & \tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22}\end{bmatrix} \\ \text{avec}\quad \mathbf{b}_0 &= -(\mathbf{F}_{(0)}^T\mathbf{F}_{(0)})^{-1}\mathbf{G}_{(0)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j\mathbf{T}_{(0)}^{-1}\end{aligned} \tag{3.26}$$

[2] Si $\mathbf{A}'^T\mathbf{A}' = \mathbf{A}^T\mathbf{A} \pm \mathbf{d}\mathbf{d}^T = \mathbf{A}^T\mathbf{A} \pm \sum_{r=1}^R \boldsymbol{d}_r\boldsymbol{d}_r^T$, alors $\mathbf{A}'$ s'obtient en effectuant $R$ mises à jour de Cholesky de rang 1 successives : $\forall r \in [\![1; R]\!]$ $\mathbf{A} = \text{cholupdate}(\mathbf{A}, \boldsymbol{d}_r, \pm)$, puis $\mathbf{A}' = \mathbf{A}$ (on note $\boldsymbol{d}_r$ la $r^{\text{e}}$ colonne de $\mathbf{d}$).

[3] En pratique, l'atome $\sigma_x\mathbf{h}_j$ étant composé de $m$ colonnes, sa position au sein de $\mathbf{G}_{(1)}$ s'exprime en fonction des indices de $(m(o_j - 1) + 1)$ à $mo_j$.

### Mise à jour de F lors de la naissance de $\mathbf{h}_j$

Dans le cas d'une naissance, nous avons vu que $\mathbf{G}_{(1)} = \widetilde{\mathbf{G}}_j$. On peut alors directement utiliser la relation (3.26) qui se factorise pour donner la formule de mise à jour suivante :

$$\begin{array}{rcl} \mathbf{F}_{(1)} = \widetilde{\mathbf{F}}_j & = & \text{cholupdate}(\overline{\mathbf{F}}_{(0)}, \mathbf{v}, +) \quad \Leftrightarrow \quad \mathbf{F}_{(1)}^T\mathbf{F}_{(1)} = \overline{\mathbf{F}}_{(0)}^T\overline{\mathbf{F}}_{(0)} + \mathbf{v}\mathbf{v}^T \\ \text{avec} \quad \overline{\mathbf{F}}_{(0)} & = & \begin{bmatrix} \mathbf{F}_{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad\qquad \text{et} \quad \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_0\mathbf{U} \\ \sigma_x^{-1}\mathbf{U}^{-1} \end{bmatrix} \end{array} \tag{3.27}$$

où $\mathbf{U}$ est une racine carré symétrique de $\mathbf{T}_{(0)}$(dim. $m \times m$).

La matrice $\mathbf{F}_{(1)}$ peut donc être calculée à partir de $\mathbf{F}_{(0)}$ et $\mathbf{T}_{(0)}$, elle-même calculée à partir de $\mathbf{F}_{(0)}$ et $\mathbf{G}_{(0)}$. Contrairement à [GILC11] où $\mathbf{T}$ est un scalaire, nous avons étendu les calculs au cas où $\mathbf{T}$ est une matrice de dimension $m$. L'existence d'une racine carré $\mathbf{U}$ (dim. $m \times m$) de $\mathbf{T}_{(0)}$ est justifiée car $\mathbf{T}_{(0)} = \sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\mathbf{B}_{(0)}^{-1}\mathbf{h}_j$ est définie positive. La racine $\mathbf{U}$ choisie devra être symétrique[4] afin que le résultat de l'équation (3.27) soit juste : $\mathbf{U} = \mathbf{U}^T$.

### Mise à jour de F lors de la mort de $\mathbf{h}_j$

*Cas 1 :* $\mathbf{G}_{(1)} = \widetilde{\mathbf{G}}_j$ *et* $\mathbf{F}_{(1)} = \widetilde{\mathbf{F}}_j$. La mise à jour de $\mathbf{F}$ se déduit facilement de la relation (3.26). En effet, on a $\mathbf{F}_{(0)}^T\mathbf{F}_{(0)} = \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{11} - \mathbf{b}_0\mathbf{T}_{(0)}\mathbf{b}_0^T$, dont le dernier terme se factorise aisément à l'aide des blocs $\tilde{\mathbf{f}}_{11}$, $\tilde{\mathbf{f}}_{12}$ et $\tilde{\mathbf{f}}_{22}$. On obtient :

$$\begin{array}{rcl} \mathbf{F}_{(0)} & = & \text{cholupdate}(\tilde{\mathbf{f}}_{11}, \mathbf{v}_1, -) \Leftrightarrow \mathbf{F}_{(0)}^T\mathbf{F}_{(0)} = \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{11} - \mathbf{v}_1\mathbf{v}_1^T \\ \text{avec } \mathbf{v}_1 & = & \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12}(\tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22})^{-1/2} (= \mathbf{b}_0\mathbf{U}) \end{array} \tag{3.28}$$

*Cas 2 :* $\mathbf{G}_{(1)} \neq \widetilde{\mathbf{G}}_j$ *et* $\mathbf{F}_{(1)} \neq \widetilde{\mathbf{F}}_j$. Le « désordre » apporté par la matrice $\mathbf{G}_{(1)}$ induit uniquement un désordre similaire pour la matrice $\mathbf{F}_{(1)}$. En écrivant $\widetilde{\mathbf{G}}_j = \mathbf{G}_{(1)}\mathbf{P}$, où $\mathbf{P}$ est une matrice de permutation[5] définie à partie de $o_j$ telle que : $\mathbf{G}_{(1)}\mathbf{P} = [\mathbf{G}_a \ \sigma_x\mathbf{h}_j \ \mathbf{G}_b]\mathbf{P} = [\mathbf{G}_a \ \mathbf{G}_b \ \sigma_x\mathbf{h}_j] = \widetilde{\mathbf{G}}_j$, on peut montrer la relation suivante entre $\widetilde{\mathbf{F}}_j$ et $\mathbf{F}_{(1)}$ :

$$\widetilde{\mathbf{F}}_j^T\widetilde{\mathbf{F}}_j = (\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})$$

Cependant,$\widetilde{\mathbf{F}}_j \neq \mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P}$ car $\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P}$ n'est pas triangulaire. Cette relation nous permet néanmoins de retrouver, à partir de $\mathbf{F}_{(1)}$, les termes nécessaires pour calculer $\mathbf{F}_{(0)}$ via l'équation (3.28), en notant la matrice $\mathbf{F}_{(1)}$ en blocs de la façon suivante :

$$\begin{array}{rcl} \mathbf{F}_{(1)} & = & \begin{bmatrix} \mathbf{f}_{aa} & \mathbf{f}_{am} & \mathbf{f}_{ab} \\ \mathbf{0} & \mathbf{f}_{mm} & \mathbf{f}_{mb} \\ \mathbf{0} & \mathbf{0} & \mathbf{f}_{bb} \end{bmatrix} \quad \Rightarrow \quad \tilde{\mathbf{f}}_{11} = \begin{bmatrix} \mathbf{f}_{aa} & \mathbf{f}_{ab} \\ \mathbf{0} & \mathbf{f}_* \end{bmatrix}, \text{ où } \mathbf{f}_* = \text{cholupdate}(\mathbf{f}_{bb}, \mathbf{f}_{mb}^T, +) \\ & = & \begin{bmatrix} \mathbf{f}_a & \mathbf{f}_m & \mathbf{f}_b \end{bmatrix} \quad \Rightarrow \quad \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12} = \begin{bmatrix} \mathbf{f}_a & \mathbf{f}_b \end{bmatrix}^T \mathbf{f}_m, \text{ et } \tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22} = \mathbf{f}_m^T\mathbf{f}_m \end{array} \tag{3.29}$$

[4] La factorisation de Cholesky de $\mathbf{T}$ donne une matrice triangulaire non symétrique, et ne pourra donc pas être utilisée pour calculer $\mathbf{U}$.

[5] La multiplication à droite par $\mathbf{P}$ permute les colonnes ; la multiplication à gauche par $\mathbf{P}^T$ permute les lignes.

### Une nouvelle stratégie pour les étapes de mort réduisant les instabilités numériques

Nous avons remarqué lors de nos tests que la stratégie d'échantillonnage de la mort d'un atome $\mathbf{h}_j$ définie précédemment pouvait apporter des problèmes de stabilité numérique.

Actuellement, la stratégie pour décider de la mort d'un atome $\mathbf{h}_j$ est de calculer $\Pr(q_j = 1 \,|\, -)$ (3.19) à partir des matrices $(\mathbf{T}_{(1)}, \boldsymbol{\gamma}_{(1)})$ (3.23), elles-mêmes calculées grâce aux matrices $(\mathbf{F},\mathbf{G})$ courante, soit $(\mathbf{F}_{(1)},\mathbf{G}_{(1)})$ (3.24). Cependant, nous avons mis en exergue des cas où la matrice $\mathbf{T}_{(1)}$ se rapproche de la matrice nulle, provoquant des grandes instabilités numériques : dans ces exemples, les $m = 2$ valeurs propres de $\mathbf{T}_{(1)}$ sont très faibles et de signes différents (bien qu'en théorie la matrice $\mathbf{T}$ est définie-positive de par sa définition), menant à un déterminant $|\mathbf{T}_{(1)}|$ négatif alors que nous devons en prendre le logarithme dans le calcul de $\Pr(q_j = 1 \,|\, -)$ via l'équation (3.23). Nous avons montré empiriquement que ce mauvais comportement numérique arrivait systématiquement quand la variance *a priori* des amplitudes $\sigma_x^2$ est relativement élevée :

$$\mathbf{T}_{(1)} \underset{\sigma_x^2 \to \infty}{\to} \mathbf{0}$$

Ce problème ne se posait pas dans le cas de [GILC11], où le problème de déconvolution aveugle imposait de fixer la variance *a priori* des amplitudes à 1, pour éviter les problèmes d'ambiguïté d'échelle. Dans notre cas, la variance des amplitudes $\sigma_x^2$ est inconnue, et peut soit être fixée à une valeur arbitraire, soit être échantillonnée dans un cadre non supervisé (voir en § 3.4). L'éventualité d'un $\sigma_x^2$ devenant trop élevée doit donc être considérée.

Nous proposons donc une nouvelle stratégie : comme $\Pr(q_j = 1 \,|\, -) = 1 - \Pr(q_j = 0 \,|\, -)$, nous calculons $\Pr(q_j = 0 \,|\, -)$ à partir de l'équation (3.23), comme dans le cas d'une naissance ; cette dernière fait intervenir les matrices $(\mathbf{T}_{(0)}, \boldsymbol{\gamma}_{(0)})$. Contrairement au cas d'une naissance où les matrices courantes sont $(\mathbf{F}_{(0)}, \mathbf{G}_{(0)})$, nécessaires pour le calcul de $(\mathbf{T}_{(0)}, \boldsymbol{\gamma}_{(0)})$ via (3.24), nous disposons seulement de $(\mathbf{F}_{(1)}, \mathbf{G}_{(1)})$ dans le cas d'une mort. Néanmoins, $\mathbf{G}_{(0)}$ se calcule à moindre coût à partir de $\mathbf{G}_{(1)}$, et grâce aux relations mises en exergue en équations (3.26) et (3.29), les matrices $(\mathbf{T}_{(0)}, \boldsymbol{\gamma}_{(0)})$ peuvent directement être récupérées à partir de $\mathbf{F}_{(1)}$ :

$$\begin{array}{rcccll} \mathbf{T}_{(0)}^{-1} & = & \sigma_x^2(\tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22}) & = & \sigma_x^2\mathbf{f}_m^T\mathbf{f}_m & \\ \mathbf{b}_0 & = & \sigma_x\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12} & = & \sigma_x \begin{bmatrix} \mathbf{f}_a & \mathbf{f}_b \end{bmatrix}^T \mathbf{f}_m & \text{(d'après (3.26),(3.29))} \\ \boldsymbol{\gamma}_{(0)} & = & (\mathbf{b}_0\mathbf{T}_{(0)})^T\mathbf{G}_{(0)}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} + \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} & & & \text{(d'après (3.24),(3.26))} \end{array} \tag{3.30}$$

Cette nouvelle stratégie évite donc les problèmes d'instabilités numériques apportés par $\mathbf{T}_{(1)}$ ; en effet, de tels problèmes n'ont jamais été remarqués pour la matrice $\mathbf{T}_{(0)}$. Notons par ailleurs que le coût de calcul de notre stratégie est similaire voir légèrement plus faible que celui de la stratégie initiale proposée par [GILC11].

### Bilan : pseudo-code pour l'échantillonnage de $q_j$

L'échantillonnage de la loi *a posteriori* marginalisée conditionnelle en $q_j$ peut donc s'effectuer à moindre coût, grâce à l'introduction des facteurs de Cholesky et à leur mise à jour efficace. Nous récapitulons l'implémentation de cet échantillonneur dans le pseudo-code 1, avec notre nouvelle stratégie pour l'événement de mort. Il faut maintenant adapter ces astuces calculatoires pour l'échantillonnage des paramètres non linéaires $\boldsymbol{\nu}_j$.

**Pseudo-code 1** Échantillonnage de $q_j$ selon $\Pr(q_j \,|\, -)$.

**Entrée:** $\boldsymbol{y}$, $q_j$, $\mathbf{F}$, $\mathbf{G}$, $\boldsymbol{\Sigma}^{-1}$, $\sigma_x^2$, $\mathbf{h}_j$, $\boldsymbol{o}$
  **si** $q_j = 0$ **alors**
    % Calcul de $\Pr(q_j = 0 \,|\, -)$
    $\boldsymbol{\alpha} \leftarrow \mathbf{F}\mathbf{G}^T$, $\boldsymbol{\beta} \leftarrow \boldsymbol{\alpha}\boldsymbol{\Sigma}^{-1}\mathbf{h}_j$
    $\mathbf{T}_{(0)} \leftarrow \sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j - \boldsymbol{\beta}^T\boldsymbol{\beta}$
    $\boldsymbol{\gamma}_{(0)} \leftarrow \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - \boldsymbol{\beta}^T(\boldsymbol{\alpha}\boldsymbol{\Sigma}^{-1}\boldsymbol{y})$
    $\eta \leftarrow \log(\sigma_x^{2m}|\mathbf{T}_{(0)}|) - \boldsymbol{\gamma}_{(0)}^T\mathbf{T}_{(0)}^{-1}\boldsymbol{\gamma}_{(0)} + 2\log(\frac{1-\lambda}{\lambda})$
    Tirer $u \sim \mathcal{U}([0;1[)$
    **si** $u > (1+\exp(-\eta/2)^{-1})$ **alors**
      % Naissance acceptée
      $q_j \leftarrow 1$, $L \leftarrow L+1$
      $\mathbf{b}_0 \leftarrow -\mathbf{F}^T\boldsymbol{\beta}\mathbf{T}_{(0)}^{-1}$
      $\mathbf{G} \leftarrow [\mathbf{G}, \sigma_x\mathbf{h}_j]$
      $\mathbf{F} \leftarrow \text{cholupdate}([\mathbf{F}\ \mathbf{0}; \mathbf{0}\ \mathbf{0}], [\mathbf{b}_0\mathbf{T}_{(0)}^{1/2}; \sigma_x^{-1}\mathbf{T}_{(0)}^{-1/2}], +)$ % Notations Matlab
      % Mise à jour du vecteur d'ordre $\boldsymbol{o}$
      $o_j \leftarrow L$
    **fin si**
  **sinon**
    % Calcul de $\Pr(q_j = 0 \,|\, -)$
    % Récupération de $\mathbf{T}_{(0)}$ et $\boldsymbol{\gamma}_{(0)}$
    $\boldsymbol{a}_j \leftarrow [m(o_j-1)+1 : mo_j]$ % Position associée : $\sigma_x\mathbf{h}_j = \mathbf{G}(:, \boldsymbol{a}_j)$
    $\mathbf{e} \leftarrow \mathbf{F}^T\mathbf{F}(:, \boldsymbol{a}_j)$ % $= [\mathbf{f}_a\ \mathbf{f}_m\ \mathbf{f}_b]^T\mathbf{f}_m$ ; Variable intermédiaire
    $\mathbf{T}_{(0)}^{-1} \leftarrow \sigma_x^2\mathbf{e}(:, \boldsymbol{a}_j)$% $= \sigma_x^2\mathbf{f}_m^T\mathbf{f}_m$
    $\mathbf{b}_0 \leftarrow \sigma_x\mathbf{e}(:, \overline{\boldsymbol{a}_j})$ % $= \sigma_x[\mathbf{f}_a\ \mathbf{f}_b]^T\mathbf{f}_m$ ; $\overline{\boldsymbol{a}_j}$ : tous les indices sauf $\boldsymbol{a}_j$
    $\boldsymbol{\gamma}_{(0)} \leftarrow (\mathbf{h}_j^T + \mathbf{T}_{(0)}^T\mathbf{b}_0^T\mathbf{G}_{(0)}^T)\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$
    $\eta \leftarrow \log(\sigma_x^{2m}|\mathbf{T}_{(0)}|) - \boldsymbol{\gamma}_{(0)}^T\mathbf{T}_{(0)}^{-1}\boldsymbol{\gamma}_{(0)} + 2\log(\frac{1-\lambda}{\lambda})$
    Tirer $u \sim \mathcal{U}([0;1[)$
    **si** $u < (1+\exp(-\eta/2)^{-1})$ **alors**
      % Mort : mise à jour de $\mathbf{G}$ et $\mathbf{F}$
      $\mathbf{G}(:, \boldsymbol{a}_j) \leftarrow \varnothing$, $\boldsymbol{i}_b \leftarrow mo_j{+}1 : mL$
      **si** $o_j < L$ **alors**
        $\mathbf{F}(\boldsymbol{i}_b, \boldsymbol{i}_b) \leftarrow \text{cholupdate}(\mathbf{F}(\boldsymbol{i}_b, \boldsymbol{i}_b), \mathbf{F}(\boldsymbol{a}_j, \boldsymbol{i}_b), +)$
      **fin si**
      $\mathbf{f} \leftarrow \mathbf{F}(\overline{\boldsymbol{a}_j}, \overline{\boldsymbol{a}_j})$ % On retire les lignes et colonnes d'indices $\boldsymbol{a}_j$
      $\mathbf{F} \leftarrow \text{cholupdate}(\mathbf{f}, \mathbf{b}_0\mathbf{T}_{(0)}^{-1/2}, -)$
      % Mise à jour du vecteur d'ordre $\boldsymbol{o}$ : on enlève 1 à tous les $o_i > o_j$
      $\boldsymbol{o}(i \,|\, (o_i > o_j)) \leftarrow$ - -1, $o_j \leftarrow 0$
      $q_j \leftarrow 0$, $L \leftarrow L-1$
    **fin si**
  **fin si**

### 3.3.3 Astuces calculatoires pour l'échantillonnage des $\boldsymbol{\nu}_j$

Pour rappel, nous avons vu en § 3.3.1 que l'échantillonnage des paramètres non linéaires se fait via une étape de Métropolis-Hastings, où la proposition $\boldsymbol{\nu}'_j$ est acceptée suivant une probabilité d'acceptation $\alpha$ définie en (3.20) : $\alpha = 1$ quand $q_j = 0$ et $\alpha = \exp(-\Psi_{\boldsymbol{\nu}_j}/2)$ avec $\Psi_{\boldsymbol{\nu}_j} = \Psi(\mathbf{B}^{-1}_{(1,\boldsymbol{\nu}'_j)}, \mathbf{B}^{-1}_{(1,\boldsymbol{\nu}_j^{(t)})}; \boldsymbol{y})$ quand $q_j = 1$. Dans ce second cas, l'échantillonnage des paramètres non linéaires fait donc intervenir, tout comme pour celui des paramètres de Bernoulli, le calcul de la fonction $\Psi(\cdot, \cdot; \boldsymbol{y})$, que l'utilisation directe des matrices $\mathbf{B}$ rend particulièrement coûteux en calcul.

Par ailleurs, dans le cas où $q_j = 1$ et où la proposition de « vie » est acceptée (soit $\boldsymbol{\nu}_j^{(t+1)} \leftarrow \boldsymbol{\nu}'_j$), l'atome $\mathbf{h}_j = \mathbf{h}(\boldsymbol{\nu}_j^{(t)})$ évolue vers $\mathbf{h}'_j = \mathbf{h}(\boldsymbol{\nu}'_j)$ au sein de la matrice $\overline{\mathbf{H}}$. Les matrices ($\mathbf{F}$,$\mathbf{G}$), que nous avons définies en § 3.3.2, dépendent de $\mathbf{h}_j$ par définition quand $q_j = 1$ ; il est donc nécessaire de les mettre à jour après une évolution de $\mathbf{h}_j$, afin qu'elles puissent être utilisées pour l'échantillonnage de $q_{j+1}$. Dans le cas où $q_j = 0$, la proposition est toujours acceptée, mais n'a pas d'influence sur les matrices $(\mathbf{F}, \mathbf{G})$ puisque le $\mathbf{h}_j$ n'est pas dans le dictionnaire réduit $\overline{\mathbf{H}}$.

À partir de ces observations, il est donc nécessaire d'étendre les astuces calculatoires vues en § 3.3.2 pour l'échantillonnage des paramètres de Bernoulli à celui des paramètres non linéaires, uniquement quand $q_j = 1$ :

1. pour réduire le coût de calcul de $\Psi_{\boldsymbol{\nu}_j}$ en évitant l'utilisation directe des matrices $\mathbf{B}$ ;
2. pour pouvoir mettre à jour le facteur de Cholesky $\mathbf{F}$ dans le cas où la « Vie » est acceptée.

Nous proposons alors une extension reposant directement sur les calculs vus précédemment, en considérant que la réalisation d'une « Vie » consiste en la « Mort » de l'atome $\mathbf{h}_j$ suivi de la « Naissance » de l'atome $\mathbf{h}'_j$.

#### Simplification de $\Psi_{\boldsymbol{\nu}_j}$ grâce à $(\mathbf{T}_{(0,\boldsymbol{\nu}_j^{(t)})}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}_j^{(t)})})$ et $(\mathbf{T}_{(0,\boldsymbol{\nu}'_j)}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}'_j)})$

En introduisant la matrice $\mathbf{B}_{(0)}$, on montre que $\Psi_{\boldsymbol{\nu}_j}$ peut se calculer à moindre coût (*i.e.* sans passer par la matrice $\mathbf{B}$) car il s'exprime en fonction de matrices $\mathbf{T}_{(0)}$ et $\gamma_{(0)}$ définies en (3.22) :

$$
\begin{aligned}
\Psi_{\boldsymbol{\nu}_j} &= \log\left(\frac{|\mathbf{B}_{(1,\boldsymbol{\nu}'_j)}|}{|\mathbf{B}_{(1,\boldsymbol{\nu}_j^{(t)})}|}\frac{|\mathbf{B}_{(0)}|}{|\mathbf{B}_{(0)}|}\right) + \boldsymbol{y}^T\left(\mathbf{B}^{-1}_{(1,\boldsymbol{\nu}'_j)} - \mathbf{B}^{-1}_{(0)} + \mathbf{B}^{-1}_{(0)} - \mathbf{B}^{-1}_{(1,\boldsymbol{\nu}_j^{(t)})}\right)\boldsymbol{y} \\
&= \Psi(\mathbf{B}^{-1}_{(1,\boldsymbol{\nu}'_j)}, \mathbf{B}^{-1}_{(0)}; \boldsymbol{y}) - \Psi(\mathbf{B}^{-1}_{(1,\boldsymbol{\nu}_j^{(t)})}, \mathbf{B}^{-1}_{(0)}; \boldsymbol{y}) \quad \text{d'après (3.19)} \\
&= \log\frac{|\mathbf{T}_{(0,\boldsymbol{\nu}'_j)}|}{|\mathbf{T}_{(0,\boldsymbol{\nu}_j^{(t)})}|} + \boldsymbol{\gamma}^T_{(0,\boldsymbol{\nu}'_j)}\mathbf{T}_{(0,\boldsymbol{\nu}'_j)}\boldsymbol{\gamma}_{(0,\boldsymbol{\nu}'_j)} - \boldsymbol{\gamma}^T_{(0,\boldsymbol{\nu}_j^{(t)})}\mathbf{T}_{(0,\boldsymbol{\nu}_j^{(t)})}\boldsymbol{\gamma}_{(0,\boldsymbol{\nu}_j^{(t)})} \quad \text{d'après (3.23)}
\end{aligned}
$$

où $(\mathbf{T}_{(0,\boldsymbol{\nu}_j^{(t)})}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}_j^{(t)})})$ et $(\mathbf{T}_{(0,\boldsymbol{\nu}'_j)}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}'_j)})$ ont la même expression analytique (3.22) ; le premier dépend de $\mathbf{h}_j = \mathbf{h}(\boldsymbol{\nu}^{(t)})$ tandis que le second dépend de $\mathbf{h}'_j = \mathbf{h}(\boldsymbol{\nu}'_j)$. En suivant les résultats vus précédemment, ces matrices peuvent se calculer en fonction de matrices ($\mathbf{F}$,$\mathbf{G}$) via l'équation (3.24), si bien que le calcul de $\Psi_{\boldsymbol{\nu}_j}$ ne fait plus intervenir de matrice $\mathbf{B}$ et se fait donc à moindre coût.

#### Calcul des $(\mathbf{T}_{(0)}, \boldsymbol{\gamma}_{(0)})$ grâce au calcul de $(\mathbf{F}_{(0)}$,$\mathbf{G}_{(0)})$

Nous avons vu en § 3.3.2 que les matrices $(\mathbf{T}_{(0,\boldsymbol{\nu}_j)}$,$\boldsymbol{\gamma}_{(0,\boldsymbol{\nu}_j)})$ pouvaient se calculer de deux façons à partir des facteurs de Cholesky : soit en fonction de $(\mathbf{F}_{(0)}, \mathbf{G}_{(0)})$ par définition (3.24),

soit à partir de $(\mathbf{F}_{(1)}, \mathbf{G}_{(1)})$ comme nous l'avons vu avec notre nouvelle stratégie pour l'étape de « Mort » (voir (3.30)).

Au moment du calcul de $\Psi_{\boldsymbol{\nu}_j}$, nous avons accès aux matrices $(\mathbf{F}, \mathbf{G})$ correspondant à l'état courant $(q_j = 1, \mathbf{h}_j = \mathbf{h}(\boldsymbol{\nu}_j^{(t)}))$, que nous notons $(\mathbf{F}_{(1)}, \mathbf{G}_{(1)})$ ; ces dernières permettent donc le calcul de $(\mathbf{T}_{(0,\boldsymbol{\nu}_j^{(t)})}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}_j^{(t)})})$. Cependant, les matrices $(\mathbf{F}, \mathbf{G})$ correspondant à l'état proposé $(q_j = 1, \mathbf{h}'_j = \mathbf{h}(\boldsymbol{\nu}'_j))$, que nous notons $(\mathbf{F}'_{(1)}, \mathbf{G}'_{(1)})$, sont inconnues ; les matrices $(\mathbf{T}_{(0,\boldsymbol{\nu}'_j)}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}'_j)})$ ne peuvent donc pas être directement calculées.

Nous proposons donc de calculer les matrices $(\mathbf{F}_{(0)}, \mathbf{G}_{(0)})$ en effectuant la « Mort » de l'atome $\mathbf{h}_j$ comme décrit en équations (3.28) et (3.29) à partir des matrices $(\mathbf{F}_{(1)}, \mathbf{G}_{(1)})$, afin de rendre possible le calcul de $(\mathbf{T}_{(0,\boldsymbol{\nu}'_j)}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}'_j)})$ via l'équation (3.24). Notons que l'on peut alors calculer $(\mathbf{T}_{(0,\boldsymbol{\nu}_j)}, \boldsymbol{\gamma}_{(0,\boldsymbol{\nu}_j)})$ à partir de $(\mathbf{F}_{(0)}, \mathbf{G}_{(0)})$ de la même façon.

*Remarque : En remarquant que* $\Psi(\mathbf{B}_{(1)}^{-1}, \mathbf{B}_{(0)}^{-1}; \boldsymbol{y}) = -\Psi(\mathbf{B}_{(0)}^{-1}, \mathbf{B}_{(1)}^{-1}; \boldsymbol{y})$*, le scalaire* $\Psi_{\boldsymbol{\nu}_j}$ *peut aussi s'écrire en fonction des matrices* $(\mathbf{T}_{(1)}, \boldsymbol{\gamma}_{(1)})$*. Cependant, il serait alors nécessaire de calculer les matrices* $(\mathbf{F}'_{(1)}, \mathbf{G}'_{(1)})$ *; par ailleurs, nous avons vu précédemment que le calcul des matrices* $\mathbf{T}_{(1)}$ *pouvait apporter des problèmes d'instabilité numérique. Notre approche est donc préférable.*

#### Mise à jour des matrices $(\mathbf{F}, \mathbf{G})$ à partir de $(\mathbf{F}_{(0)}, \mathbf{G}_{(0)})$

En cas d'acceptation de la « Vie », il est nécessaire de mettre à jour les matrices $(\mathbf{F}, \mathbf{G})$ courantes pour prendre en compte le remplacement de l'atome $\mathbf{h}_j$ par $\mathbf{h}'_j$. Les matrices $(\mathbf{F}'_{(1)}, \mathbf{G}'_{(1)})$ sont alors calculées en réalisant la « Naissance » de l'atome $\mathbf{h}'_j$ comme décrit en équation (3.27) à partir des matrices $(\mathbf{F}_{(0)}, \mathbf{G}_{(0)})$ précédemment calculées pour le calcul de $\Psi_{\boldsymbol{\nu}_j}$. Notons qu'il est aussi nécessaire de mettre à jour le *vecteur d'ordre* $\boldsymbol{o}$ : « mort » de $o_j$ (mise à zéro et décalage des suivants) puis « naissance » de $o_j$ qui prend la valeur courante de $L$.

#### Bilan : pseudo-code pour l'échantillonnage des $\boldsymbol{\nu}_j$

L'échantillonnage des paramètres non linéaires $\boldsymbol{\nu}_j$ que nous proposons s'inscrit donc dans la logique de celui des paramètres de Bernoulli. Le calcul de la probabilité d'acceptation $\alpha = \exp(-\Psi_{\boldsymbol{\nu}_j}/2)$ s'effectue à moindre coût grâce à l'utilisation du facteur de Cholesky $\mathbf{F}$ courant après l'échantillonnage de $q_j$. Puis nous mettons ce dernier à jour si la « Vie » est réalisée afin qu'il puisse être utilisé pour l'échantillonnage du paramètre de Bernoulli suivant $q_{j+1}$.

Le pseudo-code 2 récapitule les différentes étapes lors de l'échantillonnage de $\boldsymbol{\nu}_j$ via une étape de Métropolis-Hastings.

### 3.3.4 Bilan sur l'échantillonnage de la loi *a posteriori* marginalisée dans le cadre supervisé

Nous avons donc un algorithme d'échantillonnage de type Gibbs-hybride de la loi *a posteriori* marginale du modèle BGE dans le cadre supervisé $p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y})$ correspondant à l'algorithme Tab. 3.8 (p. 115). Nous avons mis en place une implémentation efficace des deux étapes d'échantillonnage, récapitulée dans les pseudo-codes 1 (échantillonnage de $q_j$) et 2 (échantillonnage de $\boldsymbol{\nu}_j$). Nous nous sommes inspirés des travaux de [GILC11] pour un cas particulier de modèle BG linéaire.

**Pseudo-code 2** Échantillonnage de $\boldsymbol{\nu}_j$ selon $p(\boldsymbol{\nu}_j \,|\, -)$ (étape de Métropolis-Hasting) dans le cas d'une loi de proposition symétrique et d'une loi *a priori* non informative sur les $\boldsymbol{\nu}_j$.

**Entrée:** $q_j$, $\mathbf{F}$, $\mathbf{G}$, $\boldsymbol{\Sigma}^{-1}$, $\sigma_x^2$, $\mathbf{h}_j$, $\boldsymbol{o}$
**Entrée:** proposition$(\cdot)$, $\mathbf{h}(\cdot)$ % permettent de proposer $\boldsymbol{\nu}'_j$ et de calculer $\mathbf{h}'_j$.
  Tirer $\boldsymbol{\nu}'_j \sim$ proposition$(\boldsymbol{\nu}_j)$
  **si** $q_j = 0$ **alors**
    $\boldsymbol{\nu}_j \leftarrow \boldsymbol{\nu}'_j$ % On accepte toujours
  **sinon**
    % $q_j = 1$ : étape MH pour échantillonnage de $\boldsymbol{\nu}_j$
    $\mathbf{h}'_j \leftarrow \mathbf{h}(\boldsymbol{\nu}'_j)$
    % Calcul de $\mathbf{G}_{(0)}$ et $\mathbf{F}_{(0)}$ (comme une mise à jour lors d'une mort)
    $\mathbf{G}_{(0)} \leftarrow \mathbf{G}$, $\mathbf{G}_{(0)}(:, \boldsymbol{a}_j) \leftarrow \varnothing$
    $\mathbf{e} \leftarrow \mathbf{F}^T\mathbf{F}(:, \boldsymbol{a}_j)$ % $= [\mathbf{f}_a\ \mathbf{f}_m\ \mathbf{f}_b]^T\mathbf{f}_m$ ; Variable intermédiaire
    $\mathbf{T}_{(0)}^{-1} \leftarrow \sigma_x^2\mathbf{e}(:, \boldsymbol{a}_j)$% $= \sigma_x^2\mathbf{f}_m^T\mathbf{f}_m$
    $\mathbf{b}_0 \leftarrow \sigma_x\mathbf{e}(:, \overline{\boldsymbol{a}_j})$ % $= \sigma_x[\mathbf{f}_a\ \mathbf{f}_b]^T\mathbf{f}_m$ ; $\overline{\boldsymbol{a}_j}$ : tous les indices sauf $\boldsymbol{a}_j$
    $\mathbf{F}_{(0)} \leftarrow \mathbf{F}$, $\boldsymbol{i}_b \leftarrow mo_j{+}1{:}mL$ % initialisation
    **si** $o_j < L$ **alors**
      $\mathbf{F}_{(0)}(\boldsymbol{i}_b, \boldsymbol{i}_b) \leftarrow$ cholupdate$(\mathbf{F}_{(0)}(\boldsymbol{i}_b, \boldsymbol{i}_b), \mathbf{F}_{(0)}(\boldsymbol{a}_j, \boldsymbol{i}_b), +)$
    **fin si**
    $\mathbf{f} \leftarrow \mathbf{F}_{(0)}(\overline{\boldsymbol{a}_j}, \overline{\boldsymbol{a}_j})$ % On retire les lignes et colonnes d'indices $\boldsymbol{a}_j$
    $\mathbf{F}_{(0)} \leftarrow$ cholupdate$(\mathbf{f}, \mathbf{b}_0\mathbf{T}_{(0)}^{-1/2}, -)$
    % Calcul des différents de $\mathbf{T}_{(0)}$ et $\boldsymbol{\gamma}_{(0)}$
    $\boldsymbol{\alpha} \leftarrow \mathbf{F}_{(0)}\mathbf{G}_{(0)}^T$, $\boldsymbol{\beta}' \leftarrow \boldsymbol{\alpha}\boldsymbol{\Sigma}^{-1}\mathbf{h}'_j$
    $\boldsymbol{\gamma}_{(0)} \leftarrow (\mathbf{h}_j^T + \mathbf{T}_{(0)}^T\mathbf{b}_0^T\mathbf{G}_{(0)}^T)\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$
    $\mathbf{T}'_{(0)} \leftarrow \sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j'^T\boldsymbol{\Sigma}^{-1}\mathbf{h}'_j - \boldsymbol{\beta}'^T\boldsymbol{\beta}'$
    $\boldsymbol{\gamma}'_{(0)} \leftarrow \mathbf{h}_j'^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - \boldsymbol{\beta}'^T(\boldsymbol{\alpha}\boldsymbol{\Sigma}^{-1}\boldsymbol{y})$
    % Calcul de la probabilité d'acceptation
    $\alpha \leftarrow \left(\frac{|\mathbf{T}'_{(0)}|}{|\mathbf{T}_{(0)}|}\right)^{-1/2}\exp\left(\frac{1}{2}(\boldsymbol{\gamma}'^T_{(0)}\mathbf{T}'^{-1}_{(0)}\boldsymbol{\gamma}'_{(0)} - \boldsymbol{\gamma}_{(0)}^T\mathbf{T}_{(0)}^{-1}\boldsymbol{\gamma}_{(0)})\right)$
    Tirer $u \sim \mathcal{U}([0;1[)$
    **si** $u < \alpha$ **alors**
      % Acceptation de la vie : mise à jour comme une naissance à partir de $\mathbf{F}_{(0)}$ et $\mathbf{G}_{(0)}$
      $\boldsymbol{\nu}_j \leftarrow \boldsymbol{\nu}'_j$
      $\mathbf{G} \leftarrow [\mathbf{G}_{(0)}\ \sigma_x\mathbf{h}'_j]$
      $\mathbf{b}'_0 \leftarrow -\mathbf{F}_{(0)}^T\mathbf{F}_{(0)}\mathbf{G}_{(0)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}'_j\mathbf{T}'^{-1}_{(0)}$
      $\mathbf{F} \leftarrow$ cholupdate$([\mathbf{F}_{(0)}\ \mathbf{0}; \mathbf{0}\ \mathbf{0}], [\mathbf{b}'_0\mathbf{T}'^{1/2}_{(0)}; \sigma_x^{-1}\mathbf{T}'^{-1/2}_{(0)}], +)$ % Notations Matlab
      % Mise à jour de $\boldsymbol{o}$ : mort de $q_j$ en position $o_j$ puis naissance de $q_j$ en position $o_j = L$
      $\boldsymbol{o}(i \,|\, (o_i > o_j)) \leftarrow$ - -1 % (mort)
      $o_j \leftarrow L$ % (naissance)
    **fin si**
  **fin si**

D'une part, nous avons étendu ces calculs pour l'échantillonnage de $q_j$ aux spécificités de notre problème : bruit non indépendant de matrice de covariance $\mathbf{\Sigma}$ et amplitudes $\boldsymbol{x}_j$ multidimensionnelles de variance $\sigma_x^2$ inconnue. Nos apports reposent à la fois sur l'extension des calculs dans ce cas particulier, mais aussi sur une stratégie différente pour l'étape de « Mort » pour minimiser des problèmes d'instabilité numérique.

D'autre part, nous avons exploité les calculs et résultats obtenus pour l'échantillonnage des $q_j$ pour une implémentation à moindre coût de l'échantillonnage des paramètres non linéaires $\boldsymbol{\nu}_j$ suivant leur loi marginale conditionnelle, ce qui n'avait pas encore été réalisé à notre connaissance. Cette exploitation garantie un enchaînement efficace de l'échantillonnage des $q_j$ et des $\boldsymbol{\nu}_j$.

Nous allons montrer dans la section suivante que nous pouvons exploiter l'échantillonnage des paramètres $q_j$ et $\boldsymbol{\nu}_j$ suivant la loi marginalisée, mais cette fois-ci *dans le cadre non supervisé*.

## 3.4 Échantillonnage exploitant la marginalisation dans le cadre non supervisé

Dans la section précédente, nous avons vu comment échantillonner la loi *a posteriori* marginalisée dans le cadre supervisé, c'est-à-dire que nous avons supposé les hyperparamètres $\boldsymbol{\theta} = (\lambda, \sigma_x^2, \mathbf{\Sigma})$ connus et fixés. Nous considérons maintenant le cadre non supervisé, où les hyperparamètres sont inconnus. Notons que si le bruit est supposé *i.i.d.*, sa matrice de covariance $\mathbf{\Sigma}$ s'écrit $\sigma_\epsilon^2 \mathbf{I}_N$ et l'hyperparamètre à échantillonner est donc la variance $\sigma_\epsilon^2$. Dans le cas contraire, on conserve la matrice de covariance $\mathbf{\Sigma}$ mais elle sera supposée connue. Nous présentons donc ici l'échantillonnage de l'ensemble des paramètres $\boldsymbol{\theta} = [\lambda, \sigma_x^2, \sigma_\epsilon^2]$.

L'échantillonnage suivant la loi jointe $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$, proposé par [BC06], ne pose pas de problème en théorie mais souffre des lenteurs de convergence de l'algorithme de Gibbs. *L'objectif de cette section est donc de proposer un échantillonneur bénéficiant de la marginalisation des amplitudes $\boldsymbol{x}$ pour l'échantillonnage des paramètres $\boldsymbol{q}$ et $\boldsymbol{\nu}$, comme présenté lors de la section précédente, pour une accélération de la convergence.*

Un premier échantillonneur serait donc celui de la loi marginalisée suivant les amplitudes $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$, dont l'algorithme est donné en Tableau 3.9. Aux étapes d'échantillonnage des paramètres de Bernoulli $\boldsymbol{q}$ et des paramètres non linéaires $\boldsymbol{\nu}_j$ déjà décrites en § 3.3, nous devons ajouter l'échantillonnage des hyperparamètres $(\lambda, \sigma_x^2, \sigma_\epsilon^2)$ suivant leur loi *a posteriori* marginalisée conditionnelle qui s'exprime à partir de l'expression de $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ (3.18) :

$$
\begin{aligned}
p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y}) &\propto |\mathbf{B}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \mathbf{B}^{-1} \boldsymbol{y}\right) \lambda^L (1-\lambda)^{J-L} p(\boldsymbol{\nu}) p(\lambda) p(\sigma_x^2) p(\sigma_\epsilon^2) \\
p(\lambda \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \sigma_x^2, \sigma_\epsilon^2, \boldsymbol{y}) &\propto \lambda^L (1-\lambda)^{J-L} p(\lambda) \quad (\propto p(\lambda \,|\, \boldsymbol{q})) \\
p(\sigma_x^2 \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \lambda, \sigma_\epsilon^2, \boldsymbol{y}) &\propto |\mathbf{B}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \mathbf{B}^{-1} \boldsymbol{y}\right) p(\sigma_x^2) \quad (\propto p(\sigma_x^2 \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \sigma_\epsilon^2)) \\
p(\sigma_\epsilon^2 \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \lambda, \sigma_x^2, \boldsymbol{y}) &\propto |\mathbf{B}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \mathbf{B}^{-1} \boldsymbol{y}\right) p(\sigma_\epsilon^2) \quad (\propto p(\sigma_\epsilon^2 \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \sigma_x^2))
\end{aligned}
$$

La loi marginalisée conditionnelle en $\lambda$ revêt la même expression que celle de la loi jointe (3.7), puisque celle-ci ne dépend pas des amplitudes. Pour une loi *a priori* beta $\lambda \sim \mathcal{B}e(\alpha, \beta)$, alors $(\lambda \,|\, -) \sim \mathcal{B}e(L+\alpha, J-L+\beta)$ (voir le Tableau 3.3 p. 104 pour plus de détails).

À l'opposé, l'expression des lois marginalisées conditionnelles en $\sigma_\epsilon^2$ et $\sigma_x^2$ deviennent plus complexes car elles dépendent non-linéairement de la matrice $\mathbf{B}$, laquelle dépend de $\sigma_\epsilon^2$ et $\sigma_x^2$ (pour rappel, $\mathbf{B} = \sigma_\epsilon^2 \mathbf{I}_N + \sigma_x^2 \overline{\mathbf{H}} \overline{\mathbf{H}}^T$). Par conséquent, aucune loi *a priori* $p(\sigma_x^2)$ et $p(\sigma_\epsilon^2)$ ne permet d'écrire les lois marginalisées conditionnelles comme des lois usuelles facilement simulables numériquement. Pour échantillonner ces variances, on doit alors faire appel à une étape de

Métropolis-Hastings, comme pour l'échantillonnage des paramètres non linéaires : une variance $\sigma'^2$ ($\sigma_x^2$ ou $\sigma_\epsilon^2$) est proposée et acceptée suivant une probabilité d'acceptation calculée comme le rapport des lois cibles - et des lois de propositions si elles ne sont pas symétriques (pour plus de détails sur cette étape de Métropolis-Hastings, voir en Annexe B.7). Cet échantillonnage de type *Acceptation/Rejet* est cependant moins efficace que l'échantillonnage direct selon la loi cible, effectué pour l'échantillonnage de la loi jointe en Tableau 3.3.

Itération $(t+1)$, à partir des échantillons courants $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{\nu}^{(t)}, \boldsymbol{\theta}^{(t)})$ :

Générer pour tout $j \in [\![1; J]\!]$,

$$\begin{array}{rcll} q_j^{(t+1)} & \sim & (q_j \,|\, \boldsymbol{q}_{-j}, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y}) & : \text{ loi de Bernoulli (voir (3.19))} \\ \boldsymbol{\nu}_j^{(t+1)} & \sim & (\boldsymbol{\nu}_j \,|\, \boldsymbol{q}, \boldsymbol{\nu}_{-j}, \boldsymbol{\theta}, \boldsymbol{y}) & : \text{ étape Métropolis-Hastings (voir (3.20))} \end{array}$$

Puis, générer les hyperparamètres :

$$\begin{array}{rcll} \lambda^{(t+1)} & \sim & (\lambda \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta}_{-\lambda}, \boldsymbol{y}) & : \mathcal{B}e(L+\alpha, J-L+\beta) \\ \sigma_x^{2\,(t+1)} & \sim & (\sigma_x^2 \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \lambda, \sigma_\epsilon^2, \boldsymbol{y}) & : \text{ étape Métropolis-Hastings} \\ \sigma_\epsilon^{2\,(t+1)} & \sim & (\sigma_\epsilon^2 \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta}_{-\sigma_\epsilon^2}, \boldsymbol{y}) & : \text{ étape Métropolis-Hastings} \end{array}$$

Tab. 3.9 – Échantillonneur de la loi *a posteriori* marginalisée dans le cadre non supervisé.

Dans cette section, nous proposons alors d'utiliser le principe du *Partially Collapsed Gibbs Sampler* (PCGS), introduit par [vDP08] et utilisé par [GILC11] dans le cadre du modèle BG classique. Dans notre cas, il permet de justifier l'échantillonnage de la loi *a posteriori* jointe $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ tout en exploitant un échantillonnage des lois marginales conditionnelles en les paramètres $q_j$ et $\boldsymbol{\nu}_j$ ; les hyperparamètres de variance sont donc directement tirés selon leur loi jointe conditionnelle, évitant les étapes de Métropolis-Hastings.

On commencera donc par présenter le PCGS pour notre modèle BGE : après avoir décrit l'échantillonneur, nous justifions sa validité en suivant les principes de [vDP08]. Nous justifions le choix du PCGS plutôt que l'échantillonnage de la loi marginalisée faisant intervenir des étapes MH pour l'échantillonnage des variances et enfin nous rentrerons dans les détails de son implémentation en faisant notamment intervenir le facteur de Cholesky $\mathbf{F}$.

Dans une seconde partie, nous donnerons quelques conseils quand au choix des *a priori* des hyperparamètres $\lambda$ et $\sigma_x^2$, résultant d'observations de simulations effectuées pendant nos travaux.

### 3.4.1 Échantillonnage de type *Partially Collapsed Gibbs Sampler* pour le modèle BGE

Le principe du *Partially collapsed Gibbs Sampler* (PCGS), proposé par [vDP08] et utilisé par [GILC11], consiste à simuler une variable auxiliaire telle que la loi cible conditionnellement à cette variable soit facilement simulable. Nous avons vu que les variances $\sigma_x^2$ et $\sigma_\epsilon^2$ ne sont pas directement simulables suivant leur loi marginalisée conditionnelle ; elles le sont cependant suivant leur loi jointe conditionnelle (voir § 3.2.1). De plus, nous savons simuler les amplitudes $\boldsymbol{x}$ suivant leur loi *a posteriori* conditionnelle, puisque cette dernière est une loi gaussienne multidimensionnelle dont nous connaissons l'expression de la moyenne $\boldsymbol{\mu}$ et de sa matrice de covariance $\boldsymbol{\Gamma}$ (voir (3.5) et (3.12)). Ainsi, le principe du PCGS dans notre cas est le suivant :

1. Tirer $\overline{\boldsymbol{x}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$, $\overline{\boldsymbol{x}}^C \sim \delta_{\overline{\boldsymbol{x}}^C}(\mathbf{0})$ - $\boldsymbol{x}$ est la variable auxiliaire ;
2. Tirer $\sigma_x^2$ et $\sigma_\epsilon^2$ suivant leur loi jointe conditionnelle $p(\sigma^2 \,|\, \boldsymbol{x}, -)$ (voir Tableau 3.3 p. 104).

L'algorithme de type PCGS que nous proposons est donné en Tableau 3.10. *Ce dernier échantillonne la loi* a posteriori *jointe* $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ *dans un cadre non supervisé, tout en bénéficiant de l'échantillonnage des paramètres d'intérêt* $(\boldsymbol{q}, \boldsymbol{\nu})$ *suivant leur loi conditionnelle marginalisée.* Cet échantillonneur est très similaire à celui proposé par [GILC11] pour le modèle BG linéaire, à la différence près de l'échantillonnage des paramètres non linéaires $\boldsymbol{\nu}_j$, qui se fait suivant la loi *a posteriori* marginalisée (et aussi de l'échantillonnage de la variances des amplitudes, fixée à 1 dans [GILC11]). Il est donc nécessaire de montrer la validité de cet algorithme, *i.e.* qu'il s'agit bien d'un algorithme MCMC dont la convergence vers la loi cible est assurée.

Itération $(t+1)$, à partir des échantillons courants $(\boldsymbol{q}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{\nu}^{(t)}, \boldsymbol{\theta}^{(t)})$ :

Générer pour tout $j \in [\![1; J]\!]$,

$$
\begin{array}{rcll}
q_j^{(t+1)} & \sim & (q_j \,|\, \boldsymbol{q}_{-j}, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y}) & : \text{ loi de Bernoulli (voir (3.19))} \\
\boldsymbol{\nu}_j^{(t+1)} & \sim & (\boldsymbol{\nu}_j \,|\, \boldsymbol{q}, \boldsymbol{\nu}_{-j}, \boldsymbol{\theta}, \boldsymbol{y}) & : \text{ étape Métropolis-Hastings (voir (3.20))}
\end{array}
$$

Puis, générer les hyperparamètres :

$$
\begin{array}{rcll}
\lambda^{(t+1)} & \sim & (\lambda \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta}_{-\lambda}, \boldsymbol{y}) & : \mathcal{B}e(L+\alpha, J-L+\beta) \\
\rightarrow \boldsymbol{x}^{(t+1)} & \sim & (\boldsymbol{x} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \lambda, \sigma_\epsilon^2, \boldsymbol{y}) & : \overline{\boldsymbol{x}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma}) \text{ et } \overline{\boldsymbol{x}}^C \sim \delta_{\overline{\boldsymbol{x}}^C}(\boldsymbol{0})) \text{ (3.12)} \\
\sigma_x^{2(t+1)} & \sim & (\sigma_x^2 \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \lambda, \sigma_\epsilon^2, \boldsymbol{y}) & : \mathcal{IG}(\alpha_x + Lm/2, \beta_x + \|\boldsymbol{x}\|^2/2) \\
\sigma_\epsilon^{2(t+1)} & \sim & (\sigma_\epsilon^2 \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\theta}_{-\sigma_\epsilon^2}, \boldsymbol{y}) & : \mathcal{IG}(\alpha_\epsilon + N/2, \beta_\epsilon + \|\boldsymbol{y} - \mathbf{H}(\boldsymbol{\nu})\boldsymbol{x}\|^2/2)
\end{array}
$$

Tab. 3.10 – Échantillonneur proposé de type PCGS de la loi *a posteriori* jointe dans le cadre non supervisé.

### Validité de l'échantillonneur de type PCGS proposé

Pour démontrer que cet algorithme satisfait bien les conditions d'obtenir un PCGS valide et efficace, nous exploitons pour l'échantillonnage de la loi $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ les principes de « marginalisation » et de « trimming » donnés par [vDP08]. Cette démonstration s'inspire directement de celle donnée par [GILC11] pour leur PCGS dans le cas de leur modèle BG linéaire.

À partir de l'échantillonneur de Gibbs de la loi jointe donné en Tab. 3.5, on peut utiliser la technique de « marginalisation », qui consiste à échantillonner plusieurs fois certaines variables par itération pour obtenir un type particulier d'échantillonneur de Gibbs, toujours valide. Ici, cela consistera à remplacer d'échantillonnage d'une amplitude $\boldsymbol{x}_j \sim (\boldsymbol{x}_j \,|\, q_j, \boldsymbol{x}{-}j, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y})$ par celui de l'ensemble des amplitudes $\boldsymbol{x} \sim (\boldsymbol{x} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y})$, donnant l'ordre d'échantillonnage suivant :

- Pour tout $j = 1, \ldots, J$, échantillonner $(q_j, \boldsymbol{\nu}_j, \boldsymbol{x})$,
- puis échantillonner $(\lambda, \sigma_x^2, \sigma_\epsilon^2)$.

Ensuite, on remarque que l'échantillonnage de $(q_j, \boldsymbol{\nu}_j, \boldsymbol{x})$ peut se faire d'une façon différente en appliquant la règle de Bayes : $p(q_j, \boldsymbol{\nu}_j, \boldsymbol{x} \,|\, \boldsymbol{y}, -) = p(\boldsymbol{x} \,|\, \boldsymbol{y}, q_j, \boldsymbol{\nu}_j, -) p(q_j, \boldsymbol{\nu}_j \,|\, \boldsymbol{y}, -)$, soit :

1. Tirer $(q_j, \boldsymbol{\nu}_j)$ suivant la loi marginale en échantillonnant d'abord directement $q_j \sim \Pr(q_j \,|\, -)$ puis $\boldsymbol{\nu}_j$ suivant une étape MH ;
2. tirer $\boldsymbol{x}$ suivant $p(\boldsymbol{x} \,|\, \boldsymbol{y}, q_j, \boldsymbol{\nu}_j, -)$.

Il suffit ensuite d'utiliser la technique de « trimming » visant à supprimer les échantillons inutiles. Ici, les $\boldsymbol{x}$ tirés aux $(J-1)$ premières étapes sont inutiles, puisque la séquences de Bernoulli

ainsi que les paramètres non linéaires peuvent être tirés suivant la loi marginale. On obtient donc l'échantillonneur que nous avons proposé en Tableau 3.10, qui est donc un PCGS valide.

La loi cible du PCGS est la loi jointe $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$, contrairement à la première alternative d'échantillonnage exploitant une étape de Métropolis-Hastings dont la loi cible est la loi marginalisée. Nous comparons donc les deux approches dans le paragraphe suivant.

#### Comparaison de l'échantillonneur PCGS avec celui de la loi marginale

Si les deux alternatives sont valides sur le plan théorique, choisir l'une ou l'autre relève des compromis suivants :

| Gibbs - loi marginale | vs. | PCGS |
|---|---|---|
| (+) échantillonnage de la loi marginale | | (-) échantillonnage de la loi jointe |
| (-) acceptation/rejet de $(\sigma_x^2\sigma_\epsilon^2)$ | | (+) échantillonnage direct de $(\sigma_x^2\sigma_\epsilon^2)$ |
| (=) échant. de $q_j$ et $\boldsymbol{\nu}_j$ suivant leur loi marginale conditionnelle | | |

Notons d'abord que l'estimation des paramètres au sens de l'EAP ne sera pas affectée par le choix de la loi cible (jointe ou marginalisée), puisqu'elle s'effectue en calculant la moyenne des échantillons marginalement par rapport aux autres paramètres.

Du point de vue des paramètres d'intérêt $(\boldsymbol{q}, \boldsymbol{\nu})$, les stratégies d'échantillonnage sont équivalentes, puisqu'ils sont tous les deux échantillonnés suivant leur loi conditionnelle marginalisée suivant les amplitudes $\boldsymbol{x}$. Dans les deux cas, la convergence de l'échantillonneur devrait donc être atteinte beaucoup plus rapidement que par celui de la loi jointe. L'influence de la différence de stratégie pour l'échantillonnage des variances $(\sigma_x^2, \sigma_\epsilon^2)$ devrait donc être minime.

Nous préfèrerons, dans le cadre de nos travaux, utiliser l'algorithme de type PCGS que nous avons présenté en Tableau 3.10. En effet, il permet une simulation directe des variances $(\sigma_x^2, \sigma_\epsilon^2)$ qui ne nécessite par le choix d'une loi de proposition adaptée permettant une convergence rapide de l'étape MH de l'échantillonneur de la loi marginalisée en Tableau 3.9.

Bien que nous ne l'ayons pas testé en pratique, nous intuitons que l'implémentation du PCGS est sûrement plus rapide que celle de l'échantillonneur de la loi marginale ; nous allons en effet voir, dans le paragraphe suivant, comment exploiter le facteur de Cholesky $\mathbf{F}$ pour une implémentation efficace.

### 3.4.2 Astuces calculatoires pour l'implémentation du PCGS

L'implémentation de l'algorithme de type PCGS ne doit pas remettre en cause les astuces calculatoires que nous avons développées en § 3.3 pour l'échantillonnage des $(q_j, \boldsymbol{\nu}_j)$ dans le cadre supervisé. Notamment, le facteur de Cholesky $\mathbf{F}$, permettant un calcul efficace des différentes probabilités de transition, dépend des hyperparamètres de variances $(\sigma_x^2, \sigma_\epsilon^2)$ (pour rappel, $\mathbf{F} = \text{chol}(\mathbf{C}^{-1})$ où $\mathbf{C} = \mathbf{I}_{mL} + \frac{\sigma_x^2}{\sigma_\epsilon^2}\overline{\mathbf{H}}^T\overline{\mathbf{H}}$ dans le cas du bruit *i.i.d.*). Une modification de ces variances doit donc entraîner la mise à jour des variables qui en dépendent.

Par ailleurs, nous allons voir que le calcul des vecteurs et matrices $(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ suivant les expressions en équation (3.5) pour l'échantillonnage des amplitudes $\boldsymbol{x}$ peut être facilité par l'utilisation du facteur de Cholesky $\mathbf{F}$.

Le pseudo-code 3 récapitule l'implémentation efficace des différentes étapes de l'échantillonnage des hyperparamètres en prenant en compte les relations ci-dessous sur le facteur de Cholesky. Notons que dans les équations suivantes, nous gardons l'expression complète de la matrice

de covariance du bruit $\mathbf{\Sigma}$ plutôt que son équivalent dans le cas *i.i.d.* $\sigma_\epsilon^2\mathbf{I}_N$, pour pouvoir appliquer ces résultats dans le cas où le bruit n'est pas *i.i.d.* et où $\mathbf{\Sigma}$ est supposée connue.

**Simulation des amplitudes $\boldsymbol{x}$.** La simulation des amplitudes $\boldsymbol{x} \sim (\boldsymbol{x} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta}, \boldsymbol{y})$ s'effectue en dissociant $\boldsymbol{x} = (\overline{\boldsymbol{x}}, \overline{\boldsymbol{x}}^C)$ :

1. $\overline{\boldsymbol{x}}^C \sim \delta_{\overline{\boldsymbol{x}}^C}(\mathbf{0})$, *i.e.* $\overline{\boldsymbol{x}}^C = \mathbf{0}$ puisque les $q_j$ associés sont nuls.
2. $\overline{\boldsymbol{x}} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Gamma})$ avec $\mathbf{\Gamma}^{-1} = \sigma_x^{-2}\mathbf{I}_{mL} + \overline{\mathbf{H}}\mathbf{\Sigma}^{-1}\overline{\mathbf{H}}$ et $\boldsymbol{\mu} = \mathbf{\Gamma}\overline{\mathbf{H}}^T\mathbf{\Sigma}^{-1}\boldsymbol{y}$ (voir équation (3.12)).

La matrice $\mathbf{\Gamma}$ est en fait très proche de la matrice $\mathbf{C}^{-1}$ ($\mathbf{\Gamma} = \sigma_x^2\mathbf{C}^{-1}$) et il est alors immédiat d'établir l'expression de $\mathbf{\Gamma}$ et $\boldsymbol{\mu}$ en fonction de la factorisation de Cholesky de $\mathbf{C}^{-1} = \mathbf{F}^T\mathbf{F}$ :

$$\begin{aligned} \mathbf{\Gamma} &= \sigma_x^2\mathbf{C}^{-1} = (\sigma_x\mathbf{F})^T(\sigma_x\mathbf{F}) \\ \boldsymbol{\mu} &= \sigma_x\mathbf{F}^T\mathbf{F}\mathbf{G}^T\mathbf{\Sigma}^{-1}\boldsymbol{y} \end{aligned}$$

De ce fait, il est immédiat de tirer $\overline{\boldsymbol{x}} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Gamma})$ à partir d'un bruit blanc gaussien $\boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{mL})$ puisqu'il suffit de calculer $\overline{\boldsymbol{x}} = \mu + \sigma_x\mathbf{F}\boldsymbol{u}$.

**Mise à jour de $(\mathbf{F}, \mathbf{G})$ après échantillonnage des variances.** La relation entre la matrice $\mathbf{C}$ et les variances étant non linéaire et complexe, le seul moyen que nous avons trouvé pour mettre à jour le facteur de Cholesky $\mathbf{F}$ est de le recalculer complètement à partir de sa définition :

$$\mathbf{F} = \text{chol}(\mathbf{C}^{-1}) \quad \text{avec} \quad \mathbf{C} = \mathbf{I}_{mL} + \sigma_x^2\overline{\mathbf{H}}^T\mathbf{\Sigma}^{-1}\overline{\mathbf{H}}$$

Ce calcul, plus lourd que les mises à jour de Cholesky de rang $m$ effectuées pour les naissances/vie/mort, ne devrait pas cependant trop alourdir le coût de calcul, puisqu'il ne sera effectué qu'une seul fois par itération (contre $J$ fois pour les étapes de naissances/vie/mort). Notons que la matrice $\mathbf{G} = \sigma_x\overline{\mathbf{H}}$ est facilement mise à jour puisque linéaire en $\sigma_x$.

---

**Pseudo-code 3** Échantillonnage de $\overline{\boldsymbol{x}}$ et des hyperparamètres $\boldsymbol{\theta} = (\lambda, \sigma_x^2, \sigma_\epsilon^2)$ (PCGS).

---

**Entrée:** $\boldsymbol{y}$, $L = \|\boldsymbol{q}\|$, $\mathbf{F}$, $\mathbf{G}$, $\lambda$, $\mathbf{\Sigma} = \sigma_\epsilon^2\mathbf{I}_N$, $\sigma_x^2$, $\overline{\mathbf{H}}$
**Entrée:** $(\alpha, \beta, \alpha_x, \beta_x, \alpha_\epsilon, \beta_\epsilon)$ paramètres des lois *a priori* sur les hyperparamètres.
% Échantillonnage de $\lambda$ : *a priori*, $\lambda \sim \mathcal{B}e(\alpha, \beta)$
$\lambda \sim \mathcal{B}e(L + \alpha, J - L + \beta)$
% Tirage des amplitudes non-nulles $\overline{\boldsymbol{x}} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Gamma})$
$\boldsymbol{\mu} \leftarrow \sigma_x\mathbf{F}^T\mathbf{F}\mathbf{G}^T\mathbf{\Sigma}^{-1}\boldsymbol{y}$
$\boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{mL})$
$\overline{\boldsymbol{x}} \leftarrow \boldsymbol{\mu} + \sigma_x\mathbf{F}^T\boldsymbol{u}$
% Échantillonnage de $\sigma_x^2$ : *a priori*, $\sigma_x^2 \sim \mathcal{IG}(\alpha_x, \beta_x)$
$\sigma_{x0}^2 \leftarrow \sigma_x^2$ % sauvergarde
$\sigma_x^2 \sim \mathcal{IG}(\alpha_x + \frac{1}{2}mL, \beta_x + \frac{1}{2}\|\overline{\boldsymbol{x}}\|^2)$
% Échantillonnage de $\sigma_\epsilon^2$ : *a priori*, $\sigma_\epsilon^2 \sim \mathcal{IG}(\alpha_\epsilon, \beta_\epsilon)$
$\sigma_\epsilon^2 \sim \mathcal{IG}(\alpha_\epsilon + \frac{1}{2}N, \beta_\epsilon + \frac{1}{2}\|\boldsymbol{y} - \overline{\mathbf{H}}\overline{\boldsymbol{x}}\|)$
% Mises à jour
$\mathbf{G} \leftarrow \frac{\sigma_x}{\sigma_{x0}}\mathbf{G}$
$\mathbf{\Sigma} \leftarrow \sigma_\epsilon^2\mathbf{I}_N$
$\mathbf{C} \leftarrow \mathbf{I}_{mL} + \mathbf{G}^T\mathbf{\Sigma}^{-1}\mathbf{G}$
$\mathbf{F} \leftarrow \text{chol}(\mathbf{C}^{-1})$

---

Nous disposons donc d'une implémentation efficace de l'échantillonneur de type PCGS. Nous donnons maintenant quelques conseils sur les réglages des lois *a priori* des hyperparamètres.

### 3.4.3 Conseils pour le réglage des lois *a priori* pour les hyperparamètres

Au cours des différents tests de notre échantillonneur PCGS, nous avons étudié différents réglages des lois *a priori* sur les hyperparamètres. Nous en avons déduit les recommandations suivantes sur le réglage des lois *a priori* sur les hyperparamètres $\lambda$ et $\sigma_x^2$.

#### Hyperparamètre $\lambda$ : un *a priori* non-informatif

Pour rappel, le paramètre $\lambda$ est la probabilité *a priori* des variables de Bernoulli $q_j$ d'être égales à 1, soit d'avoir une amplitude $\boldsymbol{x}_j$ non nulle : il s'agit de l'*a priori* influençant la parcimonie de la solution. Plus $\lambda$ est faible, plus la solution est parcimonieuse *a priori* ; une première intuition serait donc d'imposer à $\lambda$ un *a priori* favorisant les valeurs faibles. Dans le cadre du modèle BGE, $\lambda$ est aussi la probabilité *a priori* de l'existence de $\nu_j$ dans l'intervalle $\mathcal{I}_j$. Nous rappelons ici la loi *a priori* générique que nous donnons à l'hyperparamètre $\lambda$ ainsi que la loi *a posteriori* conditionnelle associée :

$$\begin{array}{lcl} (\lambda \,|\, \alpha, \beta) & \sim & \mathcal{B}e(\alpha, \beta) \\ (\lambda \,|\, \boldsymbol{y}, \boldsymbol{q}, \alpha, \beta) & \sim & \mathcal{B}e(L + \alpha, J - L + \beta),\ L = \|\boldsymbol{q}\| \end{array}$$

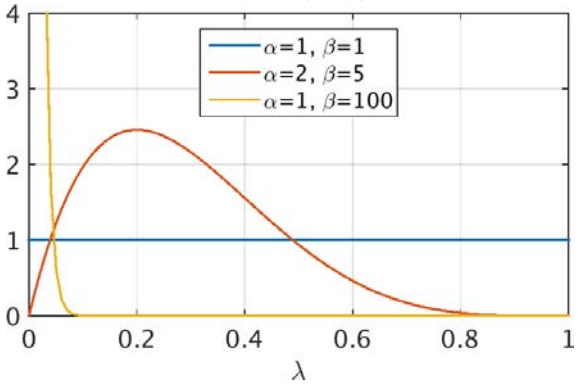
La Figure 3.6(a) donne l'allure des lois *a priori* sur $\lambda$ pour différentes valeurs de $(\alpha, \beta)$. On reconnaît notamment en bleu la loi uniforme sur $]0;1[$ pour $\alpha = \beta = 1$ ; cet *a priori* est non-informatif, dans le sens où il ne favorise aucune valeur de $\lambda$. Au contraire, les deux autres *a priori*, ($\alpha = 2$, $\beta = 5$) en rouge et ($\alpha = 1, \beta = 100$) en jaune, favorisent davantage les faibles valeurs de $\lambda$ et donc une solution plus parcimonieuse. L'influence de ces *a priori* s'observe sur les lois *a posteriori* conditionnelles associées, tracées en Figures 3.6(b-c-d) pour différentes valeurs de $L$ avec $J = 100$. On remarque que plus l'ordre du modèle est élevé ($L \to J$), plus la loi *a posteriori* favorise les $\lambda$ proches de 1 (et inversement). Ce comportement est cependant limité avec l'*a priori* ($\alpha = 1, \beta = 100$) en Figures 3.6(d) où la probabilité de simuler un $\lambda$ supérieur à 0.6 est presque nulle, même pour $L = 100$. Ce réglage de l'*a priori* semble donc cohérent avec un *a priori* de parcimonie de la solution[6].

Par ailleurs, le comportement classique de notre échantillonneur est de faire *naître* un grand nombre de $q_j = 1$ dès les premières itérations, menant rapidement à $L = \|\boldsymbol{q}\|$ élevé et donc à une faible parcimonie. Cela entraîne un ralentissement considérable du temps de calcul nécessaire pour une itération (les facteurs de Cholesky étant de dimension $mL \times mL$, les calculs associés et leur mise à jour sont donc plus coûteux). Limiter la valeur maximale de $\lambda$ grâce à un réglage adéquat de la loi *a priori* permet alors de limiter le degré maximal de parcimonie $L$, et donc d'éviter un ralentissement trop brutal de l'échantillonneur.

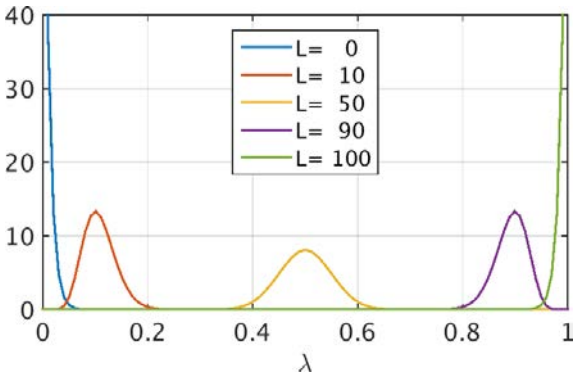
Cependant, ces *naissances* groupées en début d'échantillonnage permettent une exploration très efficace de l'espace des solutions, puisqu'un grand nombre de $q_j$ sont alors égaux à 1. Nous avons remarqué qu'en pratique, cette phase à faible parcimonie permettait une convergence plus rapide (en terme d'itérations) vers la loi cible. Par conséquent, limiter le degré maximal de parcimonie peut amener à une convergence plus lente de l'échantillonneur. Pour ces raisons, nous conseillons donc un *a priori* non-informatif sur l'hyperparamètre $\lambda$, comme la loi uniforme sur $]0;1[$.

[6] La loi Beta-rectangulaire permettrait de définir un support $[a;b]$ plus petit que $]0;1[$ pour $\lambda$, en interdisant notamment des valeurs supérieures à un certain seuil $b$. Cependant, cette loi n'est pas conjuguée avec la loi *a posteriori* en $\lambda$ et ne peut donc pas être utilisée par notre échantillonneur.

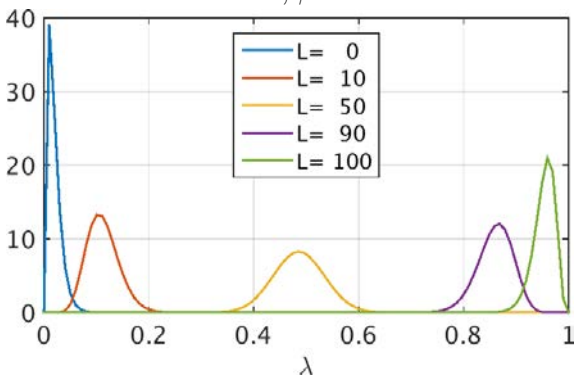
(a) Loi *a priori* $\lambda \sim \mathcal{B}e(\alpha, \beta)$




(b) Loi *a posteriori* conditionnelle $p(\lambda \,|\, -)$ $\alpha = 1,\ \beta = 1$




(c) Loi *a posteriori* conditionnelle $p(\lambda \,|\, -)$ $\alpha = 2,\ \beta = 5$




(d) Loi *a posteriori* conditionnelle $p(\lambda \,|\, -)$ $\alpha = 1,\ \beta = 100$

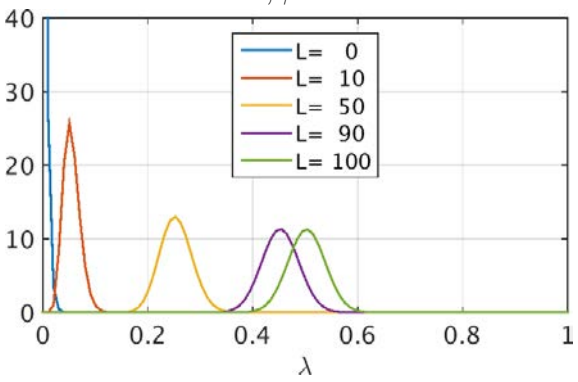



FIG. 3.6 – Allure des lois *a priori* et *a posteriori* conditionnelles pour $\lambda$. Fig.(a) allure de la loi *a priori* $\lambda \sim \mathcal{B}e(\alpha, \beta)$ pour différentes valeurs de $(\alpha, \beta)$. Fig.(b-d) chaque figure représente, pour un couple de paramètres *a priori* $(\alpha, \beta)$, l'allure de la loi *a posteriori* conditionnelle $\lambda \sim \mathcal{B}e(L+\alpha, J-L+\beta)$ pour différentes valeurs de $L = \sum_j q_j$ qui ici peut avoir pour valeur maximale $J = 100$.

**Variances $\sigma_x^2$ : choisir une loi inverse-gamma non dégénérée**

Rappelons les lois *a priori* et *a posteriori* conditionnelles que nous utilisons pour la variance du bruit $\sigma_x^2$, données dans le Tableau 3.3 p. 104 :

$$\begin{aligned}(\sigma_x^2 \,|\, \alpha_x, \beta_x) &\sim \mathcal{IG}(\alpha_x, \beta_x) \\ (\sigma_x^2 \,|\, \boldsymbol{q}, \boldsymbol{x}, \boldsymbol{y}, \alpha_x, \beta_x) &\sim \mathcal{IG}(\alpha_x + \frac{1}{2}mL, \beta_x + \frac{1}{2}\|\boldsymbol{x}\|^2)\end{aligned}$$

et étudions le réglage des paramètres $(\alpha_x, \beta_x)$ de la loi *a priori* inverse-gamma. Un réglage couramment utilisé est $\alpha_x = \beta_x = 0$, ce qui correspond à un loi non informative de Jeffreys pour une variable gaussienne : $p(\sigma_x^2) \propto \sigma_x^2$.

Cependant, dans notre cas, nous recommandons plutôt l'utilisation d'une loi inverse-gamma non-dégénérée $(\alpha_x, \beta_x > 0)$, pour deux principales raisons.

1. Si on choisit une loi *a priori* de Jeffreys $(\alpha_x = \beta_x = 0)$, alors la loi *a posteriori* conditionnelle en $\sigma_x^2$ est aussi une loi de Jeffreys dans le cas $L = \|\boldsymbol{q}\| = 0$ (*i.e.* toutes les amplitudes sont nulles). Cette loi est impropre à la simulation. Notons cependant que dans l'application à l'analyse spectrale que nous ferons en § 3.5.1, la présence de l'amplitude de la

fréquence nulle $x_0$ donnera une loi *a posteriori* conditionnelle propre, rendant possible l'échantillonnage avec un *a priori* de Jeffreys.

2. Comme on peut le voir en Figure 3.7, l'*a priori* de Jeffreys en $\sigma_x^{-2}$ autorise et même favorise des valeurs de la variance tendant vers 0, contrairement aux autres lois *a priori* inverse-gamma représentées. Cet *a priori* autorise donc la variance à prendre une valeur arbitrairement petite. Or, une variance arbitrairement petite pour les amplitudes entraîne leur faible influence sur la loi *a posteriori* ; en effet, un atome de faible amplitude peut être confondu avec le bruit. Nous avons d'ailleurs observé un comportement particulier de l'échantillonneur, quand les données ne contiennent que du bruit dans le cadre de l'analyse spectrale. Avec un *a priori* de Jeffreys, la variance $\sigma_x^2$ des amplitudes finit généralement par tendre vers 0 et presque l'intégralité des $q_j$ deviennent égaux à 1, donnant une solution avec quasiment aucune amplitude nulle, mais toutes très faibles. Nous avons aussi repéré ce comportement lorsque nous échantillonnons la variance du bruit. Ce comportement est lié au fait que nous échantillonnons la loi marginalisée en les amplitudes. Il nous semble donc important que la variance des amplitudes soit significative ; un atome $\mathbf{h}_j$ est *présent* dans le signal si son amplitude est significative et ne peut pas être confondue avec le bruit. Il s'agit d'un compromis, puisque nous voulons être capable de détecter des atomes de faible amplitude, mais ne pas les confondre avec le bruit. Le choix d'une loi *a priori* inverse-gamma non-dégénérée permet d'éviter à la variance $\sigma_x^2$ de tendre vers 0.

Nous illustrons nos divers conseils dans les tests que nous effectuons dans la suite de ce chapitre, à la fois pour $\lambda$ et pour $\sigma_x^2$.

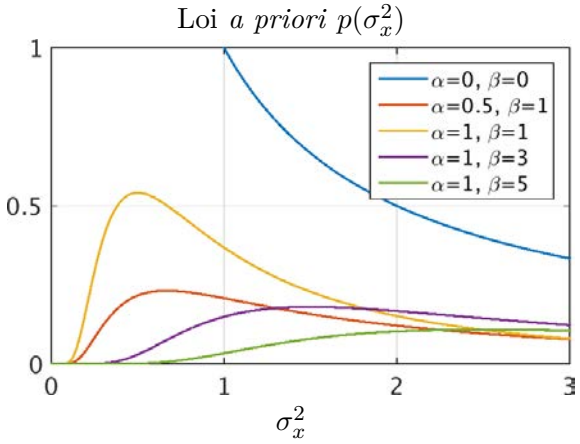



FIG. 3.7 – Allure de la loi *a priori* $\sigma_x^2 \sim \mathcal{IG}(\alpha, \beta)$ pour différentes valeurs de $\alpha, \beta$. Le couple $\alpha_x = \beta_x = 0$ (courbe en bleu) correspond à une loi de Jeffreys : $p(\sigma_x^2) \propto \sigma_x^{-2}$.

## 3.5 Application du modèle BGE et de l'échantillonneur associé aux données tests

Dans cette section, nous proposons de comparer les échantillonneurs sur les données tests présentées en § 1.5. Plus précisément, on compare l'échantillonneur de la loi jointe Gibbs-hybride proposé par [BC06] à l'échantillonneur exploitant la marginalisation des amplitudes pour l'échantillonnage des $\{q_j, \boldsymbol{\nu}_j\}_j$ que nous proposons.

*Notons que dans un souci de clarté, nous appellerons le premier « échantillonneur de la loi jointe » et le second « échantillonneur de la loi marginale », bien que cette dernière dénomination soit inexacte dans le cadre non supervisé, puisque le PCGS échantillonne la loi jointe.*

La majeure partie des résultats présentés dans cette section sont obtenus à partir des données tests en analyse spectrale (voir § 1.5.1). Nous mettons en avant une nouvelle stratégie de détection, dans le cas où des phénomènes d'*alternance/anti-corrélation* entre deux séquences de Bernoulli voisines apparaissent. Nous approfondissons les résultats présentés sur les mêmes données dans l'article de conférence [BCBB16], pour obtenir une meilleure compréhension du comportement de notre échantillonneur. Enfin, une étude statistique sur les résultats de l'application des échantillonneurs aux données tests de déconvolution impulsionnelle (voir § 1.5.2) sera réalisée, afin de montrer les atouts que ce type de méthode peut avoir vis-à-vis des méthodes déterministes.

### 3.5.1 Modèle BGE pour un problème d'analyse spectrale

Pour rappel, le signal test $\boldsymbol{y}$ pour l'analyse spectrale en échantillonnage irrégulier est constitué de la somme de $L = 5$ sinusoïdes et d'un bruit additif de RSB=5dB. Ses $N = 250$ échantillons, ainsi que la fenêtre spectrale associée, sont représentés en Figure 1.3 (p. 30). La grille de recherche est constituée de $J = 100$ intervalles $\mathcal{I}_j$, si bien que la largeur de l'intervalle $\Delta$ respecte la résolution induite par l'échantillonnage : on ne peut pas espérer trouver plus d'un pic dans chaque intervalle. Nous nous référons à la section 1.5.1 pour plus de détails. Nous donnons ici des détails spécifiques à cette application : le modèle choisi, l'adaptation à la présence de la composante continue, et le réglage des paramètres nécessaires.

#### Choix du modèle

En suivant la formulation des problèmes d'APNL, nous modélisons le signal $\boldsymbol{y}$ sous la forme :

$$\boldsymbol{y} = x_0 \mathbf{1}_N + \sum_{j=1}^{J} \mathbf{h}(\nu_j)\boldsymbol{x}_j + \boldsymbol{\epsilon}$$

Parmi les modèles de sinusoïdes équivalents (voir Tableau 1.1), nous choisissons le modèle en *cosinus et sinus*, soit $\mathbf{h}(\boldsymbol{\nu}_j) = [\cos(2\pi\nu_j \boldsymbol{t}) \ \sin(2\pi\nu_j \boldsymbol{t})]$, qui a l'avantage de n'utiliser que des variables réelles et de ne présenter qu'un seul paramètre non linéaire : la fréquence. Les amplitudes sont donc bi-dimensionnelles ($m = 2$) et il n'y a pas de paramètre de forme ($p = 1$). De plus, nous considérons que le bruit est *i.i.d.* de variance $\sigma_\epsilon^2$.

Les paramètres (et hyperparamètres) du modèle BGE pour notre modèle d'analyse spectrale sont donc $(\boldsymbol{q}, x_0, \boldsymbol{x}, \boldsymbol{\nu}, \lambda, \sigma_x^2, \sigma_\epsilon^2)$. Les variables de Bernoulli $q_j$ indiquent la présence de la fréquence $\nu_j$ d'amplitude $\boldsymbol{x}_j$ dans l'intervalle $\mathcal{I}_j$. Mis à part la présence de la composante continue $x_0$ dont nous parlons au paragraphe suivant, le modèle BGE étendu et les échantillonneurs associés s'appliquent donc simplement à ce problème d'analyse spectrale.

#### Adaptation à la présence de la composante continue

Une particularité majeure est le besoin d'estimer la composante continue $x_0$ du signal $\boldsymbol{y}$ :

$$\boldsymbol{y} = x_0 \mathbf{1}_N + \sum_{j=1}^{J} \mathbf{h}(\nu_j)\boldsymbol{x}_j + \boldsymbol{\epsilon}$$

En pratique, on travaille sur un signal dont on a retranché la valeur moyenne. Cependant, l'estimation de cette valeur moyenne peut être erronée à cause de l'échantillonnage irrégulier et il est donc nécessaire de conserver la composante continue $x_0$ dans le modèle.

Dans le cas de l'échantillonneur Gibbs-hybride de la loi jointe, soit $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, x_0, \boldsymbol{\theta} \,|\, \boldsymbol{y})$, [Bou05] propose d'attribuer à la composante continue une loi *a priori* gaussienne centrée de variance $\sigma_x^2$ identique à celle des autres amplitudes, soit $x_0 \sim \mathcal{N}(0, \sigma_x^2)$. La loi *a posteriori* conditionnelle $p(x_0 \,|\, \boldsymbol{y}, -)$ est alors une loi gaussienne qui peut s'échantillonner directement au sein de l'algorithme de Gibbs.

Dans le cas de l'échantillonnage que nous proposons, exploitant la marginalisation des amplitudes, nous gardons la même loi *a priori* sur la composante continue, soit $x_0 \sim \mathcal{N}(0, \sigma_x^2)$. En associant à $x_0$ l'atome $\mathbf{h}_0 = \mathbf{1}_N$ et en réécrivant le vecteur d'amplitude comme $\boldsymbol{x} = [x_0, \boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_J^T]^T$, la marginalisation de la loi cible $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ suivant $\boldsymbol{x}$ (qui contient maintenant $x_0$) donne le même résultat que celui donné en équation (3.18), à la différence près que le dictionnaire réduit $\overline{\mathbf{H}}$ contient l'atome $\mathbf{h}_0 = \mathbf{1}_N$ et est donc maintenant de dimension $(N \times 2L{+}1)$. En prenant en compte cette modification du dictionnaire réduit, l'ensemble des calculs pour l'échantillonnage des paramètres $q_j$ et $\nu_j$ selon leur loi conditionnelle marginalisée reste inchangé. En conséquence, l'étape d'échantillonnage de $\overline{\boldsymbol{x}}$ précédant l'échantillonnage des hyperparamètres de variances lors du PCGS est aussi inchangée ; $\overline{\boldsymbol{x}}$ contiendra toujours la composante continue $x_0$.

Enfin, dans les deux cas, la loi *a posteriori* conditionnelle en $\sigma_x^2$ servant à son échantillonnage dans le cadre non supervisé est maintenant toujours propre, même si on utilise une loi *a priori* de Jeffreys pour $\sigma_x$, puisque la présence de la composante continue permet d'écrire :

$$(\sigma_x^2 \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{y}) \;\sim\; \mathcal{IG}(\underbrace{\alpha_x + \tfrac{1}{2}(2L+1)}_{\geq \alpha_x + 1/2}, \underbrace{\beta_x + \tfrac{1}{2}\|\boldsymbol{x}\|^2}_{\geq \beta_x + x_0^2/2})$$

**Réglage des paramètres nécessaires**

Finalement, pour les deux échantillonneurs du modèle BGE, il est nécessaire de régler les paramètres en lien avec la loi de proposition $Q(\nu_j' \,|\, \boldsymbol{\nu}_j^{(t)})$ des fréquences $\nu_j'$ lors de l'étape MH. Pour rappel, cette dernière est une loi de mélange entre une marche aléatoire gaussienne de variance instrumentale $\sigma_{\text{rw}}^2$ avec une probabilité $p_{\text{rw}}$, et une simulation suivant la loi *a priori* $\mathcal{U}(\mathcal{I}_j)$ avec une probabilité de $(1 - p_{\text{rw}})$ (voir équation (3.16)). Nous donnons donc les valeurs choisies pour $\sigma_{\text{rw}}$ et $p_{\text{rw}}$ :

$$\sigma_{\text{rw}} = 0.1\Delta \quad \text{et} \quad p_{\text{rw}} = 0.5$$

*Dans le cadre supervisé,* nous fixons les hyperparamètres $(\lambda, \sigma_x^2, \sigma_\epsilon^2)$ aux valeurs suivantes :

$$\lambda_0 = 0.05 \qquad \sigma_{x,0}^2 = 0.25 \qquad \sigma_{\epsilon,0}^2 = 0.415$$

La variance du bruit choisie $\sigma_{\epsilon,0}^2$ est la vraie variance utilisée pour simuler le bruit du signal. Les valeurs choisies pour $\lambda_0$ et $\sigma_{x,0}^2$ se basent sur des considérations statistiques ; $\lambda_0$ peut être approchée par $L/J$, et $\sigma_{x,0}^2$ peut être approchée par la variance empirique des amplitudes théoriques[7].

*Dans le cadre non supervisé,* nous utilisons plusieurs réglages pour les lois *a priori* sur les hyperparamètres. Pour différencier ces différents réglages, nous utilisons une dénomination à 3 chiffres de type « *** » où chaque chiffre indique quel réglage est utilisé pour l'hyperparamètre concerné dans l'ordre $\lambda$, $\sigma_x^2$ et $\sigma_\epsilon^2$. De façon générale, « 0 » signifiera que le paramètre est fixé, « 1 » sera relié à un *a priori* non-informatif, et « 2 » à un *a priori* plus informatif. Les lois employées suivant le code « *** » sont davantage détaillées dans le Tableau 3.11.

[7] En ne prenant pas en compte $x_0$ dans les calculs, alors $\sigma_{x,0}^2 \approx \sum_{\ell=1}^{5} \boldsymbol{x}_\ell(1)^2 + \boldsymbol{x}_\ell(2)^2/(L-1) = 0.29$. Une approche proposée par [Bou05] donne $\sigma_{x,0}^2 \approx LP_s = 0.22$, où $P_s$ est la puissance moyenne du signal non bruité. nous avons donc arrondi à 0.25.

| Valeur de x | 0 | 1 | 2 |
|---|---|---|---|
| x** : $\lambda$ | $\lambda = \lambda_0$ | $\lambda \sim \mathcal{U}([0;1[)$ | $\lambda \sim \mathcal{B}e(1,100)$ |
| *x* : $\sigma_x^2$ | $\sigma_x^2 = \sigma_{x,0}^2$ | $\sigma_x^2 \sim \mathcal{IG}(0,0)$ | $\sigma_x^2 \sim \mathcal{IG}(0.5,1)$ |
| **x : $\sigma_\epsilon^2$ | $\sigma_\epsilon^2 = \sigma_{\epsilon,0}^2$ | $\sigma_\epsilon^2 \sim \mathcal{IG}(0,0)$ | $\times$ |

TAB. 3.11 – Récapitulatif des réglages de lois *a priori* des hyperparamètres selon le code « *** ».

Dans tous les cas, les échantillonneurs sont lancés sur 10000 itérations, et nous conservons les 5000 dernières pour effectuer nos estimations.

### 3.5.2 Étude d'une réalisation typique des échantillonneurs et proposition d'une nouvelle stratégie d'estimation

En Figure 3.8, nous observons une réalisation typique dans le *cadre supervisé* des échantillonneurs de la loi jointe et de la loi marginalisée. Nous faisons les remarques suivantes :

Fig.(a) : Nous observons les échantillons des séquences de Bernoulli $\boldsymbol{q}^{(t)}$, centrées autour des indices $j$ des fréquences de simulation[8]. Dans les deux cas, on obtient généralement $q_j^{(t)} = 1$ autour des indices théoriques. Quand les fréquences de simulation sont proches d'une séparation entre deux intervalles (comme c'est le cas ici des deux premières fréquences), les échantillons de la loi jointe ont tendance à rester bloqués dans un seul intervalle, ce dernier n'étant pas forcément le bon ; notre illustration en est un exemple ($q_{46}^{(t)} = 1$ alors que la fréquence théorique est en $j = 45$). Au contraire, les échantillons de la loi marginale passent plus aisément de la configuration $(q_{45} = 1, q_{46} = 0)$ à $(q_{46} = 0, q_{45} = 1)$. On parlera alors d'un phénomène d'*alternance* ou d'*anti-corrélation*.

Fig.(b) : Ce zoom au niveau des premiers échantillons de la séquence de Bernoulli nous permet d'illustrer le comportement classique des échantillonneurs : de nombreuses « naissances » lors des premières itérations, suivies de « morts ». Notons que le retour à un $\|\boldsymbol{q}^{(t)}\|$ bas est plus rapide pour l'échantillonneur marginalisé que pour l'échantillonneur joint.

Fig.(c) : Ici nous observons les estimations au sens de l'EAP de la séquence de Bernoulli. Dans le cas de l'échantillonneur de la loi jointe, 5 fréquences sont bien détectées avec une probabilité supérieure à $\alpha = 95\%$ ; cependant, la première se trouve dans un intervalle voisin. Dans le cas de l'échantillonneur de la loi marginale, l'évolution des fréquences dans deux intervalles voisins pour les deux premières met en échec leur détection par le seuillage $\alpha = 0.95$, puisqu'on obtient une estimation de $\widehat{\boldsymbol{q}}$ proche de 50% dans les deux intervalles.

Fig.(d) : Cette figure montre les fréquences estimées dans des intervalles de largeur $\Delta$ centrés sur les fréquences théoriques. Les barres d'erreur à $\pm\widehat{\sigma}_{\nu_{j,\alpha}}$, qui peuvent être calculées via l'équation (3.14), sont représentées, ainsi que les histogrammes[9] des échantillons utilisés pour les calculer. L'estimation conditionnellement aux échantillons tels que $\boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha$ met en échec l'estimation des 3 fréquences détectées pour le cas de l'échantillonneur de la loi marginale, car pour presque tous les échantillons $(t)$, $(q_{40}$ ou $q_{41})$ et $(q_{45}$ ou $q_{46})$ sont égaux à 1. En définissant $\overline{\boldsymbol{q}}_\alpha$ la réduction de $\boldsymbol{q}$ aux indices tels que $\widehat{q}_{j,\alpha} = 1$, on peut cependant définir l'estimation conditionnellement aux échantillons tels que $\overline{\boldsymbol{q}}_\alpha^{(t)} = 1$, qui permet de détecter les 3 premières fréquences (en bleu clair).

[8] Les $\widehat{\boldsymbol{q}}$ estimés sont proches de zéro pour la plupart des intervalles sauf autour des fréquences de simulation.
[9] Ces histogrammes ont été normalisés pour que la plus grande valeur vaille 1.

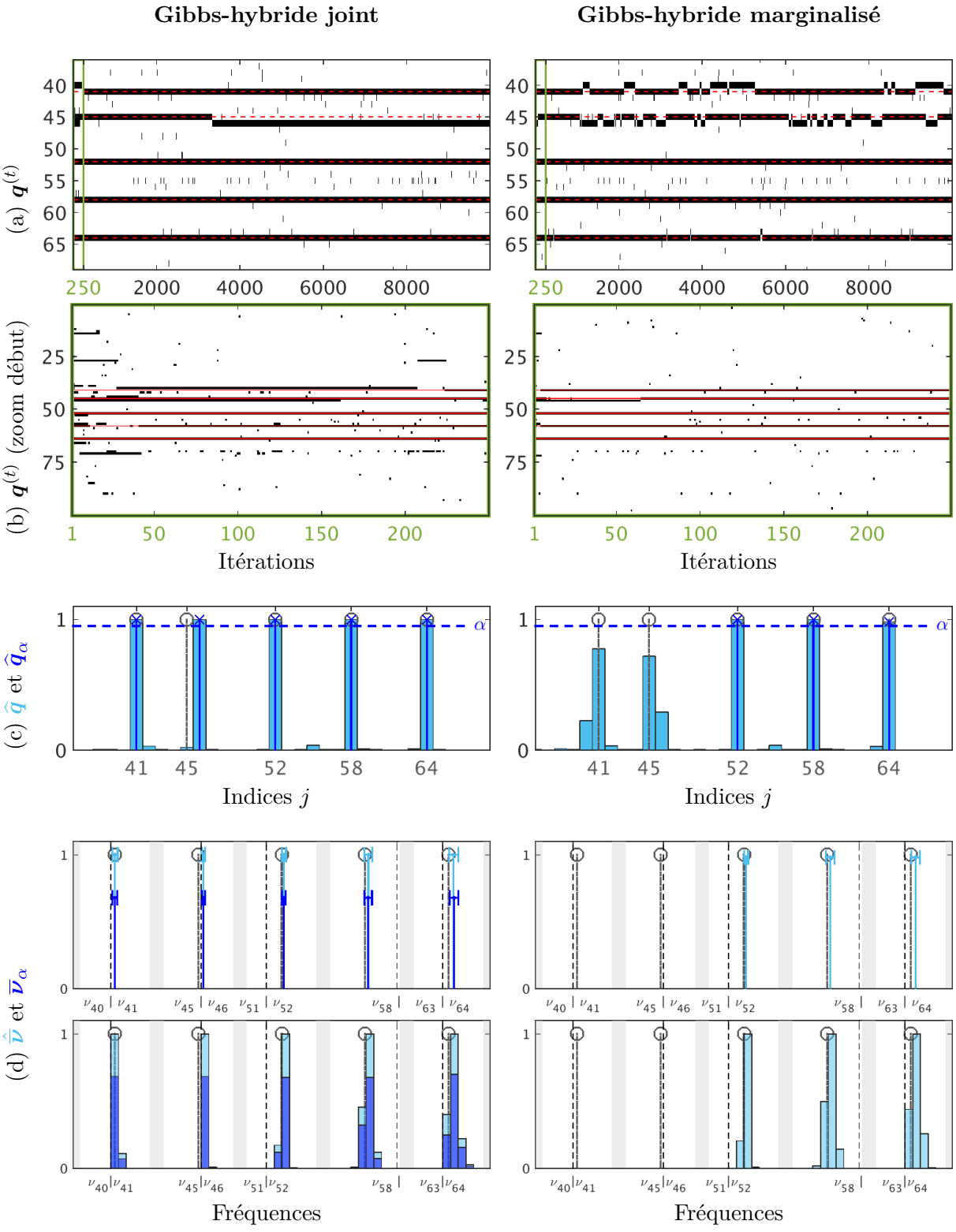



FIG. 3.8 – Comparaison des échantillonneurs Gibbs-hybride joint (gauche) et Gibbs-hybride marginalisé (droite) dans le cadre supervisé. (a) Échantillons $\boldsymbol{q}^{(t)}$ sur les $T$ =10000 itérations, zoomé autour des indices (ordonnées) théoriques. En noir : $q_j^{(t)} = 1$ ; en blanc $q_j^{(t)} = 0$ ; en rouge indices théoriques. (b) Échantillons $\boldsymbol{q}^{(t)}$ sur les 250 premières itérations, pour tous les indices. (c) Estimation de $\widehat{\boldsymbol{q}}$ et de $\widehat{\boldsymbol{q}}_\alpha$ avec $\alpha = 0.95$. (d) Estimation des fréquences : $\widehat{\overline{\boldsymbol{\nu}}}_\alpha$ conditionnellement à $\widehat{\boldsymbol{q}}_\alpha$, et $\widehat{\overline{\boldsymbol{\nu}}}$ cond. à $\overline{\boldsymbol{q}}^{(t)} = 1$, avec barre d'erreur à $\pm\widehat{\sigma}_{\nu_{j,\alpha}}$ (en haut) et histogramme des échantillons utilisés (en bas).

Ces premières observations semblent plutôt favoriser l'échantillonnage de la loi jointe, puisqu'elle permet de détecter les 5 fréquences, même si une se trouve dans un intervalle voisin de la fréquence théorique. Cependant, même si l'estimation de $\widehat{\boldsymbol{q}}$ dans le cas marginal suggère la possible existence de 4 fréquences aux indices (40,41,45,46) avec une faible probabilité (inférieure à $\alpha = 0.95$), le comportement d'*alternance* des variables de Bernoulli $(q_{40}^{(t)}, q_{41}^{(t)})$ et $(q_{45}^{(t)}, q_{46}^{(t)})$ suggère qu'il en existe seulement deux, la première oscillant entre les intervalles $(\mathcal{I}_{40}, \mathcal{I}_{41})$ et la seconde entre $(\mathcal{I}_{45}, \mathcal{I}_{46})$. On peut d'ailleurs vérifier en Figure 3.9 que les fréquences associées à chaque intervalle se trouvent très proches de la séparation entre les deux intervalles.

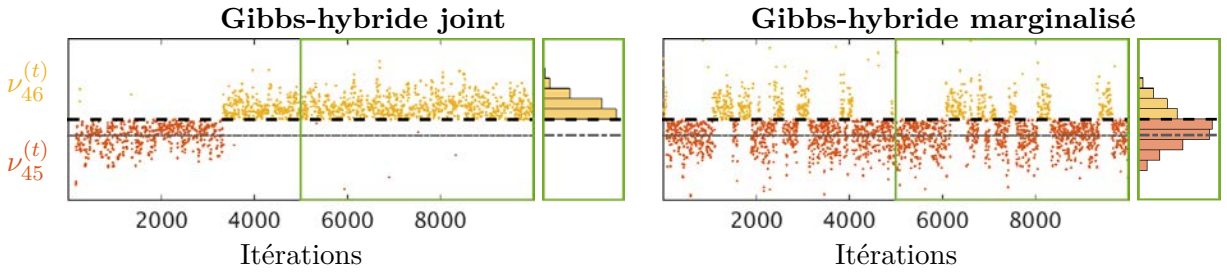



FIG. 3.9 – Échantillons d'une fréquence dans la cadre supervisé : $\nu_{45}^{(t)}$ et $\nu_{46}^{(t)}$. En tirets : la séparation de la grille. En pointillés : la position de la fréquence théorique. À droite de chaque figure est représenté l'histogramme des échantillons sur les 5000 dernières itérations.

Dans le cas où des estimations $(\widehat{q}_j; \widehat{q}_{j+1})$ voisines donnent des probabilités basses, il est donc nécessaire de pouvoir distinguer le cas où il s'agit de deux fréquences de faibles probabilités, ou d'une seule fréquence dont la valeur varie dans les deux intervalles voisins. L'estimation $\widehat{\boldsymbol{q}}$ n'est pas suffisante ; nous nous basons alors sur ces deux critères pour détecter le cas où une seule fréquence existe :

1. les séquences $q_j^{(t)}$ et $q_{j+1}^{(t)}$ doivent être suffisamment *anti-corrélées*, *i.e.* elles doivent montrer le comportement d'*alternance* entre les deux configurations où l'une est nulle quand l'autre ne l'est pas.

2. les fréquences associées $(\nu_j^{(t)}, \nu_{j+1}^{(t)})$ doivent être suffisamment proches de la séparation entre les deux intervalles.

À partir de ces observations, nous proposons la nouvelle stratégie d'estimation proposée en Tableau 3.12.

La Figure 3.10 illustre les nouvelles estimations de la séquence de Bernoulli et des fréquences grâce à cette nouvelle stratégie pour les échantillons de la loi marginale : les 5 fréquences sont cette fois-ci correctement détectées, et la barre d'erreur à $\pm\sigma$ comprend la fréquence théorique à chaque fois. Ce comportement d'*alternance* entre deux intervalles voisins est naturel quand la fréquence théorique se trouve près d'une séparation d'intervalle, puisque les échantillons d'une fréquence peuvent se retrouver dans l'un ou l'autre des intervalles. L'échantillonneur de la loi marginale donne donc de meilleurs résultats sur cet exemple.

1. Pour tous les $j \in [\![1; J-1]\!]$, on calcule pour les $T$ échantillons,

$$q^{(t)}_{(j,j+1)} = q^{(t)}_j \oplus q^{(t)}_{j+1} = q^{(t)}_j + q^{(t)}_{j+1} - q^{(t)}_j q^{(t)}_{j+1} \in \{0,1\}$$

2. Si le rapport de la somme des $q^{(t)}_{(j,j+1)}$ (nombre d'échantillons anticorrélés) sur le nombre total d'échantillons $T$ est supérieur à $\alpha$, les intervalles $\mathcal{I}_j$ et $\mathcal{I}_{j+1}$ sont fusionnés pour donner $\mathcal{I}_{(j,j+1)} = \mathcal{I}_j \cup \mathcal{I}_{j+1}$, associé à la séquence $q^{(t)}_{(j,j+1)}$ et à la fréquence $\nu^{(t)}_{(j,j+1)}$ uniquement définie quand $q^{(t)}_{(j,j+1)} = 1$ :

$$\text{si } q^{(t)}_{(j,j+1)} = 1 \quad : \quad \nu^{(t)}_{(j,j+1)} \quad = \quad \nu^{(t)}_i \,|\, q^{(t)}_i = 1,\ i \in \{j, j+1\}$$

Les estimateurs sont ensuite calculés suivant la procédure habituelle, mais sur un nombre possiblement restreint d'intervalles : estimation de $\widehat{\boldsymbol{q}}$, seuillage pour obtenir $\widehat{\boldsymbol{q}}_\alpha$ et estimation des fréquences conditionnellement à $\widehat{\boldsymbol{q}}_\alpha$.

TAB. 3.12 – Nouvelle stratégie d'estimation pour l'échantillonneur de Gibbs marginalisé : fusion des intervalles anticorrélés.

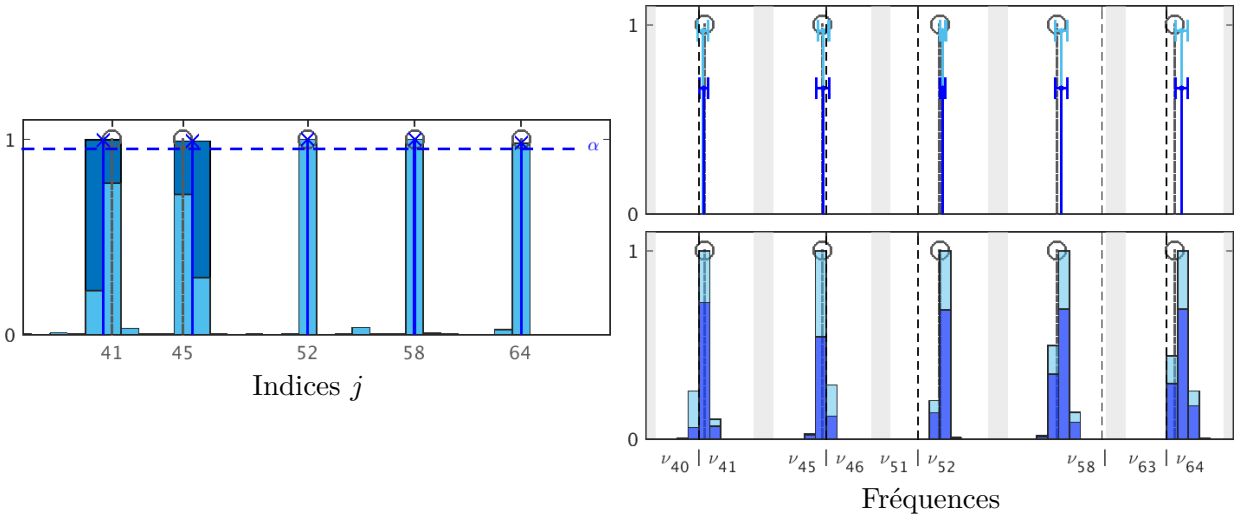



FIG. 3.10 – Résultats de la nouvelle stratégie d'estimation à partir d'échantillons de la loi marginale. À gauche : estimation de la séquence de Bernoulli - en bleu clair $\widehat{\boldsymbol{q}}$ et en bleu foncé $\widehat{\boldsymbol{q}}$ après fusion des intervalles anti-corrélés. À droite : estimation des fréquences conditionnellement à $\widehat{\boldsymbol{q}}_\alpha$ en bleu foncé et conditionnellement aux $\overline{\boldsymbol{q}}^{(t)}_\alpha = 1$ en bleu clair (en haut), et histogrammes des échantillons ayant servi aux estimations (en bas).

### 3.5.3 Comparaison des échantillonneurs sur 50 graines

Afin de confirmer les observations faites précédemment, nous avons lancé chacun des deux échantillonneurs avec 50 initialisations différentes de la graine aléatoire (nous parlerons de 50 graines), dans le *cadre supervisé* et *non supervisé*, en utilisant le réglage non-informatif des lois *a priori* « 111 ».

Dans ce but, nous caractérisons de « Correct Estimate » (Estimation correcte/CE) un résultat tel que :

1. toutes les fréquences simulées sont détectées avec une probabilité supérieure à 95% (soit à partir de $\widehat{\boldsymbol{q}}_\alpha$ avec $\alpha = 0.95$ estimé sur les 5000 dernières itérations et après possible fusion des intervalles anti-corrélés exploitant la procédure en Tableau 3.12) ;
2. aucune erreur en fréquence $|\nu_\ell^{\text{th}} - \widehat{\nu}_\ell|$ ne dépasse $2\sigma$, où $\sigma^2 = \widehat{\sigma}^2_{\nu_{\ell,\alpha}}$ est la variance estimée correspondant à la fréquence $\widehat{\nu}_\ell$.

Dans le Tableau 3.13, différentes quantités illustrent le comportement moyen de nos échantillonneurs sur les 50 graines.

| | Jointe | | Marginale | |
|---|---|---|---|---|
| | Supervisé | Non-sup. | Supervisé | Non-sup. |
| CE : pourcentage | 74 | 72 | 100 | 100 |
| CE : itérations | 2919 | 2636 | 328 | 284 |
| CE : temps (en s) | 38 | 36 | 6 | 5 |
| CE : MSE ($\times 10^{-8}$) | 1.24 | 1.38 | 1.29 | 1.21 |
| Temps total (en s) | 131 | 135 | 183 | 192 |
| Estimation $\widehat{\lambda}$ | 0.050 | 0.067 | 0.050 | 0.067 |
| Estimation $\widehat{\sigma}_x$ | 0.500 | 0.448 | 0.500 | 0.449 |
| Estimation $\widehat{\sigma}_\epsilon$ | 0.644 | 0.645 | 0.644 | 0.645 |

TAB. 3.13 – Comparaison des résultats des échantillonneurs de la loi jointe et marginalisée sur 50 initialisations différentes. Les valeurs sont données en moyenne sur les 50 initialisations différentes. La première ligne du tableau donne le pourcentage de CE, puis les suivantes donnent, en moyenne sur les chaînes donnant un CE, le nombre d'itérations et le temps de calcul associé nécessaire pour l'atteindre (calculés en déterminant s'il y a CE sur 100 itérations toutes les 100 itérations), ainsi que l'erreur quadratique moyenne (MSE) des fréquences estimées. Le temps total correspond au temps moyen mis pour effectuer les 10000 itérations.

D'abord, l'échantillonneur de la loi marginale donne 100% de CE tandis que celui de la loi jointe échoue 25% du temps, que cela soit dans le cadre supervisé ou non. En effet, la tendance de ce dernier à rester bloqué dans des modes locaux, voisins des intervalles théoriques, peut donner une erreur trop grande pour l'estimation de la fréquence, comme nous l'avons vu en Figure 3.8. Ensuite, le temps ainsi que le nombre d'itérations nécessaires pour atteindre le CE sont beaucoup plus faibles pour l'échantillonneur de la loi marginale. L'erreur de précision est d'ailleurs légèrement plus faible pour ce dernier dans le cas non supervisé. Enfin, les hyperparamètres estimés dans le cadre non supervisé sont très similaires quel que soit l'échantillonneur ;

ils sont aussi très proches des valeurs données dans le cadre supervisé. On en conclut donc que le choix de l'échantillonneur de la loi marginale converge plus rapidement (en nombre d'itérations et en temps de calcul) et donne de meilleurs résultats.

### 3.5.4 Étude du réglage des lois *a priori* des hyperparamètres

En observant uniquement les résultats de l'échantillonneur de la loi marginale, nous étudions maintenant l'influence du réglage des lois *a priori* des hyperparamètres. Dans le Tableau 3.14, on peut d'abord remarquer que quel que soit le réglage employé, le taux de CE est de 100%. Pour une loi *a priori* en $\lambda$ favorisant la parcimonie (réglages « 2** »), le nombre d'itérations et le temps nécessaires pour atteindre le CE sont légèrement plus élevés que pour une loi non informative (réglages « 1** »). Cela confirme nos indications données en § 3.4.3 : une loi uniforme est préférable à un *a priori* fort de parcimonie. De même, une loi *a priori* en $\sigma_x^2$ non dégénérée (réglages « *2* ») permet d'atteindre le CE légèrement moins rapidement que la loi *a priori* de Jeffreys (réglages « *1* »). Néanmoins, l'utilisation du réglage (« *2* ») permet d'éviter à la variance $\sigma_x^2$ de devenir trop faible tout en restant efficace. Dans les deux cas, nos indications sur le réglage des lois *a priori* sur les hyperparamètres (de type « 12* ») se montreront plus utiles dans le cas de l'application à la détection d'exoplanètes au Chapitre 4, qui est un problème bien plus difficile que ce problème test d'analyse spectrale.

Il est intéressant de comparer les valeurs des hyperparamètres estimés suivant le réglage. Les réglages de type (« 2** ») donnent un $\widehat{\lambda}$ plus faible que (« 1** ») ; les réglages de type (« *2* ») donnent un $\widehat{\sigma}_x^2$ plus élevé que (« *1* »). Cela est en concordance avec la loi *a priori* associée. Notons que l'estimation de la variance du bruit, toujours très proche de la variance de simulation, est très similaire quel que soit le réglage des lois *a priori* pour les autres hyperparamètres.

| | Réglages des hyperparamètres | | | | |
|---|---|---|---|---|---|
| | « 000 » | « 111 » | « 211 » | « 121 » | « 221 » |
| CE : pourcentage | 100 | 100 | 100 | 100 | 100 |
| CE : itérations | 328 | 284 | 328 | 332 | 490 |
| CE : temps (en s) | 6 | 5 | 7 | 6 | 9 |
| CE : MSE ($\times 10^{-8}$) | 1.29 | 1.21 | 1.33 | 1.25 | 1.26 |
| Temps total (en s) | 183 | 192 | 218 | 177 | 174 |
| Estimation $\widehat{\lambda}$ | 0.050 | 0.067 | 0.031 | 0.063 | 0.031 |
| Estimation $\widehat{\sigma}_x$ | 0.500 | 0.449 | 0.466 | 0.615 | 0.625 |
| Estimation $\widehat{\sigma}_\epsilon$ | 0.644 | 0.645 | 0.647 | 0.646 | 0.647 |

Tab. 3.14 – Comparaison des résultats des échantillonneurs de la loi marginalisée sur 50 initialisations pour différents réglages des lois *a priori* sur les hyperparamètres. Les valeurs sont données en moyenne sur les 50 initialisations différentes. La première ligne du tableau donne le pourcentage de CE, puis les suivantes donnent, en moyenne sur les chaînes donnant un CE, le nombre d'itérations et le temps de calcul associé nécessaire pour l'atteindre, ainsi que l'erreur quadratique moyenne (MSE) des fréquences estimées. Le temps total correspond au temps moyen mis pour effectuer les 10000 itérations.

Enfin, nous avons remarqué que le phénomène d'*alternance* (entre des $q_j^{(t)}$ voisins quand la fréquence théorique se trouve entre les deux intervalles) se retrouve renforcé dans le cas d'un échantillonnage non supervisé, comme on peut le voir en Figure 3.11(a) (réglage « 111 »), à comparer avec la Figure 3.8(a-droite) dans le cas supervisé. La Figure 3.11(b) montre aussi l'histogramme des « échantillons » de l'ordre du modèle[10] $L^{(t)}$ pour les différents réglages des lois *a priori* sur les hyperparamètres « 000-111-211-121-221 ». Il est intéressant de remarquer que pour les réglages qui favorisent *a priori* la parcimonie (tous sauf « 111 »), $L^{(t)} = 5$ occupe plus de la majorité des itérés ; tandis que pour le réglage le moins informatif « 111 », le pourcentage de $L^{(t)} = 5$ passe sous la barre des 50% et le pourcentage $L^{(t)} = 6$ augmente. Par ailleurs, cela montre l'importance d'étudier les échantillons $\boldsymbol{q}^{(t)}$ plutôt que ceux de l'ordre du modèle $L^{(t)} = \|\boldsymbol{q}^{(t)}\| = \|\boldsymbol{x}\|_0$, puisque les premiers sont beaucoup plus informatifs en terme de détection.

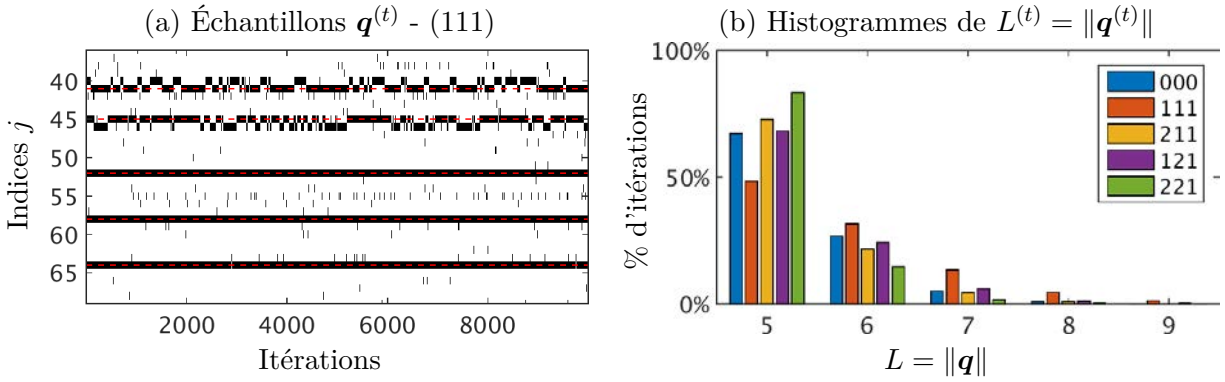



FIG. 3.11 – Échantillons de la séquence de Bernoulli selon la loi marginale dans le cadre non-supervisé et histogramme de $L^{(t)} = \|\boldsymbol{q}^{(t)}\|$ pour plusieurs réglages des lois *a priori* des hyperparamètres.

### 3.5.5 Application sur les données de déconvolution impulsionnelle

Pour compléter notre étude, nous avons aussi testé nos échantillonneurs sur les données tests en déconvolution impulsionnelle décrites en § 1.5.2. Pour rappel, elles sont composées de 200 modèles tests de $L = 4$ échos aléatoirement positionnés, avec 3 niveaux de bruit différents.

Pour chacun de ces 3×200 signaux, nous avons mis en œuvre les échantillonneurs de Gibbs-hybride pour la loi jointe et la loi marginale, dans le cadre non supervisé en prenant des *a priori* de type (réglage de type « 121 »)[11] et ce sur 10000 itérations. L'estimation de la séquence $\widehat{\boldsymbol{q}}_\alpha$ est réalisée pour 2 valeurs de $\alpha$ différentes : 0.8 et 0.95.

Pour comparer avec les résultats obtenus au Chapitre 2, nous avons calculé pour chaque méthode et chaque niveau de détection $\alpha$ le pourcentage d'ELR (« Estimation exacte des localisations » défini en § 1.5.2) ; pour rappel, cet indice est égal à 1 uniquement si $L = 4$ échos sont retrouvés et si les erreurs de localisations avec les échos théoriques sont inférieures à $\Delta/2$. Le Tableau 3.15 récapitule ces résultats.

[10] L'ordre du modèle est en fait calculé à chaque itération à partir de l'échantillon $\boldsymbol{q}^{(t)}$ : $L^{(t)} = \|\boldsymbol{q}^{(t)}\|$.

[11] Notons que la loi *a priori* de Jeffreys pour la variance des amplitudes $\sigma_x^2$ (« *1* ») est impropre pour l'échantillonnage ; en pratique, nous avons pu la mettre en œuvre et elle n'a pas posé de problème ; les résultats sont légèrement moins bon qu'avec une loi plus informative pour ces données.

| | | Loi Jointe | | Loi Marginale | |
|---|---|---|---|---|---|
| | $\alpha$ | 0.80 | 0.95 | 0.80 | 0.95 |
| RSB | 10 | 61 | **47.5** | **68.5** | 45.5 |
| | 20 | 57.5 | 56 | **97.5** | **96** |
| | 30 | 10 | 11 | **99.5** | **99.5** |

TAB. 3.15 – Test sur les données de déconvolution : comparaison des pourcentages d'ELR sur 200 signaux entre les échantilonneurs de la loi jointe et marginale dans le cadre non supervisé. Les séquences de Bernoulli $\widehat{\boldsymbol{q}}_\alpha$ sont détectées avec deux seuils $\alpha$ : 0.8 et 0.95. Les résultats sont observés pour 3 niveaux de bruits différents. Les pourcentages en gras mettent en lumière quelle méthode donne le meilleur pourcentage d'ELR (entre loi jointe et marginale).

Notons d'abord que le niveau de détection $\alpha = 0.95$ est ici trop strict ; $\alpha = 0.8$ donne de meilleurs résultats. Dans tous les cas, le pourcentage d'ELR est meilleur pour l'échantillonneur de la loi marginale que pour celui de la loi jointe pour $\alpha = 0.8$, notamment pour des niveaux de bruits moins élevés (RSB=20-30dB).

Le comportement de l'échantillonneur de la loi jointe pour un bruit très faible (RSB=30dB) semble contre-intuitif, puisque les résultats ne dépassent pas 20% d'ELR (contre 99% pour la loi marginale). Une observation des échantillons $\boldsymbol{q}^{(t)}$ nous montre que ces mauvaises estimations sont dues au blocage dans des modes locaux, comme évoqué précédemment, mais aussi à des variables de Bernoulli voisines restant toutes égales à 1 alors qu'une seule le devrait, ce qui est sûrement dû à une loi *a posteriori* trop piquée. Un tel comportement s'observe uniquement dans le cas de la loi jointe, car les amplitudes de ces $\{q_j\}_j$ voisins peuvent être très faibles sans pour autant faire passer $q_j$ à 0

Dans le cas d'un niveau de bruit plus élevé (RSB=10dB), les résultats entre loi jointe et loi marginale sont à peu près équivalents. Les cas d'échec se produisent notamment quand les localisations des échos sont trop proches. Nous avons aussi observé des cas où deux paramètres de Bernoulli $q_{j_1}^{(t)}$ et $q_{j_2}^{(t)}$ *non-voisins* présentaient le même comportement anti-corrélé que celui observé dans le cas de fréquences voisines précédemment. De telles alternances mettent en exergue des lois bimodales, pour lesquelles l'une ou l'autre des localisations peut être présente presque sûrement. Nous reviendrons plus en détails et illustrerons ce phénomène lors de nos tests sur la détection d'exoplanète dans le chapitre suivant.

Enfin, nous pouvons comparer ces résultats avec ceux obtenus dans le Chapitre 2 avec la résolution de type $\ell_0$ sur dictionnaire interpolé dans un cadre déterministe (voir Tableau 2.4 p. 78). Pour chaque niveau de bruit, les pourcentages d'ELR sont légèrement plus élevés pour l'échantillonneur marginalisé avec $\alpha = 0.8$ que pour la meilleure approximation linéaire pour la résolution $\ell_0$ (10dB : 68.5%/68%, 20 dB : 97.5%/94%, 30dB : 99.5%/88%). Notons qu'en particulier pour RSB=30dB, la solution $\ell_0$ sur dictionnaire interpolé souffrait de l'erreur d'approximation, alors que le modèle BGE modélise exactement le signal. De plus, ces résultats sont issus d'une étude statistique avec un critère binaire strict ; pour un unique signal, le résultat des résolutions $\ell_0$ donne une unique interprétation, alors que celui de l'échantillonneur permet d'observer des probabilités de détection à travers la moyenne $\widehat{\boldsymbol{q}}$. Par exemple, pour RSB=10dB et $\alpha = 0.7$, le pourcentage d'ELR grimpe à 72.5% pour l'échantillonneur. Au-delà des performances de détection, l'échantillonnage de la loi *a posteriori* en les paramètres d'intérêt permettent donc une meilleure compréhension du résultat, grâce notamment au niveau de confiance de détection donné par les $\widehat{q}_j$, et les barres d'erreur sur les localisations en prenant la variance des échantillons.

### 3.5.6 Bilan sur les tests du modèle BGE et des échantillonneurs associés

Ces tests confirment donc bien que l'échantillonneur de la loi marginale[12] permet d'obtenir une convergence plus rapide ; en terme d'itérations mais aussi en coût de calcul, grâce aux astuces calculatoires que nous avons développées. Par ailleurs, il permet d'obtenir le comportement d'*alternance/anti-corrélation* de la séquence de Bernoulli quand une localisation théorique se trouve proche de la séparation entre deux intervalles. Ce comportement permet en fait de dépasser la limite imposée par le découpage en $J$ intervalles de l'espace des localisations, puisque les localisations peuvent évoluer naturellement entre deux intervalles grâce à ces alternances. Nous avons alors donné une nouvelle stratégie de détection, se basant sur la fusion des intervalles voisins présentant des variables de Bernoulli fortement anti-corrélées. Les résultats de détection sont très satisfaisants pour les deux jeux de données pour l'échantillonnage de la loi marginale, dans les cadres supervisé et non supervisé.

## 3.6 Conclusion

Dans ce chapitre, nous sommes revenus sur le modèle Bernoulli-Gaussien étendu proposé par [BC06] dans le cadre de problèmes d'analyse spectrale en échantillonnage irrégulier, pour l'adapter au cadre plus général des problèmes d'APNL et notamment proposer un échantillonneur plus efficace de la loi *a posteriori*. Ce dernier exploite la marginalisation des amplitudes pour l'échantillonnage des paramètres d'intérêt que sont les paramètres de Bernoulli $\boldsymbol{q}$ et les paramètres non linéaires $\boldsymbol{\nu}$, dans les cadres supervisé et non supervisé. Sur le plan théorique, nos contributions majeures résident dans l'adaptation des astuces calculatoires de [GILC11] pour la mise en œuvre de l'échantillonnage des $q_j$ pour un modèle BG classique à atomes mono-dimensionnels aux spécificités de notre problème : d'abord la présence des paramètres non linéaires $\boldsymbol{\nu}$, mais aussi celle d'un bruit non indépendant et d'amplitudes multi-dimensionnelles de variance inconnue. Nous proposons aussi une stratégie différente pour réduire les instabilités numériques. Nos tests sur des données tests d'analyse spectrale et de déconvolution impulsionnelle nous ont permis de confirmer les atouts de notre échantillonneur vis-à-vis de celui de la loi jointe. Sur le plan pratique, nous avons aussi proposé une nouvelle stratégie de détection, plus adaptée à l'échantillonnage que nous proposons et au problème d'APNL.

Le modèle BGE associé à l'échantillonneur exploitant la marginalisation des amplitudes permet donc une résolution « optimale » du problème d'APNL. Notamment, les phénomènes d'*alternance* entre les variables de Bernoulli voisines remarqués quand la localisation théorique est proche de la séparation de deux intervalles, permet de dépasser le découpage de l'espace des localisations en $J$ intervalles. Dans ce cas, les échantillons de la localisation peuvent évoluer dans les deux intervalles voisins autour de la localisation théorique comme si cette dernière se trouvait au milieu d'un intervalle. De plus, les échantillons des divers paramètres permettent une meilleure interprétation des résultats, notamment à travers la séquence moyenne $\widehat{\boldsymbol{q}}$ et les barres d'erreur sur les localisations calculées à partir des variances empiriques des échantillons.

On peut cependant relever quelques limites pour le modèle BGE et l'échantillonneur proposé, notamment au niveau de la stratégie de détection et de la possibilité de marginalisation.

Notre stratégie de détection, reposant sur la fusion des intervalles voisins anti-corrélés, fonctionne plutôt bien dans le cas du problème d'analyse spectrale grâce notamment à un découpage cohérent de l'espace des localisations. En effet, la largeur $\Delta$ des intervalles $\mathcal{I}_j$ a été choisie en

[12] Pour rappel, dans le cadre non supervisé, il s'agit l'échantillonneur de type PCGS de la loi jointe exploitant la loi marginale pour les paramètres d'intérêt.

lien avec la résolution spectrale (on ne cherche pas à dépasser la résolution du périodogramme - voir 1.5.1), si bien que les échantillons d'une fréquence ne devrait pas se répandre dans un intervalle de largeur supérieur à $\Delta$, soit dans plus de deux intervalles $\mathcal{I}_j$. Dans d'autres applications (et notamment nous le verrons pour la détection d'exoplanètes étudiée au chapitre suivant), un tel découpage cohérent des intervalles peut être impossible, et des échantillons d'une localisation peuvent se retrouver dans plus de deux intervalles voisins, mettant en échec notre stratégie de détection. De même, deux localisations théoriques pourraient se retrouver dans le même intervalle $\mathcal{I}_j$ rendant impossible leur détection simultanée. Le choix d'un découpage cohérent est donc une vraie problématique, à adapter à l'application en jeu. Par ailleurs, des échantillons peuvent aussi amener à des lois bimodales à la fois pour les échantillons $q_j^{(t)}$ et pour les localisations $\nu_j^{(t)}$ ; nous l'avons remarqué succinctement dans le cas de la déconvolution impulsionnelle et nous y reviendrons plus en détails au chapitre suivant. Des stratégies de détection plus adaptées à ces problèmes pourraient donc être recherchées ; pour l'instant, une étude attentive des échantillons - plutôt qu'une détection automatique - permet de nous alerter sur la présence de tels phénomènes.

Par ailleurs, la marginalisation des amplitudes $\boldsymbol{x}$, qui permet un meilleur échantillonnage de la loi *a posteriori*, est seulement possible quand les amplitudes $\boldsymbol{x}$ ne sont pas contraintes, *i.e.* $\boldsymbol{x}_j \in \mathbb{R}^m$. Dans le cas contraire, par exemple dans le cas d'amplitudes positives, la marginalisation devient impossible. D'autres techniques pourraient alors être envisagées pour accélérer l'échantillonnage des paramètres du modèle BGE, comme par exemple l'échantillonnage par K-uplets mis en œuvre par [GILC11] dans le cas du modèle BG classique.

L'application à la détection d'exoplanètes du modèle BGE et de l'échantillonneur proposé sera traitée au chapitre suivant. Nous verrons que cette application est plus difficile que celles traitées dans ce chapitre ; notamment, le modèle de bruit sera plus compliqué et des paramètres de forme seront à prendre en compte.

# Chapitre 4

# Application à la détection d'exoplanètes

**Sommaire**


**4.1 Détection d'exoplanètes par l'analyse des vitesses radiales . . . . . 148**
4.1.1 Principe de la détection d'exoplanètes par l'analyse des vitesses radiales : formulation d'un problème inverse . . . . . 150
4.1.2 Méthodes pour l'estimation des paramètres orbitaux . . . . . 155
4.1.3 L'activité stellaire, un véritable enjeu . . . . . 158
4.1.4 Le nombre d'exoplanètes . . . . . 161
4.1.5 Bilan : la parcimonie, une alternative . . . . . 162
**4.2 Modélisation BGE et échantillonneur associé pour la détection d'exoplanètes 163**
4.2.1 Reparamétrisation du modèle . . . . . 164
4.2.2 Reformulation en APNL via le découpage du domaine des localisations . . . . . 165
4.2.3 Modèle BGE et échantillonnage associé . . . . . 166
4.2.4 Détection des exoplanètes via l'estimation des paramètres . . . . . 169
4.2.5 Bilan . . . . . 170
**4.3 Tests sur des données simulées sans activité stellaire . . . . . 171**
4.3.1 Les données . . . . . 171
4.3.2 Un premier essai : algorithmes gloutons avec descente de gradient . . . . . 172
4.3.3 Exploitation du modèle BGE et de l'échantillonneur associé . . . . . 174
4.3.4 Bilan . . . . . 182
**4.4 Tests sur des données avec activité stellaire . . . . . 182**
4.4.1 Les données . . . . . 182
4.4.2 Réglages du modèle BGE et de l'échantillonneur associé . . . . . 183
4.4.3 Résultats . . . . . 184
**4.5 Premiers tests sur des données réelles . . . . . 189**
4.5.1 Système HD40307 . . . . . 189
4.5.2 Système CoRoT-7 . . . . . 191
**4.6 Conclusion . . . . . 194**


Nous proposons dans ce chapitre d'exploiter nos travaux réalisés dans le cadre des problèmes d'APNL à un problème d'actualité en astrophysique : la détection d'exoplanètes par la méthode des vitesses radiales. Nous commençons par introduire en § 4.1 la problématique de la détection d'exoplanètes dans son contexte astrophysique : quel est le principe de cette méthode de détection, comment les signaux de vitesses radiales sont mesurés, comment les astrophysiciens traitent leurs données et quels sont les grands enjeux actuels. Notons que cette étude n'est pas à visée exhaustive et les concepts présentés sont souvent simplifiés, mais elle vise à regrouper les

diverses informations pour comprendre le problème en jeu. Elle nous permet ainsi de montrer en quoi le modèle BGE et l'échantillonneur associé est adapté à ce problème et est une alternative intéressante et un complément aux méthodes actuellement utilisées par les astrophysiciens. Après avoir décrit l'adaptation de notre méthode à ce problème en § 4.2, nous montrons en § 4.3 son efficacité sur des signaux simulés de vitesse radiale avec un bruit indépendant (sans activité stellaire). Nous confirmerons l'utilité de notre méthode sur des données plus difficiles avec un bruit corrélé (activité stellaire simulée) en § 4.4 mais aussi sur de réelles mesures de vitesses radiales en § 4.5. Ces travaux ont été menés en collaboration avec Michaël Bazot[1].

## 4.1 Détection d'exoplanètes par l'analyse des vitesses radiales

Une *exoplanète* est une planète orbitant autour d'une autre étoile que notre Soleil. Même si les planètes[2] de notre système solaire sont pour la plupart connues depuis l'Antiquité, l'existence d'exoplanètes orbitant autour d'une étoile similaire à la nôtre a été remise en cause jusqu'au début des années 90. En effet, l'éloignement des étoiles (la plus proche de nous, *Proxima du Centaure*, se trouve à plus de 4,23 années lumières de la Terre) ainsi que la différence considérable de luminosité entre une étoile et sa planète font des exoplanètes des cibles difficiles pour l'imagerie directe. Cependant, les exoplanètes passionnent autant les astrophysiciens que le grand public, que ce soit pour la possibilité de découvrir une planète où une vie extra-terrestre serait présente, ou tout du moins possible (les plus ambitieux pourront alors imaginer leur colonisation) et surtout afin de mieux comprendre les mécanismes de formation de notre système solaire. Ces motivations ont amené à la découverte de 51 Peg b[3], la première exoplanète gravitant autour d'une étoile similaire au Soleil, 51 Peg, en 1995 [MQ95]. Cette première découverte enthousiasma la communauté, si bien que le nombre d'exoplanètes découvertes à ce jour est de 3824[4]. Ces détections, toujours plus nombreuses, sont effectuées grâce à diverses méthodes que les astrophysiciens ne cessent d'améliorer. Se basant sur des observations indépendantes, elles permettent une meilleure robustesse des résultats. Nous récapitulons trois méthodes de détection, schématisées en Figure 4.1, dont les deux premières sont les plus prolifiques :

- *L'analyse de vitesses radiales* : la vitesse radiale d'une étoile (vitesse projetée sur l'axe Étoile-Observateur) étant influencée par la rotation des exoplanètes autour d'elle, son analyse permet de détecter ses compagnons planétaires. Via l'effet Doppler-Fizeau, elle est reliée à des décalages du spectre de l'étoile et est donc mesurée depuis le sol à l'aide de spectrographes à haute résolution, comme ELODIE ou HARPS. C'est grâce à cette méthode que la première exoplanète 51 Peg b a été découverte.
- *La photométrie et les transits* : cette méthode repose aussi sur l'observation de l'étoile hôte, dont la luminosité diminue lorsque l'exoplanète passe entre l'étoile et l'observateur, provoquant une *éclipse primaire*. La surveillance de l'intensité lumineuse de l'étoile par photométrie permet donc de repérer de tels phénomènes. Cette méthode a été particulièrement prolifique ces dernières années, avec le lancement des satellites CoRoT et Kepler.
- *L'imagerie directe* : cette dernière est la plus intuitive mais est limitée par la faible distance angulaire entre l'étoile et l'exoplanète ainsi que la différence de luminosité entre les deux corps. Les mesures sont effectuées à l'aide d'instruments comme SPHERE au VLT ( *Very Large Telescope*, Chili), ou ceux qui équipent les télescopes Keck et Gemini.

[1] Associé de recherche postdoctoral depuis 2015 à NYU Abou Dabi en physique stellaire.

[2] La définition d'une planète en soi peut toujours évoluer, comme peut l'illustrer par exemple le retrait du titre de planète à Pluton en 2006. Voir [Per11] pour plus d'explications.

[3] Les exoplanètes sont souvent nommées en lien avec leur étoile ; ainsi 51 Peg b correspond à la première planète de l'étoile 51 Peg (diminutif de 51 Pegasi) ; les suivantes seront nommées 51 Peg c, 51 Peg d... La lettre « a » est réservée à l'étoile.

[4] D'après l'Encyclopédie des planètes extra-solaires `http://exoplanet.eu/catalog/`, consultée le 11/09/2018.

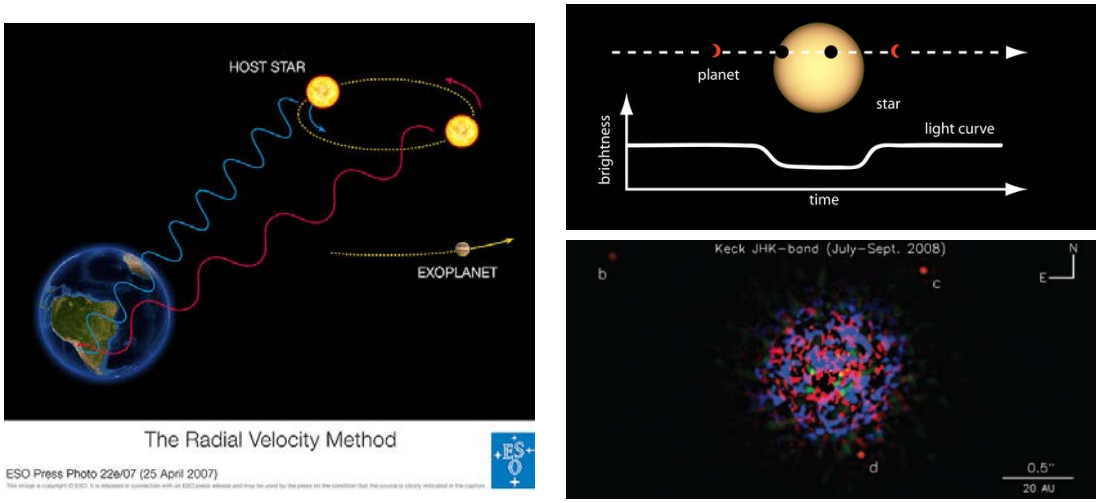



FIG. 4.1 – Différentes méthodes de détection d'exoplanètes. À gauche : vitesses radiales. Source : European Southern Observatory (ESO). En haut à droite : méthode des transits. Source : [Web]. En bas à droite : imagerie directe. Source : [MMB+08].

Ces méthodes se complètent, comme cela est illustré dans [SBA+18]. Par exemple, la méthode des vitesses radiales permet de connaître la masse minimale[5] de la planète, tandis que la méthode des transits permet de connaître son rayon ; le suivi du système par les deux méthodes permet alors d'estimer la densité de la planète et donc de savoir s'il s'agit plutôt d'une planète tellurique comme la Terre ou d'une gazeuse comme Jupiter. L'imagerie directe permet l'observation directe du spectre de la planète et facilite donc l'évaluation de sa composition. Aussi, chaque méthode subit des biais observationnels importants ; l'imagerie directe ne peut observer que des planètes suffisamment éloignées de leur étoile ; la détection par vitesses radiales est plus sensibles aux planètes massives ayant une forte influence gravitationnelle sur l'étoile ; la méthode des transits est limitée aux planètes proches de leur étoile (petite période de rotation) et aux systèmes correctement orientés par rapport au satellite. Souvent, les planètes détectées par transits avec un satellite comme Kepler sont considérées comme des candidates, jusqu'à ce qu'elles soient confirmées[6] par des observations de vitesses radiales. Pour palier les biais observationnels et obtenir des mesures plus précises et ciblées, les astrophysiciens développent des instruments toujours plus performants. On peut citer TESS et PLATO pour la détection par transits, ou encore ESPRESSO et SPIRou pour la détection par vitesses radiales.

Le principal objectif des astrophysiciens était d'abord de découvrir un grand nombre d'exoplanètes afin de mieux comprendre les différents types de planètes existant. 51 Peg b était d'ailleurs un type de planète inédit puisque non présent dans le système solaire : un « Jupiter chaud », une géante gazeuse proche de son étoile. Au fur et à mesure des découvertes et de l'amélioration des instruments de mesure et des méthodes de traitement de données, les travaux se sont davantage concentrés sur la recherche de planètes telluriques (de masse et de rayon similaires à ceux de la Terre afin de garantir une planète non gazeuse) dans le double objectif de mieux contraindre la formation de notre planète mais aussi de découvrir des mondes habitables. Pour ce dernier objectif, la recherche de telles planètes dans la *zone habitable*, notamment caractérisée par la distance entre la planète et son étoile, est particulièrement intensive. Jusqu'à

[5] On parle de masse minimale car la méthode donne seulement accès à la masse de la planète pondérée par le sinus de l'angle d'inclinaison du plan orbital par rapport à l'angle de visée.

[6] L'étude des vitesses radiales de l'étoile permet en effet de calculer la masse du corps en transit, et donc de déterminer s'il s'agit bien d'une planète ou d'un autre corps céleste.

maintenant, de telles découvertes sont très peu nombreuses et notamment réalisées grâce aux méthodes de type *transit/photométrie*.

Pour nos travaux, nous nous intéressons à la méthode de détection par analyse des vitesses radiales. En effet, des efforts considérables sont actuellement réalisés dans la communauté astrophysique pour lui permettre de détecter des planètes de type Terre ou résidant dans la zone habitable de l'étoile, comme nous allons le voir dans cette section. Nous revenons d'abord sur le modèle reliant la vitesse radiale d'une étoile et les paramètres orbitaux de ses exoplanètes, pour formuler le problème de détection comme un problème inverse. Nous décrivons ensuite les méthodes utilisées par les astrophysiciens pour estimer les paramètres orbitaux, ainsi que les principaux enjeux actuels : la prise en compte de l'activité stellaire et l'estimation du nombre d'exoplanètes. Cette première section introductive nous permettra ainsi de mieux comprendre le problème, pour lequel les méthodes de résolution que nous avons développées dans nos travaux (principalement le modèle BGE du chapitre 3) représentent une véritable alternative aux méthodes utilisées par les astrophysiciens.

Notons qu'une grande partie des informations données dans cette section provient du livre [Per11] ainsi que de thèses récentes sur le sujet [Héb15, Hay15, Har17, Sul17][7].

### 4.1.1 Principe de la détection d'exoplanètes par l'analyse des vitesses radiales : formulation d'un problème inverse

Le principe de la détection d'exoplanètes par l'analyse des vitesses radiales repose sur la relation physique qui existe entre la vitesse radiale d'une étoile et des paramètres caractéristiques de chaque exoplanète (tels que sa période de rotation) notés $\boldsymbol{\Theta}$. Une telle relation permet de formuler le problème de détection comme un problème inverse, puisqu'il s'agit alors d'estimer les paramètres $\boldsymbol{\Theta}$ qui sont « responsables » du signal mesuré de vitesses radiales $\boldsymbol{y}$ via le modèle physique.

Nous expliquons ici la nature des paramètres orbitaux ainsi que le modèle physique permettant de les relier à la vitesse radiale. Par ailleurs, nous décrivons brièvement le principe de mesure de cette dernière, qui s'effectue par spectroscopie.

#### Modèle reliant les paramètres orbitaux et la vitesse radiale

Pour relier la vitesse radiale d'une étoile (sa vitesse dans la direction observateur-étoile) aux paramètres orbitaux, il est nécessaire de comprendre les lois physiques régissant la trajectoire de l'étoile et des planètes : la rotation de planètes autour d'une étoile est due aux interactions gravitationnelles que les unes exercent sur les autres.

Plus précisément, dans le cas d'un système à deux corps (1 seule planète en orbite), la planète mais aussi l'étoile orbitent autour du centre de masse du système en suivant des trajectoires elliptiques coplanaires dont le centre de masse est un foyer ; la période et la forme des deux orbites sont similaires, mais de tailles différentes. La trajectoire d'une planète consiste donc en une ellipse dans le « plan orbital », comme illustrée en Figure 4.2. Elle est paramétrée par plusieurs paramètres, parmi lesquels nous décrivons :

- la période de l'orbite planétaire $P \in \mathbb{R}_+$, soit le temps mis par l'exoplanète pour effectuer une révolution autour de l'étoile, généralement exprimée en jours ;
- l'excentricité de l'orbite planétaire $e \in [0; 1[$. Pour $e > 0$, on parlera de *trajectoire elliptique* ; pour $e = 0$, de *trajectoire circulaire* ; dans un cas général, de *trajectoire képlérienne*.

[7] De plus, notons l'aide non négligeable apportée par Michaël Bazot et Baptiste Klein, ce dernier étant actuellement en thèse à l'IRAP sur la caractérisation des systèmes planétaires de naines rouges actives avec SPIRou.

– le temps $T^{\mathrm{P}} \in [0; P]$ et l'argument $\omega^0 \in [0; 2\pi]$ de passage au périastre. Le périastre est le point de la trajectoire excentrique le plus proche du centre de masse du système, et n'a donc de sens que pour une trajectoire excentrique. Dans le cas d'une trajectoire circulaire, $\omega^0$ peut être considéré comme nul et $T^{\mathrm{P}}$ correspondra à un temps de référence pour la définition de l'orbite.
– $a$ le demi-grand axe de l'ellipse et $i$ l'inclinaison de l'orbite.

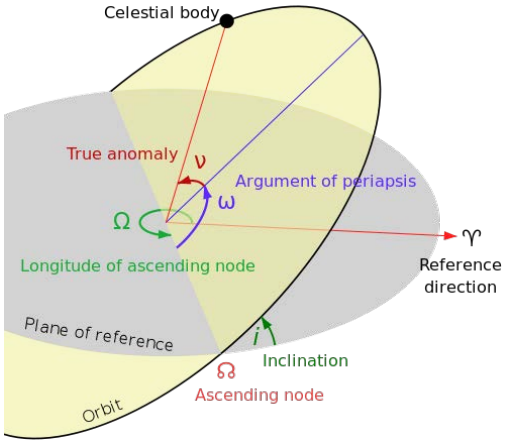



FIG. 4.2 – Illustration de l'orbite d'une exoplanète. Crédits : Nathan Hara et Jacques Laskar - ESEP.

La vitesse radiale $v(t)$ de l'étoile peut alors être modélisée comme la somme de $v_0$, la vitesse radiale du centre de masse du système, et d'une composante dépendant des paramètres orbitaux de la planète (voir [Per11] pour plus de détails). En notant $\boldsymbol{\nu} = (P, e, T^{\mathrm{P}}, \omega^0)$, on a :

$$\begin{array}{rcl} v(t; \boldsymbol{\nu}, K, v_0) & = & v_0 + K\, h_{\mathrm{kep}}(t; \boldsymbol{\nu}) \\ \text{où} \quad h_{\mathrm{kep}}(t; \boldsymbol{\nu}) & = & \cos(\rho(t; P, e, T^{\mathrm{P}}) + \omega^0) + e\cos(\omega^0) \end{array} \tag{4.1}$$

où $t$ représente le temps et où l'anomalie vraie $\rho(t; P, e, T^{\mathrm{P}})$ vérifie :

$$\tan\left(\frac{\rho(t; P, e, T^{\mathrm{P}})}{2}\right) = \sqrt{\frac{1+e}{1-e}}\,\tan\left(\frac{E(t; P, e, T^{\mathrm{P}})}{2}\right) \tag{4.2}$$

et l'anomalie excentrique $E(t; P, e, T^{\mathrm{P}})$ satisfait l'équation de Kepler :

$$E(t; P, e, T^{\mathrm{P}}) - e\sin(E(t; P, e, T^{\mathrm{P}})) = 2\pi\frac{t - T^{\mathrm{P}}}{P} \tag{4.3}$$

La « semi-amplitude » $K \in \mathbb{R}_+$ s'exprime en fonction des paramètres orbitaux :

$$K = \frac{2\pi}{P}\frac{a\sin i}{(1-e^2)^{1/2}} \tag{4.4}$$

Par la troisième loi de Kepler, elle est aussi reliée à la masse des deux systèmes ; c'est donc l'estimation conjointe des paramètres orbitaux et de la semi-amplitude qui permet d'estimer la masse minimale de la planète.

Notons que quand l'excentricité $e$ est nulle (trajectoire circulaire), le modèle se réduit à une sinusoïde de période $P$, d'amplitude $K$ et de phase $\omega'$ :

$$v(t; P, \omega', K) = v_0 + K\cos\left(\frac{2\pi t}{P} + \omega'\right)$$

Dans le cas d'un système à $L > 1$ planètes (système à $L+1$ corps), la modélisation devient plus compliquée, notamment à cause des interactions gravitationnelles possibles entre les planètes. Cependant, quand celles-ci peuvent être négligées, la vitesse radiale de l'étoile peut être modélisée comme la somme des contributions de chaque exoplanète suivant toujours une trajectoire elliptique de paramètres $\boldsymbol{\nu}_\ell = (P_\ell, e_\ell, \omega_\ell^0, T_\ell^{\mathrm{P}})$ (voir [Per11]) :

$$v(t; \boldsymbol{\nu}, \boldsymbol{K}, v_0) = v_0 + \sum_{\ell=1}^{L} K_\ell \; h_{\mathrm{kep},\ell}(t; \boldsymbol{\nu}_\ell) \tag{4.5}$$

Comme le montre l'équation (4.1), l'amplitude de la contribution d'une exoplanète à la vitesse radiale d'une étoile est fonction de la semi-amplitude $K$. Comme nous connaissons son expression en fonction des masses des astres, du demi-grand axe $a$ et d'autres paramètres orbitaux par les lois de Kepler (voir (4.4) et [Per11]), on peut facilement obtenir un ordre de grandeur des semi-amplitudes $K$ attendues suivant le type de planète étudiée. Plus la planète sera massive et/ou proche de son étoile, et plus sa contribution sera importante. Le Tableau 4.1, tiré de [Hay15], donne des ordres de grandeurs de la contribution en vitesse radiale attendue pour des planètes similaires à celles du système solaire (de type Jupiter, Neptune ou Terre). Pour des planètes de type Jupiters chauds (masse similaire à celle de Jupiter mais plus proche de l'étoile), la contribution $K$ sera de l'ordre de la dizaine de $\mathrm{m.s}^{-1}$ tandis que des planètes similaires à la Terre ne contribueraient que d'environ un dixième de $\mathrm{m.s}^{-1}$. La recherche de planètes similaires à la Terre autour d'une étoile similaire au Soleil nécessite donc de pouvoir mesurer la vitesse radiale des étoiles avec une précision de l'ordre du dixième de $\mathrm{m.s}^{-1}$. Cela a donc mené au développement de spectrographes de précision croissante, comme l'explique le paragraphe suivant.

| Masse de la planète | Distance à l'étoile $a$ | Période de l'orbite $P$ | Semi-amplitude $K$ |
|---|---|---|---|
| Jupiter (317 $M_\oplus$) | 5 UA | 5 ans | 12 $\mathrm{m.s}^{-1}$ |
| Jupiter chaud | 1 UA | 1 an | 28 $\mathrm{m.s}^{-1}$ |
| Neptune (17 $M_\oplus$) | 0.1 UA | 36 jours | 5 $\mathrm{m.s}^{-1}$ |
| Super-Terre (5 $M_\oplus$) | 0.1 UA | 36 jours | 1.4 $\mathrm{m.s}^{-1}$ |
| Terre (1 $M_\oplus$) | 1 UA | 1 an | 0.09 $\mathrm{m.s}^{-1}$ |

TAB. 4.1 – Ordre de grandeur de la contribution $K$ d'une planète à la vitesse radiale d'une étoile similaire au Soleil, suivant sa masse, sa distance à l'étoile et sa période. Les dénominatifs *Jupiter*, *Neptune* et *Terre* sont uniquement relatifs à la masse de la planète concernée (et non à la distance $a$ ni à la période $P$). Tableau tiré de [Hay15] (p.4).

*Remarque : Par la suite, $\boldsymbol{\Theta} = \{v_0, \boldsymbol{P}, \boldsymbol{e}, \boldsymbol{T}^P, \boldsymbol{\omega}^0, \boldsymbol{K}\}$ regroupe l'ensemble des paramètres à estimer. Par abus de langage, on les nommera paramètres orbitaux (bien que les semi-amplitudes et la vitesse radiale du centre de masse n'en soient pas en théorie).*

### Mesurer la vitesse radiale grâce à la spectroscopie

La vitesse radiale d'une étoile peut être mesurée grâce à l'effet Doppler-Fizeau, reliant la vitesse radiale d'un objet avec des décalages de son spectre d'absorption. Le spectre d'une étoile est un continuum de longueurs d'onde $\lambda$ présentant des bandes d'absorption qui sont caractéristiques des espèces chimiques constituant l'étoile. Lorsque l'étoile s'éloigne de l'observateur terrestre, son spectre est alors décalé vers le rouge (*red-shift*), tandis que lorsqu'elle se rapproche son spectre est décalé vers le bleu (*blue-shift*), comme schématisé en Figure 4.1. On compare alors le décalage des *raies d'absorption* du spectre mesuré $\lambda_m(t)$ avec celles d'un spectre de référence $\lambda_0$. Dans le cas d'une étoile proche où la vitesse totale relative de l'observateur peut

être négligée par rapport à la vitesse de la lumière $c$, la vitesse radiale peut être approchée de la façon suivante :

$$v_{\text{mes}}(t) = c\,\frac{\lambda_m(t) - \lambda_0}{\lambda_0} \tag{4.6}$$

reliant directement le décalage en longueur d'onde à la vitesse radiale. En pratique, la corrélation entre le spectre de référence et le spectre mesuré est calculé pour plusieurs décalages de $\lambda_0$, donnant la *Cross Correlation Function* (CCF) comme cela est illustré en Figure 4.3 ; celle-ci est approchée avec une gaussienne, dont la moyenne est directement reliée à la vitesse radiale.

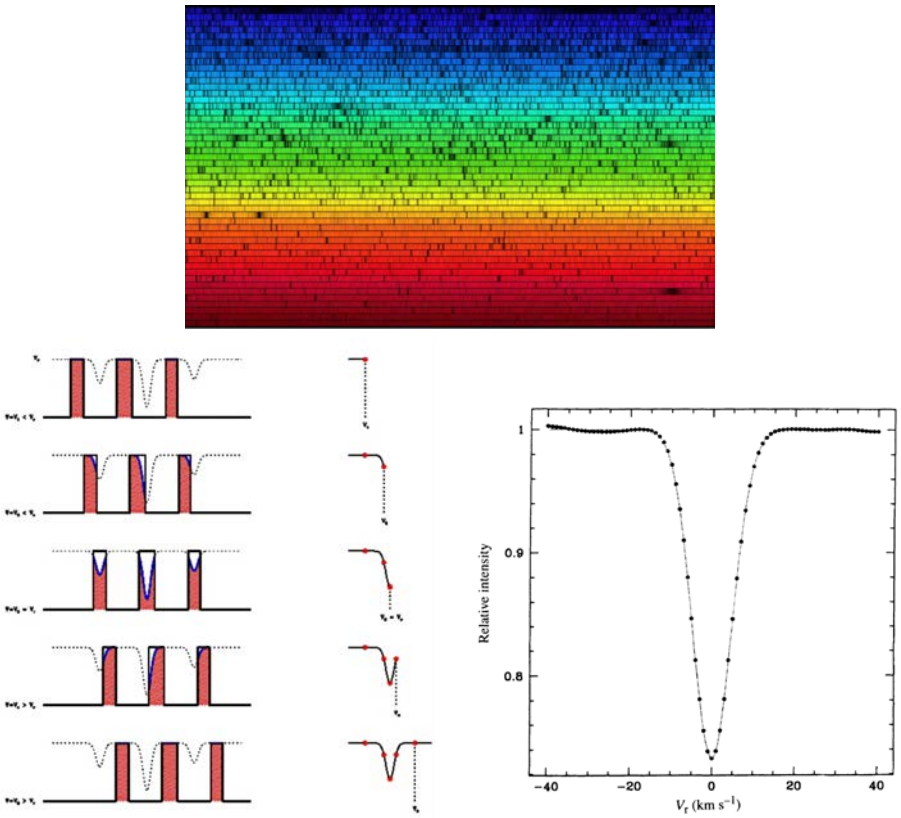



FIG. 4.3 – Illustration de la mesure de la vitesse radiale. En haut, spectre du soleil entre 392 nm (bleu) et 692 nm (rouge). En bas à gauche, corrélation entre un spectre de référence et un spectre mesuré. En bas à droite, un exemple de CCF. Source : [Har17] (Fig. 1.5, p.8).

Les mesures de vitesses radiales par spectroscopie sont effectuées depuis des télescopes sur Terre à l'aide de spectrographes à haute résolution. En effet, ces derniers doivent être capables de discerner de très faibles décalages de longueurs d'onde pour obtenir une mesure précise de la vitesse radiale. De nombreux instruments existent, mais actuellement, le plus précis d'entre eux est HARPS pour *High Accuracy Radial velocity Planet Searcher* installé à l'observtoire de la Silla, au Chili ; il atteint une précision de l'ordre de 1 m.s$^{-1}$ [MPQ+03]. Comme nous l'avons vu précédemment, les astrophysiciens cherchent à détecter des planètes ayant une contribution en vitesse radiale toujours plus faible, actuellement de l'ordre du cm.s$^{-1}$. Le nouveau spectrographe ESPRESSO, *Echelle Spectrograph for Rocky Exoplanet and Stable Spectroscopic Observations*, installé au *Very Large Telescope* (VLT) au Chili, dont les premières lumières datent de novembre 2017, permettra d'atteindre une telle précision et en théorie de détecter des planètes similaires à la Terre grâce à la méthode des vitesses radiales. Une autre stratégie est actuellement suivie : plutôt que de rechercher des planètes autour d'étoiles similaires au Soleil (émettant principalement dans le visible) les recherches peuvent se réorienter vers les « naines rouges », des étoiles

plus petites et plus froides, qui représentent environ 80% de notre voisinage stellaire. En effet, leur température moins élevée entraîne une zone habitable plus proche, dans laquelle une exoplanète aurait donc une signature plus grande en vitesse radiale qu'une exoplanète de même masse et de même distance à une étoile de type Soleil. Cependant, elles émettent principalement dans l'infra-rouge à cause de leur faible température et la mesure de leur vitesse radiale nécessite donc des spectrographes adaptés. Le spectropolarimètre SPIRou, *SPectro-Polarimètre InfraRouge*, permettra d'observer ce type d'étoiles en étudiant leur spectre dans l'infra-rouge proche. L'instrument est le fruit d'une collaboration internationale dont l'IRAP est l'investigateur principal et est installé au CFHT (Canada-France-Hawaï Télescope). Ses premières mesures scientifiques sont imminentes. Étant équipé d'un polarimètre, il permettra également d'étudier le champ magnétique des étoiles observées.

Enfin, les mesures étant effectuées depuis la Terre, elles sont sujettes à un échantillonnage irrégulier, dû à l'alternance jour/nuit, à la météo, à la saison ou encore à la disponibilité du télescope, de même type que celui dont nous avons parlé en § 1.1.3 pour l'analyse spectrale de signaux irrégulièrement échantillonnés. Un exemple de signal mesuré avec les erreurs de mesures liées à l'instrument est donné en Figure 4.4 ; on peut voir que les mesures sont effectuées de façon irrégulière sur plusieurs années, et présentent des périodes de forte densité de mesures (ici vers 2002-2003), tandis qu'il existe aussi des points isolés.

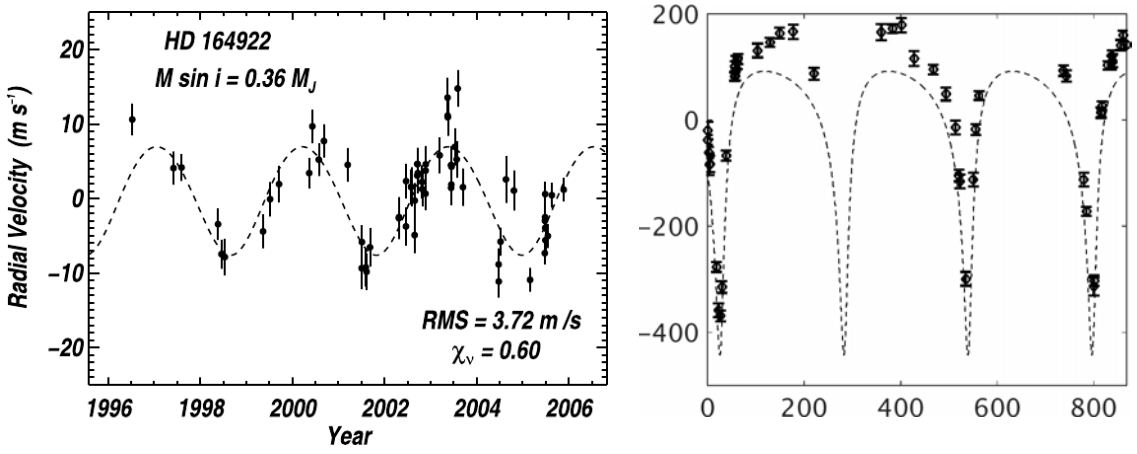



FIG. 4.4 – À gauche, illustration d'un signal de vitesse radiale pour une orbite circulaire. Les mesures sont effectuées sur plusieurs années. Source : [BWM+06]. À droite, illustration d'un signal de vitesse radiale pour une orbite elliptique d'excentricité $e = 0.67$ (abscisse en jours).

### Formulation du problème de détection d'exoplanètes comme un problème inverse non linéaire

La vitesse radiale mesurée $\boldsymbol{y} = v_{\text{mes}}(\boldsymbol{t})$ grâce à l'effet Doppler via l'équation (4.6) peut donc être reliée au modèle képlérien $\boldsymbol{v}(\boldsymbol{\nu}, \boldsymbol{K}, v_0) = v(\boldsymbol{t}; \boldsymbol{\nu}, \boldsymbol{K}, v_0)$ de l'équation (4.5) calculé en les instants d'échantillonnage $\boldsymbol{t} = [t_1, \ldots, t_N]^T$ irréguliers. Le problème de détection d'exoplanètes peut donc s'écrire comme un problème inverse parcimonieux non linéaire tel que décrit au chapitre 1 :

$$\text{Estimer } L \text{ faible, } v_0 \text{ et les } \{\boldsymbol{\nu}_\ell, K_\ell\}_\ell, \text{ tels que } \boldsymbol{y} = v_0 + \sum_{\ell=1}^{L} K_\ell \, \mathbf{h}_{\text{kep}}(\boldsymbol{\nu}_\ell) + \boldsymbol{\epsilon} \tag{4.7}$$

où $\boldsymbol{\epsilon}$ regroupe les erreurs de mesures, mais aussi des erreurs de modèle et des signaux parasites, autres que l'influence des exoplanètes. L'ordre du modèle $L$, qui ici représente le nombre d'exo-

planètes du système, peut être supposé faible, car le contraire serait physiquement peu probable. Un *a priori* de parcimonie peut donc être pris en compte.

Notons que ce problème inverse est très similaire au problème d'analyse spectrale en échantillonnage irrégulier décrit en § 1.1.3, sur lequel nous avons testé nos diverses méthodes de résolution d'APNL en § 2.6 et § 3.5. En plus de l'échantillonnage irrégulier spécifique aux mesures effectuées en astrophysique, nous avons vu que pour des orbites circulaires, le modèle képlérien (4.1) revient exactement au modèle *cosinus+phase* de l'analyse spectrale (voir Tableau 1.1 p.6). Le problème de détection d'exoplanètes se différencie de l'analyse spectrale par la présence des képlériennes (il ne s'agit pas de sinusoïdes pures mais de signaux périodiques) et donc par la présence des paramètres de forme $\boldsymbol{\upsilon} = (\boldsymbol{e}, \boldsymbol{T}^{\mathrm{P}}, \boldsymbol{\omega}^0)$, le paramètre de localisation étant la période $P$. De plus, pour des trajectoires elliptiques, le calcul de $\mathbf{h}(\boldsymbol{\nu}_\ell)$ ne peut pas se faire analytiquement ; il est nécessaire au préalable de résoudre numériquement[8] l'équation de Kepler (4.3). Par ailleurs, le bruit $\boldsymbol{\epsilon}$ pourra comprendre une partie corrélée due à l'activité stellaire ; nous y reviendrons en § 4.1.3.

Les paragraphes suivants tentent de résumer les méthodes utilisées par les astrophysiciens pour estimer les paramètres orbitaux, ainsi que les deux grands enjeux de la communauté : la prise en compte de l'activité stellaire et l'estimation du nombre d'exoplanètes.

### 4.1.2 Méthodes pour l'estimation des paramètres orbitaux

Nous présentons ici les principales méthodes utilisées par les astrophysiciens pour estimer les paramètres orbitaux, en supposant le nombre $L$ d'exoplanètes connu (nous verrons en § 4.1.4 comment est actuellement estimé le nombre $L$ d'exoplanètes). Pour un nombre fixé d'exoplanètes $L$, $5L+1$ paramètres sont à estimer : $\{P_\ell, e_\ell, \omega^0_\ell, T^{\mathrm{P}}_\ell, K_\ell\}_\ell$ ainsi que la vitesse du centre de masse $v_0$. Nous distinguons 3 types de méthodes que nous décrivons rapidement : les méthodes à base de périodogrammes et algorithmes gloutons, celles se basant sur l'échantillonnage stochastique MCMC et celles se basant sur le problème d'optimisation avec *a priori* de parcimonie grâce à une pénalisation $\ell_1$.

**Périodogrammes et sélections itératives de type glouton/greedy.** L'estimation des paramètres orbitaux par périodogrammes fut le premier type de méthodes utilisé pour chercher des périodicités dans un signal de vitesse radiale. Les extensions des périodogrammes classiques à l'échantillonnage irrégulier de [Lom76, Sca82] (périodogrammes de Lomb-Scargle / LS) permettent en effet de rechercher des sinusoïdes dans des signaux irrégulièrement échantillonnés, et peuvent donc être utilisés pour rechercher des exoplanètes à orbite circulaire. Le périodogramme le plus utilisé par la communauté est le *Generalised Lomb-Scargle periodogram* (GLS) de [ZK09], qui généralise la prise en compte des erreurs indépendantes sur les mesures $(\sigma_n^2)_n$ et de la composante continue $v_0$ (déjà introduit par [FM81]) en proposant une formulation analytique et performante du périodogramme, dans le cas d'une orbite circulaire. Typiquement, il est calculé sur une grille fine de périodes $P$ afin de sélectionner celle qui le maximise.

Le *périodogramme Képlérien* correspond à l'extension du périodogramme aux trajectoires elliptiques. Il a été introduit par [Cum04] puis généralisé par [ZK09]. Il représente un coût calculatoire plus important, car il prend en compte la dépendance non linéaire en minimum 3 paramètres : la période $P$ mais aussi l'excentricité $e$ et le temps de passage au périastre $T^{\mathrm{P}}$. Les

[8] L'équation de Kepler (4.3) peut être résolue à l'aide de méthode de Newton : en écrivant $E = E(t; P, e, T^{\mathrm{P}})$, $M = 2\pi(t - T^{\mathrm{P}})/P$ et en initialisant $E_0 = M$, on calcule les termes de la suite $E_{k+1} = E_k - (E_k - e\sin E_k - M)/(1 - e\cos E_k)$ jusqu'à atteindre un critère d'arrêt. Typiquement, sa résolution ne prend pas plus d'une dizaine d'itération, sauf pour des excentricité très élevée (e>0.95).

calculs s'effectuent soit en discrétisant sur une grille à plusieurs dimensions les paramètres non linéaires pour choisir le triplet optimal [ZK09], soit en utilisant un algorithme d'optimisation non linéaire, de type *Levenberg-Marquardt*, initialisé avec un périodogramme LS [Cum04]. La discrétisation 3D de la première technique représente un coût calculatoire considérable si l'on souhaite avoir une précision suffisante, tandis que l'optimisation non linéaire de la seconde risque de tomber dans un minimum local. Pour palier ces problèmes, les travaux de Baluev (voir [Bal14a] par exemple) proposent des stratégies d'exploration intelligente de l'espace des paramètres. Les *algorithmes génétiques* ont aussi été largement utilisés par la communauté astrophysique (voir par exemple [PCM+07]).

Les périodogrammes, qu'ils soient pour une trajectoire circulaire ou elliptique, ne permettent la sélection que d'une planète à la fois, en cherchant à faire correspondre au mieux une sinusoïde/képlérienne avec le signal $\boldsymbol{y}$. Pour la recherche de systèmes multi-planétaires, une stratégie de type algorithme glouton, souvent de type *Matching Pursuit* et appelée technique de « prewhitening » par les astrophysiciens [GD73], est généralement adoptée : le périodogramme est calculé, les paramètres orbitaux (dont la période $P$) le maximisant sont sélectionnés et permettent de calculer le résidu (différence entre le signal et le modèle estimé), et la sélection se répète itérativement sur le résidu. De nombreuses variantes existent ; par exemple, la sélection des périodes par un périodogramme de Lomb-Scargle et l'estimation des autres paramètres avec une optimisation (*Levenberg-Marquardt* [Cum04] ou algorithme génétique) à période fixée pour calculer le résidu. Comme tout algorithme glouton, de telles méthodes souffrent de problèmes de fausses détections. [RAR16] montre par exemple qu'une précédente détection grâce à de telles méthodes étaient en fait due à la fenêtre spectrale liée à l'échantillonnage irrégulier. Les tentatives d'améliorations que nous présenterons brièvement en § 4.3.2, basées sur nos travaux précédents [Bou14], ne seront pas suffisantes pour résoudre ces problèmes.

**Échantillonnage stochastique MCMC.** La communauté astrophysique s'est tournée en parallèle vers une formulation bayésienne du problème d'estimation des paramètres orbitaux, permettant une exploration stochastique de l'espace des paramètres, et une évaluation de l'incertitude sur les paramètres estimés. La majeure partie des travaux repose sur l'échantillonnage stochastique de la loi *a posteriori* des paramètres orbitaux connaissant les vitesses radiales $p(\boldsymbol{P}, \boldsymbol{e}, \boldsymbol{K}, \boldsymbol{\omega}^0, \boldsymbol{T}^{\mathrm{P}}, v_0 \,|\, \boldsymbol{y})$, où $(\boldsymbol{P}, \boldsymbol{e}, \boldsymbol{K}, \boldsymbol{\omega}^0, \boldsymbol{T}^{\mathrm{P}})$ sont de dimension $L$ fixée [For06]. L'espace des paramètres étant vaste ($5L+1$ paramètres, sans compter de possibles hyperparamètres) et donc difficile à échantillonner, les astrophysiciens se sont tournés vers des algorithmes MCMC évolués permettant une convergence rapide et une exploration la plus efficace possible de l'espace des paramètres. Nous avons répertorié trois principaux types d'algorithmes MCMC :

1. le *Parallel Tempering*, notamment dans les travaux de Gregory [Gre05, Gre07a, Gre07b]. Le principe est de lancer plusieurs chaînes de Markov en parallèle, l'une échantillonnant la loi cible, et les autres échantillonnant une loi similaire, mais dont la vraisemblance sera adaptée avec des *températures différentes* : plus la température sera élevée et plus les propositions seront acceptées facilement. Des échanges entre les états des différentes chaînes permettent alors d'explorer de façon plus efficace l'espace des paramètres.
2. les échantillonneurs de Métropolis-Hastings adaptatifs et le `emcee`. Des échantillonneurs de type Métropolis-Hastings adaptatif (voir [HST01]) ont été utilisés par [Tuo12, TAEG+13] et permettent d'adapter la loi de proposition en fonction des échantillons précédemment tirés et d'obtenir des taux d'acceptations corrects. Le `emcee`, *Ensemble samplers with affine invariance*, un échantillonneur initialement proposé par [GW10] et mis à la disposition de la communauté en Python par [FMHLG13], est l'échantillonneur le plus utilisé par la communauté. Il permet d'échantillonner l'espace des paramètres de façon efficace en proposant des modifications affines des lois de propositions, diminuant la corrélation entre

les échantillons.

3. le *Diffusive Nested Sampling* (DNS) introduit par [S+06] et utilisé pour la détection d'exoplanètes par [Bre14] et [FHB+16]. Il s'agit d'un algorithme d'échantillonnage stochastique qui n'échantillonne pas directement la loi *a posteriori*, mais plutôt un mélange des lois *a priori* contraintes par la vraisemblance. L'idée est de tirer des échantillons suivant la loi *a priori* puis de ne garder que les échantillons de vraisemblance supérieure à un seuil, qui est augmenté au fur et à mesure. Après une étape de pondération, les échantillons sont supposés suivre la loi *a posteriori*. Les auteurs prônent l'utilisation de cette méthode car elle permet une exploration plus efficace de l'espace des paramètres, notamment dans le cas de lois bimodales.

Notons que quel que soit l'échantillonneur utilisé, il s'agit ici de considérer un modèle paramétrique, où le nombre d'exoplanètes $L$ est fixé. Le problème de « label switching » (permutation des paramètres) se pose : permuter les paramètres $\boldsymbol{\Theta}_\ell$ entre les planètes n'a pas d'influence sur le signal modélisé. Cela a notamment une conséquence sur les périodes $(P_\ell)_\ell$ ; un paramètre $P_\ell$ peut correspondre à des exoplanètes distinctes suivant les itérations considérées. La méthodologie pour éviter des problèmes est soit de contraindre chaque période $P_\ell$ dans des intervalles distincts grâce à une loi *a priori* appropriée (comme le propose [Tuo12]), soit de ré-ordonner les exoplanètes après échantillonnage [Gre07b]. Enfin, les algorithmes d'échantillonnage MCMC sont la plupart du temps initialisés par une détection des périodes avec une approche de type périodogramme/glouton [HCCQ+14] ; le plus souvent avec périodogramme GLS de [ZK09]. Notons que normalement, les résultats de ces échantillonneurs ne devraient pas dépendre de l'initialisation.

**$\ell_1$-periodogram.** Tandis que la plupart des méthodes de traitement des vitesses radiales sont paramétriques, ce n'est que très récemment que les concepts d'approximation parcimonieuse sur un modèle non-paramétrique ont été introduits pour la détection d'exoplanètes. Ainsi, [HBLC16] a abordé le problème de détection d'exoplanètes comme un problème d'approximation parcimonieuse linéaire, de façon très similaire au problème de l'analyse spectrale en échantillonnage irrégulier étudié par [Bou05]. Un modèle de trajectoire circulaire est utilisé et les périodes des exoplanètes sont discrétisées sur une grille arbitrairement fine, si bien que l'estimation des périodes (ou plutôt des fréquences dans ce modèle) est effectuée en résolvant le problème d'optimisation avec pénalisation $\ell_1$ vue en § 1.2.2 (p.11). La différence majeure avec les travaux de [Bou05] et le logiciel `Sparspec`[9] [BCB07] est la prise en compte des travaux récents de [TBSR13, TBR13] sur l'utilisation de la norme atomique pour obtenir une estimation continue des paramètres continus. L'utilisation de la norme atomique n'étant pas envisageable pour des signaux irrégulièrement échantillonnés, [HBLC16] utilise les travaux de [TBR13] préconisant qu'une discrétisation très fine du paramètre continu, alliée à un lissage de type *Moving-Average* est aussi efficace que l'utilisation de la norme atomique. La solution lissée du problème avec pénalisation $\ell_1$ est alors présentée comme un *périodogramme*, en lien avec le GLS amplement utilisé par la communauté. Cette solution souffre néanmoins des mêmes problèmes que nous avons déjà évoqués en § 1.2.2, dus à la forte corrélation du dictionnaire et entraînant de possibles fausses détections. Par ailleurs, il est pour le moment impossible de prendre en compte des modèles képlériens ($e > 0$) ; les harmoniques des signaux périodiques peuvent donc être à tort considérés comme des exoplanètes.

**Bilan.** Nous avons vu diverses méthodes utilisées par la communauté astrophysique pour l'estimation des paramètres orbitaux avec le nombre de planètes $L$ fixé. Actuellement, les astrophysiciens s'attaquent à deux problèmes majeurs qui sont présents quelle que soit la méthode

[9] Logiciel disponible à l'adresse suivante : `http://www.ast.obs-mip.fr/Softwares`.

considérée : (i) la prise en compte de l'activité stellaire, pouvant ajouter au signal de vitesses radiales du bruit corrélé et même des périodicités dues à l'activité ou à la vitesse de rotation de l'étoile et non aux planètes et (ii) la détermination du nombre $L$ d'exoplanètes. Ces deux enjeux majeurs sont présentés dans les paragraphes suivants.

### 4.1.3 L'activité stellaire, un véritable enjeu

Nous avons vu en § 4.1.1 que la vitesse radiale d'une étoile était mesurée en la reliant avec des décalages de raies d'absorption du spectre de l'étoile. Ces décalages de raies sont principalement dus, via l'effet Doppler, au déplacement par rapport à l'observateur de l'étoile, attirée par une/des exoplanète(s), entraînant le modèle képlérien de la vitesse radiale vu en équation (4.5). Cependant, des mouvements relatifs à l'étoile elle-même, indépendants de l'effet gravitationnel des exoplanètes, peuvent aussi entraîner des décalages, voir des déformations des raies d'absorption de l'étoile via l'effet Doppler. Ces phénomènes, que nous nommerons par la suite *activité stellaire* ou *jitter*, sont récapitulés en Figure 4.5 et seront rapidement décrits au paragraphe suivant. Leur influence sur la vitesse radiale mesurée par spectroscopie peut varier de quelques $cm.s^{-1}$ à plusieurs centaines de $m.s^{-1}$ et donc être du même ordre de grandeur que les contributions apportées par l'influence gravitationnelle des exoplanètes. Il a même été remarqué que la contribution de l'activité stellaire dans le signal de vitesses radiales pouvait imiter la signature d'une exoplanète, de par son caractère généralement pseudo-périodique [QHS+01, RAO+15]. Au début des recherches d'exoplanètes, cette activité stellaire ne gênait pas particulièrement leur détection car les spectrographes n'était pas suffisamment précis pour y être sensible. Cependant, elle représente maintenant une pollution majeure des vitesses radiales pour la détection d'exoplanètes similaires à la Terre nécessitant des spectrographes de précision de l'ordre du $cm.s^{-1}$, comme ESPRESSO vu précédemment ; mais aussi pour les recherches autour d'étoiles plus actives, comme cela peut être le cas des naines rouges observées par SPIRou. Il s'agit donc d'un véritable enjeu d'actualité pour la communauté astrophysique.

Nous récapitulons ici les processus gouvernant l'activité stellaire, les indicateurs permettant de mieux la contraindre et enfin les diverses méthodes utilisées actuellement par les astrophysiciens pour différencier l'activité stellaire de la contribution des exoplanètes.

**Différents types d'activité stellaire.** Nous nous basons sur la Figure 4.5 ainsi que sur la description associée de [Héb15] pour décrire succinctement les différents types d'activité stellaire et comprendre en quoi il est difficile d'obtenir une modélisation analytique de telles activités, similaire à celle que l'on peut avoir pour la contribution des exoplanètes :

- les pulsations sont des modes propres d'oscillations acoustiques excités par la convection turbulente à la surface de l'étoile.
- les granulations sont des phénomènes convectifs à la surface de l'étoile, organisés dans des cellules convectives, où la matière chaude remonte vers la surface tandis que la matière froide, plus dense, redescend. Ces déplacements de matière sont présents sur l'ensemble de la surface de l'étoile, et entraînent globalement un décalage vers le bleu du spectre de l'étoile.
- les *taches stellaires* (*dark spot* ou *bright plage*) sont des tâches (sombres ou brillantes) qui apparaissent à la surface de l'étoile et peuvent durer quelques jours à quelques mois. C'est en majeure partie la rotation de l'étoile sur elle-même qui induit, par effet Doppler, un décalage et une déformation des raies spectrales, et ainsi une contribution sur la vitesse radiale.
- le cycle magnétique de l'étoile, durant généralement plusieurs années (inversion des pôles magnétiques du Soleil tous les 11 ans par exemple), influe notamment sur l'apparition des tâches stellaires.

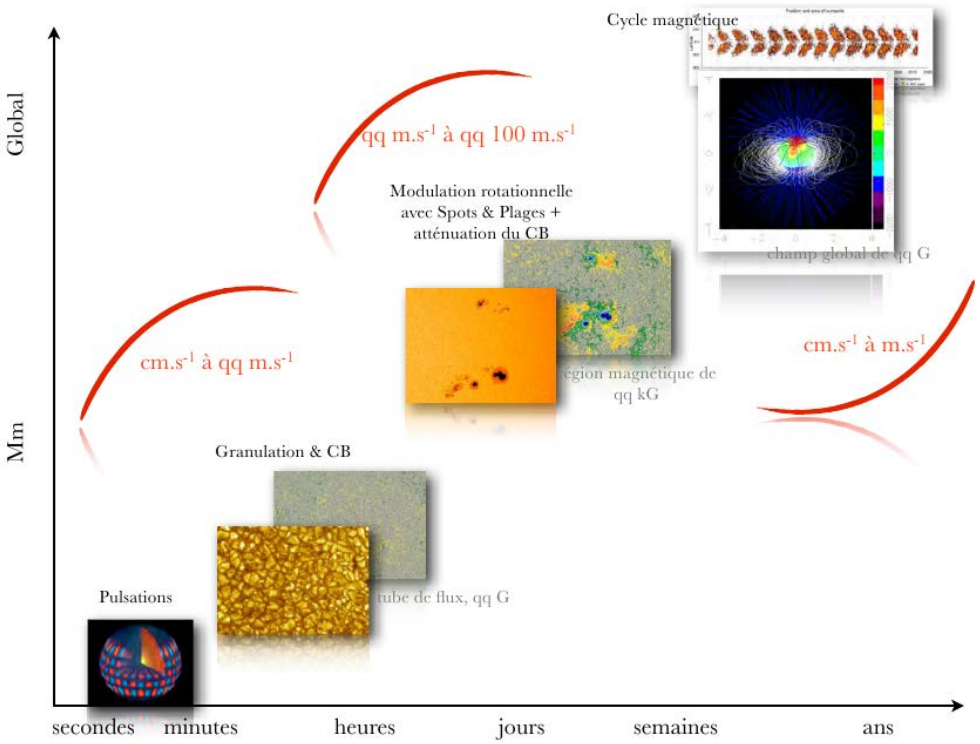



FIG. 4.5 – Représentation des différentes sources d'activité stellaire en fonction de l'ordre de grandeur de la source (ordonnées) et de la durée de vie/d'influence des évènements en question (abscisses), avec un ordre de grandeur de l'influence sur la vitesse radiale de l'étoile mesurée par spectroscopie. Source : [Héb15] (Fig. 1.15, p.27).

L'influence de ces phénomènes d'activité stellaire sur la mesure de la vitesse radiale est particulièrement difficile à modéliser, pour plusieurs raisons : (i) l'éloignement des étoiles empêche de distinguer deux points distincts à leur surface ; les connaissances sur leur activité viennent donc de l'observation du Soleil, qui ne peut se généraliser à l'ensemble des étoiles, et (ii) ces phénomènes, notamment l'apparition et la disparition des tâches, sont de durée finie (contrairement à la contribution des exoplanètes supposée à l'état stationnaire).

Cependant, certaines informations peuvent être obtenues à partir d'observations ne dépendant pas de la présence des exoplanètes ; par exemple, l'activité stellaire peut induire des déformations de la Cross-Correlation Function (CCF) permettant de mesurer la vitesse radiale, illustrées en Figure 4.6. De telles déformations sont mesurées avec des indicateurs tels que le « BIS SPAN » (indicateur utilisant le bissecteur) et le FWHM (Largeur à mi-hauteur). L'indicateur $\log(R'_{HK})$ donne quant à lui une indication sur l'activité magnétique de l'étoile.

**Modélisation adoptée.** Un modèle analytique de l'activité stellaire n'étant pas réalisable, la contribution de l'activité stellaire au signal de vitesses radiales mesurées se retrouvent donc dans le bruit. On écrit alors :

$$\boldsymbol{y} = \boldsymbol{v}_{\text{kep}} + \boldsymbol{\epsilon}_{\text{mes}} + \boldsymbol{\epsilon}_{\text{act}}$$

où $\boldsymbol{v}_{\text{kep}}$ représente la contribution des exoplanètes, $\boldsymbol{\epsilon}_{\text{mes}}$ les erreurs de mesure de la vitesse radiale, et $\boldsymbol{\epsilon}_{\text{act}}$ les contributions provenant de l'activité stellaire.

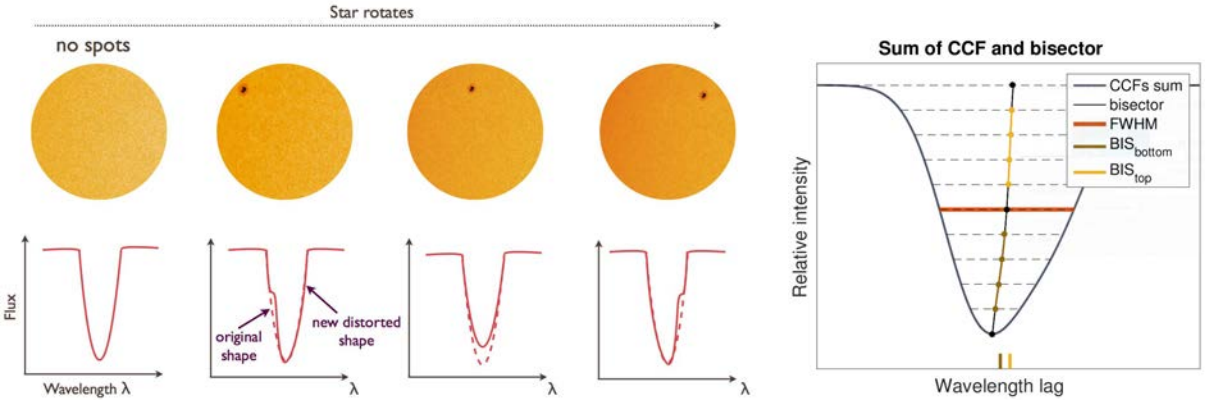



FIG. 4.6 – Illustration de l'influence de l'activité stellaire sur la CCF et des indicateurs. À gauche : déformation de la CCF suivant l'évolution d'une tache à la surface de l'étoile. Source : [Hay15] (Fig. 1.4, p.15). À droite : illustration des indicateurs FWHM et BIS d'une CCF déformée. Source : [Har17] (Fig. 1.11, p.15).

La première solution choisie par les astrophysiciens est de considérer que le bruit total $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}_{\text{mes}} + \boldsymbol{\epsilon}_{\text{act}}$ est gaussien indépendant non-identiquement distribué, mais de variance plus forte que les erreurs de mesures pour prendre en compte l'activité stellaire. Cela revient à écrire :

$$\epsilon_n \sim \mathcal{N}(0, \sigma_n^2 + \sigma_{\text{jit}}^2)$$

où les variances $\sigma_n^2$ sont relatives aux erreurs de mesures à chaque instant $t_n$ tandis que la variance $\sigma_{\text{jit}}^2$ ($\cdot_{\text{jit}}$ pour jitter) ne dépend pas des échantillons et est relative à l'activité (voir [FBH11] par exemple).

Cependant, les contributions liées à l'activité stellaire ne sont pas indépendantes ; leurs variations sont liées à l'accroissement et à la disparition des zones actives de l'étoile, le tout modulé par la période de rotation de l'étoile. Elles induisent des corrélations entre les différents instants de mesures $t_n$ de la vitesse radiale, rendant l'hypothèse d'un bruit indépendant caduc. Pour modéliser un bruit corrélé, un des moyens les plus populaires dans la communauté astrophysique est d'utiliser un *processus gaussien* stationnaire, en se basant sur les travaux de [RW06]. Pour résumer, il s'agit de modéliser le bruit lié à l'activité $\boldsymbol{\epsilon}_{\text{act}}$ comme un processus gaussien, toujours de moyenne nulle, mais dont la matrice de covariance $\boldsymbol{\Sigma}_{\text{act}}$ n'est plus diagonale et est gouvernée par une fonction de covariance $R(\cdot)$. On écrit alors :

$$\boldsymbol{\epsilon}_{\text{act}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\text{act}}) \quad \text{avec} \quad \boldsymbol{\Sigma}_{\text{act}}(m,n) = R(t_m, t_n; \boldsymbol{\theta}_{\text{act}})$$

où les $\boldsymbol{\theta}_{\text{act}}$ sont les hyperparamètres de la fonction de covariance. Grâce à cette modélisation, les $\boldsymbol{\theta}_{\text{act}}$ sont suffisants pour caractériser la matrice de covariance $\boldsymbol{\Sigma}_{\text{act}}$ ; il n'est pas nécessaire de connaître ses $N \times N$ composantes. Différentes fonctions de covariance sont utilisées dans la littérature pour modéliser l'activité stellaire, dont un aperçu est donné dans [RAO+15]. Les fonctions de covariance de type *quasi-périodique*, utilisées par [HCCQ+14], sont de la forme[10] :

$$R(t_m, t_n; \boldsymbol{\theta}_{\text{act}}) = \theta_{\text{act},1}^2 \exp\left( -\frac{(t_m - t_n)^2}{2\,\theta_{\text{act},2}^2} - \frac{\sin^2\left(\frac{\pi(t_m - t_n)}{\theta_{\text{act},3}}\right)}{2\,\theta_{\text{act},4}^2} \right) \tag{4.8}$$

[10] Le noyau quasi-périodique de l'équation (4.8) garde la même forme suivant les articles où il est utilisé, mais varie généralement au niveau de l'expression des $\boldsymbol{\theta}_{\text{act}}$, qui peuvent être mis ou pas au carré, divisés par 2, etc... La forme mentionnée ici est celle de [NFB+18] et est différente de celles de [HCCQ+14] et [RAO+15].

Ici, les paramètres $(\theta_{\text{act},3}, \theta_{\text{act},4})$ sont liés respectivement à la période et la durée des variations périodiques induites par l'activité stellaire, et $\theta_{\text{act},2}$ est lié au temps d'évolution des variations d'activité stellaire. Les paramètres $\boldsymbol{\theta}_{\text{act}}$ pouvant être interprétés en terme d'activité stellaire, les astrophysiciens tentent de les contraindre via des observations indépendantes de l'influence des exoplanètes, comme les indicateurs d'activité stellaire ($\log(\text{R'}_{HK})$, FWHM, BIS) [RAO+15] ou des mesures photométriques de la luminosité de l'étoile [HCCQ+14].

Finalement, le bruit total $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}_{\text{mes}} + \boldsymbol{\epsilon}_{\text{act}}$ est donc modélisé comme un bruit gaussien centré de matrice de covariance $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{mes}} + \sigma^2_{\text{jit}}\mathbf{I}_N + \boldsymbol{\Sigma}_{\text{act}}(\boldsymbol{\theta}_{\text{act}})$, où $\boldsymbol{\Sigma}_{\text{mes}}$ est la matrice de covariance diagonale des erreurs de mesures $\sigma^2_n$.

Enfin, les divers indicateurs d'activité ($\log(\text{R'}_{HK})$, FWHM, BIS) sont parfois très corrélés avec le signal mesuré de vitesses radiales. Ces indicteurs étant indépendants de la contribution des exoplanètes, il est suggéré de les incorporer linéairement dans le modèle, en y associant aussi des tendances temporelles linéaires et quadratiques (voir [HBLC16] par exemple). Le modèle devient alors :

$$\boldsymbol{y} = \boldsymbol{v}_{\text{kep}} + \boldsymbol{\epsilon}_{\text{mes}} + \boldsymbol{\epsilon}_{\text{act}} + \mathbf{H}_{\text{indic,trend}}\boldsymbol{x}_{\text{indic,trend}}$$

où $\mathbf{H}_{\text{indic,trend}}$ regroupe les colonnes des indicateurs ainsi que les tendances linéaires et quadratiques (colonnes $\boldsymbol{t}$,$\boldsymbol{t}^2$) et $\boldsymbol{x}_{\text{indic,trend}}$ sont les amplitudes linéaires associées.

Par ailleurs, il existe d'autres moyens de prendre en compte de l'activité stellaire, comme la modélisation d'un bruit corrélé via un modèle de type *Moving Average* de [Tuo12, TJJ+13], ou encore la modélisation conjointe de la contribution des exoplanètes et de l'activité stellaire par des képlériennes apodisées [Gre16]. Nous ne détaillons pas d'avantage ces méthodes.

### 4.1.4 Le nombre d'exoplanètes

L'estimation du nombre d'exoplanètes dont la signature est présente dans le signal de vitesses radiales $\boldsymbol{y}$ reste un des problèmes majeurs que rencontre la communauté astrophysique. Chaque type de méthode d'estimation des paramètres orbitaux est associée à une technique d'estimation de l'ordre du modèle. Nous résumons brièvement les trois qui nous semblent les plus importantes.

**Probabilité de fausse alerte (FAP).** Alliée aux méthodes de type périodogrammes et algorithmes gloutons, la FAP permet de vérifier si l'ajout d'une période détectée par périodogramme est suffisamment significative, en calculant la probabilité que le pic du périodogramme soit du à une réalisation particulière du bruit et non à la présence d'une planète. Chaque formulation du périodogramme est généralement donnée avec une expression de la FAP associée ; nous citerons en particulier celles de [Bal08] dans le cas de trajectoire circulaire, et de [Bal14b] dans le cas de trajectoire elliptique.

**Vraisemblance marginale.** La majeure partie de la communauté astrophysique s'intéresse désormais au calcul de la *vraisemblance marginale* (*marginal likelihood*) (ML), notée $\mathcal{Z}$, correspondant à la probabilité d'obtenir la réalisation des données $\boldsymbol{y}$ conditionnellement au modèle (comprenant le nombre d'exoplanètes). Plus précisément, cela correspond à la vraisemblance du nombre d'exoplanètes marginalisée par rapport aux autres paramètres, ce qui s'écrit :

$$\mathcal{Z}_L = \underbrace{p(\boldsymbol{y} \mid L)}_{\text{ML}} = \int_{\boldsymbol{\Theta}_{(L)}} \underbrace{p(\boldsymbol{y} \mid \boldsymbol{\Theta}_{(L)}, L)}_{\text{vraisemblance}} \underbrace{p(\boldsymbol{\Theta}_{(L)} \mid L)}_{\textit{a priori}} \, \mathrm{d}\boldsymbol{\Theta}_{(L)} \tag{4.9}$$

où $\boldsymbol{\Theta}_{(L)}$ regroupe tous les paramètres orbitaux associés aux $L$ exoplanètes du modèle, ainsi que les possibles paramètres liés au bruit et hyperparamètres. Ainsi, pour comparer un modèle à $L$

et $L+1$ exoplanètes, on peut calculer le rapport des lois *a posteriori* du modèle (*Posterior odds ratio*/POR), qui s'écrit, suivant la règle de Bayes :

$$\underbrace{\frac{\Pr(L+1\,|\,\boldsymbol{y})}{\Pr(L\,|\,\boldsymbol{y})}}_{\text{posterior odds ratio}} = \underbrace{\frac{\mathcal{Z}_{L+1}}{\mathcal{Z}_L}}_{\text{facteur de Bayes}} \frac{\Pr(L+1)}{\Pr(L)} \tag{4.10}$$

Si l'on suppose un *a priori* identique pour tous les ordres de modèle $L$, alors le *Posterior odds ratio* se ramène au facteur de Bayes (*Bayes Factor*/BF), comparant les vraisemblances marginales des deux modèles. Dans la communauté astrophysique, le modèle à $L+1$ exoplanètes est préféré au modèle à $L$ exoplanètes, uniquement si le facteur de Bayes est supérieur à 150 (d'après [Jef61]). Cela traduit le fait que la ML d'un modèle à $L+1$ exoplanètes est de toute façon plus élevée que celle d'un modèle à $L$ exoplanètes puisqu'il influence le modèle tout en autorisant des degrés de liberté supplémentaires. Le calcul de $\mathcal{Z}$ reste cependant une tâche mathématiquement très complexe, puisqu'il s'agit d'intégrer la loi jointe $p(\boldsymbol{y}, \boldsymbol{\Theta}_{(L)})$ sur tout l'espace des paramètres $\boldsymbol{\Theta}_{(L)}$. De nombreuses méthodes sont utilisées dans la communauté exoplanète (voir par exemple [FG07] pour une comparaison de ces méthodes), parmi lesquelles nous citons les trois suivantes :

1. l'algorithme *One-Block Metropolis-Hastings* (OBMH) de [CJ01] est un des premiers outils à avoir été utilisé par la communauté astrophysique [Tuo12, HCCQ+14].
2. les échantillonneurs de type *Importance Sampling* gagnent actuellement en popularité (voir par exemple [NFB+18]) ; ils permettent d'approcher l'intégrale en équation (4.9) en échantillonnant une loi plus simple $g(\boldsymbol{\Theta}_{(L)})$ (dont on connaît la constante de normalisation) que la loi *a posteriori* (par exemple, une loi normale multidimensionnelle dont la moyenne et la matrice de covariance sont estimés à partir d'échantillons de la loi *a posteriori*). En omettant la notation de l'ordre du modèle $L$ et en notant $\boldsymbol{\Theta}^{(i)}$ les $N^I$ échantillons de $g(\boldsymbol{\Theta})$ et $f_{\boldsymbol{y}}(\boldsymbol{\Theta}) = p(\boldsymbol{y}\,|\,\boldsymbol{\Theta})p(\boldsymbol{\Theta})/g(\boldsymbol{\Theta})$, on a en effet :
$$\mathcal{Z} = \int f_{\boldsymbol{y}}(\boldsymbol{\Theta})g(\boldsymbol{\Theta})\,\mathrm{d}\boldsymbol{\Theta} = \mathbb{E}\{f_{\boldsymbol{y}}(\boldsymbol{\Theta})\} \quad \Rightarrow \quad \widehat{\mathcal{Z}} \approx \frac{1}{N^I} \sum_{\boldsymbol{\Theta}^{(i)} \sim g(\boldsymbol{\Theta})} f_{\boldsymbol{y}}(\boldsymbol{\Theta}^{(i)})$$
3. enfin, les échantillonneurs de type *Nested-Sampling*, que nous avons déjà évoqués en § 4.1.2 offrent un cadre particulièrement bien adapté au calcul de la vraisemblance marginale, puisqu'ils permettent d'associer un *niveau de vraisemblance* avec un *volume de l'espace* a priori. Voir [BPC11] pour plus de détails.

**Trans-dimensionnal MCMC** : les algorithmes d'échantillonnage stochastique sont étendus en rajoutant l'ordre du modèle $L$ aux paramètres à échantillonner. L'échantionnage se faisant dans un espace de paramètres dont la dimension est variable, des étapes de type *Naissance/Mort* [Ste00] sur le principe des sauts réversible/*Reversible-Jump* [Gre95] sont alors ajoutées à l'algorithme d'échantillonnage. Dans le cadre de la détection d'exoplanètes, cet échantillonnage de type Reversible-Jump a d'abord été considéré comme très difficile (voir [FG07]) ; actuellement, c'est au sein d'algorithme de type *Diffusive Nested-Sampling* (DNS) vus précédemment que ce changement de dimension est pris en compte (voir [Bre14]). L'estimation de l'ordre du modèle se fait alors simplement en moyennant sur l'ensemble des échantillons tirés par l'échantillonneur. Cette approche a aussi été généralisée par [FHB+16] à la prise en compte de l'activité stellaire via un processus gaussien.

### 4.1.5 Bilan : la parcimonie, une alternative

Comme nous l'avons vu, la détection d'exoplanètes par la méthode des vitesses radiales reste un problème difficile, notamment dans l'optique de pouvoir détecter des planètes de faibles

masses, similaires à la Terre. L'amélioration des spectrographes à haute résolution permet en théorie de pouvoir détecter de telles planètes, mais elle se heurte à la présence de l'activité stellaire, pouvant imiter la signature d'exoplanètes. Aujourd'hui, les nombreuses méthodes de traitement des vitesses radiales reposent pour la plupart sur l'utilisation d'un modèle *paramétrique*, fixant le nombre d'exoplanètes pour l'estimation des paramètres orbitaux puis décidant du nombre d'exoplanètes *a posteriori* en comparant les modèles estimés. De telles stratégies sont très coûteuses et reposent sur des critères de décision généralement empiriques. De plus, notons que les méthodes développées par les astrophysiciens sont parfois difficiles à reproduire et rarement comparées entre elles sur les mêmes jeux de données. Cependant, des *challenges* entre plusieurs équipes d'astrophysiciens ont récemment été lancés pour comparer leurs méthodes de traitement de données sur des données simulées : le *RV Challenge* de [Dum16, DBD+17] vise à évaluer la prise en compte de l'activité stellaire, et [NFB+18] compare les différentes méthodes de calcul de la *vraisemblance marginale* pour l'estimation du nombre d'exoplanètes. Cela montre à la fois la volonté de la communauté astrophysique de comparer leurs diverses méthodes de façon cohérente sur des simulations, mais aussi l'importance de ces enjeux.

La reformulation du problème inverse d'estimation des paramètres orbitaux en problème d'approximation parcimonieuse, en discrétisant sur une grille le paramètre de localisation (ici la période $P$ de rotation des exoplanètes[11]), permet en théorie de passer outre cette comparaison de modèle ; les amplitudes non-nulles résultant de l'estimation permettent en effet de décider quelles périodes (et donc quelles exoplanètes) sont présentes dans le signal. À notre connaissance, seul [HBLC16], avec sa résolution du problème d'APL via pénalisation $\ell_1$, utilise la parcimonie pour l'estimation des paramètres orbitaux, mais avec des excentricités nulles. Le problème, très similaire à l'analyse spectrale en échantillonnage irrégulier, souffre pourtant des mêmes problèmes de corrélation du dictionnaire inhérents aux approches sous-optimales déjà discutés au Chapitre 1.

Nous proposons donc, dans la suite de ce chapitre, de formaliser le problème de détection d'exoplanètes en problème d'APNL que nous proposons de résoudre à l'aide du modèle Bernoulli-Gaussien étendu et de l'échantillonneur associé étudié au chapitre 3. Nous verrons en effet que le modèle BGE s'adapte particulièrement bien à la détection d'exoplanètes, et permet de répondre aux principaux enjeux d'actualité dans ce domaine, à savoir l'estimation du nombre d'exoplanètes et la prise en compte de l'activité stellaire. Le modèle BGE est donc présenté comme une alternative et un complément aux méthodes présentées dans cette section, pouvant permettre de confirmer, mettre en doute ou infirmer l'existence d'exoplanètes détectées avec d'autres méthodes.

## 4.2 Modélisation BGE et échantillonneur associé pour la détection d'exoplanètes

Le problème de la détection d'exoplanètes par la méthode des vitesses radiales peut donc être considéré comme un problème inverse non linéaire défini au Chapitre 1. Il s'agit de trouver le nombre $L$ d'exoplanètes, dont les périodes de rotation $P$ jouent le rôle de paramètre de localisation, et les autres paramètres orbitaux - en dehors de l'amplitude $K$ - jouent celui des paramètres de forme. Nos travaux sur la reformulation en APNL et les méthodes de résolution associées semblent donc bien adaptés à ce problème. Plus particulièrement, le modèle Bernoulli-Gaussien étendu et l'échantillonnage de la loi *a posteriori* que nous avons proposé au Chapitre 3, semble le plus prometteur, puisqu'il donne déjà de bons résultats sur les problèmes d'analyse spectrale en échantillonnage irrégulier (qui peut être vu comme un cas particulier de la détection

[11] Les périodes de rotation des exoplanètes étant globalement liées à leur éloignement vis-à-vis de l'étoile hôte, elles sont généralement bien distinctes - deux exoplanètes ne peuvent pas se trouver sur la même orbite.

d'exoplanètes où ces dernières seraient d'excentricité nulle) et nous avons vu que la prise en compte de paramètres de forme, ainsi qu'un bruit non indépendant, ne posaient pas de problème en théorie. Notons que nos travaux sur l'interpolation de dictionnaire, présentés au Chapitre 2, s'appliquaient déjà difficilement au problème d'analyse spectrale ; leur application à la détection d'exoplanètes semble très délicate ou du moins demanderait une étude beaucoup plus poussée. Aussi nous n'avons pas poursuivi cette piste.

Par conséquent, nous proposons d'employer le modèle BGE et l'échantillonnage associé pour aborder le problème de détection d'exoplanètes. Dans ce but, il est d'abord nécessaire de formuler ce problème en APNL : nous revenons d'abord sur le modèle du problème inverse que nous reformulons de façon légèrement différente de (4.1) afin de minimiser le nombre de paramètres non linéaires, puis nous discutons du domaine de localisation $\mathcal{I}_P = [P^{\min}; P^{\max}]$ et de son découpage, nécessaire pour la reformulation en APNL. À partir de ces choix, l'utilisation du modèle BGE et de l'échantillonneur associé peut être envisagée : après un rapide retour sur les lois *a priori* données aux paramètres non linéaires et sur la loi de proposition pour leur échantillonnage, nous adaptons notre méthode à la nécessité de prendre en compte l'activité stellaire.

### 4.2.1 Reparamétrisation du modèle

Comme défini précédemment en (4.1), l'atome continu du problème inverse non linéaire $\boldsymbol{h}_{\text{kep}}(\cdot)$ est d'amplitude scalaire $K$ et dépend non-linéairement de quatre paramètres non linéaires $(P, e, \omega^0, T^{\text{P}})$. Nous proposons la reparamétrisation suivante, déjà utilisée dans [Bou14] et équivalente à la modélisation initiale, mais présentant seulement 3 paramètres non linéaires $(P, e, \phi)$ :

$$\begin{aligned} K\ \boldsymbol{h}_{\text{kep}}(P, e, \omega^0, T^{\text{P}}) &= \mathbf{h}(P, e, \phi)\ \boldsymbol{x} \\ \text{avec} \quad \mathbf{h}(P, e, \phi) &= \begin{bmatrix} e + \cos \boldsymbol{\rho}(P, e, \phi) & \sin \boldsymbol{\rho}(P, e, \phi) \end{bmatrix} \\ \boldsymbol{x} &= \begin{bmatrix} K \cos \omega^0 \\ -K \sin \omega^0 \end{bmatrix} \in \mathbb{R}^2 \ (m = 2) \\ \phi &= \frac{T^{\text{P}}}{P} \in [0; 1] \end{aligned} \tag{4.11}$$

Cette reparamétrisation consiste en deux modifications : (i) la non-linéarité en $\omega^0$ est supprimée grâce au développement du $\cos(\boldsymbol{\rho}+\omega^0)$ en équation (4.1) et au changement de variables $(K, \omega^0) \leftrightarrow \boldsymbol{x}$ et (ii) la dépendance de $T^{\text{P}}$ en $P$ est annulée par le changement de variables $T^{\text{P}} \in [0; P] \leftrightarrow \phi = \frac{T^{\text{P}}}{P} \in [0; 1]$. Notons que les amplitudes $\boldsymbol{x}$ comme définies en (4.11) peuvent appartenir à tout l'espace $\mathbb{R}^2$ et permettent donc leur marginalisation dans le cadre d'un modèle BG. Ainsi formulé, le modèle ne présente donc que trois paramètres non linéaires contre quatre précédemment, au prix du passage à des amplitudes bi-dimensionnelle $\boldsymbol{x}$, mais cela ne pose pas de problème pour l'utilisation du modèle BGE et des algorithmes associés.

Cependant, notons que dans le cas d'une orbite circulaire ($e = 0$), le modèle s'écrit de la façon suivante :

$$\mathbf{h}(P, e = 0, \phi)\ \boldsymbol{x} = \begin{bmatrix} \cos(2\pi(\frac{\boldsymbol{t}}{P} - \phi)) & \sin(2\pi(\frac{\boldsymbol{t}}{P} - \phi)) \end{bmatrix}\ \boldsymbol{x}$$

Dans ce cas, nous avons une perte d'identifiabilité, car de nombreux modèles sont équivalents pour différentes valeurs de $(\boldsymbol{x}, \phi)$. Par exemple, à $P$ fixé, $(\boldsymbol{x}, \phi = 1/2)$ donne le même modèle que $(-\boldsymbol{x}, \phi = 0)$. On pourrait donc sans perte d'information considérer que $\phi = 0$ dans le cas d'une orbite circulaire, mais nous verrons qu'une telle dépendance entre les paramètres ne sera pas prise en compte dans notre algorithme d'échantillonnage. Cela ne posera pas de problèmes puisque l'on pourra considérer, lors de l'étape d'estimation, que $\phi = 0$ si $e = 0$.

En utilisant cette reparamétrisation pour (4.1) et en notant $\boldsymbol{\nu}_\ell = (P_\ell, e_\ell, \phi_\ell)$, $x_0 = v_0$, le problème inverse non linéaire s'écrit alors de la façon suivante :

$$\text{Estimer } L \text{ faible et les } \{\boldsymbol{\nu}_\ell, \boldsymbol{x}_\ell\}_\ell, \text{ tels que } \boldsymbol{y} = x_0 + \sum_{\ell=1}^{L} \mathbf{h}(\boldsymbol{\nu}_\ell)\boldsymbol{x}_\ell + \boldsymbol{\epsilon}$$

Les paramètres de forme de ce problème sont donc $\boldsymbol{\upsilon}_\ell = (e_\ell, \phi_\ell)$ et le paramètre de localisation $\nu_\ell = P_\ell$ (nous utiliserons donc la notation $P_\ell$ pour le paramètre de localisation dans ce chapitre).

Notons qu'à notre connaissance, cette reparamétrisation du problème n'est pas utilisée dans la communauté astrophysique. Une autre linéarisation du modèle, très proche de la nôtre, est généralement utilisée (voir par exemple [WH09]). Elle consiste à récupérer dans la composante continue $x_0$ à la fois la vitesse $v_0$ mais aussi l'ensemble des composantes continues $K_\ell e_\ell \cos \omega^0{}_\ell$, si bien que les amplitudes $\boldsymbol{x}$ s'exprimeraient toujours de la même façon en fonction de $(K, \omega^0)$ mais l'atome continu serait $\mathbf{h}(\boldsymbol{\nu}) = [\cos \boldsymbol{\rho}(\boldsymbol{\nu}) \;\; \sin \boldsymbol{\rho}(\boldsymbol{\nu})]$. Nous préférons cependant notre paramétrisation car elle permet de considérer $x_0$ indépendamment des autres paramètres.

Une autre paramétrisation est conseillée par [For06] et est notammment utilisée pour les méthodes à base d'échantillonnage MCMC (par exemple [HCCQ+14]). Elle consiste en le changement de variables $(c, s) = (\sqrt{e}\cos\omega^0, \sqrt{e}\sin\omega^0) \leftrightarrow (e, \omega^0)$ et permettrait notamment d'éviter un biais d'estimation sur les excentricités. Cependant, une telle paramétrisation conserve 4 paramètres non linéaires $(P, T^{\mathrm{P}}, c, s)$ et de plus empêche l'exploitation de la marginalisation des amplitudes, puisque ces dernières, qui sont alors les semi-amplitudes $K_\ell$, sont positives.

### 4.2.2 Reformulation en APNL via le découpage du domaine des localisations

Pour reformuler le problème inverse non linéaire (4.2.1) que nous avons décrit précédemment, il est nécessaire de découper le domaine des localisations $\mathcal{I}_\nu$ en $J$ intervalles $\mathcal{I}_j$. Le paramètre de localisation étant la période $P$ pour ce problème, le domaine $\mathcal{I}_\nu$ s'écrit $\mathcal{I}_P = [P^{\min}; P^{\max}]$. En supposant le découpage effectué, on obtient alors le problème d'APNL suivant :

$$\begin{aligned} &\text{Estimer } x_0,\ \boldsymbol{x} \textit{ parcimonieux} \text{ et } \boldsymbol{\nu} \in \mathfrak{I} \text{ tels que } \boldsymbol{y} = x_0 + \sum_{j=1}^{J} \mathbf{h}(\boldsymbol{\nu}_j)\boldsymbol{x}_j + \boldsymbol{\epsilon} \\ &\text{où} \quad \boldsymbol{\nu} \in \mathfrak{I} \quad \Leftrightarrow \quad \forall j \in [\![1; J]\!],\ \nu_j = P_j \in \mathcal{I}_j \text{ et } \boldsymbol{\upsilon}_j = (e_j, \phi_j) \in \mathcal{I}_{\boldsymbol{\upsilon}} = [0; 1[^2 \end{aligned} \tag{4.12}$$

Nous détaillons maintenant le choix des bornes $(P^{\min}, P^{\max})$ pour le domaine des localisations, ainsi que le découpage choisi. Ces choix doivent en effet être effectués judicieusement puisqu'on ne pourra détecter qu'une exoplanète par intervalle $\mathcal{I}_j$.

**Bornes du domaine des localisations.** Le choix des bornes $P^{\min}$ et $P^{\max}$ va dépendre des données de vitesses radiales utilisées. La précision actuelle des spectrographes hautes résolutions permettent *a priori* de détecter des planètes dont la période vaut au minimum 1 jour ; c'est souvent la limite minimal prise par les astrophysiciens. Pour la période maximale, elle sera souvent prise de l'ordre d'une ou deux fois l'intervalle de temps des mesures, soit $P^{\max} \propto T$ où $T = t(N) - t(1)$.

**Découpage de $\mathcal{I}_P$.** Une fois $P_{\min}$ et $P_{\max}$ déterminés, il est nécessaire de choisir un découpage judicieux du domaine des localisations $\mathcal{I}_P$. Au regard des discussions effectuées pour le découpage de l'espace des fréquences en analyse spectrale avec échantillonnage irrégulier (voir § 1.5.1), il semblerait naturel de considérer la fréquence $f = \frac{1}{P}$ plutôt que la période pour le problème (4.12).

Nous pourrions alors, comme en § 1.5.1, choisir un découpage de $\mathcal{I}_f = [1/P^{\max}; 1/P^{\min}]$ en $J$ intervalles $\mathcal{I}_j$ de largeur régulière $\Delta$, telle que cette dernière soit au maximum de la taille de la largeur du lobe principale de la fenêtre spectrale $\Delta \leq \Delta\max = 2/T$ ; les intervalles $\mathcal{I}_j$ ne pourraient alors contenir qu'une seule fréquence - et donc une seule exoplanète.

Cependant, la communauté astrophysique utilise le plus souvent une reparamétrisation en logarithme de la période, que ce soit pour les estimations via périodogramme ou par échantillonnage MCMC. Nous interprétons ce choix comme un compromis entre *résolution numérique* et *résolution physique*. En effet, la résolution que donne la fenêtre spectrale n'est pas nécessairement corrélée à la « distance en période » possible entre 2 planètes. Une illustration simple de ce compromis peut s'observer sur le système planétaire que nous connaissons le mieux : le système solaire. Imaginons que des extraterrestres veulent détecter des « exo »planètes autour de l'étoile Soleil, en mesurant la vitesse radiale de cette dernière depuis leur propre planète. En prenant des valeurs extremum $P_{\min} = 1$ jours et $P_{\max} = 60000$ jours encadrant les périodes des planètes du système solaire ($P = 58$ jours pour Mercure, $P = 59860$ jours pour Neptune), on peut construire une grille de $J = 100$ points (taille précédemment utilisée pour le BGE dans le cadre de l'analyse spectrale au Chapitre 3), d'au moins 3 manières différentes : régulière en périodes, en fréquences ou en logarithme. La répartition des périodes des planètes du système solaire sur de telles grilles (rangée dans l'ordre croissant en période) est représentée en Figure 4.7. Parmi les 3 propositions de grilles, seule la grille en logarithme permet de séparer les 8 planètes dans 8 intervalles différents ; la grille en période n'est pas assez résolue pour les petites périodes, tandis que la grille en fréquence n'est pas assez résolue pour les périodes les plus grandes. Pour que la grille en fréquence sépare correctement les 8 périodes, il serait nécessaire de considérer un découpage en plus de $J = 60000$ intervalles ; une exploitation de notre modèle BGE est inenvisageable dans ces conditions.

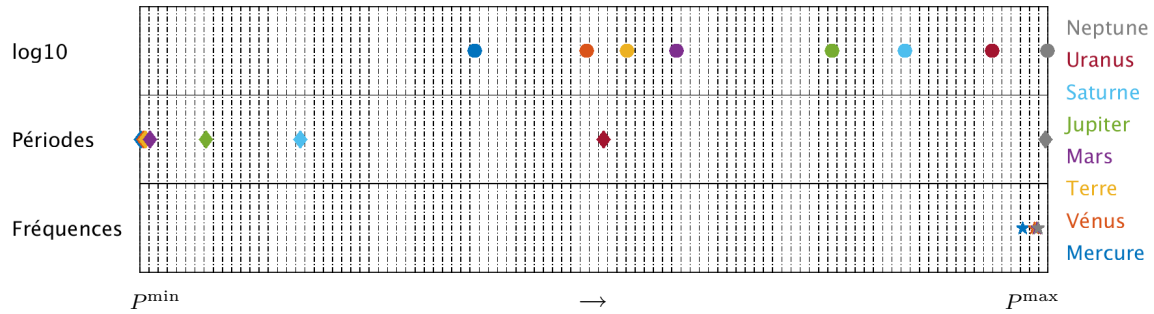



FIG. 4.7 – Répartition des planètes du système solaire sur une grille de $J = 100$ points entre $P_{\min} = 1$ jours et $P_{\max} = 60000$ jours, suivant les différentes répartitions : régulière en log10, périodes et fréquences.

Ces observations nous ont conduit à choisir une grille régulière en logarithme, ce qui revient à considérer le logarithme de la période dans le problème (4.12).

*Remarque : Par la suite, nous continuerons à parler du paramètre de localisation avec la dénomination de « période » bien qu'en pratique, le paramètre échantillonné sera un logarithme de période.*

### 4.2.3 Modèle BGE et échantillonnage associé

À partir du problème d'APNL formalisé en (4.12), la description du modèle BGE se fait sans souci. On introduit la séquence de Bernoulli $\boldsymbol{q}$ telle que chaque $q_j$ indique la présence ou l'absence d'une exoplanète dans l'intervalle de période $\mathcal{I}_j$. Les paramètres non linéaires $\boldsymbol{\nu}_j = (P_j, e_j, \phi_j)$

sont supposés mutuellement indépendants[12] et leur loi *a priori* est une loi uniforme sur leur intervalle de définition comme vu dans la description théorique du modèle BGE et n'introduisent pas d'hyperparamètres supplémentaires. Ainsi, la période[13] $P_j$ suit la loi *a priori* $\mathcal{U}(\mathcal{I}_j)$, tandis que les paramètres de forme $e_j$ et $\phi_j$ suivent la loi *a priori* $\mathcal{U}([0;1[)$. Les amplitudes $\boldsymbol{x}_j$ associées aux paramètres $q_j$ suivent une loi *a priori* Bernoulli-Gaussienne d'hyperparamètres $(\lambda, \sigma_x^2)$, avec l'extension à la composante continue $x_0$ effectuée lors de notre application à l'analyse spectrale en § 3.5.1. Pour pouvoir exploiter la marginalisation des amplitudes, nous supposerons que le bruit $\boldsymbol{\epsilon}$ suive une loi gaussienne de matrice de covariance $\boldsymbol{\Sigma}$, qui pourra être reliée avec l'activité stellaire - nous en parlons plus en détails par la suite.

La loi *a posteriori* jointe à partir de laquelle seront échantillonnés puis estimés les paramètres d'intérêt est donc, dans le cadre non supervisé, $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \lambda, \sigma_x^2, \boldsymbol{\Sigma} \,|\, \boldsymbol{y})$. La mise en œuvre de l'échantillonnage se fera suivant l'algorithme de type PCGS (Tab. 3.10 p. 128) exploitant la marginalisation des amplitudes $\boldsymbol{x}$ pour les paramètres $q_j$ et $\boldsymbol{\nu}_j$. Pour l'échantillonnage des $\boldsymbol{\nu}_j$ (qui pour rappel s'effectue suivant une étape de Métropolis-Hastings), nous conservons la loi de proposition décrite en (3.16) : les $(P'_j, e'_j, \phi'_j)$ sont proposés simultanément selon une loi de mélange entre marche aléatoire autour de l'échantillon courant et tirage suivant leur loi *a priori*. Notons qu'il est alors nécessaire de régler les paramètres $p_{\text{rw}}$ (probabilité d'effectuer la marche aléatoire plutôt que le tirage suivant la loi *a priori*) et les variances des marches aléatoires, pouvant être différente pour chaque paramètre : $(\sigma^2_{P,\text{rw}}, \sigma^2_{e,\text{rw}}, \sigma^2_{\phi,\text{rw}})$. Sous ces conditions, la probabilité d'acceptation a toujours la même expression qu'en (3.20) (rapport des lois *a posteriori* conditionnelles marginalisées).

Notons que l'on pourrait considérer des lois *a priori* différentes pour les paramètres de forme (par exemple, [Tuo12] utilise une loi *a priori* gaussienne tronquée en 0 pour favoriser *a priori* les faibles excentricités) ; il faudrait alors prendre en compte ces lois *a priori* dans la probabilité d'acceptation, ce qui se fait sans soucis. De plus, la proposition simultanée de $(P, e, \phi)'_j$ suivant la marche aléatoire ou la loi *a priori* pourrait être remplacée par une proposition indépendante de chaque paramètre suivant la loi de mélange, autorisant par exemple $P'_j$ a être proposé suivant la marche aléatoire et $(e'_j, \phi'_j)$ suivant leur loi *a priori*. Une telle loi de proposition est toujours symétrique et l'expression de la probabilité d'acceptation reste inchangée. En pratique, les résultats obtenus avec ces deux stratégies étaient similaires.

Le PCGS s'applique donc sans modification au modèle BGE des exoplanètes, à la différence près qu'il est nécessaire de prendre en compte la présence d'activité stellaire.

### Prise en compte de l'activité stellaire dans le bruit

Nous avons vu en § 4.1.3 que le bruit $\boldsymbol{\epsilon}$ présent dans le signal de vitesses radiales $\boldsymbol{y}$ peut se modéliser comme un bruit gaussien centré dont la matrice de covariance peut se décomposer en 3 termes :

$$\begin{aligned}
\boldsymbol{\Sigma} &= \boldsymbol{\Sigma}(\sigma^2_{\text{jit}}, \boldsymbol{\theta}_{\text{act}}) = \boldsymbol{\Sigma}_{\text{mes}} + \boldsymbol{\Sigma}_{\text{jit}}(\sigma^2_{\text{jit}}) + \boldsymbol{\Sigma}_{\text{act}}(\boldsymbol{\theta}_{\text{act}}) \\
\text{avec} \quad \boldsymbol{\Sigma}_{\text{mes}} &= \text{diag}(\sigma^2_1, \ldots, \sigma^2_N) \quad \text{(erreurs de mesure, supposées connues)} \\
\boldsymbol{\Sigma}_{\text{jit}}(\sigma^2_{\text{jit}}) &= \sigma^2_{\text{jit}} \mathbf{I}_N \quad \text{(jitter d'activité)} \\
\boldsymbol{\Sigma}_{\text{act}}(\boldsymbol{\theta}_{\text{act}}) &: \text{bruit corrélé induit par l'activité stellaire, modélisé par} \\
&\quad \text{un processus gaussien de paramètres } \boldsymbol{\theta}_{\text{act}} \text{ (voir (4.8)).}
\end{aligned}$$

[12] L'hypothèse d'indépendance entre $(P, e, \phi)$ est réalisée pour des raisons pratiques. En théorie, la période $P$ et l'excentricité $e$ dépendent physiquement l'une de l'autre, car il est généralement peu probable d'avoir de fortes excentricités pour des petites périodes (le système planétaire ne répondrait pas à des critères de stabilité). De même, plus l'excentricité $e$ est faible et moins la phase $\phi$ a une signification physique.

[13] Pour rappel, en pratique nous travaillons en logarithme de période, si bien que la loi *a priori* n'est pas uniforme en période mais bien en logarithme.

Suivant notre connaissance de l'activité de l'étoile, nous distinguons ici 3 possibilités :

1. Absence d'activité stellaire : $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{mes}}$, on échantillonne alors $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \lambda, \sigma_x^2 \,|\, \boldsymbol{y}, \boldsymbol{\Sigma}_{\text{mes}})$ (pas d'échantillonnage lié au bruit).
2. Activité stellaire correctement modélisée par le *jitter* : $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{mes}} + \sigma_{\text{jit}}^2 \mathbf{I}_N$. On échantillonne alors $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \lambda, \sigma_x^2, \sigma_{\text{jit}}^2 \,|\, \boldsymbol{y}, \boldsymbol{\Sigma}_{\text{mes}})$ si $\sigma_{\text{jit}}^2$ est inconnue, $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \lambda, \sigma_x^2 \,|\, \boldsymbol{y}, \boldsymbol{\Sigma}_{\text{mes}}, \sigma_{\text{jit}}^2)$ sinon.
3. Activité stellaire modélisée par un *jitter* et un processus gaussien : $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{mes}} + \sigma_{\text{jit}}^2 \mathbf{I}_N + \boldsymbol{\Sigma}_{\text{act}}(\boldsymbol{\theta}_{\text{act}})$. On échantillonne alors $p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \lambda, \sigma_x^2, \sigma_{\text{jit}}^2, \boldsymbol{\theta}_{\text{act}} \,|\, \boldsymbol{y}, \boldsymbol{\Sigma}_{\text{mes}})$.

Pour le premier cas, nous utiliserons l'échantillonnage de type PCGS (Tab. 3.10 p. 128) proposé au chapitre précédent, mais sans l'étape d'échantillonnage de la variance du bruit. Notons que la présence du bruit non identiquement distribué justifie la présence de la matrice $\boldsymbol{\Sigma}$ dans nos calculs en § 3.3. Le deuxième et le troisième cas nécessitent des adaptations.

En effet, même en l'absence de bruit corrélé (deuxième cas), la variance du jitter d'activité $\sigma_{\text{jit}}^2$ ne peut pas s'échantillonner directement suivant sa loi conditionnelle, comme c'était le cas avec la variance $\sigma_\epsilon^2$ du bruit gaussien centré *i.i.d.* dans le chapitre précédent. Si l'on reprend l'expression de la loi *a posteriori* jointe (pour rappel, les amplitudes $\boldsymbol{x}$ ont été tirées au préalable suivant leur loi *a posteriori* conditionnelle afin de pouvoir échantillonner $\sigma_x^2$ suivant la loi jointe), on a :

$$
\begin{aligned}
p(\sigma_{\text{jit}}^2 \,|\, \boldsymbol{y}, \boldsymbol{x}, -) &\propto |\boldsymbol{\Sigma}|^{-1/2} \exp(-\frac{1}{2}\|\boldsymbol{\zeta}\|^2_{\boldsymbol{\Sigma}^{-1}})\; p(\sigma_{\text{jit}}^2) \\
\text{Or} \quad \boldsymbol{\Sigma} &= \text{diag}(\sigma_{\text{jit}}^2 + \sigma_1^2, \dots, \sigma_{\text{jit}}^2 + \sigma_N^2) \\
|\boldsymbol{\Sigma}| &= \prod_{n=1}^{N} (\sigma_{\text{jit}}^2 + \sigma_n^2), \quad \text{polynôme d'ordre N en } \sigma_{\text{jit}}^2 \\
\text{Et donc} \quad p(\sigma_{\text{jit}}^2 \,|\, \boldsymbol{y}, \boldsymbol{x}, -) &\propto \prod_{n=1}^{N} (\sigma_{\text{jit}}^2 + \sigma_n^2)^{-1/2} \exp\left(-\frac{1}{2}\sum_{n=1}^{N} \frac{\zeta(n)^2}{\sigma_{\text{jit}}^2 + \sigma_n^2}\right)\; p(\sigma_{\text{jit}}^2)
\end{aligned}
$$

Il est donc, à notre connaissance, impossible de trouver une loi *a priori* pour $\sigma_{\text{jit}}^2$ conjuguée à cette loi *a posteriori*, à cause de la présence des erreurs de mesures $\sigma_n^2$. Une possibilité est d'ignorer la connaissance que l'on a des erreurs de mesures et de les considérées *i.i.d.*, ce qui revient à retourner au cas du bruit gaussien centré *i.i.d.* de variance $\sigma_\epsilon^2$, dont l'échantillonnage a déjà été décrit au chapitre précédent (voir Tableau 3.3). Néanmoins, si l'on souhaite échantillonner la variance du jitter tout en gardant la variance des erreurs de mesure, la solution est d'utiliser une étape de Métropolis-Hastings comme pour les paramètres non linéaires.

En présence de bruit corrélé (troisième cas), la dépendance de la matrice de covariance $\boldsymbol{\Sigma}_{\text{act}}$ en les paramètres du processus gaussien $\boldsymbol{\theta}_{\text{act}}$ est fortement non linéaire (voir des exemples en § 4.1.3) et il est donc impossible d'échantillonner directement les paramètres $\boldsymbol{\theta}_{\text{act}}$ suivant leur loi conditionnelle : une étape de Métropolis-Hastings est donc aussi nécessaire.

Ainsi, dans les deuxième et troisième cas, il faut remplacer l'échantillonnage direct de la variance du bruit dans l'algorithme (Tab. 3.10) par l'échantillonnage des hyperparamètres du bruit via une étape de Métropolis-Hastings - dans le troisième cas, on peut d'abord échantillonner $\sigma_{\text{jit}}^2$ puis $\boldsymbol{\theta}_{\text{act}}$, ou les deux simultanément. Pour simplifier, nous proposons de prendre pour chaque paramètre une loi *a priori* uniforme sur son ensemble de définition[14] et une loi de proposition de type loi de mélange entre marche aléatoire et loi *a priori*, comme celle utilisée pour

[14] Pour la variance du jitter, on prendra $\sigma_{\text{jit}}^2 \sim \mathcal{U}([0; V_{\max}[)$, où le choix de la variance maximale $V_{\max}$ se fait en fonction des erreurs de mesures $\boldsymbol{\Sigma}_{\text{mes}} = (\sigma_n^2)_n$, typiquement $V_{\max} = 2\max_n(\sigma_n^2)$. Pour chaque $(\theta_{\text{act,i}})_{i=1..4}$, on choisira $\theta_{\text{act,i}} \sim \mathcal{U}([v_{i,\min}; v_{i,\max}[)$, où les bornes des intervalles seront choisies suivant les connaissances *a priori* des astrophysiciens sur ces paramètres, *i.e.* sur l'activité stellaire de l'étoile.

l'échantillonnage des paramètres non linéaires (voir (3.16)). Notons qu'il faudra alors choisir la probabilité de marche aléatoire $p_{\text{rw}}^{\text{act}}$ ainsi que les variances instrumentales des marches aléatoires pour chaque paramètre. Ces paramètres seront alors échantillonnés directement suivant la loi *a posteriori* marginale ; la loi de proposition étant symétrique, la probabilité d'acceptation prend donc la même forme que celle pour les paramètres non linéaires :

$$\alpha \;=\; \left(\frac{|\mathbf{B}'|}{|\mathbf{B}|}\right)^{-1/2} \exp(-\frac{1}{2}\boldsymbol{y}^T(\mathbf{B}'^{-1} - \mathbf{B}^{-1})\boldsymbol{y}) = \exp(-\Psi(\mathbf{B}'^{-1}, \mathbf{B}^{-1}; \boldsymbol{y})/2)$$

où la notation $\mathbf{B}'$ désigne la matrice $\mathbf{B}$ dépendant des propositions ($\sigma_{\text{jit}}^{2\,\prime}$, $\boldsymbol{\theta}_{\text{act}}{}'$, ou les deux). Cependant, contrairement à la probabilité d'acceptation pour l'échantillonnage des paramètres non linéaires, nous ne pouvons pas utiliser les astuces calculatoires proposées en § 3.3.3, car la différence entre les matrices $\mathbf{B}'$ et $\mathbf{B}$ est ici de rang plein et ne permet pas les réductions de dimension ou l'utilisation du facteur de Cholesky $\mathbf{F}$. Il est donc nécessaire de calculer les matrices $\mathbf{B}$ et $\mathbf{B}'$. Néanmoins, cette étape étant effectuée une seule fois par itération, l'augmentation du coût de calcul qui en découle reste faible.

**Prise en compte de l'activité stellaire dans le dictionnaire réduit.**

Nous avons aussi vu en § 4.1.3 que l'activité stellaire pouvait être prise en compte à l'aide d'indicateurs, réunis dans la matrice $\mathbf{H}_{\text{indic,trend}}$ avec les tendances linéaires et quadratiques. Bien que nous ne l'utiliserons pas ici, notons qu'une telle prise en compte des indicateurs dans notre modèle ne posent aucune difficulté en pratique. En effet, il suffit de rajouter la matrice $\mathbf{H}_{\text{indic,trend}}$ au dictionnaire réduit $\overline{\mathbf{H}}$ (comme nous l'avions fait avec la colonne $\mathbf{1}$ pour la composante continue $x_0$) et les amplitudes $\boldsymbol{x}_{\text{indic,trend}}$ à $\overline{\boldsymbol{x}}$. On peut alors donner aux amplitudes $\boldsymbol{x}_{\text{indic,trend}}$ une loi *a priori* gaussienne centrée de variances $\sigma^2_{\text{indic,trend}}$, possiblement différentes entre chaque indicateur et de $\sigma_x^2$ (et que nous supposerons connues dans ce cas). L'adaptation des calculs pour l'échantillonnage des lois conditionnelles marginales en § 3.3 ainsi que ceux pour l'échantillonnage des hyperparamètres (notamment $\sigma_x^2$) en § 3.4 (et paragraphe précédent) ne posent pas de difficulté.

### 4.2.4 Détection des exoplanètes via l'estimation des paramètres

Enfin, la détection d'exoplanètes peut s'effectuer via l'estimation des paramètres à partir des échantillons tirés suivant la loi *a posteriori* et peut suivre la procédure décrite au Chapitre 3. Ainsi, à partir des échantillons de la séquence de Bernoulli $\boldsymbol{q}^{(t)}$, on calcule la moyenne empirique $\widehat{\boldsymbol{q}}$ à partir de laquelle, après possible fusion des intervalles anti-corrélés (voir Tableau 3.12), la détection des exoplanètes s'effectue en calculant la séquence binaire $\widehat{\boldsymbol{q}}_\alpha$ par seuillage de $\widehat{\boldsymbol{q}}$ : une exoplanète sera considérée comme détectée dans l'intervalle de période $\mathcal{I}_j$ si $\widehat{q}_{j,\alpha} = 1$, avec la probabilité $\widehat{q}_j$.

Les paramètres non linéaires $(P, e, \phi)$ seront estimés via la moyenne empirique des échantillons $\boldsymbol{\nu}^{(t)}$ vérifiant $\boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha$, tout en fournissant des barres d'erreur via la variance de ces échantillons. Après estimation analytique des amplitudes $\widehat{\boldsymbol{x}}$ via l'expression de la moyenne/du maximum de leur loi *a posteriori* conditionnelle en (3.12), on peut remonter aux paramètres orbitaux originaux des exoplanètes $(P, e, T^{\text{P}}, \omega^0, K)$ via les changements de variables effectués en § 4.2.1.

Notons toutefois que par la suite, nous mettrons en évidence des cas où la stratégie d'estimation ne sera pas suffisante ; une étude plus approfondie des échantillons est souvent nécessaire pour avoir une meilleure compréhension du résultat.

Pour appréhender la convergence de nos échantillonneurs, nous serons souvent amenés à lancer nos échantillonneurs sur plusieurs initialisations différentes de la graine aléatoire, afin d'obtenir plusieurs chaînes de Markov. Bien que l'on puisse calculer le MPSRF (présenté en § 3.1.3) de certains paramètres pour évaluer la convergence de l'échantillonnage, nous préférons introduire ici la notion d'*histogramme cumulé* pour les paramètres entiers ($\widehat{\boldsymbol{q}}$ et $\widehat{L}$). Pour la séquence de Bernoulli, il s'agira de cumuler sur l'ensemble des chaînes les estimations au sens de l'EAP $\widehat{q}_j$ pour chaque indice $j$. Pour l'ordre du modèle, on cumulera sur l'ensemble des chaînes les histogrammes des échantillons $L^{(t)} = \|\boldsymbol{q}\|^{(t)}$. Ces illustrations permettent d'observer le comportement global de ces paramètres sur l'ensemble des chaînes.

Rappelons que la communauté astrophysique cherche à estimer le nombre d'exoplanètes à partir de quantités telles que la vraisemblance marginale, le facteur de Bayes (BF) et le *Posterior odds ratio* (POR) (voir § 4.1.4). Le calcul de ces quantités est un véritable défi vu les modèles utilisés [NFB+18]. Pour notre modèle Bernoulli Gaussien étendu, ces quantités peuvent être estimées très simplement à partir des échantillons de $L^{(t)} = \|\boldsymbol{q}^{(t)}\|$. En effet, la probabilité *a posteriori* $\Pr(L \,|\, \boldsymbol{y})$ peut être estimée comme le pourcentage d'échantillons $L^{(t)}$ égaux à $L$ :

$$\Pr(L \,|\, \boldsymbol{y}) \approx \frac{1}{T} \operatorname{Card} \Big\{ t \in [\![1; T]\!],\ L^{(t)} = L \Big\}$$

Le POR se déduit simplement comme le rapport de $\Pr(L+1 \,|\, \boldsymbol{y})$ sur $\Pr(L \,|\, \boldsymbol{y})$. Ensuite, pour calculer le BF, il est nécessaire de calculer la probabilité *a priori* du modèle $\Pr(L)$. Or $L$ suit une loi binomiale car il s'agit de la somme de variables de Bernoulli ; on a alors :

$$\Pr(L) = \Pr(\|\boldsymbol{q}\| = L) = \Pr(\sum_{j=1}^{J} q_j = L) = \binom{J}{L} \lambda^L (1-\lambda)^{J-L}$$

Étant dans un cadre non supervisé, nous utiliserons $\widehat{\lambda}$ (estimateur de l'EAP par moyenne des $\lambda^{(t)}$) pour calculer les probabilités de modèles $\Pr(L)$ et $\Pr(L+1)$. Le facteur de Bayes $\mathcal{Z}_{L+1}/\mathcal{Z}_L$ se calculera donc à partir du POR et des probabilités de modèle à partir de l'équation (4.10). On notera le BF et le POR entre les modèles $L+1$ et $L$ : «$L+1$ vs. $L$».

### 4.2.5 Bilan

Nous avons donc pu constater que le modèle BGE et l'algorithme d'échantillonnage proposé s'adapte sans grande difficulté au problème de détection d'exoplanètes à partir des mesures de vitesses radiales. Il s'agit d'un problème plus difficile à résoudre que les problèmes d'analyse spectrale en échantillonnage irrégulier déjà étudiés en § 3.5, notamment à cause des *paramètres non linéaires de formes* que sont l'excentricité $e$ et la phase $\phi$, mais aussi à cause de la présence de bruit corrélé dans le signal $\boldsymbol{y}$, dû à l'activité stellaire.

Dans les sections suivantes, nous allons montrer l'efficacité du modèle BGE sur des signaux de vitesses radiales simulés, d'abord sans puis avec activité stellaire. Nous testerons enfin notre méthode sur des données réelles.

## 4.3 Tests sur des données simulées sans activité stellaire

### 4.3.1 Les données

Nous présentons ici un système stellaire étudié par les astrophysiciens, que nous utilisons pour nos simulations. Depuis l'article où est étudié le système planétaire, nous récupérons les données ayant servi à l'estimation des paramètres orbitaux : les instants d'échantillonnage $\boldsymbol{t}$ ainsi que les erreurs de mesures $(\sigma_n^2)_n$ de la vitesse radiale. Nous récupérons aussi les paramètres orbitaux estimés par les astrophysiciens, notés $\boldsymbol{\Theta}^{\text{th}}$, ce qui nous permet de générer un signal de vitesses radiales simulées $\boldsymbol{y}$ :

$$\boldsymbol{y} = \boldsymbol{y}^{\text{th}} + \boldsymbol{\epsilon}$$

où $\boldsymbol{y}^{\text{th}}$ est le modèle képlérien (4.11) calculé avec les paramètres orbitaux $\boldsymbol{\Theta}^{\text{th}}$, et $\boldsymbol{\epsilon}$ est un bruit gaussien centré indépendant et non identiquement distribué de variances $(\sigma_n^2)_n$ connues (matrice de covariance diagonale).

#### Le système HD40307

L'étoile HD40307 (qui sera appelée HD40307 par la suite) a été plusieurs fois étudiée par les astrophysiciens. Lors de son étude à partir de données provenant du téléscope HARPS, [TAEG+13] a ainsi rapporté la possible existence de six exoplanètes en rotation autour de cette étoile. C'est le nombre « élevé » d'exoplanètes qui nous a motivé à tester nos algorithmes sur ce système planétaire.

Pour simuler notre signal de vitesses radiales $\boldsymbol{y}$, nous utilisons les paramètres orbitaux donnés par [TAEG+13], dont les plus importants (périodes, excentricités et semi-amplitudes) répertoriés dans le Tableau 4.2. Le signal simulé $\boldsymbol{y}$ (en noir) ainsi que la fenêtre spectrale associée aux instants d'échantillonnage $\boldsymbol{t}$ sont représentés en Figure 4.8. La grille de recherche que nous utiliserons pour la méthode BGE est constituée de $J = 100$ points, régulièrement espacés en logarithme. Les périodes sont recherchées entre $P^{\min} = 1$ jour et $P^{\max} = 3000$ jours, ce qui correspond à l'intervalle de recherche utilisé par [TAEG+13]. La grille de recherche ainsi que la position des périodes $\boldsymbol{P}^{\text{th}}$ dans les intervalles sont représentés en Figure 4.9 : chaque période tombe bien dans un intervalle différent.

| **Périodes** $P$ (jours) | **Excentricités** $e \in [0; 1[$ | **Semi-amplitudes** $K$ (m/s) |
|---|---|---|
| 4.31 | 0.20 | 1.94 |
| 9.62 | 0.06 | 2.45 |
| 20.43 | 0.07 | 2.75 |
| 34.62 | 0.15 | 0.84 |
| 51.76 | 0.02 | 1.09 |
| 197.80 | 0.29 | 0.95 |

Tab. 4.2 – Paramètres orbitaux (périodes, excentricités et semi-amplitudes) des 6 exoplanètes reportées par [TAEG+13].

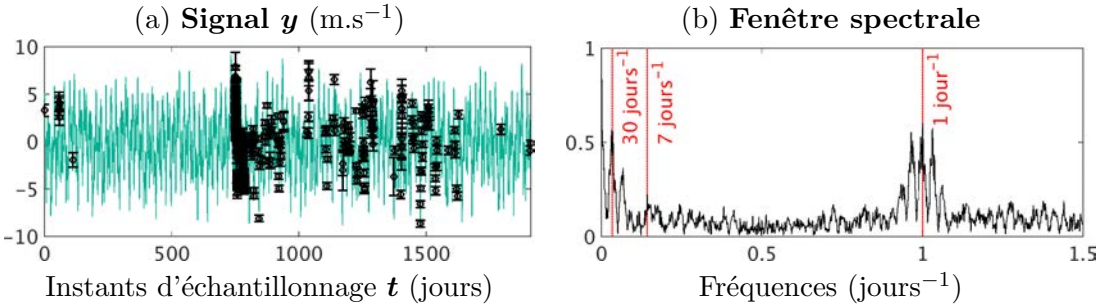



FIG. 4.8 – HD40307 (a) En vert, modèle képlérien ($\boldsymbol{y}^{\text{th}}$) simulé à partir des paramètres orbitaux estimés par les astrophysiciens. En noir, échantillons $y_n$ simulés avec du bruit aux instants $t_n$. (b) Fenêtre spectrale liées aux instants d'échantillonnage $\boldsymbol{t}$.

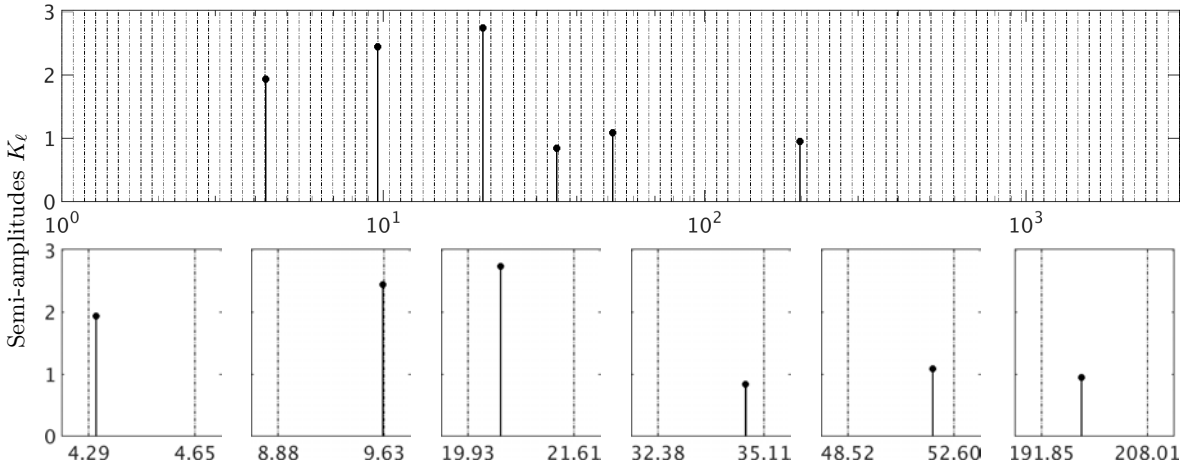



FIG. 4.9 – Représentation des périodes et des semi-amplitudes utilisées pour la simulation sur la grille de recherche de $J = 100$ points régulière en logarithme et zoom autour des périodes de simulation.

### 4.3.2 Un premier essai : algorithmes gloutons avec descente de gradient

Nous revenons dans un premier temps sur un type de méthodes que nous avons essayé de développer pour résoudre le problème de détection d'exoplanètes, avant (voir [Bou14]) et pendant ce projet de thèse : une extension des algorithmes gloutons avec descente de gradient pour l'estimation des paramètres non linéaires. Notons que ce type de méthode s'inscrit directement dans la continuité des premières méthodes utilisées par les astrophysiciens : la détection par périodogrammes avec sélection gloutonne.

Dans un premier temps, les algorithmes gloutons classiques peuvent être envisagés. Il faudrait alors construire un dictionnaire $\mathbf{H}$ en discrétisant les paramètres non linéaires du modèle (ici $(P, e, \phi)$) sur une grille arbitrairement fine pour construire des atomes $\mathbf{h}_j$. Les $L$ exoplanètes recherchées seraient alors détectées en sélectionnant itérativement l'atome $\mathbf{h}_{j_\ell}$ qui maximiserait (par exemple) la corrélation avec le résidu. Cependant, la construction d'une telle grille induit nécessairement une erreur de précision et un coût calculatoire important ; en discrétisant $(P, e, \phi)$ sur des grilles de (seulement) 1000 points chacun, cela revient (déjà) à construire un dictionnaire de plus d'un milliard atomes ! De plus, dans les approches gloutonnes classiques (de type MP, OMP, ...), les paramètres déjà estimés à l'itération précédente ne sont jamais remis en cause, puisque la sélection d'un atome revient à figer les paramètres sur la grille.

Cela nous a motivé à adopter la stratégie suivante, qui s'inscrit directement dans l'approche d'APNL présentée en § 1.3.2 et repose sur les principes des algorithmes gloutons présentés en § 1.2.2 :

- On conserve la même modélisation en APNL du problème vue en § 4.2.2, soit le découpage du domaine des localisations en $J$ intervalles $\mathcal{I}_j$ tels que $P_j \in \mathcal{I}_j$ et $(e_j, \phi_j) \in \mathcal{I}_{\boldsymbol{v}}$. On exploite donc le dictionnaire continu $\mathbf{H}(\boldsymbol{\nu})$ composé des $J$ atomes continus $\mathbf{h}(\boldsymbol{\nu}_j)$.
- À l'itération $\ell$, l'étape de sélection par un algorithme glouton revient alors à effectuer, pour chacun des intervalles $\mathcal{I}_j$, une optimisation locale en $\boldsymbol{\nu}_j = (P, e, \phi)_j$ du critère (par exemple, $\mathcal{J}(\boldsymbol{\nu}_j) = \|\boldsymbol{r}^{\ell-1} - \mathbf{h}(\boldsymbol{\nu}_j)\widehat{\boldsymbol{x}}_j^{\mathrm{MC}}\|^2$ dans le cas du MP) donnant les $\widehat{\boldsymbol{\nu}}_j$ optimaux. On sélectionne alors l'indice $j_\ell$ du $\widehat{\boldsymbol{\nu}}_j$ le plus optimal.

Ces étapes d'optimisation locale est une alternative réalisable à la grille 3D en les paramètres orbitaux. De plus, nous avions remarqué dans des travaux antérieurs (voir [Bou14]) que le critère était plus doux en $e$ et $\phi$ qu'il ne pouvait l'être en la période, laissant présager un nombre faible de minima locaux.

Par ailleurs, l'utilisation de ces étapes d'optimisation locale peut aussi permettre la remise en cause des paramètres orbitaux des exoplanètes estimés à l'étape précédente. En effet, on peut considérer deux types de sélection à l'itération $\ell$ (pour rappel, $\Omega^\ell$ est le support de la solution à l'itération $\ell$) :

$$
\begin{aligned}
\textit{Sélection} \quad &: \quad \Omega^\ell = \Omega^{\ell-1} \cup \{j_\ell\} \text{ avec } j_\ell = \arg\min_{j \notin \Omega^{\ell-1}} \mathcal{J}(j) \text{ avec } \mathcal{J}_1(j) \text{ ou } \mathcal{J}_2(j) \\
\textit{type MP} \quad \mathcal{J}_1(j) &= \min_{\boldsymbol{\nu}_j} \|\boldsymbol{r}^{\ell-1} - \mathbf{h}(\boldsymbol{\nu}_j)\widehat{\boldsymbol{x}}_j^{\mathrm{MC}}\|^2 \quad s.t.\ \boldsymbol{\nu}_j \in \mathfrak{I}_j \\
\textit{type OLS} \quad \mathcal{J}_2(j) &= \min_{\boldsymbol{\nu}_j, \overline{\boldsymbol{\nu}}} \|\boldsymbol{y} - [\overline{\mathbf{H}}_{\ell-1}(\overline{\boldsymbol{\nu}}),\ \mathbf{h}(\boldsymbol{\nu}_j)]\ \widehat{\boldsymbol{x}}^{\mathrm{MC}}\|^2 \ s.t.\ \boldsymbol{\nu}_j \in \mathfrak{I}_j \text{ et } \overline{\boldsymbol{\nu}}_i \in \mathfrak{I}_{j_i} \text{ avec } j_i \in \Omega^{\ell-1}
\end{aligned}
$$

où $\overline{\mathbf{H}}_{\ell-1}$ est le dictionnaire réduit à partir du support de l'étape précédente $\Omega^{\ell-1}$. La méthode de *type MP* ne remet pas en cause les paramètres orbitaux estimés aux itérations précédentes, au contraire de la méthode de *type OLS*[15].

Nous avons étudié de nombreuses variantes de ces types de méthodes ; par exemple, en choisissant de ne remettre en cause que les $(e, \phi)$ des étapes précédentes et en gardant fixées les périodes ; on peut aussi jouer sur l'étape de mise à jour des algorithmes gloutons. Ces variantes permettent en fait d'établir un compromis. En effet, plus on remet en cause les paramètres orbitaux déjà estimés et :

- moins la méthode sera sensible aux fausses détections caractéristiques des algorithmes glouton ;
- plus la méthode sera sensible aux minima locaux et lourde d'un point de vue calculatoire.

Nous avons remarqué que ce type d'algorithme glouton pouvait très bien fonctionner sur un exemple d'analyse spectral en échantillonnage irrégulier, comme nous l'avons rapporté dans [BC17]. Cependant, l'application à la détection d'exoplanètes semble bien plus problématique. Nous avons ainsi testé, sur plusieurs jeux de données, un certain nombre de variantes dont les suivantes :

- initialisation multiple pour les optimisations locales en $(P, e, \phi)$ ;
- utilisation d'algorithmes génétiques (via la toolbox de Matlab) pour éviter les minima locaux ;
- remplacement du modèle en $(P, e, \phi)$ par un modèle à harmoniques (réduction à un seul paramètre non linéaire, la période) ;

[15] Cette approche est de type OLS car la sélection s'effectue sur le signal $\boldsymbol{y}$ et non sur le résidu de l'itération précédente.

– sélection des périodes à l'aide d'un périodogramme de Lomb-Scargle (les périodes sont recherchées sur une grille très fine et $e$ et $\phi$ sont fixés à zéro) puis mise à jour du résidu sur modèle képlérien et optimisation locale à initialisations multiples et/ou algorithme génétique...

Malgré nos différentes tentatives, nous n'avons pas réussi à obtenir des résultats satisfaisants. Par exemple, pour les données relatives à l'étoile HD40307 présentées en § 4.3.1, nous présentons en Figure 4.10 un « histogramme » des périodes estimées par un algorithme de type OLS sur 100 réalisations différentes du bruit. On remarque que les 3 périodes d'amplitudes les plus élevées sont facilement retrouvées contrairement aux trois plus grandes. Sur l'ensemble des 100 réalisations du bruit, l'algorithme glouton de type OLS a un taux de « bonne détection » (l'erreur entre les périodes estimées et les périodes réelles ne dépassent pas la largeur d'intervalle $\Delta$ pour une grille de $J = 100$ intervalles) de seulement 24%. Par ailleurs, nous avons toujours supposé que le nombre d'exoplanètes du système était connu, évitant le problème du choix du critère d'arrêt.

Ainsi, la sélection itérative des exoplanètes par les algorithmes gloutons est un véritable problème, puisque une planète détectée ne sera jamais remise en cause. Il faudrait *a minima* une remise en cause des périodes déjà détectées avec un algorithme de type SBR [SIBD11], mais cela semble délicat à cause de la présence des paramètres de forme ; nous avons donc choisi de ne pas poursuivre cette piste.

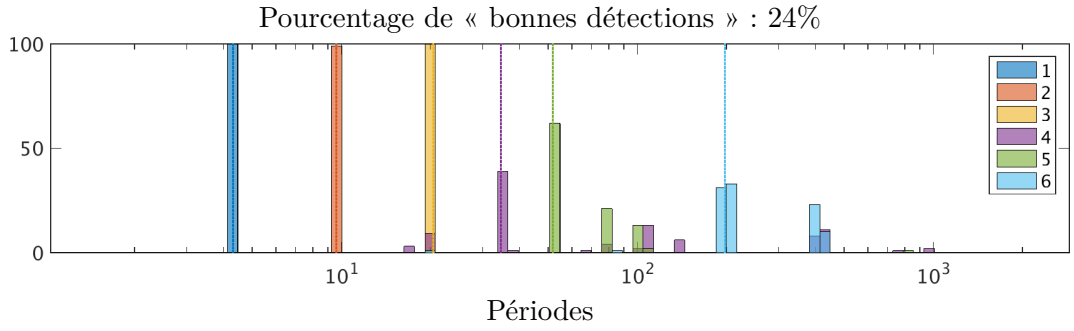



FIG. 4.10 – « Histogramme » de répartition des périodes estimées par un algorithme glouton avec descente de gradient de type OLS avec initialisations multiples, pour 100 réalisations du bruit sur le modèle HD40307. Chaque couleur correspond à une des 6 exoplanètes ; les traits en pointillés indiquent les périodes de simulation.

### 4.3.3 Exploitation du modèle BGE et de l'échantillonneur associé

Nos données étant simulées sans activité stellaire, nous ne modélisons ni n'échantillonnons de paramètres liés à l'activité. La loi cible est donc $p(\boldsymbol{q}, \boldsymbol{P}, \boldsymbol{e}, \boldsymbol{\phi}, \boldsymbol{x}, \lambda, \sigma_x^2 \,|\, \boldsymbol{y}, \boldsymbol{\Sigma}_{\text{mes}})$, où $\boldsymbol{\Sigma}_{\text{mes}}$ est la matrice diagonale de covariance du bruit regroupant la variance connue $\sigma_n^2$ du bruit de mesure pour chaque échantillon. Elle est échantillonnée sur 50000 itérations avec notre échantillonneur de type PCGS exploitant la marginalisation des amplitudes. Les estimations sont effectuées sur les 5000 dernières itérations. Nous récapitulons dans le Tableau 4.3 les différents réglages de notre échantillonneur. Ces derniers reprennent en grande partie ceux utilisés pour l'analyse spectrale au chapitre précédent. Néanmoins, nous verrons que les résultats de détection sont suffisamment bons. *Notons que le seuil de détection $\alpha$ est fixé à 0.85.*

| Variable Période : $X = \log_{10}(P)$, grille en log de périodes avec $\Delta = 0.0351$ | | | |
|---|---|---|---|
| Échantillonnage de $(X, e, \phi)_j$ : étape MH simultanée, $p_{\text{rw}} = 0.5$ | | | |
| | $X_j$ | $e_j$ | $\phi_j$ |
| *A priori* | $\mathcal{U}(\mathcal{I}_j)$ | $\mathcal{U}([0;1[)$ | $\mathcal{U}([0;1[)$ |
| Proposition | $\sigma_{X,\text{rw}} = 0.1\Delta$ | $\sigma_{e,\text{rw}} = 0.1$ | $\sigma_{\phi,\text{rw}} = 0.1$ |
| Échantillonnage de $(\lambda, \sigma_x^2)$ | | | |
| | $\lambda$ | $\sigma_x^2$ | |
| *A priori* | $\mathcal{U}([0;1[)$ | Jeffreys : $\propto 1/\sigma_x^2$ | |

TAB. 4.3 – Réglages de l'échantillonneur pour application à HD40307 sur données simulées sans activités.

**Résultats typiques.** La Figure 4.11 illustre l'échantillonnage de la loi cible sur une chaîne de 50000 échantillons. À gauche sont représentés les échantillons de la séquence de Bernoulli. Après quelques centaines d'itérations avec un grand nombre de $q_j$ égaux un 1 (voir zoom en Figure 4.12(a)), la séquence de Bernoulli se stabilise sur les indices théoriques. À droite, on peut observer à titre d'exemple l'évolution de la période et de l'excentricité dans l'intervalle 66 correspondant à la localisation de la sixième exoplanète du modèle de simulation. Les deux paramètres orbitaux évoluent autour de la valeur de simulation indiquée en rouge.

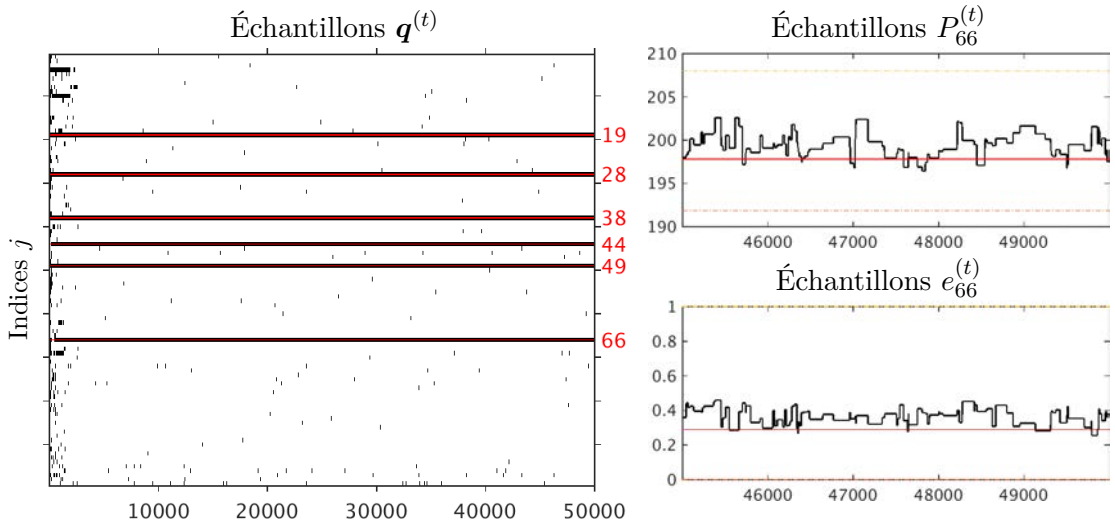



FIG. 4.11 – HD40307 - Échantillonnage des paramètres d'intérêt. Itérations en abscisses. À gauche : échantillons de la séquence de Bernoulli sur l'ensemble des 50000 itérations, indices de la grille en ordonnées, le noir correspond à $q_j = 1$. Les traits rouges horizontaux indiquent les indices des périodes de simulation. À droite : échantillons de la période (haut) et de l'excentricité (bas) sur les 5000 dernières itérations, pour l'indice 66 (la période de la 6$^{\text{e}}$ exoplanète se trouve dans $\mathcal{I}_{66}$).

La Figure 4.13(a) donne la moyenne $\widehat{\boldsymbol{q}}$ de la séquence de Bernoulli sur les 5000 dernières itérations. On constate que la valeur de $\widehat{\boldsymbol{q}}$ est proche de 1 pour chaque indice théorique, tandis que les autres valeurs sont proches de zéro ; le seuillage par $\alpha = 0.85$ donne une séquence binaire $\widehat{\boldsymbol{q}}_\alpha$ détectant correctement les 6 exoplanètes. La Figure 4.13(b) donne l'*histogramme cumulé* de 10 estimations $\widehat{\boldsymbol{q}}$ obtenues à partir de 10 initialisations différentes des graines aléatoires pour l'échantillonnage. Les périodes théoriques sont correctement retrouvées sur les 10 chaînes.

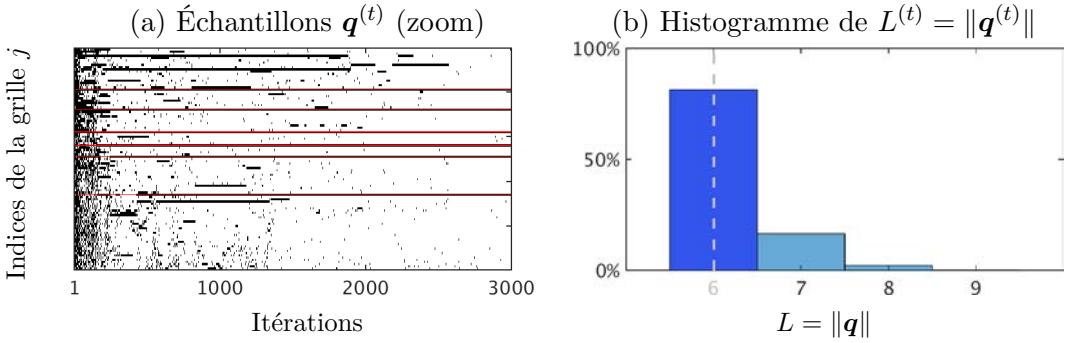



FIG. 4.12 – HD40307 - Évolution et estimation de l'ordre du modèle $L$. (a) Zoom sur les 3000 premières itérations $\boldsymbol{q}^{(t)}$ montrant l'évolution $L^{(t)}$ d'une valeur élevée au début qui décroît rapidement. (b) Histogramme de $L^{(t)}$ sur les 5000 dernières itérations. On a toujours en bleu foncé les échantillons tels que $\boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha$.

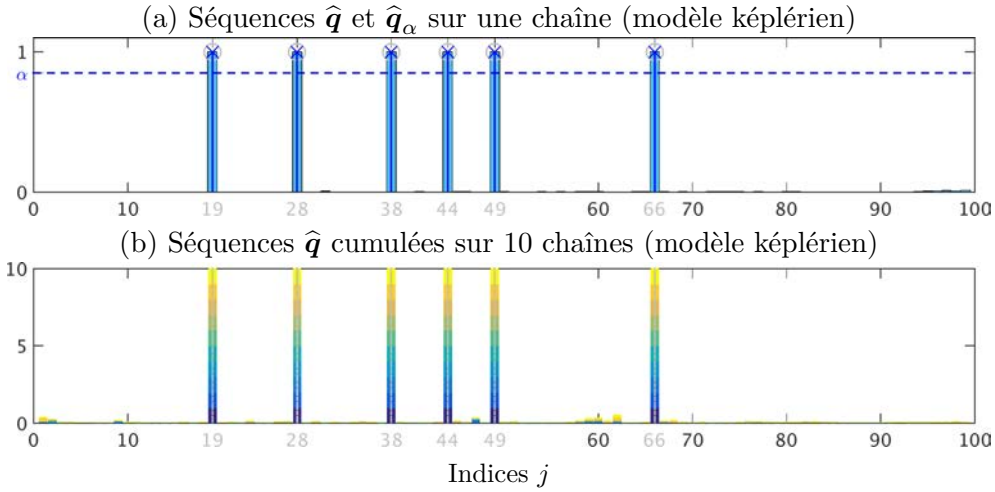



FIG. 4.13 – HD40307 - Histogrammes des séquences de Bernoulli $\boldsymbol{q}^{(t)}$. (a) Séquence moyenne $\widehat{\boldsymbol{q}}$ et seuillage $\widehat{\boldsymbol{q}}_\alpha$ avec $\alpha = 0,85$. Valeurs théoriques en grises et valeurs estimées en bleu. (b) Séquences moyennes $\widehat{\boldsymbol{q}}$ cumulées sur les 10 chaînes lancées en parallèle (une couleur par chaîne).

Nous pouvons observer en Figure 4.14 les résultats d'estimation des paramètres orbitaux $(P, e)$, conditionnellement à la séquence binaire détectée $\widehat{\boldsymbol{q}}_\alpha$. Les taux d'acceptation des étapes de Métropolis-Hastings sont très faibles (entre 0.02 et 5%) et on remarque que les échantillons sont très localisés, notamment pour ceux relatifs aux petites périodes (cela se remarque aussi en Figure 4.11). On pourrait certainement améliorer les performances de notre échantillonnage. D'une part, nous avons choisi en Tableau 4.3 un écart-type instrumental $\sigma_{X,\mathrm{rw}}$ pour l'échantillonnage des périodes identiques quel que soit l'intervalle. Il serait certainement plus correct qu'il soit adapté à l'intervalle, notamment en le choisissant plus faible pour les faibles périodes/hautes fréquences (cela reviendrait à choisir des écart-types identiques « en fréquence » pour suivre les variations de la fenêtre spectrale). D'autre part, on pourrait adapter l'ensemble des écart-types instrumentaux pour chaque paramètre non linéaire ($P$, $e$, $\phi$) afin d'augmenter le taux d'acceptation. Une valeur environnant les 20% est souvent recommandée pour assurer une convergence plus rapide (voir par exemple [Rob96]). Néanmoins, malgré notre réglage non optimisé des échantillonneurs, les erreurs d'estimation des paramètres orbitaux sont très faibles ; notamment, les échantillons des périodes sont localisées (très fortement) autour des périodes théoriques.

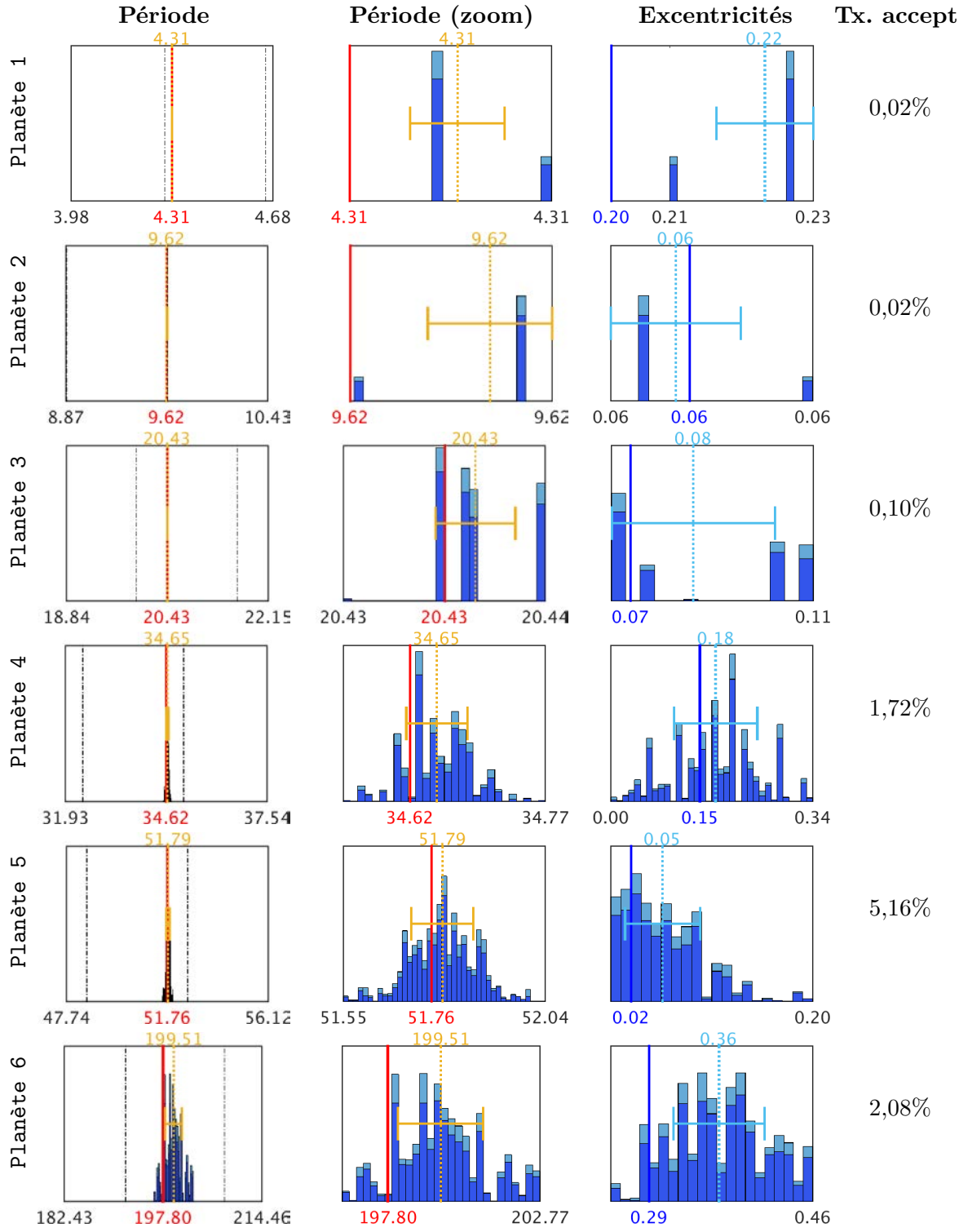



FIG. 4.14 – HD40307 - Histogrammes des échantillons des périodes et des excentricités. Chaque ligne correspond à une exoplanète. Pour chaque figure, la valeur indiquée colorée indiquée en bas correspond à la valeur théorique, tandis que la valeur indiquée en haut est la valeur estimée. La barre d'erreur tracée est de largeur $2\widehat{\sigma}$ où $\widehat{\sigma}^2$ est la variance estimée à partir des échantillons. Le taux d'acceptation est indiqué dans la dernière colonne. À gauche : histogrammes des $P_j^{(t)}$ dans les intervalles $\mathcal{I}_j$ des périodes théoriques, centrés sur la période théorique et de largeur $2\Delta$. Au centre : identique à celle de gauche mais avec zoom. À droite : histogrammes des $e_j^{(t)}$.

En Figure 4.15. nous observons les échantillons des hyperparamètres $(\lambda, \sigma_x^2)$ avec leur histogrammes associés[16]. On ne distingue par de corrélation notable entre les deux paramètres. Les valeurs estimées sont très proches des valeurs attendues (ces dernières ont été calculées en suivant les mêmes réflexions que celles menées en § 3.5.1). Notons que pour l'ensemble des paramètres estimés nous avons fait apparaître la barre d'erreur à $\pm\widehat{\sigma}$, où $\widehat{\sigma}^2$ est la variance empirique des échantillons ayant servis à l'estimation.

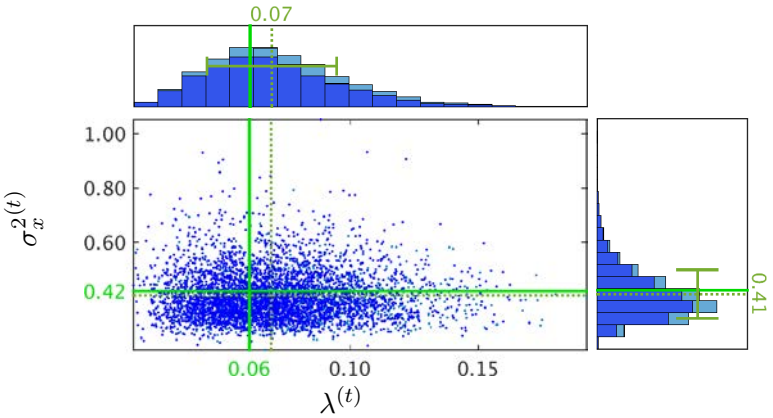



FIG. 4.15 – HD40307 - Échantillons et histogrammes des hyperparamètres $\lambda$ et $\sigma_x^2$. En bleu clair : ensemble des 5000 derniers échantillons. En bleu foncé : échantillons tels que $\boldsymbol{q}^{(t)} = \widehat{\boldsymbol{q}}_\alpha$. En vert trait plein : valeurs « attendues ». En vert pointillés : valeurs estimées.

Enfin, nous avons tracé en Figure 4.12(b) l'histogramme des échantillons $L^{(t)} = \|\boldsymbol{q}^{(t)}\|$. L'ordre du modèle, qui représente un véritable challenge pour les astrophysiciens, est ici correctement retrouvé à partir de cet histogramme, puisque la probabilité d'avoir $L = 6$ dépasse les 50%. Nous verrons durant la section suivante que ce ne sera pas toujours le cas. Notons que le facteur de Bayes (BF) « 6 vs. 5 » est infini car aucun échantillon ne donne $L^{(t)} < 6$. Le BF « 7 vs. 6 » est estimé à 0.1976.

**Modèle circulaire et présence d'harmoniques.** Nous avons observé le comportement de notre échantillonneur si un modèle de trajectoire circulaire est choisi à la place des trajectoires elliptiques ; cela revient à échantillonner la loi $p(\boldsymbol{q}, \boldsymbol{P}, \boldsymbol{x}, \lambda, \sigma_x^2 \,|\, \boldsymbol{y}, \boldsymbol{e} = \boldsymbol{0}, \boldsymbol{\phi} = \boldsymbol{0})$ et rechercher uniquement des sinusoïdes, comme en analyse spectrale. La Figure 4.16 représente les séquences moyennes $\widehat{\boldsymbol{q}}$ pour une chaîne et cumulés sur 10 chaînes. En plus des indices correspondant aux périodes de simulations, on retrouve des variables de Bernoulli élevées (supérieur à 50%) pour les indices (10,23,57), correspondant respectivement à des périodes (en moyennant sur les 5000 derniers échantillons) de 2.1559, 6.2706 et 113.2594 jours. Ces fausses détections, que l'on retrouve bien sur les 10 chaînes, correspondent en fait aux harmoniques présentes dans le signal, dues au fait que les képlériennes simulées aux périodes $\boldsymbol{P}^{\text{th}}$ ne sont pas sinusoïdales. La « fausse » exoplanète à $P_{10} \approx 2.1559$ jours correspond par exemple à la seconde harmonique de la première exoplanète : $P_1^{\text{th}} = 4,31$ jours $\approx 2P_{10}$. De telles associations ne sont pas évidentes pour les autres fausses détections aux indices (23,57), qui représentent sûrement des mélanges des diverses harmoniques dues aux 6 képlériennes. Cela confirme la nécessité d'utiliser le modèle képlérien plutôt que le modèle circulaire.

[16] En pratique, nous avons normalisé le signal $\boldsymbol{y}$ si bien que les valeurs des échantillons relatifs aux amplitudes - $\sigma_x^2$ mais aussi $\boldsymbol{x}$ par exemple - ne correspondent pas aux valeurs physiques.

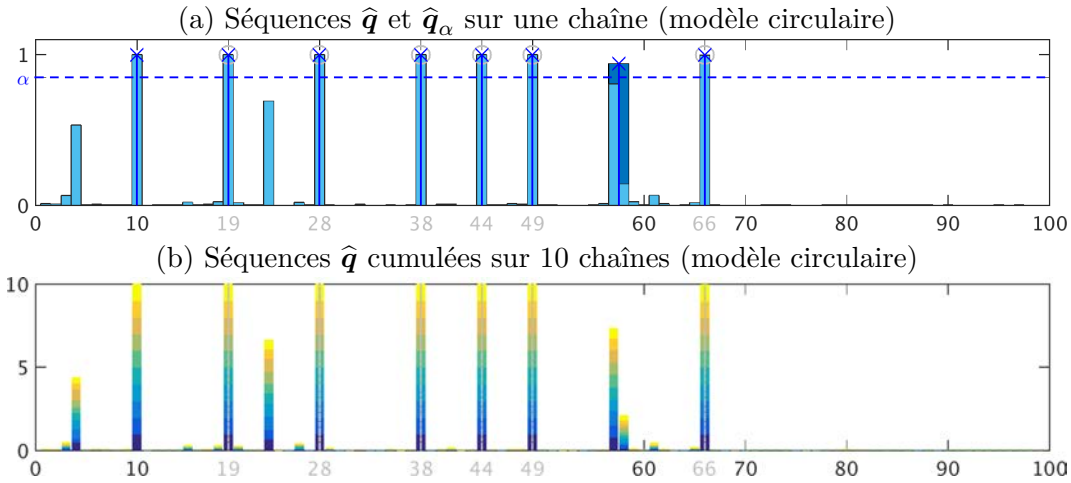



FIG. 4.16 – HD40307 - Histogrammes des séquences de Bernoulli $\boldsymbol{q}^{(t)}$ pour un modèle circulaire ($e_j = 0$). En haut : séquence moyenne $\widehat{\boldsymbol{q}}$ et seuillage $\widehat{\boldsymbol{q}}_\alpha$ avec $\alpha = 0,85$. Valeurs théoriques en grises et valeurs estimées en bleu. En bas : séquences moyennes $\widehat{\boldsymbol{q}}$ cumulées sur les 10 chaînes lancées en parallèle (une couleur par chaîne).

**Résultats statistiques.** Afin de mieux évaluer l'efficacité de notre méthode pour la détection d'exoplanètes, nous avons testé notre échantillonneur (avec 10 graines différentes pour l'initialisation aléatoire) de façon identique aux paragraphes précédents, mais sur 100 différentes réalisations du bruit. Ce test statistique est équivalent à celui effectué avec les algorithmes gloutons en § 4.3.2. Pour pouvoir comparer, nous conservons ici le terme de « bonne détection », défini précédemment, à savoir : sur les 5000 dernières itérations, le $\widehat{\boldsymbol{q}}_\alpha$ avec $\alpha = 0.85$ correspond à la séquence de Bernoulli théorique, et l'écart entre les périodes[17] estimées et les théoriques ne dépassent pas $\Delta$, la largeur des intervalles.

La Figure 4.17 représente la carte de localisation (graines vs. réalisation du bruit) des bonnes et des mauvaises détections. Sur les 10 graines d'initialisations, au maximum 2 donnent des mauvaises détections sur l'ensemble des 100 réalisations. Plus précisément :

- sur les 1000 tests, on obtient 98,9% de bonnes détections.
- sur les 100 réalisations du bruit, 91 réalisent une bonne détection sur les 10 graines.
- les histogrammes cumulés normalisés par 10 (le nombre de graines) donnent un $\widehat{\boldsymbol{q}}_\alpha$ correspondant à la séquence de Bernoulli théorique pour les 100 réalisations du bruit.

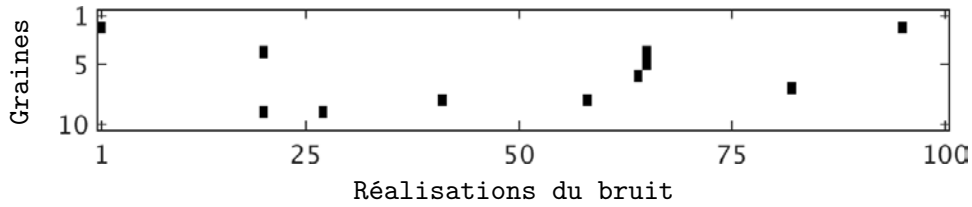



FIG. 4.17 – Carte de localisation des bonnes (en blanc) et des mauvaises détections (en noir), sur l'ensemble des 1000 chaînes calculées (100 réalisations du bruit × 10 initilisations aléatoires).

Les mauvaises détections arrivent dans deux cas. Soit un $q_j$ n'appartenant pas à la séquence théorique s'allume uniquement à la fin de l'échantillonnage (le fait de prendre uniquement les 5000 dernières itérations, pour des raisons pratiques, n'est certainement pas idéal ici), soit un grand nombre de $q_j$ sont égaux à 1 ; la chaîne n'a pas eu le temps de converger. On peut d'ailleurs

[17] Pour rappel, nous parlons de période bien qu'en pratique nous travaillons en logarithme de période.

confirmer ce comportement en calculant le MPSRF (voir en § 3.1.3 p.105) du nombre $L = \sum_j q_j$, sur les dix initialisations des graines pour chaque réalisation du bruit. Sur les 100 réalisations du bruit,

- 89 atteignent le seuil de 1.1 au bout des 50000 itérations, le maximum se situant à 30.
- 94 atteignent le seuil si on retire les chaînes initialisées aux graines correspondant aux mauvaises détections ; le maximum se situe alors à 2,3.

Cela confirme la bonne convergence des chaînes de Markov.

Pour comparer avec les résultats obtenus avec les algorithmes gloutons en Figure 4.10, nous avons réalisé en Figure 4.18 le même type d'histogramme des périodes estimés pour les 100 réalisations du bruit, en prenant les résultats de la première graine d'initialisation. On remarque ici que les 6 exoplanètes sont correctement localisées sur l'ensemble des 100 réalisations du bruit, au contraire de notre algorithme glouton.

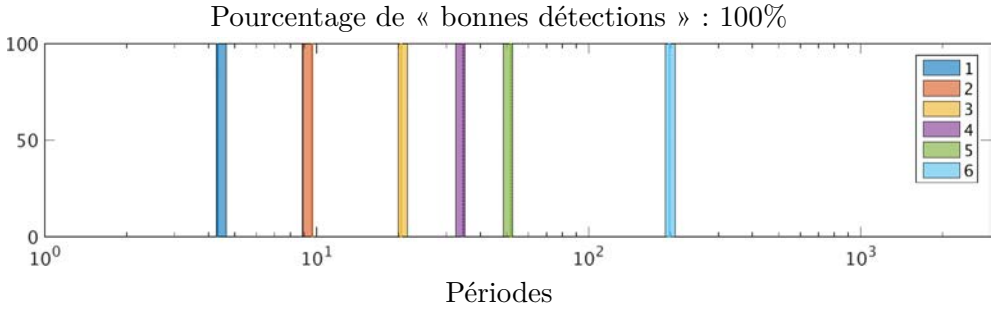



FIG. 4.18 – « Histogramme » de répartition des périodes estimées par le BGE, pour 100 réalisations du bruit sur le modèle HD40307. Voir Figure 4.10 pour plus d'explications.

Enfin, nous présentons en Figure 4.19 les résultats d'estimation des paramètres orbitaux, pour les périodes et les excentricités. Pour ce faire, nous avons pris les résultats d'estimation des périodes et des excentricités des 989/1000 chaînes qui donnent une « bonne détection ». Pour les estimations des périodes, nous obtenons le même comportement que pour les histogrammes des échantillons en Figure 4.14 : plus on va vers les petites périodes (hautes fréquences), plus les périodes estimées sont concentrées autour de leur moyenne. Les périodes estimées se retrouvent d'ailleurs dans des intervalle de largeur équivalente à 1 jour$^{-1}$, tout comme pour les histogrammes des estimations. En moyenne, les erreurs d'estimation en périodes sont très faibles, notamment si l'on compare les périodes théoriques avec la moyenne des périodes estimées.

L'estimation des excentricités est aussi correcte, quoique moins localisée. On remarque un comportement intéressant au niveau de la 5$^e$ exoplanète : son excentricité théorique étant très proche de 0 (0.02), les estimations donnent globalement des valeurs toutes supérieures (en moyenne 0.08, allant au delà de 0.2). Cela se rapporte à un problème déjà posé par les astrophysiciens, qui est l'estimation souvent sur-estimée des excentricités des exoplanètes (voir [Har17]). Pour notre méthode, cela vient sûrement du fait que l'excentricité est un paramètre positif et donc très peu d'échantillons inférieurs à 0.02 peuvent être tirés. Une reparamétrisation proposée par [HBLC16] pourrait permettre de résoudre ce problème.

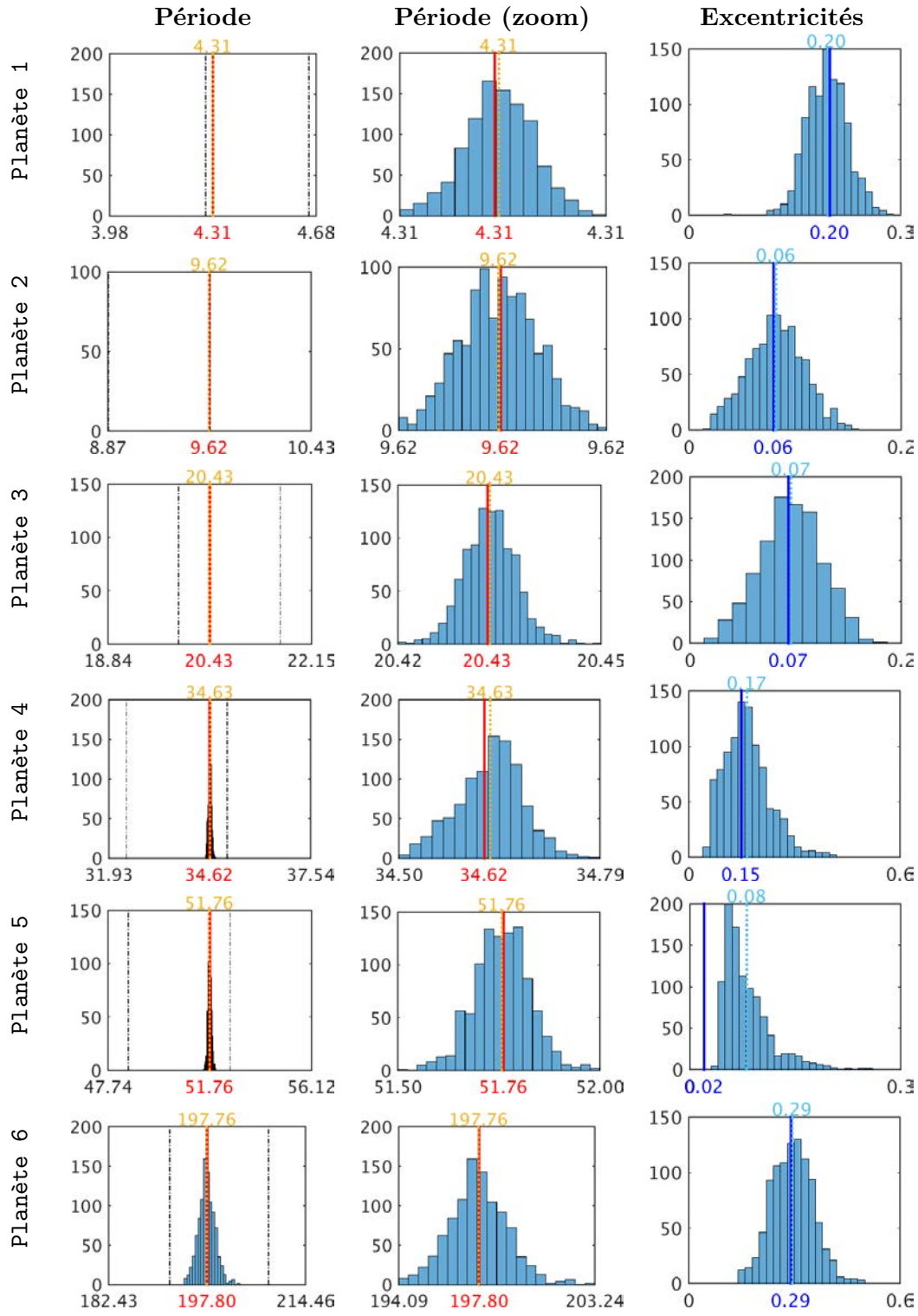



FIG. 4.19 – HD40307 - Histogrammes des estimations des périodes et des excentricités des 1000 chaînes. Chaque ligne correspond à une exoplanète. Pour chaque figure, la valeur indiquée colorée indiquée en bas correspond à la valeur théorique, tandis que la valeur indiquée en haut est la valeur estimée. À gauche : histogrammes des périodes estimées dans les intervalles $\mathcal{I}_j$ des périodes théoriques, centrés sur la période théorique et de largeur $2\Delta$. Au centre : identique à celle de gauche mais avec zoom. À droite : histogrammes des excentricités estimées.

### 4.3.4 Bilan

Nous avons pu voir dans cette section que le modèle BGE s'adapte très bien à l'estimation de paramètres orbitaux à partir de signaux de vitesses radiales simulés sans activité stellaire. L'échantillonnage des paramètres orbitaux est réalisé de façon très basique (*a priori* uniforme, loi de proposition simple) et pourrait être amélioré en prenant des *a priori* plus proches des connaissances sur les paramètres, mais aussi en cherchant à améliorer les lois de propositions, notamment au niveau de la période où celle-ci gagnerait à être adaptée à la largeur de l'intervalle en fréquence. Une reparamétrisation du modèle képlérien pourrait aussi aider à éviter une surestimation des excentricités. Les exoplanètes simulées ont cependant toutes été retrouvées avec succès. Les sections suivantes étudient le comportement du modèle BGE en présence d'activité simulée, et pour des signaux de vitesses radiales réels.

## 4.4 Tests sur des données avec activité stellaire

Nous présentons dans cette section des résultats sur des données simulées avec activité stellaire. La simulation d'activité stellaire est en soi un problème difficile ; dans le *RV challenge* de [Dum16], le logiciel SOAP 2.0 [DBS14] est utilisé afin de simuler une activité stellaire réaliste prenant en compte des effets de granulation, d'apparition et de disparition de tâches, ou encore d'activité magnétique. Les signaux de vitesses radiales simulés avec cet outil sont alors très réalistes ; peut-être trop pour le cadre de cette thèse, où il est alors nécessaire d'avoir une bonne compréhension de l'activité stellaire pour correctement régler, par exemple, les *Processus Gaussien* et indicateurs stellaires utilisés pour filtrer l'activité.

Nous avons donc décidé de montrer les résultats à partir des 6 signaux de vitesses radiales simulés donnés par [NFB+18] dans le cadre de son challenge sur l'estimation de la *Marginal Likelihood*. Les directives pour la prise en compte de l'activité stellaire sont précises et ne nécessitent donc pas une étude approfondie. Nous présentons d'abord les données récupérées, les réglages du modèle BGE et de l'échantillonneur, et enfin les résultats de nos estimations.

### 4.4.1 Les données

Les données simulées par [NFB+18] sont composées de 6 signaux de vitesses radiales, contenant chacun la contribution képlérienne de $L = 2$ exoplanètes, ainsi qu'une contribution d'activité stellaire. Les données $\boldsymbol{y}$ sont simulées en sommant la contribution des exoplanètes avec un bruit d'activité suivant une loi normale centrée de matrice de covariance $\boldsymbol{\Sigma}_{\text{act}}$, qui s'écrit :

$$\boldsymbol{\Sigma}_{\text{act}}(m,n) = \delta_{m,n}(\sigma_{\text{jit}}^2 + \sigma_n^2) + R(t_m, t_n; \boldsymbol{\theta}_{\text{act}}) \tag{4.13}$$

où $\sigma_n^2$ sont les erreurs de mesures associées aux instants $t_n$ donnés avec les mesures $y_n$, $\delta_{m,n} = 1$ uniquement si $m = n$ (et 0 sinon) et $R(t_m, t_n; \boldsymbol{\theta}_{\text{act}})$ correspond à la fonction de covariance *quasi-périodique* donnée en équation (4.8). Les hyperparamètres $(\sigma_{\text{jit}}^2, \boldsymbol{\theta}_{\text{act}})$ sont les mêmes pour les 6 systèmes et sont récapitulés en Tableau 4.4. Ce dernier donne aussi la moyenne $\bar{\sigma}$ des erreurs de mesures $(\sigma_n)_n$. Les principaux paramètres orbitaux, donnés par l'auteur, sont répertoriés dans le Tableau 4.4. Un degré de difficulté est donné aux diverses exoplanètes, suivant le rapport $K/\sigma$ comparant la contribution de l'exoplanète avec la contribution de l'activité. Ainsi, *facile* correspond à $K/\sigma > 1$, *difficile* à $K/\sigma \approx 1$ et « *impossible* » (traduction de *impractical*) à $K/\sigma < 1$.

| **Système** | **Détectabilité** | **Périodes** $P$ (jours) | **Excentricités** $e \in [0;1[$ | **Semi-amplitudes** $K$ ($\mathrm{m.s^{-1}}$) | **Erreurs moyennes** $\bar{\sigma}$ ($\mathrm{m.s^{-1}}$) |
|---|---|---|---|---|---|
| 1 | facile | 12.1 | 0.08 | 1.86 | 0.904 |
| | facile | 42.4 | 0.04 | 2.44 | |
| 2 | facile | 15.96 | 0.05 | 2.12 | 0.897 |
| | difficile | 120.5 | 0.31 | 1.36 | |
| 3 | difficile | 40.4 | 0.1 | 1.25 | 0.900 |
| | difficile | 91.9 | 0.1 | 1.19 | |
| 4 | facile | 169.1 | 0.22 | 1.58 | 0.897 |
| | « impossible » | 23.45 | 0.04 | 0.74 | |
| 5 | difficile | 31.1 | 0.04 | 0.75 | 0.903 |
| | « impossible » | 10.9 | 0.02 | 0.67 | |
| 6 | *identique à (3)* | | | | 0.907 |

| **Bruit** | **Jitter** | **Processus Gaussien** | | | |
|---|---|---|---|---|---|
| | $\sigma_{\mathrm{jit}}$ ($\mathrm{m.s^{-1}}$) | $\theta_{\mathrm{act},1}$ ($\mathrm{m.s^{-1}}$) | $\theta_{\mathrm{act},2}$ (jours) | $\theta_{\mathrm{act},3}$ (jours) | $\theta_{\mathrm{act},4}$ |
| | 0.6 | $\sqrt{3}$ | 50 | 20 | 0.5 |

Tab. 4.4 – Paramètres orbitaux des 6 signaux de vitesses radiales simulés par [NFB+18]. Les erreurs de mesures des systèmes 3 et 6 sont différentes, bien que du même ordre de grandeur ; le deux systèmes diffèrent donc par l'échantillonnage du signal et par une réalisation du bruit différente.

Les différents signaux ainsi que les fenêtres spectrales associées sont donnés en Figure 4.20. Ils sont tous constitués de $N = 200$ instants d'échantillonnage $t_n$, légèrement différents suivant le système considéré. Dans tous les cas, une pause entre les instants ∼225 jours et ∼375 jours est observée. Les fenêtres spectrales possèdent des lobes secondaires importants aux alentours de 1 jour (alternance jour/nuit) et de 1 an (alternance des saisons), ce qui est classique pour des mesures en astrophysique.

### 4.4.2 Réglages du modèle BGE et de l'échantillonneur associé

Nous donnons les réglages de notre modèle BGE et de l'échantillonneur associé (de type PCGS exploitant la marginalisation des amplitudes) pour la prise en compte de l'activité. Comme indiqué par [NFB+18], nous cherchons les périodes entre $P^{\min} = 1$ jour et $P^{\max} = 10000$ jours. Nous conservons une grille de $J = 100$ intervalles régulièrement espacés en logarithme donnant une largeur d'intervalle de $\Delta = 0.0404$ en logarithme (tout comme précédemment, nous parlerons de période alors que nous travaillerons en logarithme de période). Une telle grille permet de correctement séparer les périodes des deux exoplanètes pour tous les systèmes à étudier.

Nous conservons le même paramétrage[18] pour l'échantillonnage des paramètres orbitaux $(P, e, \phi)_j$ que dans le Tableau 4.3 sur les signaux simulés sans activités stellaires. Nous modifions cependant l'*a priori* sur la variance des amplitudes $\sigma_x^2$, en suivant les recommandations que nous avons données en § 3.4.3.

[18] Notons que [NFB+18] recommande des *a priori* différents des nôtres pour les paramètres orbitaux ; les différentes équipes devaient supposer les mêmes *a priori* afin de rendre plus aisée la comparaison entre les divers résultats. Notre objectif n'étant pas de comparer nos résultats avec les leurs, mais bien d'évaluer l'efficacité de notre échantillonneur pour un signal simulé avec activité stellaire, nous conservons les mêmes *a priori* que précédemment.

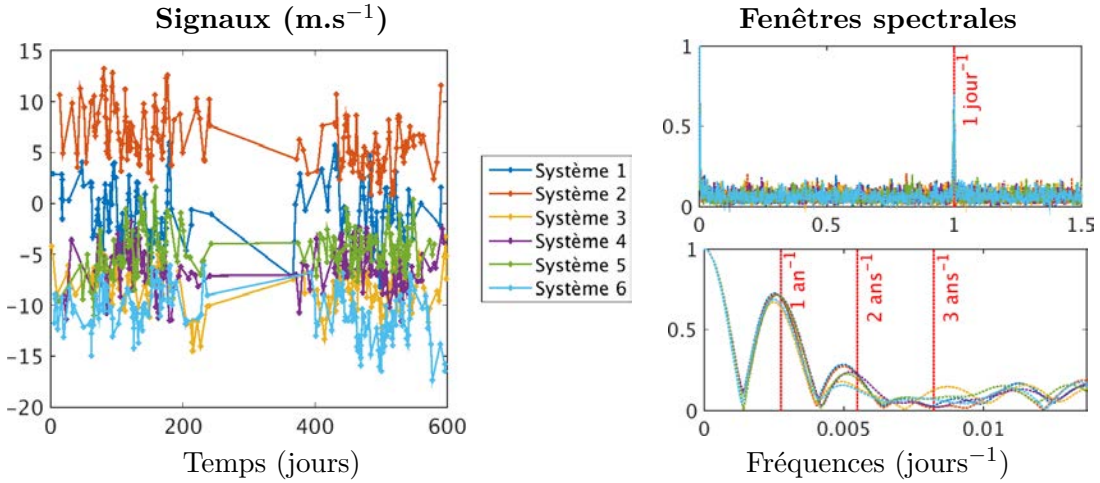



FIG. 4.20 – Les six différents systèmes simulés par [NFB+18]. À gauche : signaux de vitesses radiales $\boldsymbol{y}$. À droite : fenêtres spectrales associées aux instants d'échantillonnage $\boldsymbol{t}$.

Comme nous échantillonnons la matrice de covariance du bruit (voir paragraphe suivant), nous remplaçons l'*a priori* de Jeffreys par un *a priori* Inverse-Gamma :

$$\sigma_x^2 \sim \mathcal{IG}(\alpha_x, \beta_x)$$

où les hyper-hyperparamètres $(\alpha_x, \beta_x)$ sont réglés pour donner une valeur minimale *a priori* cohérente pour $\sigma_x^2$. Avec une loi *a priori* de Jeffreys, la variance des amplitudes $\sigma_x^2$ a en effet tendance à tendre vers 0, autorisant des planètes de très faibles amplitudes à être détectées ; cela mène généralement à un blocage de l'échantillonnage dans un mode où l'ordre du modèle $L$ est très élevé, ce qui est contraire à l'*a priori* de parcimonie que nous voulons imposer.

La plus grande modification par rapport aux tests réalisés en § 4.3 est la prise en compte de l'activité stellaire dans le modèle. Nous la modélisons par un processus gaussien de matrice de covariance $\boldsymbol{\Sigma}_{\text{act}}$ pour laquelle nous prenons le même modèle (4.13) que celui ayant servi à la simulation des signaux, dépendant des paramètres $\boldsymbol{\theta}_{\text{act}}$. Les erreurs de mesures $\sigma_n^2$ sont prises égales à celles données par l'article. Dans [NFB+18], les valeurs de $\boldsymbol{\theta}_{\text{act}}$ et $\sigma_{\text{jit}}^2$ sont données et il est demandé de les fixer à ces valeurs et de ne pas les échantillonner. Nous avons choisi ici de les échantillonner comme décrit en § 4.2.3 : leur échantillonnage s'effectue conjointement via une étape de Métropolis-Hastings. Nous prenons une loi *a priori* uniforme sur un petit intervalle autour de la valeur théorique pour $(\theta_{\text{act},2}, \theta_{\text{act},3}, \theta_{\text{act},3})$ et une loi *a priori* uniforme sur un intervalle plus grand pour les variances d'amplitudes $(\theta_{\text{act},1}, \sigma_{\text{jit}}^2)$. Les propositions s'effectuent via une loi de mélange entre marche aléatoire ($p_{\text{rw}} = 0.5$) et loi *a priori*. La loi cible de l'échantillonneur est donc $p(\boldsymbol{q}, \boldsymbol{P}, \boldsymbol{e}, \boldsymbol{\phi}, \boldsymbol{x}, \lambda, \sigma_x^2, \sigma_{\text{jit}}^2, \boldsymbol{\theta}_{\text{act}} \,|\, \boldsymbol{y})$.

### 4.4.3 Résultats

Nous avons lancé, pour chacun des 6 signaux, 10 chaînes avec différentes initialisations de la graine aléatoire sur 50000 itérations. Les estimations sont réalisées sur les 5000 dernières itérations.

**Résultats généraux à partir des histogrammes cumulés.** En Figure 4.21, nous pouvons observer les histogrammes cumulés des $\widehat{\boldsymbol{q}}$ et des $L^{(t)}$, calculés sur les 5000 dernières itérations, sur les 10 graines pour chaque système.

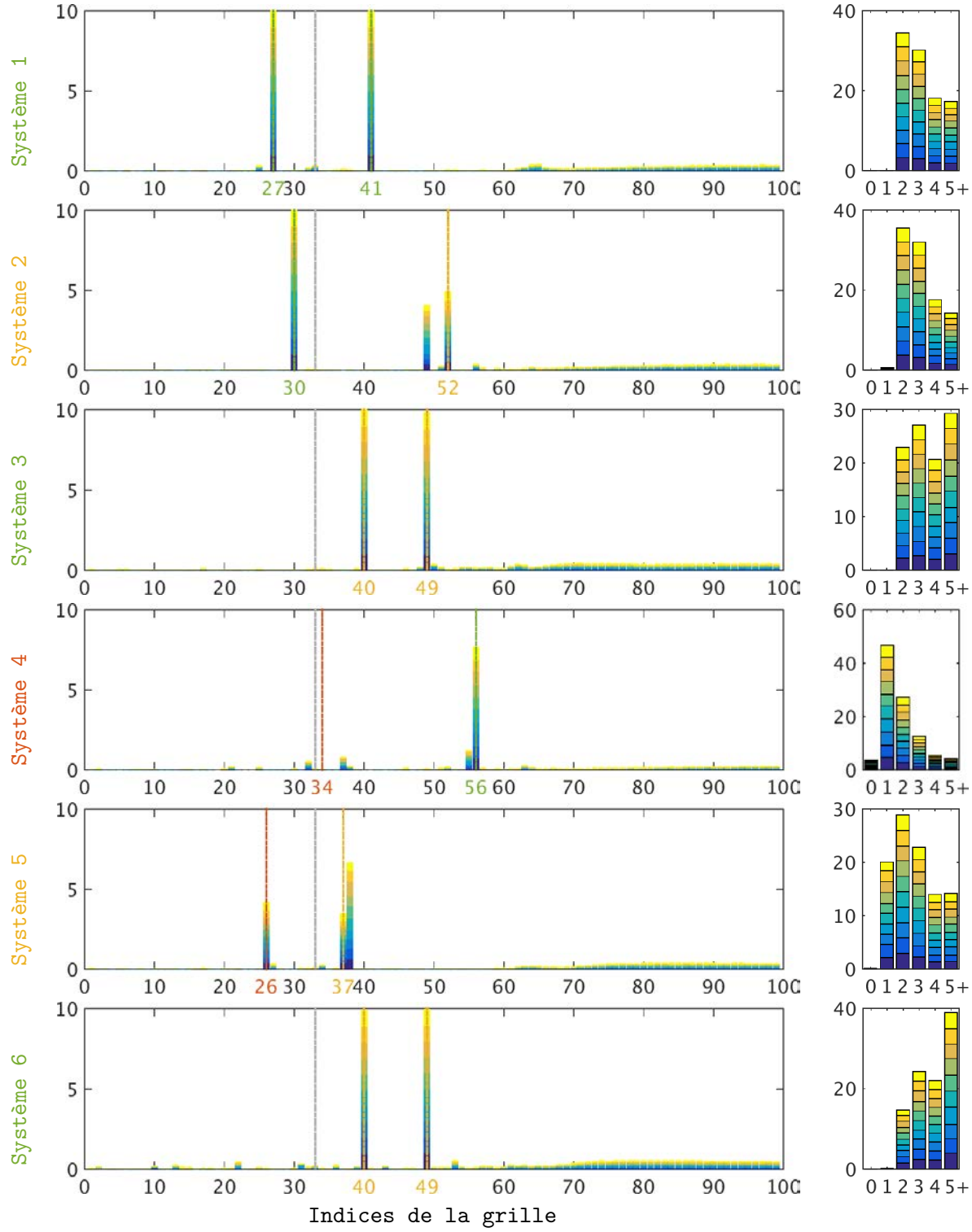



FIG. 4.21 – Histogrammes cumulés des $\widehat{\boldsymbol{q}}$ et des $\widehat{L}$ sur les 10 graines aléatoires pour chacun des 6 systèmes de [NFB+18]. À gauche : séquences $\widehat{\boldsymbol{q}}$ cumulées sur les 10 chaînes. En pointillés colorés sont indiquées les indices des périodes théoriques ; les couleurs sont relatives à la difficulté : **facile**, **difficile** et **« impossible »**. Le trait en pointillés gris représente $\theta_{\text{act},3}$, relatif à la période de rotation de l'étoile. À droite : ordres du modèle $\widehat{L}$ cumulés sur les 10 chaînes. L'ordonnée est indiquée en pourcentage, et nous avons regroupé dans « 5+ » l'ensemble des $\widehat{L}$ supérieur à 5.

Les estimations pour les systèmes 1, 3 et 6 s'en sortent très bien, puisque les 2 planètes sont à chaque fois détectées presque sûrement pour les 10 chaînes. Ce résultat est d'autant plus satisfaisant que les systèmes 3 et 6 sont supposés *difficiles* par [NFB+18]. Cependant, quand on regarde les histogrammes cumulés des $L^{(t)}$, on remarque dans les 3 cas que la probabilité d'avoir un système à $\widehat{L} = 3$ (et plus) exoplanètes est plus élevée que celle d'en avoir 2, alors que seulement 2 exoplanètes sont présentes dans les signaux. En effet, alors que les séquences de Bernoulli aux indices théoriques sont presque toujours égales à 1, permettant leur détection via l'histogramme $\widehat{\boldsymbol{q}}$, au moins une autre variable de Bernoulli $q_j$ est souvent égal à 1 ; mais cette dernière change régulièrement d'intervalle, si bien qu'on ne l'observe pas dans l'histogramme (ou alors avec une très faible probabilité, comme on peut le voir pour les derniers intervalles). Cela montre que l'espace des paramètres est bien balayé par l'échantillonneur. Ces observations confirment l'intérêt d'une observation de l'estimation $\widehat{\boldsymbol{q}}$.

Pour les trois autres systèmes, les résultats de détection sont moins concluants. Pour le système 2, l'exoplanète la moins difficile est correctement détectée ; pour la seconde, la probabilité ne semble pas suffisante (inférieure à 50% pour toutes les chaînes) et une troisième exoplanète semble ressortir à 3 intervalles de distance, toujours avec une faible probabilité. Pour le système 5, aucune exoplanète ne peut être détectée presque sûrement ; trois indices ressortent avec des probabilités comprises entre 30% et 70%. Ces deux systèmes méritent donc une étude plus approfondie des échantillons, que nous menons au paragraphe suivant.

Enfin, pour le système 4, une seule exoplanète est détectée ; la planète correspondant à $q_{34}$ reste indétectable par notre algorithme. En fait, on constate sur la Figure 4.21 que la période de cette planète et le paramètre $\theta_{\text{act},3}$ sont très proches. Or ce dernier correspond à la *pseudo-période* du processus gaussien ; on peut donc en déduire que la modélisation du bruit via un processus gaussien aura tendance à capturer des contributions des képlériennes.

**Étude plus approfondie pour les systèmes 2 et 5 : des lois *a posteriori* bimodales.** Dans les cas des systèmes 2 et 5, les moyennes de la séquence de Bernoulli sont plus élevées aux alentours des indices théoriques, mais pas suffisamment pour donner lieux à des bonnes détections. Cependant, on peut observer que ces niveaux bas de $\widehat{\boldsymbol{q}}$ sont en fait dus à un caractère bimodal des échantillons. La Figure 4.22 illustre ce phénomène autour de la deuxième période de ces deux systèmes sur une chaîne.

Les figures (a)-(b) montrent l'évolution de la séquence de Bernoulli sur les 5000 dernières itérations, pour les indices d'intérêt ($\{49, 51, 52\}$ pour le système 2 et $\{37, 38\}$ pour le système 5). Dans les deux cas, on observe le phénomène d'*alternance/anti-corrélation* pour lequel nous avons donné une nouvelle stratégie de détection en Tableau 3.12 (p. 139). On peut calculer la séquence *fusionnée* (représentée au dessus des échantillons dans les figures (a)-(b)) :

$$q_{\Omega}^{(t)} = \bigoplus_{j \in \Omega} q_j^{(t)}$$

Dans les deux cas, la séquence fusionnée $q_{\Omega}^{(t)}$ est égale à 1 pour presque toutes les itérations : 92,1% pour le système 2 et 97,8% pour le système 5. Un seuil $\alpha = 0.85$ autoriserait donc leur fusion. Pour le système 2 celle-ci n'est pas effectuée car les intervalles concernés ne sont pas voisins, au contraire, du système 5 ; on aurait donc dû avoir une fusion d'intervalle pour ce dernier.

Les histogrammes des échantillons des périodes aux indices $\Omega$ sont représentés en figures (c)-(d) et nous permettent de comprendre les raisons de l'échec de fusion pour le système 5. En effet, on n'observe pas un seul mode comme dans le cas des fusions observées au chapitre précédent, mais deux, dont l'un est centré dans l'intervalle $\mathcal{I}_{37}$ et l'autre dans $\mathcal{I}_{38}$. Notons d'ailleurs que certains échantillons de l'intervalle $\mathcal{I}_{38}$ font en fait partie du mode centré dans l'intervalle $\mathcal{I}_{37}$.

Dans les deux cas, ce comportement d'*alternance/anti-corrélation* n'illustre pas une période oscillant autour de la séparation entre deux intervalles, mais bien un caractère bimodal de la loi *a posteriori* puisqu'on obtient dans les deux cas, deux modes bien séparés. De telles observations nous amènent à conclure, dans les deux cas, à la présence presque sûrement (supérieure à 90%) d'une seule exoplanète, mais dont la localisation est incertaine, dans l'un ou l'autre des deux modes disjoints.

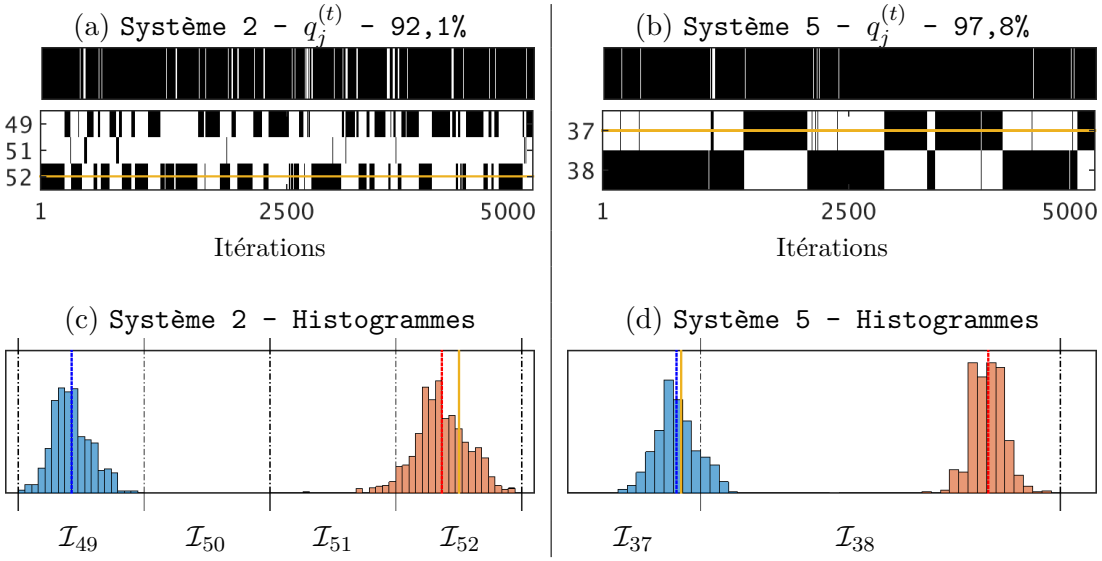



FIG. 4.22 – Observation du caractère bimodal des échantillons pour les sytèmes 2 (à gauche) et 5 (à droite) de [NFB+18]. En haut : évolution de la séquence de Bernoulli sur les 5000 dernières itérations pour les indices d'intérêt. Au dessus est réprésenté la fusion des $q_j$ sur les indices d'intérêt (on a toujours le noir pour $q_j = 1$). En bas : histogrammes des périodes dans les intervalles d'intérêt, en regroupant par couleur les deux modes (rouge et bleu), dont les moyennes sont représentées par des traits verticaux. Le trait jaune vertical indique la valeur théorique.

### Bilan sur ces tests

Ces tests et notre étude plus approfondie pour les systèmes 2 et 5 ont permis d'observer le comportement de notre échantillonneur sur modèle BGE pour des signaux simulés avec activités stellaires. Sur les 6 systèmes testés, nous n'obtenons aucune fausse détection (exoplanète non-présente dans le système). Plus précisément,

- pour 3 systèmes (1-3-6), les deux exoplanètes ont été correctement localisées.
- pour 1 système (2), la première exoplanète est correctement localisée, et la localisation de la seconde est bimodale dans des intervalles non-voisins mais peu éloignés.
- pour 1 système (5), la première exoplanète est détectée et correctement localisée avec une probabilité de 50% environ (ce qui corrobore son caractère « impossible » supposé par [NFB+18]), et la localisation de la seconde est bimodale dans des intervalles non-voisins mais peu éloignés.
- pour 1 système (4), le seconde exoplanète est correctement localisée tandis que la première n'est pas détectée, car sa période est très proche de la composante semi-périodique $\theta_{\text{act},3}$ du processus-gaussien utilisé pour générer l'activité stellaire.

Notons que notre échantillonneur a un très bon comportement puisque les conclusions sont les mêmes sur les dix graines, ce qui tend à montrer qu'il a bien convergé. Les résultats d'estimation sont récapitulés dans le Tableau 4.5.

| **Système** | **Détectabilité** | **Périodes** (jours) | **Périodes estimées** (jours) | **Probabilité de détection** (%) |
|---|---|---|---|---|
| 1 | facile | 12.1 | 12.11 | 100 |
| | facile | 42.4 | 42.20 | 100 |
| 2 | facile | 15.96 | 15.96 | 100 |
| | difficile | 120.5 | 90.47 ou 118.95 | 92.1 (44.07 ou 55.93) |
| 3 | difficile | 40.4 | 40.15 | 100 |
| | difficile | 91.9 | 92.06 | 98.02 |
| 4 | facile | 169.1 | 171.59 | 85.76 |
| | « impossible » | 23.45 | ∅ | ∅ |
| 5 | difficile | 31.1 | 31.06 ou 33.67 | 97.8 (45.77 ou 54.23) |
| | « impossible » | 10.9 | 10.92 | 41.32 |
| 6 | difficile | 40.4 | 40.56 | 96.56 |
| | difficile | 91.9 | 91.74 | 99.94 |

TAB. 4.5 – Récapitulatif des exoplanètes détectées pour les 6 systèmes avec le modèle BGE sur une chaîne d'échantillons. Périodes estimées ainsi que probabilités de présence des exoplanètes.

Nous avons ainsi montré dans cette section que le modèle BGE ainsi que l'échantillonneur proposé permet de détecter des exoplanètes à partir de signaux (simulés) de vitesses radiales perturbés par de l'activité stellaire. Ces résultats montrent aussi l'importance d'une étude des échantillons afin de déterminer de possibles modes bimodaux et le danger de l'utilisation de processus gaussien pouvant être amené à cachée une partie de la contribution périodique des exoplanètes dans le signal de vitesses radiales.

Enfin, nous avons pu calculer les facteurs de Bayes (BF) et les «Posterior odds ratio» (POR) pour chaque système ; ces derniers sont récapitulés dans le Tableau 4.6. Ces facteurs sont infinis quand aucun échantillon avec $L^{(t)} = L$ n'existe au contraire d'échantillons $L^{(t)} = L + 1$. C'est notamment le cas pour les facteurs « 2 vs. 1 » pour les systèmes 1 et 3, pour lesquels on a bien détecté deux exoplanètes. Cependant, le BF « 2 vs. 1 » du système 6 est plus faible (50), alors que les deux exoplanètes sont clairement détectées dans l'histogramme cumulé des $\widehat{\boldsymbol{q}}$ en Figure 4.21. On en conclut que, pour notre méthode, l'observation de ces facteurs apporte très peu d'information, contrairement à celle des histogrammes cumulés et plus généralement à l'étude des échantillons comme effectuée précédemment.

| | Système 1 | | Système 2 | | Système 3 | | Système 4 | | Système 5 | | Système 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BF | POR | BF | POR | BF | POR | BF | POR | BF | POR | BF | POR |
| 1 vs 0 | . | . | $+\infty$ | $+\infty$ | . | . | 9.2 | 13 | 60 | 117 | $+\infty$ | $+\infty$ |
| 2 vs 1 | $+\infty$ | $+\infty$ | 40 | 56 | $+\infty$ | $+\infty$ | 0.61 | 0.57 | 1.1 | 1.4 | 50 | 89 |
| 3 vs 2 | 0.87 | 0.93 | 0.81 | 0.84 | 1 | 1.2 | 0.63 | 0.44 | 0.82 | 0.78 | 1.2 | 1.7 |
| 4 vs 3 | 0.75 | 0.63 | 0.67 | 0.55 | 0.78 | 0.76 | 0.83 | 0.46 | 0.79 | 0.6 | 0.87 | 0.92 |
| 5 vs 4 | 0.74 | 0.52 | 0.7 | 0.48 | 0.82 | 0.66 | 0.89 | 0.41 | 0.83 | 0.52 | 0.83 | 0.73 |

TAB. 4.6 – Récapitulatif des « facteurs de Bayes » (BF) et des «Posterior odds ratio» (POR) de notre modèle pour chaque système de [NFB+18].

## 4.5 Premiers tests sur des données réelles

Nous présentons les résultats de premiers tests du modèle BGE et de l'échantillonneur associé sur des données réelles de vitesses radiales de deux étoiles : le système HD40307, qui nous a servi de base pour les tests sur données simulées en § 4.3 et le système CoRoT-7, popularisé grâce à la découverte d'une Super-Terre en orbite autour de l'étoile.

### 4.5.1 Système HD40307

Le système HD40307 a déjà été présenté en § 4.3, en se basant sur l'étude de [TAEG+13]. Ces derniers ont basé leur étude sur les vitesses radiales mesurées à l'aide du spectrographe HARPS, dont les valeurs sont mises à disposition dans l'article. Nous analyserons les deux signaux de vitesses radiales qui sont donnés dans l'article en Tableau A.2. Le premier a été calculé à partir de l'ensemble du spectre de l'étoile ; on le notera $\boldsymbol{y}_F$ (pour *Full spectra*). Le second a été calculé à partir du spectre « rouge » de l'étoile afin de réduire l'effet de son émission chromosphérique (voir Section 6. de l'article) ; on le notera $\boldsymbol{y}_R$ (pour *Red spectra*). À partir de l'article, on récupère aussi les erreurs de mesures $\sigma_n$ et les instants d'échantillonnage $t_n$ qui sont les mêmes que ceux utilisés en § 4.3 ; la fenêtre spectrale est donc déjà représentée en Figure 4.8(b).

Pour rappel, six exoplanètes ont été détectées par [TAEG+13], grâce à des méthodes de type périodogramme et échantillonnage stochastique, et à une prise en compte de l'activité stellaire avec un modèle de type « Moving Average ». Les valeurs estimées des paramètres orbitaux principaux ont été données en Tableau 4.2.

Pour l'exploitation du modèle BGE, nous gardons la même grille de recherche que celle utilisée en § 4.3 pour laquelle les périodes détectées par [TAEG+13] se trouvent dans les intervalles d'indices $j =$(19, 28, 38, 44, 49, 66), ainsi que la plupart des réglages de l'échantillonneur donné en Tableau 4.3. Pour filtrer la potentielle activité stellaire contenue par les signaux, nous pouvons être amenés à utiliser un processus gaussien, similaire à celui utilisé en § 4.4 pour les données simulées avec activité stellaire. Il sera donc nécessaire d'échantillonner les hyperparamètres du bruit $\boldsymbol{\theta}_{\text{act}}$ et $\sigma^2_{\text{jit}}$, auxquels nous donnons des lois *a priori* uniformes sur des intervalles assez grands, n'ayant pas de connaissance *a priori* sur leur valeur. De même, en suivant les recommandations données en § 3.4.3, la variance des amplitudes $\sigma_x^2$ suit une loi *a priori* inverse gamma, réglée de la même façon qu'en § 4.4.

Avec l'échantillonneur de type PCGS (Tab. 3.10), nous avons échantillonné sur 200 000 itérations deux types de loi cible : $p(\boldsymbol{q}, \boldsymbol{P}, \boldsymbol{e}, \boldsymbol{\phi}, \boldsymbol{x}, \lambda, \sigma_x^2, \sigma^2_{\text{jit}}, \boldsymbol{\theta}_{\text{act}} \,|\, \boldsymbol{y})$ (avec processus gaussien) et $p(\boldsymbol{q}, \boldsymbol{P}, \boldsymbol{e}, \boldsymbol{\phi}, \boldsymbol{x}, \lambda, \sigma_x^2, \sigma^2_{\text{jit}} \,|\, \boldsymbol{y})$ (sans processus gaussien), et ce avec 10 initialisations différentes de graines aléatoires.

En Figure 4.25 sont représentés les histogrammes cumulés des moyennes $\widehat{\boldsymbol{q}}$ calculées sur les 10 000 derniers échantillons, pour les deux signaux ($\boldsymbol{y}_F$ et $\boldsymbol{y}_R$) et pour les deux types de loi cible (avec ou sans processus gaussien).

D'abord, notons que lorsqu'on n'utilise pas de processus gaussien pour modéliser l'activité, de nombreux pics apparaissent dans les histogrammes cumulés, notamment pour le signal $\boldsymbol{y}_F$, tandis que la plupart de ces pics disparaissent quand le processus gaussien est employé. Cela dénote clairement la présence d'un bruit stellaire que le jitter seul ne peut prendre en compte.

À partir des résultats avec processus gaussien, on remarque que les 3 premières exoplanètes (indices $j = 19, 28, 38$) sont détectées avec une probabilité de presque 100%, pour les deux signaux. Cependant, les 3 planètes suivantes (indices $j = 44, 49, 66$) ne ressortent que faiblement dans l'histogramme cumulé pour $\boldsymbol{y}_R$ et elles en sont presque absentes pour $\boldsymbol{y}_F$.

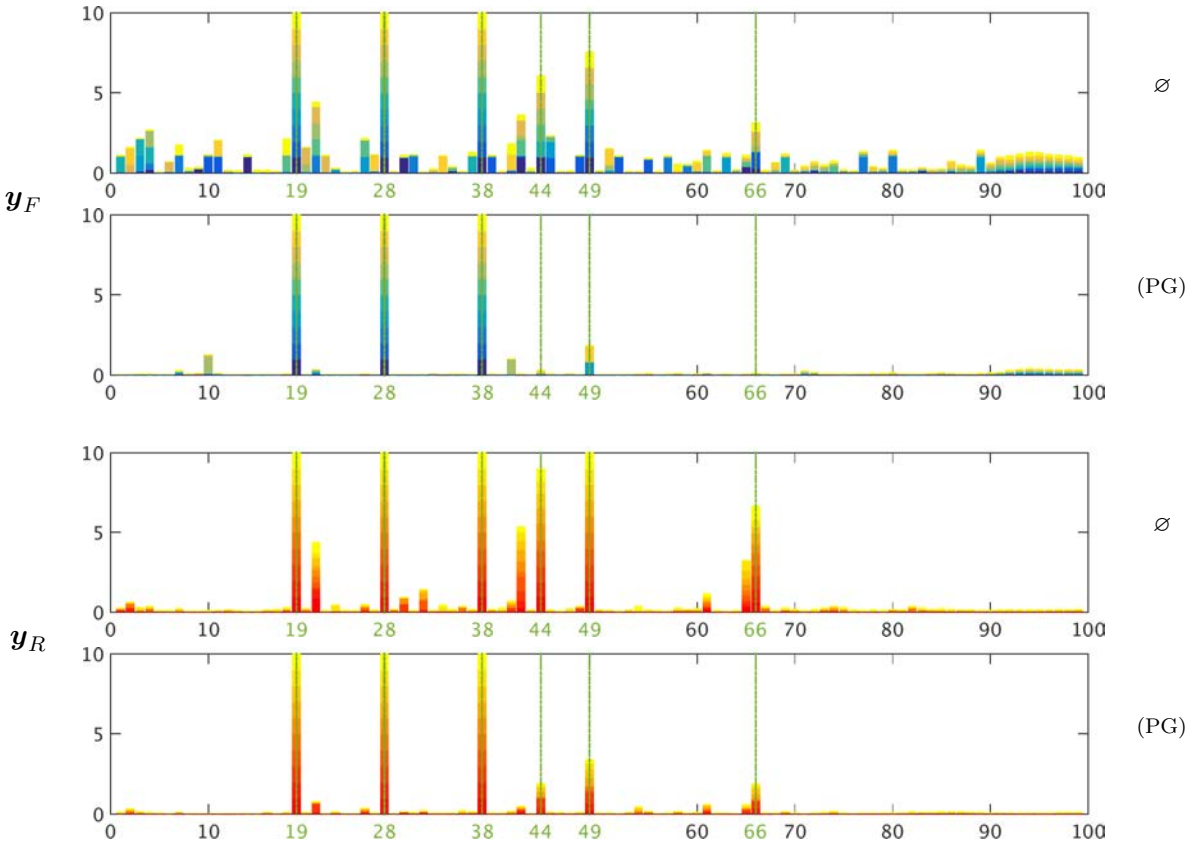



FIG. 4.23 – HD40307 - Données réelles. Histogramme cumulé des $\widehat{\boldsymbol{q}}$ sur les 10 graines aléatoires. Pour chaque signal ($\boldsymbol{y}_F$ et $\boldsymbol{y}_R$), on donne les résultats avec exploitation d'un processus gaussien (PG) ou sans (∅). Les lignes verticales indiquent les indices correspondant aux périodes des exoplanètes détectées par [TAEG+13].

Notamment, la planète HD40307 g proposée par [TAEG+13], qui serait dans la « zone habitable » de l'étoile et qui correspond à l'indice $j = 44$ de notre grille, n'excède pas les 25% de probabilité pour le signal $\boldsymbol{y}_R$.

Au regard des résultats de notre méthode, on peut donc émettre des doutes quant à l'existence de ces trois planètes. Cependant, notons que le réglage de notre méthode n'est pas idéal. En effet, d'une part nous ne prenons en compte aucune information *a priori* sur l'activité stellaire - lois *a priori* très peu informatives sur les hyperparamètres. Notons tout de même que d'après [TAEG+13], HD40307 est une étoile très peu active. D'autre part, les paramètres des lois de proposition ne sont pas réglés de façon optimale, dans le sens où les taux d'acceptation sont très faibles. Il se pourrait donc que ces résultats soient dus à une convergence non atteinte (bien qu'on retrouve le même comportement sur les séquences $\boldsymbol{q}$ pour les différentes chaînes). Des conclusions plus définitives nécessiteraient une étude plus approfondie de ces données, en collaboration avec des astrophysiciens connaissant les propriétés de cette étoile.

### 4.5.2 Système CoRoT-7

Le système CoRoT-7 est devenu populaire grâce à la découverte autour de cette étoile de l'une des premières « Super-Terre », CoRoT-7 b, annoncée par [LRS+09]. Elle a été détectée par la méthode des *transits* à partir des observations effectuées avec le satellite CoRoT. Afin de mieux étudier le système (et notamment de déterminer la masse des planètes détectées par transits), une campagne de mesures de la vitesse radiale de l'étoile CoRoT-7 a ensuite été menée avec le télescope HARPS ; c'est sur ces données que nous allons tester nos méthodes.

L'étude des vitesses radiales de CoRoT-7 a été réalisée par plusieurs équipes. Par exemple, [HCCQ+14] s'est concentré sur la modélisation de l'activité de l'étoile via un processus gaussien, en réglant ses hyperparamètres via une étude d'indicateurs d'activité indépendants des contributions planétaires. L'étoile CoRoT-7 est en effet une étoile relativement active. Nous nous basons ici sur les travaux de [FHB+16] ; ces derniers utilisent la méthode de « Trans-dimensionnal MCMC » (*Diffusive Nested Sampling* avec étapes de *Naissance/Mort*, brièvement présentée en § 4.1.2) associée à la prise en compte de l'activité stellaire avec un processus gaussien. Les données de vitesses radiales mesurées, ainsi que les temps d'échantillonnage et les erreurs de mesures sont accessibles en ligne[19]. Le signal $\boldsymbol{y}$ est composé de $N = 177$ échantillons et la fenêtre spectrale associée à l'échantillonnage est représentée en Figure 4.24, où l'on peut remarquer les forts lobes secondaires aux multiples de 1 jour$^{-1}$. Les auteurs de [FHB+16] ont confirmé l'existence des planètes CoRoT-7b ($P \approx 0.85$ jour) et CoRoT-7c ($P \approx 3.70$ jours) déjà annoncée par d'autres publications. Ils relèvent aussi la présence de deux périodicités potentiellement dues à des planètes, aux périodes $P \approx 2$ jours et $P \approx 9$ jours.

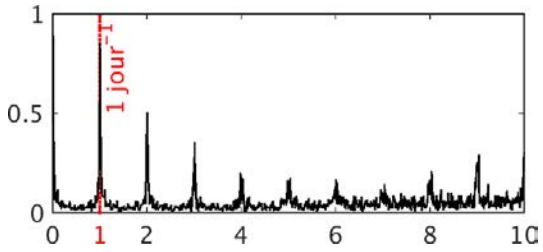



FIG. 4.24 – CoRoT-7 - Données réelles. Fenêtre spectrale associée à l'échantillonnage du signal.

Pour analyser le signal $\boldsymbol{y}$, nous exploitons donc le modèle BGE avec modélisation de l'activité stellaire par un processus gaussien, comme pour HD40307. L'échantillonneur associé a donc pour loi cible $p(\boldsymbol{q}, \boldsymbol{P}, \boldsymbol{e}, \boldsymbol{\phi}, \boldsymbol{x}, \lambda, \sigma_x^2, \sigma_{\text{jit}}^2, \boldsymbol{\theta}_{\text{act}} \,|\, \boldsymbol{y})$. La loi *a priori* pour la variance $\sigma_x^2$ est une loi inverse gamma. La largeur des intervalles $\Delta$ dépend de la définition de la grille de recherche des périodes. Comme [FHB+16], nous recherchons les périodes entre $P^{\min} = 0.1$ jour et $P^{\max} = 1000$ jours ; la grille est constituée de $J = 100$ points régulièrement espacés en logarithme, si bien que $\Delta = 0.0404$. Notons que les périodes relevées par [FHB+16] se trouvent dans les intervalles $\mathcal{I}_j$ d'indices $j = (24, 33, 39, 49)$ de notre grille (voir Tableau 4.7).

| CoRoT-7 | b | ? | c | ? |
|---|---|---|---|---|
| Périodes (jours) | 0.85 | 2 | 3.7 | 9 |
| Indices $j$ | 24 | 33 | 39 | 49 |

TAB. 4.7 – Correspondance entre les périodicités repérées par [FHB+16] et les intervalles de la grille pour le modèle BGE.

[19] Le lien est donné dans l'article : `https://github.com/j-faria/exoBD-CoRoT7`

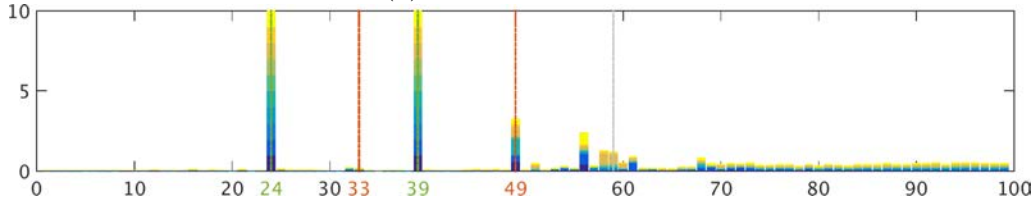
Pour les hyperparamètres du bruit $\sigma^2_{\text{jit}}$ et $\boldsymbol{\theta}_{\text{act}}$, nous prenons une loi *a priori* uniforme, dans des intervalles correspondant aux intervalles de définition pris par [FHB+16]. La proposition de ces hyperparamètres s'effectuera par une loi de mélange entre marche aléatoire et loi *a priori*.

*Notons que contrairement aux tests précédents, nous avons fait ici un réel effort sur le réglage des variances instrumentales des marches aléatoires des étapes de Métropolis-Hastings (pour les paramètres non linéaires $(\boldsymbol{P}, \boldsymbol{e}, \boldsymbol{\phi})$ et les hyperparamètres du bruit $(\sigma^2_{jit}, \boldsymbol{\theta}_{act})$) pour atteindre un taux d'acceptation de l'ordre de 20 à 30%.*

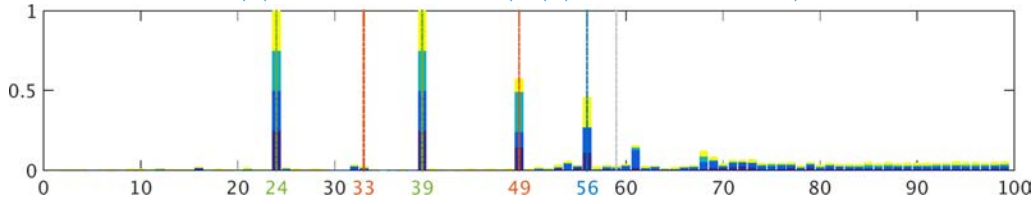
L'échantillonnage de la loi cible avec l'échantillonneur de type PCGS a été effectué sur 500 000 itérations, et ce avec 10 initialisations différentes de graines aléatoires. Les moyennes $\widehat{\boldsymbol{q}}$ sont calculées sur les 10 000 derniers échantillons.

En Figure 4.25(a), on peut observer l'histogramme cumulé des $\widehat{\boldsymbol{q}}$ pour les dix graines d'initialisation. On y retrouve bien les périodes associées aux exoplanètes CoRoT-7b et CoRoT-7c avec une probabilité de presque 100%. Cependant, il n'y a aucune trace de la possible planète autour de $P = 2$ jours ($j = 33$). Une composante périodique ressort bien autour de la période $P = 9$ jours ($j = 49$) mais avec une faible probabilité, inférieure à 50%. De plus, de faibles pics ressortent aussi aux alentours de $j = 59$, qui correspond à l'intervalle où se trouve la période de rotation de l'étoile ($P_{\text{rot}} \approx 22.3$ jours, voir [FHB+16]). Enfin, les $\widehat{\boldsymbol{q}}$ semblent différents pour les périodes au delà de l'indice $j = 48$ selon la graine aléatoire.

(a) Toutes les graines




(b) Graines 1-2-5-10 ($R(\tau)$ quasi-périodique)




(c) Graines 3-4-6-7-9 ($R(\tau)$ atténuée)

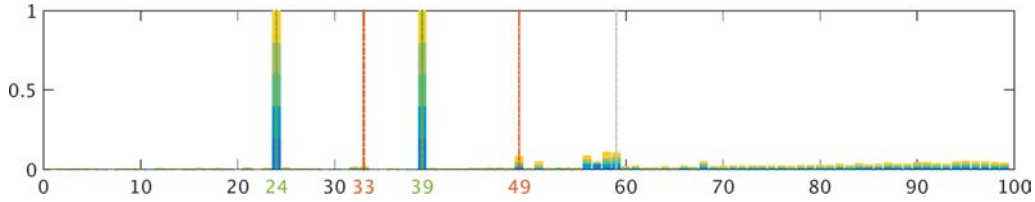



FIG. 4.25 – CoRoT-7 - Données réelles. Histogramme cumulé des $\widehat{\boldsymbol{q}}$ (a) sur les 10 graines aléatoires (b) sur les graines (1-2-5-10) et (c) sur les graines (3-4-6-7-9). Les lignes verticales vertes indiquent les indices correspondant aux périodes des exoplanètes détectées par [FHB+16], et les lignes rouges indiquent leurs détections incertaines. La ligne grise correspond à l'indice associé à la période de rotation de l'étoile $P_{\text{rot}}$.

En Figure 4.26, les histogrammes cumulés des échantillons des hyperparamètres $(\lambda, \sigma_x^2, \sigma_{\text{jit}}^2)$ sont représentés ; les échantillons semblent suivre la même loi quelle que soit la graine observée.

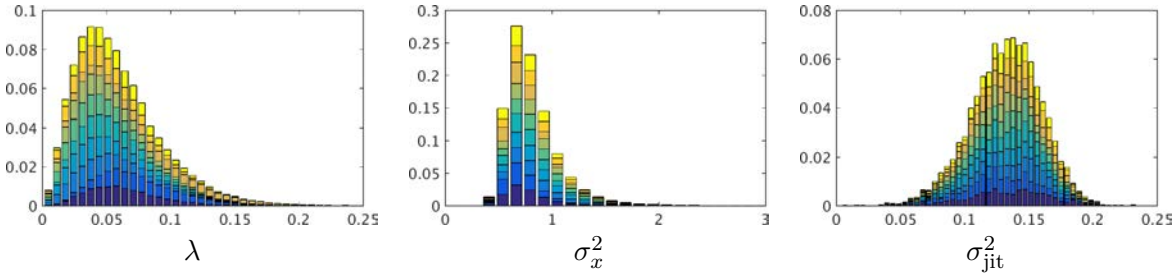



FIG. 4.26 – CoRoT-7 - Données réelles. Histogrammes cumulés des hyperparamètres ($\lambda$, $\sigma_x^2$,$\sigma_{\text{jit}}^2$).

Cependant, il n'en va pas de même pour les paramètres $\boldsymbol{\theta}_{\text{act}}$ du processus gaussien ; en effet, suivant les graines observées, les valeurs estimées peuvent être très différentes. En Figure 4.27, nous avons représenté la fonction de covariance du processus gaussien (4.8), exprimée en les paramètres $\widehat{\boldsymbol{\theta}}_{\text{act}}$ estimés par EAP sur les 10 000 dernières itérations. On distingue très clairement 3 « modes » différents, que nous avons regroupés par couleur ; en rouge, la fonction de covariance s'amortit très vite (en moins de dix jours) tandis qu'en bleu, elle s'amortit plus lentement et présente des pics de périodicité de l'ordre de 22 jours, correspondant à la valeur de $\widehat{\theta}_{\text{act},3}$ estimée pour ces graines. Ce résultat est très intéressant, car la période de rotation $P_{\text{rot}}$ de l'étoile CoRoT-7 a été estimée à 22.3 jours environ. La périodicité que l'on retrouve dans la fonction de covariance et pour le paramètre $\theta_{\text{act},3}$ semble donc être liée à la période de rotation de l'étoile. Enfin, notons que la fonction de covariance en jaune correspond à une unique graine pour laquelle l'étude des échantillons permet de déterminer que l'échantillonneur n'a pas convergé.

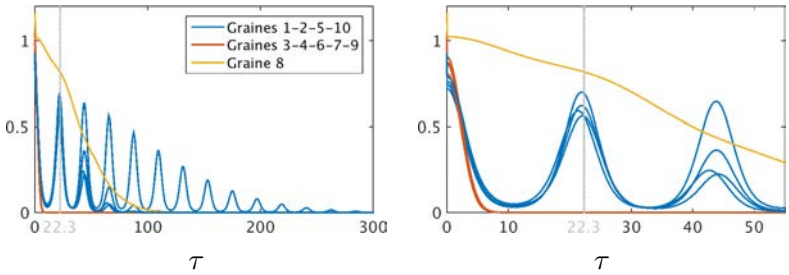



FIG. 4.27 – CoRoT-7 - Données réelles. Fonctions de covariance $R(\tau)$ exprimée en les paramètres $\boldsymbol{\theta}_{\text{act}}$ estimés ; à droite, zoom sur les faibles décalages.

Au regard des fonctions de covariance, on peut donc penser que la loi échantillonnée est bimodale ; dans un cas, le processus gaussien identifie une fonction quasi-périodique de période d'environ 22 jours, et dans un autre cas une fonction atténuée. Cela est sans doute dû au fait que le modèle de processus gaussien utilisé est très souple et peut s'adapter à la fois à des phénomènes quasi-périodiques ou à des fonctions de covariance amorties. En Figures 4.25(b-c), nous avons alors à nouveau tracé les histogrammes cumulés des moyennes $\widehat{\boldsymbol{q}}$, mais en séparant les graines avec fonction de covariance quasi-périodique (Figure(b)) de celles avec fonction de covariance atténuée (Figure(c)) (les histogrammes sont normalisés par le nombre de graines pris en compte). Dans le premier cas (Figure(b)), la périodicité autour de $P = 9$ jours (intervalle $j = 49$) ressort davantage. Par ailleurs, on voit aussi apparaître un pic au niveau de $j = 56$ ($P \approx 17$ jours). Une observation des échantillons $\boldsymbol{q}^{(t)}$ pour les indices $j = 49$ et $j = 56$ nous a permis de détecter un comportement anti-corrélé entre ces deux intervalles : pour chaque graine considérée, la moyenne de la séquence fusionnée $q_{49,56}^{(t)} = q_{49}^{(t)} \oplus q_{56}^{(t)}$ est supérieure à 85%, laissant

penser à une distribution bimodale.

Dans le second cas (Figure(c)), le processus gaussien ne modélise pas de phénomène quasi-périodique dans l'activité de l'étoile, et des pics, certes de très faibles amplitudes, apparaissent dans les histogrammes autour des indices correspondant à la période de rotation de l'étoile ($P_{\text{rot}}$ ≈22.3 jours, $j = 59$). Cela laisse penser que l'activité quasi-périodique identifiée dans les autres graines est modélisée ici par les atomes du dictionnaire, sous forme képlérienne. Dans ce cas, le pic à $j = 49$ ($P = 9$ jours) a une très faible probabilité.

Pour conclure sur ces observations, deux modes distincts ressortent de nos échantillons. Le premier possède deux exoplanètes, correspondant à CoRoT-7b et à CoRoT-7c, avec une modélisation de l'activité sans réelle périodicité apparante, mais des planètes de très faible probabilité à une période de l'ordre de la période de rotation de l'étoile. Le second fait apparaître, en plus des deux précédentes exoplanètes, une troisième périodicité se trouvant soit aux alentours de $P$=9 jours ($j = 49$) soit de $P$=17 jours ($j = 56$), tout en modélisation une activité avec une périodicité autour de 22 jours, proche de la période de rotation de l'étoile. Notons que ce comportement peut être contre intuitif, puisque on est souvent amené à considérer que le processus gaussien peut absorber une partie des périodicités du signal à cause de $\theta_{\text{act},3}$, tandis qu'ici, la convergence de $\theta_{\text{act},3}$ vers la période de rotation de l'étoile amène l'apparition d'une composante képlérienne pas très éloignée. De plus, ce comportement bimodale avait aussi été observé par [FHB+16] dans leurs résultats de simulation. Il est pour le moment difficile de conclure sur la possible existence d'une exoplanète aux périodes proches de $P$=9 jours ou $P$=17 jours (mais pas aux deux), mais une étude plus approfondie des échantillons pourrait peut-être apporter davantage d'informations.

## 4.6 Conclusion

Dans ce chapitre, nous avons pu nous intéresser à un problème passionnant d'astrophysique : la détection d'exoplanètes, ces mondes extrasolaires que le grand public s'imagine coloniser et que les astrophysiciens étudient pour mieux comprendre la formation et le fonctionnement de notre système solaire. Après une première étude qui nous a permis de comprendre comment aborder le problème et quels en sont les enjeux actuels, nous avons pu adapter et exploiter notre modèle BGE et l'échantillonneur associé développé au Chapitre 3 pour proposer une alternative inédite aux méthodes actuelles utilisées par les astrophysiciens. Nos tests numériques, réalisés sur des données simulées, sans et avec activité stellaire, mais aussi sur des données réelles, ont permis de montrer les atouts de notre méthode, notamment au niveau du problème de l'estimation du nombre d'exoplanètes. Nous montrons en effet que l'estimation seule de l'ordre du modèle $L$ (via par exemple le calcul du facteur de Bayes) n'est pas suffisante ; l'étude plus approfondie des échantillons et des séquences de Bernoulli $\widehat{\boldsymbol{q}}$ et des paramètres orbitaux $\widehat{\boldsymbol{\nu}}$ estimés permet de donner des niveaux de confiance pour chaque exoplanète détectée. Au contraire, les méthodes de type *comparaison des vraisemblances marginales* ne permettent de comparer que les probabilités des nombres d'exoplanètes. Par ailleurs, la prise en compte d'un modèle de processus gaussien pour modéliser l'activité stellaire ne pose pas de problème pour notre modèle BGE et l'algorithme d'échantillonnage proposé. Des astrophysiciens avertis peuvent donc régler les paramètres du processus via des informations indépendantes des contributions des exoplanètes (comme la période de rotation de l'étoile et les indicateurs d'activité).

Cette application du modèle BGE à un problème concret nous a aussi permis de confirmer le bon comportement de nos algorithmes d'échantillonnage, puisque même sur des données difficiles avec activité stellaire, les échantillons de dix chaines différentes menaient aux mêmes conclusions.

Cependant, quelques pistes restent à explorer pour améliorer l'interprétation des échantillons. L'aspect bimodal de la loi *a posteriori*, observé dans le cas des données simulées avec activité stellaire, nécessite une étude approfondie des échantillons pour être remarqué. Nous n'avons pas à l'heure actuelle de méthode automatique pour détecter un tel phénomène. On pourrait construire des outils permettant de constater plus aisément les corrélations entre les échantillons de Bernoulli, mais aussi entre les paramètres de localisation, afin de pouvoir distinguer les phénomènes de modes multiples et les paramètres de localisation répartis sur plusieurs intervalles.

Le découpage de la grille est d'ailleurs une limite de notre méthode pour ce problème. En effet, le découpage naturel régulier en fréquence n'est pas envisageable, puisqu'il nécessiterait un trop grand nombre d'intervalles afin de respecter la résolution « physique » en période (pas plus d'une exoplanète par intervalle). Des études pourraient être menées pour observer le comportement de notre algorithme quand deux exoplanètes se retrouvent dans le même intervalle. Le choix que nous avons fait d'utiliser un découpage en logarithme a montré son intérêt dans les différents tests que nous avons effectués.

Notons que les performances de notre échantillonneur, telles que présentées dans ce chapitre, pourraient être améliorées sans grande difficulté. En effet, nous avons choisi, à la fois pour les paramètres non linéaires, mais aussi pour les paramètres liés à l'activité stellaire, des lois *a priori* et des lois de proposition très simples (lois uniformes, et lois de mélange entre marches aléatoires et lois *a priori*). Les variances instrumentales des marches aléatoires entraînent de plus des taux d'acceptation particulièrement bas, de l'ordre de 1% (sauf pour les tests réalisés sur les mesures de vitesses radiales de CoRoT-7). Bien que cela n'ait pas gêné les bonnes performances de notre échantillonneur dans ce chapitre, un meilleur choix pour ces lois permettraient une convergence plus rapide de l'échantillonnage. Notamment, les astrophysiciens proposent une grande variété de lois *a priori* pour les paramètres du processus gaussien, qui peuvent être intégrées sans difficulté à notre algorithme. Par ailleurs, l'échantillonnage du paramètre de localisation devrait être adapté au découpage de l'espace en logarithme que nous utilisons.

Enfin, les résultats obtenus sur les données réelles, en particulier pour CoRoT-7, montrent que notre méthode est très prometteuse pour aider les astrophysiciens à analyser ces données. Des compétences astrophysiques supplémentaires seraient nécessaires pour une analyse plus approfondie et une conclusion plus définitive sur le nombre d'exoplanètes détectées ainsi que sur les paramètres orbitaux associés.

# Conclusion générale

## Contributions

Cette thèse visait à résoudre de façon « optimale » les problèmes inverses non linéaires parcimonieux présentés en chapitre 1, en prenant en compte de façon exacte la parcimonie via l'introduction de variables binaires et en estimant de façon continue les paramètres non linéaires.

Au chapitre 2, nous avons abordé ce problème dans un cadre déterministe en se basant sur le principe de l'interpolation de dictionnaire permettant de reformuler le problème en APLsc. Notre contribution majeure est la résolution de ce problème d'APLsc en exploitant la pseudo-norme $\ell_0$ grâce à une reformulation en MIP. Avant nos travaux, seules des résolutions sous-optimales (pénalisation $\ell_1$, algorithmes gloutons) étaient employées pour ce type de problème. Pour atteindre ce but, nous avons étudié dans un cadre général le principe de l'interpolation de dictionnaire, ce qui nous a permis de mettre en lumière les difficultés présentes pour des atomes continus à amplitudes réelles et complexes. Notamment, nous montrons la nécessité de prendre en compte des contraintes de disjonction dans le cas d'amplitudes à valeurs réelles, que nous sommes parvenus à prendre en compte dans la reformulation MIP. Nous proposons aussi des améliorations pour deux interpolations de dictionnaire précédemment proposées dans la littérature. Pour l'interpolation polaire, la contrainte quadratique généralement employée peut être remplacée à l'aide de contraintes linéaires, permettant une accélération non négligeable des calculs (dans le cadre de la reformulation MIP, on passe d'un MIQCP à un MIQP). Pour l'approximation par K-SVD, nous avons proposé une méthodologie correcte pour définir un espace de contraintes convexe associé à l'espace image ; surtout, nous démontrons que cette approximation peut se faire de façon continue. Nous montrons ainsi le véritable potentiel de cette interpolation de dictionnaire, notamment dans un cadre étendu de l'interpolation via notre application à l'analyse spectrale, où des paramètres de forme peuvent être pris en compte. Une comparaison de l'approche $\ell_0$ que nous proposons avec la pénalisation $\ell_1$ pour des problèmes de déconvolution impulsionnelle avec interpolation polaire a été exposée dans [BC18].

Les travaux présentés au chapitre 3 se situent dans un cadre stochastique, où la parcimonie est modélisée à l'aide de variables binaires suivant *a priori* une loi de Bernoulli. Une extension du modèle Bernoulli-Gaussien permet de prendre en compte les paramètres non linéaires dans le modèle bayésien. Nous proposons alors d'exploiter la marginalisation des amplitudes pour obtenir un échantillonneur plus efficace de la loi *a posteriori*. Dans le cadre supervisé, nous avons adapté les astuces calculatoires permettant une mise en œuvre performante de l'échantillonnage des variables de Bernoulli suivant la loi *a posteriori* marginalisée au cas où le bruit n'est pas forcément *i.i.d.* et où les amplitudes peuvent être vectorielles. Surtout, ces astuces ont été adaptées pour l'échantillonnage suivant la loi marginalisée des paramètres non linéaires. Dans le cadre non supervisé, nous proposons un algorithme de type PCGS pour échantillonner la loi jointe tout en bénéficiant de la marginalisation des amplitudes pour les paramètres d'intérêt. Les

tests réalisés sur des données d'analyse spectrale en échantillonnage irrégulier (voir [BCBB16]) ont permis de montrer les atouts des échantillonneurs proposés, pour lequels nous proposons un schéma d'estimation plus adapté. Notons que des travaux récents [BHF17, HFR18] ont exploité un type de modèle Bernoulli-Gaussien étendu avec des méthodes de résolution sous optimales, mais sans exploiter la répartition des paramètres de localisation sur une grille (la loi *a priori* de chaque paramètre de localisation est uniforme sur tout le domaine des localisations). Nous sommes convaincus que le modèle BGE proposé dans ce manuscrit est plus intéressant, car il est exempté des problèmes d'identifiabilité des paramètres de localisation grâce à l'exploitation de la grille, et cette dernière n'est pas limitante puisque l'échantillonneur proposé permet de s'en libérer (les échantillons peuvent passer d'un intervalle à un autre sans difficulté).

Enfin, nous avons montré en chapitre 4 l'intérêt du modèle BGE et de l'échantillonneur associé pour un problème d'actualité en astrophysique : la détection d'exoplanètes par la méthode des vitesses radiales. Ce dernier peut en effet être formulé comme un problème inverse non linéaire parcimonieux. Nous montrons que le modèle BGE peut facilement s'adapter à la prise en compte de l'activité stellaire via la modélisation du bruit par un processus gaussien, qui est une des particularités de ce problème. Les tests que nous avons réalisés sur des données simulées, à la fois avec et sans activité stellaire, révèlent les atouts du modèle BGE pour ce problème, notamment pour l'estimation du nombre d'exoplanètes qui représente un véritable enjeu pour les astrophysiciens. En effet, une étude plus détaillée des échantillons des variables de Bernoulli permet de mettre en évidence le caractère bimodal des lois *a posteriori*, invisible lorsque l'on observe uniquement l'ordre du modèle. Une telle étude est nécessaire pour tenter d'attribuer une probabilité d'existence aux exoplanètes repérées. Notre étude du problème nous permet donc de proposer notre méthode comme une véritable alternative aux méthodes existant actuellement pour la détection d'exoplanètes, ce que confirment les tests effectués sur les données réelles.

## Perspectives

Les résultats concluants obtenus avec les méthodes proposées dans les cadres déterministe et stochastique permettent d'envisager plusieurs perspectives.

La formulation en MIP des problèmes d'APLsc issus de l'interpolation de dictionnaire permet de retrouver le support optimal, mais au prix d'un coût de calcul important (une exploration exhaustive est quant à elle inenvisageable), comme nous l'avons vu lors des tests. Des adaptations spécifiques des algorithmes de résolution des MIP aux spécificités des problèmes parcimonieux sont actuellement envisagées pour accélérer ces résolutions [MBN19] dans le cas des APL ; une étude pour les APLsc serait intéressante. Par ailleurs, nous avons vu que le critère de choix de l'ordre du modèle $L$ avec la résolution de type $\ell_0$ souffre de l'erreur d'approximation dans le cas de signaux peu bruités, amenant à l'estimation d'un support contenant des fausses détections. On pourrait alors envisager soit de prendre en compte ces erreurs d'approximation dans le critère de choix de $L$, soit de définir des interpolations de dictionnaire minimisant cette erreur. Nous avons également montré que l'erreur d'approximation n'est pas le seul caractère limitant de l'interpolation de dictionnaire ; la relaxation convexe de l'espace image, nécessaire à la formulation des problèmes en optimisation convexe, est préjudiciable à une bonne identification du support. L'approximation par K-SVD, qui minimise l'erreur d'approximation pour un ordre $K$ donné (nombre de colonnes des atomes interpolés), représente alors une vraie piste de recherche. Nous pensons qu'une étude mathématique plus approfondie de la K-SVD permettrait de proposer une interpolation de dictionnaire donnant à la fois une faible erreur d'approximation et exploitant une faible relaxation de l'espace image. De même, le principe de l'interpolation de dictionnaire

par approximation Minimax, proposé par [PPZ17], pourrait être approfondi pour proposer une approximation minimisant l'erreur maximale tout en bénéficiant d'une faible relaxation de l'espace image.

À propos du modèle BGE et de l'échantillonneur associé, nous envisageons deux pistes de recherche. D'une part, la marginalisation des amplitudes que nous avons proposée pour accélérer la convergence de l'échantillonneur n'est pas possible pour des atomes pour lesquels les amplitudes sont sous contraintes, par exemple de positivité ; en effet, l'intégration analytique de la loi *a posteriori* n'est pas réalisable dans ce cas. Un échantillonnage par K-uplets, comme proposé par [GILC11] dans le cas d'un modèle BG linéaire, pourrait être envisagé ; il faudrait alors l'étendre à la présence des paramètres non linéaires. D'autre part, nous avons mis en exergue des lois *a posteriori* bimodales pour les séquences de Bernoulli ; cela montre certes le bon comportement de notre échantillonneur qui ne reste pas bloqué dans un mode local, mais ces modes ne sont actuellement repérables que grâce à une étude attentive des échantillons. Le développement de méthodes automatiques de détection de tels phénomènes, à partir par exemple de calculs d'anti-corrélation entre les composantes de la séquence de Bernoulli, pourrait alors être envisagé.

Notons que nos différentes méthodes (formulation MIP avec interpolation de dictionnaire dans le cadre déterministe et modèle BGE avec l'échantillonneur exploitant la marginalisation des amplitudes dans le cadre stochastique) gagneraient à être comparées avec d'autres méthodes permettant l'estimation du paramètre continu ; comme par exemple des méthodes paramétriques de type processus ponctuels marqués [MCM+16]. Ces derniers sont souvent associés à des algorithmes d'échantillonnage de type RJ-MCMC dont il est difficile de vérifier la convergence et dont l'analyse des échantillons ne peut se faire généralement qu'à travers le maximiseur de la loi *a posteriori*.

Enfin, l'application courante de notre méthode de type BGE pour la détection d'exoplanètes pourrait bénéficier de nombreuses améliorations. Notamment, des lois *a priori* plus physiques sur les paramètres (telles que proposées dans [NFB+18] ou [FHB+16]) et des lois de proposition mieux réglées pour les étapes de Métropolis-Hastings, permettront d'assurer une meilleure convergence de l'échantillonneur, pour l'application à des données réelles. De plus, en se basant sur l'analyse d'indicateurs d'activité réalisée par des experts en astrophysique, de meilleurs réglages pourront être envisagés pour les lois *a priori* sur les paramètres des processus gaussiens, permettant de prendre en compte l'activité stellaire. Le choix du découpage du domaine des localisations mériterait de plus une étude plus approfondie. Il serait enfin intéressant de comparer notre méthode sur les divers challenges proposés dans la communauté astrophysique sur la détection d'exoplanètes. En définitive, notre méthode est une alternative à celles habituellement utilisées par la communauté astrophysique. L'analyse des échantillons qu'elle fournit apporte des informations supplémentaires aux astrophysiciens pour leur permettre d'améliorer leur analyse des vitesses radiales et on peut l'espérer, contribuer à la détection de nouvelles exoplanètes.

# Annexes

# Annexe A

# Annexe du chapitre 2

## A.1 Démonstration du théorème 1 p. 61 sur l'expression des paramètres $r$ et $\theta$ pour l'interpolation polaire

La démonstration suivante permet de prouver cette propriété, aussi bien pour les fonctions à valeurs réelles que complexes. Ici, nous notons $< \boldsymbol{a}, \boldsymbol{b} >= \boldsymbol{a}^\dagger \boldsymbol{b}$ le produit scalaire dans les complexes.

**Démonstration 1** *La démonstration se déroule en plusieurs étapes.*

*P1.* *(H1) + (H2)* $\Leftrightarrow \mathfrak{Re}(< \boldsymbol{h}_0, \boldsymbol{h}_p >) = \mathfrak{Re}(< \boldsymbol{h}_0, \boldsymbol{h}_m >)$.

$$\begin{array}{ll} & \|\boldsymbol{h}_0 - \boldsymbol{h}_m\|_2^2 = \|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2^2 \ \textit{(H2)} \\ \Leftrightarrow & \|\boldsymbol{h}_0\|_2^2 + \|\boldsymbol{h}_m\|_2^2 - 2\mathfrak{Re}(< \boldsymbol{h}_0, \boldsymbol{h}_m >) = \|\boldsymbol{h}_0\|_2^2 + \|\boldsymbol{h}_p\|_2^2 - 2\mathfrak{Re}(< \boldsymbol{h}_0, \boldsymbol{h}_p >) \\ \Leftrightarrow & \mathfrak{Re}(< \boldsymbol{h}_0, \boldsymbol{h}_m >) = \mathfrak{Re}(< \boldsymbol{h}_0, \boldsymbol{h}_p >) \ \textit{car (H1)} \end{array}$$

*P2.* *Orthogonalité sur les produits scalaires quels que soient* $(r, \theta)$ *:*
$\mathfrak{Re}(< \boldsymbol{c}, \boldsymbol{v} >) = \mathfrak{Re}(< \boldsymbol{u}, \boldsymbol{v} >) = 0$.
*Notons d'abord que* $\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}$ *peuvent s'écrire de la façon suivante, en inversant la matrice* $\mathbf{A}(r, \theta)$*, avec* $\alpha = (1 + \cos\theta)^{-1}$*,* $\beta = (2\sin\theta)^{-1}$ *et* $\boldsymbol{h}_s = \boldsymbol{h}_p + \boldsymbol{h}_m$ *:*

$$\boldsymbol{c} = \frac{\alpha}{2}\boldsymbol{h}_s + (1-\alpha)\boldsymbol{h}_0, \ \boldsymbol{u} = \frac{\alpha}{2r}(2\boldsymbol{h}_0 - \boldsymbol{h}_s), \ \boldsymbol{v} = \frac{\beta}{r}(\boldsymbol{h}_p - \boldsymbol{h}_m)$$

*Puis :*

$$\begin{array}{rcl} \mathfrak{Re}(< \boldsymbol{h}_p - \boldsymbol{h}_m, \boldsymbol{h}_0 >) & = & \mathfrak{Re}(< \boldsymbol{h}_p, \boldsymbol{h}_0 >) - \mathfrak{Re}(< \boldsymbol{h}_m, \boldsymbol{h}_0 >) = 0 \ \textit{car (P1)} \\ \mathfrak{Re}(< \boldsymbol{h}_p - \boldsymbol{h}_m, \boldsymbol{h}_s >) & = & \mathfrak{Re}(< \boldsymbol{h}_p - \boldsymbol{h}_m, \boldsymbol{h}_p - \boldsymbol{h}_m >) \\ & = & \|\boldsymbol{h}_p\|^2 - \|\boldsymbol{h}_m\|^2 + \mathfrak{Re}(< \boldsymbol{h}_p, \boldsymbol{h}_m > - < \boldsymbol{h}_m, \boldsymbol{h}_p >) \\ & = & \mathfrak{Re}(< \boldsymbol{h}_p, \boldsymbol{h}_m > - < \boldsymbol{h}_m, \boldsymbol{h}_p >) \ \textit{car (H1)} \\ & = & \mathfrak{Re}(< \boldsymbol{h}_p, \boldsymbol{h}_m > -(< \boldsymbol{h}_p, \boldsymbol{h}_m >)^*) = 0 \end{array}$$

*Enfin :*

$$\begin{array}{rcl} \mathfrak{Re}(< \boldsymbol{c}, \boldsymbol{v} >) & = & \mathfrak{Re}(< \frac{\alpha}{2}\boldsymbol{h}_s + (1-\alpha)\boldsymbol{h}_0, \frac{\beta}{r}(\boldsymbol{h}_p - \boldsymbol{h}_m) >) \\ & = & \frac{\alpha\beta}{2r}\mathfrak{Re}(< \boldsymbol{h}_s, (\boldsymbol{h}_p - \boldsymbol{h}_m) >) + \frac{(1-\alpha)\beta}{r}\mathfrak{Re}(< \boldsymbol{h}_0, (\boldsymbol{h}_p - \boldsymbol{h}_m) >) \\ & = & 0 \\ \mathfrak{Re}(< \boldsymbol{u}, \boldsymbol{v} >) & = & \mathfrak{Re}(< \frac{\alpha}{2r}(2\boldsymbol{h}_0 - \boldsymbol{h}_s), \frac{\beta}{r}(\boldsymbol{h}_p - \boldsymbol{h}_m) >) \\ & = & \frac{\alpha\beta}{r^2}\mathfrak{Re}(< \boldsymbol{h}_0, \boldsymbol{h}_p - \boldsymbol{h}_m >) - \frac{\alpha\beta}{2r^2}\mathfrak{Re}(< \boldsymbol{h}_s, \boldsymbol{h}_p - \boldsymbol{h}_m >) \\ & = & 0 \end{array}$$

*P3. Calcul de* $\|f(\varphi)\|_2^2$.

$$\begin{aligned}\|f(\varphi)\|_2^2 &= \|\boldsymbol{c} + r\cos\varphi\boldsymbol{u} + r\sin\varphi\boldsymbol{v}\|_2^2 \\ &= \|\boldsymbol{c}\|_2^2 + r^2\cos^2\varphi\|\boldsymbol{u}\|_2^2 + r^2\sin^2\varphi\|\boldsymbol{v}\|_2^2 + 2r\cos\varphi\mathfrak{Re}(<\boldsymbol{c},\boldsymbol{u}>) \ \dots \\ &\quad \dots + 2r\sin\varphi\mathfrak{Re}(<\boldsymbol{c},\boldsymbol{v}>) + 2r^2\cos\varphi\sin\varphi\mathfrak{Re}(<\boldsymbol{u},\boldsymbol{v}>) \\ &= \|\boldsymbol{c}\|_2^2 + r^2\cos^2\varphi\|\boldsymbol{u}\|_2^2 + r^2\sin^2\varphi\|\boldsymbol{v}\|_2^2 + 2r\cos\varphi\mathfrak{Re}(<\boldsymbol{c},\boldsymbol{u}>) \ \textit{car (P2)}\end{aligned}$$

*P4.* $\forall\varphi,\ \|f(\varphi)\|_2^2 = R^2 \Leftrightarrow \mathfrak{Re}(<\boldsymbol{c},\boldsymbol{u}>) = 0$ *et* $\|\boldsymbol{u}\|_2 = \|\boldsymbol{v}\|_2$

$$\begin{aligned}\forall\varphi,\ \|f(\varphi)\|_2^2 = R^2 &\Leftrightarrow \forall\varphi,\ \tfrac{\partial\|f(\varphi)\|^2}{\partial\varphi} = 0 \\ &\Leftrightarrow \forall\varphi,\ 2r^2\cos\varphi\sin\varphi(\|\boldsymbol{v}\|_2^2 - \|\boldsymbol{u}\|_2^2) - 2r\sin\varphi\mathfrak{Re}(<\boldsymbol{c},\boldsymbol{u}>) = 0 \\ &\Leftrightarrow \forall\varphi\neq 0,\ r\cos\varphi(\|\boldsymbol{v}\|_2^2 - \|\boldsymbol{u}\|_2^2) - \mathfrak{Re}(<\boldsymbol{c},\boldsymbol{u}>) = 0 \\ &\Leftrightarrow \mathfrak{Re}(<\boldsymbol{c},\boldsymbol{u}>) = 0 \ \textit{et}\ \|\boldsymbol{u}\|_2 = \|\boldsymbol{v}\|_2\end{aligned}$$

*P5. Calcul de* $(r,\theta)$ *en fonction de* $\rho = \|\boldsymbol{u}\|_2$.

$$\begin{aligned}\textit{Rappel : } \boldsymbol{h}(\varphi) &= \boldsymbol{c} + r\cos\varphi\boldsymbol{u} + r\sin\varphi\boldsymbol{v} \\ \textit{Alors : } \mathfrak{Re}(<\boldsymbol{h}_0 - \boldsymbol{h}_p, \boldsymbol{h}_m - \boldsymbol{h}_p>) &= \mathfrak{Re}(<r(1-\cos\theta)\boldsymbol{u} - r\sin\theta\boldsymbol{v}, -2r\sin\theta\boldsymbol{v}>) \\ &\quad \textit{(par définition)} \\ &= 2r^2\rho^2\sin^2\theta \ \textit{(car (P2))} \\ \|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2^2 &= \|r(1-\cos\theta)\boldsymbol{u} - r\sin\theta\boldsymbol{v}\|_2^2 \ \textit{(par définition)} \\ &= \rho^2r^2(1-\cos\theta)^2 + \rho^2r^2\sin^2\theta + 2\gamma\mathfrak{Re}(<\boldsymbol{u},\boldsymbol{v}>) \\ &= \rho^2r^2(1 - 2\cos\theta + \cos^2\theta + \sin^2\theta) \\ &= 2\rho^2r^2(1-\cos\theta) \\ \|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2^2 &= \|-2r\sin\theta\boldsymbol{v}\|_2^2 = 4r^2\rho^2\sin^2\theta \\ \frac{\mathfrak{Re}(<\boldsymbol{h}_0 - \boldsymbol{h}_p, \boldsymbol{h}_m - \boldsymbol{h}_p>)}{\|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2\|\boldsymbol{h}_m - \boldsymbol{h}_p\|_2} &= \frac{2r^2\rho^2\sin^2\theta}{2\sin\theta\rho^2r^2\sqrt{2(1-\cos\theta)}} = \frac{\sin\theta}{\sqrt{2(1-\cos\theta)}} \\ &= \frac{\sin(2\theta_0)}{\sqrt{2(1-\cos(2\theta_0))}} \quad (\textit{avec } \theta = 2\theta_0) \\ &= \frac{2\sin(\theta_0)\cos(\theta_0)}{\sqrt{2(2\sin^2(\theta_0))}} = \cos\theta_0 = \cos\frac{\theta}{2} \\ \textit{Donc,}\quad \theta &= 2\arccos\left(\frac{\mathfrak{Re}(<\boldsymbol{h}_0 - \boldsymbol{h}_p, \boldsymbol{h}_m - \boldsymbol{h}_p>)}{\|\boldsymbol{h}_0 - \boldsymbol{h}_p\|_2\|\boldsymbol{h}_m - \boldsymbol{h}_p\|_2}\right) \\ \textit{et}\quad r &= \frac{\|\boldsymbol{h}_0 - \boldsymbol{h}_p\|}{\rho\sqrt{2(1-\cos\theta)}}\end{aligned}$$

## A.2 Démonstration du théorème 2 p. 68 sur la K-SVD continue

Rappelons ici le théorème 2 de la SVD continue, présenté en §2.4.2.

**Théorème (K-SVD continue)** *Soit* $\mathbf{F}_P$ *la matrice de dimension* $N \times P$ *dont les colonnes sont constituées des vecteurs* $[\boldsymbol{h}(\delta_p)]_{p=1..P}$, *avec* $\delta_p \in \mathcal{I}_\Delta^P$ *et* $P > N$. *Soit* $\mathbf{U}_P\mathbf{S}_P\mathbf{V}_P^T = \mathbf{F}_P$ *sa décomposition en valeurs singulières compacte (SVDc), où les matrices* $\mathbf{U}_P$, $\mathbf{S}_P$ *et* $\mathbf{V}_P$ *sont respectivement de dimension* $N \times N$, $N \times R$ *et* $R \times P$, *où* $R$ *est le nombre de valeurs singulières non nulles de* $\mathbf{F}_P$ *et où les valeurs singulières sont rangées par ordre décroissant dans* $\mathbf{S}_P$.
*Alors, on a les résultats suivants :*

1. $\lim_{P\to\infty}\mathbf{U}_P = \mathbf{U}_\infty$, *où* $\mathbf{U}_\infty$ *vient de la SVD de la matrice* $\mathbf{B}_\infty = \mathbf{U}_\infty\mathbf{S}_\infty\mathbf{U}_\infty^T$, *telle que* $\mathbf{B}_\infty = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \boldsymbol{h}(\delta)\,\boldsymbol{h}(\delta)^T\,d\delta$.
2. $\boldsymbol{h}(\delta) \approx \mathbf{U}_{K\infty}\Psi(\delta)$ *est l'approximation limite quand* $P \to \infty$ *de* $\boldsymbol{h}(\delta)$ *via K-SVD, où* $\mathbf{U}_{K\infty}$ *contient les* $K$ *premières colonnes de* $\mathbf{U}_\infty$ *et* $\Psi(\delta) = \mathbf{U}_{K\infty}^T\boldsymbol{h}(\delta)$.

**Démonstration.** Considérons la matrice $\mathbf{A}_P = \mathbf{F}_P\mathbf{F}_P^T = \mathbf{U}_P\mathbf{S}_P^2\mathbf{U}_P^T$, de dimension $N \times N$, et la matrice $\mathbf{B}_P = \frac{\Delta}{P}\mathbf{A}_P$. Un élément $B_P(m,n)$ en ligne $m$ et colonne $n$ de cette matrice s'écrit :

$$\begin{aligned} B_P(m,n) &= \frac{\Delta}{P}\sum_{p=1}^{P}\mathbf{F}_P(m,p)\mathbf{F}_P^T(p,n) = \frac{\Delta}{P}\sum_{p=1}^{P}\mathbf{F}_P(m,p)\mathbf{F}_P(n,p) \\ &= \frac{\Delta}{P}\sum_{p=1}^{P} h_m(\delta_p)\, h_n(\delta_p), \quad \text{avec } h_i(\cdot) \text{ le i}^{\text{e}} \text{ élement de } \boldsymbol{h}(\cdot) \\ &= \frac{\Delta}{P}\sum_{p=1}^{P} h_{m\cdot n}(\delta_p) \qquad (\text{on note } h_{m\cdot n}(\delta_p) = h_m(\delta_p)\, h_n(\delta_p)) \\ &= \frac{\Delta}{P}\sum_{p=1}^{P} h_{m\cdot n}(-\frac{\Delta}{2} + p\frac{\Delta}{P}) \quad \text{car } \delta_p \in \mathcal{I}_\Delta^P \end{aligned}$$

Quand $P \to \infty$, on reconnaît la définition de l'intégrale de Riemann :

$$\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h_{m\cdot n}(\delta)\mathrm{d}\delta = \lim_{P\to\infty}\frac{\Delta}{P}\sum_{p=1}^{P} h_{m\cdot n}(-\frac{\Delta}{2} + p\frac{\Delta}{P})$$

Alors,

$$\mathbf{B}_P \underset{P\to\infty}{\longrightarrow} \mathbf{B}_\infty, \text{ avec } \mathbf{B}_\infty(m,n) = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h_{m\cdot n}(\delta)\mathrm{d}\delta = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h_m(\delta)h_n(\delta)\ \mathrm{d}\delta$$

On considère alors la SVD de cette matrice $\mathbf{B}_\infty$ :

$$\mathbf{B}_\infty = \mathbf{U}_\infty\mathbf{S}_\infty\mathbf{U}_\infty^T$$

où les matrices sont de tailles $N \times N$ et $\mathbf{U}_\infty$ unitaire. À partir de la SVD compacte de la matrice $\mathbf{A}_P$, on peut aussi écrire une SVD de la matrice $\mathbf{B}_P$ :

$$\mathbf{B}_P = \mathbf{U}_P\mathbf{S}'_P\mathbf{U}_P^T$$

avec $\mathbf{S}'_P = \frac{\Delta}{P}\mathbf{S}_P^2$. Or, les valeurs singulières de $\mathbf{B}_P$ étant toutes strictement positives (car on a pris la SVD compacte de $\mathbf{F}$) et rangée dans un ordre précis (décroissant), alors cette SVD de $\mathbf{B}_P$ est unique[1], à une multiplication d'une colonne de $\mathbf{U}_P$ près par (-1) . Par unicité de la décomposition et comme $\mathbf{B}_P \to \mathbf{B}_\infty$ quand $P \to \infty$, alors on peut conclure que $\mathbf{U}_P \to \mathbf{U}_\infty$ à une multiplication des colonnes de $\mathbf{U}_\infty$ près par (-1).
En revenant à la SVDc de départ : $\mathbf{F}_P = \mathbf{U}_P\mathbf{S}_P\mathbf{V}_P^T$, on peut écrire :

$$\boldsymbol{h}(\delta_p) = \mathbf{F}_P(:,p) = \mathbf{U}_P\mathbf{S}_P\boldsymbol{V}_P^T(:,p) = \mathbf{U}_P\Psi_N(\delta_p)$$

avec $\Psi_N(\delta_p) \in \mathbb{R}^N$ ici. $\mathbf{U}_P$ étant unitaire et donc inversible, on peut alors écrire l'interpolateur en dimension $N$ :

$$\Psi_N(\delta_p) = \mathbf{U}_P^T\boldsymbol{h}(\delta_p)$$

En faisant tendre $P \to \infty$, alors $\delta_p \to \delta$ et par continuité on obtient l'interpolateur continu sur $\mathbb{R}^N$ :

$$\Psi_N(\delta) = \mathbf{U}_\infty^T\boldsymbol{h}(\delta) \in \mathbb{R}^N$$

Finalement, on peut en déduire l'approximation par K-SVD continue de $\boldsymbol{h}(\delta)$, en tronquant $\mathbf{U}_\infty$ en ne conservant que ses $K$ premières colonnes : $\mathbf{U}_{K\infty} = \mathbf{U}_\infty(:,1:K)$ :

$$\begin{gathered} \Psi(\delta) = \Psi_K(\delta) = \mathbf{U}_{K\infty}^T\boldsymbol{h}(\delta) \in \mathbb{R}^K \\ \boldsymbol{h}(\delta) \approx \mathbf{U}_{K\infty}\Psi(\delta) \end{gathered}$$

[1] On supposera que les valeurs singulières de $\mathbf{B}_P$ ne sont pas dégénérées.

## A.3 Expressions analytiques de $\mathbf{B}_\infty$ pour l'approximation par K-SVD continue

Nous donnons dans cette annexe les expressions des matrices $\mathbf{B}_\infty$ pour l'approximation par K-SVD continue dans le cas de l'analyse spectrale. Pour rappel, nous nous plaçons à un indice $j$ quelconque, que nous ne précisons pas (notamment, $\ddot{\nu} = \ddot{\nu}_j$). On définit les fonctions suivantes :

$$f_1(u,v) = \frac{\sin(\frac{\Delta}{2}\,u)\cos(\ddot{\nu}\,u+v)}{u} \quad \text{et} \quad f_0(u) = f_1(u,0) \tag{A.1}$$

qui nous permettront d'écrire simplement les résultats des intégrations.

### A.3.1 Analyse spectrale - modèle *cosinus et sinus*

Dans le cas du modèle *cosinus et sinus*, il est nécessaire de calculer les matrices $\mathbf{B}^c_\infty$ et $\mathbf{B}^s_\infty$, définies ligne $m$ colonne $n$ de la façon suivante :

$$\begin{aligned}
\mathbf{B}^c_\infty(m,n) &= \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h^c_m(\delta)h^c_n(\delta)\,\mathrm{d}\delta = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \cos(2\pi t_m(\ddot{\nu}+\delta))\cos(2\pi t_n(\ddot{\nu}+\delta))\,\mathrm{d}\delta \\
\mathbf{B}^s_\infty(m,n) &= \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h^s_m(\delta)h^s_n(\delta)\,\mathrm{d}\delta = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \sin(2\pi t_m(\ddot{\nu}+\delta))\sin(2\pi t_n(\ddot{\nu}+\delta))\,\mathrm{d}\delta
\end{aligned}$$

Ces intégrales se calculent analytiquement grâce aux formules de trigonométrie :

$$\begin{aligned}
\mathbf{B}^c_\infty(m,n) &= \begin{cases} f_0(t_m-t_n)+f_0(t_m+t_n) & \text{si} \quad t_m \neq t_n \\ \dfrac{\Delta}{2}+f_0(t_m+t_n) & \text{si} \quad t_m = t_n \end{cases} \\
\mathbf{B}^s_\infty(m,n) &= \begin{cases} f_0(t_m-t_n)-f_0(t_m+t_n) & \text{si} \quad t_m \neq t_n \\ \dfrac{\Delta}{2}-f_0(t_m+t_n) & \text{si} \quad t_m = t_n \end{cases}
\end{aligned}$$

### A.3.2 Analyse spectrale - modèle *cosinus+phase*

Dans le cas du modèle *cosinus+phase*, il est nécessaire de calculer les matrices $\mathbf{B}_\infty$ définies ligne $m$ colonne $n$ de la façon suivante :

$$\begin{aligned}
\mathbf{B}_\infty(m,n) &= \int_{\mathcal{I}_\theta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h_m(\delta,\phi)h_n(\delta,\phi)\,\mathrm{d}\delta\,\mathrm{d}\phi \\
&= \int_{\mathcal{I}_\theta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \cos(2\pi t_m(\ddot{\nu}+\delta)+\phi)\,\cos(2\pi t_n(\ddot{\nu}+\delta)+\phi)\,\mathrm{d}\delta\,\mathrm{d}\phi
\end{aligned}$$

où $\mathcal{I}_\theta$ est un intervalle de largeur $\theta = \pi$ ou $\theta = 2\pi$ au choix, que l'on note $\mathcal{I}_\theta = [\theta_1;\theta_2]$ ($\theta = \theta_2 - \theta_1$). Cette intégrale se calcule via les formules de trigonométrie :

$$\mathbf{B}_\infty(m,n) = \begin{cases} \theta\cdot f_0(t_m-t_n)+\sin\theta\cdot f_1(t_m+t_n,\theta_1+\theta_2) & \text{si} \quad t_m \neq t_n \\ \theta\cdot\Delta+\sin\theta\cdot f_1(t_m+t_n,\theta_1+\theta_2) & \text{si} \quad t_m = t_n \end{cases}$$

# Annexe B

# Annexe du chapitre 3

Dans cette annexe, nous donnons les démonstrations de résultats donnés en Chapitre 3 sur le modèle BGE.

## B.1 Rappel de quelques théorèmes sur les matrices

**Lemme d'inversion matricielle.** Soit $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ des matrices de dimensions telles que les multiplications matricielles considérées sont possibles.

$$\begin{array}{l}\text{Si } \mathbf{A} \text{ et } \mathbf{C} \text{ sont inversibles, alors :} \\ (\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{D}) \text{ inversible } \Leftrightarrow\ (\mathbf{C}-\mathbf{D}\mathbf{A}^{-1}\mathbf{B}) \text{ inversible} \\ (\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{D})^{-1} = \mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}(\mathbf{C}-\mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{D}\mathbf{A}^{-1}\end{array} \tag{B.1}$$

**Lemme d'inversion matricielle par blocs.** Soit $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ des matrices de dimensions telles que les multiplications matricielles considérées sont possibles.

$$\begin{array}{l}\text{Si } \mathbf{A} \text{ et } \mathbf{C} \text{ sont inversibles, alors :} \\ \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{D} & \mathbf{C}\end{pmatrix}^{-1} = \begin{pmatrix}\mathbf{A}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} + \begin{pmatrix}-\mathbf{A}^{-1}\mathbf{B}\\ \mathbf{I}\end{pmatrix}\left(\mathbf{C}-\mathbf{D}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\begin{pmatrix}-\mathbf{D}\mathbf{A}^{-1} & \mathbf{I}\end{pmatrix}\end{array} \tag{B.2}$$

**Généralisation du lemme du déterminant matriciel.** Soit $\mathbf{A}$ une matrice carrée inversible, et $\mathbf{U}, \mathbf{V}$ des matrices possiblement rectangulaires, de dimensions telles que les opérations considérées sont possibles. Alors :

$$|\mathbf{A}+\mathbf{U}\mathbf{V}^T| = |\mathbf{A}| \cdot |\mathbf{I}+\mathbf{V}^T\mathbf{A}^{-1}\mathbf{U}| \tag{B.3}$$

où $|\cdot|$ représente ici le déterminant de la matrice considérée.

## B.2 Calcul de la loi marginale

Pour le calcul de la loi marginale, nous passons par celui de la loi conditionnelle en les amplitudes. Nous nous plaçons ici dans un cadre général où l'on peut considérer le modèle BG étendu ou non : les matrices $\mathbf{H}$, $\overline{\mathbf{H}}$ et $\mathbf{h}$ peuvent ou non dépendre de paramètres non linéaires. Nous donnons ici la loi *a posteriori* jointe qui nous servira pour le calcul de la loi marginale :

$$p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y}) \propto \underbrace{p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\Sigma})\; p(\overline{\boldsymbol{x}} \,|\, \overline{\boldsymbol{q}}, \sigma_x^2)}_{f_1(\overline{\boldsymbol{x}})}\; \underbrace{p(\overline{\boldsymbol{x}}^C \,|\, \overline{\boldsymbol{q}}^C)}_{f_2(\overline{\boldsymbol{x}}^C)}\; \Pr(\boldsymbol{q} \,|\, \lambda)\; p(\boldsymbol{\nu})\; p(\boldsymbol{\theta}) \tag{B.4}$$

Notons que la fonction $f(\boldsymbol{x}) = f_1(\overline{\boldsymbol{x}})f_2(\overline{\boldsymbol{x}}^C)$ répertorie l'ensemble des facteurs qui dépendent de $\boldsymbol{x}$. Le calcul de la loi marginale se fait donc par intégration de la fonction $f(\boldsymbol{x})$ :

$$p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y}) = \int_{\boldsymbol{x}} p(\boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \propto \Pr(\boldsymbol{q} \,|\, \lambda) \; p(\boldsymbol{\nu}) \; p(\boldsymbol{\theta}) \int_{\boldsymbol{x}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \tag{B.5}$$

**Calcul de $p(\overline{\boldsymbol{x}} \,|\, -)$.** À partir de l'équation (B.4), on déduit que $p(\overline{\boldsymbol{x}} \,|\, -)$ est proportionnelle à $p(\boldsymbol{y} \,|\, \boldsymbol{q}, \boldsymbol{\nu}, \boldsymbol{x}, \boldsymbol{\Sigma}) \; p(\overline{\boldsymbol{x}} \,|\, \boldsymbol{q}, \sigma_x^2)$. Grâce aux expressions de la vraisemblance (3.3) et des lois *a priori* en $\boldsymbol{x}$ (3.2), on en déduit la relation suivante :

$$\begin{aligned} p(\overline{\boldsymbol{x}} \,|\, -) \propto f_3(\overline{\boldsymbol{x}}) \quad \text{avec } f_3(\overline{\boldsymbol{x}}) \quad &= \quad \exp(-\frac{1}{2}\|\boldsymbol{y} - \overline{\mathbf{H}}\overline{\boldsymbol{x}}\|^2_{\boldsymbol{\Sigma}^{-1}}) \exp(-\frac{1}{2\sigma_x^2}\|\overline{\boldsymbol{x}}\|^2) \\ &= \quad \exp\left(-\frac{1}{2}(\|\boldsymbol{y} - \overline{\mathbf{H}}\overline{\boldsymbol{x}}\|^2_{\boldsymbol{\Sigma}^{-1}} + \frac{\|\overline{\boldsymbol{x}}\|^2}{\sigma_x^2})\right) \\ &\propto \quad f_1(\overline{\boldsymbol{x}}) \quad \text{(les constantes de normalisation n'apparaissent pas.)} \end{aligned}$$

On cherche à simplifier le terme entre parenthèses :

$$\begin{aligned} \|\boldsymbol{y} - \overline{\mathbf{H}}\overline{\boldsymbol{x}}\|^2_{\boldsymbol{\Sigma}^{-1}} + \frac{\|\overline{\boldsymbol{x}}\|^2}{\sigma_x^2} \quad &= \quad \|\boldsymbol{y}\|^2_{\boldsymbol{\Sigma}^{-1}} - 2\boldsymbol{y}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}\overline{\boldsymbol{x}} + \overline{\boldsymbol{x}}^T\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}\overline{\boldsymbol{x}} + \frac{\overline{\boldsymbol{x}}^T\overline{\boldsymbol{x}}}{\sigma_x^2} \\ &= \quad \overline{\boldsymbol{x}}^T(\sigma_x^{-2}\mathbf{I}_{mL} + \overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}})\overline{\boldsymbol{x}} - 2\boldsymbol{y}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}\overline{\boldsymbol{x}} + \|\boldsymbol{y}\|^2_{\boldsymbol{\Sigma}^{-1}} \\ &= \quad \|\overline{\boldsymbol{x}} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Gamma}^{-1}} + \rho \end{aligned}$$

Avec :

$$\boxed{\begin{aligned} \boldsymbol{\Gamma}^{-1} &= \sigma_x^{-2}\mathbf{I}_{mL} + \overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}} \quad (\text{dim. } mL \times mL) \\ \boldsymbol{\mu} &= \boldsymbol{\Gamma}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \quad (\text{dim. } mL \times 1) \\ \rho &= \|\boldsymbol{y}\|^2_{\boldsymbol{\Sigma}^{-1}} - \|\boldsymbol{\mu}\|^2_{\boldsymbol{\Gamma}^{-1}} \quad (\text{dim. } 1) \end{aligned}} \tag{B.6}$$

On peut donc revenir à la loi conditionnelle :

$$\begin{aligned} p(\overline{\boldsymbol{x}} \,|\, -) \quad &\propto \quad \exp\left(-\frac{1}{2}(\|\overline{\boldsymbol{x}} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Gamma}^{-1}} + \rho)\right) \\ &\propto \quad \exp\left(-\frac{1}{2}(\|\overline{\boldsymbol{x}} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Gamma}^{-1}})\right) \end{aligned}$$

Et on peut donc en déduire que la loi conditionnelle *a posteriori* en $\overline{\boldsymbol{x}}$ est une loi gaussienne de moyenne $\boldsymbol{\mu}$ et de covariance $\boldsymbol{\Gamma}$.

$$\boxed{(\overline{\boldsymbol{x}} \,|\, -) \quad \sim \quad \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})}$$

**Propriétés de la covariance $\boldsymbol{\Gamma}$.**

Inversible : $\boldsymbol{\Gamma}^{-1}$ est la somme d'une matrice définie-positive ($\sigma_x^{-2}\mathbf{I}_{mL}$) et d'une matrice positive ($\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}$) ; elle est donc définie-positive, et inversible d'inverse $\boldsymbol{\Gamma}$.

Symétrique : $\boldsymbol{\Gamma}^{-1}$ est symétrique donc $\boldsymbol{\Gamma}$ l'est aussi.

**Simplification de $\rho$ et définition de B.** En développant l'expression de $\rho$ et $\boldsymbol{\mu}$ dans (B.6), on obtient :

$$\begin{aligned} \rho \quad &= \quad \|\boldsymbol{y}\|^2_{\boldsymbol{\Sigma}^{-1}} - \|\boldsymbol{\mu}\|^2_{\boldsymbol{\Gamma}^{-1}} \\ &= \quad \boldsymbol{y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - (\boldsymbol{\Gamma}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y})^T\boldsymbol{\Gamma}^{-1}(\boldsymbol{\Gamma}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}) \\ &= \quad \boldsymbol{y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - \boldsymbol{y}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}\boldsymbol{\Gamma}^T\boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \quad \text{(par symétrie de } \boldsymbol{\Sigma}^{-1} \text{ et } \boldsymbol{\Gamma}) \\ &= \quad \boldsymbol{y}^T(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}\boldsymbol{\Gamma}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1})\boldsymbol{y} \end{aligned}$$

On obtient :

$$\boxed{\begin{array}{rcl}\rho & = & \boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y} \\ \mathbf{B}^{-1} & = & \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}\boldsymbol{\Gamma}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1} \quad (\text{dim. } N \times N)\end{array}}$$

On utilise ensuite le lemme d'inversion matricielle (B.1) sur $\mathbf{B}^{-1}$ afin d'obtenir $\mathbf{B}$ :

$$\begin{array}{rcl}\mathbf{B} & = & (\mathbf{B}^{-1})^{-1} = (\underbrace{\boldsymbol{\Sigma}^{-1}}_{\mathbf{a}} - \underbrace{\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}}_{\mathbf{b}}\underbrace{\boldsymbol{\Gamma}}_{\mathbf{c}^{-1}}\underbrace{\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}}_{\mathbf{d}})^{-1} \quad (\mathbf{a},\mathbf{c}) \text{ inversibles} \\ & = & \mathbf{a}^{-1} + \mathbf{a}^{-1}\mathbf{b}(\mathbf{c} - \mathbf{d}\mathbf{a}^{-1}\mathbf{b})^{-1}\mathbf{d}\mathbf{a}^{-1} \\ & = & \boldsymbol{\Sigma} + \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}(\boldsymbol{\Gamma}^{-1} - \overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}})^{-1}\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma} \\ & = & \boldsymbol{\Sigma} + \overline{\mathbf{H}}(\boldsymbol{\Gamma}^{-1} - \overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}})^{-1}\overline{\mathbf{H}}^T \quad \text{or} \quad \boldsymbol{\Gamma}^{-1} = \sigma_x^{-2}\mathbf{I}_{mL} + \overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}} \quad \text{d'après } (B.6) \\ & = & \boldsymbol{\Sigma} + \overline{\mathbf{H}}(\sigma_x^{-2}\mathbf{I}_{mL})^{-1}\overline{\mathbf{H}}^T = \boldsymbol{\Sigma} + \sigma_x^2\overline{\mathbf{H}}\overline{\mathbf{H}}^T\end{array}$$

$$\boxed{\mathbf{B} = \boldsymbol{\Sigma} + \sigma_x^2\overline{\mathbf{H}}\overline{\mathbf{H}}^T \quad (\text{dim. } N \times N)} \tag{B.7}$$

**Déterminant de B.** En utilisant la généralisation du déterminant matriciel (B.3) :

$$\begin{array}{rcl}|\mathbf{B}| & = & |\boldsymbol{\Sigma} + \sigma_x^2\overline{\mathbf{H}}\overline{\mathbf{H}}^T| = |\underbrace{\boldsymbol{\Sigma}}_{\mathbf{a}} + \underbrace{(\sigma_x\overline{\mathbf{H}})}_{\mathbf{u}}\underbrace{(\sigma_x\overline{\mathbf{H}})^T}_{\mathbf{v}^T}| \\ & = & |\mathbf{a}| \cdot |\mathbf{I} + \mathbf{v}^T\mathbf{a}^{-1}\mathbf{u}| \\ & = & |\boldsymbol{\Sigma}| \cdot |\mathbf{I} + \sigma_x^2\overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}}| = |\boldsymbol{\Sigma}| \cdot |(\sigma_x^2)\underbrace{(\sigma_x^{-2}\mathbf{I} + \overline{\mathbf{H}}^T\boldsymbol{\Sigma}^{-1}\overline{\mathbf{H}})}_{\boldsymbol{\Gamma}^{-1}(\text{dim. } mL \times mL)}| \\ & = & (\sigma_x^2)^{mL}|\boldsymbol{\Sigma}||\boldsymbol{\Gamma}^{-1}|\end{array}$$

$$\boxed{|\mathbf{B}| = \sigma_x^{2mL}|\boldsymbol{\Sigma}||\boldsymbol{\Gamma}^{-1}|} \tag{B.8}$$

**Expression de $f_3(\overline{\boldsymbol{x}})$.** On peut déduire des résultats précédents l'expression de $f_3(\overline{\boldsymbol{x}})$ en fonction de $\mathbf{B}$ :

$$\boxed{f_3(\overline{\boldsymbol{x}}) = \exp\left(-\frac{1}{2}\|\overline{\boldsymbol{x}} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Gamma}^{-1}}\right)\exp\left(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y}\right)} \tag{B.9}$$

**Marginalisation.** Nous avons vu en équation (B.5) que la marginalisation de la loi *a posteriori* se faisait en intégrant la fonction $f(\boldsymbol{x}) = f_1(\overline{\boldsymbol{x}})f_2(\overline{\boldsymbol{x}}^C)$ par rapport à $\boldsymbol{x}$. Nous allons utiliser les résultats précédents pour réaliser cette intégration.

$$\begin{array}{rcl}\text{On a} \quad f_1(\overline{\boldsymbol{x}}) & = & p(\boldsymbol{y}\,|\,\boldsymbol{q},\overline{\boldsymbol{x}},\boldsymbol{\nu},\boldsymbol{\Sigma})\,p(\overline{\boldsymbol{x}}\,|\,\overline{\boldsymbol{q}},\sigma_x^2) = \alpha_3 f_3(\overline{\boldsymbol{x}}) \\ \text{avec} \quad \alpha_3 & = & (2\pi)^{-N/2}|\boldsymbol{\Sigma}|^{-1/2}(2\pi)^{-mL/2}\sigma_x^{-mL} \quad \text{constante de normalisation} \\ \text{et} \quad f_2(\overline{\boldsymbol{x}}) & = & p(\overline{\boldsymbol{x}}^C\,|\,\overline{\boldsymbol{q}}^C) = \delta_{\mathbf{0}}(\overline{\boldsymbol{x}}^C) \\ \text{Alors} \quad \int_{\boldsymbol{x}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} & = & \int_{\overline{\boldsymbol{x}}}\int_{\overline{\boldsymbol{x}}^C} f_1(\overline{\boldsymbol{x}})f_2(\overline{\boldsymbol{x}}^C)\,\mathrm{d}\overline{\boldsymbol{x}}\,\mathrm{d}\overline{\boldsymbol{x}}^C = \int_{\overline{\boldsymbol{x}}} f_1(\overline{\boldsymbol{x}})\,\mathrm{d}\overline{\boldsymbol{x}}\int_{\overline{\boldsymbol{x}}^C} f_2(\overline{\boldsymbol{x}}^C)\,\mathrm{d}\overline{\boldsymbol{x}}^C \\ \int_{\overline{\boldsymbol{x}}^C} f_2(\overline{\boldsymbol{x}}^C)\,\mathrm{d}\overline{\boldsymbol{x}}^C & = & \int_{\overline{\boldsymbol{x}}^C} \delta_{\mathbf{0}}(\overline{\boldsymbol{x}}^C)\,\mathrm{d}\overline{\boldsymbol{x}}^C = 1 \quad \text{(propriété du Dirac)} \\ \int_{\overline{\boldsymbol{x}}} f_1(\overline{\boldsymbol{x}})\,\mathrm{d}\overline{\boldsymbol{x}} & = & \alpha_3\int_{\overline{\boldsymbol{x}}} f_3(\overline{\boldsymbol{x}})\,\mathrm{d}\overline{\boldsymbol{x}} \\ & = & \alpha_3\exp(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y})\int_{\overline{\boldsymbol{x}}}\exp(-\frac{1}{2}\|\overline{\boldsymbol{x}} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Gamma}^{-1}})\,\mathrm{d}\overline{\boldsymbol{x}} \quad \text{(voir éq. (B.9))} \\ & = & \alpha_3\exp(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y})(2\pi)^{mL/2}|\boldsymbol{\Gamma}|^{1/2} \\ & = & \exp(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y})(2\pi)^{mL/2}|\boldsymbol{\Gamma}|^{1/2}(2\pi)^{-N/2}|\boldsymbol{\Sigma}|^{-1/2}(2\pi)^{-mL/2}\sigma_x^{-mL} \\ & = & (2\pi)^{-N/2}\exp(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y})|\mathbf{B}|^{-1/2} \quad \text{(d'après éq. (B.8))}\end{array}$$

Par conséquent, on obtient donc, à partir de (B.5) :

$$\boxed{p(\boldsymbol{q},\boldsymbol{\nu},\boldsymbol{\theta}\,|\,\boldsymbol{y}) \;\propto\; |\mathbf{B}|^{-1/2}\exp(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y})\,\Pr(\boldsymbol{q}\,|\,\lambda)\,p(\boldsymbol{\nu})\,p(\boldsymbol{\theta})} \tag{B.10}$$

## B.3 Lois *a posteriori* conditionnelles en $q_j$

Dans cette partie, on se place à un indice $j$ de la grille. On introduit la fonction suivante :

$$\Psi(\mathbf{A}_1,\mathbf{A}_2;\boldsymbol{y}) = \log\frac{|\mathbf{A}_1|}{|\mathbf{A}_2|} + \Big(\boldsymbol{y}^T(\mathbf{A}_1-\mathbf{A}_2)\boldsymbol{y}\Big)$$

À partir de l'expression de la loi *a posteriori* marginalisée (B.10) et de l'expression de $\Pr(\boldsymbol{q}\,|\,\lambda)$ donnée en (3.2), on peut extraire la loi *a posteriori* marginalisée conditionnelle en $q_j$ en ne conservant que les termes dépendant de ce dernier notamment $\mathbf{B}$ et $L=\sum_i q_i$ :

$$\begin{aligned}\Pr(q_j\,|\,-) &\propto |\mathbf{B}|^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y}\right)\,\lambda^{\sum_i q_i}(1-\lambda)^{-\sum_i q_i}\\ &= \alpha_1\,f_4(q_j) \quad \text{avec } f_4(q_j) = |\mathbf{B}_{(q_j)}|^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}_{(q_j)}^{-1}\boldsymbol{y}\right)\,\left(\frac{\lambda}{1-\lambda}\right)^{q_j}\end{aligned}$$

où $\mathbf{B}_{(0)}$ (resp. $\mathbf{B}_{(1)}$) correspond à la matrice $\mathbf{B}$ avec $q_j=0$ (resp. $q_j=1$), et où $\alpha_1$ regroupe tous les termes ne dépendant pas de $q_j$.
Comme $\Pr(q_j\,|\,-)+\Pr(1-q_j\,|\,-)=1$, on peut alors calculer la constante de normalisation $\alpha_1$ :

$$\begin{aligned}1 &= \Pr(q_j\,|\,-)+\Pr(1-q_j\,|\,-)\\ &= \alpha_1\,f_4(q_j)+\alpha_1\,f_4(1-q_j)\\ \Rightarrow\ \alpha_1 &= (f_4(q_j)+f_4(1-q_j))^{-1}\\ \text{et donc}\quad \Pr(q_j\,|\,-) &= \frac{f_4(q_j)}{f_4(q_j)+f_4(1-q_j)} = \left(1+\frac{f_4(1-q_j)}{f_4(q_j)}\right)^{-1}\end{aligned}$$

À partir de là, on obtient alors facilement le résultat suivant faisant appel à la fonction $\Psi$ :

$$\boxed{\begin{aligned}\Pr(q_j\,|\,-) &= \Big(1+\exp(-(\Psi_{q_j}+\eta_j)/2)\Big)^{-1} \quad \text{avec} \quad \eta_j = 2(-1)^{q_j}\log\left(\frac{1-\lambda}{\lambda}\right)\\ &\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{et} \quad \Psi_{q_j} = \Psi(\mathbf{B}_{(1-q_j)}^{-1},\mathbf{B}_{(q_j)}^{-1};\boldsymbol{y})\end{aligned}} \tag{B.11}$$

## B.4 Simplification de $\Psi_{q_j}$ grâce aux matrices T et $\boldsymbol{\gamma}$

**Relation entre $\mathbf{B}_{(q_j)}$ et $\mathbf{B}_{(1-q_j)}$.** Les matrices $\mathbf{B}_{(q_j)}$ et $\mathbf{B}_{(1-q_j)}$ sont très similaires, puisqu'elles ne diffèrent qu'en la valeur de $q_j$. En effet, on peut écrire :

$$\begin{aligned}\mathbf{B} &= \boldsymbol{\Sigma}+\sigma_x^2\overline{\mathbf{H}\mathbf{H}}^T = \boldsymbol{\Sigma}+\sigma_x^2\sum_{\ell\,|\,q_\ell=1}\mathbf{h}_\ell\mathbf{h}_\ell^T\\ \text{donc}\quad \mathbf{B}_{(q_j)} &= \boldsymbol{\Sigma}+\sigma_x^2\sum_{\ell\neq j\,|\,q_\ell=1}\mathbf{h}_\ell\mathbf{h}_\ell^T + q_j\sigma_x^2\mathbf{h}_j\mathbf{h}_j^T\\ \mathbf{B}_{(1-q_j)} &= \boldsymbol{\Sigma}+\sigma_x^2\sum_{\ell\neq j\,|\,q_\ell=1}\mathbf{h}_\ell\mathbf{h}_\ell^T + (1-q_j)\sigma_x^2\mathbf{h}_j\mathbf{h}_j^T\\ \Rightarrow\quad \mathbf{B}_{(1-q_j)}-\mathbf{B}_{(q_j)} &= (-1)^{q_j}\sigma_x^2\mathbf{h}_j\mathbf{h}_j^T\end{aligned}$$

$$\boxed{\mathbf{B}_{(1-q_j)}-\mathbf{B}_{(q_j)} = (-1)^{q_j}\sigma_x^2\mathbf{h}_j\mathbf{h}_j^T} \tag{B.12}$$

Cette relation va nous servir à simplifier l'expression de $\Psi_{q_j}$.

**Différence des inverses $\mathbf{B}_{(q_j)}^{-1} - \mathbf{B}_{(1-q_j)}^{-1}$.** À partir de l'équation (B.12), on peut déduire la différence des inverses grâce au lemme d'inversion matricielle (B.1).

$$\begin{aligned}
\mathbf{B}_{(1-q_j)}^{-1} &= (\mathbf{B}_{(q_j)} + (-1)^{q_j}\sigma_x^2\mathbf{h}_j\mathbf{h}_j^T)^{-1} = (\underbrace{\mathbf{B}_{(q_j)}}_{\mathbf{a}} - \underbrace{\mathbf{h}_j}_{\mathbf{b}}\underbrace{\mathbf{D}}_{\mathbf{c}^{-1}}\underbrace{\mathbf{h}_j^T}_{\mathbf{d}})^{-1} \quad \text{avec} \quad \mathbf{D} = -(-1)^{q_j}\sigma_x^2\mathbf{I}_m \\
&= \mathbf{a}^{-1} + \mathbf{a}^{-1}\mathbf{b}(\mathbf{c} - \mathbf{d}\mathbf{a}^{-1}\mathbf{b})^{-1}\mathbf{d}\mathbf{a}^{-1} \\
&= \mathbf{B}_{(q_j)}^{-1} + \mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j(\mathbf{D}^{-1} - \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j)^{-1}\mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1} \quad \text{avec} \quad \mathbf{D}^{-1} = -(-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m \\
&= \mathbf{B}_{(q_j)}^{-1} - \mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j\underbrace{((-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j)}_{\mathbf{T}_{(q_j)}}{}^{-1}\mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}
\end{aligned}$$

On en déduit :

$$\begin{aligned}
\boldsymbol{y}^T(\mathbf{B}_{(1-q_j)}^{-1} - \mathbf{B}_{(q_j)}^{-1})\boldsymbol{y} &= \boldsymbol{y}^T(-\mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j\mathbf{T}_{(q_j)}^{-1}\mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1})\boldsymbol{y} \\
&= -\underbrace{(\mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\boldsymbol{y})^T}_{\boldsymbol{\gamma}_{(q_j)}^T}\mathbf{T}_{(q_j)}^{-1}\underbrace{(\mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\boldsymbol{y})}_{\boldsymbol{\gamma}_{(q_j)}}
\end{aligned}$$

On a donc :

$$\boxed{\boldsymbol{y}^T(\mathbf{B}_{(1-q_j)}^{-1} - \mathbf{B}_{(q_j)}^{-1})\boldsymbol{y} = -\boldsymbol{\gamma}_{(q_j)}^T\mathbf{T}_{(q_j)}^{-1}\boldsymbol{\gamma}_{(q_j)}}$$

avec

$$\boxed{\begin{aligned}
\mathbf{T}_{(q_j)} &= (-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j \quad (\text{dim. } m \times m) \\
\boldsymbol{\gamma}_{(q_j)} &= \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\boldsymbol{y} \quad (\text{dim. } m \times 1)
\end{aligned}} \tag{B.13}$$

**Rapport des déterminants de $\mathbf{B}_{(q_j)}$ et $\mathbf{B}_{(1-q_j)}$.** En utilisant le résultat (B.12), on peut calculer le déterminant de la matrice $\mathbf{B}_{(1-q_j)}$ en fonction de celui de $\mathbf{B}_{(q_j)}$ grâce à la généralisation du lemme du déterminant matriciel (B.3).

$$\begin{aligned}
|\mathbf{B}_{(1-q_j)}| &= |\underbrace{\mathbf{B}_{(q_j)}}_{\mathbf{a}} + \underbrace{(-1)^{q_j}\sigma_x^2\mathbf{h}_j}_{\boldsymbol{u}}\underbrace{\mathbf{h}_j^T}_{\mathbf{v}^T}| \\
&= |\mathbf{a}| \cdot |\mathbf{I} + \mathbf{v}^T\mathbf{a}^{-1}\mathbf{u}| \\
&= |\mathbf{B}_{(q_j)}| \cdot |\mathbf{I}_m + \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}((-1)^{q_j}\sigma_x^2\mathbf{h}_j)| \\
&= |\mathbf{B}_{(q_j)}| \cdot |(-1)^{q_j}\sigma_x^2\mathbf{I}_m||\underbrace{(-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j}_{=\mathbf{T}_{(q_j)} \text{ d'après (B.13)}}| \\
&= |\mathbf{B}_{(q_j)}||\mathbf{T}_{(q_j)}|((-1)^{q_j}\sigma_x^2)^m
\end{aligned}$$

$$\boxed{\frac{|\mathbf{B}_{(1-q_j)}|}{|\mathbf{B}_{(q_j)}|} = (-1)^{mq_j}\sigma_x^{2m}|\mathbf{T}_{(q_j)}|} \tag{B.14}$$

**Réécriture de $\Psi_{q_j}$.** À partir de (B.11) et des résultats précédents, on obtient donc :

$$\boxed{\Psi_{q_j} = \log\left((-1)^{mq_j}\sigma_x^{2m}|\mathbf{T}_{(q_j)}|\right) - \boldsymbol{\gamma}_{(q_j)}^T\mathbf{T}_{(q_j)}^{-1}\boldsymbol{\gamma}_{(q_j)}}$$

## B.5 Simplification des matrices T et $\gamma$ avec la factorisation de Cholesky F

**Définition de G, C et F.** En introduisant la matrice $\mathbf{G} = \sigma_x \overline{\mathbf{H}}$, on peut réécrire la matrice $\mathbf{B}$ :

$$\mathbf{B} = \boldsymbol{\Sigma} + \sigma_x^{-2}\overline{\mathbf{H}}\overline{\mathbf{H}}^T = \boldsymbol{\Sigma} + \mathbf{G}\mathbf{G}^T$$

On applique alors à $\mathbf{B}$ le lemme d'inversion matricielle (B.1) :

$$\begin{aligned}
\mathbf{B}^{-1} &= (\underbrace{\boldsymbol{\Sigma}}_{\mathbf{a}} + \underbrace{\mathbf{G}}_{\mathbf{b}}\underbrace{\mathbf{I}}_{-\mathbf{c}^{-1}}\underbrace{\mathbf{G}^T}_{\mathbf{d}})^{-1} \\
&= \mathbf{a}^{-1} + \mathbf{a}^{-1}\mathbf{b}(\mathbf{c} - \mathbf{d}\mathbf{a}^{-1}\mathbf{b})^{-1}\mathbf{d}\mathbf{a}^{-1} \\
&= \boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1}\mathbf{G}(-\mathbf{I} - \mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\boldsymbol{\Sigma}^{-1} \\
&= \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{G}\underbrace{(\mathbf{I} + \mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G})^{-1}}_{=\mathbf{C}^{-1}}\mathbf{G}^T\boldsymbol{\Sigma}^{-1}
\end{aligned}$$

On a donc :

$$\boxed{\begin{aligned}
\mathbf{G} &= \sigma_x\overline{\mathbf{H}} \\
\mathbf{B} &= \boldsymbol{\Sigma} + \mathbf{G}\mathbf{G}^T \\
\mathbf{B}^{-1} &= \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{G}\mathbf{C}^{-1}\mathbf{G}^T\boldsymbol{\Sigma}^{-1} \\
\mathbf{C} &= \mathbf{I}_{mL} + \mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G}
\end{aligned}} \tag{B.15}$$

La matrice $\mathbf{C}^{-1}$ étant définie positive ($\mathbf{C}$ est la somme de $\mathbf{I}_{mL}$ définie positive et de $\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G}$ semi-définie positive), on peut utiliser sa factorisation de Cholesky :

$$\boxed{\begin{array}{c}\mathbf{C}^{-1} = \mathbf{F}^T\mathbf{F} \\ \text{où } \mathbf{F} \text{ est triangulaire supérieure (dim. } mL \times mL)\end{array}} \tag{B.16}$$

**Simplication de $\mathbf{T}, \boldsymbol{\gamma}$.** Les matrices $\mathbf{T}$ et $\boldsymbol{\gamma}$ peuvent donc maintenant être exprimées à l'aide de la matrice triangulaire supérieure $\mathbf{F}$. Notons que la dépendance de $\mathbf{B}$ en $q_j$ et $\boldsymbol{\nu}_j$ se transmet à $\mathbf{G}$, $\mathbf{C}$ et donc $\mathbf{F}$ :

$$\begin{aligned}
\mathbf{T}_{(q_j)} &= (-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\mathbf{h}_j \\
&= (-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{G}_{(q_j)}\mathbf{C}_{(q_j)}^{-1}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1})\mathbf{h}_j \\
&= (-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j - \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\mathbf{G}_{(q_j)}\mathbf{F}_{(q_j)}^T\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j \\
&= (-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j - (\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j)^T(\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j) \\
\boldsymbol{\gamma}_{(q_j)} &= \mathbf{h}_j^T\mathbf{B}_{(q_j)}^{-1}\boldsymbol{y} \\
&= \mathbf{h}_j^T(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{G}_{(q_j)}\mathbf{C}_{(q_j)}^{-1}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1})\boldsymbol{y} \\
&= \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\mathbf{G}_{(q_j)}\mathbf{F}_{(q_j)}^T\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \\
&= \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - (\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j)^T(\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y})
\end{aligned}$$

$$\boxed{\begin{aligned}
\mathbf{T}_{(q_j)} &= (-1)^{q_j}\sigma_x^{-2}\mathbf{I}_m + \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j - (\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j)^T(\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j) \\
\boldsymbol{\gamma}_{(q_j)} &= \mathbf{h}_j^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} - (\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\mathbf{h}_j)^T(\mathbf{F}_{(q_j)}\mathbf{G}_{(q_j)}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y})
\end{aligned}} \tag{B.17}$$

## B.6 Mise à jour du facteur de Cholesky F

Ici, nous suivons les conventions de mises à jour de la matrice $\mathbf{G}$ définie en § 3.3.2 (p.119). Nous rappelons la notation $\widetilde{\mathbf{G}}_j$ définie en (3.25) correspondant à la matrice $\mathbf{G}_{(1)}$ telle que les

atomes $\sigma_x \mathbf{h}_j$ se trouvent en dernière position ; $\widetilde{\mathbf{F}}_j$ et $\widetilde{\mathbf{C}}_j$ s'expriment en fonction de $\widetilde{\mathbf{G}}_j$. Dans cette section, nous démontrons les divers résultats pour la mise à jour de la matrice de Cholesky $\mathbf{F}$ (à partir de la page 119).

**Relation de base pour les mises à jour de Cholesky (éq. 3.26).** À partir de la définition de $\mathbf{C}$ en éq. B.15, on peut écrire :

$$\begin{aligned} \mathbf{C}_{(0)} &= \mathbf{I}_{D_0} + \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{G}_{(0)} \text{ (dim. } D_0 \times D_0) \\ \widetilde{\mathbf{C}}_j &= \mathbf{I}_{D_1} + \widetilde{\mathbf{G}}_j^T \boldsymbol{\Sigma}^{-1} \widetilde{\mathbf{G}}_j \text{ (dim. } D_1 \times D_1) \end{aligned}$$

où $D_1 = D_0 + m$. En représentant la matrice $\widetilde{\mathbf{C}}_j$ par blocs et en utilisant la définition de $\mathbf{G}_{(1)} = \begin{bmatrix} \mathbf{G}_{(0)} & \sigma_x \mathbf{h}_j \end{bmatrix}$, on obtient :

$$\begin{aligned} \widetilde{\mathbf{C}}_j &= \mathbf{I}_{D_1} + \begin{bmatrix} \mathbf{G}_{(0)}^T \\ \sigma_x \mathbf{h}_j^T \end{bmatrix} \boldsymbol{\Sigma}^{-1} \begin{bmatrix} \mathbf{G}_{(0)} & \sigma_x \mathbf{h}_j \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I}_{D_0} + \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{G}_{(0)} & \sigma_x \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \\ \sigma_x \mathbf{h}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{G}_{(0)} & \mathbf{I}_{mL} + \sigma_x^2 \mathbf{h}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C}_{(0)} & \sigma_x \mathbf{a}_2 \\ \sigma_x \mathbf{a}_2^T & \mathbf{a}_1 \end{bmatrix} \quad \text{avec} \quad \mathbf{a}_1 = \mathbf{I}_{mL} + \sigma_x^2 \mathbf{h}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \quad \text{et} \quad \mathbf{a}_2 = \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \end{aligned}$$

On applique alors le lemme d'inversion matricielle par blocs (B.2) :

$$\begin{aligned} \widetilde{\mathbf{C}}_j^{-1} &= \begin{bmatrix} \mathbf{C}_{(0)} & \sigma_x \mathbf{a}_2 \\ \sigma_x \mathbf{a}_2^T & \mathbf{a}_1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{C}_{(0)}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\sigma_x \mathbf{C}_{(0)}^{-1} \mathbf{a}_2 \\ \mathbf{I} \end{bmatrix} \left[ \mathbf{a}_1 - \sigma_x^2 \mathbf{a}_2^T \mathbf{C}_{(0)}^{-1} \mathbf{a}_2 \right]^{-1} \begin{bmatrix} -\sigma_x \mathbf{a}_2^T \mathbf{C}_{(0)}^{-1} & \mathbf{I} \end{bmatrix} \\ \text{Or} \quad \mathbf{a}_1 - \sigma_x^2 \mathbf{a}_2^T \mathbf{C}_{(0)}^{-1} \mathbf{a}_2 &= \mathbf{I}_{mL} + \sigma_x^2 \mathbf{h}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j - (\sigma_x \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j)^T \mathbf{C}_{(0)}^{-1} \sigma_x \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \\ &= \sigma_x^2 (\sigma_x^{-2} \mathbf{I}_{mL} + \mathbf{h}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j - (\mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j)^T \mathbf{C}_{(0)}^{-1} \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j) \\ &= \sigma_x^2 \mathbf{T}_{(0)} \quad \text{d'après} \quad (B.17) \\ \text{et on définit} \quad \mathbf{a}_3 &= -\mathbf{C}_{(0)}^{-1} \mathbf{a}_2 \\ &= -\mathbf{C}_{(0)}^{-1} \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \\ \text{On a alors :} \quad \widetilde{\mathbf{C}}_j^{-1} &= \begin{bmatrix} \mathbf{C}_{(0)}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \sigma_x \mathbf{a}_3 \\ \mathbf{I} \end{bmatrix} \left[ \sigma_x^2 \mathbf{T}_{(0)} \right]^{-1} \begin{bmatrix} \sigma_x \mathbf{a}_3^T & \mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C}_{(0)}^{-1} + \mathbf{a}_3 \mathbf{T}_{(0)}^{-1} \mathbf{a}_3^T & \sigma_x^{-1} \mathbf{a}_3 \mathbf{T}_{(0)}^{-1} \\ \sigma_x^{-1} \mathbf{T}_{(0)}^{-1} \mathbf{a}_3^T & \sigma_x^{-2} \mathbf{T}_{(0)}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C}_{(0)}^{-1} + \mathbf{b}_0 \mathbf{T}_{(0)} \mathbf{b}_0^T & \sigma_x^{-1} \mathbf{b}_0 \\ \sigma_x^{-1} \mathbf{b}_0^T & \sigma_x^{-2} \mathbf{T}_{(0)}^{-1} \end{bmatrix} \\ \text{avec} \quad \mathbf{b}_0 &= \mathbf{a}_3 \mathbf{T}_{(0)}^{-1} = -\mathbf{C}_{(0)}^{-1} \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \mathbf{T}_{(0)}^{-1} \end{aligned}$$

Par conséquent, en factorisant $\mathbf{C}_{(0)}^{-1} = \mathbf{F}_{(0)}^T \mathbf{F}_{(0)}$ et en posant $\widetilde{\mathbf{F}}_j = \begin{bmatrix} \tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{12} \\ \mathbf{0} & \tilde{\mathbf{f}}_{22} \end{bmatrix}$, où $\tilde{\mathbf{f}}_{11}$(dim. $D_0 \times D_0$) et $\tilde{\mathbf{f}}_{22}$(dim. $mL \times mL$) sont des matrices triangulaires supérieures, on retrouve bien la relation de base pour les mises à jour de Cholesky (3.26) :

$$\begin{aligned} \widetilde{\mathbf{F}}_j^T \widetilde{\mathbf{F}}_j &= \begin{bmatrix} \mathbf{F}_{(0)}^T \mathbf{F}_{(0)} + \mathbf{b}_0 \mathbf{T}_{(0)} \mathbf{b}_0^T & \sigma_x^{-1} \mathbf{b}_0 \\ \sigma_x^{-1} \mathbf{b}_0^T & \sigma_x^{-2} \mathbf{T}_{(0)}^{-1} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{f}}_{11}^T \tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{11}^T \tilde{\mathbf{f}}_{12} \\ (\tilde{\mathbf{f}}_{11}^T \tilde{\mathbf{f}}_{12})^T & \tilde{\mathbf{f}}_{12}^T \tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T \tilde{\mathbf{f}}_{22} \end{bmatrix} \\ \text{avec} \quad \mathbf{b}_0 &= -(\mathbf{F}_{(0)}^T \mathbf{F}_{(0)})^{-1} \mathbf{G}_{(0)}^T \boldsymbol{\Sigma}^{-1} \mathbf{h}_j \mathbf{T}_{(0)}^{-1} \end{aligned} \tag{B.18}$$

**Mise à jour de F lors de la naissance de $\mathbf{h}_j$ (éq.(3.27)).** Ce résultat se démontre aisément en développant la factorisation $\overline{\mathbf{F}}_{(0)}^T\overline{\mathbf{F}}_{(0)} + \mathbf{v}\mathbf{v}^T$.

**Mise à jour de F lors de la mort de $\mathbf{h}_j$ (cas 1 - éq.(3.28)).**

$$\begin{array}{rrcl}
\text{D'après (B.18), on a} & \mathbf{F}_{(0)}^T\mathbf{F}_{(0)} & = & \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{11} - \mathbf{b}_0\mathbf{T}_{(0)}\mathbf{b}_0^T \\
\text{mais aussi} & \mathbf{T}_{(0)}^{-1} & = & \sigma_x^2(\tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22}) \\
\text{et} & \mathbf{b}_0 & = & \sigma_x\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12} \\
\text{par conséquent, on a donc} & \mathbf{b}_0\mathbf{T}_{(0)}\mathbf{b}_0^T & = & \mathbf{v}_1\mathbf{v}_1^T \\
\text{avec} & \mathbf{v}_1 & = & \mathbf{b}_0\mathbf{U},\ \mathbf{U} = \mathbf{T}_{(0)}^{1/2} = \sigma_x(\tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22})^{1/2} \\
 & & = & \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12}(\tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22})^{1/2}
\end{array}$$

**Mise à jour de F lors de la mort de $\mathbf{h}_j$ (cas 2 - éq.(3.29)).** Dans ce cas, la matrice $\mathbf{G}_{(1)}$ s'écrit $\mathbf{G}_{(1)} = [\mathbf{G}_a \quad \sigma_x\mathbf{h}_j \quad \mathbf{G}_b]$ avec $\mathbf{G}_a(\text{dim. } N \times a)$ et $\mathbf{G}_b(\text{dim. } N \times b)$. On peut alors définir la matrice $\mathbf{P}$ suivante :

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_a & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \\ \mathbf{0} & \mathbf{I}_b & \mathbf{0} \end{bmatrix} \quad \text{telle que} \quad \mathbf{G}_{(1)}\mathbf{P} = [\mathbf{G}_a \quad \mathbf{G}_b \quad \sigma_x\mathbf{h}_j] = \widetilde{\mathbf{G}}_j$$

La matrice $\mathbf{P}$ est une matrice de permutation : on a $\mathbf{P}^{-1} = \mathbf{P}^T$. Notons que multiplier à droite par $\mathbf{P}$ permute les colonnes (comme c'est le cas ici), tandis que multiplier à gauche permute les lignes. On démontre alors la relation $\widetilde{\mathbf{F}}_j^T\widetilde{\mathbf{F}}_j = (\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})$ :

$$\begin{array}{rrcl}
 & \widetilde{\mathbf{C}}_j & = & \mathbf{I} + \widetilde{\mathbf{G}}_j^T\mathbf{\Sigma}^{-1}\widetilde{\mathbf{G}}_j = \mathbf{I} + (\mathbf{G}_{(1)}\mathbf{P})^T\mathbf{\Sigma}^{-1}\mathbf{G}_{(1)}\mathbf{P} \\
 & & = & \mathbf{P}^T(\mathbf{I} + \mathbf{G}_{(1)}^T\mathbf{\Sigma}^{-1}\mathbf{G}_{(1)})\mathbf{P} \\
 & & = & \mathbf{P}^T\mathbf{C}_{(1)}\mathbf{P} \\
\text{Alors} & \widetilde{\mathbf{C}}_j^{-1} & = & (\mathbf{P}^T\mathbf{C}_{(1)}\mathbf{P})^{-1} = \mathbf{P}^{-1}\mathbf{C}_{(1)}^{-1}\mathbf{P}^{T^{-1}} \\
 & & = & \mathbf{P}^T\mathbf{C}_{(1)}^{-1}\mathbf{P} \\
\text{Soit} & \widetilde{\mathbf{F}}_j^T\widetilde{\mathbf{F}}_j & = & \mathbf{P}^T\mathbf{F}_{(1)}^T\mathbf{F}_{(1)}\mathbf{P} = (\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})
\end{array}$$

On peut alors écrire $\mathbf{F}_{(1)}$ en blocs de dimensions $(a, m, b)$ pour pouvoir appliquer les permutations.

$$\text{En posant} \quad \mathbf{F}_{(1)} = \begin{bmatrix} \mathbf{f}_{aa} & \mathbf{f}_{am} & \mathbf{f}_{ab} \\ \mathbf{0} & \mathbf{f}_{mm} & \mathbf{f}_{mb} \\ \mathbf{0} & \mathbf{0} & \mathbf{f}_{bb} \end{bmatrix} \quad \text{alors} \quad \mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P} = \begin{bmatrix} \mathbf{f}_{aa} & \mathbf{f}_{ab} & \mathbf{f}_{am} \\ \mathbf{0} & \mathbf{f}_{bb} & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_{mb} & \mathbf{f}_{mm} \end{bmatrix}$$

On a donc $\widetilde{\mathbf{F}}_j \neq \mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P}$ car cette dernière n'est pas triangulaire supérieure. Cependant, pour mettre à jour $\mathbf{F}$ vers $\mathbf{F}_{(0)}$, il est uniquement nécessaire d'avoir accès au bloc $\tilde{\mathbf{f}}_{11}$ de $\widetilde{\mathbf{F}}_j$ et aux valeurs de $\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12}$ et $(\tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22})$ pour calculer $\mathbf{v}_1$, comme nous l'avons vu dans le 1er cas de mort de l'atome $\mathbf{h}_j$.

Pour récupérer $\tilde{\mathbf{f}}_{11}$, on repart de l'écriture par blocs $(a, m, b)$ de $\mathbf{F}_{(1)}$ et on calcule le produit $(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})$ :

$$\begin{array}{rcl}
(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P}) & = & \begin{bmatrix} \mathbf{f}_{aa}^T\mathbf{f}_{aa} & \mathbf{f}_{aa}^T\mathbf{f}_{ab} & \mathbf{f}_{aa}^T\mathbf{f}_{am} \\ \mathbf{f}_{ab}^T\mathbf{f}_{aa} & \mathbf{f}_{ab}^T\mathbf{f}_{ab} + \mathbf{f}_{bb}^T\mathbf{f}_{bb} + \mathbf{f}_{mb}^T\mathbf{f}_{mb} & \mathbf{f}_{ab}^T\mathbf{f}_{am} + \mathbf{f}_{mb}^T\mathbf{f}_{mm} \\ \mathbf{f}_{am}^T\mathbf{f}_{aa} & \mathbf{f}_{am}^T\mathbf{f}_{ab} + \mathbf{f}_{mm}^T\mathbf{f}_{mb} & \mathbf{f}_{am}^T\mathbf{f}_{am} + \mathbf{f}_{mm}^T\mathbf{f}_{mm} \end{bmatrix} \\
\text{Ce qui est égal à : } \widetilde{\mathbf{F}}_j^T\widetilde{\mathbf{F}}_j & = & \begin{bmatrix} \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12} \\ \tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22} \end{bmatrix}
\end{array}$$

Le produit $\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{11}$ s'identifie donc avec les 4 blocs colorés en verts de $(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})$. Or, $\mathbf{f}_{bb}$ étant une matrice triangulaire supérieure de rang plein et $\mathbf{f}_{mb}^T$ une matrice de $m$ colonne, on peut définir le facteur de Cholesky $\mathbf{f}_*$ triangulaire supérieure telle que $\mathbf{f}_*^T\mathbf{f}_* = \mathbf{f}_{bb}^T\mathbf{f}_{bb} + \mathbf{f}_{mb}^T\mathbf{f}_{mb}$, soit $\mathbf{f}_* = \text{cholupdate}(\mathbf{f}_{bb}, \mathbf{f}_{mb}^T, +)$. Par unicité de la factorisation de Cholesky, on peut alors en déduire en identifiant les blocs colorés en vert que $\tilde{\mathbf{f}}_{11} = \begin{bmatrix}\mathbf{f}_{aa} & \mathbf{f}_{ab}\\ \mathbf{0} & \mathbf{f}_*\end{bmatrix}$, la première conclusion de l'équation (3.29).

Pour récupérer les valeurs de $\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12}$ et $(\tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22})$, on écrit cette fois ci la matrice $\mathbf{F}_{(1)}$ en 3 blocs colonnes : $\mathbf{F}_{(1)} = \begin{bmatrix}\mathbf{f}_a & \mathbf{f}_m & \mathbf{f}_b\end{bmatrix}$. On a alors :

$$\begin{aligned}
(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{P}^T\mathbf{F}_{(1)}\mathbf{P}) &= (\mathbf{F}_{(1)}\mathbf{P})^T(\mathbf{F}_{(1)}\mathbf{P})\\
&= \begin{bmatrix}\mathbf{f}_a & \mathbf{f}_b & \mathbf{f}_m\end{bmatrix}^T\begin{bmatrix}\mathbf{f}_a & \mathbf{f}_b & \mathbf{f}_m\end{bmatrix}\\
&= \begin{bmatrix}\mathbf{f}_a^T\mathbf{f}_a & \mathbf{f}_a^T\mathbf{f}_b & \mathbf{f}_a^T\mathbf{f}_m\\ \mathbf{f}_b^T\mathbf{f}_a & \mathbf{f}_b^T\mathbf{f}_b & \mathbf{f}_b^T\mathbf{f}_m\\ \mathbf{f}_m^T\mathbf{f}_a & \mathbf{f}_m^T\mathbf{f}_b & \mathbf{f}_m^T\mathbf{f}_m\end{bmatrix}\\
\text{Ce qui est égal à : } \widetilde{\mathbf{F}}_j^T\widetilde{\mathbf{F}}_j &= \begin{bmatrix}\tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{11}^T\tilde{\mathbf{f}}_{12}\\ \tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{11} & \tilde{\mathbf{f}}_{12}^T\tilde{\mathbf{f}}_{12} + \tilde{\mathbf{f}}_{22}^T\tilde{\mathbf{f}}_{22}\end{bmatrix}
\end{aligned}$$

En identifiant les termes de couleurs identiques, on obtient la deuxième conclusion de l'équation (3.29).

## B.7 Cadre non supervisé : échantillonnage des variances avec une étape de Métropolis-Hastings

Tout comme dans le cas des paramètres non linéaires $\boldsymbol{\nu}$ que nous ne pouvions échantillonner directement à cause de la non-linéarité du modèle, nous pouvons échantillonner les variances du bruit et des amplitudes avec une étape de Métropolis-Hastings. Pour chacune des variances, $\sigma_x^2$ et $\sigma_\epsilon^2$, que nous nommerons ici $\sigma^2$, cela revient à :

1. Proposer une variance candidate $\sigma'^2$ suivant une loi de proposition $Q(\sigma'^2 \,|\, \sigma^2)$.
2. Accepter la proposition avec la probabilité :

$$\alpha = \frac{p(\sigma'^2 \,|\, -)}{p(\sigma^2 \,|\, -)}\frac{Q(\sigma^2 \,|\, \sigma'^2)}{Q(\sigma'^2 \,|\, \sigma^2)} = \frac{|\mathbf{B}'|^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}'^{-1}\boldsymbol{y}\right)}{|\mathbf{B}|^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{y}^T\mathbf{B}^{-1}\boldsymbol{y}\right)}\frac{p(\sigma'^2)}{p(\sigma^2)}\frac{Q(\sigma^2 \,|\, \sigma'^2)}{Q(\sigma'^2 \,|\, \sigma^2)}$$

où $\mathbf{B}$ (resp. $\mathbf{B}'$) correspond à la matrice pour la valeur du paramètre $\sigma^2$ (resp. $\sigma'^2$). Afin de calculer la probabilité d'acceptation $\alpha$, on peut calculer les matrices $\mathbf{B}$ et $\mathbf{B}'$ à l'aide de leur définition ($\mathbf{B} = \sigma_\epsilon^2\mathbf{I}_N + \sigma_x^2\overline{\mathbf{H}}\overline{\mathbf{H}}^T$). Cependant, ces deux matrices sont très similaires, puisqu'elles ne diffèrent qu'en la valeur de $\sigma^2$ :

$$\begin{aligned}
\text{Pour } \sigma^2 = \sigma_x^2 \;:\; \mathbf{B}' - \mathbf{B} &= (\sigma'^2 - \sigma^2)\overline{\mathbf{H}}\overline{\mathbf{H}}^T \quad (\text{dim. } N \times N)\\
\text{Pour } \sigma^2 = \sigma_\epsilon^2 \;:\; \mathbf{B}' - \mathbf{B} &= (\sigma'^2 - \sigma^2)\mathbf{I}_N \quad (\text{dim. } N \times N)
\end{aligned}$$

On peut alors utiliser les astuces calculatoires vues en § 3.3 (basées sur les lemmes d'inversion et de déterminant matriciels) afin de pouvoir réécrire la probabilité d'acceptation $\alpha$ en fonction de deux matrices $(\mathbf{T}, \boldsymbol{\gamma})$, elles-mêmes calculables l'aide des matrices courantes $\mathbf{F}$ et $\mathbf{G}$. Contrairement au cas des naissances et des morts où les matrices $\mathbf{T}$,$\boldsymbol{\gamma}$ étaient de dimensions $m$, les matrices $\mathbf{T}$,$\boldsymbol{\gamma}$ seront ici de dimensions plus élevées :

$$\begin{array}{lcllll} \text{Pour } \sigma^2 = \sigma_x^2 & : & \mathbf{T} & (\text{dim. } mL \times mL) & \text{et} & \boldsymbol{\gamma} \quad (\text{dim. } mL \times 1) \\ \text{Pour } \sigma^2 = \sigma_\epsilon^2 & : & \mathbf{T} & (\text{dim. } N \times N) & \text{et} & \boldsymbol{\gamma} \quad (\text{dim. } N \times 1) \end{array}$$

Le calcul de la probabilité d'acceptation devrait cependant être plus efficace qu'en calculant les matrices $\mathbf{B}$ et $\mathbf{B}'$.
Enfin, notons que l'on peut aussi proposer simultanément les deux variances pour n'avoir qu'une seule étape de Métropolis-Hastings. La probabilité d'acceptation est alors :

$$\alpha = \frac{p({\sigma'}_x^2, {\sigma'}_\epsilon^2 \,|\, -)}{p(\sigma_x^2, \sigma_\epsilon^2 \,|\, -)} \frac{Q(\sigma_x^2, \sigma_\epsilon^2 \,|\, {\sigma'}_x^2, {\sigma'}_\epsilon^2)}{Q({\sigma'}_x^2, {\sigma'}_\epsilon^2 \,|\, \sigma_x^2, \sigma_\epsilon^2)} = \frac{|\mathbf{B}'|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T {\mathbf{B}'}^{-1} \boldsymbol{y}\right)}{|\mathbf{B}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \mathbf{B}^{-1} \boldsymbol{y}\right)} \frac{p({\sigma'}_x^2) p({\sigma'}_\epsilon^2)}{p(\sigma_x^2) p(\sigma_\epsilon^2)} \frac{Q(\sigma_x^2, \sigma_\epsilon^2 \,|\, {\sigma'}_x^2, {\sigma'}_\epsilon^2)}{Q({\sigma'}_x^2, {\sigma'}_\epsilon^2 \,|\, \sigma_x^2, \sigma_\epsilon^2)}$$

Cependant, on ne peut plus profiter de l'avantage calculatoire proposé ci-dessus, car la différence des deux matrices $(\mathbf{B}', \mathbf{B})$ s'écrit $\mathbf{B}' - \mathbf{B} = ({\sigma'}_\epsilon^2 - \sigma_\epsilon^2)\mathbf{I}_N + ({\sigma'}_x^2 - \sigma_x^2)\overline{\mathbf{H}}\overline{\mathbf{H}}^T$ et ne peut plus s'écrire sous la forme $\mathbf{a} + \mathbf{b}\mathbf{c}^{-1}\mathbf{d}$ permettant l'utilisation du lemme de l'inversion matricielle.

# Table des figures


1.1 Influence de l'échantillonnage régulier ou irrégulier sur la fenètre spectrale. . . . . 7
1.2 Représentation de la discrétisation du domaine des localisations pour la reformulation en APL, et de sa partition pour la reformulation en APNL. . . . . 23
1.3 Signal irrégulièrement échantillonnés et son contenu spectral. . . . . 30
1.4 (a) Ondelette utilisée. (b) Exemple de signal simulé avec RSB=10dB ; signal en noir, bruit en gris, positions réelles du maximum des échos signalées avec les ronds rouges. . . . . 31
1.5 Schéma représentant l'échantillonnage des données et la grille. En noir, la grille de recherche entre 0 et $J-1=99$. En bleu, l'écho $\boldsymbol{h}(0)$ et en rouge $\boldsymbol{h}(99)$, où $\nu \in \mathcal{G}$. En jaune, l'écho $\boldsymbol{h}(43.7)$ où $\nu \notin \mathcal{G}$. En-dessous de chaque écho est représentée sa durée énergétique. . . . . 32

2.1 Représentation de la partition du domaine de localisation pour la reformulation en APNL spécifique au cadre de l'interpolation de dictionnaire. . . . . 35
2.2 Illustration de l'espace image $\mathcal{S}^+$ et de sa relaxation convexe $\mathcal{H}^+$ pour les différentes interpolations de dictionnaire. . . . . 43
2.3 Illustration des erreurs d'approximation pour différentes approximations. À gauche : erreur d'approximation moyenne en fonction de la taille $\Delta$ des intervalles, pour l'approximation discrète/APL (BP), les approximations de Taylor à l'ordre 1 (T1) et 2 (T2) et pour l'interpolation polaire (P). Crédit : [ETS11]. À droite : erreur d'approximation en fonction de $\tau = \delta \in \mathcal{I}_\Delta$ pour un intervalle donné, et pour les approximations Minimax, Taylor à l'ordre 1, l'interpolation polaire et l'approximation par K-SVD. Crédit : [PPZ17]. . . . . 46
2.4 Illustration de l'espace image $\mathcal{S}_j$ (en bleu) et des relaxations convexes $\mathcal{H}_j^+$ et $\mathcal{H}_j^-$ (en rouge) pour les différentes approximations linéaires. La relaxation convexe $\mathcal{H}_j$ de $\mathcal{S}_j$ est représentée en jaune. . . . . 49
2.5 Illustration du principe de l'interpolation polaire. . . . . 60
2.6 Illustration en 3D des arcs interpolateurs proposés par [Eka12] et [FDJ15]. . . . 62
2.7 Comparaison des espaces $\mathcal{S}^+$ et $\mathcal{H}^+$ de l'interpolation polaire pour deux $\theta$ différents. 63
2.8 Évolution de $\theta$ en fonction de $\Delta$. . . . . 64
2.9 Représentation des différentes relaxations convexes de l'espace $\mathcal{S}^+$ à $\alpha > 0$ fixé dans le cas de l'interpolation polaire. De gauche à droite : de la relaxation la plus grande à l'enveloppe convexe. . . . . 64
2.10 Illustration de l'espace $\mathcal{C}_0^+$ dans le plan $(\widetilde{x}_1, \widetilde{x}_k)$ pour 3 cas différents, chacun construit à partir $P=5$ points $\boldsymbol{z}_p$. . . . . 67
2.11 Représentations des points $\boldsymbol{z}_p$ en discrétisant de plus en plus finement l'intervalle $\mathcal{I}_\Delta$. (a) On augmente progressivement $P$. (b) Pour $P=65$, on regarde la répartition des points suivant les $\delta_p$ auxquels ils correspondent. Modèle utilisé pour les tests en § 2.5. . . . . 69

2.12 Illustration de l'espace $\mathcal{S}^{\mathbb{R}}$ pour $K = 2$ (gauche) et $K = 3$ (droite) et pour $\Delta = \Delta_0$, $4\Delta_0$ et $8\Delta_0$. Les lignes bleues (resp. rouges) correspondent à des dilatations positives (resp. négatives) de $\Psi(\delta)$. . . . . 71
2.13 Erreur d'approximation théorique moyenne pour chaque interpolation testée, pour $\Delta$ variant de $\Delta_0/10$ à $10\Delta_0$. En bleu (BP) : sur grille ($\delta = 0$). En rouge (T1) : approximation de Taylor. En jaune (Po) : interpolation polaire. En violet (K2) : approximation par 2-SVD continue. En vert (K3) : approximation par 3-SVD continue. . . . . 73
2.14 Résultats d'estimation sur le signal en Figure 1.4(b) avec RSB=10dB pour les méthodes d'approximation parcimonieuse : $\mathcal{P}^{\mathcal{D}}_{2+1}$, $\mathcal{P}^{\mathcal{D}}_{2/0}$, $\mathcal{P}^{\mathcal{C}}_{2+1}$ et $\mathcal{P}^{\mathcal{C}}_{2/0}$. Les cercles rouges montrent les vraies localisations des pics et les croix bleues leur positions estimées. Pour chaque méthode, le résidu est tracé en gris, et sa norme $\zeta^2$ et l'ASD associé sont donnés. La figure du bas montre les résultats des 4 estimations en zoomant autour du 4$^{\text{e}}$ pic. . . . . 76
2.15 Représentation du temps moyen de calcul pour la résolution $\ell_0$ en fonction du pourcentage d'ELR. . . . . 79
2.16 Illustrations pour l'interpolation polaire en analyse spectrale. . . . . 84
2.17 Approximation K-SVD. Allure des interpolateurs $\Psi^c(\delta_j)$ (rouge) et des espaces $\mathcal{S}^+$ (bleu) et $\mathcal{S}^-$ (jaune) autour d'une fréquence $f_0$, pour $J = 100$ et $J = 200$ (et donc deux $\Delta$ différents). Première et deuxième ligne : projection dans les espaces $(\widetilde{x}_1, \widetilde{x}_2)$ et $(\widetilde{x}_1, \widetilde{x}_3)$. . . . . 85
2.18 Pourcentage de non-injectivité (K=2) et de relaxation des contraintes (K=3) pour l'espace des contraintes avec approximation par K-SVD suivant le nombre d'intervalles $J$. . . . . 85
2.19 Représentation des espaces $\mathcal{S}^+_\pi \cup \mathcal{S}^-_\pi$ et $\mathcal{S}^+_{2\pi}$ pour plusieurs tailles de grille $J$, autour de la fréquence $f_0$. (a) En bleu $\mathcal{S}^+_\pi$ et en rouge $\mathcal{S}^-_\pi$. (b) $\mathcal{S}^+_{2\pi}$ : En vert, points correspondant à $\phi \leq \pi$ et en jaune points correspondants à $\phi \geq \pi$. (c) Les cercles en jaune, rouge et violet représentent respectivement $\Psi(-\frac{\Delta}{2}, \phi)$, $\Psi(0, \phi)$ et $\Psi(\frac{\Delta}{2}, \phi)$ pour $\phi$ variant de 0 à $2\pi$. (d) Proposition de plan séparateur reposant sur les cercles $\Psi(\pm\frac{\Delta}{2}, \phi)$ pour contraindre l'espace $\mathcal{S}^+_{2\pi}$. (e-f) Espace $\mathcal{S}^+_{2\pi}$ pour $J = 100$ - cercles et plans séparateurs. (g) Pourcentage de points appartenant à l'espace décrit par les deux plans selon $J$. . . . . 88
2.20 Erreur d'approximation moyenne $\xi_J$ pour différentes valeurs de taille de grille $J$ et pour les différentes approximations linéaires proposées. BP signifie sur grille ($\delta_j = 0$). . . . . 91
2.21 Représentation du temps de calcul pour la résolution $\ell_0$ en fonction de l'erreur moyenne sur l'estimation des fréquences. . . . . 91

3.1 Illustration du modèle Bernoulli-Gaussien pour des amplitudes bidimensionnelles. 98
3.2 Illustration de l'estimation au sens de l'EAP de la séquence de Bernoulli $\widehat{\boldsymbol{q}}$ et de la séquence détectée $\widehat{\boldsymbol{q}}_\alpha$ avec un seuil $\alpha = 0.8$. . . . . 100
3.3 Exemples typiques de configurations favorisant les modes locaux pour l'échantillonneur de Gibbs. . . . . 105
3.4 Illustration du modèle Bernoulli-Gaussien étendu. . . . . 107
3.5 Illustration de l'estimation séquentielle de la séquence de Bernoulli et des paramètres de localisation. . . . . 109
3.6 Allure des lois *a priori* et *a posteriori* conditionnelles pour $\lambda$. . . . . 132
3.7 Allure de la loi *a priori* $\sigma_x^2$. . . . . 133

3.8 Comparaison des échantillonneurs Gibbs-hybride joint et Gibbs-hybride marginalisé dans le cadre supervisé sur une chaîne de Markov. . . . . 137
3.9 Échantillons d'une fréquence dans la cadre supervisé. . . . . 138
3.10 Résultats de la nouvelle stratégie d'estimation à partir d'échantillons de la loi marginale. . . . . 139
3.11 Échantillons de la séquence de Bernoulli selon la loi marginale dans le cadre non-supervisé et histogramme de $L^{(t)} = \|\boldsymbol{q}^{(t)}\|$ pour plusieurs réglages des lois *a priori* des hyperparamètres. . . . . 142

4.1 Différentes méthodes de détection d'exoplanètes. . . . . 149
4.2 Illustration de l'orbite d'une exoplanète. Crédits : Nathan Hara et Jacques Laskar - ESEP. . . . . 151
4.3 Illustration de la mesure de la vitesse radiale. . . . . 153
4.4 À gauche, illustration d'un signal de vitesse radiale pour une orbite circulaire. Les mesures sont effectuées sur plusieurs années. Source : [BWM+06]. À droite, illustration d'un signal de vitesse radiale pour une orbite elliptique d'excentricité $e = 0.67$ (abscisse en jours). . . . . 154
4.5 Représentation des différentes sources d'activité stellaire en fonction de l'ordre de grandeur de la source (ordonnées) et de la durée de vie/d'influence des évènements en question (abscisses), avec un ordre de grandeur de l'influence sur la vitesse radiale de l'étoile mesurée par spectroscopie. Source : [Héb15] (Fig. 1.15, p.27). . 159
4.6 Illustration de l'influence de l'activité stellaire sur la CCF et des indicateurs. . . 160
4.7 Répartition des planètes du système solaire sur une grille de $J = 100$ points entre $P_{\min} = 1$ jours et $P_{\max} = 60000$ jours, suivant les différentes répartitions : régulière en log10, périodes et fréquences. . . . . 166
4.8 HD40307 (a) En vert, modèle képlérien ($\boldsymbol{y}^{\text{th}}$) simulé à partir des paramètres orbitaux estimés par les astrophysiciens. En noir, échantillons $y_n$ simulés avec du bruit aux instants $t_n$. (b) Fenêtre spectrale liées aux instants d'échantillonnage $\boldsymbol{t}$. 172
4.9 Représentation des périodes et des semi-amplitudes utilisées pour la simulation sur la grille de recherche de $J = 100$ points régulière en logarithme et zoom autour des périodes de simulation. . . . . 172
4.10 « Histogramme » de répartition des périodes estimées par un algorithme glouton avec descente de gradient de type OLS avec initialisations multiples, pour 100 réalisations du bruit sur le modèle HD40307. Chaque couleur correspond à une des 6 exoplanètes ; les traits en pointillés indiquent les périodes de simulation. . 174
4.11 HD40307 - Échantillonnage des paramètres d'intérêt. . . . . 175
4.12 HD40307 - Évolution et estimation de l'ordre du modèle $L$ . . . . . 176
4.13 HD40307 - Histogrammes des séquences de Bernoulli $\boldsymbol{q}^{(t)}$. . . . . 176
4.14 HD40307 - Histogrammes des échantillons des périodes et des excentricités. . . . 177
4.15 HD40307 - Échantillons et histogrammes des hyperparamètres $\lambda$ et $\sigma_x^2$. . . . . 178
4.16 HD40307 - Histogrammes des séquences de Bernoulli $\boldsymbol{q}^{(t)}$ pour un modèle circulaire. . . . . 179
4.17 Carte de localisation des bonnes (en blanc) et des mauvaises détections (en noir), sur l'ensemble des 1000 chaînes calculées (100 réalisations du bruit × 10 initilisations aléatoires). . . . . 179
4.18 « Histogramme » de répartition des périodes estimées par le BGE, pour 100 réalisations du bruit sur le modèle HD40307. Voir Figure 4.10 pour plus d'explications. . . . . 180
4.19 HD40307 - Histogrammes des estimations des périodes et des excentricités des 1000 chaînes. . . . . 181

4.20 Les six différents systèmes simulés par [NFB+18]. À gauche : signaux de vitesses radiales $\boldsymbol{y}$. À droite : fenêtres spectrales associées aux instants d'échantillonnage $\boldsymbol{t}$. 184
4.21 Histogrammes cumulés des $\widehat{\boldsymbol{q}}$ et des $\widehat{L}$ sur les 10 graines aléatoires pour chacun des 6 systèmes. . . . . 185
4.22 Observation du caractère bimodal des échantillons pour les sytèmes 2 et 5. . . . 187
4.23 HD40307 - Données réelles. Histogramme cumulé des $\widehat{\boldsymbol{q}}$ sur les 10 graines aléatoires. Pour chaque signal ($\boldsymbol{y}_F$ et $\boldsymbol{y}_R$), on donne les résultats avec exploitation d'un processus gaussien (PG) ou sans ($\varnothing$). Les lignes verticales indiquent les indices correspondant aux périodes des exoplanètes détectées par [TAEG+13]. . . . . 190
4.24 CoRoT-7 - Données réelles. Fenêtre spectrale associée à l'échantillonnage du signal. . . . . 191
4.25 CoRoT-7 - Données réelles. Histogramme cumulé des $\widehat{\boldsymbol{q}}$ (a) sur les 10 graines aléatoires (b) sur les graines (1-2-5-10) et (c) sur les graines (3-4-6-7-9). Les lignes verticales vertes indiquent les indices correspondant aux périodes des exoplanètes détectées par [FHB+16], et les lignes rouges indiquent leurs détections incertaines. La ligne grise correspond à l'indice associé à la période de rotation de l'étoile $P_{\text{rot}}$. 192
4.26 CoRoT-7 - Données réelles. Histogrammes cumulés des hyperparamètres ($\lambda$, $\sigma_x^2$,$\sigma_{\text{jit}}^2$). . . . . 193
4.27 CoRoT-7 - Données réelles. Fonctions de covariance $R(\tau)$ exprimée en les paramètres $\boldsymbol{\theta}_{\text{act}}$ estimés ; à droite, zoom sur les faibles décalages. . . . . 193

# Liste des tableaux


1.1 Récapitulatif des trois modèles de sinusoïdes pour de l'analyse spectrale, dans le cadre des problèmes inverses non linéaires. . . . . 6
1.2 Valeurs de simulation des fréquences et des amplitudes pour le signal $\boldsymbol{y}$. . . . . 30

2.1 Récapitulatif des propriétés des différentes interpolations de dictionnaire existantes. 41
2.2 Réglage du *big-M* suivant les différents cas d'interpolation. . . . . 75
2.3 Pourcentage d'ELR et ASD moyen (et écart-type) calculés sur l'ensemble des 200 tests pour $\mathcal{P}^{\mathcal{C}}$. . . . . 77
2.4 Pourcentage d'ELR et ASD moyen (et écart-type) calculés sur l'ensemble des 200 tests pour les différentes approximations linéaires avec résolution $\ell_0$. . . . . 78
2.5 Récapitulatif des diverses dimensions pour la résolution $\ell_0$ des problèmes d'approximation parcimonieuse sur dictionnaires interpolés. . . . . 90

3.1 Algorithme d'échantillonnage de Gibbs classique. . . . . 101
3.2 Algorithme d'échantillonnage de Gibbs pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x} \,|\, \boldsymbol{y})$. . . . . 103
3.3 Récapitulatif des correspondances entre les loi *a priori* sur les hyperparamètres et les lois *a posteriori* conditionnelles qui en résultent dans le cadre du modèle Bernoulli-Gaussien. . . . . 104
3.4 Algorithme d'échantillonnage de Gibbs pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ dans le cadre non supervisé. . . . . 104
3.5 Algorithme d'échantillonnage de Gibbs-hybride pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$ dans le cadre non supervisé. . . . . 110
3.6 Algorithme d'échantillonnage de Gibbs avec étapes de Reversible-Jump (RJGS) pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu} \,|\, \boldsymbol{y})$ . . . . . 111
3.7 Algorithme d'échantillonnage de Gibbs avec étapes de Métropolis-Hastings (MHGS) pour générer des échantillons de la loi *a posteriori* $p(\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{\nu} \,|\, \boldsymbol{y})$ . . . . . 112
3.8 Échantillonneur de la loi *a posteriori* marginalisée $p(\boldsymbol{q}, \boldsymbol{\nu} \,|\, \boldsymbol{y}, \boldsymbol{\theta})$. . . . . 115
3.9 Échantillonneur de la loi *a posteriori* marginalisée dans le cadre non supervisé. . 127
3.10 Échantillonneur proposé de type PCGS de la loi *a posteriori* jointe dans le cadre non supervisé. . . . . 128
3.11 Récapitulatif des réglages de lois *a priori* des hyperparamètres. . . . . 136
3.12 Nouvelle stratégie d'estimation pour l'échantillonneur de Gibbs marginalisé : fusion des intervalles anticorrélés. . . . . 139
3.13 Comparaison des résultats des échantillonneurs de la loi jointe et marginalisée sur 50 initialisations différentes. . . . . 140
3.14 Comparaison des résultats des échantillonneurs de la loi marginalisée sur 50 initialisations pour différents réglages des lois *a priori* sur les hyperparamètres. . . 141

3.15 Test sur les données de déconvolution : comparaison des pourcentages d'ELR entre les échantilonneurs de la loi jointe et marginale dans le cadre non supervisé. 143

4.1 Ordre de grandeur de la contribution $K$ d'une planète à la vitesse radiale d'une étoile similaire au Soleil, suivant sa masse, sa distance à l'étoile et sa période. Les dénominatifs *Jupiter*, *Neptune* et *Terre* sont uniquement relatifs à la masse de la planète concernée (et non à la distance $a$ ni à la période $P$). Tableau tiré de [Hay15] (p.4). . . . . 152
4.2 Paramètres orbitaux (périodes, excentricités et semi-amplitudes) des 6 exoplanètes reportées par [TAEG+13]. . . . . 171
4.3 Réglages de l'échantillonneur pour application à HD40307 sur données simulées sans activités. . . . . 175
4.4 Paramètres orbitaux des 6 signaux de vitesses radiales simulés par [NFB+18]. Les erreurs de mesures des systèmes 3 et 6 sont différentes, bien que du même ordre de grandeur ; le deux systèmes diffèrent donc par l'échantillonnage du signal et par une réalisation du bruit différente. . . . . 183
4.5 Récapitulatif des exoplanètes détectées pour les 6 systèmes avec le modèle BGE sur une chaîne d'échantillons. Périodes estimées ainsi que probabilités de présence des exoplanètes. . . . . 188
4.6 Récapitulatif des « facteurs de Bayes » (BF) et des «Posterior odds ratio» (POR) de notre modèle pour chaque système de [NFB+18]. . . . . 188
4.7 Correspondance entre les périodicités repérées par [FHB+16] et les intervalles de la grille pour le modèle BGE. . . . . 191

# Bibliographie


[AD99] C. Andrieu and A. Doucet. Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC. *IEEE Transactions on Signal Processing*, 47(10) :2667–2676, 1999. 103

[Bab12] P. Babu. *Spectral analysis of nonuniformly sampled data and applications.* PhD thesis, Uppsala universitet, 2012. 7

[Bal08] R.V. Baluev. Assessing the statistical significance of periodogram peaks. *Monthly Notices of the Royal Astronomical Society*, 385(3) :1279–1285, 2008. 161

[Bal14a] R.V. Baluev. Keplerian periodogram for Doppler exoplanet detection : optimized computation and analytic significance thresholds. *Monthly Notices of the Royal Astronomical Society*, 446(2) :1478–1492, 2014. 156

[Bal14b] R.V. Baluev. Keplerian periodogram for Doppler exoplanet detection : optimized computation and analytic significance thresholds. *Monthly Notices of the Royal Astronomical Society*, 446(2) :1478–1492, 2014. 161

[BC05] S. Bourguignon and H. Carfantan. Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data. In *2005 13th IEEE Statistical Signal Processing Workshop (SSP)*, pages 811–816. IEEE, 2005. 8, 97, 104

[BC06] S. Bourguignon and H. Carfantan. Spectral analysis of irregularly sampled data using a Bernoulli-Gaussian model with free frequencies. In *2006 Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, pages 516–519, Toulouse, France, mai 2006. 25, 27, 28, 29, 74, 96, 106, 107, 108, 109, 113, 114, 126, 133, 144

[BC13] S. Bourguignon and H. Carfantan. *Estimation de spectres de raies pour des signaux irrégulièrement échantillonnés en Astrophysique.* Traité IC2, série Signal et image, Lavoisier-Hermes, 2013. 7, 13

[BC17] M. Boudineau and H. Carfantan. Approximation parcimonieuse avec dictionnaire continu pour l'analyse spectrale de données irrégulièrement échantillonnées. In *26e colloque GRETSI*, 2017. 29, 80, 173

[BC18] M. Boudineau and H. Carfantan. An L0 solution to sparse approximation problems with continuous dictionaries. In *2018 Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, pages 4539–4543, 2018. 28, 72, 73, 75, 197

[BCB07] S. Bourguignon, H. Carfantan, and T. Böhm. SparSpec : a new method for fitting multiple sinusoids with irregularly sampled data. *Astronomy and Astrophysics*, 462(1) :379–387, 2007. 157

[BCBB16] M. Boudineau, H. Carfantan, S. Bourguignon, and M. Bazot. Sampling schemes and parameter estimation for nonlinear Bernoulli-Gaussian sparse models. In *2016 IEEE Statistical Signal Processing Workshop (SSP)*, pages 1–5, 2016. 28, 29, 96, 134, 198

[BCBC04] T. Böhm, C. Catala, L. Balona, and B. Carter. Spectroscopic monitoring of the Herbig Ae star HD 104237-I. Multiperiodic stellar oscillations. *Astronomy and Astrophysics*, 427(3) :907–922, 2004. 5

[BCI07] S. Bourguignon, H. Carfantan, and J. Idier. Minimisation de critères de moindres carrés pénalisés par la norme L1 dans le cas complexe. In *21*$^{e}$ *colloque GRETSI*, 2007. 12

[BG98] S.P. Brooks and A. Gelman. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7 :434–455, 1998. 104, 105

[BHF17] M-A. Badiu, T.L. Hansen, and B.H. Fleury. Variational Bayesian inference of line spectra. *IEEE Transactions on Signal Processing*, 65(9) :2247–2261, 2017. 25, 97, 198

[Bix12] R.E. Bixby. A brief history of linear and mixed-integer programming computation. *Documenta Mathematica*, pages 107–121, 2012. 14

[BNCM16] S. Bourguignon, J. Ninin, H. Carfantan, and M. Mongeau. Exact sparse approximation problems via mixed-integer programming : Formulations and computational performance. *IEEE Transactions on Signal Processing*, 64(6) :1405–1419, 2016. 13, 14, 17, 27, 53, 54, 92

[Bou05] S. Bourguignon. *Analyse spectrale à haute résolution de signaux irrégulièrement échantillonnés*. Thèse de doctorat, Université Paul Sabatier – Toulouse III, décembre 2005. 7, 18, 80, 97, 100, 102, 103, 106, 110, 112, 113, 114, 135, 157

[Bou14] M. Boudineau. Modélisation Bernoulli Gaussienne étendue pour la recherche de périodicités dans un signal irrégulièrement échantillonné : application à la détection d'exoplanètes. Rapport de stage de fin d'étude, École Centrale de Nantes, septembre 2014. 25, 27, 96, 106, 110, 111, 113, 156, 164, 172, 173

[BPC11] B.J. Brewer, L.B. Pártay, and G. Csányi. Diffusive nested sampling. *Statistics and Computing*, 21(4) :649–656, 2011. 162

[Bre14] B.J. Brewer. Inference for trans-dimensional Bayesian models with diffusive nested sampling. *arXiv preprint arXiv :1411.3921*, 2014. 157, 162

[BSCI11] S. Bourguignon, C. Soussen, H. Carfantan, and J. Idier. Sparse deconvolution : Comparison of statistical and deterministic approaches. In *2011 IEEE Statistical Signal Processing Workshop (SSP)*, pages 317–320. IEEE, 2011. 13

[BT09] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1) :183–202, 2009. 12

[BTR13] B.N. Bhaskar, G. Tang, and B. Recht. Atomic norm denoising with applications to line spectral estimation. *IEEE Transactions on Signal Processing*, 61(23) :5987–5999, 2013. 22

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004. 35, 36, 37, 53

[BWM$^{+}$06] R. P. Butler, J. T. Wright, G. W. Marcy, D. A. Fisher, S. S. Vogt, C. G. Tinney, B. D. Carter, J. A. Johnson, C. McCarthy, and A. J. Penny. Catalog of nearby exoplanets. *The Astrophysical Journal*, 646(1) :505–522, juillet 2006. 154, 219

[CBIS13] E. Carcreff, S. Bourguignon, J. Idier, and L. Simon. Resolution enhancement of ultrasonic signals by up-sampled sparse deconvolution. In *Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, 2013. 4, 31

[CBL89] S. Chen, S.A. Billings, and W. Luo. Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, 50(5) :1873–1896, 1989. 11

[CCL96] Q. Cheng, R. Chen, and T.-H. Li. Simultaneous wavelet estimation and deconvolution of reflection seismic signals. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2) :377–384, mars 1996. 4, 18, 19, 97, 99, 101

[CGI96] F. Champagnat, Y. Goussard, and J. Idier. Unsupervised deconvolution of sparse spike trains using stochastic approximation. *IEEE Transactions on Signal Processing*, 44(12) :2988–2998, 1996. 19, 97, 114

[Cha93] F. Champagnat. *Déconvolution impulsionnelle et extensions pour la caractérisation des milieux inhomogènes en échographie*. Thèse de doctorat, Paris 11, 1993. 97

[CJ01] S. Chib and I. Jeliazkov. Marginal likelihood from the Metropolis-Hastings output. *Journal of American Statistical Association*, 96(453) :270–281, mars 2001. 162

[CJPT13] E. Chouzenoux, A. Jezierska, J-C. Pesquet, and H. Talbot. A majorize-minimize subspace approach for L2-L0 image regularization. *SIAM Journal on Imaging Sciences*, 6(1) :563–591, 2013. 13

[CRPW12] V. Chandrasekaran, B. Recht, P.A. Parrilo, and A.S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6) :805–849, 2012. 21, 22

[CRT06] E.J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles : Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2) :489–509, 2006. 13

[Cum04] A. Cumming. Detectability of Extrasolar Planets in Radial Velocity Surveys. *Monthly Notices of the Royal Astronomical Society*, 345 :1165–1176, novembre 2004. 155, 156

[DBD+17] X. Dumusque, F. Borsa, M. Damasso, R.F. Díaz, P.C. Gregory, N.C. Hara, A. Hatzes, V. Rajpaul, M. Tuomi, S. Aigrain, et al. Radial-velocity fitting challenge-II. First results of the analysis of the data set. *Astronomy and Astrophysics*, 598 :A133, 2017. 163

[DBS14] X. Dumusque, I. Boisse, and N.C. Santos. SOAP 2.0 : a tool to estimate the photometric and radial velocity variations induced by stellar spots and plages. *The Astrophysical Journal*, 796(2) :132, 2014. 182

[DDDM04] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11) :1413–1457, 2004. 12

[DHD12] A. Drémeau, C. Herzet, and L. Daudet. Boltzmann machine and mean-field approximation for structured sparse decompositions. *IEEE Transactions on Signal Processing*, 60(7) :3425–3438, 2012. 18

[DK15] A. Drémeau and F. Krzakala. Phase recovery from a bayesian point of view : the variational approach. In *2015 Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, pages 3661–3665. IEEE, 2015. 97

[DMTC09] N. Dobigeon, S. Moussaoui, J-Y. Tourneret, and C. Carteret. Bayesian separation of spectral sources under non-negativity and full additivity constraints. *Signal Processing*, 89(12) :2657–2669, 2009. 97

[DP17] V. Duval and G. Peyré. Sparse spikes super-resolution on thin grids II : the continuous basis pursuit. *Inverse Problems*, 33(9) :095008, 2017. 45

[DT10] N. Dobigeon and J-Y. Tourneret. Bayesian orthogonal component analysis for sparse representation. *IEEE Transactions on Signal Processing*, 58(5) :2675–2685, 2010. 97

[Dub96] F. Dublanchet. *Contribution de la méthodologie bayésienne à l'analyse spectrale de raies pures et à la goniométrie haute résolution.* Thèse de doctorat, Université de Paris Sud, Orsay, octobre 1996. 25, 97

[Dum16] X. Dumusque. Radial velocity fitting challenge-I. Simulating the data set including realistic stellar radial-velocity signals. *Astronomy and Astrophysics*, 593 :A5, 2016. 163, 182

[DZ02] X. Descombes and J. Zerubia. Marked point process in image analysis. *IEEE Signal Processing Magazine*, 19(5) :77–84, 2002. 4

[Eka12] C. Ekanadham. *Continuous basis pursuit and its applications.* PhD thesis, New York University, 2012. 59, 61, 62, 217

[Ela10] M. Elad. *Sparse and Redundant Representations : From Theory to Applications in Signal and Image Processing.* Springer, 2010. 11

[Eng13] C.A. Engelbrecht. Extracting oscillation frequencies from data : various approaches. *Proceedings of the International Astronomical Union*, 9(S301) :77–84, 2013. 7

[ETS11] C. Ekanadham, D. Tranchina, and E. P. Simoncelli. Recovery of sparse translation-invariant signals with continuous basis pursuit. *IEEE Transactions on Signal Processing*, 59(10), 2011. 24, 27, 28, 34, 37, 38, 41, 42, 43, 45, 46, 53, 56, 59, 60, 61, 62, 63, 86, 92, 217

[FBH11] F. Feroz, S.T. Balan, and M.P. Hobson. Detecting extrasolar planets from stellar radial velocities using Bayesian evidence. *Monthly Notices of the Royal Astronomical Society*, 415(4) :3462–3472, 2011. 160

[FCPC13] A. Florescu, E. Chouzenoux, J-C. Pesquet, and S. Ciochina. A constrained optimization approach for complex sparse perturbed models. 2013. 45

[FDJ15] K. Fyhn, M. F. Duarte, and S. H. Jensen. Compressive Parameter Estimation for Sparse Translation-Invariant Signals Using Polar Interpolation. *IEEE Transactions on Signal Processing*, 63(4), 2015. 24, 28, 29, 34, 37, 39, 40, 44, 45, 47, 49, 50, 51, 52, 56, 58, 59, 60, 61, 62, 76, 80, 217

[FG07] E. B. Ford and P. C. Gregory. Bayesian Model Selection and Extrasolar Planet Detection. In G. J. Babu and E. D. Feigelson, editors, *Statistical Challenges in Modern Astronomy IV*, volume 371 of *Astronomical Society of the Pacific Conference Series*, page 189, novembre 2007. 162

[FHB+16] J.P. Faria, R.D. Haywood, B.J. Brewer, P. Figueira, M. Oshagh, A. Santerne, and N.C. Santos. Uncovering the planets and stellar activity of CoRoT-7 using only radial velocities. *Astronomy and Astrophysics*, 588 :A31, 2016. 157, 162, 191, 192, 194, 199, 220, 222

[FJA05] L. Fathi, G. Jourdain, and M. Arndt. A Bernoulli-Gaussian model for multipath channel estimation in downlink WCDMA. In *2005 13th IEEE Statistical Signal Processing Workshop (SSP)*, pages 651–656. IEEE, 2005. 97

[FL12] A. Fannjiang and W. Liao. Coherence pattern–guided compressive sensing with unresolved grids. *SIAM Journal on Imaging Sciences*, 5(1) :179–202, 2012. 40

[FM81] S. Ferraz-Mello. Estimation of periods from unequally spaced observations. *The Astrophysical Journal*, 86 :619, 1981. 7, 155

[FMC17] F. Fereidoony, S.A. Mirtaheri, and S. Chamaani. M-Sequence Sensor and Continuous Basis Estimator for Wall Parameter Estimation Utilizing Through the Wall Sensing. *IEEE Sensors Journal*, 17(13) :4083–4091, 2017. 44, 46, 62

[FMHLG13] D. Foreman-Mackey, D.W. Hogg, D. Lang, and J. Goodman. emcee : the MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125(925) :306, 2013. 156

[For06] E.B. Ford. Improving the efficiency of Markov chain Monte Carlo for analyzing the orbits of extrasolar planets. *The Astrophysical Journal*, 642(1) :505, 2006. 156, 165

[Fos95] G. Foster. The cleanest Fourier spectrum. *The Astrophysical Journal*, 109 :1889–1902, 1995. 25

[Fuc04] J-J Fuchs. On sparse representations in arbitrary redundant bases. *IEEE Transactions on Information Theory*, 50(6) :1341–1344, 2004. 57

[Fuc05] J-J. Fuchs. Recovery of exact sparse representations in the presence of bounded noise. *IEEE Transactions on Information Theory*, 51(10) :3601–3608, 2005. 13

[GD73] D.F. Gray and K. Desikachary. A new approach to periodogram analyses. *The Astrophysical Journal*, 181 :523–530, 1973. 7, 156

[GILC11] D. Ge, J. Idier, and E. Le Carpentier. Enhanced sampling schemes for MCMC based blind Bernoulli-Gaussian deconvolution. *Signal Processing*, 91(4) :759–772, avril 2011. 96, 97, 103, 105, 106, 114, 117, 120, 121, 124, 127, 128, 144, 145, 199

[Gre95] P.J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4) :711–732, décembre 1995. 96, 110, 111, 162

[Gre05] P. C. Gregory. A Bayesian analysis of extrasolar planet data for HD 73526. *The Astrophysical Journal*, 631 :1198–1214, octobre 2005. 156

[Gre07a] P. C. Gregory. A Bayesian Kepler periodogram detects a second planets in HD 208487. *Monthly Notices of the Royal Astronomical Society*, 374(4) :1321–1333, février 2007. 156

[Gre07b] P. C. Gregory. A Bayesian periodogram finds evidence for three planets in HD 11964. *Monthly Notices of the Royal Astronomical Society*, 381(4) :1607–1616, novembre 2007. 156, 157

[Gre16] P.C. Gregory. An apodized Kepler periodogram for separating planetary and stellar activity signals. *Monthly Notices of the Royal Astronomical Society*, 458(3) :2604–2633, 2016. 161

[GW10] J. Goodman and J. Weare. Ensemble samplers with affine invariance. *Communications in applied mathematics and computational science*, 5(1) :65–80, 2010. 156

[Han92] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34(4) :561–580, 1992. 12

[Har17] N. Hara. *Outils pour l'analyse de données de vitesses radiales*. Thèse de doctorat, Université Pierre et Marie Curie-Paris VI, 2017. 150, 153, 160, 180

[Hay15] R.D. Haywood. *Hide and Seek : Radial-Velocity Searches for Planets around Active Stars*. PhD thesis, University of St Andrews, 2015. 150, 152, 160, 222

[HBFR] T.L. Hansen, M.A. Badiu, B.H. Fleury, and B.D. Rao. A sparse Bayesian learning algorithm with dictionary parameter estimation. In *2014 8th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. 44

[HBLC16] N.C. Hara, G. Boué, J. Laskar, and A.C.M. Correia. Radial velocity data analysis with compressed sensing techniques. *Monthly Notices of the Royal Astronomical Society*, 464(1) :1220–1246, 2016. 22, 157, 161, 163, 180

[HCCQ+14] R.D. Haywood, A. Collier Cameron, D. Queloz, SCC Barros, M. Deleuil, R. Fares, M. Gillon, A.F. Lanza, C. Lovis, C. Moutou, et al. Planets and stellar activity : hide and seek in the CoRoT-7 system. *Monthly Notices of the Royal Astronomical Society*, 443(3) :2517–2531, 2014. 157, 160, 161, 162, 165, 191

[HD10] C. Herzet and A. Drémeau. Bayesian pursuit algorithms. In *2010 18th European Signal Processing Conference*, pages 1474–1478. IEEE, 2010. 11

[Héb15] E. Hébrard. *Magnétométrie stellaire et imagerie Zeeman-Doppler appliquées à la recherche d'exoplanètes par mesures vélocimétriques*. Thèse de doctorat, Université Paul Sabatier – Toulouse III, 2015. 150, 158, 159, 219

[HFR18] T.L. Hansen, B.H. Fleury, and B.D. Rao. Superfast line spectral estimation. *IEEE Transactions on Signal Processing*, 66(10) :2511–2526, 2018. 25, 97, 198

[HST01] H. Haario, E. Saksman, and J. Tamminen. An adaptative Metropolis algorithm. *Bernoulli*, 7(2) :223–242, avril 2001. 156

[Jef61] H. Jeffreys. *Theory of Probability*. Oxford, England, Third edition, 1961. 162

[JTD11] P. Jain, A. Tewari, and I.S. Dhillon. Orthogonal matching pursuit with replacement. In *Advances in Neural Information Processing Systems*, pages 1215–1223, 2011. 11

[KHN15] G. Kail, F. Hlawatsch, and C. Novak. SMLR-Type Blind Deconvolution of Sparse Pulse Sequences Under a Minimum Temporal Distance Constraint. *IEEE Transactions on Signal Processing*, 63(18) :4838–4853, 2015. 19, 25

[KM82] J. Kormylo and J.M. Mendel. Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes. *IEEE Transactions on Information Theory*, 28 :482–488, 1982. 17, 18, 19, 97, 113

[KT09] M. Kowalski and B. Torrésani. Sparsity and persistence : mixed norms provide simple signal models with dependent coefficients. *Signal Image and Video Processing*, 3(3) :251–264, 2009. 12

[KTHD12] G. Kail, J. Tourneret, F. Hlawatsch, and N. Dobigeon. Blind Deconvolution of Sparse Pulse Sequences Under a Minimum Distance Constraint : A Partially Collapsed Gibbs Sampler Method. *IEEE Transactions on Signal Processing*, 60(6) :2727–2743, juin 2012. 19, 27, 97

[KYHP14] K.C. Knudson, J. Yates, A. Huk, and J.W. Pillow. Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 1215–1223. Curran Associates, Inc., 2014. 24, 28, 34, 39, 41, 42, 43, 65, 66, 67, 69

[KYO18] Y. Koyano, K. Yatabe, and Y. Oikawa. Infinite-dimensional svd for revealing microphone array's characteristics. *Applied Acoustics*, 129 :116–125, 2018. 69

[LB04] P. Lenz and M. Breger. Period04 : A software package to extract multiple frequencies from real data. *Proceedings of the International Astronomical Union*, 2004(IAUS224) :786–790, 2004. 25

[Lom76] N.R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and space science*, 39(2) :447–462, 1976. 7, 155

[LRS+09] A. Léger, D. Rouan, J. Schneider, P. Barge, M. Fridlund, B. Samuel, M. Ollivier, E. Guenther, M. Deleuil, H.J. Deeg, et al. Transiting exoplanets from the CoRoT space mission-VIII. CoRoT-7b : the first super-Earth with measured radius. *Astronomy and Astrophysics*, 506(1) :287–302, 2009. 191

[LWK94] J.S. Liu, W.H. Wong, and A. Kong. Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes. *Biometrika*, 81(1) :27–40, 1994. 113

[MBN19] R.B. Mhenni, S. Bourguignon, and J. Ninin. Global Optimization for Sparse Solutions to Least Squares Problems. Communication privée, 2019. 15, 198

[MCM+16] C. Meillier, F. Chatelain, O. Michel, R. Bacon, L. Piqueras, R. Bacher, and H. Ayasso. SELFI : an object-based, Bayesian method for faint emission line source detection in MUSE deep field data cubes. *Astronomy and Astrophysics*, 588 :A140, 2016. 4, 199

[MD16] D. Mo and M.F. Duarte. Compressive line spectrum estimation with clustering and interpolation. In *2016 Annual Conference on Information Science and Systems (CISS)*, pages 572–577. IEEE, 2016. 46

[MMB+08] C. Marois, B. Macintosh, T. Barman, B. Zuckerman, I. Song, J. Patience, D. Lafrenière, and R. Doyon. Direct imaging of multiple planets orbiting the star hr 8799. *Science*, 322(5906) :1348–1352, 2008. 149

[MMSC] H. Mortada, V. Mazet, C. Soussen, and C. Collet. Separation of delayed parameterized sources. In *Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO)*. 45, 46

[MPQ+03] M. Mayor, F. Pepe, D. Queloz, F. Bouchy, G. Rupprecht, G. Lo Curto, G. Avila, W. Benz, J.-L. Bertaux, X. Bonfils, T. Dall, H. Dekker, B. Delabre, W. Eckert, M. Fleury, A. Gilliotte, D. Gojak, J. C. Guzman, D. Kohler, J.-L. Lizon, A. Longinotti, C. Lovis, D. Megevand, L. Pasquini, J. Reyes, J.-P. Sivan, D. Sosnowska, R. Soto, S. Udry, A. van Kesteren, L. Weber, and U. Weilenmann. Setting New Standards with HARPS. *The Messenger*, 114 :20–24, décembre 2003. 153

[MQ95] M. Mayor and D. Queloz. A Jupiter-mass companion to a solar-type star. *Nature*, 378 :355–359, novembre 1995. 5, 148

[MZ93] S. Mallat and Z. Zhang. Matching Pursuits With Time-Frequency Dictionaries. *IEEE Transactions on Signal Processing*, 41(12), 1993. 10

[Nat95] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2) :227–234, 1995. 10

[NFB+18] B.E. Nelson, E.B. Ford, J. Buchner, R. Cloutier, R.F. Díaz, J.P. Faria, V.M. Rajpaul, and S. Rukdee. Quantifying the Evidence for a Planet in Radial Velocity Data. *arXiv preprint arXiv :1806.04683*, 2018. 29, 160, 162, 163, 170, 182, 183, 184, 185, 186, 187, 188, 199, 220, 222

[OPT00] M.R. Osborne, B. Presnell, and B.A. Turlach. A new approach to variable selection in least squares problems. *IMA journal of numerical analysis*, 20(3) :389–403, 2000. 12

[OSK94] M. S. O'Brien, A. N. Sinclair, and S. M. Kramer. Recovery of a sparse spike time series by L1 norm deconvolution. *IEEE Transactions on Signal Processing*, 42(12) :3353–3365, 1994. 8

[PCM+07] F. Pepe, A.C.M. Correia, M. Mayor, O. Tamuz, J. Couetdic, W. Benz, J-L. Bertaux, F. Bouchy, J. Laskar, C. Lovis, et al. The HARPS search for southern extra-solar

planets-VIII. $\mu$ arae, a system with four planets. *Astronomy and Astrophysics*, 462(2) :769–776, 2007. 156

[Per11] M. Perryman. *The exoplanet handbook*. Cambridge university press, 2011. 148, 150, 151, 152

[PPZ17] T. Passarin, D.R. Pipa, and M. Zibetti. A minimax dictionary expansion for sparse continuous reconstruction. In *Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO)*, pages 2136–2140. IEEE, 2017. 34, 41, 42, 44, 46, 66, 89, 93, 199, 217

[PRK93] Y.C. Pati, R. Rezaiifar, and P.S. Krishnaprasad. Orthogonal matching pursuit : Recursive function approximation with applications to wavelet decomposition. In *1993 Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pages 40–44. IEEE, 1993. 11

[PSC17] Y. Park, W. Seong, and Y. Choo. Compressive time delay estimation off the grid. *Journal of Acoustical Society America*, 141(6) :EL585–EL591, 2017. 22

[QHS+01] D. Queloz, G.W. Henry, J.P. Sivan, S.L. Baliunas, J.L. Beuzit, R.A. Donahue, M. Mayor, D. Naef, C. Perrier, and S. Udry. No planet for HD 166435. *Astronomy and Astrophysics*, 379(1) :279–287, 2001. 158

[RAO+15] V. Rajpaul, S. Aigrain, M.A. Osborne, S. Reece, and S. Roberts. A Gaussian process framework for modelling stellar activity signals in radial velocity data. *Monthly Notices of the Royal Astronomical Society*, 452(3) :2269–2291, 2015. 158, 160, 161

[RAR16] V. Rajpaul, S. Aigrain, and S. Roberts. Ghost in the time series : no planet for Alpha Cen B. *Monthly Notices of the Royal Astronomical Society*, 456 :L6–L10, février 2016. 156

[RLD87] D.H. Roberts, J. Lehár, and J.W. Dreher. Time series analysis with Clean-Part one-Derivation of a spectrum. *The Astrophysical Journal*, 93 :968, 1987. 7

[Rob96] C.P. Robert. *Méthodes de Monte Carlo par chaînes de Markov*. Economica, 1996. 109, 176

[RW06] C.E. Rasmussen and C.K.I. Williams. *Gaussian process for machine learning*. MIT press, 2006. 160

[S+06] J. Skilling et al. Nested sampling for general Bayesian computation. *Bayesian analysis*, 1(4) :833–859, 2006. 157

[SBA+18] A. Santerne, B. Brugger, D. J. Armstrong, V. Adibekyan, J. Lillo-Box, H. Gosselin, A. Aguichine, J.-M. Almenara, D. Barrado, S. C. C. Barros, D. Bayliss, I. Boisse, A. S. Bonomo, F. Bouchy, D. J. A. Brown, M. Deleuil, E. Delgado Mena, O. Demangeon, R. F. Díaz, A. Doyle, X. Dumusque, F. Faedi, J. P. Faria, P. Figueira, E. Foxell, H. Giles, G. Hébrard, S. Hojjatpanah, M. Hobson, J. Jackman, G. King, J. Kirk, K. W. F. Lam, R. Ligi, C. Lovis, T. Louden, J. McCormac, O. Mousis, J. J. Neal, H. P. Osborn, F. Pepe, D. Pollacco, N. C. Santos, S. G. Sousa, S. Udry, and A. Vigan. An Earth-sized exoplanet with a Mercury-like composition. *Nature Astronomy*, 2 :393–400, mai 2018. 149

[SBFA15] E. Soubies, L. Blanc-Féraud, and G. Aubert. A Continous Exact l0 penalty (CEL0) for least squares regularized problem. *SIAM Journal on Imaging Sciences*, 8 :1574–1606, 2015. 13

[SBFA17] E. Soubies, L. Blanc-Féraud, and G. Aubert. A unified view of exact continuous penalties for \ell_2-\ell_0 minimization. *SIAM Journal of Optimization*, 27(3) :2034–2060, 2017. 13

[Sca82] J.D. Scargle. Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263 :835–853, 1982. 7, 155

[SIBD11] C. Soussen, J. Idier, D. Brie, and J. Duan. From Bernoulli-Gaussian deconvolution to sparse signal restoration. *IEEE Transactions on Signal Processing*, 59(10) :4572–4584, octobre 2011. 11, 17, 18, 174

[Ste00] M. Stephens. Bayesian analysis of mixture models with an unknown number of components-an alternative to reversible jump methods. *Annals of Statistics*, pages 40–74, 2000. 162

[Sul17] S. Sulis. *Statistical methods using hydrodynamic simulations of stellar atmospheres for detecting exoplanets in radial velocity data.* Thèse de doctorat, Université Côte d'Azur, 2017. 150

[TAEG+13] M. Tuomi, G. Anglada-Escudé, E. Gerlach, H. R. A. Hones, A. Reiners, E. J. Rivera, S. S. Vogt, and R. P. Butler. Habitable-zone super-Earth cadidate in a six-planet system around the K2.5V star HD 40307. *Astronomy and Astrophysics*, 549 :A48, janvier 2013. 156, 171, 189, 190, 220, 222

[TBR13] G. Tang, B. Narayan Bhaskar, and B. Recht. Sparse recovery over continuous dictionaries-just discretize. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 1043–1047. IEEE, 2013. 22, 157

[TBSR13] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *IEEE Transactions on Information Theory*, 59(11) :7465–7490, 2013. 21, 22, 157

[TJJ+13] M. Tuomi, HRA Jones, J.S. Jenkins, C.G. Tinney, R.P. Butler, S.S. Vogt, J.R. Barnes, R.A. Wittenmyer, S. O'Toole, J. Horner, et al. Signals embedded in the radial velocity noise-Periodic variations in the $\tau$ Ceti velocities. *Astronomy and Astrophysics*, 551 :A79, 2013. 161

[Tro04] J.A. Tropp. Greed is good : algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10), 2004. 10, 13

[Tro06] J.A. Tropp. Just relax : Convex programming methods for identifying sparse signals in noise. *IEEE Transactions on Information Theory*, 52(3) :1030–1051, 2006. 13

[Tuo12] M. Tuomi. Evidence for 9 planets in the HD 10180 system. *Astronomy and Astrophysics*, 543 :A52, juillet 2012. 156, 157, 161, 162, 167

[TW10] J. A. Tropp and S. J. Wright. Computational methods for sparse solution of linear inverse problems. *Proceedings of the IEEE*, 98(6) :948–958, 2010. 11

[vDP08] D. A. van Dyk and T. Park. Partially Collapsed Gibbs Samplers : Theory and Methods. *Journal of American Statistical Association*, 103 :790–796, 2008. 28, 96, 114, 127, 128

[VR01] M. C. W. Van Rossum. A novel spike distance. *Neural Computation*, 13(4) :751–763, 2001. 32

[Web] Site Web. https ://www.quora.com/What-are-the-most-prominent-methods-for-detecting-exoplanets. 149

[WH09] J.T. Wright and A.W. Howard. Efficient fitting of multiplanet keplerian models to radial velocity and astrometry data. *Astronomy and Astrophysics Supplement Series*, 182(1) :205, 2009. 165

[WS01] R. Waagepetersen and D. Sorensen. A tutorial on Reversible Jump MCMC with a view toward applications in QTL-mapping. *International Statistical Review*, 69(1) :49–61, avril 2001. 111

[Zha11] T. Zhang. Adaptive forward-backward greedy algorithm for learning sparse representations. *IEEE Transactions on Information Theory*, 57(7) :4689–4708, 2011. 11

[ZK09] M. Zechmeister and M. Kürster. The generalised Lomb-Scargle periodogram-a new formalism for the floating-mean and Keplerian periodograms. *Astronomy and Astrophysics*, 496(2) :577–584, 2009. 155, 156, 157

## Résumé

Cette thèse s'intéresse à la résolution de problèmes inverses non linéaires exploitant un a priori de parcimonie ; plus particulièrement, des problèmes où les données se modélisent comme la combinaison linéaire d'un faible nombre de fonctions non linéaires en un paramètre dit de « localisation » (par exemple la fréquence en analyse spectrale ou le décalage temporel en déconvolution impulsionnelle). Ces problèmes se reformulent classiquement en un problème d'approximation parcimonieuse linéaire (APL) en évaluant les fonctions non linéaires sur une grille de discrétisation arbitrairement fine du paramètre de localisation, formant ainsi un « dictionnaire discret ». Cependant, une telle approche se heurte à deux difficultés majeures. D'une part, le dictionnaire provenant d'une telle discrétisation est fortement corrélé et met en échec les méthodes de résolution sous-optimales classiques comme la pénalisation L1 ou les algorithmes gloutons. D'autre part, l'estimation du paramètre de localisation, appartenant nécessairement à la grille de discrétisation, se fait de manière discrète, ce qui entraîne une erreur de modélisation. Dans ce travail nous proposons des solutions pour faire face à ces deux enjeux, d'une part via la prise en compte de la parcimonie de façon exacte en introduisant un ensemble de variables binaires, et d'autre part via la résolution « optimale » de tels problèmes sur « dictionnaire continu » permettant l'estimation continue du paramètre de localisation. Deux axes de recherches ont été suivis, et l'utilisation des algorithmes proposés est illustrée sur des problèmes de type déconvolution impulsionnelle et analyse spectrale de signaux irrégulièrement échantillonnés. Le premier axe de ce travail exploite le principe « d'interpolation de dictionnaire », consistant en une linéarisation du dictionnaire continu pour obtenir un problème d'APL sous contraintes. L'introduction des variables binaires nous permet de reformuler ce problème sous forme de « programmation mixte en nombres entiers » (Mixed Integer Programming – MIP) et ainsi de modéliser de façon exacte la parcimonie sous la forme de la « pseudo-norme L0 ». Différents types d'interpolation de dictionnaires et de relaxation des contraintes sont étudiés afin de résoudre de façon optimale le problème grâce à des algorithmes classiques de type MIP. Le second axe se place dans le cadre probabiliste Bayésien, où les variables binaires nous permettent de modéliser la parcimonie en exploitant un modèle de type Bernoulli-Gaussien. Ce modèle est étendu (modèle BGE) pour la prise en compte de la variable de localisation continue. L'estimation des paramètres est alors effectuée à partir d'échantillons tirés avec des algorithmes de type Monte Carlo par Chaîne de Markov. Plus précisément, nous montrons que la marginalisation des amplitudes permet une accélération de l'algorithme de Gibbs dans le cas supervisé (hyperparamètres du modèle connu). De plus, nous proposons de bénéficier d'une telle marginalisation dans le cas non supervisé via une approche de type « Partially Collapsed Gibbs Sampler. » Enfin, nous avons adapté le modèle BGE et les algorithmes associés à un problème d'actualité en astrophysique : la détection d'exoplanètes par la méthode des vitesses radiales. Son efficacité sera illustrée sur des données simulées ainsi que sur des données réelles.



**Mots clés** : Problèmes inverses, Approximations parcimonieuses, Variables binaires, Dictionnaire continu, Modèle Bernoulli-Gaussien, Estimation



## Abstract

This thesis deals with solutions of nonlinear inverse problems using a sparsity prior ; more specifically when the data can be modelled as a linear combination of a few functions, which depend non-linearly on a "location" parameter, i.e. frequencies for spectral analysis or time-delay for spike train deconvolution. These problems are generally reformulated as linear sparse approximation problems, thanks to an evaluation of the nonlinear functions at location parameters discretised on a thin grid, building a "discrete dictionary". However, such an approach has two major drawbacks. On the one hand, the discrete dictionary is highly correlated ; classical sub-optimal methods such as L1-penalisation or greedy algorithms can then fail. On the other hand, the estimated location parameter, which belongs to the discretisation grid, is necessarily discrete and that leads to model errors. To deal with these issues, we propose in this work an exact sparsity model, thanks to the introduction of binary variables, and an optimal solution of the problem with a "continuous dictionary" allowing a continuous estimation of the location parameter. We focus on two research axes, which we illustrate with problems such as spike train deconvolution and spectral analysis of unevenly sampled data. The first axis focusses on the "dictionary interpolation" principle, which consists in a linearisation of the continuous dictionary in order to get a constrained linear sparse approximation problem. The introduction of binary variables allows us to reformulate this problem as a "Mixed Integer Program" (MIP) and to exactly model the sparsity thanks to the "pseudo-norm L0". We study different kinds of dictionary interpolation and constraints relaxation, in order to solve the problem optimally thanks to MIP classical algorithms. For the second axis, in a Bayesian framework, the binary variables are supposed random with a Bernoulli distribution and we model the sparsity through a Bernoulli-Gaussian prior. This model is extended to take into account continuous location parameters (BGE model). We then estimate the parameters from samples drawn using Markov chain Monte Carlo algorithms. In particular, we show that marginalising the amplitudes allows us to improve the sampling of a Gibbs algorithm in a supervised case (when the model's hyperparameters are known). In an unsupervised case, we propose to take advantage of such a marginalisation through a "Partially Collapsed Gibbs Sampler." Finally, we adapt the BGE model and associated samplers to a topical science case in Astrophysics : the detection of exoplanets from radial velocity measurements. The efficiency of our method will be illustrated with simulated data, as well as actual astrophysical data.



**Keywords** : Inverse problems, Sparse Approximations, Binary variables, Continuous dictionary, Bernoulli-Gaussian model, Estimation