UNIVERSITÉ DE TOULON

ÉCOLE DOCTORALE 548 — MER ET SCIENCES

IM2NP, UMR CNRS 7334, équipe Signal & Tracking

THÈSE

présentée par :

Léontine SÉGAL

soutenue le : **24 octobre 2025**

pour obtenir le grade de Docteur en Automatique, Signal, Productique, Robotique

Vers une estimation robuste des propriétés physico-chimiques des nuages moléculaires interstellaires en l'absence de vérité terrain

Thèse dirigée par :

M. Antoine ROUEFF	Professeur des universités, Université de Toulon
M. Jérôme PETY	Astronome, IRAM/LUX

Jury :

M. Jean-Yves TOURNERET, président	Professeur des universités, INP-ENSEEIHT
Mme. Frédérique CONRUYT-MOTTE, rapportrice	Directrice de recherche HDR, IPAG
M. Loïc DENIS, rapporteur	Professeur des universités, Université Jean Monnet
Mme. Adeline PAIEMENT-BONNEAU, examinatrice	Maître de conférences HDR, Université de Toulon
Mme. Romane LE GAL, examinatrice	Astronome adjointe, IPAG
Mme. Maryvonne GERIN, invitée	Directrice de recherche HDR, LUX

im2np cnrs anr iram Institut de Radioastronomie Millimétrique ORION-B

Contents

Remerciements **6**

Acronyms **8**

Notation **9**

List of Figures **11**

List of Tables **12**

1 Introduction **13**
1.1 Contexte scientifique et objectives de la thèse . . . 13
1.2 Scientific context and thesis objectives . . . 15
References . . . 17

2 Radio observations of molecular clouds **19**
2.1 Molecular clouds . . . 19
2.1.1 From dark to molecular clouds . . . 19
2.1.2 The Orion B giant molecular cloud (GMC) . . . 20
2.1.3 Cloud constituents and observations . . . 20
2.2 The ORION-B dataset . . . 24
2.3 Rotational spectrum analysis . . . 29
2.3.1 Collisional processes . . . 29
2.3.2 Radiative processes . . . 30
2.4 Different molecular tracers as complementary probes of the gas . . . 33
References . . . 36

3 Statistical inference **39**
3.1 Introduction to estimation theory . . . 40
3.2 Cramér-Rao bound (CRB) . . . 45
3.2.1 Introduction . . . 45
3.2.2 Definition and role in estimation theory . . . 45
3.2.3 Practical use of the CRB . . . 48
3.3 Maximum likelihood estimator (MLE) . . . 49
3.3.1 Asymptotic properties and implementation . . . 49
3.3.2 Performance analysis of the MLE on simulated dataset . . . 51
3.3.3 From simulated data to observation analysis . . . 51
References . . . 53

Appendix 3 . . . 54
3.A Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model . . . 54

4 A heterogeneous cloud model to probe dense cores **93**
4.1 A sandwich to model dense core line of sight . . . 94
4.1.1 Model statement . . . 94
4.1.2 Application to the Horsehead nebula . . . 95
4.2 Influence of the spectral sampling on the estimation precision . . . 96
4.2.1 Experiment setup . . . 96
4.2.2 On the S/N saturation . . . 97
4.2.3 Impact of the sampling phase on the CRB . . . 98
4.2.4 Impact of the channel spacing on the CRB . . . 99
References . . . 105
Appendix 4 . . . 105
4.A Change of variables on the CRB . . . 105
4.B Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead Nebula . . . 106

5 Parameter estimation under multiplicative noise **137**
5.1 Measurement model and statistical description of the multiplicative noise . . . 139
5.1.1 Measurement model . . . 139
5.1.2 Specifying the multiplicative noise . . . 139
5.2 Considering the multiplicative noise as unknown and deterministic . . . 141
5.2.1 FIM and CRB . . . 142
5.3 Considering the multiplicative noise as random . . . 145
5.3.1 Hybrid approach . . . 145
5.3.1.1 FIM and Hybrid CRB (HCRB) . . . 146
5.3.2 Marginal approach . . . 148
5.3.2.1 FIM and marginal CRB . . . 149
5.4 Neglecting the multiplicative noise . . . 149
5.4.1 Likelihood under the misspecified model . . . 150
5.4.2 Knowledgeable case . . . 150
5.4.2.1 Asymptotic behavior of the MLE . . . 150
5.4.2.2 Unbiasedness of the MLE . . . 151
5.4.2.3 Misspecified CRB (MCRB) . . . 152
5.4.3 Naive case . . . 152
5.5 Quantitative comparison of the different CRBs in the case of a sight line toward a simulated dense core . . . 153
5.6 Generalizing the comparison on actual GMC observations . . . 157
5.7 Summary and conclusions . . . 157
References . . . 158
Appendix 5 . . . 160
5.A Computation of the Kullback-Leibler divergence (KLD) . . . 160
5.B Derivative of composite function . . . 160

5.C Calculation of the FIM under the hybrid approach 161
5.D Sherman-Morrison formula 163
5.E Calculation of the MCRB 163
5.F Quantification of the precision loss under the misspecified model . . . 165

6 Conclusion and perspectives **171**
6.1 On the road to a reliable estimate of the physicochemical characteristics of interstellar molecular clouds, despite the lack of reference data 171
6.2 Perspectives 173
6.3 Vers des estimations fiables des caractéristiques physico-chimiques des nuages interstellaires en l'absence de vérité terrain 175
6.4 Perspectives 177

Contributions **179**

Abstract & Résumé **181**

Remerciements

Je remercie chaleureusement l'ensemble des membres du jury pour avoir accepté d'évaluer ce travail. Merci à Loïc Denis, professeur des universités au laboratoire Hubert Curien (Université de Saint-Etienne) et à Frédérique Motte, directrice de recherche au CNRS à l'Institut de Planétonologie et d'Astrophysique de Grenoble (IPAG), pour l'avoir rapporté. Merci à Romane Le Gal, astronome adjointe à l'IPAG et à Adeline Paiement, maître de conférence au Laboratoire d'Informatique et des Systèmes (LIS), pour l'avoir examiné. Merci à Jean-Yves Tourneret, professeur des université à l'ENSEEIHT pour avoir présidé ce jury. Merci à vous tous d'avoir témoigné tant de bienveillance et d'intérêt pour ces travaux lors de nos échanges.

Cette thèse fût menée au sein d'ORION-B[1], équipe d'astrophysiciens et chercheurs en traitement de l'information.

Je souhaite exprimer ma plus profonde gratitude envers mes deux directeurs de thèse, Antoine Roueff, professeur des universités à l'IM2NP (Université de Toulon), et Jérôme Pety, astronome à l'Institut de RadioAstronomie Millimétrique de Grenoble (IRAM), qui ont rendu possible de réaliser cette thèse interdisciplinaire avec l'encadrement le plus épanouissant que je n'aurais imaginé pouvoir recevoir. Je vous remercie de m'avoir donné la chance de vivre cette aventure. Apprendre et travailler à vos côtés, et être ainsi témoin de votre curiosité insatiable, de votre grande sensibilité, ainsi que de votre discernement, m'est inestimable.

Je remercie chaleureusement l'ensemble des membres du consortium pour m'avoir intégrée si vite parmi eux, et ainsi me faire entrevoir les horizons fascinants de leurs domaines d'études respectifs.

Mes pensées vont tout particulièrement à Maryvonne Gérin, directrice de recherche au LUX, qui m'a guidée avec bienveillance durant cette thèse, et à Evelyne Roueff, professeur émérite à l'Observatoire de Paris, qui, toutes deux, par leur regard et expertise, notamment en chimie, ont guidé ces travaux et m'ont permis de mieux comprendre, et, ainsi, m'émerveiller face à certains de nos résultats.

I warmly thank Javier Goicoechea, researcher at the Instituto de Física Fundamental (CSIC), for having tought me radiative transfer theory and for having, ever since, carefully followed my research work. Discovering the intricate mechanisms of our universe through your enthusiastic and passionate eyes has made me want to delve into it even further.

Mon affection va enfin à mes compères doctorants et post-doctorants Jan, Miriam, Pierre,

[1]Malheureusement pas dans le nuage interstellaire à proprement parler, mais au sein de la collaboration éponyme.

Lucas, Ivana, Antoine Z. et Helena. Lors de nos grandes réunions semestrielles, votre bienveillance, votre bonne humeur et vos conseils m'ont toujours redonné du baume au cœur et permis de tracer de plus belles perspectives pour la suite.

Cette thèse s'est déroulée au sein de l'équipe Signal & Tracking (STr) de l'IM2NP, et ponctuée par de brefs séjours à l'IRAM.

Je suis profondément reconnaissante envers l'IRAM, et notamment son directeur Karl Schuster, pour avoir accepté de financer ma thèse. Je remercie les membres de cet institut qui, malgré mes venues rares et fugaces, m'ont toujours accueillie avec chaleur.

Mes pensées vont désormais aux membres de ce qui fût comme ma deuxième famille durant ces trois ans, l'équipe STr : Gilles et Jean, Claude et Annie-Claude, Bruno et Bernard, et mes consœurs doctorantes Lisa et Lucie. Je ne m'attendais pas à rencontrer tant de personnes qui, en si peu de temps, deviendraient si importantes dans ma vie. Merci pour votre joie de vivre et les souvenirs que j'emporte avec moi; Notamment des plantes (le bégonia bambou, le patchouli, le doliprane et le petit thym), des chats (Gros Mouchou, Mira et Masha), des tortues, de notre "stagiaire" et mascotte Moka, des livres et du Métal Hurlant...

Merci à Valérie et Mina pour vos éclats de rires contagieux, votre générosité et "petites" attentions au quotidien, qui, j'en suis certaine, y sont pour beaucoup dans la réussite de nombreuses thèses menées à l'IM2NP.

Je remercie chaleureusement Jérôme Mars, professeur des universités au GIPSA-LAB (ENSE3), qui fût mon enseignant de géosciences en école d'ingénieurs, qui depuis, continue de me donner de précieux conseils et de son temps avec beaucoup de gentillesse.

Je remercie mon professeur d'imagerie hyperspectrale, Jocelyn Chanussot, directeur de recherches au sein de l'INRIA, qui, lorsque je lui ai exprimé mon souhait de réaliser mon stage de fin d'études dans le domaine de l'astrophysique, m'orienta vers Jérôme et Antoine.

Mes pensées vont enfin à ma famille. À votre amour du théâtre, de la musique, du cirque, et du dessin, qui me donne un tel souffle pour toute chose que j'entreprends. Vous observer écrire, écouter, répéter, *jouer*, c'est-à-dire apprendre continuellement, constitue indéniablement ce qui a rendu pour moi passion et travail acharné indissociable.

Mes chers grands-parents, Jeanne et Alain, je vous remercie de m'avoir donné la joie immense d'assister à ma soutenance. Mon tendre papi et Docteur Ségal, à travers ta présence notamment, il y a avait aussi Sacha et Jankel.

Mes plus tendres pensées à mes grands-parents et arrière-grands-parents, que j'aime ou que j'aurais voulu connaître.

À Mo, Godot et Blaise.

Acronyms

CMB Cosmic Microwave Background
CRB Cramér-Rao Bound
FWHM Full Width at Half Maximum
FIM Fisher Information Matrix
GMC Giant Molecular Cloud
HCRB Hybrid Cramér-Rao Bound
IR InfraRed
IRAM Institut de RadioAstronomie Millimétrique
IRAM-30m The IRAM 30-meter telescope
ISM Interstellar Medium
KLD Kullback-Leibler Divergence
LS(E) Least Squares (Estimator)
MCRB Misspecified Cramér-Rao Bound
ML(E) Maximum Likelihood (Estimator)
MSE Mean Square Error
MVU(E) Minimum Variance unbiased (Estimator)
NLL Negative Log-Likelihood
ORION-B Outstanding Radio-Imaging of OrioN B
p.d.f. Probability Density Function
PDR PhotoDissociation Region
PI Principal Investigator
S/N Signal-to-Noise Ratio
UV UltraViolet
WHISPER Wideband High-resolution IRAM-30m Surveys at two Positions with Emir Receivers

Notation

Physical constants

h The Planck constant, $6.626070150 \times 10^{-34}$ J Hz^{-1}.
k The Boltzmann constant, $1.380648520 \times 10^{-23}$ J/K.
c Speed of light, 2.99792458×10^{5} km s^{-1}.

Physical quantities

A_V Visual extinction, expressed in magnitude (mag).
C_V Centroid velocity (km s^{-1}).
dV Velocity resolution (km s^{-1}).
$n_{\rm col}$ Volume density of colliders, expressed in molecule per cm^3 (cm^{-3}).
$n_{\rm crit}$ Critical volume density (cm^{-3}).
$n_{\rm crit}^{\rm eff}$ Effective critical volume density (cm^{-3}).
$n_{\rm H_2}$ Volume density of H_2 (cm^{-3}).
N(X) Column density of the species X (cm^{-2}).
$T_{\rm b}$ Brightness temperature, expressed in Kelvin (K).
$T_{\rm CMB}$ Temperature of the Cosmic Microwave Background (CMB), ~ 2.73 K.
$T_{\rm ex}$ Excitation temperature (K).
$T_{\rm kin}$ Kinetic temperature (K).
$T_{\rm peak}$ Peak intensity (K).
τ Opacity or optical depth (no unit).
σ_V Velocity dispersion (km s^{-1}).
$d\nu$ Frequency resolution (kHz).

Mathematical notation

$\boldsymbol{I}_N$ Identity matrix of size N.
$\boldsymbol{0}_N$ Vector of zeros of size N.
exp(.) Exponential function.
ln(.) Natural logarithm, i.e., $\forall y \in \mathbb{R}$, $y = \ln(\exp y)$.
log(.) Decimal logarithm, i.e., $\forall y \in \mathbb{R}^{+*}$, $\log(y) = \ln(y)/\ln 10$.
$.^\top$ Transposition function.
$\det(\boldsymbol{A})$ Determinant of the matrix $\boldsymbol{A}$.
$\mathrm{tr}(\boldsymbol{A})$ Trace of the matrix $\boldsymbol{A}$.
$\mathcal{N}(\mu, \sigma^2)$ Gaussian distribution with a mean of μ and a variance of σ^2.
Log-$\mathcal{N}(\mu, \sigma^2)$ Lognormal distribution with μ and σ^2 such that if $X \sim$ Log-$\mathcal{N}(\mu, \sigma^2)$, then $\ln X \sim \mathcal{N}(\mu, \sigma^2)$.

$p_{\boldsymbol{X}}$ Probability density function (p.d.f.) of $\boldsymbol{X}$.
$\mathbb{E}_{\boldsymbol{X}}$ Statistical expectation over $\boldsymbol{X}$ defined by $\mathbb{E}_{\boldsymbol{X}}\left[f(\boldsymbol{X})\right] = \int f(\boldsymbol{x})\, p_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$. Simply denoted by $\mathbb{E}$ when the context avoids any ambiguity.
$\mathbb{V}[\boldsymbol{X}]$ Variance of X defined by $\mathbb{V}\left[X\right] = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big]$.
$\boldsymbol{\mathcal{I}}$ Fisher information matrix.
$\boldsymbol{\mathcal{B}}$ Cramér-Rao bound.

Variables
$\boldsymbol{Y} \in \mathbb{R}^L$ The set of L molecular lines. Written in lower case, $\boldsymbol{y}$ denotes a particular realization of the random vector $\boldsymbol{Y}$.
$\boldsymbol{X} \in \mathbb{R}^N$ A molecular line (in K) belonging to $\boldsymbol{Y}$, composed of N velocity channels. Written in lower case, $\boldsymbol{x}$ denotes a measured spectrum, a particular realization of $\boldsymbol{X}$.
$\boldsymbol{s} \in \mathbb{R}^N$ Physical model of molecular line parameterized by $\boldsymbol{\theta}$.
$\boldsymbol{B} \in \mathbb{R}^N$ The additive noise (in K) such that $\boldsymbol{B} = (B_1 \cdots B_N)^\top$ where $B \in \boldsymbol{B}$ are independent white Gaussian noises.
$C \in \mathbb{R}$ The calibration noise (no unit). Written in lower case, c denotes the calibration noise when it treated as a deterministic parameter.
$\boldsymbol{\theta} \in \mathbb{R}^M$ The deterministic parameter to estimate made of M physical, chemical and kinematic parameters.
$\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}$ Maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$. Simply denoted by $\widehat{\boldsymbol{\theta}}$ when the context is unambiguous.
$\boldsymbol{\theta}_{\mathrm{tr}} \in \mathbb{R}^M$ The true value of $\boldsymbol{\theta}$.
$\boldsymbol{\theta}_{\mathrm{ps}} \in \mathbb{R}^M$ The pseudo-true value of $\boldsymbol{\theta}$, i.e., the solution of $\widehat{\boldsymbol{\theta}}$ obtained for a misspecified measurement model.
$\sigma_b^2 \boldsymbol{I}_N \in \mathbb{R}^{N \times N}$ Covariance matrix of the additive noise $\boldsymbol{B}$.
$\sigma_c^2 \in \mathbb{R}$ Variance of the calibration noise C. One has $\sigma_c^2 = \{0.05, 0.1\}$ for molecular lines at the 3 mm and 1 mm atmospheric window, respectively.
$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} \in \mathbb{R}^{M \times M}$ Covariance matrix of the estimator $\widehat{\boldsymbol{\theta}}$.
$\mu_c \in \mathbb{R}$ Mean of the calibration noise C. In this thesis, $\mu_c = 1$ regardless of the considered molecular line.
$\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}} \in \mathbb{R}^M$ Mean of the estimator $\widehat{\boldsymbol{\theta}}$.

List of Figures

2.1 An overview of the Orion complex in the visible light . . . 21
2.2 Molecular lines measured toward the Horsehead Nebula from the WHISPER dataset at the 3 mm band . . . 23
2.3 Electromagnetic spectrum diagram . . . 24
2.4 Spectrum of $C^{18}O(1-0)$ toward a dense core of the Horsehead Nebula . . . 25
2.5 Representation of the three frequency tunings observed within the ORION-B project . . . 26
2.6 Two representations of the $C^{18}O(1-0)$ line cube towards the Horsehead Nebula 27
2.7 A summary of the ORION-B dataset . . . 28
2.8 Illustration of an optically thick and thin molecular line . . . 31
2.9 variations with respect to the ^{12}CO column density of the brightness temperature (T_b), the T_{ex}/T_{kin} ratio, the opacity (τ), and the n_{H_2}/n_{crit}^{eff} ratio . . . 34
2.10 variations with respect to the HCO^+ column density of the brightness temperature (T_b), the T_{ex}/T_{kin} ratio, the opacity (τ), and the n_{H_2}/n_{crit}^{eff} ratio . . . 35
2.11 Ratios of the brightness temperature of the ratio $^{12}CO(2-1)$ over $^{12}CO(1-0)$ and of $HCO^+(3-2)/HCO^+(1-0)$ as a function of the kinetic temperature and volume density of the gas . . . 36

3.1 Illustration of the dependence of the probability density function (p.d.f.) on the unknown parameter . . . 41
3.2 Illustration of the characteristics of the estimator's performance in the Gaussian scalar case . . . 44
3.3 Illustration of the geometrical interpretation of the components of the CRB matrix in the 2D case . . . 47
3.4 Illustration of the two-step procedure implemented to estimate the MLE . . . 50
3.5 Illustration of the estimator's performance analysis through Monte Carlo experiment on a simulated dataset . . . 52

4.1 Sandwich model of the line of sight . . . 94
4.2 Illustration of the sensitivity of the estimation precision to the phase sampling through the $HCO^+(1-0)$ line . . . 100
4.3 Degradation of the estimation precision in function of the frequency channel spacing . . . 101
4.4 Degradation of the estimation precision in function of the frequency channel spacing (focus between 20 and 100 kHz) . . . 101
4.5 Illustration of the frequency sampling on the molecular line profiles . . . 104

5.1 Comparison between the lognomal and normal distribution 141
5.2 Composite image of the Horsehead Nebula from the ORION-B dataset and spectra toward the dense core "DC 2" 154
5.3 Comparison of the different CRBs of the logarithm of the kinetic temperature on simulated dataset 155
5.4 Comparison of the different CRBs of the logarithm of the volume density on simulated dataset 156

List of Tables

2.1 Critical densities of ^{12}CO and of HCO^+ lines at $T_{kin} = 15\,K$ 34

4.1 Parameters used to create the dataset for the experiment of Sect. 4.2 97
4.2 Characteristics of the studied molecular lines in the experiment of Sect. 4.2 . 97
4.3 Velocity dispersion (σ_V) of the sandwich converted in frequency 98
4.4 Mean and standard deviation of the degradation of the estimation precision at the spectral sampling of 60 kHz for all the molecular lines 102
4.5 Mean and standard deviation of the degradation of the estimation precision at the spectral sampling of 120 kHz for all the molecular lines 102
4.6 Mean and standard deviation of the degradation of the estimation precision at the spectral sampling of 240 kHz for all the molecular lines 103

5.1 Parameters used to generate a dataset with a dense core 154
5.2 Characteristics of the studied molecular lines used to generate the dataset . . 154

Chapter 1

Introduction

Contents

1.1 Contexte scientifique et objectives de la thèse **13**
1.2 Scientific context and thesis objectives **15**
References **17**

1.1 Contexte scientifique et objectives de la thèse

Contexte scientifique — Les étoiles naissent au sein des régions denses et froides des nuages moléculaires géants, structures les plus massives du milieu interstellaire. Ces régions, appelées *cœurs denses* et caractérisées par des densités typiques de 10^{4-6} particules par cm^3 à ~ 10 K (Draine 2011, Chapter 32; Bergin and Tafalla 2007), ne sont qualifiées de "denses" que relativement aux autres phases du milieu interstellaire, et demeurent trop diffuses pour pouvoir être étudiées en laboratoire, où l'ultravide est atteint pour une densité de $10^7\,\mathrm{cm}^{-3}$ à 300 K (V. V. Rao *et al.* 1998). L'étude des processus physico-chimiques impliqués dans la formation stellaire au sein des cœurs denses repose ainsi essentiellement sur leur *observation*. Notre compréhension de la formation des étoiles progresse à mesure que les méthodes d'acquisition et d'analyse de ces observations se perfectionnent. Les progrès technologiques réalisés au cours des dernières décennies permettent d'acquerir des jeux de données dont les champs de vue s'étendent sur plusieurs parsecs[1] et cela, à haute résolution spatiale et spectrale. Le Large Program ORION-B[2] (PIs: Jérôme Pety[3] & Maryvonne Gerin[4]) a sondé avec le télescope de 30 m de l'IRAM[5] (IRAM-30m) le nuage d'Orion B sur un champ de vue de $5\,\mathrm{deg}^2$ avec une résolution angulaire de $\sim 25''$. Cette campagne d'observation a produit un cube de données composé de 1074×758 pixels (ou lignes de visée), chacun composé de $\sim 200\,000$ canaux fréquentiels couvrant un bande de 72 à 116 GHz. Analyser une telle quantité de données nécessite de développer des méthodes statistiques spécifiques, tenant

[1] Un parsec (pc) correspond à la distance parcourue par la lumière en ~ 3.26 années.
[2] Outstanding Radio-Imaging of OrioN B
[3] IRAM, Grenoble (France) and LUX, Paris (France)
[4] LUX, Paris (France)
[5] Institut de RadioAstronomie Millimétrique (France, Espagne, Allemagne)

compte des connaissances physiques et chimiques du milieu. Un premier enjeu abordé dans cette thèse est de réaliser un traitement faiblement supervisé de ces données multi-pixels. De plus, l'observation du nuage dans une large bande spectrale implique que plusieurs espèces moléculaires peuvent être détectées en chacun des pixels. Un tel spectre est qualifié de *multi-raies*. Le second enjeu consiste à développer des outils d'analyse permettant d'extraire l'information contenue dans ces observations multi-raies.

Le consortium ORION-B — Le consortium ORION-B fût créé en 2014 pour relever ces défis à travers la collaboration d'astrophysiciens, d'astrochimistes et de scientifiques des données. Depuis sa création, le consortium a réuni l'expertise de physiciens de l'IRAM, du LUX, LAB[6], CSIC[7], IRAP[8], JPL[9], des universités de Cardiff et UC Chile[10], ainsi que des statisticiens de l'institut Fresnel[11] et des laboratoires IM2NP[12], GIPSA-LAB[13], INRIA[14], et CRIStAL[15]. Le jeu de données ORION-B acquis entre 2013 et 2020 nous sert de modèle pour développer des techniques d'analyse d'observations de nuages moléculaires géants. Les sujets d'interêt sont aussi variés que les processus physico-chimiques qui sont à l'œuvre dans les nuages, comme l'étude du champ magnétique (Běslić *et al.* 2024), des rétroactions stellaires (Palud 2023), des propriétés chimiques des cœurs denses jusqu'aux bords diffus et chauds du nuage (Santa-Maria *et al.* 2023). Sonder ces phénomènes nécessitent d'*inférer* leurs propriétés physico-chimiques sous-jacentes. Les approches variées mises en œuvre par les membres du consortium (algorithmes d'apprentissage (Einig *et al.* 2024), inférence Bayésienne (Palud *et al.* 2023), et borne de Cramér-Rao (Roueff *et al.* 2021)) nous permettent d'évaluer la pertinence de nos résultats, alors que nous ne disposons d'aucune *vérité terrain*.

Ma thèse s'inscrit dans le cadre des recherches menées par cette collaboration internationale et pluridisciplinaire. Elle a été financée par l'IRAM grâce à la subvention ANR DAOISM (ANR-21-CE31-0010), et réalisée sous la supervision d'Antoine Roueff, enseignant-chercheur à l'IM2NP en traitement de l'information, et Jérôme Pety, radioastronome sénior à l'IRAM. Ma thèse s'est déroulée au sein de l'équipe Signal & Tracking (STr) de l'IM2NP. De plus, j'ai été attentivement suivie par Maryvonne Gerin, astrophysicienne dont les recherches portent plus précisemment sur la chimie du milieu interstellaire. Avant de débuter cette thèse, j'ai eu la chance de rejoindre le consortium dans le cadre de mon stage de fin d'études d'ingénieurs. Lors de ce stage, j'ai également été encadrée durant deux semaines par Javier R. Goicoechea, astrochimiste au CSIC.

Objectifs de la thèse — Lors de cette thèse, j'ai co-développé une approche statistique permettant d'obtenir des champs de propriétés de nuages moléculaires (entre autres, de la

[6]Laboratoire d'Astrophysique de Bordeaux (France)
[7]Spanish National Research Council, Madrid (Espagne)
[8]Institut de Recherche en Astronomie et Planétologie, Toulouse (France)
[9]Jet Propulsion Laboratory, California (US)
[10]Pontificia Universidad Católica of Chile, Santiago
[11]Marseille (France)
[12]Institut Matériaux Microélectronique Nanosciences de Provence, Toulon (France)
[13]Grenoble Images Parole Signal Automatique (France)
[14]Institut National de Recherche en Informatique et Automatique, Grenoble (France)
[15]Centre de Recherche en Informatique, Signal et Automatique de Lille (France)

température cinétique, de la densité volumique, de la pression thermique et de la cinématique du gaz) à partir des données ORION-B. Mon approche a consisté à construire un modèle adéquat basé sur les connaissances physico-chimiques du milieu interstellaire, afin d'estimer les propriétés du nuage en l'ajustant aux observations. Partant du modèle le plus simple, une première étape de la thèse a été de progressivement le compléxifier, tout en injectant les hypothèses chimiques permettant de diminuer les résidus et de conserver des précisions d'estimation acceptables. La validation du modèle s'est notamment appuyée sur les résultats obtenus dans les lignes de visée provenant de la nébuleuse de la Tête de Cheval, une région d'Orion B particulièrement étudiée par les astrophysiciens.

Organisation du manuscrit — Les Chapitres 2 et 3 introduisent les concepts astrophysiques et statistiques adoptés dans cette thèse afin de relever les défis précedemment évoqués. Les Chapitres 4 et 5 présentent les contributions majeures obtenues lors de cette thèse. Le Chapitre 2 fournit un aperçu des processus physico-chimiques à l'œuvre dans le milieu interstellaire, permettant d'appréhender la nature des données *millimétriques* d'ORION-B. Le Chapitre 3 aborde les concepts statistiques misent en œuvre pour inférer les caractéristiques du nuage et quantifier leur précision d'estimation. Le Chapitre 4 présente la méthode d'estimation proposée et son application sur une région d'Orion-B pour un sous-ensemble de raies moléculaires. Ce chapitre comporte également une étude de la précision d'estimation en fonction de la résolution spectrale à temps d'intégration constant, en vue d'orienter la conception des spectromètres de la génération suivante de détécteurs du télescope IRAM-30m. Dans le Chapitre 5, nous proposons une méthodologie permettant de quantifier la perte de précision induite par la présence des erreurs de calibrations affectant les observations. Enfin, les conclusions et perspectives de ce travail sont traitées dans le Chapitre 6.

1.2 Scientific context and thesis objectives

Scientific context — Star formation occurs in the dense, cold regions of giant molecular clouds (GMCs) that are the most massive structures within the interstellar medium (ISM). These regions, referred to as *dense cores* and characterized by a typical density of 10^{4-6} particles per cm^3 at ~ 10 K (Draine 2011, Chapter 32; Bergin and Tafalla 2007), are only qualified as "dense" in relation to the other phases of the ISM, but remain too diffuse for their conditions to be reproduced in laboratory on Earth, where the ultrahigh vaccum is obtained for a density of $10^7\,\mathrm{cm}^{-3}$ at 300 K (V. V. Rao *et al.* 1998). The study of the physical and chemical processes involved in star formation therefore relies essentially on *observations*. More precisely, our understanding of star formation progresses as the methods for acquiring and analyzing observations evolve. Technological advances in recent decades make it possible to acquire both spatially and spectrally well-resolved dataset on fields of view spanning several dozens of parsecs[16]. For instance, the ORION-B[17] Large Program

[16] One parsec (pc) corresponds to the distance light travels in ~ 3.26 years.
[17] Outstanding Radio-Imaging of OrioN B

(PIs: Jérôme Pety[18] & Maryvonne Gerin[19]) surveyed with the IRAM[20]-30m telescope the Orion B cloud over a field of view of $5\,\mathrm{deg}^2$ with an angular resolution of $\sim 25''$. This observation campaign produced a data cube that consists of 1074×758 pixels, i.e., lines of sight, each made of $\sim 200\,000$ frequency channels between 72 and 116 GHz. Analyzing such a large amount of data requires developing dedicated statistical methods that take into account both physical and chemical knowledge. The first challenge addressed in this thesis is to develop an automated pipeline with minimal human intervention for analysing large-scale, multi-pixel datasets. Moreover, observing over a large bandwidth implies that in each pixel, several molecular species can be detected. Such spectrum is referred to as *multi-line* spectrum. The second challenge lies in the development of analysis tools tailored to retrieve information from multi-line observations.

The ORION-B consortium — The ORION-B consortium was created in 2014 to meet these challenges through the collaboration of astrophysicists, astrochemists and data scientists. Since its creation, the consortium has gathered the expertise of the physicists from the IRAM, LUX, LAB[21], CSIC[22], IRAP[23], JPL[24] laboratories, the UC Chile[25] and Cardiff University, and data scientists from the Fresnel[26] institute, IM2NP[27], GIPSA-LAB[28], INRIA[29], and CRIStAL[30] laboratories. The ORION-B dataset acquired between 2013 and 2020 serves us as a template of massive star forming GMCs to develop analysis techniques of GMC observations. The scientific subjects of interest concern the diversity of processes that take place within GMC, e.g., magnetic field (Běslić *et al.* 2024), impact of the stellar feedbacks (Palud 2023), and the chemistry evolution from dense cores to the edges of the cloud (Santa-Maria *et al.* 2023). Tracing such phenomena required to *infer* the underlying physicochemical properties of the medium. The diverse approaches explored by the team members (e.g., machine learning algorithms (Einig *et al.* 2024), Bayesian inference (Palud *et al.* 2023) and Cramér-Rao lower bound (Roueff *et al.* 2021)) helps us to evaluate the relevance of the results, when no *ground truth* is available.

This thesis is part of the research carried out by this international, multidisciplinary collaboration. My thesis has been funded by IRAM from the ANR DAOISM grant (ANR-21-CE31-0010), and has been carried out under the co-supervision of Antoine Roueff, professor in data sciences at IM2NP and Jérôme Pety, senior radio-astronomer at IRAM. My thesis was carried out at the IM2NP within the Signal & Tracking (STr) research team. Moreover, I have been closely supervised by Maryvonne Gerin, astrophysicist dedicated to the study of

[18] IRAM, Grenoble (France) and LUX, Paris (France)
[19] LUX, Paris (France)
[20] Institut de RadioAstronomie Millimétrique (France, Spain, and Germany)
[21] Laboratoire d'Astrophysique de Bordeaux (France)
[22] Spanish National Research Council, Madrid (Spain)
[23] Institut de Recherche en Astronomie et Planétologie, Toulouse (France)
[24] Jet Propulsion Laboratory, California (US)
[25] Pontificia Universidad Católica of Chile, Santiago
[26] Marseille (France)
[27] Institut Matériaux Microélectronique Nanosciences de Provence, Toulon (France)
[28] Grenoble Images Parole Signal Automatique (France)
[29] Institut National de Recherche en Informatique et Automatique, Grenoble (France)
[30] Centre de Recherche en Informatique, Signal et Automatique de Lille (France)

the ISM chemistry. Before beginning this thesis, I had the opportunity to join the ORION-B consortium through my end-of-studies engineering internship. During two weeks of this internship, I have also been mentored by Javier R. Goicoechea, astrochemist at CSIC.

Thesis objectives — In this thesis, I co-developed a statistical approach for deriving fields of molecular cloud properties (e.g., kinetic temperature, volume density, thermal pressure and gas velocity), through the analysis of the ORION-B dataset. My approach consists in constructing an adequate model based on physical and chemical knowledge of the ISM to retrieve the cloud properties from the fit of data. During the first part of this thesis, starting from the simplest model, I carefully increased the complexity of the model and add chemical knowledge to reduce the residue while maintaining a reasonable estimation precision. The model building was also driven from results obtained in lines of sight from the Horsehead Nebula, a region particularly well studied by astrophysicists.

Organization of this manuscript — In the Chapter 2 and Chapter 3 are introduced the astrophysical and statistical concepts that will be used throughout this thesis to meet these goals, whereas in Chapter 4 and Chapter 5 are presented the main contributions of this thesis. Chapter 2 provides an overview of the physicochemical processes that take place in the ISM required to understand the nature of the ORION-B dataset, which consists in *millimeter-wave* observations. Chapter 3 details the statistical concepts adopted to estimate the cloud properties and quantify the estimation precision. Chapter 4 presents the proposed inference method and its application to part of Orion B for a particular set of lines. This chapter also studies the evolution of the estimation precision when increasing the spectral resolution at constant integration time, in order to guide the definition of the spectrometers for the next generation of receivers at the IRAM-30m telescope. In Chapter 5, we provide a methodology to quantify the precision loss due to the presence of calibration errors that affect the observations. Finally, conclusions and perspectives are addressed in Chap 6.

References

Bergin, E. A. and M. Tafalla (Sept. 2007). "Cold Dark Clouds: The Initial Conditions for Star Formation". In: *Annual Review of A&A* 45.1, pp. 339–396. DOI: 10.1146/annurev.astro.45.071206.100404.

Běslić, I. *et al.* (2024). "The magnetic field in the Flame nebula". In: *A&A* 684, A212. DOI: 10.1051/0004-6361/202348376.

Draine, Bruce T. (2011). "Physics of the Interstellar and Intergalactic Medium". In: Princeton University Press.

Einig, L. *et al.* (2024). "Quantifying the informativity of emission lines to infer physical conditions in giant molecular clouds - I. Application to model predictions". In: *A&A* 691, A109. DOI: 10.1051/0004-6361/202451588.

Palud, P. (Dec. 2023). "Sampling methods for statistical inference of non-linear inverse problems : spatial distribution of physico-chemical properties of the interstellar medium". Theses. Centrale Lille Institut. URL: https://theses.hal.science/tel-04594046.

Palud, P. *et al.* (2023). "Neural network-based emulation of interstellar medium models". In: *A&A* 678, A198. DOI: 10.1051/0004-6361/202347074.

Roueff, A. *et al.* (2021). "$C^{18}O$, ^{13}CO, and ^{12}CO abundances and excitation temperatures in the Orion B molecular cloud - Analysis of the achievable precision in modeling spectral lines within the approximation of the local thermodynamic equilibrium". In: *A&A* 645, A26. DOI: 10.1051/0004-6361/202037776.

Santa-Maria, M. G. *et al.* (2023). "HCN emission from translucent gas and UV-illuminated cloud edges revealed by wide-field IRAM 30 m maps of the Orion B GMC - Revisiting its role as a tracer of the dense gas reservoir for star formation". In: *A&A* 679, A4. DOI: 10.1051/0004-6361/202346598.

V. V. Rao, T. B. Gosh, and K. L. Chopra (1998). "Vacuum Science and Technology". In: First edition. Allied, pp. 5–6.

Chapter 2

Radio observations of molecular clouds

Contents

2.1 Molecular clouds **19**
2.1.1 From dark to molecular clouds 19
2.1.2 The Orion B giant molecular cloud (GMC) 20
2.1.3 Cloud constituents and observations 20
2.2 The ORION-B dataset **24**
2.3 Rotational spectrum analysis **29**
2.3.1 Collisional processes 29
2.3.2 Radiative processes 30
2.4 Different molecular tracers as complementary probes of the gas **33**
References **36**

Here are introduced the astrophysical concepts that will serve as the basis for modeling the multi-line observations of the ORION-B dataset. Section 2.1 introduces the notion of Giant Molecular Clouds (GMCs) and of its two main components: interstellar dust and molecular gas, as well as the millimeter-wave observations that are aimed to be modeled in this thesis. Section 2.3 provides an overview of the radiative transfer principles required to understand the nature of the ORION-B dataset. Finally, Section 2.4 shows how different rotational lines deliver complementary information on the emitting gas.

2.1 Molecular clouds

2.1.1 From dark to molecular clouds

Molecular clouds were first detected before being characterized. William Hershel reported in Herschel 1785 the presence of dark patches in the optical sky, which therefore seemed punctuated with empty regions. Over the next two centuries, it was debated whether these dark nebulæ were empty space repelling stars, or some matter absorbing the light from the stellar background. Edward Emerson Barnard convinced himself that these dark objects were indeed "obscuring masses" when inspecting his own photographs of the Milky Way between 1904 and 1919. He compiled 182 of these dark objects in Barnard 1919, including Barnard 33 that belongs to the Orion constellation, also known as the Horsehead Nebula

(see Fig. 2.1c). The Horsehead Nebula occupies a very special place in this thesis because it will serve as a template of dense core regions to develop and benchmark our model of observations.

Considering the ISM as the birthplace of stars was proposed by Spitzer in 1941 (Spitzer 1941), and was explicitly addressed by Whipple (Whipple 1946). Driven by the Lyman Spitzer's intuition, Bok and Reilly 1947 investigated particular dark, compact, round-shape sources referred to as "globules", whose features were then suspected to form stars through gravitational collapse. Yun and Clemens 1990 confirmed this conjecture by detecting at *infrared* (IR) wavelengths young stellar objects inside globules. Indeed, the exploration of the sky at *radio* wavelengths shorter than 1 cm that began in the late 1930s made it possible to capture the photons emitted by the dark clouds themselves. Since then, observations from IR to (sub)millimeter wavelengths revealed that dark clouds are in fact made of a mixture of solid *dust* and *molecular species* in the gas phase. Since then, dark clouds are also referred to as molecular clouds.

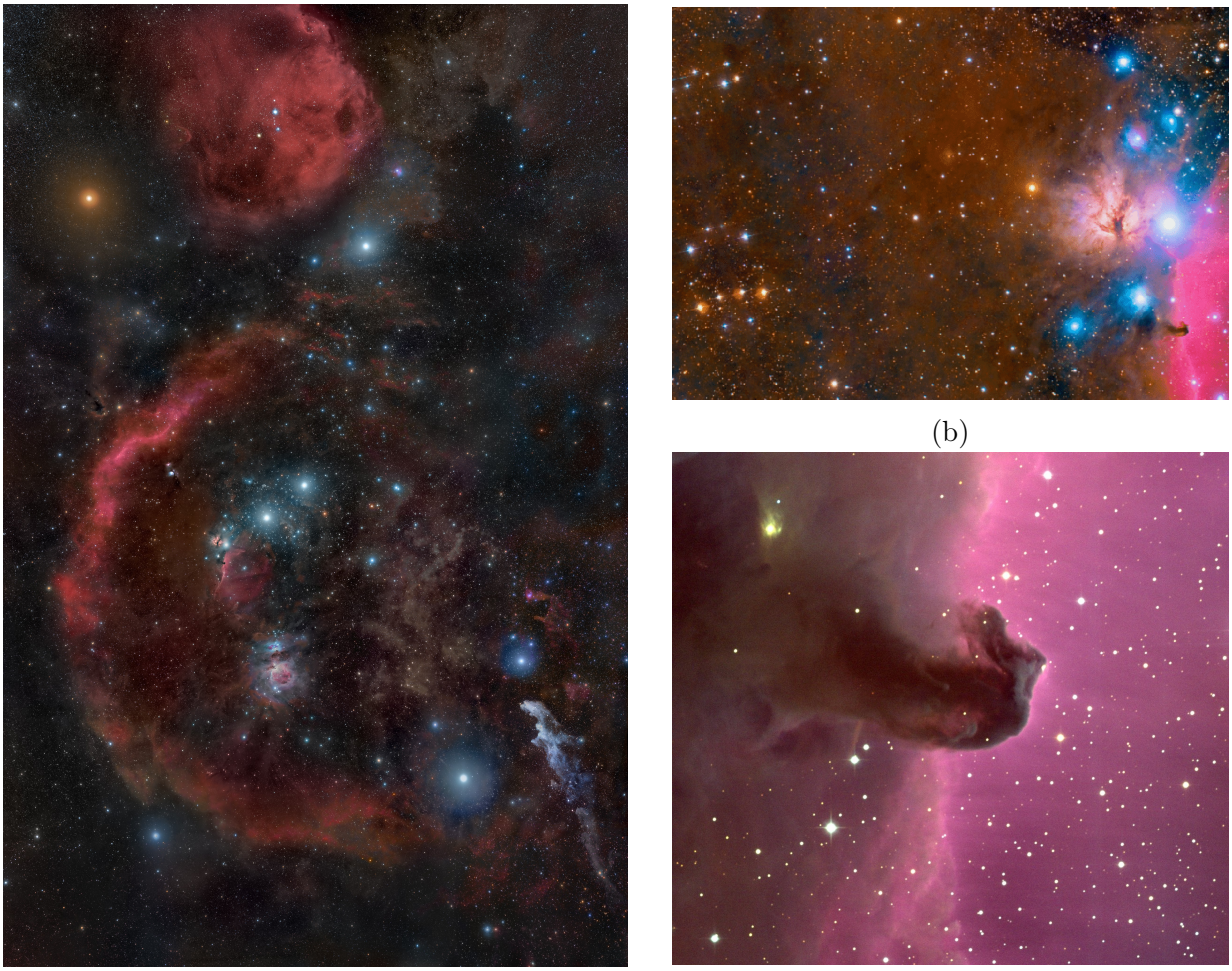
2.1.2 The Orion B giant molecular cloud (GMC)

Molecular clouds are found in the ISM under various sizes, shapes, densities, and masses. Orion B is a *giant* molecular cloud (GMC), a terminology that generally refers to particularly large clouds whose size spans a dozen of parsecs $(2-20)$ with a total mass of $10^{3-5.3}$ solar masses (Draine 2011, Chapter 32). The Orion B GMC, along with Orion A, belongs to the Orion complex of GMCs in the Orion constellation. The Orion complex, located about $1000-1400$ ly away from Earth, is among the closest[1] active star-forming regions from the Solar System. This proximity makes this region of the sky remarkable, even at visible wavelengths, and has enabled astronomers to study it in greater depth over time. The projected landscape of the Orion complex, shown in visible light in Fig. 2.1a, extends over ~ 390 ly (Burton *et al.* 1992) and reveals the diversity of interacting (ionized, atomic, and molecular) phases in the ISM, which therefore makes it an ideal laboratory for investigating the underlying physicochemical processes that take place there.

2.1.3 Cloud constituents and observations

Dust and visual extinction (A_V) — Interstellar dust is made of tiny (from $\sim 0.1-100$ nm) solid grains. While the dust phase only represents about 1% of the gas mass in the ISM (Lequeux *et al.* 2002), it plays a decisive role in the thermodynamics and chemical evolution of GMCs. The fact that GMCs appear as dark patches in visible observations is directly linked to the existence of dust grains along the line of sight. These grains absorb the far-ultraviolet (far-UV) and optical photons from nearby stars and re-emit them in the far-IR. GMCs thus appear dark in far-UV and visible observations while they shine in the far-IR. The *visual extinction*, denoted by A_V, quantifies the attenuation of the visible light. This latter quantity, expressed in magnitude on a logarithmic scale, is to first order proportional to the quantity of matter along the line of sight (see, e.g., Draine 2011, p. 237). The proportionality coefficient depends on the properties (shape, size, and nature) of the

[1] With, e.g., the Taurus molecular cloud that is ~ 141 pc away from Earth (Zucker *et al.* 2019).



(a) (b) (c)

Figure 2.1: An overview of the Orion complex in the visible light. a) The Orion complex (credit: Rogelio Bernal Andreo). b) Region of the Orion B cloud studied in this thesis (credit: Leonardo Orazi). c) The Horsehead Nebula (credit: N.A.Sharp/NOIRLab/NSF/AURA/, https://noirlab.edu/public/images/noao-02188/).

dust population along the line of sight. This coefficient is often assumed constant in the Milky Way. As dust grains are solid, they mostly emit a continuum spectrum whose shape is a slightly modified blackbody. The emission peak in the far-IR thus solely depends on the dust typical temperature that is $\sim 15\,\mathrm{K}$ in GMCs. Since all the far-UV energy is converted into far-IR energy, the fit of the spectro-energy-distribution of the dust continuum allows astronomers to get values for the far-UV field and the total amount of dust, which can be converted to the visual extinction. Moreover, the most energetic photons from high-mass stars will be able to ionize the dust that thus releases free energetic electrons in the gas. These electrons will then heat the gas through random collisions. This process, referred to as the photoelectric effect, is one of the main source of heating of molecular clouds. Finally, dust has an important role on the evolution of the molecular chemistry for two main reasons. First, dust grains act as catalysts of the chemistry. As the gas is diffuse and cold, the number

of collisions and their collisional energy are low. When colliding with dust grains, atoms and molecules can stick on them. For instance, the formation of molecular hydrogen (H_2) happens on dust grains in our galaxy. When the volume density of the gas increases, the collision time gets shorter and more and more species stick on dust grains, react together in the surface, forming a layer of ice. This phenomenon depletes the gas of its more abundant molecules such as ^{12}CO. When some source of energy is provided to the ice mantle (e.g., collision with cosmic rays), this activates a rich chemistry, and the formation of complex organic molecules. The attenuation of the far-UV field prevents the molecules to be dissociated and the atoms from being ionized. On the other hand, when a star is born inside a molecular cloud, its radiative energy will first evaporate the ice and release complex molecular species in the gas, thus contributing to enrich it.

Molecular species — Reflecting the overabundance of hydrogen (H) and helium (He) in the ISM ($\sim$ 73% and 25% of the total mass, respectively) that stems from the Big Bang, the gas phase of GMCs is mainly composed of hydrogen in molecular form (H_2) and He. In photodissociation regions (PDRs), where UV photons prevent the formation of H_2, hydrogen is found in neutral (H I) and ionized (H II) form as well. The advances in radio astronomy since the 1940s have made it possible to detect hundreds of other molecules and their isotopologues[2] in the ISM (McGuire 2018). Among the most studied ones is carbon monoxide (^{12}CO; Wilson *et al.* 1970) and its isotopologues (^{13}CO, C^{18}O; Penzias *et al.* 1971), followed by ammonia (NH_3; Cheung *et al.* 1968), hydrogen cyanide (HCN; Snyder and Buhl 1971), and formylium (HCO^+; Buhl and Snyder 1970; Klemperer 1970). Nevertheless, these species represent only a minor fraction of the gas. For instance, about one molecule of ^{12}CO can be found for 10 000 of H_2 (e.g., Tielens 2005, Sect. 10.4.3).

Whereas solid dust show a continuum spectrum, gaseous molecular species exhibit a discrete one made of emission and absorption lines, referred to as *molecular lines*. These lines occur at molecule-specific frequencies, so their detection alone delivers information on the chemical composition of the medium. To illustrate typical molecular signatures that can be acquired in molecular clouds, Fig. 2.2 shows two spectra at the 3 mm band from the WHISPER[3] (PI: Jérôme Pety) survey of the Horsehead Nebula. The analysis carried out in Guzman Veloso 2013 revealed the presence of $\sim$ 70 molecules toward the two studied lines of sight. One notes that these spectra only consist in emission lines[4], even though they appear against the cosmic microwave background (CMB). Indeed, the CMB temperature ($T_{\rm CMB} \sim 2.73$ K; Fixsen 2009) is too low to excite the molecules of the molecular gas, whose typical temperature is $\sim 10 - 50$ K (Draine 2011, Chapter 1).

[2]Molecular species that differ only in their mass number A (number of protons Z and neutrons), i.e., consider $^{A_1}_{Z}$X and $^{A_2}_{Z}$X with $A_1 \neq A_2$. For instance, ^{12}CO and ^{13}CO are isotopologues.

[3]Wideband High-resolution IRAM-30m Surveys at two Positions with Emir Receivers (Guzmán, V. *et al.* 2011; Pety *et al.* 2012; Gratier *et al.* 2013)

[4]Absorption lines can be observed in sight lines where cold molecular gas lies in front of strong sources, such as galactic center regions, also known as *quasars* (e.g., Liszt, H. *et al.* 2018).

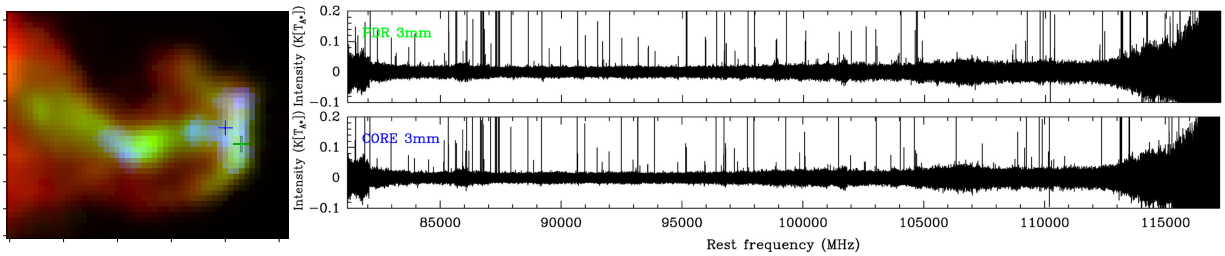


Figure 2.2: Molecular lines measured toward the Horsehead Nebula from the WHISPER dataset at the 3 mm band. Left panel: Composite image of the Horsehead Nebula from the ORION-B dataset made of the normalized integrated intensities of the first rotational transition of ^{12}CO (red channel), C^{18}O (green channel) and H^{13}CO$^+$ (blue channel). The blue and green cross indicate the location of the "CORE" and the "PDR" sight lines, respectively. The corresponding spectra from the WHISPER dataset are shown beside (credit: Guzman Veloso 2013, Chapter 2).

An emission line at the frequency $\nu_{\rm up\to low}$ results from the radiative de-excitation of molecular species relaxing from an energy level $E_{\rm up}$ to a lower one $E_{\rm low}$, such that

$$\nu_{\rm up\to low} = \Delta E/h, \tag{2.1}$$

where $\Delta E = E_{\rm up} - E_{\rm low}$ and h is the Planck constant. The molecule may first have been excited by absorbing a photon, which corresponds to a radiative excitation, or by colliding with another particle (mainly H_2, He, and electrons). Radiative excitation dominates either in diffuse gas or in the dense cloud regions exposed to a strong incident radiation field (e.g., PDR). Collisional excitation rather dominates in dense far-UV shielded regions.

Collisions are driven by the random motions associated with the kinetic energy of the molecules, which is quantified by the *kinetic temperature* of the gas, denoted by $T_{\rm kin}$ (K). Molecular hydrogen, as the most abundant species of the medium, is the main partner the molecules collide with. Depending on the energy variation of the transition, given by ΔE, the molecule's energy state can change through three types of transitions: rotational transition ($\Delta E \sim$ meV, corresponding to $\nu_{\rm up\to low}$ in the (sub)millimeter wavelengths), vibrational transition (a fraction of eV, in the IR), and electronic transition (a few eV, in the far-UV; Lequeux *et al.* 2002). Hence, the kind of transitions that can be observed depends on the kinetic temperature of the gas and the molecule. For instance, ^{12}CO undergoes its first rotational transition, denoted by ^{12}CO($J = 0 \to 1$)[5], for a frequency of 115.271 GHz, which corresponds to a variation of energy between the upper and lower levels of $\Delta E/k \sim 5.5$ K (k is the Boltzmann constant). which is only a few Kelvin higher than $T_{\rm CMB}$. In contrast, about 3080 K is required to reach its first vibrational transition (based on Tielens 2005, Sect. 2.1.3). The diagram shown in Fig. 2.3 summarizes the relation between an object temperature and the type of spectroscopy used to study it. Thus, *rotational spectroscopy*, which is the analysis of molecular lines at millimeter wavelengths (from 1 mm to 10 mm, i.e., $30 \lesssim \nu_{\rm up\to low} \lesssim 300$ GHz), is especially suitable for observing cold molecular clouds. Hereafter, the rotational transition ($J = J_{\rm up} \to J_{\rm low}$) is simply denoted by ($J_{\rm up} - J_{\rm low}$). In

[5] J is the rotational quantum number. Here, $J_{\rm up} = 1$ and $J_{\rm low} = 0$ is the fundamental level.

addition to revealing the medium composition, detecting different molecular lines (whether of the same species or not) can therefore provide information on the different gas phases encountered along the line of sight, since they are emitted by gas cells of different physical conditions. Observing and analyzing the H_2 emission lines excited by the cold molecular gas would provide the best representative conditions of the gas. However, symmetrical molecules such as H_2 have no permanent dipole moment, and therefore the only accessible rotational transitions are the very faint quadrupolar lines with $\Delta J = 2$, which probe lukewarm[6] gas. Consequently, to retrieve the gas properties from millimeter observations, one has to focus on other molecules that compose the medium in smaller proportions. These molecules are therefore qualified as molecular *tracers*.

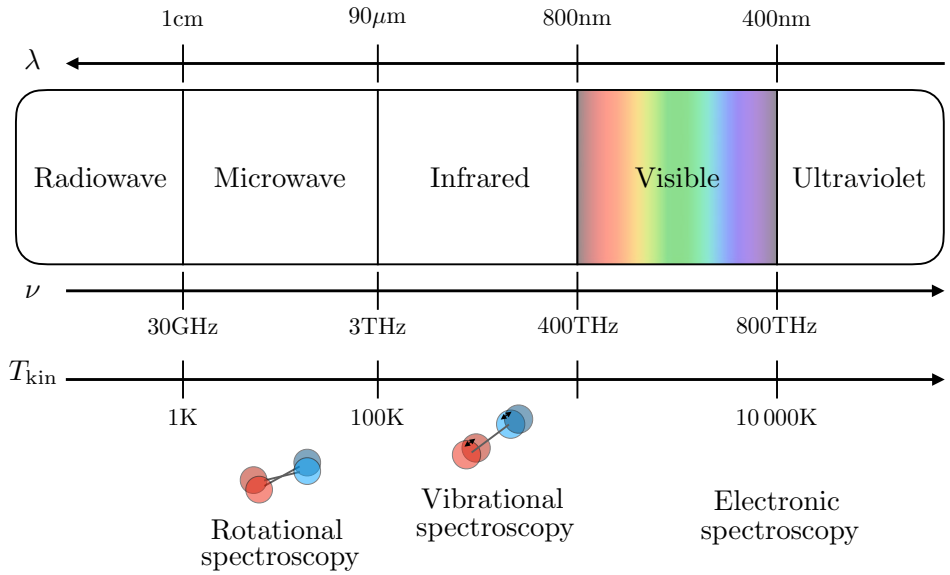


Figure 2.3: Electromagnetic spectrum diagram.

2.2 The ORION-B dataset

The observable: a molecular line and its interpretation — As stated earlier, in millimeter radio astronomy, we observe the intensity emitted by the interstellar gas as a function of the gas velocity along the line of sight plus some noise. This observable is named a spectrum. For instance, Fig. 2.4 shows the $C^{18}O(1-0)$ molecular line toward a dense core of the Horsehead Nebula. The typical profile of the emission line is Gaussian. The characteristics of this Gaussian are interpreted as follows. The position of the Gaussian corresponds to the typical velocity of the gas along the line of sight, referred to as the *centroid velocity* (C_V). The Gaussian width is related to the typical velocity dispersion (σ_V) of the gas along the line of sight. The latter quantity is related to the full width at half maximum (FWHM) by

$$\mathrm{FWHM} = \sqrt{8\ln(2)}\,\sigma_V .$$

[6] A thermally unstable phase of the interstellar gas, at intermediate temperature between the cold and warm neutral medium (Kalberla and Haud 2018).

The integrated area, or the *integrated intensity*, is related to the amount of $C^{18}O$ molecules along the line of sight. If we know the amount of H_2 molecules along the same line of sight, we will be able to compute the abundance of $C^{18}O$ relative to H_2. Finally, the peak intensity is related to the excitation of the species at the considered transition by collisions.

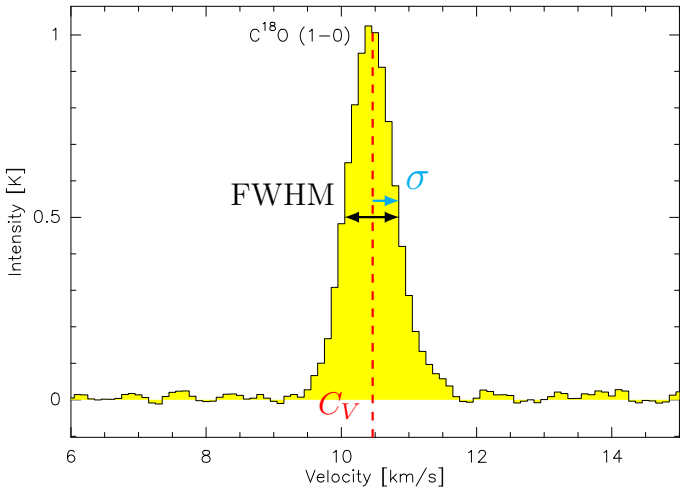


Figure 2.4: Spectrum of $C^{18}O(1-0)$ toward a dense core of the Horsehead Nebula.

The ORION-B project — The science goals of the ORION-B project is to contribute to elucidate the link between star formation and the structure (dynamics, physics, and chemistry) of the host parent. To do this, the consortium has searched to build an unbiased dataset with molecular tracers of all the typical components (diffuse gas, filaments, and dense cores) of a GMC. Almost all the 3 mm atmospheric window was observed in 3 different frequency tunings shown in Fig. 2.5. Each tuning delivers two spectra of an instantaneous bandwidth of about $\sim 8\,\mathrm{GHz}$, separated by about $\sim 8\,\mathrm{GHz}$. The three tunings thus covers a total bandwidth about 40 GHz, sampled every 200 kHz. The IRAM receivers are still monopixel ones. In order to map the sky, the telescope thus sweeps the sky while integrating the signal and dumping the resulting spectrum at a regular time interval (typically every 0.25 seconds). This observing mode is named "On-The-Fly". In $\sim$ 850 hours of telescope times scheduled from 2013 to 2020, the ORION-B consortium succeeded to observe $5\,\mathrm{deg}^2$ toward the Orion B cloud. The angular resolution of the IRAM-30m varies from $36''$ at 72 GHz to $21''$ at 116 GHz. One fully sampled image thus contains about 1 million of pixels, and the dataset contains about 200 000 such images. At 24 images per second, this would produce a movie of about 2h20. Nevertheless, this large quantity of data is limited by the spatial median noise level that varies from 0.1 to 0.5 K as a function of the observed frequency. Clear signal is only detected on about 1% of the frequency channels, corresponding to the brightest molecular lines observed in this GMC.

The definition of a line cube — As most of the channels contain mostly noise, line cubes are built around the observed frequencies of the brightest lines. A cube can be thought as a set of spectra such as the one presented in Fig. 2.4, each spectrum corresponding to the line

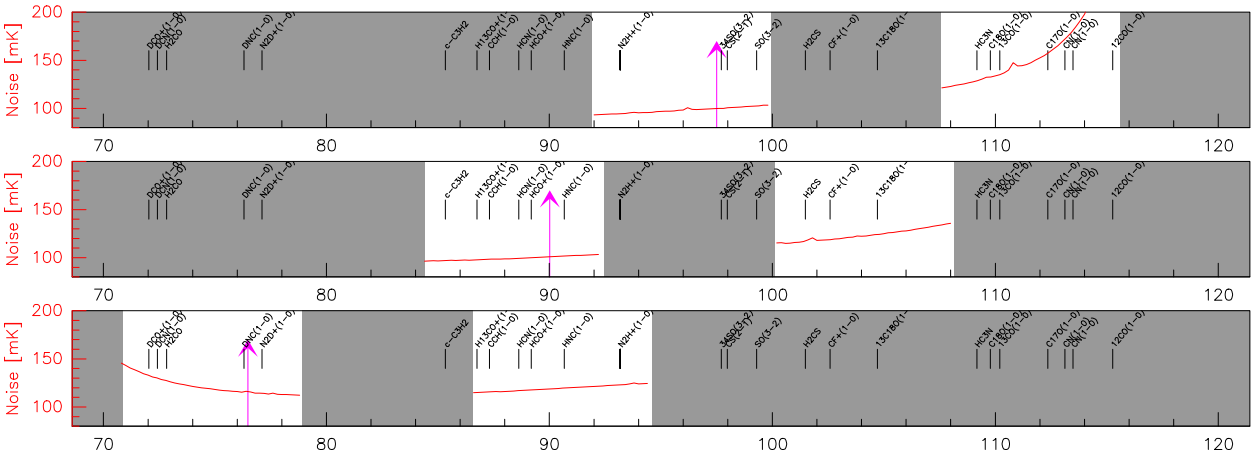


Figure 2.5: Representation of the three frequency tunings observed within the ORION-B project. The grey rectangle shows the 3 mm atmospheric window. The white rectangles show the frequency intervals where the signal is recorded at a spectral resolution of 200 kHz for each of the three tuning. Each vertical line shows the rest frequency of a relatively bright line. The red curves show the typical sensitivity as a function of frequency (expressed in GHz).

of sight at the associated pixel. The top part of Fig. 2.6 shows as a map of spectra where each spectrum is represented at its position on the sky. The line cube can also be thought as a set of images regularly sampled in velocity. The bottom part of Fig. 2.6 shows such a channel map. Figure 2.6 only displays a small region of the full field of view, toward the Horsehead Nebula. The ORION-B project produced about 30 such line cubes.

The map of the line integrated profile — The spatial distribution of the signal integrated along the velocity axis is referred to as map of the line integrated profile. Figure 2.7 shows these maps for 16 different molecular lines. In order to get a consistent view of the relative intensities at the same pixels, we define a velocity interval where there is significant signal on the highest signal-to-noise ratio (S/N) cube, i.e., the ^{12}CO$(1-0)$ line cube, and we use this velocity interval for all the other lines. The velocity width of the corresponding mask cube is shown on the bottom row next to the map obtained by integrated at a frequency where only noise is expected. For completeness, Fig. 2.7 also shows the spatial distribution of the dust visual extinction (A_V) and the far-UV radiation field, denoted by G_0, measured from the dust continuum emission. The S/N varies by order of magnitudes from one line to another. Beyond this effect, some bright lines such as ^{12}CO$(1-0)$ are well detected on almost all the observed field of view, while others bright lines such as $N_2H^+(1-0)$ are only detected on a small fraction of the pixels revealing a filamentary structure. This reflects the internal structure of GMC that is made of dense cores linked by a filamentary network embedded into more diffuse gas. Besides, one notes that the lines are overall more luminous where the gas is highly illuminated in far-UV photons (large values of G_0).

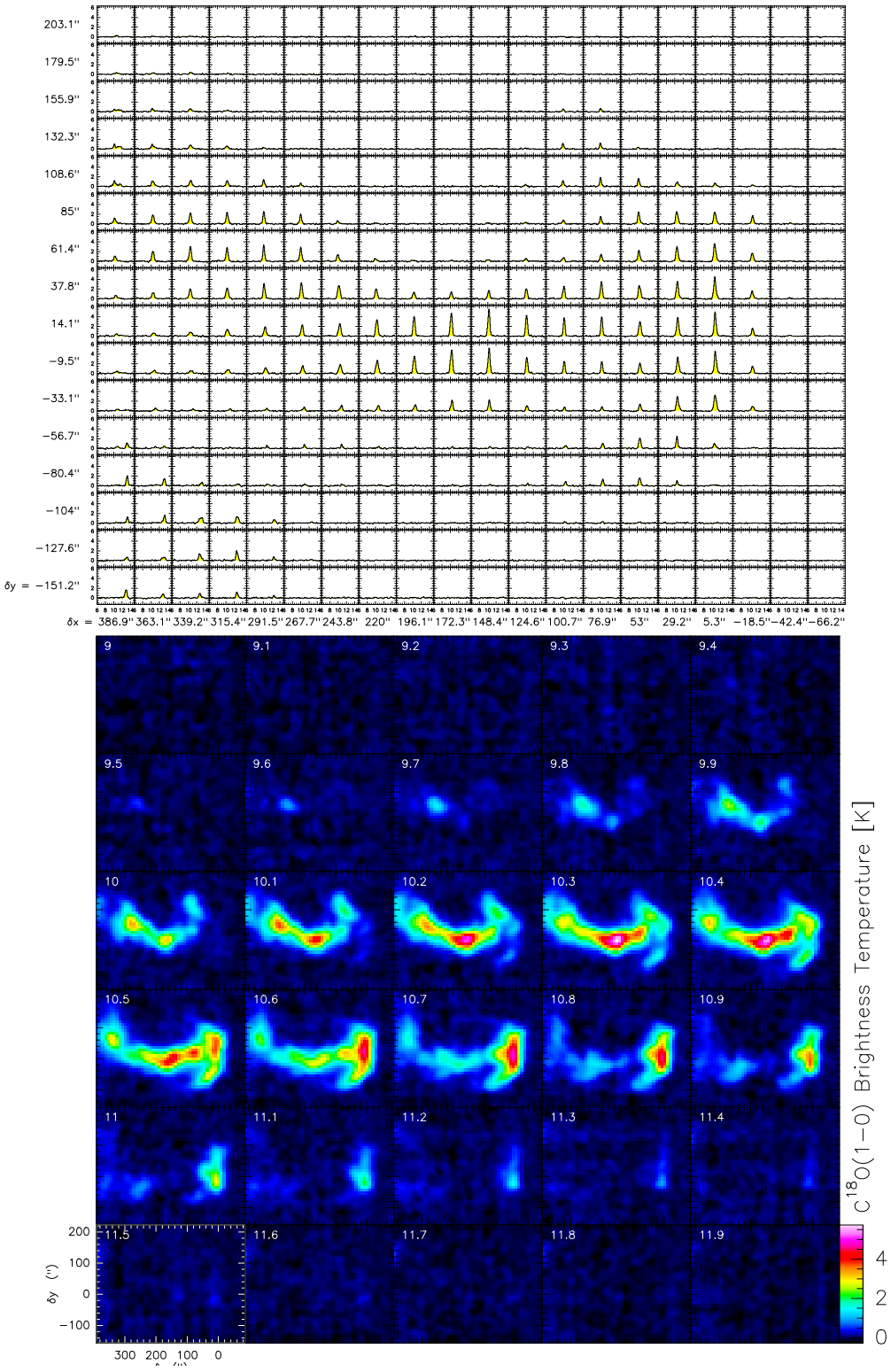


Figure 2.6: Two representations of the $C^{18}O(1-0)$ line cube towards the Horsehead Nebula. Top panel: Map of spectra where each spectrum is shown at its actual location on the sky. Bottom panel: Channel maps where each image shows the spatial distribution of the line at the velocity ($\mathrm{km\,s^{-1}}$) reported in the top left corner of the image.

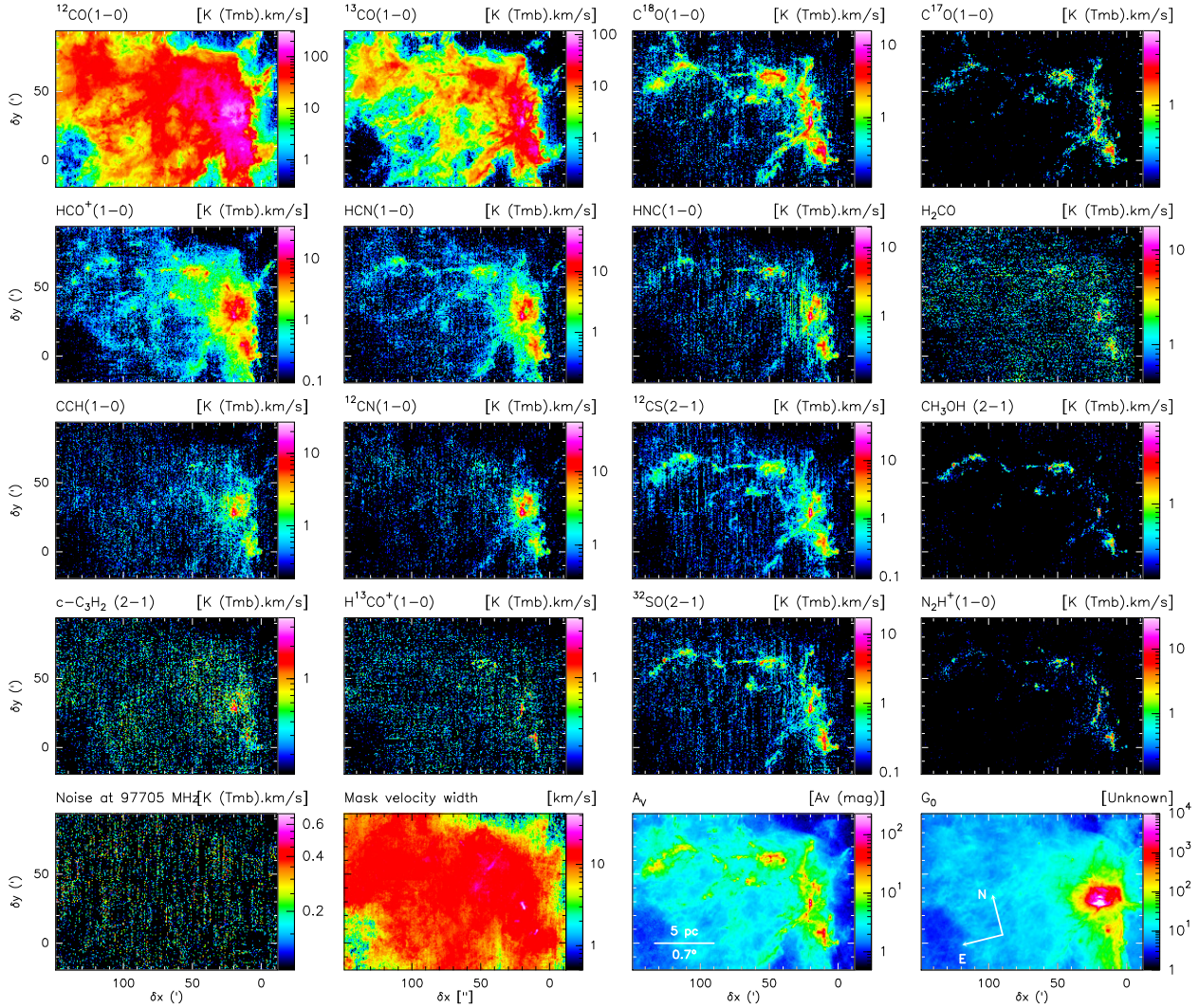


Figure 2.7: A summary of the ORION-B dataset. The four top rows show maps of the integrated intensity of twelve 3 mm lines observed with the IRAM-30m as part of the ORION-B Large program towards the Orion B GMC. The left panel of the bottom row shows the maps of integrated intensity towards a portion of the spectra without detected line, for reference. All the lines have been integrated on the same velocity interval that depends on the position of the sky. The second panel of the bottom row shows the spatial distribution of the mask spectral width used to compute the line integrated intensity. The last two panels show the maps of visual extinction (A_V) and far UV illumination (G_0) inferred from the dust continuum emission observed as part of the Herschel Gould Belt survey. The North and East are plotted on the bottom right panel. The FoV is about 18 pc × 13 pc for a source distance of 400 pc.

2.3 Rotational spectrum analysis

Molecular line profiles result from photon emissions and absorptions that occur when the species changes from an energy level to another. Thus, the knowledge of the population of the energy levels, i.e., the number of molecules found in the different energy levels, is the key information for characterizing a spectrum. The population of energy levels mainly stems from two mechanisms: collisional and radiative processes.

2.3.1 Collisional processes

Inelastic collisional processes correspond to energy transfers that occur when two species collide. Collisions are mainly due to the most abundant species in the gas such as H I, H_2, He, and electrons. These species are referred to as *colliders* whose density is denoted by $n_{\rm col}$ ($\rm cm^{-3}$). The denser the gas, the more important are the collisional processes with respect to the radiative ones. For each species and transition, a critical density $n_{\rm crit}$($\rm cm^{-3}$) is defined as the density of colliders beyond which the collisional processes dominate over the radiative ones. When $n_{\rm col} \gg n_{\rm crit}$, the number of collisions is such that the depopulation of an energy level is "instantly" compensated by its repopulation and in this case, the gas has reached the Local Thermodynamic Equilibrium (LTE). Then, the ratio between the level populations of an upper level ($E_{\rm up}$) and a lower one ($E_{\rm low}$), denoted by $n_{\rm up}$ and $n_{\rm low}$, respectively, only depends on the kinetic temperature ($T_{\rm kin}$), i.e.,

$$\frac{n_{\rm up}}{n_{\rm low}} = \frac{g_{\rm up}}{g_{\rm low}} \exp\left(-\frac{h\nu_{\rm up\to low}}{kT_{\rm kin}}\right),$$

where $g_{\rm up}$ and $g_{\rm low}$ are the degeneracy of the upper level and the lower level, respectively. When LTE is not yet reached, $n_{\rm up}/n_{\rm low}$ allows us to define a so-called *excitation temperature* $T_{\rm ex}$ (K) such that

$$\frac{n_{\rm up}}{n_{\rm low}} = \frac{g_{\rm up}}{g_{\rm low}} \exp\left(-\frac{h\nu_{\rm up\to low}}{kT_{\rm ex}}\right), \tag{2.2}$$

which is useful to characterize the degree of excitation in the non-LTE regime since it quantifies how the energy levels are relatively populated. Depending on the $T_{\rm ex}$ value, the following regimes are defined.

Thermal regime — When $T_{\rm ex} > 0\,\rm K$, the lower energy level is more populated than the upper one. It corresponds to the thermal emission regime. In particular, $T_{\rm ex} \ll T_{\rm kin}$ corresponds to a sub-thermal regime and $T_{\rm ex} \gg T_{\rm kin}$ corresponds to a supra-thermal regime. When $n_{\rm col} \gg n_{\rm crit}$, LTE is reached and then, $T_{\rm ex} \to T_{\rm kin}$.

Population inversion — Otherwise, when $T_{\rm ex} < 0\,\rm K$, the upper energy level is more populated than the lower one. This phenomenon is referred to as *population inversion*. The LASER[7] and MASER[8] processes occur in this regime where an overpopulation of the species is kept in an excited state.

[7]Light Amplification by Stimulated Emission of Radiation
[8]Microwave Amplification by Stimulated Emission of Radiation

2.3.2 Radiative processes

Radiative processes refer to the interactions between the gas and photons, including absorption, emission, and scattering processes. Let us consider a homogeneous slab of gas composed of the molecule X (see Fig. 2.8a). The intensity received by the radio telescope, expressed in K, at the frequency $\nu_{\mathrm{up}\to\mathrm{low}}$ is given by

$$s(\nu_{\mathrm{up}\to\mathrm{low}}) = \left(J\left(T_{\mathrm{ex}}, \nu_{\mathrm{up}\to\mathrm{low}}\right) - J\left(T_{\mathrm{CMB}}, \nu_{\mathrm{up}\to\mathrm{low}}\right)\right)\left(1 - \exp\left(-\tau_{\nu_{\mathrm{up}\to\mathrm{low}}}\right)\right), \tag{2.3}$$

where the intensity function J (K) is defined by

$$\begin{aligned} J(T,\nu) &= \frac{c^2}{2\,k\nu^2} B(T,\nu) \\ &= \frac{h\nu}{k} \frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1}. \end{aligned} \tag{2.4}$$

In the latter formula, $B(T,\nu)$ is the spectral distribution of a blackbody at the temperature T, T_{ex} is the excitation temperature of the line, and $T_{\mathrm{CMB}} \simeq 2.73\,\mathrm{K}$ is the temperature of the cosmic microwave background (CMB). In Eq. (2.3), $\tau_{\nu_{\mathrm{up}\to\mathrm{low}}}$ denotes the *opacity* of the line, a quantity that allows us to characterize the absorption and emission processes that can occur when a photon emitted at $\nu_{\mathrm{up}\to\mathrm{low}}$ is propagating through the gas. It is defined by (private communication with Javier R. Goicoechea; Mangum and Shirley 2015)

$$\tau_{\nu_{\mathrm{up}\to\mathrm{low}}} = 10^{10}\, \frac{A_{\mathrm{up,low}}}{8\pi\,\mathrm{FWHM}} \left(\frac{c}{\nu}\right)^3 \frac{g_{\mathrm{up}}}{g_{\mathrm{low}}}\, N_{\mathrm{low}} \left(1 - \exp\left(-\frac{h\nu_{\mathrm{up}\to\mathrm{low}}}{kT_{\mathrm{ex}}}\right)\right), \tag{2.5}$$

where the FWHM ($\mathrm{km\,s^{-1}}$) is the full width at half maximum, $A_{\mathrm{up,low}}$ ($\mathrm{s^{-1}}$) is the Einstein coefficient for spontaneous emission, which is related to the typical lifetime of the species in the upper energy level (E_{up}) before spontaneously it de-excites to the lower energy one (E_{low}). In Eq. (2.5), N_{low} ($\mathrm{cm^{-2}}$) is the *column density* of X in E_{low}. It corresponds to the volume density n_{low} of X in E_{low}, integrated along the line of sight L, i.e.,

$$N_{\mathrm{low}} = \int_0^L n_{\mathrm{low}}(l)\,\mathrm{d}l. \tag{2.6}$$

The term 10^{10} in Eq. (2.5) solely is a conversion factor since the introduced quantities are in cgs units.

Opacity (τ) — As shown in Eq. (2.5), the opacity contains the information desired about the population of energy levels of the considered molecule.

- When $\tau \gg 1$, the molecular line is qualified as *optically thick*. As $(1 - \exp(-\tau)) \underset{\tau\to\infty}{\simeq} 1$, the received intensity defined in Eq. (2.3) yields to

  $$s(\nu_{\mathrm{up}\to\mathrm{low}}) \simeq J(T_{\mathrm{ex}}, \nu_{\mathrm{up}\to\mathrm{low}}) - J(T_{\mathrm{CMB}}, \nu_{\mathrm{up}\to\mathrm{low}}).$$

  The medium then behaves as a blackbody of temperature T_{ex} because of line trapping (see Fig. 2.8b). The only radiation that reaches the observer is that coming from the

surface of the medium. The measured spectrum therefore no longer contains any information on the gas density but only about the excitation temperature. Furthermore, when $n_{\rm col} \gg n_{\rm crit}$, LTE is reached such that $T_{\rm ex} \to T_{\rm kin}$. The received intensity then allows us to directly access to the kinetic temperature of the gas, i.e.,

$$s(\nu_{\rm up\to low}) \simeq J(T_{\rm kin}, \nu_{\rm up\to low}) - J(T_{\rm CMB}, \nu_{\rm up\to low}). \tag{2.7}$$

Thus, an optically thick line under LTE proves to be an efficient thermometer of molecular gas.

- When $\tau \ll 1$, the line is referred to as *optically thin*. As $(1 - \exp -\tau) \underset{\tau\to 0}{\simeq} \tau$, the received intensity yields to

$$s(\nu_{\rm up\to low}) \simeq \left(J(T_{\rm ex}, \nu_{\rm up\to low}) - J(T_{\rm CMB}, \nu_{\rm up\to low})\right) \tau_{\nu_{\rm up\to low}}. \tag{2.8}$$

In the intermediate case, photons emitted by the rear of the cloud are more likely to be absorbed by the medium than those emitted near the front surface, as they must travel a longer distance through the medium before escaping from it. The signal received thus comes from the gas outskirts close to the observer (see Fig. 2.8c). Thus, the lower the opacity, the deeper the cloud is explored.

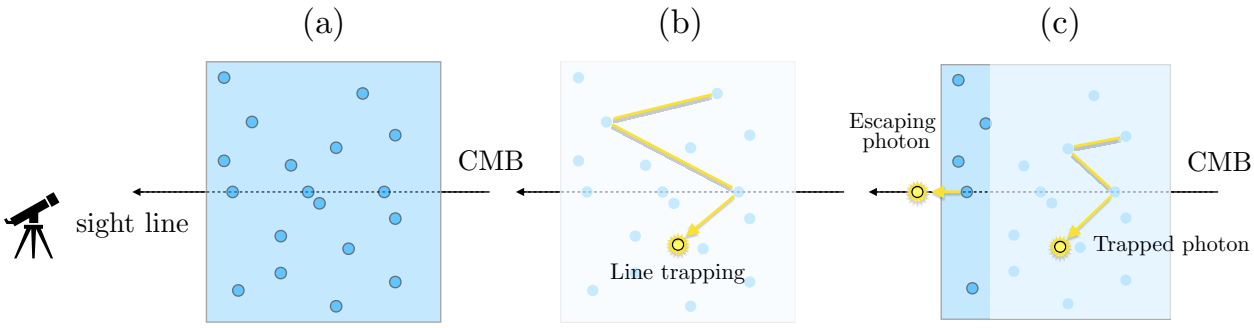


Figure 2.8: Illustration of an optically thick and thin molecular line. a) A molecular cloud composed of the species X (blue dots), observed at the line frequency $\nu_{\rm up\to low}$. b) Optically thick line whose spectrum appeared as a blackbody ($\tau \gg 1$). c) Optically thin line ($\tau \ll 1$).

The Doppler effect in turbulent gas — A given species emits photons at well-defined frequencies, referred to as rest frequencies. However, when the gas cell that emits the photons has a velocity ΔV ($\rm km\,s^{-1}$) with respect to the observer, the received photons seem have been emitted at a shifted frequency (redshifted for a gas cell moving away from the observer). This Doppler effect can be written in the radio convention of velocity as

$$\nu_{\rm red} = \nu_{\rm up\to low}\left(1 - \frac{\Delta V}{c}\right), \tag{2.9}$$

where c is the speed of the light. For a microturbulent gas, the gas velocity follows a Gaussian distribution centered on the systemic velocity of the source, referred to as the *centroid velocity* (C_V), with a typical *velocity dispersion* σ_V. The Doppler effect therefore broadens the natural

opacity profile (Ψ) of the line as the photons can be emitted at frequencies slightly lower or higher than the systemic redshifted frequency. By taking into account this phenomenon, the intensity received by the observer around the frequency $\nu_{\mathrm{up}\to\mathrm{low}}$ is given by

$$s(\nu_{\mathrm{up}\to\mathrm{low}}) = \left(J\left(T_{\mathrm{ex}}, \nu_{\mathrm{up}\to\mathrm{low}}\right) - J\left(T_{\mathrm{CMB}}, \nu_{\mathrm{up}\to\mathrm{low}}\right)\right)\left(1 - \exp\left(-\Psi_{\nu_{\mathrm{up}\to\mathrm{low}}}\right)\right). \tag{2.10}$$

Using the Doppler effect in the radio convention (Eq. (2.9)), it is possible to translate the frequency axis into a velocity one. This translation eases the comparison of several transitions of the same species or even of lines from different species. The opacity profile then reads

$$\Psi_{\nu_{\mathrm{up}\to\mathrm{low}}}(V) = \tau_{\nu_{\mathrm{up}\to\mathrm{low}}} \exp\left(-\frac{(V - C_V)^2}{2\sigma_V^2}\right). \tag{2.11}$$

Effective critical density ($n_{\mathrm{crit}}^{\mathrm{eff}}$) — According to the above, the higher the opacity of the line, the larger portion of this molecular gas behaves as a cavity from which the photons struggle to escape. In this cavity, radiative excitation processes are important so that a lower quantity of colliders n_{col} is sufficient for LTE to be achieved. In addition, the more microturbulent the medium, the more significant the Doppler effect, and thus the easier a photon will escape from the cavity because it will be seen by the other gas cells of the cavity at a shifted frequency. The effective critical density $n_{\mathrm{crit}}^{\mathrm{eff}}$ (cm^{-3}) allows us to take into account the effects of line-trapping and line-opacity. For a homogeneous slab of gas without microturbulence, the *effective critical* density of a considered molecular line is given by

$$n_{\mathrm{crit}}^{\mathrm{eff}} = n_{\mathrm{crit}} \frac{1 - \exp\left(-3\,\tau_{\nu_{\mathrm{up}\to\mathrm{low}}}\right)}{3\,\tau_{\nu_{\mathrm{up}\to\mathrm{low}}}}. \tag{2.12}$$

Summary — The excitation temperature (T_{ex}), the opacity ($\tau_{\nu_{\mathrm{up}\to\mathrm{low}}}$), the velocity dispersion (σ_V), and the effective critical density ($n_{\mathrm{crit}}^{\mathrm{eff}}$) are thus the important parameters to characterize the gas properties from the received intensity by the observer. Specific combinations of $\{T_{\mathrm{ex}}, \tau, \sigma_V, n_{\mathrm{crit}}^{\mathrm{eff}}\}$ allow us to probe different information about the physical state of the cloud.

RADEX: A non-LTE radiative transfers software — RADEX (van der Tak *et al.* 2010) is a Fortran program that uses the escape probability formalism to compute the emerging radiation from a homogeneous medium of simple geometry. The non-LTE population of the energy levels of the emitting species is computed by solving the statistical balance between radiative and collisional excitation and de-excitation of the important energy levels. This requires collisional coefficient tables that summarize for each molecule the quantum physics underlying the collision excitation. Moreover, the escape probability formalism assumes that the population of the energy levels only depends on the local property of the gas, e.g., its kinetic temperature and its volume density. This decouples the solving of the statistical balance from the global radiative transfer problem. This assumption is well adapted to the ISM because the large velocity gradients in hypersonic turbulence produce a large Doppler effect implying that the photons are no longer absorbed by the medium but escaping from it.

The input parameters of RADEX are the kinetic temperature of the gas, the CMB temperature, the volume density of the important colliders (e.g., H_2, He, and electrons), the column density of the considered molecule, the FWHM of the molecular line, the geometry (including the escape probability detailed method), and the tables of collisional coefficients. RADEX delivers the line brightness (T_b), opacity (τ), and excitation temperature (T_ex).

2.4 Different molecular tracers as complementary probes of the gas

Here, we illustrate how the analysis of diverse molecular tracers allows an astronomer to recover the physical characteristics of the gas $n_{\mathrm{H_2}}$ and T_kin. The methodology is illustrated through the study of the transitions $(1-0)$, $(2-1)$ of ^{12}CO and $(1-0)$, $(3-2)$ of HCO^+. We also neglect the Doppler effect. As $h\,\nu_{\mathrm{up}\to\mathrm{low}} \ll k T_\mathrm{ex}$ for all the considered lines[9], the intensity function is approximately equal to the excitation temperature, i.e.,

$$\begin{aligned} J(T_\mathrm{ex},\nu) &= \frac{h\nu}{k}\frac{1}{\exp\left(\frac{h\nu}{kT_\mathrm{ex}}\right)-1} \\ &\approx T_\mathrm{ex}. \end{aligned} \tag{2.13}$$

From Eq. (2.3), the intensity received by the observer can therefore be expressed as a temperature by

$$T_\mathrm{b} = \left(T_\mathrm{ex} - T_\mathrm{CMB}\right)\left(1-\exp\left(-\tau\right)\right), \tag{2.14}$$

where T_b is called the brightness temperature. This is the *Rayleigh-Jeans* approximation. To characterize the received intensity following the physical state of a gas, the following quantities are computed and plotted, as a function of the column density of the species:

- the brightness temperature T_b,
- the ratio $T_\mathrm{ex}/T_\mathrm{kin}$, the distance from LTE (see Sect. 2.3.1),
- the opacity τ,
- the ratio $n_\mathrm{col}/n_\mathrm{crit}^\mathrm{eff}$, which quantifies whether collisional processes dominate over the radiative ones. As the main collider is $n_{\mathrm{H_2}}$, when $n_{\mathrm{H_2}}/n_\mathrm{crit}^\mathrm{eff} \gg 1$, the collisional processes dominate the radiative ones.

To simulate a typical GMC, we choose $n_{\mathrm{H_2}} = 10^4\,\mathrm{cm^{-3}}$ and $T_\mathrm{kin} \in [15, 20, 40]$ K. The Table 2.1 presents the critical densities of the studied species and transitions, for these kinetic temperatures. From these values, we compute the effective critical densities using Eq. (2.12). The chosen column density ranges, i.e., $N(^{12}\mathrm{CO}) \in [10^{16}, 10^{19}]\,\mathrm{cm^{-2}}$ and $N(HCO^+) \in [10^{11.2}, 10^{14.2}]\,\mathrm{cm^{-2}}$, reflect the typical abundance ratio $N(^{12}\mathrm{CO})/N(HCO^+) = 10^{4.68}$. Besides, we use the uniform sphere geometry of RADEX as the escape probability method.

[9] For instance, at the highest transition frequency $\nu_{\mathrm{up}\to\mathrm{low}} = 230.5$ GHz of ^{12}CO$(2-1)$, $h\,\nu_{\mathrm{up}\to\mathrm{low}}/k\,T_\mathrm{ex} = 0.27 \ll 1$.

Table 2.1: Critical densities of ^{12}CO$(1-0)$, $(2-1)$ and of HCO$^+(1-0)$, $(3-2)$ at $T_{\rm kin} = 15\,$K.

Species	*Transition*	*Frequency* [GHz]	$n_{\rm crit}$ [cm^{-3}]
^{12}CO	$(1-0)$	115.2712	2.0×10^3
	$(2-1)$	230.5380	1.6×10^4
HCO$^+$	$(1-0)$	89.1885	1.5×10^5
	$(3-2)$	267.5573	2.0×10^7

^{12}CO analysis — Figure 2.9 shows the impact of the kinetic temperature on the brightness of the ^{12}CO lines (1$^{\rm st}$ column panels), their $T_{\rm ex}/T_{\rm kin}$ ratio (2$^{\rm nd}$ column panels), their opacity (3$^{\rm nd}$ column panels), and the $n_{\rm H_2}/n^{\rm eff}_{\rm crit}$ ratio (4$^{\rm th}$ column panels). All these values are computed for $n_{\rm H_2} = 10^4\,$cm^{-3} and as a function of the column density. For $T_{\rm kin} = 15\,$K, the brightness limit of both lines is reached for $N(^{12}{\rm CO}) \sim 10^{17}\,$cm^{-2}. The corresponding opacities are larger than 10, implying a largely optically thick medium, and $n_{\rm H_2}/n^{\rm eff}_{\rm crit} \sim 10^2$. As the collisional processes completely dominate the radiative ones, the gas easily reaches LTE. The medium is thus observed as a blackbody of temperature $T_{\rm kin}$: $T_{\rm b} \approx T_{\rm ex} - T_{\rm CMB}$. In other words, the brightness temperature of ^{12}CO is an efficient thermometer of the gas kinetic temperature.

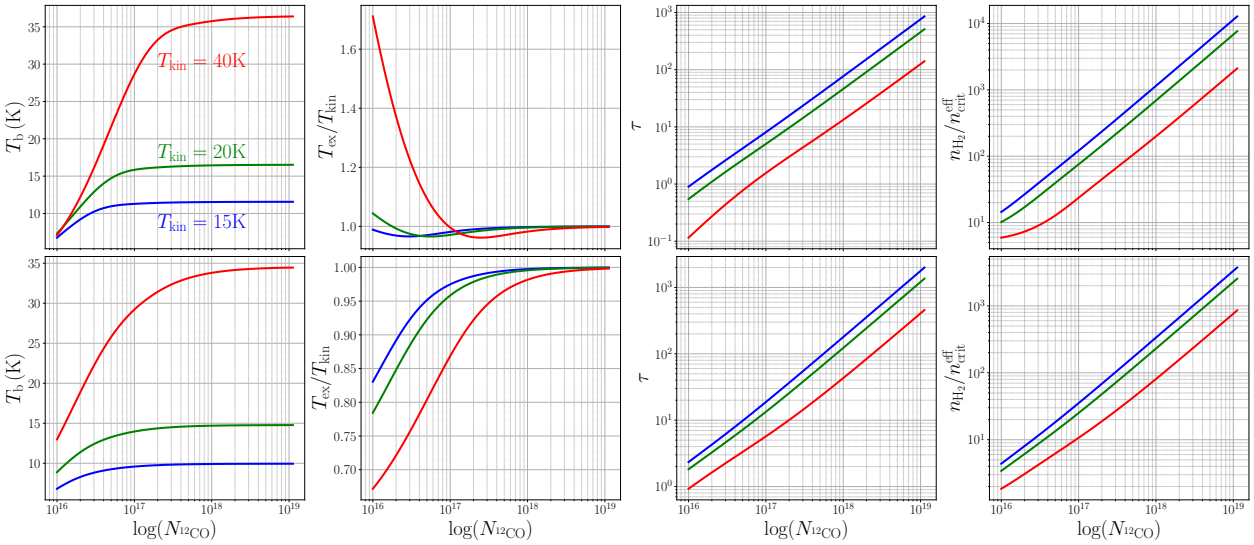


Figure 2.9: From left to right, variations with respect to the ^{12}CO column density of the brightness temperature ($T_{\rm b}$), the $T_{\rm ex}/T_{\rm kin}$ ratio, the opacity (τ), and the $n_{\rm H_2}/n^{\rm eff}_{\rm crit}$ ratio. The computation are done for $n_{\rm H_2} = 10^4\,$cm^{-3}. The top row panels concern ^{12}CO$(1-0)$, and the bottom ones, ^{12}CO$(2-1)$.

HCO^+ analysis — Figure 2.10 shows the variations with the column density of the same quantities as Figure 2.9, but for the HCO^+ lines. The lines of HCO^+ behave differently from the ^{12}CO ones. For $N(HCO^+) \in [10^{13}, 10^{14.2}]$ cm^{-2}, the brightness temperature of the two HCO^+ lines increases quickly with the column density. The lines are optically thick, implying $T_b \approx T_{ex} - T_{CMB}$, but the ratios of n_{H_2}/n_{crit}^{eff} is between 1 and 10^2 for the $(1-0)$ line and lower than 1 for the $(3-2)$ line, for almost all the values of column densities. These values are insufficient for LTE to be fully reached, i.e., $T_{ex} \neq T_{kin}$, and the brightness temperature still increases with the column density. Although the HCO^+ lines are optically thick, their higher critical density than those of the ^{12}CO lines maintain the non-LTE regime of the medium. While the $HCO^+(1-0)$ and $(3-2)$ are not thermometers of the gas, they provide precious information on the volume density of the main collider, i.e., n_{H_2}.

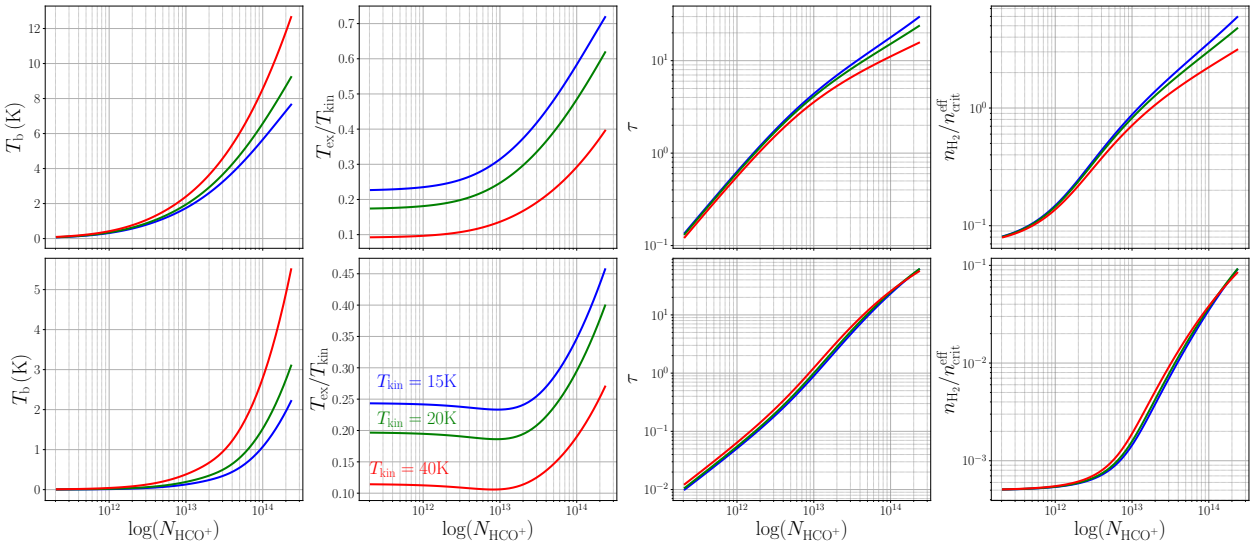


Figure 2.10: From left to right, variations with respect to the HCO^+ column density of the brightness temperature (T_b), the T_{ex}/T_{kin} ratio, the opacity (τ), and the n_{H_2}/n_{crit}^{eff} ratio. The computation are done for $n_{H_2} = 10^4$ cm^{-3}. The top row panels concern $HCO^+(1-0)$, and the bottom ones, $HCO^+(3-2)$.

Complementary behaviors of ^{12}CO and HCO^+ — In order to highlight the distinct behaviors of ^{12}CO and HCO^+, the left and right panels of Fig. 2.11 shows the $\frac{T_b(^{12}CO(2-1))}{T_b(^{12}CO(1-0))}$ and $\frac{T_b(HCO^+(3-2))}{T_b(HCO^+(1-0))}$ ratios as a function of n_{H_2} and T_{kin}. The brightness ratio of the ^{12}CO transitions shows constant T_{kin} lines regardless of the volume density while the ratio of the HCO^+ transitions shows almost constant n_{H_2} lines regardless of the gas kinetic temperature. These opposite behaviors confirm the previous observations: ^{12}CO informs us about the temperature while HCO^+ is more likely to inform us about the volume density. Thus, the study of these two different species allows us to obtain complementary information about the cloud.

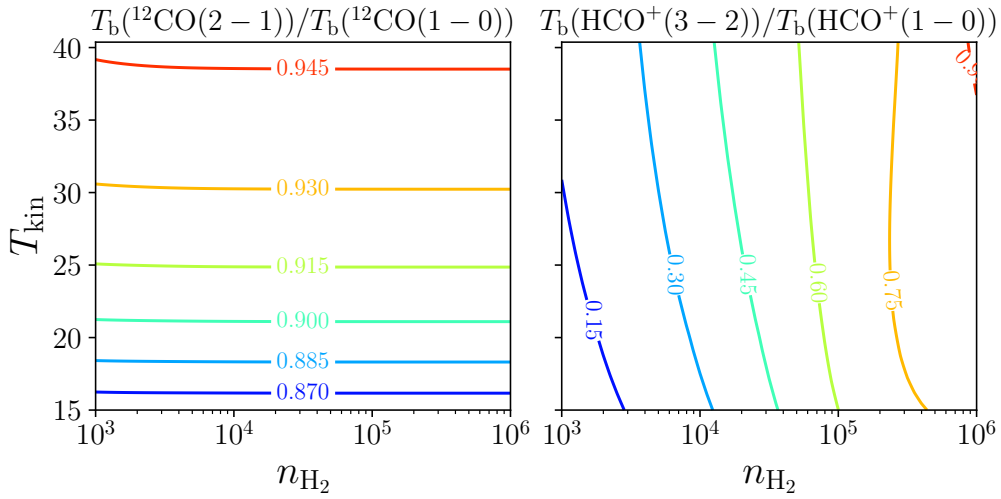


Figure 2.11: Ratios of the brightness temperature of $^{12}CO(2-1)$ over $^{12}CO(1-0)$ (left panel) and of $HCO^+(3-2)/HCO^+(1-0)$ (right panel) as a function of the kinetic temperature (T_{kin}) and volume density (n_{H_2}) of the gas. This figure illustrates that the low rotational lines of ^{12}CO are efficient thermometers of the gas while the low ones of HCO^+ are sensitive probes of the volume density. The computation are done for $N(^{12}CO) = 10^{19}\,cm^{-2}$ and $N(HCO^+) = 10^{14.32}\,cm^{-2}$.

References

Barnard, E. E. (Jan. 1919). "On the dark markings of the sky, with a catalogue of 182 such objects." In: *The Astrophysical Journal* 49, pp. 1–24. DOI: 10.1086/142439.

Bok, Bart J. and Edith F. Reilly (Mar. 1947). "Small Dark Nebulae." In: *The Astrophysical Journal* 105, p. 255. DOI: 10.1086/144901.

Buhl, D. and L. E. Snyder (1970). "Unidentified Interstellar Microwave Line". In: *Nature* 228.5268, pp. 267–269. DOI: 10.1038/228267a0.

Burton, W. B., B. G. Elmegreen, and R. Genzel (1992). "The galactic interstellar medium". In: *Saas-Fee Advanced Course 21: The Galactic Interstellar Medium*. Ed. by D. Pfenniger and P. Bartholdi. Springer-Verlag, p. 356.

Cheung, A. C. *et al.* (Dec. 1968). "Detection of NH_3 Molecules in the Interstellar Medium by Their Microwave Emission". In: *Phys. Rev. Lett.* 21 (25), pp. 1701–1705. DOI: 10.1103/PhysRevLett.21.1701.

Draine, Bruce T. (2011). "Physics of the Interstellar and Intergalactic Medium". In: Princeton University Press.

Fixsen, D. J. (Nov. 2009). "The temperature of the cosmic microwave background". In: *The Astrophysical Journal* 707.2, pp. 916–920. ISSN: 1538-4357. DOI: 10.1088/0004-637x/707/2/916.

Gratier, P. *et al.* (2013). "The IRAM-30 m line survey of the Horsehead PDR - III. High abundance of complex (iso-)nitrile molecules in UV-illuminated gas". In: *A&A* 557, A101. DOI: 10.1051/0004-6361/201321031.

Guzman Veloso, V. (Nov. 2013). "Physical and Chemical Conditions in the Horsehead Photodissociation Region". Doctoral Dissertation. Université Pierre et Marie Curie - Paris VI. URL: https://theses.hal.science/tel-00950116.

Guzmán, V. *et al.* (2011). "H2CO in the Horsehead PDR: photo-desorption of dust grain ice mantles". In: *A&A* 534, A49. DOI: 10.1051/0004-6361/201117257.

Herschel, W. (Jan. 1785). "On the Construction of the Heavens." In: *Philosophical Transactions of the Royal Society of London Series I* 75, pp. 213–266.

Kalberla, P. M. W. and U. Haud (2018). "Properties of cold and warm H I gas phases derived from a Gaussian decomposition of HI4PI data". In: *A&A* 619, A58. DOI: 10.1051/0004-6361/201833146.

Klemperer, W. (1970). "Carrier of the Interstellar 89.190 GHz Line". In: *Nature* 227.5264, pp. 1230–1230. DOI: 10.1038/2271230a0.

Lequeux, J., E. Falgarone, and C. Ryter (2002). "Le milieu interstellaire". In: EDP Sciences. "Savoirs Actuels", pp. 74–81, 167.

Liszt, H., Gerin, M., and Grenier, I. (2018). "Molecular gas and dark neutral medium in the outskirts of Chamaeleon". In: *A&A* 617, A54. DOI: 10.1051/0004-6361/201833167.

Mangum, J G. and Y. L. Shirley (2015). "How to Calculate Molecular Column Density". In: *The Astronomical Society of the Pacific* 127. DOI: 10.1086/680323.

McGuire, Brett A. (Dec. 2018). "2018 Census of Interstellar, Circumstellar, Extragalactic, Protoplanetary Disk, and Exoplanetary Molecules". In: *The Astrophysical Journal* 239.2, 17, p. 17. DOI: 10.3847/1538-4365/aae5d2.

Penzias, A. A., K. B. Jefferts, and R. W. Wilson (May 1971). "Interstellar $^{12}C^{16}O$, $^{13}C^{16}O$, and $^{12}C^{18}O$". In: *The Astrophysical Journal* 165, p. 229. DOI: 10.1086/150893.

Pety, J. *et al.* (2012). "The IRAM-30 m line survey of the Horsehead PDR - II. First detection of the l-C3H+ hydrocarbon cation". In: *A&A* 548, A68. DOI: 10.1051/0004-6361/201220062.

Snyder, L. E. and D. Buhl (1971). "Observations of radio emission from interstellar hydrogen cyanide". In: *Astrophysical Journal, vol. 163, p. L47* 163, p. L47.

Spitzer Jr., L. (May 1941). "The Dynamics of the Interstellar Medium. I. Local Equilibrium." In: *The Astrophysical Journal* 93, p. 369. DOI: 10.1086/144273.

Tielens, A. G. G. M. (2005). *The Physics and Chemistry of the Interstellar Medium*. Cambridge University Press.

van der Tak, F. F. S. *et al.* (Oct. 2010). *Radex: Fast Non-LTE Analysis of Interstellar Line Spectra*. Astrophysics Source Code Library, record ascl:1010.075.

Whipple, Fred L. (July 1946). "Concentrations of the Interstellar Medium." In: *The Astrophysical Journal* 104, p. 1. DOI: 10.1086/144829.

Wilson, R. W., K. B. Jefferts, and A. A. Penzias (July 1970). "Carbon Monoxide in the Orion Nebula". In: *The Astrophysical Journal* 161, p. L43. DOI: 10.1086/180567.

Yun, J. L. and Dan P. Clemens (Dec. 1990). "Star Formation in Small Globules: Bart BOK Was Correct!" In: *The Astrophysical Journal* 365, p. L73. DOI: 10.1086/185891.

Zucker, C. *et al.* (July 2019). "A Large Catalog of Accurate Distances to Local Molecular Clouds: The Gaia DR2 Edition". In: *The Astrophysical Journal* 879.2, p. 125. DOI: 10.3847/1538-4357/ab2388.

Chapter 3

Statistical inference

Contents

3.1 Introduction to estimation theory **40**
3.2 Cramér-Rao bound (CRB) **45**
3.2.1 Introduction 45
3.2.2 Definition and role in estimation theory 45
3.2.3 Practical use of the CRB 48
3.3 Maximum likelihood estimator (MLE) **49**
3.3.1 Asymptotic properties and implementation 49
3.3.2 Performance analysis of the MLE on simulated dataset 51
3.3.3 From simulated data to observation analysis 51
References **53**
Appendix 3 . . . . . **54**
3.A Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model 54

In the previous chapter, we presented the ORION-B dataset that consists in multi-line observations and illustrated how these data depend upon some parameters that characterize the underlying physics, chemistry and kinematics of the gas. Our aim is to *infer* these parameters of interest from the data. Moreover, in the astrophysical problem under consideration where no ground truth is available while the sought quantities may span several orders of magnitude[1], quantifying the confidence in the estimation results is as important as the inference task itself. This chapter presents the tools used in this thesis to meet these two aims. Section 3.1 sets the estimation problem under consideration, by introducing the key concepts of statistical inference, and provides an overview of the existing methods to solve it. Section 3.2 introduces the statistical tool used to provide confidence intervals for the estimation results. Section 3.3 details the estimator implemented in this thesis and provides an analysis of its performance.

General conventions — In this manuscript, scalars are written in the regular font, while vectors and matrices are denoted in the bold font. All vectors are column vectors (e.g., $\boldsymbol{x} = (x_1 \cdots x_N)^\top$, where $^\top$ denotes the transpose of $\boldsymbol{x}$). Capital letters are dedicated

[1] e.g., $n_{\mathrm{H_2}} \sim 10^{2-6}\,\mathrm{cm}^{-3}$ from diffuse gas to dense cores.

to random variables while their particular realizations are written in lowercase. $\mathbb{R}$ is the set of real numbers. The notation $\mathbb{R}^{r\times c}$ is used when introducing real-valued matrices composed of r rows and c columns, while $\mathbb{R}^r$ is used when dealing with vectors.

3.1 Introduction to estimation theory

Problem statement — To introduce the estimation problem under consideration, let us consider a single molecular line $\boldsymbol{x} = (x_1 \cdots x_N)^\top$ measured over N velocity channels toward one line of sight. Based on Sect. 2.3, one can construct a physical model $\boldsymbol{s} = (s_1 \cdots s_N)^\top$, such that in absence of noise[2] and modeling error, one would observe

$$\boldsymbol{x} = \boldsymbol{s}(\boldsymbol{\theta}),$$

where $\boldsymbol{\theta} \in \mathbb{R}^M$ is the vector of unknown parameters that characterize the gas. In particular, when considering a homogeneous gas, $\boldsymbol{s}(\boldsymbol{\theta})$ can be expressed with Eq. (2.10) as a function of five parameters, i.e., $\boldsymbol{\theta} = (T_{\rm kin}\ n_{\rm H_2}\ N_{\rm mol}\ \sigma_V\ C_V)^\top$, where $T_{\rm kin}$ is the kinetic temperature of the gas, $n_{\rm H_2}$ is the volume density of $\rm H_2$, $N_{\rm mol}$ is the column density of the considered molecule, σ_V is the velocity dispersion, and C_V is the centroid velocity of the line. However, measurements are subject to diverse disturbances which, due to their unpredictability, are considered as *random*. For instance, the atmospheric contribution can be taken into account with an additive white[3] Gaussian noise $\boldsymbol{B} = (B_1 \cdots B_N)^\top$, whose randomness is highlighted with the uppercase notation. The measurements $\boldsymbol{x}$ are then considered as the *realizations* of the random vector $\boldsymbol{X}$, defined by the probabilistic measurement model

$$\boldsymbol{X} = \boldsymbol{s}(\boldsymbol{\theta}) + \boldsymbol{B}, \tag{3.1}$$

where $\boldsymbol{B} \sim \mathcal{N}(\boldsymbol{0}_N,\ \sigma_b^2 \boldsymbol{I}_N)$, i.e., follows the normal distribution of mean $\boldsymbol{0}_N$, the null vector, and of covariance matrix $\sigma_b^2 \boldsymbol{I}_N$, where σ_b^2 is the standard deviation of the additive noise, and $\boldsymbol{I}_N$ is the identity matrix of size N. Our aim is to build a function that allows us to infer $\boldsymbol{\theta}$ from the measurements $\boldsymbol{x}$, i.e., an *estimator* of $\boldsymbol{\theta}$, denoted by $\widehat{\boldsymbol{\theta}}$,

$$\begin{array}{ccc} \mathbb{R}^N & \longrightarrow & \mathbb{R}^M \\ \boldsymbol{x} & \longmapsto & \widehat{\boldsymbol{\theta}}(\boldsymbol{x}). \end{array}$$

Since $\boldsymbol{B}$ is random, $\boldsymbol{X}$ is characterized by its *probability density function* (p.d.f.). This p.d.f. is denoted by $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$ to highlight its dependence on the unknown parameter $\boldsymbol{\theta}$. Indeed, one expects to obtain different distributions of the measurements $\boldsymbol{x}$ when observing various environments. To illustrate this, let us consider an observer who wants to infer $\theta = \log(n_{\rm H_2})$ from x, the line intensity measured at the rest frequency channel of $\rm H^{13}CO^+(1-0)$, a tracer of denser gas than its isotopologue $\rm HCO^+$. Based on the measurement model defined in Eq. (3.1), we show in Fig. 3.1a $p_X(x|\theta)$ when observing toward two environments that only differ in their volume density: a relatively diffuse gas ($\theta_1 = 3.5$) and a denser one ($\theta_2 = 5$). When observing toward diffuse gas, no signal is received since $s(\theta_1) \approx 0$. Based

[2] We refer to noise any source that does not depend on the physical parameter of interest $\boldsymbol{\theta}$.

[3] A random process of mean null whose samples are uncorrelated.

on Eq. (3.1), one therefore recovers in $p_X(x|\theta_1)$ (see the red distribution) the probabilistic model of the additive noise, i.e., a Gaussian distribution of zero mean[4]. Toward the dense gas, however, the species can be detected so higher intensity values for x are acquired (see the blue distribution). Nevertheless, recovering an accurate estimation of θ from a single realization x

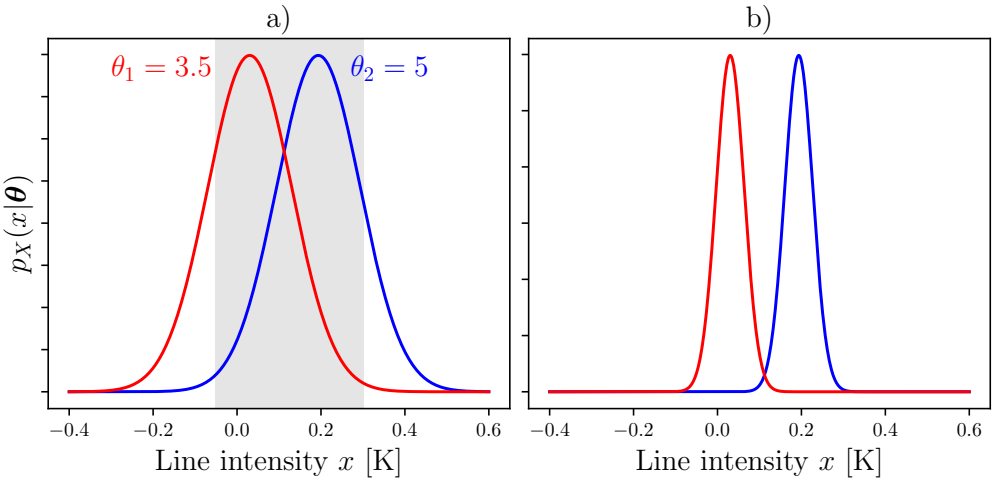


Figure 3.1: a) Dependence of the p.d.f. of X on the unknown parameter $\theta = \log n_{\mathrm{H_2}}$, based on the measurement model defined in Eq. (3.1). The p.d.f. $p_X(x|\theta_1)$ and $p_X(x|\theta_2)$ are plotted in red and blue, respectively. The gray region emphasizes the overlapping between the two distributions. b) Same layout than the panel (a) but for a S/N increased by a factor of 3.

of $p_X(x|\theta)$ is challenging because the two distributions are overlapping (as emphasized by the gray ambiguity area in Fig. 3.1a). To gain in estimation accuracy, a solution is to improve the S/N by, for instance, increasing the integration time. As shown in Fig. 3.1b, increasing the S/N by a factor of 3 reduces the noise level, and therefore the p.d.f. dispersion, sufficiently to remove any ambiguity in the estimation of θ. However, as discussed in Roueff *et al.* 2024, Sect. 5.2.2, increasing the integration time to achieve a better S/N for a given molecular line proves to be more expensive than providing additional observations of different molecular tracers.

Furthermore, in the case of the Orion B observation campaign, besides considering several molecular lines, we investigate the gas kinematics, i.e., the line dispersion (σ_V) and the centroid velocity (C_V), as well as physicochemical properties. We therefore examine the whole spectrum $\boldsymbol{x} = (x_1 \cdots x_N)^\top$ described by the measurement model defined in Eq. (3.1).

Constructing an estimator — Several techniques exist to build an estimator $\widehat{\boldsymbol{\theta}}(\boldsymbol{x})$. An intuitive one is to seek the value of $\boldsymbol{\theta}$ that minimizes the reconstruction error of $\boldsymbol{x}$ with the physical model $\boldsymbol{s}(\boldsymbol{\theta})$, i.e.,

$$\widehat{\boldsymbol{\theta}}_{\mathrm{LS}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{\theta}} \sum_{n=1}^{N} (x_n - s_n(\boldsymbol{\theta}))^2. \tag{3.2}$$

[4]Subtracting the baseline centers the emission around zero in signal-free spectral regions so that the additive noise contribution can induce negative intensity values.

One recognizes in Eq. (3.2) the distance between the measurements and the physical model. This approach leads to the *least squares* (LS) estimator. Formulated by Carl Friedrich Gauss in 1809, the LS method has notably enabled observers to deduce the trajectory of asteroids and planets (Gauss 1857). Its ease of implementation lies in the fact that no information on the measurements is required, making it particularly suitable when the probabilistic measurement model can not be formulated. When the latter can be found, another possibility is to implement the *maximum likelihood* estimator (MLE). As seen above, the set of measurements obtained for a given value of $\boldsymbol{\theta}$ can be described by $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$. Since $\boldsymbol{\theta}$ is unknown and $\boldsymbol{x} = (x_1 \cdots x_N)^\top$ have been measured, the MLE proposes to seek the most *likely* value of $\boldsymbol{\theta}$ underlying $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$, evaluated for the particular measurements $\boldsymbol{x}$. From this perspective, $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$ is no longer called the p.d.f. of the measurements $\boldsymbol{x}$ for a given $\boldsymbol{\theta}$, but is referred to as the *likelihood* function of $\boldsymbol{\theta}$ for a given $\boldsymbol{x}$. Thus, the MLE provides the value of $\boldsymbol{\theta}$ that maximizes the likelihood, i.e.,

$$\widehat{\boldsymbol{\theta}}_{\text{MLE}}(\boldsymbol{x}) = \arg\max_{\boldsymbol{\theta}} p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}).$$

In the case where measurements are Gaussian and statistically independent (see Eq. (3.1)), the likelihood evaluated for $\boldsymbol{x}$ reads

$$p_{\boldsymbol{X}}(\boldsymbol{x}|\theta) = \prod_{n=1}^{N} p_{X_n}(x_n|\boldsymbol{\theta}), \tag{3.3}$$

where

$$p_{X_n}(x_n|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\,\sigma_b^2}} \exp\left\{-\frac{1}{2\,\sigma_b^2}(x_n - s_n(\boldsymbol{\theta}))^2\right\}.$$

One then expresses Eq. (3.3) as a more convenient expression to manipulate, known as the *negative log-likelihood* (NLL), i.e.,

$$\begin{aligned} -\ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}) &= -\sum_{n=1}^{N} \ln p_{X_n}(x_n|\boldsymbol{\theta}) \\ &= \text{cst} + \frac{1}{2\,\sigma_b^2} \sum_{n=1}^{N} (x_n - s_n(\boldsymbol{\theta}))^2, \end{aligned} \tag{3.4}$$

such as the MLE estimator reads

$$\widehat{\boldsymbol{\theta}}_{\text{MLE}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{\theta}} \{-\ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})\}, \tag{3.5}$$

which means that in the Gaussian case, the MLE and the LSE are equivalent, and sometimes referred to as the χ^2 minimization technique.

In this thesis, the unknown parameter is considered *deterministic*, i.e., we assume that a *true* value of $\boldsymbol{\theta}$ exists, denoted by $\boldsymbol{\theta}_{\text{tr}}$, and we seek for an estimator such that $\widehat{\boldsymbol{\theta}}(\boldsymbol{x}) = \boldsymbol{\theta}_{\text{tr}}$. This perspective is a *frequentist* one. Other approaches exist to build an estimator. For instance, when the observer has some *prior* knowledge about the unknown parameter (e.g., a constrained range for n_{H_2}), he can adopt a *Bayesian* approach, which consists in refining

his knowledge about the unknown parameter according to the measurements. In such a perspective, he considers the unknown parameter as a random variable $\boldsymbol{\Theta}$ to quantify his *uncertainty* before and after having measured $\boldsymbol{x}$. We shall consider this Bayesian point of view when taking into account in the probabilistic measurement model the calibration uncertainty as a *nuisance*[5] parameter (see Chap. 5). Furthermore, when the distribution of $\boldsymbol{X}$ cannot be approximated properly by a parametric family of p.d.f., but a labelled dataset of k samples $(\boldsymbol{x}_k, \boldsymbol{\theta}_k)$ is available, one can perform regression with machine learning algorithms such as random forest or neural network.

Estimator's performance characterization — To choose the most appropriate estimator for a given problem among the various existing ones, one needs to define a criterion that characterizes the estimator's performance *statistically*, i.e., by taking into account all the measurements that can be obtained. Thus, while the estimation task is done on the measurements $\boldsymbol{x}$, the estimator's performance analysis is carried out with the probabilistic measurement model that describes them. One therefore examines $\widehat{\boldsymbol{\theta}}$ as a function of the random vector $\boldsymbol{X}$, i.e.,

$$\begin{array}{ccc} \mathbb{R}^N & \longrightarrow & \mathbb{R}^M \\ \boldsymbol{X} & \longmapsto & \widehat{\boldsymbol{\theta}}(\boldsymbol{X}). \end{array}$$

Hence, although $\widehat{\boldsymbol{\theta}}$ is a deterministic function, $\widehat{\boldsymbol{\theta}}(\boldsymbol{X})$ is random and is thus characterized by its p.d.f. $p_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})$, where we hereafter denote $\widehat{\boldsymbol{\theta}}(\boldsymbol{X})$ by $\widehat{\boldsymbol{\theta}}$ to keep the notation simple. For instance, in the Gaussian case where $\widehat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}})$, the distribution of the estimator reads

$$p_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \frac{1}{\sqrt{\det\left(2\pi\,\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}\right)}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}-\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}}\right)^{\top} \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}^{-1}\left(\boldsymbol{\theta}-\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}}\right)\right\}. \tag{3.6}$$

Here, $p_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})$ is completely characterized by its two first statistical moments: its mean $\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}}$ and its covariance matrix $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$. Statistical moments are computed with the *expectation*, denoted by $\mathbb{E}$,

$$\mathbb{E}\left[f\left(\widehat{\boldsymbol{\theta}}\right)\right] = \int f\left(\boldsymbol{\theta}\right)\, p_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}, \tag{3.7}$$

where f is a given function. As Eq. (3.7) shows, the expectation is performed over all the values of $\boldsymbol{\theta}$ weighted by their probability of occurrence $p_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})$. The k-order moment are obtained for $f(\boldsymbol{\theta}) = \boldsymbol{\theta}^k$. In particular, when f is the identity function ($k = 1$), one obtains the mean of $\widehat{\boldsymbol{\theta}}$, i.e., $\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}} = \int \boldsymbol{\theta}\, p_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}$. The estimator's covariance $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ quantifies the mean squared *dispersion* from the mean $\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}}$, i.e.,

$$\begin{aligned} \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} &= \mathbb{E}\left[(\widehat{\boldsymbol{\theta}}-\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}})(\widehat{\boldsymbol{\theta}}-\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}})^{\top}\right] \\ &= \underbrace{\mathbb{E}\left[\widehat{\boldsymbol{\theta}}\,\widehat{\boldsymbol{\theta}}^{\top}\right]}_{2^{\text{nd}}\text{ moment of }\widehat{\boldsymbol{\theta}}} - \boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}}\,\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}}^{\top}, \end{aligned}$$

[5] An unknown parameter that is not of interest when estimating the parameters of the physical model.

and, as such called the second *central* moment of $\widehat{\boldsymbol{\theta}}$. This latter quantity characterizes the estimator's *precision*, and it is desired to be as small as possible. In theory, the estimator's performance is characterized by the p.d.f. of $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{\text{tr}}$. In practice, we limit the characterization of this p.d.f. with its first two moments. The first one, i.e., the mean of the estimation error is referred to as the estimator's *biais*, and reads

$$\begin{aligned}\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_{\text{tr}}} &= \mathbb{E}\left[\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{\text{tr}}\right] \\ &= \mathbb{E}\left[\widehat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_{\text{tr}} \\ &= \boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}} - \boldsymbol{\theta}_{\text{tr}}.\end{aligned}$$

The bias therefore quantifies how far the estimator is from the true value in *average*. In general, one searches for an *unbiased* estimator, i.e., of bias null, so that $\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}} = \boldsymbol{\theta}_{\text{tr}}$. The second moment of the estimation error is referred to as the *mean squared error* (MSE) since it measures the average of the squared dispersion from the true value, i.e.,

$$\begin{aligned}\text{MSE} &= \mathbb{E}\left[(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{\text{tr}})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{\text{tr}})^\top\right] \\ &= \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} + \boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_{\text{tr}}}\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_{\text{tr}}}^\top. \end{aligned} \tag{3.8}$$

As shown in Eq. (3.8), the MSE delivers a measure of both the precision ($\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$) and the bias ($\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_{\text{tr}}}$) of the estimator. One therefore considers that it quantifies the overall performance of $\widehat{\boldsymbol{\theta}}$, which we refer to as the estimator's *accuracy*. Figure 3.2 summarizes the notion of bias and precision in the Gaussian scalar case. In 1D, the covariance of $\widehat{\theta}$ is called the *variance* ($\mathbb{V}_{\widehat{\theta}}$) while the dispersion ($\sigma_{\widehat{\theta}}$) refers to the square root of the variance, i.e., $\sigma_{\widehat{\theta}}^2 = \mathbb{V}_{\widehat{\theta}}$. In this thesis, we refer to the dispersion as the estimator's *precision*.

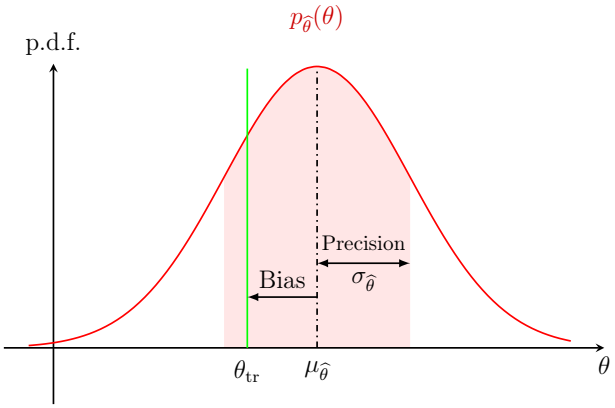


Figure 3.2: Illustration of the characteristics of the estimator's performance in the Gaussian scalar case. The *bias* of the estimator $\widehat{\theta}$ quantifies the distance between the estimator's mean ($\mu_{\widehat{\theta}}$) and the true value (θ_{tr}). The estimator's *precision* ($\sigma_{\widehat{\theta}}$) corresponds to the root square of the variance ($\mathbb{V}_{\widehat{\theta}}$).

Since the MSE takes into account both the bias and the precision of the estimator, one might consider to search for the estimator that minimizes the MSE. However, as discussed in

Kay 1993, Sect. 2.4, the bias generally depends on the unknown parameter $\boldsymbol{\theta}$, which makes the construction of the optimal estimator in sense of the MSE unrealizable. To overcome this issue, a standard solution is to restrict oneself to the family of unbiased estimators, i.e., $\widehat{\boldsymbol{\theta}}$ such that $\boldsymbol{\mu}_{\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_{\text{tr}}} = 0$. As Eq. (3.8) shows, minimizing the MSE then boils down to minimizing the estimator's covariance $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$, i.e., to search for the minimum variance unbiased estimator (MVUE). However, demonstrating that a given estimator is indeed the MVUE is challenging since the existence of such estimator is not guaranteed. Fortunately, one can at least derive lower *bounds* on the variance of a given family of estimators.

3.2 Cramér-Rao bound (CRB)

In the following, we introduce the lower bound used in this thesis, known as the Cramér-Rao (lower) Bound (CRB), which provides a lower bound on the variance of any unbiased estimator.

3.2.1 Introduction

The Cramér-Rao bound (CRB) and its inverse, the *Fisher information* are detailed in every textbook of statistical inference (e.g., Kay 1993; Van Trees 2001). Indeed, they are commonly used in practice because they are relatively simple to compute and, as we shall illustrate it in this section, they allow us to optimize the system's performance and inform us whether all available information has been extracted from data.

3.2.2 Definition and role in estimation theory

Fisher information and Cramér-Rao bound — Let us consider an observation $\boldsymbol{x}$ drawn from a parametric model $\boldsymbol{X} \sim p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\theta_1 \cdots \theta_M)^\top$, meaning that the physical model $\boldsymbol{s}(\boldsymbol{\theta})$ depends on M unknown parameters. The Fisher information, denoted by $\boldsymbol{\mathcal{I}}$, is the matrix of size $M \times M$ whose term i, j is given by

$$[\boldsymbol{\mathcal{I}}(\boldsymbol{\theta})]_{i,j} = -\mathbb{E}\left[\frac{\partial^2 \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right] \quad \forall i, j \in \{1, ..., M\}^2, \tag{3.9}$$

where the expectation is performed over $p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta})$. Under the following mild regularity conditions,

- the support of $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$ is independent of $\boldsymbol{\theta}$,
- $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$ is differentiable and integrable, and these two operators can commute,

a fundamental property can be demonstrated. For any unbiased estimator $\widehat{\boldsymbol{\theta}}$, its covariance matrix $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ verifies (see the proof in Appendix 3A of Kay 1993)

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} \geq \boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta}), \tag{3.10}$$

where the inverse of the fisher information matrix (FIM) is called the Cramér-Rao (lower) Bound (CRB), and is denoted by $\boldsymbol{\mathcal{B}}(\boldsymbol{\theta})$, i.e.,

$$\boldsymbol{\mathcal{B}}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta}). \tag{3.11}$$

An unbiased estimator that reaches this bound, i.e., $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} = \boldsymbol{\mathcal{B}}(\boldsymbol{\theta})$, is referred to as an *efficient* estimator. Such an estimator provides the minimum variance among all unbiased estimators, i.e., the MVU estimator introduced in the previous section. Because the inequality defined in Eq. (3.10) concerns matrices, its interpretation is not trivial. Let us therefore give more insight about its meaning.In the case of a single unknown parameter θ ($M = 1$), the variance of $\widehat{\theta}$ simply verifies

$$\mathbb{V}_{\widehat{\theta}} \geq 1/\mathcal{I}(\theta),$$

which means that the more Fisher information, the better the precision. When $M > 1$, Eq. (3.10) means that for any vector $\boldsymbol{v} \in \mathbb{R}^M$, one has

$$\boldsymbol{v}^\top \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}\, \boldsymbol{v} \geq \boldsymbol{v}^\top \boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\, \boldsymbol{v}.$$

For instance, assuming that θ_1 (the first component of $\boldsymbol{\theta}$) is the parameter of interest, choosing $\boldsymbol{v} = (1\,0\cdots 0)^\top$ allows one to derive a bound on the variance of interest[6].

$$\mathbb{V}_{\widehat{\theta}_1} \geq \left[\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right]_{1,1},$$

where $\left[\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right]_{1,1}$ is the first diagonal term of the inverse of the FIM. If all the unknown parameters are of interest, one can choose to focus on the trace inequality,

$$\mathrm{tr}\left(\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}\right) \geq \mathrm{tr}\left(\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right),$$

which is obtained for $\boldsymbol{v}^\top = (1\cdots 1)^\top$ and corresponds to

$$\sum_{m=1}^{M} \mathbb{V}_{\widehat{\theta}_m} \geq \sum_{m=1}^{M} \left[\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right]_{m,m}.$$

However, the trace may not be suitable when the set of unknown parameters is heterogeneous in terms of units, as in our case where $\boldsymbol{\theta}$ is composed of diverse physical properties (expressed in K, cm^{-2}, and cm^{-3}) and kinematics ($\mathrm{km\,s}^{-1}$). Furthermore, focusing on the diagonal terms of the CRB implies neglecting the information provided by the off-diagonal terms, which for efficient estimators, characterize the correlation between $\widehat{\theta}_i$ and $\widehat{\theta}_j$ ($i \neq j$). To take into account these correlations, one can consider the determinant inequality

$$\det\left(\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}\right) \geq \det\left(\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right).$$

The determinant can be interpreted as the hypervolume bounding the region of confidence characterized by the covariance matrix $\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})$. In the 2D case ($M = 2$), for instance, it can be shown that the surface of confidence of a Gaussian random vector with a mean $\boldsymbol{\theta}$ and a

[6]Note that $\left[\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right]_{1,1} \neq 1/\left[\boldsymbol{\mathcal{I}}(\boldsymbol{\theta})\right]_{1,1}$

covariance $\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})$ is an ellipse centered on $\boldsymbol{\theta}$ as illustrated in Fig. 3.3. Here, $\boldsymbol{\theta} = (\theta_1\, \theta_2)^\top$ such that

$$\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{B}(\theta_1) & \left[\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right]_{2,1} \\ \left[\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right]_{1,2} & \mathcal{B}(\theta_2) \end{pmatrix}, \tag{3.12}$$

where the diagonal term $\mathcal{B}(\theta_i) = \left[\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})\right]_{i,i}$ provides the lower bound on the variance $\mathbb{V}_{\widehat{\theta}_i}$ on each component of $\boldsymbol{\theta}$, i.e.,

$$\mathbb{V}_{\widehat{\theta}_i} \geq \mathcal{B}(\theta_i) \qquad \forall i \in \{1, 2\}.$$

In this thesis, since the physical parameters are all of interest, we shall extract from the inverse of the FIM its diagonal terms $\mathcal{B}(\theta_i)$. We shall refer to these terms as the CRB of the considered parameter.

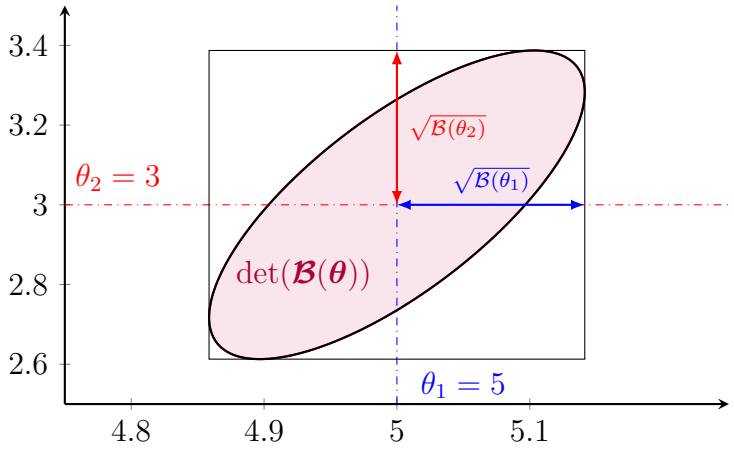


Figure 3.3: Illustration of the geometrical interpretation of the components of the CRB matrix $\boldsymbol{\mathcal{B}}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})$ in the 2D case, where $\boldsymbol{\theta} = (\theta_1\, \theta_2)^\top$. The diagonal terms of the $\boldsymbol{\mathcal{B}}(\boldsymbol{\theta})$, $\mathcal{B}(\theta_1)$ and $\mathcal{B}(\theta_2)$, allow one to construct the confidence intervals given by the rectangle of length $2 \times \sqrt{\mathcal{B}(\theta_1)}$ and width $2 \times \sqrt{\mathcal{B}(\theta_2)}$, while the determinant of the CRB, $\det\left(\boldsymbol{\mathcal{B}}(\boldsymbol{\theta})\right)$, provides a tighter confidence region corresponding to the ellipse's surface.

Other performance bounds — As mentioned in the introduction of this section, the CRB is one of the most popular bound because can be easily computed. However, its main drawback lies in the fact that it is not always reachable, i.e., there may be no estimator that reaches it. Indeed, as Eq. (3.3.1) shows, the FIM is the opposite of the expectation of the Hessian evaluated at the true parameter $\boldsymbol{\theta}$, which implies that it only *locally* characterizes the p.d.f. of the measurements without taking into account its global shape. To overcome this limitation, other performance bounds have been derived, such as the Bhattacharyya and Barankin bounds (Bhattacharyya 1946, Barankin 1949). These other bounds provide higher, and thus more powerful lower bounds. However, their derivation is more complex because they impose more restrictions on the class of estimators. These constraints are imposing either zero higher-order derivatives of the bias, or zero bias for several test points or over an interval. As these additional constraints are complicated to compute, the associated bounds are less used than the CRB.

3.2.3 Practical use of the Cramér-Rao bound

Woodward 1953 computed the CRB in the echolocalisation (i.e., RADAR[7]) case, and showed that the estimation precision of an object distance is proportional to the S/N and inversely proportional to the frequency bandwidth. At this time, one knew that to detect more distant sources, one has to increase the S/N and therefore the emitted power. However, it was unknown that increasing the bandwidth is crucial as well. This particular study of Woodward can thus be considered as a turning point from the energy era (maximization of the S/N) to the *information* era (maximization of the Fisher information). A first fundamental use of the CRB is therefore to guide one in the conception a system, since it quantifies the system's sensitivity with respect to the user's task, i.e., parameter estimation. Of course, such an optimization can also be performed using a given estimator $\widehat{\boldsymbol{\theta}}$. The main interest of using the CRB instead of a given $\widehat{\boldsymbol{\theta}}$ is that the optimization is then *independent* of the choice of the estimator. It depends solely upon the quantity of information contained in the measurements. In Roueff *et al.* 2024, Sect. 5, we investigate the sensitivity of the first two transition lines of ^{13}CO with respect to the volume density ($n_{\mathrm{H_2}}$), the kinetic temperature (T_{kin}), and the molecule's column density. In particular, we obtain high CRB values for T_{kin} in physical regimes where the lines are optically thick and at LTE, and should thus be robust thermometers. The CRB actually allows us to detect that there is a degeneracy between kinetic temperature and volume density on the line shape of the low J lines of CO (see Fig. 10 of Roueff *et al.* 2024). The estimation of T_{kin} is therefore degraded. It then becomes obvious why observing higher J transitions of CO allows the astrophysicists to disentangle the two excitation effects.

As discussed in Sect. 3.1, a modeler has several approaches to build an estimator. Once he chose one, he can also use the CRB to benchmark his estimator's performance. If the (unbiased) estimator's variance reaches the CRB, the modeler can be sure that it is useless to search for another (unbiased) estimator since none will have a smaller variance. One can argue that some biased estimator with a lower MSE might exist. However, an efficient, unbiased estimator is usually sufficient to meet the user's needs. On the other hand, if the estimator's variance is, e.g., ten times higher than the CRB, it might suggest that a more precise estimator exists. The CRB therefore can help the modeler to evaluate whether it is worth searching for a more precise (unbiased) estimator, bearing in mind that it may be more complex to implement.

Ideally, one derives the formal expression of the CRB, as did Woodward for the RADAR. In this thesis, although we construct a parametric physical model $\boldsymbol{s}(\boldsymbol{\theta})$, we have, however, no parametric expression of the excitation temperature (T_{ex}) and of the opacity (τ) as the function of the unknown parameter $\boldsymbol{\theta}$, but derive these two quantities numerically with RADEX (see the end of Sect. 2.3.2). To derive the CRB, we therefore first compute the FIM, where the gradient over $\boldsymbol{\theta}$ are derived with a finite difference technique, before inverting it numerically. As addressed in the next section, we chose an estimator of $\boldsymbol{\theta}$ that is *asymptotically* efficient. We thus use the CRB to provide confidence intervals on our

[7] RAdio Detection And Ranging

estimation results $\widehat{\theta}_i \in \widehat{\boldsymbol{\theta}}$, such as

$$\boxed{\theta_i \simeq \widehat{\theta}_i \pm \sqrt{\mathcal{B}(\widehat{\theta}_i)},} \tag{3.13}$$

meaning that we assume $\mathcal{B}(\widehat{\theta}_i) \simeq \mathcal{B}(\theta_i)$. This choice allows us to not only provide the spatial maps of the estimated parameters, but also the maps of *uncertainties* on these parameters. In the next section, we detail how we implement the chosen estimator, which is the MLE (introduced in Sect. 3.1).

3.3 Maximum likelihood estimator (MLE)

3.3.1 Asymptotic properties and implementation

Asymptotic properties — We choose to implement the MLE ($\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}$) because of the following advantageous theoretical properties. The MLE is an asymptotically[8] unbiased and asymptotically efficient estimator. Its asymptotic regime is Gaussian, i.e.,

$$\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}} \sim \mathcal{N}(\boldsymbol{\theta}_{\mathrm{tr}},\, \boldsymbol{\mathcal{B}}(\boldsymbol{\theta}_{\mathrm{tr}})),$$

which makes the Eq. (3.13) suitable for providing confidence intervals when the Gaussian regime is reached. In other words, the MLE asymptotically converges toward the MVU. In addition, one can show that if an efficient unbiased estimator exists, the MLE provides it (Kay 1993).

Implementation — From the practical point of view a first challenge is to find the minimum of the NLL, which is a function of the outputs of the "black box" radiative transfer code RADEX (see Sect. 2.3.2). We therefore implemented an iterative optimization process to estimate $\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}$. As detailed in the Appendix C of Ségal *et al.* 2024, we estimate the MLE through a two-step procedure: we first perform a grid search to obtain an approximated solution for the MLE, before refining this solution with a gradient algorithm. The model of cloud developed in this thesis has 14 unknown parameters (see Sect. 4.5.3 of Ségal *et al.* 2024), excluding an exhaustive exploration of the associated 14−dimensional grid. As illustrated in Fig. 3.4, the NLL, i.e., $-\ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$, may have local minima. To avoid getting trapped in a local minimum, we randomly spread over the parameter space several initial guesses, referred to as *walkers* (denoted by $\theta_{\mathrm{init},1}$ and $\theta_{\mathrm{init},2}$ in Fig. 3.4a). The solution ($\theta_{\mathrm{end},2}$ in Fig. 3.4a) is then used to initialize the gradient descent (GD) algorithm (θ_{init} in Fig. 3.4b). During the GD, the estimated value of $\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}$ is updated following

$$\widehat{\boldsymbol{\theta}}^{i+1}_{\mathrm{MLE}} = \widehat{\boldsymbol{\theta}}^{i}_{\mathrm{MLE}} - \alpha \underbrace{\boldsymbol{\mathcal{I}}^{-1}(\widehat{\boldsymbol{\theta}}^{i}_{\mathrm{MLE}})}_{\text{inverse of the FIM}} \nabla_{\boldsymbol{\theta}} \{ \underbrace{-\ln p_{\boldsymbol{X}}(\boldsymbol{x}|\widehat{\boldsymbol{\theta}}^{i}_{\mathrm{MLE}})}_{\text{NLL}} \}, \tag{3.14}$$

where i is the iteration index, and α is a hyperparameter that allows us to adjust the step size at each iteration. Compared to a Newton-Raphson method, the *scoring* method (Kay

[8] When the number of samples is large. In our case, the samples correspond to velocity channels.

1993, Chap. 7) consists in replacing in Eq. (3.14) the Hessian matrix by the FIM. Indeed, the FIM is closely related to the Hessian matrix, i.e.,

$$[\boldsymbol{\mathcal{I}}(\boldsymbol{\theta})]_{i,j} = -\mathbb{E}\Bigg[\underbrace{\frac{\partial^2 \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}}_{\text{Hessian matrix component}}\Bigg] \quad \forall i,j \in \{1,...,M\}^2,$$

but the expectation calculus allows one to remove some random terms. Thus, the FIM provides an approximation of the Hessian which has the advantage of being less sensitive to the measurement noise, and can therefore be better conditioned. In particular, its eigenvalue cannot be negative, which is not the case of the Hessian matrix. When the FIM is, nevertheless, still ill-conditioned[9], we replace it by the Moore Penrose pseudo-inverse (see details in Appendix B.3 of Roueff *et al.* 2024 provided in Appendix 3.A of this manuscript). The closer the solution of the grid search is to the global minimum, the faster the convergence of the GD algorithm. The GD outcome therefore highly depends on the step sizes chosen to sample the grid. The user must adjust these step sizes according to the computation time that is considered sustainable. Further details about the settings for our NLL minimization are provided in Appendix C of Ségal *et al.* 2024.

Since the unbiased, efficient, and Gaussian behavior of the MLE is only guaranteed in the asymptotic regime (i.e., large number of velocity channels), one needs to analyze the estimator's performance through Monte Carlo experiments, at least for the physical regime of interest.

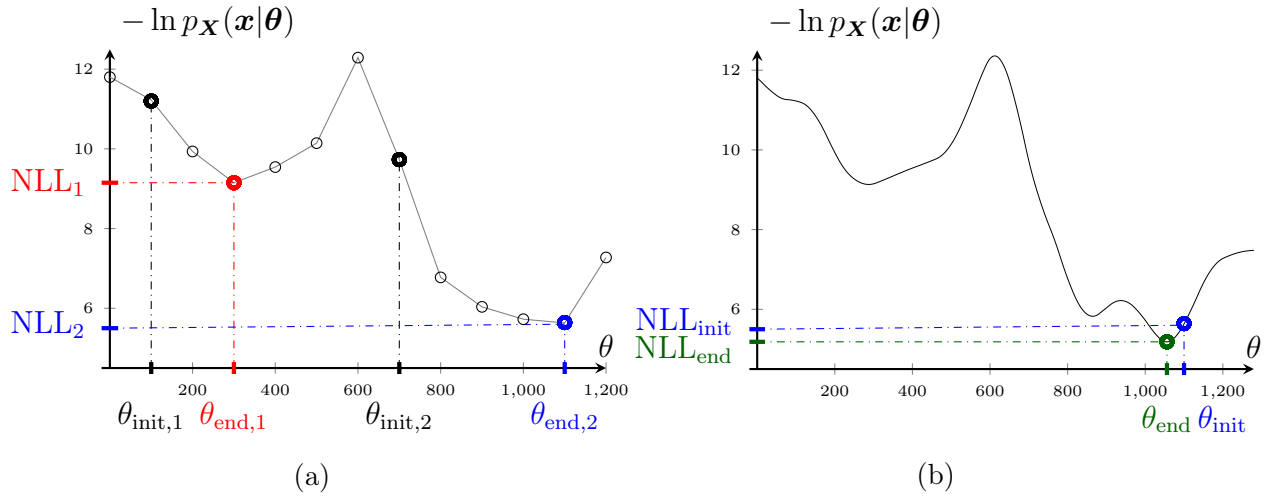


Figure 3.4: Illustration of the two-step procedure implemented to estimate the MLE: a grid search (panel a), followed by a gradient descent algorithm (panel b). The solution obtained at the end of the grid search ($\theta_{\mathrm{end},\,2}$ in panel a) is used to initialize the gradient algorithm (θ_{init} in panel b).

[9] In this thesis, when the ratio between the highest and the lowest eigenvalue is higher than 10^9.

3.3.2 Performance analysis of the MLE on simulated dataset

Monte Carlo experiment consists in analyzing the estimator's performance on simulated datasets. This allows one to quantify the bias and the variance of the estimator for a fixed *ground truth* ($\boldsymbol{\theta}_{\text{tr}}$). Indeed, to create a simulated dataset, one draws a large number P of samples $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_P\}$ from a given statistical model, i.e., $\boldsymbol{X} \sim p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}_{\text{tr}})$, where the chosen $\boldsymbol{\theta}_{\text{tr}}$ characterizes a physical regime of interest. In theory, this performance analysis has to be reproduced for all physical regimes of interest, but it may require too much computation time. Figure 3.5 illustrates how we interpret Monte Carlo results when analyzing our estimator's performance. When estimating $\boldsymbol{\theta} = (\theta_1 \cdots \theta_M)^\top$, i.e., M parameters, one can examine the estimation results for two particular components of $\boldsymbol{\theta}$, e.g., θ_1 and θ_2. In Fig. 3.5, each gray point corresponds to an estimation result $\{\widehat{\theta}_{1,p}, \widehat{\theta}_{2,p}\}$ obtained for the sample $\boldsymbol{x}_p$. We first assess the unbiasedness of the estimator by computing the empirical mean of the scatter plot, denoted by $\overline{\boldsymbol{\mu}}$,

$$\overline{\boldsymbol{\mu}} = \frac{1}{P}\begin{pmatrix} \sum_{p=1}^{P} \widehat{\theta}_{1,p} \\ \sum_{p=1}^{P} \widehat{\theta}_{2,p} \end{pmatrix}$$

which is displayed by the black cross in Fig. 3.5. The empirical covariance of the scatter plot, denoted by $\overline{\boldsymbol{\Sigma}}$, reads

$$\overline{\boldsymbol{\Sigma}} = \frac{1}{P-1}\sum_{p=1}^{P}\left(\widehat{\boldsymbol{\theta}}_{1\text{-}2,p} - \overline{\boldsymbol{\mu}}\right)\left(\widehat{\boldsymbol{\theta}}_{1\text{-}2,p} - \overline{\boldsymbol{\mu}}\right)^\top,$$

where $\widehat{\boldsymbol{\theta}}_{1\text{-}2,p} = \left(\widehat{\theta}_{1,p}\ \widehat{\theta}_{2,p}\right)^\top$. The covariance matrix of $\overline{\boldsymbol{\mu}}$, denoted by $\boldsymbol{\Sigma}_{\overline{\boldsymbol{\mu}}}$, is then given by

$$\boldsymbol{\Sigma}_{\overline{\boldsymbol{\mu}}} = \frac{1}{P}\overline{\boldsymbol{\Sigma}}. \tag{3.15}$$

This covariance matrix $\boldsymbol{\Sigma}_{\overline{\boldsymbol{\mu}}}$ allows us to construct a confidence ellipse on the estimator's empirical mean $\overline{\boldsymbol{\mu}}$, centered on $\overline{\boldsymbol{\mu}}$ (see the black ellipse in Fig. 3.5). For the case considered in Fig. 3.5, we observe that the ground truth is contained within this confidence ellipse. We therefore do not detect any estimator's bias with this simulation.

To assess the efficiency of the estimator, we draw a confidence ellipse centered on the ground truth $\{\theta_{\text{tr},1}, \theta_{\text{tr},2}\}$ whose shape is computed from the CRB matrix (plotted in purple in Fig. 3.5), as previously illustrated in Fig. 3.3. We then examine the proportion of samples within this ellipse. For the chosen 3σ dispersion, this proportion is 99% for an efficient estimator.

We perform such a Monte Carlo simulations in Sect. 3.3 of Ségal *et al.* 2025, which is provided in Appendix 5.F of this manuscript. In these simulations, the measured proportion of points within the CRB ellipse lies between 87% and 98% depending on the pair of parameters $\{\theta_1, \theta_2\}$ considered, which means that the estimator is "close" to the efficient regime.

3.3.3 From simulated data to observation analysis

Analyzing the estimator's performance on simulated data through Monte Carlo experiment allows us to assess that the implemented estimator is unbiased and efficient for a physical

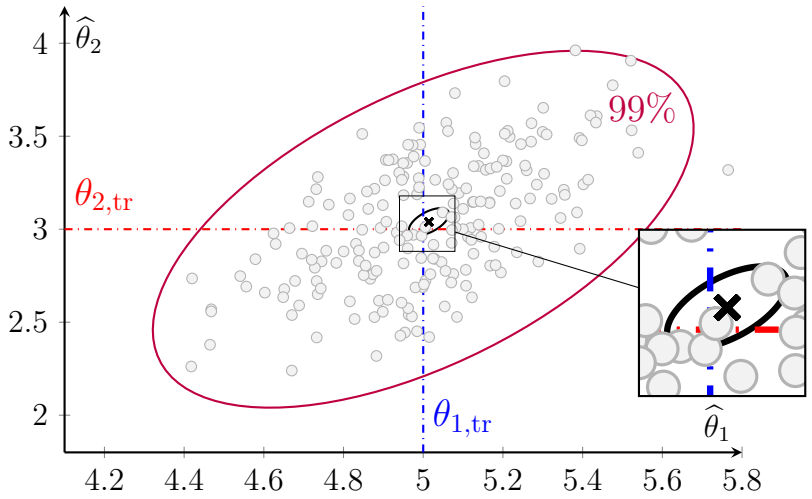


Figure 3.5: Illustration of the estimator's performance analysis through Monte Carlo experiment on a simulated dataset. Each point is an estimation $\widehat{\boldsymbol{\theta}}_p$ obtained for a simulated observation $\boldsymbol{x}_p$ drawn from $p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}_{\mathrm{tr}})$, where the ground truth $\boldsymbol{\theta}_{\mathrm{tr}}$ is namely composed of $\theta_{\mathrm{tr},1}$ and $\theta_{\mathrm{tr},2}$. The purple ellipse is centered on $\{\theta_{\mathrm{tr},1}, \theta_{\mathrm{tr},2}\}$, and its shape is computed from the CRB matrix (defined in Eq. (3.11)) with a dispersion of 3σ. The proportion of points within this ellipse of confidence is 99%. The estimator's empirical mean is indicated by the black cross. The confidence region for this mean is given by the black ellipse, computed from the empirical mean's covariance matrix (see Eq. (3.15)).

regime of interest, and within the assumption of the considered measurement model. When analyzing real observations, an additional difficulty, beyond the estimation task itself, is that the measurement model makes physicochemical assumptions that are too simple in relation to reality, which remains partly hidden with the limited amount of available observations. In particular, to reduce the number of unknowns, one can consider that all molecular species are bathed in the same cell of gas, i.e., that the cloud is homogeneous, which implies estimating only one kinetic temperature and volume density for all species. Such assumptions allow us to decrease the degree of freedom of the measurement model and therefore reduce the *variance* of the unknown parameters of interest, but at the expense of the apparition of estimation *biases*. The challenge for observers and modelers is therefore to identify the best physical assumptions to reach a suitable *trade-off* between the estimator's bias and variance, and according to the science purposes.

In Roueff *et al.* 2024, we investigate and quantify the estimator's bias and variance induced by two types of physical choices made when analyzing several molecular lines: 1) the selected molecular tracers and 2) the abundances ratios[10] assumed between these tracers. A first study is carried out on the real observations toward the Horsehead nebula, by assuming the cloud homogeneous. As Figs. 4, C.3, C.4 of this article show, we observe that the estimations of the column densities, the kinetic temperature, and the volume density depend significantly upon the subsets of lines considered and the injected priors on molec-

[10]The abundance of a species X is given by the ratio between its column density and the column density of a species of reference. In this thesis, H_2 is our species of reference.

ular abundances. As each molecular line can probe a specific physicochemical regime (see Sect. 2.4), such biased estimates may suggest the presence of *heterogeneous* physicochemical properties of the gas along the line of sight, i.e., assuming the cloud homogeneous induces unrealistic physical values with overly optimistic intervals of confidence. This conjecture is confirmed in Sect. 4 of Roueff *et al.* 2024, in which the estimator's performance is analyzed through Monte Carlo experiments on simulated data generated from distributions that differ purposely from the measurement model.

The main result reported in Roueff *et al.* 2024 was obtained a little before 2022, and actually motivated this thesis: since a homogeneous model is not suitable to describe the multi-molecular line observations of GMCs, we decided to increase the complexity of the physical model accordingly. In the next chapter, we propose a multilayer model of cloud in order to take into account the heterogeneity of the line of sight. We provide in Appendix 3.A the article of Roueff *et al.* 2024, which is the transition to the model on which I dedicated myself during this thesis. The article of Roueff *et al.* 2024 thus represents a significant part of my formation in both astrophysical and statistical fields, in particular, by checking the calculations of the FIM and CRB (see Appendix A of 3.A).

References

Barankin, E. W. (1949). "Locally Best Unbiased Estimates". In: *The Annals of Mathematical Statistics* 20.4. DOI: 10.1214/aoms/1177729943.

Bhattacharyya, A. (1946). "On a Measure of Divergence between Two Multinomial Populations". In: *Sankhyā: The Indian Journal of Statistics (1933-1960)* 7.4, pp. 401–406. ISSN: 00364452.

Gauss, C. F. (1857). *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*. Trans. by C. H. Davis. English translation of *Theoria motus corporum coelestium* (1809). Boston: Little, Brown and Company. URL: https://archive.org/details/motionofheavenly00gausrich.

Kay, S. (1993). *Fundamentals of statistical signal processing: estimation theory*. USA: Prentice-Hall, Inc. ISBN: 0133457117.

Roueff, A. *et al.* (2024). "Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model - Application to the estimation of the gas temperature and volume density". In: *A&A* 686, A255. DOI: 10.1051/0004-6361/202449148.

Ségal, L. *et al.* (2024). "Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead nebula". In: *A&A* 692, A160. DOI: 10.1051/0004-6361/202451567.

Ségal, L. *et al.* (Aug. 2025). "De la possibilité d'ignorer un bruit de calibration lors de la construction d'un estimateur du maximum de vraisemblance en radio-astronomie". In: *GRETSI : Groupe de Recherche et d'Etudes de Traitement du Signal et des Images 2025*. Starsbourg, France. URL: https://hal.science/hal-05332727.

Van Trees, H. L. (2001). In: *Detection, Estimation, and Modulation Theory: Detection, Estimation, and Linear Modulation Theory*. John Wiley & Sons, Ltd. Chap. 2, pp. 19–165. DOI: https://doi.org/10.1002/0471221082.ch2.

Woodward, P. (1953). "Theory of radar information". In: *Transactions of the IRE Professional Group on Information Theory* 1.1, pp. 108–113. DOI: 10.1109/TIT.1953.1188560.

Appendix 3

3.A Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model

A&A, 686, A255 (2024)
https://doi.org/10.1051/0004-6361/202449148
© The Authors 2024

Astronomy & Astrophysics

Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model

Application to the estimation of the gas temperature and volume density

Antoine Roueff[1], Jérôme Pety[2,3], Maryvonne Gerin[3], Léontine E. Ségal[2,1], Javier R. Goicoechea[4], Harvey S. Liszt[5], Pierre Gratier [6], Ivana Beslic[3], Lucas Einig[2,7], Mathilde Gaudel[3], Jan H. Orkisz[8], Pierre Palud[9,10], Miriam G. Santa-Maria[4], Victor de Souza Magalhaes[2], Antoine Zakardjian[11], Sébastien Bardeau[2], Emeric Bron[10], Pierre Chainais[9], Simon Coudé[15,16], Karine Demyk[11], Viviana V. Guzman[12], Annie Hughes[11], David Languignon[10], François Levrier[13], Dariusz C. Lis[14], Jacques Le Bourlot[10], Franck Le Petit[10], Nicolas Peretto[17], Evelyne Roueff[10], Albrecht Sievers[2], and Pierre-Antoine Thouvenin[9]

(Affiliations can be found after the references)

Received 2 January 2024 / Accepted 19 February 2024

ABSTRACT

Context. Robust radiative transfer techniques are requisite for efficiently extracting the physical and chemical information from molecular rotational lines.

Aims. We study several hypotheses that enable robust estimations of the column densities and physical conditions when fitting one or two transitions per molecular species. We study the extent to which simplifying assumptions aimed at reducing the complexity of the problem introduce estimation biases and how to detect them.

Methods. We focus on the CO and HCO^+ isotopologues and analyze maps of a 50 square arcminutes field. We used the RADEX escape probability model to solve the statistical equilibrium equations and compute the emerging line profiles, assuming that all species coexist. Depending on the considered set of species, we also fixed the abundance ratio between some species and explored different values. We proposed a maximum likelihood estimator to infer the physical conditions and considered the effect of both the thermal noise and calibration uncertainty. We analyzed any potential biases induced by model misspecifications by comparing the results on the actual data for several sets of species and confirmed with Monte Carlo simulations. The variance of the estimations and the efficiency of the estimator were studied based on the Cramér-Rao lower bound.

Results. Column densities can be estimated with 30% accuracy, while the best estimations of the volume density are found to be within a factor of two. Under the chosen model framework, the peak ^{12}CO (1 − 0) is useful for constraining the kinetic temperature. The thermal pressure is better and more robustly estimated than the volume density and kinetic temperature separately. Analyzing CO and HCO^+ isotopologues and fitting the full line profile are recommended practices with respect to detecting possible biases.

Conclusions. Combining a non-local thermodynamic equilibrium model with a rigorous analysis of the accuracy allows us to obtain an efficient estimator and identify where the model is misspecified. We note that other combinations of molecular lines could be studied in the future.

Key words. line: profiles – radiative transfer – methods: data analysis – methods: statistical – ISM: clouds – ISM: general

1. Introduction

Recent progress in developing receiver and backend technologies has now made it possible for a significant proportion of a giant molecular cloud to be imaged with enough sensitivity to simultaneously map dozens of molecular lines. Nishimura et al. (2017) and Watanabe et al. (2017) observed the W3(OH) and W51 molecular clouds with the Nobeyama 45m telescope over a significant fraction of the 87–112 GHz spectral range, while Barnes et al. (2020) studied the W49N massive star-forming region with the IRAM 30m telescope in two spectral ranges: 86.1–99.9 GHz and 101.8–115.6 GHz. We studied the Orion B star-forming cloud with the IRAM 30 m telescope within the framework of the project named Outstanding Radio-Imaging of OrioN-B (ORION-B, PI: Jérôme Pety & Maryvonne Gerin). The resulting data cube covers 13×18 pc on the sky and a bandwidth of 40 GHz at a typical resolution of 50 mpc and 0.6 km s^{-1} and at a typical sensitivity of 0.1 K in the main beam temperature scale, as described by Pety et al. (2017) in their discussion of results from a subset of 5.6×7.5 pc. These scales are derived assuming a distance to the Orion B cloud of 400 pc, as discussed in Pety et al. (2017) and confirmed by more recent analyses based on *Gaia* distances who found a distance between 397 and 410 pc (Zucker et al. 2021; Cao et al. 2023). Extracting physical information from such large spectroscopic datasets requires the use of robust and fast radiative transfer methods that can be applied with limited human supervision. Large imaging datasets with thousands of pixels usually include bright spectral lines in a specific spectral range, for instance, the 3 mm spectral window. The number of mapped lines for a given species is then usually limited to one. Extracting quantitative information therefore represents a different challenge as compared to studies of specific objects where several transitions of a few molecular species can be obtained on a restricted field of view. For this reason, and

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

because the method is simple and fast, the local thermodynamic equilibrium (LTE) framework is often used to provide estimations of molecular column densities and excitation temperature over the covered large scale maps. As shown by Roueff et al. (2021), this simple hypothesis may lead to bias in the derived parameters as the excitation temperatures of the three main CO isotopologues are different, the excitation temperature of ^{12}CO being typically twice as large as that of ^{13}CO and C^{18}O. In general, the excitation temperature of the three CO isotopologues are different with $T_{\rm ex}(^{12}{\rm CO}) \geq T_{\rm ex}(^{13}{\rm CO}) \geq T_{\rm ex}({\rm C}^{18}{\rm O})$. If it is not taken properly into account, this difference may induce significant errors in the derivation of molecular column densities and relative abundances.

Furthermore, the hypothesis that the level populations of the considered species are at thermal equilibrium is often proven to be invalid because the volume densities are too low for collisions to be fully effective in balancing the populations of the energy levels (see, e.g., Shirley 2015). Non-LTE excitation and radiative transfer models have thus been developed for decades to analyze the molecular line emission and have been widely used in different contexts, most often by adding a priori information on some properties of the considered source (e.g., its structure) to limit the parameter space that is meant to be explored. When studying a specific object, an analytical model of the geometry, density structure, and velocity field of the object is often specified. The information in the line intensities is then used to fit the velocity field, volume density, temperature structure, and the molecular column densities. Additional hypotheses on the temperature can also be made, either by using additional information (e.g., the kinetic temperature equal to the dust temperature) or by fixing the kinetic temperature to some given value. For instance, Tafalla et al. (2021) used an isothermal model with simple parametric laws for the variation of the density, velocity dispersion, and molecular abundances as a function of the position to analyze their survey of the Perseus molecular cloud. Outputs of thermo-chemical models such as the Meudon PDR code (Le Petit et al. 2006) can also be used as inputs for radiative transfer calculations. Using the volume density, kinetic temperature, and relative abundances of CS and CO isotopologues from a plane-parallel PDR model of the Horsehead nebula, Goicoechea et al. (2006) computed the expected molecular emission in an edge-on cloud with a nonlocal and non-LTE Monte Carlo radiative transfer model, and set the derived constraints on the sulfur abundance. These examples show that this approach works well for known sources with many observations across a broad range of wavelengths accessible because it requires substantial a priori information.

For sources with less available information, especially when a simple geometry cannot be specified, fitting molecular line emission may lead to ambiguities and uncertainties in the derived parameters. In their analysis of the (1 – 0) lines of ^{12}CO, ^{13}CO, and C^{18}O in the Orion A molecular cloud, Castets et al. (1990) showed that the correlation between the estimated kinetic temperature and the volume density may lead to ambiguities and large uncertainties in these parameters. For this particular cloud, using the peak ^{12}CO intensity as a constraint for the kinetic temperature offers relatively poor results because different values of the kinetic temperature ought to be used for the different CO isotopologues. Castets et al. (1990) assumed that these different values of the kinetic temperature are spatially uniform across the cloud and deduced them from observations of the ammonia inversion lines towards small sub-regions of the map. Hence, these examples show that even with a non-LTE model, the inversion of the radiative transfer equation is not solely based on the information provided by the considered molecular lines; rather, it makes use of additional a priori information, which may be derived from the analysis of additional datasets or from a geometric or theoretical model. The uncertainties on the derived parameters are then more difficult to evaluate as they depend both on the quality of the data and the quality of the model, including the additional a priori assumptions.

The question of accuracy of the retrieval of physical conditions from a set of observed noisy lines was discussed by Tunnard & Greve (2016) for the observations of CO or HCN lines in distant galaxies. In this study, the authors used relatively optimistic observation conditions with a noise level set either to 0 or to 10% of the line intensity. They showed that the accuracy of the estimated parameters (density, temperature, and pressure) is not better than half a dex (i.e., a factor of three) and that using isotopologue lines helps to increase the accuracy of the retrieval. For the considered species and lines, they show that better results are obtained when the isotopologue abundance ratio is not fixed because introducing a fixed value may bias the results and lead to satisfactory fits even though the fitted parameters have significant offsets from their true values. Tunnard & Greve (2016) analyzed CO and HCN separately and did not attempt to combine the information from these two species.

To study the bias in the gas column density determination based on CO (1 – 0) rotational line emission, Teng et al. (2023) discussed how the (1 – 0), (2 – 1), and (3 – 2) CO isotopologue emission, and the CO column density in three nearby galaxies depend on the physical conditions. Using a large grid of RADEX models, they derived (for each pixel) the volume density, kinetic temperature, CO column density per unit line width, and the isotopic ratio between ^{12}CO and ^{13}CO and between ^{13}CO and C^{18}O. An additional constraint on the path length along the line of sight was added to remove models with low volume density and very large CO column densities. In these models, the CO abundance relative to H_2 is fixed. The fitted results are used to estimate the H_2 column density from the CO column density and the conversion factor between the ^{12}CO (1 – 0) intensity and the H_2 column density for each pixel. They showed that the variations of this conversion factor are mainly related to variations of the CO optical depth and secondly to variations of the kinetic temperature. The necessity to constrain the length along the line of sight illustrates once again the presence of ambiguities in the derivation of the physical conditions and the possibility of non physical results. An anti-correlation between the volume density and kinetic temperature is clearly seen in their likelihood distributions.

As in the case of LTE models, non-LTE models may also suffer from bias induced by a priori information, but such effects are not sufficiently documented in studies of non-LTE radiative transfer modeling. The present paper focuses on the quantitative uncertainties obtained for the volume density and kinetic temperature derived for the Horsehead nebula, when analyzing the 3 mm spectrum of the different isotopologues of CO and HCO^+. As shown by Bron et al. (2018) the combination of the (1 – 0) lines of the three main CO isotopologues and of HCO^+ is sufficient to identify the different regimes of volume density and FUV illumination across the mapped area. It is now necessary to transform this qualitative study into quantitative maps of physical parameters, using non-LTE radiative transfer models.

Non-LTE models can be developed for different source geometries. The simplest formulation, which is implemented in local escape probability codes such as RADEX (van der Tak et al. 2007), uses a uniform 1D source description. It solves the combined radiative transfer and molecular excitation in a

local way, where local means that molecules interact with the local radiation field and are fully decoupled from other radiation emission along any other position of the cloud. The emergent radiation results from the combination of all locally produced photons. More sophisticated radiative transfer codes can treat the non-local interaction between the excitation of molecular levels at a given position of the cloud and the radiation field from any other cloud position. for instance the MOLPOP-CEP code (Asensio Ramos & Elitzur 2018) or codes based on Monte Carlo methods (Bernes 1979; Goicoechea et al. 2006; Brinch & Hogerheijde 2010). These more sophisticated models naturally take into account the presence of density, kinetic temperature, or relative abundance gradients along the line of sight. Such non-local methods are however more complicated to implement, they are numerically heavier, and they require more a priori information about the structure of the considered cloud. They provide more accurate results than non-LTE local models, especially in the case when radiative effects become dominant; for instance, in the case of the radiative transport of optically thick lines in a medium where the volume density is much lower than the critical density.

The article is organized as follows. Section 2 presents the tools that will be used in this paper: the RADEX radiative transfer code, the Cramér-Rao bound (CRB) used to infer a reference precision, and a maximum likelihood estimator used to fit simultaneously several lines for each line of sight. After a succinct presentation of the data, Sect. 3 offers a comparison of the estimated parameters when fitting the data under different assumptions. We discuss the tradeoff between variance and bias on two experiments in Sect. 4. Section 5 presents a parametric study of the achievable precision on the column densities, kinetic temperature, volume density, and thermal pressure. Section 6 discusses the best strategy to maximize the precision without biasing the results. Appendices A and B present details about the computation of the Cramér-Rao bound and the maximum likelihood estimator. Appendices C and D show supplemental figures related to the discussion in Sects. 3 and 5.

2. Methods

This section introduces the astrophysical framework (the non-LTE escape probability radiative transfer code RADEX and the associated concept of critical density) and the statistical method (the Cramér-Rao bound and our maximum likelihood fitter) used throughout this paper.

2.1. RADEX: A non-LTE radiative transfer model to estimate the gas volume density and kinetic temperature

The non-LTE radiative transfer code RADEX (van der Tak et al. 2007) uses an escape probability formalism to compute the emerging radiation from an homogeneous medium of simple geometry. In this article, we use the uniform sphere geometry. Below, we clarify the associated notations and equations for a single species.

As recalled recently in Roueff et al. (2021), the brightness temperature, s, at observed frequency ν around the rest frequency of a line (ν_l) can be written as

$$s(\nu) = \{J(T_{\mathrm{ex},l}, \nu_l) - J(T_{\mathrm{CMB}}, \nu)\}\,[1 - \exp(-\Psi_l(\nu))]\,. \tag{1}$$

In this equation

Table 1. Origin of the collision coefficients used in this article.

Species	Collider	Database	Reference
^{12}CO	H_2	LAMDA [(1)]	Yang et al. (2010)
^{13}CO	H_2	LAMDA	Yang et al. (2010)
C^{18}O	H_2	LAMDA	Yang et al. (2010)
HCO$^+$	H_2	EMAA [(2)]	Denis-Alpizar et al. (2020)
HCO$^+$	e^-	EMAA	Fuente et al. (2008)
H^{13}CO$^+$	H_2	EMAA	Denis-Alpizar et al. (2020)
H^{13}CO$^+$	e^-	EMAA	Fuente et al. (2008)

Notes. [(1)]Leiden Atomic and Molecular Database (LAMDA) at https://home.strw.leidenuniv.nl/~moldata/. [(2)]Excitation of Molecules and Atoms for Astrophysics (EMAA) at https://emaa.osug.fr.

- J is defined as

$$J(T, \nu) = \frac{c^2}{2k\nu^2} B(T, \nu) = \frac{h\nu}{k} \frac{1}{\exp\frac{h\nu}{kT} - 1}, \tag{2}$$

  where $B(T, \nu)$ is the spectral distribution of the radiation of a blackbody at temperature, T, and h, k are the Planck and Boltzmann constants, respectively.
- As we assume that the only background source of emission is the cosmic microwave background (CMB), we compute J at its temperature ($T_{\mathrm{CMB}} = 2.73$ K).
- $T_{\mathrm{ex},l}$ is the excitation temperature related to the ratio of the population in the upper (n_{up}) and lower (n_{low}) levels of the line, l,

$$\frac{n_{\mathrm{up}}}{n_{\mathrm{low}}} = \frac{g_{\mathrm{up}}}{g_{\mathrm{low}}} \exp\left[-\frac{h\nu_l}{kT_{\mathrm{ex},l}}\right]. \tag{3}$$

  The population of the different energy levels is locally computed by solving the statistical balance between radiative and collisional excitation and de-excitation of the important energy levels. Table 1 lists the origin of the collisional coefficients used in this article.
- The term $[1 - \exp(-\Psi(\nu))]$ in Eq. (1) represents the emission (or absorption) by the emitting (or absorbing) medium along the line of sight (LoS). The function Ψ is the opacity profile that corresponds to the integrated opacity through the whole slab. The RADEX code computes the radiative transfer assuming that the opacity profile has a rectangular shape with a specified full width at half maximum (FWHM). To be able to fit the line profiles, we assume that each line l is described by a Gaussian profile

$$\Psi_l(\nu) = \Psi_l(V) = \tau_l \exp\left(-\frac{(V - C_V)^2}{2\sigma_V^2}\right), \tag{4}$$

  where τ_l is the line opacity computed by RADEX, C_V and σ_V are the systemic velocity and the velocity dispersion of the source along the studied LoS. In this equation, the velocity and frequency are related by the Doppler formula expressed in the radio convention as $\nu_l = \nu\left(1 - \frac{V}{c}\right)$. The FWHM used by RADEX is related to σ_V through

$$\sigma_V = \frac{\mathrm{FWHM}}{\sqrt{8\ln 2}}. \tag{5}$$

The difference between the integrated opacities of the rectangular and Gaussian profiles is small, using

$$\int \Psi_l(V)\mathrm{d}V = \sqrt{\frac{2\pi}{8\ln 2}}\,\tau_l\,\mathrm{FWHM} \sim 1.06\,\tau_l\,\mathrm{FWHM}. \quad (6)$$

Furthermore, assuming that $J(T_{\rm CMB},\nu) = J(T_{\rm CMB},\nu_l)$ for ν close to $\nu_l^{\rm red}$, Eq. (1) can be simplified as

$$s(V) = \{J(T_{{\rm ex},l},\nu_l) - J(T_{\rm CMB},\nu_l)\}\,[1-\exp(-\Psi_l(V))]\,. \quad (7)$$

The physical conditions are thus characterized by five unknown parameters: the kinetic temperature, $T_{\rm kin}$, the volume density, $n_{\rm H_2}$, the column density, N, of the considered species, the velocity dispersion, σ_V, and the mean velocity, C_V, along the considered LoS. The emerging spectrum is computed through Eqs. (4) and (7), where the excitation temperature, $T_{{\rm ex},l}$, and opacity, τ_l, are computed with the RADEX code. RADEX uses $\{T_{\rm kin}, n_{\rm H_2}, N, \mathrm{FWHM}\}$ as input parameters plus two additional parameters that we fixed as follows.

– The ortho-to-para ratio (OPR) of the molecular hydrogen is defined as

$$n(\mathrm{H_2}) = n(\mathrm{p-H_2}) + n(\mathrm{o-H_2}) \quad \text{and} \quad \mathrm{OPR} = \frac{n(\mathrm{o-H_2})}{n(\mathrm{p-H_2})}. \quad (8)$$

Introducing the numerical values of the level energies of molecular hydrogen, we obtain the thermal equilibrium value as

$$\mathrm{OPR} = \frac{9e^{-\frac{170.502}{T_{\rm kin}}} + 21e^{-\frac{1015.153}{T_{\rm kin}}} + 33e^{-\frac{2503.870}{T_{\rm kin}}}}{1 + 5e^{-\frac{509.850}{T_{\rm kin}}} + 9e^{-\frac{1681.678}{T_{\rm kin}}}}. \quad (9)$$

This equation assumes that the H_2 level population is at thermal equilibrium and that $T_{\rm kin} < 500\,\mathrm{K}$.

– The electron density $n(\mathrm{e^-})$ is derived from the gas volume density with

$$x(\mathrm{e^-}) = \frac{n(\mathrm{e^-})}{n(\mathrm{H_2})} = \max\left(2.10^{-4}\sqrt{\frac{100}{n(\mathrm{H_2})}}, 10^{-8}\right). \quad (10)$$

This formula takes into account the decrease of the ionization as a function of density due to recombination. It includes a lower limit consistent with chemical model predictions in dense and dark regions (Bron et al. 2021).

Inelastic collisions with electrons contribute to the excitation of molecular rotational levels as long as the electron fraction remains higher than the ratio of the downward collision rates with H_2 and with electrons. For HCO^+, collisions with electrons will be effective for an electron fraction larger than about 3×10^{-5} at a kinetic temperature of 30 K. With the adopted scaling law (Eq. (10)) this corresponds to densities lower than $\sim$4000 $\mathrm{cm^{-3}}$. For CO, only collisions with neutral species are included, because collisions with electrons are only efficient compared to collisions with H_2 or helium for high dipole molecules like HCO^+ or HCN at the typical electron fractions found in GMCs (Liszt 2012; Goldsmith & Kauffmann 2017; Santa-Maria et al. 2023).

2.2. The notion of critical density to help with the interpretation of the results

The critical density is the volume density at which (de-) excitations by collisions equal radiative (de-)excitations. It is thus a crucial characteristic of each line, which allows us to assess the typical density above which the LTE regime is reached. As both collisional and radiative excitation processes contribute, the critical density is usually defined in the context of an optically thin line, $n_{\rm crit}^{\rm thin}$, and then modified to take into account radiative excitation (Shirley 2015). The effective critical density $n_{\rm crit}^{\rm eff}$, including line trapping effects, can be calculated for each line with

$$n_{\rm crit}^{\rm eff} = n_{\rm crit}^{\rm thin}\,\beta \quad \text{where} \quad \beta = \frac{1.5}{\tau}\left[1 - \frac{2}{\tau^2} + \left(\frac{2}{\tau} + \frac{2}{\tau^2}\right)e^{-\tau}\right]. \quad (11)$$

The value $n_{\rm crit}^{\rm thin}$ is computed as the sum of the Einstein coefficients, A, divided by the sum of collisional de-excitation rates. The parameter β is the average escape probability in the case of a uniform sphere (the default case in RADEX, see van der Tak et al. 2007).

2.3. The Cramér-Rao bound to provide a reference precision

Once the physical model and its associated assumptions are clearly stated, the Cramér-Rao bound delivers a lower bound of the variance of the estimation of each fitted parameter. This means that all the information will have been extracted from the data when we find an unbiased estimator of any given physical parameter whose variance is equal to the Cramér-Rao bound.

Based on a physical model and the associated assumptions, one can calculate the Fisher matrix (defined in Eq. (A.6)) and then numerically compute its inverse to get the Cramér-Rao bound matrix

$$\mathcal{B}(\boldsymbol{\theta}) = \boldsymbol{I}_F^{-1}(\boldsymbol{\theta}). \quad (12)$$

Noting θ_1, θ_2, ... the unknown parameters of the problem, the ith diagonal term of the CRB matrix provides a lower bound on any unbiased estimator $\widehat{\theta_i}$ of θ_i (see, e.g., Roueff et al. 2021, for an example)

$$\mathrm{var}\left(\widehat{\theta_i}\right) \geq [\mathcal{B}(\boldsymbol{\theta})]_{ii}\,. \quad (13)$$

Thus, $\mathcal{B}^{1/2}(\theta_i) = \sqrt{[\mathcal{B}(\boldsymbol{\theta})]_{ii}}$ is equivalent to a standard deviation, which can be interpreted as a precision of reference for each physical environments (e.g., diffuse gas, dense core, etc.) which takes into account the statistics of the noise that disturbs the observation. The CRB is a local property. For example, for a Gaussian noise, the CRB is the curvature of the negative log-likelihood (NLL) at its minimum. Thus, the CRB is useful as a first guess of the precision that can be expected, but it is not sensitive to the presence of other local minima, which may induce "abnormal" estimations in low signal-to-noise ratio (S/N) cases (Garthwaite et al. 1995).

Appendix A details the computation of the Fisher matrix for our use of the RADEX model, when the unknown parameters are $\boldsymbol{\theta} = [\log T_{\rm kin}, \log n_{\rm H_2}, \log N, \sigma_V, C_V]$. Section 5 analyzes the evolution of $\mathcal{B}^{1/2}(\log N)$, $\mathcal{B}^{1/2}(\log n_{\rm H_2})$ and $\mathcal{B}^{1/2}(T_{\rm kin})$ as a function of the volume density and the column density, while the other parameters remain fixed. As the thermal pressure may be better constrained than the volume density or kinetic temperature,

it is useful to also compute the Cramér-Rao bound for the thermal pressure. Formally, since $\log P_{\rm th} = \log T_{\rm kin} + \log(n_{\rm H_2})$, one simply has

$$\mathcal{B}(\log P_{\rm th}) = \mathcal{B}(\log T_{\rm kin}) + \mathcal{B}(\log n_{\rm H_2}) + 2\mathcal{B}(\log T_{\rm kin}, \log n_{\rm H_2}), \tag{14}$$

where $\mathcal{B}(\log T_{\rm kin}, \log n_{\rm H_2})$ is the off-diagonal term of the CRB matrix. This term can be interpreted as the covariance between the $\log T_{\rm kin}$ and $\log n_{\rm H_2}$ estimations.

2.4. A maximum likelihood estimator to fit the data

In addition to the RADEX and noise models, we need an algorithm to fit the data and deliver the associated estimation $\widehat{\boldsymbol{\theta}}$ of the physical parameters $\boldsymbol{\theta}$. Each pixel is fitted independently from all the other ones. We will classically fit the observed data $\boldsymbol{x}$ by looking for the estimated $\widehat{\boldsymbol{\theta}}$ that maximizes the log-likelihood function

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \left(\sum_{l=1}^{L} \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x}_l) \right), \tag{15}$$

where l is the line index of the L observed lines. For a given pixel, our algorithm can be decomposed in four steps, as described below.

Signal detection and estimation of the centroid velocity. We first compute the noise level of the spectrum in channels devoid of signal. Then, we detect the presence of a peak in the intensity for each analyzed line. The initial centroid velocity for all the lines is estimated as the peak location on the summation of ^{13}CO and C^{18}O line spectra. This is a tradeoff between fully optically thin lines, and one with high S/N. We then decide whether the pixel of the image is further fitted. Since all species are assumed to share the same centroid velocity, we only require that more than half of the lines are detected. For example, for the set $\mathcal{S}\left\{\mathrm{C^{18}O}, \mathrm{H^{13}CO^{+}}\right\}$, we only require that both C^{18}O lines are detected.

Initialization of the fit estimation. We initialize the estimation of the other physical parameters on the point that minimizes the negative log-likelihood (NLL) on a grid that regularly samples the fitted parameters. Here it is crucial to take into account the a priori constraints that are assumed, for instance, the same kinetic temperature and volume density (see Appendix B.2 for details). This step yields an initial estimation of all parameters $\boldsymbol{\theta}$. It may also induce some implicit a priori information depending on the chosen grid boundaries.

Maximization of the likelihood. We then refine our initial estimation with a standard iterative gradient descent. The iterator is stopped when one of the following conditions is achieved: 1) The difference of NLL between two consecutive iterations is smaller than 10^{-6}; or 2) the number of iterations reaches 30; or 3) the NLL minimum does not decrease for five consecutive iterations. This may happen because of numerical instability.

Estimation of the confidence interval. Once the physical parameters have been obtained, we compute the Cramér-Rao bound for each parameter $\mathcal{B}(\widehat{\theta_i})$ to provide a reference interval of confidence. Although we replaced θ_i by $\widehat{\theta_i}$, this interval of confidence remains mostly valid as long as the model is valid, and as soon as the estimator is efficient (i.e., it is unbiased and the variance is equal to the Cramér-Rao bound).

Table 2. Properties of the observed lines.

Line [(1)]	ν GHz	$E_{\rm up}$ K	A s^{-1}	dV [(2)] km s^{-1}	Noise [(3)] mK
^{12}CO (1 – 0)	115.3	5.5	7.203e-08	0.5	91
^{13}CO (1 – 0)	110.2	5.3	6.294e-08	0.1	74
^{13}CO (2 – 1)	220.4	15.9	6.038e-07	0.1	155
C^{18}O (1 – 0)	109.8	5.3	6.266e-08	0.1	72
C^{18}O (2 – 1)	219.6	15.8	6.011e-07	0.1	121
HCO$^+$ (1 – 0)	89.2	4.3	4.187e-05	0.1	318
H^{13}CO$^+$ (1 – 0)	86.8	4.2	3.853e-05	0.5	45

Notes. [(1)]Line frequencies, upper energy levels and Einstein A coefficients are taken from the Cologne Data Base for Molecular Spectroscopy (CDMS) (Müller et al. 2001). [(2)]Channel spacing after resampling. [(3)]Median noise σ_b in the Horsehead after resampling and smoothing. Angular resolution after smoothing is 31″.

Additional implementation details about these implementations are provided in Appendix B.

3. Quantifying the variability of the estimation of physical parameters on actual data

In this prototype study, we used the RADEX model to fit the spectra of a combination of the (1 – 0) and (2 – 1) transitions for some of the following species: ^{12}CO, ^{13}CO, C^{18}O, HCO$^+$, and H^{13}CO$^+$. Table 2 lists the properties of each line (frequency, upper energy level and Einstein coefficient) as well as the properties of the observed spectra (spectral spacing and thermal noise). The studied field of view covers about 50 square arcminutes towards the Horsehead Pillar at the South West of Orion B, which is illuminated by the σ-Ori star.

3.1. Data

3.1.1. Acquisition and data reduction

The data were obtained at the IRAM 30m telescope. The (1 – 0) lines were acquired in 2013-2020 in the context of the ORION-B large program (Pety et al. 2022; Gerin et al. 2023), using a combination of the EMIR receivers and FTS spectrometers. The (2 – 1) lines towards the Horsehead pillar were acquired in 2006 (PI: N. Peretto) using the ABCD generation of receivers and the VESPA auto-correlator. They have been first presented and used in Roueff et al. (2021).

The data reduction is described in Pety et al. (2017) and Roueff et al. (2021). It uses the standard methods provided by the `GILDAS`[1]`/CLASS` software. The data cubes were gridded on the same astrometric grid. They were smoothed to the lowest achieved angular resolution (i.e., 31″) in order to ease the multi-line fit towards the same LoS without having to take care of different angular resolutions. The velocity spacings are listed in Table 2. We used two velocity spacings (0.1 and 0.5 km s^{-1}) depending on the native spectral resolution of the spectrometer used for each line.

3.1.2. Presentation of the data

Figure 1 summarizes the data with a combination of images of the peak and integrated intensities as well as percentiles of the

[1] See `http://www.iram.fr/IRAMFR/GILDAS` for more information about the GILDAS software (Pety 2005).

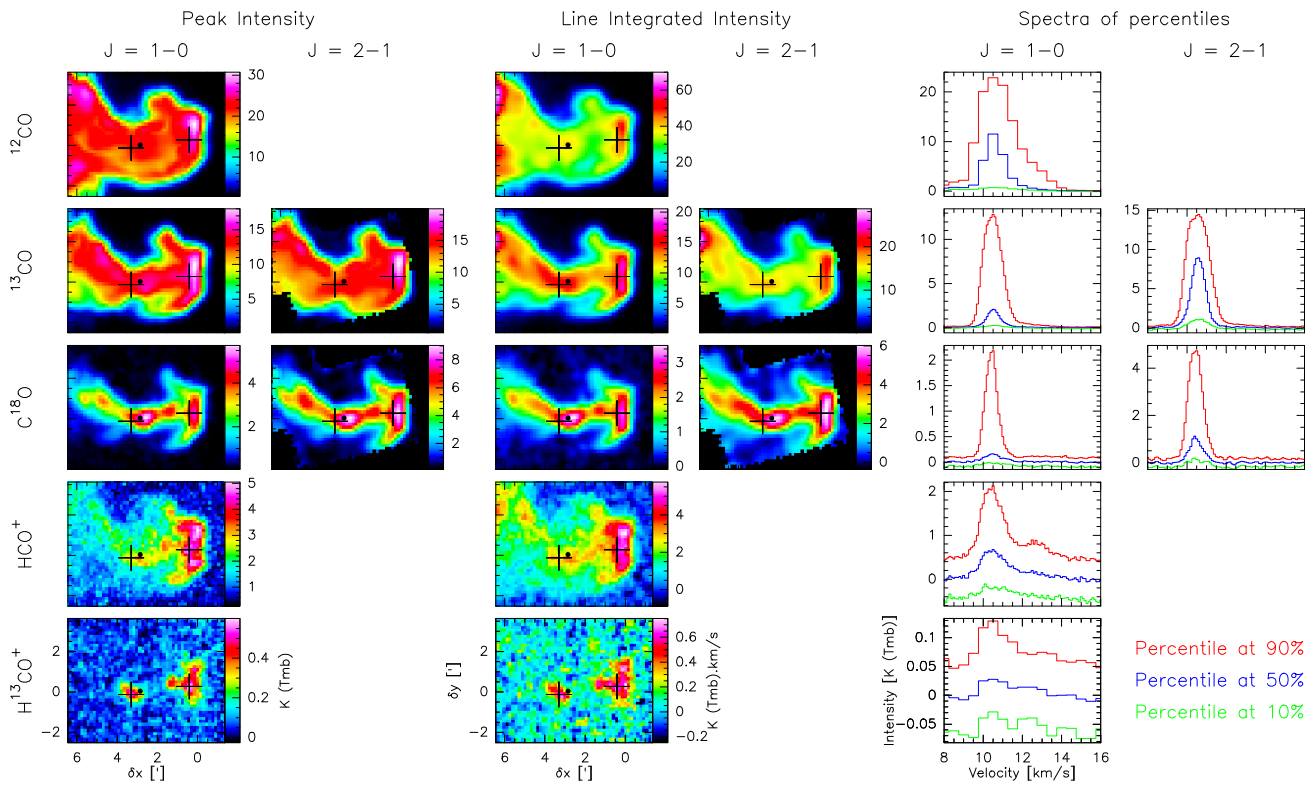


Fig. 1. Presentation of the data towards the Horsehead pillar. Left and middle panels: spatial distribution of the peak intensity and of the integrated intensity. The used projection is the Azimuthal one, centered on RA= $05^h40^m54.27^s$, Dec= $-02°28'00.00''$ in the ICRS frame, and with a position angle of 14°. The two crosses show the positions of the two dense cores whose coordinates are in Table 3. The black dot shows the position of the line of sight studied in more detailed in Sect. 4. Right panel: the green, blue, and red spectra show the spectra of percentiles at 10, 50, and 90%, respectively, computed over the associated field of view.

Table 3. Coordinates of lines of sight marked on Fig. 1.

Line of sight	RA	Dec
Dense core #1	$05^h40^m55.60^s$	$-02°27'38.01''$
Dense core #2	$05^h41^m07.21^s$	$-02°27'19.37''$
Studied position	$05^h41^m05.31^s$	$-02°27'16.88''$

line intensities as a function of the velocity among the considered pixels. The spatial shape of the Horsehead pillar varies with the species. The ^{12}CO (1 – 0) line is detected towards most of the field of view, while the rarer CO isotopologues better show the spine of the Horsehead. The H^{13}CO$^+$ (1 – 0) line is mainly detected towards the two known dense cores of the Horsehead (e.g. Ward-Thompson et al. 2006). Table 3 lists the specified locations on the image. The two crosses show the positions of these two dense cores. The position of the western dense core is the Horsehead dense core position defined in Pety et al. (2007) and extensively studied in the framework of the Horsehead WHISPER project (see, e.g., Guzmán et al. 2014). The cross for the eastern dense core points towards the local maximum of the H^{13}CO$^+$ (1 – 0) line emission. The HCO$^+$ (1 – 0) line exhibits a mixed behavior with some filamentary structure embedded in more diffuse emission that has a similar shape as the ^{12}CO (1 – 0) line.

The line profiles mostly show a main velocity component centered near $V_{LSR} \sim 10.5\,km\,s^{-1}$. The ^{12}CO, HCO$^+$, and H^{13}CO$^+$ (1 – 0) lines exhibit a second velocity component or wing around $V_{LSR} \sim 13.0\,km\,s^{-1}$. As the presence of a single velocity component is required to be consistent with the hypothesis that the lines are emitted by the same layer of molecular gas along the LoS, we only concentrate on the $10.5\,km\,s^{-1}$ velocity component. Moreover this $10.5\,km\,s^{-1}$-velocity component is clearly detected for all the lines.

The ^{13}CO (2 – 1) and C^{18}O (2 – 1) lines have similar properties as their (1 – 0) transitions regarding their spatial and spectral shapes, but with minor differences due to excitation effects. Finally, most of the observed lines have high S/N with the noteworthy exception of H^{13}CO$^+$ (1 – 0).

3.2. Fitting assumptions

3.2.1. Simplifying physical and chemical assumptions

As detailed in Sect. 2.1, the considered non-LTE model is characterized by five unknown parameters (n_{H_2}, T_{kin}, N, σ_V and C_V) per modeled species. The purpose of this article is to analyze the possibility of estimating those parameters as a function of the LoS from observations of few lines (1 or 2), from several species. We consider here different sets of species containing from one to three species among: ^{12}CO, ^{13}CO, C^{18}O, HCO$^+$, and H^{13}CO$^+$. Since the number of unknown parameters is too

Table 4. Sets of species, the assumed chemistry, the number of available lines, and the number of unknown parameters.

Set of species	Relative abundance	# lines	# unknowns
$\mathcal{S}\left\{^{13}\mathrm{CO}\right\}$ or $\mathcal{S}\left\{\mathrm{C}^{18}\mathrm{O}\right\}$	–	2	5
$\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$	free	3	6
$\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+; \mathcal{A}\right\}$	$\frac{N(^{13}\mathrm{CO})}{N(\mathrm{HCO}^+)} = 10^{2.88}$	3	5
$\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+; \mathcal{A}^\star\right\}$	$\frac{N(^{13}\mathrm{CO})}{N(\mathrm{HCO}^+)} = 10^{3.18}$	3	5
$\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C}^{18}\mathrm{O}\right\}$	free	4	6
$\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C}^{18}\mathrm{O}; \mathcal{A}\right\}$	$\frac{N(^{13}\mathrm{CO})}{N(\mathrm{C}^{18}\mathrm{O})} = 10^{0.90}$	4	5
$\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C}^{18}\mathrm{O}; \mathcal{A}^\star\right\}$	$\frac{N(^{13}\mathrm{CO})}{N(\mathrm{C}^{18}\mathrm{O})} = 10^{1.20}$	4	5
$\mathcal{S}\left\{\mathrm{C}^{18}\mathrm{O}, \mathrm{H}^{13}\mathrm{CO}^+\right\}$	free	3	6
$\mathcal{S}\left\{\mathrm{C}^{18}\mathrm{O}, \mathrm{H}^{13}\mathrm{CO}^+; \mathcal{A}\right\}$	$\frac{N(\mathrm{C}^{18}\mathrm{O})}{N(\mathrm{H}^{13}\mathrm{CO}^+)} = 10^{3.78}$	3	5
$\mathcal{S}\left\{\mathrm{C}^{18}\mathrm{O}, \mathrm{H}^{13}\mathrm{CO}^+; \mathcal{A}^\star\right\}$	$\frac{N(\mathrm{C}^{18}\mathrm{O})}{N(\mathrm{H}^{13}\mathrm{CO}^+)} = 10^{3.48}$	3	5
$\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$	$\frac{N(^{12}\mathrm{CO})}{N(^{13}\mathrm{CO})} = 10^{1.80}$	4	6
$\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{C}^{18}\mathrm{O}\right\}$	$\frac{N(^{12}\mathrm{CO})}{N(^{13}\mathrm{CO})} = 10^{1.80}$	5	6

Notes. When more than one species is fitted, we assume that they share the values of $(n_{\mathrm{H}_2}, T_{\mathrm{kin}}, C_V$, and $\sigma_V)$.

large compared to the available data, assumptions are required in order to reduce the number of unknowns with respect to the number of fitted lines. Table 4 lists the cases considered in this section. For all species sets, we assume that the gas emitting the different lines share the same kinetic temperature, volume density, centroid velocity and velocity dispersion. In other words, we assume that their emissions come from the same gas cell in the cloud. We also often assume that at least one of the column density ratios is known and thus fixed. We thus have between 5 and 6 unknowns. The first five ones are the kinetic temperature, volume density, centroid velocity and velocity dispersion, and the column density of either ^{13}CO or C^{18}O. For some sets, the sixth unknown is either the $\frac{N(\mathrm{HCO}^+)}{N(^{13}\mathrm{CO})}$ or the $\frac{N(\mathrm{H}^{13}\mathrm{CO}^+)}{N(\mathrm{C}^{18}\mathrm{O})}$ relative abundance.

Considering a single species with two lines, for instance, the ^{13}CO or C^{18}O isotopologues, allows us to get an idea of their column densities without the assumption of co-location. For sets containing two species, we associate species that share a good probability of being emitted from similar regions: $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$ for diffuse regions, $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C}^{18}\mathrm{O}\right\}$ for denser filaments, and $\mathcal{S}\left\{\mathrm{C}^{18}\mathrm{O}, \mathrm{H}^{13}\mathrm{CO}^+\right\}$ for dense cores. We add the ^{12}CO peak value to check its influence.

The $(1-0)$ lines of the HCO$^+$ isotopologues have higher critical densities than the CO isotopologues' ground-state transitions because the dipole moments of these molecular ions is much larger than the ones of neutral species (Shirley 2015). These lines may thus bring additional information towards the two dense cores of the Horsehead pillar. We use the two main isotopologues to assess the impact of their different optical depths and S/Ns.

Since this paper focuses on a region with $n_{\mathrm{H}_2} \gtrsim 10^3$ cm^{-3}, we only use the spectrum of ^{12}CO at its peak value. Indeed, the line profile delivered by the non-LTE model with a simple Gaussian opacity profile does not suitably describe the profile of extremely optically thick lines such as ^{12}CO $(1-0)$ (Leung & Liszt 1976).

We add a $\mathcal{A}$ symbol in the set name when we fix the relative abundance. In this case, the relative abundance is defined by using the typical isotopic ratios for Orion ^{12}C/^{13}C = 63, ^{16}O/^{18}O = 510, N(^{13}CO)/N(H_2)=2.3×10^{-6} and N(HCO$^+$)/N(H_2)=3×10^{-9} (Gerin et al. 2019; Roueff et al. 2021, and references therein).

$$\begin{aligned} N(^{12}\mathrm{CO})/N(^{13}\mathrm{CO}) &= 10^{1.8} \sim 60, \\ N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O}) &= 10^{0.9} \sim 8, \\ N(^{13}\mathrm{CO})/N(\mathrm{HCO}^+) &= 10^{2.88} \sim 800, \\ N(\mathrm{C}^{18}\mathrm{O})/N(\mathrm{H}^{13}\mathrm{CO}^+) &= 10^{3.78} \sim 6000. \end{aligned} \tag{16}$$

We also varied the last three abundance ratios by ±0.3 dex in order to assess the impact of variations of the assumed chemistry. These cases are marked with the $\mathcal{A}^\star$ symbol.

3.2.2. Noise model

The data are affected by thermal noise and calibration uncertainty. We thus use the following model of data

$$x(v) = c\, s(v) + b(v), \tag{17}$$

where x is the observed data, s the source intensity, b is the thermal noise, and c represents the calibration uncertainty. We assume that b and c are independent random variables with normal distribution: $b \sim \mathcal{N}(0, \sigma_b^2)$ and $c \sim \mathcal{N}(1, \sigma_c^2)$ (see Einig et al. 2023, for more details).

The standard deviations associated to the calibration uncertainty are fixed as

$$\sigma_c = \begin{cases} 5\% & \text{for 3 mm lines,} \\ 10\% & \text{for 1 mm lines.} \end{cases} \tag{18}$$

The median values of the thermal standard deviation measured on the data are listed in Table 2.

3.2.3. Choice of likelihood and S/N saturation

When quantifying the theoretical precision of estimations (see Sect. 5), we take into account both the thermal noise and calibration uncertainty. However, we neglect the calibration uncertainty and we instead set the S/N to a maximum value of 10, when fitting the data with the maximum likelihood estimator.

On one hand, including the multiplicative factor in the estimation leads to fits where the potential shortcomings of the physical assumptions are incorrectly attributed to calibration errors, because the fit then cares too much about the shape of the spectra, and too little about its amplitude. This is counterproductive for real data analysis. We thus choose the standard negative likelihood when fitting data $(\boldsymbol{x})$, i.e.,

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x}) = -\sum_n \frac{(x_n - s_n)^2}{2\sigma_b^2}, \tag{19}$$

where s_n is a sampled version of Eq. (7) and thus is a function of the unknown parameters. On the other hand, we artificially set σ_b such that the S/N, defined as

$$S/N = \frac{\max_v s(v)}{\sigma_b}, \tag{20}$$

has a maximum value of 10. This artificial saturation is required to ensure that lines with a S/N between 3 and 10 (e.g., H^{13}CO$^+$ $(1-0)$) have a non negligible weight compared to lines with S/N of 100 (e.g., the ^{13}CO lines). The value of 10 also ensures that calibration errors that are lower than 10% are considered as noise, and contribute to the error budget together with the additive noise.

3.3. Results

We systematically check here the stability of the fitted parameters as a function of the set of species fitted and the associated a priori hypotheses. In the following figures, we select one parameter at a time. We look at chemical (column densities and relative abundances, i.e., ratios of column densities) and physical (kinetic temperature, volume density, and thermal pressure) parameters. We focus on reliable estimations defined as

$$\mathcal{B}^{1/2}(\widehat{\theta}) \leq 0.3 \quad \text{for} \quad \theta = \log N_i,\ \log T_{\rm kin},\ \text{and}\ \log n_{\rm H_2}. \tag{21}$$

For each parameter, we present two figures: 1) the maps of the fitted parameter on the Horsehead pillar data for different sets of species and/or a priori (e.g., the top of Fig. 2); 2) joint histograms of this parameter estimations. The bottom of Fig. 2 compares the estimations of $N(^{13}\mathrm{CO})$ obtained for 7 different sets of species and/or a priori defined in Table 4: $\mathcal{S}\left\{^{13}\mathrm{CO}\right\}$, $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO^+}; \mathcal{A}\right\}$, $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO^+}; \mathcal{A}^*\right\}$, $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}; \mathcal{A}\right\}$, $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}; \mathcal{A}^*\right\}$, $\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{HCO^+}\right\}$ and $\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$. There are 21 pairs for 7 different fits. The figure thus appears as a lower triangle of size 6×6. For conciseness sake, we only show in the main paper these two kinds of figures for the case of the column density of ^{13}CO. For the other parameters, we put these figures in Appendix C, and we synthesize the joint histograms by their distance in log-space to the line of slope 1. More precisely, we compute the mean and the standard deviation of $\left|\widehat{\theta_x} - \widehat{\theta_y}\right| / \sqrt{2}$, where θ_x and θ_y are the estimated parameters plotted along the x and y axes of the histograms. We present these values in a tabular way where we put in the same lower triangle the values of the mean and standard deviation, color-coded by the root mean square error, RMSE $= \sqrt{\mathrm{mean}^2 + \mathrm{sdev}^2}$. Figure 3 shows the associated "tables" for $N(^{13}\mathrm{CO})$, $N(\mathrm{C^{18}O})$, $N(\mathrm{HCO^+})$, and $N(\mathrm{H^{13}CO^+})$, $T_{\rm kin}$, and $n_{\rm H_2}$. A non zero mean indicates a systematic shift compared to the line of slope 1, while a large standard deviation indicates a large dispersion of the distances to the line whose slope is the mean, and a large RMSE a large dispersion of the distances to the line of slope 1. Quantitatively the values can also be interpreted with the help of Table 8. Values of ±0.1, ±0.3, and ±0.6 correspond to multiplicative uncertainty of the order of 20%, a factor 2, and 4 respectively.

3.3.1. One detailed case: The column density of ^{13}CO

The top of Fig. 2 shows the maps of the column densities of ^{13}CO fitted on the Horsehead pillar under different hypotheses. In this case, almost all pixels deliver a column density with a good CRB reference precision whatever the set of species and of a priori used during the fit. However, the shape of the maps of the column density varies depending on the chosen set of lines. If we take the case where we only fit the ^{13}CO lines as reference (panel a), we see different behaviors. First, adding the $\mathrm{C^{18}O}$ lines increases $N(^{13}\mathrm{CO})$ towards the two dense cores or the surrounding filaments depending on the assumed chemistry. Second, adding the $\mathrm{HCO^+}$ lines decreases $N(^{13}\mathrm{CO})$ towards the dense core in the Horsehead throat but increases it towards the dense core on the top of the head. The effect is more or less pronounced depending on the assumed chemistry. Third, adding the peak temperature of the ^{12}CO $(1-0)$ line regularizes the overall shape of the column density map.

These trends are quantitatively confirmed by the joint histograms shown in the bottom of Fig. 2. The most consistent values of the column densities are obtained in the cases $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO^+}; \mathcal{A}\right\}$ and $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO^+}; \mathcal{A}^\star\right\}$ where only the assumed chemistry is changed. This is followed by the pairs ($\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$ and $\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{HCO^+}\right\}$), and ($\mathcal{S}\left\{^{13}\mathrm{CO}\right\}$ and $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO^+}, \mathcal{A}\right\}$). The less consistent values of the column densities are obtained for the pairs ($\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}; \mathcal{A}\right\}$ and $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}; \mathcal{A}^\star\right\}$).

Overall, the column densities of ^{13}CO are defined within a typical multiplicative accuracy of 30%. Adding $\mathrm{C^{18}O}$ (resp. $\mathrm{HCO^+}$) gives more weight to the dense cores (resp. more diffuse envelope) when fitting the column density of ^{13}CO.

3.3.2. Column densities and relative abundances

The top row of Fig. 3 and the associated figures in Appendix C synthesize the results on column densities. The column densities of $\mathrm{C^{18}O}$ are mostly consistent within 15%. That is a factor ~ 2 better than ^{13}CO. Accordingly, the maps of $N(\mathrm{C^{18}O})$ show spatial variations that are similar in all tested cases (see Fig. C.1). The noteworthy exception are the couples ($\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}; \mathcal{A}\right\}$ and $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}; \mathcal{A}^\star\right\}$), probably because the column density of $\mathrm{C^{18}O}$ is sensitive to the choice of the $N(^{13}\mathrm{CO})/N(\mathrm{C^{18}O})$ abundance ratio.

The column densities of $\mathrm{HCO^+}$ and $\mathrm{H^{13}CO^+}$ are more scattered and biased depending on the fitted case. The worst case is $N(\mathrm{H^{13}CO^+})$ because the associated line has the lowest S/N of all. But $N(\mathrm{HCO^+})$ also shows a large scatter. We relate this to a possible evolution of the $\mathrm{HCO^+}$ chemistry between dense cores and more diffuse gas, implying an evolution of its relative abundance.

Figure 4 illustrates the stability of the relative abundances depending on the case. The relative abundance of $\mathrm{C^{18}O}$ and ^{13}CO is pretty insensitive to the presence or absence of the peak intensity of the ^{12}CO $(1-0)$ line during the fit. In contrast, the relative abundances of (^{13}CO and $\mathrm{HCO^+}$) or ($\mathrm{C^{18}O}$ and $\mathrm{H^{13}CO^+}$) are sensitive (both in bias and scatter) to the presence or absence of the peak intensity of the ^{12}CO $(1-0)$ line. This may be due to the fact that the peak intensity mostly constrains the kinetic temperature. As the CO isotopologue line intensities are mostly sensitive to the thermal pressure in the considered parameter space, the constraint on the temperature has an effect on the volume density determination. With their high dipole moment (thus higher critical densities) the $\mathrm{HCO^+}$isotopologue excitation is sensitive to the volume density.

3.3.3. Kinetic temperature

The determination of the kinetic temperature presents large variations depending on the chosen set of lines. As shown in Fig. C.3, using both the $(1-0)$ and $(2-1)$ lines of ^{13}CO and $\mathrm{C^{18}O}$ enables a correct determination of the kinetic temperature in the high column density filament. In addition, the choice of the fixed relative abundance between these species has little impact on the $T_{\rm kin}$ estimation. In contrast, the same set of lines leads to unrealistic high values of $T_{\rm kin}$ outside the filament. The combination of either ^{13}CO and $\mathrm{HCO^+}$ or $\mathrm{C^{18}O}$ and $\mathrm{H^{13}CO^+}$ performs poorly. The determination of the kinetic temperature is sensitive to the assumption made on the relative abundances. The $T_{\rm kin}$ estimations seem to be biased as the correlation with the determinations from ^{13}CO and $\mathrm{C^{18}O}$ shows some scatter and presents a significant offset.

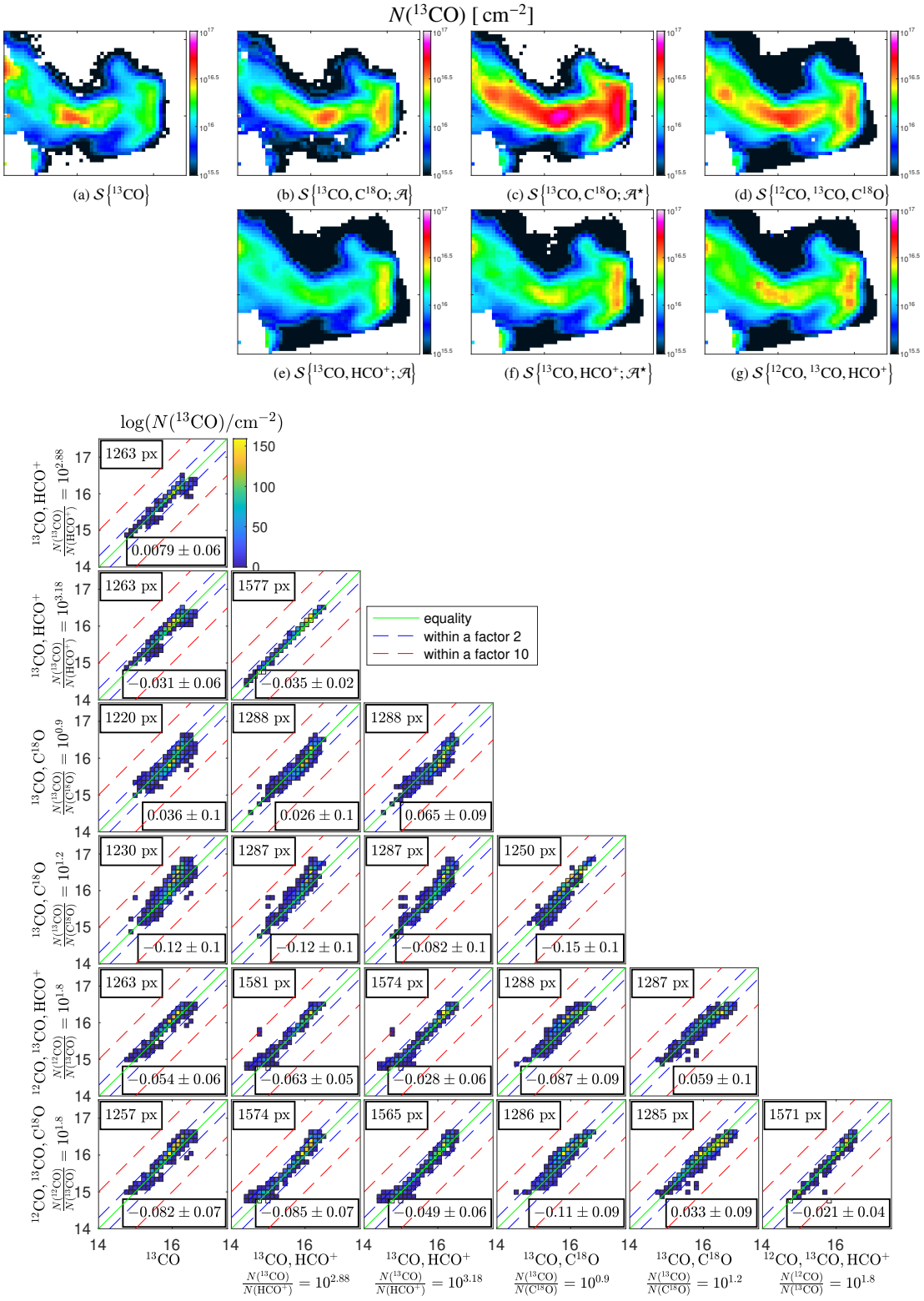


Fig. 2. Comparison of the estimated column density of ^{13}CO on the Horsehead pillar data for different sets of species and/or a priori. In all cases $T_{\rm kin}$, $n_{\rm H_2}$, C_V and σ_V are assumed to be the same for all lines. Top: maps of estimated $N(^{13}\mathrm{CO})$. Table 4 lists the details of each studied case. Bottom: joint histograms of $\log N(^{13}\mathrm{CO})$ estimations as a function of different sets of species and there associated a priori hypotheses, shown as the column and row labels. In each panel, the green, blue, and red lines show the identity function, and ratios of values within a factor 2 or 10, respectively. The top left and bottom right legends give the number of pixels used to compute the histogram, and the mean and standard deviation of the distance to the green line.

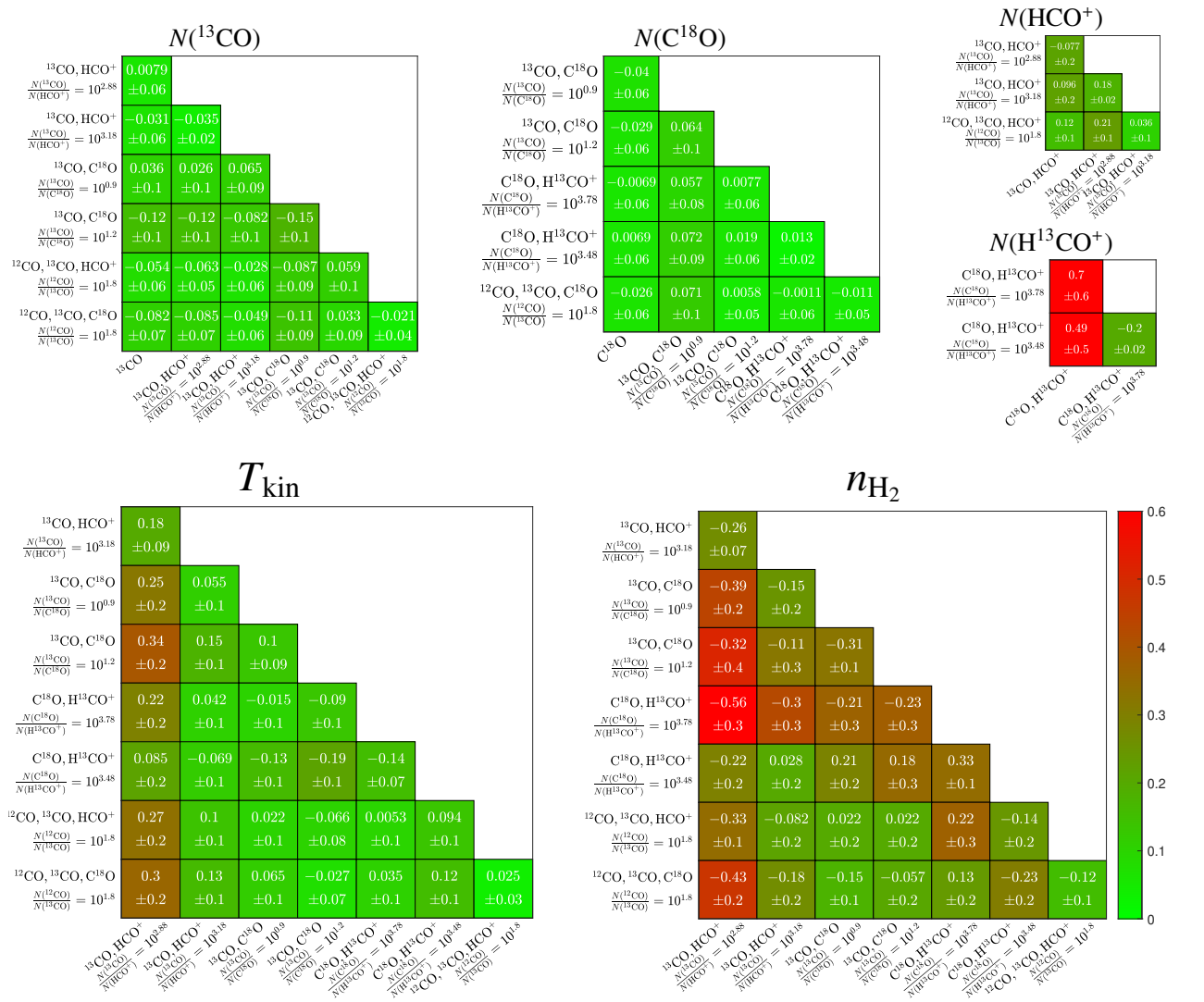


Fig. 3. Mean and standard deviation of the distance in log-space between the estimations (i.e., $\left|\widehat{\theta_x} - \widehat{\theta_y}\right| / \sqrt{2}$ where $\theta = \log N$, or $\log T_{kin}$ or $\log n_{H_2}$) as a function of different sets of species and their associated a priori hypotheses, shown as the column and row labels. In all cases T_{kin}, n_{H_2}, C_V and σ_V are assumed to be the same for all lines. Table 4 lists the details. The cells are color-coded with the values of the root mean square error = $\sqrt{\text{mean}^2 + \text{sdev}^2}$. From left to right, the top graphs show the results for $N(^{13}CO)$, $N(C^{18}O)$, $N(HCO^+)$, and $N(H^{13}CO^+)$, and the bottom graphs the results for the kinetic temperature and volume density. This figure summarizes the joint histograms shown in Figs. 2, C.1–C.4.

Adding ^{12}CO to the set of lines improves the T_{kin} estimation because this sets an upper limit to the value of T_{kin} of about 30 K. The resulting temperature maps are more extended and present smooth variations. The two sets perform about equally well and lead to similar estimations across the map with no systematic difference. However, using the peak of the ^{12}CO (1 − 0) emission introduces a bias in the estimation as this line is produced in the external regions of the considered sight-line, where the kinetic temperature may not be representative of the conditions along the full LoS.

3.3.4. Volume density

The volume density is the most difficult parameter to estimate. The scatter can reach a factor of ten and different sets of lines can lead to determinations with a systematic offset of up to 0.5 dex, a factor of three. In the high column density regions, the sets combining $C^{18}O$ and $H^{13}CO^+$ perform well, but the absolute value of the volume density scales with the fixed relatives abundance. Indeed, the determination are well correlated but offset from each other by 0.34 dex, which corresponds to the change in the fixed relative abundances. A similar effect is seen for the sets combining ^{13}CO and HCO^+, which leads to well correlated but offsets density estimations. For pixels where both sets can be used and the hypotheses on abundances are consistent, it is encouraging to see that ^{13}CO and HCO^+, and $C^{18}O$ and $H^{13}CO^+$ lead to similar estimations of the volume density with no significant offset, even if a significant scatter of about 0.2 dex remains.

The estimations of the volume density made from the sets combining CO isotopologues only, seem to reach a maximum value near $10^{4.3}$ cm^{-3} (see Fig. C.4). This may be related to the limited range of kinetic temperature and the lower critical density of the considered CO isotopologue transitions as

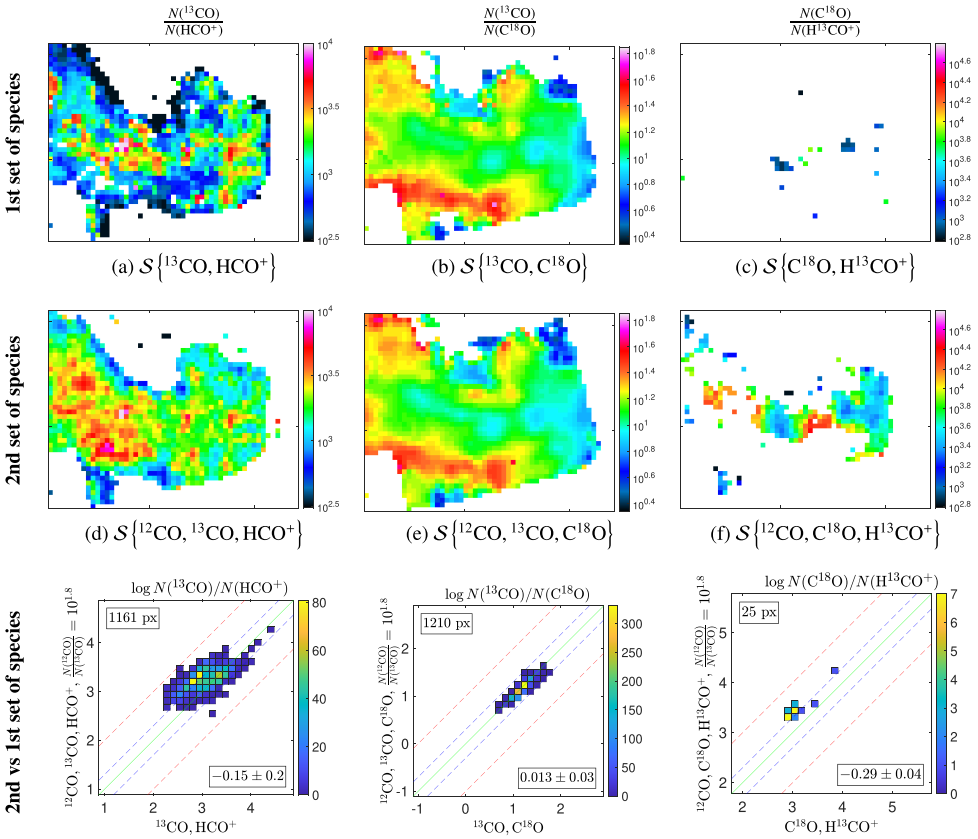


Fig. 4. Maps and joint histograms of relative abundances for $\log\{N(^{13}CO)/N(HCO^+)\}$ (left), $\log\{N(^{13}CO)/N(C^{18}O)\}$ (middle), and $\log\{N(C^{18}O)/N(H^{13}CO^+)\}$ (right). The layout of the figure is similar to Fig. 2.

compared with those of the high dipole moment species HCO^+ and $H^{13}CO^+$.

3.3.5. Thermal pressure

The estimation of the thermal pressure is more accurate and less biased than that of T_{kin} and n_{H_2} separately (see Fig. C.5). In particular the hypothesis on relative abundance has little impact on the estimation of the thermal pressure, while we have seen that it has a strong influence on the volume density. Estimations made with or without ^{12}CO are consistent in most cases. The thermal pressure for the pixels associated with dense cores where $C^{18}O$ and $H^{13}CO^+$ are detected with high S/N is somewhat higher when using $C^{18}O$ and $H^{13}CO^+$ only as compared with sets including ^{12}CO but the effect is small. The better accuracy of the thermal pressure estimations can be understood by looking at the shape of the NLL function in the (n_{H_2}, T_{kin}) parameter space. This is discussed in Sect. 4.

3.4. Intermediate summary

In this section, we have systematically compared the results of line fits with the RADEX non-LTE radiative transfer model, on the same dataset with the same fitting algorithm. The main difference between the different experiments has been the number and the set of fitted lines. The fitted parameters have been the column densities of the ^{13}CO and HCO^+ isotopologues, and the associated kinetic temperature and volume density. We have also studied the derived relative abundances and thermal pressure.

This systematic comparison emphasizes the potential variability of the estimated parameters depending on the chosen set of fitted species and lines. The joint histograms of the same fitted parameters as a function of the fitting assumptions show not only important scatter around the line of slope 1 but also biases with respect to this line. The maps of fitted parameters show regular patterns, suggesting that the noise level is not the cause of the found variations. The volume density is the least accurate estimated parameter, followed by the kinetic temperature. The thermal pressure estimates are more accurate than the volume density and kinetic temperature. The column densities and, above all, the ratio of column densities are the most accurate fitted parameters.

We do not address here whether the quality of the fit allows us to define which set of species and associated assumptions best fits the data for two reasons. First, we here explore the impact of the typical analysis of line data where the best solution is selected by minimizing the negative log-likelihood of the fit for a fixed set of species. Second, fitting a subset of lines necessarily leads to smaller χ^2, but this does not imply that estimations are better. Third, comparing the quality of fit performed with different numbers of unknown parameters requires more advanced statistical tools than just the χ^2 value to select the "best" estimations. For instance, when the abundance are estimated instead of being fixed, the χ^2 is necessarily smaller, but this does not imply that the estimation is better, notably because of the induced estimation variance. This topic will be studied in a future paper.

Table 5. Input physical parameters for the two experiments

Exp.	$N(^{12}CO)$ cm^{-2}	$N(^{13}CO)$ cm^{-2}	$N(C^{18}O)$ cm^{-2}	$N(HCO^+)$ cm^{-2}	$N(H^{13}CO^+)$ cm^{-2}	T_{kin} K	$n(H_2)$ cm^{-3}	σ_V $km\,s^{-1}$	C_V $km\,s^{-1}$
# 1	$10^{18.15}$	$10^{16.35}$	$10^{15.45}$	$10^{13.47}$	$10^{11.67}$	27.1	$10^{3.65}$	0.32	10.5
# 2	idem	idem	idem	idem	idem	22.1	$10^{4.65}$	idem	idem

Table 6. Line characteristics for experiment #1.

Species	Line	σ_b K	T_{peak} K	Peak S/N –	τ –	T_{ex} K	n^{thin}_{crit} cm^{-3}		n^{eff}_{crit} cm^{-3}	
							o – H_2	p – H_2	o – H_2	p – H_2
^{12}CO	(1 – 0)	2.3	23	10	50	27	$10^{3.34}$	$10^{3.35}$	$10^{1.81}$	$10^{1.82}$
^{13}CO	(1 – 0)	1.5	15	10	1.0	26	$10^{3.28}$	$10^{3.29}$	$10^{3.12}$	$10^{3.13}$
^{13}CO	(2 – 1)	1.4	14	10	3.9	19	$10^{3.77}$	$10^{3.81}$	$10^{3.3}$	$10^{3.35}$
$C^{18}O$	(1 – 0)	0.36	3.6	10	0.11	38	$10^{3.28}$	$10^{3.29}$	$10^{3.26}$	$10^{3.27}$
$C^{18}O$	(2 – 1)	0.48	4.8	10	0.78	14	$10^{3.77}$	$10^{3.81}$	$10^{3.65}$	$10^{3.69}$
HCO^+	(1 – 0)	0.31	3.1	10	12	6.1	$10^{5.26}$	$10^{5.3}$	$10^{4.36}$	$10^{4.39}$
$H^{13}CO^+$	(1 – 0)	0.10	0.15	1.5	0.41	3.2	$10^{5.23}$	$10^{5.26}$	$10^{5.16}$	$10^{5.19}$

Notes. The peak S/N is defined in Eq. (20).

Table 7. Line characteristics for experiment #2.

Species	Line	σ_b K	T_{peak} K	Peak S/N –	τ –	T_{ex} K	n^{thin}_{crit} cm^{-3}		n^{eff}_{crit} cm^{-3}	
							o – H_2	p – H_2	o – H_2	p – H_2
^{12}CO	(1 – 0)	1.9	19	10	71	22	$10^{3.33}$	$10^{3.35}$	$10^{1.65}$	$10^{1.67}$
^{13}CO	(1 – 0)	1.2	12	10	1.1	22	$10^{3.27}$	$10^{3.29}$	$10^{3.11}$	$10^{3.13}$
^{13}CO	(2 – 1)	1.6	16	10	3.2	21	$10^{3.77}$	$10^{3.81}$	$10^{3.36}$	$10^{3.4}$
$C^{18}O$	(1 – 0)	0.26	2.5	10	0.14	23	$10^{3.27}$	$10^{3.29}$	$10^{3.25}$	$10^{3.26}$
$C^{18}O$	(2 – 1)	0.55	5.5	10	0.43	21	$10^{3.76}$	$10^{3.8}$	$10^{3.7}$	$10^{3.74}$
HCO^+	(1 – 0)	0.88	8.8	10	4.2	12	$10^{5.25}$	$10^{5.28}$	$10^{4.75}$	$10^{4.69}$
$H^{13}CO^+$	(1 – 0)	0.10	0.54	5.4	0.18	6.3	$10^{5.21}$	$10^{5.25}$	$10^{5.18}$	$10^{5.22}$

Notes. The peak S/N is defined in Eq. (20).

4. On the importance of prior assumptions on the bias-variance tradeoff

In the previous section, we tried to estimate the gas volume density and kinetic temperature in addition to the molecular column densities, when fitting different combinations of the CO and HCO^+ main isotopologue lines together. This requires some additional a priori knowledge linking the parameters so that the number of unknowns becomes smaller than the number of constraints, and the associated confidence interval become "reasonably" small. In particular, we make the hypothesis that the same region along the LoS dominates the emission of all lines, with a unique volume density and kinetic temperature. We also make chemical assumptions in the form of fixed relative abundances. However, a priori knowledge may turn into estimation biases if the assumptions happen to be incorrect. This problem is the well-known variance vs. bias tradeoff. In this section, we will study this tradeoff on simulated data in order to identify potential sources of bias.

4.1. Description of the experiments and associated line characteristics

To do this, we choose input physical conditions based on the estimations computed in Sect. 3 on actual data towards the Horsehead nebula. We choose a pixel at the edge of the central dense core (black point in Fig. 1), for which the fit on actual data shows different values of kinetic temperature and volume density, depending on the set of species studied. In particular, analyzing just the ^{13}CO and $C^{18}O$ lines provides nonphysical solutions (large kinetic temperature and small volume density compared to standard dense core values). Meaningful values for kinetic temperature and volume density are only obtained when adding different species.

Table 5 lists the parameters for two experiments sampling typical conditions for this pixel. In experiment #1, the gas is warm (~27 K) and moderately dense ($\sim 4.5\times10^3$ cm^{-3}). In experiment #2, the gas is colder (~22 K) and denser ($\sim 4.5\times10^4$ cm^{-3}). Tables 6 and 7 list the characteristics of the lines for experiment #1 and #2, respectively. In this section, calibration errors

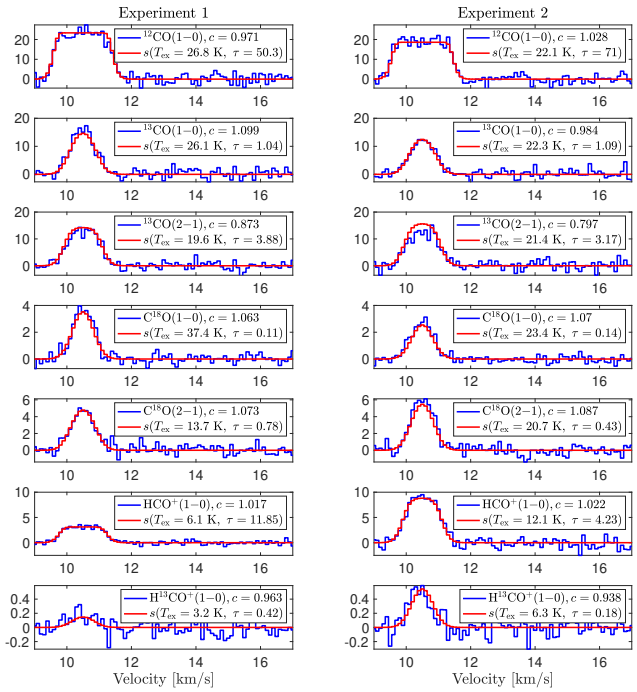


Fig. 5. Comparison of the ideal (in red) line profiles with noisy (in blue) ones for experiment #1 and #2 (see Table 5). In the legends, c is the estimated calibration factor (see Eq. (17)).

and thermal noise are treated as follows. On one hand, when generating the spectra, we use the model described in Sect. 3 to simulate data with calibration errors and thermal noise. We directly use a thermal noise level that will give a maximum S/N value of 10, except for the low S/N line $H^{13}CO^+$. On the other hand, when fitting the spectra, we use the likelihood that only takes into account the additive noise as in the fit of the actual data in Sect. 3. As explained there, this allows the low S/N lines to contribute to the fit.

Figure 5 compares the associated typical profiles of the spectra. For each experiment, we have computed 200 Monte-Carlo realizations with white Gaussian drawings for the thermal noise and calibration error. We fitted the obtained spectra with the maximum likelihood estimator described in Sect. 2.4. We thus sample the probability distribution function of the fitted parameters.

In experiment #1, the peak S/N of $H^{13}CO^+$ (1 – 0) is low (∼1.5) and this line has a probability of detection of 63%, whereas all the lines are detected in experiment #2, i.e., the peak $S/N \gtrsim 5$.

The ^{12}CO (1 – 0) line has a large opacity ($\tau > 50$) in both experiments and its excitation temperature is equal to the kinetic temperature. This can be understood as a consequence of the significantly large volume density of the gas with respect to the line's effective critical density. Very large opacities deliver flat-top line profiles that are not found in actual observations of ^{12}CO (1 – 0) lines in interstellar clouds. The absence of flat topped ^{12}CO profiles shows that even ^{12}CO probes into the cloud instead of just its surface because of the large velocity gradients along the LoS. This example shows that just fitting the integrated line emission is not sufficient. Estimating the goodness of fit requires a comparison between the modeled and measured line profiles. For this reason, we only used the peak emission of the ^{12}CO (1 – 0) line, but we did fit the whole profile for the other species, including ^{13}CO and C^{18}O.

The ^{13}CO and C^{18}O lines have different excitation regimes depending on the experiment. In experiment #2, where the volume density is one order of magnitude larger than the (thin and effective) critical densities, both the (1 – 0) and (2 – 1) lines are close to the local thermodynamic equilibrium (LTE), i.e., their excitation temperature is close to the kinetic temperature. However, in experiment #1, where the volume density is of the same order of magnitude as the (thin and effective) critical densities, only the ^{13}CO (1 – 0) line is close to LTE. The ^{13}CO and C^{18}O (2 – 1) lines are sub-thermally excited, and the C^{18}O (1 – 0) line is supra-thermally excited. Such a regime of suprathermal excitation of the low-J CO lines has already been pointed out by Leung & Liszt (1976). The HCO^+ and $H^{13}CO^+$ (1 – 0) lines are sub-thermally excited in both experiments.

The C^{18}O (1 – 0) and (2 – 1) lines, as well as the $H^{13}CO^+$ (1 – 0) line, are optically thin in both experiments. The ^{13}CO lines have similar opacities (1 and 3 – 4 for the (1 – 0) and (2 – 1) lines) in both experiments. The opacity of the HCO^+ (1 – 0) line is three times larger in experiment #1 than in experiment #2 where the lowest energy levels are much less populated because the density is higher.

4.2. Structure of the figures

Figures 6 to 8 show the comparison between the Cramér-Rao bound reference precision and the parameters fitted on Monte-Carlo drawings of noisy RADEX models for experiments #1 and #2. In the three figures, the top row describes the RADEX input parameters. The next rows show the results obtained when fitting different sets of lines. We first show the results when the assumptions are consistent between the model used to generate the data and the fitted spectra. We then show the results when these fitting assumptions are "biased" (e.g., simplified) with respect to the (true) generative models. In these cases, the bottom row lists the values of the parameters of the generative models, which are modified compared to the top row, i.e., these changes are ignored when fitting the spectra.

The first column describes the set of lines used and the kind of assumptions made (biased or unbiased) during the fits. The next two columns then compare the estimated column densities for both experiments. The estimated column densities are shown as scatter plots, except for the second row of Fig. 6 where only one species is fitted. In this specific case, we compare the histogram from the Monte-Carlo experiment (in green) with the normal density of an unbiased efficient estimator (i.e., with a mean equal to the true input value, and a variance equal to the CRB of $\log N$) in red. The last two columns show the scatter plots of the estimated kinetic temperature and volume density.

In these figures, the green dots correspond to the estimated parameters obtained on the 200 independent realizations of the Monte-Carlo experiments. The red ellipse is a confidence region centered on the true parameters. Its size and orientation is computed from the Cramér-Rao bound matrix. The percentage of realizations that lie in this confidence ellipse is mentioned in the legend of the graph. Any unbiased efficient estimator will deliver 99% ± 1% of its estimations inside the confidence ellipse.

When the temperature and volume density vary among the species, the red cross indicates the true values for ^{13}CO and/or C^{18}O, and the black cross indicates the true values for the other species (^{12}CO or HCO^+). This allows us to visualize the effect of the biased a priori knowledge on the fit. The axes and associated uncertainties are computed on the logarithm of the estimated

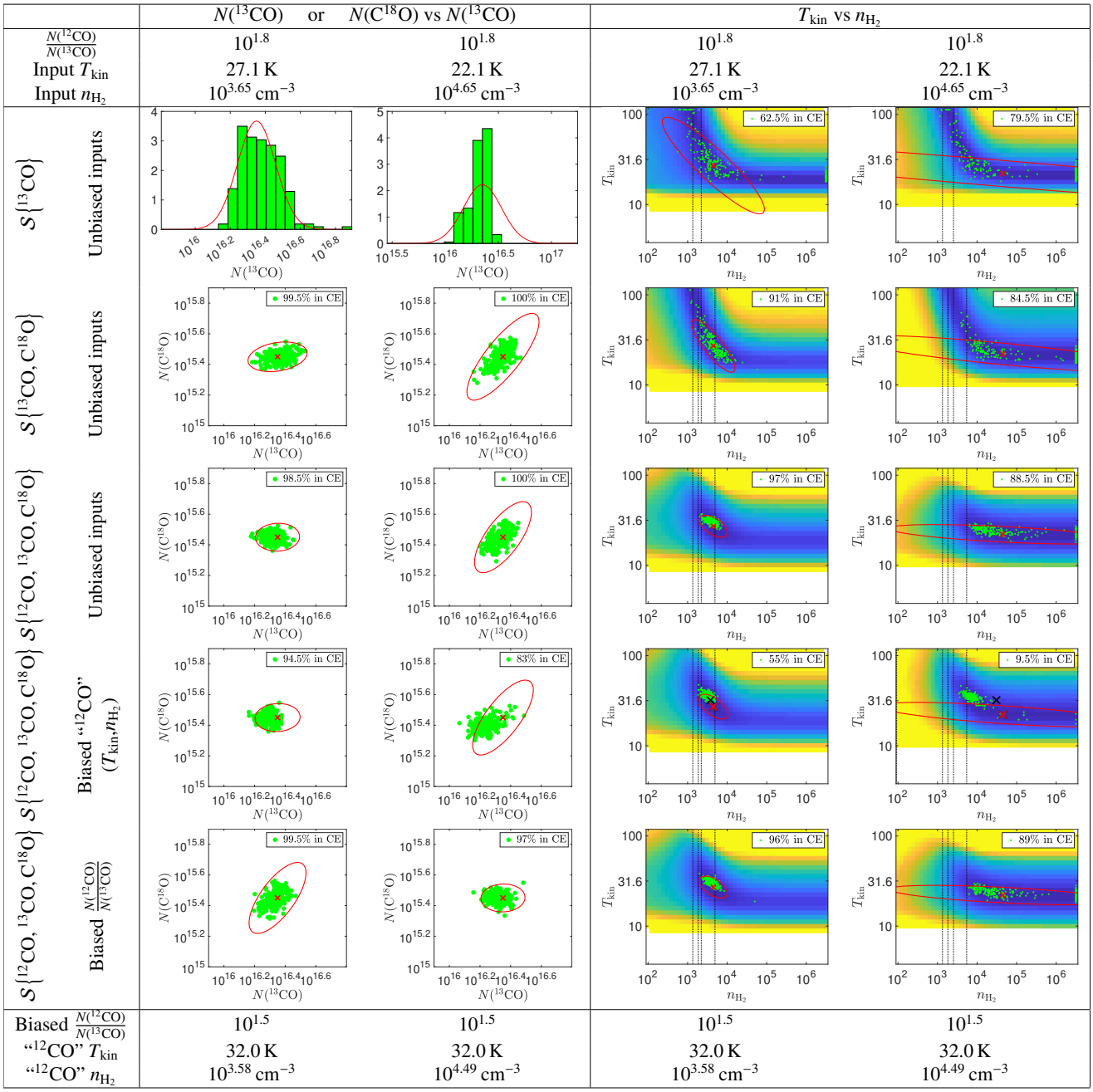


Fig. 6. Comparison of the performances of our maximum likelihood estimator for the estimation of the column densities (Cols. #2 and #3) and the couple $(T_{\rm kin}, n_{\rm H_2})$ (Cols. #4 and #5) and for two examples of kinetic temperature and volume density. The first example (Cols. #2 and #4) at medium density and warm temperature. The second one (Cols. #3 and #5) at high density and colder temperature. In the first and second rows we estimate the parameters with the ^{13}CO lines, and the (^{13}CO, C^{18}O) lines, respectively. In the last three rows we estimate the parameters with the lines of the three main CO isotopologues. In the last two rows we biased one of the parameters for the ^{12}CO RADEX model to test the effect of incorrect physical assumptions in the estimator. The kinetic temperature and volume density of the ^{12}CO RADEX model is biased in the fourth row, while the ratio of abundances of ^{12}CO and ^{13}CO is biased to a smaller values than assumed in the estimator on the fifth row. In all panels, the modeled and biased values of the estimated parameters are shown with the red and black crosses, respectively. The red ellipses, computed with the Cramér-Rao bound, are confidence ellipses (CE) of efficient estimations with a probability of 99%. The green points shows the estimated values for 200 simulated spectra. The color images on Cols. #3 and #4 show the values of the negative log-likelihood for the couple $(T_{\rm kin}, n_{\rm H_2})$ that our estimator assumes to be common to all analyzed species. The vertical dotted line are the effective critical density of the different transitions.

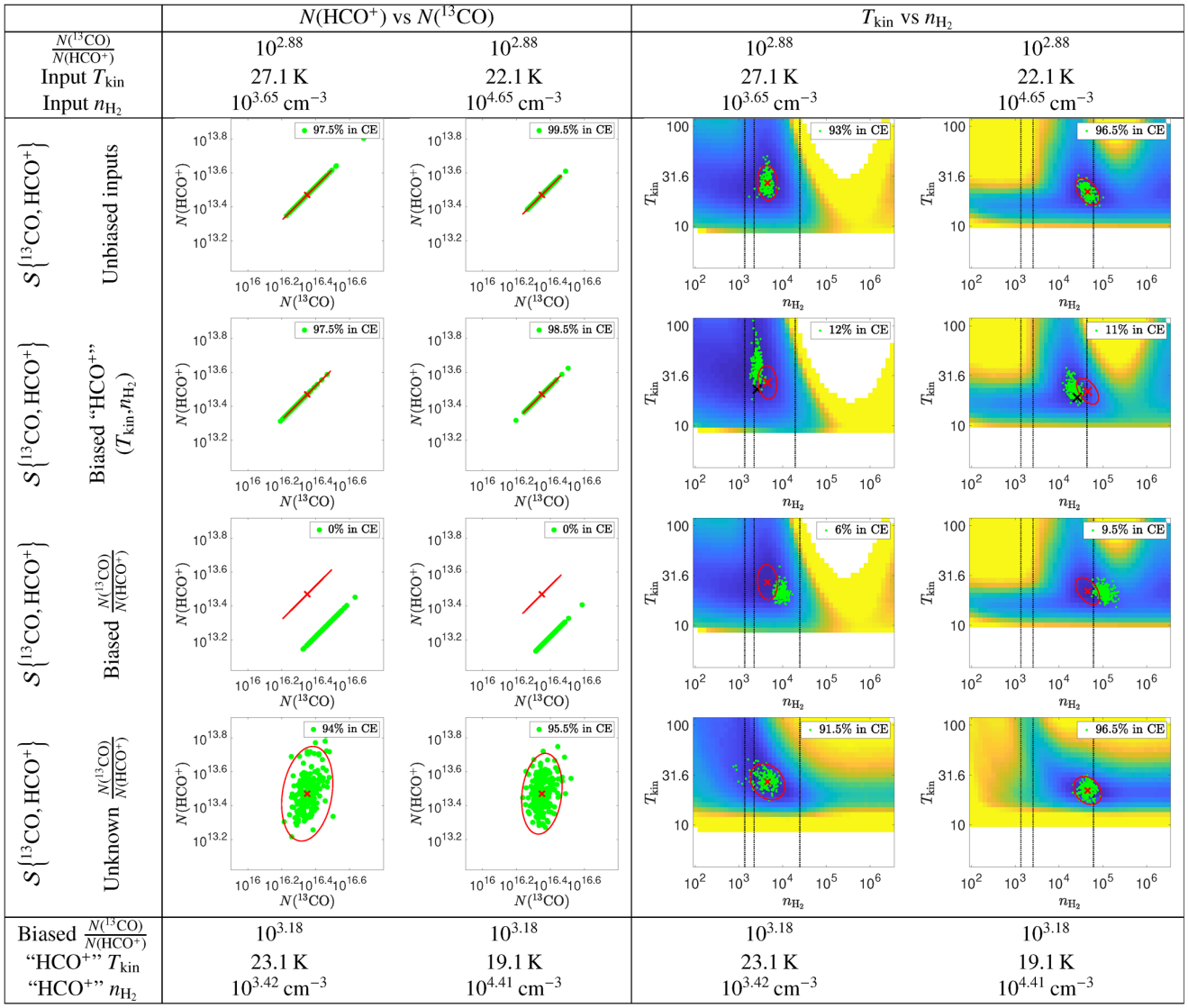


Fig. 7. Same as Fig. 6 except that 1) the couple of species used to estimate the parameters are ^{13}CO and HCO$^+$, and 2) the biases on the RADEX model are imposed on the HCO$^+$ species.

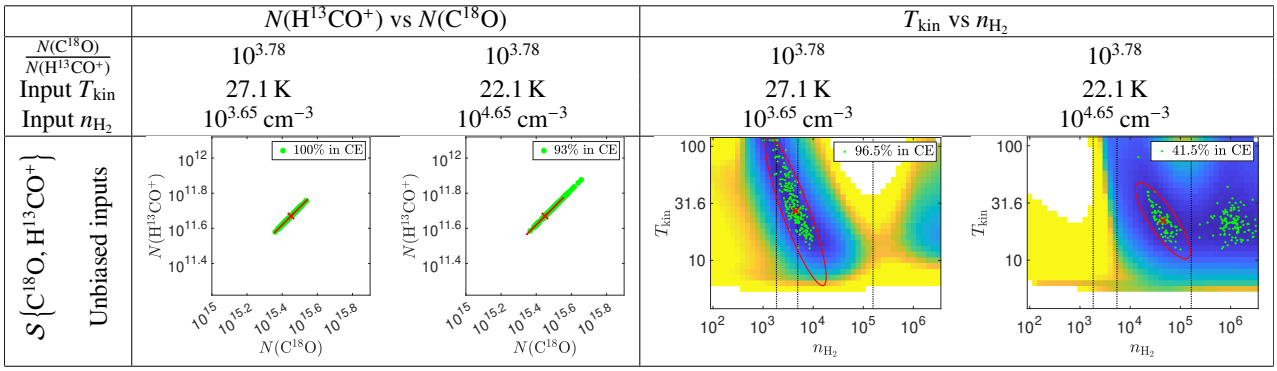


Fig. 8. Same as Fig. 6 except that 1) the couple of species used to estimate the parameters are C^{18}O and H^{13}CO$^+$, and 2) the RADEX model are not biased.

parameters. In Figs. 7 and 8, where the abundance ratios of the column densities is assumed fixed during the fit, the estimated values lie on a straight lines, and the confidence ellipses become straight lines as well. In other words, the estimation of these column densities are perfectly correlated, as expected.

The $T_{\rm kin}$ vs. $n_{\rm H_2}$ scatter plots are overlaid on the variations of the negative log-likelihood as a function of $T_{\rm kin}$ and $n_{\rm H_2}$. In all generality, the negative log-likelihood variations lie in a space of dimension larger than two (the true value of estimated parameters). However, as explained in Appendix B.2, the variations of the negative log-likelihood can be projected on the 2D sub-space for the couple of parameters $(T_{\rm kin}, n_{\rm H_2})$ because the maximum likelihood estimator assumes that all species have the same $T_{\rm kin}$ and $n_{\rm H_2}$. The images thus show the projection of the 4D matrix of negative log-likelihood onto the plane $(n_{\rm H_2}, T_{\rm kin})$ by choosing the column density (N) and velocity dispersion (σ_V) that maximize it for each value of $(n_{\rm H_2}, T_{\rm kin})$.

This projection of the negative log-likelihood is shown for the ideal realization of the RADEX model without Gaussian noise. While the shape of the negative log-likelihood varies slightly with the noise realization, its minimum location that yields the values of the estimated parameters may change drastically depending on the specific noise realization. This explains why the estimated parameters shown as green dots explore the blue valleys in these images. Finally, the effective critical densities of the different lines are overlaid as dotted vertical lines.

4.3. Results for the column densities

Figures 6 to 8 show that the column densities are derived with a low variance (high certainty). Biased hypotheses on kinetic temperature and volume densities do not influence the column density values much. In contrast, biased hypotheses on abundance ratios bias the derived column densities.

For most of the cases, more than 95% of the estimated column densities (^{13}CO, C^{18}O, HCO$^+$, or H^{13}CO$^+$) are in the confidence ellipse, illustrating the good performance of the Maximum Likelihood Estimator (MLE). The variance on the column density reduces as the number of assumptions implemented in the estimator increases. For some scenario (e.g., first row of Fig. 6, experiment #2), enforcing the volume density to be smaller than $10^{6.5}$ cm^{-3} provides column density estimations with smaller variance than the one predicted by the CRB.

The effects of biased assumptions depend on the case. On one hand, assuming a too low kinetic temperature or a too high volume density for ^{12}CO does slightly bias the estimated column densities of ^{13}CO and C^{18}O (fifth row of Fig. 6), but they remain inside the CRB ellipses. Similarly, assuming an incorrect abundance ratio for $N(^{12}\mathrm{CO})/N(^{13}\mathrm{CO})$ does not impact the values of the column densities of ^{13}CO or C^{18}O much. This is probably due to the fact that only the peak intensity of the ^{12}CO $(1-0)$ line is used to constrain the kinetic temperature. On the other hand, assuming an abundance ratio for $N(^{13}\mathrm{CO})/N(\mathrm{HCO^+})$ twice as low as the true values artificially increases the column density of ^{13}CO and decreases the column density of HCO$^+$. The bias introduced on $N(^{13}\mathrm{CO})$ is smaller than the one on $N(\mathrm{HCO^+})$.

4.4. Results for the kinetic temperature and the volume density

The situation for the physical conditions (the kinetic temperature and volume density) is different. In this section, we first compare the shape of the negative log-likelihood as a function of the kinetic temperature and volume density in the different cases. We then analyze the variances and biases of these parameters for the two experiments separately.

4.4.1. Shape of the negative log-likelihood image

The shape of the NLL significantly depends on the combination of lines/species used, in particular on their effective critical densities. When fitting only the CO isotopologue lines (Fig. 6), the minimum of the NLL has the shape of a valley at approximately constant thermal pressure ($T_{\rm kin} \propto 1/n_{\rm H_2}$) below a density of $\sim 10^4$ cm^{-3} and constant temperature above this density. When fitting $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO^+}\right\}$ (Fig. 7), the shape of the NLL minimum (blue) region is more complex. For experiment #1, where the modeled gas has $T_{\rm kin} = 27.1$ K and $n_{\rm H_2} = 10^{3.65}$ cm^{-3} (i.e., a thermal pressure of $\sim 1.2 \times 10^5$ K cm^{-3}), it is bounded at an approximately low constant temperature of ~ 15 K and at an approximately constant pressure of $\sim 2 \times 10^5$ K cm^{-3}. However, the chance of getting a second local NLL minimum at a $n_{\rm H_2} \gtrsim 10^6$ cm^{-3} exists and it increases in experiment #2, where the modeled gas has $T_{\rm kin} = 22.1$ K and $n_{\rm H_2} = 10^{4.65}$ cm^{-3} (i.e., a thermal pressure of $\sim 10^6$ K cm^{-3}). This implies that some ambiguity between a low and a high volume density solution can occur. The shape is qualitatively similar for $\mathcal{S}\left\{\mathrm{C^{18}O}, \mathrm{H^{13}CO^+}\right\}$. However, many estimations are localized on the second local maxima in experiment #2 because the two local maxima gets closer to each other compared to experiment #1 (see Cols. 3 and 4 of Fig. 8).

4.4.2. Experiment #1

The maximum likelihood estimator (MLE) only becomes efficient (most of the dots gather inside the CRB ellipse) and the variation of the estimation decreases quickly when relevant information is delivered by adding lines.

In Fig. 6, the confidence ellipse for $n_{\rm H_2}$ covers 2 orders of magnitude when studying $\mathcal{S}\left\{^{13}\mathrm{CO}\right\}$, one order of magnitude when studying $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$, and a factor 1.9 when studying $\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$. The confidence ellipse for $T_{\rm kin}$ covers about 1 order of magnitude when studying either $\mathcal{S}\left\{^{13}\mathrm{CO}\right\}$ or $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$, and a factor 1.4 when studying $\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$. Adding the information about the peak temperature of a very optically thick line at LTE as ^{12}CO $(1-0)$ allows us to constrain the temperature.

A combination of optically thick and thin lines improves the determination of the density by an order of magnitude (case $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$). But the volume density only becomes well constrained at the same time as the temperature when adding the fit of the ^{12}CO $(1-0)$ peak temperature. Figure 7 indicates that the fitting of $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO^+}\right\}$ delivers small uncertainty intervals compared to, e.g., the fitting of $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C^{18}O}\right\}$. Increased precision here is not a question of the S/N, as the S/N of C^{18}O is higher than that of HCO$^+$, but of difference in the critical density. Indeed, the effective critical density of the HCO$^+$ $(1-0)$ line is more than one order of magnitude greater than those for the two ^{13}CO lines (see vertical lines).

Having lines of different critical densities is, however, not always sufficient. In Fig. 8, the uncertainties of the kinetic temperature and volume density are about one order of magnitude

each when fitting $\mathcal{S}\left\{C^{18}O, H^{13}CO^+\right\}$ whose critical densities differ by more than one order of magnitude. In this case, all lines are optically thin, implying a degeneracy between temperature and density, the pressure being relatively well constrained. We verified that the $(T_{\rm kin}, n_{\rm H_2})$ degeneracy remains when increasing the S/N of the $H^{13}CO^+$ line by decreasing the noise level σ_b by a factor 10.

Assuming that the spatial portion of the LoS that dominates the emission of the lines is identical for the ^{12}CO line and its rare isotopologues slightly biases the estimation to high kinetic temperatures and low volume densities (row 4 of Fig. 6). This confirms that the kinetic temperature is above all constrained by the peak temperature of the ^{12}CO (1 − 0) line. In the test where the generative model assumes that the HCO^+ line emission comes from gas at a slightly different temperature and volume density than the gas producing the ^{13}CO emission (2nd row of Fig. 7), the kinetic temperature becomes highly biased. As we know that dense cores and filaments are surrounded by a more diffuse, lower density and warmer envelope, a multi-layer model could be more suited in this case. This will be the subject of another article.

Assuming an $N(^{13}CO)/N(HCO^+)$ abundance ratio lower than reality by a factor 2 positively biases the estimated volume density by a factor 2 as shown in the fourth row of Fig. 7. The main difference with the impact of biases on the column densities is that a majority of the estimations lies outside the CRB ellipses.

4.4.3. Experiment #2

Several significant differences exist between Experiment #1 and Experiment #2. While the accuracy of the temperature increases significantly when adding the constraint from the ^{12}CO (1 − 0) peak temperature, no information on the volume density can be recovered in experiment #2, when using only the CO isotopologue lines. The CRB interval of confidence for $n_{\rm H_2}$ covers the full range of explored densities. This is due to the fact that all lines get close to the LTE state (see the excitation temperature in Table 7) where only the kinetic temperature determines the excitation of the species.

The fifth column of Fig. 7 shows that improving the $n_{\rm H_2}$ estimation requires us to analyze a line or species with a higher effective critical density than the considered CO isotopologue lines. But here again, this is only true when some of the fitted lines are optically thin and others are optically thick in order to lift the degeneracy between kinetic temperature (kinetic energy of collisions) and volume density (number of collisions) during the excitation of the species. Moreover, Fig. 8 shows that there exist ambiguities between the two local minima of the NLL at the true volume density of the generative model or at a much higher density value.

Just like experiment #1, an underestimation of the a priori knowledge on the abundance ratio between ^{13}CO and ^{12}CO leads to an overestimation of $n_{\rm H_2}$, but the induced bias, which can be observed in the fourth row of Fig. 6, is much more important here: almost a factor 10. Not only is the volume density biased, but even worse, the observed variance is small. This means that the observer obtains a region with low spatial variance for the estimated parameters when the maximum likelihood estimator is applied to several pixels, thus giving an incorrect impression of the robustness of the results. Such an example shows that a physical or chemical misspecification in the hypotheses of the model used to fit the data can lead to incorrect estimates of the fitted parameters, most notably the kinetic temperature and/or the volume density.

One possible explanation of this bias is that since $n_{\rm H_2} \leq n_{\rm crit}^{\rm eff}$, the non-LTE excitation model feeds the line upper level with more collisions to compensate for the underestimation of $N(HCO^+)$ in order to produce a sufficiently bright HCO^+ (1 − 0) line.

4.5. Intermediate summary

We have studied the estimation of the column densities, kinetic temperatures, and volume densities when fitting different combinations of the CO and HCO^+ isotopologue lines with the RADEX escape probability model, on two example lines of sight. While the Cramér-Rao bound is a lower limit for the variance of any unbiased estimator, we here confirm that this lower limit can be reached by a "standard" maximum likelihood estimator, provided that the assumptions used during the fitting process are consistent (i.e., unbiased) with respect to reality (generative model).

Estimations of the column density are efficient on the considered examples, even when the uncertainty (CRB) on $n_{\rm H_2}$ or $T_{\rm kin}$ are large. This means that the escape probability models provide robust estimations of column densities allowing for the production of maps in a wide variety of situations, which can be quantitatively characterized a priori with $\mathcal{B}^{1/2}(\log N) < 0.1$. Assuming an incorrect abundance ratio for, e.g., $N(^{13}CO)/N(HCO^+)$ biases the estimation of $\log(N(HCO^+))$ more than the estimation of $\log(N(^{13}CO))$.

Likewise, we confirm that CRB precisions for the volume density and kinetic temperature can be achieved by a "standard" maximum likelihood estimator when $\mathcal{B}^{1/2}(\log T_{\rm kin}) < 0.1$, and $\mathcal{B}^{1/2}(\log n_{\rm H_2}) < 0.1$. A good estimation of the kinetic temperature and volume density requires a combination of high-enough S/N lines of different opacities (thick and thin) and effective critical densities. On one hand, this is delivered by the fit of $\mathcal{S}\left\{^{12}CO, ^{13}CO, C^{18}O\right\}$ in experiment #1, where $n_{\rm H_2} = 10^{3.65}\,{\rm cm^{-3}}$ and $T_{\rm kin} = 27.1\,{\rm K}$. In this case, a perfect knowledge of $N(^{12}CO)/N(^{13}CO)$ is not required, but an incorrect assumption about the gas temperature that emits the ^{12}CO line biases both the kinetic temperature and the volume density. This is consistent with the fact that the ^{12}CO peak is mostly useful to regularize the estimation of $T_{\rm kin}$. On the other hand, fitting $\mathcal{S}\left\{^{12}CO, ^{13}CO, C^{18}O\right\}$ in experiment #2, where $n_{\rm H_2} = 10^{4.65}\,{\rm cm^{-3}}$ and $T_{\rm kin} = 22.1\,{\rm K}$, fails to constrain the volume density, in agreement with the fact that the low-J lines of the CO isotopologues have almost reached LTE.

Fitting a set of lines that have different critical densities enlarges the range of $(T_{\rm kin}, n_{\rm H_2})$ values for which the kinetic temperature and volume density can be estimated as long as some of the lines are optically thin and others optically thick. An incorrect assumption on the abundance ratio between ^{13}CO and HCO^+ biases the estimation of the temperature and density because the corresponding (1 − 0) lines have different effective critical densities.

5. Parametric study of the reference precisions on the column densities, kinetic temperature, volume density and thermal pressure

The Cramér-Rao bound gives the reference precision that can be achieved by an unbiased and efficient estimator for a given physical model (here RADEX) and observational setup (combination

Table 8. Considered precisions on $\log x$, and x, plus the associated color codes systematically used in Sect. 5.

$\mathcal{B}^{1/2}(\log x) = \sigma$	0.03	0.06	0.1	0.3	0.6	1	2
$[10^{-\sigma}, 10^{+\sigma}]$	[0.93, 1.07]	[0.87, 1.15]	[0.79, 1.25]	[0.5, 2.0]	[0.25, 4.0]	[0.1, 10]	[0.01, 100]
Color code	Light green	Green	Yellow	Orange	Red	Pink	Grey

of lines and noise levels). In the previous section, we studied two specific combinations of physical conditions for a molecular cloud. We now extend this quantitative study to a large domain of physical conditions that can be constrained with a suitable precision for different combinations of species, associated lines, and a priori knowledge linking them.

We will restrict the combination of species to the CO and HCO^+ main isotopologues and the combination of lines to the $(1-0)$ transition for all species, plus the $(2-1)$ transition for the ^{13}CO and $C^{18}O$ species. In more details, we will study the information that can be extracted in the following sets of observations: 1) We analyze in depth the simplest case, i.e., when only the $(1-0)$ and $(2-1)$ lines of ^{13}CO are available. 2) We then quantify the gain in precision when adding the same lines of the rarer $C^{18}O$ isotopologue. 3) We also compare the benefits of two different combinations of one CO isotopologue with one isotopologue of HCO^+. We will combine $C^{18}O$ and $H^{13}CO^+$ that are easy to observe in dense cores, but also ^{13}CO and $H^{12}CO^+$ that are better suited to probe the lower volume density envelope. 4) We quantify the gain in precision when adding the constraint from the peak temperature of the ^{12}CO $(1-0)$ lines to the previous combinations of species and lines. 5) We then quantify the loss of information when only the lowest J line of ^{13}CO and $C^{18}O$ are available. This is of particular interest for the ORION-B project that mapped a large fraction of the Orion B cloud between 70 and 116 GHz. 6) We finally check the potential advantage of fixing the relative abundances.

For these studies, we will explore the variations of the precision in the two-dimensional space: column density of molecular hydrogen ($10^{20} \leq N(H_2) \leq 10^{24}\,cm^{-2}$) vs. its volume density ($10^2 \leq n_{H_2} \leq 10^6\,cm^{-3}$). The other parameters to generate the model and compute the CRB are as follows.

Physical parameters. A kinetic temperature of 22 K, a velocity dispersion of $0.32\,km\,s^{-1}$, and a centroid velocity of $0\,km\,s^{-1}$.

Chemical abundances. We use those defined in Eq. (16) in Sect. 3, plus the following ^{13}CO abundance relative to molecular hydrogen

$$\frac{N(^{13}CO)}{N(H_2)} = 2.5 \times 10^{-6} = 10^{-5.6}. \quad (22)$$

The range of studied column densities for the different species has been adapted so that it fits the range of molecular hydrogen column densities, given the chosen relative abundances.

5.1. Typical figure layout

The first column of Fig. 9, as well as Figs. 11–13 compare the CRB reference precisions when combining different low-J lines of different species. This precision is plotted for different values of column densities (y-axes) and volume densities (x-axes). The values of the parameters that are fixed when generating the RADEX spectra are listed in the bottom row. The reference precision depends on the a priori assumptions that are taken into account by the estimator. These are listed in the second row from the top. From top to bottom, the panels show the precisions on the species column densities, and gas kinetic temperature, volume density, and thermal pressure.

For any parameter x, $\mathcal{B}^{1/2}(\log x)$ is equal to the standard deviation of any unbiased and efficient estimators of $\log x$. Table 8 lists the relationship between $\mathcal{B}^{1/2}(\log x)$ and the associated relative interval of confidence on x and the adopted color code. We are going from a $\pm1\sigma$ interval to a multiplicative factor interval of $[10^{-\sigma}, 10^{+\sigma}]$. A $\pm1\sigma$ value of 0.03 gives a relative confidence interval of the same order of magnitude as the calibration error (5-10%). A $\pm1\sigma$ value of 0.1 corresponds to a multiplicative uncertainty of the order of 20%. This is considered a good precision for the column density and kinetic temperature. Values of 0.3 and 0.6 for $\pm1\sigma$ give a precision within a factor of 2 and 4, respectively. Getting a reliable estimate of the volume density at this precision level would already be a great achievement. Values of 1 and 2 for $\pm1\sigma$ deliver an uncertainty that covers between two and four orders of magnitude. The associated parameters are too uncertain to be useful.

In order to emphasize these different ranges of relative precisions, the maps of $\mathcal{B}^{1/2}(\log x)$ are systematically color-coded throughout the paper as follows: green means good to excellent relative precisions, yellow and orange standard relative precisions, and red to pink poor relative precisions. The grey color is used to indicate the parameter space where the precision is so low that we consider that the lines deliver no information on the considered parameter.

5.2. A simple case: ^{13}CO (1-0) and (2-1)

Figure 9 compares the CRB reference precisions with the main characteristics of the $(1-0)$ and $(2-1)$ lines for ^{13}CO. The precisions depend much on the considered physical parameter.

5.2.1. Precision of the column density

The column density is the easiest parameter to estimate. It can be estimated within a factor of two for a large variety of column and volume densities. However the range of column densities with a good precision decreases from three orders of magnitude at low volume density to 0.5 order of magnitude at high volume density.

A precision better than 25% can only be reached for volume densities lower than $5 \times 10^3\,cm^{-3}$ and a column density between $10^{21.5}$ and $10^{22.5}\,cm^{-2}$. The latter conditions imply $\tau_{(1-0)} < 10$, $\tau_{(2-1)} > 1$, and $n_{H_2} < 10\,n_{crit}^{eff}$ for the two lines. The influence of the opacity can be interpreted as a trade-off between having a sufficiently high S/N while avoiding a loss of information on the column density because the line intensities saturate. The influence of n_{H_2}/n_{crit}^{eff} is more complex but can be related to the radiative pumping effect at high column densities and low volume densities.

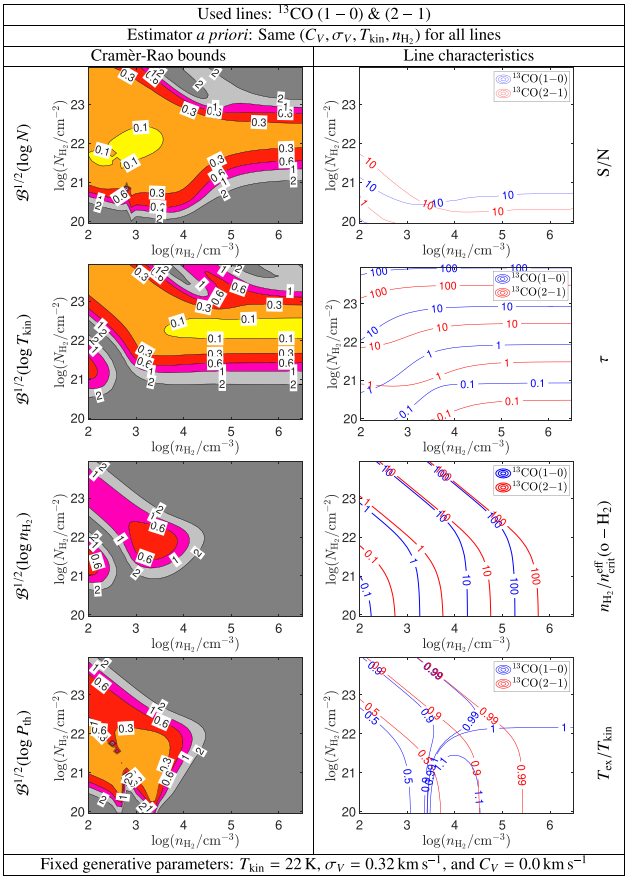


Fig. 9. Comparison of the CRB reference precisions (left column) on the column density, kinetic temperature, volume density, and thermal pressure, when fitting the ^{13}CO (1 – 0) and ^{13}CO (2 – 1) lines, with these lines main characteristics (right column): S/Ns, opacities, ratios of the volume density to the effective critical densities, and ratios of the excitation temperature over the kinetic temperature. The blue and red curves show the results for the (1 – 0) and (2 – 1) lines, respectively. The values of the parameters that are fixed when generating the RADEX spectra are listed in the bottom row. The a priori knowledge that are used when computing the CRB precisions are listed in the second row.

5.2.2. Precision on the kinetic temperature

The kinetic temperature can be estimated within a factor of two in a narrower range of conditions that typically spans one order of magnitude for the column density around $N(\mathrm{H_2}) = 10^{22.3}\,\mathrm{cm^{-2}}$ when the volume density is larger than the optically thin critical density $\sim 10^3\,\mathrm{cm^{-3}}$.

A precision better than 25% can only be reached in a different parameter space from the one that maximimizes the precision on the column density. It corresponds to a column density $10^{22.1} < N(\mathrm{H_2}) < 10^{22.6}\,\mathrm{cm^{-2}}$ for $n_{\mathrm{H_2}} > 10^{3.6}\,\mathrm{cm^{-3}}$. This implies similar conditions for the line opacities ($\tau_{(1-0)} < 10$, $\tau_{(2-1)} > 1$), but a different condition on the line critical density ($n_{\mathrm{H_2}} > 5\,n_{\rm crit}^{\rm eff}$).

For high opacity, Eq. (1) simplifies at its peak value to $s \simeq J(T_{\rm ex}) - J(T_{\rm CMB})$. Moreover, when $n_{\mathrm{H_2}}$ is sufficiently above $n_{\rm crit}^{\rm eff}$, the local thermodynamic equilibrium is reached, and $T_{\rm ex} = T_{\rm kin}$ for both lines. It thus makes sense that the estimation of $T_{\rm kin}$ can be accurate when these conditions are met. It is counter-intuitive that the precision decreases when the ^{13}CO column density increases above $10^{17}\,\mathrm{cm^{-2}}$ and $n_{\mathrm{H_2}} \gtrsim 10^4\,\mathrm{cm^{-3}}$. Indeed, the approximation $s \simeq J(T_{\rm kin}) - J(T_{\rm CMB})$ is expected to be valid under these conditions, and the gas is expected to be at local thermodynamic equilibrium.

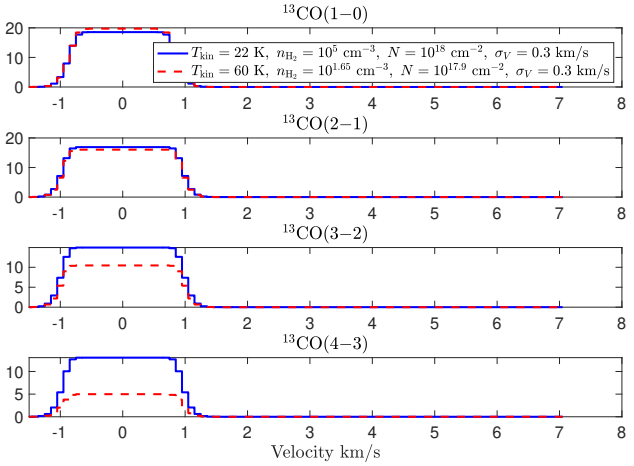


Fig. 10. Comparison of the spectral shape of the (1 – 0) (top) to (4 – 3) (bottom) lines of ^{13}CO for two different combinations of kinetic temperature, volume density, and column density. The plain blue and dashed red histograms show the spectra for (22 K, 10^5 cm^{-3}, 10^{18} cm^{-2}) and (60 K, $10^{1.65}$ cm^{-3}, $10^{17.9}$ cm^{-2}), respectively. The lines share the same centroid velocity C_V = 0 km s^{-1} and velocity dispersion σ_V = 0.32 km s^{-1}. The two sets of physical conditions deliver degenerate spectral profiles within the noise level for the (1 – 0) and (2 – 1) lines. Additional information, e.g., the (3 – 2) or (4 – 3) lines, is needed to lift this degeneracy.

To understand this phenomenon, we compare in Fig. 10 the spectra of the lines for the (1 – 0) to (4 – 3) transitions of ^{13}CO for two different pairs of ($T_{\rm kin}$,$n_{\mathrm{H_2}}$) values, (22 K, 10^5 cm^{-3}) and (60 K, $10^{1.65}$ cm^{-3}), and logarithms of the column densities that differ by the precision we aim for, i.e., 0.1: $N(^{13}\mathrm{CO}) = 10^{18}\,\mathrm{cm^{-2}}$ or $10^{17.9}\,\mathrm{cm^{-2}}$. The (1 – 0) and (2 – 1) line have similar profiles within the calibration uncertainty for these two sets of conditions, implying uncertain kinetic temperatures and column densities. An ambiguity therefore remains between a high density, low kinetic temperature case, and a low density, high kinetic temperature case, even for a large molecular column density[2]. Table 9 lists the main parameters of the lines shown in Fig. 10. It shows that only $T_{\rm ex}$ are significantly different between the blue and red profiles, and LTE is not reached for the red profiles.

This example shows the power of the Cramér-Rao bound. It quantifies the precision that any unbiased efficient estimator will reach for a given model, independent on the physical insight we can have. Indeed, assuming that the peak temperature of the spectra is a good approximation of the kinetic temperature for optically thick lines is an a priori knowledge that would change the model of Eq. (1), and thus the CRB computation. When the RADEX model is used for fitting noisy data without this a priori knowledge, the precision on the kinetic temperature and column density decreases significantly. Two other possibilities to increase the precision would be to average many observations of the (1 – 0) or (2 – 1) lines to decrease the noise level and the uncertainty of the calibration or to add the observations of the

[2] The fact that at least one high-J CO line is required to break the degeneracy between these two regimes is true even though flat-top CO line profiles are absent in actual observations.

Table 9. Characteristics of ^{13}CO lines shown in Fig. 10.

Line	τ	T_{ex}	T_{peak}	n^{thin}_{crit}	n^{eff}_{crit}
	–	K	K	cm^{-3}	cm^{-3}
(1 – 0)	50	22.0	18.5	$10^{3.29}$	$10^{1.77}$
(2 – 1)	140	22.0	16.9	$10^{3.81}$	$10^{1.84}$
(3 – 2)	174	22.0	15	$10^{4.24}$	$10^{2.17}$
(4 – 3)	135	22.0	13	$10^{4.56}$	$10^{2.6}$
For T_{kin} = 22 K, n_{H_2} = 10^5 cm^{-3}, N = 10^{18} cm^{-2}.					
(1 – 0)	43	23.2	19.7	$10^{3.27}$	$10^{1.81}$
(2 – 1)	134	21.1	16	$10^{3.82}$	$10^{1.87}$
(3 – 2)	185	17.2	10.5	$10^{4.25}$	$10^{2.15}$
(4 – 3)	132	12.8	5	$10^{4.56}$	$10^{2.62}$
For T_{kin} = 60 K, n_{H_2} = $10^{1.65}$ cm^{-3}, N = $10^{17.9}$ cm^{-2}.					

Notes. n^{thin}_{crit} and n^{eff}_{crit} are the critical density of p – H_2. In both case σ_V = 0.3 km s^{-1}.

(3 – 2) or (4 – 3) lines with enough S/N because their profiles are much more sensitive to the variations of the physical conditions. A gain of a factor 10 on the precision of the estimation of σ_b is approximately required to gain the same factor on the estimation of the kinetic temperature. Adding lines is thus a better strategy than increasing the integration time because the required integration time to actually detect a new line is often significantly lower than the additional time needed to improve the S/N and calibration uncertainty on an already detected line.

5.2.3. Precision of the volume density or thermal pressure

The precision of the volume density is at best a factor of 4 for a small region around $n_{H_2} \sim 10^{3.5}$ cm^{-3} and $N(H_2) \sim 10^{22}$ cm^{-2}. As expected, when $T_{ex} \simeq T_{kin}$, the local thermal equilibrium is reached, and the line loses any sensitivity to the volume density. When $n_{H_2} \lesssim n^{eff}_{crit}$, the line becomes sensitive either to the volume density or to the thermal pressure. It is well known that when the lines become optically thin, the excitation is controlled by the product of the volume density and kinetic temperature, namely, the thermal pressure. The independent determination of the volume density and the kinetic temperature is difficult (the shape of their CRB are similar at low densities), while the thermal pressure can be determined to within a factor of two. It can therefore be interesting to derive thermal pressure maps of molecular clouds, which can be compared with maps of the magnetic or turbulent pressure.

5.2.4. Effective critical density versus excitation temperature

While the ratios of the volume density to the effective critical density show similar behaviors for the (1 – 0) and (2 – 1) lines, the ratios of the line excitation temperatures to the gas kinetic temperature behave quite differently for $n_{H_2} \gtrsim 10^{3.3}$ cm^{-3} and $N(H_2) \leq 10^{22}$ cm^{-2}. In this region $T_{ex}(1-0) \gtrsim T_{kin}$. This is the sign that the J = 1 energy level becomes super-thermally excited. While this inversion of population effect is moderate, it happens in a region of (T_{kin}, n_{H_2}) that is well represented in the Horsehead pillar.

5.3. One or two CO isotopologues

Figure 11 compares the precision that can be reached when studying either a single CO isotopologue (either ^{13}CO or C^{18}O) or their combinations. In all three cases we use their (1 – 0) and/or (2 – 1) lines, but we make no assumption on their abundances.

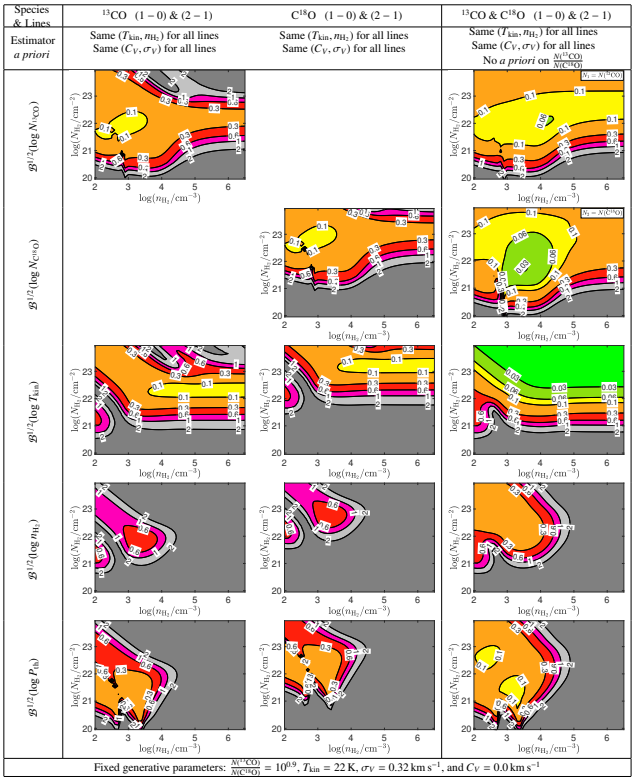


Fig. 11. Comparison of the CRB precisions when combining the first two J transitions of one or two CO isotopologues. These reference precisions are plotted for different values of column densities (y-axes) and volume densities (x-axes). The values of the parameters that are fixed when generating the RADEX spectra are listed in the bottom row. The precision depends on the a priori knowledges that are taken into account by the estimator. These are listed in the second row. The first two columns show the precisions for a single CO isotopologue (^{13}CO or C^{18}O, while the last column shows the precisions when studying the ^{13}CO and C^{18}O isotopologues simultaneously. From top to bottom, the panels show the precisions on the ^{13}CO and C^{18}O column densities, kinetic temperature, volume density, and thermal pressure. The CRB precisions are color coded as listed in Table 8.

The two first columns compare the cases when using the two lines of either ^{13}CO or C^{18}O. As C^{18}O and ^{13}CO have similar values of E_{up}, ν_l and A (see Table 2), the shapes of the CRB variations in the $N(H_2), n_{H_2}$ space are similar, but the results for the reference precisions with C^{18}O are vertically shifted by the fixed abundance ratio $N(^{13}CO)/N(C^{18}O) = 10^{0.9}$ with respect to the reference precisions with ^{13}CO. For a given $N(H_2)$ column density and noise level, this makes two major differences. First, the C^{18}O lines will have lower S/N. This difference will only have an impact when the S/N becomes lower than 10 as this is our noise saturation level. Second, the C^{18}O lines will often be optically thin when the ^{13}CO lines will have a higher opacity, approaching the optically thick regime. These two effects imply that the C^{18}O lines will be more suited to constrain gas with larger column densities than the ^{13}CO lines for the same velocity dispersion.

When we assume that the emission of ^{13}CO and C^{18}O is produced in gas with the same n_{H_2}, T_{kin}, and velocity dispersion, the combinations of different opacities and S/Ns for the same

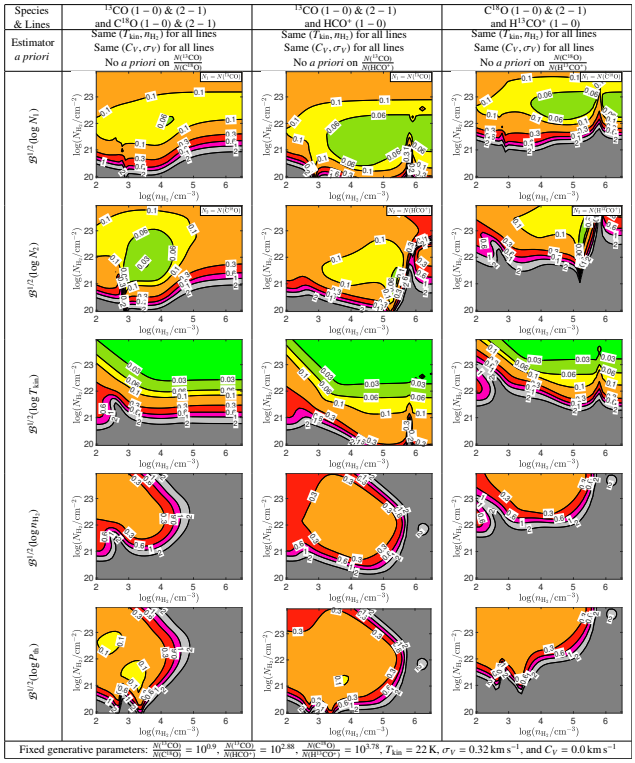


Fig. 12. Comparison of the CRB reference precisions for different combinations of two species and all their available lines. The layout of this figure is identical to Fig. 11, except for the combination of studied lines: the first column shows the precision when the (1 – 0) and (2 – 1) lines of the ^{13}CO and C^{18}O isotopologues are used. The last two columns study a combination of the one CO isotopologue with one HCO$^+$ isotopologue: ^{13}CO and H^{12}CO$^+$ on the second column and C^{18}O and H^{13}CO$^+$ on the third column.

rotational transition of two isotopologues will considerably modify the shape of the precision in the $N(H_2)$ vs. n_{H_2} space. This corresponds to the physical situation where ^{13}CO and C^{18}O are co-located spatially. The last column of Fig. 11 shows the CRB reference precisions when using both isotopologues together.

The most spectacular change on precision happens for the kinetic temperature. The low precision for $T_{\rm kin}$ at $N(H_2) \simeq 10^{23.6}$ cm^{-2} and $n_{H_2} \simeq 10^5$ cm^{-3} is now replaced by an excellent precision. Adding an isotopologue therefore helps to resolve the ambiguity between low density-high temperature and high density-low temperature gas as discussed above. As the C^{18}O lines have smaller opacities than the ^{13}CO ones, the associated effective critical densities $n_{\rm crit}^{\rm eff}$ are closer to the optically thin critical densities, and we can expect that n_{H_2} is more precisely estimated. Indeed, the best precision achieved on n_{H_2} increased by a factor 2.

5.4. Two CO isotopologues versus one CO plus one HCO$^+$ isotopologues

Figure 12 compares the CRB reference precisions when studying two molecular species among four. The first column shows the precision for the two low-J lines of ^{13}CO and C^{18}O, while the last two columns combines the two low-J lines of one of the CO isotopologues with the (1 – 0) line of one of the HCO$^+$ isotopologues.

As the effective critical density, $n_{\rm crit}^{\rm eff}$, of HCO$^+$ (resp. H^{13}CO$^+$) is significantly larger than the one of ^{13}CO (resp. C^{18}O, see Table 6), we expect a significant gain in precision for $n_{H_2} \gtrsim 10^4$ cm^{-3} compared to the estimations using only the ^{13}CO and C^{18}O lines. This is indeed the case for $N(H_2)$ column densities above $10^{21.1}$ cm^{-2} and below $10^{22.6}$ cm^{-2} for the $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$ set of species, and for $N(H_2)$ column densities above $10^{22.2}$ cm^{-2} for the $\mathcal{S}\left\{\mathrm{C}^{18}\mathrm{O}, \mathrm{H}^{13}\mathrm{CO}^+\right\}$ set. This is due to the lower abundance of C^{18}O and H^{13}CO$^+$ that turns into lower opacities, and thus a greater sensitivity to physical conditions at large column densities. For the $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$ set of species, the precision for the kinetic temperature, volume density and pressure is now reasonable for almost all the space of covered $N(H_2)$, n_{H_2} values. A precision better than 25% seems reachable for the conditions encountered in the Horsehead nebula, which corresponds to $\log(n_{H_2}) = 4 \pm 0.5$ cm^{-3} and $T_{\rm kin} \in [15, 40]$ K as shown in Fig. 16.

5.5. Adding the constraint from the ^{12}CO (1–0) peak temperature

Figure 13 compares the same combinations of species ($\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C}^{18}\mathrm{O}\right\}$ or $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$) and their associated available transitions with the additional information coming from the peak temperature (line temperature at the velocity channel that maximizes the intensity) of the ^{12}CO (1 – 0) line.

As expected by the fact that the ^{12}CO (1 – 0) line is highly optically thick, this additional constraint enables one to reach an estimation of $T_{\rm kin}$ within 10% for almost all the covered space of physical conditions. Adding this constraint also has a significant impact on the n_{H_2} accuracy for the $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{C}^{18}\mathrm{O}\right\}$ set. The impact remains modest for the other set, $\mathcal{S}\left\{^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$.

5.6. Only using the (1–0) lines

Roueff et al. (2021) showed in the LTE framework that using the two lowest J transition for ^{13}CO or C^{18}O considerably increases the space of $N(H_2)$, n_{H_2} values for which the CRB reference precision is reasonable. We find a similar effect when only studying ^{13}CO or C^{18}O alone or even their combination. However this result is linked to the association of species (^{13}CO and C^{18}O) having similar collisional and radiative rate coefficients. We here consider the benefit of using only the (1 – 0) lines or both the (1 – 0) and (2 – 1) lines of ^{13}CO and C^{18}O when studying the associations with one of the isotopologues of HCO$^+$ that have a much higher dipole moment, and thus different radiative rate coefficients.

Columns 1 and 2 (resp. 4 and 5) of Fig. 14 allow us to quantify the benefit of using the (1 – 0) and (2 – 1) lines over only using the (1 – 0) line for the $\mathcal{S}\left\{\mathrm{C}^{18}\mathrm{O}, \mathrm{H}^{13}\mathrm{CO}^+\right\}$ (resp. $\mathcal{S}\left\{^{12}\mathrm{CO}, ^{13}\mathrm{CO}, \mathrm{HCO}^+\right\}$) set. In contrast with the cases using only the CO isotopologues, the loss of precision remains weak when combining a ground state transition from a CO isotopologue with the same line from an HCO$^+$ isotopologue. The region of $N(H_2)$, n_{H_2} values for which the CRB stays relatively constant only slightly decreases. This suggests that getting several low-J lines of selected species with 1) different excitation requirements and 2) a sufficiently well understood chemistry, would be a good surrogate to observations of higher J transitions of the same species, which lie in atmospheric windows that require more stringent weather conditions to get good observations.

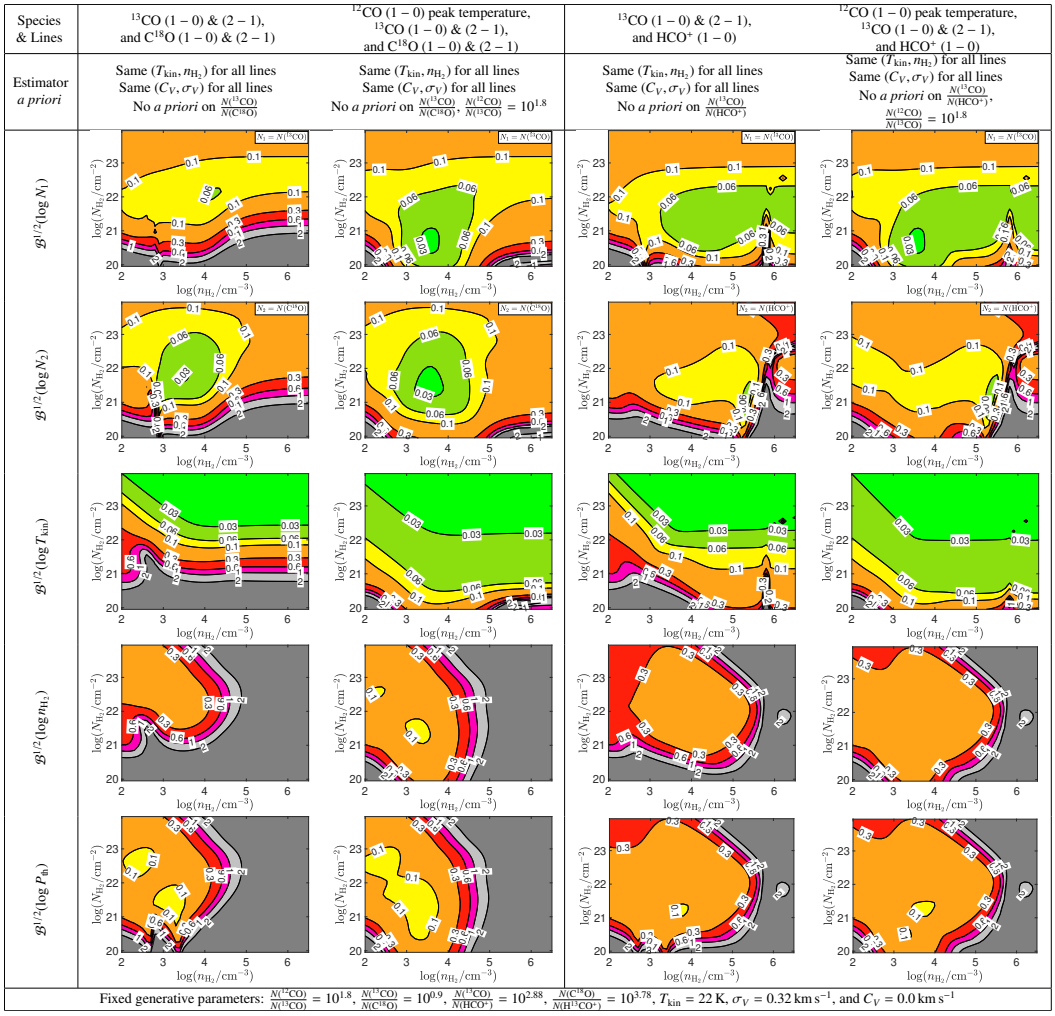


Fig. 13. Comparison of the CRB reference precisions for all the available lines of different combinations of two species, plus the constraint from ^{12}CO (1 – 0) peak temperature. The reminder of the layout is identical to Fig. 12.

5.7. Fixing the relative abundances of the CO and HCO$^+$ *isotopologues*

Up to now, we only studied cases where the relative abundance of the CO and HCO$^+$ isotopologues have been unconstrained. We finally study the impact of fixing the abundance of either ^{13}CO/HCO$^+$ or C^{18}O/H^{13}CO$^+$ on the CRB reference precision. Columns #2 and #3 (resp. #5 and #6) of Fig. 14 allow us to quantify the benefit of fixing the relative abundances at constant number of transitions for the $\mathcal{S}\left\{C^{18}O, H^{13}CO^+\right\}$ (resp. $\mathcal{S}\left\{^{12}CO, ^{13}CO, HCO^+\right\}$) set. The increase of precision is important for the column densities as expected because information on the abundances has been added and we thus have two measurements at different opacities of the same quantity (the column density). However, the increase of the precision is also obvious for kinetic temperature, volume density, and thermal pressure, in the case of the $\mathcal{S}\left\{C^{18}O, H^{13}CO^+\right\}$ set. For the $\mathcal{S}\left\{^{12}CO, ^{13}CO, HCO^+\right\}$ case, only the volume density and thermal pressure show an increased precision as the kinetic temperature is mostly constrained by the ^{12}CO (1 – 0) peak intensity. However, these gains in precision require an accurate a priori on the abundance ratio. Otherwise it induces biases as shown in Sect. 4. Leaving the relative abundance as free as possible is thus a better solution as long as their knowledges is imprecise.

6. Astrophysical consequences and perspectives

In the previous sections we have analyzed the accuracy that can be achieved on the column density and physical conditions, depending on the choice of considered spectral lines, the S/N, and domain of parameters. Figures 15 and 16 summarize the fits obtained on molecular maps of the Horsehead nebula, in the framework of a single emitting layer with uniform physical conditions and fixed or free relative abundances. Three sets of molecular lines are used: ^{13}CO and HCO$^+$, ^{13}CO and C^{18}O, and C^{18}O and H^{13}CO$^+$. The first one is adapted to more diffuse lines of sight, while the third one is more suited to denser ones. Each set of estimations has been performed with or without the additional constraint from the peak ^{12}CO intensity. The used abundance ratio values, when fixed, are the ones listed in Eq. (16).

6.1. Column density

The temperature information brought by the ^{12}CO (1 – 0) peak intensity is essential to keep the kinetic temperature in an

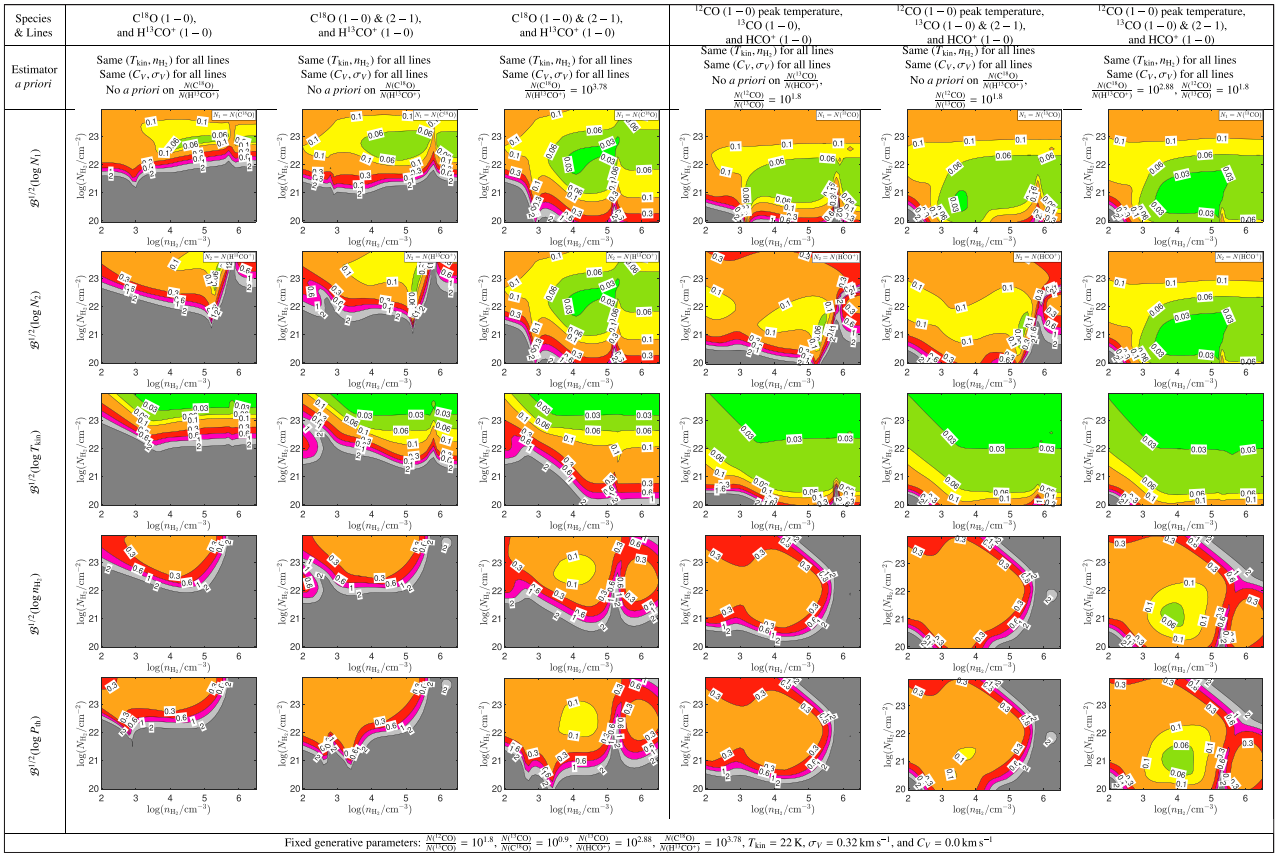


Fig. 14. Comparison of the CRB reference precisions when using either only (1 – 0) lines or adding the (2 – 1) lines for ^{13}CO and C^{18}O and influence of the a priori on the abundance ratio.

acceptable range, below 100 K. However this constraint may bias the results on column densities in the presence of a temperature gradient along the LoS. As the temperature probed by the ^{12}CO (1 – 0) line corresponds to the outer layers where the line is emitted, it may be invalid for the deeper and more shielded layers.

A comparison of the middle column of rows a and c in Fig. 15 shows that the ^{13}CO and C^{18}O column densities derived with or without using the ^{12}CO peak temperature agree well with each other. This suggests that this bias is acceptable for the CO isotopologues. The flattening of the N(C^{18}O) vs. N(^{13}CO) histogram for low values may reflect the effect of noise in the relatively faint C^{18}O lines since a positive fluctuation of the noise may lead to an apparently stronger line given the relatively low line intensity in these regions, ≤ 0.5 K km s^{-1}.

The situation is more complex for the high dipole moment species HCO$^+$ and H^{13}CO$^+$. For both species, the scatter is quite large reaching about 0.5 dex or a factor of three. The flattening of the relation at low column densities in the case of N(HCO$^+$) vs. N(C^{18}O), and the presence of high HCO$^+$ column density pixels for moderate values of N(^{13}CO) in the case of N(HCO$^+$) vs. N(^{13}CO), probably indicate a bias induced by our hypotheses in addition to the effect of noise. For instance, in regions with relatively faint ^{13}CO emission, the hypothesis of optically thick ^{12}CO (1 – 0) emission may not be fully valid leading to a bias in the assumed kinetic temperature. When the ^{13}CO (1 – 0) line is not fully optically thick, its peak temperature is expected to be lower than the kinetic temperature, which implies that the kinetic temperature could be biased towards somewhat lower values.

An independent result is that the H^{13}CO$^+$ column densities derived without fixing the abundance ratio are somewhat higher (by a factor of about two) than when fixing it. The scatter thus probably indicates real variation of the relative abundances of these two species.

6.2. Kinetic temperature

The effect of the assumptions used to derive the kinetic temperature is clearly seen in rows b and c of Fig. 16. Without the information from ^{12}CO(1 – 0), the fitted physical conditions are tightly correlated. This reflects the shape of the log-likelihood function shown in Fig. 6 to 8. Indeed the strong degeneracy between a low density/high temperature and high density/low temperature models implies that the derived parameters explore the valley of low negative log-likelihood in the ($n_{\rm H_2}$,$T_{\rm kin}$) space, depending on the line intensities, line profiles, and noise level of each pixel.

The situation is slightly better when combining the (1 – 0) and (2 – 1) lines of both ^{13}CO and C^{18}O because the lower opacity of the C^{18}O lines as compared to ^{13}CO implies a higher sensitivity of the level population to the H$_2$ density up to the C^{18}O (2 – 1) critical density of $\sim 10^4$ cm^{-3}.

The combination of C^{18}O and H^{13}CO$^+$ data also leads to an exploration of the shape of the log-likelihood function

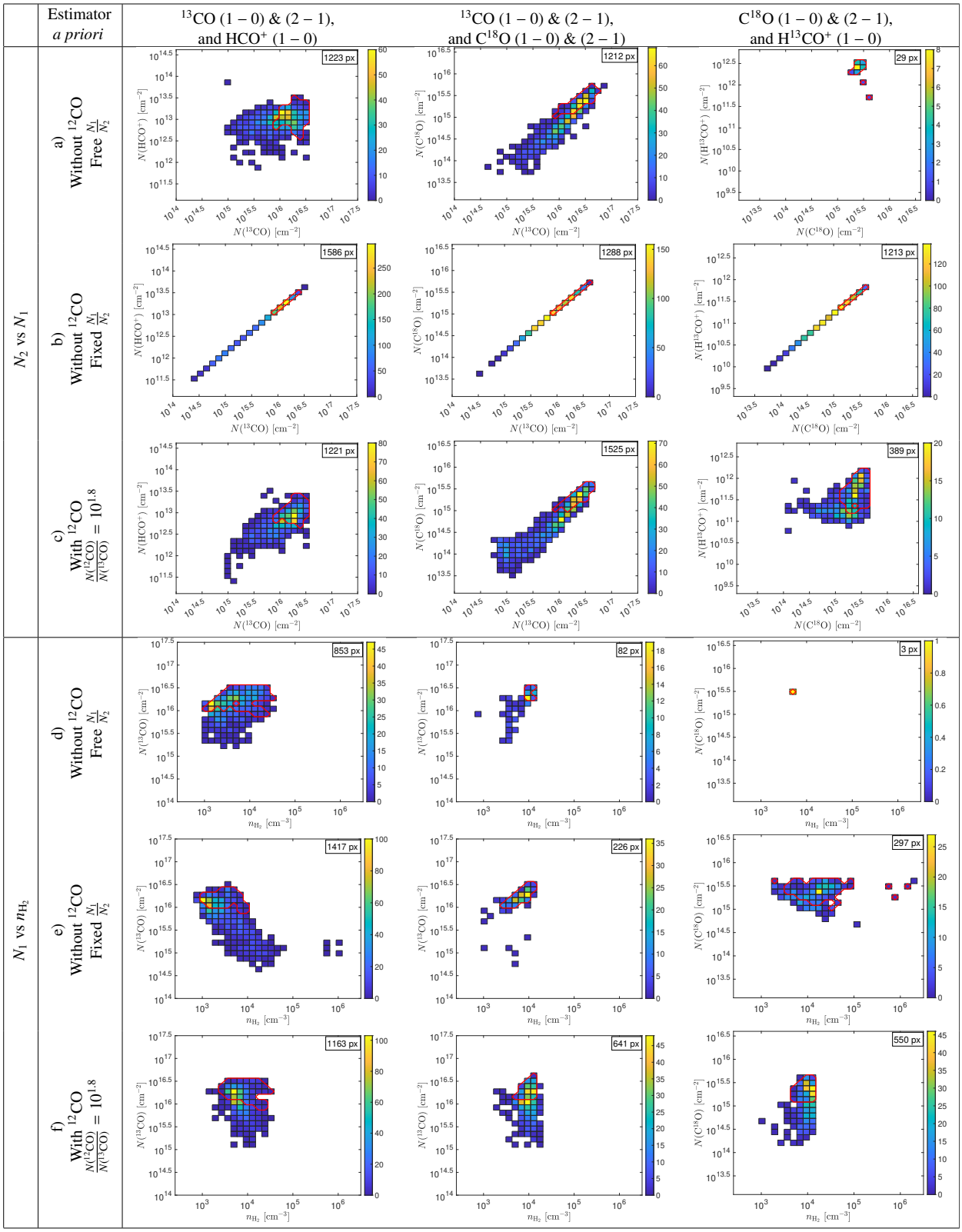


Fig. 15. Comparison of the column density and physical conditions obtained towards the Horsehead nebula using different sets of molecular lines. The red contour delineates the pixels associated with dense cores and where $H^{13}CO^+$ (1 – 0) is detected at the 3σ level. When abundances or abundance ratios are fixed, the values are given in Eq. (16). Only estimations with a good reference precision (CRB) have been kept (see Eq. (21)).

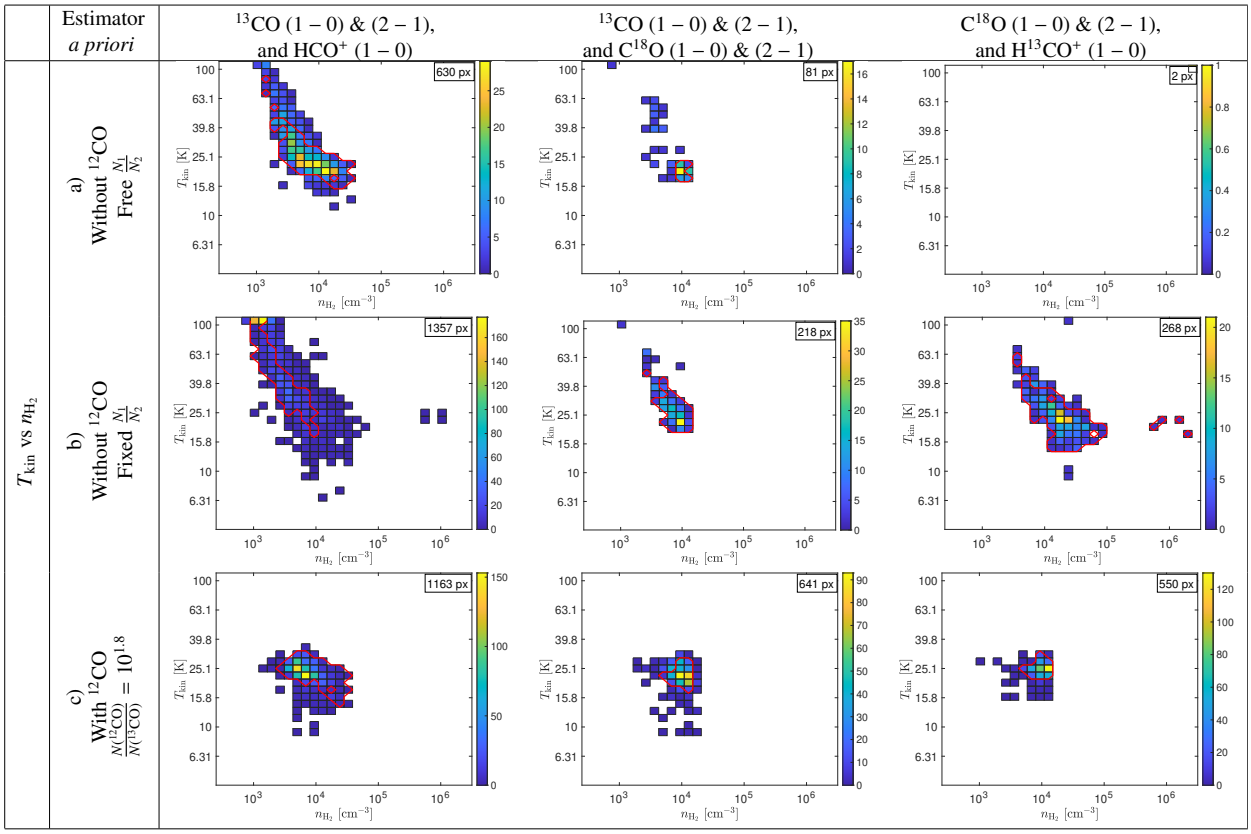


Fig. 16. Continuation of Fig. 15 showing the ($n_{\rm H_2}$,$T_{\rm kin}$) estimations.

by the different pixels. The results towards dense core pixels reach 10^6 cm^{-3}, a somewhat unrealistically high value at the spatial scale probed by the studied observations, 0.05–0.1 pc. Since these pixels have high C^{18}O column densities and a good CRB, this incorrect result suggests a model mismatch, or an impossibility to separate the different minima of the negative log-likelihood function.

When the kinetic temperature is constrained by the ^{12}CO peak temperature, the range of derived temperatures and densities becomes smaller and more consistent between the different ensembles of molecular lines. However, the small scatter of the kinetic temperature map may be directly related to the hypotheses (single layer, use of ^{12}CO) and may not reflect the real physical conditions.

6.3. Volume density

The bias and variance analyses presented in Sects. 4 and 5 have shown that it is difficult to obtain a good precision on molecular gas density because 1) of degeneracies between the volume density, the kinetic temperature, and the column densities, or 2) of the potential presence of more than one local minimum in the log-likelihood function. These degeneracies are clearly seen in the plots displaying the variation of the derived column densities as a function of the density when ^{12}CO (1 – 0) is not used (see Figs. 15d and e). The anti-correlation between the density and column density when using the ^{13}CO and HCO$^+$ lines seems nonphysical, as high column density regions are usually associated with high density regions. The lower density pixels appear to have a high kinetic temperature, a consequence of the degeneracy between the density and kinetic temperature. The same behavior is seen for the two other sets of lines, with densities shifted towards somewhat higher values as these sets of lines include transitions with higher effective critical densities than the first set of lines.

Adding a constraint on the kinetic temperature from ^{12}CO breaks the degeneracy. The derived densities occupy a rather narrow range of values, between $\sim 3\times10^3$ cm^{-3} and $\sim 3\times10^4$ cm^{-3} for all sets of lines. There is no difference in the volume density between the pixels associated with the dense cores (defined as regions with strong H^{13}CO$^+$ emission) and pixels elsewhere in the map, while it would have been expected that the density increases in the dense cores.

When looking at the volume density maps displayed in Fig. C.4 more differences appear. Using the ^{12}CO peak temperature is essential for the first set of lines as the kinetic temperatures (and densities) become too high (resp. low) without this constraint. The combination of ^{13}CO and C^{18}O lines performs quite well without ^{12}CO in the high column density region of the horsehead neck, but many pixels remain with nonphysical estimations. For these two sets of lines, using ^{12}CO without adding a constraint on the relative abundances appears to be the best choice. Because the H^{13}CO$^+$(1 – 0) emission is only detected in restricted areas, the different hypotheses have a strong impact on the derived parameters as some hypotheses allow for the use of pixels with marginal detections while others do not. Fixing the relative abundances between C^{18}O and H^{13}CO$^+$ allows us to obtain reasonable values for the density and kinetic temperature in the H^{13}CO$^+$ cores with a good CRB. This avoids the risk of

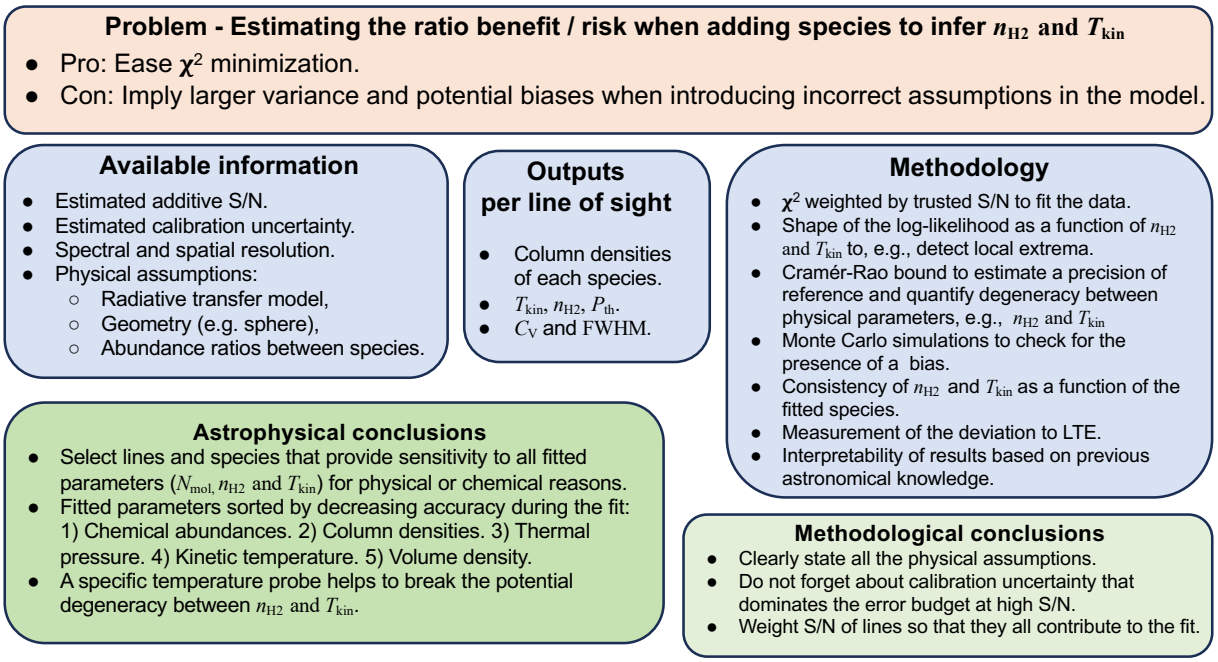


Fig. 17. Methodological and astrophysical summary.

biasing the temperature by using a temperature probe associated with the cloud envelope. However, the analysis is limited to the small set of pixels with a good S/N detection of $H^{13}CO^+(1-0)$.

6.4. Pressure

Although the estimations of density and temperature may be biased, the estimation of the thermal pressure is expected to be more accurate as the minimum of the negative log-likelihood function is nearly coincident with the locus of constant thermal pressure. Nevertheless, a detailed analysis of the accuracy of the pressure determination remains necessary, as the different degeneracies described above still play a role. In addition, the determination of the thermal pressure results from a fit of the molecular line emission under some hypotheses (e.g., uniform medium, co-localized molecular species, etc). Therefore, it is recommended to clearly describe the assumptions leading to the estimation of the parameters of interest as different sets of hypotheses may lead to different values. Overall, the CO isotopologue lines can be used to obtain rather good estimations of the thermal pressure, an interesting quantity for the energy budget of molecular gas when all hypotheses are clearly stated as recommended above.

6.5. Generalization

From the analysis of the Horsehead nebula, we can draw some recommendations for the analysis of the selected molecular emission lines: low J lines of the CO and HCO^+ isotopologues using radiative transfer models of a uniform cloud. Figure 17 summarizes these recommendations. These lines are among the strongest seen in Galactic molecular clouds and nearby galaxies. It is therefore interesting to optimize their analysis and extract most of information from such observations.

In many conditions, even the excitation of the CO isotopologues departs from the local thermodynamic equilibrium (see Fig. D.1), and a non-LTE model like RADEX should be preferred for the analysis of the molecular line emission. However, it is important that the hypotheses involved in the non-LTE calculations are clearly spelled out.

The determination of column densities is accurate for ^{13}CO and $C^{18}O$. It does not need additional hypotheses than the presence of all species in the same volume. Column densities for the higher dipole moment species HCO^+ and $H^{13}CO^+$ are more sensitive to the set of hypotheses as these species are sub-thermally excited in most of the explored parameter space, hence their emission is sensitive to the physical conditions. The lower S/N of the faint emission lines also implies a lower accuracy of the determined parameters.

The determination of the molecular hydrogen volume density and kinetic temperature is more difficult to achieve with a good precision. As discussed in Sect. 4.4.3, because of the strong degeneracy between the density and kinetic temperature in determining the CO isotopologue excitation, an independent constraint on the kinetic temperature is needed. Using the peak value of the ^{12}CO $(1-0)$ line works well for that purposes and allows us to at least partially break the degeneracy. This constraint is useful for the determination of the kinetic temperature but ambiguities may remain for the molecular hydrogen density. Given the achievable S/N and calibration accuracy, different physical conditions may lead to very similar emerging spectra as illustrated in Fig. 10 and discussed in Sect. 5.2. Adding a line from a high dipole moment species is useful for constraining the volume density, provided the line is observed with a good S/N and this species abundance relative to that of one CO isotopologue can be used to constrain the fit. Otherwise, adding another line is less beneficial because of the degeneracy between density and column density. An error in fixing the relative abundance of the high dipole moment species relative to the CO isotopologue leads to a bias in the derivation of the density but this bias remains tolerable when compared to the uncertainties when only one CO isotopologue is used.

Using different molecular lines from different species brings more constraints for deriving the physical conditions and column densities, at the expense of the need for a strong hypothesis: the presence of all species within the same volume of gas with uniform conditions. Such a strong hypothesis may be far from the reality of interstellar clouds which exhibit gradients of density, temperature and molecular abundances across their volumes. In particular, towards lines of sight where dense cores are present, the range of physical conditions is broad and the uniform cloud hypothesis is expected to lead to poor results. One example of such poor results is the determination of the kinetic temperature using the ^{12}CO $(1-0)$ line which does not differentiate pixels within dense cores from pixels outside cores. In fact, it is known from more detailed analyses that the kinetic temperature decreases from the extended molecular cloud envelopes to cores (e.g., Hocuk et al. 2017; Rodríguez-Baras et al. 2021). Instead of uniform models, more sophisticated geometrical layouts are needed for specific environments. For instance, dense cores that are embedded in filaments, and in the molecular gas envelope clearly need a multi-layer model, in which a cold and dense region is surrounded by a warmer and more diffuse envelope. This will be the subject of another paper. Another case which would deserve a multi-layer model could be a dense, non edge-on, UV illuminated photodissociation region (PDR) where a strong temperature gradient is present due to the attenuation of the UV radiation when entering the molecular gas. Such PDRs often have a nearly isobaric equation of state (e.g., Hernández-Vera et al. 2023, for the Horsehead nebula), hence an isobaric model would be better suited than a uniform model with a single density and temperature value.

For all cases, having a temperature probe independent of the lines used for the column density and volume density estimations has been shown to be useful for breaking the degeneracies in the excitation conditions. The peak temperature of the ^{12}CO $(1-0)$ line is adequate for the bulk of the molecular emission, but overestimates the temperature of the dense and shielded regions. In these regions, the ratio of the $(1,1)$ and $(2,2)$ inversion lines of NH_3 near 23 GHz is known to perform well as a temperature probe (Maret et al. 2009), but getting such data requires combining observations from different telescopes. Efforts should be made to find another, easily accessible, temperature probe based on molecular lines, that could be used in the analysis of well shielded regions. Hacar et al. (2020) proposed to use the ratio of the HCN and HNC intensities to probe the kinetic temperature between about 15 and 60 K. However in their extensive analysis of the Orion B cloud, Santa-Maria et al. (2023) showed that the HCN/HNC ratio is more sensitive to the radiation field than to the kinetic temperature. Another possibility to probe the kinetic temperature of shielded regions could be using the dust temperature, but this also requires observations over a wide range of wavelengths to break the degeneracy between the dust temperature and column density. Rodríguez-Baras et al. (2021) have used radiative transfer calculations and showed that the difference between the kinetic temperature and the dust temperature remains lower than 1 K along lines of sight with an extinction larger than 8 magnitudes (4 mag on each side from the cloud edge) for the conditions corresponding to the sources probed in the GEMS IRAM large program, dense cores with an external FUV radiation field between 5 and 60 times the Inter-Stellar Radiation Field and a thermal pressure of 5×10^5 K cm^{-3}). Further studies over a wider parameter space and including the comparison of the predicted dust temperature and its estimation from the fit of the spectral energy distribution are needed to establish the conditions where the dust temperature is a good proxy of the kinetic temperature, and thus the bias which could be introduced by this hypothesis.

7. Conclusion

In this article, we study the fit of multi-species molecular lines with a non-LTE radiative transfer model and we propose a method to test the derived results, including the possibility of identifying potential model misspecifications. Figure 17 summarizes our main recommendations and their implications for further studies. We list a few general points below.

- While LTE models are simple to run and interpret, the excitation of CO isotopologues does not always reach this limit and non-LTE models should be preferred, especially when the analysis of CO isotopologue lines is combined with lines from higher dipole moment species.
- It is important to take into account the S/N of the transition studied, including the thermal noise and calibration uncertainties, when fitting radiative transfer models to observed molecular emission. This can be done for instance by limiting the S/N ratio of the strongest lines to mimic the effect of calibration uncertainties.
- When fitting radiative transfer models to observations, it is important to consider the information from the full line profile and not just from the integrated intensity, because this allows us to consider the broadening of the line profile when the opacity increases.
- Because different physical conditions may produce very similar line profiles for some transitions, it is important to account for such potential degeneracies when exploring the parameter space.
- Combining molecular species with different excitation requirements, for instance, high dipole (high critical density) and low dipole moment species (low critical density), allows for the degeneracy in the physical condition parameter space to be broken, provided the relative abundances of the considered species are known from independent observations or from models.
- The escape probability model is adequate for studying large scale molecular line maps. For detailed studies of specific objects of known geometry, more sophisticated models may be required to better take into account the radiative coupling along the line of sight.

Based on the precision analysis, we make the following additional recommendations.

- To increase the precision of the estimations, it is more interesting to combine lines with different excitation requirements than to increase the S/N ratio of a targeted spectral line. This is because the needed integration time to detect a new line if often significantly lower than the additional time needed to improve the S/N ratio on an already detected line.
- With the availability of broad band receivers, combining lines from different species accessible with the same receiver usually allows for a more efficient use of the observing time than accessing lines from different energy levels, which often require stringent weather conditions to get good observations.
- The thermal pressure of the emitting medium is often determined with a better precision than the kinetic temperature of volume density taken separately. It can therefore be interesting to derive thermal pressure maps of molecular clouds, which can be compared with maps of the magnetic or turbulent pressure.

This work is focused on the combination of CO and HCO^+ isotopologue lines. Other high-dipole moment species could be considered in the future. For instance, CS, CN, HNC or N_2H^+ have bright lines that are easily detectable in large scale maps of molecular clouds. While this study is also focused on a single homogeneous layer, it is clear that this hypothesis is not fully suited to lines of sight towards dense cores. In a following paper, we will study a multi-layer model and compare its performances with the homogeneous model.

Acknowledgements. This work is based on observations carried out under project numbers 019-13, 022-14, 145-14, 122-15, 018-16, and finally the large program number 124-16 with the IRAM 30m telescope. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain). This work was supported by the French Agence Nationale de la Recherche through the DAOISM grant ANR-21-CE31-0010, and by the Programme National "Physique et Chimie du Milieu Interstellaire" (PCMI) of CNRS/INSU with INC/INP, co-funded by CEA and CNES. This project has received financial support from the CNRS through the MITI interdisciplinary programs. JRG and MGSM thank the Spanish MCINN for funding support under grant PID2019-106110G-100. Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). D.C.L. was supported by USRA through a grant for SOFIA Program 09-0015. This research has made use of spectroscopic and collisional data from the EMAA (https://emaa.osug.fr and https://dx.doi.org/10.17178/EMAA) and LAMDA (https://home.strw.leidenuniv.nl/~moldata/) databases. EMAA is supported by the Observatoire des Sciences de l'Univers de Grenoble (OSUG). The LAMDA database is supported by the Netherlands Organization for Scientific Research (NWO), the Netherlands Research School for Astronomy (NOVA), and the Swedish Research Council.

References

Asensio Ramos, A., & Elitzur, M. 2018, A&A, 616, A131
Barnes, A. T., Kauffmann, J., Bigiel, F., et al. 2020, MNRAS, 497, 1972
Bernes, C. 1979, A&A, 73, 67
Brinch, C., & Hogerheijde, M. R. 2010, A&A, 523, A25
Bron, E., Daudon, C., Pety, J., et al. 2018, A&A, 610, A12
Bron, E., Roueff, E., Gerin, M., et al. 2021, A&A, 645, A28
Cao, Z., Jiang, B., Zhao, H., & Sun, M. 2023, ApJ, 945, 132
Castets, A., Duvert, G., Dutrey, A., et al. 1990, A&A, 234, 469
Denis-Alpizar, O., Stoecklin, T., Dutrey, A., & Guilloteau, S. 2020, MNRAS, 497, 4276
Einig, L., Pety, J., Roueff, A., et al. 2023, A&A, 677, A158
Fuente, A., Garcia-Burillo, S., Usero, A., et al. 2008, A&A, 492, 675
Garthwaite, P. H., Jolliffe, I. T., & Jones, B. 1995, Statistical Inference (London: Prentice Hall Europe)
Gerin, M., Liszt, H., Neufeld, D., et al. 2019, A&A, 622, A26
Gerin, M., Pety, J., Roueff, A., Bron, E., & Orion-B Team. 2023, in Physics and Chemistry of Star Formation: The Dynamical ISM Across Time and Spatial Scales, 230
Goicoechea, J. R., Pety, J., Gerin, M., et al. 2006, A&A, 456, 565
Goldsmith, P. F., & Kauffmann, J. 2017, ApJ, 841, 25
Guzmán, V. V., Pety, J., Gratier, P., et al. 2014, Faraday Discuss., 168, 103
Hacar, A., Bosman, A. D., & van Dishoeck, E. F. 2020, A&A, 635, A4
Hernández-Vera, C., Guzmán, V. V., Goicoechea, J. R., et al. 2023, A&A, 677, A152
Hocuk, S., Szűcs, L., Caselli, P., et al. 2017, A&A, 604, A58
Le Petit, F., Nehmé, C., Le Bourlot, J., & Roueff, E. 2006, ApJS, 164, 506
Leung, C. M., & Liszt, H. S. 1976, ApJ, 208, 732
Liszt, H. S. 2012, A&A, 538, A27
Maret, S., Faure, A., Scifoni, E., & Wiesenfeld, L. 2009, MNRAS, 399, 425
Müller, H. S. P., Thorwirth, S., Roth, D. A., & Winnewisser, G. 2001, A&A, 370, L49
Nishimura, Y., Watanabe, Y., Harada, N., et al. 2017, ApJ, 848, 17
Pety, J. 2005, in SF2A-2005: Semaine de l'Astrophysique Francaise, eds. F. Casoli, T. Contini, J. M. Hameury, & L. Pagani, 721
Pety, J., Goicoechea, J. R., Hily-Blant, P., Gerin, M., & Teyssier, D. 2007, A&A, 464, L41
Pety, J., Guzmán, V. V., Orkisz, J. H., et al. 2017, A&A, 599, A98
Pety, J., Gerin, M., Bron, E., et al. 2022, in Eur. Phys. J. Web Conf. 265, 00048
Rodríguez-Baras, M., Fuente, A., Riviére-Marichalar, P., et al. 2021, A&A, 648, A120
Roueff, A., Gerin, M., Gratier, P., et al. 2021, A&A, 645, A26
Santa-Maria, M. G., Goicoechea, J. R., Pety, J., et al. 2023, A&A, 679, A4
Shirley, Y. L. 2015, PASP, 127, 299
Stoica, P., & Moses, R. 2005, Spectral Analysis of Signals (New Jersey: Prentice Hall)
Tafalla, M., Usero, A., & Hacar, A. 2021, A&A, 646, A97
Teng, Y.-H., Sandstrom, K. M., Sun, J., et al. 2023, ApJ, 950, 119
Tunnard, R., & Greve, T. R. 2016, ApJ, 819, 161
van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, A&A, 468, 627
Ward-Thompson, D., Nutter, D., Bontemps, S., Whitworth, A., & Attwood, R. 2006, MNRAS, 369, 1201
Watanabe, Y., Nishimura, Y., Harada, N., et al. 2017, ApJ, 845, 116
Yang, B., Stancil, P. C., Balakrishnan, N., & Forrey, R. C. 2010, ApJ, 718, 1062
Zucker, C., Goodman, A., Alves, J., et al. 2021, ApJ, 919, 35

[1] Université de Toulon, Aix Marseille Univ, CNRS, IM2NP, Toulon, France
e-mail: antoine.roueff@univ-tln.fr
[2] IRAM, 300 rue de la Piscine, 38406 Saint-Martin-d'Hères, France
[3] LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, 75014 Paris, France
[4] Instituto de Física Fundamental (CSIC). Calle Serrano 121, 28006, Madrid, Spain
[5] National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
[6] Laboratoire d'Astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, Allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[7] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, GIPSA-Lab, Grenoble 38000, France
[8] Chalmers University of Technology, Department of Space, Earth and Environment, 412 93 Gothenburg, Sweden
[9] Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRIStAL, 59651 Villeneuve d'Ascq, France
[10] LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, 92190 Meudon, France
[11] Institut de Recherche en Astrophysique et Planétologie (IRAP), Université Paul Sabatier, Toulouse cedex 4, France
[12] Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile
[13] Laboratoire de Physique de l'Ecole normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, Sorbonne Paris Cité, Paris, France
[14] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[15] Department of Earth, Environment, and Physics, Worcester State University, Worcester, MA 01602, USA
[16] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA, 02138, USA
[17] School of Physics and Astronomy, Cardiff University, Queen's buildings, Cardiff CF24 3AA, UK

Appendix A: Fisher information matrix

To compute Cramér-Rao bound associated to the unknown parameters $\boldsymbol{\theta}$, we need to calculate the Fisher matrix of the L observed lines $\boldsymbol{x}_{1:L} = \{\boldsymbol{x}_1 ..., \boldsymbol{x}_L\}$. Nevertheless, since the noise on the different lines are statistically independent, the Fisher matrix of $\boldsymbol{x}_{1:L}$ is simply the sum of the Fisher matrix of each line

$$\boldsymbol{I}_F(\boldsymbol{\theta}; \boldsymbol{x}_{1:L}) = \sum_{l=1}^{L} \boldsymbol{I}_F(\boldsymbol{\theta}, \boldsymbol{x}_l). \tag{A.1}$$

Below we detail the calculation of the Fisher matrix for a single line and the vector of unknown parameter is simply $\boldsymbol{\theta} = [\log T_{\rm kin}, \log n_{\rm H_2}, \log N, \sigma_V, C_V]$. The generalization of these computations to the case where we also fit the logarithm of some of the column densities; for instance, $\boldsymbol{\theta} = [\log T_{\rm kin}, \log n_{\rm H_2}, \log N(^{13}{\rm CO}), \log N({\rm C}^{18}{\rm O}), \sigma_V, C_V]$, is straightforward. For didactic reasons, the presentation is organized in two parts. In Sect. A.1, the multiplicative noise is omitted. It is taken into account in Sect. A.2.

Appendix A.1. Case without multiplicative noise

In this section, we assume only additive noise:

$$x_n = s_n + b_n. \tag{A.2}$$

where b_n is a white Gaussian noise with variance $\sigma^2_{b,l}$ and s_n is a sampled version of Eq. (7):

$$s_n = \{J(T_{{\rm ex},l}, \nu_l) - J(T_{\rm CMB}, \nu_l)\}\,[1 - \exp(-\Psi_n)]\,. \tag{A.3}$$

Thus, Ψ_n is a sampled version of Eq. (4):

$$\Psi_n = \tau_l \exp\left(-\frac{(V_n - C_V)^2}{2\sigma_V^2}\right), \tag{A.4}$$

where V_n is the velocity at channel n. Assuming that the noise is white, the log-likelihood of the sampled observation at line l is

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x}_l) = cte - \sum_n \frac{(x_n - s_n)^2}{2\sigma^2_{b,l}}, \tag{A.5}$$

and the Fisher matrix $\boldsymbol{I}_F$ (see Stoica & Moses 2005) is

$$\forall (i,\, j) \quad \boldsymbol{I}_F(\boldsymbol{\theta}, \boldsymbol{x}_l) = \frac{1}{\sigma^2_{b,l}} \sum_n \frac{\partial s_n}{\partial \theta_i} \frac{\partial s_n}{\partial \theta_j}. \tag{A.6}$$

Thus, to get $\boldsymbol{I}_F(\boldsymbol{\theta}, \boldsymbol{x}_l)$, we simply need to evaluate the gradient $\frac{\partial s_n}{\partial \theta_i}$, where θ_i is the ith component of $\boldsymbol{\theta}$. This is done in the following.

Appendix A.1.1. Calculation of $\frac{\partial s_n}{\partial \log T_{\rm kin}}$

We note that

$$\frac{\partial s_n}{\partial \log T_{\rm kin}} = \frac{\partial s_n}{\partial T_{\rm kin}} \ln(10) T_{\rm kin}. \tag{A.7}$$

Moreover, based on Eq. (A.3), we have

$$\begin{aligned}\frac{\partial s_n}{\partial T_{\rm kin}} = \; & \frac{\partial J(T_{{\rm ex},l}, \nu_l)}{\partial T_{{\rm ex},l}} \frac{\partial T_{{\rm ex},l}}{\partial T_{\rm kin}} [1 - \exp(-\Psi_n)] \\ & + \{J(T_{{\rm ex},l}, \nu_l) - J(T_{\rm CMB}, \nu_l)\} \frac{\partial \Psi_n}{\partial T_{\rm kin}} \exp(-\Psi_n).\end{aligned} \tag{A.8}$$

In practice, we numerically estimated $\frac{\partial T_{{\rm ex},l}}{\partial T_{\rm kin}}$ with a finite difference technique using RADEX with a step of $T_{\rm kin}/1024$.

Appendix A.1.2. Calculation of $\frac{\partial J(T,\nu)}{\partial T}$

$$J(T, \nu) = \frac{h\nu}{k} \frac{1}{\exp \frac{h\nu}{kT} - 1}. \tag{A.9}$$

Thus

$$\frac{\partial J(T,\nu)}{\partial T} = \frac{h\nu}{k} \frac{\frac{h\nu}{kT^2} \exp \frac{h\nu}{kT}}{\left(\exp \frac{h\nu}{kT} - 1\right)^2} = \frac{h^2\nu^2}{k^2T^2} \frac{\exp \frac{h\nu}{kT}}{\left(\exp \frac{h\nu}{kT} - 1\right)^2}, \tag{A.10}$$

and finally

$$\frac{\partial J(T,\nu)}{\partial T} = \frac{h^2\nu^2}{k^2T^2} \frac{1}{\exp \frac{h\nu}{kT} - 2 + \exp -\frac{h\nu}{kT}}. \tag{A.11}$$

Appendix A.1.3. Calculation of $\frac{\partial \Psi_n}{\partial T_{\rm kin}}$

Based on Eq. (A.4),

$$\frac{\partial \Psi_n}{\partial T_{\rm kin}} = \frac{\partial \tau_l}{\partial T_{\rm kin}} \exp\left(-\frac{(V_n - C_V)^2}{2\sigma_V^2}\right), \tag{A.12}$$

where the term $\frac{\partial \tau_l}{\partial T_{\rm kin}}$ is also numerically estimated with finite difference as in Eq. (A.8).

Appendix A.1.4. Calculation of $\frac{x_n}{\partial \log n_{\rm H2}}$ and $\frac{x_n}{\partial \log N}$

The calculation of $\frac{\partial s_n}{\partial \log n_{\rm H_2}}$ and $\frac{\partial s_n}{\partial \log N}$ are straightforward adaptations from $\frac{x_n}{\partial \log N}$. We only need to specify that the finite difference technique is applied for both $\log n_{\rm H_2}$ and $\log N$ with a step of 0.001.

Appendix A.1.5. Calculation of $\frac{\partial s_n}{\partial \sigma_V}$

To get

$$\begin{aligned}\frac{\partial s_n}{\partial \sigma_V} = \; & \frac{\partial J(T_{{\rm ex},l}, \nu_l)}{\partial T_{{\rm ex},l}} \frac{\partial T_{{\rm ex},l}}{\partial \sigma_V} [1 - \exp(-\Psi_n)] \\ & + \{J(T_{{\rm ex},l}, \nu_l) - J(T_{\rm CMB}, \nu_l)\} \frac{\partial \Psi_n}{\partial \sigma_V} \exp(-\Psi_n),\end{aligned} \tag{A.13}$$

we need to detail the calculation of $\frac{\partial \Psi_n}{\partial \sigma_V}$.

Appendix A.1.6. $\frac{\partial \Psi_n}{\partial \sigma_V}$

Based on Eq. (A.4), one has

$$\begin{aligned}\frac{\partial \Psi_n}{\partial \sigma_V} = \; & \frac{\partial \tau_l}{\partial \sigma_V} \exp\left(-\frac{(V_n - C_V)^2}{2\sigma_V^2}\right) \\ & + \tau_l \frac{(V_n - C_V)^2}{\sigma_V^3} \exp\left(-\frac{(V_n - C_V)^2}{2\sigma_V^2}\right),\end{aligned} \tag{A.14}$$

where the term $\frac{\partial \tau_l}{\partial \sigma_V}$ is numerically estimated with a finite difference technique with a step of FWHM/128.

Appendix A.1.7. Gradient calculation along Δ_V

$$\frac{\partial s_n}{\partial \Delta_V} = (J(T_{{\rm ex},l}, \nu_l) - J(T_{\rm CMB}, \nu_l)) \frac{\partial \Psi_n}{\partial \Delta_V} \exp(-\Psi_n). \tag{A.15}$$

Appendix A.1.8. $\partial \Psi_n / \partial \Delta_V$

Based on Eq. (A.4), one has

$$\frac{\partial \Psi_n}{\partial \Delta_V} = \tau_l \frac{V_n - C_V}{\sigma_V^2} \exp\left(-\frac{(V_n - C_V)^2}{2\sigma_V^2}\right). \tag{A.16}$$

Appendix A.2. Case with multiplicative noise

Let us now take into account the multiplicative noise

$$\forall n \quad x_n = c.s_n + b_n, \tag{A.17}$$

where in addition of b_n, one has $c \sim \mathcal{N}(1, \sigma^2_{c,l})$, which is assumed independent b_n. The column vector $\boldsymbol{x} = [x_1,\ x_2,\ ...]^T$ is the sum of two Gaussian random vectors, thus it remains a Gaussian vector. Its mean is $\boldsymbol{s} = [s_1,\ s_2,\ ...]^T$ and it is straightforward to show that is covariance matrix is

$$\boldsymbol{\Sigma} = \sigma^2_{b,l}\, \boldsymbol{I}_d + \sigma^2_{c,l}\, \boldsymbol{s}\boldsymbol{s}^T, \tag{A.18}$$

where $\boldsymbol{I}_d$ is the identity matrix. The log-likelihood of $\boldsymbol{x}$ is

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x}_l) = cte - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{x}-\boldsymbol{s})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{s}), \tag{A.19}$$

which leads, after tedious but straightforward calculations, to

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{x}_l) = cte - \frac{1}{2}\left(\log\sigma^2_l + \frac{\boldsymbol{x}^T\boldsymbol{x}}{\sigma^2_{b,l}} + \frac{\boldsymbol{s}^T\boldsymbol{s}}{\sigma^2_l} - \frac{(\boldsymbol{x}^T\boldsymbol{s})^2\sigma^2_{c,l}}{\sigma^2_b\sigma^2_l} - \frac{2\boldsymbol{x}^T\boldsymbol{s}}{\sigma^2_l}\right), \tag{A.20}$$

where $\sigma^2_l = \sigma^2_{b,l} + \sigma^2_{c,l}(\boldsymbol{s}^T\boldsymbol{s})$. The Slepian-Bang formula Stoica & Moses (2005) gives us the following relations,

$$[\boldsymbol{I}_F]_{ij} = \frac{\partial \boldsymbol{s}^T}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{s}}{\partial\theta_j} + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right), \tag{A.21}$$

which leads, after tedious but straightforward calculations to

$$\begin{aligned}[\boldsymbol{I}_F]_{ij} &= \left(\frac{\sigma^2_l+\sigma^4_{c,l}(\boldsymbol{s}^T\boldsymbol{s})}{\sigma^2_{b,l}\sigma^2_l}\right)\left(\frac{\partial\boldsymbol{s}^T}{\partial\theta_i}\frac{\partial\boldsymbol{s}}{\partial\theta_j}\right)\\ &+ \left(\frac{\sigma^2_{b,l}\sigma^4_{c,l}-\sigma^6_{c,l}(\boldsymbol{s}^T\boldsymbol{s})}{\sigma^2_{b,l}\sigma^4_l} - \frac{\sigma^2_{c,l}}{\sigma^2_{b,l}\sigma^2_l}\right)\left(\boldsymbol{s}^T\frac{\partial\boldsymbol{s}}{\partial\theta_i}\right)\left(\boldsymbol{s}^T\frac{\partial\boldsymbol{s}}{\partial\theta_j}\right).\end{aligned} \tag{A.22}$$

Appendix B: Details on the used maximum likelihood estimator

The following subsections provide details on the implementation of the maximum likelihood estimator we use in this article. We first explain how we localize and detect the Gaussian opacity profile (Eq. (4)) on high S/N and optically thin lines. This allows us to provide an initial estimation of the centroid velocity. We then discuss 1) the initialization of the other parameters and 2) the iterative algorithm to refine their estimation.

Appendix B.1. Signal detection and initial estimation of the centroid velocity

We start by computing the noise standard deviation on baseline channels by using the range of channels that are devoid of signal on the studied field of view. These channels can be selected manually from raw data (see the right part of Fig. 1). We then estimate the centroid velocity C_V. To do this, we first increase the S/N by adding all the lines from $C^{18}O$ and ^{13}CO after resampling them on the same velocity grid. Although the ^{12}CO (1 − 0) line is bright, its high opacity makes it mostly sensitive to the kinematics of surface of the cloud instead of the bulk of the flow. On the opposite, the S/N on HCO^+ and $H^{13}CO^+$ is significantly lower, which make them useless to estimate centroid velocity. We then apply a matched filter by convolving the summed spectra with a Gaussian shape of FWHM = 0.35 km/s. This shape corresponds to an optically thin approximation of a line. Selecting the velocity where the intensity of the filtered signal is maximum thus delivers an initial estimation of the centroid velocity C_V for all the lines.

Afterwards, we additionally check the presence of the velocity component in a majority of the studied lines. More precisely, the pixel is further processed only when more than half of the lines are detected. Else, the estimated parameters are nullified. To do this, we integrate the profile within a velocity interval of 1 km s^{-1} centered on the estimated centroid velocity, C_V. The line is considered detected when the signal is larger than $3\sqrt{N}\,\sigma_b$, where N is the number of velocity channels in 1 km s^{-1}, and σ_b the associated noise standard deviation.

Table B.1. Minimum and maximum parameter bounds, where estimations are saturated to prevent RADEX from crashing

$T_{\rm kin}$ [K]	$\log(n_{\rm H_2}/\,{\rm cm}^3)$	$\log(N(^{13}{\rm CO})/\,{\rm cm}^2)$	σ_V [km s^{-1}]
[4, 113]	[2, 6.5]	[13, 18.6]	[0.25, 2]

Appendix B.2. Initializing the fit estimation

The signal detection step provides an initial estimate of the centroid velocity for all the lines. In order to initialize the other parameters, we compute the negative log-likelihood (NLL) on a fixed grid of models computed with RADEX. This grid sample the parameter space in logarithmic steps for all parameters whose range of plausible values cover at least one order of magnitude. In practice, only the velocity dispersion is sampled linearly. In details,

- The velocity dispersion, FWHM, is sampled linearly between 0.25 and 2 km s^{-1} in steps of 0.05 km s^{-1}.
- The kinetic temperature, $T_{\rm kin}$, is sampled between 4 and 113 K in logarithmic steps of 0.05.
- The H_2 volume density, $n(H_2)$ is sampled between 10^2 and $10^{6.5}$ cm^{-3} in logarithmic steps of 0.1.
- $N(^{13}CO)$ is sampled between 10^{13} and 10^{18} cm^{-2} in logarithmic steps of 0.05.
- The column densities of ^{12}CO, $C^{18}O$, HCO^+, and $H^{13}CO^+$ are sampled with the same grid as ^{13}CO, shifted to take into account the relative abundances defined in Eq. (16).

Using this grid, for each species, we thus get a 4D tensor of negative log-likelihood (NLL) values. The way we explore this tensor, depends on the a priori information that are available. For instance, let us consider the $\mathcal{S}\left\{^{13}{\rm CO}, {\rm HCO}^+\right\}$ estimation problem (see Table 4). We need to estimate the column density of these two species independently with the constraint that the triplet ($T_{\rm kin}$, $n_{\rm H_2}$, σ_V) is identical for all species. On the other hand, with $\mathcal{S}\left\{^{13}{\rm CO}, {\rm HCO}^+; \mathcal{A}\right\}$, the situation is slightly different since we need to impose the abundance ratio between ^{13}CO and HCO^+. Adapting the implemented estimator to every possible a priori knowledge is thus required. At the end of this process, we get an initial estimation noted $\boldsymbol{\theta}^{(0)}$.

Appendix B.3. Iterative gradient descent

To refine our estimation of the negative log-likelihood, we use the following iterative estimation in which the Hessian is the Fisher matrix starting from $\boldsymbol{\theta}^{(0)}$

$$\widehat{\boldsymbol{\theta}}^{(i+1)} = \widehat{\boldsymbol{\theta}}^{(i)} - \boldsymbol{I}_F^{-1}(\widehat{\boldsymbol{\theta}}^{(i)})\nabla_{\boldsymbol{\theta}}(\widehat{\boldsymbol{\theta}}^{(i)}), \tag{B.1}$$

where i is the ith iteration, $\boldsymbol{I}_F(\boldsymbol{\theta})$ is the Fisher matrix as a function of $\boldsymbol{\theta}$, and $\nabla_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is the gradient of the negative log-likelihood

$$\forall j \quad [\nabla_{\boldsymbol{\theta}}(\boldsymbol{\theta})]_j = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \sum_{n=1}^{K} \frac{\partial s_{n,l}}{\partial \theta_j} (s_{n,l} - x_{n,l}), \tag{B.2}$$

where j is the index of the estimated parameter. The inversion of Fisher matrix in Eq. (B.1) may be numerically unstable. When the ratio between the highest and the smallest eigenvalue is higher than 10^9, we use the Moore Penrose pseudo-inverse instead of the inverse in Eq. (B.1).

Furthermore, when Eq. (B.1) would lead to parameters out of the initial grid, we saturate these parameters to prevent RADEX from crashing (see Table B.1). For example, we impose $n_{\mathrm{H_2}}$ to be smaller than $10^{6.5}$ cm^{-3}. While a volume density at $10^{6.5}$ cm^{-3} cannot be trusted in this case, the column density may still be reliable. These conditions add some a priori knowledge that is not available in the computation of the Cramér-Rao bound.

Appendix C: Supplemental figures on data fit

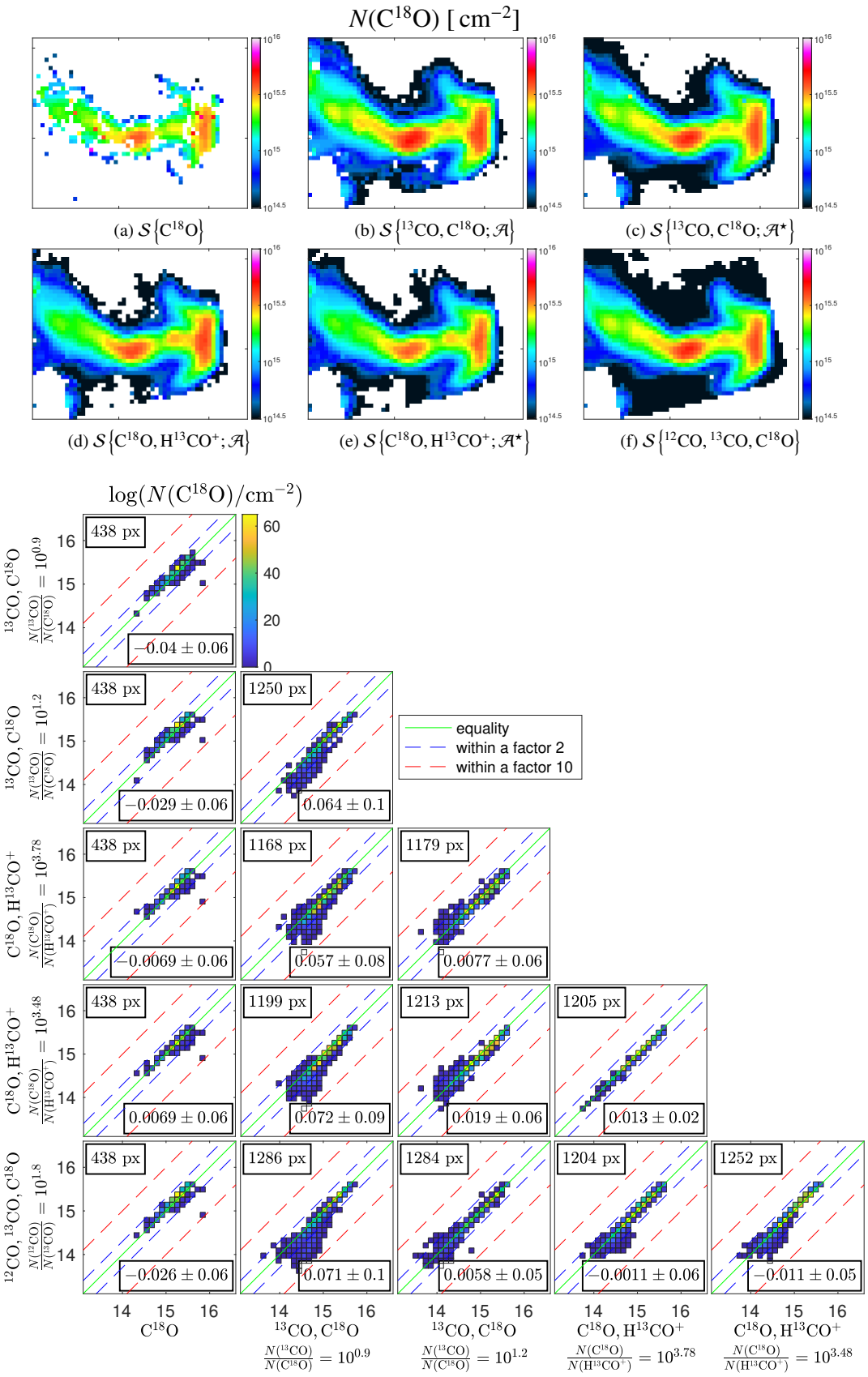


Fig. C.1. Same as Fig. 2 but for $N(C^{18}O)$.

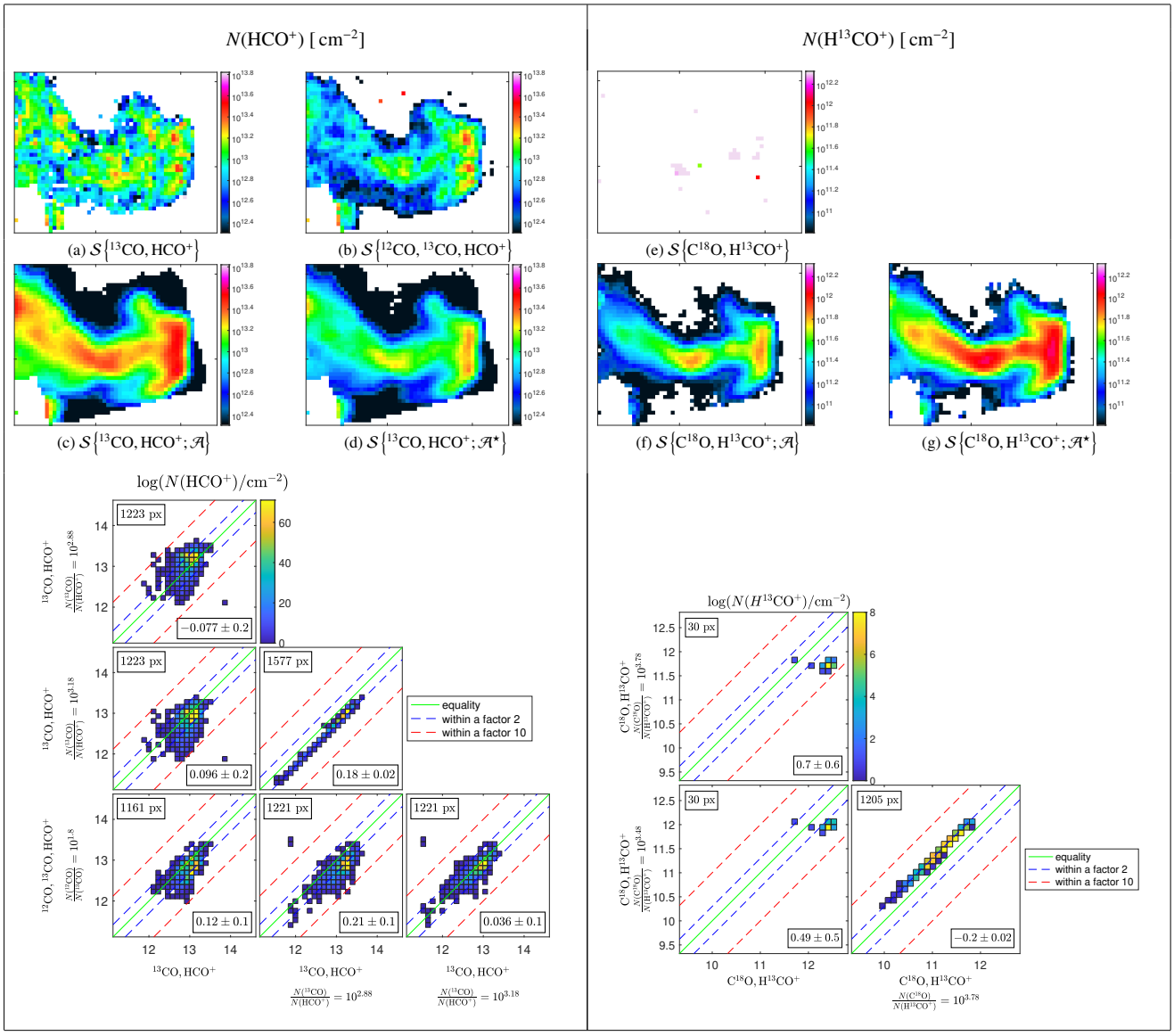


Fig. C.2. Same as Fig. 2 but for $N(\mathrm{HCO^+})$ **(left)** and $N(\mathrm{H^{13}CO^+})$ **(right)**.

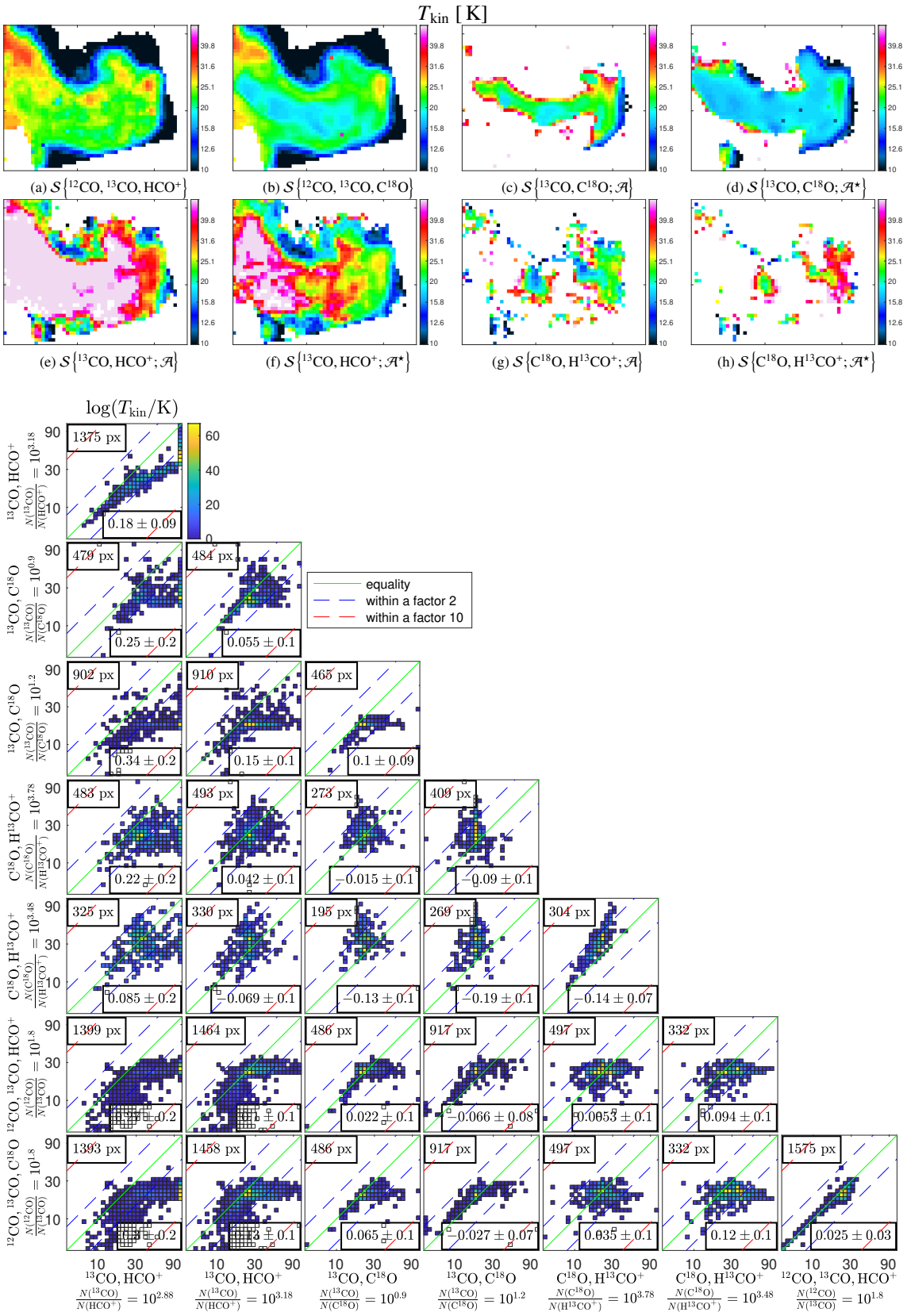


Fig. C.3. Same as Fig. 2 but for $T_{\rm kin}$.

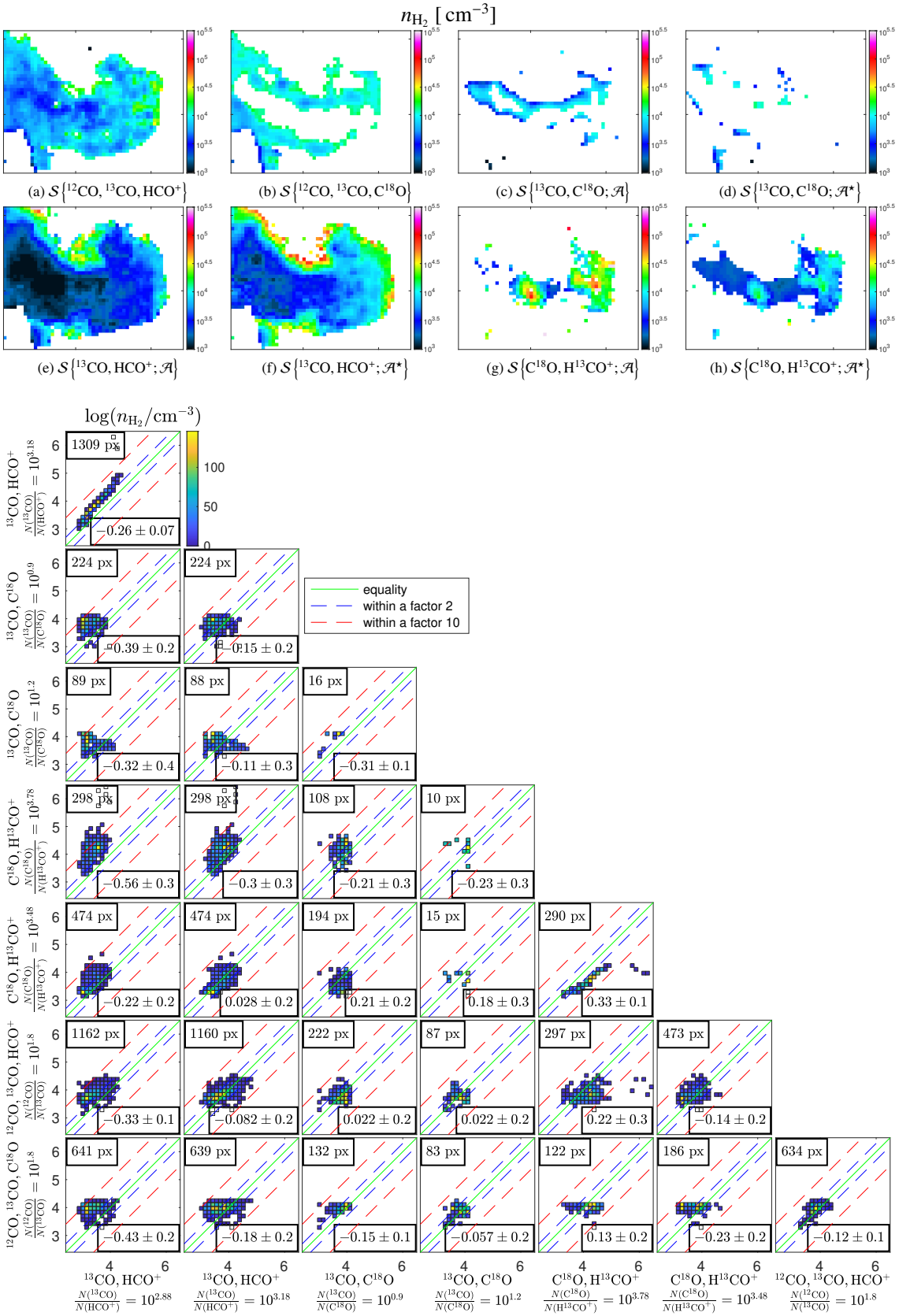


Fig. C.4. Same as Fig. 2 but for n_{H_2}.

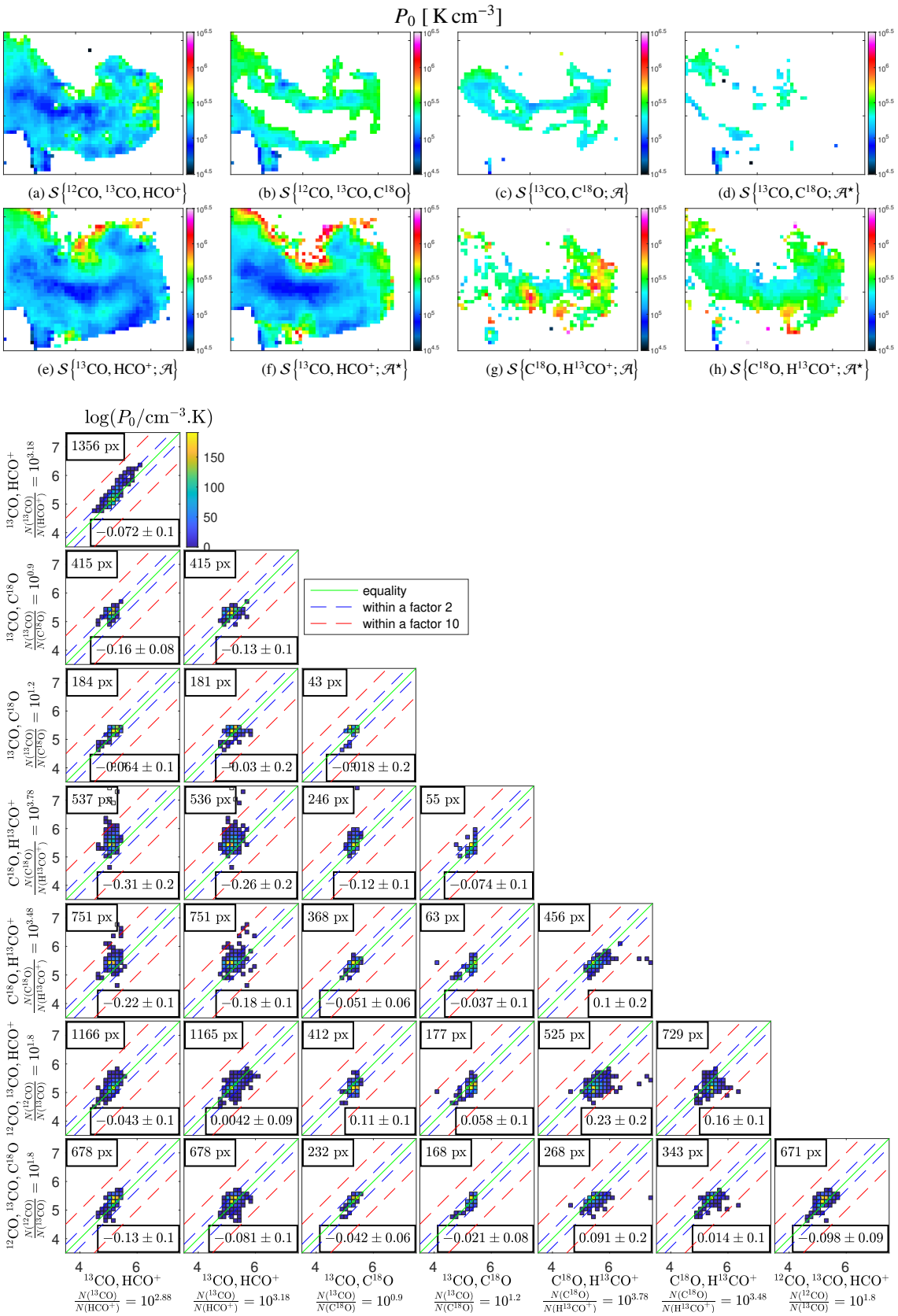


Fig. C.5. Same as Fig. 2 but for $P_{\rm th}$.

Appendix D: Supplemental figure for the parametric study of the reference precisions

Figure D.1 shows the main characteristics (signal-to-noise ratios, opacities, ratios of the volume density to the effective critical densities, and ratios of the excitation temperature over the kinetic temperature) of all the lines used in this paper for the same range of volume and column densities.

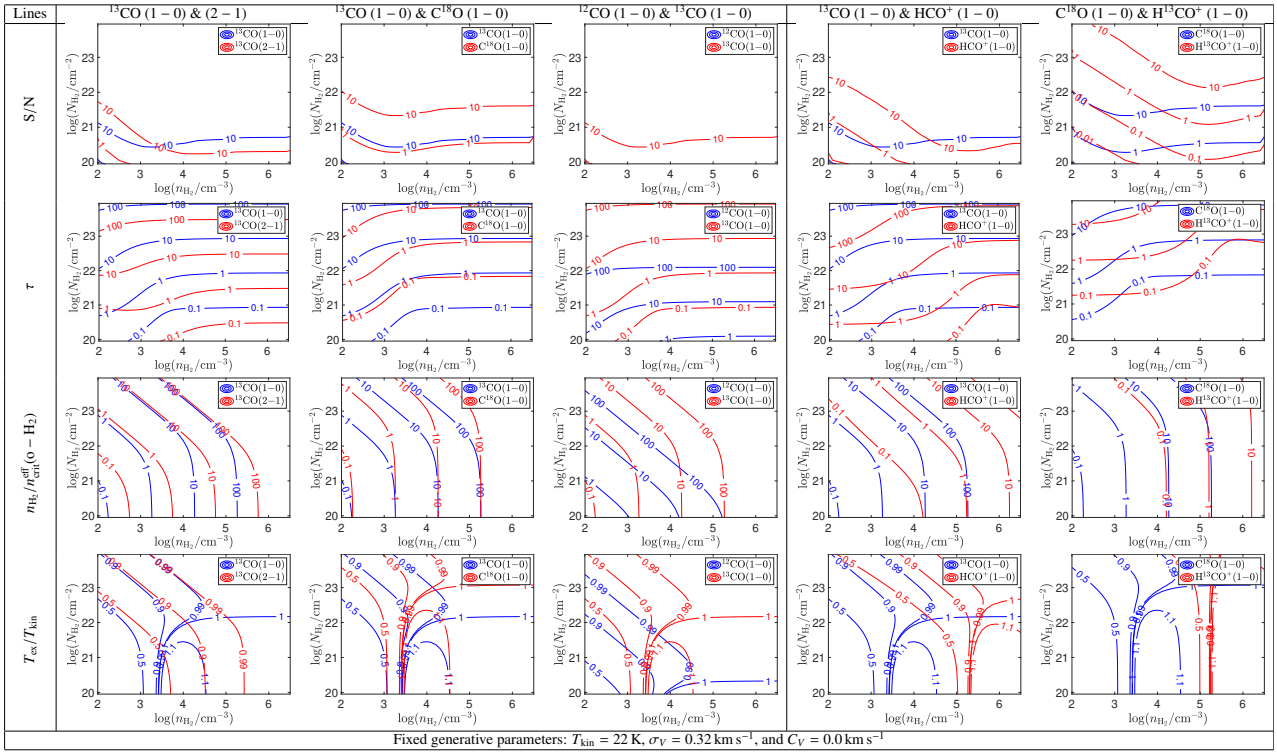


Fig. D.1. Line main characteristics: S/Ns, opacities, ratios of the volume density to the effective critical densities, and ratios of the excitation and kinetic temperatures. The blue and red curves show the results for the (1 – 0) and (2 – 1) lines, respectively. The values of the parameters that are fixed when generating the RADEX spectra are listed in the bottom row, except for the abundances that are given in Eq. (16).

Chapter 4

A heterogeneous cloud model to probe dense cores

Contents

4.1 A sandwich to model dense core line of sight **94**
4.1.1 Model statement 94
4.1.2 Application to the Horsehead nebula 95
4.2 Influence of the spectral sampling on the estimation precision **96**
4.2.1 Experiment setup 96
4.2.2 On the S/N saturation 97
4.2.3 Impact of the sampling phase on the CRB 98
4.2.4 Impact of the channel spacing on the CRB 99
References **105**
Appendix 4 . . . . . **105**
4.A Change of variables on the CRB 105
4.B Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead Nebula 106

Roueff *et al.* 2024 shed light on the fact that a homogeneous model of cloud is not capable of extracting robust estimates of the gas properties from multi-line data, because molecular tracers probe different physical regimes along the line of sight. In this chapter, I present in Sect. 4.1 the *multilayer* model of cloud that we developed in this thesis to overcome this issue. This model allows one to take into account potential variations of the gas characteristics (i.e., $T_{\rm kin}$, $n_{\rm H_2}$, N, and kinematics). While it remains a simple description of the cloud, it allows one to namely quantify the contribution of a given layer to the total emission, and to investigate the evolution of chemistry along the line of sight.

Moreover, we investigate in Sect. 4.2 the gain in the estimation precision that could be obtained by improving the resolution of the data, toward a typical dense core line of sight.

4.1 A sandwich to model dense core line of sight

4.1.1 Model statement

The studied molecular lines were carefully chosen as they probe either warm, diffuse gas (e.g., ^{12}CO), or cold, dense core (e.g., H^{13}CO$^+$), or gas at intermediate conditions (e.g., ^{13}CO, C^{18}O, HCO$^+$). To take into account the information contained by these complementary tracers, we proposed a "sandwich" model of the line of sight, i.e., made of an inner slab of gas, surrounded by an outer envelope. As illustrated in Figure 4.1, from the observer point of view, the received photons can therefore originate from 4 regions of the line of sight: the foreground layer, the inner layer, the background layer, and from the Cosmic Microwave Background (CMB). The outer and inner layer $i \in \{1, 2\}$ is characterized by a set of physicochemical parameters and kinematics $\boldsymbol{\theta}_i = \{T_{\mathrm{kin},i},\, n_{\mathrm{H_2},i},\, N_i, \mathrm{FWHM}_i, C_{V,i}\}$. Compared to the homogeneous model developed in Roueff *et al.* 2024, the sandwich model has twice more parameters.

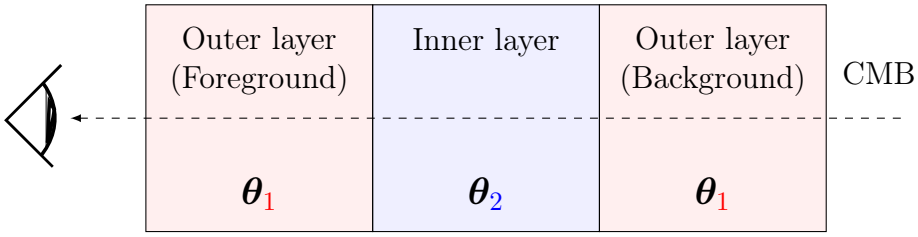


Figure 4.1: Sandwich model of the line of sight. The cloud is made of an inner layer, surrounded by an outer envelope, which thus consists in a foreground and background layer. The inner and outer layers are characterized by the set of 5 parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively.

Then, photons emitted by a given layer can be absorbed by the layer itself, but also by the layers in front of it. To derive the total amount of photons that indeed escape from the cloud, we thus extend the radiative transfer equation defined by Eq. (2.10) of Chapter 2, as the sum of all the layer's contribution. The line intensity s, function of the velocity channel V, thus reads

$$s(V) = s_{\mathrm{fore}}(V) + s_{\mathrm{inner}}(V) + s_{\mathrm{back}}(V) + s_{\mathrm{CMB}}(V), \tag{4.1}$$

with

$$\begin{aligned}
s_{\mathrm{fore}} &= J(T_{\mathrm{ex},1})\,(1 - \exp(-\Psi_1))\,, && (4.2)\\
s_{\mathrm{inner}} &= J(T_{\mathrm{ex},2})\,(1 - \exp(-\Psi_2))\exp(-\Psi_1)\,, && (4.3)\\
s_{\mathrm{back}} &= J(T_{\mathrm{ex},1})\,(1 - \exp(-\Psi_1))\exp(-\Psi_1 - \Psi_2)\,, && (4.4)\\
s_{\mathrm{CMB}} &= J(T_{\mathrm{CMB}})\,(\exp(-2\Psi_1 - \Psi_2)) - J(T_{\mathrm{CMB}}), && (4.5)
\end{aligned}$$

where J is the intensity function defined in Eq. (2.4), Ψ is the opacity profile defined in Eq. (2.11). These latter quantities are functions of the excitation temperature (T_{ex}) and the opacity (τ) of the line, respectively, which are obtained with RADEX for a given set of inputs $\{T_{\mathrm{kin}},\, n_{\mathrm{H_2}},\, N, \mathrm{FWHM}\}$. We provided in Appendix B of Ségal *et al.* 2024 the formalism

generalized to a $M-$layer cloud. We stress that this formalism is an approximation of the exact radiative transfer, which neglects the scattering of photons between layers. This approximation is consistent with the use of the escape probability formalism for the radiative transfer inside each individual layer.

4.1.2 Application to the Horsehead nebula

We used this model to fit the $(1-0)$ and $(2-1)$ lines of the CO and HCO^+ isotopologues toward the translucent, filamentary, and dense core lines of sight toward the Horsehead pillar. The details are available in the article by Ségal *et al.* 2024, which is provided in Appendix 4.B. We here only summarize the main astrophysical findings.

- A precise modeling of the chemistry and physics toward filaments and dense cores requires a heterogeneous model composed of, at least an inner dense layer surrounded by two outer translucent ones.

- The proposed model succeeds in reproducing integrated intensities within 20% for all lines. The associated bias is small for the ^{12}CO and $C^{18}O$ lines.

- We quantitatively confirm that the different lines of the CO and HCO^+ isotopologues are sensitive to different depths along the line of sight. Indeed, the emission from the $^{12}CO(1-0)$ line is dominated by the foreground layer (85%). The integrated intensity of both the lines of ^{13}CO comes from all the three layers at approximately the same proportion. The inner layer has the largest contribution to the total integrated line intensity of the $C^{18}O(1-0)$ and $C^{18}O(2-1)$ lines, of about 60%. The emission from the $HCO^+(1-0)$ and $H^{13}CO^+(1-0)$ lines is dominated by the outer layers (foreground and background), but the inner layer also contributes significantly. Moreover, the emission from these two lines is sub-thermalized with an excitation temperature lower than 10 K.

- The deduced chemical properties show 1) the depletion of CO isotopologues toward the dense core lines of sight and 2) the increase of the ^{12}CO and HCO^+ column density as a function of the visual extinction A_V or the ratio of the UV radiation field over the visual extinction, i.e., G_0/A_V, in the warm translucent sight lines. The basic abundances of ^{12}CO, ^{13}CO, and $C^{18}O$ defined in Eq. (24) are reached at A_V between 3 and 7 mag, respectively. The typical mean abundance of HCO^+ with respect to H_2 is $\sim 1.3\times10^{-9}$, which is about two times lower than the usual value in diffuse clouds (Gerin *et al.* 2019). The estimated column density ratio $N(^{13}CO)/N(C^{18}O)$ in the envelope increases with decreasing visual extinction and reaches 25 in the pillar outskirts.

- In addition to confirming the general rotation of the Horsehead pillar around its axis, we found the signature of accreting motions of the foreground layer onto the inner denser gas.

- While the inferred $T_{\rm kin}$ of the envelope varies from 25 to 40 K, that of the inner layer drops to $\sim 15\,$K in the western dense core. The estimated $n_{\rm H_2}$ in the inner layer is

$\sim 3 \times 10^4\,\mathrm{cm}^{-3}$ toward the filament, and it increases by a factor of ten toward dense cores. The thermal pressure of the outer layer shows a decreasing east-west gradient.

4.2 Influence of the spectral sampling on the estimation precision

The IRAM's current spectrometers are capable of sampling the full bandwidth provided by its receivers (i.e., 8 GHz × 2 sidebands × 2 polarizations) at a typical frequency spacing of 200 kHz. This sampling corresponds to a velocity resolution of about 0.5 and 0.25 $\mathrm{km\,s}^{-1}$ at the 3 and 1 mm bandwidth. This channel spacing barely enables us to resolve molecular lines with a FWHM of $\sim 1\,\mathrm{km\,s}^{-1}$, which is a typical value for local GMCs. Here, we investigate the evolution of the estimation precision of the sandwich's parameters as a function of the spectral resolution. Such study will enable us to quantitatively answer the two following questions:

1. Is there a significant gain in accuracy when fitting line profiles instead of fitting integrated intensities?
2. Is there an optimal spectral resolution?

4.2.1 Experiment setup

To this end, we study the impact of the spectral sampling on the CRB of the parameters of interest for a set of spectra simulated with the sandwich. The parameters chosen are listed in Table 4.1 and correspond to the values fitted toward the dense core line of sight called DC 2 in Ségal *et al.* 2024 . We will compute the *marginal* CRB, which takes into account both the additive noise and the calibration uncertainty (see Sect. 5.3.2 for details). We vary two sampling parameters. In a first experiment, we keep the channel spacing constant at 200 kHz and modify the sampling phase so that the line maximum lies between the center of a channel and one of its edges. In a second experiment, we also vary the channel spacing from 20 to 400 kHz. We start from a noise-free spectrum sampled at 2 kHz, i.e., at least 10 times better than the sampling range considered in both experiments. We also operate at constant integration time. The additive noise rms σ_b is given by

$$\sigma_b \propto \frac{T_{\mathrm{sys}}}{\sqrt{d\nu\, dt}},$$

where T_{sys} is the system temperature, $d\nu$ is the frequency channel spacing, and dt is the integration time. Furthermore, following the radio convention for velocities, the relationship between the frequency and velocity spacing ($d\nu$ and dV, respectively) is given by

$$|dV| = c\,\frac{d\nu}{\nu}, \tag{4.6}$$

where c is the speed of light, and ν the rest frequency of the line under consideration. Operating at constant integration time means co-varying the noise level and spectral resolution

as follows

$$\sigma_b \sqrt{d\nu} \propto \text{cst} \quad \Leftrightarrow \quad \sigma_b \sqrt{\nu\,|dV|} \propto \text{cst}. \tag{4.7}$$

Our starting point for σ_b is the noise levels measured for the ORION-B dataset (see Table 4.2). Finally, Table 4.3 gives the velocity dispersion (σ_V) of the sandwich expressed in frequency with Eq. (4.6), at the rest frequency of each molecular line. These quantities provide useful reference points for interpreting the results.

Table 4.1: Parameters used to create the molecular lines.

Parameter θ	*Units*	*Sandwich*	
		Outer Layer	*Inner Layer*
T_{kin}	K	28.41	13.14
$\log(n_{\text{H}_2}/\,\text{cm}^{-3})$	–	3.87	5
$\log(N_{^{12}\text{CO}}/\,\text{cm}^{-2})$	–	17.25	17.91
$\log(N_{^{13}\text{CO}}/\,\text{cm}^{-2})$	–	15.95	16.11
$\log(N_{\text{C}^{18}\text{O}}/\,\text{cm}^{-2})$	–	14.90	15.21
$\log(N_{\text{HCO}^+}/\,\text{cm}^{-2})$	–	12.72	13.23
$\log(N_{\text{H}^{13}\text{CO}^+}/\,\text{cm}^{-2})$	–	11.42	11.43
FWHM	$\text{km}\,\text{s}^{-1}$	0.85	0.42
C_V	$\text{km}\,\text{s}^{-1}$	10.46	10.29

Table 4.2: Characteristics of the studied molecular lines used to derive the constants in Eq. (4.7). The additive noise dispersion σ_b are the ones measured toward the DC 2 pixel, and $\sigma_b^\star$ derived from σ_b adjusted by the S/N saturation.

Species	*Transition*	*Frequency* [GHz]	σ_b [mK]	$\sigma_b^\star$ [mK]	dV [$\text{km}\,\text{s}^{-1}$]
^{12}CO	$(1-0)$	115.271202	89	2154	0.5
^{13}CO	$(1-0)$	110.201354	79	1378	0.1
	$(2-1)$	220.398684	151	1394	0.1
C^{18}O	$(1-0)$	109.782173	83	357	0.1
	$(2-1)$	219.560354	125	649	0.1
HCO$^+$	$(1-0)$	89.188525	276	277	0.1
H^{13}CO$^+$	$(1-0)$	86.754288	40	54	0.5

4.2.2 On the S/N saturation

We want to quantify the effect of spectral resolution on the estimation precision under conditions as close to reality as possible. We therefore saturate the S/N of the brightest lines to ensure that the faintest ones also contribute to the MLE, and to obtain an estimation precision not overly optimistic. Roueff *et al.* 2024 and Ségal *et al.* 2024 saturated the S/N peak at 10. At constant integration time and variable spectral resolution ($d\nu$), and by

Table 4.3: Velocity dispersion (σ_V) of the sandwich converted in frequency with Eq. (4.6) for all the considered molecular lines. The velocity dispersion is $0.36\,\mathrm{km\,s^{-1}}$ and $0.18\,\mathrm{km\,s^{-1}}$ in the outer and inner layer, respectively.

Molecular line	*Frequency* [GHz]	σ_V [kHz]	
		Outer Layer	*Inner Layer*
$H^{13}CO^+(1-0)$	86.754288	104	52
$HCO^+(1-0)$	89.188525	107	53
$C^{18}O(1-0)$	109.782173	132	65
$^{13}CO(1-0)$	110.201354	133	66
$^{12}CO(1-0)$	115.271202	139	69
$C^{18}O(2-1)$	219.560354	264	131
$^{13}CO(2-1)$	220.398684	265	131

assuming that the intensity peak is resolved, we obtain

$$S/N_{\mathrm{peak}} \propto \frac{1}{\sqrt{d\nu}}.$$

The S/N_{peak} saturation after modification of the spectral resolution would therefore produce threshold effects that will blur the interpretation of the results. Since the algorithm[1] that derives the sampled spectra preserves the line flux independently of the spectral sampling, a better way to balance the contributions of the different lines to the likelihood consists in saturating the S/N over the integrated line area. Indeed, at constant integration time and variable spectral resolution, one now has

$$S/N_{\mathrm{area}} \propto \mathrm{cst},$$

where the constant depends only on the fixed spectral range over which the line is integrated. The S/N_{area} saturation will therefore always have the same impact on the weight of the different lines in the likelihood, regardless of the spectral resolution. This solution is chosen for the following experiments and will be adopted in future studies.

4.2.3 Impact of the sampling phase on the CRB

Figure 4.2 shows the impact of the sampling grid phase on the CRB of the fitted parameters. The sampling grid is modified for all the lines used in the fit, but the upper panels of Fig. 4.2 show only the $HCO^+(1-0)$ line at 2 kHz ($\boldsymbol{s}(\boldsymbol{\theta})$ in Fig. 4.2a) and 0.5 $\mathrm{km\,s^{-1}}$ (Fig. 4.2b) channel spacing. In Fig. 4.2b, the green and red spectra show the profile of the $HCO^+(1-0)$ line when the actual peak of the line corresponds to the center of the final channel or one of its edges, respectively. The resampling algorithm keeps the integrated intensity constant. The middle and bottom panels show the variation in estimation precision normalized by its value at 20 kHz sampling, i.e., $\sqrt{\mathcal{B}(\theta)}/\sqrt{\mathcal{B}_{20\,\mathrm{kHz}}(\theta)}$, as a function of sampling phase, expressed

[1] To conserve the line flux from sampling, we apply a mean filter whose window size is $dV/0.001$.

in velocity. The middle and bottom lines show the normalized estimation precision for the outer and inner layers, respectively. As shown in Appendix 4.A, one can derive the CRB of θ, i.e., $\sqrt{\mathcal{B}(\theta)}$, from the one of $\log\theta$, i.e., $\sqrt{\mathcal{B}(\log\theta)}$. To ease the result interpretation, we therefore perform this change of variables for $T_{\rm kin}$, $n_{\rm H_2}$, and $N(^{13}{\rm CO})$ of both the sandwich's layers. The vertical dotted lines indicate the sampling phase for the green and red spectra.

Normalized precision values are always greater than 1, implying a degradation of the estimate as sampling becomes coarser. Overall, the precision on the inner layer parameters is more degraded than that of outer layer parameters. As expected, the parameters most affected are the kinematic ones (centroid velocity and line width), whose accuracy degradation ranges from 5 to 100%.

The degradation of the estimation precision on centroid velocity and line width are sometimes correlated and sometimes anti-correlated. The degradation is difficult to predict, as it depends on the asymmetrical shape of the line. The most surprising result is that the precision of column density estimation is degraded by around 25%.

As the effect of phase sampling can be significant, we will take it into account in the following experiment.

4.2.4 Impact of the channel spacing on the CRB

Figure 4.3 shows the degradation of the estimation precision (i.e., $\sqrt{\mathcal{B}(\theta)}/\sqrt{\mathcal{B}_{20\,\rm kHz}(\theta)}$) as a function of frequency channel spacing ($d\nu$) for the outer layer (first column) and inner layer (second column) parameters. The first row concerns the physical parameters $T_{\rm kin}$, $n_{\rm H_2}$ and $N(^{13}{\rm CO})$, while the second row concerns the kinematic parameters, i.e., the FWHM and C_V. To take into account the influence of the sampling phase on the CRB (discussed in Sect. 4.2.3), we plot the associated range of CRB values for each considered spectral spacing. Figure 4.3 shows results for spectral spacing between 20 and 400 kHz, while Fig. 4.4 focuses on the range between 20 and 100 kHz.

The CRB of all parameters decreases as spectral resolution improves, even though the noise level per channel increases. This results from the fact that line shapes contain information that is not resolved when spectral sampling is coarse. To illustrate these effects, Fig. 4.5 shows the line profiles modeled in the left-hand column and the line profiles resampled at 60, 120 and 240 kHz in the right-hand column. The gas velocity dispersions in the inner and outer layers are 0.18 and 0.36 $\rm km\,s^{-1}$, respectively. This translates into a spectral dispersion of around $60-130$ and $120-265$ kHz for the 3 and 1 mm lines, respectively (see Table 4.3). Indeed, the Nyquist sampling criterion indicates that a spectral resolution of 60 kHz barely resolves the contribution of the 3 mm lines to the inner layer (30 kHz would theoretically be necessary).

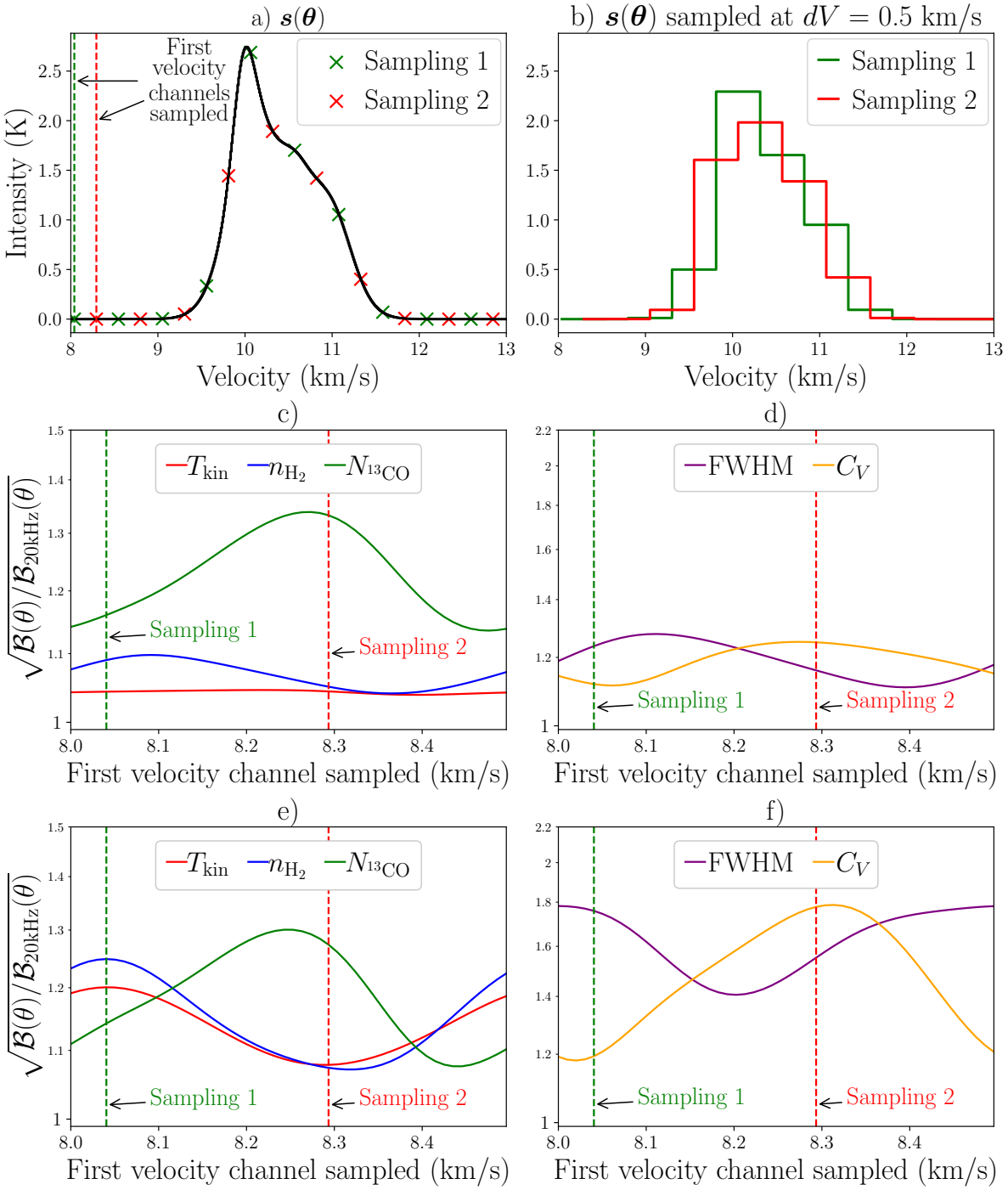


Figure 4.2: Illustration of the sensitivity of the estimation precision to the phase sampling through the $HCO^+(1-0)$ line. First row panels: a) The spectrum $\boldsymbol{s}(\boldsymbol{\theta})$ at a resolution of $d\nu = 20$ kHz, is sampled at the velocity resolution $dV = 0.5\,\mathrm{km\,s^{-1}}$ (or the spectral resolution $d\nu \sim 150\,\mathrm{kHz}$), with two samplings slightly shifted from each other (Sampling 1 and 2 indicated by the green and red crosses, respectively). b) The corresponding spectra reconstructed. Second row panels: c-d) Degradation of the precision of the estimation ($\sqrt{\mathcal{B}(\theta)}$) with respect to the precision of reference ($\sqrt{\mathcal{B}_{20\,\mathrm{kHz}}(\theta)}$) in function of the phase sampling for the outer layer's parameters. Third row panels: e-f) Same layout than the second row ones but for the inner layer.

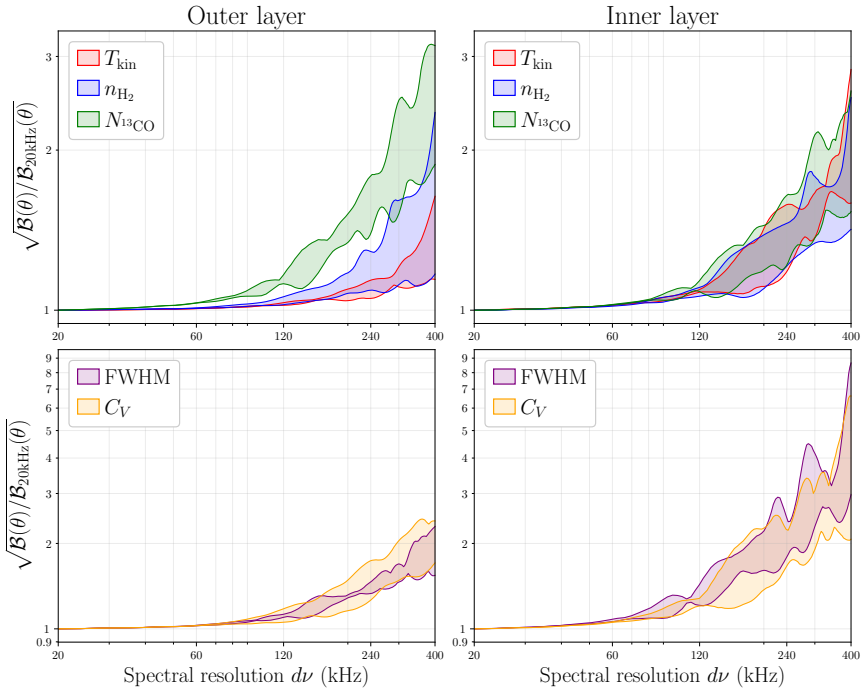


Figure 4.3: Degradation of the estimation precision $\sqrt{\mathcal{B}(\theta)}$ in function of the frequency channel spacing $(d\nu)$ with respect of the estimation precision of reference obtained at $d\nu = 20\,\text{kHz}$, i.e., $\sqrt{\mathcal{B}_{20\,\text{kHz}}(\theta)}$. First row, from left to right: Physical parameters of the outer and the inner layer. Second row: Same layout than the first row but for the gas kinematics.

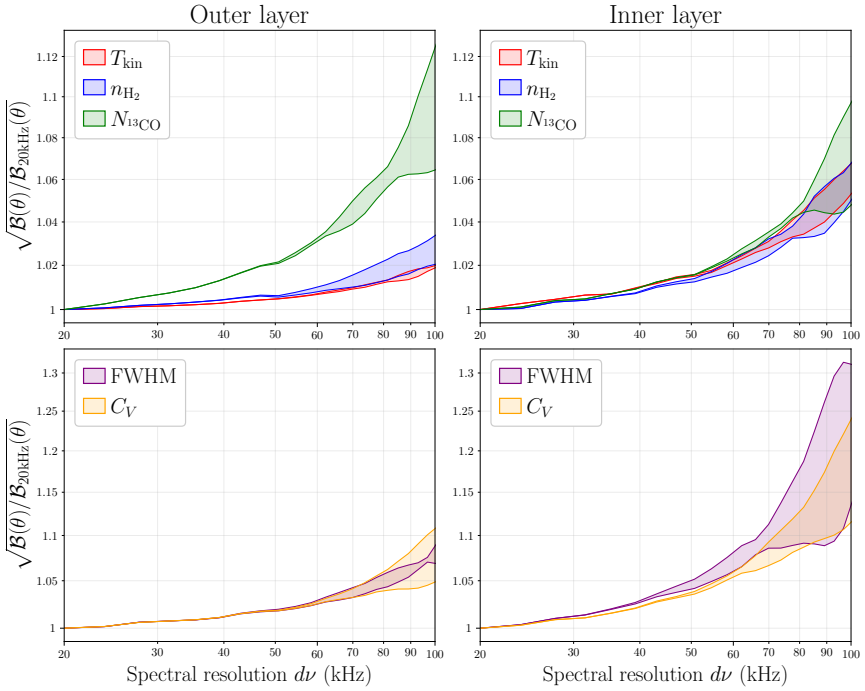


Figure 4.4: Same plots than in Fig. 4.3, but focused on the frequency sampling range between 20 and 100 kHz.

The rate of improvement in estimation precision becomes less than a few percents for all parameters when $d\nu \leq 60\,\mathrm{kHz}$. To quantify these results, Tables 4.4 to 4.6 show the degradation in precision of the parameters (T_{kin}, $n_{\mathrm{H_2}}$, and $N(^{13}\mathrm{CO})$, C_V, and FWHM) for three spectral resolutions, i.e., 60, 120, and 240 kHz.

The parameters whose precision is most affected by the spectral resolution deterioration are the centroid velocity and velocity dispersion (expressed as a FWHM), as expected. Kinetic temperature and volume density are the least affected by spectral resolution, probably because line excitation conditions play a more important role in overall amplitude than in line shape. The most surprising result for an astronomer is the sensitivity of the ^{12}CO column density precision to the spectral resolution. In particular, full line integration (e.g., using 400 kHz spectral resolution) gives results that are a factor of two more uncertain than estimates made with the best spectral resolution. The improved S/N ratio of the integrated line profile does not compensate for the loss of line shape information. This observation may come from a correlation between column density estimation of FWHM whose estimations are both related to the total flux of the signal.

In summary, the study of the evolution of parameter precision as a function of the spectral resolution suggests that a good profile resolution (Nyquist sampling) for most of the lines is important to extract all the information contained in the data. Furthermore, to be complete, this study should also take into account the baseline adjustment that occurs during data reduction, and which could bias the zero-reference level of intensities.

Table 4.4: Mean and standard deviation of the degradation of the estimation precision $(\sqrt{\mathcal{B}(\theta)}/\sqrt{\mathcal{B}_{20\,\mathrm{kHz}}(\theta)})$ at the spectral sampling $d\nu = 60\,\mathrm{kHz}$ for all the molecular lines.

Parameter θ	*Units*	*Sandwich*	
		Outer Layer	*Inner Layer*
T_{kin}	–	1.01 ± 0.00	1.02 ± 0.00
$n_{\mathrm{H_2}}$	$\mathrm{cm^{-3}}$	1.01 ± 0.00	1.02 ± 0.00
$N_{^{13}\mathrm{CO}}$	$\mathrm{cm^{-2}}$	1.03 ± 0.00	1.02 ± 0.00
FWHM	$\mathrm{km\,s^{-1}}$	1.02 ± 0.00	1.07 ± 0.01
C_V	$\mathrm{km\,s^{-1}}$	1.02 ± 0.00	1.05 ± 0.00

Table 4.5: Mean and standard deviation of the degradation of the estimation precision $(\sqrt{\mathcal{B}(\theta)}/\sqrt{\mathcal{B}_{20\,\mathrm{kHz}}(\theta)})$ at the spectral sampling $d\nu = 120\,\mathrm{kHz}$ for all the molecular lines.

Parameter θ	*Units*	*Sandwich*	
		Outer Layer	*Inner Layer*
T_{kin}	–	1.03 ± 0.00	1.09 ± 0.00
$n_{\mathrm{H_2}}$	$\mathrm{cm^{-3}}$	1.04 ± 0.01	1.08 ± 0.01
$N_{^{13}\mathrm{CO}}$	$\mathrm{cm^{-2}}$	1.14 ± 0.01	1.11 ± 0.02
FWHM	$\mathrm{km\,s^{-1}}$	1.11 ± 0.00	1.31 ± 0.08
C_V	$\mathrm{km\,s^{-1}}$	1.12 ± 0.04	1.25 ± 0.03

Table 4.6: Mean and standard deviation of the degradation of the estimation precision $(\sqrt{\mathcal{B}(\theta)}/\sqrt{\mathcal{B}_{20\,\mathrm{kHz}}(\theta)})$ at the spectral sampling $d\nu = 240\,\mathrm{kHz}$ for all the molecular lines.

Parameter θ	*Units*	*Sandwich*	
		Outer Layer	*Inner Layer*
T_{kin}	$-$	1.09 ± 0.03	1.36 ± 0.14
$n_{\mathrm{H_2}}$	cm^{-3}	1.19 ± 0.07	1.33 ± 0.09
$N_{^{13}\mathrm{CO}}$	cm^{-2}	1.54 ± 0.14	1.39 ± 0.15
FWHM	$\mathrm{km\,s}^{-1}$	1.40 ± 0.04	2.15 ± 0.27
C_V	$\mathrm{km\,s}^{-1}$	1.47 ± 0.17	1.96 ± 0.20

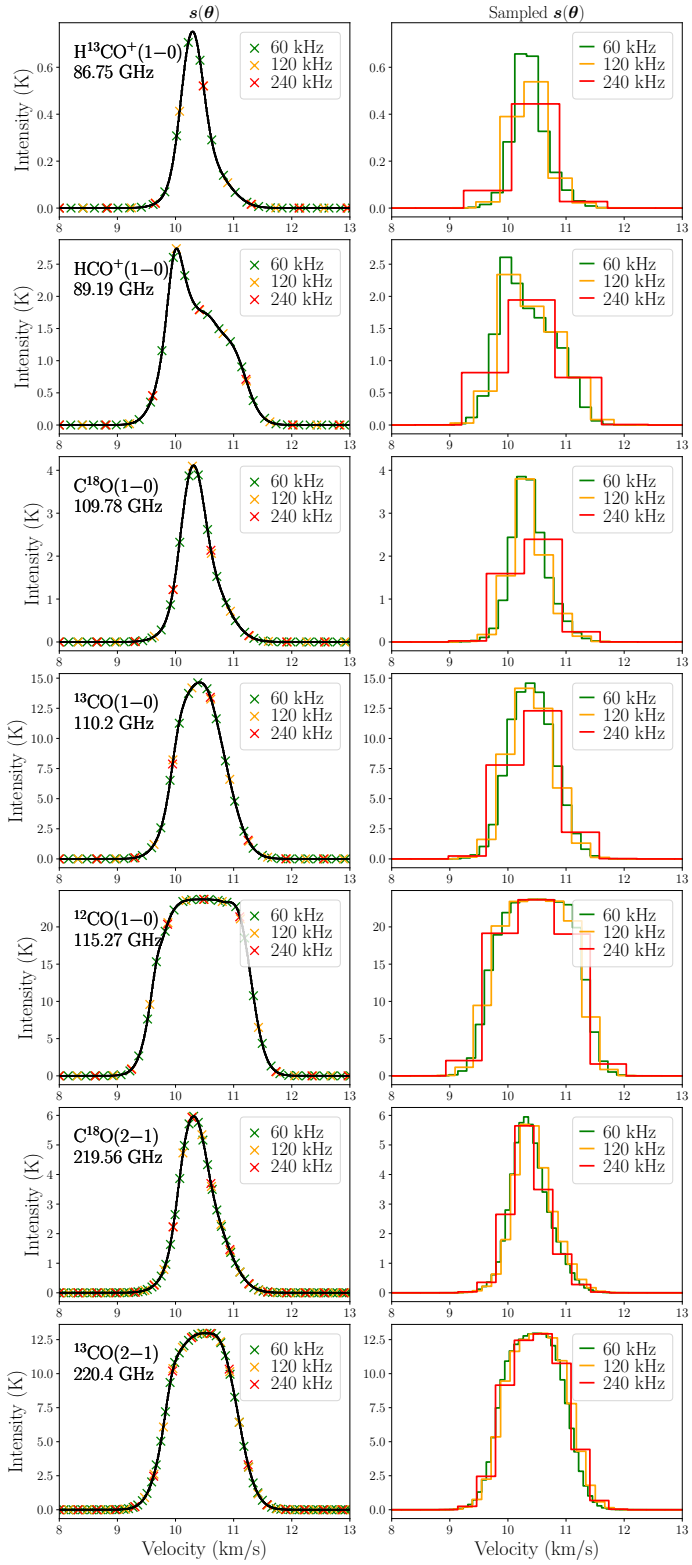


Figure 4.5: Illustration of the frequency sampling on the molecular line profiles. Each row concerns a molecular line. In the left panel is shown the line at 2 kHz, denoted by $\boldsymbol{s}(\boldsymbol{\theta})$, and at 60 (green), 120 (orange), and 240 kHz (red). The corresponding reconstructed spectra are shown in the right panel. Molecular lines are sorted in increasing order of rest frequency, from the top to the bottom (see the value printed at top left of each panel).

References

Gerin, M. *et al.* (2019). "Molecular ion abundances in the diffuse ISM: CF+, HCO+, HOC+, and C3H+". In: *A&A* 622, A26. DOI: 10.1051/0004-6361/201833661.

Kay, S. (1993). *Fundamentals of statistical signal processing: estimation theory*. USA: Prentice-Hall, Inc. ISBN: 0133457117.

Roueff, A. *et al.* (2024). "Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model - Application to the estimation of the gas temperature and volume density". In: *A&A* 686, A255. DOI: 10.1051/0004-6361/202449148.

Ségal, L. *et al.* (2024). "Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead nebula". In: *A&A* 692, A160. DOI: 10.1051/0004-6361/202451567.

Appendix 4

4.A Change of variables on the CRB

We consider having estimated the vector of physical parameters $\boldsymbol{\theta} = (\theta_1 \cdots \theta_M)^\top$, and having derived the corresponding CRB matrix $\boldsymbol{\mathcal{B}}(\boldsymbol{\theta})$, where the diagonal terms provide the estimation precision on each parameter. We now want to get the CRB on $\boldsymbol{\alpha} = \boldsymbol{f}(\boldsymbol{\theta}) \in \mathbb{R}^M$, denoted by $\boldsymbol{\mathcal{B}}(\boldsymbol{\alpha})$, such that

$$\forall i \in \{1, ..., M\}, \quad f_i(\theta) = \begin{cases} 10^\theta & \text{if } \theta \in \{\log T_{\rm kin},\, \log n_{\rm H_2},\, \log N(^{13}\text{CO})\}, \\ \theta & \text{if } \theta \in \{\text{FWHM},\, C_V\}, \end{cases} \tag{4.8}$$

in order to derive the estimation precision on, e.g., $T_{\rm kin}$ instead of $\log T_{\rm kin}$. Following Kay 1993, Sect. 3.8, one has

$$\boldsymbol{\mathcal{B}}(\boldsymbol{\alpha}) = \frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\, \boldsymbol{\mathcal{B}}(\boldsymbol{\theta})\, \frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^\top,$$

where $\frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is the Jacobian matrix. For the function $\boldsymbol{f}$ defined in Eq. (4.8), the Jacobian matrix is of shape $M \times M$, diagonal and its diagonal terms read

$$\left[\frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]_{i,i} = \begin{cases} \ln 10 \times 10^{\theta_i} & \text{if } \theta_i \in \{\log T_{\rm kin},\, \log n_{\rm H_2},\, \log N(^{13}\text{CO})\}, \\ 1 & \text{if } \theta_i \in \{\text{FWHM},\, C_V\}. \end{cases}$$

Hence, the CRB on, e.g., $T_{\rm kin}$, denoted by $\mathcal{B}(T_{\rm kin})$ is expressed from the one of $\log T_{\rm kin}$, denoted by $\mathcal{B}(\log T_{\rm kin})$, as follows

$$\mathcal{B}(T_{\rm kin}) = (\ln 10 \times T_{\rm kin})^2\, \mathcal{B}(\log T_{\rm kin}).$$

The estimation precision is thus given by $\sqrt{\mathcal{B}(T_{\rm kin})} = (\ln 10 \times T_{\rm kin})\, \sqrt{\mathcal{B}(\log T_{\rm kin})}$. Considering the degradation of the precision $\sqrt{\mathcal{B}(\theta)}/\sqrt{\mathcal{B}_{20\,\rm kHz}(\theta)}$, as done in Sect. 4.2, is thus equivalent to considering $\sqrt{\mathcal{B}(\log\theta)}/\sqrt{\mathcal{B}_{20\,\rm kHz}(\log\theta)}$.

4.B Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead Nebula

A&A, 692, A160 (2024)
https://doi.org/10.1051/0004-6361/202451567
© The Authors 2024

Astronomy & Astrophysics

Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead nebula

Léontine Ségal[1,2], Antoine Roueff[2,⋆], Jérôme Pety[1,3], Maryvonne Gerin[3], Evelyne Roueff[9], Javier R. Goicoechea[4], Ivana Bešlic[3], Simon Coudé[14,15], Lucas Einig[1,7], Helena Mazurek[3], Jan H. Orkisz[1], Pierre Palud[8,9], Miriam G. Santa-Maria[17,4], Antoine Zakardjian[10], Sébastien Bardeau[1], Emeric Bron[9], Pierre Chainais[8], Karine Demyk[10], Victor de Souza Magalhaes[5], Pierre Gratier[6], Viviana V. Guzman[11], Annie Hughes[10], David Languignon[9], François Levrier[12], Jacques Le Bourlot[9], Franck Le Petit[9], Dariusz C. Lis[13], Harvey S. Liszt[5], Nicolas Peretto[16], Albrecht Sievers[1], and Pierre-Antoine Thouvenin[8]

(Affiliations can be found after the references)

Received 18 July 2024 / Accepted 23 September 2024

ABSTRACT

Context. Dense and cold molecular cores and filaments are surrounded by an envelope of translucent gas. Some of the low-*J* emission lines of CO and HCO^+ isotopologues are more sensitive to the conditions either in the translucent environment or in the dense and cold one because their intensities result from a complex interplay of radiative transfer and chemical properties of these heterogeneous lines of sight (LoSs).

Aims. We extend our previous single-zone modeling with a more realistic approach that introduces multiple layers to take account of possibly varying conditions along the LoS. We used the IRAM-30m data from the ORION-B large program toward the Horsehead nebula in order to demonstrate our method's capability and effectiveness.

Methods. We propose a cloud model composed of three homogeneous slabs of gas along each LoS, representing an outer envelope and a more shielded inner layer. We used the non-LTE radiative transfer code RADEX to model the line profiles from the kinetic temperature ($T_{\rm kin}$), the volume density ($n_{\rm H_2}$), kinematics, and chemical properties of the different layers. We then used a fast and robust maximum likelihood estimator to simultaneously fit the observed lines of the CO and HCO^+ isotopologues. To limit the variance on the estimates, we propose a simple chemical model by constraining the column densities.

Results. A single-layer model cannot reproduce the spectral line asymmetries that result from a combination of different radial velocities and absorption effects among layers. A minimal heterogeneous model (three layers only) is sufficient for the Horsehead application, as it provides good fits of the seven fitted lines over a large part of the studied field of view. The decomposition of the intensity into three layers allowed us to discuss the distribution of the estimated physical or chemical properties along the LoS. About 80% of the ^{12}CO integrated intensity comes from the outer envelope, while ∼55% of the integrated intensity of the (1 − 0) and (2 − 1) lines of $C^{18}O$ comes from the inner layer. For the lines of the ^{13}CO and the HCO^+ isotopologues, integrated intensities are more equally distributed over the cloud layers. The estimated column density ratio $N(^{13}CO)/N(C^{18}O)$ in the envelope increases with decreasing visual extinction, and it reaches 25 in the pillar outskirts. While the inferred $T_{\rm kin}$ of the envelope varies from 25 to 40 K, that of the inner layer drops to ∼15 K in the western dense core. The estimated $n_{\rm H_2}$ in the inner layer is $\sim 3 \times 10^4$ cm^{-3} toward the filament, and it increases by a factor of ten toward dense cores.

Conclusions. Our proposed method correctly retrieves the physical and chemical properties of the Horsehead nebula. It also offers promising prospects for less supervised model fits of wider-field datasets.

Key words. methods: data analysis – methods: statistical – ISM: abundances – ISM: clouds – ISM: lines and bands – photon-dominated region (PDR)

1. Introduction

Star formation takes place in giant molecular clouds (GMCs) with an overall small efficiency (Shu et al. 1987; Lada & Lada 2003; Kennicutt & Evans 2012; Schinnerer & Leroy 2024, and references therein). Current star formation theory assumes that a bottleneck comes from the formation of dense cores of density $\gtrsim 10^4$ cm^{-3} (e.g., Bergin & Tafalla 2007), and recent observations suggest that the galactic environment of a GMC impacts its star formation efficiency (Schinnerer & Leroy 2024).

Although GMCs show an internal structure, which can be probed by observing various molecular lines, a typical representation of a GMC is still lacking. This is significant, as it will benchmark the general understanding of the molecular interstellar medium (ISM) in the Milky Way and nearby galaxies. To achieve this representation, realistic large-scale maps of the cloud physical parameters (column densities, volume density, kinetic temperature, and kinematics) are needed for a representative sample of GMCs in the Milky Way. While the current generation of millimeter observatories, such as the IRAM-30m telescope, NOEMA, or ALMA, enables surveys of tens of molecular lines over wide fields of view (e.g., Pety et al. 2017; Kauffmann et al. 2017; Evans et al. 2020; Barnes et al. 2020; Tafalla et al. 2021; Dame & Lada 2023), there is still a lack of fully relevant methods to easily extract the physical and chemical information that these lines encode.

⋆ Corresponding author; antoine.roueff@univ-tln.fr

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

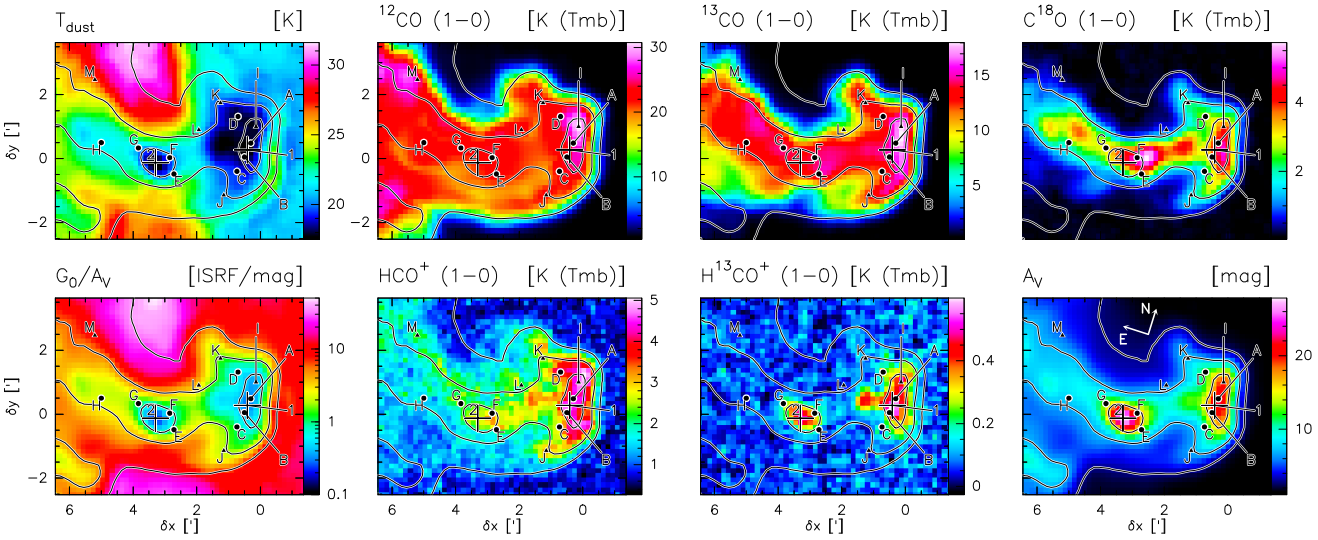


Fig. 1. Spatial distribution of the peak main beam temperature (T_{mb}) of the (1 − 0) line of the CO and HCO^+ isotopologues as well as of the dust temperature (T_{dust}), visual extinction (A_V), and the ratio of the far UV field over the visual extinction (G_0/A_V) toward the Horsehead nebula. The used projection is the Azimuthal one rotated by 14° around the center of coordinates RA= $05^h40^m54.27^s$, Dec= −02°28′00.00″ in the International Celestial Reference System. Contours of the dust visual extinction at 3, 7, and 16 magnitudes are overlaid on each panel. The crosses and circles show lines of sight that are studied in more detail in this paper (see Table 2 for the corresponding coordinates).

It is thus important to develop fast and rigorous methods to infer 2D maps of cloud physical parameters from modern-day large-scale, well-resolved multi-line maps. In order to accurately fit the observed line intensities and profiles, these methods must take into account that lines of sight (LoSs) contain gas components with different physical conditions and chemical compositions, which often interact radiatively. Previous studies based on coupled non-local thermodynamic equilibrium (non-LTE) radiative transfer and chemical models were designed to fit high angular resolution multi-line observations toward either a single line of sight of interest or narrow fields of view (see, e.g., Goicoechea et al. 2006, 2009, toward the Horsehead PDR). With the ability to conduct large mapping surveys of molecular gas lines at high spatial resolution, the main challenge now is to develop robust methods for much larger areas.

The simplest modeling geometry to do this with is a homogeneous LoS composed of a single component. However, the line emission maps of Pety et al. (2017); Kauffmann et al. (2017); Barnes et al. (2020) suggest that the emission comes from gas with different physical conditions, at least along the LoSs toward filaments and dense cores. Roueff et al. (2024) document inconsistencies in the inferred gas conditions when simultaneously analyzing different molecular lines with a non-LTE radiative transfer cloud model of a homogeneous slab of gas. For instance, simultaneously fitting different sets of lines (such as ^{12}CO, ^{13}CO, and HCO^+ on the one hand and $C^{18}O$ and $H^{13}CO^+$ on the other) in the same field of view yields drastic differences in inferred volume densities up to one order of magnitude. This suggests that model misspecifications prevent us from deciphering the information provided by emission lines. In this study, we develop a three-layer model to consider the heterogeneous conditions along the LoSs toward filaments or dense cores while keeping the model as simple as possible. This multi-layer cloud model is based on routinely used radiative transfer methods. In spite of its simplicity, it allows us to analyze the contribution of the different layers to the line emission in addition to delivering an estimation of the physical quantities with their confidence interval.

The region under study is the Horsehead nebula (Barnard 33), located in the Orion B cloud. This region is an ideal candidate for evaluating the model's performance for several reasons. First, this nebula hosts a diversity of physico-chemical regimes such as photon dissociation regions (PDRs); translucent regions ($T_{kin} > 30$ K, $n_{H_2} \lesssim 10^4$ cm^{-3}); and dense cores ($T_{kin} < 15$ K, $n_{H_2} \gtrsim 10^5$ cm^{-3}). Corresponding density and temperature gradients (e.g., within the PDR Hernández-Vera et al. (2023)) and steep changes of the gas chemical state detected along the LoS suggest that a standard radiative transfer analysis within a homogeneous cloud model is a too simple description for this region. Second, the Horsehead nebula has been the subject of numerous studies, from the analysis of its kinematics (e.g., Hily-Blant et al. 2005; Gaudel et al. 2023) to the dissection of the chemistry and structure of its PDR (e.g., Gerin et al. 2009; Guzmán et al. 2011, 2012; Pety et al. 2012; Gratier et al. 2013; Guzmán et al. 2013, 2014, 2015; Fuente et al. 2017; Hernández-Vera et al. 2023; Abergel et al. 2024). Applying our model to this familiar, albeit complex object, is therefore an opportunity to better diagnose the relevance of the proposed technique.

The paper is organized as follows. Section 2 briefly describes the observations. The radiative transfer and noise models as well as the fitting method and the estimation of the confidence intervals are described in Sect. 3. The physical and chemical assumptions are justified in Sect. 4. Section 5 presents the results on the intensity decomposition, the gas velocity field, the estimated chemical abundances, and the physical conditions (kinetic temperature and volume density). Section 6 discusses the estimated abundances. Finally, Sect. 7 summarizes our findings.

2. Observations

In this study, we use the same data as in Roueff et al. (2024). As they are fully described there, we only summarize useful information here. The field of view contains the Horsehead nebula. Figure 1 shows the spatial distribution of various tracers of

Table 1. Characteristics of the studied molecular lines.

Species	Transition	Frequency (GHz)	Vel. Res. (km s^{-1})	Ang. Res. ($''$)	σ_b [(a)] (mK)
^{12}CO	(1 – 0)	115.271202	0.5	31	91
^{13}CO	(1 – 0)	110.201354	0.1	31	74
	(2 – 1)	220.398684	0.1	31	155
C^{18}O	(1 – 0)	109.782173	0.1	31	72
	(2 – 1)	219.560354	0.1	31	121
HCO$^+$	(1 – 0)	89.188525	0.1	31	318
H^{13}CO$^+$	(1 – 0)	86.754288	0.5	31	45

Notes. [(a)]Median of standard deviations of signal-free channels, after resampling and smoothing (see Sect. 3.4). This corresponds to the typical baseline noise level at the angular and spectral resolutions listed in the table.

the interstellar gas and dust. The data include the low-J transitions of the isotopologues of carbon monoxide (^{12}CO, ^{13}CO, and C^{18}O) and formylium (HCO$^+$and H^{13}CO$^+$), observed with the IRAM-30m telescope, as part of the ORION-B large program (PIs: J. Pety & M. Gerin). Table 1 lists the properties of the studied emission lines.

The data also include maps of the dust visual extinction (A_V) and of dust temperature ($T_{\rm dust}$) derived by Lombardi et al. (2014) and Pety et al. (2017) from the Herschel Gould Belt Survey data (André et al. 2010). We used the far UV illumination parameter (G_0) derived by Santa-Maria et al. (2023) to compute a map of the ratio G_0/A_V. Assuming that there is some relation between the volume and column densities, G_0/A_V is a proxy of the ratio of the far UV field by the volume density, which controls to first order the chemistry of PDRs. In this study, the spatial distribution of A_V, $T_{\rm dust}$, and G_0/A_V only serve i) as references to locate the different media that are present in the Horsehead nebula, and ii) to compute the chemical abundances. In particular, they are not taken into account as prior knowledge to characterize the gas physical properties along each LoS.

We identify in Fig. 1 two dense cores where $A_V > 16$ mag and $T_{\rm dust} < 20$ K. These cores correspond to the LoSs where the H^{13}CO$^+$ (1 – 0) line is clearly detected. The maximum of the H^{13}CO$^+$ (1 – 0) intensity also corresponds to the maximum of the visual extinction or gas column density. In each core, the C^{18}O (1 – 0) emission peaks are offset compared to the peaks of the H_2 column density. This might be due to C^{18}O depletion from the gas phase, as CO is known to freeze on grain mantles when the dust grains are cold enough ($T_{\rm dust} \leq 15$ K, e.g., Tafalla et al. 2002; Hily-Blant et al. 2005).

Outside these dense cores, the line emission from C^{18}O (1 – 0) is mostly associated with a filamentary structure where $A_V > 7$ mag. The ^{12}CO and ^{13}CO (1 – 0) emission is relatively bright down to $A_V \simeq 3$ mag. Conversely, the emission of these two lines shows no obvious spatial pattern where the two dense cores are seen at high visual extinction. The HCO$^+$ (1 – 0) emission fills most of the regions with $A_V > 3$ mag. But it shows a large east-west emission gradient, probably related to the far UV illumination from the σ Ori star that excites the IC 434 HII region (see, e.g., Ochsendorf et al. 2014). Another striking feature in Fig. 1 is the variation of the signal-to-noise ratio (S/N) by more than one order of magnitude as a function of the emission line. The ^{12}CO and ^{13}CO (1 – 0) lines have a maximum S/N of 369 and 242, respectively, while the HCO$^+$ and H^{13}CO$^+$ (1 – 0) maximum S/N are 18.4 and 12.5, respectively. This mostly reflects the relative line intensities as broadband receivers deliver similar receiver noise over their bandpass.

Table 2. Coordinates of the lines of sight marked in Fig. 1. DC and LoS refer to dense core and line of sight, respectively.

Name	RA	Dec	$(\delta x, \delta y)$
DC 1	05^{h}40^{m}55.6^s	−02°27′38″	(+025″, +17″)
LoS A	05^{h}40^{m}54.8^s	−02°27′28″	(+018″, +29″)
LoS B	05^{h}40^{m}56.2^s	−02°27′50″	(+030″, +03″)
LoS C	05^{h}40^{m}57.5^s	−02°28′13″	(+044″, −24″)
LoS D	05^{h}40^{m}55.7^s	−02°26′33″	(+042″, +79″)
DC 2	05^{h}41^{m}07.2^s	−02°27′19″	(+198″, −08″)
LoS E	05^{h}41^{m}05.3^s	−02°27′49″	(+163″, −29″)
LoS F	05^{h}41^{m}05.3^s	−02°27′17″	(+171″, +02″)
LoS G	05^{h}41^{m}08.8^s	−02°26′45″	(+230″, +20″)
LoS H	05^{h}41^{m}13.2^s	−02°26′18″	(+300″, +30″)
LoS I	05^{h}40^{m}53.7^s	−02°27′03″	(+008″, +60″)
LoS J	05^{h}40^{m}59.8^s	−02°28′50″	(+069″, −69″)
LoS K	05^{h}40^{m}57.4^s	−02°26′01″	(+075″, 105″)
LoS L	05^{h}41^{m}00.9^s	−02°26′40″	(+116″, +54″)
LoS M	05^{h}41^{m}13.2^s	−02°24′40″	(+324″, 126″)

Table 2 lists the coordinates of LoSs marked in Fig. 1. These LoSs are studied in greater detail in Sect. 5. The positions DC 1 and DC 2 represent the centers of the eastern and western dense cores that belong to the Horsehead nebula. The LoSs (A, B, C, D) and (E, F, G, H) allowed us to study variations inside or in the outskirt of these two dense cores. There is an embedded young stellar object toward LoS I. Finally, the LoSs (J, K, L, M) are representative of the more translucent gas around the Horsehead pillar.

The visible extinction estimated from the dust spectral energy distribution is a measure of all hydrogen atoms present along each LoS. We needed to correct it for two effects in order to meaningfully match it with the results of our line fitting method. First, Pety et al. (2017) estimate that about 1 mag of visual extinction is coming from atomic hydrogen. Second, the southwestern edge of the Orion B GMC has typically two velocity components centered on 5 and 10.5 km s^{-1} (Pety et al. 2017), while the method proposed here assumes that only one velocity component is present. Appendix A details how we corrected the visual extinction to match the amount of molecular gas associated with the main velocity component at 10.5 km s^{-1}. From this point on, we only use this corrected visual extinction, which we call "molecular" visual extinction.

3. Methods

The emission from the CO and HCO$^+$ isotopologues seems to be sensitive to the different gas regimes that are present along the LoS throughout the Horsehead nebula. Cold and dense gas in cores ($T_{\rm kin} \sim 15$ K, $n_{\rm H_2} \geq 10^5$ cm^{-3}, Ward-Thompson et al. 2006) is surrounded by gas with intermediate temperature and density in filaments ($T_{\rm kin} \in [20, 40]$ K, $n_{\rm H_2} \sim 10^4$ cm^{-3}), and an envelope of warm and translucent gas ($T_{\rm kin} \in [40, 100]$ K, $n_{\rm H_2} \sim 10^3$ cm^{-3}, see, e.g., Goicoechea et al. 2006).

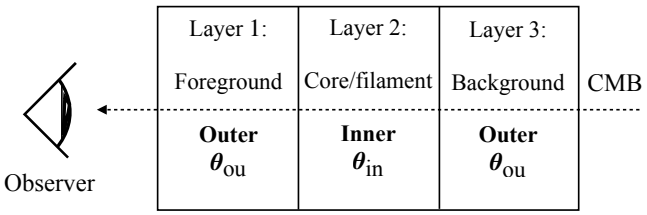


Fig. 2. Sketch of the multilayer model of a LoS. Top: line of sight decomposed into three layers: the foreground, inner, and background layers. Bottom: inner slab (layer 2) surrounded by two outer slabs that share the same physical and chemical state. Each region has homogeneous physical and chemical conditions (noted $\boldsymbol{\theta}_{\rm in}$ and $\boldsymbol{\theta}_{\rm ou}$, respectively).

Roueff et al. (2024) show the benefits of simultaneous fitting CO and HCO^+ isotopologue lines to probe the Horsehead nebula environments, but highlight the limitations and biases induced by a homogeneous model for describing these very molecular lines. This paper studies how these limitations are overcome by fitting the same lines with a heterogeneous model. In this section we introduce the model for an heterogeneous line of sight and the methods to fit it efficiently.

3.1. Modeling a heterogeneous line of sight

We aim to model the emission from a dense core or a filament surrounded by translucent gas. To do this, we model each LoS as three successive layers starting from the observer as illustrated in the top part of Fig. 2: the foreground layer (layer 1), the inner layer (layer 2), the background layer (layer 3). The contribution from the Cosmic Microwave Background (CMB) is considered separately as it is not emitted by the ISM cloud. We nickname this specific multilayer structure "sandwich model" because the foreground and background layers are part of the same outer layer and share the same physical and chemical conditions (see the bottom part of Fig. 2). In contrast, the inner layer has physical and chemical conditions that are distinct from those of the outer layer.

Following Myers et al. (1996), we assume 1) that the line excitation is local and the level populations in a slab do not depend on the physical conditions of other layers, but 2) that the emission of one layer is absorbed by the layers located between it and the observer. In other words, we neglect the non-local radiative coupling between layers, and the scattering of line photons of the core by the envelope, that is, the absorption and reemission of line photons coming from the core. This is a good approximation for CO and for optically thin lines, but it will be less exact for very subthermal ($T_{\rm ex} \ll T_{\rm kin}$) optically thick lines. These assumptions are consistent with the escape probability radiative transfer method that is used to model the emission from each layer (see, e.g., van der Tak et al. 2007). This approximation allows us to easily decompose the contribution of the different layers to the observed spectrum.

The brightness temperature s at observed frequency ν can be written as

$$s(\nu) = s_{\rm fore}(\nu) + s_{\rm inner}(\nu) + s_{\rm back}(\nu) + s_{\rm CMB}(\nu), \tag{1}$$

where $s_L(\nu)$ is the brightness temperature of the layer $L \in \{\rm fore, inner, back\}$ attenuated by the gas in front of it. For any layer L, it can be written as

$$s_L(\nu) = J(T_{{\rm ex},L}, \nu)\,\left[1 - \exp\left(-\Psi_L(\nu)\right)\right] \exp\left(-\sum_{i=1}^{L-1} \Psi_i(\nu)\right), \tag{2}$$

where $\Psi_i(\nu)$ is the opacity profile of layer i and the specific intensity of layer L, expressed in temperature units, is

$$J(T_{{\rm ex},L}, \nu) = \frac{h\nu}{k} \frac{1}{\exp\left(\frac{h\nu}{kT_{{\rm ex},L}}\right) - 1}. \tag{3}$$

The constants h and k in Eq. (3) are the Planck and Boltzmann constants, respectively. The specific intensity of layer L, $J(T_{{\rm ex},L}, \nu)$, is affected by the layer itself through the term $[1 - \exp(-\Psi_L(\nu))]$, and by the gas of the layers separating the layer L from the observer through the term $\exp\left(-\sum_{i=1}^{L-1} \Psi_i(\nu)\right)$. The excitation temperature $T_{{\rm ex},L}$ between the upper and lower energy levels is related to the ratio of the populations of these energy levels through

$$\frac{n_{\rm up}}{n_{\rm low}} = \frac{g_{\rm up}}{g_{\rm low}} \exp\left(-\frac{h\nu}{kT_{{\rm ex},L}}\right), \tag{4}$$

where $g_{\rm up}$ and $g_{\rm low}$ are the degeneracies of the associated levels (Draine 2011, Chap. 7, Eq. (7.8)).

Each molecular transition provides a spectrum around the rest frequency ν_l of the emission line l in the source frame. When dealing with several molecular lines simultaneously, it is more convenient to consider opacity profiles as a function of velocity V rather than frequency ν. These are related through the Doppler effect expressed with the radio convention as

$$\frac{\nu}{\nu_l} = 1 - \frac{V}{c}, \tag{5}$$

where c is the velocity of light. Assuming that the distribution of the gas velocities, which results from turbulent and thermal motions, is a Gaussian, centered on the centroid velocity $C_{V,L}$, and of velocity dispersion $\sigma_{V,L}$, we can express the opacity profile for layer L and line l as a function of the velocity as

$$\Psi_{l,L}(V) = \tau_{l,L} \exp\left[-\frac{(V - C_{V,L})^2}{2\sigma_{V,L}^2}\right], \tag{6}$$

where $\tau_{l,L}$ is the optical depth at the center of line l for the layer L. We assume that the centroid velocity $C_{V,L}$ and velocity dispersion $\sigma_{V,L}$ are identical for all molecular lines in this layer. They thus only depend on the layer L. The full width at half maximum (FWHM) is related to the velocity dispersion σ_V by

$$\text{FWHM} = \sqrt{8\ln 2}\,\sigma_V. \tag{7}$$

Finally, as the lines observed in the ISM have narrow widths, we can assume that $J(T_{{\rm ex},L}, \nu) \approx J(T_{{\rm ex},L}, \nu_l)$. For the sake of readability, we hereafter note $J(T_{{\rm ex},L}, \nu)$ as $J(T_{{\rm ex},L})$ and $\Psi_L(V) = \Psi_L$.

In summary, the spectrum observed at the output of the simplified sandwich model can be written as

$$s(V) = s_{\rm fore}(V) + s_{\rm inner}(V) + s_{\rm back}(V) + s_{\rm CMB}(V), \tag{8}$$

with

$$s_{\rm fore} = J(T_{{\rm ex},1})\left[1 - \exp\left(-\Psi_1\right)\right], \tag{9}$$
$$s_{\rm inner} = J(T_{{\rm ex},2})\left[1 - \exp\left(-\Psi_2\right)\right]\exp\left(-\Psi_1\right), \tag{10}$$
$$s_{\rm back} = J(T_{{\rm ex},1})\left[1 - \exp\left(-\Psi_1\right)\right]\exp\left(-\Psi_1 - \Psi_2\right), \tag{11}$$
$$s_{\rm CMB} = J(T_{\rm CMB})\left[\exp\left(-2\Psi_1 - \Psi_2\right)\right] - J(T_{\rm CMB}), \tag{12}$$

where the $J(T_{\rm ex})$ and Ψ terms depend on the line l, and the Ψ terms in addition depends on the velocity V, but $J(T_{\rm CMB})$ only

Table 3. References of the rate coefficients used in this article to compute the collisional excitation.

Species	Collider	Database	Reference
^{12}CO	H_2	LAMDA [(1)]	Yang et al. (2010)
^{13}CO	H_2	LAMDA	Yang et al. (2010)
C^{18}O	H_2	LAMDA	Yang et al. (2010)
HCO$^+$	H_2	EMAA [(2)]	Denis-Alpizar et al. (2020)
HCO$^+$	e^-	EMAA	Fuente et al. (2008)
H^{13}CO$^+$	H_2	EMAA	Denis-Alpizar et al. (2020)
H^{13}CO$^+$	e^-	EMAA	Fuente et al. (2008)

Notes. [(1)]Leiden Atomic and Molecular Database (LAMDA) at https://home.strw.leidenuniv.nl/~moldata/. [(2)]Excitation of Molecules and Atoms for Astrophysics (EMAA) at https://emaa.osug.fr.

slowly depends on the frequency. In these equations, we used the symmetry of the slabs to set $\Psi_1 = \Psi_3$ (see Fig. 2). Moreover, the contribution from the CMB has a different expression as compared to the other layers. The CMB is a continuum emission from a source located at an "infinite" distance from the observer. This CMB emission is thus absorbed by the inner and outer layers in front of it (i.e., $J(T_{\rm CMB})\,[\exp(-2\Psi_1 - \Psi_2)]$). However, the calibrated spectrum obtained with the IRAM-30m radiotelescope results from the subtraction of the observation toward a reference LoS devoid of signal coming from the source (named OFF-position) from the observation toward the ON-source LoS. Assuming that the CMB emission is homogeneous and isotropic on the sky, the signal from the OFF-source LoS only contains the CMB emission. Furthermore, we assume that the CMB along the OFF-source LoS is not affected by any absorption. The ON-OFF process thus effectively subtracts $J(T_{\rm CMB})$ from the observed line signal. We fold this into our model by subtracting $J(T_{\rm CMB})$ from the CMB contribution to the modeled line emission (i.e., $s_{\rm CMB}$ in Eq. (12)). Finally, we use as CMB temperature $T_{\rm CMB} \simeq 2.73$ K (Mather et al. 1994). Appendix B generalizes this formalism to an arbitrary number of layers.

3.2. Using the RADEX code to compute the emission and opacity of each layer

We used the non-LTE RADEX code (van der Tak et al. 2007) to compute the excitation temperature ($T_{\rm ex}$) and opacity (τ) for each layer and each line. RADEX allows one to compute large grids of models in reasonnable times. To do this, RADEX uses a local approach, namely the escape probability radiative transfer method, to derive the populations of the different energy levels of the chosen molecular species. The references of the collisional excitation rate coefficients are listed in Table 3. We assume that the collisional partners are molecular hydrogen for all species, plus electrons for the high-dipole moment species HCO$^+$ and H^{13}CO$^+$(Goldsmith & Kauffmann 2017). The detailed prescriptions for the ortho-to-para ratio of the molecular hydrogen and for the electron density can be found in Sect. 2.1 of Roueff et al. (2024).

We have chosen to use the escape probability for a static uniform sphere. A spherical geometry appears better suited than a plane parallel geometry, because it allows photons to escape in all directions, and molecular clouds have a complex self-similar structure. In the chosen multilayer model, the velocity field is independently fitted in each layer and the layers are not radiatively coupled (i.e., the radiation field of a layer does not contribute to the level population of the gas in another layer). Therefore, we have chosen to consider each layer as a uniform medium. Using the expanding sphere geometry is less appropriate because this physical description requires no discontinuity for the centroid velocity between the layers. The expanding sphere model also implicitly assumes that the outer layer has a larger velocity dispersion than the inner layer because the expansion velocity field reaches higher values further away from the center. For a single-layer model, the influence of the choice of the escape probability law has been discussed by van der Tak et al. (2007). At low column densities, the uniform sphere model and the expanding spherical model lead to similar emergent intensities within about 10% while the plane parallel slab predicts higher line opacities and intensities. The difference is more significant at high column densities. This means that the choice of the escape probability law impacts the estimation of the physical conditions in a systematic way, but the difference between the estimations remain moderate (less than 50% for the density) for lines with low to moderate opacities (less than ∼10). Anyway, the proposed method can be used with another choice of the escape probability, depending on the science context.

3.3. Estimated input parameters

For each species and each layer, the physical conditions are characterized by five unknown parameters: the kinetic temperature ($T_{\rm kin}$), the volume density of molecular hydrogen ($n_{\rm H_2}$), the column density (N) of the considered species, the velocity dispersion (σ_V) expressed as a full width at half maximum (FWHM), and the centroid velocity (C_V). Since the foreground and background layers are part of the same outer layer, they are characterized by the same five parameters. For each species, the sandwich model is therefore characterized by ten unknown parameters,

$$\begin{aligned}\boldsymbol{\theta} = \{\, &\log T_{\rm kin,\,ou}, \log n_{\rm H_2,\,ou}, \log N_{\rm ou}, {\rm FWHM}_{\rm ou}, C_{V,\,\rm ou},\\ &\log T_{\rm kin,\,in}, \log n_{\rm H_2,\,in}, \log N_{\rm in}, {\rm FWHM}_{\rm in}, C_{V,\,\rm in}\}. \end{aligned} \quad (13)$$

We assumed that the five species are co-located inside each layer. The species thus share the same $T_{\rm kin}$, $n_{\rm H_2}$, C_V, and FWHM per layer.

Constraining the column density of each species in each layer is complex because we only have one or two transitions per species. On the one hand, estimating all these column densities simultaneously is inefficient because there is not enough information to get accurate estimations. On the other hand, Roueff et al. (2024) show that fixing the ratio of column densities can bias the estimation of both the column densities and the other physical parameters such as the volume density when the injected constraints are invalid. To work around this issue, we discuss in Sect. 4 the choices of chemical assumptions made to decrease the number of fitted parameters.

3.4. Noise model

A noise model is required 1) to quantify the precision of the estimates, and 2) to derive the maximum likelihood estimator (MLE), which is used to fit the data. We design two slightly different noise models to meet the specific needs of each goal. From the first perspective, we take all noise sources into account in order to avoid underestimating the achievable precision (i.e., estimating too small error bars). In particular, our model takes

into account not only the thermal additive noise, but also a calibration multiplicative noise that has a significant effect at high S/N (see Einig et al. 2023). However, when deriving the MLE, we omit the calibration multiplicative noise because Roueff et al. (2024) observe that taking it into account leads to correct results on Monte Carlo simulations, but poorer and inconsistent fits on actual data. They suggest that, in the case of real data, this additional flexibility of the fit is used to deal not only with calibration errors but also with model misspecifications, the latter effect being more significant than the former one. We now detail the full noise model used to quantify the precision on estimates. An observed spectrum $\boldsymbol{x}_l$ for line l is a realization of the signal (noted $\boldsymbol{s}_l$) from Eq. (8) regularly sampled in velocity affected by an additive thermal noise $\boldsymbol{b}_l$ and a multiplicative calibration noise c

$$\boldsymbol{x}_l = c\,\boldsymbol{s}_l + \boldsymbol{b}_l. \tag{14}$$

The additive noise $\boldsymbol{b}_l$ is drawn from a centered Gaussian distribution with standard deviation $\sigma_{b,l}$, whose typical values are listed in Table 1. The calibration noise is drawn from a Gaussian distribution of mean 1.0 and standard deviation $\sigma_{c,l}$ with

$$\sigma_{c,l} = \begin{cases} 5\% & \text{for 3 mm lines,} \\ 10\% & \text{for 1 mm lines.} \end{cases} \tag{15}$$

The value quoted at 3 mm was estimated on the ORION-B dataset in Appendix D of Einig et al. (2023).

3.5. Saturation of the S/N during the fit

Our proposed estimate $\widehat{\boldsymbol{\theta}}$ of the vector of unknowns $\boldsymbol{\theta}$ is obtained by minimizing the negative log-likelihood (NLL):

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\theta} \left[-\log \mathcal{L}(\boldsymbol{\theta}) \right], \tag{16}$$

where

$$-\log \mathcal{L}(\boldsymbol{\theta}) = \frac{1}{2} \sum_l \frac{\|\boldsymbol{x}_l - \boldsymbol{s}_l\|^2}{\sigma_{b,l}^2} + \text{constant}. \tag{17}$$

In this equation, $\boldsymbol{x}_l$ and $\boldsymbol{s}_l$ are the observed and modeled spectra for line l, and $\sigma_{b,l}^2$ is the associated noise level. The NLL, $-\log \mathcal{L}$, is a function of the physical input parameters $\boldsymbol{\theta}$, and we search for the value $\widehat{\boldsymbol{\theta}}$ that minimizes it. Following Roueff et al. (2024), we only take into account the additive thermal noise in this expression. Indeed, implementing a MLE that takes into account the calibration noise leads to better performances on Monte Carlo simulations, especially when the S/N is high. However, we observe that it results on actual data in artificially large calibration factors to compensate for the oversimplified model assumptions, and thus in unrealistic estimations of the physical and chemical parameters (see Sect. 3.4).

We also saturate the peak S/N of all lines to a value of ten by adjusting the $\sigma_{b,l}$ value in the likelihood function when needed. This saturation ensures that the HCO^+ isotopologues have similar weights in the NLL as the CO isotopologues and thus significantly contribute to the estimation of the parameters. This saturation is needed because the use of broadband receivers means that the integration time is the same for all lines, resulting in different S/N for lines of different intensities. We are thus limited by the line that has the minimum S/N, that is the $H^{13}CO^+$ (1 − 0) line in our case. Figure 3 compares some observed spectra with their saturated noise versions. This figure illustrates the potential effect of the S/N saturation on the line profiles. We actually just adjust the $\sigma_{b,l}$ value during the fit without changing the data themselves. In each panel, the top left and right numbers give the relative weight of a difference of 1 K in any given channel, that is,

$$\left(\frac{1\,\mathrm{K}}{\sigma_{b,l}}\right)^2, \tag{18}$$

on the fit χ^2 value before and after the S/N saturation, respectively. Before the saturation (left numbers), the CO isotopologues completely dominate the variations of the χ^2. In contrast, after saturation (right numbers), the HCO^+ isotopologues can significantly influence the variations of the χ^2.

3.6. Constrained optimization

Compared to Roueff et al. (2024), the challenge in this study is to estimate about twice as many parameters by going from a homogeneous model to a heterogeneous one. To deal with this challenge, our optimizer starts with a grid search followed by a gradient descent algorithm. The grid of RADEX models is computed once and for all, while RADEX is run on the fly to evaluate the line models during the gradient descent. In addition, we chose to optimize the likelihood under the constraint that the column density ratios can vary within some intervals based on a priori astrophysical knowledge. In our implementation, the column density ratios are estimated within these constraints on abundance ratios during the grid search with a finite resolution of 0.05 in the logarithm of the abundance ratio (that is a factor of 1.12 in the abundance ratio). The column density ratios are then fixed during the gradient descent to the optimum value found during the grid search. This choice allows us to regularize our gradient, and thus to avoid convergence issues that occur when trying to apply a gradient descent to refine the estimations of the column density ratios. Indeed, the more unknowns, the higher the probability of obtaining singular Fisher matrices. Appendix C details this optimization process.

3.7. Removing inaccurate estimations with the Cramér-Rao bound

Once the fit has converged (i.e., after the grid search and the gradient descent), we filter out estimations with too large $[-\sigma, +\sigma]$ confidence intervals. It happens that the square root of the Cramér-Rao bound (CRB) provides a lower bound on the standard deviation of estimated parameters for any unbiased estimators (Garthwaite et al. 1995). Mathematically speaking,

$$\sigma\left(\widehat{\theta_i}\right) \geq \left[\mathrm{CRB}(\boldsymbol{\theta})\right]_{ii}^{1/2} \quad \text{with} \quad \mathrm{CRB}(\boldsymbol{\theta}) = \boldsymbol{I}_F^{-1}(\boldsymbol{\theta}), \tag{19}$$

where $\boldsymbol{I}_F$ is the Fisher information matrix defined by

$$\forall (i,j) \quad [\boldsymbol{I}_F(\boldsymbol{\theta})]_{ij} = \mathbb{E}\left[\frac{\partial \log \mathcal{L}(\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \log \mathcal{L}(\boldsymbol{\theta})}{\partial \theta_j} \right]. \tag{20}$$

In the previous equation, $\mathbb{E}$ is the statistical expectation among the different realizations of $\boldsymbol{x}$, defined in Eq. (14). As in Roueff et al. (2024), and since the true value of $\boldsymbol{\theta}$ is unknown, we compute $\mathrm{CRB}(\widehat{\boldsymbol{\theta}})$ instead of $\mathrm{CRB}(\boldsymbol{\theta})$ to provide reference precisions

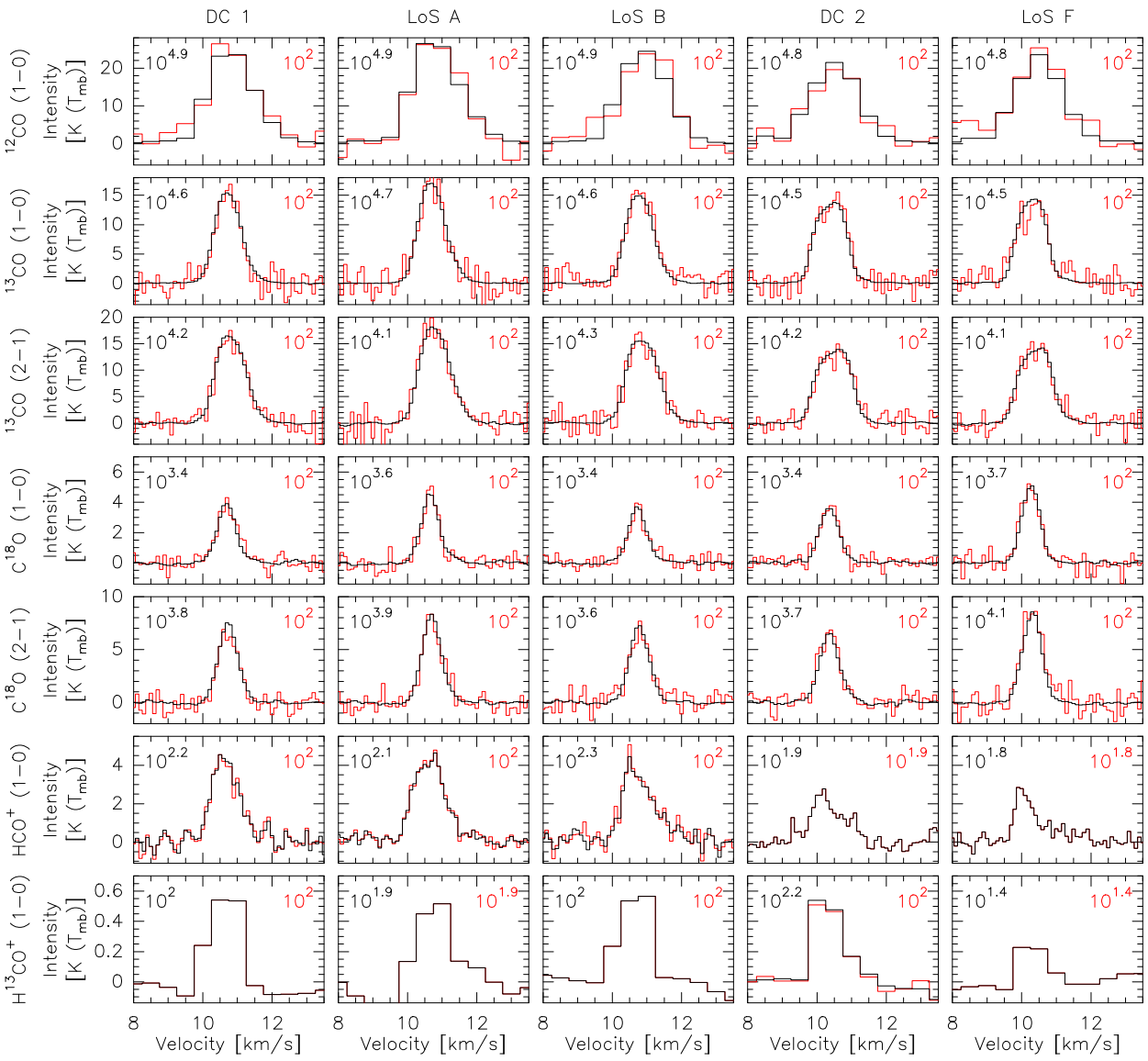


Fig. 3. Comparison of five observed spectra (black histograms) and the same ones plus the required amount of noise to saturate their peak S/N to ten (red histograms). The positions of the spectra are marked on Fig. 1 and listed in Table 2. Each row corresponds to a given species and transition, while each column corresponds to a given position. In each panel, the top left and right numbers give the relative weight of a difference of 1 K in any given channel on the fit χ^2 value before and after the S/N saturation, respectively. The ^{12}CO (1 − 0) and H^{13}CO$^+$ (1 − 0) lines were observed with a resolution about 5 times coarser than the other lines. When the peak S/N is lower than 10, the red spectrum is hidden under the black one.

of our estimations. Therefore, if $\widehat{\theta_i}$ is the estimate of $\theta_i \in \boldsymbol{\theta}$ (see Eq. (13)), we only retain $\widehat{\theta_i}$ for LoSs that comply with

$$\begin{aligned}
&\mathrm{CRB}\left(\widehat{C_V}\right) \leq 0.1 \text{ km s}^{-1},\\
&\mathrm{CRB}^{1/2}\left(\widehat{\mathrm{FWHM}}\right) \leq 0.1 \text{ km s}^{-1},\\
&\mathrm{CRB}^{1/2}\left(\widehat{\log T_{\mathrm{kin}}}\right) \leq 0.1,\\
&\mathrm{CRB}^{1/2}\left(\widehat{\log N}\right) \leq 0.3,\\
&\mathrm{CRB}^{1/2}\left(\widehat{\log n_{\mathrm{H_2}}}\right) \leq 2.0,\\
&\mathrm{CRB}^{1/2}\left(\widehat{\log P_{\mathrm{ther}}}\right) \leq 2.0,
\end{aligned} \tag{21}$$

where P_{ther} is the thermal pressure. When computing these CRB reference precisions, we take into account both the thermal and calibration noises to avoid underestimating the credibility interval (see Sect. 3.4). In the previous section, we have noted that some abundance ratios are only estimated during the grid search within astrophysically chosen fixed bounds, in order to stabilize the optimization problem. As the estimated abundance ratios are constrained within 12%, we simply neglect the increase of uncertainty due to this procedure. We therefore only have access to the accuracy on the estimation of one column density per layer with a Fisher information matrix of size 10×10, and we have no access to the accuracy on the estimation of the ratio of the column densities.

The gradients required to calculate the Fisher matrix and the CRB for any multilayer model, including the sandwich model, are given in Appendix D.

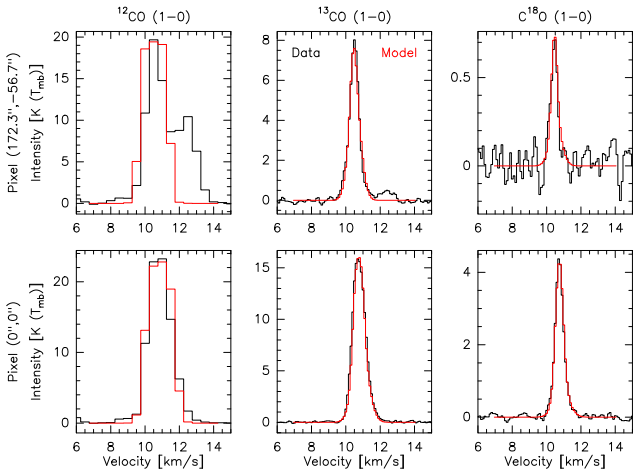


Fig. 4. Comparison of the observed (black) and modeled (red) $(1-0)$ line of the ^{12}CO, ^{13}CO, and C^{18}O species for two different LoSs.

4. Physical and chemical assumptions

Having a large number of fitted parameters leads to large uncertainties. To reduce these uncertainties, we need to add constraints on estimated parameters with physical and chemical assumptions. Roueff et al. (2024) discuss in depth the possible occurrence of biases linked to these additional assumptions, and highlight that a single unreasonable constraint can dramatically bias the results. This section details and justifies the proposed hypotheses. First, we explain how we use the information carried by the ^{12}CO $(1-0)$ line of high opacity. Then, we compare the sandwich model with a simple homogeneous model consisting of a single layer, to illustrate the gain in signal reconstruction quality with the latter model. We continue by discussing the differences in the kinematic properties of the different layers, namely their velocity centroids and velocity dispersions. Finally, we introduce the hypotheses on the chemical abundances in the inner and outer layers.

4.1. Specific treatment for the ^{12}CO (1–0) line

Figure 4 compares the observed and modeled spectra of the $(1-0)$ lines of the three main CO isotopologues for two lines of sight. The main difference between these LoSs is the existence of a second velocity component between 12 and 14 km s^{-1} that contributes more than 10% of the total integrated intensity of the ^{12}CO $(1-0)$ line and much less than this for the ^{13}CO and C^{18}O $(1-0)$ lines. While the model does a correct reconstruction of the main velocity component for the ^{12}CO $(1-0)$ line when we only fit the peak intensity of this line, the second velocity component perturbs the fit when we try to fit the full ^{12}CO $(1-0)$ profile. Since at this stage our algorithm has not been designed to fit multiple velocity components, we thus fit only the ^{12}CO $(1-0)$ peak temperature as proposed by Roueff et al. (2024). Fitting multiple velocity components will be the subject of a forthcoming paper.

4.2. On the importance of a heterogeneous model

The simplest possible geometry is to assume a homogeneous cloud composed of a single layer emitting in front of the CMB. However, the maps of the line emission (see Fig. 1) suggest that the emission comes from gas with different physical conditions, at least along the LoSs toward filaments and dense cores. We therefore introduce a three-layer model to take into account this inhomogeneity while keeping the model as simple as possible.

In this section we compare the integrated intensity computed on the data with that derived from the molecular line fitting with two different models. The first one is a homogeneous model described with

$$s = s_{\rm fore} + s_{\rm CMB}. \tag{22}$$

The second one is a three-layer model described with

$$s = s_{\rm fore} + s_{\rm inner} + s_{\rm back} + s_{\rm CMB}. \tag{23}$$

The components used in these equations are defined in Eq. (9) to (12).

To simplify the comparison, we kept in each experiment the same fixed relative abundances, defined by using the following isotopic ratios for Orion ^{12}C/^{13}C= 63, ^{16}O/ ^{18}O= 510, and N(HCO$^+$)/N(H$_2$)=3 × 10^{-9} and $N(^{13}{\rm CO})/N({\rm H_2}) = 2.3 \times 10^{-6}$ (Gerin et al. 2019; Roueff et al. 2024, and references therein). Those give

$$\begin{aligned} &N(^{12}{\rm CO})/N(^{13}{\rm CO}) = 10^{1.8} \sim 63,\\ &N(^{13}{\rm CO})/N({\rm C}^{18}{\rm O}) = 10^{0.9} \sim 8,\\ &N(^{13}{\rm CO})/N({\rm HCO}^+) = 10^{2.88} \sim 800,\\ &N({\rm C}^{18}{\rm O})/N({\rm H}^{13}{\rm CO}^+) = 10^{3.78} \sim 6000. \end{aligned} \tag{24}$$

We refer to these column density ratios collectively as "basic" in the remainder of this work. In contrast, the centroid velocities, velocity dispersions, column densities, kinetic temperatures, and volume densities were estimated independently for the inner and outer layers.

Figure 5 compares the reconstruction of all the line profiles for the best fits obtained with either the homogeneous or the sandwich model. The total integrated line is plotted as a function of the visual extinction between 10 and 27 mag, where the probability of having different gas components along a single LoS is high. Intensities are averaged over all pixels that belong to a 1 mag sliding window of the visual extinction. The colored areas around the relative error curves show the associated standard deviations for all the pixels belonging to the sliding window. The vertical width of this $\pm\sigma$-interval mostly decreases from low to high visual extinction because the number of pixels quickly decreases with A_V.

The first obvious feature in Fig. 5 is that the sandwich model better reconstructs the line total integrated intensities than the homogeneous model. This is expected since the number of fitted parameters doubles from 5 to 10 when going from a homogeneous to a sandwich model. More interestingly, the sandwich model better reconstructs the lines that mostly originate from one type of medium with specific physical conditions: either translucent or dense and cold gas. The most striking example is the H^{13}CO$^+$ line whose relative error decreases from $\sim -80\%$ to less than -25% for $A_V > 15$ mag. The second example is the C^{18}O $(2-1)$ line whose typical relative error decreases from $\sim -20\%$ to less than $\sim 10\%$ over the plotted A_V range. The second clear feature is that the sandwich model allows to decrease the intensities of the lines that are known to be optically thick (namely ^{12}CO $(1-0)$ and HCO$^+$ $(1-0)$) and thus mostly sensitive to the external layer, while it increases the intensity of the optically thinnest line (e.g., H^{13}CO$^+$ $(1-0)$). Finally, the sandwich model becomes unbiased (i.e., its relative error is compatible with zero within the $\pm\sigma$-interval) for large ranges

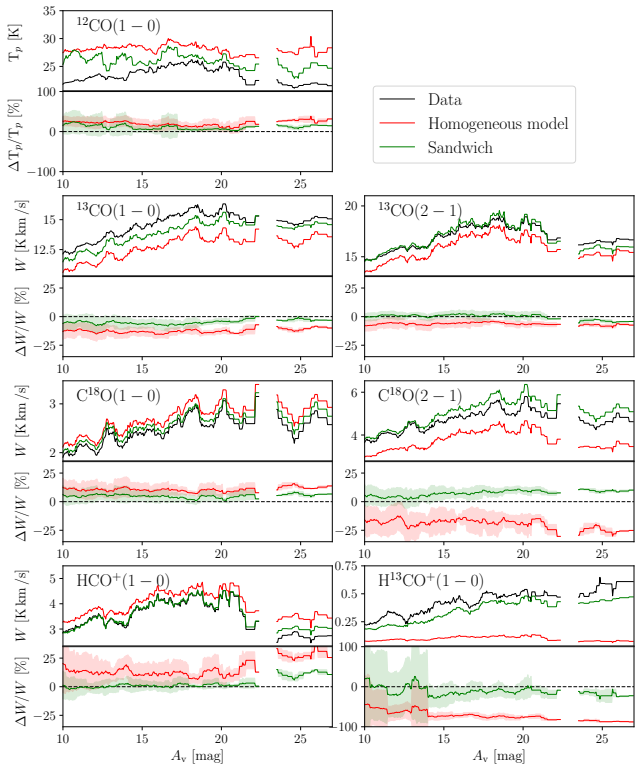


Fig. 5. Comparison of the reconstruction of all integrated line intensities for the best fit obtained with either a homogeneous (in red) or a simple sandwich model (in green). The integrated line intensity or peak temperature are plotted as a function of the visual extinction (A_V). More precisely, the integrated line intensities (zero-order moment) or peak temperature are averaged over all the pixels whose visual extinction belongs to the $[A_V - 0.5, A_V + 0.5]$ mag interval, where A_V is plotted on the x-axis of the panels. The lack of values for $A_V \in [23, 24]$ mag results from the fact that there is no pixel in this range of A_V over the studied field of view. In each panel, the top one shows the line integrated intensity or peak intensity, while the bottom one shows the reconstruction relative error in percentage. The colored areas on the relative error panels show the standard deviation over all the pixels that belong to the 1 mag-A_V sliding window. The horizontal dashed line indicates a perfect reconstruction. While the simple sandwich model already better fit the data than the homogeneous model, the reconstruction will still be improved by refining the assumptions.

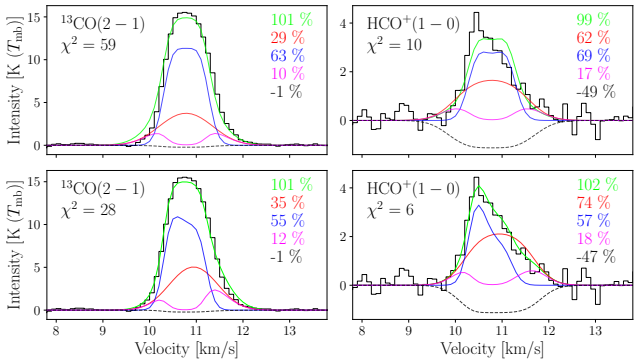


Fig. 6. Comparison of the data and the fit of the sandwich model when either the centroid velocities of all layers are assumed to be identical (top row) or they can have different values in the inner and outer layers (bottom row). The left and right columns show this comparison for the ^{13}CO (2 − 1) and HCO$^+$ (1 − 0) lines, respectively. These spectra correspond to the LoS B (see Table 2). Data and fitted spectra are shown in black and green lines, respectively. The contributions of the different layers are shown as colored lines: red for the foreground layer (see Eq. (9)), blue for the inner layer (see Eq. (10)), pink for the background layer (see Eq. (11)), and dashed black for the CMB (see Eq. (12)).

of A_V and for many lines. The most obvious example is the ^{13}CO (2 − 1) line. Whatever the A_V range, the absolute value of the relative reconstruction error almost always remains below 10%, a value consistent with the fact that we saturate the peak S/N to 10. Hereafter, we only discuss variants of the sandwich model.

4.3. On the importance of having independent centroid velocities to reproduce the spectrum asymmetries

A straightforward way to decrease the number of fitted parameters would be to use the same velocity distribution (centroid velocity and velocity dispersion) for the inner and outer layers. However, it is well known that the lines mostly tracing the dense and shielded gas (e.g., the C^{18}O lines) are narrower than lines (e.g., the ^{12}CO lines) probing the surrounding lower density, warmer, and more turbulent gas. This result is for example highlighted by Orkisz et al. (2019); Roueff et al. (2021); Gaudel et al. (2023). We here explore the importance of estimating independently the centroid velocities for the inner and outer layers. To do this, we compare the reconstruction of the spectra for the LoS B that shows asymmetric line profiles for, e.g., the HCO$^+$ (1 − 0) line (see Table 2 and Fig. 3).

Figure 6 shows the fit of the ^{13}CO (2 − 1) and HCO$^+$ (1 − 0) lines for a sandwich model with the same centroid velocity for all the layers (top row) or with independent centroid velocities for the inner and outer layers (bottom row). The conditions for the fit of the other parameters are unchanged between these two experiments. In particular, the centroid velocity for each layer is identical for the seven fitted lines, and the column density ratios are fixed to their basic values listed in Eq. (24).

Fitting independently the centroid velocities of the inner and outer layers better reproduces the wings of the ^{13}CO (2 − 1) line, explaining the decrease by more than a factor of two of the χ^2 value for this line. But, more importantly, it is the only way to reproduce the large asymmetry of the HCO$^+$ (1 − 0) line. Indeed, when using the same centroid velocity for all layers, the contribution of each layer shows a symmetrical profile. In contrast, allowing the simultaneous fit of the outer and inner centroid velocities delivers a symmetrical profile contribution of the foreground layer (red spectrum), and asymmetrical ones for all other layers (inner, background, and CMB). These asymmetrical line profiles are not just the result of adding symmetrical components shifted in velocity. Instead, marked asymmetries mainly result from the absorption of the intrinsically symmetrical emission term (J in Eq. (3)) of one layer by the velocity-shifted absorption of another layer in front of it. Taking into account different centroid velocities for the inner and outer layers therefore allowed us to model this coupling between emission and absorption.

Table 4 compares the fitted parameters for the inner and outer layers for these two experiments. The differences are of the order of 10%, except for the volume densities of the inner and outer layers, which increase by 0.05 to 0.1 dex (12 to 25%). When comparing the homogeneous model to the sandwich model in Sect. 4.2, the centroid velocities of the inner and outer layers are fitted independently. This is probably a reason why the

Table 4. Fitted parameters when assuming that the inner and outer layers have the same or different centroid velocities and velocity dispersions.

Parameter	Unit	$C_{V,\mathrm{in}} = C_{V,\mathrm{ou}}$		Free C_V	
		Inner	Outer	Inner	Outer
C_V	km s^{-1}	10.78 ± 0.01	10.78 ± 0.01	10.76 ± 0.01	10.94 ± 0.03
FWHM	km s^{-1}	0.59 ± 0.03	1.22 ± 0.11	0.54 ± 0.02	1.10 ± 0.08
T_{kin}	K	22 ± 3	30 ± 4	21 ± 2	30 ± 3
$\log(n_{\mathrm{H_2}}/\mathrm{cm^{-3}})$	–	5.36 ± 0.41	3.78 ± 0.14	5.41 ± 0.41	3.93 ± 0.08
$\log(N_{^{13}\mathrm{CO}}/\mathrm{cm^{-2}})$	–	16.36 ± 0.03	15.45 ± 0.06	16.33 ± 0.03	15.54 ± 0.05

HCO^+ (1 − 0) integrated line profile is much better reconstructed with the sandwich model than with the homogeneous model.

Hereafter, we only discuss models where the centroid velocities for the inner and outer layers are independently estimated.

4.4. On the importance of fitting the N($C^{18}O$)/N(^{13}CO) column density ratio

Roueff et al. (2024) show that it is in principle better to avoid constraining the column densities of the studied species because this lowers potential biases. However, fitting here one column density per species and per layer would imply to estimate ten additional parameters (five species for the inner, and five for the outer layers). Such a fit would lead to large uncertainties because of the limited number of measured transitions for the different species. We therefore search for the most suitable chemical hypotheses for the Horsehead nebula to limit the number of fitted column densities.

We start from the experiment described in Sect. 4.2, namely, fitting a sandwich model with column density ratios following the basic abundance ratios given in Eq. (24). While the relative error on the integrated intensity is within ±10% for the ^{13}CO and C^{18}O lines for $A_V > 10$ mag, the fit almost systematically underestimates the integrated intensity of the ^{13}CO lines, and overestimates that of the C^{18}O lines, i.e., the relative errors are anticorrelated for these two species (see the green curves of Fig. 5). This behavior that is more salient for the (1 − 0) lines, questions the value chosen for the basic column density ratio. We therefore carried out a series of experiments in which we refitted the spectra over the whole field of view for several values of the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio. For each pixel, we thus obtained 14 solutions for which the only difference is the value of the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio that goes from 5.62 to 25.12 in multiplicative steps of 1.12[1]. In all of these experiments, we used the same value of the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio for both the inner and outer layers. As the solutions are obtained with the same data and same number of fitted parameters, comparing the NLL values (Eq. (17)) allowed us to select the best value of the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratios for each pixel. As this manual approach is akin to a fit of the column density ratio, we incorporated it in the grid search step of the proposed MLE estimator for the remainder of this work.

Figure 7 compares the reconstruction of the ^{13}CO and C^{18}O line integrated intensities obtained with either a fixed basic $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio for all pixels, or an estimated $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio per pixel. The reconstructed integrated intensities are shown as a function of the visual extinction in the [4, 13] magnitude range, where we expect deviations of the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio from its basic value because of the selective photodissociation of C^{18}O combined with enrichment

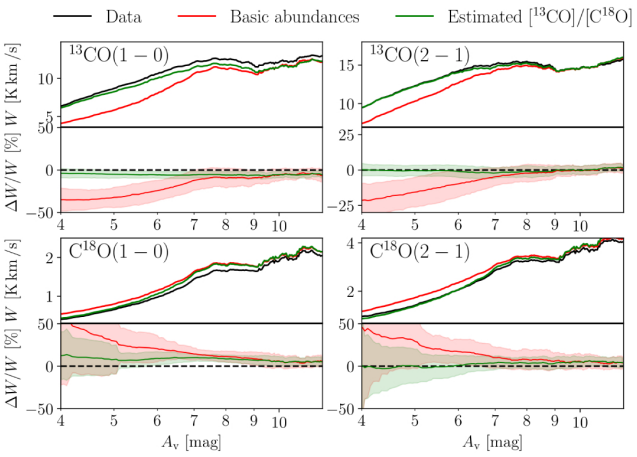


Fig. 7. Comparison of the reconstruction of the ^{13}CO and C^{18}O line integrated intensities for the sandwich model fit obtained with either a fixed relative abundance of [^{13}CO]/[C^{18}O] (in red) or an estimated relative abundance of this ratio (in green). The figure layout is similar to Fig. 5.

of ^{13}CO by isotopic fractionation (Langer et al. 1984; Kong et al. 2015). We do not show the results for lower A_V values because the C^{18}O lines have either very low S/N or are not detected at all, implying that the data are no longer sensitive to the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio. We find that fitting the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ ratio considerably decreases the negative (resp. positive) bias when reconstructing the ^{13}CO (resp. C^{18}O) integrated line intensity. In particular, the (2 − 1) lines now have unbiased fitted integrated intensities.

Hereafter, we continue to fit the $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ column density ratio.

4.5. On the importance of allowing the N(HCO^+)/N(^{13}CO) column density ratio to vary in the outer translucent layer

In this subsection we first show that fixing the value of the $N(\mathrm{HCO}^+)/N(^{13}\mathrm{CO})$ column density ratio to its basic value listed in Eq. (24) leads to an ambiguity in the estimation of the volume density of the inner layer because the NLL shows two local minima of similar depth around 10^4 and 10^6 cm^{-3}. The estimator thus oscillates between these two values depending on the noise level. We therefore considered fitting the column densities of the HCO^+ isotopologues in addition to the column densities of the CO isotopologues (see Sect. 4.4). However, fitting the column densities of both HCO^+ and $H^{13}CO^+$ for all lines of sight will result in a high variance because of the higher number of unknowns and relatively low S/N of the HCO^+ isotopologue lines. As a compromise, we thus propose 1) to take $N(\mathrm{HCO}^+)/N(^{12}\mathrm{CO}) = N(\mathrm{H}^{13}\mathrm{CO}^+)/N(^{13}\mathrm{CO})$ for the outer layer, as this is plausible for translucent gas following the analysis described in Sect. 4.5.2, and 2) to keep the basic abundance ratios in the inner layer made of colder, denser gas.

4.5.1. On the impact of the N(HCO^+)/N(^{13}CO) value on the number of local minima of the NLL

The model described in Sect. 4.4 fits the column densities of ^{13}CO and C^{18}O, but it fixes the column densities of ^{12}CO, HCO^+, and $H^{13}CO^+$ by assuming that they have constant abundances relative to ^{13}CO. Fitting all the lines of sight toward the Horsehead nebula with this model shows that the fit oscillates

[1] The basic value of this ratio is 7.94 (see Eq. (24)).

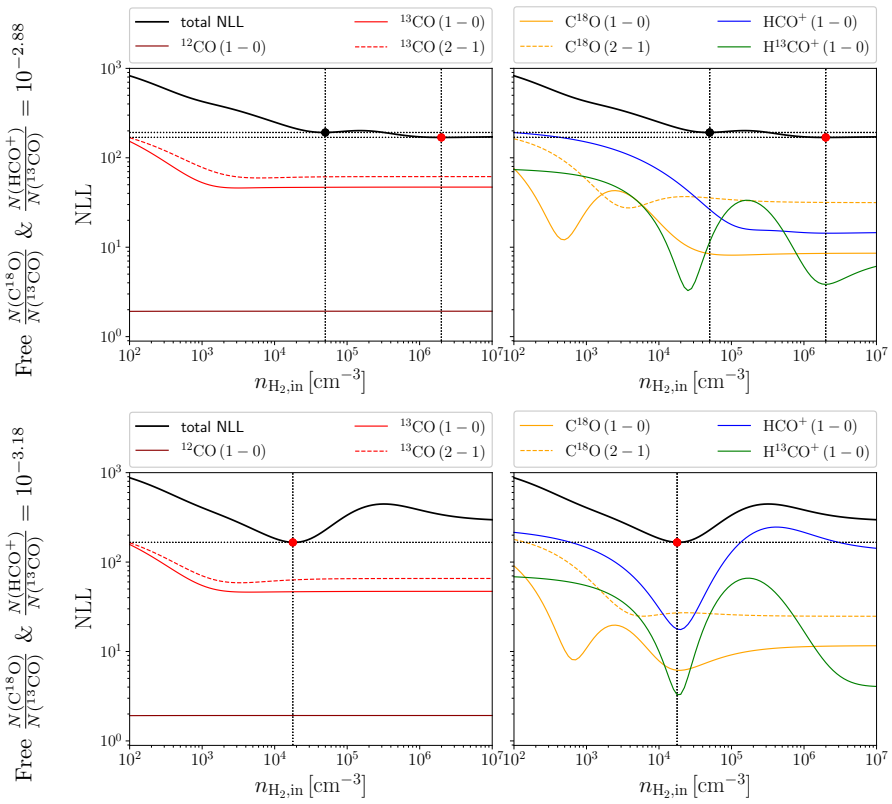


Fig. 8. Comparison of the variations of the NLL as a function of the volume density of the inner layer for the LoS A. We used a sandwich model where $N(^{13}CO)$ and $N(C^{18}O)$ were fitted, while the column densities of ^{12}CO, HCO^+, and $H^{13}CO^+$ were computed assuming fixed abundances relative to ^{13}CO. The top and bottom rows show the NLL variations for two values of $N(HCO^+)/N(^{13}CO)$ varying by a factor of 2.0. The left panel shows the NLL variations for the ^{12}CO and ^{13}CO lines, while the right panel show that of the $C^{18}O$, HCO^+, and $H^{13}CO^+$ lines. In both panels, the black plain lines shows the variations of the NLL when fitting all the lines. The top row shows two local minima at similar NLL values while the bottom row only shows one local minimum. The vertical and horizontal black dotted lines intersect at the lowest NLL value, indicated by the red circle, and at the second local minimum that is indicated by the black circle, when it exists.

between two solutions for the inner layers: one with a density around 10^4 cm^{-3}, and another one with the density around 10^6 cm^{-3}. This situation often happens when fitting low-J lines of CO and HCO^+ isotopologues (see, for instance, Fig. 10 Roueff et al. 2024). As previous studies of the Horsehead nebula suggest that the volume density should be about $\sim 2 \times 10^5$ cm^{-3} (Pety et al. 2005; Habart et al. 2005; Goicoechea et al. 2006; Pety et al. 2007), it is tempting to limit the range of volume densities to this upper value, but it is better to solve this issue based on chemical assumptions.

To understand why the fit fluctuates between two values of the volume density separated by two orders of magnitude, we have selected one line of sight for which the volume density of the inner layer is fitted to 10^6 cm^{-3}, namely LoS A. The top row of Fig. 8 shows the variations of the NLL as a function of the volume density of the inner layer, while all other estimated parameters remain fixed at the values found during the fit. The minimum of the NLL for this LoS happens at 2×10^6 cm^{-3}, but a second local minimum occurs at 5×10^4 cm^{-3} with only an increase of the NLL of 14%. The NLL of all optically thick (^{12}CO, ^{13}CO, and HCO^+) lines show a decrease followed by a plateau when the volume density increases. A plateau implies that the line intensity is no longer sensitive to the volume density of the inner layer. The NLL of the two optically thin lines ($C^{18}O$ and $H^{13}CO^+$) show one or two more or less deep wells at various relatively low $n_{H_2,in}$ values followed by a plateau at higher $n_{H_2,in}$ values. Only the (1 − 0) lines of HCO^+ and $H^{13}CO^+$ have not yet reached the NLL plateau at $n_{H_2,in} \sim 10^4$ cm^{-3}. The variations of the NLL of these two lines above 10^4 cm^{-3} thus discriminate between the fitted $n_{H_2,in}$ values. This is consistent with the higher critical density of the HCO^+ isotopologue lines ($\sim 10^{5\pm0.3}$ cm^{-3}) than those of the CO isotopologue lines ($\sim 10^{3.3\pm0.2}$ cm^{-3}, see Tables 6 and 7 of Roueff et al. 2024).

Next, we decreased the abundance of HCO^+ with respect to ^{13}CO by a factor of two. The bottom row of Fig. 8 shows the variations of the NLL as a function of the volume density in this case. The NLL variations are significant only for the HCO^+ (1 − 0) line, which now shows a deep well aligned with the much more optically thin $H^{13}CO^+$ (1 − 0) line at $n_{H_2,in} \sim 1.8 \times 10^4$ cm^{-3}. This translates into a marked well of the total NLL at the same value of $n_{H_2,in}$. For $n_{H_2,in} \gtrsim 10^4$ cm^{-3}, the fitted value of $n_{H_2,in}$ is governed by the variations of the NLL of the (1 − 0) lines of the two main HCO^+ isotopologues. Varying the abundance of HCO^+ relative to ^{13}CO yields or lifts a degeneracy of the solution for $n_{H_2,in}$. It is thus important to fit the column density of HCO^+ and $H^{13}CO^+$ in order to avoid biasing the value of the fitted volume densities. However, fitting the column density of HCO^+ and $H^{13}CO^+$ in both layers increases considerably the number of unknowns to be estimated. We thus turn to chemical considerations to assess whether we can make a more relevant hypothesis than assuming constant abundances of HCO^+ and $H^{13}CO^+$ relative to ^{12}CO and ^{13}CO respectively.

4.5.2. The N(HCO$^+$)/N(^{12}CO)=N(H^{13}CO$^+$)/N(^{13}CO) hypothesis for the outer layer

The outer layer consists primarily of translucent gas with moderate temperatures in the presence of ultraviolet radiation. The forward fractionation reaction involved in the

$$^{13}C^+ + {}^{12}CO \rightleftharpoons {}^{13}CO + {}^{12}C^+ + \Delta E\,(\sim 35\,\mathrm{K}) \qquad (25)$$

equilibrium is moderately efficient in lukewarm gas (Langer 1977). The left-to-right reaction is privileged when the gas kinetic temperature is $\lesssim 35$ K because the right-to-left reaction is endothermic. Moreover, selective photodissociation tends to

Table 5. Assumptions on column density ratios or relative abundances for the basic and improved sandwich models.

Column density ratio		Basic model	Improved model	
		Inner & outer layers	Inner layer	Outer layer
Assumed	$^{12}CO/^{13}CO$	$10^{1.8} \sim 63$	$10^{1.8}$	$[10^{1.3}, 10^{1.9}] \sim [20, 80]$
	$^{13}CO/C^{18}O$	$10^{0.9} \sim 7.9$	$10^{0.9}$	$[10^{0.75}, 10^{1.4}] \sim [5.6, 25.1]$
	$^{13}CO/HCO^+$	$10^{2.88} \sim 759$	$[10^{2.88}, 10^{3.78}]$	$[10^{2.88}, 10^{3.78}]$
	$^{13}CO/H^{13}CO^+$	$10^{4.68}$	$^{12}CO/HCO^+$	$^{12}CO/HCO^+$
Derived	$^{12}CO/C^{18}O$	$10^{2.7} \sim 500$	$10^{2.7}$	$[10^{2.05}, 10^{3.3}]$
	$C^{18}O/H^{13}CO^+$	$10^{3.78}$	$[10^{3.78}, 10^{4.68}]$	$[10^{2.78}, 10^{4.93}]$
	$^{12}CO/HCO^+$	$10^{4.68}$	$[10^{4.68}, 10^{5.58}]$	$[10^{4.18}, 10^{5.68}]$

Notes. We write X/Y the column density ratio $N(X)/N(Y)$, for each couple of species $\{X, Y\}$. For the improved model, intervals correspond of the explored ranges when estimating the corresponding ratios. Moreover, $N(^{13}CO)/N(H^{13}CO^+)$ is deduced from $N(^{13}CO)/N(H^{13}CO+) = N(^{12}CO)/N(HCO+)$ (see Sect. 4.5.2).

destroy more efficiently ^{13}CO and $C^{18}O$ than ^{12}CO. The actual chemical balance between these two opposite trends requires a detailed radiative transfer treatment which can be performed for example by the Meudon PDR code.

We now discuss the main formation and destruction pathways for HCO^+ and $H^{13}CO^+$. Both molecular ions are essentially formed through two reactions that involve $^{12}CO^+$ or $^{13}CO^+$. The first reaction,

$$^{12}CO^+ + H_2 \rightarrow HCO^+ + H,$$
$$^{13}CO^+ + H_2 \rightarrow H^{13}CO^+ + H,$$

has the same reaction rate coefficient for both isotopologues, $k = 7.5 \times 10^{-10}$ cm^3 s^{-1} (Scott et al. 1997). The second reaction also have the same reaction rate coefficient, $k = 1.7 \times 10^{-9}$ cm^3 s^{-1} (computed with the Langevin approximation Gioumousis & Stevenson 1958):

$$^{12}CO + H_3^+ \rightarrow HCO^+ + H_2,$$
$$^{13}CO + H_3^+ \rightarrow H^{13}CO^+ + H_2.$$

In addition, both molecular ions are principally destroyed through dissociative recombination with a reaction rate coefficient of 1.4×10^{-6} cm^3 s^{-1} at 30 K (Hamberg et al. 2014):

$$HCO^+ + e^- \rightarrow {}^{12}CO + H,$$
$$H^{13}CO^+ + e^- \rightarrow {}^{13}CO + H.$$

With an electronic fraction of about 10^{-4}–10^{-6} in translucent environments, as found in Bron et al. (2021), the destruction rate of HCO^+ and $H^{13}CO^+$ is in the $(10^{-10}$–$10^{-12}) \times n_{H_2}$ s^{-1} range.

The HCO^+ and $H^{13}CO^+$ destruction rates via the isotopic exchange reactions (Mladenović & Roueff 2014),

$$^{13}CO + HCO^+ \rightleftharpoons H^{13}CO^+ + {}^{12}CO + \Delta E(= 17.8\,\mathrm{K}), \quad (26)$$

have a significantly smaller value of a few $10^{-14} \times n_{H_2}$ s^{-1}, as the forward reaction rate coefficient is 6.5×10^{-10} cm^3 s^{-1} at 30 K, and the relative abundance of ^{12}CO is a few 10^{-5}. In the conditions where the main formation channel of HCO^+ and $H^{13}CO^+$ is via the reactions of ^{12}CO and ^{13}CO with H_3^+, we obtain the $HCO^+/H^{13}CO^+ = {}^{12}CO/{}^{13}CO$ relation at steady state.

We checked how the 0D models used in Bron et al. (2021), which sample a range of densities and illuminations appropriate to translucent conditions, probe the $N(^{12}CO)/N(HCO^+)$ ratio as a function of the $N(^{13}CO)/N(H^{13}CO^+)$ ratio. We introduced selective photodissociation effects in CO isotopologues by introducing different approximate analytical expressions of the photodissociation rates obtained from Fig. 9 of Visser et al. (2009), in the chemical network, including the various isotopic exchange reactions of carbon and oxygen. Given these assumptions, we have computed 4000 models that uniformly sample the parameter space appropriate to translucent environments as in (Bron et al. 2021), and filtered the simulations that lead to $N(^{13}CO) < 10^{14}$ cm^{-2}. Figure 9 shows the joint histograms of $N(HCO^+)/N(^{12}CO)$ vs $N(H^{13}CO^+)/N(^{13}CO)$ in log-space for the 3224 remaining models. The models sample more than three orders of magnitude for these abundance ratios but the histogram is relatively well approximated by a line of slope one going through the basic abundance ratio listed in Eq. (24).

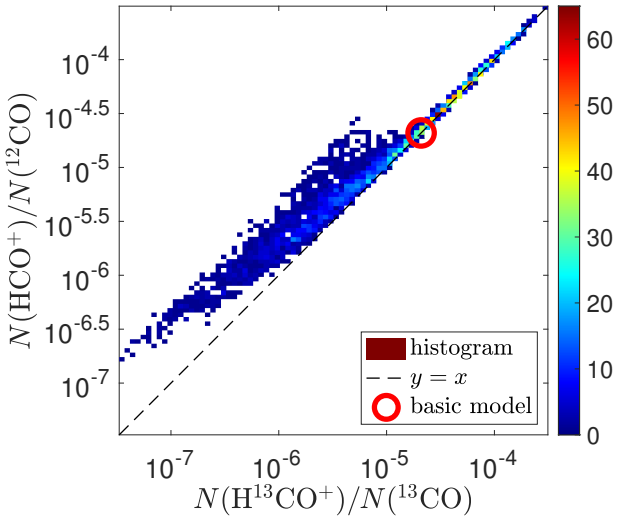


Fig. 9. Joint histogram of $N(HCO^+)/N(^{12}CO)$ versus $N(H^{13}CO^+)/N(^{13}CO)$ in log space for the 3224 chemical models sampling various physical conditions in translucent gas. The red circle shows the ratio derived from the basic abundances listed in Eq. (24). The dashed blue line has a slope of 1.0.

For the sake of clarity, Table 5 compares the chemical assumptions and their implications for column density ratios for the two main chemical models presented so far.

4.5.3. Intermediate summary

In the remainder of this study, we only consider the following hypotheses that we refer to as the "improved" model. We fit the peak intensity of the ^{12}CO (1 − 0) line and the line profiles of all the other lines. We consider a sandwich model (foreground, inner, background, and CMB) in which the foreground and background layers (collectively called the outer layer) share the same physical and chemical properties. In contrast, the physical properties (velocity centroid and velocity dispersion, kinetic temperature, and volume density) are estimated independently for the inner and outer layers. In each layer, all the lines share the same physical and chemical properties, that is, their emissions are considered co-located inside each layer.

From the chemical viewpoint, the inner and outer layers are modeled differently. In the inner dense layer, we estimate the ^{13}CO and HCO$^+$ column densities, and we derive the column densities of the other species by assuming that they scale with the elemental abundances of the corresponding carbon and oxygen isotopes, that is,

$$N(^{12}\mathrm{CO}) = 63\,N(^{13}\mathrm{CO}) \quad (27)$$

$$N(\mathrm{C}^{18}\mathrm{O}) = N(^{13}\mathrm{CO})/7.9, \quad (28)$$

$$N(\mathrm{H}^{13}\mathrm{CO}^+) = N(\mathrm{HCO}^+)/63. \quad (29)$$

In the outer translucent layer, we fit the column density of ^{12}CO, ^{13}CO, C^{18}O, and HCO$^+$, and we deduce the column density of H^{13}CO$^+$ with the constraint

$$\frac{N(\mathrm{H}^{13}\mathrm{CO}^+)}{N(^{13}\mathrm{CO})} = \frac{N(\mathrm{HCO}^+)}{N(^{12}\mathrm{CO})}. \quad (30)$$

We thus have 14 parameters to estimate: four physical parameters ($n_{\mathrm{H_2}}$, T_{kin}, C_V, FWHM) per layer plus four and two column densities for the outer and inner layers, respectively.

5. Results

Our goal in this section is two-fold. We first want to check whether the used hypotheses allow us to estimate chemical and physical parameters that are consistent with previous findings. Second, most of the previous results about the Horsehead pillar were derived with a single-layer model, while the sandwich model allows us to characterize the chemical and physical structure of the gas along the LoS. We first focus on the structure along the LoS (Sect. 5.1 and 5.2): We study whether the integrated intensity of the different lines is dominated by one of the three layers or whether it is distributed between the different layers. We search for the origin of the asymmetry of the line profiles. We compare the line excitation temperature with the CMB. We then look at the abundances of the CO and HCO$^+$ isotopologues along the LoS as a function of H_2 column density and far UV radiation field (Sect. 5.3): we try to quantitatively relate the variations of the abundances with known chemical processes such as photodissociation, isotopic fractionation, or freezing onto dust grains. Finally, we investigate the relationship of the derived physical parameters with the star formation activity (Sects. 5.4 and 5.5): we search for a kinematic signature of infall or expansion of the gas onto the filaments and dense cores. We discuss whether the molecular gas properties (kinetic temperature, volume density, and thermal pressure) are affected by the nearby presence of a PDR or of a young stellar object.

5.1. Contribution of the different layers to the integrated intensity

For each studied line, Fig. 10 shows 1) the relative error of the integrated intensity estimated by the model with respect to the observed value, integrated between 8.0 and 13.5 km s^{-1} (first column), and 2) the contributions of the different layers (foreground, inner, background, and CMB, from the second to the fifth column) to the line integrated intensity. While our estimator only fits the ^{12}CO (1 − 0) peak intensity, we still provide the relative error and contributions of each layer to the integrated intensity of this line. We have filtered out sight-lines where the relative error on the integrated line intensity is larger than 40% or the visual extinction is lower than 3 mag for all lines, except for the faint H^{13}CO$^+$ (1 − 0) line where we used a threshold of 10 mag. The percentage listed in the top right corner of each panel is the mean±rms value computed over the selected sight-lines for the associated map. As expected, a negative contribution is ascribed to the CMB layer because of the ON-OFF observation procedure that implies the subtraction of a black body emission at 2.73 K as indicated in Eq. (8). Moreover, the sum of contributions of the layers to the integrated intensities can be larger than 100% when the estimated signal exceeds the observed integrated line intensity.

The integrated intensity of all the lines is reconstructed within 20% overall (= $\sqrt{\mathrm{mean}^2 + \mathrm{rms}^2}$). The rms of the relative error on the integrated intensity decreases when the S/N increases, except for the ^{12}CO (1 − 0) line. The ^{13}CO (2 − 1) and C^{18}O (1 − 0) and (2 − 1) lines are reconstructed with much less bias (±1%) than the ^{12}CO, ^{13}CO, HCO$^+$, and H^{13}CO$^+$ (1 − 0) lines. The bias for the ^{12}CO (1 − 0) line can be explained by the fact that we only use the peak temperature in the fit.

The contributions of the different layers to the integrated intensity depend much on the considered species. The ^{12}CO (1 − 0) integrated intensity mainly comes from the foreground layer (85%) with a second contribution from the inner layer (22%), while the integrated intensity of the two C^{18}O lines mainly comes from the inner layer (57%) with an approximately equal contributions from the foreground and background layers (∼24% each). The ^{13}CO lines have an intermediate behavior. The ^{13}CO (1 − 0) integrated intensity comes from the foreground, inner, and background layers at approximately equal shares. The foreground layer contributes more than the background layer for the ^{13}CO (2 − 1) line. This is consistent with the higher optical depth for the (2 − 1) line as compared with the (1 − 0) one.

The (1 − 0) line of the HCO$^+$ isotopologues shows a very different behavior that is mostly related to the fact that its excitation temperature is close to the CMB temperature. While the subtraction of the CMB has a small effect on the CO isotopologues, it is significant for the HCO$^+$ and H^{13}CO$^+$ lines and can lead to a significant negative contribution to the total integrated line intensity, beyond −50%. The foreground, inner, and background layers must therefore provide contributions significantly larger than the detected integrated line intensity to compensate, implying that the percentage of the combined layer contributions is larger than 100%. The inner layer contribution to the (1 − 0) line is about 50% for both HCO$^+$ and H^{13}CO$^+$. However, the foreground contribution largely dominates the background contribution for the HCO$^+$ optically thick line, while both layers have approximately equal contributions for the optically thin H^{13}CO$^+$ line.

Looking at the spatial distribution of the contributions for the ^{13}CO and C^{18}O lines, the inner layer contributes much more than the outer layers when the visual extinction is larger than

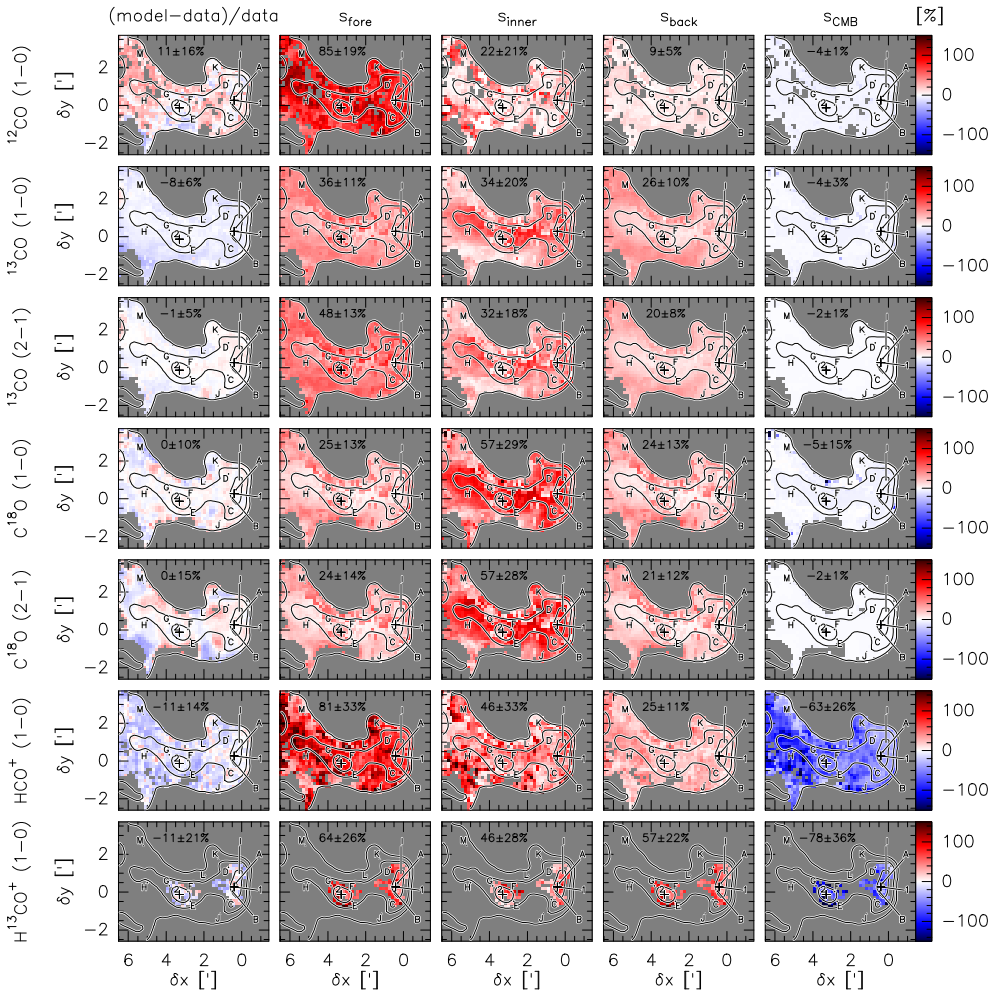


Fig. 10. Maps of the contributions of the different layers to the line intensity integrated between 8 and 13.5 km s^{-1}. First column: relative error of the model with respect of the data in percentage. Second to fifth columns: contribution of the foreground, inner, background, and CMB layers to the line integrated intensity in percentage of that measured on the data. Each row corresponds to a given line: from top to bottom, ^{12}CO (1 – 0), ^{13}CO (1 – 0) and (2 – 1), C^{18}O (1 – 0) and (2 – 1), HCO$^+$ (1 – 0), and H^{13}CO$^+$ (1 – 0). The percentage on the top center of each panel is the mean±rms value computed over the associated maps. Positive and negative relative errors or contributions are shown in red and blue, respectively.

7 mag, and vice-versa. We relate this to the change of geometry of the Horsehead pillar that happens approximately at this visual extinction. Above 7 mag, the geometry is mostly face-on and the inner layer dominates for optically thin lines. Below 7 mag, the geometry becomes edge-on, and the contribution of the inner layer decreases and sometimes vanishes compared to that of the outer layers.

5.2. Contribution of the different layers to the spectra

To get a better understanding of how the different layers contribute to the observed spectra, Figs. 11, H.1, and H.2 show the modeled decomposition of spectra selected around the dense cores 1, and 2, and then in the outskirts of the Horsehead nebula, respectively. In all cases, the background emission is attenuated by the inner layer in its foreground. Hence, the contribution of the background layer to the emergent spectra mostly matters in the line wings. Similarly, the emission from the inner layer is absorbed by the foreground layer, and this radiative interaction allows one to fit the asymmetries of the profiles of the optically thick lines of ^{12}CO, ^{13}CO, and HCO$^+$.

Figure 11 first shows the decomposition at the peak of the H^{13}CO$^+$ emission (DC 1) for the dense core located at the top of the Horsehead. Positions A and B are two nearby LoSs that have already been used to justify some of the assumptions made in our modeling (see Sects. 4.3 and 4.5). Positions C and D are located at the edge of the dense core away from the σ Ori star that photo-excites the Horsehead PDR, where the HCO$^+$ (1 – 0) peak emission abruptly decreases from above 3.5 K to below 2.5 K on Fig. 1. The inner layer emission contributes the most to the integrated line intensity near the dense core projected center (e.g., positions 1 and A). This is true even for the ^{12}CO (1 – 0) line where the inner emission contributes about 20% of the total integrated line intensity. When going further away from the dense core center (positions C and D), the inner layer continues to provide the dominant contribution for the optically thin C^{18}O lines, while the contributions from the foreground and background layers have either similar integrated line intensities as that from the inner layer, or even dominate for the ^{12}CO or HCO$^+$ lines. Position B seems to be an intermediate case.

Figure H.1 shows the decomposition at the peak of the H^{13}CO$^+$ emission (DC 2) for the dense core located in the

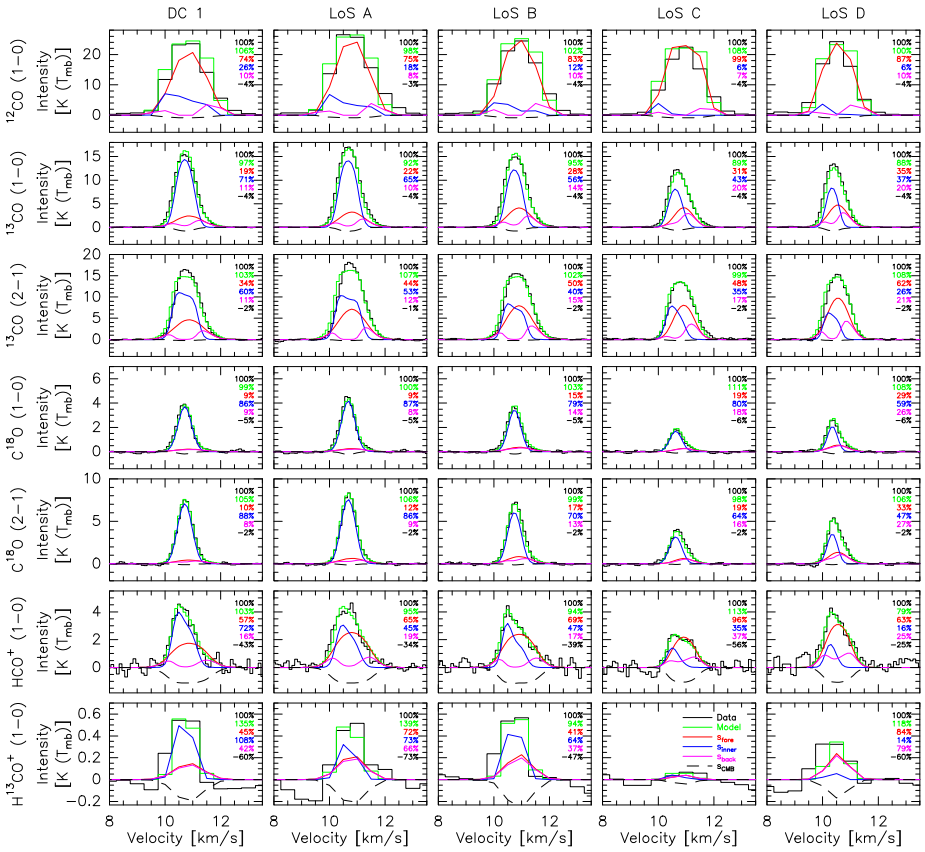


Fig. 11. Modeled decomposition of some spectra around the first dense core. Each column corresponds to the spectra for one of the lines of sight listed in Table 2. Each row corresponds to a given line, from top to bottom: ^{12}CO (1 – 0), ^{13}CO (1 – 0) and (2 – 1), C^{18}O (1 – 0) and (2 – 1), HCO$^+$ (1 – 0), and H^{13}CO$^+$ (1 – 0). We show the decomposition of the full spectrum of ^{12}CO (1 – 0) even though we only used its peak intensity during the fit. In each panel, the data and model spectra are shown as the black and green histograms, respectively. The contributions of the foreground, inner, background layers are shown as the plain lines in red, blue, pink, respectively, and the CMB layer is displayed as the dashed black line.

neck of the Horsehead. Positions E and G are located where the H^{13}CO$^+$ (1 – 0) peak emission abruptly decreases from above 0.25 to below 0.10 K. Position F is extensively studied in Roueff et al. (2024). Position H is located at a LoS where the C^{18}O (1 – 0) emission drops below 3 K. In contrast with the analysis of the previous dense core, the inner layer contributes less to the overall integrated line intensity of the optically thick lines at position 2 near the dense core projected center than at position F located further away. Positions E and G at the edge of the dense core shows a strong asymmetry of the most optically thick lines for which the foreground and background contributions may even dominate the integrated line intensity. This effect is less prominent for the position H located at the transition between the C^{18}O filament and its environment.

Figure H.2 shows the decomposition for LoSs located at the outskirt of Horsehead pillar. There is an embedded young stellar object at position I. Positions J and K are located in the Horsehead mane and muzzle where the C^{18}O (1 – 0) peak intensity is still ∼2 K, but the H^{13}CO$^+$ (1 – 0) is not clearly detected anymore. Finally, positions L and M are located where the ^{13}CO (1 – 0) peak intensity drops below 10 K. Except for position I that has a similar behavior as the positions near the dense core 1, positions J to M have no clearly detected H^{13}CO$^+$, and faint HCO$^+$ and C^{18}O lines. The optically thick lines of HCO$^+$ and ^{12}CO at the K and L positions have inner and outer contributions whose radiative interaction (absorption) almost cancel each other. This is accompanied by an inversion of the centroid velocities of the two components (see Sect. 5.4). Finally, position M shows a spectral decomposition where the fit seems to have inverted the properties of the inner and outer layers as compared with the rest of the map. The integrated line intensity is completely dominated by the outer layers, and the associated linewidth is much wider than for the previous positions. We interpret this as the transition between the face-on and edge-on geometry at the outskirts of the Horsehead pillar. In the edge-on geometry the chemistry is better represented by assumptions enforced in the outer layer. Position M is a clear example of the strong coupling between the chosen chemical hypotheses as a function of the geometry and the inferred kinematics properties.

5.3. Chemical abundances

5.3.1. Relative abundances

Figure 12 shows the spatial distribution of ratios of different column densities derived for the outer layer. These ratios are normalized by the basic value derived from elemental abundances of carbon and oxygen isotopes (see Eq. (24)) and previous studies of the targeted molecules as explained in Sect. 4.2. By

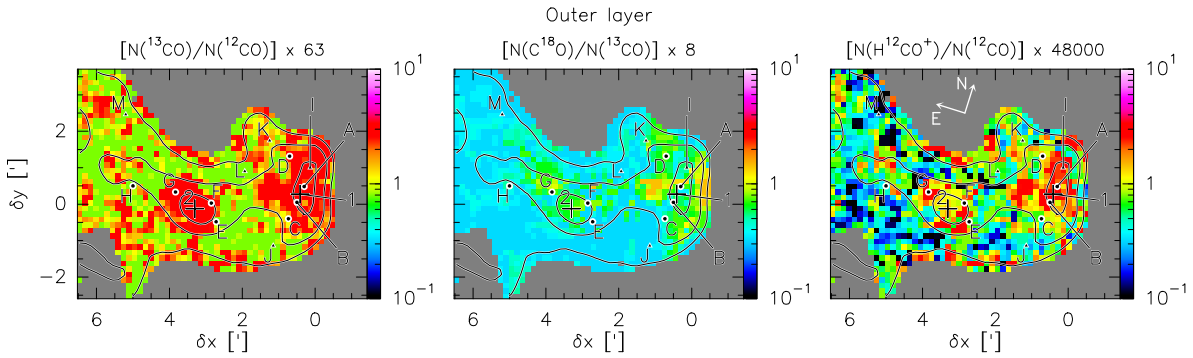


Fig. 12. Maps of the ratios of the column densities for the outer layer. These ratios have been normalized by the basic values established using the elemental isotopic ratios of carbon and oxygen (see Eq. (24)). By construction $N(\mathrm{HCO^+})/N(^{12}\mathrm{CO}) = N(\mathrm{H}^{13}\mathrm{CO^+})/N(^{13}\mathrm{CO})$ (see Sect. 4.5). Contours are drawn at visual extinctions of 3, 7, and 16 mag.

construction, these basic values are enforced in the inner layer, and $N(\mathrm{HCO^+})/N(^{12}\mathrm{CO})$=$N(\mathrm{H}^{13}\mathrm{CO^+})/N(^{13}\mathrm{CO})$ in both layers (see Sect. 4.5).

The mean value of the $N(^{13}\mathrm{CO})/N(^{12}\mathrm{CO})$ ratio in the outer layer is ∼1.6 times the "basic" value used in the inner layer. Moreover, $^{13}\mathrm{C^+}$ is detected by Pabst et al. (2017), and the temperature remains moderate in the outer layer, as shown in Sect. 5.5. All this allows the forward reaction of the fractionation reaction (Eq. (25)) to predominate in the outer layer (see also Liszt 2007; Röllig & Ossenkopf 2013). The $N(^{13}\mathrm{CO})/N(^{12}\mathrm{CO})$ ratio is also enhanced toward the DC 2 LoS, which is in line with the $\mathrm{H}^{13}\mathrm{CO^+}$ detection in that environment.

Two effects help explain the spatial variations of the $N(\mathrm{C}^{18}\mathrm{O})/N(^{13}\mathrm{CO})$ ratio in the outer layer. First, the abundance of $^{13}\mathrm{CO}$ is enhanced due to the $\mathrm{C^+}$ fractionation described above. Second, $\mathrm{C}^{18}\mathrm{O}$ is more easily photodissociated than $^{13}\mathrm{CO}$ (a process named selective photodissociation). Indeed, the shift between the $\mathrm{C}^{18}\mathrm{O}$ and $^{12}\mathrm{CO}$ pre-dissociated UV absorption transitions is significantly larger than the one between the $^{13}\mathrm{CO}$ and $^{12}\mathrm{CO}$ transitions. Shielding effect by $^{12}\mathrm{CO}$ are thus reduced for $\mathrm{C}^{18}\mathrm{O}$ with respect to $^{13}\mathrm{CO}$. Both effects are important over most of the field of view, in particular for visual extinctions lower than 7 mag.

The spatial variations of the $\mathrm{HCO^+}$ abundance relative to CO are larger than those of the two other abundance ratios. The $N(\mathrm{HCO^+})/N(^{12}\mathrm{CO})$ and $N(\mathrm{H}^{13}\mathrm{CO^+})/N(^{13}\mathrm{CO})$ ratios are one to four times larger than the basic value toward dense core LoSs and two to ten times lower than this value for LoSs where $3 \leq A_V \leq 7$ mag. The $\mathrm{HCO^+}$ abundance relative to CO reaches values within a factor of two from the basic value in regions where A_V is around 10 mag.

5.3.2. Column densities and mean abundances relative to $\mathrm{H_2}$

For each species, Fig. 13 shows the inner, outer (foreground plus background), and total (inner plus twice outer) column densities, as well as its mean abundance along the LoS defined as

$$\langle X \rangle = \frac{N_{\mathrm{tot}}(X)}{N_{\mathrm{tot}}(\mathrm{H_2})}, \tag{31}$$

where

$$N_{\mathrm{tot}}(X) = \sum_L N_L(X), \quad \text{and} \quad N_{\mathrm{tot}}(\mathrm{H_2}) = \sum_L N_L(\mathrm{H_2}). \tag{32}$$

Following (Pety et al. 2017), we estimated the total column density of $\mathrm{H_2}$ from the "molecular" visual extinction (defined in Sect. 2 and Appendix A) as

$$N_{\mathrm{tot}}(\mathrm{H_2}) = 0.9 \times 10^{21}\ \mathrm{cm^{-2}}\ \frac{A_V^{\mathrm{mol}}}{1\ \mathrm{mag}}. \tag{33}$$

We have also added the ratio of column densities in the inner and outer layers, $N_{\mathrm{in}}/N_{\mathrm{ou}}$ ratio for each species in this figure to illustrate whether one layer dominates or whether the mean abundance results from just an average of both layers. Indeed, if we define the abundance of a species X with respect to $\mathrm{H_2}$ in the layer L as

$$[X]_L = \frac{N_L(X)}{N_L(\mathrm{H_2})}, \tag{34}$$

it is straightforward to show that

$$\frac{\langle X \rangle}{[X]_{\mathrm{in}}} = \frac{1 + 2\,\frac{N_{\mathrm{ou}}(X)}{N_{\mathrm{in}}(X)}}{1 + 2\,\frac{[X]_{in}}{[X]_{ou}}\,\frac{N_{\mathrm{ou}}(X)}{N_{\mathrm{in}}(X)}}. \tag{35}$$

In this equation, the mean abundance of species X converges to $[X]_{\mathrm{in}}$ when $N_{\mathrm{ou}}(X) \ll N_{\mathrm{in}}(X)$, and to $[X]_{\mathrm{ou}}$ when $N_{\mathrm{ou}}(X) \gg N_{\mathrm{in}}(X)$. The convergence rate depends on the value of the $[X]_{\mathrm{in}}$ / $[X]_{\mathrm{ou}}$ ratio.

The maps of the total column density for the three main CO isotopologues clearly show the filamentary structure inside the Horsehead pillar. From this viewpoint, they look more like the column density of the inner layer than that of the outer layer. This is consistent with the fact that the ratio of the inner and outer column densities is much larger than one for extinctions above 7 mag, and around or lower than one when $3 \leq A_V \leq 7$ mag. Even though the $^{12}\mathrm{CO}\,(1-0)$ line is highly optically thick, the simultaneous fit of different chemical species allowed us to recover reasonable column densities of $^{12}\mathrm{CO}$ for the inner layer.

In contrast, the maps of the total column density for the two main $\mathrm{HCO^+}$ isotopologues show a more balanced combination of column densities from the inner and outer layers, in agreement with the less contrasted maps of the ratio of the inner and outer column densities. For $\mathrm{HCO^+}$, the column density of the inner layer delineates the filamentary structure with a peak of the column density toward the dense cores. The map of the outer column density shows less structure. The total column density map of $\mathrm{H}^{13}\mathrm{CO^+}$ clearly shows higher column densities in the regions of higher extinctions. Overall the relative abundance seems to be somewhat higher along the LoSs toward dense cores and fairly constant elsewhere given the limited S/N of the line.

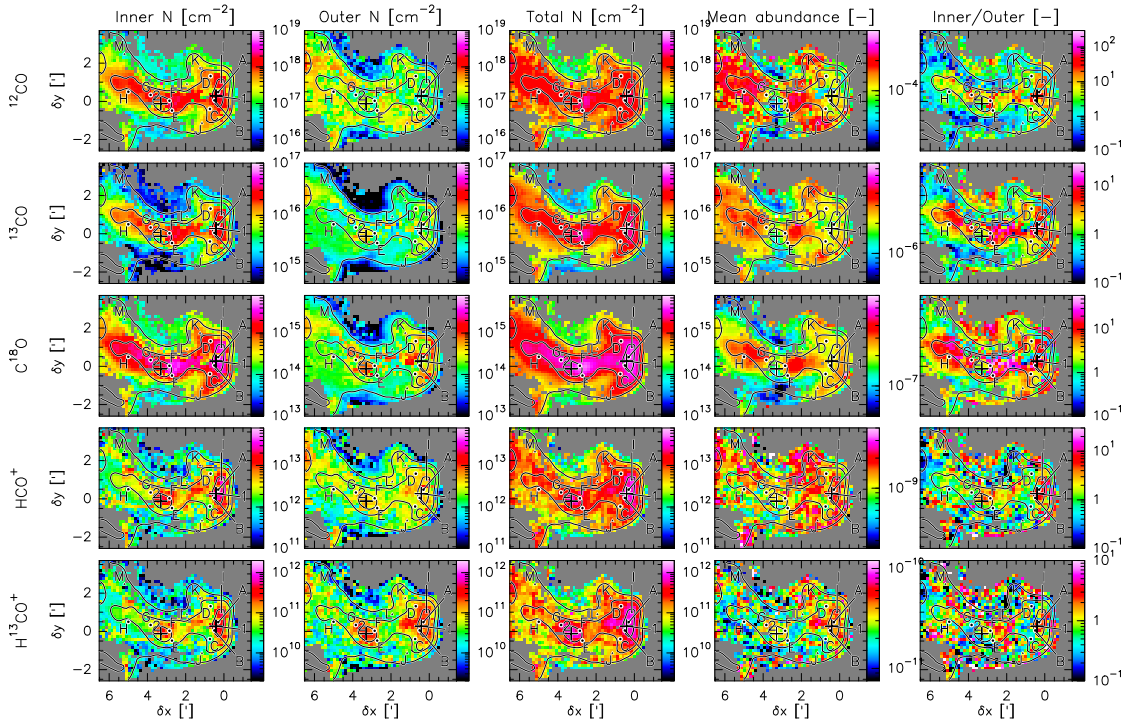


Fig. 13. Maps of the chemical properties for ^{12}CO, ^{13}CO, C^{18}O, and HCO$^+$, from top to bottom. The three first columns show the column densities for the inner, outer (foreground or background), and total (inner plus twice outer) column densities, while the last column shows the mean abundance, namely, the total column density divided by the H_2 column density deduced from the visual extinction. The last column is the ratio of the inner by the outer column density. Contours are drawn at visual extinctions of 3, 7, and 16 mag.

The most striking feature of the mean abundance maps is the depletion of the CO isotopologues toward the two dense cores. This is obvious for the dense core in the Horsehead neck. For the dense core at the top of the Horsehead, it is clearer for the mean abundance of ^{12}CO than for those of ^{13}CO and C^{18}O. To quantitatively confirm these trends, Fig. 14 shows the joint histograms of the column densities and mean abundances as a function of the "molecular" visual extinction. We have fitted independently in each range of A_V a power-law to the scatter plots of $N_{\rm tot}$ versus A_V. We have used the CRB values to take into account the uncertainty on $N_{\rm tot}$ and we have assumed that the visual extinction are perfectly known. Appendix E explains the computation of the uncertainty on $N_{\rm tot}$, and Appendix F details the fit algorithm. For the sake of reproducibility, the values of the fitted coefficients are listed in Table 6 and shown as the black lines in the third column of Fig. 14. The abundance of C^{18}O with respect to H_2 reaches 2.0×10^{-7} at the DC 1 position and 1.3×10^{-7} at DC 2. As the used "basic" abundance of C^{18}O is 2.84×10^{-7}, this implies depletion factors of 1.4 for DC 1 and 2.2 for DC 2. Such values are in agreement with abundances and depletion factors obtained in other studies of molecular cores. For instance, Tafalla et al. (2002) found a strong depletion of C^{18}O in the inner regions of five Taurus starless cores (less than 100″ or 0.07 pc from their center) using a reference C^{18}O abundance of 1.7×10^{-7}. In a more recent study of four dense cores located in infrared dark clouds, Feng et al. (2020) also measure depletion factors of a few for C^{18}O.

As the far UV radiation field plays an important role at least in the eastern part of the Horsehead pillar, Fig. H.3 shows the joint histograms of the column densities and mean abundances as a function of the G_0/A_V ratio. We use this ratio as a proxy of the ratio of the far UV field over the volume density that controls to first order the chemistry of PDRs. Using this proxy assumes that there is some relation between the volume and column densities.

The mean abundances of ^{12}CO and ^{13}CO increases up to a visual extinction of about 3 mag. At this visual extinction, they reach the values corresponding to the elemental abundance of the associated carbon isotope. For $A_V \gtrsim 7$ mag depletion sets in, and the column density does not increase much anymore, which leads to a decreased abundance. The abundance of C^{18}O increases up to $A_V \sim 7$ mag, and then decreases. Looking at the G_0/A_V ratio, the column densities increase when G_0/A_V decreases from 15 to 1.5 ISRF mag^{-1}, and then they remain mostly constant. When $G_0/A_V \leq 4.75$ ISRF mag^{-1}, the inner column density clearly dominates the column density budget. Above this value the total column density is dominated by the outer layer for ^{12}CO and ^{13}CO, but it remains dominated by the inner layer for C^{18}O.

The column densities of the two HCO$^+$ isotopologues studied increase more steadily with increasing visual extinction or decreasing G_0/A_V values. The associated mean abundances are quite scattered. Their typical values seem to be lower than the basic values (see Eq. (24)) by a factor between two and three (or between 0.3 and 0.5 dex). For both isotopologues, the total column density includes contributions from both layers.

Looking at the inner and outer column densities of the CO isotopologues as a function of G_0/A_V, we can distinguish another regime. For $1.5 \leq G_0/A_V \leq 4.75$ ISRF mag^{-1}, the column densities of CO isotopologues clearly decrease with increasing G_0/A_V in the inner layer, while they seem to saturate in the outer layer. The opposite behavior is seen for $4.75 \leq G_0/A_V$ ISRF mag^{-1}, namely the column densities of the

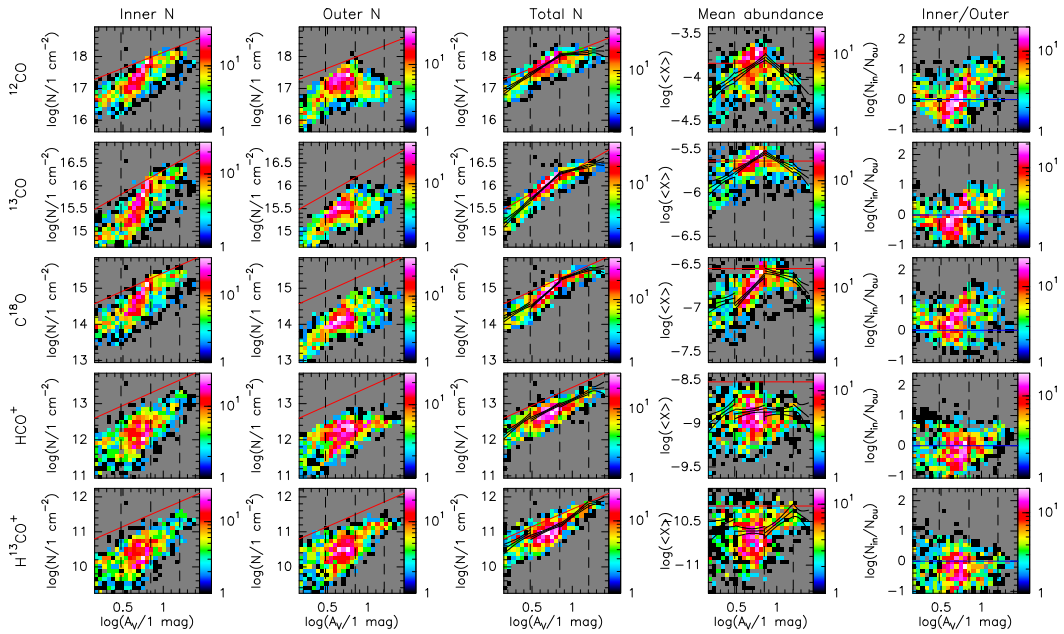


Fig. 14. Joint histograms of the molecular column densities, abundances, or their ratios as a function of the visual extinction for ^{12}CO, ^{13}CO, C^{18}O, HCO$^+$, and H^{13}CO$^+$, from top to bottom. The three first columns show the column densities for the inner, outer (foreground or background), and total (inner plus twice outer) column densities, while the last column shows the mean abundance, namely, the total column density divided by the H_2 column density deduced from the visual extinction. The dashed vertical lines mark the visual extinction at 3, 7, and 16 mag. The red plain lines show the species basic abundance. The black lines show the piecewise power-law fits of the column densities within the A_V ranges specified above (see Table 6) and the $\pm 3\sigma$ intervals of the fits. These fits can be directly converted to an abundance vs visual extinction relation.

Table 6. Parameters of the piecewise power-law fits ($\log(N_{\rm tot}) = \alpha\ [\log(A_V) - \log(A_V)_0] + \beta$) for each studied species.

	$0.15 \le \log(A_V) \le 0.48$		$0.48 \le \log(A_V) \le 0.85$		$0.85 \le \log(A_V) \le 1.2$		$1.2 \le \log(A_V) \le 1.4$	
$\log(A_V)_0$	0.315		0.665		1.025		1.3	
^{12}CO	$\begin{pmatrix} 1.65 \\ 17.18 \end{pmatrix}$	$\begin{pmatrix} 0.01904 \\ 0.00019 \\ -0.00084 \end{pmatrix}$	$\begin{pmatrix} 1.63 \\ 17.75 \end{pmatrix}$	$\begin{pmatrix} 0.00430 \\ 0.00005 \\ -0.00019 \end{pmatrix}$	$\begin{pmatrix} 0.20 \\ 18.10 \end{pmatrix}$	$\begin{pmatrix} 0.00749 \\ 0.00010 \\ 0.00044 \end{pmatrix}$	$\begin{pmatrix} -0.03 \\ 18.18 \end{pmatrix}$	$\begin{pmatrix} 0.16839 \\ 0.00054 \\ 0.00552 \end{pmatrix}$
^{13}CO	$\begin{pmatrix} 1.56 \\ 15.42 \end{pmatrix}$	$\begin{pmatrix} 0.01287 \\ 0.00013 \\ -0.00054 \end{pmatrix}$	$\begin{pmatrix} 1.60 \\ 15.98 \end{pmatrix}$	$\begin{pmatrix} 0.00155 \\ 0.00002 \\ -0.00007 \end{pmatrix}$	$\begin{pmatrix} 0.45 \\ 16.36 \end{pmatrix}$	$\begin{pmatrix} 0.00293 \\ 0.00003 \\ 0.00011 \end{pmatrix}$	$\begin{pmatrix} 0.35 \\ 16.50 \end{pmatrix}$	$\begin{pmatrix} 0.04363 \\ 0.00014 \\ 0.00133 \end{pmatrix}$
C^{18}O	$\begin{pmatrix} 1.26 \\ 14.36 \end{pmatrix}$	$\begin{pmatrix} 0.01556 \\ 0.00017 \\ -0.00063 \end{pmatrix}$	$\begin{pmatrix} 1.96 \\ 14.84 \end{pmatrix}$	$\begin{pmatrix} 0.00258 \\ 0.00003 \\ -0.00013 \end{pmatrix}$	$\begin{pmatrix} 0.79 \\ 15.38 \end{pmatrix}$	$\begin{pmatrix} 0.00401 \\ 0.00005 \\ 0.00016 \end{pmatrix}$	$\begin{pmatrix} -0.01 \\ 15.53 \end{pmatrix}$	$\begin{pmatrix} 0.05222 \\ 0.00017 \\ 0.00164 \end{pmatrix}$
HCO$^+$	$\begin{pmatrix} 1.67 \\ 12.40 \end{pmatrix}$	$\begin{pmatrix} 0.02108 \\ 0.00020 \\ -0.00096 \end{pmatrix}$	$\begin{pmatrix} 1.16 \\ 12.78 \end{pmatrix}$	$\begin{pmatrix} 0.00584 \\ 0.00006 \\ -0.00019 \end{pmatrix}$	$\begin{pmatrix} 1.03 \\ 13.13 \end{pmatrix}$	$\begin{pmatrix} 0.00787 \\ 0.00009 \\ 0.00015 \end{pmatrix}$	$\begin{pmatrix} 0.22 \\ 13.42 \end{pmatrix}$	$\begin{pmatrix} 0.15840 \\ 0.00049 \\ 0.00516 \end{pmatrix}$
H^{13}CO$^+$	$\begin{pmatrix} 1.43 \\ 10.74 \end{pmatrix}$	$\begin{pmatrix} 0.03755 \\ 0.00038 \\ -0.00172 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ 11.03 \end{pmatrix}$	$\begin{pmatrix} 0.00791 \\ 0.00008 \\ -0.00026 \end{pmatrix}$	$\begin{pmatrix} 1.75 \\ 11.49 \end{pmatrix}$	$\begin{pmatrix} 0.01084 \\ 0.00013 \\ 0.00039 \end{pmatrix}$	$\begin{pmatrix} 0.35 \\ 11.84 \end{pmatrix}$	$\begin{pmatrix} 0.09248 \\ 0.00029 \\ 0.00235 \end{pmatrix}$

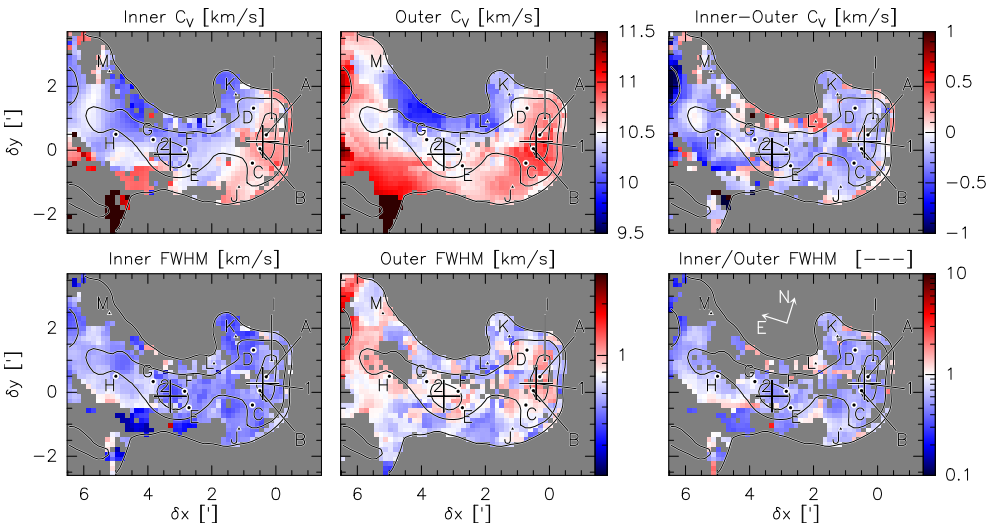
Notes. In each sub-table, we provide on the left α and β coefficents and on the right var(α), var(β) and cov(α,β). The error bars of the line fit are given by $\sqrt{\mathrm{var}(\alpha)\,(\log(A_V) - \log(A_V)_0)^2 + \mathrm{var}(\beta) + 2\,\mathrm{cov}(\alpha,\beta)\,(\log(A_V) - \log(A_V)_0)}$.

CO isotopologues do not change much with increasing G_0/A_V in the inner layer, while the column densities of the outer layer decrease. This behavior may be caused by the increased photodissociation of the CO isotopologues in a higher far UV radiation field.

5.4. Kinematics

In Fig. 15, we compare the spatial distribution of the centroid velocity and velocity dispersion between the inner and outer layers. The centroid velocity of the inner and outer layers displays to first order an increasing gradient from north to south of the Horsehead. Hily-Blant et al. (2005) already described this behavior and interpreted this as a global rotation of the Horsehead pillar around its east-west axis.

We concentrate on LoSs where $A_V \ge 7$ mag. The difference of the centroid velocities for the inner and outer layers confirms that the centroid velocity of the outer layer follows that of the inner layer, shifted by about -0.25 km s^{-1}. To check whether this


Inner C_V [km/s]
Outer C_V [km/s]
Inner−Outer C_V [km/s]
Inner FWHM [km/s]
Outer FWHM [km/s]
Inner/Outer FWHM [−−−]
δy ['] δx ['] δx ['] δx

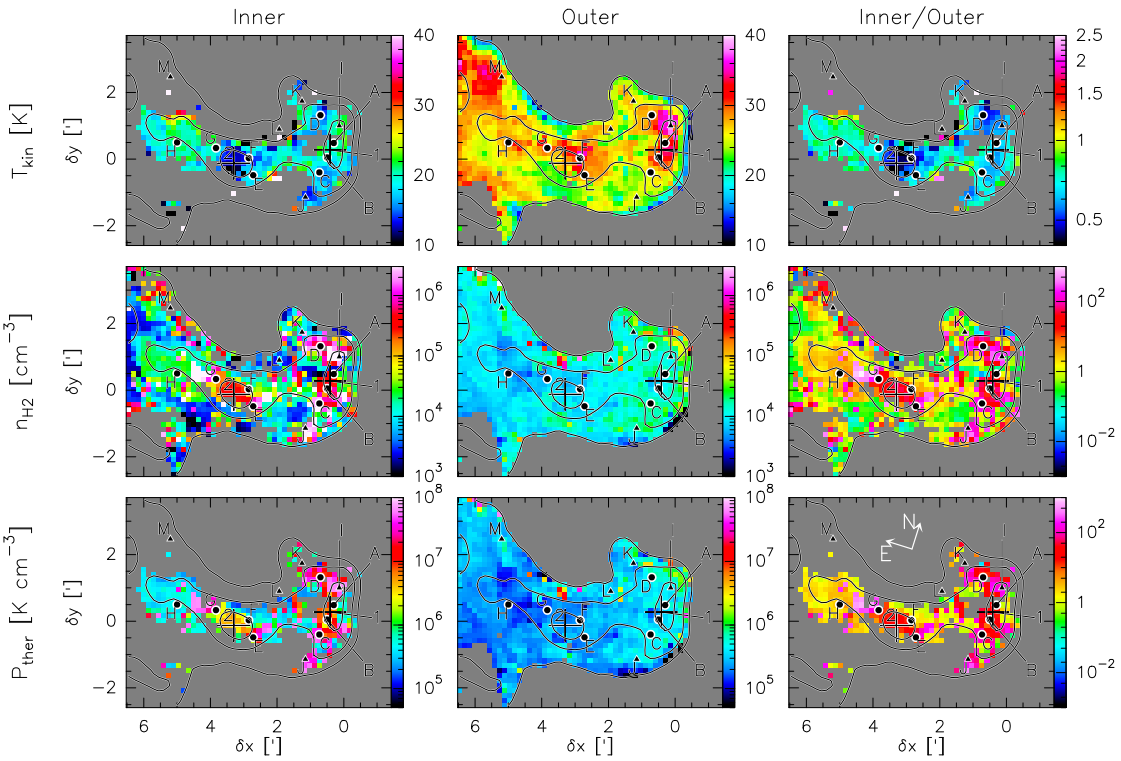


Fig. 16. Maps of the gas physical conditions. From top to bottom: kinetic temperature, volume density, and thermal pressure. The first two columns display the values for the inner and outer (foreground or background) layers, while the last column shows their ratio.

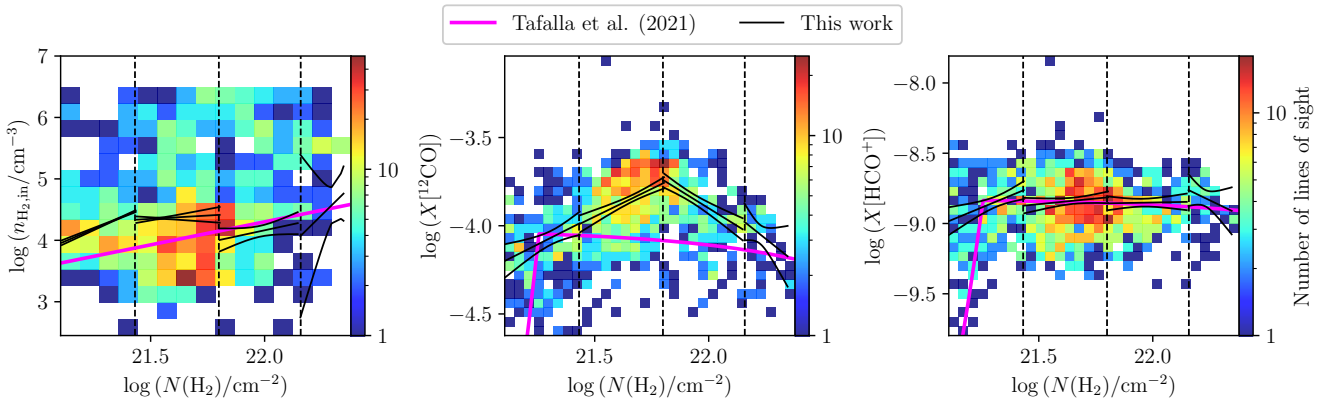


Fig. 17. Comparison between the sandwich and TUH models. Left: volume density of the inner layer $n_{\mathrm{H_2,\,in}}$ estimated with the sandwich model as a function of the H_2 column density. Middle and right: comparison of the mean abundance estimated with the sandwich model as a function of the H_2 column density and the model proposed by Tafalla et al. (2021). The second and third panels show this comparison for ^{12}CO and HCO^+, respectively. In all panels, joint histograms show our data. The black lines show the power law fits and their uncertainty intervals. The pink lines show the Tafalla et al.'s model. The dashed vertical lines mark the $N(H_2)$ corresponding to a visual extinction at 3, 7, and 16 mag.

and apply a model of the abundance of molecular species as a function of the H_2 column density (a summary is given in Appendix G). We refer to this model as the TUH one from this point on. The left panel of Fig. 17 overlays TUH model onto the joint distributions of the estimated volume density of the inner layer as a function of the H_2 column density. The two models are consistent for $\log(N(H_2)/\,\mathrm{cm^{-2}}) > 21.8$, but they show differences of about a factor of two for lower column densities.

The middle and right panels of Fig. 17 overlay the TUH model onto the joint distributions of the estimated mean abundance of ^{12}CO and HCO^+ versus H_2 column density toward the Horsehead nebula. These joint distributions are identical to these displayed in Fig. 14, except that the molecular visual extinctions were scaled to the H_2 column density using Eq. (33). For an easier comparison, we also show the fits listed in Table 6.

The ^{12}CO mean abundance profile that we derive toward the Horsehead nebula is different from TUH one. Our power-law fit

reaches its maximum for $\log(N(H_2)/\,\mathrm{cm}^{-2}) = 21.8$, while TUH model starts to decline at $\log(N(H_2)/\,\mathrm{cm}^{-2}) \sim 21.3$. Moreover, the maximum abundance of ^{12}CO is two to three times higher than that predicted by the TUH model.

In contrast, the agreement between our estimated HCO^+ mean abundance profile and TUH model is good for $21.4 < \log(N(H_2)/\,\mathrm{cm}^{-2}) < 22.2$. Both profiles show a slow decrease as $N(H_2)$ increases. Above $\log(N(H_2)/\,\mathrm{cm}^{-2}) = 22.2$, our estimation yields a steeper fall of the HCO^+ abundance than TUH model, but TUH model still lies within our fit confidence interval.

For both species, TUH model predicts a sharp drop of the abundance below $\log(N(H_2)/\,\mathrm{cm}^{-2}) \lesssim 21.3$ or $A_V = 2$ mag, resulting from the molecule photodissociation. The Horsehead nebula contains few LoSs in this A_V regime at the very outskirts of the pillar. It is thus difficult to draw a robust conclusion for this regime.

7. Conclusions

In this paper, we have proposed a multilayer model that allows us 1) to fit the low−J lines of different species toward a heterogeneous LoS and 2) to recover information about the chemical abundances, the kinematics, and the physical state (kinetic temperature and volume density) of the gas at different depths along the LoS. To do this, we introduced a fast and efficient maximum likelihood estimator that simultaneously fits all the lines available for each pixel independently.

We also took special care in selecting adequate physical and chemical assumptions in order to avoid biasing the results. Following the recommendations of Roueff et al. (2024), we not only searched for small and unbiased fit residuals but also coherent physical and chemical results. More precisely, in addition to fixing the LoS geometry to three layers (similar to a core surrounded by an envelope), we approximated a centroid velocity gradient along the LoS by allowing the layers to have different velocities in order to be able to reproduce the observed line asymmetries. We also proposed a varying chemistry along the LoS in order to take into account the different processes (e.g., fractionation or molecular freeze-out) at play when the gas is more or less exposed to far UV or more or less dense and cold. While we assumed that the relative abundances ^{12}CO/^{13}CO, C^{18}O/^{13}CO, and H^{13}CO$^+$/HCO$^+$ scale with the elemental abundances of the corresponding carbon and oxygen isotopes in the inner dense layer, we fit the column densities of ^{12}CO, ^{13}CO, C^{18}O, and HCO^+ in the outer translucent layer, and we deduced the column density of H^{13}CO$^+$ with the constraint $N(H^{13}CO^+)/N(^{13}CO) = N(HCO^+)/N(^{12}CO)$.

We used this model to fit the $(1-0)$ and $(2-1)$ lines of the CO and HCO^+ isotopologues toward the translucent, filamentary, and dense core LoSs in the Horsehead pillar. Our main astrophysical findings are as follows:

- A precise modeling of the chemistry and physics toward filaments and dense cores requires a heterogeneous model composed of at least an inner dense layer surrounded by two outer translucent layers.
- The proposed model succeeds in reproducing integrated intensities within 20 % for all lines. The associated bias is small for the ^{13}CO and C^{18}O lines.
- We quantitatively confirm that the different lines of the CO and HCO^+ isotopologues are sensitive to different depths along the LoS: The emission from the ^{12}CO $(1-0)$ line is dominated by the foreground layer (85%). The integrated intensity of the ^{13}CO $(1-0)$ and $(2-1)$ lines comes from all three layers at approximately the same share. The inner layer has the largest contribution to the total integrated line intensity of the C^{18}O $(1-0)$ and $(2-1)$ lines, about 60%. The emission from the HCO^+ and H^{13}CO$^+$ $(1-0)$ lines is dominated by the outer layers (foreground and background), but the inner layer also contributes significantly. Moreover, the emission from the HCO^+ and H^{13}CO$^+$ $(1-0)$ lines is sub-thermalized with an excitation temperature lower than 10 K.
- Deduced chemical properties clearly show 1) the depletion of CO isotopologues toward the dense core LoSs and 2) the increase of the ^{12}CO and HCO^+ column density as a function of A_V or G_0/A_V in the warm translucent LoSs. The basic abundances of ^{12}CO, ^{13}CO, and C^{18}O defined in Eq. (24) are reached at visual extinction between 3 and 7 mag, respectively. The typical mean abundance of HCO^+ with respect to H_2 is $\sim 1.3 \times 10^{-9}$, which is about two times lower than the usual value in diffuse clouds (Gerin et al. 2019). The estimated column density ratio $N(^{13}CO)/N(C^{18}O)$ in the envelope increases with decreasing visual extinction and reaches 25 in the pillar outskirts.
- In addition to confirming the general rotation of the Horsehead pillar around its axis, we found the signature of accreting motions of the foreground layer onto the inner denser gas.
- While the inferred $T_{\rm kin}$ of the envelope varies from 25 to 40 K, that of the inner layer drops to ∼15 K in the western dense core. The estimated $n_{\rm H_2}$ in the inner layer is $\sim 3 \times 10^4$ cm^{-3} toward the filament, and it increases by a factor of ten toward dense cores. The thermal pressure of the outer layer shows a decreasing east-west gradient.

Except for the abundance ratios in the inner layer that were fixed to ease the estimation in the dense core LoSs, the parameters were estimated from the data with confidence intervals. It will be useful to analyze other datasets in order to check whether the used physical and chemical hypotheses depend on the considered region. If they indeed depend on the physical gas conditions or region, trying to process large maps would imply automatically selecting the most suitable assumptions for a given LoS. This would probably require some insight into the dominant chemical processes as a function of the typical volume density and kinetic temperature. Processing large maps will also require dealing with spectra that present multiple velocity components.

Acknowledgements. We thank the referee for useful comments that helped us improve the manuscript. This work is based on observations carried out under project numbers 019-13, 022-14, 145-14, 122-15, 018-16, and finally the large program number 124-16 with the IRAM 30m telescope. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain). This work received support from the French Agence Nationale de la Recherche through the DAO-ISM grant ANR-21-CE31-0010, and from the Programme National "Physique et Chimie du Milieu Interstellaire" (PCMI) of CNRS/INSU with INC/INP, co-funded by CEA and CNES. M.G.S.M. and J.R.G. thank the Spanish MICINN for funding support under grant PID2019-106110GB-I00. M.G.S.M acknowledges support from the NSF under grant CAREER 2142300. Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). D.C.L. acknowledges financial support from the National Aeronautics and Space Administration (NASA) Astrophysics Data Analysis Program (ADAP).

References

Abergel, A., Misselt, K., Gordon, K. D., et al. 2024, A&A, 687, A4
André, P., Men'shchikov, A., Bontemps, S., et al. 2010, A&A, 518, L102
Barnes, A. T., Kauffmann, J., Bigiel, F., et al. 2020, MNRAS, 497, 1972
Bergin, E. A., & Tafalla, M. 2007, ARA&A, 45, 339
Bron, E., Roueff, E., Gerin, M., et al. 2021, A&A, 645, A28

Dame, T. M., & Lada, C. J. 2023, ApJ, 944, 197
Denis-Alpizar, O., Stoecklin, T., Dutrey, A., & Guilloteau, S. 2020, MNRAS, 497, 4276
Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Princeton University Press)
Einig, L., Pety, J., Roueff, A., et al. 2023, A&A, 677, A158
Evans, Neal J., I., Kim, K.-T., Wu, J., et al. 2020, ApJ, 894, 103
Feng, S., Li, D., Caselli, P., et al. 2020, ApJ, 901, 145
Fuente, A., Garcia-Burillo, S., Usero, A., et al. 2008, A&A, 492, 675
Fuente, A., Goicoechea, J. R., Pety, J., et al. 2017, ApJ, 851, L49
Garthwaite, P. H., Jolliffe, I. T., & Jones, B. 1995, Statistical Inference (London: Prentice Hall Europe)
Gaudel, M., Orkisz, J. H., Gerin, M., et al. 2023, A&A, 670, A59
Gerin, M., Goicoechea, J. R., Pety, J., & Hily-Blant, P. 2009, A&A, 494, 977
Gerin, M., Liszt, H., Neufeld, D., et al. 2019, A&A, 622, A26
Gioumousis, G., & Stevenson, D. P. 1958, J. Chem. Phys., 29, 294
Goicoechea, J. R., Pety, J., Gerin, M., et al. 2006, A&A, 456, 565
Goicoechea, J. R., Pety, J., Gerin, M., Hily-Blant, P., & Le Bourlot, J. 2009, A&A, 498, 771
Goldsmith, P. F., & Kauffmann, J. 2017, ApJ, 841, 25
Gratier, P., Pety, J., Guzmán, V., et al. 2013, A&A, 557, A101
Guzmán, V., Pety, J., Goicoechea, J. R., Gerin, M., & Roueff, E. 2011, A&A, 534, A49
Guzmán, V., Pety, J., Gratier, P., et al. 2012, A&A, 543, L1
Guzmán, V. V., Goicoechea, J. R., Pety, J., et al. 2013, A&A, 560, A73
Guzmán, V. V., Pety, J., Gratier, P., et al. 2014, Faraday Discuss., 168, 103
Guzmán, V. V., Pety, J., Goicoechea, J. R., et al. 2015, ApJ, 800, L33
Habart, E., Abergel, A., Walmsley, C. M., Teyssier, D., & Pety, J. 2005, A&A, 437, 177
Hamberg, M., Kashperka, I., Thomas, R. D., et al. 2014, J. Phys. Chem. A, 118, 6034
Hernández-Vera, C., Guzmán, V. V., Goicoechea, J. R., et al. 2023, A&A, 677, A152
Hily-Blant, P., Teyssier, D., Philipp, S., & Güsten, R. 2005, A&A, 440, 909
Kauffmann, J., Goldsmith, P. F., Melnick, G., et al. 2017, A&A, 605, L5
Kay, S. M. 1997, Fundamentals of Statistical Signal Processing: Estimation Theory (Prentice Hall)
Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531
Kong, S., Lada, C. J., Lada, E. A., et al. 2015, ApJ, 805, 58
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Langer, W. D. 1977, ApJ, 212, L39
Langer, W. D., Graedel, T. E., Frerking, M. A., & Armentrout, P. B. 1984, ApJ, 277, 581
Liszt, H. S. 2007, A&A, 476, 291
Lombardi, M., Bouy, H., Alves, J., & Lada, C. J. 2014, A&A, 566, A45
Mather, J. C., Cheng, E. S., Cottingham, D. A., et al. 1994, ApJ, 420, 439
Mladenović, M., & Roueff, E. 2014, A&A, 566, A144
Myers, P. C., Mardones, D., Tafalla, M., Williams, J. P., & Wilner, D. J. 1996, ApJ, 465, L133
Ochsendorf, B. B., Cox, N. L. J., Krijt, S., et al. 2014, A&A, 563, A65
Orkisz, J. H., Peretto, N., Pety, J., et al. 2019, A&A, 624, A113
Pabst, C. H. M., Goicoechea, J. R., Teyssier, D., et al. 2017, A&A, 606, A29
Pety, J., Teyssier, D., Fossé, D., et al. 2005, A&A, 435, 885
Pety, J., Goicoechea, J. R., Hily-Blant, P., Gerin, M., & Teyssier, D. 2007, A&A, 464, L41
Pety, J., Gratier, P., Guzmán, V., et al. 2012, A&A, 548, A68
Pety, J., Guzmán, V. V., Orkisz, J. H., et al. 2017, A&A, 599, A98
Röllig, M., & Ossenkopf, V. 2013, A&A, 550, A56
Roueff, A., Gerin, M., Gratier, P., et al. 2021, A&A, 645, A26
Roueff, A., Pety, J., Gerin, M., et al. 2024, A&A, 686, A255
Santa-Maria, M. G., Goicoechea, J. R., Pety, J., et al. 2023, A&A, 679, A4
Schinnerer, E., & Leroy, A. K. 2024, arXiv e-prints [arXiv:2403.19843]
Scott, G. B., Fairley, D. A., Freeman, C. G., et al. 1997, J. Chem. Phys., 106, 3982
Shu, F. H., Adams, F. C., & Lizano, S. 1987, ARA&A, 25, 23
Tafalla, M., Myers, P. C., Caselli, P., Walmsley, C. M., & Comito, C. 2002, ApJ, 569, 815
Tafalla, M., Usero, A., & Hacar, A. 2021, A&A, 646, A97
Tafalla, M., Usero, A., & Hacar, A. 2023, A&A, 679, A112
van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, A&A, 468, 627
Visser, R., van Dishoeck, E. F., & Black, J. H. 2009, A&A, 503, 323
Ward-Thompson, D., Nutter, D., Bontemps, S., Whitworth, A., & Attwood, R. 2006, MNRAS, 369, 1201
Yang, B., Stancil, P. C., Balakrishnan, N., & Forrey, R. C. 2010, ApJ, 718, 1062

[1] IRAM, 300 rue de la Piscine, 38406 Saint Martin d'Hères, France
[2] Université de Toulon, Aix Marseille Univ, CNRS, IM2NP, Toulon, France
[3] LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, 75014 Paris, France
[4] Instituto de Física Fundamental (CSIC), Calle Serrano 121, 28006, Madrid, Spain
[5] National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
[6] Laboratoire d'Astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, Allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[7] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, GIPSA-Lab, Grenoble 38000, France
[8] Univ. Lille, CNRS, Centrale Lille, UMR 9189 – CRIStAL, 59651 Villeneuve d'Ascq, France
[9] LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, 92190 Meudon, France
[10] Institut de Recherche en Astrophysique et Planétologie (IRAP), Université Paul Sabatier, Toulouse cedex 4, France
[11] Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile
[12] Laboratoire de Physique de l'Ecole normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, Sorbonne Paris Cité, Paris, France
[13] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[14] Department of Earth, Environment, and Physics, Worcester State University, Worcester, MA 01602, USA
[15] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA, 02138, USA
[16] School of Physics and Astronomy, Cardiff University, Queen's buildings, Cardiff CF24 3AA, UK
[17] Department of Astronomy, University of Florida, PO Box 112055, Gainesville, FL 32611, USA

Appendix A: Corrections to the measured visible extinction

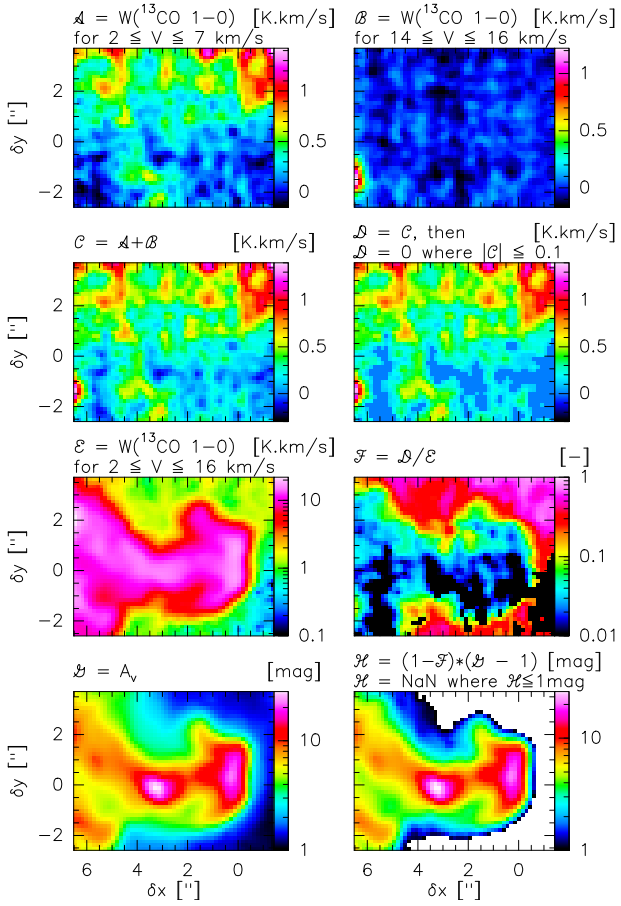


Fig. A.1. Steps to correct the visual extinction for 1) the line integrated emission that arises outside of the $[7, 14\ \mathrm{km\,s^{-1}}]$, and 2) the contribution from atomic hydrogen.

The visible extinction estimated from the dust spectral energy distribution is a measure of all hydrogen atoms present along each line of sight. In order to deduce the amount of molecular hydrogen associated to the modeled lines, we thus needed to correct it for two effects that are illustrated in Fig. A.1.

First, Pety et al. (2017) show that the southwestern edge of the Orion B GMC has on average two velocity components: the brightest one centered around 10.5 $\mathrm{km\,s^{-1}}$ and a fainter one centered around 5 $\mathrm{km\,s^{-1}}$. Each of them has a typical FWHM of 4 $\mathrm{km\,s^{-1}}$. The model fitted in this paper still has limitations to be able to fit multiple velocity components. We thus restricted the studied velocity interval around the bright component, namely, between 7 and 14 $\mathrm{km\,s^{-1}}$. We used the method presented in Einig et al. (2023) to define a position-position-velocity mask of LoS containing some signal on the ^{13}CO(1 − 0) line. We choose this line as the best compromise between optical depth and S/N. We estimated that the relevant velocity interval is $[2, 16]$ $\mathrm{km\,s^{-1}}$. It is thus obvious that part of the estimated visual extinction is associated with line emission between $[2, 7]$ $\mathrm{km\,s^{-1}}$ and $[14, 16]$ $\mathrm{km\,s^{-1}}$. We thus computed the ^{13}CO(1 − 0) integrated intensity inside these two intervals. This integrated intensity map has negative values in regions where the S/N is low. An estimation based on the negative intensity values leads us to conclude that we can only trust the integrated intensity inside the above velocity intervals when it is larger than 0.1 $\mathrm{K\,km\,s^{-1}}$. We thus set the integrated intensity to 0 when it is lower than 0.1 $\mathrm{K\,km\,s^{-1}}$. We then computed the intensity integrated inside $[2, 16]$ $\mathrm{km\,s^{-1}}$, and the contribution of the $[2, 7]$ $\mathrm{km\,s^{-1}}$ and $[14, 16]$ $\mathrm{km\,s^{-1}}$ velocity intervals to the ^{13}CO(1 − 0) intensity integrated inside $[2, 16]$ $\mathrm{km\,s^{-1}}$. This allowed us to compute the fraction of the ^{13}CO(1 − 0) integrated intensity coming from the $[7, 14]$ $\mathrm{km\,s^{-1}}$ velocity interval.

Second, part of the hydrogen may still be in atomic form. Looking at the HI emission toward the Orion B GMC, Pety et al. (2017) estimate that about 1 mag of visual extinction is coming from atomic hydrogen. We thus subtracted 1 mag to the total dust extinction, and we applied the previously computed fraction to this dust extinction corrected for the atomic contribution. This gives the amount of visual extinction coming from molecular gas in the $[7, 14]$ $\mathrm{km\,s^{-1}}$ velocity interval. As the low A_V part is still uncertain, we additionally nullified all the LoSs whose visual extinction associated with molecular gas is lower than 1 mag. In this study, we use this "molecular" visual extinction.

Appendix B: Radiative transfer equation for a M-layer cloud

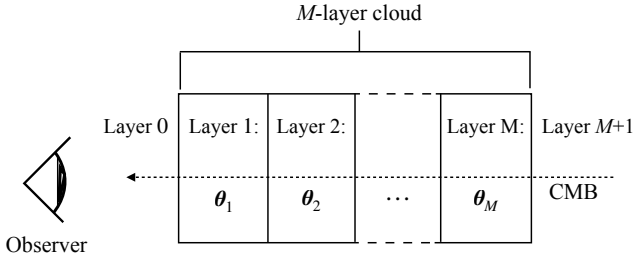


Fig. B.1. Sketch of a M-layer model of a line of sight. Each slab has homogeneous physical and chemical conditions (characterized by the parameter vectors $\boldsymbol{\theta}_i$, for $1 \leq i \leq M$). The emission from the CMB (layer $M+1$) happens at infinity, and it is assumed homogeneous and isotropic.

Let a M-layer cloud model of the LoS as depicted in B.1. A simple generalization of the computation for a sandwich model (including the ON-OFF calibration procedure) yields

$$s_M = \sum_{i=1}^{M} J(T_{\mathrm{ex},i}) \left[1 - \exp\left(-\Psi_i\right)\right] \exp\left(-\sum_{j=1}^{i-1} \Psi_j\right) - J(T_{\mathrm{CMB}}) \left[1 - \exp\left(-\sum_{k=1}^{M} \Psi_k\right)\right].$$

Assuming that there exist a layer 0 of opacity $\Psi_0 = 0$, we can rewrite this equation as

$$s_M = \sum_{m=0}^{M} \left[J(T_{\mathrm{ex},m+1}) - J(T_{\mathrm{ex},m})\right] \exp\left(-\sum_{k=0}^{m} \Psi_k\right), \tag{B.1}$$

where we have written $J(T_{\mathrm{CMB}}) = J(T_{\mathrm{ex},0}) = J(T_{\mathrm{ex},M+1})$ only for mathematical convenience (i.e., it is not physically interpretable). This alternate expression simplifies the computation of the Fisher matrix and Cramér-Rao bound.

Indeed, developing Eq. (B.1), we have

$$\begin{aligned} s_M = \; & J(T_{\mathrm{ex},1})\left[1-\exp\left(-\Psi_1\right)\right]\exp\left(-\Psi_0\right) \\ & + J(T_{\mathrm{ex},2})\left[1-\exp\left(-\Psi_2\right)\right]\exp\left(-\Psi_0-\Psi_1\right) \\ & + \cdots \\ & + J(T_{\mathrm{ex},M})\left[1-\exp\left(-\Psi_M\right)\right]\exp\left(-\sum_{j=0}^{M-1}\Psi_j\right) \\ & - J(T_{\mathrm{CMB}})\left[1-\exp\left(-\sum_{k=0}^{M}\Psi_k\right)\right]. \end{aligned}$$

Reorganizing the terms yields

$$\begin{aligned} s_M = \; & J(T_{\mathrm{ex},1})\exp\left(-\Psi_0\right) - J(T_{\mathrm{ex},1})\exp\left(-\Psi_1-\Psi_0\right) \\ & + J(T_{\mathrm{ex},2})\exp\left(-\Psi_0-\Psi_1\right) \\ & - J(T_{\mathrm{ex},2})\exp\left(-\Psi_2-\Psi_0-\Psi_1\right) \\ & + \cdots \\ & + J(T_{\mathrm{ex},M})\left(-\sum_{j=0}^{M-1}\Psi_j\right) - J(T_{\mathrm{ex},M})\exp\left(-\sum_{j=0}^{M}\Psi_j\right) \\ & - J(T_{\mathrm{CMB}}) + J(T_{\mathrm{CMB}})\exp\left(-\sum_{k=0}^{M}\Psi_k\right), \end{aligned} \tag{B.2}$$

and

$$\begin{aligned} s_M = \; & J(T_{\mathrm{ex},1})\exp\left(-\Psi_0\right) - J(T_{\mathrm{ex},1})\exp\left(-\Psi_1-\Psi_0\right) \\ & + J(T_{\mathrm{ex},2})\exp\left(-\Psi_0-\Psi_1\right) \\ & - J(T_{\mathrm{ex},2})\exp\left(-\Psi_2-\Psi_0-\Psi_1\right) \\ & + \cdots \\ & + J(T_{\mathrm{ex},M})\left(-\sum_{j=0}^{M-1}\Psi_j\right) - J(T_{\mathrm{ex},M})\exp\left(-\sum_{j=0}^{M}\Psi_j\right) \\ & - J(T_{\mathrm{ex},0}) + J(T_{\mathrm{ex},M+1})\exp\left(-\sum_{k=0}^{M}\Psi_k\right). \end{aligned} \tag{B.3}$$

By factorizing each term $\exp\left(-\sum_{k=0}^{m}\Psi_k\right)$ for $m \leqslant M$ in Eq. (B.3), we get

$$\begin{aligned} s_M = \; & \left[J(T_{\mathrm{ex},1}) - J(T_{\mathrm{ex},0})\right]\exp\left(-\Psi_0\right) \\ & + \left[J(T_{\mathrm{ex},2}) - J(T_{\mathrm{ex},1})\right]\exp\left(-\Psi_0-\Psi_1\right) \\ & + \cdots \\ & + \left[J(T_{\mathrm{ex},M+1}) - J(T_{\mathrm{ex},M})\right]\exp\left(-\sum_{k=0}^{m}\Psi_k\right). \end{aligned}$$

This form can be rewritten as Eq. (B.1).

Appendix C: Negative log-likelihood minimization

One technical challenge of the proposed sandwich model is to define an algorithm that converges within a reasonable amount of time to the parameter vector $\boldsymbol{\theta}$ that minimizes the NLL:

$$-\log\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{2}\sum_{l}\frac{\|\boldsymbol{x}_l - \boldsymbol{s}_l\|^2}{\sigma_{b,l}^2} + \text{constant}, \tag{C.1}$$

where the sum is performed on all the observed lines $\boldsymbol{x}_l$ and the sandwich modeled lines $\boldsymbol{s}_l(\boldsymbol{\theta})$, parametrized by the unknown vector $\boldsymbol{\theta}$ (see Eq. (1)). The challenge here is that the number of unknowns is much larger than in Roueff et al. (2024) for the same amount of observed information.

Lines of sight are processed independently of each other. The main steps of the minimization of Eq. C.1 are similar to these described in Roueff et al. (2024): An initial estimation of the centroid velocity is followed by a grid search, and finally by a gradient descent.

C.1. Initial estimation of the centroid velocity

We sum the lines of $C^{18}O$ and ^{13}CO to get a bright and approximately thin velocity spectrum. The velocity at which the intensity of the summed spectrum is maximum provides an initial estimation of the centroid velocity C_V for all the lines and for both the inner and outer layers.

C.2. Initial optimization with a grid search

We start by searching inside a grid in order to find a good first approximation of the vector $\boldsymbol{\theta}$ that minimizes the NLL. We detail here the case of the "improved" model (see Sect. 4.5.3). We need to explore a six-dimentional space $\left\{T_{\mathrm{kin}}, n_{\mathrm{H_2}}, N(^{13}\mathrm{CO}), N(\mathrm{HCO^+}), \mathrm{FWHM}, C_V\right\}$ for the inner layer, and a nine-dimentional space $\left\{T_{\mathrm{kin}}, n_{\mathrm{H_2}}, N(^{12}\mathrm{CO}), N(^{13}\mathrm{CO}), N(\mathrm{C^{18}O}), N(\mathrm{HCO^+}), \mathrm{FWHM}, C_V\right\}$ for the outer layer. Performing an exhaustive grid search is not an option because it would be too time-consuming: Testing ten values per dimension requires 10^{15} calculations of Eq. C.1. We thus used an iterative stochastic search in a gridded space, that is, a given number of "walkers" randomly explore the 15−dimensional space and converge to their closest local minima. Increasing the number of walkers allows the probability of getting trapped in a local minimum to be decreased, while increasing the number of walker steps is useful to ensure convergence.

To save CPU time, the opacity (τ_l) and excitation temperature ($T_{\mathrm{ex},l}$) of each molecular species l are pre-calculated with RADEX on fixed grids. As the sampling of parameter space is identical for the inner and outer layers, only one four-dimentional grid per species needs to be pre-computed. This avoids an excessive load on the RAM memory. More precisely, for each molecular line l, we pre-computed τ_l and $T_{\mathrm{ex},l}$ on the following regularly sampled grid:

- The H_2 volume density $n_{\mathrm{H_2}}$ is sampled between 10^2 and $10^{6.5}$ cm^{-3} in logarithmic steps of 0.1 (46 values).
- The kinetic temperature T_{kin} is sampled between 4 and 113 K in logarithmic steps of 0.05 (30 values).
- The column density of ^{13}CO is sampled between 10^{13} and 10^{18} cm^{-2} in logarithmic steps of 0.05 (102 values).
- The other column densities $N(^{12}CO)$, $N(C^{18}O)$, $N(HCO^+)$, $N(H^{13}CO^+)$ are sampled on the same grid except that they are shifted by the value fixed by the abundance ratios mentioned in Eq. (24).
- The FWHM is sampled linearly between 0.25 and 2 km s^{-1} in steps of 0.05 km s^{-1} (36 values).

During the random walk, the column densities are probed under the following constraints. For the inner layer,

$$\begin{aligned} & N(^{12}\mathrm{CO})/N(^{13}\mathrm{CO}) = 10^{1.8} \sim 63, \\ & N(^{13}\mathrm{CO})/N(\mathrm{C^{18}O}) = 10^{0.9} \sim 7.9, \\ & 10^{2.88} \sim 759 \leq N(^{13}\mathrm{CO})/N(\mathrm{HCO^+}) \leq 10^{3.78} \sim 6000, \\ & \text{and} \quad N(\mathrm{HCO^+})/N(^{12}\mathrm{CO}) = N(\mathrm{H^{13}CO^+})/N(^{13}\mathrm{CO}). \end{aligned} \tag{C.2}$$

For the outer layer, the constraints are less severe:

$$\begin{aligned} & 10^{1.3} \sim 20 \leq N(^{12}\mathrm{CO})/N(^{13}\mathrm{CO}) \leq 10^{1.9} \sim 80, \\ & 10^{0.75} \sim 5.6 \leq N(^{13}\mathrm{CO})/N(\mathrm{C^{18}O}) \leq 10^{1.4} \sim 25.1, \\ & 10^{2.88} \sim 759 \leq N(^{13}\mathrm{CO})/N(\mathrm{HCO^+}) \leq 10^{3.78} \sim 6000, \\ & \text{and} \quad N(\mathrm{HCO^+})/N(^{12}\mathrm{CO}) = N(\mathrm{H^{13}CO^+})/N(^{13}\mathrm{CO}). \end{aligned} \tag{C.3}$$

Empirically, we found that the following four-stage algorithm provides a satisfactory stochastic optimization in a reasonable amount of time (less than one minute per pixel on a 2024 laptop). At each stage, we select 10 % of the walkers to save CPU time.

1. We initialize the centroid velocities of the inner and outer layers to the value estimated in App. C.1, and we fix them to these values during stage 1 and 2. We then randomly draw the

$\boldsymbol{\theta}$ position of 10^6 walkers in the remaining 13−dimensional space, and we select the 10^5 walkers located at positions with the lowest NLL values. Indeed, these walkers are the ones who have the highest chance to be located close to the minimum.
2. Each selected walker moves in a randomly chosen direction with a randomly chosen step size to explore its vicinity. Its $\boldsymbol{\theta}$ position is then updated, if and only if the new position decreases the NLL value. Ten such steps are made during this exploratory stage, and at the end we select the 10^4 walkers whose position corresponds to the lowest NLL values.
3. Each selected walker now makes 10^3 steps. In this stage, the two dimensions of centroid velocities are added to the search space. The size of the step along the centroid velocity axes is set to ±10 m/s. At the end of the this stage, we select the 10^3 walkers located at positions with the lowest NLL values.
4. Each of the 10^3 remaining walkers makes 5000 new steps to finalize the convergence.

Obviously, these settings can be adjusted differently. And this algorithm does not provide any guarantee of convergence. To boost confidence in the estimated $\boldsymbol{\theta}$, one can run the procedure several times to check that the result remains unchanged. However, checking the number of walkers that have converged to the minimum NLL value found already gives a useful indication.

C.3. Final optimization with a gradient descent

The solution provided by the grid search can be refined with a gradient descent algorithm for which the Hessian matrix is replaced by the Fisher matrix. At each iteration i, the vector $\boldsymbol{\theta}$ is updated such as

$$\widehat{\boldsymbol{\theta}}^{i+1} = \widehat{\boldsymbol{\theta}}^{i} - \alpha \left(\boldsymbol{I}_F(\widehat{\boldsymbol{\theta}}^{i})\right)^{-1} \nabla_{\boldsymbol{\theta}} - \log \mathcal{L}(\widehat{\boldsymbol{\theta}}^{i}), \tag{C.4}$$

where $\boldsymbol{I}_F(\boldsymbol{\theta})$ is the Fisher Information Matrix computed at $\boldsymbol{\theta}$ without considering calibration noise (see Eq. (D.5)), $\nabla_{\boldsymbol{\theta}} - \log \mathcal{L}$ is the gradient

$$\nabla_{\boldsymbol{\theta}} - \log \mathcal{L}(\boldsymbol{\theta}) = \sum_l \frac{1}{\sigma_{b,l}^2} \frac{\partial \boldsymbol{s}_l}{\partial \boldsymbol{\theta}}^{\top} (\boldsymbol{x}_l - \boldsymbol{s}_l), \tag{C.5}$$

and $\alpha \in [0, 1]$ is a parameter to control the convergence speed. To estimate α, we compute $\mathcal{L}(\boldsymbol{\theta})$ for $\alpha = \{0.1, 0.5, 1\}$ and we make a quadratic fit to get an estimation of α that minimizes the NLL in the interval $[0.1, 1]$. The gradient descent algorithm is stopped when the difference of NLL between two consecutive iterations is smaller than 10^{-6} (our criterion convergence), or when the number of iterations reaches 30 (which is enough for the considered problem). During the gradient descent, the abundance ratios remain unchanged. If one decides to apply the gradient on all the searched column density, the constraint from Eq. C.2 and C.3 would be difficult to satisfy, but more important, the Fisher matrix could become singular.

Appendix D: Fisher matrix calculation

We compute the Fisher matrix for the M-layer model described in the previous section rather than to the limiting case of the sandwich model.

D.1. Fisher matrix as a function of gradients $\frac{\partial s_M}{\partial \theta_i}$

To compute the Fisher matrix for several observed lines $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_L$, we just need to sum the Fisher matrices of the different molecular lines

$$\boldsymbol{I}_F(\boldsymbol{\theta}; \sum_l \boldsymbol{x}_l) = \sum_l \boldsymbol{I}_F(\boldsymbol{\theta}, \boldsymbol{x}_l). \tag{D.1}$$

This section explains the computations required to yield the Fisher matrix for a single line $\boldsymbol{x}_l$.

As discussed in Sect. 3.6, we consider the case where abundance ratios with respect to ^{13}CO are fixed. The vector of unknowns can be written as

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M], \tag{D.2}$$

where each layer of index m is characterized by the vector of physical parameters

$$\boldsymbol{\theta}_m = [\log n_{\mathrm{H_2},m}, \log T_{\mathrm{kin},m}, \log N(^{13}\mathrm{CO})_m, \mathrm{FWHM}_m, C_{V,m}]. \tag{D.3}$$

The Fisher matrix is therefore of size $5M \times 5M$.

In their Appendix A.2, Roueff et al. (2024) showed that, when the calibration noise is taken into account (see Eq. (14)), the term i, j of the Fisher matrix can be written as

$$\begin{aligned}[\boldsymbol{I}_F(\boldsymbol{\theta}; \boldsymbol{x}_l)]_{ij} &= \left(\frac{\sigma_l^2 + \sigma_{c,l}^4 (\boldsymbol{s}_M^T \boldsymbol{s}_M)}{\sigma_{b,l}^2 \sigma_l^2}\right) \left(\frac{\partial \boldsymbol{s}_M^T}{\partial \theta_i} \frac{\partial \boldsymbol{s}_M}{\partial \theta_j}\right) \\ &+ \left(\frac{\sigma_{b,l}^2 \sigma_{c,l}^4 - \sigma_{c,l}^6 (\boldsymbol{s}^T \boldsymbol{s})}{\sigma_{b,l}^2 \sigma_l^4} - \frac{\sigma_{c,l}^2}{\sigma_{b,l}^2 \sigma_l^2}\right) \left(\boldsymbol{s}_M^T \frac{\partial \boldsymbol{s}_M}{\partial \theta_i}\right) \left(\boldsymbol{s}_M^T \frac{\partial \boldsymbol{s}_M}{\partial \theta_j}\right),\end{aligned} \tag{D.4}$$

where $\boldsymbol{s}_M$ is defined in Eq. (B.1), and $\sigma_l^2 = \sigma_{b,l}^2 + \sigma_{c,l}^2 (\boldsymbol{s}^T \boldsymbol{s})$. When the calibration noise is neglected, Eq. (D.4) simplifies into

$$[\boldsymbol{I}_F(\boldsymbol{\theta}; \boldsymbol{x}_l)]_{ij} = \frac{1}{\sigma_{b,l}^2} \left(\frac{\partial \boldsymbol{s}_M^T}{\partial \theta_i} \frac{\partial \boldsymbol{s}_M}{\partial \theta_j}\right). \tag{D.5}$$

In both cases, only the gradients $\frac{\partial \boldsymbol{s}_M^T}{\partial \theta_i}$ for all unknown parameters θ_i are needed to complete the computations of the Fisher matrix.

D.2. Computation of $\frac{\partial s_M}{\partial theta_i}$

We differentiate Eq. (B.1) to obtain for each component $\theta_i \in \boldsymbol{\theta}_i$, and for $1 \leq i \leq M$

$$\begin{aligned}\frac{\partial \boldsymbol{s}_M}{\partial \boldsymbol{\theta}_i} &= \frac{\partial J(T_{\mathrm{ex},i})}{\partial \theta_i} \exp\left(-\sum_{j=0}^{i-1} \Psi_j\right) \\ &- \frac{\partial J(T_{\mathrm{ex},i})}{\partial \theta_i} \exp\left(-\sum_{j=0}^{i} \Psi_j\right) \\ &- \sum_{m=i}^{M} \left[J(T_{\mathrm{ex},m+1}) - J(T_{\mathrm{ex},m})\right] \exp\left(-\sum_{k=0}^{m} \Psi_k\right) \frac{\partial \Psi_i}{\partial \theta_i}.\end{aligned}$$

Finally,

$$\begin{aligned}\frac{\partial \boldsymbol{s}_M}{\partial \boldsymbol{\theta}_i} &= \frac{\partial J(T_{\mathrm{ex},i})}{\partial \theta_i} \left[1 - \exp(-\Psi_i)\right] \exp\left(-\sum_{j=0}^{i-1} \Psi_j\right) \\ &+ \frac{\partial \Psi_i}{\partial \theta_i} \sum_{m=i}^{M} \left[J(T_{\mathrm{ex},m}) - J(T_{\mathrm{ex},m+1})\right] \exp\left(-\sum_{k=0}^{m} \Psi_k\right),\end{aligned} \tag{D.6}$$

where the terms $\frac{\partial J(T_{\mathrm{ex},i})}{\partial \theta_i}$ and $\frac{\partial \Psi_i}{\partial \theta_i}$ are expressed in D.2.1 and D.2.2 below.

D.2.1. Computation of $\frac{\partial J(T_{\mathrm{ex},i})}{\partial \theta_i}$

One has

$$\frac{\partial J(T_{\mathrm{ex},i})}{\partial \theta_i} = \frac{\partial J(T_{\mathrm{ex},i})}{\partial T_{\mathrm{ex},i}} \frac{\partial T_{\mathrm{ex},i}}{\partial \theta_i}.$$

Then,

$$\frac{\partial J(T_{\mathrm{ex},i})}{\partial \theta_i} = \begin{cases} 0 & \text{if} \quad \theta_i = C_{V,i}, \\ \left(\frac{h\nu}{kT_{\mathrm{ex},i}}\right)^2 \frac{\exp\left(\frac{h\nu}{kT_{\mathrm{ex},i}}\right)}{\left[\exp\left(\frac{h\nu}{kT_{\mathrm{ex},i}}\right)-1\right]^2} \frac{\partial T_{\mathrm{ex},i}}{\partial \theta_i} & \text{else.} \end{cases} \tag{D.7}$$

D.2.2. Computation of $\frac{\partial \Psi_i}{\partial \theta_i}$

One has $\frac{\partial \Psi_i}{\partial \theta_i} = \frac{\partial \Psi_i}{\partial \tau_{\mathrm{i}}} \frac{\partial \tau_{\mathrm{i}}}{\partial \theta_i}$. Then,

$$\frac{\partial \Psi_i}{\partial \theta_i} = \begin{cases} \left[\frac{\partial \tau_{\mathrm{i}}}{\partial \theta_i} + \tau_{\mathrm{i}} \frac{(V-C_{V,i})^2}{\sigma^3_{V,i}}\right] \exp\left(\frac{-(V-C_{V,i})^2}{2\sigma^2_{V,i}}\right) & \text{if} \quad \theta_i = \sigma_{V,i}, \\ \tau_{\mathrm{i}} \frac{(V-C_{V,i})}{\sigma^2_{V,i}} \exp\left(\frac{-(V-C_{V,i})^2}{2\sigma^2_{V,i}}\right) & \text{else if} \quad \theta_i = C_{V,i}, \\ \frac{\partial \tau_{\mathrm{i}}}{\partial \theta_i} \exp\left(\frac{-(V-C_{V,i})^2}{2\sigma^2_{V,i}}\right) & \text{else.} \end{cases} \tag{D.8}$$

We now detail how the terms $\frac{\partial T_{\mathrm{ex},i}}{\partial \theta_i}$ and $\frac{\partial \tau_{\mathrm{i}}}{\partial \theta_i}$ are computed. For $\theta_i \in \left\{\log n_{\mathrm{H}_2,i}, \log N(^{13}\mathrm{CO})_i\right\}$, $\frac{\partial T_{\mathrm{ex},i}}{\partial \theta_i}$ and $\frac{\partial \tau_{\mathrm{i}}}{\partial \theta_i}$ are numerically estimated with RADEX by a finite difference technique with a step of 0.001 as it done in Roueff et al. (2024).

For $\theta_i = \log T_{\mathrm{kin},i}$, one has

$$\log T_{\mathrm{kin},i} = \frac{\ln(T_{\mathrm{kin},i})}{\ln(10)}.$$

Then,

$$\frac{\partial T_{\mathrm{ex},i}}{\partial \log T_{\mathrm{kin},i}} = \ln(10)\ T_{\mathrm{kin},i} \frac{\partial T_{\mathrm{ex},i}}{\partial T_{\mathrm{kin},i}} \tag{D.9}$$

and

$$\frac{\partial \tau_{\mathrm{i}}}{\partial \log T_{\mathrm{kin},i}} = \ln(10)\ T_{\mathrm{kin},i} \frac{\partial \tau_{\mathrm{i}}}{\partial T_{\mathrm{kin},i}}, \tag{D.10}$$

where $\frac{\partial T_{\mathrm{ex},i}}{\partial T_{\mathrm{kin},i}}$ and $\frac{\partial \tau_{\mathrm{i}}}{\partial T_{\mathrm{kin},i}}$ are numerically estimated with RADEX by a finite difference technique with a step of $T_{\mathrm{kin},i}/1024$.

For $\theta_i = \sigma_{V,i}$, one has

$$\sigma_{V,i} = \frac{\mathrm{FWHM}_i}{\sqrt{8\ln 2}}.$$

Thus,

$$\frac{\partial T_{\mathrm{ex},i}}{\partial \sigma_{V,i}} = \sqrt{8\ln 2}\, \frac{\partial T_{\mathrm{ex},i}}{\partial \mathrm{FWHM}_i} \quad \text{and} \quad \frac{\partial \tau_{\mathrm{i}}}{\partial \sigma_{V,i}} = \sqrt{8\ln 2}\, \frac{\partial \tau_{\mathrm{i}}}{\partial \mathrm{FWHM}_i}, \tag{D.11}$$

where the terms $\frac{\partial T_{\mathrm{ex},i}}{\partial \mathrm{FWHM}_i}$ and $\frac{\partial \tau_{\mathrm{i}}}{\partial \mathrm{FWHM}_i}$ are numerically estimated with RADEX by a finite difference technique with a step of $\mathrm{FWHM}_i/128$.

For $\theta_i = C_{V,i}$, one has

$$\frac{\partial T_{\mathrm{ex},i}}{\partial C_{V,i}} = 0$$

and

$$\frac{\partial \tau_{\mathrm{i}}}{\partial C_{V,i}} = 0.$$

D.3. Application to the sandwich model

In the case of the sandwich model, $M = 3$ and $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_3$. One thus has

$$\frac{\partial \boldsymbol{s}_M}{\partial \boldsymbol{\theta}_{\mathrm{ou}}} = \frac{\partial \boldsymbol{s}_M}{\partial \boldsymbol{\theta}_1} + \frac{\partial \boldsymbol{s}_M}{\partial \boldsymbol{\theta}_3} \quad \text{and} \quad \frac{\partial \boldsymbol{s}_M}{\partial \boldsymbol{\theta}_{\mathrm{in}}} = \frac{\partial \boldsymbol{s}_M}{\partial \boldsymbol{\theta}_2},$$

where $\boldsymbol{\theta}_{\mathrm{ou}}$ and $\boldsymbol{\theta}_{\mathrm{in}}$ are the vector of unknowns of the outer and inner layers respectively.

Appendix E: Cramér-Rao lower bounds for derived quantities

E.1. Thermal pressure

Following Roueff et al. (2024), we compute the CRB on the logarithm of the thermal pressure $P_{\mathrm{ther},m}$ for each layer m of the sandwich model as

$$\begin{aligned} \mathrm{CRB}(\log P_{\mathrm{ther},m}) \quad &= \mathrm{CRB}(\log T_{\mathrm{kin},m}) \\ &+ \mathrm{CRB}(\log n_{\mathrm{H}_2,m}) \\ &+ 2\,\mathrm{CRB}(\log T_{\mathrm{kin},m}, \log n_{\mathrm{H}_2,m}), \end{aligned} \tag{E.1}$$

where $\mathrm{CRB}(\log T_{\mathrm{kin},m}, \log n_{\mathrm{H}_2,m})$ is the off-diagonal term of the Fisher matrix.

E.2. Centroid velocity difference $\Delta C_V = C_{V,in} - C_{V,ou}$

Let $C_{V,\mathrm{in}}$ and $C_{V,\mathrm{ou}}$, the centroid velocities of the inner and outer layers, respectively. The CRB on $\Delta C_V = C_{V,\mathrm{in}} - C_{V,\mathrm{ou}}$ is

$$\mathrm{CRB}(\Delta C_V) = \mathrm{CRB}(C_{V,\mathrm{in}}) + \mathrm{CRB}(C_{V,\mathrm{ou}}) - 2\,\mathrm{CRB}(C_{V,\mathrm{in}}, C_{V,\mathrm{ou}}), \tag{E.2}$$

where $\mathrm{CRB}(C_{V,\mathrm{in}}, C_{V,\mathrm{ou}})$ is the off-diagonal term of the Fisher matrix.

E.3. Total column density

We derive the logarithm of the total column density of a species X, noted $\log N_{\mathrm{tot}}(X)$, from the estimation $\log N_m(\mathrm{X})$ with

$$\log N_{\mathrm{tot}}(X) = \log \sum_m 10^{\log N_m(\mathrm{X})}. \tag{E.3}$$

This section details how we compute the Cramér-Rao lower bound for $\log N_{\mathrm{tot}}(X)$. The formula is generalized for any multilayer model in E.3.1 and we deal with the sandwich model case in E.3.2.

E.3.1. Case of an M-layer cloud

Each layer m of the multilayer model is characterized by a five-parameter vector

$$\boldsymbol{\theta}_m = \{\log n_{\mathrm{H}_2,m},\ \log T_{\mathrm{kin},m},\ \log N(^{13}\mathrm{CO})_m,\ \mathrm{FWHM}_m, C_{V,m}\}. \tag{E.4}$$

The Fisher matrix dimension is $5M \times 5M$. As explains by Kay (1997, Sect. 3.8, Eq. 3.30), one can then compute the lower bound on the variance of $\alpha = g(\boldsymbol{\theta})$ with

$$\mathrm{var}\,(\widehat{\alpha}) \geq \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathrm{CRB}(\boldsymbol{\theta}) \frac{\partial g(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}}, \tag{E.5}$$

where $\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is the Jacobian matrix. Thus, for the total column density with Eq. (E.3), one has $g(\boldsymbol{\theta}) = \log \sum_{m=0}^{M} 10^{\log N_m(\mathrm{X})}$, and

$$\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1} \cdots \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_{5M}} \right], \tag{E.6}$$

where

$$\frac{\partial g(\boldsymbol{\theta})}{\partial \theta_i} = \begin{cases} \frac{10^{\log N_m(\mathrm{X})}}{\sum_{m=0}^{M} 10^{\log N_m(\mathrm{X})}} & \text{if} \quad \theta_i = \log N_m(\mathrm{X}), \forall m \in [\![1, M]\!], \\ 0 & \text{else.} \end{cases} \tag{E.7}$$

E.3.2. Application to the sandwich model

Applying the result of the previous section leads to

$$\log N_{\mathrm{tot}}(X) = \log\left(2 \times 10^{\log N_1(\mathrm{X})} + 10^{\log N_2(\mathrm{X})}\right), \tag{E.8}$$

and

$$\frac{\partial g(\boldsymbol{\theta})}{\partial \theta_i} = \begin{cases} 2 \times \frac{10^{\log N_1(\mathrm{X})}}{\sum_{m=0}^{M} 10^{\log N_m(\mathrm{X})}} & \text{if} \quad \theta_i = \log N_1(\mathrm{X}), \\ \frac{10^{\log N_2(\mathrm{X})}}{\sum_{m=0}^{M} 10^{\log N_m(\mathrm{X})}} & \text{if} \quad \theta_i = \log N_2(\mathrm{X}), \\ 0 & \text{else.} \end{cases} \tag{E.9}$$

E.4. FWHM$_{in}$/FWHM$_{ou}$

We let $\mathrm{FWHM_{in}}$ and $\mathrm{FWHM_{ou}}$ be the FWHM of the inner and the outer layers, respectively. We use Eq. (E.5) to compute the CRB on the FWHM ratio. In this case, $g(\boldsymbol{\theta}) = \mathrm{FWHM_{in}}/\mathrm{FWHM_{ou}}$ and

$$\frac{\partial g(\boldsymbol{\theta})}{\partial \theta_i} = \begin{cases} 1/\mathrm{FWHM_{ou}} & \text{if} \quad \theta_i = \mathrm{FWHM_{in}}, \\ -\mathrm{FWHM_{in}}/\mathrm{FWHM_{ou}^2} & \text{if} \quad \theta_i = \mathrm{FWHM_{ou}}, \\ 0 & \text{else.} \end{cases} \tag{E.10}$$

Appendix F: Piecewise power-law fit of N_{tot} as a function of A_v

We aim at fitting power-laws through the total column density as a function A_V for different ranges of A_V. This means to make a linear regression in the log-log space, that is, fitting

$$y_n = \alpha\,(x_n - x_0) + \beta + b_n, \tag{F.1}$$

where $x = \log(A_v)$, $y = \log N_{\mathrm{tot}}$, and x_0 is the center of the $\log A_v$ bin and b_n is an additive noise drawn from a white Gaussian distribution. Equation (F.1) can be written:

$$\boldsymbol{y} = \boldsymbol{M}\,\boldsymbol{z} + \boldsymbol{b}, \tag{F.2}$$

where $\boldsymbol{y} = [y_1, ..., y_N]^T$, $\boldsymbol{z} = [\alpha, \beta]^T$, $\boldsymbol{b} = [b_1, ..., b_N]^T$ and

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{m}_1 \\ ... \\ \boldsymbol{m}_N \end{pmatrix} \quad \text{with} \quad \boldsymbol{m}_n = (x_n - x_0,\ 1). \tag{F.3}$$

Moreover, to take into account the uncertainty on $\log N_{\mathrm{tot}}$, we assume that the variance of b_n is proportional to the CRB of $\log N_{\mathrm{tot}}$. Thus, $\boldsymbol{B} \sim \mathcal{N}(0, \lambda \boldsymbol{R})$ where λ is unknown and $\boldsymbol{R}$ is a diagonal matrix with the CRB of $\log N_{\mathrm{tot}}$. The weighted least square estimator of $\boldsymbol{z}$ is then

$$\widehat{\boldsymbol{z}} = (\boldsymbol{M}^T \boldsymbol{R}^{-1} \boldsymbol{M})^{-1} \boldsymbol{M}^T \boldsymbol{R}^{-1} \boldsymbol{y}, \tag{F.4}$$

and λ can be estimated with

$$\widehat{\lambda} = \frac{1}{N} \sum_n \frac{\left(y_n - \boldsymbol{m}_n^T \widehat{\boldsymbol{z}}\right)^2}{R_n}, \tag{F.5}$$

and the estimation of the covariance of $\boldsymbol{z}$ is

$$\widehat{\boldsymbol{\Gamma}} = \widehat{\lambda}\left(\boldsymbol{M}^T \boldsymbol{R}^{-1} \boldsymbol{M}\right)^{-1}. \tag{F.6}$$

For a given x, the estimation of y is

$$\widehat{y} = \widehat{\alpha}\,(x - x_0) + \widehat{\beta} \tag{F.7}$$

and its variance is

$$\mathrm{var}(\widehat{y}) = \mathrm{var}(\widehat{\alpha})\,(x - x_0)^2 + \mathrm{var}(\widehat{\beta}) + 2\mathrm{cov}(\widehat{\alpha}, \widehat{\beta})\,(x - x_0) \tag{F.8}$$

where $\mathrm{var}(\widehat{\alpha}) = \left[\widehat{\boldsymbol{\Gamma}}\right]_{1,1}$ $\mathrm{var}(\widehat{\beta}) = \left[\widehat{\boldsymbol{\Gamma}}\right]_{2,2}$ $\mathrm{cov}(\widehat{\alpha}, \widehat{\beta}) = \left[\widehat{\boldsymbol{\Gamma}}\right]_{1,2}$.

Appendix G: TUH model in a nutshell

Table G.1. Values of the parameters in the TUH model of the abundance profile as a function of the H_2 column density.

Species	X_0	A_0/ mag	n_{fr}/ cm^{-3}
^{12}CO	9.5×10^{-5}	2	10^5
HCO^+	1.5×10^{-9}	2	2×10^5

For CO and HCO^+, (Tafalla et al. 2021) assume that the abundance averaged along the LoS, $\langle X \rangle$, is defined as

$$\langle X \rangle\,[N(\mathrm{H_2})] = X_0 \times f_{\mathrm{out}}\,[N(\mathrm{H_2})] \times f_{\mathrm{in}}\,[N(\mathrm{H_2})]\,, \tag{G.1}$$

where X_0 is a scaling constant, f_{out} mostly models the photodissociation of the species X by the far UV field in translucent gas, and f_{in} mostly models the molecular freeze-out onto dust grains in cold dense cores. They use the following simple analytic expressions

$$f_{\mathrm{out}}\,[N(\mathrm{H_2})] = \begin{cases} \exp\left[6\,(A_V - A_0)\right] & \text{if } A_V \leq A_0, \\ 1 & \text{if } A_V > A_0, \end{cases} \tag{G.2}$$

and

$$f_{\mathrm{in}}\,[N(\mathrm{H_2})] = \exp\left(-\frac{n_{\mathrm{H_2}}}{n_{\mathrm{fr}}}\right), \tag{G.3}$$

where A_0 and n_{fr} are the visual extinction and volume densities, which characterize the transition zone for the photodissociation and the molecular freeze-out, respectively. They estimate the volume density with

$$n_{\mathrm{H_2}} = 2 \times 10^4\ \mathrm{cm^{-3}} \left(\frac{N(\mathrm{H_2})}{10^{22}\ \mathrm{cm^{-2}}}\right)^{0.75}. \tag{G.4}$$

Appendix H: Supplementary figures

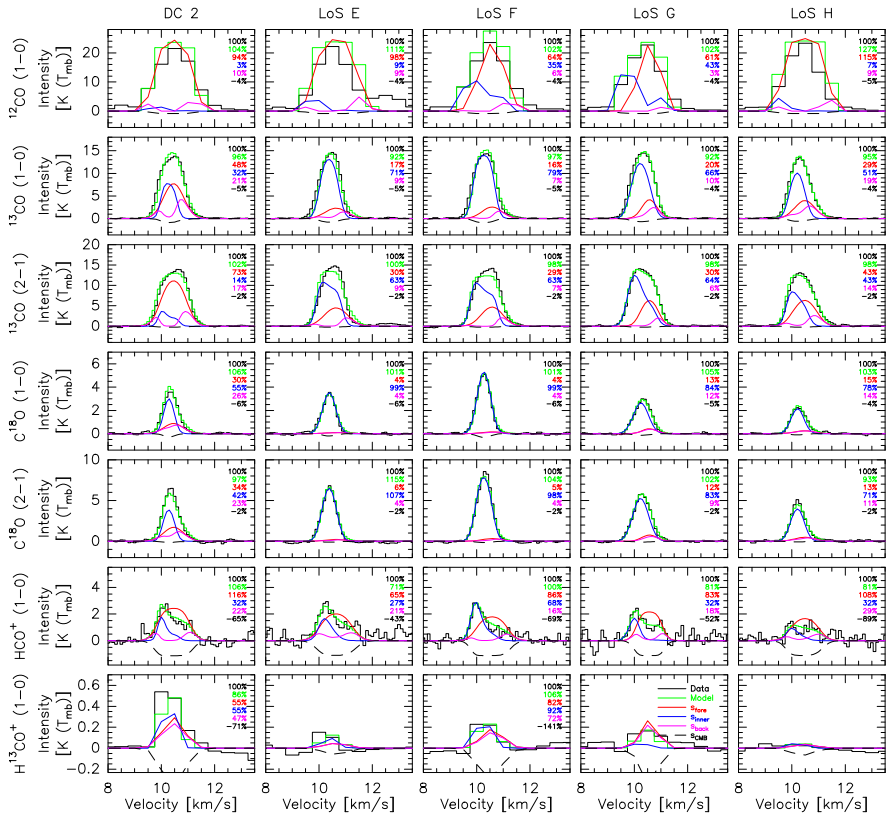


Fig. H.1. Same as Fig. 11 but for LoSs around the dense core 2.

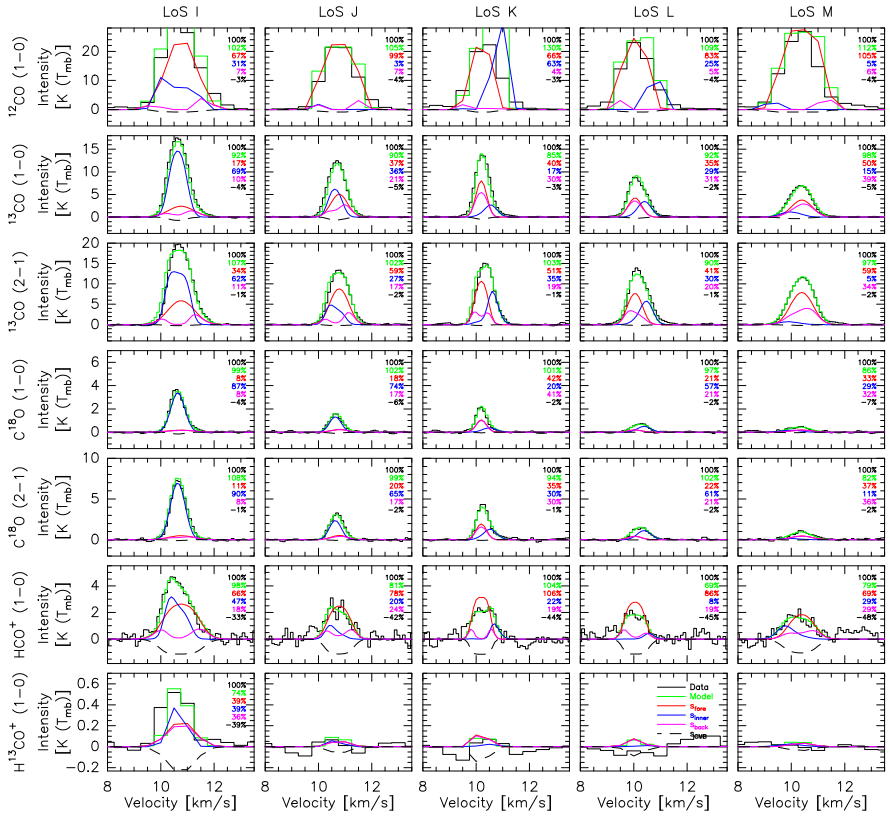


Fig. H.2. Same as Fig. 11 but for LoSs at the outskirt of the Horsehead nebula.

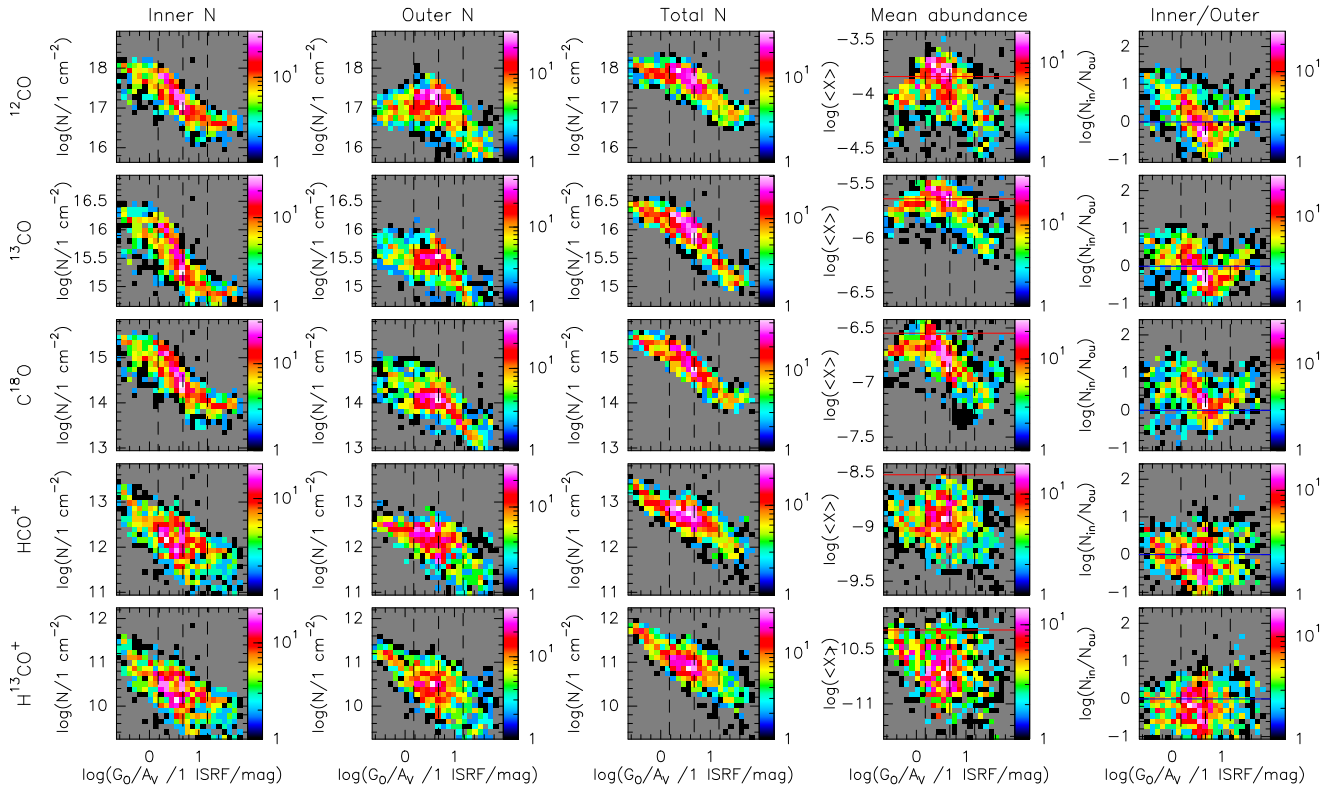


Fig. H.3. Joint histograms of the column densities, their abundances, or their ratios as a function of the visual extinction for the studied species. The dashed vertical lines mark the G_0/A_V values at 1.5, 4.75, and 15 ISRF/ mag. The remainder of the figure layout is identical to Fig. 14.

Chapter 5

Parameter estimation under multiplicative noise

Contents

5.1 Measurement model and statistical description of the multiplicative noise **139**
5.1.1 Measurement model . . . 139
5.1.2 Specifying the multiplicative noise . . . 139
5.2 Considering the multiplicative noise as unknown and deterministic . . . **141**
5.2.1 FIM and CRB . . . 142
5.3 Considering the multiplicative noise as random . . . **145**
5.3.1 Hybrid approach . . . 145
5.3.1.1 FIM and Hybrid CRB (HCRB) . . . 146
5.3.2 Marginal approach . . . 148
5.3.2.1 FIM and marginal CRB . . . 149
5.4 Neglecting the multiplicative noise . . . **149**
5.4.1 Likelihood under the misspecified model . . . 150
5.4.2 Knowledgeable case . . . 150
5.4.2.1 Asymptotic behavior of the MLE . . . 150
5.4.2.2 Unbiasedness of the MLE . . . 151
5.4.2.3 Misspecified CRB (MCRB) . . . 152
5.4.3 Naive case . . . 152
5.5 Quantitative comparison of the different CRBs in the case of a sight line toward a simulated dense core . . . **153**
5.6 Generalizing the comparison on actual GMC observations . . . **157**
5.7 Summary and conclusions . . . **157**
References . . . **158**
Appendix 5 . . . **160**
5.A Computation of the Kullback-Leibler divergence (KLD) . . . 160
5.B Derivative of composite function . . . 160
5.C Calculation of the FIM under the hybrid approach . . . 161
5.D Sherman-Morrison formula . . . 163
5.E Calculation of the MCRB . . . 163
5.F Quantification of the precision loss under the misspecified model . 165

Studies such as Roueff *et al.* 2021, 2024 shed light on the heterogeneous nature of GMCs along the line of sight. To improve our description of the medium, we extended the radiative transfer equation to multiple layers of gas and evaluated in Ségal *et al.* 2024 the particular "sandwich" configuration for analyzing the Horsehead Nebula observations (see Chap. 4). Although the "sandwich" remains a simplistic description of GMC dense core, the science results in Ségal *et al.* 2024 are promising ones. From the earliest studies Roueff *et al.* 2021, 2024 to this latter one, the calibration uncertainty, which can be seen as a multiplicative noise, nevertheless remains neglected when fitting the model to the measurements. Here, we ask whether taking this multiplicative noise into account during the fit could improve the estimate accuracy of the physical parameters of interest?

To illustrate our goals, let us first consider a simple astrophysical measurement x,

$$x = c\, s(\boldsymbol{\theta}) + b, \tag{5.1}$$

where s is the underlying physical model parameterized by $\boldsymbol{\theta}$, b is an additive noise, and c is an unknown calibration scalar. Throughout this thesis, we adopt a frequentist point of view for the physical parameter $\boldsymbol{\theta}$, i.e., by treating it as *deterministic* unknown. Because of the additive noise b, the measurement x is uncertain and is therefore treated as a *random* variable whose statistical properties derive from the random variable associated with b and the term $c\, s(\boldsymbol{\theta})$, where $s(\boldsymbol{\theta})$ is deterministic.

From this point, several approaches can be considered to treat c. First, this calibration factor can be merely neglected during the fit by assuming that $c = 1$. This is the approach followed in the studies cited above. A second approach consists in considering c as a deterministic unknown to estimate in addition to $\boldsymbol{\theta}$ during the fit. A third approach is a Bayesian one, where c is considered as a *random* variable described by some *a priori* p.d.f. (for instance, with a mean of 1 and a standard deviation σ_c). In this chapter, we propose to compare the precision obtained on the estimation of $\boldsymbol{\theta}$ in these three approaches. As addressed in Sect. 3.2, the CRB will serve as a *reference* precision. We shall also check whether neglecting the calibration factor during the fit biases our estimator of $\boldsymbol{\theta}$.

After introducing the measurement model and specifying the calibration factor in Sect. 5.1, we first study the precision when considering c as a deterministic unknown in Sect. 5.2 (second approach mentioned above). In Sect. 5.3 is addressed the Bayesian approach, where c is considered random. The first approach is studied in Sect. 5.4, where we show that neglecting the calibration noise during the fit does not bias the estimation of θ, before computing the corresponding reference precision. The different approaches are then compared in Sect. 5.5 on a simulated dataset. Finally, Appendix 5.F contains a GRETSI article in which further investigations of the estimator's performance are carried out on the Horsehead Nebula dataset.

5.1 Measurement model and statistical description of the multiplicative noise

5.1.1 Measurement model

Here, we remind the measurement model used to analyze the ORION-B dataset. Let us denote $\boldsymbol{y} = \{\boldsymbol{x}_1 \cdots \boldsymbol{x}_L\}$ the set of L molecular lines, where each molecular line $\boldsymbol{x} = (x_1 \cdots x_N)^\top \in \boldsymbol{y}$ is measured over N velocity channels toward one line of sight. As proposed in Einig *et al.* 2023, we assume that the observation $\boldsymbol{x}$ is a realization of the random variable $\boldsymbol{X}$,

$$\boldsymbol{X} = C\,\boldsymbol{s}(\boldsymbol{\theta}) + \boldsymbol{B}, \tag{5.2}$$

where $\boldsymbol{s} = (s_1 \cdots s_N)^\top$ is the spectrum built with a physical model (e.g., the sandwich), function of the unknown parameter $\boldsymbol{\theta} \in \mathbb{R}^M$ that characterizes the cloud. The randomness of $\boldsymbol{X}$ originates from two statistically independent sources: an additive noise $\boldsymbol{B} = (B_1 \cdots B_N)^\top \in \mathbb{R}^N$ and a multiplicative noise $C \in \mathbb{R}$. The additive noise is dominated by the atmospheric contribution, which is added to the signal before it reaches the detector, and then removed during the processing. The variable $B_n \in \boldsymbol{B}$ is a white Gaussian noise, i.e., $B_n \sim \mathcal{N}(0,\, \sigma_b^2)$ (follows a Gaussian distribution with zero mean and variance σ_b^2) such that $\boldsymbol{B} \sim \mathcal{N}(\boldsymbol{0}_N,\, \sigma_b^2 \boldsymbol{I}_N)$, where $\boldsymbol{0}_N = (0 \cdots 0)^\top$ is of size N and $\boldsymbol{I}_N$ is the identity matrix of size N. The measurement chain is made of successive components, each of them applying an uncertain multiplicative gain. The multiplicative noise C is therefore referred to as the *calibration* noise.

As explained in Appendix D of Einig *et al.* 2023, the observer can repeatedly point the telescope at a reference line of sight that is used as a primary calibrator. This specific line of sight is chosen to achieve a high signal-to-noise ratio (S/N) in a short integration time so that the additive noise can be neglected compared to the multiplicative noise. Moreover, the signal measured toward the reference line of sight is assumed to be constant during the experiment. When these two assumptions are fulfilled, the variation of the integrated emission is due only to calibration errors. Repeating this process dozens of times (under various weather and observation conditions) allows one to estimate the empirical p.d.f. of the calibration noise, and in particular its standard deviation. Nevertheless, the absolute calibration accuracy of the telescope is more challenging to estimate since no source of human-controlled power exists yet outside the atmosphere. In this thesis, we will assume that the absolute calibration of the IRAM-30m is perfect, i.e., the mean of the calibration noise is equal to 1.

5.1.2 Specifying the multiplicative noise

To explain the proposed statistical description of the calibration noise, let us consider a single molecular line measurement $\boldsymbol{x}$, drawn from Eq. (5.2). Following the reasoning presented in Galton 1879, the logarithm of the calibration noise, which can be read as the sum of statistically independent random processes, shows a distribution converging to a normal distribution under the central limit theorem. The lognormal distribution would therefore appear as a natural choice for describing multiplicative variable such as our calibration noise.

Einig *et al.* 2023 used the ORION-B observation strategy to estimate the standard deviation of the calibration noise, denoted by σ_c. They find $\sigma_c = \{0.05, 0.1\}$ for molecular lines at the 3 mm and 1 mm atmospheric window, respectively. The mean of the calibration noise, denoted by μ_c, is assumed to be 1 as the absolute calibration of the IRAM-30m is considered perfect. Figure 5.1, (a) displays in solid line lognormal distributions of the calibration noise, i.e., $C \sim \text{Log-}\mathcal{N}(\mu, \sigma^2)$, denoted by p, and defined by

$$p(c) = \frac{1}{c\,\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln c - \mu)^2}{2\sigma^2}\right\},$$

where the mean μ and the standard deviation σ are chosen such that $\mathbb{E}[C] = \mu_c$ and $\mathbb{V}[C] = \sigma_c^2$, where $\mathbb{V}[C] = \mathbb{E}[(C - \mathbb{E}[C])^2]$ is the variance of C. One has (see e.g., Wikipedia contributors 2025)

$$\mu = \ln\left(\frac{\mu_c^2}{\sqrt{\mu_c^2 + \sigma_c^2}}\right) \quad \text{and} \quad \sigma = \sqrt{\ln\left(1 + \frac{\sigma_c^2}{\mu_c^2}\right)}. \tag{5.3}$$

We observe that normal distributions, denoted by q, with mean μ_c and standard deviation σ_c, i.e.,

$$q(c) = \frac{1}{\sigma_c\sqrt{2\pi}} \exp\left\{-\frac{(c - \mu_c)^2}{2\sigma_c^2}\right\},$$

seem close to the lognormal distributions (see the dashed curves in Fig. 5.1a). In the following, we propose to quantify the relevance of the Gaussian approximation of the lognormal distribution using the *Kullback-Leibler* divergence (KLD; Kullback and Leibler 1951). Assuming that observations originate from the *true* distribution p, but are described with an approximated distribution q, the KLD, denoted by $\mathcal{K}[p \,\|\, q]$, which reads

$$\mathcal{K}[p \,\|\, q] = \int p(c) \log_2\left(\frac{p(c)}{q(c)}\right) \mathrm{d}c, \tag{5.4}$$

allows one to quantify the divergence between the distributions q and p. In Eq. (5.4), we choose the base 2 logarithm instead of the natural logarithm, since it offers an intuitive interpretation of the KLD as the number of supplementary bits required to encode a message drawn from p with a code designed for q, as explained in Shannon 1948 for the field of the information theory. One notes that the KLD is an asymmetrical metric, i.e., $\mathcal{K}[p \,\|\, q] \neq \mathcal{K}[q \,\|\, p]$. Didactic applications of the KLD are presented in the book Burnham and Anderson 2002, Sect. 2.1.1, for the cases of measuring the divergence between a gamma distribution and a Weibull, lognormal, inverse Gaussian and F distribution.

In our problem, we estimate the integral in Eq. (5.4) using a simulated random variable C as detailed in Appendix 5.A. Figure 5.1, (b) shows $\mathcal{K}[p \,\|\, q]$ as a function of the standard deviation σ_c. We measure for the particular values characterizing the ORION-B dataset $\sigma_c = 0.05$ and $\sigma_c = 0.1$, a KLD of ~ 0.003 and ~ 0.011 bits, respectively. The KLD increases linearly beyond $\sigma_c \simeq 0.3$. In Ségal *et al.* 2025 (see Sect. 5.F), we illustrate that the asymmetric profile of the lognormal distribution becomes apparent for $\sigma_c \gtrsim 0.3$. For analyzing the ORION-B dataset, we thus choose the Gaussian distribution to describe the calibration noise.

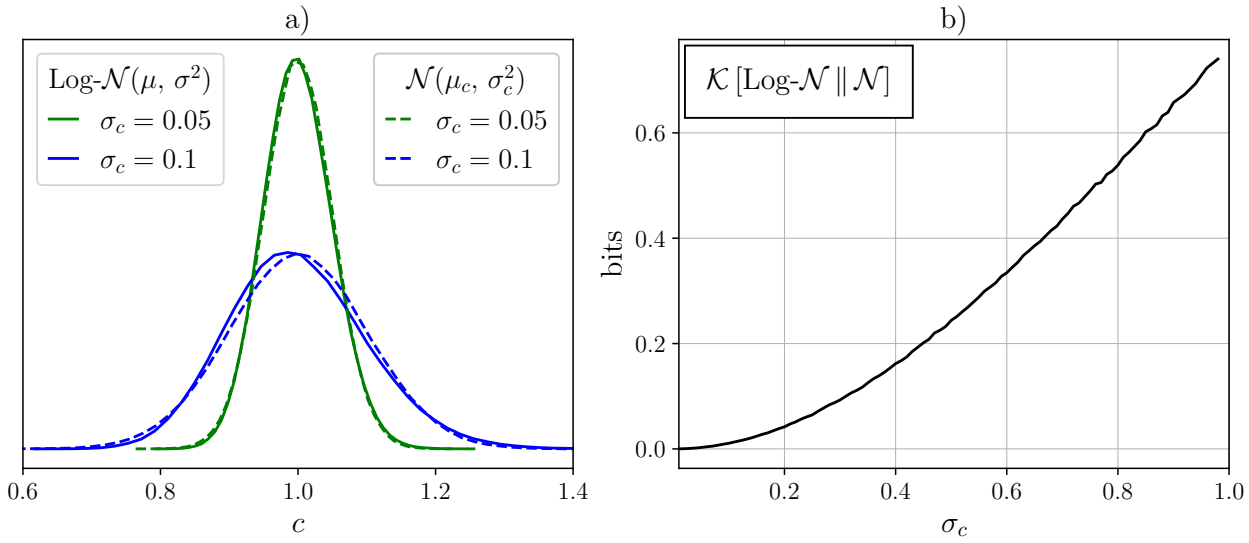


Figure 5.1: Comparison between the lognomal and normal distribution. a) Solid curves: Log-normal distributions (Log-$\mathcal{N}$) with mean μ and standard deviation σ derived from Eq. (5.3), with $\mu_c = 1$ and $\sigma_c = \{0.05, 0.1\}$ (in green and blue, respectively). Dashed curves: Normal distributions ($\mathcal{N}$) with mean $\mu_c = 1$ and standard deviation $\sigma_c = \{0.05, 0.1\}$ (in green and blue, respectively). b) Kullback-Leibler divergence ($\mathcal{K}$) between Log-$\mathcal{N}$ and $\mathcal{N}$. Each point has been computed with a sample of size $K = 10^6$ (see Appendix 5.A).

5.2 Considering the multiplicative noise as unknown and deterministic

In this section, we examine a first possible approach for constructing the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$. In particular, we compute the CRB as the reference precision of the estimator. Since molecular lines are assumed to be statistically independent, the FIM of the set is obtained by adding the FIM of all molecular lines. The case where the calibration noises of some molecular lines are correlated is beyond the scope of this thesis but remains worth considering regarding the application (e.g., measuring simultaneously multiple molecular lines over the same bandwidth). Hereafter, the vector of calibration noises of all the molecular lines in the set is denoted by $\boldsymbol{C} = (C_1 \cdots C_L)^\top$.

Although $\boldsymbol{C}$ is assumed random in the measurement model, we consider in this first approach no knowledge about the calibration noises, which are therefore treated as *deterministic* unknowns. The vector of the parameters that we wish estimating is then deterministic, and denoted by

$$(\boldsymbol{\theta}^\top \boldsymbol{c}^\top)^\top = (\theta_1 \cdots \theta_M \, c_1 \cdots c_L)^\top \in \mathbb{R}^{M+L},$$

where the deterministic nature of the calibration noises is emphasized by using lowercase notation. Let us consider the molecular line $\boldsymbol{x} \in \boldsymbol{y}$, bearing the multiplicative noise c, the corresponding component of $\boldsymbol{c}$. The additive noise being Gaussian, the measurements are normally distributed, i.e., $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ with mean $\boldsymbol{\mu}_1$ and covariance matrix $\boldsymbol{\Sigma}_1$,

$$\boldsymbol{\mu}_1(\boldsymbol{\theta}, c) = c\,\boldsymbol{s}(\boldsymbol{\theta}) \quad \text{and} \quad \boldsymbol{\Sigma}_1 = \sigma_b^2\,\boldsymbol{I}_N, \tag{5.5}$$

and described by the p.d.f., parametrized by $\boldsymbol{\theta}$ and c,

$$\begin{aligned} p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) &= \frac{1}{\sqrt{\det\left(2\pi\,\boldsymbol{\Sigma}_1\right)}} \exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}_1^\top \boldsymbol{\Sigma}_1^{-1}\boldsymbol{\varepsilon}_1\right\} \\ &= \frac{1}{\sqrt{2\pi\,\sigma_b^2}^N} \exp\left\{-\frac{1}{2\,\sigma_b^2}\,\boldsymbol{\varepsilon}_1^\top\boldsymbol{\varepsilon}_1\right\}, \end{aligned} \tag{5.6}$$

where $\boldsymbol{\varepsilon}_1(\boldsymbol{\theta}, c) = \boldsymbol{x} - \boldsymbol{\mu}_1(\boldsymbol{\theta}, c)$. When evaluated for a given observation $\boldsymbol{x}$ and as a function of $\boldsymbol{\theta}$ and c, this p.d.f. is called the *likelihood* function. Then follows a more convenient expression to manipulate, known as the *log-likelihood* function:

$$\ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) = \text{cst} - \frac{1}{2\,\sigma_b^2}\,\boldsymbol{\varepsilon}_1^\top\boldsymbol{\varepsilon}_1.$$

Besides, the log-likelihood gradient can be written

$$\nabla_{\boldsymbol{\theta},\boldsymbol{c}} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) = \left(\nabla_{\boldsymbol{\theta}}^\top \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) \quad \nabla_{\boldsymbol{c}}^\top \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)\right)^\top, \tag{5.7}$$

where $\nabla_{\boldsymbol{\theta},\boldsymbol{c}}$ denotes the gradient operator over the unknown parameters $\boldsymbol{\theta}$ and $\boldsymbol{c}$, i.e.,

$$\nabla_{\boldsymbol{\theta},\boldsymbol{c}} = \left(\frac{\partial}{\partial\theta_1} \cdots \frac{\partial}{\partial\theta_M}\, \frac{\partial}{\partial c_1} \cdots \frac{\partial}{\partial c_L}\right)^\top.$$

5.2.1 Fisher information matrix and Cramér-Rao bound

Case of a single molecular line — Following standard statistical methods as the one detailed in Kay 1993, chapter 3.7, the FIM of the considered molecular line in this first approach, denoted by $\boldsymbol{\mathcal{I}}_{1,\boldsymbol{X}}$, is the matrix of size $(M+L)\times(M+L)$, defined by

$$\boldsymbol{\mathcal{I}}_{1,\boldsymbol{X}}(\boldsymbol{\theta}, \boldsymbol{c}) = \mathbb{E}\left[\nabla_{\boldsymbol{\theta},\boldsymbol{c}} \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}, c)\, \nabla_{\boldsymbol{\theta},\boldsymbol{c}}^\top \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}, c)\right], \tag{5.8}$$

where $\mathbb{E}\left[f(\boldsymbol{X})\right]$ is the expectation of $f(\boldsymbol{x})$ evaluated over all the realizations $\boldsymbol{x}$, i.e.,

$$\mathbb{E}\left[f(\boldsymbol{X})\right] = \int f(\boldsymbol{x})\, p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)\, \mathrm{d}\boldsymbol{x}.$$

Since the molecular lines have different calibration noise, one notes that

$$\forall l \in [\![1, L]\!] \text{ such that } c_l \neq c, \quad \frac{\partial \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)}{\partial c_l} = 0.$$

The explicit form of the log-likelihood gradient defined in Eq. (5.7) is therefore

$$\nabla_{\boldsymbol{\theta},\boldsymbol{c}} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) = \Big(\underbrace{\nabla_{\boldsymbol{\theta}}^\top \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)}_{M} \quad \underbrace{0 \cdots 0 \quad \tfrac{\partial \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)}{\partial c} \quad 0 \cdots 0}_{L}\Big)^\top, \tag{5.9}$$

where

$$\nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) = \frac{c}{\sigma_b^2}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\boldsymbol{\varepsilon}_1 \quad \text{and} \quad \frac{\partial \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)}{\partial c} = \frac{1}{\sigma_b^2}\,\boldsymbol{s}^\top\boldsymbol{\varepsilon}_1. \tag{5.10}$$

The calculation details are provided in Appendix 5.B. From Eq. (5.8) and Eq. (5.9), the FIM is

$$\boldsymbol{\mathcal{I}}_{1,\boldsymbol{X}}(\boldsymbol{\theta},\boldsymbol{c}) = \mathbb{E}\left[\begin{array}{c|ccccccc} \boldsymbol{A}_{1,\boldsymbol{X}} & 0 & \cdots & 0 & \boldsymbol{E}_{1,\boldsymbol{X}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & & 0 & & \cdots & 0 \\ \vdots & \vdots & & & \vdots & & & \vdots \\ 0 & & & & 0 & & & \\ \boldsymbol{E}_{1,\boldsymbol{X}}^\top & 0 & \cdots & 0 & d_{1,\boldsymbol{X}} & 0 & \cdots & 0 \\ 0 & & & & 0 & & & \\ \vdots & \vdots & & & \vdots & & & \vdots \\ 0 & 0 & \cdots & & 0 & & \cdots & 0 \end{array}\right] \begin{array}{l} \left.\begin{array}{l}\\\\\\\\\end{array}\right\} M \\ \left.\begin{array}{l}\\\\\\\\\\\\\\\\\end{array}\right\} L \end{array}, \tag{5.11}$$

$$\underbrace{\qquad}_{M} \qquad \underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{L}$$

where $\boldsymbol{A}_{1,\boldsymbol{X}} \in \mathbb{R}^{M\times M}$ reads

$$\begin{aligned} \boldsymbol{A}_{1,\boldsymbol{X}} &= \nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta},c)\, \nabla_{\boldsymbol{\theta}}^\top \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta},c) \\ &= \frac{c^2}{\sigma_b^4}\, \nabla_{\boldsymbol{\theta}} \boldsymbol{s}\, \boldsymbol{\varepsilon}_1\, \boldsymbol{\varepsilon}_1^\top \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s}, \end{aligned} \tag{5.12}$$

and $\boldsymbol{E}_{1,\boldsymbol{X}} \in \mathbb{R}^M$ is

$$\begin{aligned} \boldsymbol{E}_{1,\boldsymbol{X}} &= \nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta},c)\, \frac{\partial^\top \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta},c)}{\partial c} \\ &= \frac{c}{\sigma_b^4}\, \nabla_{\boldsymbol{\theta}} \boldsymbol{s}\, \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^\top \boldsymbol{s}, \end{aligned} \tag{5.13}$$

and $d_{1,\boldsymbol{X}} \in \mathbb{R}$ is

$$\begin{aligned} d_{1,\boldsymbol{X}} &= \frac{\partial \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta},c)}{\partial c}\, \frac{\partial^\top \ln p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta},c)}{\partial c} \\ &= \frac{1}{\sigma_b^4}\, \boldsymbol{s}^\top \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^\top \boldsymbol{s}. \end{aligned} \tag{5.14}$$

Since $\mathbb{E}\left[\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^\top\right]$ is the covariance matrix $\boldsymbol{\Sigma}_1 = \sigma_b^2\, \boldsymbol{I}_N$ defined in Eq. (5.5), performing the expectation on Eq. (5.12), Eq. (5.13), and Eq. (5.14) leads to

$$\mathbb{E}\left[\boldsymbol{A}_{1,\boldsymbol{X}}\right] = \frac{c^2}{\sigma_b^2}\, \nabla_{\boldsymbol{\theta}} \boldsymbol{s}\, \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s}, \tag{5.15}$$

$$\mathbb{E}\left[\boldsymbol{E}_{1,\boldsymbol{X}}\right] = \frac{c}{\sigma_b^2}\, \nabla_{\boldsymbol{\theta}} \boldsymbol{s}\, \boldsymbol{s}, \tag{5.16}$$

and

$$\mathbb{E}\left[d_{1,\boldsymbol{X}}\right] = \frac{1}{\sigma_b^2}\, \boldsymbol{s}^\top \boldsymbol{s}. \tag{5.17}$$

Case of multiple molecular lines — Let us consider the set of L molecular lines $\boldsymbol{y} = \{\boldsymbol{x}_1 \cdots \boldsymbol{x}_L\}$ as a particular realization of $\boldsymbol{Y} = \{\boldsymbol{X}_1 \cdots \boldsymbol{X}_L\}$. As the molecular lines are assumed to be statistically independent, the FIM of the set, denoted by $\boldsymbol{\mathcal{I}}_1$, is obtained by summing the FIMs of all the lines, i.e.,

$$\boldsymbol{\mathcal{I}}_1(\boldsymbol{\theta}, \boldsymbol{c}) = \sum_{\boldsymbol{X} \in \boldsymbol{Y}} \boldsymbol{\mathcal{I}}_{1,\boldsymbol{X}}(\boldsymbol{\theta}, \boldsymbol{c}),$$

where $\boldsymbol{\mathcal{I}}_{1,\boldsymbol{X}}$ is defined in Eq. (5.11). The FIM is then

$$\boldsymbol{\mathcal{I}}_1(\boldsymbol{\theta}, \boldsymbol{c}) = \mathbb{E}\left[\begin{array}{c|ccc} \sum_{\boldsymbol{X} \in \boldsymbol{Y}} \boldsymbol{A}_{1,\boldsymbol{X}} & \boldsymbol{E}_{1,\boldsymbol{X}_1} & \cdots & \boldsymbol{E}_{1,\boldsymbol{X}_L} \\ \hline \boldsymbol{E}_{1,\boldsymbol{X}_1}^\top & d_{1,\boldsymbol{X}_1} & 0 \cdots & 0 \\ & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \boldsymbol{E}_{1,\boldsymbol{X}_L}^\top & 0 & \cdots 0 & d_{1,\boldsymbol{X}_L} \end{array}\right],$$

where, from Eq. (5.15), Eq. (5.16), and Eq. (5.17), one obtains

$$\mathbb{E}\left[\sum_{\boldsymbol{X} \in \boldsymbol{Y}} \boldsymbol{A}_{1,\boldsymbol{X}}\right] = \sum_{l=1}^{L} \frac{c_l^2}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \, \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s}_l,$$

and

$$\forall l \in [\![1, L]\!], \quad \mathbb{E}\left[\boldsymbol{E}_{1,\boldsymbol{X}_l}\right] = \frac{c_l}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \, \boldsymbol{s}_l, \quad \text{and} \quad \mathbb{E}\left[d_{1,\boldsymbol{X}_l}\right] = \frac{1}{\sigma_{b,l}^2} \boldsymbol{s}_l^\top \boldsymbol{s}_l.$$

Cramér-Rao bound — The Cramér-Rao bound (CRB), denoted by $\boldsymbol{\mathcal{B}}_1$, is the inverse matrix of the FIM (Kay 1993, chapter 3), i.e., $\boldsymbol{\mathcal{B}}_1(\boldsymbol{\theta}, \boldsymbol{c}) = \boldsymbol{\mathcal{I}}_1^{-1}(\boldsymbol{\theta}, \boldsymbol{c})$. In this first approach where multiplication noises are estimated in addition of $\boldsymbol{\theta}$, the CRB can be written as composed of four block matrices,

$$\boldsymbol{\mathcal{B}}_1(\boldsymbol{\theta}, \boldsymbol{c}) = \begin{pmatrix} \boldsymbol{\mathcal{B}}_1(\boldsymbol{\theta}) & [\boldsymbol{\mathcal{B}}_1]_{1,2} \\ [\boldsymbol{\mathcal{B}}_1]_{1,2}^\top & \boldsymbol{\mathcal{B}}_1(\boldsymbol{c}) \end{pmatrix},$$

where $\boldsymbol{\mathcal{B}}_1(\boldsymbol{\theta}) \in \mathbb{R}^{M \times M}$, $\boldsymbol{\mathcal{B}}_1(\boldsymbol{c}) \in \mathbb{R}^{L \times L}$, and $[\boldsymbol{\mathcal{B}}_1]_{1,2} \in \mathbb{R}^{M \times L}$. The matrix $\boldsymbol{\mathcal{B}}_1(\boldsymbol{\theta})$ corresponds to the CRB on the physical parameter $\boldsymbol{\theta}$ that interests us.

5.3 Considering the multiplicative noise as random

In the previous section, we adopted a frequentist perspective on the calibration noise, i.e., we considered it as a deterministic parameter to estimate in addition of the physical parameters of interest. Another approach consists in taking into account the available *a priori* knowledge about the calibration noise: as addressed in Sect. 5.1, it behaves as a Gaussian variable of mean 1, with a standard deviation that can be measured, i.e., $C \sim \mathcal{N}(\mu_c = 1, \sigma_c^2)$. This is the Bayesian perspective that is adopted in the section. In Sect. 5.3.1, we first compute a CRB that is said to be *hybrid* (HCRB) because the parameters of the physical model remain determinist. While the HCRB might provide a lower bound on the precision of $\boldsymbol{\theta}$ than the CRB, it is unclear whether an unbiased estimator exists whose the covariance matrix actually reaches the HCRB. Fortunately, the calibration noise is a *nuisance* parameter and we shall demonstrate that in our case, its estimation is not required to recover information on the physical parameters. To do this, we *marginalize* it in Sect. 5.3.2 and we compute the CRB associated with the p.d.f. of the pair of variables $(\boldsymbol{X}, C)$ integrated over C. Not only the corresponding CRB will give a reference precision on $\boldsymbol{\theta}$ that takes into account the *a priori* information on C, but more importantly, the asymptotic behavior of the MLE remains well established in this deterministic framework. The FIMs obtained under the "hybrid" and "marginal" approaches will be denoted by $2a$ and $2b$, respectively.

5.3.1 Hybrid approach

In this approach, the vector of parameters to estimate is so-called hybrid because is composed of both deterministic (i.e., $\boldsymbol{\theta}$) and random (i.e., $\boldsymbol{C}$) parameters:

$$(\boldsymbol{\theta}^\top \boldsymbol{C}^\top)^\top = (\theta_1 \cdots \theta_M \, C_1 \cdots C_L)^\top \in \mathbb{R}^{M+L}.$$

Joint p.d.f. — Following the approach adopted in the previous section, let us start by considering a single molecular line $\boldsymbol{x} \in \boldsymbol{y}$, with the corresponding multiplicative noise $C \sim \mathcal{N}(\mu_c = 1,\, \sigma_c^2)$. Throughout this manuscript, we preferred to retain μ_c as a variable rather than set it at 1, because it helped us to check the homogeneity of the expressions when carrying out the calculations, and it makes them usable in other contexts. The measurements are described by the *joint* p.d.f., denoted by $p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta})$, which can be written with the Bayes' formula as follows

$$p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta}) = p_{\boldsymbol{X}|C}(\boldsymbol{x}|\boldsymbol{\theta}, c) \times p_C(c),$$

where $p_C(c)$ denotes the *prior* p.d.f. that encodes our knowledge about the multiplicative noise before any measurement, and $p_{\boldsymbol{X}|C}(\boldsymbol{x}|\boldsymbol{\theta}, c)$ is the *conditional* p.d.f. that corresponds to the p.d.f. when $\boldsymbol{x}$ and c are known. One notes that $p_{\boldsymbol{X}|C}(\boldsymbol{x}|\boldsymbol{\theta}, c)$ is equal to $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)$, the p.d.f. introduced in the previous approach and defined in Eq. (5.6). The log-likelihood function can therefore be expressed as follows

$$\ln p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta}) = \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) + \ln p_C(c), \tag{5.18}$$

where

$$\ln p_C(c) = \text{cst} - \frac{1}{2\,\sigma_c^2}(c - \mu_c)^2. \tag{5.19}$$

5.3.1.1 Fisher information matrix and Hybrid Cramér-Rao bound (HCRB)

Case of a single molecular line — In this approach, the FIM corresponding to the considered molecular line, denoted by $\boldsymbol{\mathcal{I}}_{2a,\boldsymbol{X}}$, reads

$$\boldsymbol{\mathcal{I}}_{2a,\boldsymbol{X}}(\boldsymbol{\theta}, C) = \mathbb{E}\left[\nabla_{\boldsymbol{\theta},c} \ln p_{\boldsymbol{X},C}(\boldsymbol{X}, C|\boldsymbol{\theta})\, \nabla_{\boldsymbol{\theta},c}^{\top} \ln p_{\boldsymbol{X},C}(\boldsymbol{X}, C|\boldsymbol{\theta})\right],$$

with

$$\mathbb{E}\left[f(\boldsymbol{X}, C)\right] = \int_{\boldsymbol{X}} \int_{C} f(\boldsymbol{x}, c)\, p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}c.$$

One notes that the measurement model defined in Eq. (5.2), can be rewritten

$$\boldsymbol{X} = C\, \boldsymbol{s}(\boldsymbol{\theta}) + \boldsymbol{B} \quad \Rightarrow \quad \begin{pmatrix} \boldsymbol{X} \\ C \end{pmatrix} = \underbrace{\begin{pmatrix} \boldsymbol{s}(\boldsymbol{\theta}) & \boldsymbol{I}_N \\ 1 & \boldsymbol{0}_N^{\top} \end{pmatrix}}_{\text{deterministic}} \begin{pmatrix} C \\ \boldsymbol{B} \end{pmatrix}. \tag{5.20}$$

Computing the expectation of $(\boldsymbol{X}\, C)^{\top}$ thus involves to perform the expectation over the vector $(C\, \boldsymbol{B})^{\top}$. One then has

$$\mathbb{E}\begin{bmatrix} \boldsymbol{X} \\ C \end{bmatrix} = \begin{pmatrix} \mu_c\, \boldsymbol{s}(\boldsymbol{\theta}) \\ \mu_c \end{pmatrix}.$$

To derive the expression of $\boldsymbol{\mathcal{I}}_{2a,\boldsymbol{X}}$, we detail below the term $\nabla_{\boldsymbol{\theta},c} \ln p_{\boldsymbol{X},C}(\boldsymbol{X}, C|\boldsymbol{\theta})$. From Eq. (5.18), one has

$$\nabla_{\boldsymbol{\theta},c} \ln p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta},c} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) + \nabla_{\boldsymbol{\theta},c} \ln p_C(c), \tag{5.21}$$

where $\nabla_{\boldsymbol{\theta},c} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)$ is detailed in Eq. (5.9). Moreover, from Eq. (5.19), one notes that

$$\forall \theta \in \left(\boldsymbol{\theta}^{\top} \boldsymbol{c}^{\top}\right)^{\top} \text{ such that } \theta \neq c, \quad \frac{\partial \ln p_C(c)}{\partial \theta} = 0.$$

The Eq. (5.21) thus reads

$$\nabla_{\boldsymbol{\theta},c} \ln p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta}) = \left(\nabla_{\boldsymbol{\theta}}^{\top} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) \quad 0 \cdots 0 \quad \tfrac{\partial \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta},c)}{\partial c} + \tfrac{\partial \ln p_C(c)}{\partial c} \quad 0 \cdots 0\right)^{\top},$$

where $\nabla_{\boldsymbol{\theta}}^{\top} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c)$ and $\frac{\partial \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta},c)}{\partial c}$ are both defined in Eq. (5.10), and

$$\frac{\partial \ln p_C(c)}{\partial c} = \frac{\mu_c - c}{\sigma_c^2}.$$

The FIM is finally given by

$$\boldsymbol{\mathcal{I}}_{2a,\boldsymbol{X}}(\boldsymbol{\theta}, \boldsymbol{C}) = \mathbb{E}\left[\begin{array}{c|ccccccc} \boldsymbol{A}_{2a,\boldsymbol{X}} & 0 & \cdots & 0 & \boldsymbol{E}_{2a,\boldsymbol{X}} & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & & 0 & & \cdots & 0 \\ \vdots & \vdots & & & \vdots & & & \vdots \\ 0 & & & & 0 & & & \\ \boldsymbol{E}_{2a,\boldsymbol{X}}^{\top} & 0 & \cdots & 0 & d_{2a,\boldsymbol{X}} & 0 & \cdots & 0 \\ 0 & & & & 0 & & & \\ \vdots & \vdots & & & \vdots & & & \vdots \\ 0 & 0 & \cdots & & 0 & & \cdots & 0 \end{array}\right],$$

where

$$\boldsymbol{A}_{2a,\boldsymbol{X}} = \frac{C^2}{\sigma_b^4}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\boldsymbol{\mathcal{E}}_1\boldsymbol{\mathcal{E}}_1^\top\nabla_{\boldsymbol{\theta}}^\top\boldsymbol{s}, \tag{5.22}$$

$$\boldsymbol{E}_{2a,\boldsymbol{X}} = \frac{C}{\sigma_b^4}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\boldsymbol{\mathcal{E}}_1\boldsymbol{\mathcal{E}}_1^\top\boldsymbol{s} + \frac{C\,(\mu_c - C)}{\sigma_b^2\sigma_c^2}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\boldsymbol{\mathcal{E}}_1, \tag{5.23}$$

and

$$d_{2a,\boldsymbol{X}} = \frac{1}{\sigma_b^4}\,\boldsymbol{s}^\top\boldsymbol{\mathcal{E}}_1\boldsymbol{\mathcal{E}}_1^\top\boldsymbol{s} + 2\,\frac{(\mu_c - C)}{\sigma_b^2\sigma_c^2}\,\boldsymbol{s}^\top\boldsymbol{\mathcal{E}}_1 + \frac{(\mu_c - C)^2}{\sigma_c^4}. \tag{5.24}$$

Performing the expectation on Eq. (5.22), Eq. (5.23), and Eq. (5.24), leads to

$$\mathbb{E}\left[\boldsymbol{A}_{2a,\boldsymbol{X}}\right] = \frac{(\sigma_c^2 + \mu_c^2)}{\sigma_b^2}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\nabla_{\boldsymbol{\theta}}^\top\boldsymbol{s}, \tag{5.25}$$

$$\mathbb{E}\left[\boldsymbol{E}_{2a,\boldsymbol{X}}\right] = \frac{\mu_c}{\sigma_b^2}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\boldsymbol{s}, \tag{5.26}$$

and

$$\mathbb{E}\left[d_{2a,\boldsymbol{X}}\right] = \frac{1}{\sigma_b^2}\,\boldsymbol{s}^\top\boldsymbol{s} + \frac{1}{\sigma_c^2}. \tag{5.27}$$

The calculation details are provided in Appendix 5.C.

Case of multiple molecular lines — It is straightforward to show that the FIM when considering the set of molecular lines, denoted by $\boldsymbol{\mathcal{I}}_{2a}$, is

$$\boldsymbol{\mathcal{I}}_{2a}(\boldsymbol{\theta}, \boldsymbol{C}) = \mathbb{E}\left[\begin{array}{c|ccccc} \sum_{\boldsymbol{X}\in\boldsymbol{Y}}\boldsymbol{A}_{2a,\boldsymbol{X}} & \boldsymbol{E}_{2a,\boldsymbol{X}_1} & & \cdots & & \boldsymbol{E}_{2a,\boldsymbol{X}_L} \\ \hline \boldsymbol{E}_{2a,\boldsymbol{X}_1}^\top & d_{2a,\boldsymbol{X}_1} & 0 & \cdots & 0 \\ & 0 & & \ddots & \vdots \\ \vdots & \vdots & \ddots & & 0 \\ \boldsymbol{E}_{2a,\boldsymbol{X}_L}^\top & 0 & \cdots & 0 & d_{2a,\boldsymbol{X}_L} \end{array}\right].$$

The FIM components are obtained from Eq. (5.25), Eq. (5.26) and Eq. (5.27), i.e.,

$$\mathbb{E}\left[\sum_{\boldsymbol{X}\in\boldsymbol{Y}}\boldsymbol{A}_{2a,\boldsymbol{X}}\right] = \sum_{l=1}^{L}\frac{\left(\sigma_{c,l}^2 + \mu_{c,l}^2\right)}{\sigma_{b,l}^2}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}_l\,\nabla_{\boldsymbol{\theta}}^\top\boldsymbol{s}_l,$$

$$\forall l \in [\![1, L]\!], \quad \mathbb{E}\left[\boldsymbol{E}_{2a,\boldsymbol{X}_l}\right] = \frac{\mu_{c,l}}{\sigma_{b,l}^2}\,\nabla_{\boldsymbol{\theta}}\boldsymbol{s}_l\,\boldsymbol{s}_l, \quad \text{and} \quad \mathbb{E}\left[d_{2a,\boldsymbol{X}_l}\right] = \frac{1}{\sigma_{b,l}^2}\,\boldsymbol{s}_l^\top\boldsymbol{s}_l + \frac{1}{\sigma_{c,l}^2}.$$

Hybrid Cramér-Rao bound (HCRB) — The HCRB, defined by $\boldsymbol{\mathcal{B}}_{2a}(\boldsymbol{\theta}, \boldsymbol{C}) = \boldsymbol{\mathcal{I}}_{2a}^{-1}(\boldsymbol{\theta}, \boldsymbol{C})$, can be written as

$$\boldsymbol{\mathcal{B}}_{2a}(\boldsymbol{\theta}, \boldsymbol{C}) = \begin{pmatrix} \boldsymbol{\mathcal{B}}_{2a}(\boldsymbol{\theta}) & [\boldsymbol{\mathcal{B}}_{2a}]_{1,2} \\ [\boldsymbol{\mathcal{B}}_{2a}]_{1,2}^{\top} & \mathcal{B}_{2a}(\boldsymbol{C}) \end{pmatrix},$$

where the block matrix $\boldsymbol{\mathcal{B}}_{2a}(\boldsymbol{\theta}) \in \mathbb{R}^{M \times M}$ concerns the deterministic part of the vector of unknowns that interests us, while $\mathcal{B}_{2a}(\boldsymbol{C}) \in \mathbb{R}$ concerns the random one. For the sake of convenience, we shall hereafter refer to the submatrix $\boldsymbol{\mathcal{B}}_{2a}(\boldsymbol{\theta})$ as the HCRB.

5.3.2 Marginal approach

As the calibration noise is a nuisance parameter whose value is superfluous, we propose to discard it by describing the measurements with the marginal p.d.f. $p(\boldsymbol{x}|\boldsymbol{\theta}) = \int_C p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta})\,\mathrm{d}c$. Obtaining the analytical expression of the marginal p.d.f. is often intractable since the joint p.d.f. may be difficult to express in the first place. However, in our case, the marginal p.d.f. can be readily found: as the additive and multiplicative noises are independent and Gaussian, Eq. (5.20) shows that the pair of variables $(\boldsymbol{X}, C)$ is also Gaussian, and so is the marginal p.d.f. The marginalization step allows us to recover a deterministic estimation problem from the hybrid one. A theoretical advantage is that the associated CRB, which we call the *marginal* CRB from this point on, is known to be asymptotically tight (e.g., Cramér 1946, Chapter 32.5), while Noam and Messer 2009 show that the HCRB is generally lower than the marginal CRB. A practical advantage is that reducing the number of unknowns implies a lower-dimensional FIM, which can help to limit the numerical issues associated with its inversion.

Marginal p.d.f. — Let us consider a single molecular line $\boldsymbol{x} \in \boldsymbol{y}$ with a corresponding multiplicative noise C. Having demonstrated that the measurements are normally distributed, i.e., $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, we simply need to calculate the mean and covariance from Eq. (5.2) while taking account that $C \sim \mathcal{N}(\mu_c, \sigma_c^2)$:

$$\boldsymbol{\mu}_2(\boldsymbol{\theta}) = \mu_c\, \boldsymbol{s}(\boldsymbol{\theta}) \quad \text{and} \quad \boldsymbol{\Sigma}_2(\boldsymbol{\theta}) = \sigma_c^2\, \boldsymbol{s}(\boldsymbol{\theta})\boldsymbol{s}(\boldsymbol{\theta})^{\top} + \sigma_b^2\, \boldsymbol{I}_N. \tag{5.28}$$

Thus, the marginal p.d.f. is

$$p(\boldsymbol{x}|\boldsymbol{\theta}) = \frac{1}{\sqrt{\det\left(2\pi\, \boldsymbol{\Sigma}_2\right)}} \exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}_2^{\top}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\varepsilon}_2\right\}, \tag{5.29}$$

where $\boldsymbol{\varepsilon}_2(\boldsymbol{\theta}) = \boldsymbol{x} - \boldsymbol{\mu}_2(\boldsymbol{\theta})$. The determinant of $\boldsymbol{\Sigma}_2$ can be detailed by recognizing in Eq. (5.28) the form $\det\left(\boldsymbol{I}_N + \boldsymbol{u}\boldsymbol{v}^{\top}\right) = 1 + \boldsymbol{u}^{\top}\boldsymbol{v}$. (e.g., Petersen and Pedersen 2008, "*The Matrix Cookbook*", Eq. (24)). One thus has

$$\det\left(\boldsymbol{\Sigma}_2\right) = \sigma_b^{2N}\left(1 + \frac{\sigma_c^2}{\sigma_b^2}\,\boldsymbol{s}^{\top}\boldsymbol{s}\right).$$

In Eq. (5.29), the inverse of the covariance matrix can be found by using the *Sherman-Morrison* formula (defined in Eq. (5.47) of Appendix 5.D), leading to

$$\boldsymbol{\Sigma}_2^{-1} = \frac{1}{\sigma_b^2}\left(\boldsymbol{I}_N - \frac{\sigma_c^2\, \boldsymbol{s}\boldsymbol{s}^{\top}}{\sigma_b^2 + \sigma_c^2\, \boldsymbol{s}^{\top}\boldsymbol{s}}\right).$$

5.3.2.1 Fisher information matrix and Marginal Cramér-Rao bound

Case of a single molecular line — In this approach, the FIM of the considered molecular line, denoted by $\boldsymbol{\mathcal{I}}_{2b,\boldsymbol{X}}$, can be expressed by using the *Slepian-Bangs* formula (Stoica and Moses 2005, Eq. (B.3.3)) as

$$\boldsymbol{\mathcal{I}}_{2b,\boldsymbol{X}}(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}}\boldsymbol{s}\left(\frac{1}{\sigma_b^2}\left(\mu_c^2 + \frac{\sigma_c^4\boldsymbol{s}^\top\boldsymbol{s}}{\sigma_c^2\boldsymbol{s}^\top\boldsymbol{s}+\sigma_b^2}\right)\boldsymbol{I}_N + \left(\frac{\sigma_b^2\sigma_c^4 - \sigma_c^6\boldsymbol{s}^\top\boldsymbol{s}}{\sigma_b^2\left(\sigma_c^2\boldsymbol{s}^\top\boldsymbol{s}+\sigma_b^2\right)^2} - \frac{\sigma_c^2\mu_c^2}{\sigma_b^2\left(\sigma_c^2\boldsymbol{s}^\top\boldsymbol{s}+\sigma_b^2\right)}\right)\boldsymbol{s}\boldsymbol{s}^\top\right)\nabla_{\boldsymbol{\theta}}^\top\boldsymbol{s}. \tag{5.30}$$

The detailed demonstration can be found in Appendix A of Roueff *et al.* 2024, Eq. (A.22).

Case of multiple molecular lines — As carried out in the previous approaches, the FIM when considering the set of molecular lines, denoted by $\boldsymbol{\mathcal{I}}_{2b}$, is then:

$$\boxed{\boldsymbol{\mathcal{I}}_{2b}(\boldsymbol{\theta}) = \sum_{\boldsymbol{X}\in\boldsymbol{Y}} \boldsymbol{\mathcal{I}}_{2b,\boldsymbol{X}}(\boldsymbol{\theta}),}$$

where $\boldsymbol{\mathcal{I}}_{2b,\boldsymbol{X}}$ is defined in Eq. (5.30).

Marginal Cramér-Rao bound — The *marginal* CRB, is given by $\boldsymbol{\mathcal{B}}_{2b}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{I}}_{2b}^{-1}(\boldsymbol{\theta})$.

5.4 Neglecting the multiplicative noise

The third approach that we examine merely consists in ignoring the presence of the multiplicative noise. Indeed, this hypothesis is well adapted when the relative calibration error is smaller than the additive S/N. In this situation, the data are generated with a *true* measurement model that takes into account both the multiplicative and additive noise, i.e.,

$$\boldsymbol{X} = C\,\boldsymbol{s}(\boldsymbol{\theta}) + \boldsymbol{B}, \tag{5.31}$$

but are analyzed with a *misspecified* model, in which the calibration noise is not present, i.e.,

$$\boldsymbol{X} = \boldsymbol{s}(\boldsymbol{\theta}) + \boldsymbol{B}. \tag{5.32}$$

We have concluded the previous section by stating that the best way of treating the true model Eq. (5.31) is to follow the "marginal" approach, in which the unbiased MLE asymptotically yields to the *true* value of $\boldsymbol{\theta}$, hereafter denoted by $\boldsymbol{\theta}_{\mathrm{tr}}$. In Sect. 5.4.1, we shall first construct the MLE associated with the misspecified model Eq. (5.32), denoted by $\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}$. The latter estimator will converge to a *pseudo* solution of $\boldsymbol{\theta}$, denoted by $\boldsymbol{\theta}_{\mathrm{ps}}$. Two subsidiary perspectives can be envisaged.

Knowledgeable case — In the first one, a knowledgeable modeler is aware that Eq. (5.32) is only an approximation of the true generative model Eq. (5.31). He will thus analyze the actual performance of $\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}$ through a Monte-Carlo simulation on a dataset generated with the true model. He will more importantly ask whether the estimator is biased, i.e., how far is $\boldsymbol{\theta}_{\mathrm{ps}}$ from $\boldsymbol{\theta}_{\mathrm{tr}}$, besides examining the variance of $\widehat{\boldsymbol{\theta}}_{\mathrm{MLE}}$.

Naive case — In the second one, the modeler is naive in the sense that he is not aware that the data he tries to model are affected by the calibration disturbance. He will thus evaluate the estimator performance through a Monte-Carlo simulation on a dataset generated with the misspecified model Eq. (5.32), instead of the true one.

We shall here discuss these two situations in turn. In Sect. 5.4.2, we will consider a knowledgeable modeler to show that $\widehat{\boldsymbol{\theta}}_{\text{MLE}}$ is asymptotically unbiased, i.e., $\boldsymbol{\theta}_{\text{ps}} = \boldsymbol{\theta}_{\text{tr}}$. Then, we shall compute a lower bound on the estimator variance that takes into account the knowledge that Eq. (5.32) is only an approximated model of the data. This bound is known as the *misspecified* CRB (MCRB), and we shall denote it by $\boldsymbol{\mathcal{B}}_{3a}$. In Sect. 5.4.3, we shall compute the standard CRB, denoted by $\boldsymbol{\mathcal{B}}_{3b}$, as a naive modeler would. It will allow us to investigate in the next section how a naive modeler is over-optimistic about the precision of the estimation results when using the $\boldsymbol{\mathcal{B}}_{3b}$ as the reference one.

5.4.1 Likelihood under the misspecified model

When the measurements are assumed to be generated by the misspecified model Eq. (5.32) instead of the true one Eq. (5.31), the randomness of the considered molecular line solely results from the Gaussian additive noise, i.e., $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}_3,\, \boldsymbol{\Sigma}_3)$ where

$$\boldsymbol{\mu}_3(\boldsymbol{\theta}) = \boldsymbol{s}(\boldsymbol{\theta}) \quad \text{and} \quad \boldsymbol{\Sigma}_3 = \sigma_b^2\, \boldsymbol{I}_N. \tag{5.33}$$

The p.d.f. when omitting the multiplicative noise, denoted by $p_{\boldsymbol{X},\not\phi}(\boldsymbol{x}|\boldsymbol{\theta})$, then reads

$$p_{\boldsymbol{X},\not\phi}(\boldsymbol{x}|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\,\sigma_b^2}^N} \exp\{-\frac{1}{2\,\sigma_b^2}\,\boldsymbol{\varepsilon}_3^\top \boldsymbol{\varepsilon}_3\},$$

where $\boldsymbol{\varepsilon}_3(\boldsymbol{\theta}) = \boldsymbol{x} - \boldsymbol{\mu}_3(\boldsymbol{\theta})$. Consequently, the set of independent molecular lines $\boldsymbol{Y} = \{\boldsymbol{X}_1 \cdots \boldsymbol{X}_L\}$ is assumed to be described by the following p.d.f., that is said to be *misspecified*,

$$p_{\boldsymbol{Y},\not\phi}(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{\boldsymbol{X}\in\boldsymbol{Y}} p_{\boldsymbol{X},\not\phi}(\boldsymbol{x}|\boldsymbol{\theta}). \tag{5.34}$$

The log-likelihood gradient then reads

$$\nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{Y},\not\phi}(\boldsymbol{y}|\boldsymbol{\theta}) = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2}\, \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l\, (\boldsymbol{x}_l - \boldsymbol{s}_l)\,. \tag{5.35}$$

5.4.2 Knowledgeable case

5.4.2.1 Asymptotic behavior of the MLE

Let $\widehat{\boldsymbol{\theta}}_{\text{MLE}}$ be the MLE built from $p_{\boldsymbol{Y},\not\phi}(\boldsymbol{y}|\boldsymbol{\theta})$. Huber 1967, White 1982, and Akaike 1973 obtained three main results: 1) $\widehat{\boldsymbol{\theta}}_{\text{MLE}}$ converges to a normal distribution, $\mathcal{N}(\boldsymbol{\theta}_{\text{ps}}, \boldsymbol{\mathcal{B}}_{3a})$, when the number of independent samples[1] N tends to infinity, 2) $\boldsymbol{\theta}_{\text{ps}}$ is the value that minimizes

[1] In our case, the number of velocity channels.

the KLD between the p.d.f. of the true model given by Eq. (5.29) and the p.d.f. of the assumed model, here given by Eq. (5.34), and, 3) the covariance matrix $\boldsymbol{\mathcal{B}}_{3a}$ reads (e.g., Ren *et al.* 2015; Richmond and Horowitz 2013)

$$\boldsymbol{\mathcal{B}}_{3a} = \boldsymbol{G}^{-1}\boldsymbol{F}\boldsymbol{G}^{-1}, \tag{5.36}$$

where

$$\boldsymbol{F} = \mathbb{E}\left[\nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{Y},\phi}(\boldsymbol{Y}|\boldsymbol{\theta})\, \nabla_{\boldsymbol{\theta}}^{\top} \ln p_{\boldsymbol{Y},\phi}(\boldsymbol{Y}|\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{ps}}}\right] \tag{5.37}$$

and

$$\boldsymbol{G} = -\mathbb{E}\left[\nabla_{\boldsymbol{\theta}}^{2} \ln p_{\boldsymbol{Y},\phi}(\boldsymbol{Y}|\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{ps}}}\right]. \tag{5.38}$$

In the latter equations, the expectation is performed over the measurements drawn from the true (marginal) distribution $p_{\boldsymbol{Y}}(\boldsymbol{Y}|\boldsymbol{\theta}_{\text{tr}})$. The matrix $\boldsymbol{\mathcal{B}}_{3a}$ is known as the *misspecified*[2] CRB (MCRB). In absence of misspecification, we find the standard CRB expression $\boldsymbol{\mathcal{B}}_{3a} = \boldsymbol{\mathcal{B}}_{2b}$ since $F = G$.

5.4.2.2 Unbiasedness of the MLE

We here show that $\widehat{\boldsymbol{\theta}}_{\text{MLE}}$ is asymptotically unbiased, i.e., $\boldsymbol{\theta}_{\text{ps}} = \boldsymbol{\theta}_{\text{tr}}$, We first consider the case of a single molecular line before generalizing the result to the case of several ones.

Case of a single molecular line — As a reminder of Sect. 5.3.2, the measurement model is

$$\boldsymbol{X} = C\,\boldsymbol{s}(\boldsymbol{\theta}_{\text{tr}}) + \boldsymbol{B},$$

where

$$\boldsymbol{\mu}_2(\boldsymbol{\theta}_{\text{tr}}) = \mu_c\,\boldsymbol{s}(\boldsymbol{\theta}_{\text{tr}}) \quad \text{and} \quad \boldsymbol{\Sigma}_2(\boldsymbol{\theta}_{\text{tr}}) = \sigma_c^2\,\boldsymbol{s}(\boldsymbol{\theta}_{\text{tr}})\boldsymbol{s}(\boldsymbol{\theta}_{\text{tr}}) + \sigma_b^2\,\boldsymbol{I}_N. \tag{5.39}$$

The asymptotic mean of $\widehat{\boldsymbol{\theta}}_{\text{MLE}}$, $\boldsymbol{\theta}_{\text{ps}}$, verifies (White 1982)

$$\boldsymbol{\theta}_{\text{ps}} = \arg\min_{\boldsymbol{\theta}} \mathcal{K}\left[p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}_{\text{tr}}) \,\|\, p_{\boldsymbol{X},\phi}(\boldsymbol{X}|\boldsymbol{\theta})\right],$$

where the KLD, denoted by $\mathcal{K}\left[p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}_{\text{tr}}) \,\|\, p_{\boldsymbol{X},\phi}(\boldsymbol{X}|\boldsymbol{\theta})\right]$, quantifies how much $p_{\boldsymbol{X},\phi}(\boldsymbol{x}|\boldsymbol{\theta})$ differs from $p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}_{\text{tr}})$. As the two distributions are Gaussian, it can be expressed as follows (Duchi 2007, Sect. 9)

$$\begin{aligned}\mathcal{K}\left[p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}_{\text{tr}}) \,\|\, p_{\boldsymbol{X},\phi}(\boldsymbol{X}|\boldsymbol{\theta})\right] = \frac{1}{2}\left(\text{tr}\left(\boldsymbol{\Sigma}_3^{-1}\boldsymbol{\Sigma}_2\right) - N + \ln\left(\frac{\det\boldsymbol{\Sigma}_3}{\det\boldsymbol{\Sigma}_2}\right)\right) \\ + \frac{1}{2}\left(\boldsymbol{\mu}_3 - \boldsymbol{\mu}_2\right)^{\top}\boldsymbol{\Sigma}_3^{-1}\left(\boldsymbol{\mu}_3 - \boldsymbol{\mu}_2\right),\end{aligned}$$

where $\{\boldsymbol{\mu}_3,\, \boldsymbol{\Sigma}_3\}$ and $\{\boldsymbol{\mu}_2,\, \boldsymbol{\Sigma}_2\}$ are defined in Eq. (5.33) and Eq. (5.39), respectively. Since $\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Sigma}_3$ do not depend on $\boldsymbol{\theta}$, the KLD can thus be rewritten as

$$\mathcal{K}\left[p_{\boldsymbol{X}}(\boldsymbol{X}|\boldsymbol{\theta}_{\text{tr}}) \,\|\, p_{\boldsymbol{X},\phi}(\boldsymbol{X}|\boldsymbol{\theta})\right] = \text{cst} + \frac{1}{2\,\sigma_b^2}\big(\boldsymbol{s}(\boldsymbol{\theta}) - \mu_c\,\boldsymbol{s}(\boldsymbol{\theta}_{\text{tr}})\big)^{\top}\big(\boldsymbol{s}(\boldsymbol{\theta}) - \mu_c\,\boldsymbol{s}(\boldsymbol{\theta}_{\text{tr}})\big).$$

As $\mu_c = 1$, the KLD, which is here a squared norm, is therefore minimized for

$$\boxed{\boldsymbol{\theta}_{\text{ps}} = \boldsymbol{\theta}_{\text{tr}}.}$$

[2]The qualifer *modified* can also be encountered (e.g., D'Andrea *et al.* 1994; Gini 1998).

Case of multiple molecular lines — Since the molecular lines are independent and each described by a Gaussian distribution, both the p.d.f. of the set $\boldsymbol{Y}$ and the misspecified p.d.f., denoted by $p_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}_{\text{tr}})$ and $p_{\boldsymbol{Y},\not{c}}(\boldsymbol{y}|\boldsymbol{\theta})$, respectively, are Gaussian. Following the same reasoning as above, the KLD can therefore be expressed as follows

$$\mathcal{K}\left[p_{\boldsymbol{Y}}(\boldsymbol{Y}|\boldsymbol{\theta}_{\text{tr}}) \,\|\, p_{\boldsymbol{Y},\not{c}}(\boldsymbol{Y}|\boldsymbol{\theta})\right] = \text{cst} + \frac{1}{2}\left(\boldsymbol{\mu}_{3,\boldsymbol{Y}} - \boldsymbol{\mu}_{2,\boldsymbol{Y}}\right)^{\top} \boldsymbol{\Sigma}_{3,\boldsymbol{Y}}^{-1}\left(\boldsymbol{\mu}_{3,\boldsymbol{Y}} - \boldsymbol{\mu}_{2,\boldsymbol{Y}}\right),$$

where the term

$$\boldsymbol{\mu}_{3,\boldsymbol{Y}} - \boldsymbol{\mu}_{2,\boldsymbol{Y}} = \left(\left(\boldsymbol{s}_1(\boldsymbol{\theta}) - \mu_{c,1}\,\boldsymbol{s}_1(\boldsymbol{\theta}_{\text{tr}})\right)^{\top} \quad \cdots \quad \left(\boldsymbol{s}_L(\boldsymbol{\theta}) - \mu_{c,L}\,\boldsymbol{s}_L(\boldsymbol{\theta}_{\text{tr}})\right)^{\top}\right)^{\top}$$

is the sole function of $\boldsymbol{\theta}$, since $\boldsymbol{\Sigma}_{3,\boldsymbol{Y}}$ is explained by

$$\boldsymbol{\Sigma}_{3,\boldsymbol{Y}} = \begin{pmatrix} \boldsymbol{\Sigma}_{3,\boldsymbol{X}_1} & 0 & \cdots & 0 \\ 0 & & \ddots & \vdots \\ \vdots & \ddots & & 0 \\ 0 & \cdots & 0 & \boldsymbol{\Sigma}_{3,\boldsymbol{X}_L} \end{pmatrix},$$

where $\forall l \in [\![1, L]\!]$, $\boldsymbol{\Sigma}_{3,\boldsymbol{X}_l} = \sigma_{b,l}^2\,\boldsymbol{I}_N$.

5.4.2.3 Misspecified CRB (MCRB)

We here detail the MCRB to quantify how neglecting the multiplicative noise in the measurement model deteriorates the asymptotic precision of the MLE. The MCRB, denoted by $\boldsymbol{\mathcal{B}}_{3a}$, is defined in Eq. (5.36), Eq. (5.38), and Eq. (5.38). Besides, based on the result obtained in Sect. 5.4.2.2, these equations can be evaluated at $\boldsymbol{\theta}_{\text{ps}} = \boldsymbol{\theta}_{\text{tr}}$. Thus, the matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ read

$$\boxed{\boldsymbol{F} = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \left(\sigma_{c,l}^2\,\boldsymbol{s}_l \boldsymbol{s}_l^{\top} + \sigma_{b,l}^2\,\boldsymbol{I}_N\right) \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{tr}}}}$$

and

$$\boxed{\boldsymbol{G} = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{tr}}}.}$$

The calculation details are provided in Appendix 5.E. These two matrices allow us to compute $\boldsymbol{\mathcal{B}}_{3a}$ with Eq. (5.36).

5.4.3 Naive case

In this section, we compute the CRB when the calibration noise is neglected in the measurement model while this misspecification is not taken into account.

Fisher information matrix — In this situation, the FIM, denoted by $\boldsymbol{\mathcal{I}}_{3b}$, reads

$$\boldsymbol{\mathcal{I}}_{3b}(\boldsymbol{\theta}) = \mathbb{E}\left[\nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{Y},\not{C}}(\boldsymbol{Y}|\boldsymbol{\theta})\, \nabla_{\boldsymbol{\theta}}^{\top} \ln p_{\boldsymbol{Y},\not{C}}(\boldsymbol{Y}|\boldsymbol{\theta})\right],$$

where $\nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{Y},\not{C}}(\boldsymbol{y}|\boldsymbol{\theta})$ is defined in Eq. (5.35). The explicit form of $\boldsymbol{\mathcal{I}}_{3b}$ is then

$$\begin{aligned}\boldsymbol{\mathcal{I}}_{3b}(\boldsymbol{\theta}) &= \sum_{l=1}^{L} \mathbb{E}\left[\frac{1}{\sigma_{b,l}^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \, (\boldsymbol{x}_l - \boldsymbol{s}_l)\,(\boldsymbol{x}_l - \boldsymbol{s}_l)^{\top} \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l\right] \\ &= \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \underbrace{\mathbb{E}\left[(\boldsymbol{x}_l - \boldsymbol{s}_l)\,(\boldsymbol{x}_l - \boldsymbol{s}_l)^{\top}\right]}_{=\boldsymbol{\Sigma}_3} \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l,\end{aligned}$$

where $\boldsymbol{\Sigma}_3$ is defined in Eq. (5.33). Finally,

$$\boxed{\boldsymbol{\mathcal{I}}_{3b}(\boldsymbol{\theta}) = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \, \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l.}$$

Cramér-Rao bound — The corresponding CRB is then obtained by $\boldsymbol{\mathcal{B}}_{3b}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{I}}_{3b}^{-1}(\boldsymbol{\theta})$.

5.5 Quantitative comparison of the different CRBs in the case of a sight line toward a simulated dense core

In the previous section, we derived the formulas of the CRBs associated with three different classes of estimators of the physical parameter $\boldsymbol{\theta}$ in the presence of the calibration noise C: the CRB when both $\boldsymbol{\theta}$ and C are estimated, the marginal CRB when only $\boldsymbol{\theta}$ is estimated since the nuisance parameter C is discarded, the MCRB when a modeler choose to neglect the calibration noise while knowing that is actually perturbs the measurements ("knowledgeable" case), and the standard CRB when a modeler is not even aware that the data is affected by the calibration noise ("naive" case). In this section, we compare the quantitative values of these CRBs for a set of parameters that is typical of a dense core sight line in a molecular cloud.

These typical dense core conditions derived from the estimates obtained in Ségal *et al.* 2024 when analyzing the Horsehead Nebula. More especially, we choose the estimates obtained for the so-called "DC 2" pixel (see the red marker on Fig. 5.2a) because this sight line belongs to the most shielded region of the field of view, where the $\mathrm{H^{13}CO^{+}}(1-0)$ line that probes the dense and cold gas, is well detected (see Fig. 5.2d). The estimated parameters obtained when fitting the sandwich model to the spectra measured at DC 2 are listed in Table 5.1. In the vicinity of this typical dense core regime, we investigate the CRB variations on the two most challenging parameters to probe, i.e., the kinetic temperature ($T_{\mathrm{kin,\,in}}$) and the volume density ($n_{\mathrm{H_2,\,in}}$) in the inner layer. In the following are shown the CRB maps obtained when exploring the regime in which $T_{\mathrm{kin,\,in}} \in [4.16,\, 37.0]$ K and $\log\left(n_{\mathrm{H_2,\,in}}/\,\mathrm{cm^{-3}}\right) \in [4,\, 6]$, while all the other parameters are set at the values listed

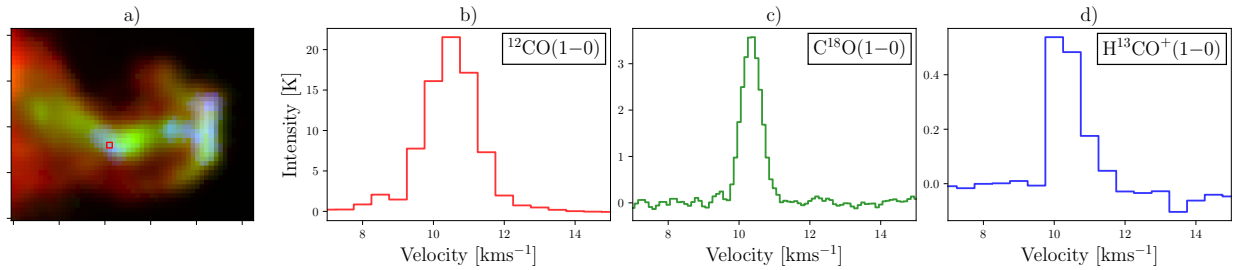


Figure 5.2: Composite image of the Horsehead Nebula from the ORION-B dataset and spectra toward the dense core "DC 2". a) Normalized integrated intensities of the first rotational transition of ^{12}CO (red channel), C^{18}O (green channel) and H^{13}CO$^+$ (blue channel). The red square indicates the location of the pixel DC 2. b, c, d) Measured spectrum at DC 2 of $^{12}\mathrm{CO}(1-0)$, $\mathrm{C}^{18}\mathrm{O}(1-0)$, and $\mathrm{H}^{13}\mathrm{CO}^+(1-0)$, respectively.

Table 5.1: Parameters used to generate a dataset with a dense core based on the estimation obtained when fitting the sandwich model to the pixel DC 2 of the Horsehead Nebula dataset.

Parameter	*Units*	*Layer*	
θ		*Inner*	*Outer*
T_{kin}	K	13.14	28.41
$\log\left(n_{\mathrm{H}_2}/\,\mathrm{cm}^{-3}\right)$	–	5.	3.87
$\log\left(N_{^{12}\mathrm{CO}}/\,\mathrm{cm}^{-2}\right)$	–	17.91	17.25
$\log\left(N_{^{13}\mathrm{CO}}/\,\mathrm{cm}^{-2}\right)$	–	16.11	15.95
$\log\left(N_{\mathrm{C}^{18}\mathrm{O}}/\,\mathrm{cm}^{-2}\right)$	–	15.21	14.90
$\log\left(N_{\mathrm{HCO}^+}/\,\mathrm{cm}^{-2}\right)$	–	13.23	12.72
$\log\left(N_{\mathrm{H}^{13}\mathrm{CO}^+}/\,\mathrm{cm}^{-2}\right)$	–	11.43	11.42
FWHM	$\mathrm{km\,s}^{-1}$	0.42	0.85
C_V	$\mathrm{km\,s}^{-1}$	10.29	10.46

Table 5.2: Characteristics of the studied molecular lines used to generate the dataset.

Species	*Transition*	*Frequency* [GHz]	*Vel. Res.* [$\mathrm{km\,s}^{-1}$]	σ_b [mK]
^{12}CO	$(1-0)$	115.271202	0.5	89
^{13}CO	$(1-0)$	110.201354	0.1	79
	$(2-1)$	220.398684	0.1	151
C^{18}O	$(1-0)$	109.782173	0.1	83
	$(2-1)$	219.560354	0.1	125
HCO$^+$	$(1-0)$	89.188525	0.1	276
H^{13}CO$^+$	$(1-0)$	86.754288	0.5	40

in Table 5.1. For each molecular line, we simulate an additive noise drawn from a normal distribution of zero mean and a standard deviation σ_b. As processed in Ségal *et al.* 2024, the value of σ_b is set to either the empirical noise measured on DC 2 for each line (see Table 5.2) or $T_{\text{peak}}/10$, in order to saturate the S/N and thus ensure that all lines contribute to the fit (see details in Ségal *et al.* 2024).

When comparing two points of view A and B, we compute the associated loss of precision on the estimation of $\theta \in \boldsymbol{\theta}$ as follows,

$$\boxed{\epsilon_{A,B}(\theta) \triangleq \frac{\sqrt{\mathcal{B}_B(\theta)} - \sqrt{\mathcal{B}_A(\theta)}}{\sqrt{\mathcal{B}_A(\theta)}}.} \tag{5.40}$$

Because of the asymmetry of the latter equation, $\epsilon_{A,B}(\theta)$ is interpreted as the precision loss implied by choosing the point of view B rather than A. Figure 5.3 and 5.4 compare the deterministic, misspecified, marginal and naive CRB for the logarithm of the kinetic temperature and the volume density, respectively. In these figures, the top row shows the square root of the CRB maps ($\sqrt{\mathcal{B}(\theta)}$) ordered by decreasing value from left to right. The bottom shows the relative loss of precision between two successive CRB when going from right to left. In all panels, the red square indicates the simulated dense core regime and the corresponding value is reported at the upper left.

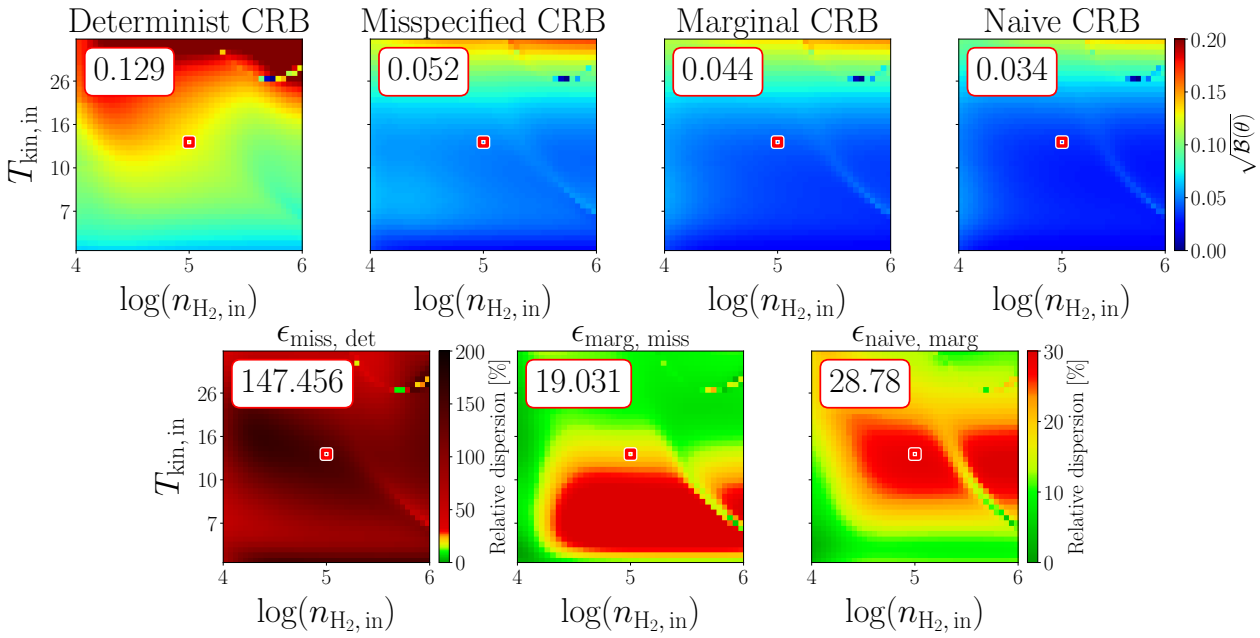


Figure 5.3: Comparison of the deterministic, misspecified, marginal, and naive CRB for the logarithm of the kinetic temperature $\log(T_{\text{kin, in}})$. The top row shows the root square of the CRB values order by decreasing value from left to right. The bottom shows the relative loss of precision between two successive CRB when going from right to left. The red square indicates the simulated dense core regime and the corresponding value is reported at the upper left.

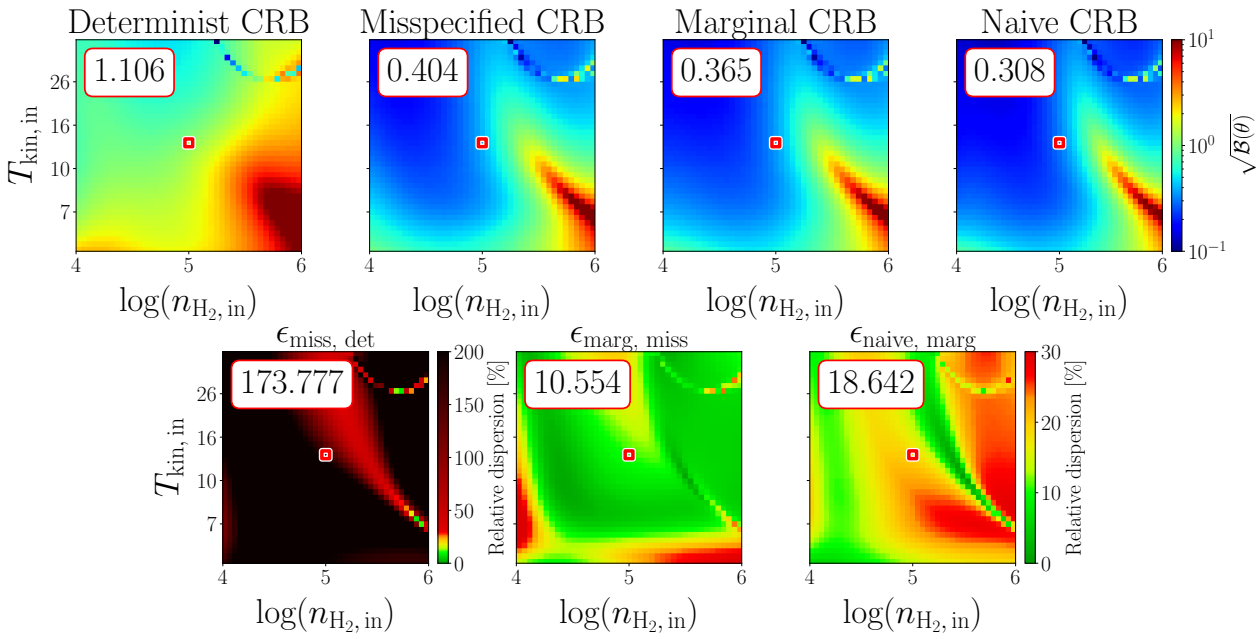


Figure 5.4: Same legend as in Fig. 5.3, but for the logarithm of the volume density $\log(n_{\mathrm{H_2,\,in}})$.

For a dataset that is contaminated by calibration noise, the best solution would be to build a marginal estimator as it would take into account all the information we have on the calibration noise (i.e., its mean and its dispersion). That's why the marginal CRB has the second smallest value. Although they neglect the calibration noise, the naive and knowledgeable modelers get on average the same physical estimations as these of the marginal estimator because we showed that neglecting the calibration noise gives an unbiased MLE (see Sect. 5.4.2.2). However, the associated precision described by the CRBs are different. On one hand, the naive modeler CRB provides an *over-optimistic* interval of confidence by $20-30\%$ compared to the marginal precision. On the other hand, the knowledgeable modeler (misspecified) CRB provides a realistic interval of confidence, but his estimator's performance is expected to be degraded by $10-20\%$ with respect to the marginal precision.

Finally, trying to determine both the physical parameters and the multiplicative noise without any *a priori* information (deterministic case) will lead to a significant loss of precision compared to the misspecified point of view. This is expected as there can be much ambiguity between the calibration noise and the actual output of the sandwich model.

In Roueff *et al.* 2024 and Ségal *et al.* 2024, we neglected the calibration noise (using the misspecified MLE) and computed the marginal CRB to provide the confidence interval on the estimations. The results of this chapter suggest that using the marginal MLE instead of the misspecified MLE could improve the precision by 10% on the logarithm of the kinetic temperature, and 20% on the logarithm of the volume density. However, the underlying assumption made during this investigation is that our physical model, $\boldsymbol{s}(\boldsymbol{\theta})$, perfectly describes the radiative and chemical processes, although we had to approximate them to keep the problem tractable. The uncertainty on the amplitude of the line brightness, i.e., $c\,\boldsymbol{s}(\boldsymbol{\theta})$, therefore results not only from the calibration noise c, but also from the approximated physical model. Under these conditions, using the marginal MLE estimator that takes into account

the calibration noise means that we do not completely trust anymore the amplitude of the line, with the high risk of confusing an approximation problem in ISM physics or chemistry with a calibration problem. When analyzing real GMC observations, we indeed observed implausible physical and chemical results when we tried to take into account the calibration noise. In other words, using a marginal MLE estimator is counter-productive because the hoped-for gain in precision actually results from a trade-off between our physical model approximation and the calibration noise. Neglecting the calibration noise therefore appears as the right choice in our case where the physics and chemistry have to be approximated. However, the misspecified CRB (instead of the naive CRB or the marginal CRB) needs to be used to provide realistic intervals of confidence. That's what we shall do in future studies.

5.6 Generalizing the comparison on actual GMC observations

In Ségal *et al.* 2025 (see Appendix 5.F), we generalize the previous study to actual observations toward the Horsehead Nebula that we model with the sandwich. Figure 4 of this study shows the loss of precision on $T_{\rm kin}$ and $n_{\rm H_2}$ that can be induced by neglecting the calibration noise instead of following the marginal approach. The relative deviation in the estimation precision reaches $\sim 13\%$ for the densest regions, and at most 30% in diffuse ones where the sandwich model is overparameterized and therefore less suited to such observations. We then assess the performance of the MLE constructed when neglecting the calibration noise for the simulated dense core pixel studied in Sect. 5.5. Figure 5 of Ségal *et al.* 2025 shows that for this pixel, the estimator is efficient and unbiased. Moreover, the marginal CRB and MCRB provide the same precision on the estimated parameters.

5.7 Summary and conclusions

In this chapter, we have calculated the CRBs on the physical parameters associated with three different points of view about the multiplicative calibration noise: 1) the CRB when considering the multiplicative noise as a deterministic parameter that is estimated in addition of the physical one, 2) the marginal CRB when considering the multiplicative noise as a random parameter that is marginalized from the measurement model, and 3) the MCRB when the calibration noise is neglected, although it affects the observations. We have compared these CRBs through Monte-Carlo simulations for a typical dense core line of sight and actual observations toward the Horsehead Nebula, which exhibits a diversity of physical regimes (e.g., translucent gas, filaments) in addition of the dense core one. These studies show that for a calibration error of dispersion less than 10%, purposely neglecting the calibration noise leads to 1) unbiased estimation results, and 2) a loss of precision between 10 and 30% that is deemed insignificant for the physical problem under consideration. The latter results confirm our choice to neglect the multiplicative noise for analyzing the Horsehead Nebula dataset in Roueff *et al.* 2021, 2024 and with the sandwich model in Ségal *et al.* 2024. In practical terms, this result is important for the MLE implementation, because estimations prove to be more stable when the multiplicative noise is neglected, which might suggest that introducing the

calibration factor as an additional degree of freedom yields to our model overfitting. Using the marginal CRB to quantify the estimation precision, as done in Roueff *et al.* 2021, 2024 and Ségal *et al.* 2024, is therefore the preferred choice. In the situation where the MCRB would indicate a loss of precision deemed significant, we would consider another perspective about the calibration noise (e.g., the hybrid one). The presented methodology therefore should be pursued when studying other datasets with other multilayer model configurations, molecular lines, and physical models.

References

Akaike, H. (1973). "Information Theory and an Extension of the Maximum Likelihood Principle". In: *Proceeding of the Second International Symposium on Information Theory*. Budapest: Akadémiai Kiadó.

Bartlett, M. S. (1951). "An Inverse Matrix Adjustment Arising in Discriminant Analysis". In: *The Annals of Mathematical Statistics* 22.1, pp. 107–111. DOI: 10.1214/aoms/1177729698.

Burnham, K. P. and D. R. Anderson (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Verlag, pp. 54–57.

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton Mathematical Series. Princeton University Press.

D'Andrea, A. N., U. Mengali, and R. Reggiannini (1994). "The modified Cramer-Rao bound and its application to synchronization problems". In: *IEEE Transactions on Communications* 42.234, pp. 1391–1399. DOI: 10.1109/TCOMM.1994.580247.

Duchi, J. (2007). "Derivations for linear algebra and optimization". In: *Berkeley, California* 3.1, pp. 2325–5870.

Einig, L. *et al.* (2023). "Deep learning denoising by dimension reduction: Application to the ORION-B line cubes". In: *A&A* 677, A158. DOI: 10.1051/0004-6361/202346064.

Galton, F. (1879). "XII. The geometric mean, in vital and social statistics". eng. In: *Proceedings of the Royal Society of London* 29.196, pp. 365–367. ISSN: 0370-1662.

Gini, F. (1998). "A radar application of a modified Cramer-Rao bound: parameter estimation in non-Gaussian clutter". In: *IEEE Transactions on Signal Processing* 46.7, pp. 1945–1953. DOI: 10.1109/78.700966.

Huber, P.J. (1967). "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions". In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press* 1, pp. 221–233.

Kay, S. (1993). *Fundamentals of statistical signal processing: estimation theory*. USA: Prentice-Hall, Inc. ISBN: 0133457117.

Kullback, S. and R. A. Leibler (1951). "On Information and Sufficiency". In: *Annals of Mathematical Statistics* 22, pp. 79–86.

Noam, Y. and H. Messer (2009). "Notes on the Tightness of the Hybrid CramÉr–Rao Lower Bound". In: *IEEE Transactions on Signal Processing* 57.6, pp. 2074–2084. DOI: 10.1109/TSP.2009.2015113.

Petersen, K. B. and M. S. Pedersen (Oct. 2008). *The Matrix Cookbook*. Version 20081110.

Ren, C. *et al.* (2015). "Performance bounds under misspecification model for MIMO radar application". In: *2015 23rd European Signal Processing Conference (EUSIPCO)*, pp. 514–518. DOI: 10.1109/EUSIPCO.2015.7362436.

Richmond, Christ D. and Larry L. Horowitz (2013). "Parameter bounds under misspecified models". In: *2013 Asilomar Conference on Signals, Systems and Computers*, pp. 176–180. DOI: 10.1109/ACSSC.2013.6810254.

Roueff, A. *et al.* (2021). "$C^{18}O$, ^{13}CO, and ^{12}CO abundances and excitation temperatures in the Orion B molecular cloud - Analysis of the achievable precision in modeling spectral lines within the approximation of the local thermodynamic equilibrium". In: *A&A* 645, A26. DOI: 10.1051/0004-6361/202037776.

Roueff, A. *et al.* (2024). "Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model - Application to the estimation of the gas temperature and volume density". In: *A&A* 686, A255. DOI: 10.1051/0004-6361/202449148.

Ségal, L. *et al.* (2024). "Toward a robust physical and chemical characterization of heterogeneous lines of sight: The case of the Horsehead nebula". In: *A&A* 692, A160. DOI: 10.1051/0004-6361/202451567.

Ségal, L. *et al.* (Aug. 2025). "De la possibilité d'ignorer un bruit de calibration lors de la construction d'un estimateur du maximum de vraisemblance en radio-astronomie". In: *GRETSI : Groupe de Recherche et d'Etudes de Traitement du Signal et des Images 2025*. Starsbourg, France. URL: https://hal.science/hal-05332727.

Shannon, C. E. (1948). "A mathematical theory of communication". In: *The Bell System Technical Journal* 27.4, pp. 623–656. DOI: 10.1002/j.1538-7305.1948.tb00917.x.

Stoica, P. and R. L. Moses (2005). *Spectral analysis of signals*. Vol. 452. Prentice Hall, Upper Saddle River, N.J., p. 360.

White, H. (1982). "Maximum Likelihood Estimation of Misspecified Models". In: *Econometrica* 50.1, pp. 1–25. ISSN: 00129682, 14680262.

Wikipedia contributors (2025). *Log-normal distribution — Wikipedia, The Free Encyclopedia*. Accessed: 2025-06-07. URL: https://en.wikipedia.org/wiki/Log-normal_distribution.

Appendix 5

5.A Computation of the Kullback-Leibler divergence (KLD)

Here, we detail how we compute the Kullback-Leibler divergence (KLD), denoted by $\mathcal{K}[p \,\|\, q]$, between the (true) lognormal distribution p and the (assumed) normal distribution q for the random variable C. As a reminder of Sect. 5.1.2, this KLD reads

$$\mathcal{K}[p \,\|\, q] = \int p(c) \log_2 \left(\frac{p(c)}{q(c)} \right) \mathrm{d}c, \tag{5.41}$$

where p is Log-$\mathcal{N}(\mu, \sigma^2)$ and q is $\mathcal{N}(\mu_c, \sigma_c^2)$. The mean μ and the standard deviation σ are chosen such that $\mathbb{E}[C] = \mu_c$ and $\mathbb{V}[C] = \sigma_c^2$, i.e.,

$$\mu = \ln \left(\frac{\mu_c^2}{\sqrt{\mu_c^2 + \sigma_c^2}} \right) \quad \text{and} \quad \sigma = \sqrt{\ln \left(1 + \frac{\sigma_c^2}{\mu_c^2} \right)}. \tag{5.42}$$

One recognizes in Eq. (5.41) the statistical expectation of $\log_2(p/q)$ performed over the true distribution p. Based on the law of large numbers, the expectation can be approximated by the empirical mean. Thus, let us generate K samples $c^1, \cdots, c^K$ drawn from the true distribution, i.e., $C \sim$ Log-$\mathcal{N}(\mu, \sigma^2)$. On then has

$$\begin{aligned} \mathcal{K}[p \,\|\, q] &= \int p(c) \log_2 \left(\frac{p(c)}{q(c)} \right) \mathrm{d}c \\ &= \mathbb{E}\left[\log_2 \left(\frac{p(C)}{q(C)} \right) \right] \\ &\approx \frac{1}{K} \sum_{k=1}^{K} \log_2 \left(\frac{p(c^k)}{q(c^k)} \right). \end{aligned}$$

5.B Derivative of composite function

Let $\boldsymbol{h} \circ \boldsymbol{f}$ the composite function defined such as

$$\begin{array}{ccccc} \mathbb{R}^M & \longrightarrow & \mathbb{R}^N & \longrightarrow & \mathbb{R}^q \\ \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \cdots \\ \theta_M \end{pmatrix} & \longmapsto & \boldsymbol{f}(\boldsymbol{\theta}) = \begin{pmatrix} f_1(\boldsymbol{\theta}) \\ \cdots \\ f_N(\boldsymbol{\theta}) \end{pmatrix} & \longmapsto & \boldsymbol{h}(\boldsymbol{f}(\boldsymbol{\theta})) = (\boldsymbol{h} \circ \boldsymbol{f})(\boldsymbol{\theta}) = \begin{pmatrix} h_1(\boldsymbol{f}(\boldsymbol{\theta})) \\ \cdots \\ h_q(\boldsymbol{f}(\boldsymbol{\theta})) \end{pmatrix} \end{array}$$

The gradient of $\boldsymbol{h} \circ \boldsymbol{f}$ with respect of $\boldsymbol{\theta}$ is

$$\nabla_{\boldsymbol{\theta}} (\boldsymbol{h} \circ \boldsymbol{f})(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} \boldsymbol{f}(\boldsymbol{\theta}) \, \nabla_{\boldsymbol{y}} \boldsymbol{h}(\boldsymbol{y})|_{\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{\theta})}. \tag{5.43}$$

Application — Let us consider

$$\begin{array}{ccc} \mathbb{R}^N & \longrightarrow & \mathbb{R}^N \\ \boldsymbol{y} = \begin{pmatrix} y_1 \\ \cdots \\ y_N \end{pmatrix} & \longmapsto & \boldsymbol{h}(\boldsymbol{y}) = \boldsymbol{y}^\top \boldsymbol{y} \end{array}$$

Bearing in mind that $\boldsymbol{y}^\top \boldsymbol{y} = \operatorname{tr}\left(\boldsymbol{y}\boldsymbol{y}^\top\right)$ and $\nabla_{\boldsymbol{y}} \operatorname{tr}\left(\boldsymbol{y}\boldsymbol{y}^\top\right) = 2\,\boldsymbol{y}$ (e.g., Petersen and Pedersen 2008, "*The Matrix Cookbook*", Eq. (17) and Eq. (115)), Eq. (5.43) leads to

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} (\boldsymbol{h} \circ \boldsymbol{f})(\boldsymbol{\theta}) &= \nabla_{\boldsymbol{\theta}} \boldsymbol{f}(\boldsymbol{\theta})\, \nabla_{\boldsymbol{y}} (\boldsymbol{y}^\top \boldsymbol{y})|_{\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{\theta})} \\ &= \nabla_{\boldsymbol{\theta}} \boldsymbol{f}(\boldsymbol{\theta})\, \nabla_{\boldsymbol{y}} \operatorname{tr}\left(\boldsymbol{y}\boldsymbol{y}^\top\right)|_{\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{\theta})} \\ &= 2\nabla_{\boldsymbol{\theta}} \boldsymbol{f}(\boldsymbol{\theta})\, \boldsymbol{y}|_{\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{\theta})} \\ &= 2\nabla_{\boldsymbol{\theta}} \boldsymbol{f}(\boldsymbol{\theta})\, \boldsymbol{f}(\boldsymbol{\theta}). \end{aligned}$$

In Sect. 5.2.1, where considering $\ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) = \text{cst} - \frac{1}{2\,\sigma_b^2}\, \boldsymbol{\varepsilon}_1^\top \boldsymbol{\varepsilon}_1$ with $\boldsymbol{\varepsilon}_1(\boldsymbol{\theta}) = \boldsymbol{x} - c\,\boldsymbol{s}(\boldsymbol{\theta})$, the gradient over $\boldsymbol{\theta}$ is therefore

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}, c) &= -\frac{1}{\sigma_b^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{\varepsilon}_1\, \boldsymbol{\varepsilon}_1 \\ &= \frac{c}{\sigma_b^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}\, \boldsymbol{\varepsilon}_1. \end{aligned}$$

5.C Calculation of the FIM under the hybrid approach

In this section, we detail the calculations of the FIM under the hybrid approach, i.e., when the multiplicative noise is assumed random while being estimating. As a reminder of Sect. 5.3.1.1, the FIM when considering a single molecular line $\boldsymbol{X} \in \boldsymbol{Y}$ with the multiplicative noise C, denoted $\boldsymbol{\mathcal{I}}_{2a,\boldsymbol{X}}$, is defined by

$$\boldsymbol{\mathcal{I}}_{2a,\boldsymbol{X}}(\boldsymbol{\theta}, \boldsymbol{C}) = \mathbb{E}\left[\begin{array}{c|ccccccc} \boldsymbol{A}_{2a,\boldsymbol{X}} & 0 & \cdots & 0 & \boldsymbol{E}_{2a,\boldsymbol{X}} & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & & 0 & & \cdots & 0 \\ \vdots & \vdots & & & \vdots & & & \vdots \\ 0 & & & & 0 & & & \\ \boldsymbol{E}_{2a,\boldsymbol{X}}^\top & 0 & \cdots & 0 & d_{2a,\boldsymbol{X}} & 0 & \cdots & 0 \\ 0 & & & & 0 & & & \\ \vdots & \vdots & & & \vdots & & & \vdots \\ 0 & 0 & \cdots & & 0 & & \cdots & 0 \end{array}\right]. \tag{5.44}$$

In Eq. (5.44), $\mathbb{E}$ denotes the expectation computed over the considered random variables $\boldsymbol{X}$ and C. Let us first show that $\mathbb{E}\left[f(\boldsymbol{X}, C)\right] = \mathbb{E}_C\left[\mathbb{E}_{\boldsymbol{X}|C}\left[f(\boldsymbol{X}, C)\right]\right]$ where $\mathbb{E}_{\boldsymbol{X}|C}$ is the expectation performed over $\boldsymbol{X}$ when C is known, and $\mathbb{E}_C$ is the expectation performed over

C. One has

$$\begin{aligned}
\mathbb{E}\left[f(\boldsymbol{X}, C)\right] &= \int_C \int_{\boldsymbol{X}} f(\boldsymbol{x}, c) \times p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}c \\
&= \int_C \int_{\boldsymbol{X}} f(\boldsymbol{x}, c) \times p_{\boldsymbol{X}|C}(\boldsymbol{x}|\boldsymbol{\theta}, c) \times p_C(c) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}c \\
&= \int_C \left(\int_{\boldsymbol{X}} f(\boldsymbol{x}, c) \times p_{\boldsymbol{X}|C}(\boldsymbol{x}|\boldsymbol{\theta}, c) \, \mathrm{d}\boldsymbol{x} \right) p_C(c) \, \mathrm{d}c \\
&= \int_C \mathbb{E}_{\boldsymbol{X}|C}\left[f(\boldsymbol{X}, C)\right] \times p_C(c) \, \mathrm{d}c \\
&= \mathbb{E}_C\left[\mathbb{E}_{\boldsymbol{X}|C}\left[f(\boldsymbol{X}, C)\right]\right].
\end{aligned} \tag{5.45}$$

Explicit form of $\mathbb{E}\left[\boldsymbol{A}_{2a,\boldsymbol{X}}\right]$ — The term $\boldsymbol{A}_{2a,\boldsymbol{X}}$ is given by Eq. (5.22), i.e.,

$$\boldsymbol{A}_{2a,\boldsymbol{X}} = \frac{C^2}{\sigma_b^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^\top \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s}.$$

Using Eq. (5.45), one obtains

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{A}_{2a,\boldsymbol{X}}\right] &= \frac{1}{\sigma_b^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \mathbb{E}\left[C^2 \, \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^\top\right] \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s} \\
&= \frac{1}{\sigma_b^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \mathbb{E}_C\left[C^2 \underbrace{\mathbb{E}_{\boldsymbol{X}|C}\left[\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^\top\right]}_{\boldsymbol{\Sigma}_1 = \sigma_b^2 \boldsymbol{I}_N} \right] \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s} \\
&= \frac{1}{\sigma_b^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \mathbb{E}_C\left[C^2\right] \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s} \\
&= \frac{(\sigma_c^2 + \mu_c^2)}{\sigma_b^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s}.
\end{aligned}$$

Then,

$$\boxed{\mathbb{E}\left[\boldsymbol{A}_{2a,\boldsymbol{X}}\right] = \frac{(\sigma_c^2 + \mu_c^2)}{\sigma_b^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \nabla_{\boldsymbol{\theta}}^\top \boldsymbol{s}.}$$

Explicit form of $\mathbb{E}\left[\boldsymbol{E}_{2a,\boldsymbol{X}}\right]$ — From Eq. (5.23), one has

$$\boldsymbol{E}_{2a,\boldsymbol{X}} = \frac{C}{\sigma_b^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_1^\top \boldsymbol{s} + \frac{C\left(\mu_c - C\right)}{\sigma_b^2 \sigma_c^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s} \, \boldsymbol{\varepsilon}_1. \tag{5.46}$$

Since $\mathbb{E}_{\boldsymbol{X}|C}[\mathcal{E}_1] = 0$, applying the expectation to Eq. (5.46) leads to

$$\begin{aligned}\mathbb{E}[\boldsymbol{E}_{2a,\boldsymbol{X}}] &= \mathbb{E}_C\left[\mathbb{E}_{\boldsymbol{X}|C}[\boldsymbol{E}_{2a,\boldsymbol{X}}]\right] \\ &= \mathbb{E}_C\left[\mathbb{E}_{\boldsymbol{X}|C}\left[\frac{C}{\sigma_b^4}\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\mathcal{E}_1\mathcal{E}_1^\top\boldsymbol{s}\right]\right] \\ &= \frac{1}{\sigma_b^4}\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\mathbb{E}_C\left[C\,\underbrace{\mathbb{E}_{\boldsymbol{X}|C}\left[\mathcal{E}_1\mathcal{E}_1^\top\right]}_{\boldsymbol{\Sigma}_1=\sigma_b^2\,\boldsymbol{I}_N}\right]\boldsymbol{s} \\ &= \frac{\mu_c}{\sigma_b^2}\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\boldsymbol{s}.\end{aligned}$$

Then,

$$\boxed{\mathbb{E}[\boldsymbol{E}_{2a,\boldsymbol{X}}] = \frac{\mu_c}{\sigma_b^2}\nabla_{\boldsymbol{\theta}}\boldsymbol{s}\,\boldsymbol{s}.}$$

Explicit form of $\mathbb{E}[d_{2a,\boldsymbol{X}}]$ — From Eq. (5.24), one has

$$d_{2a,\boldsymbol{X}} = \frac{1}{\sigma_b^4}\boldsymbol{s}^\top\mathcal{E}_1\mathcal{E}_1^\top\boldsymbol{s} + 2\,\frac{(\mu_c - C)}{\sigma_b^2\sigma_c^2}\,\boldsymbol{s}^\top\mathcal{E}_1 + \frac{(\mu_c - C)^2}{\sigma_c^4}.$$

From the above, one has

$$\mathbb{E}[d_{2a,\boldsymbol{X}}] = \frac{1}{\sigma_b^4}\boldsymbol{s}^\top\mathbb{E}_C\left[\underbrace{\mathbb{E}_{\boldsymbol{X}|C}\left[\mathcal{E}_1\mathcal{E}_1^\top\right]}_{\boldsymbol{\Sigma}_1=\sigma_b^2\,\boldsymbol{I}_N}\right]\boldsymbol{s} + \frac{2}{\sigma_b^2\sigma_c^2}\,\mathbb{E}_C\left[(\mu_c - C)\,\boldsymbol{s}^\top\underbrace{\mathbb{E}_{\boldsymbol{X}|C}[\mathcal{E}_1]}_{=0}\right] + \frac{1}{\sigma_c^4}\underbrace{\mathbb{E}_C\left[(\mu_c - C)^2\right]}_{=\sigma_c^2}.$$

Then,

$$\boxed{\mathbb{E}[d_{2a,\boldsymbol{X}}] = \frac{1}{\sigma_b^2}\boldsymbol{s}^\top\boldsymbol{s} + \frac{1}{\sigma_c^2}.}$$

5.D Sherman-Morrison formula

Let $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ be a nonsingular matrix, and $\boldsymbol{u} \in \mathbb{R}^N$ and $\boldsymbol{v} \in \mathbb{R}^N$ such that $\left(\boldsymbol{A} + \boldsymbol{u}.\boldsymbol{v}^\top\right)$ is nonsingular. The *Sherman-Morrison* formula is defined by (e.g., Bartlett 1951)

$$\left(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u}}. \tag{5.47}$$

5.E Calculation of the MCRB

In this section, we detail the calculations of the modified CRB (MCRB) associated to the approach where the calibration noises are neglected. As a reminder of Sect. 5.4, let us

consider the set of molecular lines $\boldsymbol{Y} = \{\boldsymbol{X}_1 \cdots \boldsymbol{X}_L\}$ that is assumed to be described by the misspecified p.d.f. $p_{\boldsymbol{Y},\not\phi}(\boldsymbol{y}|\boldsymbol{\theta})$ such that

$$\nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{Y},\not\phi}(\boldsymbol{y}|\boldsymbol{\theta}) = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \, (\boldsymbol{x}_l - \boldsymbol{s}_l) \,. \tag{5.48}$$

The MCRB, denoted $\boldsymbol{\mathcal{B}}_{3a}$, reads

$$\boldsymbol{\mathcal{B}}_{3a} = \boldsymbol{G}^{-1} \boldsymbol{F} \boldsymbol{G}^{-1}, \tag{5.49}$$

where

$$\boldsymbol{F} = \mathbb{E}\left[\nabla_{\boldsymbol{\theta}} \ln p_{\boldsymbol{Y},\not\phi}(\boldsymbol{Y}|\boldsymbol{\theta}) \, \nabla_{\boldsymbol{\theta}}^{\top} \ln p_{\boldsymbol{Y},\not\phi}(\boldsymbol{Y}|\boldsymbol{\theta}) \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm ps}} \right]$$

and

$$\boldsymbol{G} = -\mathbb{E}\left[\nabla_{\boldsymbol{\theta}}^2 \ln p_{\boldsymbol{Y},\not\phi}(\boldsymbol{Y}|\boldsymbol{\theta}) \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm ps}} \right]. \tag{5.50}$$

From the result obtained in Sect. 5.4.2.2, we hereafter evaluate these latter expressions at $\boldsymbol{\theta}_{\rm ps} = \boldsymbol{\theta}_{\rm tr}$. Let us first derive the explicit form of $\boldsymbol{F}$. From Eq. (5.48), one has

$$\begin{aligned}
\boldsymbol{F} &= \mathbb{E}\left[\left(\sum_{i=1}^{L} \frac{1}{\sigma_{b,i}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_i \, (\boldsymbol{X}_i - \boldsymbol{s}_i) \right) \left(\sum_{j=1}^{L} \frac{1}{\sigma_{b,j}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_j \, (\boldsymbol{X}_j - \boldsymbol{s}_j) \right)^{\top} \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm tr}} \right] \\
&= \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \underbrace{\mathbb{E}\left[(\boldsymbol{X}_l - \boldsymbol{s}_l) (\boldsymbol{X}_l - \boldsymbol{s}_l)^{\top} \right]}_{=\boldsymbol{\Sigma}_{3,l}(\boldsymbol{\theta}_{\rm t})} \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm tr}} \qquad (5.51) \\
&\quad + \sum_{i=1}^{L} \sum_{\substack{j=1 \\ j\neq i}}^{L} \frac{1}{\sigma_{b,i}^2} \frac{1}{\sigma_{b,j}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_i \underbrace{\mathbb{E}\left[(\boldsymbol{X}_i - \boldsymbol{s}_i) (\boldsymbol{X}_j - \boldsymbol{s}_j)^{\top} \right]}_{=0} \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_j \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm tr}}. \qquad (5.52)
\end{aligned}$$

Since $\mu_c = 1$ for all the molecular lines, one recognizes in Eq. (5.51) the covariance matrix of the true distribution, i.e., $\boldsymbol{\Sigma}_{2,l}(\boldsymbol{\theta}_{\rm tr}) = \sigma_{c,l}^2 \, \boldsymbol{s}_l(\boldsymbol{\theta}_{\rm tr}) \boldsymbol{s}_l^{\top}(\boldsymbol{\theta}_{\rm tr}) + \sigma_{b,l}^2 \, \boldsymbol{I}_N$. Moreover, the term given by Eq. (5.52) is null since the molecular lines are independent. One finally obtains

$$\boxed{\boldsymbol{F} = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^4} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \left(\sigma_{c,l}^2 \, \boldsymbol{s}_l \boldsymbol{s}_l^{\top} + \sigma_{b,l}^2 \, \boldsymbol{I}_N \right) \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm tr}}.}$$

Let us now detail $\boldsymbol{G}$, defined in Eq. (5.50). From Eq. (5.48), one finds

$$\nabla_{\boldsymbol{\theta}}^2 \ln p_{\boldsymbol{Y},\not\phi}(\boldsymbol{x}|\boldsymbol{\theta}) = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \left(\nabla_{\boldsymbol{\theta}}^2 \boldsymbol{s}_l \, (\boldsymbol{x}_l - \boldsymbol{s}_l) - \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l \right). \tag{5.53}$$

Then, performing the expectation on Eq. (5.53) leads to

$$\boldsymbol{G} = -\sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}}^2 \boldsymbol{s}_l \underbrace{\mathbb{E}\left[\boldsymbol{X}_l - \boldsymbol{s}_l \right]}_{=0} \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm tr}} + \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm tr}}.$$

Finally,

$$\boxed{\boldsymbol{G} = \sum_{l=1}^{L} \frac{1}{\sigma_{b,l}^2} \nabla_{\boldsymbol{\theta}} \boldsymbol{s}_l \nabla_{\boldsymbol{\theta}}^{\top} \boldsymbol{s}_l \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\rm tr}}.}$$

5.F Quantification of the precision loss under the misspecified model

De la possibilité d'ignorer un bruit de calibration lors de la construction d'un estimateur du maximum de vraisemblance en radio-astronomie

Léontine SÉGAL[1,2] Antoine ROUEFF[2] Claude JAUFFRET[2] Jérôme PETY[1,3] Maryvonne GERIN[3] et ORION-B [4]

[1]Institut de Radioastronomie Millimétrique, 300 rue de la Piscine, 38400 Saint-Martin-d'Hères, France

[2]Université de Toulon, Aix Marseille Univ, CNRS, IM2NP, Toulon, France

[3]LUX, France

[4]https://www.iram.fr/~pety/ORION-B/team.html

Résumé – Nous présentons une méthodologie permettant de quantifier la perte de précision sur l'estimation de paramètres d'intérêt lorsqu'un bruit de calibration est négligé. L'approche repose sur l'utilisation des bornes de Cramér-Rao (BCR) et BCR modifiée (BCRM), cette dernière étant une borne inférieure sur la matrice de covariance d'un estimateur construit avec un modèle mal spécifié. La méthode est illustrée sur un cas théorique simple avant d'être appliquée au cas de données astrophysiques, où on estime la température cinétique et la densité volumique du gaz interstellaire. La perte de précision mesurée, inférieure à 13%, est négligeable compte-tenu de la complexité du problème astrophysique considéré.

Abstract – We propose a method for quantifying the loss of precision in the estimation of parameters of interest when a calibration noise is neglected. The method is based on the use of the Cramér-Rao Bound (CRB) and the modified CRB, this latter being a lower bound on the covariance matrix of an estimator built with a misspecified model. The approach is illustrated on a simple theoretical case before being applied to the case of astrophysical data, where we estimate the kinetic temperature and the volume density of the interstellar medium. The accuracy loss, less than 13%, is negligeable given the complexity of the problem under consideration.

1 Introduction

On s'intéresse aux émissions du gaz composant les nuages interstellaires, régions où se forment les étoiles. Ces nuages moléculaires froids ($\leq 100\,\mathrm{K}$) sont observés dans le domaine des ondes millimétriques. Chaque espèce chimique émet un spectre d'émission, ou « raie moléculaire », observé sur N canaux fréquentiels. Comme proposé dans [2], les observations se modélisent par

$$\boldsymbol{x} = c\,\boldsymbol{s}\,(\boldsymbol{\theta}) + \boldsymbol{b} \tag{1}$$

où $\boldsymbol{x} = (x_1 \cdots x_N)^\top$ est entâché de deux bruits statistiquement indépendants : un bruit additif $\boldsymbol{b} = (b_1 \cdots b_N)^\top$ tel que $b_{n\in[\![1,N]\!]} \overset{i.i.d.}{\sim} \mathcal{N}(0,\,\sigma_b^2)$ et un bruit de calibration multiplicatif $c \in \mathbb{R}$. Le signal d'intérêt $\boldsymbol{s} = (s_1 \cdots s_N)^\top$ est fonction du paramètre $\boldsymbol{\theta} \in \mathbb{R}^d$ caractérisant le milieu. Le bruit additif est dominé par l'interaction entre l'atmosphère et les photons avant qu'ils ne parviennent au télescope terrestre. Le bruit de calibration résulte de la stratégie d'observation et des incertitudes liées aux éléments successifs de la chaîne de mesure. Dans d'autres travaux, ce terme multiplicatif est aussi fonction de $\boldsymbol{\theta}$ et contribue donc à caractériser la source [8].

Pour le jeu de données acquis dans le cadre de la campagne d'observation du Large Program ORION-B (P.I. : J. Pety & M. Gerin), on mesure une dispersion du bruit de calibration $\sigma_c \sim 0.05$ à 0.1 selon la fenêtre atmosphérique considérée (de $3\,\mathrm{mm}$ à $1\,\mathrm{mm}$) [2]. Pour analyser ces données ORION-B, nous avons construit dans [12] un estimateur du maximum de vraisemblance (MV) de $\boldsymbol{\theta}$ en négligeant le bruit de calibration, car de faible dispersion. Dans cet article, nous présentons une méthodologie afin de quantifier la perte de précision sur l'estimation de $\boldsymbol{\theta}$ induite par ce choix.

Ayant négligé le bruit de calibration, nous considérons la borne de Cramér-Rao Modifiée (BCRM) [3]. En effet, comme illustré dans les travaux en géolocalisation [7] et applications RADAR [9], la BCRM permet de prendre en compte les erreurs de modèle dans la quantification des performances d'estimation. La BCR hybride [6], adaptée à l'estimation jointe de paramètres déterministes et aléatoires, aurait pu être considérée. Toutefois, le paramètre de nuisance aléatoire c peut être marginalisé pour de faibles σ_c, afin de n'estimer que le paramètre d'intérêt déterministe $\boldsymbol{\theta}$. La borne de Cramér-Rao (BCR) sur ce modèle marginalisé nous fournit une précision de référence. Notre analyse de précision repose sur la comparaison de la BCRM à cette BCR.

Cet article est construit comme suit. La Sec. 2 présente une étude de précision théorique menée dans le cadre d'un modèle simplifié de Eq. (1) permettant d'effectuer des calculs analytiques. La Sec. 3 présente l'application de la méthodologie aux observations astrophysiques du projet ORION-B, dans le cadre de l'estimation de la température cinétique et de la densité volumique du gaz interstellaire. Les resultats obtenus et perspectives sont discutés en Sec. 4.

2 Étude théorique

Dans cette étude, on montre que la distribution du bruit de calibration peut s'approximer par une loi normale pourvu que l'écart-type soit inférieur à ~ 0.3. Ensuite, nous calculons analytiquement la perte de précision relative induite lorsque ce bruit est négligé pour un modèle simplifié de Eq. (1).

2.1 Approximation Gaussienne de c

Comme le montre [4], la loi log-normale, notée log-$\mathcal{N}$, est appropriée pour décrire un bruit multiplicatif tel que c. La Fig. 1a présente trois densités log-$\mathcal{N}$ de paramètres $\{\mu,\, \sigma^2\}$ variables, indiqués en légende. Une loi normale $\mathcal{N}(\mu_c, \sigma_c^2)$ avec $\sigma_c = 0.1$ et $\mu_c = 1$ se confond à la log-$\mathcal{N}$ de plus faible dispersion. L'approximation Gaussienne devient critiquable pour $\sigma_c \gtrsim 0.3$, compte-tenu de l'asymétrie apparente de la log-$\mathcal{N}$ (courbe orange continue). La caractérisation du bruit de calibration faite par [2] indique que $0.05 \leq \sigma_c \leq 0.1$ pour les données ORION-B. Dans cette étude, nous considérerons donc que $c \sim \mathcal{N}(\mu_c = 1, \sigma_c^2)$ et gardons à l'esprit que le modèle de mesures construit n'est valable que pour $\sigma_c \leq 0.3$.

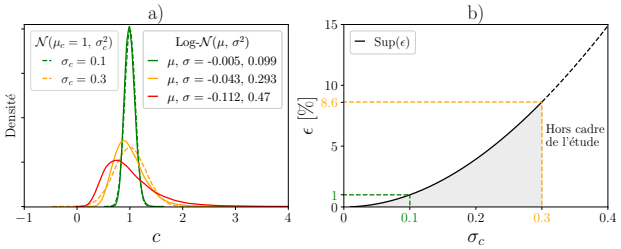


FIGURE 1 : a) Log-$\mathcal{N}$ (courbes continues) et $\mathcal{N}$ (discontinues). b) Écart relatif ϵ sur la précision d'estimation (%) (défini Sec. 2.2.3) en fonction de σ_c pour différentes valeurs de $\{N, \sigma_b, \theta\}$ et $\mu_c = 1$ (région grisée). La borne supérieure Sup(ϵ) est en noir (continue).

2.2 Illustration sur un modèle simplifié

2.2.1 Description du modèle

A titre didactique, on considère le modèle suivant

$$\boldsymbol{x} = c\,\theta \mathbf{1}_N + \boldsymbol{b} \tag{2}$$

avec $\mathbf{1}_N = (1 \cdots 1)^\top$ de taille N, $\theta \in \mathbb{R}$, $c \sim \mathcal{N}(\mu_c,\, \sigma_c^2)$ et $\boldsymbol{b} \sim \mathcal{N}(\mathbf{0}_N,\, \sigma_b^2 \boldsymbol{I}_N)$ où $\mathbf{0}_N = (0 \cdots 0)^\top$ de taille N et $\boldsymbol{I}_N$ la matrice identité de taille N. Les bruits $\boldsymbol{b}$ et c étant Gaussiens et indépendants, le couple d'éléments aléatoires $(\boldsymbol{X}, C)$ suit aussi une distribution Gaussienne. La marginalisation $p(\boldsymbol{x}|\theta) = \int p_{\boldsymbol{X},C}(\boldsymbol{x}, c|\theta)\, dc$, souvent difficile à calculer analytiquement, s'exprime simplement ici car $\boldsymbol{X}$ est la somme de deux vecteurs aléatoires Gaussiens. Dans notre cas, la distribution marginale est

$$p(\boldsymbol{x}|\theta) = \frac{1}{\sqrt{\det(2\pi\boldsymbol{\Sigma})}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)\right\} \tag{3}$$

de moyenne $\boldsymbol{\mu} = \mu_c \theta \mathbf{1}_N$ et de matrice de covariance $\boldsymbol{\Sigma} = \sigma_c^2\, \theta^2 \mathbf{1}_N \mathbf{1}_N^\top + \sigma_b^2 \boldsymbol{I}_N$. De plus, on a

$$\det(\boldsymbol{\Sigma}) = \sigma_b^{2(N-1)} \gamma^2, \quad \text{avec} \quad \gamma^2 = \sigma_c^2 N \theta^2 + \sigma_b^2 \tag{4}$$

et la formule de *Sherman-Morrison* [1] permet de calculer

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_b^2}\left(\boldsymbol{I}_N - \frac{\sigma_c^2 \theta^2}{\gamma^2} \mathbf{1}_N \mathbf{1}_N^\top\right) \tag{5}$$

et ainsi d'avoir une formulation explicite de la vraisemblance.

2.2.2 Calcul des précisions sur l'estimation de θ

La BCR est une borne inférieure de la matrice de covariance d'un estimateur sans biais [5]. Notée $\mathcal{B}$, elle est l'inverse de la matrice d'information de Fisher, soit $\mathcal{B}(\theta) = \mathcal{I}(\theta)^{-1}$, où

$$\mathcal{I}(\theta) = \mathbb{E}\left[\frac{\partial \ln p(\boldsymbol{x}|\theta)}{\partial \theta} \frac{\partial \ln p(\boldsymbol{x}|\theta)}{\partial \theta}^\top\right]. \tag{6}$$

Pour le modèle Eq. (2), $\mathcal{I}(\theta) = \frac{N}{\gamma^4}\left(\mu_c^2 \gamma^2 + 2\,\sigma_c^4 N \theta^2\right)$ et

$$\mathcal{B}(\theta) = \frac{\gamma^2}{N} \frac{1}{\mu_c^2 + \frac{2\,\sigma_c^4 N \theta^2}{\gamma^2}}. \tag{7}$$

Ignorer la présence du bruit de calibration, comme cela a été fait dans [11, 12], équivaut à attribuer aux mesures la distribution approchée $p_{\not c}(\boldsymbol{x}|\theta)$,

$$p_{\not c}(\boldsymbol{x}|\theta) = \frac{1}{\sqrt{(2\pi\sigma_b^2)^N}} \exp\left\{-\frac{1}{2\sigma_b^2}\sum_{n=1}^{N}(x_n - \theta)^2\right\}. \tag{8}$$

Pour $\mu_c = 1$, l'approximation de Eq. (3) par (8) n'induit pas de biais d'estimation. La BCRM construite à partir de $p_{\not c}(\boldsymbol{x}|\theta)$, notée $\mathcal{B}_M(\theta)$, est définie alors par $\mathcal{B}_M(\theta) = A^{-1} J A^{-1}$ [10, 9], où

$$A = -\mathbb{E}_{p(\boldsymbol{x}|\theta)}\left[\frac{\partial^2 \ln p_{\not c}(\boldsymbol{x}|\theta)}{\partial^2 \theta}\right] \tag{9}$$

et

$$J = \mathbb{E}_{p(\boldsymbol{x}|\theta)}\left[\frac{\partial \ln p_{\not c}(\boldsymbol{x}|\theta)}{\partial \theta} \frac{\partial \ln p_{\not c}(\boldsymbol{x}|\theta)}{\partial \theta}^\top\right]. \tag{10}$$

Ici, $A = -\frac{N}{\sigma_b^2}$ et $J = \frac{N}{\sigma_b^4}\left(\sigma_c^2 \theta^2 N + \sigma_b^2\right)$, d'où

$$\mathcal{B}_M(\theta) = \frac{\gamma^2}{N}. \tag{11}$$

Comme montré dans [6], $\mathcal{B}_M(\theta) \geq \mathcal{B}(\theta)$:

$$\mathcal{B}(\theta) = \mathcal{B}_M(\theta) \frac{1}{\mu_c^2 + \frac{2\,\sigma_c^4 N \theta^2}{\gamma^2}}, \quad \mu_c = 1 \quad \text{et} \quad \frac{2\,\sigma_c^4 N \theta^2}{\gamma^2} \geq 0.$$

2.2.3 Perte de précision sur l'estimation de θ

On quantifie la perte de précision sur l'estimation de θ avec l'écart relatif

$$\epsilon(\tau) \triangleq \frac{\sqrt{\mathcal{B}_M} - \sqrt{\mathcal{B}}}{\sqrt{\mathcal{B}}} \tag{12}$$

Ici, $\epsilon(\tau) = \sqrt{\mu_c^2 + \frac{2\,\sigma_c^2}{1+\tau}} - 1$, où $\tau = \frac{\sigma_b^2}{\sigma_c^2 N \theta^2}$. La fonction $\epsilon(\tau)$, strictement décroissante sur $\mathbb{R}^*$, est majorée par

$$\text{Sup}(\epsilon) = \lim_{\tau \to 0} \epsilon(\tau) = \sqrt{\mu_c^2 + 2\,\sigma_c^2} - 1. \tag{13}$$

Pour ce modèle particulier Eq. (2), la perte de précision induite par la négligence du bruit de calibration est essentiellement régie par la dispersion σ_c. On retrouve en particulier, que lorsque $2\,\sigma_c^2 \ll \mu_c^2$, l'écart relatif tend à s'annuler, traduisant l'équivalence des deux modèles. La Fig. 1b montre en fonction de σ_c l'écart relatif ϵ, pour $\mu_c^2 = 1$. On constate que l'écart maximal, Sup(ϵ), demeure inférieur à 10% pour $\sigma_c \leq 0.3$. L'écart relatif augmente à mesure que σ_c augmente. On observe ainsi un lien entre la perte de précision et la qualité de l'approximation Gaussienne d'une loi log-$\mathcal{N}$.

3 Étude sur les données astrophysiques

Après avoir illustré notre méthodologie sur le modèle théorique, nous examinons une partie des données ORION-B, correspondant aux observations de la nébuleuse de la Tête de Cheval [12].

3.1 Analyse du milieu interstellaire

Notre compréhension de la formation stellaire s'appuie essentiellement sur l'ajustement d'un modèle physique aux observations, ce qui permet d'estimer notamment la température cinétique (T en K) et la densité volumique (n en cm^{-3}) du gaz. L'observation de la Tête de Cheval sur une large bande fréquentielle a permis de mesurer une vingtaine de raies moléculaires par pixel. Comme le montre la Fig. 2a, observer simultanément différentes raies moléculaires permet de sonder l'évolution des propriétés du gaz le long de la ligne de visée.

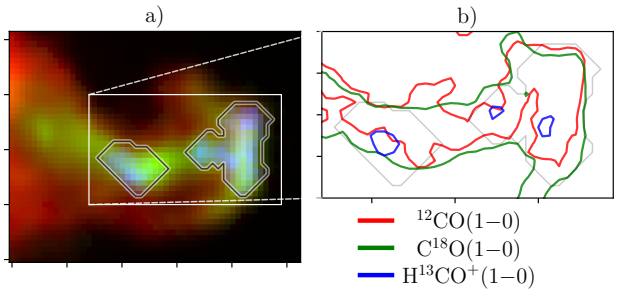


FIGURE 2 : a) Émissions intégrées normalisées ($\sum_n x_n / \max \boldsymbol{x}$) des raies $(1-0)$ de ^{12}CO (canal rouge), C^{18}O (vert) et H^{13}CO$^+$ (bleu) pour les lignes de visées où RSB ≥ 3. Les régions les plus denses sont délimitées en noir. b) Zones où le RSB $\geq 250, 50, 10$ pour ^{12}CO (rouge), C^{18}O (vert), et H^{13}CO$^+$ (bleu), respectivement.

L'émission de la raie $^{12}\mathrm{CO}(1-0)$ (en rouge) permet de sonder l'enveloppe chaude et diffuse ($\geq 30\,\mathrm{K}$, $\leq 10^3\,\mathrm{cm}^{-3}$). L'émission de $\mathrm{C}^{18}\mathrm{O}(1-0)$ (en vert) provient d'une région plus compacte d'aspect filamentaire, de température et densité intermédiaires. La raie $\mathrm{H}^{13}\mathrm{CO}^+(1-0)$ (en bleu) n'est détectée que dans deux régions de la nébuleuse, appelées « cœurs denses » car la densité y est la plus élevée ($\geq 10^5\,\mathrm{cm}^{-3}$). Ces régions sont aussi les plus froides ($\sim 5-15\,\mathrm{K}$) car protégées des radiations extérieures. C'est au sein des cœurs denses que se formeront des étoiles.

Le code de transfert radiatif RADEX [13] permet de simuler une raie moléculaire à partir des composantes du paramètre $\boldsymbol{\theta}$ (dont T et n) d'un milieu homogène. Afin de tenir compte de l'hétérogénéité dans la ligne de visée, nous avons proposé dans [12] un modèle de nuage à 3 couches, disposées en « sandwich » (voir Fig. 3). Cette approche permet de décrire les régions denses de la nébuleuse par un cœur dense et froid, enveloppé dans du gaz plus diffus et chaud ([12], Fig. 16). Comme l'illustre la Fig. 2b, le rapport signal sur bruit (RSB), défini par $\frac{\max \boldsymbol{x}}{\sigma_b}$, diffère de plus d'un ordre de grandeur pour différentes raies moléculaires au sein d'un même pixel. Cette disparité est problématique car la simplicité du modèle sandwich pour décrire le nuage induit inévitablement des erreurs de modèle. Nous modérons donc la confiance attribuée aux raies de fort RSB au bénéfice des raies plus bruitées, ces dernières pouvant sonder les cœurs denses du nuage. Pour cela, nous saturons le RSB à 10 en augmentant la variance du bruit additif dans le calcul de la vraisemblance. Les données restent intouchées.

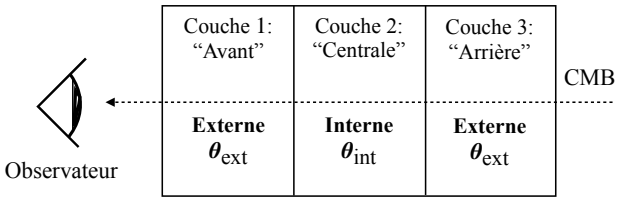


FIGURE 3 : Modèle « sandwich » de la ligne de visée. L'émission mesurée résulte des contributions de l'avant et arrière de la couche externe, de la couche interne et du fond diffus cosmologique (Cosmic Microwave Background (CMB)). Les couches externe et interne sont respectivement caractérisées par $\boldsymbol{\theta}_{\mathrm{ext}} = (T_{\mathrm{ext}}, n_{\mathrm{ext}}, \cdots)$ et $\boldsymbol{\theta}_{\mathrm{int}} = (T_{\mathrm{int}}, n_{\mathrm{int}}, \cdots)$.

Le bruit de calibration ayant été négligé dans nos travaux [11, 12], nous examinons ici les conséquences de ce choix à l'aide de la méthodologie présentée Sec. 2.

3.2 Quantification de la perte de précision

Le modèle sandwich construit dans [12], nécessite l'estimation de 10 inconnues physico-chimiques supplémentaires, en plus de celle des températures et densités volumiques. Dans cet article, nous nous limitons à l'examen de T et n des couches externe et interne du nuage, indicées respectivement ext et int (voir Fig.3). La Fig. 4 montre l'écart relatif ϵ (cf. Eq.(12)) pour les estimations de $\log(T)$ et $\log(n)$ obtenues dans [12].

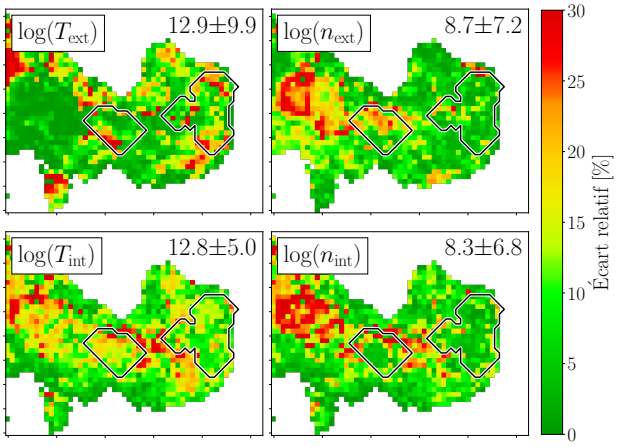


FIGURE 4 : Écart relatif ϵ (%) sur la précision d'estimation de $\log(T)$ et $\log(n)$ dans les couches externe (1ère ligne) et interne (2ème ligne). La moyenne et l'écart-type indiqués sont calculés dans les zones denses, délimitées en noir.

Nous observons un écart relatif des précisions d'estimation $\sim 13\%$ pour les zones les plus denses. Cet écart atteint au plus 30% dans les régions plus diffuses, où le modèle à trois couches est le moins adapté. Ces précisions asymptotiques étant fournies par les BCRs, il convient d'étudier l'efficacité de l'estimateur utilisé, qui néglige le bruit de calibration. Cette analyse est menée dans la section suivante.

3.3 Performances de l'estimateur sur simulations Monte-Carlo

Afin de construire un jeu de spectres correspondant à l'emission provenant d'un cœur dense et froid ($n_{\text{int}}, T_{\text{int}} = 10^5\,\text{cm}^{-3}, 13\,\text{K}$) enveloppé par un gaz chaud et diffus ($n_{\text{ext}}, T_{\text{ext}} = 10^{3.8}\,\text{cm}^{-3}, 28\,\text{K}$), nous utilisons le modèle de mesures Eq. (1), où $\boldsymbol{s}$ est décrit par le sandwich (voir [12], Eq. 8). La Fig. 5 présente sous forme d'histogrammes les résultats d'estimation obtenus par Monte-Carlo avec l'estimateur du MV.

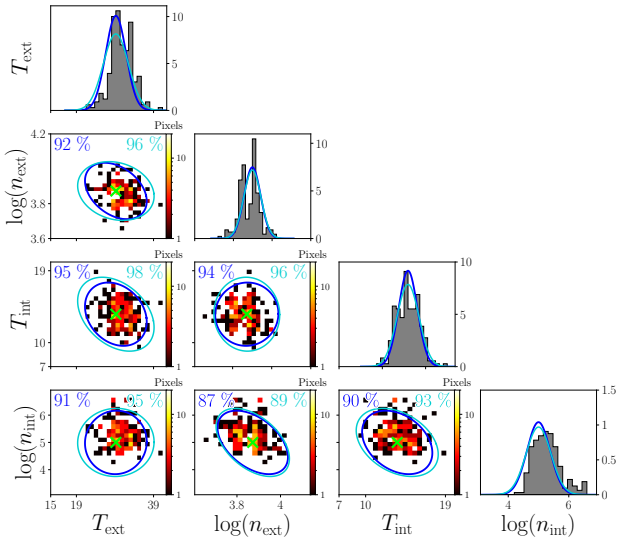


FIGURE 5 : Histogrammes des estimations sur jeu de données simulées (200 réalisations). La valeur vraie est indiquée par la croix verte. Les ellipses de confiance à 99% décrites par la BCR et BCRM sont respectivement tracées en bleu et cyan. Les proportions (%) des estimations contenues dans ces ellipses sont indiquées en haut de chaque panneau.

Comme obtenu dans les études précédentes, les BCR et BCRM sont proches (voir ellipses, Fig.5). La proportion de points contenus dans les ellipses est de 89 à 98% alors qu'elle serait de 99% pour un estimateur efficace. Les histogrammes traduisent un comportement Gaussien de l'estimateur, excepté celui de $\log(n_{\text{int}})$, où un deuxième mode apparaît à $n_{\text{int}} \sim 10^6\,\text{cm}^{-3}$.

4 Conclusion et perpectives

La méthode présentée dans ce travail a permis de quantifier que pour une faible dispersion ($\sigma_c \leq 0.1 \times \mu_c$), négliger un bruit de calibration lors de la construction de l'estimateur du MV induit une perte de précision $\lesssim 13\%$ sur l'estimation des paramètres d'intérêt. Compte-tenu des ordres de grandeur de ces variables astrophysiques et de la complexité du problème considéré, cette perte semble négligeable. Ce constat repose toutefois sur l'analyse des cas particuliers du modèle théorique Eq. (2) et d'un modèle physique adapté aux données radio-astronomiques. Il serait donc intéressant d'examiner le cas de systèmes de mesures différents. En particulier, mener à bien le calcul de marginalisation lorsque le bruit de calibration présente une forte dispersion telle que l'approximation Gaussienne n'est plus vérifiée (≥ 0.3) constituera un défi important.

5 Remerciements

Ce travail a été soutenu par l'Agence Nationale de la Recherche sous la subvention DAOISM (ANR-21-CE31-0010).

Références

[1] M. S. BARTLETT : An inverse matrix adjustment arising in discriminant analysis. *The Annals of Mathematical Statistics*, 22(1):107–111, 1951.

[2] L. EINIG et le consortium ORION-B : Deep learning denoising by dimension reduction : Application to the orion-b line cubes. *A&A*, 677:A158, 2023.

[3] F. GINI, R. REGGIANNINI et U. MENGALI : The modified cramer-rao bound in vector parameter estimation. *IEEE Trans. on Communications*, 46(1):52–60, 1998.

[4] N L. JOHNSON, S. KOTZ et N. BALAKRISHNAN : *Continuous multivariate distributions*, volume 1, page 209. Wiley New York, 1972.

[5] S. M. KAY : *Fundamentals of statistical signal Proc. : estimation theory*, chapitre 3. Prentice-Hall, Inc., USA, 1993.

[6] Y. NOAM et H. MESSER : Notes on the tightness of the hybrid cramér - rao lower bound. *IEEE Trans. on Signal Proc.*, 57(6):2074–2084, 2009.

[7] L. ORTEGA, J. VILÀ-VALLS et E. CHAUMETTE : Theoretical evaluation of the gnss synchronization performance degradation under interferences. *In ION GNSS+ 2022*.

[8] P. PALUD et le consortium ORION-B : Mélange de bruits et échantillonnage de posterior non log-concave. *In Actes du GRETSI 2022*.

[9] C. REN, M. N. El KORSO, J. GALY, E. CHAUMETTE, P. LARZABAL et A. RENAUX : Performance bounds under misspecification model for mimo radar application. *In 2015 23rd EUSIPCO*, pages 514–518, 2015.

[10] C. D. RICHMOND et L. L. HOROWITZ : Parameter bounds under misspecified models. *In 2013 Asilomar Conference on Signals, Systems and Computers*, pages 176–180, 2013.

[11] A. ROUEFF et le consortium ORION-B : Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model - Application to the estimation of the gas temperature and volume density. *A&A*, 686:A255, 2024.

[12] L. SÉGAL et le consortium ORION-B : Toward a robust physical and chemical characterization of heterogeneous lines of sight : The case of the horsehead nebula. *A&A*, 692:A160, 2024.

[13] F. F. S. van der TAK, J. H. BLACK, F. L. SCHÖIER, D. J. JANSEN et E. F. van DISHOECK : A computer program for fast non-LTE analysis of interstellar line spectra. With diagnostic plots to interpret observed line intensity ratios. *A&A*, 468(2):627–635, juin 2007.

Chapter 6

Conclusion and perspectives

Contents

6.1 On the road to a reliable estimate of the physicochemical characteristics of interstellar molecular clouds, despite the lack of reference data **171**
6.2 Perspectives **173**
6.3 Vers des estimations fiables des caractéristiques physico-chimiques des nuages interstellaires en l'absence de vérité terrain **175**
6.4 Perspectives **177**

6.1 On the road to a reliable estimate of the physicochemical characteristics of interstellar molecular clouds, despite the lack of reference data

With the current and next generation of spectrometers in the millimeter wavelength range, observing vast regions of giant molecular clouds (GMCs) through numerous molecular tracers of gas will become routine. Moreover, the study of molecular lines first provides information on the amount of matter, morphology, and kinematics of the medium, and then on how competition between turbulence, magnetic field, and gravitation structures the gas into dense filaments and cores. Beyond these initial considerations, the extreme conditions (low temperatures and very low densities) that prevail in these clouds imply that the emission of molecular lines is out of local thermodynamic equilibrium. Line emission is therefore sensitive to the number and energy of collisions, and hence to the gas volume density and kinetic temperature.

Obtaining wide-field maps of the kinematics and physicochemical conditions of the gas (density, kinetic temperature, molecular abundances) could change our understanding of the formation of stars and their planetary systems. To achieve this goal, however, we need to infer the physicochemical quantities we are looking for from noisy observations, over millions of pixels, and which are subject to calibration uncertainty. What's more, there is no direct ground truth, since it is (still) impossible to make in-situ measurements, while remote observations can only inform us about part of the complexity of the physicochemical

phenomena at work.

In these conditions, the two main challenges are 1) to assess how much information can really be extracted from the data, and 2) to determine the reliability of the estimated quantities. This requires not only a detailed understanding of the most significant physicochemical processes for obtaining a reasonable model of line emission, but also the introduction of state-of-the-art inference concepts to reveal the most robust results and their limits. My thesis proposes some keys to realizing part of this program.

I first co-developed an inference method to estimate from complementary molecular lines the following parameters of interest: the kinetic temperature ($T_{\rm kin}$), the volume density ($n_{\rm H_2}$), and thermal pressure ($P_{\rm th}$) of the gas, the column densities of the considered molecules (N), centroid velocity (C_V), and the velocity dispersion (σ_V) along the considered line of sight. Monte-Carlo simulations were used to study whether the proposed estimator is unbiased and efficient. When this was the case, I used the Cramér-Rao bound (CRB) to estimate confidence intervals on the different parameters.

In the absence of ground truth and in view of limiting the number of unknown parameters, the approach was to start by considering the simplest possible model of the environment (in particular, a single homogeneous layer of gas per line of sight), and then to enrich the model until the residues no longer showed any obvious bias. The evolution of the model took into account pre-existing knowledge of molecular clouds, such as their most plausible structure along a line of sight, or the chemical links between the species whose lines were studied. Once the model was established, I interpreted the results obtained within their confidence intervals.

More precisely, Roueff *et al.* 2024 used a homogeneous cloud model on a portion of Orion B, the Horsehead nebula, to reproduce the lines of the CO and HCO^+ isotopologues. This study detected the presence of biases on $T_{\rm kin}$, $n_{\rm H_2}$, and column densities deduced as a function of the subset of lines selected and the knowledge injected on chemical abundances. An in-depth Monte-Carlo study showed that the model is too simple to describe the multi-line observations considered, and that the model *misspecifications* lead to the biases observed on the real data. Indeed, the molecular tracers studied probe different parts of the line of sight characterized by a specific chemistry: either the dense, cold gas, or its warmer, translucent environment.

Ségal *et al.* 2024 propose a multilayer model to take into account the heterogeneity of the physicochemical conditions of the gas along the line of sight. For example, the lines of sight toward dense cores are represented by a so-called "sandwich", where the dense, cold core is surrounded by warmer, translucent gas. To limit the variance on the estimates, we also propose a simple chemical model that depends on the layer: for the inner layer, we fix the relative abundances of the CO isotopologues to typical ISM values (i.e., $[^{12}CO] \,/\, [^{12}CO] \sim 63$ and $[^{12}CO] \,/\, [C^{18}O] \sim 7.9$), while all the CO column densities are fitted in the translucent outer layer. For the HCO^+ isotopologues, the column density of $H^{12}CO^+$ is fitted in the two layers while the column density of $H^{13}CO^+$ is derived from other column densities through $N(H^{13}CO^+) = N(H^{12}CO^+) \times N(^{13}CO)/N(^{12}CO)$. This allowed me to obtain the following results.

1. A single-layer model cannot reproduce the spectral line asymmetries that result from a combination of different radial velocities and absorption effects among layers. A mini-

mal heterogeneous model (three layers only) is sufficient for the case of the Horsehead nebula, as it fits the seven considered molecular lines over a large part of the field of view.

2. The decomposition of the intensity into three layers allowed us to discuss the distribution of the estimated physical or chemical properties along the line of sight. About 80% of the integrated intensity of ^{12}CO comes from the outer envelope, while $\sim 55\%$ of the integrated intensity of the $(1-0)$ and $(2-1)$ lines of C^{18}O originates from the inner layer. For the lines of ^{13}CO and the HCO$^+$ isotopologues, integrated intensities are more equally distributed over the cloud layers.

3. The estimated column density ratio $N(^{13}\mathrm{CO})/N(\mathrm{C}^{18}\mathrm{O})$ in the envelope increases with decreasing visual extinction, and it reaches 25 in the pillar outskirts. The typical mean abundance of HCO$^+$ with respect to $\mathrm{H_2}$ is $\sim 1.3 \times 10^{-9}$, which is about two times lower than the usual value in diffuse clouds.

4. Both the abundances of ^{12}CO and HCO$^+$ increase almost linearly with the $\mathrm{H_2}$ column density, i.e., $[^{12}\mathrm{CO}] \propto N(\mathrm{H_2})^{0.64}$ until $N(\mathrm{H_2}) = 6.4 \times 10^{21}\,\mathrm{cm^{-2}}$, and then decreases. This results from the freezing of molecules on dust grains. Besides, $[\mathrm{HCO^+}] \propto N(\mathrm{H_2})^{0.67}$ until $N(\mathrm{H_2}) = 2.7 \times 10^{21}\,\mathrm{cm^{-2}}$, and then stays approximately constant.

5. In addition to confirming the general rotation of the Horsehead pillar around its axis, we found the signature of accreting motions of the foreground layer onto the inner denser gas.

6. While the inferred T_{kin} of the envelope varies from 25 to 40 K, that of the inner layer drops to $\sim$ 15 K in the western dense core. The estimated $n_{\mathrm{H_2}}$ in the inner layer is $\sim 3 \times 10^4\,\mathrm{cm^{-3}}$ toward the filament, and it increases by a factor of ten toward dense cores. The thermal pressure of the outer layer shows a decreasing east-west gradient.

In these two studies, calibration noise was neglected during the estimation task, while Einig *et al.* 2023 estimated calibration uncertainties of the order of 5% and 10% at the 3 and 1 mm band, respectively. To confirm the validity of this choice, I carried out an exhaustive study of the different ways of taking calibration noise into account. When the calibration uncertainty is less than 10%, deliberately neglecting calibration noise leads 1) to unbiased estimation results, and 2) to a loss of accuracy of between 10 and 30%, which is deemed insignificant for the astrophysical problem under consideration. The latter results confirm our decision to neglect multiplicative noise here. In future works, we nevertheless advise to compute the misspecified CRB to improve the relevance of the confidence intervals.

Finally, I also used the CRB to study the potential gain in precision of the estimates if all the molecular lines were well resolved in velocity (i.e., at 0.1 km/s).

6.2 Perspectives

I now describe three perspectives of my thesis.

Chemistry analysis of the Horsehead nebula — So far, I have only considered 7 molecular lines from the ORION-B dataset. However, a dozen other molecular species have been detected, such as CS, SO, CCH, C_3H_2, and N_2H^+. It is not easy to use them to constrain physical parameters (i.e., $T_{\rm kin}$ and $n_{\rm H_2}$), as their chemistry (and therefore their relative abundance) is difficult to model simply. On the other hand, it is now possible to fix the physical parameters to estimate the associated column densities. The information contained in these faint emission lines could thus make a significant contribution to improving our understanding of chemistry of simple molecules in the ISM.

From a thousand to a million pixels — Until now, detailed studies of physicochemical conditions have mostly been carried out on just a few pixels. This is due to the difficulty of obtaining robust estimates without strong human supervision. The analysis carried out on the Horsehead nebula concerns around 1500 pixels, whereas the ORION-B data contains almost 1 million. Beyond the computation time required, the main difficulty in generalizing the developed analysis approach comes from the assumption that only one velocity component is present along each line of sight. The Horsehead nebula region was chosen precisely because it respects this assumption. This is not the case for the rest of the Orion B cloud, where most lines of sight have *multiple* velocity components.

Fitting components whose supports are disjoint along the velocity axis is trivial. It then remains to *partition* the Orion B cloud into regions with one, two, ... velocity components. This problem can be reformulated as a *model selection* one based on minimum description length (MDL). This involves choosing the best description of the data from a point of view of the information theory. The challenge arises at transition regions where two distinct velocity components gradually merge into a single velocity component, so that only one potentially asymmetric molecular line is observed. Solving this problem requires characterization of the fusion process. Although we have so far taken a pixel-by-pixel approach to estimating the spatial distribution of physicochemical parameters, the use of spatial information may prove to be useful in identifying the fusion mechanism.

From Orion B to other GMCs in the Milky Way — The aim of the ORION-B program is to calibrate our algorithms on a nearby GMC using lightly supervised techniques, before applying them to more distant GMCs. Deploying these methodologies on more distant clouds will require taking into account the loss of spatial resolution. Zakardjian *et al.* 2025 study this problem at totally different spatial scales, as they focus on extragalactic data. Nevertheless, this study underlines that success depends on the selection of a simplified representation of the mixing of physical conditions within an observed beam. This means that searching for a spatially coherent representation within a beam is too optimistic. So we shall have to find another statistical characterization of the observed map. For example, we could try to partition the cube data into different mixtures of the physical regimes of interest (e.g., diffuse and translucent gas, filaments, dense cores, etc.). Such a treatment would open the way to a new statistical description of the data.

6.3 Vers des estimations fiables des caractéristiques physico-chimiques des nuages interstellaires en l'absence de vérité terrain

Les générations actuelle et future de spectromètres dans le domaine millimétrique rendront courante l'observation de vastes régions de nuages interstellaires à l'aide de nombreux traceurs moléculaires. Par ailleurs, l'étude de raies moléculaires fournit d'abord des informations sur la quantité de matière, la morphologie et la cinématique des nuages, mais aussi sur la manière dont la compétition entre la turbulence, le champ magnétique et la gravitation structure le gaz en filaments et cœurs denses. Au-delà de ces considérations initiales, les conditions extrêmes (basses températures et très faibles densités) qui règnent dans les nuages impliquent que l'émission des raies moléculaires se trouve hors de l'équilibre thermodynamique local. Les raies d'émission sont ainsi sensibles au nombre et à l'énergie des collisions, et donc à la densité volumique et température cinétique du gaz.

Obtenir des cartes de grand champ de vue de la cinétique et des propriétés physico-chimiques du gaz (densité, température cinétique, abondances moléculaires) pourrait faire évoluer notre compréhension de la formation des étoiles et de leurs systèmes planétaires. Pour atteindre cet objectif, il est toutefois nécessaire d'inférer ces caractéristiques physico-chimiques d'intérêt à partir d'observations bruitées, soumises notamment à des incertitudes d'étalonnage, et cela, pour des millions de pixels. Enfin, nous ne disposons pas de vérité terrain directe compte-tenu de l'impossibilité d'effectuer (à ce jour) des mesures *in situ*, tandis que les observations à distance ne peuvent nous renseigner que sur une partie de la complexité des phénomènes physico-chimiques qui sont à l'œuvre.

Dans ces conditions, les deux principaux défis consistent à 1) évaluer la quantité d'information pouvant réellement être extraite des données, et 2) déterminer la fiabilité des quantités estimées. Cela nécessite non seulement une compréhension approfondie des processus physico-chimiques les plus importants, afin d'obtenir une modélisation raisonnable de raies d'émission, mais aussi d'introduire des concepts d'estimation de pointe permettant d'identifier les résultats les plus fiables et de révéler leurs limites. Ma thèse propose quelques clés pour réaliser ces objectifs.

Dans un premier temps, j'ai co-développé une méthodologie permettant d'estimer à partir de raies moléculaires complémentaires les paramètres d'intérêt suivants : les températures cinétiques (T_{kin}), les densités volumiques ($n_{\mathrm{H_2}}$), et les pressions thermiques du gaz, les densités de colonne des molécules considérées (N), ainsi que des controïdes de vitesse (C_V) et dispersions de vitesse (σ_V) au sein d'une ligne de visée donnée. Des simulations Monte-Carlo ont été menées afin de déterminer si l'estimateur proposé est non biaisé et efficace. Dans le cas échéant, la borne de Cramér-Rao (BCR) a été utilisée afin d'estimer les barres d'erreur sur les différents paramètres.

En l'absence de vérité terrain et afin de limiter le nombre de paramètres inconnus, l'approche fût de commencer par la modélisation la plus simple possible de l'environnement (c'est-à-dire, une couche de gaz homogène par ligne de visée), puis de l'enrichir progressivement jusqu'à ce que les résidus ne présentent plus aucun biais évident. L'évolution du modèle intègre les connaissances *a priori* dont nous disposons sur les nuages moléculaires,

comme leur structure la plus plausible le long de la ligne de visée, ou encore les abondances moléculaires reliant les espèces dont les raies sont étudiées. Une fois le modèle établi, j'ai pu intepréter les résultats d'estimation obtenus compte-tenu de leurs intervalles de confiance.

Plus précisemment, un modèle homogène de nuage fût utilisé dans Roueff *et al.* 2024 sur une portion d'Orion B, la nébuleuse de la Tête de Cheval, pour reconstruire les raies d'émission des isotopologues de CO et HCO^+. Cette étude a mis en évidence la présence de biais dans les estimations de $T_{\rm kin}$, $n_{\rm H_2}$, et des densités de colonne, selon le sous-ensemble de raies sélectionnées et des connaissances injectées sur les abondances chimiques. Une étude Monte-Carlo approfondie a montré qu'un modèle homogène est trop simple pour décrire les observations multi-raies considérées, et que cette *approximation* pourrait mener aux biais observés sur les données réelles. En effet, les traceurs moléculaires étudiés sondent différentes régions de la ligne de visée, chacune caractérisée par une chimie spécifique : on y trouve du gaz dense et froid, mais aussi translucide et chaud.

Un modèle multicouche est donc proposé dans Ségal *et al.* 2024 pour tenir compte de l'hétérogénéité des conditions physico-chimiques du gaz le long de la ligne de visée. Par exemple, les lignes de visée en direction des cœurs denses sont représentées par un modèle qualifié de "sandwich", où le noyau dense et froid est encapsulé par une couche de gaz chaude et translucide. Par ailleurs, afin de limiter la variance sur les estimations, nous avons proposé un modèle chimique simple qui dépend de la couche de gaz considérée : pour la couche interne, les abondances relatives des isotopologues de CO ont été fixées aux valeurs typiques du milieu interstellaire (i.e., $[^{12}CO]\,/\,[^{12}CO] \sim 63$ et $[^{12}CO]\,/\,[C^{18}O] \sim 7.9$), tandis qu'elles sont toutes estimées dans la couche externe translucide. Concernant les isotopologues de HCO^+, la densité de colonne de $H^{12}CO^+$ est estimée dans les deux couches, tandis que celle de $H^{13}CO^+$ est déduite telle que $N(H^{13}CO^+) = N(H^{12}CO^+) \times N(^{13}CO)/N(^{12}CO)$. Cette modélisation multicouche a permis d'obtenir les résultats suivants.

1. Un modèle homogène ne peut reproduire les asymétries des raies résultant d'une combinaison de vitesses radiales différentes et d'effets d'absorption entre les couches. Un modèle hétérogène minimaliste (trois couches seulement) est suffisant pour le cas de la nébuleuse de la Tête de Cheval, car s'ajuste aux sept raies d'émission moléculaires étudiées sur une grande partie du champ de vue.

2. La décomposition de l'intensité en trois couches nous a permis d'évaluer la distribution des propriétés physiques et chimiques estimées le long de la ligne de visée. Près de 80% de l'intensité intégrée de ^{12}CO provient de l'avant de la couche externe, tandis que $\sim 55\%$ de l'intensité intégrée des transitions $(1-0)$, $(2-1)$ de $C^{18}O$ provient de la couche interne. Concernant les raies de ^{13}CO et des isotopologues de HCO^+, l'intensité intégrée se trouve plus également répartie sur l'ensemble des couches du nuage.

3. Le ratio des densités de colonne $N(^{13}CO)/N(C^{18}O)$ dans l'enveloppe augmente lorsque que l'extinction visuelle diminue, et atteint 25 dans les régions périphériques du nuage. L'abondance moyenne typique de HCO^+ par rapport à H_2 est de $\sim 1.3 \times 10^{-9}$, soit environ 2 fois moins que la valeur couramment retrouvée dans les nuages diffus.

4. Les abondances de ^{12}CO et HCO^+ augmentent toutes deux de façon presque linéaire en fonction de la densité de colonne de H_2, i.e., $[^{12}CO] \propto N(H_2)^{0.64}$ jusqu'à $N(H_2) =$

$6.4 \times 10^{21}\,\mathrm{cm}^{-2}$, puis décroît. Cela peut résulter du gèle des molécules à l'état gazeux sur les grains de poussière froids. Par ailleurs, $[\mathrm{HCO}^+] \propto N(\mathrm{H_2})^{0.67}$ jusqu'à $N(\mathrm{H_2}) = 2.7 \times 10^{21}\,\mathrm{cm}^{-2}$, puis reste approximativement constante.

5. En plus de confirmer la rotation globale de la Tête de Cheval autour de son axe, nous avons detecté la signature d'un phénomène d'accrétion de la couche externe avant du nuage par la couche interne plus dense.

6. Tandis que T_{kin} estimée dans l'enveloppe varie de 25 à 40 K, la température cinétique dans la couche interne atteint ~ 15 K dans le cœur dense le plus à l'ouest. La densité volumique estimée dans la couche interne est de $\sim 3 \times 10^4\,\mathrm{cm}^{-3}$ en direction des filaments, et augmente d'un facteur 10 vers les cœurs denses. La pression thermale de la couche externe présente un gradient décroissant d'est en ouest.

Dans ces deux études, le bruit de calibration a été négligé lors de la tâche d'estimation, alors que Einig *et al.* 2023 a évalué les incertitudes de calibration de l'ordre de 5% et 10% pour respectivement les fenêtres atmosphériques à 3 et 1 mm. Pour confirmer la validité de ce choix, j'ai examiné de façon exhaustive les différentes façons dont ce bruit de calibration peut être traité. Lorsque son écart-type demeure inférieur à 10%, omettre délibérément le bruit de calibration mène à 1) un estimateur non biaisé des paramètres physiques d'intérêt, et 2) une perte de précision allant de 10 à 30% sur leur estimation, ce qui est considéré acceptable pour le problème astrophysique considéré. Ces derniers résultats confirment notre décision de négliger ce bruit multiplicatif de notre modèle de mesures. Pour de futurs travaux, nous recommandons toutefois de calculer la BCR *modifiée*, qui fournit des intervalles de confiance prenant en compte de cette simplification. Enfin, la BCR fût par ailleurs utilisée pour évaluer le gain sur la précision d'estimation si toutes les raies moléculaires étaient bien résolues en vitesse (i.e., à 0.1 km/s).

6.4 Perspectives

J'aborde désormais trois perspectives pour ces travaux de thèse.

Analyse chimique de la nébuleuse de la Tête de Cheval — Jusqu'à présent, je n'ai considéré que 7 raies moléculaires du jeu de données ORION-B, alors qu'une dizaine d'espèces chimiques supplémentaires ont été détectées, comme CS, SO, CCH, C_3H_2, ou encore N_2H^+. Ces derniers traceurs moléculaires sont difficilement exploitables pour contraindre les paramètres physiques (i.e., T_{kin} et $n_{\mathrm{H_2}}$), car leur chimie (et ainsi, leurs abondances relatives) est difficile à modéliser simplement. Néanmoins, il est désormais possible de fixer les paramètres physiques pour estimer les densitées de colonne de ces espèces supplémentaires. L'information contenue dans ces raies de faible émission pourrait ainsi contribuer significativement à approfondir notre compréhension de la chimie de ces molécules simples dans le milieu interstellaire.

D'un millier à un million de pixels — Jusqu'à présent, les études approfondies des conditions physico-chimiques des nuages ont principalement été menées sur quelques pixels

seulement, et cela parce qu'il est difficile d'obtenir des estimations fiables sans une supervision humaine considérable. L'analyse de la nébuleuse de la Tête de Cheval fût menée sur 1500 pixels environ, tandis que le jeu de données ORION-B en contient près d'un million. Au-delà du temps de calcul nécessaire, la principale difficulté pour généraliser la méthode d'analyse développée provient de l'hypothèse qu'une seule composante de vitesse se trouve le long de chaque ligne de visée. La région correspondant à la nébuleuse de la Tête de Cheval fût précisemment choisie parce que cette hypothèse y est respectée. Cela n'est pas le cas pour le reste du nuage d'Orion B, où la plupart des lignes de visée montrent de *multiples* composantes de vitesse. Ajuster des composantes dont les supports sont disjoints en vitesse est trivial. Il reste ensuite à *segmenter* le nuage d'Orion B en régions à une, deux, ... composantes de vitesse. Ce problème peut être reformulé comme un problème de *sélection de modèles* basée sur la *longueur de description minimale* (*Minimum Description Length*; MDL en anglais). Il s'agit de choisir la meilleure description des données du point de vue de la théorie de l'information. La difficulté survient dans les régions transitoires où deux composantes de vitesse distinctes fusionnent progressivement en une composante, de sorte qu'une seule raie moléculaire (potentiellement asymétrique) ne soit observée. Résoudre ce problème nécessite de caractériser ce processus de fusion. Bien que nous ayons jusqu'à présent adopté une approche pixel par pixel pour estimer la distribution spatiale des paramètres physico-chimiques, l'exploitation de l'information spatiale pourrait s'avérer utile pour identifier ce mécanisme de fusion.

D'Orion B aux autres nuages géants moléculaires de la Voie Lactée — La démarche du programme ORION-B consiste à calibrer nos algorithmes d'analyse sur un nuage géant moléculaire proche, à l'aide de techniques peu supervisées, avant de les appliquer à des nuages de notre galaxie plus éloignés. Déployer la méthode proposée dans cette thèse sur des nuages plus distants nécessitera la prise en compte de dégradation de la résolution spatiale. Bien que Zakardjian *et al.* 2025 aborde ce problème à une toute autre échelle, car traite des données extragalactiques, cette dernière étude souligne que la robustesse des résultats dépend du choix d'une représentation simplifiée des conditions physiques mélangées au sein du faisceau donné. Cela signifie qu'il serait illusoire de rechercher une représentation spatialement cohérente au sein du faisceau. Nous devrons donc trouver une autre façon de caractériser statistiquement la carte d'observations. Par exemple, nous pourrions essayer de segmenter le cube de données selon différents mélanges de régimes physiques d'intérêt (e.g., gaz diffus et translucide, filaments, cœurs denses, etc.). Un tel traitement ouvrirait la voie vers une nouvelle description statistique des données.

Contributions

I have written as a first author an astrophysical article (Ségal *et al.* 2024) and a paper in the proceedings of the national GRETSI conference (Ségal *et al.* 2025). I have also co-developped another astrophysical article (Roueff *et al.* 2024). Besides, I have co-authored with minor contributions several articles from the ORION-B consortium.

I developed a public software called TOAST[1] to allow the astrophysicists to test or adapt the analysis methodology proposed in this thesis. This codes infers, among other parameters, the kinetic temperature and volume density from the simultaneous analysis of multiple molecular lines toward heterogeneous line of sight, e.g., dense core. The code is developed in PYTHON and provides tutorials to analyze the Horsehead nebula dataset studied in Roueff *et al.* 2024 and Ségal *et al.* 2024.

— International journals —

- Ségal, L. *et al.* (2024). "Toward a robust physical and chemical characterization of hetero- geneous lines of sight: The case of the Horsehead nebula". In: Astronomy & Astrophysics 692, A160. doi: 10.1051/0004-6361/202451567
- Roueff, A. *et al.* (2024). "Bias versus variance when fitting multi-species molecular lines with a non-LTE radiative transfer model - Application to the estimation of the gas temperature and volume density". In: Astronomy & Astrophysics 686, A255. doi: 10.1051/0004-6361/202449148

— National conferences with proceedings —

- Ségal, L. *et al.* (2025). "De la possibilité d'ignorer un bruit de calibration lors de la construction d'un estimateur du maximum de vraisemblance en radio-astronomie", GRETSI, 2025. Contributed talk.

— Other publications in international journals as co-author —

- Palud, P. *et al.* (2025). "BEETROOTS: Spatially regularized Bayesian inference of physical parameter maps. Application to Orion". In: Astronomy & Astrophysics 698, A311. doi: 10.1051/0004-6361/202554266

[1] mulTilayer clOud for simultAneouS analysis of mulTi-molecular lines, available at https://github.com/LeontineSegal/TOAST.git

- Zakardjian, A. *et al.* (2025). "Understanding spatially unresolved measurements of molecular line emission". In: Astronomy & Astrophysics 696, A165. doi: 10.1051/0004-6361/202451104

- Einig, L. *et al.* (2024). "Quantifying the informativity of emission lines to infer physical conditions in giant molecular clouds - I. Application to model predictions". In: Astronomy & Astrophysics 691, A109. doi: 10.1051/0004-6361/202451588

- Santa-Maria, M. G. *et al.* (2023). "HCN emission from translucent gas and UV-illuminated cloud edges revealed by wide-field IRAM 30 m maps of the Orion B GMC - Revisiting its role as a tracer of the dense gas reservoir for star formation". In: Astronomy & Astrophysics 679, A4. doi: 10.1051/0004-6361/202346598

- Palud, P. *et al.* (2023). "Neural network-based emulation of interstellar medium models". In: Astronomy & Astrophysics 678, A198. doi: 10.1051/0004-6361/202347074

- Zakardjian, A. *et al.* (2025). "Estimating the dense gas mass of molecular clouds using spatially unresolved 3 mm line observations". Astronomy & Astrophysics (forthcoming article) doi: 10.1051/0004-6361/202555123

- Bešlić, I. *et al.* (2025). "The first estimation of the ionization fraction in dense and translucent molecular gas across Orion B". Astronomy & Astrophysics, 702, A205. doi: 10.1051/0004-6361/202553706

— Conference talks without proceedings —

- Colloque du Programme Physique Chimie du Milieu Interstellaire (PCMI), Bordeaux (France), October 2024. "Towards robust estimations of molecular cloud physical conditions from multi-species lines: The case of the Horsehead nebula". Contributed talk.

- 53rd Young European Radio Astronomers Conference (YERAC), Madrid (Spain), September 2024. "A statistical approach to infer molecular cloud conditions from multi-molecular lines analysis of Horsehead nebula observations". Brief talk.

— Seminar presentations and workshops —

- Atelier sur les méthodes IA et ML en astrophysique pour l'analyse du milieu interstellaire, Bordeaux (France), October 2024. Contributed talk.

- GDR IASIS, État des lieux des méthodes d'inversion en astronomie et challenges à venir, Gif-sur-Yvette (France), June 2024. "Analyse multi-espèces du milieu interstellaire fondée sur les observations millimétriques de la nébuleuse de la Tête de Cheval". Contributed talk.

- ORION-B/DAOISM workshops (France). Every 9 months, the consortium's members gather to share our progress. LERMA, Paris, June 2022, January 2023; CRIStAL, Lille, July 2023; IRAM, Grenoble, January 2024; IRAP, Toulouse, July 2024; IM2NP, Toulon, January 2025; LUX, Paris, April 2025.

Abstract & Résumé

Abstract — Stars form in giant molecular clouds (GMCs) that have extreme combined physical conditions impossible to reproduce on Earth: a temperature between 10 and 100 K and a volume density between 100 and 10^6 cm^{-3}. Remote observation is therefore the only way to determine these physical conditions. However, under these typical conditions, molecular hydrogen, which is by far the most abundant molecular species, is mostly invisible. The properties of the gas are indirectly probed by the lines emitted by minor molecular tracers such as carbon monoxide (CO). The estimates derived from multi-species analyses necessarily rely on a priori constraints (e.g., cloud geometry) which significantly bias the results when incorrect. It is therefore necessary to develop a methodology for reliably fitting multi-species observations in order to deepen our understanding of star formation.

The proposed statistical treatment is based on estimation theory and performance bounds. In particular, the Cramér-Rao bound (CRB), which provides a lower bound of the precision for any unbiased estimator, is used here 1) to verify the efficiency of the proposed maximum likelihood estimator, 2) to provide error bars for the actual parameter estimates, 3) to quantify the loss of precision when the presence of calibration errors affecting the data is deliberately neglected in order to simplify the adjustment procedure, and 4) to study the impact of spectral resolution on estimation precision and guide requirements for future millimeter-wavelength spectrometers.

Under the guidance of astrophysicists from the ORION-B/DAOISM consortium, I applied these statistical tools to the analysis of a radiative transfer multilayer cloud model in order to study heterogeneous properties (column density, volume density, kinetic temperature, thermal pressure, and the gas kinematics) along any line of sight toward filaments and *dense cores*. Applied to low-J transitions of CO and HCO^+ isotopologues in the *Horsehead Nebula*, the main findings are: 1) a minimal heterogeneous model (only three layers) is sufficient in this case, as it fits the seven selected molecular lines over the considered field of view, 2) CO is depleted from the gas phase to dust ice for lines of sight toward the two embedded dense cores, and 3) the data show a signature of accretion motions from the foreground translucent layer toward the denser internal gas. This proof of concept, combined with the codes shared with the community, paves the way for processing larger fields of view.

Keywords — Statistical data processing, Estimation theory, Performance Bound, Inter-Stellar Medium (ISM), Radiative transfer and chemical modeling, Initial conditions for star formation.

Résumé — Les étoiles se forment dans des nuages moléculaires géants qui présentent des conditions physiques extrêmes impossibles à reproduire sur Terre : une température comprise entre 10 et 100 K et une densité volumique comprise entre 100 et 10^6 cm^{-3}. La télédétection est donc le seul moyen de déterminer ces conditions physiques. Cependant, dans ces conditions typiques, l'hydrogène moléculaire, qui est de loin l'espèce moléculaire la plus abondante, est en grande partie invisible. Les propriétés du gaz sont indirectement sondées avec les raies émises par des traceurs moléculaires mineurs tels que le monoxyde de carbone (CO). Les estimations obtenues par des analyses multi-espèces reposent nécessairement sur des contraintes a priori (par exemple, la géométrie du nuage) qui biaisent considérablement les résultats lorsqu'elles sont incorrectes. Il est donc nécessaire de développer une méthodologie permettant d'ajuster de manière fiable les observations multi-espèces afin d'approfondir notre compréhension de la formation des étoiles.

Le traitement statistique proposé repose sur la théorie de l'estimation et des bornes de performance. En particulier, la borne de Cramér-Rao (CRB), qui fournit une borne inférieure de précision pour tout estimateur non biaisé, est utilisée ici pour 1) vérifier l'efficacité de l'estimateur du maximum de vraisemblance proposé, 2) fournir des barres d'erreur sur les estimations des paramètres physiques sur données réelles, 3) quantifier la perte de précision lorsque la présence d'erreurs d'étalonnage affectant les données est délibérément négligée afin de simplifier la procédure d'ajustement du modèle aux mesures, et enfin 4) étudier l'impact de la résolution spectrale sur la précision de l'estimation et orienter la conception des futurs spectromètres aux longueurs d'onde millimétriques.

Guidée par les astrophysiciens du consortium ORION-B/DAOISM, j'ai appliqué ces outils statistiques pour analyser un modèle de transfert radiatif de nuage multi-couches afin d'étudier les propriétés hétérogènes (densité de colonne, densité volumique, température cinétique, pression thermique et cinématique du gaz) le long de lignes de visée en direction de filaments et de *cœurs denses*. Appliqué aux transitions de faible J des isotopologues de CO et HCO^+ dans la *nébuleuse de la Tête de Cheval*, les résultats principaux sont les suivants : 1) un modèle hétérogène minimal (seulement trois couches) est suffisant dans ce cas, car s'ajuste correctement aux sept raies moléculaires sélectionnées sur le champ de vue considéré, 2) le CO en phase gazeuse s'appauvrit pour les lignes de visée en direction des deux cœurs denses de cette source, suggérant que l'espèce se retrouve piégée sur les manteaux de glace des poussières, et 3) les données suggèrent un mouvement d'accrétion de la couche translucide au premier plan vers le gaz interne plus dense. Cette preuve de concept, associée aux codes partagés avec la communauté, ouvre la voie au traitement de grands champs de vue.

Mot clés — Traitement statistique de données, Théorie de l'estimation, Borne de performance, Milieu InterStellaire (MIS), Modèle de transfert radiatif et modélisation chimique, Conditions initiales de la formation des étoiles.