École Doctorale Économie, Management, Mathématiques et Physique (EM2P)

THÈSE DE DOCTORAT

de

l'Université de Cergy-Pontoise

Spécialité: Mathématiques

présentée et soutenue publiquement par

José G. GÓMEZ

le 13 novembre 2017.

Théorèmes limites pour des fonctionnelles de clusters d'extrêmes et applications

Directeur de thèse : **Paul DOUKHAN**

Jury de soutenance :

Eva Löcherbach,	Professeur, AGM, Université de Cergy Pontoise	Présidente du jury
Joseph Rynkiewicz,	Professeur, SAMM, Université Paris I	Examinateur
Johan Segers,	Professeur, ISBA, Université catholique de Louvain	Rapporteur
Patrice Bertail,	Professeur, MODAL'X, Université Paris Nanterre	Rapporteur
Paul Doukhan,	Professeur, AGM, Université de Cergy Pontoise	Directeur

Laboratoire AGM (UMR 8088)
UFR Sciences et Techniques-UCP
2 Avenue Adolphe-Chauvin
95 302 Cergy-Pontoise

École doctorale EM2P
33 Boulevard du Port
Université de Cergy-Pontoise
95 011 Cergy-Pontoise

Remerciements

Mes remerciements sincères à mon directeur de thèse, Paul Doukhan, pour m'avoir présenté un domaine de mathématiques très riche, et pour m'avoir suggéré un sujet de thèse pertinent. Paul a été mon guide tout au long de mon parcours de recherche. Je le remercie aussi pour sa patience pendant ces années de thèse et pour tous ses conseils. Je porte à son égard une grande estime et un grand respect.

Je remercie Johan Segers et Patrice Bertail pour avoir accepté de rapporter ma thèse et de participer à mon jury de thèse. Je les remercie pour leur lecture minutieuse et leurs remarques. Je souhaite spécialement exprimer ma gratitude envers le Professeur Segers pour ses conseils très utiles et pertinents lors des moments difficiles de ma thèse. Je remercie les autres membres du jury : Eva Löcherbach et Joseph Rynkiewicz, pour avoir accepté d'être examinateurs de ma thèse.

Je remercie également le Projet LABEX MME-DII (ANR11-LBX-0023-01) d'avoir accepté de financer ma thèse.

Je remercie les membres du labo AGM : Emmanuel Hebey, Vladimir Georgescu, Linda Isone, Jennifer Denis et Thomas Ballesteros pour leur efficacité. Je tiens à remercier également mes camarades de bureau : Lorenzo Pittau et Xiaoyin Li. Merci pour votre convivialité et pour votre humour, en particulier Xiaoyin : tu m'as beaucoup manqué ces dernières années de thèse ! Je remercie aussi Willian Kengne (auparavant au labo AGM et maintenant au labo THEMA) pour l'organisation (avec mon directeur) des groupes de travail de statistiques au laboratoire AGM, qui ont été une bonne source d'inspiration. Je le remercie aussi pour sa gentillesse et son attention.

Je tiens à remercier vivement le Professeur Didier Dacunha-Castelle pour toutes les réunions que nous avons eu afin de parler des statistiques et de leurs applications. Il m'a appris à réfléchir et à attaquer les problèmes d'un point de vue plus appliqué. Je le remercie chaleureusement pour son esprit très ouvert et très collaboratif avec les étudiants d'Amérique latine et d'ailleurs.

Je remercie le Professeur José R. León pour avoir appuyé mon doctorat en France et pour ses bons conseils.

Je remercie Lisandro Fermín de m'avoir proposé un stage au CIMFAV (Centro de Investigación Matemática de Fenómenos Aleatorios - Valparaiso) au Chili pour discuter des problèmes d'éoliennes dans cette région. Je remercie aussi Karine Bertin et Cristian Meza d'avoir financé une partie de ce stage. Je remercie tous les membres du CIMFAV pour l'ambiance sympathique et chaleureuse pendant mon séjour. Je remercie avec une grande joie Perrine Soret, qui était également en stage au CIMFAV, d'avoir partagé son appartement et son bon esprit pendant la moitié de mon séjour.

Je remercie les membres du Département de Mathématiques de l'Université Paris 13 (LAGA) d'avoir accepté ma candidature ATER. Je remercie spécialement Yueyun Hu, Philippe Souplet, Thomas Duyckaerts, Bérangère Delourme, Didier Gamblin et Isabelle Barbotin de m'avoir si bien accueilli au sein de leur laboratoire, dans une ambiance conviviale de travail.

Je remercie Ieva Grublyté et Jacek Leśkow pour leurs relectures et leurs commentaires concernant mon premier article.

Je remercie mes amis : Ana Karina Fermín et Erwan Le Pennec pour toute leur aide pendant mon installation et mon temps d'adaptation en France, également pour leur amitié et leurs conseils ; Mariana Garcia et Ignacio Rodriguez pour la relecture de mes textes rédigés en anglais et leurs corrections ; Alessandra Fariñas et Argenis Mendez pour les petites discutions de maths ; Diana Sandoval, Luis Paredes, Fabián Agudelo, Luisa F. Murcia, Paula Montaño, Rafael Gómez y Jairo Cugliari pour leur bon esprit ; et spécialement Mélina Gallopin d'être toujours là pour m'encourager dans le monde universitaire français, pour partager des informations pertinentes qui m'ont beaucoup aidé dans la vie quotidienne et pour sa grande amitié : *Gracias por ayudarme a empujar al burrito* Mélina !

Je remercie mes camarades et amis de l'Université Centrale du Venezuela pour le bel esprit transmis envers leurs étudiants, malgré les difficultés politiques et socio - économiques du pays.

Finalement, de la manière plus simple mais la plus profonde, je remercie ma famille au Venezuela pour tout son amour inconditionnel ; et très tendrement ma famille en France : Lysiane Courteille et mon petit Bruno, qui réchauffent d'enthousiasme mes jours d'hiver et rafraîchissent de calme mes jours d'été.

Résumé

Cette thèse traite principalement des théorèmes limites pour les processus empiriques de fonctionnelles de clusters d'extrêmes de séquences et champs aléatoires faiblement dépendants.

Des théorèmes limites pour les processus empiriques de fonctionnelles de clusters d'extrême de séries temporelles stationnaires sont donnés par Drees & Rootzén [2010] sous des conditions de régularité absolue (ou "$\beta-$mélange"). Cependant, ces conditions de dépendance de type mélange sont très restrictives : elles sont particulièrement adaptées aux modèles dans la finance et dans l'histoire, et elles sont de plus compliquées à vérifier. Généralement, pour d'autres modèles fréquemment rencontré dans les domaines applicatifs, les conditions de mélange ne sont pas satisfaites. En revanche, les conditions de dépendance faible, selon Doukhan & Louhichi [1999] et Dedecker & Prieur [2004a], sont des conditions qui généralisent les notions de mélange et d'association. Elles sont plus simple à vérifier et peuvent être satisfaites pour de nombreux modèles. Plus précisément, sous des conditions faibles, tous les processus causaux ou non causaux sont faiblement dépendants: les processus Gaussien, associés, linéaires, ARCH(∞), bilinéaires et notamment Volterra entrent dans cette liste. À partir de ces conditions favorables, nous étendons certains des théorèmes limites de Drees & Rootzén [2010] à processus faiblement dépendants. En outre, comme application des théorèmes précédents, nous montrons la convergence en loi de l'estimateur de l'extremogramme de Davis & Mikosch [2009] et l'estimateur fonctionnel de l'indice extrémal de Drees [2011] sous dépendance faible.

Nous démontrons un théorème de la valeur extrême pour les champs aléatoires stationnaires faiblement dépendants et nous proposons, sous les mêmes conditions, un critère du domaine d'attraction d'une loi d'extrêmes.

Le document se conclue sur des théorèmes limites pour les processus empiriques de fonctionnelles de clusters d'extrêmes de champs aléatoires stationnaires faiblement dépendants, et met en évidence la convergence en loi de l'estimateur d'un extremogramme de processus spatio-temporels stationnaires faiblement dépendants en tant qu'application.

Mots-clefs : dépendance faible, fonctionnelles de clusters d'extrêmes, théorème de la limite centrale, méthode de Lindeberg, indice extrémal, extremogramme.

Limit theorems for functionals of clusters of extremes and applications.

Abstract

This thesis deals mainly with limit theorems for empirical processes of extreme cluster functionals of weakly dependent random fields and sequences.

Limit theorems for empirical processes of extreme cluster functionals of stationary time series are given by Drees & Rootzén [2010] under absolute regularity (or $\beta-$mixing) conditions. However, these dependence conditions of mixing type are very restrictive: on the one hand, they are best suited for models in finance and history, and on the other hand, they are difficult to verify. Generally, for other models common in applications, the mixing conditions are not satisfied. In contrast, weak dependence conditions, as defined by Doukhan & Louhichi [1999] and Dedecker & Prieur [2004a], are dependence conditions which generalises the notions of mixing and association. These are easier to verify and applicable to a wide list of models. More precisely, under weak conditions, all the causal or non-causal processes are weakly dependent: Gaussian, associated, linear, ARCH(∞), bilinear and Volterra processes are some included in this list. Under these conveniences, we expand some of the limit theorems of Drees & Rootzén [2010] to weakly dependent processes. These latter results are used in order to show the convergence in distribution of the extremogram estimator of Davis & Mikosch [2009] and the functional estimator of the extremal index introduced by Drees [2011] under weak dependence.

We prove an extreme value theorem for weakly dependent stationary random fields and we propose, under the same conditions, a domain of attraction criteria of a law of extremes.

The document ends with limit theorems for the empirical process of extreme cluster functionals of stationary weakly dependent random fields, deriving also the convergence in distribution of the estimator of an extremogram for stationary weakly dependent space-time processes.

Keywords: weak dependence, functionals of clusters of extremes, central limit theorem, Lindeberg method, extremal index, extremogram.

Contents

Théorèmes limites pour des fonctionnelles de clusters d'extrêmes et applications

José G. GÓMEZ

13 novembre 2017.

Chapter 1

Introduction générale et résultats principaux

Cette thèse porte principalement sur des théorèmes limites pour les processus empiriques de fonctionnelles de clusters d'extrêmes de processus et champs aléatoires faiblement dépendants, au sens de Doukhan & Louhichi [1999]. Nous utilisons ces théorèmes pour valider les propriétés asymptotiques de l'estimateur de l'extremogramme de Davis & Mikosch [2009], ainsi que celles de l'estimateur fonctionnel de l'indice extrémal de Drees [2011] sous une hypothèse de dépendance faible. Nous développons également une théorie des valeurs extrêmes pour des champs aléatoires stationnaires et faiblement dépendants; nous proposons un critère permettant d'identifier le domaine d'attraction pour les lois d'extrêmes de tels modèles. Cela nous met en position de développer des théorèmes limites pour les processus empiriques de fonctionnelles de clusters d'extrêmes de champs aléatoires stationnaires faiblement dépendants. Nous montrons la convergence en loi de l'estimateur d'un extrémogramme de processus spatio-temporels stationnaires faiblement dépendants à titre d'application à l'étude d'éoliennes au Chili. Ce document se compose ainsi de cinq chapitres détaillés comme suit:

- Le chapitre 1 est une introduction générale rédigée en français et il inclut les résultats principaux.
- Le chapitre 2 a fait l'objet d'une publication acceptée: "Dependent Lindeberg CLT for the fidis of empirical processes of cluster functionals", *Statistics: A Journal of Theoretical and Applied Statistics* (2017).
- Le chapitre 3 a fait l'objet d'un article en collaboration avec Paul Doukhan: "Limit theorems for empirical processes of cluster functionals without mixing".

Actuellement en révision pour *Extremes: Statistical Theory and Applications in Science, Engineering and Economics*.

- Le chapitre 4 fait l'objet de l'article : "On extreme values in stationary weakly dependent random fields".
- Le chapitre 5 fait partie d'un article en cours de rédaction intitulé : "Extreme cluster functionals in stationary weakly dependent random fields".

Soient $X_1, \ldots, X_n$ des observations aux temps $1, \ldots, n$ d'une série temporelle stationnaire réelle $X = (X_t)_{t \in \mathbb{Z}}$, définie sur $(\Omega, \mathcal{A}, \mathbb{P})$. Notons F fonction de répartition (f.r.) commune à ses marginales. Soient $(u_k)_{k \in \mathbb{N}}$ des seuils et $(a_k > 0)_{k \in \mathbb{N}}$ des constantes de normalisation. Définissons la fonction de répartition de queue conditionnelle

$$G_n(x) := \frac{1 - F(u_n + a_n x)}{1 - F(u_n)}. \tag{1.1}$$

On suppose que la suite (G_n) converge vers une f.r. non dégénérée G, pour tout $x \geqslant 0$. Nous estimons la f.r. G par la fonction de répartition empirique de queue:

$$\widehat{G}_n(x) := \frac{1}{nF(u_n)} \sum_{i=1}^{n} \mathbb{1}_{\{X_i > u_n + a_n x\}}, \quad x \geqslant 0 \tag{1.2}$$

et nous construisons des régions de confiance, en utilisant le processus empirique de queue:

$$\bar{Z}_n(x) = \sqrt{nF(u_n)} \left(\widehat{G}_n(x) - G_n(x) \right), \quad x \geqslant 0. \tag{1.3}$$

Sous les hypothèses de $\beta-$mélange, Rootzén [1995] montre que le processus empirique $(\bar{Z}_n(x))_{x \geqslant 0}$ converge dans l'espace des fonctions càdlàg, $D([0, x_G))$, vers un processus Gaussien $(\bar{Z}(x))_{x \geqslant 0}$, où $x_G = \sup\{x : G(x) < 1\}$. Drees [2000, 2002] améliorent son résultat. Rootzén [2009] donne aussi des résultats sur ce processus empirique en considérant des seuils aléatoires et en affaiblissant les hyphothèses de dépendance sur les séquences aléatoires considérant une hypothèse de $\alpha-$mélange.

Remarquons que la fonction $\widehat{G}$ peut être réécrite, pour n assez grand, comme:

$$\widehat{G}(x) = \frac{1}{nF(u_n)} \sum_{j=1}^{m_n} f_x \left((X_{n,i})_{(j-1)r_n < i \leqslant j r_n} \right), \tag{1.4}$$

où $m_n = \lceil n/r_n \rceil := \max\{k \in \mathbb{N} : j \leqslant n/r_n\}$, $X_{n,i} = (X_i - u_n)_+/a_n$ et f_x est une fonction de vecteurs de longueur arbitraire définie par

$$f_x(x_1, \dots, x_r) = \sum_{i=1}^{r} \mathbb{1}_{\{x_i > x\}}. \tag{1.5}$$

Dans tous les travaux de Drees et Rootzén mentionnés et dans leurs références, cette fonction f_x est utilisée implicitement. En fait, f_x est l'exemple le plus classique d'une fonctionnelle de clusters "d'extrêmes". Globalement, pour un ensemble E et un "failure set" $A \subset E$, une fonctionnelle de clusters f est une fonction réelle de blocs finis $B \in \bigcup_{r=1}^{\infty} E^r$ telle que f prend la valeur 0 si aucun élément de B n'appartient à A et $f(B) = f(B^c)$, où B^c est le plus grand sous-bloc de B qui commence et se termine par un élément de A.

Les premiers travaux sur les fonctionnelles de clusters d'extrêmes sont [Rootzén et al., 1998][(1)], [Yun, 2000] et [Segers, 2003], qui développent des résultats sur les distributions de sommes de tableaux de queues ($\sum_{i=1}^{n} \phi(X_{n,i})$, avec $\phi(0) = 0$), les distributions de fonctionnelles de cluster d'extrêmes, pour des chaînes de Markov d'ordre d et les distributions de fonctionnelles de clusters d'extrêmes de suites stationnaires, respectivement. Une définition générale de fonctionnelle de clusters est donnée dans [Drees & Rootzén, 2010], qui prouve des théorèmes limites pour le processus empiriques de fonctionnelles de clusters. Soit $\mathcal{F}$ une classe de fonctionnelles de clusters. Le processus empirique $(Z_n(f))_{f \in \mathcal{F}}$ de fonctionnelles de clusters en $\mathcal{F}$ est défini par

$$Z_n(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} \left(f(Y_{n,j}) - \mathbb{E} f(Y_{n,j}) \right), \quad f \in \mathcal{F} \tag{1.6}$$

où $v_n = \mathbb{P}(X_{n,1} \in A)$ (A est le "failure set"), $Y_{n,j} := (X_{n,i})_{(j-1)r_n < i \leqslant jr_n}$ avec $1 \leqslant j \leqslant m_n$ et $\{(X_{n,i})_{1 \leqslant i \leqslant n}\}_{n \in \mathbb{N}}$ est un tableau triangulaire stationnaire par lignes tel que $\mathbb{P}(X_{n,1} \in \cdot \,| X_{n,1} \in A)$ converge vers une distribution non dégénérée.
Pour $A = \mathbb{R}^d \setminus \{0\}$ et sous les hyphothèses:

(B*) $(l_n)_{n \geqslant 1}$ et $(r_n)_{n \geqslant 1}$ sont deux suites infinies entières telles que $l_n \ll r_n \ll v_n^{-1} \ll n$, avec $l_n \longrightarrow \infty$ lorsque $n \to \infty$;

(C.1) $\mathbb{E}\left((f(Y_{n,1}) - \mathbb{E} f(Y_{n,1}))^2 \mathbb{1}_{\{|f(Y_{n,1}) - \mathbb{E} f(Y_{n,1})| > \epsilon \sqrt{nv_n}\}} \right) = o(r_n v_n)$, pour tout $\epsilon > 0$ et tout $f \in \mathcal{F}$;

(C.2) $(r_n v_n)^{-1} \mathrm{Cov}\left(f(Y_{n,1}), g(Y_{n,1}) \right) \longrightarrow c(f,g)$, pour tout $f, g \in \mathcal{F}$;

[1]Sans utiliser la terminologie "fonctionnelle de cluster"

(C.3) $\Delta_n(f) := f((X_{n,i})_{1\leqslant i\leqslant r_n}) - f((X_{n,i})_{1\leqslant i\leqslant r_n-l_n})$ est tel que

$$\begin{aligned}\mathbb{E}|\Delta_n(f)-\mathbb{E}\Delta_n(f)|^2\mathbb{I}_{\{|\Delta_n(f)-\mathbb{E}\Delta_n(f)|\leq\sqrt{nv_n}\}} &= o(r_n v_n),\\ \mathbb{P}\left(|\Delta_n(f)-\mathbb{E}\Delta_n(f)|>\sqrt{nv_n}\right) &= o(r_n/n),\end{aligned}$$

pour tout $f \in \mathcal{F}$;

(D) le tableau triangulaire est $\beta-$mélangeant, et tel que $\beta_n(l_n) = o(r_n/n)$;

(T) des conditions pour obtenir la tension ou l'equicontinuité asymptotique;

Drees & Rootzén [2010] montrent que la suite des processus empiriques Z_n converge vers un processus Gaussien centré Z, de fonction de covariance c.

Cette thèse est motivée principalement par ce résultat pour plusieurs raisons.
- D'abord, une grande partie des estimateurs utilisés dans l'analyse statistique des extrêmes peuvent être écrits en termes de fonctionnelles de clusters; les théorèmes limites des processus empiriques de fonctionnelles de clusters s'appliquent. L'estimateur de l'extrémogramme présenté dans le Chapitre 2 et l'estimateur fonctionnel d'indice extrémal présenté dans le Chapitre 3 sont des exemples d'application.
- La seconde raison est que la littérature existant sur les théorèmes limites de fonctionnelles de clusters concerne uniquement les cas mélangeants. Ces conditions de mélange sont très restrictives pour certains modèles. *Un exemple classique de processus non-mélangeant est le processus $AR(1)$, solution de la récursion:*

$$X_i = \frac{1}{b}(X_{i-1}+\xi_i), \quad i \in \mathbb{Z} \tag{1.7}$$

où $b \geqslant 2$ est une valeur entière fixée et $(\xi_i)_{i\in\mathbb{Z}}$ est une séquence de variables aléatoires indépendantes et distributées uniformément sur l'ensemble $\{0,1,\ldots,b-1\}$ ([Andrews, 1984] pour le cas $b=2$ et [Ango Nze & Doukhan, 2004] pour le cas $b>2$). Cela justifie le développement de théorèmes limites de fonctionnelles de clusters sous des conditions de dépendance moins restrictives que les conditions de mélange et motive également leur extensions dans les champs aléatoires ainsi que dans les processus spatio-temporels.

Nous utilisons des conditions de "dépendance faible" introduites par Doukhan & Louhichi [1999], généralisant les notions de mélange et d'association, qui sont plus flexibles et prennent en compte des classes plus grandes de modèles de séries temporelles, de champs aléatoires, etc.

Soit $h : E^r \subseteq (\mathbb{R}^d)^r \longrightarrow \mathbb{R}$ une fonction, avec $r \in \mathbb{N}$. On note

$$\text{Lip}(h) := \sup_{(x_1,\dots,x_r)\neq(y_1,\dots,y_r)\in E^r} \frac{|h(x_1,\dots,x_r) - h(y_1,\dots,y_r)|}{\|x_1 - y_1\| + \cdots + \|x_r - y_r\|}.$$

Un tableau triangulaire de variables aléatoires à valeurs dans E, et stationnaires par lignes $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ est (ϵ, ψ)**-faiblement dépendant** (au sens de Doukhan & Louhichi [1999]) s'il existe une fonction $\psi : \mathbb{N}^2 \times (\mathbb{R}^+)^2 \longrightarrow \mathbb{R}^+$, une suite croissante d'entiers positifs $(l_n)_{n\in\mathbb{N}}$ avec $l_n \ll n$, et une suite positive $(\epsilon_n(l))_{1\leqslant l\leqslant n, n\in\mathbb{N}}$ telle que $\epsilon_n(l_n) \underset{n\to\infty}{\longrightarrow} 0$ et

$$\left|Cov\left(h_1(X_{n,i_1},\dots,X_{n,i_u}), h_2(X_{n,j_1},\dots,X_{n,j_v})\right)\right| \leqslant \psi(u, v, Lip(h_1), Lip(h_2)) \cdot \epsilon_n(l), \quad (1.8)$$

pour tous $(u,v) \in \mathbb{N}\times\mathbb{N}$, $(i_1,\dots,i_u) \in \mathbb{N}^u$ et $(j_1,\dots,j_v) \in \mathbb{N}^v$ avec $i_1 < \cdots < i_u < i_u + l \leqslant j_1 < \cdots < j_v \leqslant n$, et pour tout couple de fonctions $(h_1,h_2) \in \Lambda^u(E) \times \Lambda^v(E)$, où $\Lambda^s(E) := \{h : E^s \longrightarrow \mathbb{R}$ Lipschitz avec $\|h\|_\infty \leqslant 1$ et $\text{Lip}(h) < \infty\}$.

Des exemples intéressants sont les fonctions $\psi(u,v,s,t) = vt$ (processus linéaires causaux), $\psi(u,v,s,t) = us + vt$ (processus linéaires non causaux), $\psi(u,v,s,t) = uvst$ (processus associés ou gaussiens), et $\psi(u,v,s,t) = us + vt + uvst$. Ces types de dépendance faible seront notées respectivement à l'aide des coefficients θ_n, η_n, κ_n et λ_n, au lieu de ϵ_n.

Remarque : Pour les tableaux triangulaires, notez que nous avons défini un coefficient de dépendance ϵ_n pour chaque $n-$ieme ligne fixée, parce que les fonctions qui apparaissent dans ces tableaux sont non uniformément Lipschitz.
Précisément, dans les applications considérées ici, les variables aléatoires normalisées sont définies comme

$$X_{n,i} = L_n(X_i, \dots, X_{i+d}),$$

où L_n est une suite de fonctions et $i = 1,\dots,n$, avec $n \in \mathbb{N}$. Dans ce cas, $\beta_n(k) \leq \beta_X(k)$ pourrait être choisi uniformément alors que dans les cas de dépendance faible on aurait par exemple $\theta_n(k) \leq \text{Lip}(L_n)\theta_X(k)$.

Un premier avantage à choisir la dépendance faible plutôt que des conditions de mélange est que les covariances de variables aléatoires sont souvent plus faciles à calculer que les coefficients de mélange. En fait, les conditions de mélange sont très difficiles et compliquées à vérifier (par exemple Doukhan [1994] donne les bornes des coefficients de mélange pour des modèles mélangeants).

Ensuite, on vient d'expliciter un exemple simple de processus autoregressif non-mélangeant en (1.7). Cependant, Doukhan & Louhichi [1999] montre que ce processus est faiblement dépendant. De plus, cette notion de dépendance faible est suffisamment large pour inclure de nombreux exemples intéressants tels que: des modèles de Markov stationnaires, bilinéaires, et plus généralement, des schémas de Bernoulli. Plus précisément, sous des hypothèses appropriées, tous les processus causaux et non-causaux sont faiblement dépendants. C'est le cas des processus Gaussiens, associés, linéaires, ARCH(∞), ou de Volterra. Pour plus de détails concernant la dépendance faible, on consultera [Dedecker et al., 2007] qui les approfondit ainsi que d'autres conditions de dépendance faible, avec de nombreux exemples et applications.

Le premier résultat de ce travail est une application de la méthode de Lindeberg dépendante de Bardet et al. [2007]:
Posons,

$$W_{n,j}(\mathbf{f}_k) := (nv_n)^{-1/2}\left(f_1(Y_{n,j}) - \mathbb{E}f_1(Y_{n,j}), \ldots, f_k(Y_{n,j}) - \mathbb{E}f_k(Y_{n,j})\right), \tag{1.9}$$

pour $1 \leqslant j \leqslant m_n$ et $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$, avec $k \geqslant 1$.

Théorème 2.1 (TLC de Lindeberg pour des fonctionnelles de clusters) *Soit $\mathcal{F}$ une classe de fonctionnelles de clusters qui satisfait les conditions (C.1) et (C.2) avec $r_n \ll v_n^{-1} \ll n$. Alors, si*

$$T_{t,m_n}(\mathbf{f}_k) := \sum_{j=1}^{m_n}\left|Cov\left(\exp(i\langle t, \sum_{s=1}^{j-1} W_{n,s}(\mathbf{f}_k)\rangle), \exp(i\langle t, W_{n,j}(\mathbf{f}_k)\rangle)\right)\right| \tag{1.10}$$

converge vers 0 lorsque $n \to \infty$, pour tout $t \in \mathbb{R}^k$ et tout $k-$uple $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$ de fonctionnelles de clusters, les lois fini-dimensionnelles (fidis) du processus empirique $(Z_n(f))_{f\in\mathcal{F}}$ de fonctionnelles de clusters convergent vers celles d'un processus Gaussien $(Z(f))_{f\in\mathcal{F}}$ centré et de fonction de covariance c.

La seule différence entre les théorèmes limites classiques de Lindeberg (dans le cas indépendant) et le Théorème 2.1 consiste à remplacer la condition d'indépendance par la convergence vers 0 de $T_{t,m_n}(\mathbf{f}_k)$. Clairement, cette expression $T_{t,m_n}(\mathbf{f}_k)$ est liée à la dépendance des variables aléatoires du tableau triangulaire $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$, qui est écrite en termes de sommes de covariances. Nous pouvons donc utiliser les coefficients de dépendance faible pour borner (1.10) et étendre le résultat pour des classes de modèles faiblement dépendants.

Proposition 2.1 *Soit $\mathcal{F}$ une classe de fonctionnelles de cluster $C-$Lipschitz (voire Définition 2.2) telle que les conditions (C.1), (C.2) et les conditions suivantes*

(C.3*) $\sqrt{Var\left(\Delta_n(f)\right)} = o\left(\frac{r_n^2}{n^2}\sqrt{nv_n}\right)$;

(C.4*) $\mathbb{E}^{1/2}\left(|f(Y_n^{(r_n-l_n)})|^2 \mathbb{1}_{\left\{|f(Y_n^{(r_n-l_n)})|>\sqrt{nv_n}\right\}}\right) = o\left(\frac{r_n^2}{n^2}\sqrt{nv_n}\right)$

soient satisfaites, avec (B). De plus, supposons qu'il existe des constantes positives C, α, ρ (ρ dépendante de n) telles que*

$$\sup_{x\in E}\sup_{1\leqslant i\leqslant n}\mathbb{P}\left(X_{n,i}\in B(x,\rho/2)\right)\leqslant C\rho^{\alpha}. \tag{1.11}$$

Alors, les répartitions fini-dimensionnelles (fidis) du processus empirique $(Z_n(f))_{f\in\mathcal{F}}$ de fonctionnelles de clusters convergent vers celles d'un processus Gaussien centré $(Z(f))_{f\in\mathcal{F}}$ de fonction de covariance c (définie en (C.2)), si le tableau triangulaire $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ satisfait au moins une des conditions suivantes de dépendance faible:

(D.1) *θ-faiblement dépendant tel que $\theta_n(k) = \mathcal{O}\left(k^{-\theta}\right)$ pour un $\theta > 1$ avec $l_n^{-\theta} = o\left(n^{-1}\right)$,*

(D.2) *η-faiblement dépendant tel que $\eta_n(k) = \mathcal{O}\left(k^{-\eta}\right)$ pour $\eta > 0$ avec $l_n^{-\eta} = o\left(r_n/n^2\right)$,*

(D.3) *κ-faiblement dépendant tel que $\kappa_n(k) = \mathcal{O}\left(k^{-\kappa}\right)$ pour $\kappa > 0$ avec $l_n^{-\kappa} = o\left(\frac{1}{r_n^u n^v}\wedge\frac{1}{n^2}\right)$, et $u \geqslant 0$ et $v \in [0,2]$,*

(D.4) *λ-faiblement dépendant tel que $\lambda_n(k) = \mathcal{O}\left(k^{-\lambda}\right)$ pour $\lambda > 0$ avec $l_n^{-\lambda} = o\left(\frac{1}{r_n^u n^v}\wedge\frac{1}{n^2}\right)$, et $u \geqslant 0$, $v \in [0,2]$.*

Nous avons mentionné jusqu'ici des théorèmes limites pour les fidis de processus empiriques de fonctionnelles de clusters. En revanche, la convergence fonctionnelle est plus délicate. En fait, les théorèmes de limite centrale fonctionnels (TLCFs) reposent généralement sur de lourds arguments de couplage sous une hypothèse de régularité absolue (voir [Doukhan et al., 1995]). Ces arguments sont plus problématiques sous hypothèses de dépendance faible. Cependant, les TLCs obtenus pour les fidis sont suffisants pour plusieurs exemples et applications. Plus particulièrement, nous utilisons la convergence des fidis de la Proposition 2.1 pour démontrer un TLC pour l'estimateur de l'extrémogramme dans la Section 2.4.

Pour la convergence fonctionnelle (ou uniforme), la τ-dépendance introduite par Dedecker & Prieur [2004a] est plus adaptée. Cette condition de dépendance est moins restrictive que la condition de mélange fort et peut-être plus simple à vérifier en plusieurs situations. En particulier, Dedecker & Prieur [2004a] montrent que le processus $AR(1)$ non-mélangeant défini en (1.7) est $\tau-$dépendant. Cette condition est aussi très puissante (voir [Dedecker & Prieur, 2004b, 2005; Dedecker et al., 2007]); elle fournit des arguments de couplage qui rendent possible de remplacer des blocs dépendants par des blocs indépendants. Nous utilisons cette dernière propriété avec les conditions (C.1)-(C.3) et les conditions (B), (C), (D.1) et (C.4) de la Section 3.3 pour démontrer la convergence des fidis du processus empirique $(Z_n(f))_{f\in\mathcal{F}}$ vers celles d'un processus Gaussien. Nous utilisons l'entropie à crochets pour définir les conditions nécessaires à la tension asymptotique, ainsi que des bornes sur les caractéristiques de dénombrement par rapport à une semi-métrique aléatoire, pour prouver l'équicontinuité asymptotique, en utilisant les résultats dans le cas indépendant, Van Der Vaart & Wellner [1996, § 2.11]. Les théorèmes de limite centrale uniformes suivent de la convergence des fidis et de la tension asymptotique, ou bien de l'équicontinuité asymptotique.

Précisément, soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité et soit $\mathcal{M}$ une sous-tribu de $\mathcal{A}$. Soit X une variable aléatoire à valeurs dans $E \subseteq \mathbb{R}^d$ et qui est $\mathbb{L}^p$-intégrable (*i.e.* X satisfait que $\|X\|_p = (\int \|x\|^p P_X(dx))^{\frac{1}{p}} < \infty$). Dedecker & Prieur [2004b] définissent le coefficient τ_p entre $\mathcal{M}$ et X:

$$\tau_p(\mathcal{M}, X) = \left\| \sup \left\{ \int h(x) P_{X|\mathcal{M}}(dx) - \int h(x) P_X(dx) : h \in \Lambda^*(E) \right\} \right\|_p , \qquad (1.12)$$

où P_X est la distribution de X, $P_{X|\mathcal{M}}$ est la probabilité conditionnelle de X sachant $\mathcal{M}$ et $\Lambda^*(E)$ dénote la classe de fonctions de Lipschitz $h : E \longrightarrow \mathbb{R}$ telles que $\mathrm{Lip}(h) \leqslant 1$. Soit $\mathbb{M} = \{(X_{n,i})_{1\le i\le n}\}_{n\in\mathbb{N}}$ un tableau triangulaire de variables aléatoires stationnaires par lignes qui prennent leurs valeurs dans E et qui sont $\mathbb{L}^p$-intégrables. Pour chaque $i \in \{1, \ldots, n\}$, $\mathcal{M}_{n,i} = \sigma(X_{n,j},\ j \le i)$, désigne la sous-tribu de $\mathcal{A}$ engendrée par les variables aléatoires normalisées $(X_{n,j})_{j\le i}$.

On dit que $\mathbb{M}$ est τ_p-faiblement dépendant, si le coefficient

$$\tau_{p,r}(k) = \max_{1\le l\le r} \frac{1}{l} \sup \left\{ \tau_p(\mathcal{M}_{n,i}, (X_{n,j_1}, \ldots, X_{n,j_l})),\ i + k \le j_1 < \cdots < j_l \le n \right\}, \quad (1.13)$$

est tel que $\lim_{k\longrightarrow\infty} \limsup_{r\longrightarrow\infty} \tau_{p,r}(k) = 0$, où la métrique sur E^l est définie par $\delta_l(x, y) = \sum_{i=1}^{l} \|x_i - y_i\|$.

En utilisant les proprietés de couplage du coefficient τ (Lemme 3.1), nous montrons l'existence d'une séquence de blocs aléatoires indépendants $(Y^*_{n,j})_{1\le j\le m_n}$ de taille $(r_n - l_n)$, tels que:

$$Y^*_{n,j} \overset{\mathcal{D}}{=} Y^{(r_n-l_n)}_{n,j} \quad et \quad \mathbb{E}\left(\delta_{r_n-l_n}(Y^{(r_n-l_n)}_{n,j},\ Y^*_{n,j})\right) \le r_n \tau_{1,r_n}(l_n), \quad j = 1, \dots, m_n, \tag{1.14}$$

lorsque les blocs $(Y_{n,j})_{1\leqslant j\leqslant m_n}$ sont p.s. bornés et si l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ est assez riche. Un résultat similaire est obtenu pour les blocs pairs et impairs (voir les Lemmes 3.3 et 3.4).

Avec ces deux lemmes et une approximation des fonctionnelles de clusters C-Lipschitz par des fonctionnelles de clusters Lipschitz, nous montrons les deux théorèmes suivants:

Théorèmes 3.2 et 3.3 *Soit $\mathcal{F}$ une classe de fonctionnelles de clusters C-Lipschitz. Supposons α, β, γ et η tels que les conditions (B), (C), (D.1) et (C.4) soient remplies. Alors,*

1. *si (T.1)-(T.4) sont satisfaites, la suite de processus empirique $(Z_n)_{n\in\mathbb{N}}$ est tendu dans $\ell^\infty(\mathcal{F})$.*

2. *si (T.1)-(T.3), (T.4') et (T.5) sont satisfaites, le processus empirique $(Z_n)_{n\in\mathbb{N}}$ est asymptotiquement équicontinu.*

Plus loin, si (C.2) et (C.3) sont remplies, alors Z_n converge vers un processus Gaussien centré Z avec une fonction de covariance c définit en (C.2).

Les théorèmes de la valeur extrême pour les séries temporelles à temps discret ou à temps continu sont généralement développés sous des hypothèses d'indépendance, ou bien sous des conditions de stationnarité et de mélange. Une référence classique pour ces résultats est Leadbetter et al. [1983]. Dans le contexte des champs aléatoires, Leadbetter & Rootzén [1998] prouvent un théorème de la valeur extrême (TVE), mais aussi, sous des conditions de mélange, ou plus précisément, sous la condition "Cw-mixing". Il s'agit de la condition habituelle de mélange fort (multipliée par le nombre de sous-blocs), où le coefficient de mélange est limité aux événements de type $E = \{\sup\{X_{\mathbf{t}} \ :\ \mathbf{t} \in A\} \leqslant u\}$, où $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$ dénote le champ aléatoire étudié et où A est un sous-ensemble de $\mathbb{Z}^d$. Cette condition de dépendance reste cependant difficile à vérifier.

En utilisant les arguments déjà mentionnés sur la dépendance faible, nous proposons donc un TVE pour des champs aléatoires stationnaires faiblement dépendants:

Théorème 4.1 *Soit* $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{N}^d\}$ *un champ aléatoire stationnaire et soit* $\mathbf{n} = (\lceil \vartheta_1(n) \rceil, \lceil \vartheta_2(n) \rceil, \dots, \lceil \vartheta_d(n) \rceil)$ *un chemin monotone croissant (défini au début de la section 4.5). Désignons* $M_{\mathbf{n}} = \sup\left\{X_{\mathbf{t}} : \mathbf{t} \in \prod_{i=1}^d \{1, 2, \dots, \lceil \vartheta_i(n) \rceil\}\right\}$ *et supposons qu'il existe des constantes* $a_{\mathbf{n}} > 0$ *et* $b_{\mathbf{n}} \in \mathbb{R}$ *telles que* $\mathbb{P}\left(a_{\mathbf{n}}^{-1}(M_{\mathbf{n}} - b_{\mathbf{n}}) \leqslant x\right) \longrightarrow G(x)$*, lorsque* $n \to \infty$*, avec* G *non-dégénérée. Alors, si* X *satisfait au moins l'une des conditions spatio-faiblement dépendants (SWD) et si* $\mathbb{P}(\sup\{X_{\mathbf{t}} : \mathbf{t} \in B\} = a_{\mathbf{n}} x + b_{\mathbf{n}}) = 0$*, pour tout couple* $(x, n) \in \mathbb{R} \times \mathbb{N}$ *et tout* $B \subset \mathbb{N}^d$*,* G *est une loi du type de valeurs extrêmes.*

Si le champ aléatoire est stationnaire et satisfait l'une des conditions SWD, nous obtenons aussi que le type de loi limite pour le maximum $M_{\mathbf{n}}$ est déterminé par le comportement de la queue de la fonction de répartition commune aux marginales du maximum des sous-blocs.

En effet, notons $M(B) = \sup\{X_{\mathbf{t}} : \mathbf{t} \in B\}$ pour $B \subseteq \mathbb{N}^d$ et $B_{j_1 j_2 \dots j_d} := \prod_{i=1}^d \{(j_i - 1)r_i + 1, \dots, j_i r_i\}$, où $r_i = r_{n_i} = o(n_i)$ et $\mathbf{n} = (n_1, \dots, n_d)$. On en déduit, par le Lemme 4.1, que l'ensemble des variables aléatoires

$$\left(M(B_{j_1, j_2, \dots, j_d})\right)_{j_1, \dots, j_d \in \{1, \dots, d\}}$$

admet 1 pour indice extrémal. Par conséquent, en utilisant cette "indépendance asymptotique des max-blocs", nous prouvons le résultat suivant:

Proposition 4.3 *Soit* $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ *un champ aléatoire stationnaire qui satisfait au moins l'une des conditions SWD. Notons* $m_i = \lceil n_i / r_i \rceil$ *(*$i \in \{1, \dots, d\}$*),* $m_{\mathbf{n}} = m_1 m_2 \cdots m_d$ *et* $\mathbf{J} := B_{11\dots1}$*. Supposons que le* $(1 - m_{\mathbf{n}}^{-1})$*−percentile* $\gamma_{\mathbf{n}}$ *de* $M(\mathbf{J})$ *satisfait:*

(TD) $\mathbb{P}(M(\mathbf{J}) > \gamma_{\mathbf{n}} + a_{\mathbf{n}} x) / \mathbb{P}(M(\mathbf{J}) > \gamma_{\mathbf{n}}) \longrightarrow H(x)$ *pour des constantes positives* $a_{\mathbf{n}}$ *et une fonction décroissante* $H(x)$ *telle que* $H(x) \underset{x \to -\infty}{\longrightarrow} \infty$ *et* $H(x) \underset{x \to \infty}{\longrightarrow} 0$*;*

Supposons $\mathbb{P}(M(B) = a_{\mathbf{n}} x + \gamma_{\mathbf{n}}) = 0$*, pour tout* $B \in \mathbb{N}^d$ *et tout couple* $(x, \mathbf{n}) \in \mathbb{R} \times \mathbb{N}^d$*. Alors,*

$$\mathbb{P}\left(a_{\mathbf{n}}^{-1}(M_{\mathbf{n}} - \gamma_{\mathbf{n}}) \leqslant x\right) \underset{\mathbf{n} \to \infty}{\longrightarrow} G(x) = \exp(-H(x)). \tag{1.15}$$

Finalement, dans le dernier chapitre, nous utilisons ces idées et les idées de la preuve des lemmes de Lindeberg indépendant et dépendant de Bardet et al. [2007] pour développer une extension du Théorème 2.1 aux champs aléatoires et aux processus spatio-temporels stationnaires faiblement dépendants. De plus, nous définissons un extrémogramme pour les processus spatio-temporels et nous utilisons le théorème (étendu) pour établir sa convergence en loi.

Chapter 2

Dependent Lindeberg CLT for the fidis of empirical processes of cluster functionals.

Drees & Rootzén [2010] have proven central limit theorems for empirical processes of extreme values cluster functionals built from β-mixing processes. However this family of β-mixing processes is quite restrictive. In this chapter, we expand some of these results, for the finite-dimensional marginal distributions (fidis), to a more general dependent processes family, known as weakly dependent processes in the sense of Doukhan & Louhichi [1999]. In this context, the central limit theorem (CLT) for the fidis of empirical processes of cluster functionals is sufficient in some applications, for instance, we show the convergence without mixing conditions of the extremogram estimator introduced by Davis & Mikosch [2009], including a small example with simulation of the extremogram of a weakly dependent random process but non mixing, to confirm the efficacy of our result.

2.1 Introduction

In light of recent developments in massive data processing via *parallel processing*, it is of interest to consider the construction of statistics as functions of data blocks. In the case of extremes (rare events), not only very little data are relevant in estimations, but they are also hidden among a large mass of "common data". Thus comes the natural idea of considering clustering of extremes, which consists here to obtain the smaller sub-block of extreme values over each block, while conveniently suppressing "common" data in each block, generally replaced by a null value. Such null values may be inoffensive mathematically, yet they are an obstacle computationally in terms of time. These and many other reasons encourage the study of extreme cluster functionals. In particular, this chapter contributes to the asymptotic behaviour of extreme cluster functionals, providing an extension of the dependence conditions for the central limit theorems of Drees & Rootzén [2010, § 2.1], *i.e.* in the case of the finite-dimensional marginal distributions (fidis) of empirical processes of extreme cluster functionals.

For this, we use the dependent Lindeberg method of Bardet et al. [2007], where central limit theorems are obtained if besides the usual Lindeberg condition, a sequence $T = T(n)$ (which summarises the dependence of the partial sums) converges to zero as n, the number of random variables, tends to infinity. This term T of dependence writes as the sum of covariances, bounded here by weak-dependence coefficients defined by Doukhan & Louhichi [1999]. Therefore, for weakly dependent random variables, we obtain CLT for the fidis of empirical processes of cluster functionals under convenient conditions for the decay rates of the weak-dependence coefficients.

This extension is motivated by several reasons. The first reason for this extension is that such weak-dependence assumptions include non-mixing models: consider e.g. the AR(1)-input, solution of the recursion

$$X_k = \frac{1}{b}\left(X_{k-1} + \xi_k\right), \quad k \in \mathbb{Z}, \tag{2.1}$$

with $b \geqslant 2$ integer, and with $(\xi_k)_{k\in\mathbb{N}}$ as a sequence of independent and uniformly distributed random variables on the set $U(b) := \{0, 1, \ldots, b-1\}$. This process is not mixing in the sense of Rosenblatt [1956], see [Andrews, 1984] for $b = 2$ and [Ango Nze & Doukhan, 2004] for $b > 2$. However Doukhan & Louhichi [1999] proved that such process is weakly dependent.

The second reason is that mixing assumptions are difficult of checking, e.g. Doukhan [1994] provides, with evident difficulties, explicit bounds of the decay of mixing sequences. On the other hand, note that an important property of associated random variables is that zero correlation implies independence (see [Newman, 1984]). Therefore, one may hope that dependence will appear in this case only through the covariance structure, which is much easier to compute that a mixing coefficient.

More generally, under weak assumptions, weak-dependence in the sense of Doukhan & Louhichi [1999], includes models like Bernoulli shifts, Markov, associated, mixing, etc.

This chapter is organised as follows: Section 2.2 is dedicated to recall the definitions of cluster functional and a general empirical process of cluster functionals. Additionally we present some normalised (extreme) random variables which are applied some cluster functionals. In Section 2.3 a general CLT for the fidis of these empirical processes based on Lindeberg method by Bardet et al. [2007] is provided. We define weak-dependence and we present examples of suitable weakly dependent random variables. Finally, we apply the initial theorem to such weakly dependent random variables, in order to derive CLT under specific decay rates assumptions for the weak-dependence coefficients and under specific conditions for the marginal distributions of the random variables and of the cluster functionals. In Section 2.4 we develop the extremogram of Davis & Mikosch [2009] as an example where convergence of the fidis of empirical processes of cluster functionals is sufficient, in fact, this provides us with a complete application. We also include a simulation study of the extremogram estimator of a non-mixing model. Proofs are given in Section 2.5.

2.2 Basic definitions

In this section, we outline some basic definitions and hypotheses that we use in this and the next chapter in order to prove limit theorems of empirical processes of (extreme) cluster functionals.

2.2.1 Cluster functionals

Roughly, an extreme cluster functional is a map that acts on blocks (arbitrary-length but not random-length) of "extremes" of random variables in such a way that for each block B, the map takes the same value in B and B^c, where B^c is the *core* of B, defined as the smaller sub-block of B that contains all the extreme values and the non-extreme values among them. Besides, this application is null for each block B that does not contains extreme values.

Formally, let $(E, \mathcal{E})$ be a measurable subspace of $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ for some $d \geqslant 1$ such that $0 \in E$. Following the deterministic definition of Drees & Rootzén [2010], we consider first the set

$$E_\cup := \bigcup_{r=1}^{\infty} E^r,$$

of E-valued sequences of finite length, which is equipped with the σ-field $\mathcal{E}_\cup$ induced by the Borel-σ-fields on E^r, for $r \in \mathbb{N}$. Let $x \in E_\cup$, then we can write $x = (x_1, \ldots, x_r)$

for some $r \in \mathbb{N}$. The **core** $x^c \in E_\cup$ of x is defined [(1)] by

$$x^c := \begin{cases} (x_{r_I}, x_{r_I+1}, \dots, x_{r_S}), & \text{if } x \neq 0_r \\ \\ 0, & \text{otherwise,} \end{cases}$$

where 0_r denotes the null element in E^r, besides $r_I := \min\{i \in \{1, \dots, r\} : x_i \neq 0\}$ (the first non-null value of the block x) and $r_S := \max\{i \in \{1, \dots, r\} : x_i \neq 0\}$ (the last non-null value of the block x).

Definition 2.1 *A* ***cluster functional*** *is a measurable map* $f : (E_\cup, \mathcal{E}_\cup) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *such that*

$$f(x) = f(x^c), \qquad \text{for all } x \in E_\cup, \quad \text{and} \quad f(0_r) = 0 \ (\forall r \geqslant 1). \tag{2.2}$$

Under the definition (2.2), it is easy to build a large amount of examples of cluster functionals. Nevertheless, the useful examples used to build estimators through these cluster functionals are functionals of the type:

$$f(x_1, \dots, x_r) = \sum_{i=1}^{r} \phi(x_i), \tag{2.3}$$

where $\phi : E \longrightarrow \mathbb{R}$ is a measurable function such that $\phi(0) = 0$. Generally speaking, these functions ϕ are indicator functions (or products of another measurable function $H : E \longrightarrow \mathbb{R}$ with an indicator function). For example, for $E = [0, \infty)$,

- if $\phi(x) = x\mathbb{1}_{(0,\infty)}(x)$, then (2.3) is the sum of excesses of the block $B = (x_1, \dots, x_r)$.
- if $\phi(x) = \mathbb{1}_{(0,\infty)}(x)$, then (2.3) is the number of excesses of the block $B = (x_1, \dots, x_r)$, where $\mathbb{1}_A(\cdot)$ denotes the usual indicator function of a subset A.

Another classic examples for $E = [0, \infty)$ are the following:
the component-wise maximum of a cluster

$$f(x_1, \dots, x_r) = \max_{1 \leqslant i \leqslant r} x_i, \tag{2.4}$$

[1]Note that the core also considers the null values that exist between the non-null values. For example: $(0, 1, 2, 0, 0, 3, 0, 1, 0, 0)^c = (1, 2, 0, 0, 3, 0, 1)$, which is the smaller sub-block of $x = (0, 1, 2, 0, 0, 3, 0, 1, 0, 0)$ which contains all non-null values as well as the null values between them.

and the duration of a cluster

$$f(x_1,\dots,x_r) = \max\{i : x_i > 0\} - \min\{i : x_i > 0\}. \tag{2.5}$$

For $E = \mathbb{R}$, the number of threshold up-crossings

$$f(x_1,\dots,x_r) = \mathbb{1}_{(0,\infty)}(x_1) + \mathbb{1}_{(-\infty,0]\times(0,\infty)}(x_1,x_2) + \cdots + \mathbb{1}_{(-\infty,0]\times(0,\infty)}(x_{r-1},x_r). \tag{2.6}$$

2.2.2 Extremes of random variables and some normalisations

On the other hand, let us consider a triangular array $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ of E-valued normalised random variables $X_{n,i}$, defined on some probability space $(\Omega,\mathcal{A},\mathbb{P})$ for $i = 1,\dots,n$ and $n \in \mathbb{N}$. We will always assume in this document that $\mathbb{M}$ is row-wise stationary, this means that $(X_{n,i})_{1\leqslant i\leqslant n}$ is stationary for each $n \in \mathbb{N}$. Additionally, we require that the sequence of conditional distributions of $X_{n,1}$, given that $X_{n,1}$ belongs to the failure set $A \subseteq E \setminus \{0\}$ (2), converges weakly to some non-degenerate limit. *i.e.* we assume that there exists a non-degenerate distribution G such that

$$\mathbb{P}(X_{n,1} \in \cdot \mid X_{n,1} \in A) \underset{n\to\infty}{\longrightarrow} G(\cdot).$$

Generally, these normalised random variables $X_{n,i}$ are built from another random process $(X_i)_{i\in\mathbb{Z}}$, in such a way that the normalisation maps all non-extreme values to zero.
For instance, let $(X_i)_{i\in\mathbb{N}}$ be a real-valued stationary time series with a marginal cumulative distribution function F and let $(u_n)_{n\in\mathbb{N}}$ be a non-decreasing sequence of thresholds such that $u_n \uparrow x_F$, where

$$x_F = \sup\{x \in \mathbb{R} : F(x) < 1\} \quad \text{and} \quad v_n = \mathbb{P}(X_1 > u_n) \underset{n\to\infty}{\longrightarrow} 0.$$

Note that the tail distribution function of X_1 may be asymptotically degenerate, which means that there exists a point $a \in \mathbb{R}$ such that

$$\overline{P}_n(x) = \mathbb{P}(X_1 - u_n > x | X_1 > u_n) \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{(-\infty,a]}(x).$$

However, if F belongs to the domain of attraction of some extreme-value distribution, then as this is proved in [Pickands, 1975], there exist $\gamma \in \mathbb{R}$ and a sequence of positive

[2] In fact, here the term "extreme" means "to belong to the failure set". Therefore, we will say in general that $X_{n,i}$ is an extreme value if and only if $X_{n,i}$ belongs to the failure set A.

constants $(a_n)_{n\in\mathbb{N}}$, depending on the sequence u_n, such that

$$P_n(x) = \mathbb{P}(X_{n,1} > x | X_1 > u_n) \underset{n\to\infty}{\longrightarrow} \begin{cases} (1+\gamma x)_+^{-1/\gamma}, & \text{if } \gamma \neq 0, \\ \exp(-x), & \text{if } \gamma = 0, \end{cases} \tag{2.7}$$

locally uniform on $(0,\infty)$, where

$$X_{n,i} = \left(\frac{X_i - u_n}{a_n}\right)_+ := \max\left\{\frac{X_i - u_n}{a_n}, 0\right\}, \qquad \text{for } 1 \leqslant i \leqslant n; \tag{2.8}$$

are the normalised excesses of X_i over u_n.

For the multidimensional case, let $\mathbb{X} = (\mathbf{X}_i)_{i\in\mathbb{N}}$ be a stationary $\mathbb{R}^d$-valued time series such that all components of $\mathbf{X}_i$ have the same marginal distribution. Since such time series $\mathbb{X}$ may exhibit dependence across coordinates and over time, if $\|\cdot\|$ denotes an arbitrary norm on $\mathbb{R}^d$, then an interesting normalisation for the study of the extreme values of $\mathbb{X}$ would be:

$$X_{n,i} = u_n^{-1}\mathbf{X}_i \mathbb{1}_{\{\|\mathbf{X}_i\| > u_n\}}, \qquad \text{for } 1 \leqslant i \leqslant n; \tag{2.9}$$

where $(u_n)_{n\in\mathbb{N}}$ is a sequence of high quantiles of the process $\mathbb{X}$.

Another example for the multidimensional case is the following. Let $d \geq 1$ and let $(\mathbf{X}_i)_{i\in\mathbb{N}}$ be a $\mathbb{R}^d$-valued random process such that $\mathbf{X}_i = (\xi_{i,1}, \ldots, \xi_{i,d})$ admit the same marginal distribution for all their coordinates. Then, in this case, an usual normalisation for $\mathbf{X}_i$ would be:

$$X_{n,i} = \left(\left(\frac{\xi_{i,1} - u_n}{a_n}\right)_+, \left(\frac{\xi_{i,2} - u_n}{a_n}\right)_+, \ldots, \left(\frac{\xi_{i,d} - u_n}{a_n}\right)_+\right), \tag{2.10}$$

where $(u_n)_{n\in\mathbb{N}}$ and $(a_n)_{n\in\mathbb{N}}$ are defined as in eqn. (2.8). Here, $X_{n,i}$ is the vector of $d-$normalised excesses over the threshold u_n for each coordinate.
An example in this frame is, with $\xi_{i,j} = X_{i+j-1}$ for $1 \leq j \leq d$, the normalisation of d consecutive excesses of a real-valued process $(X_i)_{i\in\mathbb{N}}$:

$$X_{n,i} = \left(\left(\frac{X_i - u_n}{a_n}\right)_+, \left(\frac{X_{i+1} - u_n}{a_n}\right)_+, \ldots, \left(\frac{X_{i+d-1} - u_n}{a_n}\right)_+\right). \tag{2.11}$$

This example brings information on the extremal dependence structure and fits situations like:

(i) d consecutive days of rain observed in a given city, such that the volume of precipitations may be larger than the volume of water that can be drained (through sewers, soil, rivers, etc.),

(ii) d very large claims reported to an insurance company in a very small time interval, with respect to typical cases. This may be a risk with respect to the response capacity of the insurance company, and

(iii) d consecutive days of low temperatures observed in a given city, such that the power consumption (due to the use of heating appliances, etc.) endangers the response capacity of the company in charge of the energy distribution.

2.2.3 Empirical processes of cluster functionals

We can now to define a general empirical process indexed by a family of cluster functionals. Indeed, let $\mathcal{F}$ be a class of cluster functionals and let $Y_{n,j}$ be the j-th block of r_n consecutive values of the n-th row $(X_{n,i})_{1\leqslant i\leqslant n}$. There are thus $m_n := \lceil n/r_n \rceil = \max\{j \in \mathbb{N} : j \leqslant n/r_n\}$ blocks,

$$Y_{n,j} := (X_{n,i})_{(j-1)r_n+1\leqslant i\leqslant jr_n}, \tag{2.12}$$

of length r_n, with $1 \leqslant j \leqslant m_n$. The **"empirical process Z_n of cluster functionals"** in $\mathcal{F}$, is the process $(Z_n(f))_{f\in\mathcal{F}}$ defined by

$$Z_n(f) := \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} (f(Y_{n,j}) - \mathbb{E}f(Y_{n,j})), \tag{2.13}$$

where $v_n := \mathbb{P}(X_{n,1} \in A)$ with $A \subseteq E \setminus \{0\}$ (the failure set).

In particular, observe that if the blocks $\{(Y_{n,j})_{1\leqslant j\leqslant m_n}\}_{n\in\mathbb{N}}$ are independent and if we take into account the following essential convergence assumptions:

(C.1) $\mathbb{E}\left((f(Y_{n,1}) - \mathbb{E}f(Y_{n,1}))^2 \mathbb{1}_{\{|f(Y_{n,1})-\mathbb{E}f(Y_{n,1})|>\epsilon\sqrt{nv_n}\}} \right) = o(r_n v_n)$, for all $\epsilon > 0$, and all $f \in \mathcal{F}$.

(C.2) $(r_n v_n)^{-1}\mathrm{Cov}\,(f(Y_{n,1}), g(Y_{n,1})) \longrightarrow c(f,g)$, for all $f, g \in \mathcal{F}$,

with $r_n \ll v_n^{-1} \ll n$, then the fidis of the empirical process $(Z_n(f))_{f\in\mathcal{F}}$ of cluster functionals converge to the fidis of a Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c.

Drees & Rootzén [2010] have proved central limit theorems (CLTs) for the process (2.13). In particular, the convergence of the fidis of $(Z_n(f))_{f\in\mathcal{F}}$ were proved applying the Bernstein blocks technique together with a β-mixing coupling condition, to boil down convergence of sum over the original blocks to convergence of sums over i.i.d. blocks through Eberlein [1984]'s technique involving the metric of total variation. However, the family of mixing processes is quite restrictive. We can see this through a particularly simple example: the AR(1)-process defined in (2.1), which is not even α-mixing. Therefore, the results of Drees & Rootzén [2010] can not be used here.
In our case, we solve this problem for the fidis of such empirical process through dependent Lindeberg method, developed by Bardet et al. [2007], followed by its applications to weakly dependent random sequences defined by Doukhan & Louhichi [1999].

The weak spot under these weak-dependence conditions, in the sense of Doukhan & Louhichi [1999], is that we have no coupling arguments to arrive at a uniform CLT, as Drees & Rootzén [2010] have done in their paper by using the rich coupling properties of the β-mixing processes together with Van Der Vaart & Wellner [1996]'s tightness criteria and asymptotic equicontinuity conditions.

The benefit of this work is that the convergence of the fidis is sufficient in several examples and applications. We will show a particularly useful application in Section 2.4.

2.3 Limit theorems for the fidis of empirical processes of cluster functionals

In order to adapt the dependent Lindeberg method of Bardet et al. [2007] to cluster functionals, let us write

$$W_{n,j}(\mathbf{f}_k) := (nv_n)^{-1/2}\left(f_1(Y_{n,j}) - \mathbb{E}f_1(Y_{n,j}), \ldots, f_k(Y_{n,j}) - \mathbb{E}f_k(Y_{n,j})\right), \tag{2.14}$$

for $1 \leqslant j \leqslant m_n$ and $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$. Therefore, with this notation we have derived the following result:

Theorem 2.1 (Lindeberg CLT for cluster functionals) *Let $\mathcal{F}$ be a family of cluster functionals such that the assumptions (C.1) and (C.2) hold with $r_n \ll v_n^{-1} \ll n$.*

Then, if

$$T_{t,m_n}(\mathbf{f}_k) := \sum_{j=1}^{m_n} \left| Cov\left(\exp(i\langle t, \sum_{s=1}^{j-1} W_{n,s}(\mathbf{f}_k)\rangle), \exp(i\langle t, W_{n,j}(\mathbf{f}_k)\rangle) \right) \right| \tag{2.15}$$

converges to 0 *as* $n \to \infty$*, for all* $t \in \mathbb{R}^k$ *and all k-tuple of cluster functionals* $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$*, the fidis of the empirical process* $(Z_n(f))_{f\in\mathcal{F}}$ *of cluster functionals converge to the fidis of a Gaussian process* $(Z(f))_{f\in\mathcal{F}}$ *with covariance function c.*

We have just seen that the convergence of the fidis of $(Z_n(f))_{f\in\mathcal{F}}$ to a Gaussian law is moreover obtained because $T_{t,m_n}(f_1, \ldots, f_k)$ converges to 0, as n tends to infinity, for all $t \in \mathbb{R}^k$ and all $(f_1, \ldots, f_k) \in \mathcal{F}^k$ with $k \in \mathbb{N}$. Actually this expression (2.15) is related to the dependence of the random variables of $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$, which is written in terms of sums of covariances, that can conveniently be bounded by weak-dependence coefficients defined in [Doukhan & Louhichi, 1999] and [Dedecker et al., 2007].

2.3.1 Weak Dependence

Let $f : E^r \subseteq (\mathbb{R}^d)^r \longrightarrow \mathbb{R}$ be a function, with $r \in \mathbb{N}$. As usual, we denote by

$$\text{Lip}(f) := \sup_{(x_1,\ldots,x_r)\neq(y_1,\ldots,y_r)\in E^r} \frac{|f(x_1, \ldots, x_r) - f(y_1, \ldots, y_r)|}{\|x_1 - y_1\| + \cdots + \|x_r - y_r\|}.$$

Similar to the definition of Doukhan & Louhichi [1999], we say that a triangular array of row-wise stationary E-valued random variables $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ is (ϵ, ψ)-**weakly dependent** ((ϵ, ψ)-WD) if there exist a function $\psi : \mathbb{N}^2 \times (\mathbb{R}^+)^2 \longrightarrow \mathbb{R}^+$, an infinite increasing sequence of positive integers $(l_n)_{n\in\mathbb{N}}$ with $l_n \ll n$, and a positive sequence $(\epsilon_n(l))_{1\leqslant l\leqslant n, n\in\mathbb{N}}$ such that $\epsilon_n(l_n) \underset{n\to\infty}{\longrightarrow} 0$ and

$$\left|Cov\left(f(X_{n,i_1}, \ldots, X_{n,i_u}), g(X_{n,j_1}, \ldots, X_{n,j_v})\right)\right| \leqslant \psi(u, v, Lip(f), Lip(g)) \cdot \epsilon_n(l), \tag{2.16}$$

for all $(u, v) \in \mathbb{N} \times \mathbb{N}$, all $(i_1, \ldots, i_u) \in \mathbb{N}^u$ and $(j_1, \ldots, j_v) \in \mathbb{N}^v$ with $i_1 < \cdots < i_u < i_u + l \leqslant j_1 < \cdots < j_v \leqslant n$, and for all pair of functions $(f, g) \in \Lambda^u(E) \times \Lambda^v(E)$, where $\Lambda^s(E) := \{h : E^s \longrightarrow \mathbb{R} \text{ Lipschitz with } \|h\|_\infty \leqslant 1 \text{ and } \text{Lip}(h) < \infty\}$.

We will consider in this chapter four different particular cases of functions ψ corresponding to four cases of weakly dependent sequences (for other cases, see [Doukhan & Louhichi, 1999; Dedecker et al., 2007]). These cases are defined as follows:

1. Let $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ be a causal shift triangular array, this means that there exist a sequence of measurable functions $H_n : D^{\mathbb{Z}} \longrightarrow E$, with $n \in \mathbb{N}$, and a D-valued sequence of independent and identically distributed random variables (i.i.d.r.v's) $(\xi_i)_{i\in\mathbb{Z}}$ such that $X_{n,1} = H_n(\xi_1, \xi_0, \xi_{-1}, \ldots)$ is defined almost surely for all $n \in \mathbb{N}$. For these triangular arrays, we consider then the **θ-weakly dependent causal condition**, which is defined by (2.16) with

$$\psi(u, v, Lip(f), Lip(g)) = vLip(g). \tag{2.17}$$

   In this case, we will simply write $\theta_n(l)$ instead of $\epsilon_n(l)$.

2. If $\mathbb{M}$ is a noncausal triangular array, the **η, κ, λ-weakly dependent conditions** are defined respectively by the functions

$$\psi(u, v, Lip(f), Lip(g)) = uLip(f) + vLip(g), \tag{2.18}$$
$$\psi(u, v, Lip(f), Lip(g)) = uvLip(f)Lip(g), \tag{2.19}$$
$$\psi(u, v, Lip(f), Lip(g)) = uLip(f) + vLip(g) + uvLip(f)Lip(g), \tag{2.20}$$

   such that (2.16) is satisfied, where we write $\eta_n(l)$, $\kappa_n(l)$ and $\lambda_n(l)$, respectively, instead of $\epsilon_n(l)$.

Now, we give a little list of examples of weakly dependent triangular arrays with their dependence properties.

Example 2.1 (Association models) Let $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ be an associated row-wise stationary triangular array, *i.e.* for all coordinatewise non-decreasing real-valued functions h and k,

$$\text{Cov}\left(h(X_{n,i}, i \in A), k(X_{n,i}, i \in B)\right) \geqslant 0$$

holds for all subsets A and B of $\{1, 2, \ldots, n\}$, with $n \in \mathbb{N}$. Then $\mathbb{M}$ is κ-weakly dependent such that $\kappa_n(l) = \mathcal{O}\left(\sup_{i>l} \text{Cov}(X_{n,1}, X_{n,i})\right)$ (see [Doukhan & Louhichi, 1999]).

Example 2.2 (Gaussian models) If $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ is a Gaussian triangular array and $\lim_{i\to\infty} \limsup_{n\to\infty} \text{Cov}(X_{n,1}, X_{n,i+1}) = 0$, then $\mathbb{M}$ is a κ-weakly dependent triangular array such that $\kappa_n(l) = \mathcal{O}\left(\sup_{i>l} |\text{Cov}(X_{n,1}, X_{n,i})|\right)$ (see [Doukhan & Louhichi, 1999]).

Example 2.3 (Noncausal shifts models) Let $H_n : D^{\mathbb{Z}} \longrightarrow E$ be a sequence of measurable functions and $(\xi_i)_{i\in\mathbb{Z}}$ be a D-valued sequence of i.i.d.r.v's. A noncausal shift triangular array with innovation process $(\xi_i)_{i\in\mathbb{Z}}$ is a triangular array $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ such that the entries are defined by $X_{n,i} = H_n\left(\xi_{i-j}, j \in \mathbb{Z}\right)$, with $1 \leqslant i \leqslant n$ and $n \in \mathbb{N}$. Assume now that $X_{n,i}$ is defined almost surely for some (and therefore for all) $i \in \{1, \dots, n\}$, for each $n \in \mathbb{N}$; and that the sequence $(\Delta_n(l))_{l\in\mathbb{N}}$ defined by

$$\Delta_n(l) := \mathbb{E}\left|H_n(\xi_{i-j}, j \in \mathbb{Z}) - H_n(\xi_{i-j}\mathbb{1}_{\{|j|\leqslant l\}}, j \in \mathbb{Z})\right|, \tag{2.21}$$

converges to zero as $l = l_n$ tends to infinity. Then $\mathbb{M}$ is η-weakly dependent with $\eta_n(l) \leqslant 2\Delta_n\left(\lceil l/2\rceil\right)$ (see [Dedecker et al., 2007, § 3.1.2]).

Example 2.4 (Causal shifts models) Let $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ be a causal shift triangular array with innovation process $(\xi_i)_{i\in\mathbb{Z}}$ and let $(\xi_i^*)_{i\in\mathbb{Z}}$ be an independent copy of the independent and identically distributed sequence $(\xi_i)_{i\in\mathbb{Z}}$. Assume now that the sequence $(\Delta_n^*(l))_{l\in\mathbb{N}}$ defined by

$$\Delta_n^*(l) := \mathbb{E}\left|H_n(\xi_l, \xi_{l-1}, \dots) - H_n(\xi_l, \xi_{l-1}, \dots, \xi_2, \xi_1^*, \xi_0^*, \dots)\right| \tag{2.22}$$

converges to zero as $l = l_n$ tends to infinity. Then $\mathbb{M}$ is θ-weakly dependent with $\theta_n(l) \leqslant \Delta_n^*(l)$ (see [Dedecker et al., 2007, § 3.1.4]).

Application 2.1 *Suppose that $X = (X_i)_{i\in\mathbb{Z}}$ is a causal linear process such that*

$$X_i = \sum_{j\geqslant 0} b_j \xi_{i-j} \quad i \in \mathbb{Z}. \tag{2.23}$$

Consider now the normalisation (2.8), i.e. for each $n \in \mathbb{N}$ and $i \in \{1, \dots, n\}$; $X_{n,i} = H_n(\xi_i, \xi_{i-1}, \dots) = a_n^{-1}\left(\sum_{j\geqslant 0} b_j \xi_{i-j} - u_n\right)_+$, for some constants $a_n > 0$ and $u_n \in \mathbb{R}$ such that (2.7) holds. Then, if ξ_0^2 is integrable, we obtain that

$$\Delta_n^*(l) \leqslant \frac{\sqrt{v_n}}{a_n}\left(2Var(\xi_0)\sum_{j\geqslant l-1} b_j^2\right)^{1/2}, \tag{2.24}$$

where $v_n = \mathbb{P}(X_0 > u_n)$. In particular, observe that the AR(1) - process (2.1) can be rewritten as the causal linear process $X_i = \sum_{j\geqslant 0} b^{-j-1}\xi_{i-j}$, with ξ_0 uniformly distributed on $\{0, \dots, b-1\}$. In this case X_0 is uniformly distributed over $[0, 1]$ and $a_n = 1 - u_n = v_n$, therefore $\Delta_n^*(l) \leqslant \dfrac{b^{-(l+1)}}{\sqrt{6v_n}}$.

Example 2.5 (Markov models) Let $F : (E^r, \mathcal{B}(E^r)) \times (D, \mathcal{D}) \longrightarrow (E, \mathcal{B}(E))$ be a measurable function and let $\{(X_{n,i})_{1 \leqslant i \leqslant n}\}_{n \in \mathbb{N}}$ be a triangular array with entries in E such that

$$X_{n,i} = F(X_{n,i-1}, X_{n,i-2}, \ldots, X_{n,i-r}; \xi_i), \qquad i = r+1, \ldots, n; \tag{2.25}$$

for some sequence $(\xi_i)_{i \geqslant r+1}$ of i.i.d.r.v's with values in a measurable space D and independent of $(X_{n,1}, \ldots, X_{n,r})$. Let $\tilde{Y}_{n,i} = (X_{n,i}, X_{n,i-1}, \ldots, X_{n,i-r+1})$. Then $\tilde{Y}_{n,i} = M(\tilde{Y}_{n,i-1}; \xi_i)$ defines a Markov chain for $i = r+1, \ldots, n$ and $n \in \mathbb{N}$ with

$$M(x_1, x_2, \ldots, x_r; \xi) = (F(x_1, \ldots, x_r; \xi), x_1, x_2, \ldots, x_{r-1}).$$

For each $n \in \mathbb{N}$, assume that $(X_{n,i})_{i=r+1,\ldots,n}$ is a stationary solution of (2.25). Let $\tilde{Y}_{n,r} = (X_{n,r}, \ldots, X_{n,1})$ and let $\tilde{Y}^*_{n,r} = (X^*_{n,r}, \ldots, X^*_{n,1})$ be an independent vector with the same law as $\tilde{Y}_{n,r}$. Then, if we define the recurrence

$$X^*_{n,i} = F(X^*_{n,i-1}, X^*_{n,i-2}, \ldots, X^*_{n,i-r}; \xi_i), \quad i = r+1, \ldots, n; \tag{2.26}$$

with $n \in \mathbb{N}$, we obtain that $\theta_n(l) \leqslant \tilde{\Delta}_n(l) := \mathbb{E}|X_{n,l} - X^*_{n,l}|$.
If, e.g., F is such that

$$\|F(x_1, \ldots, x_r; \xi) - F(y_1, \ldots, y_r; \xi)\|_p \leq \sum_{i=1}^{r} a_i |x_i - y_i|, \quad \text{with} \quad \sum_{i=1}^{r} a_i < 1, \tag{2.27}$$

for some $p \geqslant 1$, then $\tilde{\Delta}_n(l) \leqslant C a^l$ for some $a \in [0,1)$ and some $C > 0$. This is proved in [Dedecker et al., 2007, page 34].

Application 2.2 (Contractive Markov chains) *Let $X_{n,i} = F(X_{n,i-1}, \xi_i)$ be a Markov chain such that $F : (E, \mathcal{B}(E)) \times (D, \mathcal{D}) \longrightarrow (E, \mathcal{B}(E))$ is a measurable function which satisfies:*

$$A^p = \mathbb{E}\|F(0;\xi)\|^p < \infty, \text{ and } \|F(x;\xi) - F(y;\xi)\|^p \leq a^p \|x - y\|^p, \tag{2.28}$$

for some $a \in (0,1)$ and some $p \in [1, \infty]$. Then $(X_{n,i})_{1 \leqslant i \leqslant n}$ has a stationary solution with p-th order finite moment for each $n \in \mathbb{N}$ (see [Dedecker et al., 2007, page 35]). Moreover under this condition, $\tilde{\Delta}_n(l) = \|X^*_{n,1} - X_{n,1}\|_p \cdot a^l$.

Remark 2.1 *Let $E = \mathbb{R}$ and $F(u;\xi) = A(u) + B(u)\xi$, for suitable Lipschitz functions $A(u)$ and $B(u)$ with $u \in \mathbb{R}$. Then the corresponding iterative model (ARCH-type process) $X_{n,i} = F(X_{n,i-1}; \xi_i)$ satisfies (2.28) with $a = Lip(A) + \|\xi_1\|_p Lip(B) < 1$.*

Application 2.3 (Nonlinear AR(l)-models) *Let $r \geq 1$ and $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ be the triangular array row-wise stationary solution of some equation*

$$X_{n,i} = R(X_{n,i-1}, \ldots, X_{n,i-r}) + \xi_i,$$

for some measurable function $R : \mathbb{R}^r \longrightarrow \mathbb{R}$. If $\|\xi_1\|_p < \infty$ and

$$|R(u_1, \ldots, u_r) - R(v_1, \ldots, v_r)| \leq \sum_{i=1}^{r} a_i |u_i - v_i|, \text{ for } a_1, \ldots, a_r \geq 0, \text{ with } \sum_{i=1}^{r} a_i < 1,$$

and for all $(u_1, \ldots, u_r), (v_1, \ldots, v_r) \in \mathbb{R}^r$, then the function $G : \mathbb{R}^{r+1} \to \mathbb{R}$ defined by $G(u;\xi) = R(u) + \xi$ satisfies Condition (2.27) and therefore the sequence $(\tilde{\Delta}_n(l))_l$ is bounded above by Ca^l, for some $a \in [0, 1)$ and $C > 0$.

Under suitable assumptions, the families of causal and non-causal bilinear processes, non-causal finite order Volterra processes, causal and non-causal infinite memory processes, etc., are also weakly dependent. For more details of weak-dependence properties of these processes, see the book of Dedecker et al. [2007].

2.3.2 CLTs for the fidis

In the result below, we give a CLT for the fidis of the empirical process $(Z_n(f))_{f\in\mathcal{F}}$ of cluster functionals with weakly dependent triangular arrays. In the proof (Section 2.5), we need that the functionals $f \in \mathcal{F}$ can be approximated by a sequence of Lipschitz cluster functionals $(f_n)_{n\geqslant 0}$. In order to build such functionals f_n, we consider f inside a special subclass of cluster functionals, which we have called C-Lipschitz cluster functionals (defined below), together with certain truncation assumption (C.4) on the functionals f evaluated at the sub-blocks of length $r_n - l_n$. Moreover, we need a concentration condition (2.30) on the probability measure.

Definition 2.2 *Let f be a cluster functional. For each $r \in \mathbb{N}$, we denote $D^r(f)$ the set of discontinuities of $f|_{E^r}$ and $C^r(f) := E^r \setminus D^r(f)$. Let $\{\mathcal{C}_k^r\}_{k\in\Lambda(r)}$ be the family of distinct connected components of $C^r(f)$, where $\Lambda(r)$ denotes the index set of this family. We say then that f is **C-Lipschitz** if for each $r \in \mathbb{N}$, $D^r(f)$ is r-null Lebesgue, $\Lambda(r)$ is finite and f is Lipschitz on each component $\mathcal{C}_k^r$.*

Remark 2.2 (i) It is clear that if f is Lipschitz, then it is C-Lipschitz. Therefore the functionals defined in (2.4) and (2.5) are C-Lipschitz.
(ii) Let $\phi : E \longrightarrow \mathbb{R}$ be a function such that $\phi(0) = 0$ and let $D(\phi)$ be the set of discontinuities of ϕ. If $Card(D(\phi))$ is finite and ϕ is Lipschitz in each maximal connected component of $E \setminus D(\phi)$, then the functional defined in (2.3) is also C-Lipschitz. Similarly we can see that the functional defined in (2.6) is also C-Lipschitz.

We consider the following notations, used throughout the remainder of this paper. Let $x = (x_1, x_2, \ldots, x_r)$, we denote

$$x^{(l:k)} := \begin{cases} 0 & \text{if } r < l, \\ (x_l, \ldots, x_k) & \text{if } 1 \leqslant l \leqslant k \leqslant r, \\ (x_l, \ldots, x_r) & \text{if } k > r; \end{cases}$$

and $x^{(k)} := x^{(1:k)}$. Besides, since $\mathbb{M} = \{(X_{n,i})_{1 \leqslant i \leqslant n}\}_{n \in \mathbb{N}}$ is row-wise stationary, we can write Y_n to denote a "generic block" such that $Y_n \stackrel{\mathcal{D}}{=} Y_{n,1}$.

On the other hand, if $f \in \mathcal{F}$, we denote

$$\Delta_n(f) := f(Y_n) - f(Y_n^{(r_n - l_n)}), \tag{2.29}$$

where r_n is the length of the block Y_n and l_n is a positive integer such that $l_n < r_n$.

Proposition 2.1 *Suppose that $\mathcal{F}$ is a class of C-Lipschitz cluster functionals such that (C.1), (C.2) and the following convergence conditions*

(C.3) $\sqrt{Var\left(\Delta_n(f)\right)} = o\left(\frac{r_n^2}{n^2}\sqrt{nv_n}\right)$

(C.4) $\mathbb{E}^{1/2}\left(|f(Y_n^{(r_n-l_n)})|^2 \mathbb{1}_{\left\{|f(Y_n^{(r_n-l_n)})| > \sqrt{nv_n}\right\}}\right) = o\left(\frac{r_n^2}{n^2}\sqrt{nv_n}\right)$,

are satisfied, with $r_n, l_n \underset{n\to\infty}{\longrightarrow} \infty$ such that $l_n \ll r_n \ll v_n^{-1} \ll n$. Additionally, assume that the r.v's $(X_{n,i})_{1 \leqslant i \leqslant n}$ are such that there exist positive real constants C, α, ρ (ρ dependent of n) such that

$$\sup_{x \in E} \sup_{1 \leqslant i \leqslant n} \mathbb{P}\left(X_{n,i} \in B(x, \rho/2)\right) \leqslant C\rho^{\alpha}. \tag{2.30}$$

Then the fidis of the empirical process $(Z_n(f))_{f\in\mathcal{F}}$ of cluster functionals converge to the fidis of a Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c (defined in (C.2)), if the triangular array $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ satisfies at least one of the following weak-dependence cases:

(D.1) *θ-weakly dependent such that $\theta_n(k) = \mathcal{O}\left(k^{-\theta}\right)$ for some $\theta > 1$ and $l_n^{-\theta} = o\left(n^{-1}\right)$,*

(D.2) *η-weakly dependent such that $\eta_n(k) = \mathcal{O}\left(k^{-\eta}\right)$ for some $\eta > 0$ and $l_n^{-\eta} = o\left(r_n/n^2\right)$,*

(D.3) *κ-weakly dependent such that $\kappa_n(k) = \mathcal{O}\left(k^{-\kappa}\right)$ for some $\kappa > 0$ and $l_n^{-\kappa} = o\left(\frac{1}{r_n^u n^v} \wedge \frac{1}{n^2}\right)$, for some $u \geqslant 0$ and $v \in [0,2]$,*

(D.4) *λ-weakly dependent such that $\lambda_n(k) = \mathcal{O}\left(k^{-\lambda}\right)$ for some $\lambda > 0$ and $l_n^{-\lambda} = o\left(\frac{1}{r_n^u n^v} \wedge \frac{1}{n^2}\right)$, for some $u \geqslant 0$ and $v \in [0,2]$.*

Generally, the convergence (C.2) can be easily verified. Anyway, the limit c defined in (C.2) is not always explicit. However, through the following proposition (which is a similar result to Theorems 1 and 3 of Segers [2003]) we provide sufficient conditions to verify (C.2), and which in some situations are easier to prove. Besides, in this way we can give an alternative expression to the covariance function c, as it is shown below in Corollary 2.1.

In order to carry this out, it is necessary to consider the following assumption:

(TC) There is a sequence $W = (W_i)_{i\geqslant 1}$ of E-valued random variables such that, for all $k \in \mathbb{N}$, the joint conditional distribution

$$P_{(X_{n,i},\mathbb{1}_{\{0\}}(X_{n,i}))_{1\leqslant i\leqslant k}|X_{n,1}\neq 0}$$

converges weakly to $P_{(W_i,\mathbb{1}_{\{0\}}(W_i))_{1\leqslant i\leqslant k}}$. Moreover, for all $f \in \mathcal{F}$ and all $k \geqslant 2$,

$$\mathbb{P}(W^{(2:k)} \in D^{k-1}(f), W_i = 0,\ \forall i > k) = \mathbb{P}(W^{(k)} \in D^k(f), W_i = 0,\ \forall i > k) = 0,$$

where $W^{(k)} = (W_1, \ldots, W_k)$ and $W^{(2:k)} = (W_2, \ldots, W_k)$ for all k.

Remark 2.3 The existence of such sequence W is guaranteed in particular from Theorem 2 in [Segers, 2003] with $E = [0, \infty)$ and the normalisation (2.8). There, the author has shown that if

$$P_{((X_{n,i})_{1\leqslant i\leqslant k}|X_1>u_n)} \underset{n\to\infty}{\longrightarrow} -\log G_k,$$

where G_k is some $k-$dimensional extreme value distribution for all $k \in \mathbb{N}$, then there exists a "*tail chain*" $(U_i)_{i\in\mathbb{N}}$ such that $W_i = \max\{U_i, 0\}$ and

$$P_{((X_{n,i},\mathbb{1}_{\{0\}}(X_{n,i}))_{1\leqslant i\leqslant k}|X_1>u_n)} \underset{n\to\infty}{\longrightarrow} P_{(W_i,\mathbb{1}_{\{0\}}(W_i))_{1\leqslant i\leqslant k}} \tag{2.31}$$

for all $k \in \mathbb{N}$.

Proposition 2.2 *Suppose that the triangular array $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ satisfies at least one of the following weak-dependence conditions:*

(D.1′) θ-weakly dependent such that $\theta_n(l_n) = o\left(v_n^{p+1}\right)$

(D.2′) η-weakly dependent such that $\eta_n(l_n) = o\left(v_n^{p+1}/r_n\right)$

(D.3′) κ-weakly dependent such that $\kappa_n(l_n) = o\left(v_n^{2p+1}/r_n\right)$

(D.4′) λ-weakly dependent such that $\lambda_n(l_n) = o\left(v_n^{2p+1}/r_n\right)$,

for some $p > 0$, where (r_n) and (l_n) are positive integer sequences such that $l_n \ll r_n \ll v_n^{-1} \ll n$, with $l_n \underset{n\to\infty}{\longrightarrow} \infty$. Then, if $\mathcal{F}$ is uniformly bounded,

$$\mathbb{E}\left(f(Y_n)|Y_n \neq 0\right) = \theta_n^{-1}\mathbb{E}\left(f(Y_{n,1}) - f(Y_{n,1}^{(2:r_n)})|X_{n,1} \neq 0\right) + o(1), \quad \forall f \in \mathcal{F},$$

and

$$\theta_n := \frac{\mathbb{P}(Y_n \neq 0)}{r_n v_n} = \mathbb{P}(Y_{n,1}^{(2:r_n)} = 0|X_{n,1} \neq 0)(1 + o(1)).$$

Additionally, if the assumption (TC) is satisfied, then

$$\begin{array}{rcl} m_W & := & \sup\{i \geqslant 1 : W_i \neq 0\} < \infty, \\ \theta_n & \underset{n\to\infty}{\longrightarrow} & \theta := \mathbb{P}(W_i = 0, \forall i \geqslant 2) = \mathbb{P}(m_W = 1) > 0, \\ P_{f(Y_n)|Y_n\neq 0} & \underset{n\to\infty}{\longrightarrow} & \dfrac{1}{\theta}\left(\mathbb{P}(f(W) \in \cdot\ ) - \mathbb{P}(f(W^{(2:\infty)}) \in \cdot\ ,\ m_W \geqslant 2)\right). \end{array}$$

Corollary 2.1 *Suppose that the hypotheses from Proposition 2.1 are maintained with the failure set $A = E \setminus \{0\}$ and that the assumption (TC) is satisfied. If, additionally, for each case of weak dependence θ, η, κ and λ we request that $n^{-1}v_n^{-1-p} = \mathcal{O}(1)$, $(r_n n^{-1})^2 = \mathcal{O}(v_n^{p+1})$, $(r_n^{1-u}n^{-v} \wedge r_n n^{-2}) = \mathcal{O}(v_n^{2p+1})$ and $(r_n^{1-u}n^{-v} \wedge r_n n^{-2}) = \mathcal{O}(v_n^{2p+1})$ be fulfilled respectively for some $p > 0$, then the fidis of the empirical process $(Z_n(f))_{f\in\mathcal{F}}$ of cluster functionals converge to the fidis of a centred Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c defined by*

$$c(f,g) = \mathbb{E}\left((fg)(W) - (fg)(W^{(2:\infty)})\right). \tag{2.32}$$

2.4 Application: the extremogram

In this section, we provide an application where it suffices to consider the convergence of the finite-dimensional marginal distributions of empirical processes of cluster functionals. Specifically, we prove that under suitable distributional conditions, the extremogram estimator for weakly dependent time series is asymptotically normal.

2.4.1 The extremogram

For a strictly stationary $\mathbb{R}^d$-valued time series $(X_t)_{t\in\mathbb{Z}}$, Davis & Mikosch [2009] have defined the **extremogram** for two sets A and B, both bounded away from zero[(3)], by

$$\rho_{A,B}(h) := \lim_{x\to\infty} \mathbb{P}(x^{-1}X_h \in B | x^{-1}X_0 \in A), \qquad h = 0, 1, 2, \dots, \tag{2.33}$$

provided that the limit exists.

According to Davis & Mikosch [2009], a "natural" estimator of the extremogram based on the observations $X_1, \dots, X_n$ is:

$$\widehat{\rho}_{A,B,n}(h) := \frac{\sum_{i=1}^{n-h} \mathbb{1}_{\{u_n^{-1}X_{i+h}\in B, u_n^{-1}X_i\in A\}}}{\sum_{i=1}^{n} \mathbb{1}_{\{u_n^{-1}X_i\in A\}}}, \tag{2.34}$$

where u_n is a high quantile of the process which replaces x in the limit (2.33). Of course, the choice of such a sequence of quantiles $(u_n)_{n\in\mathbb{N}}$ is not arbitrary. In particular, a sufficient condition for the existence of the limit (2.34) for any two sets A and B, bounded away from zero, is the following convergence:

$$n\mathbb{P}(u_n^{-1}(X_1, \dots, X_h) \in \cdot\ ) \xrightarrow[n\to\infty]{vague} \mu_h(\cdot), \tag{2.35}$$

for each $h \geqslant 1$, where $(\mu_h)_{h\in\mathbb{N}}$ is a sequence of non-null Radon measures on the Borel σ-field of $\mathbb{R}^{dh} \setminus \{0\}$.
Besides, $v_n = \mathbb{P}(u_n^{-1}X_0 \in A) \xrightarrow[n\to\infty]{} 0$ with $nv_n \xrightarrow[n\to\infty]{} \infty$, in order to have consistency in the results.

Let us now define the pre-asymptotic extremogram (PA-extremogram) $\rho_{A,B,n}(h) := \mathbb{P}(u_n^{-1}X_h \in B | u_n^{-1}X_0 \in A)$, and let l be a positive integer. Then, under suitable conditions of convergence and weak-dependence,

$$\sqrt{nv_n}\left(\widehat{\rho}_{A,B,n}(h) - \rho_{A,B,n}(h)\right)_{0\leqslant h\leqslant l} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \Sigma_{A,B}\right), \tag{2.36}$$

[3]A set S is bounded away from zero if $S \subset \{y : |y| > r\}$ for some $r > 0$

where $\Sigma_{A,B}$ is defined below in (2.46).

Indeed, if for each $h \in \{1, \ldots, l\}$ with $l < r$, we define the cluster functional $f_{A,B,h} : (\mathbb{R}^d_{\cup}, \mathcal{R}_{\cup}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$f_{A,B,h}(x_1, \ldots, x_r) := \sum_{i=1}^{r-h} \mathbb{1}_{A\times B}(x_i, x_{i+h}), \tag{2.37}$$

then, by using the normalisation (2.9), we can rewrite the estimator (2.34) as:

$$\widehat{\rho}_{A,B,n}(h) = \frac{\sqrt{nv_n} Z_n(f_{A,B,h}) + m_n \mathbb{E} f_{A,B,h}(Y_{n,1}) + \sum_{j=1}^{m_n} \delta_{n,j}(f_{A,B,h}) + R_n(A,B,h)}{\sqrt{nv_n} Z_n(f_{A,A,0}) + m_n \mathbb{E} f_{A,A,0}(Y_{n,1}) + R_n(A,A,0)} \tag{2.38}$$

where

$$\delta_{n,j}(f_{A,B,h}) := \sum_{i=jr_n-h+1}^{jr_n} \mathbb{1}_{\{u_n^{-1}X_i \in A, u_n^{-1}X_{i+h}\in B\}} \tag{2.39}$$

$$R_n(A,B,h) := \sum_{i=m_n r_n+1}^{n-h} \mathbb{1}_{\{u_n^{-1}X_i \in A, u_n^{-1}X_{i+h}\in B\}}. \tag{2.40}$$

On the other hand, if $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ satisfies at least one of the following weak-dependence conditions:

(D.1″) θ-weakly dependent such that $\sum_{k=1}^{m_n-1}(1-km_n^{-1})\theta_n(l_n+(k-1)r_n) = o(nr_n^{-1}v_n^{p+2})$

(D.2″) η-weakly dependent such that $\sum_{k=1}^{m_n-1}(1-km_n^{-1})\eta_n(l_n+(k-1)r_n) = o(nr_n^{-1}v_n^{p+2})$

(D.3″) κ-weakly dependent such that $\sum_{k=1}^{m_n-1}(1-km_n^{-1})\kappa_n(l_n+(k-1)r_n) = o(nr_n^{-3}v_n^{p+2})$

(D.4″) λ-weakly dependent such that $\sum_{k=1}^{m_n-1}(1-km_n^{-1})\lambda_n(l_n+(k-1)r_n) = o(nr_n^{-3}v_n^{p+2})$,

for some $p > 0$, and

$$\sum_{k=1}^{r_n-1}\left(1-\frac{k}{r_n}\right)\rho_{A,A,n}(k) = o(nv_n) \tag{2.41}$$

where $l_n \ll r_n \ll v_n^{-1} \ll n$ and $\sqrt{nv_n} = o(r_n)$, with $l_n \underset{n\to\infty}{\longrightarrow} \infty$.
Then,

$$\sqrt{nv_n}\left(\widehat{\rho}_{A,B,n}(h) - \rho_{A,B,n}(h)\right) = Z_n(f_{A,B,h}) - \rho_{A,B,n}(h) Z_n(f_{A,A,0}) + o_P(1). \tag{2.42}$$

Now, based on the equality (2.42), we shall formalise (2.36) through the following result:

Proposition 2.3 *Let $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ be the normalised random variables defined in (2.9), built from a strictly stationary regularly varying sequence $(X_i)_{i\in\mathbb{N}}$ of $\mathbb{R}^d$-valued random vectors. Suppose that the big and small block sizes r_n and l_n are such that $l_n \ll r_n \ll v_n^{-1} \ll n$ and $\sqrt{nv_n} = o(r_n)$, with $l_n \underset{n\to\infty}{\longrightarrow} \infty$. Then, if $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ satisfies:*

(i) at least one of the weak-dependence conditions of the list (D.1″) - (D.4″);

(ii) the concentration condition (2.30) and the relation (2.41);

(iii) the convergence conditions: (C.1), (C.3) and (C.4) with $f = f_{A,B,h}$; and

(iv) if there exist the covariance functions $\sigma_{A,B}$ and $\sigma'_{A,B}$ such that:

$$\frac{1}{r_n v_n}\sum_{i=1}^{r_n-h}\sum_{j=1}^{r_n-h'}\mathbb{P}\left(X_{n,i}, X_{n,j}\in A; X_{n,i+h}, X_{n,j+h'}\in B\right) \underset{n\to\infty}{\longrightarrow} \sigma_{A,B}(h,h') \tag{2.43}$$

$$\frac{1}{r_n v_n}\sum_{i=1}^{r_n-h}\sum_{j=1}^{r_n}\mathbb{P}\left(X_{n,i}, X_{n,j}\in A; X_{n,i+h}\in B\right) \underset{n\to\infty}{\longrightarrow} \sigma'_{A,B}(h). \tag{2.44}$$

Then,

$$\sqrt{nv_n}\left(\widehat{\rho}_{A,B,n}(h) - \rho_{A,B,n}(h)\right)_{0\leqslant h\leqslant l} \underset{n\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \mathcal{N}\left(0,\Sigma_{A,B}\right), \tag{2.45}$$

where the covariance matrix $\Sigma_{A,B}$ is defined by

$$\left[\sigma_{A,B}(h,h') - \rho_{A,B}(h')\sigma'_{A,B}(h) - \rho_{A,B}(h)\sigma'_{A,B}(h') + \rho_{A,B}(h)\rho_{A,B}(h')\sigma'_{A,A}(0)\right]_{0\leqslant h,h'\leqslant l}. \tag{2.46}$$

Remark 2.4 (i) Condition (2.35) is fulfilled by the regularly varying processes with index $\alpha > 0$. For details on the interpretation of the structure of regularly varying sequences, see [Basrak & Segers, 2009].

(ii) Under suitable α-mixing conditions, Davis & Mikosch [2009] proved the convergence (2.36).

(iii) Note that the regularly varying condition is necessary to ensure the existence of the PA-extremogram for all A, B bounded away from zero. However, if the regularly varying condition is replaced in Proposition 2.3 by the existence of $\rho_{A_0,B_0}(h)$ (for all $h\in\mathbb{N}$), for some A_0 and B_0 bounded away from zero, and if the assumptions of Proposition 2.3 hold for $A = A_0$ and $B = B_0$, then we also obtain the convergence (2.45) with the covariance matrix (2.46) for $A = A_0$ and $B = B_0$.

2.4.2 Simulation study

In order to assess our results numerically, we perform a numerical simulation of a real-valued weakly dependent data, estimate its extremogram by means of (2.34) and compare it with the true pre-asymptotic extremogram $\rho_{A,B,n}(\cdot)$.

Theoretic model

Let us consider the AR(1)-process (2.1) given in the introduction. Since, as X_0 is uniformly distributed on $[0,1]$ and $X_i = b^{-i}X_0 + \sum_{s=1}^{i} b^{s-i-1}\xi_s$ for all $i \geqslant 1$, then for $A = B = (1,\infty)$ we obtain:

$$\mathbb{P}\left(\frac{X_h}{y} \in B \,\middle|\, \frac{X_0}{y} \in A\right) = \frac{1}{b^h} \sum_{j_1,\dots,j_h \in U(b)} \min\left\{1, \frac{1}{1-y}\left(1 - yb^h + \sum_{s=1}^{h} \frac{j_s}{b^{1-s}}\right)_+\right\}, \tag{2.47}$$

where $y = 1 - 1/x$. Therefore,

$$\rho_{A,B}(h) = b^{-h} \qquad \text{for } h = 0, 1, \dots; \text{ and} \tag{2.48}$$

$$\rho_{A,B,n}(h) = \frac{1}{b^h} \sum_{j_1,\dots,j_h \in U(b)} \min\left\{1, \frac{1}{v_n}\left(1 - (1-v_n)b^h + \sum_{s=1}^{h} \frac{j_s}{b^{1-s}}\right)_+\right\}, \tag{2.49}$$

for $h < n$, where $v_n := \mathbb{P}(X_0/u_n \in A) = 1 - u_n$.

Remark 2.5 In this example, the conditions (2.30) and (2.41) are trivially satisfied. On the other hand, consider the functional defined in (2.37). Then, the condition (C.3) is satisfied whenever $\sqrt{l_n/n} = o\left((r_n/n)^2\right)$. Moreover,

$$(r_n v_n)^{-1}\mathbb{E} f^3_{A,B,h}(Y_n) \underset{n\to\infty}{\longrightarrow} C_h b^{-h} \tag{2.50}$$

for some constant $C_h > 0$, which means that the condition (C.1) is met. Using (2.50) and considering $n^{3/2} = o(r_n^{3/2}(nv_n)^{\delta/4})$ for some $\delta \in (0,1]$ we verify (C.4). Finally, we show also that

$$\sigma_{A,B}(h,h') = C_{h,h'}(b^{-h} + b^{-h'}) + b^{-(h\vee h')}, \qquad \text{for } h, h' \geqslant 0, \tag{2.51}$$

$$\sigma'_{A,B}(h) = \rho_{A,B}(h) + 2C'_h + \frac{2}{b^h}\left(h + 2 - \frac{b^h - 1}{b-1}\right), \qquad \text{for } h \geqslant 0, \tag{2.52}$$

where $C_{h,h'}$ and C'_h are positive constants such that

$$C_{h,h'} := \lim_{n\to\infty} \sum_{k=1}^{r_n - h\vee h' -1} \left(1 - \frac{h \vee h' + k}{r_n}\right) b^{-k},$$
$$C'_h := \lim_{n\to\infty} \sum_{k=1}^{r_n - h -1} \left(1 - \frac{h + k}{r_n}\right) b^{-k}.$$

Remark 2.6 Observe that the equivalence (2.48) proves that the family of non-mixing processes such that the extremogram (2.33) exists for some A, B bounded away from zero, is not empty.

Experiment

In order to carry out the experiment, we have first generated from the AR(1)-process defined in (2.1) with $b = 2$, a dataset of $n = 10^4$. Besides, we have chosen $v_n = n^{-0.34}$, the probability of to obtain an excess over the high quantile $u_n = 1 - v_n$, assuming that we have previously determined that the distribution of the X_i is uniformly distributed on $[0, 1]$.

The true PA-extremogram (2.49) evaluated at the right tail $(A = B = (1, \infty))$ and its estimation through the estimator (2.34), for lags $h = 1 \ldots, 10$, are displayed respectively by the black solid line and the blue solid line, on the left side of Figure 2.1. Additionally, due to the simplicity of the AR(1)-process in this study (for the general case, see Remark 2.8), we obtain an explicit expression for the pre-asymptotic variance function:

$$\begin{aligned}\sigma^2_{A,B,n}(h) := \frac{\mathrm{Var}(f_{A,B,h}(Y_n))}{r_n v_n} &= \left(1 - \frac{h}{r_n} - r_n v_n \rho_{A,B,n}(h)\right) \rho_{A,B,n}(h) \\ &+ \frac{2}{b^h} \sum_{k=1}^{r_n-h-1} \left(1 - \frac{h+k}{r_n}\right) \frac{1}{b^k} \sum_{j_1,\ldots,j_{k+h} \in U(b)} \min\left\{1, b^k + \frac{\mu_b(j_1,\ldots,j_k)}{v_n},\right. \\ &\left. b^h + \frac{\mu_b(j_1,\ldots,j_h)}{v_n}, b^{k+h} + \frac{\mu_b(j_1,\ldots,j_{k+h})}{v_n}\right\}_+ \end{aligned} \tag{2.53}$$

with $\mu_b(j_1, \ldots, j_s) := 1 - b^s + \sum_{i=1}^{s} b^{i-1} j_i$. Notice that $\sigma^2_{A,B,n}(h) \underset{n\to\infty}{\longrightarrow} \sigma_{A,B}(h,h)$ for $h = 1, 2, \ldots$. We then use the expression (2.53) in order to provide an estimator $\widehat{\sigma}_{A,B,n}$ for the variance by simply replacing the PA-extremogram estimator for the PA-extremogram in (2.53), *i.e.*

$$\widehat{\sigma}^2_{A,B,n}(h) := \left(1 - \frac{h}{r_n} - r_n v_n \widehat{\rho}_{A,B,n}(h)\right) \widehat{\rho}_{A,B,n}(h) + \frac{2}{b^h} \sum_{k=1}^{r_n-h-1} \left(1 - \frac{h+k}{r_n}\right) \frac{1}{b^k} \sum_{j_1,\ldots,j_{k+h} \in U(b)} \min \left\{1, b^k + \frac{\mu_b(j_1,\ldots,j_k)}{v_n}, b^h + \frac{\mu_b(j_1,\ldots,j_h)}{v_n}, b^{k+h} + \frac{\mu_b(j_1,\ldots,j_{k+h})}{v_n}\right\}_+, \qquad h = 1,2,\ldots. \tag{2.54}$$

We thus obtain $(1-\alpha)$ two-sided confidence interval estimator $I(h) = (I_L(h), I_U(h))$ for $\rho_{A,B,n}(h)$, for each $h = 1,\ldots,10$, such that

$$I_L(h) := \widehat{\rho}_{A,B,n}(h) - z_{\alpha/2}(nv_n)^{-1/2}\widehat{\sigma}_{A,B,n}(h)$$
$$I_U(h) := \widehat{\rho}_{A,B,n}(h) + z_{\alpha/2}(nv_n)^{-1/2}\widehat{\sigma}_{A,B,n}(h),$$

where z_α denotes the upper $100\alpha\%$ point of the standard normal distribution. In particular, we show $100(1-\alpha)\% = 95\%$ two-sided confidence bands (red dashed lines) with $r_n = \lceil n^{1/3} \rceil$ on the same left side of Figure 2.1. As was expected for the AR(1)-process studied here, observing the confidence bands, we note that the extremal dependence vanishes as the lag h increases.

Now, let's assess the normality of the vector

$$\widehat{\rho}_{A,B,n} := (\tilde{\rho}_{A,B,n}(1), \ldots, \tilde{\rho}_{A,B,n}(10)), \tag{2.55}$$

where its coordinates are defined by $\tilde{\rho}_{A,B,n}(h) := \sqrt{nv_n}(\widehat{\rho}_{A,B,n}(h) - \rho_{A,B,n}(h))$, with $h = 1,2,\ldots,10$.
For this, we simulate $N = 100$ samplings of $n = 10^4$ from the AR(1)-process (2.1) with $b = 2$, followed by their N estimated PA-extremograms (*i.e.* N samplings of the PA-extremogram estimator). On the right side in Figure 2.1, we show the box-plots of the coordinates of the vector $\widehat{\rho}_{A,B,n}$. The blue square points correspond to the empirical mean of the corresponding coordinate. Note that the symmetry of the box-plots shows the normality on each coordinate. On the other hand, in order to assess the normality of the vector $\widehat{\rho}_{A,B,n}$, we use three tests: Mardia, Henze-Zirkler and Royston test (for the details of these tests, see [Korkmaz et al., 2014; Mardia, 1970; Henze & Zirkler, 1990; Royston, 1992]), which yield the results shown on Table 2.1.

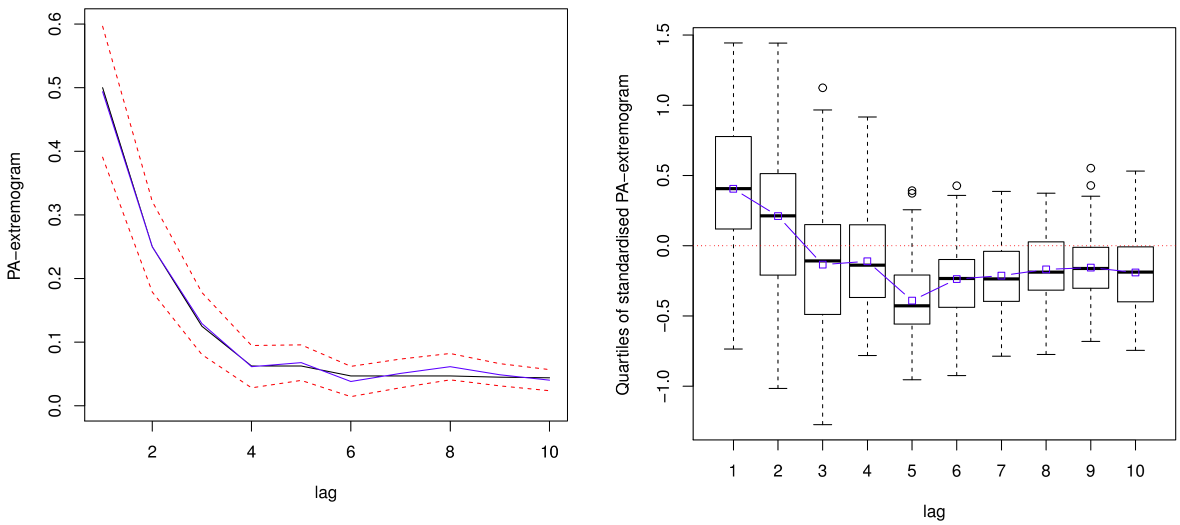


Figure 2.1: For the AR(1)-process (2.1) with $b = 2$. Left: 95% two-sided confidence bands (red dashed lines) for the PA-extremogram. The true and estimated PA-extremogram are shown respectively through the black solid line and blue solid line. Right: Box-plots of each coordinate of the PA-extremogram estimator vector $\widehat{\rho}_{A,B,n}$ defined in (2.55).

Test	Test Statistic	p-value
Mardia		
Skewness	218.9986	0.506387
Kurtosis	-0.812372	0.416578
Henze-Zirkler	0.982839	0.279119
Royston	5.846811	0.586721

Table 2.1: *Multivariate normality test results.*

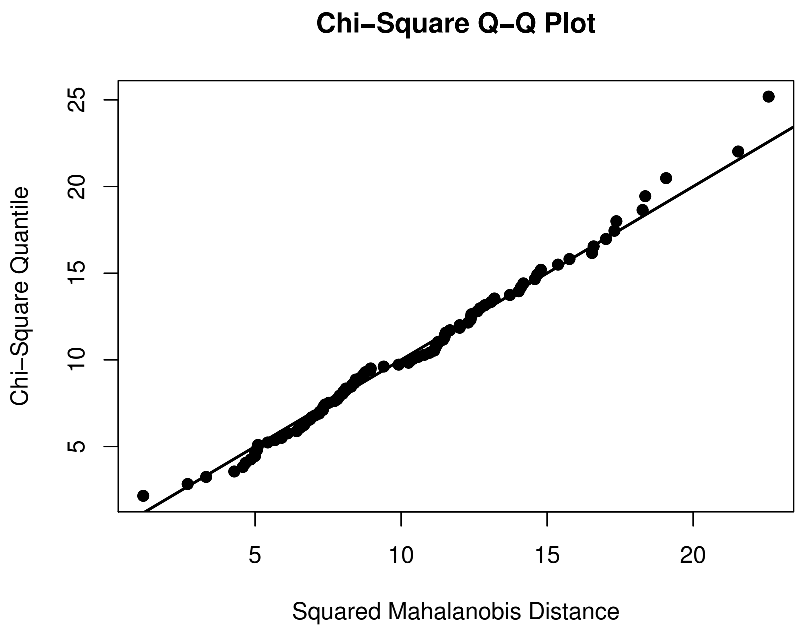


Figure 2.2: *Chi-Square Q-Q plot: Empirical quantiles of Squared Mahalanobis Distance from $\widehat{\rho}_{A,B,n}$ to $\vec{0}$ vs Chi-Square Quantiles.*

Moreover, Chi-Square Q-Q Plot for Squared Mahalanobis Distance from $\widehat{\rho}_{A,b,n}$ to $\vec{0}$ is displayed on Figure 2.2. We can see that the Q-Q Plots are, in fact, almost all on the straight line. Therefore, observing this and the p-values obtained in the three test of normality on Table 2.1, we can conclude that the vector $\widehat{\rho}_{A,B,n}$ has a gaussian behaviour, confirming our expectations from Proposition 2.3.

Remark 2.7 Notice that, given the result (2.48), if we assume that some data behaves as an AR(1)-input, then we can also estimate the b value of such AR(1)-input, from the PA-extremogram estimator through the relation:

$$\widehat{b} = \widehat{b}(h) = (\widehat{\rho}_{A,B,n}(h))^{-1/h}, \quad h = 0, 1, \ldots .$$

Remark 2.8 To conclude, observe that the asymptotic covariance $\sigma_{A,B}(h,h')$ is the infinite sum of unknown probabilities, which can not be written generally in an useful way in practice. Davis & Mikosch [2012] use bootstrap procedures to approximate the variance $\sigma^2_{A,B}(h) := \sigma_{A,B}(h,h)$ in order to construct "asymptotically correct" confidence bands for the PA-extremogram. Moreover, they prove bootstrap consistency under mixing conditions. Here we do not include this approach under weak-dependence, because it is not the aim of this work, however, this will be addressed in a forthcoming paper.

2.5 Proofs

Proof of Theorem 2.1. The proof is basically a direct application of Theorem 1 in [Bardet et al., 2007] to the random variables $(W_{n,j}(\mathbf{f}_k))_{1\leqslant j\leqslant m_n}$ defined in (2.14).

Indeed, notice that Assumption (C.1) implies that

$$B_n(\epsilon) := \sum_{j=1}^{m_n} \mathbb{E}\|W_{n,j}(\mathbf{f}_k)\|^2 \mathbb{1}_{\{\|W_{n,j}(\mathbf{f}_k)\|>\epsilon\}} \underset{n\to\infty}{\longrightarrow} 0,$$

for all k-tuple $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$, $k \in \mathbb{N}$ and $\epsilon > 0$. Note that this last statement is weaker than the assumption H_δ of Bardet et al. [2007]. However, the conclusion of Theorem 1 in [Bardet et al., 2007] continues to be fulfilled if we substitute the assumption H_δ for (C.1) (see Remark 1 in [Bardet et al., 2007]).

On the other hand, Assumption (C.2) ensures the existence of the positive matrix $\Sigma_k = (c(f_s, f_t))_{s,t=1,\ldots,k}$, because

$$\Sigma_{n,k} := \sum_{j=1}^{m_n} \left(\frac{\mathrm{Cov}(f_s(Y_{n,j}), f_t(Y_{n,j}))}{nv_n}\right)_{s,t=1,\ldots,k} \underset{n\to\infty}{\longrightarrow} \Sigma_k,$$

for all k-tuple $(f_1, \ldots, f_k) \in \mathcal{F}^k$, with $k \in \mathbb{N}$. The proof ends considering the dependence condition (2.15). □

Proof of Proposition 2.1. The proof of this proposition is based on Theorem 2.1. Therefore, we only have to prove that $T_{t,m_n}(\mathbf{f}_k) \underset{n\to\infty}{\longrightarrow} 0$, for all $t \in \mathbb{R}^k$ and all k-tuple of cluster functionals $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$, with $k \in \mathbb{N}$.

Indeed, for $j \in \{2, \ldots, m_n\}$ and $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$, notice that

$$\mathrm{Cov}\left(\exp(i\langle t, \sum_{s=1}^{j-1} W_{n,s}(\mathbf{f}_k)\rangle), \exp(i\langle t, W_{n,j}(\mathbf{f}_k)\rangle)\right)$$

can be rewritten as:

$$\mathrm{Cov}(F_j, G_j) := \mathrm{Cov}\left(F_{t,n}^{\mathbf{f}_k}(Y_{n,1}, \ldots, Y_{n,j-1}), G_{t,n}^{\mathbf{f}_k}(Y_{n,j})\right), \tag{2.56}$$

where

$$G_{t,n}^{\mathbf{f}_k}(s) = \exp(i\langle t, \sum_{l=1}^{k} \frac{f_l(s) - \mathbb{E}f_l(s)}{\sqrt{nv_n}} e_l\rangle)$$

and

$$F_{t,n}^{\mathbf{f}_k}(s_1,\ldots,s_j)=\prod_{h=1}^{j}G_{t,n}^{\mathbf{f}_k}(s_h),$$

with $e_1,\ldots,e_k$ denoting the canonical base in $\mathbb{R}^k$. Moreover, it is clear that $\|G_{t,n}^{\mathbf{f}_k}\|_\infty\leqslant 1$ and $\|F_{t,n}^{\mathbf{f}_k}\|_\infty\leqslant 1$, for all $t\in\mathbb{R}^k$ and all $\mathbf{f}_k=(f_1,\ldots,f_k)\in\mathcal{F}^k$, for any $k\geqslant 1$.

Then,

$$\begin{aligned}
\left|\mathrm{Cov}\left(F_j,G_j\right)\right| &= \left|\mathrm{Cov}\left(F_{t,n}^{\mathbf{f}_k}(Y_{n,1},\ldots,Y_{n,j-1}),G_{t,n}^{\mathbf{f}_k}(Y_{n,j})\right)\right|\\
&\leqslant \left|\mathrm{Cov}\left(F_{t,n}^{\mathbf{f}_k}(Y_{n,1},\ldots,Y_{n,j-1})-F_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,1}^{(r_n-l_n)},\ldots,Y_{n,j-1}^{(r_n-l_n)}),G_{t,n}^{\mathbf{f}_k}(Y_{n,j})\right)\right|\\
&\quad+\left|\mathrm{Cov}\left(F_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,1}^{(r_n-l_n)},\ldots,Y_{n,j-1}^{(r_n-l_n)}),G_{t,n}^{\mathbf{f}_k}(Y_{n,j})-G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,j}^{(r_n-l_n)})\right)\right|\\
&\quad+\left|\mathrm{Cov}\left(F_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,1}^{(r_n-l_n)},\ldots,Y_{n,j-1}^{(r_n-l_n)}),G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,j}^{(r_n-l_n)})\right)\right|\\
&\leqslant 2\mathbb{E}\left|F_{t,n}^{\mathbf{f}_k}(Y_{n,1},\ldots,Y_{n,j-1})-F_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,1}^{(r_n-l_n)},\ldots,Y_{n,j-1}^{(r_n-l_n)})\right| &(2.57)\\
&\quad+2\mathbb{E}\left|G_{t,n}^{\mathbf{f}_k}(Y_{n,j})-G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,j}^{(r_n-l_n)})\right| &(2.58)\\
&\quad+\left|\mathrm{Cov}\left(F_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,1}^{(r_n-l_n)},\ldots,Y_{n,j-1}^{(r_n-l_n)}),G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,j}^{(r_n-l_n)})\right)\right|, &(2.59)
\end{aligned}$$

where $\mathbf{f}_{k,\rho}^{[T]}=(f_{1,\rho}^{[T]},\ldots,f_{k,\rho}^{[T]})$ is a k-tuple of Lipschitz cluster functionals which approximates $\mathbf{f}_k$ as $\rho\downarrow 0$ and as $T\uparrow\infty$, defined as follows:

Let $f:E^r\longrightarrow\mathbb{R}$ be a C-Lipschitz cluster functional and we denote $f^{[T]}=f\vee(-T)\wedge T$ its truncation by T, for $T>0$. Consider the set of discontinuities of $f^{[T]}$, $D^r(f^{[T]})$. Now, for each $\rho>0$, let $D_\rho^r=\bigcup\{B(y,\rho/2):y\in D^r(f^{[T]})\}$ be a open covering of $D^r(f^{[T]})$ and denote $C_\rho^r:=E^r\setminus D_\rho^r$. If $\{\mathcal{C}_k^r\}_{k\in\Lambda(r)}$ is the family of distinct connected components of $C^r(f)$ and $L(f):=\max_{k\in\Lambda(r)}\mathrm{Lip}(f|_{\mathcal{C}_k^r})$, then for each $\rho\in\left(0,\frac{2T}{L(f)}\right)$, observe that $f^{[T]}\Big|_{C_\rho^r}$ is a Lipschitz functional such that

$$\mathrm{Lip}(f^{[T]}\Big|_{C_\rho^r})\leqslant\frac{2T}{\rho}.$$

Using Kirszbraun's theorem, the functional $f^{[T]}\Big|_{C_\rho} : C_\rho \longrightarrow \mathbb{R}$ has a Lipschitz extension $g_\rho : E^r \longrightarrow \mathbb{R}$ such that $\mathrm{Lip}(g_\rho) = \mathrm{Lip}(f^{[T]}\Big|_{C_\rho^r})$. Moreover, this functional is defined by

$$g_\rho(x_1,\dots,x_r) = \inf_{(y_1,\dots,y_r)\in C_\rho^r} \left\{ f|_{C_\rho^r}(y_1,\dots,y_r) + \mathrm{Lip}(f|_{C_\rho^r}) \sum_{i=1}^{r} \|x_i - y_i\| \right\}. \tag{2.60}$$

Thus we choose $f_\rho^{[T]} = g_\rho$.

In this way, we obtain bounds for $\mathrm{Lip}G_{t,n}^{\mathbf{f}_k}$ and $\mathrm{Lip}F_{t,n}^{\mathbf{f}_k}$:

$$\mathrm{Lip}G_{t,n}^{\mathbf{f}_k} \leqslant \frac{2\|t\|}{\sqrt{nv_n}} \sqrt{\sum_{l=1}^{k} \left(\mathrm{Lip}f_{l,\rho}^{[T]}\right)^2} \leqslant \frac{4\|t\|\sqrt{k}T}{\rho\sqrt{nv_n}} \quad \text{and} \quad \mathrm{Lip}F_{t,n}^{\mathbf{f}_k} \leqslant \mathrm{Lip}G_{t,n}^{\mathbf{f}_k}. \tag{2.61}$$

Let us now denote by

$$C(F,G) := \sum_{j=1}^{m_n} \left| \mathrm{Cov}\left(F_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,1}^{(r_n-l_n)},\dots,Y_{n,j-1}^{(r_n-l_n)}), G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,j}^{(r_n-l_n)}) \right) \right|,$$

the sum over $j = 1,\dots,m_n$ of the term (2.59). Then, combining (2.61) with the definition of the weak-dependence coefficients (2.17) - (2.20), we obtain bounds for $C(F,G)$ according to the respective condition of weak dependence assumed for $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$:

1. $\theta-$WD implies $C(F,G) \leqslant \dfrac{4T\|t\|\sqrt{k}}{\rho}\dfrac{n}{\sqrt{nv_n}}\theta_n(l_n)$,
2. $\eta-$WD implies $C(F,G) \leqslant \dfrac{2T\|t\|\sqrt{k}}{\rho}\dfrac{n^2}{r_n\sqrt{nv_n}}\left(1+\dfrac{r_n}{n}\right)\eta_n(l_n)$,
3. $\kappa-$WD implies $C(F,G) \leqslant \dfrac{8T^2\|t\|^2k}{\rho^2}\dfrac{n}{v_n}\kappa_n(l_n)$,
4. $\lambda-$WD implies $C(F,G) \leqslant \left(\dfrac{2T\|t\|\sqrt{k}}{\rho}\dfrac{n^2}{r_n\sqrt{nv_n}}\left(1+\dfrac{r_n}{n}\right) + \dfrac{8T^2\|t\|^2k}{\rho^2}\dfrac{n}{v_n}\right)\lambda_n(l_n)$.

On the other hand, notice that:

$$|F_{t,n}^{f}(s_1,\dots,s_p) - F_{t,n}^{f'}(s_1',\dots,s_p')| \leqslant \sum_{i=1}^{p} |G_{t,n}^{f}(s_i) - G_{t,n}^{f}(s_i')| + \sum_{i=1}^{p} |G_{t,n}^{f}(s_i') - G_{t,n}^{f'}(s_i')|.$$

Thus, the terms (2.57) and (2.58) can be bounded as follows:

$$
\begin{aligned}
2\mathbb{E}&\left|F_{t,n}^{\mathbf{f}_k}(Y_{n,1},\ldots,Y_{n,j-1}) - F_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,1}^{(r_n-l_n)},\ldots,Y_{n,j-1}^{(r_n-l_n)})\right| \\
&\qquad + 2\mathbb{E}\left|G_{t,n}^{\mathbf{f}_k}(Y_{n,j}) - G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,j}^{(r_n-l_n)})\right| \\
&\leqslant 2\mathbb{E}\left(\sum_{i=1}^{j-1}|G_{t,n}^{\mathbf{f}_k}(Y_{n,i}) - G_{t,n}^{\mathbf{f}_k}(Y_{n,i}^{(r_n-l_n)})| + \sum_{i=1}^{j-1}|G_{t,n}^{\mathbf{f}_k}(Y_{n,i}^{(r_n-l_n)}) - G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,i}^{(r_n-l_n)})|\right) \\
&\quad + 2\mathbb{E}\left(|G_{t,n}^{\mathbf{f}_k}(Y_{n,j}) - G_{t,n}^{\mathbf{f}_k}(Y_{n,j}^{(r_n-l_n)})| + |G_{t,n}^{\mathbf{f}_k}(Y_{n,i}^{(r_n-l_n)}) - G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,i}^{(r_n-l_n)})|\right) \\
&= 2\sum_{i=1}^{j}\mathbb{E}|G_{t,n}^{\mathbf{f}_k}(Y_{n,i}) - G_{t,n}^{\mathbf{f}_k}(Y_{n,i}^{(r_n-l_n)})| + 2\sum_{i=1}^{j}\mathbb{E}|G_{t,n}^{\mathbf{f}_k}(Y_{n,i}^{(r_n-l_n)}) - G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_{n,i}^{(r_n-l_n)})| \\
&= 2j\mathbb{E}|G_{t,n}^{\mathbf{f}_k}(Y_n) - G_{t,n}^{\mathbf{f}_k}(Y_n^{(r_n-l_n)})| + 2j\mathbb{E}|G_{t,n}^{\mathbf{f}_k}(Y_n^{(r_n-l_n)}) - G_{t,n}^{\mathbf{f}_{k,\rho}^{[T]}}(Y_n^{(r_n-l_n)})| \\
&= 2j\|t\|\sqrt{\sum_{l=1}^{k}\frac{\mathrm{Var}(\Delta_n(f_l))}{nv_n}} + 2j\|t\|\sqrt{\sum_{l=1}^{k}\frac{\mathrm{Var}(f_l(Y_n^{(r_n-l_n)}) - f_{l,\rho}^{[T]}(Y_n^{(r_n-l_n)}))}{nv_n}} \\
&\qquad\qquad \leqslant \frac{2j\|t\|}{\sqrt{nv_n}}\left(\sqrt{\sum_{l=1}^{k}\mathrm{Var}(\Delta_n(f_l))}\right. \\
&\qquad\qquad\quad \left. + \sqrt{\sum_{l=1}^{k}\mathbb{E}\left(|f_l(Y_n^{(r_n-l_n)})|^2\mathbb{1}_{\{|f_l(Y_n^{(r_n-l_n)})|>T\}}\right) + 4T^2k(C\rho^{\alpha})^{(r_n-l_n)\delta}}\right),
\end{aligned}
$$

for some $\delta \in (0,1)$.

Using all of the above, we obtain that

$$
\begin{aligned}
T_{,m_n}(\mathbf{f}_k) &= \sum_{j=1}^{m_n}|\mathrm{Cov}(F_j,G_j)| \\
&\leqslant \|t\|\sqrt{k}\left(1+\frac{r_n}{n}\right)\frac{n^2}{r_n^2\sqrt{nv_n}}\left(\sqrt{\mathrm{Var}(\Delta_n(f))}\right. && (2.62) \\
&\quad \left. + \sqrt{\mathbb{E}|f(Y_n^{(r_n-l_n)})|^2\mathbb{1}_{\{|f(Y_n^{(r_n-l_n)})|>T\}}} + 2T(C\rho^{\alpha})^{(r_n-l_n)\delta/2}\right) && (2.63) \\
&\quad + C(F,G). && (2.64)
\end{aligned}
$$

Finally it suffices to choose T such that $T = \mathcal{O}(\sqrt{nv_n})$ with

$$1)\ \rho = \left(\frac{r_n^6}{n^4}\right)^{\frac{1}{2+\alpha\delta(r_n-l_n)}} \frac{(\theta_n(l_n))^{\frac{2}{2+\alpha\delta(r_n-l_n)}}}{C^{(\alpha+\frac{2}{\delta(r_n-l_n)})^{-1}}},$$

$$2)\ \rho = \left(\frac{r_n}{n}\right)^{\frac{4}{2+\alpha\delta(r_n-l_n)}} \frac{(\eta_n(l_n))^{\frac{2}{2+\alpha\delta(r_n-l_n)}}}{C^{(\alpha+\frac{2}{\delta(r_n-l_n)})^{-1}}},$$

$$3)\ \rho = \left(\frac{r_n^{2+u}}{n^{2-v}}\right)^{\frac{2}{4+\alpha\delta(r_n-l_n)}} \frac{(\kappa_n(l_n))^{\frac{2}{4+\alpha\delta(r_n-l_n)}}}{C^{(\alpha+\frac{4}{\delta(r_n-l_n)})^{-1}}},$$

$$4)\ \rho = \left(\frac{r_n^{2+u}}{n^{2-v}}\right)^{\frac{2}{4+\alpha\delta(r_n-l_n)}} \frac{(\lambda_n(l_n))^{\frac{2}{4+\alpha\delta(r_n-l_n)}}}{C^{(\alpha+\frac{4}{\delta(r_n-l_n)})^{-1}}},$$

for each respective weak-dependence condition and take into account the assumptions (C.3) and (C.4) to obtain the CLT. □

***Proof of Proposition* 2.2.** Suffices to prove the following multidimensional version of Condition (6) in Segers [2003]:

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(Y_n^{(l+1:r_n)} \neq 0 \,\middle|\, X_{n,1} \neq 0\right) = 0, \tag{2.65}$$

since the rest of the proof follows the same steps of the proof of Lemma 2.5 in Drees & Rootzén [2010].

Indeed, let $f_n : E^{r_n-l} \longrightarrow [0,1]$ and $g : E \longrightarrow [0,1]$ be functions defined by $f_n(x) = 1 - \mathbb{1}_{\{0_{r_n-l}\}}(x)$ and $g(x) = 1 - \mathbb{1}_{\{0\}}(x)$. Similar to the functions defined in (2.60), we can build increasing sequences of Lipschitz functions $(f_{n,k})_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ such that $f_{n,k} \underset{k\to\infty}{\longrightarrow} f_n$ and $g_k \underset{k\to\infty}{\longrightarrow} g$ uniformly, with $\mathrm{Lip}(f_{n,k}) = \mathrm{Lip}(g_k) = v_k^{-p}$ for some $p > 0$. Suppose that the triangular array $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ is (ϵ, ψ)-weakly dependent and let $k = k_n < n$,
then

$$\begin{aligned}
&\mathbb{E}\left(f_{n,k}(Y_n^{(l+1:r_n)}) \cdot g_k(X_{n,1})\right) \\
&\quad = \mathrm{Cov}\left(f_{n,k}(Y_n^{(l+1:r_n)}), g_k(X_{n,1})\right) + \mathbb{E}(f_{n,k}(Y_n^{(l+1:r_n)}))\mathbb{E}\left(g_k(X_{n,1})\right) \\
&\quad \leqslant \psi(r_n - l, 1, \mathrm{Lip}(f_{n,k}), \mathrm{Lip}(g_k))\epsilon_n(l) + \mathbb{E}(f_{n,k}(Y_n^{(l+1:r_n)}))\mathbb{E}\left(g_k(X_{n,1})\right)
\end{aligned}$$

Thus,

$$\limsup_{n\to\infty} \mathbb{P}\left(Y_n^{(l+1:r_n)} \neq 0 \,\middle|\, X_{n,1} \neq 0 \right) \leqslant \limsup_{k\to\infty} \psi(r_n - l, 1, \mathrm{Lip}(f_{n,k}), \mathrm{Lip}(g_k)) \frac{\epsilon_n(l)}{v_n} + \limsup_{k\to\infty} r_n v_n (1 - \frac{l}{r_n}) = \limsup_{k\to\infty} \psi(r_n - l, 1, \mathrm{Lip}(f_{n,k}), \mathrm{Lip}(g_k)) \frac{\epsilon_n(l)}{v_n}.$$

Finally, if in particular the triangular array $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ is weakly-dependent in at least one of the cases of the list (D.1') - (D.4'), the limit (2.65) is proven. □

Proof of Corollary 2.1 Note that for each case of weak dependence θ, η, κ and λ if we ask that $n^{-1}v_n^{-1-p} = \mathcal{O}(1)$, $(r_n n^{-1})^2 = \mathcal{O}(v_n^{p+1})$, $(r_n^{1-u} n^{-v} \wedge r_n n^{-2}) = \mathcal{O}(v_n^{2p+1})$ and $(r_n^{1-u} n^{-v} \wedge r_n n^{-2}) = \mathcal{O}(v_n^{2p+1})$ are fulfilled respectively, and if additionally we combine this with the respective dependence condition of the list (D.1) - (D.4) of Proposition 2.1, we then obtain (D.1') - (D.4') respectively, as the weak-dependence case may be. □

Proof of the relation (2.42). Relation (2.38) implies that for each $h = 0, \ldots, l$,

$$\sqrt{nv_n}(\widehat{\rho}_{A,B,n}(h) - \rho_{A,B,n}(h)) = \frac{Z_n(f_{A,B,h}) - \frac{h\sqrt{nv_n}}{r_n}\rho_{A,B,n}(h) - \rho_{A,B,n}(h) Z_n(f_{A,A,0}) + S_n(h) + D_n(h)}{(nv_n)^{-1/2} Z_n(f_{A,A,0}) + 1 + (nv_n)^{-1} R_n(A,A,0)} + o(1),$$

where

$$S_n(h) := (nv_n)^{-1/2} \left(R_n(A,B,h) - \rho_{A,B,n}(h) \cdot R_n(A,A,0) \right); \quad \text{for } h \geqslant 0$$

$$D_n(h) := (nv_n)^{-1/2} \sum_{j=1}^{m_n} \delta_{n,j}(f_{A,B,h}) \quad \text{for } h \geqslant 1 \quad \text{and } D_n(0) = 0.$$

Note that

$$(nv_n)^{-1/2} \mathbb{E} |R_n(A,B,h)| \leqslant r_n v_n (nv_n)^{-1/2} \rho_{A,B,n}(h) \underset{n\to\infty}{\longrightarrow} 0$$

and

$$(nv_n)^{-1/2} \mathbb{E} \left| \sum_{j=1}^{m_n} \delta_{n,j}(f_{A,B,n}) \right| \leqslant \frac{\sqrt{nv_n}}{r_n} h \rho_{A,B,n}(h) \underset{n\to\infty}{\longrightarrow} 0,$$

where the last convergence to zero is because $\sqrt{nv_n} = o(r_n)$. Therefore, using Chebyshev's inequality on the random variables $S_n(h)$ and $D_n(h)$, we prove that these variables converge to zero in probability. On the other hand, by using again Chebyshev's inequality on the random variable $\zeta_n = (nv_n)^{-1/2} Z_n(f_{A,A,0})$ combined with the stationarity of the time series, the approximation of the sum of indicator functions $\phi(x_1, \ldots, x_{r_n - l_n}) = \sum_{i=1}^{r_n - l_n} \mathbb{1}_{u_n A}(x_i)$ through lipschitzian increasing functions ϕ_ρ with $\rho = v_n^p$ for some $p > 0$, the assumption (2.41), and the weak-dependence condition of the list (D.1")-(D.4") as the case may be; we can then obtain the convergence to zero in probability of such a variable ζ_n. Finally, given that $\sqrt{nv_n} = o(r_n)$, we obtain the relation (2.42). □

Proof of Proposition 2.3. First, note that we have all the conditions to obtain relation (2.42). On the other hand, observe that the weak-dependence conditions (D.1")-(D.4") are stronger than the dependence conditions (D.1)-(D.4). Moreover, as the conditions (C.1), (C.3), (C.4) and (2.30) are maintained just as in Proposition 2.1, and as we have assumed the existence of the covariance function c of (C.2) for the functionals $f_{A,B,h}$ and $f_{A,A,0}$ through relation (2.43), then we have gaussian convergence of the random variables $Z_n(f_{A,B,h})$ and $Z_n(f_{A,A,0})$.

Finally, considering the existence of the covariance functions (2.44), we obtain the result. □

Proof of the expressions (2.47), (2.48) and (2.49). Due to that $X_i = b^{-1}X_0 + \sum_{s=1}^{i} b^{s-i-1}\xi_s$ for all $i \geqslant 1$, then for $h \geqslant 0$ we have

$$\begin{aligned}
\mathbb{P}\left(\frac{X_h}{y} \in B \,\middle|\, \frac{X_0}{y} \in A\right) &= \frac{1}{1-y}\mathbb{P}(X_h > y, X_0 > y) \\
&= \frac{1}{1-y}\mathbb{P}\left(X_0 > \max\left\{y, \left(y - \sum_{s=1}^{h} \frac{\xi_s}{b^{1-s+h}}\right) b^h\right\}\right) \\
&= \frac{1}{(1-y)b^h} \sum_{j_1, j_2, \ldots, j_h \in U(b)} \mathbb{P}\left(X_0 > \max\left\{y, \left(y - \sum_{s=1}^{h} \frac{j_s}{b^{1-s+h}}\right) b^h\right\}\right) \\
&= \frac{1}{b^h} \sum_{j_1, j_2, \ldots, j_h \in U(b)} \min\left\{1, \frac{1}{1-y}\left(1 - yb^h + \sum_{s=1}^{h} \frac{j_s}{b^{1-s}}\right)_+\right\}.
\end{aligned}$$

This proves relation (2.47). On the other hand, note that $\mu_b(j_1, \ldots, j_h) := 1 - b^h + \sum_{s=1}^{h} b^{s-1} j_s \leqslant -1$ for all $(j_1, \ldots, j_h) \in U^h(b) \setminus \{(b-1, \ldots, b-1)\}$ and $\mu_b(b-1, b-$

$1, \ldots, b-1) = 0$. Then, by substituting $y = 1 - 1/x$ in (2.47) and taking the limit when $x \to \infty$, we obtain (2.48). Finally, to prove (2.49) it suffices to substitute $y = 1 - v_n$ in (2.47). □

Proof of Remark 2.5. First, as $X_i = b^{-1}X_0 + \sum_{s=1}^{i} b^{s-i-1}\xi_s$ for all $i \geqslant 1$, then for n large enough (so large that $v_n b^{l_n} < 1$),

$$\begin{aligned}
&\mathbb{P}\left(X_0, X_k, X_h, X_{k+h} \in (u_n, \infty)\right) \\
&= \mathbb{P}\left(X_0 > \max\left\{u_n,\ b^k u_n - \sum_{s=1}^{k} b^{s-1}\xi_s,\ b^h u_n - \sum_{s=1}^{h} b^{s-1}\xi_s,\ b^{k+h}u_n - \sum_{s=1}^{k+h} b^{s-1}\xi_s\right\}\right) \\
&\quad = \sum_{j_1,\ldots,j_{k+h}\in U(b)} \mathbb{P}\left(X_0 > \max\left\{u_n,\ b^k u_n - \sum_{s=1}^{k} b^{s-1}j_s,\ b^h u_n - \sum_{s=1}^{h} b^{s-1}j_s,\right.\right. \\
&\qquad\qquad \left.\left. b^{k+h}u_n - \sum_{s=1}^{k+h} b^{s-1}j_s\right\},\ (\xi_1,\ldots,x_{k+h}) = (j_1,\ldots,j_{k+h})\right) \\
&\qquad = \frac{1}{b^{k+h}} \sum_{j_1,\ldots,j_{k+h}\in U(b)} \mathbb{P}\left(X_0 > \max\left\{u_n, b^k u_n - \sum_{s=1}^{k} b^{s-1}j_s,\right.\right. \\
&\qquad\qquad \left.\left. b^h u_n - \sum_{s=1}^{h} b^{s-1}j_s, b^{k+h}u_n - \sum_{s=1}^{k+h} b^{s-1}j_s\right\}\right) \\
&\qquad = \frac{1}{b^{k+h}} \sum_{j_1,\ldots,j_{k+h}\in U(b)} \min\left\{1 - u_n,\ \left(1 + \sum_{s=1}^{k} b^{s-1}j_s - b^k u_n\right)_+,\right. \\
&\qquad\qquad \left.\left(1 + \sum_{s=1}^{h} b^{s-1}j_s - b^h u_n\right)_+,\ \left(1 + \sum_{s=1}^{k+h} b^{s-1}j_s - b^{k+h}u_n\right)_+\right\} \\
&\quad = \frac{1}{b^{k+h}} \min\left\{1 - u_n,\ b^k(1-u_n),\ b^h(1-u_n),\ b^{k+h}(1-u_n)\right\} = \frac{v_n}{b^{k+h}},
\end{aligned} \tag{2.66}$$

the penultimate equality is by the same reasoning of the previous proof, *i.e.* $\mu_b(j_1,\ldots,j_k) := 1 - b^k + \sum_{s=1}^{k} b^{s-1}j_s \leqslant -1$ for all $(j_1,\ldots,j_k) \in U^k(b) \setminus \{(b-1,\ldots,b-1)\}$ and $\mu_b(b-1, b-1, \ldots, b-1) = 0$.

Thus,

$$\begin{aligned}
&\mathrm{Var}(\Delta_n(f_{A,B,h})) = l_n v_n \rho_{A,B,n}(h)\left(1 - l_n v_n \rho_{A,B,n}(h)\right) \\
&+ 2l_n \sum_{k=1}^{l_n-1}\left(1 - \frac{k}{l_n}\right)\mathbb{P}\left(X_0, X_k, X_h, X_{k+h} \in (u_n,\infty)\right) \\
&= l_n v_n \rho_{A,B,n}(h)\left(1 - l_n v_n \rho_{A,B,n}(h)\right) + \frac{2v_n l_n}{b^h}\sum_{k=1}^{l_n-1}\left(1 - \frac{k}{l_n}\right)\frac{1}{b^k},
\end{aligned} \tag{2.67}$$

so,

$$\frac{n^2}{r_n^2}\left(\frac{\mathrm{Var}(\Delta_n(f_{A,B,h}))}{nv_n}\right)^{1/2}$$
$$= \left(\frac{l_n}{r_n}\frac{n^3}{r_n^3}\left(\rho_{A,B,h}(h)[1-l_nv_n\rho_{A,B,n}(h)] + 2\frac{1}{b^h}\sum_{k=1}^{l_n-1}\left(1-\frac{k}{l_n}\right)\frac{1}{b^k}\right)\right)^{1/2}$$

which converges to zero when $(l_n/n)^{1/2} = o((r_n/n)^2)$, and therefore (C.3) holds. On the other hand,

$$\begin{aligned}\mathbb{E}f_{A,B,h}^3(Y_n) &= \sum_{i,j,k=1}^{r_n-h}\mathbb{P}\left(X_i, X_j, X_k, X_{i+h}, X_{j+h}, X_{k+h} \in (u_n,\infty)\right)\\ &= 6\sum_{1\leqslant i<j<k\leqslant r_n-h}\mathbb{P}\left(X_i, X_j, X_k, X_{i+h}, X_{j+h}, X_{k+h} \in (u_n,\infty)\right)\\ &\quad + 6\sum_{1\leqslant i<j\leqslant r_n-h}\mathbb{P}\left(X_i, X_j, X_{i+h}, X_{j+h} \in (u_n,\infty)\right)\\ &\quad + \sum_{1\leqslant i\leqslant r_n-h}\mathbb{P}\left(X_i, X_{i+h} \in (u_n,\infty)\right),\end{aligned} \tag{2.68}$$

which, using the same method of the above proof, we have that

$$\begin{aligned}T_1 := &\sum_{1\leqslant i<j<k\leqslant r_n-h}\mathbb{P}\left(X_i, X_j, X_k, X_{i+h}, X_{j+h}, X_{k+h} \in (u_n,\infty)\right)\\ &= \sum_{j=3}^{r_n-h-1}\sum_{i=1}^{j-2}(r_n-h-j)(j-1-i)\frac{v_n}{b^{j+h}}\\ &= \frac{v_n(r_n-h)}{2b^h}\sum_{j=3}^{r_n-h-1}\left(1-\frac{j}{r_n-h}\right)\frac{(j-2)(j-1)}{b^j}\\ &\leqslant \frac{r_nv_n}{2b^h}\sum_{j=3}^{r_n-h-1}\frac{(j-2)(j-1)}{b^j} \leqslant \frac{r_nv_n}{2b^h}C_{1,h},\end{aligned} \tag{2.69}$$

$$\begin{aligned}T_2 := &\sum_{1\leqslant i<j\leqslant r_n-h}\mathbb{P}\left(X_i, X_j, X_{i+h}, X_{j+h} \in (u_n,\infty)\right)\\ &\leqslant \sum_{k=1}^{r_n-h-1}(r_n-h-k)\frac{v_n}{b^{k+h}} \leqslant \frac{r_nv_n}{b^h}C_{2,h},\end{aligned} \tag{2.70}$$

$$T_3 := \sum_{1\leqslant i\leqslant r_n-h}\mathbb{P}\left(X_i, X_{i+h} \in (u_n,\infty)\right) \leqslant \frac{r_nv_n}{b^h}, \tag{2.71}$$

where

$$C_{1,h} := \lim_{n\to\infty} \sum_{j=3}^{r_n-h-1} (j-2)(j-1)b^{-j} < \infty, \text{ and}$$
$$C_{2,h} := \lim_{n\to\infty} \sum_{k=1}^{r_n-h-1} (1-k/(r_n-h))\, b^{-h} < \infty.$$

Therefore, combining (2.68) with (2.69)-(2.71), we obtain that

$$(r_n v_n)^{-1}\mathbb{E} f^3_{A,B,h}(Y_n) \leqslant \frac{3}{b^h}C_{1,h} + \frac{6}{b^h}C_{2,h} + \frac{1}{b^h} = C_h b^{-h}, \tag{2.72}$$

where $C_h := 3C_{1,h} + 6C_{2,h} + 1$. This verify (C.2).

Additionally, note that

$$\frac{n^2}{r_n^2}\sqrt{\frac{\mathbb{E}|f_{A,B,h}(Y_n)|^2 \mathbb{1}_{\{|f_{A,B,h}(Y_n)|>\sqrt{nv_n}\}}}{nv_n}} = \frac{n^{3/2}}{r_n^{3/2}(nv_n)^{\delta/4}}\sqrt{\frac{\mathbb{E}|f_{A,B,h}(Y_n)|^{2+\delta}}{r_n v_n}}, \tag{2.73}$$

which, by (2.72), converges to zero when $n^{3/2} = o\left(r_n^{3/2}(nv_n)^{\delta/4}\right)$, for some $\delta \in (0,1]$. Hence, we have also verified (C.4).

Now, in order to calculate the covariance functions, first, we rewrite the covariances as

$$\begin{aligned}\operatorname{Cov}(f_{A,B,h}(Y_n), f_{A,B,h'}(Y_n)) &= \sum_{i=1}^{r_n-h}\sum_{j=1}^{r_n-h'} \mathbb{P}\left\{X_i, X_j, X_{i+h}, X_{j+h'} \in (0,\infty)\right\} \\ &= E_1 + E_2 + E_3,\end{aligned}$$

where

$$E_1 := (r_n - h\vee h')\mathbb{P}(X_0, X_h, X_{h'} \in (u_n,\infty)) = r_n v_n \left(1 - \frac{h\vee h'}{r_n}\right)\frac{1}{b^{h\vee h'}},$$
$$E_2 := \sum_{i=r_n-h\vee h'+1}^{r_n-h\wedge h'} \mathbb{P}(X_i, X_{i-h\wedge h'} \in (0,\infty)) = (h\vee h' - h\wedge h')v_n\rho_{A,B,n}(h\wedge h'),$$

$$
\begin{aligned}
E_3 &:= \sum_{i\neq j} \mathbb{P}(X_0, X_{j-i}, X_h, X_{j-i+h'} \in (0,\infty)) \\
&= \sum_{1\leqslant i\neq j\leqslant r_n - h\vee h'} \mathbb{P}(X_0, X_{j-1}, X_h, X_{j-i+h'} \in (0,\infty)) \\
&+ \sum_{r_n - h\vee h'+1\leqslant i\neq j\leqslant r_n - h\wedge h'} \mathbb{P}(X_0, X_{j-1}, X_h, X_{j-i+h'} \in (0,\infty)) \\
&= \sum_{k=1}^{r_n - h\vee h' -1} (r_n - h\vee h' - k) \left[\frac{v_n}{b^{k\vee h\vee k+h'}} + \frac{v_n}{b^{k\vee h'\vee k+h}}\right] \\
&+ \sum_{k=1}^{h\vee h' - h\wedge h' -1} \frac{(h\vee h' - h\wedge h' - k)v_n}{b^{k\vee h\vee k+h'}} \\
&= \left(\frac{1}{b^h} + \frac{1}{b^{h'}}\right)\left(r_n v_n^2 C_{n,h,h'} + \sum_{k=1}^{h\vee h' - h\wedge h' -1} \frac{(h\vee h' - h\wedge h' - k)}{b^k}\right)
\end{aligned}
$$

with $C_{n,h,h'} := \sum_{k=1}^{r_n - h\vee h' -1} (1 - (h\vee h' + k)/r_n)\, b^{-k}$ which converges to some positive constant $C_{h,h'}$ as $n\to\infty$. For the development of E_1, E_2 and E_3, once again, we have used $X_i = b^{-1}X_0 + \sum_{s=1}^{i} b^{s-i-1}\xi_s$ for all $i\geqslant 1$, stationarity and n so large that $v_n b^{r_n} < 1$.

Finally,

$$
\sigma_{A,B}(h,h') \quad := \quad \lim_{n\to\infty} \frac{\mathrm{Cov}(f_{A,B,h}(Y_n), f_{A,B,h'}(Y_n))}{r_n v_n} \quad = \quad \frac{1}{b^{h\vee h'}} + C_{h,h'}\left(\frac{1}{b^h} + \frac{1}{b^{h'}}\right).
$$

We obtain the expression for $\sigma'_{A,B}(h)$ similarly. □

***Proof of the expression* 2.53.** This directly follows from (2.66), because

$$
\begin{aligned}
&\mathbb{P}\left(X_0, X_k, X_h, X_{k+h} \in (u_n,\infty)\right) \\
&\qquad = \frac{1}{b^{k+h}} \sum_{j_1,\ldots,j_{k+h}\in U(b)} \min\Bigg\{1 - u_n, \left(1 + \sum_{s=1}^{k} b^{s-1} j_s - b^k u_n\right)_+, \\
&\qquad\quad \left(1 + \sum_{s=1}^{h} b^{s-1} j_s - b^h u_n\right)_+, \left(1 + \sum_{s=1}^{k+h} b^{s-1} j_s - b^{k+h} u_n\right)_+\Bigg\} \\
&= \frac{v_n}{b^{k+h}} \sum_{j_1,\ldots,j_{k+h}\in U(b)} \min\Bigg\{1, b^k + \frac{\mu_b(j_1,\ldots,j_k)}{v_n}, b^h + \frac{\mu_b(j_1,\ldots,j_h)}{v_n}, \\
&\qquad\qquad b^{k+h} + \frac{\mu_b(j_1,\ldots,j_{k+h})}{v_n}\Bigg\}_+ .
\end{aligned}
$$

□

Chapter 3

Limit theorems for empirical processes of cluster functionals without mixing.

Understanding the way extreme values do cluster in the case of time series is an essential problem for extreme values theory and risk management. Drees & Rootzén [2010] provided a deep solution to this problem.
Until now, the existing literature on functional central limit theorems for empirical processes of cluster functionals is developed under mixing conditions. However, as we have also mentioned in the previous chapter, a simple non-mixing model does not fit their assumptions. The aim of this chapter is to include more general models of weakly dependent time series in those functional central limit theorems. To that end, we relax here the mixing conditions to $\tau-$dependence, by using the rich coupling properties of Dedecker & Prieur [2004a].
To illustrate the applicability of our functional central limit theorem (FCLT) for cluster functionals, we derive a FCLT without mixing conditions for the blocks estimator of the extremal index. We also perform simulations to show the precision on our result.

3.1 Introduction

This is a well-known fact that extremes of independent sequences are isolated. This feature does not always extend to dependent situations: in this case extremes occur within clusters, as this generally happens for real phenomenas. For instance, if we observe a heavy rain day i (discounted in millimetres of precipitation over some threshold u) at a specific tropical location, this is likely that during the days $i+1, i+2, \ldots, i+T-1$, we should also observe heavy rain at this location; here T is of course a random variable. The same applies to various phenomenas, such as hot days, strong cold days, etc.

The number $\mathbb{E}[T]$ is usually called "the (extremal) mean cluster size", and it is the reciprocal value of the extremal index θ, provided that $\theta > 0$. Obviously it is important to estimate this value θ, in particular, [Drees, 2011] introduces the following functional estimator $(\widehat{\theta}^*_{n,t})_{0<t\leq 1}$ of θ:

$$\widehat{\theta}^*_{n,t} = \frac{\sum_{j=1}^{m_n} f_t(X_{n,(j-1)r_n+1}, \dots, X_{n,jr_r})}{\sum_{j=1}^{m_n} g_t(X_{n,(j-1)r_n+1}, \dots, X_{n,jr_r})}, \tag{3.1}$$

where $X_{n,1}, \dots, X_{n,n}$ denotes a collection of normalised random variables, $m_n = \max\{j \in \mathbb{N} : j \leq n/r_n\}$ with $r_n \ll n$ and

$$f_t(x_1, \dots, x_r) = \mathbb{I}_{\{\max_{1\leq i\leq r} x_i > 1-t\}}, \tag{3.2}$$

$$g_t(x_1, \dots, x_r) = \sum_{i=1}^{r} \mathbb{I}_{\{x_i > 1-t\}}. \tag{3.3}$$

As we explained in the previous chapter, both maps (3.2) and (3.3) are classical examples of "extreme" cluster functionals. They may simplify estimation procedures in extreme values theory, under weak dependence. For instance, the estimator constructed from cluster functionals of the extremogram of [Davis & Mikosch, 2009] is given in [Gómez, 2015] (Section 2.4 of this document), the tail distributions is provided in [Drees & Rootzén, 2010], etc.

Considering definitions and notations from the previous chapter, $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ continues denoting here, in this chapter, a triangular array of $E-$valued row-wise stationary normalised random variables, defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, such that the sequence of conditional distributions of $X_{n,1}$, given that $X_{n,1}$ belongs to the failure set $A \subseteq E \setminus \{0\}$, converges weakly to some non-degenerate limit. However, without loss of generality, we will always consider in this chapter $A = E \setminus \{0\}$.

Let $\mathcal{F}$ be a class of cluster functionals and let $(Z_n(f))_{f\in\mathcal{F}}$ be the empirical process of cluster functionals in $\mathcal{F}$ defined in (2.13).
From Chapter 2, we already know that the convergence of fidis of $(Z_n(f))_{f\in\mathcal{F}}$ under $\beta-$mixing conditions were proved by Drees & Rootzén [2010]. Additionally, those authors have extended the results to the uniform convergence by adding [Van Der Vaart & Wellner, 1996]'s tightness criteria and asymptotic equicontinuity conditions to the convergence of the fidis that they had obtained.

In this chapter we provide a FCLT for such empirical process $(Z_n(f))_{f\in\mathcal{F}}$ of cluster functionals to more general classes of weakly dependent processes, the classes of *τ-weakly dependent processes*, introduced by Dedecker & Prieur [2004a].

Notice that, in particular, the process (2.1) is also τ-weakly dependent, see [Dedecker & Prieur, 2004a] (*i.e.* the AR(1)-input (2.1) is also an example of a τ-weakly dependent random process which is not mixing). The same situation occurs for general causal Bernoulli shifts, Markov models, etc. It is thus important to extend the FCLT for empirical processes of cluster functional to more general classes of weakly dependent processes.
In order to proceed, we apply the coupling results of Dedecker & Prieur [2004a,b] under τ-dependence assumptions to the random blocks $(Y_{n,i})_{1\leqslant i\leqslant m_n}$, in order to build independent random blocks $(Y^*_{n,i})_{1\leqslant i\leqslant m_n}$ coupled to the original blocks $(Y_{n,i})_{1\leqslant i\leqslant m_n}$, which will allow us to use [Van Der Vaart & Wellner, 1996]'s criteria of tightness and asymptotic equicontinuity (under independence) together with the convergence of the fidis of the empirical process of cluster functionals (2.13).

Mention that Gómez [2015] proves the convergence of the fidis of the empirical process (2.13) under even a weaker conditions of weak dependence, see Chapter 2. Functional CLT are not always necessary, which make this result of an independent interest.

This chapter is organised as follows. In Section 3.2, we shortly recall some features related to τ-weak dependence, including some examples. In Section 3.3, we provide the convergence of the fidis of the empirical process (2.13) and their corollaries. Moreover, in the same section, we examine the uniform convergence of the empirical process (2.13), through conditions for asymptotic tightness and asymptotic equicontinuity. In Section 3.4 we give an application of the above functional theorem, the functional blocks estimator of the extremal index. In Section 3.5 we study the multidimensional tail empirical processes for the case of the AR(1)-inputs defined in (2.1). Also we include a simulation study of the functional blocks estimator of the extremal index for this AR(1)-inputs, in order to demonstrate the accuracy of this technique. Proofs are reported to Section 3.6.

3.2 τ-weak dependence

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $\mathcal{M}$ be a σ-algebra of $\mathcal{A}$ and let X be a random variable with values in a Polish space E endowed with its metric δ. Suppose that X is $\mathbb{L}^p$-integrable (*i.e.* X satisfies $\|\delta(X,0)\|_p = (\int \delta^p(x,0)P_X(dx))^{\frac{1}{p}} < \infty$). Dedecker & Prieur [2004b] defined the **coefficient** τ_p as:

$$\tau_p(\mathcal{M}, X) = \left\| \sup \left\{ \int h(x)P_{X|\mathcal{M}}(dx) - \int h(x)P_X(dx) : h \in \Lambda(E,\delta) \right\} \right\|_p \tag{3.4}$$

where P_X is the distribution of X, $P_{X|\mathcal{M}}$ is a conditional distribution of X given $\mathcal{M}$ and $\Lambda(E,\delta)$ denotes the class of all Lipschitz functions $h : E \longrightarrow \mathbb{R}$ such that

$$\text{Lip}(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{\delta(x,y)} \leq 1.$$

Let $\mathbb{M} = \{(X_{n,i})_{1 \leq i \leq n}\}_{n \in \mathbb{N}}$ be a triangular array of $\mathbb{L}^p$-integrable E-valued row-wise stationary random variables. For each $i \in \{1, \ldots, n\}$, we set $\mathcal{M}_{n,i} = \sigma(X_{n,j},\ j \leq i)$, the σ-algebra of $\mathcal{A}$ generated by the normalised random variables $(X_{n,j})_{j \leq i}$.

Then, we define the coefficient:

$$\tau_{p,r}(k) = \max_{1 \leq l \leq r} \frac{1}{l} \sup \left\{ \tau_p(\mathcal{M}_{n,i}, (X_{n,j_1}, \ldots, X_{n,j_l})),\ i + k \leq j_1 < \cdots < j_l \leq n \right\}, \tag{3.5}$$

where we consider here the metric

$$\delta_l(x,y) = \sum_{i=1}^{l} \delta(x_i, y_i) \tag{3.6}$$

on E^l. Moreover, we say that $\mathbb{M}$ is τ_p-weakly dependent if

$$\lim_{k \longrightarrow \infty} \limsup_{r \longrightarrow \infty} \tau_{p,r}(k) = 0. \tag{3.7}$$

3.2.1 Some examples of τ-weakly dependent processes.

We show now a non-exhaustive list of τ-weakly dependent processes.

Example 3.1 (Causal Bernoulli shifts) Let $\mathbb{M} = \{(X_{n,i})_{1 \leqslant i \leqslant n}\}_{n \in \mathbb{N}}$ be a causal shift triangular array with innovation process $(\xi_i)_{i \in \mathbb{Z}}$ and consider the sequence $(\Delta_n^*(l))_{l \geq 0}$ defined in (2.22). Then, if $\mathcal{M}_{n,i} = \sigma(X_{n,j} : j \leq i)$, the coefficient $\tau_{1,r_n}(k)$ of $\mathbb{M}$ is bounded above by $\Delta_n^*(k)$, for all $k \in \mathbb{N}$. (See [Dedecker & Prieur, 2004a]).

Remark 3.1 It is clear that the triangular array $\{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ defined from a causal linear process (2.23) in Application 2.1 is τ-weakly dependent if ξ_0^2 is integrable and RHS of (2.24) converges to zero as n tends to infinity.

Example 3.2 (Markov models) Let $\mathbb{M} = \{(X_{n,i})_{1\leqslant i\leqslant n}\}_{n\in\mathbb{N}}$ be a triangular array such that the entries $X_{n,i}$ are defined by (2.25). Assume now that for each $n \in \mathbb{N}$, $(X_{n,i})_{r+1\leqslant i\leqslant n}$ is a stationary solution of (2.25) and define $\tilde{Y}_{n,r} = (X_{n,r}, \dots, X_{n,1})$. Let now $\tilde{Y}^*_{n,r} = (X^*_{n,r}, \dots, X^*_{n,1})$ be a random variable independent of $(\tilde{Y}_{n,r}, (\xi_i)_{i\geqslant r+1})$ and distributed as $\tilde{Y}_{n,r}$. Then setting $X^*_{n,i}$ as (2.26), we have for $i = r+1, \dots, n$ that $X^*_{n,i}$ is distributed as $X_{n,i}$ and independent of $\mathcal{M}_{n,r} = \sigma(X_{n,r}, \dots, X_{n,1})$. Therefore, as in the previous example, let $(\tilde{\Delta}_{p,n}(l))_{l\geq 0}$ be a non increasing sequence defined by

$$\tilde{\Delta}_{p,n}(l) := \left(\mathbb{E}||X_{n,l} - X^*_{n,l}||^p\right)^{1/p}.$$

Then one can apply the result of Lemma 3 in [Dedecker & Prieur, 2004a], and we obtain that $\tau_{p,r_n}(k) \leq \tilde{\Delta}_{p,n}(k)$.

Remark 3.2 Notice that $\theta_n(l) \leqslant \tau_{1,n}(l) \leqslant \tilde{\Delta}_{1,n}(l) = \tilde{\Delta}_n$. Therefore, the particular case (2.27) and Applications 2.2 and 2.3 are contained in this list of examples with the same bounds presented in such frame.

3.3 Central limit theorems for empirical processes of cluster functionals

3.3.1 Convergence of fidis

The proof of the convergence of fidis of the empirical process (2.13) still relies on the Bernstein blocks technique. Namely, we extract from each block $Y_{n,j}$ with length r_n a sub-block with length $l_n = o(r_n)$. Then the remaining sub-blocks, separated by l_n variables, together with convenient conditions of τ-dependence, will allow to couple independent blocks to the original sub-blocks, and yield the CLT.
Note that to use such coupling argument, we must approximate the cluster functionals $f \in \mathcal{F}$ by sequences of Lipschitz cluster functionals $\{f_n\}_{n\geq 0}$. However, in order to make sense to this, we must assume that the functionals $f \in \mathcal{F}$ are C-Lipschitz (see Definition 2.2).

Consider now the assumptions (C.1) and (C.2) of Chapter 2. Besides, the following assumption, which guarantees that the extraction of small blocks does not interfere with the convergence of fidis of $(Z_n(f))_{f\in\mathcal{F}}$.

(C.3) For all $f \in \mathcal{F}$,

$$\begin{aligned}\mathbb{E}|\Delta_n(f) - \mathbb{E}\Delta_n(f)|^2 \mathbb{I}_{\{|\Delta_n(f)-\mathbb{E}\Delta_n(f)| \le \sqrt{nv_n}\}} &= o(r_n v_n), \\ \mathbb{P}\left(|\Delta_n(f) - \mathbb{E}\Delta_n(f)| > \sqrt{nv_n}\ \right) &= o(r_n/n).\end{aligned}$$

Additionally, in order to obtain the convergence of fidis of the empirical process of cluster functionals evaluated on the big-blocks, the following assumptions are necessary:

(B) The integer sequences l_n and r_n are such that

$$l_n \ll r_n \ll v_n^{-1} \ll n \quad \text{and} \quad (nv_n)^{-\gamma/2-\eta(1+\beta)} = o(r_n/n),$$

for some positive constants β, γ, η, where $l_n \longrightarrow \infty$ and $nv_n \longrightarrow \infty$ as $n \to \infty$;

(C) For all $f \in \mathcal{F}$, there exist positive constants C and β such that

$$\mathbb{P}\left(Y_n \in D^{r_n}_{\rho_n}(f)\right) \le C\rho_n^{\beta},$$

where, $D^r_\rho(f) = \bigcup_{x\in D^r(f)}\{y \in E^r : \delta_r(x,y) < \rho\}$ with $\rho > 0$;

(D.1) For all $n \in \mathbb{N}$ and all $i \in \{1,\ldots,n\}$, the random variables $X_{n,i}$ are almost surely bounded and the triangular array formed by these variables, $\{(X_{n,i})_{1\le i\le n}\}_{n\in\mathbb{N}}$, is τ-weakly dependent such that

$$\tau_{1,r_n}(l_n) = o\left(n^{-1}(nv_n)^{\gamma/2}\right),$$

for some $\gamma \in (0,1)$;

(C.4) For all $f \in \mathcal{F}$,

$$\mathbb{E}\left|f(Y_n)\right|^{2+\alpha} = O(r_n v_n) \text{ and } \mathbb{P}\left(|f(Y_n)| > (nv_n)^{\frac{1-\gamma}{2}-\eta}\right) = o\left(n^{-1}r_n(nv_n)^{\frac{\alpha}{2(1+\alpha)}}\right)$$

for some $\alpha > 0$, $\gamma \in (0,1)$ and $\eta \in (0, \frac{1-\gamma}{2})$.

Remark 3.3 Note that the following assumption is a slightly stronger but simplified version of Assumption (C.4), which can be useful when we want to verify this condition in particular cases:

(C.4') For all $f \in \mathcal{F}$, $\mathbb{E}\,|f(Y_n)|^{2+\alpha} = O(r_n v_n)$ for some $\alpha > \dfrac{\gamma + 2\eta}{1-\gamma-2\eta}$, with some $\gamma \in (0,1)$ and $\eta \in \left(0, \frac{1-\gamma}{2}\right)$.

On the other hand, note that Assumption (C.4) implies Assumption (C.1).

Theorem 3.1 *Let $\mathcal{F}$ be a family of C-Lipschitz cluster functionals such that (C.2) and (C.3) hold. Then, if there exist α, β, γ and η such that Assumptions (B), (C), (D.1) and (C.4) are satisfied simultaneously, the fidis of the cluster functionals empirical process $(Z_n(f))_{f\in\mathcal{F}}$ converge to the fidis of a centred Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c defined in Assumption (C.2).*

We will now give a similar result to Proposition 2.2, which extends Theorems 1 and 3 in [Segers, 2003] to a multivariate situation. Such result (Proposition 3.1) provides sufficient conditions ensuring (C.2). Additionally, an expression of c (defined in (C.2)) is provided in Corollary 3.1.
In order to carry this out, consider also here Assumption (TC), explained in Chapter 2.

Proposition 3.1 *Suppose that the r.v's $\{(X_{n,i})_{1\le i\le n}\}_{n\in\mathbb{N}}$ satisfies the following condition:*

(D.2)

$$\lim_{l\to\infty} \limsup_{n\to\infty} \frac{r_n^{1+\epsilon}}{v_n} \tau_{1,r_n}(l) = 0, \quad \textit{for some } \epsilon > 0, \tag{3.8}$$

with $r_n \ll v_n^{-1} \ll n$ and $r_n \longrightarrow \infty$, $nv_n \longrightarrow \infty$ as $n \to \infty$.

Then, if $\mathcal{F}$ is uniformly bounded,

$$\mathbb{E}\left(f(Y_n)|Y_n \neq 0\right) = \theta_n^{-1}\mathbb{E}\left(f(Y_{n,1}) - f(Y_{n,1}^{(2:r_n)})|X_{n,1} \neq 0\right) + o(1), \quad \forall f \in \mathcal{F},$$

and

$$\theta_n = \frac{\mathbb{P}(Y_n \neq 0)}{r_n v_n} = \mathbb{P}\left(Y_{n,1}^{(2:r_n)} = 0 | X_{n,1} \neq 0\right)(1 + o(1)). \tag{3.9}$$

Additionally, if the assumption (TC) is satisfied, then:

$$\begin{array}{lcl} m_W & = & \sup\{i \geq 1 / W_i \neq 0\} < \infty, \\ \theta_n & \underset{n\to\infty}{\longrightarrow} & \theta = \mathbb{P}\left(W_i = 0, \forall i \geq 2\right) = \mathbb{P}(m_W = 1) > 0, \\ P_{\{f(Y_n)|Y_n\neq 0\}} & \underset{n\to\infty}{\overset{w}{\longrightarrow}} & \dfrac{1}{\theta}\left(\mathbb{P}\left(f(W) \in \cdot\,\right) - \mathbb{P}\left(f(W^{(2:\infty)}) \in \cdot\,,\ m_W \geq 2\right)\right). \end{array}$$

Corollary 3.1 *Let $\mathcal{F}$ be a family of C-Lipschitz cluster functionals such that (C.3) holds. Suppose that there exist α, β, γ and η such that Assumptions (B), (C), (D.1) and (C.4) are satisfied simultaneously, and $r_n^{1+\epsilon} = \mathcal{O}\left((nv_n)^{1-\gamma/2}\right)$ for some $\epsilon > 0$ such that (D.2) holds. If additionally (TC) holds, then the fidis of the cluster functionals empirical process $(Z_n(f))_{f\in\mathcal{F}}$ converge to the fidis of a centred Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c defined by*

$$c(f,g) = \mathbb{E}\left((fg)(W) - (fg)(W^{(2:\infty)})\right). \tag{3.10}$$

There are many cases for which $\|f\|_\infty = \sup_{x\in E_\cup} |f(x)| < \infty$, for all $f \in \mathcal{F}$. Under this condition, it is clear that Assumptions (C.1) and (C.3) are satisfied and Assumption (C.4) is not necessary. Therefore, it is important to note the following corollary.

Corollary 3.2 *Suppose that $\mathcal{F}$ is a class of bounded C-Lipschitz cluster functionals and that there exist β, γ and η such that Assumptions (B), (C) and (D.1) are satisfied simultaneously. Then, if (C.2) holds, the fidis of the cluster functionals empirical process $(Z_n(f))_{f\in\mathcal{F}}$ converge those of a centred Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c defined in the Assumption (C.2).*
Moreover, if additionally (TC) holds and $r_n^{1+\epsilon} = \mathcal{O}\left((nv_n)^{1-\gamma/2}\right)$ for some $\epsilon > 0$ such that condition (D.2) holds, then we obtain that the covariance function c is defined as in (3.10).

3.3.2 Uniform convergence

To prove uniform convergence, we use either asymptotic tightness of Z_n in the space $\ell^\infty(\mathcal{F})$ or asymptotic equicontinuity conditions, through the results in [Van Der Vaart & Wellner, 1996, § 2.11]. However, those results use independence, therefore a coupling argument for blocks $\left((Y_{n,j})_{1\le j\le m_n}\right)_{n\in\mathbb{N}}$ is needed, as this is shown in the proofs of Theorem 3.2 and 3.3 in Section 3.6.

Asymptotic tightness

Definition 3.1 *The sequence $(Z_n)_{n\in\mathbb{N}}$ is asymptotically tight if for every $\epsilon > 0$ there exists a compact set $K \subset \ell^\infty(\mathcal{F})$ such that*

$$\limsup_{n\longrightarrow\infty} \mathbb{P}^*(Z_n \notin K^\alpha) < \epsilon, \quad \textit{for any } \alpha > 0,$$

where $K^\alpha = \{Z \in \ell^\infty(\mathcal{F}),\ d_\mathcal{F}(Z,K) < \alpha\}$ is the "$\alpha-$enlargement" around K and $\mathbb{P}^*$ denotes the outer probability.

Definition 3.2 *The bracketing number $N_{[\cdot]}(\epsilon, \mathcal{F}, L_2^n)$ is defined as the smallest number N_ϵ such that for each $n \in \mathbb{N}$ there exits a partition $(\mathcal{F}_{n,k}^\epsilon)_{1\le k\le N_\epsilon}$ of $\mathcal{F}$ such that*

$$\mathbb{E}^* \sup_{f,g\in\mathcal{F}_{n,k}^\epsilon} (f(Y_n) - g(Y_n))^2 \le \epsilon^2 r_n v_n, \qquad \text{for} \quad 1 \le k \le N_\epsilon,$$

where $\mathbb{E}^*$ denotes the outer expectation.

In order to use theorem 2.11.9 in [Van Der Vaart & Wellner, 1996] we need:

(T.1) The set $\mathcal{F}$ of cluster functionals is such that for each $f \in \mathcal{F}$, the expectation $\mathbb{E}f^2(Y_n)$ is finite for all $n \in \mathbb{N}$ and such that the envelope function satisfies:

$$F(x) = \sup_{f\in\mathcal{F}} |f(x)| < \infty, \qquad \forall x \in E_\cup.$$

(T.2) $\mathbb{E}^* \left(F(Y_n)\mathbb{I}_{\{F(Y_n)>\epsilon\sqrt{nv_n}\}} \right) = o(r_n\sqrt{v_n/n})$ for all $\epsilon > 0$.

Note that for a sequence $(h_n(\alpha))_{n\ge1}$ of monotonically increasing positive functions, the convergence of $h_n(\alpha_n)$ to zero $\forall\alpha_n \downarrow 0$ is equivalent to

$$\lim_{\alpha\downarrow0} \limsup_{n\to\infty} h_n(\alpha) = 0,$$

thus the assumptions 2 and 3 of Theorem 2.11.9 in [Van Der Vaart & Wellner, 1996] are reformulated as follows:

(T.3) There exists a semi-metric ρ on $\mathcal{F}$ such that $\mathcal{F}$ is totally bounded with respect to (w.r.t.) ρ and

$$\lim_{\alpha\downarrow0} \limsup_{n\longrightarrow\infty} \sup_{\{f,g\in\mathcal{F}/\ \rho(f,g)<\alpha\}} \frac{1}{r_n v_n} \mathbb{E}\left(f(Y_n) - g(Y_n)\right)^2 = 0.$$

(T.4)

$$\lim_{\alpha\downarrow0} \limsup_{n\longrightarrow\infty} \int_0^\alpha \sqrt{\log N_{[\cdot]}(\epsilon, \mathcal{F}, L_2^n)}d\epsilon = 0.$$

Theorem 3.2 *Let $\mathcal{F}$ be a family of C-Lipschitz cluster functionals and suppose that there exist α, β, γ and η such that Assumptions (B), (C), (D.1) and (C.4) are satisfied simultaneously. Then, if (T.1)-(T.4) hold, the empirical process $(Z_n)_{n\in\mathbb{N}}$ is asymptotically tight in $\ell^\infty(\mathcal{F})$. Further, if in addition (C.2) and (C.3) hold, then Z_n converges to a centred Gaussian process Z with covariance function c defined in (C.2).*

Asymptotic equicontinuity

Definition 3.3 *The sequence $(Z_n)_{n\in\mathbb{N}}$ is asymptotically equicontinuous w.r.t. a semi-metric ρ if for any $\epsilon > 0$ and $\eta > 0$ there exists some $\alpha > 0$ such that:*

$$\limsup_{n\longrightarrow\infty} \mathbb{P}^* \left(\sup_{\{f,g\in\mathcal{F}/\ \rho(f,g)<\alpha\}} |Z_n(f) - Z_n(g)| > \epsilon \right) < \eta.$$

We use Theorem 2.11.1 in [Van Der Vaart & Wellner, 1996] to prove asymptotic equicontinuity. In order to do this, we need to define a semi-metric ρ_n on $\mathcal{F}$ as follows.

Let $(Y^*_{n,j})_{1\le j\le m_n}$ be the independent copies of $(Y_{n,j})_{1\le j\le m_n}$, then:

$$\rho_n(f,g) = \sqrt{\frac{1}{nv_n}\sum_{j=1}^{m_n}(f(Y^*_{n,j}) - g(Y^*_{n,j}))^2}. \tag{3.11}$$

Moreover, we denote by $N(\epsilon, \mathcal{F}, \rho)$ the "covering number", that is the minimum number of balls (w.r.t. the semi-metric ρ) with radius $\epsilon > 0$ necessary to cover $\mathcal{F}$.

We set two more assumptions:

(T.4') For $k = 1, 2$ the map

$$(Y^*_{n,1}, \ldots, Y^*_{n,\lceil m_n/2\rceil}) \longmapsto \sup_{\{f,g\in\mathcal{F}/\ \rho(f,g)<\alpha\}} \sum_{j=1}^{\lceil m_n/2\rceil} e_j \left(f(Y^*_{n,j}) - g(Y^*_{n,j}) \right)^k$$

is measurable for each $\alpha > 0$, each vector $(e_1, \ldots, e_{\lceil m_n/2\rceil}) \in \{-1, 0, 1\}^{\lceil m_n/2\rceil}$ and each $n \in \mathbb{N}$.

(T.5)

$$\lim_{\alpha\downarrow 0} \limsup_{n\longrightarrow\infty} \mathbb{P}^* \left(\int_0^\alpha \sqrt{\log N(\epsilon, \mathcal{F}, \rho_n)}d\epsilon > \eta \right) = 0, \quad \forall \eta > 0.$$

Theorem 3.3 *Let $\mathcal{F}$ be a family of C-Lipschitz cluster functionals and suppose that there exist α, β, γ and η such that Assumptions (B), (C), (D.1) and (C.4) are satisfied simultaneously. Then, if (T.1)-(T.3), (T.4') and (T.5) hold, the empirical process $(Z_n)_{n\in\mathbb{N}}$ is asymptotically equicontinuous.*
Further, if in addition (C.2) and (C.3) hold, then Z_n converges to a centred Gaussian process Z with covariance function c defined in (C.2).

3.4 Application: functional blocks estimator of the extremal index

Let $(X_i)_{1\le i\le n}$ be a set of real random variables with common distribution function F that belongs to the domain of attraction of some extreme-value distribution and let $u_n = F^{\leftarrow}(1-v_n t)$ for $t \in [0,1]$ and $v_n \downarrow 0$.
Suppose that the normalised sequence $X_{n,i} = (X_i - u_n)_+ / v_n$ is such that (D.2) holds. Then Proposition 3.1 entails that there exists a number (*extremal index*) $\theta \in (0,1]$ such that

$$\theta_{n,t} = \frac{\mathbb{P}(Y_n \neq 0)}{r_n v_n t} = \frac{\mathbb{P}\left(\max_{1\le i\le r_n} X_i > u_n\right)}{r_n v_n t} \underset{n\to\infty}{\longrightarrow} \theta. \tag{3.12}$$

The above convergence is uniform with respect to $t \in (0,1]$ and u_n, v_n are implicitly functions of t.

Given the convergence (3.12), Drees [2011] suggested to estimate θ by replacing the unknown probability $\mathbb{P}\left(\max_{1\le i\le r_n} X_i > u_n\right)$ and the unknown expectation $r_n v_n t = \mathbb{E}\left(\sum_{i=1}^{r_n} \mathbb{I}_{\{X_i > u_n\}}\right)$ by an empirical counterpart of $\theta_{n,t}$:

$$\widehat{\theta}^*_{n,t} = \frac{\sum_{j=1}^{m_n} \mathbb{I}_{\{\max_{(j-1)r_n < i \le jr_n} X_i > u_n\}}}{\sum_{j=1}^{m_n} \sum_{i=(j-1)r_n+1}^{jr_n} \mathbb{I}_{\{X_i > u_n\}}}, \tag{3.13}$$

where $m_n = \lceil n/r_n \rceil$ with $1 \ll r_n \ll v_n^{-1} \ll n$ but $nv_n \longrightarrow \infty$. The estimator (3.13) is the functional version of the *blocks estimator of the extremal index*, which can be expressed in terms of two empirical processes of cluster functionals $(Z_n(f_t), Z_n(g_t))_{0\le t\le 1}$ defined in (2.13).
For this, suppose without loss of generality, that the random variables $(X_i)_{1\le i\le n}$ are uniformly distributed on $[0,1]$ (otherwise, just consider the transformation $U_i = F(X_i)$, $1 \le i \le n$, where F is the distribution function of X_1, see [Drees, 2011, page 5]. Then, using the normalisation (2.8) with $a_n = v_n = 1 - u_n$ and the blocks $(Y_{n,j})_{1\le j\le m_n}$ defined in (2.12), we can write:

$$\widehat{\theta}^*_{n,t} = \frac{m_n^{-1}\sum_{j=1}^{m_n} f_t(Y_{n,j})}{m_n^{-1}\sum_{j=1}^{m_n} g_t(Y_{n,j})} = \frac{\mathbb{E}f_t(Y_{n,1}) + (nv_n)^{1/2} m_n^{-1} Z_n(f_t)}{\mathbb{E}g_t(Y_{n,1}) + (nv_n)^{1/2} m_n^{-1} Z_n(g_t)}, \tag{3.14}$$

where

$$f_t(x_1,\dots,x_r) = \mathbb{I}_{\{\max_{1\le i\le r} x_i > 1-t\}} \tag{3.15}$$

$$g_t(x_1,\dots,x_r) = \sum_{i=1}^{r} \mathbb{I}_{\{x_i > 1-t\}}. \tag{3.16}$$

For this special case, assume:

(C.2.1) $(r_n v_n)^{-1}\mathrm{Cov}(g_s(Y_n), g_t(Y_n)) \longrightarrow c_g(s,t)$, for all $0 \le s,t \le 1$.

(C.2.2) $(r_n v_n)^{-1}\mathrm{Cov}(f_s(Y_n), g_t(Y_n)) \longrightarrow c_{fg}(s,t)$, for all $0 \le s,t \le 1$.

(T) For some bounded function $h : (0,1] \longrightarrow \mathbb{R}$ such that $\lim_{t\to 0} h(t) = 0$

$$(r_n v_n)^{-1}\mathbb{E}\left(g_s(Y_n) - g_t(Y_n)\right)^2 \le h(t-s), \quad \text{for all } 0 \le s < t \le 1,$$

for all n sufficiently large.

We relax here the β-mixing condition assuming only the τ-dependence condition, in order to extend the first three results of [Drees, 2011]:

Proposition 3.2

(2.1) Suppose that the triangular array $\{(X_{n,i})_{1\le i\le n}\}_{n\in\mathbb{N}}$ is τ-weakly dependent, such that (D.1) holds with the basic condition: $r_n \ll v_n^{-1} \ll n$, and $(nv_n)^{-a} = o(r_n/n)$ for some $a \in (\frac{1}{2}, 1-\frac{\gamma}{2})$, where $r_n \longrightarrow \infty$ and $nv_n \longrightarrow \infty$ as $n \to \infty$. Then $(Z_n(f_t))_{0\le t\le 1}$ converges weakly to $Z_f = (\sqrt{\theta}B_t)_{0\le t\le 1}$, where B_t denote a standard Brownian motion.

(2.2) Suppose additionally that (C.2.1) and (T) hold, and that $r_n = o(\sqrt{nv_n})$. Then, if furthermore there exist $\alpha > 0$ and $\eta \in (\frac{1-\gamma}{4}, \frac{1-\gamma}{2})$ such that $\mathbb{E}\,|g_t(Y_n)|^{2+\alpha} = O(r_n v_n)$ and $\mathbb{P}\left(|g_t(Y_n)| > (nv_n)^{\frac{1-\gamma}{2}-\eta}\right) = o\left(n^{-1} r_n (nv_n)^{\frac{\alpha}{2(1+\alpha)}}\right)$ for all $t \in [0,1]$, the sequence of processes $(Z_n(g_t))_{0\le t\le 1}$ converges weakly to a centred Gaussian process $(Z_g(t))_{0\le t\le 1}$ with covariance function c_g.

(2.3) Under all the hypothesis of (2.1) and (2.2), if moreover (C.2.2) holds, then $(Z_n(f_t), Z_n(g_t))_{0\le t\le 1}$ converge weakly to $(Z_f(t), Z_g(t))_{0\le t\le 1}$ with

$$\begin{aligned} Cov(Z_f(s), Z_f(t)) &= \theta(s\wedge t),\\ Cov(Z_g(s), Z_g(t)) &= c_g(s,t),\\ Cov(Z_f(s), Z_g(t)) &= c_{fg}(s,t), \quad 0 \le s,t \le 1. \end{aligned} \tag{3.17}$$

Using the same argument as in Remark 2.3, we derive explicit expressions for the covariance functions c_g and c_{fg} as functions of the "tail chains" of $(X_i)_{i\in\mathbb{N}}$. Namely, if we assume that for every $k \in \mathbb{N}$, the distribution function of $(X_1,\ldots,,X_k)$ belongs to the domain of attraction of an extreme-value distribution, then there exists a sequence $W=(W_i)_{i\in\mathbb{N}}$ such that (2.31) holds.

In this case we obtain that:

$$c_g(s,t) = s\wedge t + \sum_{i=1}^{\infty}\Big(\mathbb{P}(W_1>1-s, W_{i+1}>1-t) + \mathbb{P}(W_1>1-t, W_{i+1}>1-s)\Big)$$

$$c_{fg}(s,t) = \begin{cases} \mathbb{P}\left(W_1>1-t, \max_{j\geq 1} W_j > 1-s\right) \\ +\sum_{i=1}^{\infty}\mathbb{P}\left(W_1>1-s, W_{i+1}>1-t, \max_{j\geq 2} W_j \leq 1-s\right), & s<t, \\ t, & s\geq t. \end{cases}$$

Corollary 3.3 *If the assumptions of Proposition 3.2 - (2.3) hold, then:*

$$\left(\sqrt{nv_n}t(\widehat{\theta}^*_{n,t}-\theta_{n,t})\right)_{0<t\leq 1} \xrightarrow[n\to\infty]{w} Z = Z_f - \theta Z_g, \tag{3.18}$$

where Z is a Gaussian process such that $\mathbb{E}Z(t)=0$ and

$$c(s,t) = Cov(Z(s),Z(t)) = \theta\Big(s\wedge t - c_{fg}(s,t) - c_{fg}(t,s)\Big) + \theta^2 c_g(s,t). \tag{3.19}$$

Note that in applications to real-data, the quantile function $F^{\leftarrow}(\cdot)$ is unknown. Therefore it is better to consider the blocks estimator of the extremal index, relative to the empirical quantile function $F_n^{\leftarrow}(\cdot)$:

$$\widehat{\theta}_{n,t} = \frac{\sum_{j=1}^{m_n}\mathbb{I}_{\{\max_{(j-1)r_n<i\leq jr_n} X_i > X_{n-\lceil nv_nt\rceil:n}\}}}{\sum_{j=1}^{m_n}\sum_{i=(j-1)r_n+1}^{jr_n}\mathbb{I}_{\{X_i>X_{n-\lceil nv_nt\rceil:n}\}}}, \tag{3.20}$$

where $X_{n-i:n}$ denotes the $(i+1)$-th largest order statistics, *i.e.*: $\max_{1\leq i\leq n} X_i = X_{n:n} \geq X_{n-1:n} \geq \cdots \geq X_{n-i:n}$.

As an application of Corollary 3.3, we derive:

Corollary 3.4 *If assumptions of Proposition 3.2-(2.3) hold, then*

$$\left(\sqrt{nv_n}t(\widehat{\theta}_{n,t} - \theta_{n,\frac{1}{v_n}(1-X_{n-\lceil nv_nt\rceil:n})})\right)_{0<t\leq 1} \xrightarrow[n\to\infty]{w} Z. \tag{3.21}$$

Moreover, if for each $t_0 \in (0,1)$ and each constant $T_1 > 0$ there exists $T_2 > 0$ such that

$$\sup_{\left\{s,t\geq t_0 \;/\; |s-t|\leq \frac{T_1}{\sqrt{nv_n}}\right\}} \left| \frac{\theta_{n,s}-\theta}{\theta_{n,t}-\theta} - 1 \right| \leq \frac{T_2}{\sqrt{nv_n}}, \tag{3.22}$$

then

$$\left(\sqrt{nv_n}\, t(\widehat{\theta}_{n,t} - \theta_{n,t})\right)_{0<t\leq 1} \xrightarrow[n\to\infty]{w} Z. \tag{3.23}$$

The expression (3.21) is the empirical variant of (3.18) involving the empirical quantile:

$$\widehat{\theta}_{n,t} = \widehat{\theta}^*_{n,\frac{1}{v_n}(1-X_{n-\lceil nv_n t\rceil:n})}.$$

3.5 Example and simulations

3.5.1 Multivariate tail empirical process of an AR(1)-process

Consider the AR(1)-process (2.1). X_0 is then uniformly distributed on $[0,1]$. Set the normalised random variables $\{(X_{n,i})_{1\leq i\leq n}\}_{n\in\mathbb{N}}$ as in eqn. (2.11) with $a_n = v_n = 1-u_n$. For vectors $x = (x_1,\ldots,x_d)$ and $y = (y_1,\ldots,y_d)$ in $[0,1]^d$, denote $x \leq y$ if and only if $x_i \leq y_i$, for all $i = 1,\ldots,d$. Then

$$\begin{aligned}
&\mathbb{P}\Big(X_{n,1} > x \,|\, X_{n,1} \neq 0\Big) \\
&\quad = \frac{1}{b^d \overline{v}_n} \sum_{j_1,\ldots,j_d \in U(b)} \min\left\{ \max_{i=1,\ldots,d} \left\{ 1 - b^i + \sum_{s=1}^{i} b^{s-1} j_s + b^i v_n (1-x_i) \right\}_+ , 1 \right\} \\
&\quad \xrightarrow[n\to\infty]{} \max_{i=1,\ldots,d} \{ b^{i-d}(1-x_i) \},
\end{aligned} \tag{3.24}$$

where $\overline{v}_n = \mathbb{P}(X_{n,1} \neq 0) \sim \; v_n = \mathbb{P}(X_1 > u_n) \xrightarrow[n\to\infty]{} 0.$

Consider now the family of cluster functionals:

$$\mathcal{F} = \left\{ f_x :\; x \in [0,1]^d \right\}, \quad \text{with} \quad f_x(x_1,\ldots,x_r) = \sum_{i=1}^{r} \mathbb{I}_{\{x_i > x\}}. \tag{3.25}$$

Here the covariance function c of (C.2) writes:

$$\begin{aligned}
c(x,y) = &\min\left\{ \max_{k=1,\ldots,d} \{ b^k (1-x_k) \}, \max_{k=1,\ldots,d} \{ b^k (1-y_k) \} \right\} \\
&+ \sum_{i=1}^{\infty} H_{b,i}(x,y) + \sum_{i=1}^{\infty} H_{b,i}(y,x),
\end{aligned} \tag{3.26}$$

where, for $i \geq d$

$$H_{b,i}(x,y) = \frac{1}{b^i} \min \left\{ \max_{k=1,\dots,d} \{b^k(1-x_k)\}, \max_{k=1,\dots,d} \{b^{k+i}(1-y_k)\} \right\} \tag{3.27}$$

and for $1 \leq i < d$,

$$H_{b,i}(x,y) \\ = \frac{1}{b^i} \min \left\{ \max_{k=1,\dots,i} \{b^k(1-x_k)\}, \max_{k=i+1,\dots,d} \{b^k \min(1-x_k, 1-y_k)\}, \max_{k=d-i,\dots,d} \{b^{k+i}(1-y_k)\} \right\}.$$

Condition (C) holds with $C = 2$ and $\beta = 1$. Moreover, note that for all sufficiently large n:

$$\mathbb{E}\left(\sum_{i=1}^{r_n} \mathbb{I}_{(x,y]}\left(\frac{X_i - u_n}{a_n}\right)\right)^2 \leq \left(\frac{1+\epsilon}{e}\right)^{(1+\epsilon)} |\log(y-x)|^{-(1+\epsilon)}\ r_n v_n,$$

for some $\epsilon > 0$ and all $0 \leq x < y \leq 1$, with $y - x \leq 1/2$. Then for the same family $\mathcal{F}$, by Corollary 3.6 in [Drees & Rootzén, 2010], conditions (C.3), (T.1), (T.2) hold, and $\sup_{x\in[0,1]^d} \left|Z_n(f_x) - \tilde{Z}_n(x)\right| \underset{n\to\infty}{\longrightarrow} 0$ in outer probability, where

$$\tilde{Z}_n(x) = \frac{1}{\sqrt{nv_n}} \sum_{i=1}^{n} \left(\mathbb{I}_{\{X_{n,i}>x\}} - \mathbb{P}(X_{n,i} > x)\right).$$

Besides, conditions (T.3) and (T.4) are verified in Example 3.8 of the same paper. Clearly condition (C.4') is satisfied trivially. Hence, setting l_n and r_n such that $b^{-l_n} = o\left(n^{-2}(nv_n)^{1+\gamma/2}\right)$ and $(nv_n)^{-\gamma/2-2\eta} = o(n^{-1}r_n)$, for some $\gamma \in (0,1)$ and $\eta \in (0, \frac{1-\gamma}{2})$, and such that $r_n = o(\sqrt{nv_n})$, the conditions (B) and (D.1) are satisfied (see Application 2.1). Therefore, by Theorem 3.2, the empirical process $(\tilde{Z}_n(x))_{x\in[0,1]^d}$ converges to a centred Gaussian process Z with the covariance function in eqn. (3.26).

3.5.2 Simulation study

We study computationally the asymptotic behaviour of the blocks estimator of the extremal index, defined in (3.20), for the AR(1)-process (2.1). Since X_0 is uniformly distributed on $[0,1]$ and $X_i = \frac{X_0}{b^i} + \sum_{s=1}^{i} \frac{\xi_s}{b^{i-s+1}}$ for all $i \geq 1$, setting $u_n = 1 - v_n t$ for $t \in (0,1]$, we get a theoretical expression for the index $\theta_{n,t}$ defined in (3.12):

$$\theta_{n,t} = \frac{1}{b^{r_n} r_n v_n t} \sum_{j_1,\dots,j_{r_n}\in U(b)} \min\left\{ \max_{i=1,\dots,r_n} \left\{ 1 - b^i(1-v_n t) + \sum_{s=1}^{i} b^{s-1} j_s \right\}_+, 1 \right\}, \tag{3.28}$$

which converges to some $\theta = \theta(b) \in (0,1)$ as $n \to \infty$, provided that r_n and l_n are such that $b^{-l_n} = o(v_n^2 r_n^{-1-\epsilon})$ for some $\epsilon > 0$ and $r_n \ll v_n^{-1} \ll n$ with $r_n \longrightarrow \infty$ and $nv_n \longrightarrow \infty$ as $n \to \infty$.

This theoretical expression (3.28) of the empirical extremal index $\theta_{n,t}$ is compared with its estimator $\widehat{\theta}_{n,t}$.
For this, we simulate the $AR(1)-$processes (2.1) for $b = 2, 3$ (according to the case) and we estimate their (functional) extremal index through of the blocks estimator (3.20), to compare it with the true (functional) asymptotic extremal index (3.28). Indeed, we generate a sample of size $n = 10^4$, which adjusts appropriately to an $AR(1)$-process (2.1) for $b = 2$ ($b = 3$). Moreover, we suppose that the threshold u_n is such that $v_n = n^{-1/4}$ and we set the blocks length $r_n = \lceil \log(n) \rceil$. In Figure 3.1 - left side, we show the functions $t \mapsto \theta_{n=10^4,t}$ (blue solid line) and $t \mapsto \widehat{\theta}_{n=10^4,t}(\omega_0)$ (black solid line), for the case $b = 2$. Moreover, in the same panel, we show confidence bands with a confidence level $1 - \alpha = 0.95$, which in the case of the band delimited by the red dashed lines, we have used the estimated variance through the estimator

$$\widehat{\sigma}^2_{n,t} = \frac{1}{nv_n} \sum_{j=1}^{m_n} \left(f_t(Y_{n,j}) - \widehat{\theta}_n g_t(Y_{n,j}) \right)^2, \tag{3.29}$$

where $\widehat{\theta}_n = \int_0^1 \widehat{\theta}_{n,t} dt$. In the case of the band delimited by the blue dashed lines, we have used the estimated variance by:

$$\widehat{\sigma'}^2_{n,t} = \widehat{c}(t,t) = \widehat{\theta}_n \cdot t \cdot \left(\widehat{\theta}_n \left(1 + \frac{1}{b} + \frac{2}{b \log(n)} \right) - 1 \right), \tag{3.30}$$

which is the covariance function defined in (3.19), calculated for the AR(1)-process of this experiment, but replacing θ by the estimator $\widehat{\theta}_n$. The same description applies to the case $b = 3$, see Figure 3.1 - right side.

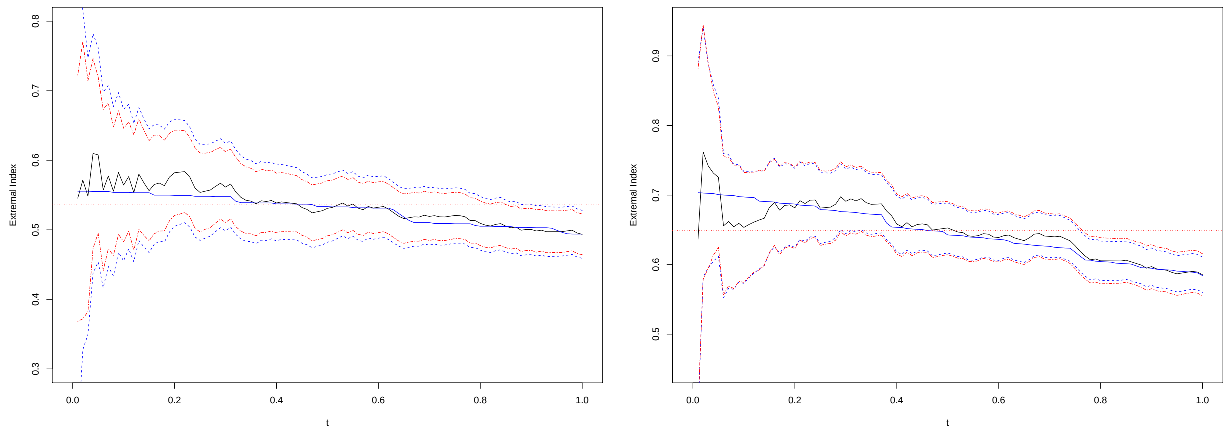


Figure 3.1: Fluctuations of the blocks estimator of the extremal index for two AR(1)-inputs. (a) On the left side, the AR(1)-input (2.1) with $b = 2$: $t \mapsto \theta_{n=10^4,t}$ is the blue solid line and $t \mapsto \widehat{\theta}_{n=10^4,t}(\omega_0)$ is the black solid line. The confidence band delimited by the red (resp. blue) dashed lines is built through the variance estimated by (3.29) (resp. (3.30)). The horizontal red dotted line is the estimated value by $\widehat{\theta}_n$. (b) On the right side, the AR(1)-input (2.1) with $b = 3$.

3.6 Proofs

In this section we prove the results given in Sections 3.3 and 3.4, as well as we justify some expressions included in Section 3.5. However, before starting the proofs, it is necessary to justify coupling arguments, which we give through Lemmas 3.3 and 3.4. Additionally, the existence of a Lipschitz approximation for bounded cluster functionals is provided by Lemma 3.5.

First of all, consider the following coupling result for multidimensional random variables. The original statement is given through Proposition 2.1 and Corollary 2.2 in [Dedecker & Prieur, 2004b].

Lemma 3.1 (Coupling) *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{M}$ a σ-algebra of $\mathcal{A}$ and X a random variable with values in a Polish space (E, δ). Assume that the integral $\int \delta(x, x_0) P_X(dx)$ is finite for some $x_0 \in E$ (and therefore for any such x_0). Assume that Ω is rich enough (this means that there exists a random variable U uniformly distributed over $[0,1]$, independent of the σ-algebra generated by X and $\mathcal{M}$). Then there exists a random variable X^*, measurable with respect to (w.r.t.) $\mathcal{M} \vee \sigma(X) \vee \sigma(U)$, independent of $\mathcal{M}$ and with the same distribution as X, and such that :*

$$\tau_1(\mathcal{M}, X) = \mathbb{E}\delta(X, X^*). \tag{3.31}$$

Lemma 3.2 *Suppose Y is a integrable random variable and X, U are random vectors such that $\sigma(U)$ is independent of $\sigma(X, Y)$. Then*

$$\mathbb{E}\left(Y|X,U\right) = \mathbb{E}\left(Y|X\right).$$

Lemma 3.3 (Coupling for the big-blocks) *Suppose that the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is rich enough. Assume that the random blocks $(Y_{n,j})_{1\le j\le m_n}$ are almost surely bounded. Then there exists a sequence $(Y^*_{n,j})_{1\le j\le m_n}$ of independent random blocks with the size $(r_n - l_n)$, such that:*

$$Y^*_{n,j} \stackrel{\mathcal{D}}{=} Y^{(r_n-l_n)}_{n,j} \quad \textit{and} \qquad \mathbb{E}\left(\delta_{r_n-l_n}(Y^{(r_n-l_n)}_{n,j},\ Y^*_{n,j})\right) \le r_n \tau_{1,r_n}(l_n), \quad j = 1,\dots,m_n. \tag{3.32}$$

Proof. Let $Y^{(r_n-l_n)}_{n,1}, \dots, Y^{(r_n-l_n)}_{n,m_n}$ be the sub-blocks extracted from the blocks $Y_{n,1}, \dots, Y_{n,m_n}$. We will build coupled blocks to those sub-blocks in a recursive way, through Lemma 3.1.
First of all, since the space $(\Omega, \mathcal{A}, \mathbb{P})$ is rich enough, we shall consider $(m_n - 1)$-independent random variables, uniformly distributed over $[0, 1]$, and independent of the random blocks $Y_{n,1}, \dots, Y_{n,m_n}$: we denote them by $U_1, \dots, U_{m_n-1}$. Apply Lemma 3.1 to $Y^{(r_n-l_n)}_{n,2}$ with the σ-algebra $\mathcal{M}^*_{n,1} = \sigma(Y^{(r_n-l_n)}_{n,1})$ and $U = U_1$, then there exists a random block $Y^*_{n,2}$ (which is a function of $Y^{(r_n-l_n)}_{n,2}$ and U_1), measurable w.r.t. $\sigma(Y^{(r_n-l_n)}_{n,1}, Y^{(r_n-l_n)}_{n,2}, U_1)$, and independent of $\mathcal{M}^*_{n,1}$ such that this block is distributed as $Y^{(r_n-l_n)}_{n,2}$ and

$$\begin{aligned}\mathbb{E}\left(\delta_{r_n-l_n}(Y^{(r_n-l_n)}_{n,2}, Y^*_{n,2})\right) &= \tau_1(\mathcal{M}^*_{n,1}\,, Y^{(r_n-l_n)}_{n,2}) \\ &= \tau_1(\mathcal{M}_{n,1}\,, Y^{(r_n-l_n)}_{n,2}) \le (r_n - l_n)\tau_{1,r_n-l_n}(l_n),\end{aligned}$$

where $\mathcal{M}_{n,j} = \sigma(X_{n,i} : i \le jr_n - l_n)$ for $j = 1, \dots, m_n$. In the same way, let $\mathcal{M}^*_{n,2} = \sigma(Y_{n,1}, Y^{(r_n-l_n)}_{n,2}, U_1)$ and consider the random variable U_2. Again, applying Lemma 3.1 to $Y^{(r_n-l_n)}_{n,3}$, there exists a random block $Y^*_{n,3}$ (function of $Y^{(r_n-l_n)}_{n,3}$ and U_2), measurable w.r.t. $\mathcal{M}^*_{n,2} \vee \sigma(Y^{(r_n-l_n)}_{n,3}) \vee \sigma(U_2)$, independent of $\mathcal{M}^*_{n,2}$ and distributed as $Y^{(r_n-l_n)}_{n,3}$ such that

$$\begin{aligned}\mathbb{E}\left(\delta_{r_n-l_n}(Y^{(r_n-l_n)}_{n,3}, Y^*_{n,3})\right) &= \tau_1(\mathcal{M}^*_{n,2}\,, Y^{(r_n-l_n)}_{n,3}) \\ &= \tau_1(\mathcal{M}_{n,2}\,, Y^{(r_n-l_n)}_{n,3}) \le (r_n - l_n)\tau_{1,r_n-l_n}(l_n),\end{aligned}$$

here the last equality is due to Lemma 3.2.
We proceed analogously until the last sub-block $Y_{n,m_n}^{(r_n-l_n)}$. In this last case, consider the σ-algebra

$$\mathcal{M}^*_{n,m_n-1} = \sigma(Y_{n,1}, \ldots, Y_{n,m_n-2}, Y_{n,m_n-1}^{(r_n-l_n)}, U_1, \ldots, U_{m_n-2})$$

and the random variable U_{m_n-1}, then Lemma 3.1 guarantees the existence of a random block Y^*_{n,m_n}, measurable w.r.t. $\mathcal{M}^*_{n,m_n-1} \vee \sigma(Y_{n,m_n}^{(r_n-l_n)}) \vee \sigma(U_{m_n-1})$, independent of $\mathcal{M}^*_{n,m_n-1}$, and distributed as $Y_{n,m_n}^{(r_n-l_n)}$. This block is such that:

$$\begin{aligned}\mathbb{E}\left(\delta_{r_n-l_n}(Y_{n,m_n}^{(r_n-l_n)}, Y^*_{n,m_n})\right) &= \tau_1(\mathcal{M}^*_{n,m_n-1}, Y_{n,m_n}^{(r_n-l_n)}) \\ &= \tau_1(\mathcal{M}_{n,m_n-1}, Y_{n,m_n}^{(r_n-l_n)}) \leq (r_n - l_n)\tau_{1,r_n-l_n}(l_n),\end{aligned}$$

by applying again Lemma 3.2 in the las equality.

Finally, set $Y^*_{n,1} = Y_{n,1}^{(r_n-l_n)}$, then by construction, the random blocks $Y^*_{n,1}, \ldots, Y^*_{n,m_n}$ are independent such that (3.32) holds. □

Lemma 3.4 (Coupling for the even and for the odd blocks) *Suppose that the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is rich enough. Assume that the random blocks $(Y_{n,j})_{1\leq j\leq m_n}$ are almost surely bounded, and consider together even and odd blocks, by using $k = 0$ or 1 according to the parity. Then, for each $k \in \{0, 1\}$, there exists a sequence $(Y^*_{n,j,k})_{1\leq j\leq \lceil m_n/2\rceil}$ of independent random blocks of size r_n, such that*

$$Y^*_{n,j,k} \stackrel{\mathcal{D}}{=} Y_{n,2j-k}, \quad \text{and} \quad \mathbb{E}\left(\delta_{r_n}(Y_{n,2j-k}, Y^*_{n,j,k})\right) \leq r_n\tau_{1,r_n}(r_n), \quad j = 1, \ldots, \lceil m_n/2\rceil. \tag{3.33}$$

Proof. The same argument as for the previous proof holds. Here the sub-blocks $(Y_{n,2j-k})_{1\leq j\leq \lceil \frac{m_n}{2}\rceil}$ are separated by r_n random variables: we set

$$\mathcal{M}_{n,j} = \sigma\left(X_{n,i} : i \leq (2j-k)r_n\right) \quad \text{and} \quad \mathcal{M}^*_{n,j} = \mathcal{M}_{n,j} \vee \sigma(U_1, \ldots, U_{j-1}),$$

for the block $Y_{n,2(j+1)-k}$. □

Lemma 3.5 (Lipschitz Approximation) *Let f be a bounded C-Lipschitz cluster functional. Then there exists a sequence $(\widetilde{f}_n)_{n\geq 1}$ of Lipschitz cluster functionals such that $\widetilde{f}_n$ converges point-wise to f on $C(f) = E_\cup \setminus \bigcup_{r=1}^{\infty} D^r(f)$.*

Proof. Let $r \in \mathbb{N}$ fixed. For each $\rho > 0$, we define the sets

$$D_\rho^r(f) = \bigcup_{x \in D^r(f)} B(x, \rho/2) \tag{3.34}$$

and $C_\rho^r(f) = E^r \setminus D_\rho^r(f)$, where $B(x,\rho) = \{y \in E^r,\ \delta_r(x,y) < \rho\}$. In order to simplify the notation, we write D_ρ and C_ρ instead of $D_\rho^r(f)$ and $C_\rho^r(f)$, respectively. Let $\{\mathcal{C}_k\}_{k\in\Lambda(r)}$ be the family of distinct components of $C^r(f)$ and define $\mathcal{C}_k(\rho) = \mathcal{C}_k \cap C_\rho$. Clearly $\{\mathcal{C}_k(\rho)\}_{k\in\Lambda(r)}$ is a partition of C_ρ. Since f is C-Lipschitz, then $f|_{C_\rho} : C_\rho \longrightarrow \mathbb{R}$ is Lipschitz with

$$\mathrm{Lip}(f|_{C_\rho}) \leq \frac{2\|f\|_\infty}{\rho}, \tag{3.35}$$

whenever $\rho \leq 2\|f\|_\infty / L(f)$, where $L(f) = \max_{k\in\Lambda(r)} \mathrm{Lip}(f|_{\mathcal{C}_k})$.

Indeed, note that if $x \in \mathcal{C}_k(\rho)$ and $y \in \mathcal{C}_l(\rho)$, with $k \neq l$, then

$$\delta_r(x,y) \geq \rho, \quad \text{entails} \quad \frac{|f(x) - f(y)|}{\delta_r(x,y)} \leq \frac{2\|f\|_\infty}{\rho}.$$

Thus

$$L_{k,l}(f|\rho) = \sup_{\substack{(x,y)\,\in\,\mathcal{C}_k(\rho)\times\mathcal{C}_l(\rho)\\ x\neq y}} \frac{|f(x) - f(y)|}{\delta_r(x,y)} \leq \frac{2\|f\|_\infty}{\rho}, \quad \text{for all } k \neq l. \tag{3.36}$$

Also, if $k = l$, then

$$L_k(f|\rho) = L_{k,k}(f|\rho) \leq \mathrm{Lip}(f|_{\mathcal{C}_k}) \leq L(f) \leq \frac{2\|f\|_\infty}{\rho}. \tag{3.37}$$

Therefore,

$$\begin{aligned} \mathrm{Lip}(f|_{C_\rho}) = \max \Bigg\{ &\max_{k\in\Lambda(r)} \sup_{x\neq y, x,y\in\mathcal{C}_k(\rho)} \frac{|f(x) - f(y)|}{\delta_r(x,y)}, \\ &\max_{k\neq l, k,l\in\Lambda(r)} \sup_{(x,y)\in\mathcal{C}_k(\rho)\times\mathcal{C}_l(\rho)} \frac{|f(x) - f(y)|}{\delta_r(x,y)} \Bigg\} \\ = \max \Big\{ &\max_{k\in\Lambda(r)} L_k(f|\rho)\ ,\ \max_{k\neq l, k,l\in\Lambda(r)} L_{k,l}(f|\rho) \Big\} \leq \frac{2\|f\|_\infty}{\rho}. \end{aligned} \tag{3.38}$$

On the other hand, using Kirszbraun's theorem (see Theorem 2.10.43 in [Federer, 1969]), for each $\rho \in (0, 2\|f\|_\infty / L(f))$ there exists a Lipschitz extension $f_\rho : E^r \longrightarrow \mathbb{R}$ of $f|_{C_\rho}$ such that $\mathrm{Lip}(f_\rho) = \mathrm{Lip}(f|_{C_\rho})$. Precisely, f_ρ is defined by

$$f_\rho(x) = \inf_{y \in C_\rho} \left\{ f|_{C_\rho}(y) + \mathrm{Lip}(f|_{C_\rho}) \delta_r(x,y) \right\}. \tag{3.39}$$

Clearly $f_\rho(x) \longrightarrow f(x)$ as $\rho \to 0$, for all $x \in E^r \setminus D^r(f)$.

To conclude, take $\rho = \rho_n < 2\|f\|_\infty / L(f)$ such that $\rho_n \downarrow 0$ as $n \to \infty$, and $\widetilde{f}_n = f_{\rho_n}$. □

Proof of Theorem 3.1. Let $(Y_{n,j})_{1 \le j \le m_n}$ be the blocks built from $(X_{n,i})_{1 \le i \le n}$. We start considering from Lemma 3.4, the independent blocks $(Y^*_{n,j,k})_{1 \le j \le \lceil \frac{m_n}{2} \rceil}$ with $k \in \{0,1\}$, such that (3.33) is satisfied. Define $\Delta^*_{n,j,k}(f) = f(Y^*_{n,j,k}) - f(Y^{*(r_n - l_n)}_{n,j,k})$ for $k = 0,1$ and $j = 1, \dots, \lceil \frac{m_n}{2} \rceil$. This is obvious that $\Delta^*_{n,j,k}(f) \overset{\mathcal{D}}{=} \Delta_{n,2j-k}(f) \overset{\mathcal{D}}{=} \Delta_n(f)$ for each j and each k, where $\Delta_{n,j}(f) = f(Y_{n,j}) - f(Y^{(r_n - l_n)}_{n,j})$ and $\Delta_n(f)$ is already defined in (2.29). Considering Assumption (C.3), we apply then [Petrov, 1975]'s theorem 1 (Section IX.1) to the i.i.d.r.v's $\widetilde{X}_{n,j} = (nv_n)^{-1/2} \Delta^*_{n,j,k}(f)$, to obtain that

$$DZ^*_{n,k}(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \left(\Delta^*_{n,j,k}(f) - \mathbb{E}\Delta^*_{n,j,k}(f) \right) = o_P(1), \qquad \text{for } k = 0,1. \tag{3.40}$$

We will prove now that

$$DZ_{n,k}(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \left(\Delta_{n,2j-k}(f) - \mathbb{E}\Delta_{n,2j-k}(f) \right) = o_P(1), \qquad \text{for } k = 0,1, \tag{3.41}$$

which immediately will imply that $DZ_n(f) = DZ_{n,0}(f) + DZ_{n,1}(f) = o_P(1)$.

Indeed, from Lemma 3.5, we consider a sequence of Lipschitz cluster functionals $(f_\rho)_{\rho > 0}$ which approximates f. In addition, for each $\rho > 0$, let $f^T_\rho = f_\rho \vee (-T) \wedge T$ be a truncation of f_ρ by T, for $T > 0$.

Then, for each $k \in \{0,1\}$, we write

$$C_{n,k}(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \mathbb{E}|f(Y_{n,2j-k}) - f(Y^*_{n,j,k})|$$
$$\leq \frac{2}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \mathbb{E}|f(Y_{n,2j-k}) - f^T_\rho(Y_{n,2j-k})| + \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \mathbb{E}|f^T_\rho(Y_{n,2j-k}) - f^T_\rho(Y^*_{n,j,k})|. \tag{3.42}$$

Firstly, observe that using the Lipschitz property of f^T_ρ, and the coupling relation (3.33), then the second right hand term of (3.42) is bounded as:

$$\frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil m_n/2 \rceil} \mathbb{E}|f^T_\rho(Y_{n,2j-k}) - f^T_\rho(Y^*_{n,j,k})| \leq \frac{2T}{\rho\sqrt{nv_n}} \sum_{j=1}^{\lceil m_n/2 \rceil} \mathbb{E}\left(\delta_{r_n}(Y_{n,2j-k}, Y^*_{n,j,k})\right)$$
$$\leq \frac{nT}{\rho\sqrt{nv_n}} \tau_{1,r_n}(r_n) \leq \frac{n}{(nv_n)^{\frac{\gamma}{2}}} \tau_{1,r_n}(r_n), \tag{3.43}$$

where we set $T = T_n = (nv_n)^{\frac{1-\gamma}{2}-\eta}$ and $\rho = \rho_n = (nv_n)^{-\eta}$, in the last inequality.

Secondly, the first term of the right hand side of (3.42) is bounded as:

$$\frac{2}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \mathbb{E}|f(Y_{n,2j-k}) - f^T_\rho(Y_{n,2j-k})| \leq \frac{n}{r_n\sqrt{nv_n}} \mathbb{E}|f(Y_n) - f^T_\rho(Y_n)|$$
$$\leq \frac{n}{r_n\sqrt{nv_n}} \left(\mathbb{E}\left(|f(Y_n) - f^T(Y_n)| \cdot \mathbb{1}_{\{|f(Y_n)|>T\}} \right) + \mathbb{E}\left(|f^T(Y_n) - f^T_\rho(Y_n)| \cdot \mathbb{1}_{\{Y_n \in D^{r_n}_\rho(f)\}} \right) \right)$$
$$\leq 2\left(\frac{\mathbb{E}|f(Y_n)|^p}{r_n v_n}\right)^{\frac{1}{p}} \left(\frac{(nv_n)^{q/2}\mathbb{P}\left(|f(Y_n)| > T\right)}{r_n v_n}\right)^{\frac{1}{q}} + \frac{2nTC\rho^\beta}{r_n\sqrt{nv_n}}$$
$$= 2\left(\frac{\mathbb{E}|f(Y_n)|^{2+\alpha}}{r_n v_n}\right)^{\frac{1}{2+\alpha}} \left(\frac{(nv_n)^{\frac{2+\alpha}{2+2\alpha}}\mathbb{P}\left(|f(Y_n)| > T\right)}{r_n v_n}\right)^{\frac{1+\alpha}{2+\alpha}} + 2C\frac{n}{r_n}(nv_n)^{-\gamma/2-\eta(1+\beta)}, \tag{3.44}$$

where $p^{-1} + q^{-1} = (2+\alpha)^{-1} + (1+\alpha)/(2+\alpha) = 1$, and we have used the assumption (C) in the second term of the last inequality of (3.44).

Combining (3.43) with (3.44) and considering the assumptions (D.1), (B) and (C.4), we obtain that

$$C_{n,k}(f) \underset{n\to\infty}{\longrightarrow} 0, \quad \text{for } k = 0,1. \tag{3.45}$$

Clearly the convergence (3.45) also holds if $C_{n,k}(f)$ is built with the sub-blocks

$$(Y_{n,2j-k}^{(r_n-l_n)}, Y_{n,j,k}^{*(r_n-l_n)})_j.$$

That is, using (C), (D.1), (B) and (C.4), we obtain similarly that

$$\widetilde{C}_{n,k}(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil m_n/2 \rceil} \mathbb{E}|f(Y_{n,2j-k}^{(r_n-l_n)}) - f(Y_{n,j,k}^{*(r_n-l_n)})| \underset{n\to\infty}{\longrightarrow} 0, \quad \text{for } k = 0, 1. \tag{3.46}$$

Then, the limits (3.45) and (3.46) together with (3.40), implies (3.41) because:

$$\begin{aligned} \mathbb{E}|DZ_{n,k}(f) - DZ_{n,k}^*(f)| &\leq \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil m_n/2 \rceil} \mathbb{E}|\Delta_{n,2j-k}(f) - \Delta_{n,j,k}^*(f)| \\ &\leq C_{n,k}(f) + \widetilde{C}_{n,k}(f) \underset{n\to\infty}{\longrightarrow} 0, \quad \text{for } k = 0, 1. \end{aligned} \tag{3.47}$$

Now, let

$$BZ_n(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} \left(f(Y_{n,j}^{(r_n-l_n)}) - \mathbb{E}f(Y_{n,j}^{(r_n-l_n)}) \right) \tag{3.48}$$

and

$$BZ_n^*(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} \left(f(Y_{n,j}^*) - \mathbb{E}f(Y_{n,j}^*) \right), \tag{3.49}$$

where $(Y_{n,j}^*)_{1\leq j\leq m_n}$ are independent blocks such that (3.32) holds.

In order to prove the convergence of fidis of Z_n to a centred Gaussian process, we just need to show that $BZ_n(f) - BZ_n^*(f) = o_P(1)$ for all $f \in \mathcal{F}$, because $Z_n(f) = BZ_n(f) + DZ_n(f) = BZ_n(f) + o_P(1)$ for all $f \in \mathcal{F}$ and the fidis of BZ_n^* converge to a centred Gaussian process with covariance function c by Assumptions (C.1) (see the last statement in Remark 3.3) and (C.2).

Indeed, using the same arguments as in (3.42)-(3.45) with the sub-blocks separated by l_n, we obtain that

$$\begin{aligned}
&\mathbb{E}|BZ_n(f) - BZ_n^*(f)| \\
&\leq \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} \left(2\mathbb{E}|f(Y_{n,j}^{(r_n-l_n)}) - f_\rho^T(Y_{n,j}^{(r_n-l_n)})| + \mathbb{E}|f_\rho^T(Y_{n,j}^{(r_n-l_n)}) - f_\rho^T(Y_{n,j}^*)| \right) \\
&\leq 4 \left(\frac{\mathbb{E}|f(Y_n^{(r_n-l_n)})|^{2+\alpha}}{r_n v_n} \right)^{\frac{1}{2+\alpha}} \left(\frac{(nv_n)^{\frac{2+\alpha}{2+2\alpha}} \mathbb{P}\left(|f(Y_n^{(r_n-l_n)})| > T\right)}{r_n v_n} \right)^{\frac{1+\alpha}{2+\alpha}} \\
&\qquad + 4C\frac{n}{r_n}(nv_n)^{-\gamma/2-\eta(1+\beta)} + 2n(nv_n)^{-\gamma/2}\tau_{1,r_n}(l_n) = o(1), \quad (3.50)
\end{aligned}$$

where the last equality in (3.50) is again by the assumptions (D.1), (B) and (C.4). Finally, the proof is complete. □

Proof of Proposition 3.1. It will suffice to show that:

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(Y_n^{(l+1:r_n)} \neq 0 \middle| X_{n,1} \neq 0\right) = 0, \quad (3.51)$$

because the end of the proof follows the same lines of the proof of Lemma 2.5 in [Drees & Rootzén, 2010].

Indeed, let $h(x) = 1 - \mathbb{1}_{\{0\}}(x)$ be a function defined on E^{r_n-l}. From Lemma 3.5, there are Lipschitz functions $(h_\rho)_{\rho>0}$ such that h_ρ converges to h on $E^{r_n-l} \setminus \{0\}$, as $\rho \downarrow 0$. Besides, these functions h_ρ are defined as in (3.39).

Thus,

$$\begin{aligned}
&\mathbb{P}\left(Y_n^{(l+1:r_n)} \neq 0 \middle| X_{n,1} \neq 0\right) = \frac{1}{v_n}\mathbb{P}\left(Y_n^{(l+1:r_n)} \neq 0, X_{n,1} \neq 0\right) \\
&= \frac{1}{v_n}\int_{\{X_{n,1}\neq 0\}} \mathbb{P}\left(Y_n^{(l+1:r_n)} \neq 0 \middle| \sigma(X_{n,1})\right) d\mathbb{P} = \frac{1}{v_n}\int_{\{X_{n,1}\neq 0\}} \mathbb{E}\left(\mathbb{I}_{\{Y_n^{(l+1:r_n)}\neq 0\}} \middle| \sigma(X_{n,1})\right) d\mathbb{P} \\
&= \frac{1}{v_n}\int_{\{X_{n,1}\neq 0\}} \mathbb{E}\left(h_\rho(Y_n^{(l+1:r_n)}) \middle| \sigma(X_{n,1})\right) d\mathbb{P} + o(1) \\
&= \frac{1}{v_n}\mathrm{Lip}(h_\rho)\int_{\{X_{n,1}\neq 0\}} \mathbb{E}\left(\frac{h_\rho(Y_n^{(l+1:r_n)})}{\mathrm{Lip}(h_\rho)} \middle| \sigma(X_{n,1})\right) d\mathbb{P} + o(1) \\
&\leq \frac{\mathrm{Lip}(h_\rho)}{v_n}\int_{\{X_{n,1}\neq 0\}} \left(\mathbb{E}\left(\frac{h_\rho(Y_n^{(l+1:r_n)})}{\mathrm{Lip}(h_\rho)} \middle| \sigma(X_{n,1})\right) - \mathbb{E}\left(\frac{h_\rho(Y_n^{(l+1:r_n)})}{\mathrm{Lip}(h_\rho)}\right)\right) d\mathbb{P} \\
&\qquad + \mathbb{P}\left(Y_n^{(l+1:r_n)} \neq 0\right) + o(1) \\
&\leq \frac{\mathrm{Lip}(h_\rho)}{v_n}(r_n - l)\tau_{p,r_n}(l) + (r_n - l)v_n + o(1) \leq 2\frac{r_n - l}{\rho v_n}\tau_{p,r_n}(l) + o(1).
\end{aligned}$$

Finally, setting $\rho = r_n^{-\alpha}$ for some $\alpha > 0$, (3.51) follows from Assumption (D.2). □

Proof of Theorem 3.2. On the one hand, note that Z_n is asymptotically tight if and only if $Z_{n,k}$, defined by

$$Z_{n,k}(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \left(f(Y_{n,2j-k}) - \mathbb{E} f(Y_{n,2j-k}) \right), \tag{3.52}$$

is asymptotically tight for each $k \in \{0, 1\}$.

On the other hand, for each $k \in \{0, 1\}$, we apply Lemma 3.4 to build independent blocks $(Y^*_{n,j,k})_{1 \le j \le \lceil \frac{m_n}{2} \rceil}$ such that (3.33) is satisfied. Now, defining

$$Z^*_{n,k}(f) = \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{\lceil \frac{m_n}{2} \rceil} \Big(f(Y^*_{n,j,k}) - \mathbb{E} f(Y^*_{n,j,k}) \Big), \quad \text{for } k = 0, 1, \tag{3.53}$$

Theorem 2.11.9 in [Van Der Vaart & Wellner, 1996] implies that $Z^*_{n,k}$ is asymptotically tight, for each $k \in \{0, 1\}$. For this, set $Z_{nj}(f) = f(Y^*_{n,j,k})$ and $\lceil \frac{m_n}{2} \rceil$ instead of m_n. Therefore, $Z_{n,k}$ is asymptotically tight for each $k \in \{0, 1\}$, because

$$\mathbb{E}|Z_{n,k}(f) - Z^*_{n,k}(f)| \le C_{n,k}(f) \underset{n\to\infty}{\longrightarrow} 0, \quad \forall f \in \mathcal{F}, \quad \forall k \in \{0, 1\}, \tag{3.54}$$

by the assumptions (C), (D.1), (B) and (C.4).

The last assertion to prove relies on Theorem 3.1. □

Proof of Theorem 3.3. Consider (T.5). Note that from the triangle inequality, Z_n is asymptotically equicontinuous if $Z_{n,k}$ from eqn. (3.52) is asymptotically equicontinuous for each $k \in \{0, 1\}$. Now, we use again Lemma 3.4 together with the assumptions (C), (D.1), (B) and (C.4), as in the previous proof, to prove that $Z_{n,k}$ is asymptotically equicontinuous iff $Z^*_{n,k}$ is asymptotically equicontinuous for each $k \in \{0, 1\}$. In fact, in this case, $Z^*_{n,k}$ is asymptotically equicontinuous by using the conditions (T1)-(T3), (T4') and (T5) with Theorem 2.11.1 in [Van Der Vaart & Wellner, 1996].

The remaining steps are the same in the proof of Theorem 2.10 in [Drees & Rootzén, 2010]. □

Proof of Proposition 3.2. The proof is a direct application of Theorem 3.1. Indeed, for the item:

2.1) Condition (C.3) is trivially satisfied and condition (C.4) is not necessary because the cluster functionals are bounded in this case (see Corollary 3.2). Condition (C) holds with $\beta = 1$ and $C = 2$. Moreover, note that the second equality of (B) is required in the proof of Theorem 3.1 to prove the convergence to zero of the second term of (3.44), which in this case, it is suffice to consider $(nv_n)^{-a} = o(r_n/n)$ as $n \to \infty$, for some $a \in (\frac{1}{2}, 1 - \frac{\gamma}{2})$. Condition (C.2) is proved in the proof of Theorem 2.1(i) of [Drees, 2011].

2.2) Condition (C.2) of Theorem 3.1 follow from condition (C.2.1). Conditions (D.1) and (C.4) are already fixed. In order to verify condition (B), note that from the assumptions of (2.1): $(nv_n)^{-a} = o(r_n/n)$ for some $a \in (\frac{1}{2}, 1 - \frac{\gamma}{2})$, therefore it is suffice to write $a = \frac{\gamma}{2} + 2\eta$.
The remaining steps for the proof of conditions (T.1)-(T.4) and (C.3) are the same as in the proof of Theorem 2.1(ii) in [Drees, 2011]. Here the assumptions (C1) and (C2) of the latter reference are replaced by our assumptions (D.1) and (B), respectively.

2.3) The steps are the same as in the proof of Theorem 2.1(iii) in [Drees, 2011], but replacing the assumptions (C3.1) and (C3.2) in his proof by our conditions (C.2.1) and (C.2.2).

The proof is complete. □

Proof of Corollaries 3.3 and 3.4. Our Proposition 3.2 is the counterpart of Theorem 2.1 in [Drees, 2011] under the tau-dependence condition. Thus replace [Drees, 2011]'s theorem 2.1 by our Proposition 3.2 in the proof of his Corollaries 2.3 and 2.4. This proves respectively our Corollaries 3.3 and 3.4. □

Proof of the expression (3.24). If $(X_i)_{i\in\mathbb{Z}}$ is the AR(1)-process, the stationary solution of (2.1), note that for each $i \in \mathbb{N}$

$$X_i = \frac{X_0}{b^i} + \sum_{s=1}^{i} \frac{\xi_s}{b^{i-s+1}}. \tag{3.55}$$

If n large enough in order that $b^d v_n < 1$, then for $x \in [0,1]^d$:

$$\begin{aligned}
\mathbb{P}\left(X_{n,1} > x, X_{n,1} \neq 0\right) &= \mathbb{P}\left(X_i > a_n x_i + u_n, \text{ for some } i = 1,\dots,d\right) \\
&= \mathbb{P}\left(X_0 > b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1}\xi_s, \text{ for some } i = 1,\dots,d\right) \\
&= \mathbb{P}\left(X_0 > \min_{i=1,\dots,d}\left\{b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1}\xi_s\right\}\right) \\
= \sum_{j_1,\dots,j_d \in U(b)} \mathbb{P}&\left(X_0 > \min_{i=1,\dots,d}\left\{b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1}j_s\right\}, (\xi_1,\dots,\xi_d) = (j_1,\dots,j_d)\right) \\
&= \frac{1}{b^d}\sum_{j_1,\dots,j_d \in U(b)} \mathbb{P}\left(X_0 > \min_{i=1,\dots,d}\left\{b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1}j_s\right\}\right) \\
&= \frac{1}{b^d}\sum_{j_1,\dots,j_d \in U(b)} \min\left\{\max_{i=1,\dots,d}\left\{1 - b^i + \sum_{s=1}^{i} b^{s-1}j_s + b^i v_n(1-x_i)\right\}_+, \quad 1\right\} \\
&= \frac{1}{b^d}\max_{i=1,\dots,d}\left\{b^i v_n(1-x_i)\right\},
\end{aligned}$$

since

$$\mu_b(j_1,\dots,j_d;i) = 1 - b^i + \sum_{s=1}^{i} b^{s-1}j_s \leq -1$$

for all $(j_1,\dots,j_d) \in U^d(b) \setminus \{(b-1,\dots,b-1)\}$ and $\mu_b(b-1,b-1,\dots,b-1) = 0$.

Thus,

$$\mathbb{P}\left(X_{n,1} > x | X_{n,1} \neq 0\right) \underset{n\to\infty}{\longrightarrow} \max_{i=1,\dots,d}\left\{b^{i-d}(1-x_i)\right\}, \tag{3.56}$$

which leads to the expression (3.24). □

Proof of the expression (3.26). Let $x, y \in [0,1]^d$. Then as before for $i \geq 1$, if n is sufficiently large such that $b^{i+d}v_n < 1$, we obtain:

$$\begin{aligned}
&\mathbb{P}\left(X_{n,1} > x, X_{n,i+1} > y\right) \\
&\quad= \mathbb{P}\left(X_k > a_n x_k + u_n, X_{i+l} > a_n y_l + u_n, \text{ for some } (k,j) \in \{1,\dots,d\}^2\right) \\
&\qquad= \mathbb{P}\Bigg(X_0 > \min_{k=1,\dots,d}\left\{b^k(a_n x_k + u_n) - \sum_{s=1}^{k} b^{s-1}\xi_s\right\}, \\
&\qquad\quad X_i > \min_{l=1,\dots,d}\left\{b^l(a_n y_l + u_n) - \sum_{s=1}^{l} b^{s-1}\xi_{s+i}\right\}\Bigg) \\
&\quad= \sum_{\substack{j_1,\dots,j_d \in U(b)\\ j_{i+1},\dots,j_{i+d} \in U(b)}} \mathbb{P}\Bigg(X_0 > \min_{k=1,\dots,d}\left\{b^k(a_n x_k + u_n) - \sum_{s=1}^{k} b^{s-1}j_s\right\}, \\
&\qquad\quad X_i > \min_{l=1,\dots,d}\left\{b^l(a_n y_l + u_n) - \sum_{s=1}^{l} b^{s-1}j_{s+i}\right\}, \\
&\qquad\qquad (\xi_1,\dots,\xi_d,\xi_{i+1},\dots,\xi_{i+d}) = (j_1,\dots,j_{i+d})\Bigg)
\end{aligned}$$

and then

$$\begin{aligned}
&\mathbb{P}(X_{n,1} > x, X_{n,i+1} > y) \\
&\quad= \frac{1}{b^d}\sum_{\substack{j_1,\dots,j_d \in U(b)\\ j_{i+1},\dots,j_{i+d} \in U(b)}} \mathbb{P}\Bigg(X_0 > \min_{k=1,\dots,d}\left\{b^k(a_n x_k + u_n) - \sum_{s=1}^{k} b^{s-1}j_s\right\}, \\
&\qquad X_i > \min_{l=1,\dots,d}\left\{b^l(a_n y_l + u_n) - \sum_{s=1}^{l} b^{s-1}j_{s+i}\right\}, (\xi_1,\dots,\xi_d) = (j_1,\dots,j_d)\Bigg) \\
&= \frac{1}{b^d}\sum_{\substack{j_1,\dots,j_d \in U(b)\\ j_{i+1},\dots,j_{i+d} \in U(b)}} \mathbb{P}\Bigg(X_0 > 1 - \min\left\{\max_{k=1,\dots,d}\left\{\mu_b(j_1,\dots,j_d;k) + b^k v_n(1-x_k)\right\}_+, 1\right\}, \\
&X_i > 1 - \min\left\{\max_{l=1,\dots,d}\left\{\mu_b(j_{i+1,\dots,j_{i+d}}) + b^l v_n(1-y_l)\right\}_+, 1\right\}, (\xi_1,\dots,\xi_d) = (j_1,\dots,j_d)\Bigg) \\
&\qquad\qquad= \frac{1}{b^d}\mathbb{P}\Bigg(X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\}, \\
&\qquad\qquad\qquad X_i > 1 - \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\}, \xi_1 = \dots = \xi_d = b-1\Bigg),
\end{aligned}$$

since $\mu_b(j_1,\dots,j_d;i) = 1 - b^i + \sum_{s=1}^{i} b^{s-1}j_s \le -1$ for all $(j_1,\dots,j_d) \in U^d(b) \setminus \{(b-1,\dots,b-1)\}$ and $\mu_b(b-1,b-1,\dots,b-1) = 0$.

Moreover, note that for $i > d$

$$
\begin{aligned}
&\mathbb{P}(X_{n,1} > x, X_{n,i+1} > y) \\
&\quad = \frac{1}{b^d}\mathbb{P}\left(X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right. \\
&\quad\quad \left. X_i > 1 - \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right)\right), \xi_1 = \ldots = \xi_d = b-1\Big) \\
&\quad = \frac{1}{b^d}\mathbb{P}\left(X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right. \\
&\quad\quad \left. X_0 > b^i - \max_{l=1,\dots,d}\left\{b^{l+i} v_n(1-y_l)\right\} + 1 - b^d - \sum_{s=d+1}^{i} b^{s-1}\xi_s\right) \\
&\quad = \frac{1}{b^d}\sum_{j_{d+1},\dots,j_i \in U(b)} \mathbb{P}\left(X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right. \\
&\left. X_0 > b^i - b^i \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\} + 1 - b^d - \sum_{s=d+1}^{i} b^{s-1} j_s, (\xi_{d+1},\dots,\xi_i) = (j_{d+1},\dots,j_i)\right) \\
&\quad = \frac{1}{b^i}\sum_{j_{d+1},\dots,j_i \in U(b)} \mathbb{P}\left(X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right. \\
&\quad\quad \left. X_0 > b^i - b^i \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\} + 1 - b^d - \sum_{s=d+1}^{i} b^{s-1} j_s\right) \\
&\quad = \frac{1}{b^i}\sum_{j_{d+1},\dots,j_i \in U(b)} \min\left\{\max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right. \\
&\quad\quad \left. \max_{l=1,\dots,d}\left\{b^d + \sum_{s=d+1}^{i} b^{s-1} j_s + b^{l+i} v_n(1-y_l) - b^i\right\}_+, 1\right\} \\
&\quad = \frac{v_n}{b^i}\min\left\{\max_{k=1,\dots,d}\left\{b^k(1-x_k)\right\}, \max_{l=1,\dots,d}\left\{b^{l+i}(1-y_l)\right\}\right\} = v_n H_{b,i}(x,y).
\end{aligned}
$$

Similarly for $1 \le i < d$, we obtain that

$$
\begin{aligned}
&\mathbb{P}(X_{n,1} > x, X_{n,i+1} > y) \\
&= \frac{v_n}{b^i}\min\left\{\max_{k=1,\dots,i}\{b^k(1-x_k)\}, \max_{k=i+1,\dots,d}\{b^k \min(1-x_k, 1-y_k)\}, \max_{k=d-i,\dots,d}\{b^{k+i}(1-y_k)\}\right\} \\
&= v_n H_{b,i}(x,y).
\end{aligned}
$$

From Lemma 5.2-(iii) in [Drees & Rootzén, 2010] we infer that $\mathbb{E}|f(Y_n)| = o(\sqrt{r_n v_n})$. Thus, for n sufficiently large:

$$\begin{aligned}
&\frac{1}{r_n v_n}\mathrm{Cov}\left(f_x(Y_n), f_y(Y_n)\right) \\
&\sim \mathbb{P}(X_{n,1} > x, X_{n,1} > y) + \sum_{i=1}^{r_n-1}\left(1 - \frac{i}{r_n}\right)\left(\mathbb{P}(X_{n,1} > x, X_{n,i+1} > y) + \mathbb{P}(X_{n,1} > y, X_{n,i+1} > x)\right) \\
&\quad \underset{n\to\infty}{\longrightarrow} \min\left\{\max_{k=1,\dots,d}\{b^k(1-x_k)\}, \max_{k=1,\dots,d}\{b^k(1-y_k)\}\right\} + \sum_{i=1}^{\infty}\left(H_{b,i}(x,y) + H_{b,i}(y,x)\right).
\end{aligned}$$

This concludes the proof. □

Proof of the relation (3.28). The proof is similar to the proof of the expression (3.24) but noting that $d = r_n$ increases when n increases. Indeed,

$$\begin{aligned}
\mathbb{P}\left(\max_{1\le i\le r_n} X_i > 1 - v_n t\right) &= \mathbb{P}\left(X_i > 1 - v_n t,\ \text{for some } i = 1,\dots,r_n\right) \\
&= \mathbb{P}\left(X_0 > b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}\xi_s,\ \text{for some } i = 1,\dots,r_n\right) \\
&= \mathbb{P}\left(X_0 > \min_{1\le i\le r_n}\left\{b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}\xi_s\right\}\right) \\
&= \sum_{j_1,\dots,j_{r_n}\in U(b)} \mathbb{P}\left(X_0 > \min_{1\le i\le r_n}\left\{b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}\xi_s\right\}, (\xi_1,\dots,\xi_{r_n}) = (j_1,\dots,j_{r_n})\right) \\
&= \frac{1}{b^{r_n}}\sum_{j_1,\dots,j_{r_n}\in U(b)} \mathbb{P}\left(X_0 > \min_{1\le i\le r_n}\left\{b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}j_s\right\}\right) \\
&= \frac{1}{b^{r_n}}\sum_{j_1,\dots,j_{r_n}\in U(b)} \min\left\{\max_{1\le i\le r_n}\left\{1 + \sum_{s=1}^{i} b^{s-1}j_s - b^i(1-v_n t)\right\}_+, 1\right\}.
\end{aligned}$$

□

Chapter 4

On extreme values in stationary weakly dependent random fields.

The existing literature on extremal types theorem for stationary random processes and fields is, until now, developed under either mixing or "Coordinatewise (Cw)-mixing" conditions. However, as we have mentioned in the previous chapters, mixing conditions is very restrictive and difficult to verify. These limitations do not escape when we consider the case of Cw-mixing of Leadbetter & Rootzén [1998]. In this context, we provide extremal types theorem for stationary random fields under a weaker and simplest to verify dependence condition, introduced by Doukhan & Louhichi [1999]. Our initial results allow us to discuss domain of attraction criteria, as in [Leadbetter & Rootzén, 1998]. Finally, we include a simulation to show the precision of the convergence given in our results.

4.1 Introduction

It is a known fact that, given random variables $X_1, \ldots, X_n$ independent and identically distributed, the distribution of the normalised maximum $M_n = \max\{X_1, \ldots, X_n\}$ (that is, $\mathbb{P}(a_n^{-1}(M_n - b_n) \leqslant x)$, for some $a_n > 0$ and $b_n \in \mathbb{R}$), converges to a non-degenerate distribution G, which is of extremal type. The same result is obtained by Leadbetter et al. [1983] for dependent stationary sequences $(X_i)_{i\in\mathbb{N}}$ under a weak mixing condition. Even more, Leadbetter et al. [1983] expanded this result to stationary processes $\{X_t,\ t \geqslant 0\}$ of continuous time, redefining the maximum as $M_T = \sup\{X_t :\ 0 \leqslant t \leqslant T\}$, again under a weak mixing condition. In the case of stationary Gaussian random fields, we must begin by mentioning Adler [1981, § 6.9], who showed that such result holds for $M_T = \sup\{X_\mathbf{t} : \mathbf{t} \in [0, T]^d\}$ ($d > 1$), under weak covariance conditions, precisely that

$$\mathbb{P}\left(\frac{M_T - b_T}{a_T} \leqslant x\right) \underset{T\to\infty}{\longrightarrow} G(x), \tag{4.1}$$

where $G(x) = \exp(-\exp(-x))$ and whose normalisations are defined by

$$a_T = (2d \ \log T)^{-1/2}$$
$$b_T = (2d \ \log T)^{1/2} + \frac{\frac{1}{2}(d-1)(\log\log T) + \log\left((2\pi)^{-1} \sqrt{\det(\Lambda)(d/\pi)^{d-1}}\right)}{(2d \ \log T)^{1/2}},$$

with Λ denoting the usual matrix of second-order spectral moments. Finally, using all those ideas, Leadbetter & Rootzén [1998] generalised the result to stationary random fields under appropriate long-range dependence restrictions, showing that, if the limit (4.1) holds for some normalisation constants $a_T > 0$ and b_T, then G is of extreme value type. Specifically, they have shown this result considering a weaker version than the usual mixing condition of Rosenblatt [1956], they have called it "*Coordinatewise (Cw) mixing*", which is the usual strong mixing condition (multiplied by the number of sub-blocks of the domain), restricted to events of type $E = \{\max\{X_{\mathbf{t}} : \ \mathbf{t} \in A\} \leqslant u\}$. However, this condition in practice continues to be very difficult to verify. For instance, this can be observed in the only example presented by Ferreira & Pereira [2008], in which is shown how to compute the extremal index of stationary random fields. There they use in the theoretical results this coefficient of Cw-mixing and the random field used in order to verify this mixing property is very very particular. Note also that, Leadbetter & Rootzén [1998] provide no example showing how to verify this condition on their paper. In fact, they leave that point open for those who want to develop techniques to calculate the decay rates of the Cw-mixing coefficients.

A natural idea to attack these limitations could be the use of the results of classical mixing theory for random fields, *e.g* [Doukhan, 1994]. Nevertheless, this continues to be a limitation since the family of mixing random processes and fields is quite restrictive. For instance, the AR(1)-process defined in (2.1) and the random field defined in (4.14) are not mixing.

In view of the above remarks, we will focus on obtaining "extremal types theorem" for the maxima $M_{(n_1,\ldots,n_d)} = \max\left\{X_{\mathbf{t}} : \mathbf{t} \in \prod_{i=1}^{d}\{1,2,\ldots,n_i\}\right\}$ of stationary random fields $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$, under a condition of dependence much easier to verify than Cw-mixing condition and much more general than mixing conditions. For this, we will again use weak dependence in the sense of Doukhan & Louhichi [1999].

This chapter is organised as follows. In Section 4.2 we recall a general result on extremal types theorem provided by Leadbetter & Rootzén [1998]. In Sectiion 4.3 we define weak-dependence for stationary random fields and we provide some examples. Besides, we show an non trivial example of a stationary weakly dependent random field which is not mixing. In Section 4.4 we set the conditions on the weak dependence coefficients in order to obtain the preliminary result (Lemma 4.1), this will be the heart of all other results. In Section 4.5 we provide extremal types theorem and we discuss domain of attraction criteria. A simulation study is included in Section 4.6 using a stationary random field non-mixing. In Section 4.7 we include the proofs.

4.2 A general result on extremal types

The following general extremal types theorem gives the general property to ensure that the maximum of dependent processes and fields have distribution of extreme value type.

Proposition 4.1 (Proposition 2.1 of Leadbetter & Rootzén [1998]) *Let M_T, with $T > 0$, be random variables such that*

$$\mathbb{P}\left(a_T^{-1}(M_T - b_T) \leqslant x\right) \underset{T\to\infty}{\longrightarrow} G(x), \tag{4.2}$$

where G is a non-degenerate distribution and $a_T > 0$, $b_T \in \mathbb{R}$ are normalisation constants. Suppose now that for each real x, $u_T = a_T x + b_T$ and

$$\mathbb{P}\left(M_{kT} \leqslant u_{kT}\right) - \mathbb{P}^k(M_{\phi_k(T)} \leqslant u_{kT}) \underset{T\to\infty}{\longrightarrow} 0 \tag{4.3}$$

holds for each $k = 1, 2, \ldots$ and some continuous strictly increasing functions $\phi_k(T) \underset{T\to\infty}{\longrightarrow} \infty$. Then G is of extreme value type.

Let $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ be a random field. Note that (4.3) is obviously true if $T = n$, $\phi_k(T) = T$ and M_{kn} is the maximum of $k-$independent random blocks $Y_B = (X_i,\ i \in B,\ |B| = n)$. Moreover, observe that the approximation (4.3) suggests decay of dependence. We thus can suppose local dependence and consider small long-range dependence conditions in order to verify such condition (4.3) and develop the extremal types theorem for stationary random fields. In particular, we will develop the results under weak-dependence conditions in the sense of Doukhan & Louhichi [1999]. These weak-dependence conditions have been already introduced in Chapter 2 for random sequences and we will extend it for random fields in the next subsection.

4.3 Weak Dependence of Random Fields

4.3.1 Definitions

Similarly to the Chapter 2, let $\Lambda^u(E)$ be the set of $\mathbb{R}$-valued functions defined on E^u with $u \in \mathbb{N}$ and $E \subseteq \mathbb{R}$, that are bounded by 1 and have a finite Lipschitz modulus $\text{Lip}(\cdot)$, *i.e.*

$$\Lambda^u(E) = \{f : E^u \longrightarrow \mathbb{R} \quad | \quad \text{Lip}(f) < \infty \quad \textit{and} \quad ||f||_\infty \leqslant 1\},$$

where

$$\text{Lip}(f) = \sup_{(x_1,\ldots,x_u)\neq(y_1,\ldots,y_u)} \frac{|f(x_1,\ldots,x_u) - f(y_1,\ldots,y_u)|}{\delta_u\left((x_1,\ldots,x_u),(y_1,\ldots,y_u)\right)},$$

with

$$\delta_r((x_1,\ldots,x_r),(y_1,\ldots,y_r)) = \sum_{i=1}^{r} |x_i - y_i|. \tag{4.4}$$

We will consider E-valued random fields over $\mathbb{Z}^d$, for some $d \in \mathbb{N}$ fixed. In this case, if we set the norm $||(k_1,\ldots,k_d)|| = \delta_d((k_1,\ldots,k_d),\mathbf{0})$ in $\mathbb{Z}^d$, we say that two finite sequences $\mathbf{I} = (\mathbf{i}_1,\ldots,\mathbf{i}_u)$ and $\mathbf{J} = (\mathbf{j}_1,\ldots,\mathbf{j}_v)$ in $\mathbb{Z}^d$ are $l-$distant if

$$\min\{||\mathbf{i}_s - \mathbf{j}_t|| : s = 1,\ldots,u; \quad t = 1,\ldots,v\} = l.$$

Definition 4.1 *Let $\psi : \mathbb{N}^2 \times (\mathbb{R}^+)^2 \to \mathbb{R}^+$ be a function and let $(\epsilon(l))_{l\geqslant 0}$ be a real positive sequence tending to zero. The random fields $X = \{X_\mathbf{t} : \mathbf{t} \in \mathbb{Z}^d\}$ is (ϵ,ψ)-weakly dependent if for any pair of $l-$distant finite sequences $\mathbf{I} = (\mathbf{i}_1,\ldots,\mathbf{i}_u)$, $\mathbf{J} = (\mathbf{j}_1,\ldots,\mathbf{j}_v)$; and any pair of functions $(f,g) \in \Lambda^u(E) \times \Lambda^v(E)$:*

$$\left|Cov(f(X_{\mathbf{i}_1},\ldots,X_{\mathbf{i}_u}),g(X_{\mathbf{j}_1},\ldots,X_{\mathbf{j}_v}))\right| \leqslant \psi\left(u,v,Lip(f),Lip(g)\right)\epsilon(l). \tag{4.5}$$

In particular,
if $\psi(u,v,x,y) = vy$, this is called θ-dependence and $\epsilon(l)$ will be denoted by $\theta(l)$,
if $\psi(u,v,x,y) = ux + vy$, this is called η-dependence and $\epsilon(l)$ will be denoted by $\eta(l)$,
if $\psi(u,v,x,y) = uvxy$, this is called κ-dependence and $\epsilon(l)$ will be denoted by $\kappa(l)$,
if $\psi(u,v,x,y) = ux + vy + uvxy$, this is called λ-dependence and $\epsilon(l)$ will be denoted by $\lambda(l)$.

4.3.2 Examples

In this subsection we give a non-exhaustive list of examples of weakly dependent random fields. In all these examples we consider a centred unit variance independent identically distributed random field $(\xi_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$.

Example 4.1 (Bernoulli shifts) Consider a function $H : \mathbb{R}^{\mathbb{Z}^d} \longrightarrow \mathbb{R}$ and define $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ as

$$X_{\mathbf{i}} = H\left(\xi_{\mathbf{i}-\mathbf{j}},\ \mathbf{j}\in\mathbb{Z}^d\right).$$

For each $l > 0$, define now $X_{\mathbf{i},l} = H\left(\xi_{\mathbf{j}-\mathbf{i}},\ \|j\| \leqslant l\right)$. Observe that $\Delta_p(l) := \|X_{\mathbf{i}} - X_{\mathbf{i},l}\|_p$ does not depend on $\mathbf{i}$. In this case, $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ is $\eta-$weakly dependent with $\eta(l) = 2\Delta_1(l/2-1)$.
In particular, if $d = 1$ and the random field is causal: $X_i = H(\xi_{i-j},\ j \geqslant 0)$, then $\eta(l) = \theta(l) = \Delta_1(l-1)$ and the weak dependence function takes the simple form $\psi(u,v,x,y) = vy$ (See also the dependence property for causal random variables 2.17 and some examples in § 2.3.1).

In order to briefly explain the technique for demonstrating weak-dependence in stationary random fields, we will use the following application. For the general case, the exact same steps can be applied, see [Doukhan & Louhichi, 1999; Doukhan & Lang, 2002; Dedecker et al., 2007].

Application 4.1 (Linear fields) *Define $X = (X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ as*

$$X_{\mathbf{i}} = \sum_{\mathbf{j}\in\mathbb{Z}^d} b_{\mathbf{j}}\xi_{\mathbf{i}-\mathbf{j}}, \tag{4.6}$$

where $\sum_{\mathbf{j}} b_{\mathbf{j}}^2 < \infty$. Then X is $\eta-$weakly dependent, where

$$\eta^2(2l) = 4\sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2. \tag{4.7}$$

Indeed, let $l > 0$ and $X_{\mathbf{i},l} = \sum_{\|\mathbf{j}\|\leqslant l} b_{\mathbf{j}}\xi_{\mathbf{i}-\mathbf{j}}$. Then, we have that

$$\Delta_2^2(l) := \mathbb{E}\left(X_{\mathbf{i},l} - X_{\mathbf{i}}\right)^2 = \mathbb{E}\left(\sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}\xi_{\mathbf{i}-\mathbf{j}}\right)^2 = \sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2\,\mathbb{E}\xi_{\mathbf{i}-\mathbf{j}}^2 = \sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2, \tag{4.8}$$

which does not depend on $\mathbf{i}$ as we expected.
On the other hand, let $\vec{X}_{\mathbf{I}} = (X_{\mathbf{i}_1},\ldots,X_{\mathbf{i}_u})$ and $\vec{X}_{\mathbf{J}} = (X_{\mathbf{j}_1},\ldots,X_{\mathbf{j}_v})$, and consider

their truncated versions $\vec{X}_{\mathbf{I},s} = (X_{\mathbf{i}_1,s}, \ldots, X_{\mathbf{i}_u,s})$ and $\vec{X}_{\mathbf{J},s} = (X_{\mathbf{j}_1,s}, \ldots, X_{\mathbf{j}_v,s})$. Then, if $\mathbf{I} = (\mathbf{i}_1, \ldots, \mathbf{i}_u)$ and $\mathbf{J} = (\mathbf{j}_1, \ldots, \mathbf{j}_v)$ are $l-$distant with $l > 2s$,

$$\mathrm{Cov}\left(f(\vec{X}_{\mathbf{I},s}), g(\vec{X}_{\mathbf{J},s})\right) = 0, \qquad \forall (f,g) \in \Lambda^u(\mathbb{R}) \times \Lambda^v(\mathbb{R}).$$

Thus,

$$\begin{aligned}
\left|\mathrm{Cov}\left(f(\vec{X}_{\mathbf{I}}), g(\vec{X}_{\mathbf{J}})\right)\right| &\leqslant \left|\mathrm{Cov}\left(f(\vec{X}_{\mathbf{I}}) - f(\vec{X}_{\mathbf{I},s}), g(\vec{X}_{\mathbf{J}})\right)\right| + \left|\mathrm{Cov}\left(f(\vec{X}_{\mathbf{I},s}), g(\vec{X}_{\mathbf{J}}) - g(\vec{X}_{\mathbf{J},s})\right)\right| \\
&\leqslant 2\|g\|_\infty \mathbb{E}\left|f(\vec{X}_{\mathbf{I}}) - f(\vec{X}_{\mathbf{I},s})\right| + 2\|f\|_\infty \mathbb{E}\left|g(\vec{X}_{\mathbf{J}}) - g(\vec{X}_{\mathbf{J},s})\right| \\
&\leqslant 2\mathrm{Lip}(f)\mathbb{E}\delta_u(\vec{X}_{\mathbf{I}}, \vec{X}_{\mathbf{I},s}) + 2\mathrm{Lip}(g)\mathbb{E}\delta_v(\vec{X}_{\mathbf{J}}, \vec{X}_{\mathbf{J},s}) \\
&= 2\mathrm{Lip}(f)\sum_{k=1}^{u} \mathbb{E}|X_{\mathbf{i}_k} - X_{\mathbf{i}_k,s}| + 2\mathrm{Lip}(g)\sum_{k=1}^{v} \mathbb{E}|X_{\mathbf{j}_k} - X_{\mathbf{j}_k,s}| \\
&\leqslant 2(u\mathrm{Lip}(f) + v\mathrm{Lip}(g))\Delta_p(s), \qquad (4.9)
\end{aligned}$$

for any $p \geqslant 1$. Therefore, it is enough to choose $\eta(l) = 2\Delta_1(l/2 - 1)$. In particular, as the series b_k is square summable, we can take $p = 2$ to obtain from (4.8) that

$$\eta^2(2l) = 4\Delta_2^2(l) = 4\sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2. \qquad (4.10)$$

Application 4.2 (Markovian fields) *Let $\mathbf{v} \in \mathbb{Z}^d$, define a shift operator $B_{\mathbf{v}}$ in the fields on $\mathbb{Z}^d$ as $(B_{\mathbf{v}} \cdot X)_{\mathbf{t}} = X_{\mathbf{t}-\mathbf{v}}$. Now, consider a finite sequence of reals $(a_j)_{j=1,\ldots,D}$ and a finite sequence $(\mathbf{v}_1, \ldots, \mathbf{v}_D) \in (\mathbb{Z}^d)^D$. A Markovian field is defined by the neighbour regression formula:*

$$X_{\mathbf{i}} = \sum_{j=1}^{D} a_j X_{\mathbf{i}-\mathbf{v}_j} + \xi_{\mathbf{i}} = (A \cdot X)_{\mathbf{i}} + \xi_{\mathbf{i}}, \qquad (4.11)$$

where $A = \sum_{j=1}^{D} a_j B_{\mathbf{v}_j}$. Assume that $a = \sum_{j=1}^{D} |a_j| < 1$, then there exists an integrable stationary solution to (4.11):

$$X_{\mathbf{i}} = \sum_{p=0}^{\infty} (A^p \cdot \xi)_{\mathbf{i}} = \sum_{p=0}^{\infty} \sum_{\substack{0 \leqslant j_1,\ldots,j_D \\ j_1+\cdots+j_D=p}} \frac{p!}{j_1! \cdots j_D!} a_1^{j_1} \cdots a_D^{j_D} \xi_{\mathbf{i}-(j_1\mathbf{v}_1+\cdots+j_D\mathbf{v}_D)}. \qquad (4.12)$$

Note that

$$\sum_{\substack{0 \leqslant j_1,\ldots,j_D \\ j_1+\cdots+j_D=p}} \frac{p!}{j_1! \cdots j_D!} |a_1^{j_1} \cdots a_D^{j_D}| = a^p,$$

therefore the process (4.11) can be rewritten as (4.6) with absolutely summable coefficients, i.e

$$X_{\mathbf{i}} = \sum_{\mathbf{k} \in \mathbb{Z}^d} b_{\mathbf{k}} \xi_{\mathbf{i}-\mathbf{k}},$$

where the series $(b_{\mathbf{k}})$ are, by definition, such that

$$|b_{\mathbf{k}}| \leqslant \sum_{p=0}^{\infty} \sum_{(j_1,\dots,j_D)\in V_{\mathbf{k},p}} \frac{p!}{j_1!\cdots j_D!} |a_1^{j_1}\cdots a_D^{j_D}|,$$

with $V_{\mathbf{k},p} := \{(j_1,\dots j_D) \in \mathbb{N}^D \ : \ j_1+\cdots+j_D = p,\ j_1\mathbf{v}_1+\cdots j_D\mathbf{v}_D = \mathbf{k}\}$. Finally, denoting $v = \max\{\|\mathbf{v}_1\|_\infty,\dots,\|\mathbf{v}_D\|_\infty\}$, observe that $V_{k,p}$ is empty if $p < \|\mathbf{k}\|_\infty/v$, and $|b_{\mathbf{k}}| \leqslant (1-a)^{-1}a^{\|\mathbf{k}\|_\infty/v}$. So that $(b_{\mathbf{k}})$ is square summable and $\eta-$weakly dependent such that (4.7) holds.

Example 4.2 (Chaotic Volterra fields) Assume that $(\xi_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ has finite moments of any order. Define $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ as

$$X_{\mathbf{i}} = \sum_{s=1}^{\infty} X_{\mathbf{i}}^{(s)} \qquad \text{where} \quad X_{\mathbf{i}}^{(s)} = \sum_{\mathbf{j}_1,\dots,\mathbf{j}_s\in\mathbb{Z}^d} a^{(s)}_{\mathbf{j}_1,\dots,\mathbf{j}_s} \xi_{\mathbf{i}-\mathbf{j}_1}\cdots\xi_{\mathbf{i}-\mathbf{j}_s}. \tag{4.13}$$

The real series $a^{(s)}_{\mathbf{j}_1,\dots,\mathbf{j}_s}$ are absolutely summable. Moreover, the field $(X_{\mathbf{i}})$ is $\eta-$weakly dependent with

$$\eta(2l) \leqslant \sum_{s=1}^{\infty}\sum_{l=1}^{s} \sum_{\substack{\|\mathbf{j}_l\|>l \\ \mathbf{j}_1,\dots,\mathbf{j}_s\in\mathbb{Z}^d}} |a^{(s)}_{\mathbf{j}_1,\dots,\mathbf{j}_s}|\mathbb{E}|\xi_{\mathbf{0}}|^s.$$

Note that if the random field is causal, that is, if the indices $\mathbf{j}_l$ are all in $\mathbb{N}^d$, then the bound holds for $\eta(l)$.

Application 4.3 (*$ARCH(\infty)-$*fields) *Let a be a real positive, $(b_{\mathbf{j}})_{\mathbf{j}\in T^+}$ a nonnegative sequence and $(\varsigma_{\mathbf{j}})_{\mathbf{j}\in T^+}$ an i.i.d nonnegative random field where $T^+ = \mathbb{N}^d\setminus\{\mathbf{0}\}$. Define $X = (X_{\mathbf{i}})_{\mathbf{i}}$ through the recurrence relation*

$$X_{\mathbf{i}} = \left(a + \sum_{\mathbf{j}\in T^+} b_{\mathbf{j}} X_{\mathbf{i}-\mathbf{j}}\right)\varsigma_{\mathbf{i}}.$$

If $c = \mathbb{E}(\varsigma_{\mathbf{0}})\sum_{\mathbf{j}\in T^+} b_{\mathbf{j}} < 1$, Doukhan et al. [2006] proved that such models have a stationary representation with the chaotic expansion

$$X_{\mathbf{i}} = a\varsigma_{\mathbf{i}} + a\sum_{l=1}^{\infty} X_{\mathbf{i}}^l, \qquad \textit{where}$$
$$X_{\mathbf{i}}^l = \sum_{\mathbf{j}_1\in T^+}\cdots\sum_{\mathbf{j}_l\in T^+} b_{\mathbf{j}_1}\cdots b_{\mathbf{j}_l}\varsigma_{\mathbf{i}-\mathbf{j}_1}\cdots\varsigma_{\mathbf{i}-(\mathbf{j}_1+\cdots+\mathbf{j}_l)}.$$

Note that the convergence of the series over l comes from $\mathbb{E}|X^l_{\mathbf{i}}| \leqslant c^l$. Now, let

$$X^{L,m}_{\mathbf{i}} = a\varsigma_{\mathbf{i}} + a\sum_{l=1}^{L}\sum_{\mathbf{j}_1\in[0:m]^d}\cdots\sum_{\mathbf{j}_l\in[0:m]^d} b_{\mathbf{j}_1}\cdots b_{\mathbf{j}_l}\varsigma_{\mathbf{i}-\mathbf{j}_1}\cdots\varsigma_{\mathbf{i}-(\mathbf{j}_1+\cdots+\mathbf{j}_l)},$$

with the notation $[0:m]=\{0,1,\ldots,m\}$. Observe that this field $(X^{L,m}_{\mathbf{i}})_{\mathbf{i}}$ is an approximation of X, such that

$$\mathbb{E}\left|X_{\mathbf{i}} - X^{L,m}_{\mathbf{i}}\right| \leqslant \sum_{l>L} c^l + \rho(m),$$

where $\rho(m) = \sum_{\mathbf{j}\notin[0:m]^d}|b_{\mathbf{j}}|$. Then, X is $\eta-$weakly dependent, with coefficient

$$\eta(l) = \min_{Lm\leqslant l}\left(\frac{c^{L+1}}{1-c} + \rho(m)\right).$$

In particular, for standard ARCH models with delay p (this is, $b_{\mathbf{j}} = 0$ for all $\|\mathbf{j}\| > p$), the coefficient $\eta(l) = c^{l/p}/(1-c)$. Besides, in the arithmetic decay case, $b_{\mathbf{j}} = C\|\mathbf{j}\|^{-a}$, $\eta(l) = Const.\left(\frac{l}{\log l}\right)^{1-a}$. In the geometric case, $b_{\mathbf{j}} = C\exp(-b\|\mathbf{j}\|)$, $\eta(l) = Const.\gamma^{\sqrt{l}}$, for $\gamma = \exp(-\sqrt{-b\log c})$.

Example 4.3 (Associated random field) Associated random fields are $\kappa-$weakly dependent, with $\kappa(l) = \sum_{\|\mathbf{j}\|>l}\mathrm{Cov}(X_{\mathbf{0}}, X_{\mathbf{j}})$. See [Doukhan & Louhichi, 1999].

For other examples of weakly dependent random fields see [Doukhan & Lang, 2002], [Doukhan & Truquet, 2007] and [Dedecker et al., 2007]. In this references they also include examples of weakly dependent random fields with weakly dependent noise $\xi = (\xi_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$.

4.3.3 A non trivial example of non-mixing weakly dependent linear field

Let $\xi = (\xi_i)_{i\in\mathbb{Z}}$ and $\varsigma = (\varsigma_j)_{j\in\mathbb{Z}}$ be two i.i.d Bernoulli sequences with the same parameter $p = 1/2$ such that ξ and ς are independents. Define $(U_i)_{i\in\mathbb{Z}}$ and $(V_j)_{j\in\mathbb{Z}}$ as

$$U_i = \sum_{k=0}^{\infty}\frac{\xi_{i-k}}{2^k}, \qquad V_j = \sum_{k=0}^{\infty}\frac{\varsigma_{j-k}}{2^k}.$$

We have already mentioned in previous chapters that U_i and V_j are two uniform processes which are weakly dependent (see [Doukhan & Louhichi, 1999] and [Dedecker

& Prieur, 2004a]) which are non-mixing (because U_0 is a deterministic function of U_i for any $i > 0$. The same for V, see [Andrews, 1984]). Consider now the random field

$$X_{(i,j)} = U_i V_j = \sum_{k \geqslant 0, l \geqslant 0} \frac{\xi_{i-k} \varsigma_{j-l}}{2^{k+l}}. \tag{4.14}$$

This is clearly a stationary linear random field with innovations $\xi_i \varsigma_j$.
Note that

$$\Delta_2^2(r) = \mathbb{E} \left| X_{(i,j)} - \sum_{0 \leqslant k+l \leqslant r} \frac{\xi_{i-k} \varsigma_{j-l}}{2^{k+l}} \right|^2 = \sum_{\substack{k+l>r \\ k \geqslant 0, l \geqslant 0}} 4^{-(k+l)},$$

therefore $\Delta_2(r) \leqslant \sqrt{2r} 2^{-r}$. Using (4.7), it is proven that the random field $(X_{(i,j)})_{i,j \in \mathbb{Z}}$ is $\eta-$weakly dependent such that $\eta(r) = 2^{2-r/2}\sqrt{r}$. However, this field is non-mixing, as this is proven in [Doukhan & Lang, 2002, § 2.2.6].

4.4 Asymptotic max-independence in stationary weakly-dependent random fields

In this section we provide the main preliminary results in order to prove extremal types theorem in the next section.
For this, consider the following notations, definitions and conditions.

Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field. For subsets B of $\mathbb{Z}^d$, we denote

$$M(B) = \sup\{X_{\mathbf{t}} : \mathbf{t} \in B\},$$

and we will write

$$M_{\mathbf{n}} = M(E_{\mathbf{n}}),$$

where $\mathbf{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{N}^d$ and $E_{\mathbf{n}} = \prod_{i=1}^d [n_i]$, with the notations: $[k] := [1 : k]$ and $[i : j] := \{i, i+1, \ldots, j-1, j\} \subset \mathbb{N}$.

Weak dependence conditions will be given here in such a way that the extremal types theorem holds for $M_{\mathbf{n}}$, *i.e.* such that any non-degenerate limit G for $M_{\mathbf{n}}$ normalised as in (4.2) must be of extreme value type. A similar result is developed in [Leadbetter & Rootzén, 1998] but under mixing conditions (more precisely, under "Cw-mixing" conditions).

We will use a version spatial of the Bernstein block technique for the proof of the main result. In order to develop this technique here, it is necessary to consider for each $i \in [d]$ a sequence $r_i := r_{n_i} = o(n_i)$ to build the lengths of sides of $d-$blocks (or $d-$grids):

$$B_{j_1 j_2 \dots j_d} := \prod_{i=1}^{d} [(j_i - 1)r_i + 1 : j_i r_i],$$

which will be used for subdivision of $E_{\mathbf{n}} = \prod_{i=1}^{d} [n_i]$. Moreover, we denote $m_i = m_{n_i} = \lceil n_i / r_i \rceil$ for $i \in [d]$ with $\lceil x \rceil := \max\{j \in \mathbb{N} : j \leqslant x\}$. Then, it is clear that $E_{\mathbf{n}}$ contains $m_{\mathbf{n}} = m_1 \cdot m_2 \cdots m_d$ complete blocks, and no more that $(m_1 + m_2 + \cdots + m_d - d + 1)$ incomplete ones.

SWD conditions. Let $l_i := l_{n_i} \underset{n_i \to \infty}{\longrightarrow} \infty$ be a sequence such that $l_i = o(r_i)$ for each $i \in [d]$. We say that a random field X satisfy at least one SWD condition if X satisfy at least one of the following dependence conditions:

1. θ-weakly dependent such that for each $i-$direction, with $i \in [d]$,

$$\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1}}{m_{i-1} \cdots m_2 m_1} \theta(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \dots, n_i) \to \infty; \tag{4.15}$$

   for some $(\alpha_1, \dots, \alpha_i) \in [1, \infty)^i \setminus \{(1, \dots, 1)\}$.

2. η-weakly dependent such that for each $i-$direction, with $i \in [d]$,

$$\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1} m_i^{\beta}}{m_{i-1} \cdots m_2 m_1} \eta(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \dots, n_i) \to \infty; \tag{4.16}$$

   for some $(\alpha_1, \dots, \alpha_i, \beta) \in [1, \infty)^{i+1} \setminus \{(1, \dots, 1)\}$.

3. κ-weakly dependent such that for each $i-$direction, with $i \in [d]$,

$$\left(\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1}}{m_{i-1} \cdots m_2 m_1} \right)^2 \kappa(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \dots, n_i) \to \infty; \tag{4.17}$$

4. λ-weakly dependent such that for each $i-$direction, with $i \in [d]$,

$$\left(\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1}}{m_{i-1} \cdots m_2 m_1} \right)^2 \lambda(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \dots, n_i) \to \infty; \tag{4.18}$$

   for some $(\alpha_1, \dots, \alpha_i) \in [1, \infty)^i \setminus \{(1, \dots, 1)\}$.
   In these items, $\mathbf{n} = (n_1, n_2, \dots, n_d) \to \infty$ means that $n_j \to \infty$ for each $j \in [d]$. Besides, we set the convention $m_0 = 1$.

Lemma 4.1 *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field that satisfy at least one SWD condition. Let $(u_{\mathbf{n}})_{\mathbf{n}\in\mathbb{N}^d}$ be a family of levels such that $\mathbb{P}(M(B) = u_{\mathbf{n}}) = 0$, for all $B \subset \mathbb{Z}^d$ and all $\mathbf{n} \in \mathbb{N}^d$. Then if $B_{\mathbf{n}} = \prod_{i=1}^d [m_i r_i]$ (this is, $B_{\mathbf{n}} = \bigcup_{j_1,\cdots,j_d} B_{j_1,\cdots,j_d}$),*

$$\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) = \mathbb{P}^{m_{\mathbf{n}}}(M(\mathbf{J}) \leqslant u_{\mathbf{n}}) + o(1) \tag{4.19}$$

as $\mathbf{n} \to \infty$, where $\mathbf{J} := B_{11\ldots 1} = \prod_{i=1}^d [r_i]$.

Proposition 4.2 *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field that satisfy at least one SWD condition. Suppose that $(u_{\mathbf{n}})_{\mathbf{n}\in\mathbb{N}^d}$ is a family of levels such that $\mathbb{P}(M(B) = u_{\mathbf{n}}) = 0$, for all $B \subset \mathbb{Z}^d$ and all $\mathbf{n} \in \mathbb{N}^d$. Let $\mathbf{I} = \prod_{i=1}^d [\lceil a_i n_i \rceil]$ (with $a_i \in (0,1]$ for all $i \in [d]$), be a $d-$sub-grid of $E_{\mathbf{n}}$, where a_i may change with $\mathbf{n}$ for all $i \in [n]$ but $a_1 a_2 \cdots a_d \longrightarrow a > 0$ as $\mathbf{n} \to \infty$. Then, under the established notation*

1. $\mathbb{P}(M(\mathbf{I}) \leqslant u_{\mathbf{n}}) - \mathbb{P}^{a m_{\mathbf{n}}}(M(\mathbf{J}) \leqslant u_{\mathbf{n}}) \longrightarrow 0$
2. $\mathbb{P}(M(\mathbf{I}) \leqslant u_{\mathbf{n}}) - \mathbb{P}^{a}(M_{\mathbf{n}} \leqslant u_{\mathbf{n}}) \longrightarrow 0.$

Note that if $\mathbb{P}(M_{\mathbf{n}} \leqslant u_{\mathbf{n}})$ has a limit $G(x)$, with $u_{\mathbf{n}} = a_{\mathbf{n}} x + b_{\mathbf{n}}$, then, from 2. in the previous proposition, $\mathbb{P}(M(\mathbf{I}) \leqslant u_{\mathbf{n}})$ has the limit $G^a(x)$, which is used to show max-stability of G in the proof of Theorem 4.1

4.5 Extremal types theorem and domain of attraction criteria

As we have said before, $\mathbf{n} \to \infty$ means $n_i \to \infty$ for each $i \in [d]$. However, the extremal types theorem for stationary random fields can be reformulated in terms of the limiting distribution of $M_{\mathbf{n}}$ as $\mathbf{n} \to \infty$ along a monotone path on the grid $\mathbb{N}^d$, *i.e.* along $\mathbf{n} = (n, \lceil \vartheta_1(n) \rceil, \ldots, \lceil \vartheta_{d-1}(n) \rceil)$ for some strictly increasing continuous functions $\vartheta_j : [1,\infty) \longrightarrow [1,\infty)$, with $j \in [d-1]$, such that $\vartheta_j(T) \longrightarrow \infty$ as $T \to \infty$, for all $j \in [d-1]$.

Theorem 4.1 *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field and suppose that $\mathbb{P}\left(a_{\mathbf{n}}^{-1}(M_{\mathbf{n}} - b_{\mathbf{n}}) \leqslant x\right) \longrightarrow G(x)$, non-degenerate as $\mathbf{n} \to \infty$ along the monotone path $\mathbf{n} = (n, \lceil \vartheta_1(n) \rceil, \ldots, \lceil \vartheta_{d-1}(n) \rceil)$ defined above. Then, if X satisfies at least one SWD condition and $\mathbb{P}(M(B) = a_{\mathbf{n}} x + b_{\mathbf{n}}) = 0$, for all $(x, n) \in \mathbb{R} \times \mathbb{N}$ and all $B \subset \mathbb{Z}^d$, G is of extreme value type.*

4.5.1 On domain of attraction criteria

The purpose of this section is to provide a characterisation of the maximum domain of attraction of a extreme value distribution G.

Since the random field is stationary, this characterisation will be an analogous version to the one already made for the cases: i.i.d. random variables, stationary sequences with non-zero extremal index, etc. (See [Leadbetter et al., 1983; Leadbetter & Rootzén, 1998]). Namely, under SWD conditions, we obtain that the type of limiting distribution for maxima is also determined by the tail behaviour of the common marginal distribution function for each either term or each maximum over (fixed) sub-blocks.

Note that from Lemma 4.1, we can deduce that the random variables

$$\left(M(B_{j_1,j_2,\ldots,j_d})\right)_{j_1,\ldots,j_d\in[d]}$$

have extremal index 1. Therefore, using this "max-block asymptotic independence", we can provide the characterisation through the tail distribution function of $M(\mathbf{J})$, where, by stationarity, $\mathbf{J} := B_{11\ldots1} = \prod_{i=1}^{d}[r_i]$ denotes "the first block" or a generic block. Let now $\gamma_{\mathbf{n}}$ be the $(1-m_{\mathbf{n}}^{-1})-$percentile of $M(\mathbf{J})$, *i.e.* $\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right) = m_{\mathbf{n}}^{-1}$.

With the above notation, we state the following proposition.

Proposition 4.3 *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field that satisfies at least one SWD condition. Suppose that*

(TD) $\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}} + a_{\mathbf{n}}x\right)/\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right) \longrightarrow H(x)$ *for some constants $a_{\mathbf{n}} > 0$ and some non-increasing function $H(x)$ such that $H(x) \underset{x\to-\infty}{\longrightarrow} \infty$ and $H(x) \underset{x\to\infty}{\longrightarrow} 0$;*

and that additionally $\mathbb{P}(M(B) = a_{\mathbf{n}}x + \gamma_{\mathbf{n}}) = 0$, for all $B \in \mathbb{Z}^d$ and all $(x,\mathbf{n}) \in \mathbb{R}\times\mathbb{N}^d$. Then

$$\mathbb{P}\left(a_{\mathbf{n}}^{-1}(M_{\mathbf{n}} - \gamma_{\mathbf{n}}) \leqslant x\right) \underset{\mathbf{n}\to\infty}{\longrightarrow} G(x) = \exp(-H(x)). \tag{4.20}$$

Remark 4.1 Note that if F is the distribution function of $M(\mathbf{J})$ for fixed $\mathbf{J}$ and if the assumptions of the previous proposition are satisfied, then the relation (4.20) implies that F belongs to the domain of attraction of G.

4.6 Simulation study

In this section, we estimate the generalised extreme value (GEV) distribution function of the non-mixing stationary random field (4.14).

For this, we simulate $N = 200$ samplings of this random field (4.14), restricted to the domain $E = E_{(10^4,10^4)} = [10^4] \times [10^4]$, and we calculate the respective N maximum values. Applying the maximum-likelihood estimator, we obtain the estimates shown in Table 4.1 with estimated parameter covariance matrix in Table 4.2.

	location (μ)	scale (σ)	shape (γ)
Estimated parameters:	0.99953743	0.00031833	-0.74907362
Standard Error Estimates:	0.00000002	0.00000002	0.005866192

Table 4.1: Maximum-likehood estimates for the parameters of the GEV df of the random field (4.14) over the block $E_{(10^4,10^4)}$.

	location	scale	shape
location	3.999999e-16	-8.108354e-24	-3.864037e-15
scale	-8.108354e-24	3.999999e-16	-8.674275e-17
shape	-3.864037e-15	-8.674275e-17	3.441221e-05

Table 4.2: Estimated parameter covariance matrix in Table 4.1.

Besides, in Figure 4.1, observe that the simulated data of the maximum of the random field taken in this experiment is well fitted to the GEV distribution function $G(x) = G(x; \mu, \sigma, \gamma)$, with the estimated parameters (μ, σ, γ) shown in Table 4.1. Note that G is here of Weibull type with parameter $\alpha = -\gamma^{-1} = 1.334982$.

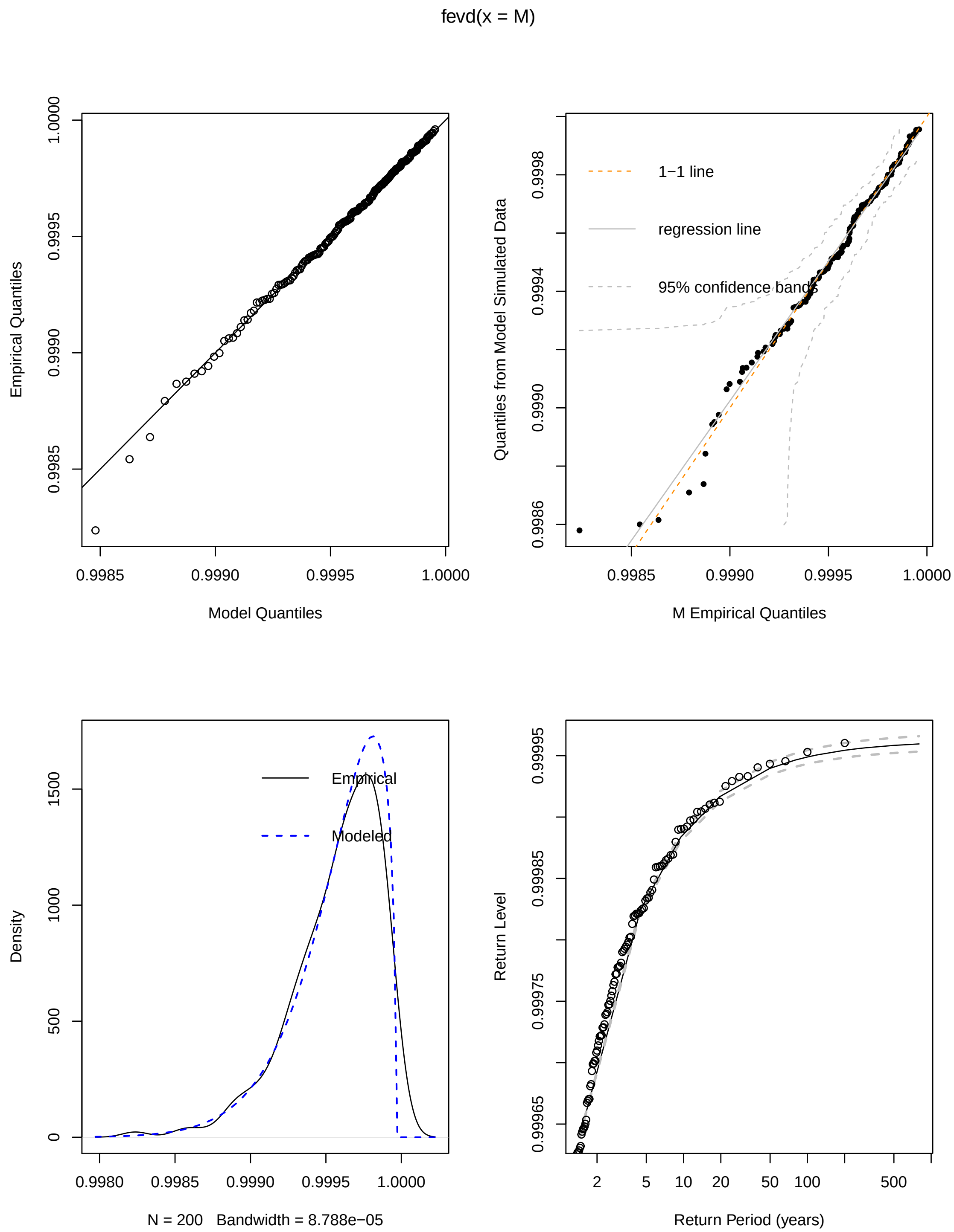


Figure 4.1: Diagnostics from the GEV df fitted to the maximum of the random field (4.14) over the block $E = E_{(10^4,10^4)}$. Quantile-quantile plot (top left), quantiles from a sample drawn from the fitted GEV df against the empirical data quantiles with 95% confidence bands (top right), density plots of empirical data and fitted GEV df (bottom left), and return level plot with 95% point-wise normal approximation confidence intervals (bottom right).

4.7 Proofs

Without loss of generality, we will consider $d = 2$ to reduce notations and simplify the proof of Lemma 4.1.

In the proof of Lemma 4.1 we approximate the indicatrices functions by Lipschitz functions in order to be able to use the weak-dependence conditions.
For this approximation, let $\mathbb{M} := \mathbb{M}_{a\times b}(\mathbb{R})$ be the set of real-matrices with a rows and b columns. For $u > 0$, let $f_{u,a,b} : \mathbb{M} \longrightarrow \mathbb{R}$ be a function defined by

$$f_{u,a,b}((x_{ij})_{ij}) = \mathbb{1}_{\left\{\max\{x_{ij}\ :\ (i,j)\in[a]\times[b]\}\leqslant u\right\}}. \tag{4.21}$$

Clearly the discontinuity points of $f_{u,a,b}$ are the matrices $(x_{ij})_{ij} \in \mathbb{M}$ such that

$$\max\{x_{ij}\ :\ (i,j) \in [a] \times [b]\} = u,$$

i.e. the set of discontinuities of $f_{u,a,b}$ is given by

$$D(f_{u,a,b}) := \left\{(x_{ij})_{ij} \in \mathbb{M}\ :\ \max\{x_{ij}\ :\ (i,j) \in [a] \times [b]\} = u\right\}.$$

Let $K_{a,b} : \mathbb{M} \longrightarrow \mathbb{R}$ be a function defined as:

$$K_{a,b}((x_{ij})_{ij}) := \mathbb{1}_{\left\{\max\{x_{ij}\ :(i,j)\in[a]\times[b]\}\leqslant -1\right\}} - \frac{1}{2}\sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-1)\mathbb{1}_{\left\{|x_{ij}|<1,\ x_{ij}\geqslant x_{kl}\ \forall (k,l)\neq(i,j)\right\}}. \tag{4.22}$$

Now, if we consider a positive sequence h_n converging to zero, as $n \to \infty$, then the sequence of Lipschitz functions

$$K_{n,a,b}(M) := K_{a,b}(h_n^{-1}(M - \mathbb{U})), \tag{4.23}$$

converges uniformly to $f_{u,a,b}$ on $\mathbb{M} \setminus D(f_{u,a,b})$, where $\mathbb{U} := (u_{ij})_{ij} \in \mathbb{M}$ is such that $u_{ij} = u$ for all $(i,j) \in [a] \times [b]$. Moreover, note that

$$Lip(K_{n,a,b}) \leqslant h_n^{-1}. \tag{4.24}$$

Proof of Lemma 4.1. Let $\mathbf{J}_i = [(i-1)r_1 + 1 : ir_1] \times [m_2 r_2]$, $\mathbf{J}'_i = [(i-1)r_1 + l_1 : ir_1 - l_1] \times [m_2 r_2]$ and $\mathbf{J}^*_i = \mathbf{J}_i \setminus \mathbf{J}'_i$.
It is evident that $B_{\mathbf{n}} = \bigcup_{i=1}^{m_1} \mathbf{J}_i$ and

$$\mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) = \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}_m) \leqslant u_{\mathbf{n}}\right)$$
$$\leqslant \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right), \tag{4.25}$$

for any $m \in [2 : m_1]$. Thus

$$\begin{aligned}
0 &\leqslant \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) \\
&= \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\ ,\ M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) > u_{\mathbf{n}}\right) \\
&\leqslant \mathbb{P}(M(\mathbf{J}^*_m) > u_{\mathbf{n}}) = \mathbb{P}(M(\mathbf{J}^*_1) > u_{\mathbf{n}}). \qquad (4.26)
\end{aligned}$$

On the other hand, observe that

$$\begin{aligned}
0 &\leqslant \mathbb{P}(M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}) - \mathbb{P}(M(\mathbf{J}_m) \leqslant u_{\mathbf{n}}) \\
&\leqslant \mathbb{P}(M(\mathbf{J}^*_m) > u_{\mathbf{n}}) = \mathbb{P}(M(\mathbf{J}^*_1) > u_{\mathbf{n}}) \qquad (4.27)
\end{aligned}$$

and that

$$\begin{aligned}
\Delta_{m,x} &:= \left|\mathrm{Cov}\left(K_{\mathbf{n},m_2r_2,(m-1)r_1}\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right), K_{\mathbf{n},m_2r_2,r_1-2l_1}\left(\mathbf{J}'_m\right)\right)\right| \\
&= \left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_k) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\mathbb{P}\left(M(\mathbf{J}'_k) \leqslant u_{\mathbf{n}}\right)\right| \\
&\quad + o(1). \qquad (4.28)
\end{aligned}$$

Using stationarity and (4.26)-(4.28), it follows that

$$\begin{aligned}
&\left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\mathbb{P}\left(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}\right)\right| \\
&= \left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\mathbb{P}\left(M(\mathbf{J}_m) \leqslant u_{\mathbf{n}}\right)\right| \\
&\leqslant \left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right)\right| \\
&+ \left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\mathbb{P}\left(M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right)\right| \\
&\quad + \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\left|\mathbb{P}\left(M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M(\mathbf{J}_m) \leqslant u_{\mathbf{n}}\right)\right| \\
&\leqslant \left[1 + \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\right]\mathbb{P}\left(M(J^*_1) > u_{\mathbf{n}}\right) + \Delta_{m,x} + o(1), \qquad (4.29)
\end{aligned}$$

which implies that

$$\begin{aligned}
\delta_{\mathbf{n}} &:= |\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) - \mathbb{P}^{m_1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}})| \\
&= \left| \sum_{j=1}^{m_1-1} \left[\mathbb{P}\left(M\left(\bigcup_{i=1}^{m_1-j+1} \mathbf{J}_i \right) \leqslant u_{\mathbf{n}} \right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m_1-j} \mathbf{J}_i \right) \leqslant u_{\mathbf{n}} \right) \mathbb{P}(M(\mathbf{J}_{m_1-j+1}) \leqslant u_{\mathbf{n}}) \right] \right. \\
&\qquad \left. \times \mathbb{P}^{j-1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) \right| \\
&\leqslant \sum_{j=1}^{m_1-1} \left[1 + \mathbb{P}\left(M\left(\bigcup_{i=1}^{j} \mathbf{J}_i \right) \leqslant u_{\mathbf{n}} \right) \right] \mathbb{P}\left(M(J_1^*) > u_{\mathbf{n}} \right) + \sum_{j=1}^{m_1-1} \Delta_{j+1,x} + o(1) \\
&\leqslant 2(m_1-1)\mathbb{P}\left\{ M(J_1^*) > u_{\mathbf{n}} \right\} + \sum_{j=1}^{m_1-1} \Delta_{j+1,x} + o(1). \qquad (4.30)
\end{aligned}$$

We will show that $\delta_{\mathbf{n}} \longrightarrow 0$ as $\mathbf{n} \to \infty$.

- Decay of the dependence term.

  First, if we suppose that X is η-weakly-dependent, we obtain for $j \in [2 : m_1]$ that

$$\begin{aligned}
\Delta_{j,x} &\leqslant \psi\left(\mathrm{Lip}(K_{\mathbf{n},m_2r_2,(j-1)r_1}), \mathrm{Lip}(K_{\mathbf{n},m_2r_2,r_1-2l_1}), m_2(j-1)r_2r_1, m_2r_2(r_1-2l_1) \right) \eta(l_1) \\
&\leqslant \frac{m_2r_2r_1(j-2l_1/r_1)}{h_{\mathbf{n}}} \eta(l_1), \qquad (4.31)
\end{aligned}$$

  thus

$$\begin{aligned}
\sum_{j=1}^{m_1-1} \Delta_{j+1,x} &\leqslant \frac{n_2n_1m_1}{h_n} \left[\frac{1-m_1^{-1}}{2} + \left(\frac{1}{m_1} - \frac{2l_1}{r_1m_1} \right) \left(1 - \frac{1}{m_1} \right) \right] \eta(l_1) \\
&\leqslant 2^{-1} n_2 n_1^{\alpha_1} m_1^{\beta} \left[1 + \frac{2}{m_1} \left(1 - \frac{2l_1}{r_1} \right) \right] \eta(l_1), \qquad (4.32)
\end{aligned}$$

  where we have set here $h_{\mathbf{n}} = n_1^{1-\alpha_1} m_1^{1-\beta}$, for some $(\alpha_1, \beta) \in [1, \infty)^2 \setminus \{(1,1)\}$

  In the same way, if we suppose that X is respectively κ, λ or $\theta-$weakly dependent, we have respectively that

$$\begin{aligned}
\sum_{j=1}^{m_1-1} \Delta_{j+1,x} &\leqslant \frac{n_2^2 n_1^2 (1-2l_1/r_1)(1-1/m_1)}{2h_{\mathbf{n}}^2} \kappa(l_1) \\
&\leqslant 2^{-1} \left(n_2 n_1^{\alpha_1} \right)^2 (1-2l_1/r_1) \kappa(l_1), \qquad (4.33)
\end{aligned}$$

$$\sum_{j=1}^{m_1-1} \Delta_{j+1,x} \leqslant \left[\frac{n_2 n_1 m_1}{2h_n} \left(1 + \frac{2(1-2l_1/r_1)}{m_1}\right) + \frac{n_2^2 n_1^2 (1-2l_1/r_1)}{2h_{\mathbf{n}}^2} \right] \lambda(l_1)$$
$$\leqslant 2^{-1} \left(n_2 n_1^{\alpha_1}\right)^2 \left[\frac{h_n m_1}{n_2 n_1} \left(1 + \frac{2(1-2l_1/r_1)}{m_1}\right) + (1-2l_1/r_1) \right] \lambda(l_1). \quad (4.34)$$

$$\sum_{j=1}^{m_1-1} \Delta_{j+1,x} \leqslant \frac{n_2 n_1}{h_n} \left(1 - \frac{2l_1}{r_1}\right) \left(1 - \frac{1}{m_1}\right) \theta(l_1)$$
$$\leqslant n_2 n_1^{\alpha_1} \left(1 - \frac{2l_1}{r_1}\right) \left(1 - \frac{1}{m_1}\right) \theta(l_1) \quad (4.35)$$

where we have set $h_{\mathbf{n}} = n_1^{1-\alpha_1}$ for κ and λ cases, for some $\alpha_1 \in (1, \infty)$. For the $\theta-$weakly dependent case, it is suffice to choose $h_{\mathbf{n}} = n_1^{1-\alpha_1}$, for some $\alpha_1 \in (1, \infty)$.

- Decay of the expectation of the excesses in the remaining blocks J^*.

We only need consider that $\mathbb{P}^{m_1}\{M(J_1^*) \leqslant u_{\mathbf{n}}\} \longrightarrow \rho$ for some $\rho \in [0,1]$. In fact, if $\rho = 1$, then $m_1 \log \mathbb{P}\{M(\mathbf{J}_1^*) \leqslant u_{\mathbf{n}}\} \longrightarrow 0$ as $\mathbf{n} \to \infty$. Therefore, $m_1 \mathbb{P}\{M(\mathbf{J}_1^*) > u_{\mathbf{n}}\} \longrightarrow 0$ as $\mathbf{n} \to \infty$. Using this and the respective SWD condition (4.15 - 4.18), it follows that $\delta_{\mathbf{n}} \longrightarrow 0$ by (4.30).

Now, we suppose that $\rho < 1$. Note that for $\mathbf{n}$ sufficiently large, we can choose $k_1 = k_{1,\mathbf{n}}$ rectangles congruent $\{\mathbf{J}_{1j}^*\}_{j \in [k_1]}$ to $\mathbf{J}_1^*$ inside $\mathbf{J}_1'$, such that all are mutually separated by at least l_1 points in the $1-$direction. Using similar arguments to (4.26)-(4.30), but in the (4.26) and (4.27) cases we only use the left hand inequality, we obtain

$$\mathbb{P}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) \leqslant \mathbb{P}^{k_1}(M(\mathbf{J}_1^*) \leqslant u_{\mathbf{n}}) + \sum_{j=1}^{k_1-1} \tilde{\Delta}_{j+1,x} + o(1) \quad (4.36)$$

where

$$\tilde{\Delta}_{j,x} = \mathrm{Cov}\left(K_{n,m_2 r_2, 2(j-1)l_1}\left(\bigcup_{i=1}^{j-1} \mathbf{J}_{1i}^* \right), K_{n,m_2 r_2, 2l_1}(\mathbf{J}_{1j}^*) \right).$$

Thus,

$$\mathbb{P}^{m_1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) \leqslant \mathbb{P}^{k_1 m_1}(M(\mathbf{J}_1^*) \leqslant u_{\mathbf{n}}) + \sum_{j=1}^{m_1-1} \Delta_{j+1,x} + o(1) \quad (4.37)$$
$$= (\rho + o(1))^{k_1} + o(1) \longrightarrow 0,$$

because $\rho < 1$. Hence $\delta_{\mathbf{n}} = \mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) + o(1)$. However, it follows similarly that

$$\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) \leqslant \mathbb{P}^{m_1}\{M(\mathbf{J}_1') \leqslant u_{\mathbf{n}}\} + \sum_{j=1}^{m_1-1} \Delta_{j+1,x} + o(1)$$

which tends to zero since (4.36) and hence (4.37) also apply with $M(\mathbf{J}_1')$ in place of $M(\mathbf{J}_1)$. In consequence, $\delta_{\mathbf{n}} \longrightarrow 0$ as $\mathbf{n} \to \infty$ whenever $\rho < 1$. Therefore, (4.30) again holds.

Now to prove that $\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) = \mathbb{P}^{m_1 m_2}(M(\mathbf{J}) \leqslant u_{\mathbf{n}}) + o(1)$, it will be suffice to show that:

$$\mathbb{P}^{m_1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) - \mathbb{P}^{m_1 m_2}(M(\mathbf{J}) \leqslant u_{\mathbf{n}}) \longrightarrow 0.$$

However, this follows by entirely similar reasoning, splitting the rectangle $\mathbf{J}_1$ into rectangles $H_i = [1 : r_1] \times [(i-1)r_2 + 1 : r_2]$, for $i \in [m_2]$, and bearing in mind that the Lipschitz approximations are $K_{\mathbf{n},r_1,(j-1)r_2}$ and $K_{\mathbf{n},r_1,r_2-2l_2}$, defined in smaller domains of size $u = r_2 r_1 (j-1)$ and $v = r_1(r_2 - 2l_2)$, respectively. □

Proof of Proposition 4.2. The steps are the same as in the proof of Proposition 3.2 in [Leadbetter & Rootzén, 1998], but replacing Lemma 3.1 in this proof by our Lemma 4.1. □

Proof of Theorem 4.1. Let $f : [1,\infty) \longrightarrow [1,\infty)$ the continuous strictly increasing function defined by

$$f(T) = T\vartheta_1(T) \cdots \vartheta_{d-1}(T),$$

and define now

$$\phi_k(T) = f^{-1}\left(\frac{1}{k} f(kT)\right), \quad k = 1, 2, 3, \ldots.$$

If we set $T^* = \phi_k(\frac{T}{k})$, note that $kf(T^*) = f(T)$. Then, for $\mathbf{n} = (n, \lceil\vartheta_1(n)\rceil, \ldots, \lceil\vartheta_{d-1}(n)\rceil)$, we set $\mathbf{I} = [\lceil a_1 n\rceil] \times [\lceil a_2\vartheta_1(n)\rceil] \times \cdots \times [\lceil a_d \vartheta_{d-1}(n)\rceil]$, where $a_1 = \dfrac{n^*}{n}$ and $a_{i+1} = \dfrac{\vartheta_i(n^*)}{\vartheta_i(n)}$ for $i = 1, \ldots, d-1$. Clearly $a_i \in (0,1]$ for all $i \in [d]$, because ϑ_i is strictly increasing for all $i \in [d]$ and $n^* = f^{-1}\left(\frac{1}{k} f(n)\right) \leqslant n$ for all $(n,k) \in \mathbb{N}^2$. Besides,

$$a_1 \cdots a_d = \frac{f(n^*)}{f(n)} = \frac{1}{k} =: a$$

Then the assumptions of Proposition 4.1 hold for this $\mathbf{I}$ with these $(a_i)_{i\in[d]}$ and a. Thus,

$$\mathbb{P}\left(M_{\mathbf{n}} \leqslant u_{\mathbf{n}}\right) - \mathbb{P}^k\left(M(\mathbf{I}) \leqslant u_{\mathbf{n}}\right) \underset{\mathbf{n}\to\infty}{\longrightarrow} 0. \tag{4.38}$$

By writing $M_{\mathbf{n}} = M_n$, $u_{\mathbf{n}} = u_n$ and $M(\mathbf{I}) = M_{n^*}$, expression (4.38) is equivalent to

$$\mathbb{P}\left(M_n \leqslant u_n\right) - \mathbb{P}^k\left(M_{n^*} \leqslant u_n\right) \underset{n\to\infty}{\longrightarrow} 0, \tag{4.39}$$

which verifies (4.3) by setting $n^* = \phi_k(n/k)$ and replacing n by kT. Finally the result now follows from Proposition 4.1. □

Proof of Proposition 4.3. It follows from Proposition 4.2 - 1 (with $a_i = 1$ for all $i \in [d]$) and the assumption (TD) that

$$\begin{aligned}
\mathbb{P}\left(M_{\mathbf{n}} \leqslant a_{\mathbf{n}}x + \gamma_{\mathbf{n}}\right) &= \mathbb{P}^{m_{\mathbf{n}}}\left(M(\mathbf{J}) \leqslant a_{\mathbf{n}}x + \gamma_{\mathbf{n}}\right) + o(1) \\
&= \left(1 - H(x)\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right)(1 + o(1))\right)^{m_{\mathbf{n}}} + o(1) \\
&= \left(1 - \frac{H(x)}{m_{\mathbf{n}}}(1 + o(1))\right)^{m_{\mathbf{n}}} + o(1) \\
&\underset{\mathbf{n}\to\infty}{\longrightarrow} \exp(-H(x)),
\end{aligned}$$

where the last equality is because $\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right) = m_{\mathbf{n}}^{-1}$. □

Chapter 5

Limit theorems for cluster functionals of stationary weakly dependent random fields.

We provide central limit theorems for the finite dimensional marginal distributions of empirical processes of cluster functionals of stationary weakly dependent random fields. We define an extremogram for space-time processes, we propose an estimator for this extremogram and we show its asymptotic behaviour in distribution. A numerical example is presented, specifically, we provide an estimation of the extremogram for Chilean wind speeds data.

5.1 Introduction

In Chapter 2, we have shown central limit theorems (CLTs) for empirical processes of cluster functionals of $\mathbb{R}^k-$ valued stationary weakly dependent random sequences. In addition, we have used this CLT in order to show asymptotic normality of the extremogram of $\mathbb{R}^k-$ valued stationary weakly dependent random sequences. Following the same spirit, we will extend these results to random fields.

Let $X = \{X_{\mathbf{t}} \; : \; \mathbf{t} \in \mathbb{N}^d\}$ be a $\mathbb{R}^k-$valued stationary random field and let

$$X_N = \left\{ X_{\mathbf{n},\mathbf{t}} \; : \; \mathbf{t} \in \prod_{i=1}^{d} [n_i] \right\}_{\mathbf{n} \in \mathbb{N}^d}$$

be the corresponding normalised random variables, such that

$$\mathbb{P}\left(X_{\mathbf{n},\mathbf{1}} \in \cdot \mid X_{\mathbf{n},\mathbf{1}} \in A\right) \underset{\mathbf{n}\to\infty}{\longrightarrow} G(\cdot), \tag{5.1}$$

where G is a non-degenerate distribution and $A \subseteq \mathbb{R}^k \setminus \{0\}$ is the failure set. Here $\mathbf{n} := (n_1, \ldots, n_d)$ and $\mathbf{1} := (1, \ldots, 1) \in \mathbb{N}^d$. Besides, $\mathbf{n} \to \infty$ means that $n_i \to \infty$ for all $i \in [d]$. As we established in Chapter 4, one denotes $[n] := [1 : n]$ and $[k : n] := \{k, k+1, \ldots, n\} \subset \mathbb{Z}$.

On the other hand, for each $i \in [d]$, let $r_i := r_{n_i}$ be a integer value such that $r_i = o(n_i)$ and $m_i := \lceil n_i / r_i \rceil := \max\{k \in \mathbb{N} : k \leqslant n_i / r_i\}$. Similarly to Chapter 2, we define the $d-$**blocks** (or simply *blocks*) of X_N by

$$Y_{\mathbf{n},j_1\ldots j_d} := \left(X_{\mathbf{n},\mathbf{t}}\right)_{\mathbf{t}\in\prod_{i=1}^d [(j_i-1)r_i+1\,:\,j_i r_i]}, \tag{5.2}$$

where $(j_1, \ldots, j_d) \in D_{\mathbf{n},d} := \prod_{i=1}^d [m_i]$. We have thus $m_1 \cdots m_d$ blocks $Y_{\mathbf{n},j_1\ldots j_d}$. Besides, by stationarity, we denote $Y_{\mathbf{n}} \stackrel{\mathcal{D}}{=} Y_{\mathbf{n},\mathbf{1}}$ (a generic block of X_N).

Definition 5.1 *Let $y = (x_{\mathbf{t}})_{\mathbf{t}\in\prod_{i=1}^d [r_i]}$ be a $d-$block. We define the* **core** *of the block y (w.r.t. the failure set A) as*

$$y^c = \begin{cases} (x_{\mathbf{t}})_{\mathbf{t}\in\prod_{i=1}^d [r_{i,I}\,:\,r_{i,S}]}, & \text{if } x_{\mathbf{t}} \in A \text{ for some } \mathbf{t} \in \prod_{i=1}^d [r_i]; \\ 0, & \text{otherwise;} \end{cases}$$

where, for each $i \in [d]$, $r_{i,I}$ and $r_{i,S}$ are defined as:

$$r_{i,I} = \min\left\{ j_i \in [r_i] \;:\; x_{(j_1,\ldots,j_i,\ldots,j_d)} \in A, \text{with } (j_1, \ldots, j_{i-1}, j_{i+1}, \ldots, j_d) \in \prod_{k\in[d]\setminus\{i\}} [r_k] \right\},$$

$$r_{i,S} = \max\left\{ j_i \in [r_i] \;:\; x_{(j_1,\ldots,j_i,\ldots,j_d)} \in A, \text{with } (j_1, \ldots, j_{i-1}, j_{i+1}, \ldots, j_d) \in \prod_{k\in[d]\setminus\{i\}} [r_k] \right\}.$$

We adjust now the definition of cluster functionals to $d-$blocks. Let $(E, \mathcal{E})$ be a measurable subspace of $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ for some $k \geqslant 1$ such that $0 \in E$. Let $\mathbb{B}_{l_1,\ldots,l_d}(E)$ be the set of $E-$valued blocks (or arrays) of size $l_1 \times l_2 \times \cdots \times l_d$, with $l_1, \ldots, l_d \in \mathbb{N}$. Consider now the set

$$E_\cup := \bigcup_{l_1,\ldots,l_d=1}^{\infty} \mathbb{B}_{l_1,\ldots,l_d}(E),$$

which is equipped with the $\sigma-$field $\mathcal{E}_\cup$ induced by the Borel$-\sigma-$fields on $\mathbb{B}_{l_1,\ldots,l_d}(E)$, for $l_1,\ldots,l_d \in \mathbb{N}$. As it was defined in Chapter 2 ([1]), a **cluster functional** is a measurable map $f : (E_\cup, \mathcal{E}_\cup) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$f(y) = f(y^c), \text{ for all } y \in E_\cup, \quad \text{and} \quad f(0) = 0. \tag{5.3}$$

Let $\mathcal{F}$ be a class of cluster functionals and let $\{Y_{\mathbf{n},j_1j_2\ldots j_d} : (j_1,\ldots,j_d) \in D_{\mathbf{n},d}\}$ be the family of blocks of size $r_1 \times r_2 \times \cdots \times r_d$ defined in (5.2). The **"empirical process $Z_\mathbf{n}$ of cluster functionals"** in $\mathcal{F}$, is the process $(Z_\mathbf{n}(f))_{f\in\mathcal{F}}$ defined by

$$Z_\mathbf{n}(f) := \frac{1}{\sqrt{|\mathbf{n}|v_\mathbf{n}}} \sum_{(j_1,\ldots,j_d)\in D_{\mathbf{n},d}} (f(Y_{\mathbf{n},j_1\ldots j_d}) - \mathbb{E}f(Y_{\mathbf{n},j_1\ldots j_d})), \tag{5.4}$$

where $|\mathbf{n}| = n_1 \cdots n_d$ and $v_\mathbf{n} := \mathbb{P}(X_{\mathbf{n},\mathbf{1}} \in A)$ with $A \subseteq E \setminus \{0\}$ denoting the failure set.

We will prove the convergence of the fidis of the empirical process (5.4) to a gaussian process using the Lindeberg method of Bardet et al. [2007], adapted to stationary weakly-dependent random fields. We use this result in order to show the convergence in distribution of the iso-extremogram estimator of stationary weakly-dependent space-time processes.

This chapter is developed as follows. In Section 5.2 we provide a CLT for the fidis of the empirical process of cluster functionals (5.4). In Section 5.3 we define the iso-extremogram and we use the theorem of the previous section in order to show, under additional suitable conditions, the asymptotic normality of the iso-extremogram estimator of stationary weakly-dependent space-time processes (or random fields). We dedicate Section 5.4 to present a numerical example, which consists of the estimation of the iso-extremogram for Chilean wind speeds data. Proofs are given in the last section (Section 5.5).

5.2 Central limit theorems

In this section we provide a central limit theorem for the fidis of the empirical process defined in (5.4). The proof consists in the same technique that Bardet et al. [2007]

[1]This definition is only a generalisation with respect to the $d-$blocks and their cores.

used in the demonstrations of their dependent and independent Lindeberg lemmas, but generalised here to random fields.

In order to develop this CLT, firstly consider the following basic assumption:

(B) The vector $\mathbf{r} = (r_1, \ldots, r_d) \in \mathbb{N}^d$ is such that $r_i \ll n_i$ and $r_i \to \infty$ for each $i \in [d]$. Besides, denoting $|\mathbf{r}| = r_1 \cdots r_d$,

$$|\mathbf{r}| \ll v_{\mathbf{n}}^{-1} \ll |\mathbf{n}|. \tag{5.5}$$

Secondly, consider the following essential convergence assumptions:

(C.1) $\mathbb{E}\left((f(Y_{\mathbf{n}}) - \mathbb{E}f(Y_{\mathbf{n}}))^2 \mathbb{1}_{\left\{|f(Y_{\mathbf{n}}) - \mathbb{E}f(Y_{\mathbf{n}})| > \epsilon\sqrt{|\mathbf{n}|v_n}\right\}} \right) = o(|\mathbf{r}|v_{\mathbf{n}})$, for all $\epsilon > 0$, and all $f \in \mathcal{F}$.

(C.2) $(|\mathbf{r}|v_{\mathbf{n}})^{-1}\mathrm{Cov}\left(f(Y_{\mathbf{n}}), g(Y_{\mathbf{n}})\right) \longrightarrow c(f,g)$, for all $f, g \in \mathcal{F}$.

Remark 5.1 We have mentioned before that $\mathbf{n} = (n_1, \ldots, n_d) \to \infty$ means $n_i \to \infty$ for each $i \in [d]$. However, as we established in Section 4.5, the limit of a sequence indexed with $\mathbf{n}$ as $\mathbf{n} \to \infty$ can be reformulated in terms of the limit of such a sequence as "$\mathbf{n} \to \infty$ *along a monotone path on the grid* $\mathbb{N}^d$", *i.e.* along $\mathbf{n} = (\lceil \vartheta_1(n) \rceil, \ldots, \lceil \vartheta_d(n) \rceil)$ for some strictly increasing continuous functions $\vartheta_i : [1, \infty) \longrightarrow [1, \infty)$, with $i \in [d]$, such that $\vartheta_i(n) \longrightarrow \infty$ as $n \to \infty$, for $i \in [d]$.

Let $Y_{\mathbf{n},j_1 \ldots j_d}$, with $(j_1, \ldots, j_d) \in D_{\mathbf{n},d}$, be the random blocks defined in (5.2). For each $k-$tuple of cluster functionals $\mathbf{f}_k = (f_1, \ldots, f_k)$ and each $(j_1, \ldots, j_d) \in D_{\mathbf{n},\mathbf{d}}$, we define the random vector:

$$W_{\mathbf{n},j_1 \ldots j_d} := \frac{1}{\sqrt{|\mathbf{n}|v_{\mathbf{n}}}} \left(f_1(Y_{\mathbf{n},j_1 \ldots j_d}) - \mathbb{E}f_1(Y_{\mathbf{n},j_1 \ldots j_d}), \ldots, f_k(Y_{\mathbf{n},j_1 \ldots j_d}) - \mathbb{E}f_k(Y_{\mathbf{n},j_1 \ldots j_d}) \right). \tag{5.6}$$

Without loss of generality and in order to simplify writing, we will consider $d = 2$ in the rest of this section.

Let $(W'_{\mathbf{n},ij})_{(i,j) \in D_{\mathbf{n},2}}$ be a sequence of zero mean independent random variables with values in $\mathbb{R}^k$, independent of the sequence $(W_{\mathbf{n},ij})_{(i,j) \in D_{\mathbf{n},2}}$ defined in (5.6), and such that $W'_{\mathbf{n},ij} \sim \mathcal{N}_k\left(\mathbf{0}, \mathrm{Cov}(W_{\mathbf{n},ij})\right)$, for all $(i,j) \in D_{\mathbf{n},2}$. Denote by $\mathcal{C}_b^3$ the set of bounded functions $\mathbb{R}^k \longrightarrow \mathbb{R}$ with bounded and continuous partial derivatives up to order 3. For $h \in \mathcal{C}_b^3$ and $\mathbf{n} = (n_1, n_2) \in \mathbb{N}^2$, define

$$\Delta_{\mathbf{n}} := \left| \mathbb{E}\left(h\left(\sum_{(i,j)\in D_{\mathbf{n},2}} W_{\mathbf{n},ij} \right) - h\left(\sum_{(i,j)\in D_{\mathbf{n},2}} W'_{\mathbf{n},ij} \right) \right) \right|. \tag{5.7}$$

The following assumption will allow us to give, in a useful and simplified form, lemmas of Lindeberg under independence and dependence.

(C.1′) It exists $\delta \in (0,1]$ such that for all $(i,j) \in D_{\mathbf{n},2}$, $\mathbb{E}\left\|W_{\mathbf{n},ij}\right\|^{2+\delta} < \infty$, for all $\mathbf{n} \in \mathbb{N}^2$ and all $k-$tuple of cluster functionals $(f_1, \ldots, f_k) \in \mathcal{F}^k$. Denote

$$A_{\mathbf{n}} := \sum_{(i,j)\in D_{\mathbf{n},2}} \mathbb{E}\left\|W_{\mathbf{n},ij}\right\|^{2+\delta}.$$

Lemma 5.1 (Lindeberg under independence) *Suppose that the blocks $(Y_{\mathbf{n},ij})_{(i,j)\in D_{\mathbf{n},2}}$ are independents and that the random variables $(W_{\mathbf{n},ij})_{(i,j)\in D_{\mathbf{n},2}}$ defined in (5.6) satisfy Assumption (C.1′). Then, for all $\mathbf{n} \in \mathbb{N}^2$:*

$$\Delta_{\mathbf{n}} \leqslant 6 \cdot \|h^{(2)}\|_\infty^{1-\delta} \cdot \|h^{(3)}\|_\infty^{\delta} \cdot A_{\mathbf{n}}.$$

Remark 5.2 Taking $\epsilon < 6\|h^{(2)}\|_\infty \left(\|h^{(3)}\|_\infty\right)^{-1}$ and using suitably the second inequality of (5.23) in the proof of Lemma 5.1, classical Lindeberg conditions may be used:

$$\Delta_{\mathbf{n}} \leqslant 2\, \|h^{(2)}\|_\infty\, B_{\mathbf{n}}(\epsilon) + \|h^{(3)}\|_\infty\, a_{\mathbf{n}} \left(\frac{4}{3}\epsilon + \sqrt{B_{\mathbf{n}}(\epsilon)} \right), \tag{5.8}$$

where

$$B_{\mathbf{n}}(\epsilon) = \sum_{(i,j)\in D_{\mathbf{n},2}} \mathbb{E}\left(\|W_{\mathbf{n},ij}\|^2 \mathbb{1}_{\{\|W_{\mathbf{n},ij}\| > \epsilon\}} \right), \quad \text{for } \epsilon > 0,\ \mathbf{n} \in \mathbb{N}^2;$$

$$a_{\mathbf{n}} = \sum_{(i.j)\in D_{\mathbf{n},2}} \mathbb{E}\|W_{\mathbf{n},ij}\|^2 < \infty, \quad \text{for } \mathbf{n} \in \mathbb{N}^2.$$

Moreover, these classical Lindeberg conditions derive the conditions from Lemma 5.1. Indeed,

$$\Delta_{\mathbf{n}} \leqslant 2\, \|h^{(2)}\|_\infty\, \epsilon^{-\delta} A_{\mathbf{n}} + \|h^{(3)}\|_\infty\, a_{\mathbf{n}} \left(\frac{4}{3}\epsilon + \epsilon^{-\delta/2}\sqrt{A_{\mathbf{n}}} \right),$$

for $0 < \delta < 1$ and $\epsilon > 0$. The proof of this remark for general independent random vectors is given in [Bardet et al., 2007, p. 165].

Remark 5.3 Observe that the assumptions (C.1) and (C.2) imply that $B_{\mathbf{n}}(\epsilon) \underset{\mathbf{n}\to\infty}{\longrightarrow} 0$ and that $a_{\mathbf{n}} = \sum_{i=1}^{k} c(f_i, f_i) < \infty$, respectively. Therefore, if the blocks $(Y_{\mathbf{n},ij})_{(i,j)\in D_{\mathbf{n},2}}$ are independent and if the assumptions (C.1) and (C.2) hold, then from Lemma 5.1, the fidis of the empirical process $(Z_{\mathbf{n}}(f))_{f\in\mathcal{F}}$ of cluster functionals converge to the fidis of a Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c.

For the dependent case, we need to consider more notations:

Let $L_i^j := \{(i,v) : v \in [j]\} \subset D_{\mathbf{n},2}$, for all $(i,j) \in D_{\mathbf{n},2}$. We set $L_i^0 = L_0^j = \emptyset$ for any $i \in [m_1]$ and any $j \in [m_2]$.

For each $k \in \mathbb{N}$, $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$, $\mathbf{t} \in \mathbb{R}^k$ and $\mathbf{n} \in \mathbb{N}^2$; define

$$T_{\mathbf{n},\mathbf{t}}(\mathbf{f}_k) := \sum_{(j_1,j_2)\in D_{\mathbf{n},2}} \left| \mathrm{Cov}\left(\exp(i\langle \mathbf{t}, \sum_{(u_1,u_2)\in D_{\mathbf{n},2}\setminus(\cup_{l=0}^{j_1-1} L_l^{m_2} \cup L_{j_1}^{j_2})} W_{\mathbf{n},u_1u_2}\rangle), \exp\left(i\langle \mathbf{t}, W_{\mathbf{n},j_1j_2}\rangle\right) \right) \right| \tag{5.9}$$

Lemma 5.2 (Dependent Lindeberg lemma) *Suppose that the r.v.'s $(W_{\mathbf{n},ij})_{(i,j)\in D_{\mathbf{n},2}}$ defined in (5.6) satisfy Assumption (C.1'). Then, for all $\mathbf{n} \in \mathbb{N}^2$, each $k \in \mathbb{N}$ and all $k-$tuple $\mathbf{f}_k = (f_1, \ldots, f_k)$ of cluster functionals:*

$$\Delta_{\mathbf{n}} \leqslant T_{\mathbf{n},\mathbf{t}}(\mathbf{f}_k) + 6\|\mathbf{t}\|^{2+\delta} A_{\mathbf{n}},$$

where $T_{\mathbf{n},\mathbf{t}}(\mathbf{f}_k)$ is defined in (5.9).

The previous lemma together with Remark 5.2 derive the following theorem.

Theorem 5.1 (Lindeberg CLT for extreme cluster functionals of random fields) *Suppose that the basic assumption (B) holds and that the assumptions (C.1) and (C.2) are satisfied. Then, if for each $k \in \mathbb{N}$, $T_{\mathbf{n},\mathbf{t}}(\mathbf{f}_k)$ converges to zero as $\mathbf{n} \to \infty$, for all $\mathbf{t} \in \mathbb{R}^k$ and all $k-$tuple $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$ of cluster functionals, the fidis of the empirical process $(Z_{\mathbf{n}}(f))_{f\in\mathcal{F}}$ of cluster functionals converge to the fidis of a Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c defined in (C.2).*

Remark 5.4 The previous theorem can be formulated for $d = 3$ as follows. Define $S_i = \{(u, v, w) : u \in [i], v \in [m_2], w \in [m_3]\} \subseteq D_{\mathbf{n},3}$, for $i \in [m_1]$, with the convention $S_0 = \emptyset$. Moreover, $L^k_{ij} = \{(i, j, w) : w \in [k]\}$, for $(i, j, k) \in D_{\mathbf{n},3}$, and $L^k_{ij} = \emptyset$ if i, j or k is zero. Then, if (B), (C.1), (C.2) are satisfied (for $d = 3$), and if for each $k \in \mathbb{N}$,

$$T^*_{\mathbf{n},\mathbf{t}}(\mathbf{f}_k) = \sum_{(j_1,j_2,j_3)\in D_{\mathbf{n},3}} \left|\mathrm{Cov}\left(\exp(i\langle \mathbf{t}, V_{\mathbf{n},j_1j_2j_3}\rangle), \exp\left(i\langle \mathbf{t}, W_{\mathbf{n},j_1j_2j_3}\rangle\right)\right)\right| \tag{5.10}$$

converges to zero as $\mathbf{n} \to \infty$, for all $\mathbf{t} \in \mathbb{R}^k$ and all $k-$tuple $\mathbf{f}_k = (f_1, \ldots, f_k) \in \mathcal{F}^k$ of cluster functionals, with

$$V_{\mathbf{n},j_1j_2j_3} := \sum_{(u_1,u_2,u_3)\in D_{\mathbf{n},3}\setminus\left(S_{j_1-1} \cup \bigcup_{l=0}^{j_2-1} L^{m_3}_{j_1l} \cup L^{j_3}_{j_1j_2}\right)} W_{\mathbf{n},u_1u_2u_3},$$

the fidis of the empirical process $(Z_\mathbf{n}(f))_{f\in\mathcal{F}}$ of cluster functionals converge to the fidis of a Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c.

Note that the expression (5.9) or (5.10) can be bounded by either the strong mixing coefficient or by the weak-dependence coefficients of Doukhan & Louhichi [1999] for stationary random fields. Some of these coefficients are mentioned in the previous chapter. However, we will not show such bounds in this document. The work in preparation [Gómez, 2017b] will include these conditions or bounds in order to obtain limit theorems of cluster functionals for specific types of stationary weakly dependent random fields.

5.3 Application: The extremogram for space - time processes

Let $X = \{X_t(\mathbf{s}) : \mathbf{s} \in \mathbb{Z}^d, t \geqslant 0\}$ be a $\mathbb{R}^k-$valued space-time process, which is stationary in both space and time. We define the **extremogram** of X for two sets A and B both bounded away from zero by

$$\rho_{A,B}(\mathbf{s}, h_t) := \lim_{x\to\infty} \mathbb{P}\left(x^{-1}X_{h_t}(\mathbf{s}) \in B \,\middle|\, x^{-1}X_0(\mathbf{0}) \in A\right), \quad (\mathbf{s}, h_t) \in \mathbb{Z}^d \times [0, \infty); \tag{5.11}$$

provided that the limit exists.

In estimating the extremogram, the limit on x in (5.11) is replaced by a high quantile $u_{\mathbf{n}}$ of the process. Defining u_n as the $(1-1/k_n)-$quantile of the stationary distribution of $\|X_t(\mathbf{s})\|$ or related quantity, with $k_n = o(n) \underset{n\to\infty}{\longrightarrow} \infty$, one can redefine (5.11) by

$$\rho_{A,B}(\mathbf{s}, h_t) := \lim_{n\to\infty} \mathbb{P}\left(u_n^{-1} X_{h_t}(\mathbf{s}) \in B \,\middle|\, u_n^{-1} X_0(\mathbf{0}) \in A\right), \quad (\mathbf{s}, h_t) \in \mathbb{Z}^d \times [0,\infty). \quad (5.12)$$

The choice of such a sequence of quantiles $(u_n)_{n\in\mathbb{N}}$ is not arbitrary. The main condition to guarantee the existence of the limit (5.12) for any two sets A and B bounded away from zero, is that it must satisfy the following convergence

$$k_n \mathbb{P}\left(u_n^{-1}\left(X_{t_1}(\mathbf{s}_1), \ldots, X_{t_p}(\mathbf{s}_p)\right) \in \,\cdot\,\right) \underset{n\to\infty}{\overset{vague}{\longrightarrow}} m_{(\mathbf{s}_1,t_1),\ldots,(\mathbf{s}_p,t_p)}(\,\cdot\,), \quad (5.13)$$

for all $(\mathbf{s}_i, t_i) \in \mathbb{Z}^d \times [0,\infty)$, $i \in [p]$, $p \in \mathbb{N}$, where

$$\left(m_{(\mathbf{s}_1,t_1),\ldots,(\mathbf{s}_p,t_p)}\right)_{(\mathbf{s_i},t_i)\in\mathbb{Z}^d\times[0,\infty),\ i\in[p],\ p\in\mathbb{N}}$$

is a collection of Radon measures on the Borel $\sigma-$field $\mathcal{B}(\overline{\mathbb{R}}^{kp} \setminus \{\mathbf{0}\})$, not all of them being the null measure, with $m_{(\mathbf{s}_1,t_1),\ldots,(\mathbf{s}_p,t_p)}(\overline{\mathbb{R}}^{kp} \setminus \mathbb{R}^{kp}) = 0$.

In this case,

$$\begin{aligned}\mathbb{P}\left(u_n^{-1} X_{h_t}(\mathbf{s}) \in B \,\middle|\, u_n^{-1} X_0(\mathbf{0}) \in A\right) &= \frac{k_n \mathbb{P}\left(u_n^{-1}\left(X_0(\mathbf{0}), X_{h_t}(\mathbf{s})\right) \in A \times B\right)}{k_n \mathbb{P}\left(u_n^{-1} X_0(\mathbf{0}) \in A\right)} \\ &\longrightarrow \frac{m_{(\mathbf{0},0),(\mathbf{s},h_t)}(A \times B)}{m_{(\mathbf{0},0)}(A)} = \rho_{A,B}(\mathbf{s}, h_t),\end{aligned}$$

provided that $m_{(\mathbf{0},0)}(A) > 0$.

Remark 5.5 The condition (5.13) is particularly satisfied if the random field $\widetilde{X}(\mathbf{s}, t) = X_t(\mathbf{s})$, for $(\mathbf{s}, t) \in \mathbb{Z}^d \times [0,\infty)$, is regularly varying. For details and examples of regularly varying random fields, see [Davis & Mikosch, 2008].

Note that the extremogram (5.12) is a function of two lags: a spatial-lag $\mathbf{s} \in \mathbb{Z}^d$ and a non-negative time-lag h_t. Due to all the spatial values that the spatial-lag $\mathbf{s}$ takes, in practice, it is very complicated to analyse the results of the estimation of such extremogram. Moreover, the calculation would be computationally very slow. In order to obtain a simpler interpretation and simplify the calculations, we will assume that the space-time process X satisfies the following "isotropy" condition:

(I) For each pair of non-negative integers h_t and h_s,

$$\mathbb{P}\left(X_0(\mathbf{0}) \in A, X_{h_t}(\mathbf{s}) \in B\right) = \mathbb{P}\left(X_0(\mathbf{0}) \in A, X_{h_t}(\mathbf{s}') \in B\right), \quad \text{for any } \mathbf{s}, \mathbf{s}' \in \mathbb{S}_{h_s}^{d-1},$$

where we denote $\mathbb{S}_h^{d-1} := \left\{\mathbf{s} \in \mathbb{Z}^d : \|\mathbf{s}\|_\infty = h\right\}$ with $h \geqslant 0$ and $\|(s_1, \ldots, s_d)\|_\infty = \max_{i=1,\ldots,d} |s_i|$.

Under this condition, the extremogram (5.12) can be redefined using only two non-negative integer lags: a spatial-lag h_s and a time-lag h_t. Indeed, under the assumption (I), we define the **iso-extremogram** of X for two sets A and B both bounded away from zero by

$$\rho^*_{A,B}(h_s, h_t) = \rho_{A,B}(h_s \cdot \vec{e}_1,\ h_t), \qquad h_s, h_t = 0, 1, 2, \ldots \tag{5.14}$$

where $\vec{e}_1 = (1, 0, 0, 0, \ldots) \in \mathbb{R}^d$ (the first element of the standard basis of $\mathbb{R}^d$).

We will propose now a estimator for the iso-extremogram. For this, consider w.l.o.g. $d = 2$, because the case $d > 2$ is similar.

Let $X_{\mathbf{n}} := \{X_t(i,j) : (i,j,t) \in [n_1] \times [n_2] \times [n_3]\}$ be the observations from a $\mathbb{R}^k-$valued space-time process X, stationary in both space and time, and which satisfies the condition (I). Let us set $n = n_1 n_2 n_3$. The sample iso-extremogram based on the observations $X_{\mathbf{n}}$ is given by

$$\widehat{\rho^*}_{A,B}(h_s, h_t) := \frac{\sum_{(j_1,j_2)\in[m_1]\times[m_2]} \sum_{t=1}^{n_3-h_t} \sum_{(i_1,i_2)\in \mathbb{S}_{h_s}(c_{j_1 j_2})} \frac{1}{\#\mathbb{S}_{h_s}(c_{j_1 j_2})} \cdot \mathbb{1}_{\left\{\frac{X_{t+h_t}(i_1,i_2)}{u_n} \in B,\ \frac{X_t(c_{j_1 j_2})}{u_n} \in A\right\}}}{\sum_{(j_1,j_2)\in[m_1]\times[m_2]} \sum_{t=1}^{n_3} \mathbb{1}_{\left\{\frac{X_t(c_{j_1 j_2})}{u_n} \in A\right\}}}, \tag{5.15}$$

for $h_s = 0, 1, 2, \ldots, \lceil 2^{-1} \min\{r_1, r_2\}\rceil - 1$, and $h_t = 0, \ldots, n-1$, where

$$c_{ij} := \left(\left\lceil \frac{(2i-1)r_1 + 1}{2} \right\rceil, \left\lceil \frac{(2j-1)r_2 + 1}{2} \right\rceil\right)$$

denotes the "centre" of the block $B_{ij} = [(i-1)r_1 + 1 : ir_1] \times [(j-1)r_2 + 1 : jr_2]$, for $(i,j) \in [m_1] \times [m_2]$. Besides, $\mathbb{S}_h(u,v) := \{(i,j) \in [n_1] \times [n_2] : \|(u,v) - (i,j)\|_\infty = h\}$ with $h \geqslant 0$ and $\#E$ denotes the cardinality of the set E. Recalling that $r_i = r_{n_i}$ and $m_i = \lceil n_i / r_i \rceil$, for $i = 1, 2, 3$.

Defining the cluster functional $f_{A,B,h_1,h_2} : \left(\bigcup_{l_1,l_2,l_3=1}^{\infty} \mathbb{B}_{l_1 l_2 l_3}(\mathbb{R}^k), \mathcal{R}_\cup\right) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, for $h_1, h_2 = 0, 1, 2, \ldots$, such that

$$f_{A,B,h_1,h_2}\left((x_{(i_1,i_2,i_3)})_{(i_1,i_2,i_3)\in[l_1]\times[l_2]\times[l_3]}\right) = \sum_{(i_1,i_2)\in\mathbb{S}_{h_1}(c)} \sum_{i_3=1}^{l_3-h_2} \frac{\mathbb{1}_{A\times B}(x_{(c,i_3)}, x_{(i_1,i_2,i_3+h_2)})}{\#\mathbb{S}_{h_1}(c)}, \tag{5.16}$$

with $c = (\lceil (l_1+1)/2 \rceil, \lceil (l_2+1)/2 \rceil) \in [l_1] \times [l_2]$ (the "centre" of the block $B = [l_1] \times [l_2]$), we can rewrite the estimator (5.15) as:

$$\widehat{\rho^*}_{A,B}(h_s, h_t) = \frac{\sum_{(j_1,j_2,j_3)\in D_{\mathbf{n},3}} f_{A,B,h_s,h_t}(Y_{\mathbf{n},j_1 j_2 j_3}) + \delta_{\mathbf{n}} + R_{A,B,h_s,h_t}}{\sum_{(j_1,j_2,j_3)\in D_{\mathbf{n},3}} f_{A,A,0,0}(Y_{\mathbf{n},j_1 j_2 j_3}) + R_{A,A,0,0}}, \tag{5.17}$$

where

$$\delta_{\mathbf{n}} := \sum_{(j_1,j_2,j_3)\in D_{\mathbf{n},3}} \sum_{(i_1,i_2)\in\mathbb{S}_{h_s}(c_{j_1 j_2})} \sum_{t=j_3 r_3 - h_t + 1}^{j_3 r_3} \frac{1}{\#\mathbb{S}_{h_s}(c_{j_1 j_2})} \mathbb{1}_{\left\{\frac{X_{t+h_t}(i_1,i_2)}{u_n}\in B,\ \frac{X_t(c_{j_1 j_2})}{u_n}\in A\right\}},$$

$$R_{A,B,h_s,h_t} := \sum_{(j_1,j_2)\in[m_1]\times[m_2]} \sum_{(i_1,i_2)\in\mathbb{S}_{h_2}(c_{j_1 j_2})} \sum_{t=m_3 r_3+1}^{n_3-h_t} \frac{1}{\#\mathbb{S}_{h_s}(c_{j_1 j_2})} \mathbb{1}_{\left\{\frac{X_{t+h_t}(i_1,i_2)}{u_n}\in B,\ \frac{X_t(c_{j_1 j_2})}{u_n}\in A\right\}}.$$

We can therefore write (5.17) in terms of empirical processes of cluster functionals (5.4) and use Lindeberg CLT for extreme cluster functionals of random fields (Theorem 5.1) together with suitable conditions of joint distributions, in order to prove the convergence in distribution of the iso-extremogram estimator.

For this, firstly we set some considerations: the normalised random variables are defined here by $X_{\mathbf{n},(i_1,i_2,t)} = u_n^{-1} X_t(i_1, i_2)$, where $\mathbf{n} = (n_1, n_2, n_3)$ and $n = n_1 n_2 n_3$; and the random blocks $(Y_{\mathbf{n},j_1 j_2 j_3})_{(j_1,j_2,j_3)\in D_{\mathbf{n},3}}$ as in (5.2). We define $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and

$$\mathcal{F}_{A,B} := \left\{ f_{A,B,h_s,h_t} : h_s, h_t \in \mathbb{N}_0 \right\}$$

as the family of cluster functionals defined in (5.16). Moreover, for the set A, bounded away from zero, let $v_n := \mathbb{P}\left(u_n^{-1} X_0(0,0) \in A\right)$.

Secondly, consider the following conditions:

(C.2') For each $h_s, h'_s, h_t, h'_t \in \mathbb{N}_0$,

$$\sum_{\mathbf{i}\in\mathbb{S}_{h_s}(c)} \sum_{\mathbf{i}'\in\mathbb{S}_{h'_s}(c)} \sum_{t=1}^{r_3-h_t} \sum_{t'=1}^{r_3-h'_t} \frac{\mathbb{P}\left(u_n^{-1}(X_t(c), X_{t'}(c)) \in A^2, (X_{t+h_t}(\mathbf{i}), X_{t'+h'_t}(\mathbf{i}')) \in B^2\right)}{rv_n \cdot \#\mathbb{S}_{h_s}(c) \cdot \#\mathbb{S}_{h'_s}(c)} \underset{\mathbf{n}\to\infty}{\longrightarrow} \sigma_{A,B}((h_s,h_t),(h'_s,h'_t)), \quad (5.18)$$

and for each $h_s, h_t \in \mathbb{N}_0$,

$$\sum_{\mathbf{i}\in\mathbb{S}_{h_s}(c)} \sum_{t=1}^{r_3-h_t} \sum_{t'=1}^{r_3} \frac{\mathbb{P}\left(u_n^{-1}(X_t(c), X_{t'}(c)) \in A^2, X_{t+h_t}(\mathbf{i}) \in B\right)}{rv_n \cdot \#\mathbb{S}_{h_s}(c)} \underset{\mathbf{n}\to\infty}{\longrightarrow} \sigma'_{A,B}(h_s,h_t), \quad (5.19)$$

where $r = r_1r_2r_3$ and $c = (\lceil (r_1+1)/2 \rceil, \lceil (r_2+1)/2 \rceil)$, is the "centre" of the block $B_{11} = [r_1] \times [r_2]$.

(C.3)

$$\sum_{(c,t),(c',t')\in C(r_1,r_2)\times[n_3]} \mathbb{P}\left(u_n^{-1}(X_t(c), X_{t'}(c')) \in \ A \times A\right) = \mathcal{O}(1),$$

where $C(r_1, r_2) := \{c_{ij} \in [n_1] \times [n_2] : (i,j) \in [m_1] \times [m_2]\}$ is set of the "centres" of the blocks $B_{ij} = [(i-1)r_1 + 1 : ir_1] \times [(j-1)r_2 + 1 : jr_2]$.

Proposition 5.1 *Assume that the following conditions hold for the $\mathbb{R}^k-$valued space-time process $X = \{X_t(i_1, i_2) : (i_1, i_2, t) \in \mathbb{Z}^2 \times [0, \infty)\}$:*

1. *The process X is stationary in both space and time and satisfies the condition (I).*

2. *The sequence (u_n) is such that (5.13) holds. Moreover satisfies that $r \ll v_n^{-1} \ll n$ and $\sqrt{nv_n} \ll r \ll nv_nr_3$, where $n = n_1n_2n_3$, $r = r_1r_2r_3$, $r_i \ll n_i$ and $r_i \underset{n_i\to\infty}{\longrightarrow} \infty$ for $i = 1, 2, 3$.*

3. *The normalised process $X_N = (X_{\mathbf{n},\mathbf{i}})$ together with the family of cluster functionals $\mathcal{F}_{A,B}$ satisfies either the assumption (C.1') or else (C.1). Moreover, for each $p \in \mathbb{N}$, the coefficient $T^*_{\mathbf{n},\mathbf{t}}(\mathbf{f}_p)$ defined in (5.10) converges to zero as $\mathbf{n} \to \infty$, for all $p-$tuple of cluster functionals $(f_1, \ldots, f_p) \in \mathcal{F}^p_{A,B}$ and all $\mathbf{t} \in \mathbb{R}^p$. The same assumption holds together with the family $\mathcal{F}_A := \{f_{A,A,0,0}\}$, which contains a single functional.*

4. *The conditions (C.2') and (C.3) hold.*

Then, for each $(L_s, L_t) \in \mathbb{N}_0 \times \mathbb{N}_0$,

$$\frac{\sqrt{n v_{\mathbf{n}}}}{r_1 r_2}\left(\widehat{\rho^*}_{A,B}(h_s,h_t) - \rho^*_{A,B,n}(h_s,h_t)\right)_{0\leqslant h_s\leqslant L_s,\ 0\leqslant h_t\leqslant L_t} \xrightarrow[\mathbf{n}\to\infty]{\mathcal{D}} \mathcal{N}(0,\Sigma_{A,B,L_s,L_t}), \tag{5.20}$$

where $\rho^*_{A,B,n}(h_s,h_t) := \mathbb{P}\left(u_n^{-1}X_{h_t}(h_s\cdot e_1)\in B\,\middle|\,u_n^{-1}X_0(\mathbf{0})\in A\right)$ *and* Σ_{A,B,L_s,L_t} *is the covariance matrix, defined by the coefficients*

$$\sigma_{\mathbf{h},\mathbf{h}'} = \sigma_{A,B}(\mathbf{h},\mathbf{h}') - \rho^*_{A,B}(\mathbf{h}')\sigma'_{A,B}(\mathbf{h}) - \rho^*_{A,B}(\mathbf{h})\sigma'_{A,B}(\mathbf{h}') + \rho^*_{A,B}(\mathbf{h})\rho^*_{A,B}(\mathbf{h}')\sigma'_{A,A}(\mathbf{0}), \tag{5.21}$$

with $\mathbf{h},\mathbf{h}'\in[0:L_t]\times[0:L_s]$.

5.4 Example: estimation of the iso-extremogram for Chilean wind speeds data

In Chile, on the hills of Valparaiso and nearby cities, the electrical and telephone cabling system is, for the most part, placed above the ground. On the other hand, electrical cuts and accidents are frequently registered in the city due to the presence of strong wind currents from the Pacific. The company that supplies electricity to the city, the insurance companies and the government are interested in predicting this type of events in order to avoid material and sometimes human losses.

This thesis is not intended to study this particular problem. However, we will use some of the data to give a numerical example of the iso-extremogram estimator. Specifically, we estimate the iso-extremogram (5.14) of the space-time process of the wind speeds of the Chilean region delimited by the longitudes: -71.92120 and -70.28467, and by the latitudes: -32.71611 and -32.71164 (see Figure 5.1).

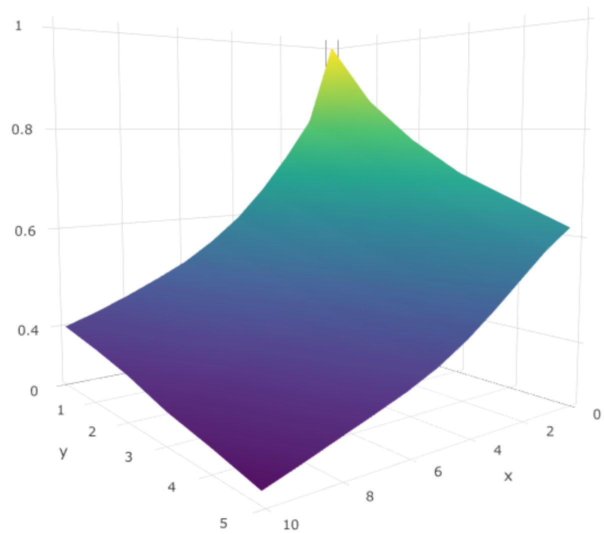
For this, we consider the Chilean wind speeds data

$$\{X_t(i,j)\in\mathbb{R} : (i,j,t)\in[154]\times[64]\times[744]\},$$

where (i,j) is the corresponding point in $\mathbb{Z}^2$ of the rescaled $(longitude, latitude)$ [*i.e.* the pair of integers (i,j) is the image of the $(longitude, latitude)$ under a bijective transformation T], and t denotes the time, which is measured in hours. Here, the grid $[154]\times[64]$ is separated by $1km$ and $n_3 = 744$ corresponds to the 744 hours of the month of June, 2009. Generally, June is the month of the year that registers strongest wind speeds on this region.

Figure 5.1: Chilean region delimited by the longitudes: −71.92120 and −70.28467, and the latitudes: −32.71611 and −32.71164.

The estimated iso-extremogram of the space-time process of the wind speeds of the Chilean region, for the sets bounded away from zero $A = B = (1, \infty)$, is shown on the Figure 5.2. Here, we have set $r_1 = r_2 = 12$ and $L_s = 5$ because the grid is not very large. Moreover, $L_t = 10$ and u_n is the estimated 0.95−quantile from the data. From the figures, one can note both a temporal and spatial decay of the iso-extremogram.



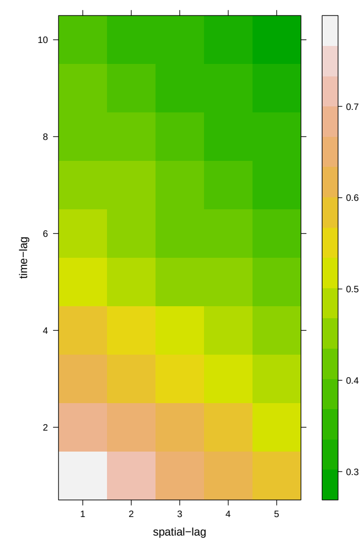


Figure 5.2: The estimated iso-extremogram from the Chilean wind speeds data. Left: Surface of the estimation with $(h_s, h_t) \in [0:5] \times [0:10]$. Right: The level-plot of the estimation for $(h_s, h_t) \in [5] \times [10]$.

5.5 Proofs

Proof of Lemma 5.1. First, notice that

$$\Delta_{\mathbf{n}} \leqslant \sum_{(i,j)\in D_{\mathbf{n},2}} \Delta_{\mathbf{n},ij}, \tag{5.22}$$

where

$$\begin{aligned}
\Delta_{\mathbf{n},ij} &:= \left| \mathbb{E}\left(h_{ij}(V_{\mathbf{n},ij} + W_{\mathbf{n},ij}) - h_{ij}(V_{\mathbf{n},ij} + W'_{\mathbf{n},ij}) \right) \right|, \quad \forall (i,j) \in D_{\mathbf{n},2}; \\
V_{\mathbf{n},ij} &:= \sum_{(u,v)\in D_{\mathbf{n},2}\setminus\left(\bigcup_{l=0}^{i-1} L_l^{m_2} \cup L_i^j\right)} W_{\mathbf{n},uv}, \quad \forall (i,j) \in D_{\mathbf{n},2} \setminus \{(m_1, m_2)\}, \\
V_{\mathbf{n},m_1m_2} &= 0; \quad \text{and} \\
h_{ij}(x) &:= \mathbb{E}\left(h\left(x + \sum_{u=0}^{i-1}\sum_{v=1}^{m_2} W'_{\mathbf{n},uv} + \sum_{v=0}^{j-1} W'_{\mathbf{n},iv} \right) \right).
\end{aligned}$$

Besides, we set the convention $W_{\mathbf{n},ij} = 0$, if $i = 0$ or $j = 0$.

Now, we will use some lines of the proof of Lemma 1 in Bardet et al. [2007].

Let $v, w \in \mathbb{R}^k$. From Taylor's formula, there exist vectors $v_{1,w}, v_{2,w} \in \mathbb{R}^k$ such that:

$$\begin{aligned} h(v+w) &= h(v) + h^{(1)}(v)(w) + \frac{1}{2}h^{(2)}(v_{1,w})(w,w) \\ &= h(v) + h^{(1)}(v)(w) + \frac{1}{2}h^{(2)}(v)(w,w) + \frac{1}{6}h^{(3)}(v_{2,w})(w,w,w), \end{aligned}$$

where, for $j = 1,2,3$, $h^{(j)}(v)(w_1, w_2, \ldots, w_j)$ stands for the value of the symmetric $j-$linear form from $h^{(j)}$ of $(w_1, \ldots, w_j)$ at v. Moreover, denote

$$\|h^{(j)}(v)\|_1 = \sup_{\|w_1\|,\ldots,\|w_j\| \leqslant 1} |h^{(j)}(v)(w_1, \ldots, w_j)| \quad \& \quad \|h^{(j)}\|_\infty = \sup_{v \in \mathbb{R}^k} \|h^{(j)}(v)\|_1.$$

Thus, for $v, w, w' \in \mathbb{R}^k$, there exists some suitable vectors $v_{1,w}, v_{2,w}, v_{1,w'}, v_{2,w'} \in \mathbb{R}^k$ such that

$$\begin{aligned} h(v+w) - h(v+w') &= h^{(1)}(v)(w-w') + \frac{1}{2}\left(h^{(2)}(v)(w,w) - h^{(2)}(v)(w',w')\right) \\ &\quad + \frac{1}{2}\left(\left(h^{(2)}(v_{1,w}) - h^{(2)}(v)\right)(w,w) - \left(h^{(2)}(v_{1,w'}) - h^{(2)}(v)\right)(w',w')\right), \end{aligned}$$

using the approximation of Taylor of order 2, and

$$\begin{aligned} h(v+w) - h(v+w') &= h^{(1)}(v)(w-w') + \frac{1}{2}\left(h^{(2)}(v)(w,w) - h^{(2)}(v)(w',w')\right) \\ &\quad + \frac{1}{6}\left(h^{(3)}(v_{2,w})(w,w,w) - h^{(3)}(v_{2,w'})(w',w',w')\right), \end{aligned}$$

using the approximation of Taylor of order 3.

Thus, $\gamma = h(v+w) - h(v+w') - h^{(1)}(v)(w-w') - \frac{1}{2}\left(h^{(2)}(v)(w,w) - h^{(2)}(v)(w',w')\right)$ satisfies:

$$\begin{aligned} |\gamma| &\leqslant \left(\left(\|w\|^2 + \|w'\|^2\right)\|h^{(2)}\|_\infty\right) \wedge \left(\frac{1}{6}\left(\|w\|^3 + \|w'\|^3\right)\|h^{(3)}\|_\infty\right) \\ &\leqslant \left(\|w\|^2\|h^{(2)}\|_\infty\right) \wedge \left(\frac{1}{6}\|w\|^3\|h^{(3)}\|_\infty\right) + \left(\|w\|^2\|h^{(2)}\|_\infty\right) \wedge \left(\frac{1}{6}\|w'\|^3\|h^{(3)}\|_\infty\right) \\ &\quad + \left(\|w'\|^2\|h^{(2)}\|_\infty\right) \wedge \left(\frac{1}{6}\|w\|^3\|h^{(3)}\|_\infty\right) + \left(\|w'\|^2\|h^{(2)}\|_\infty\right) \wedge \left(\frac{1}{6}\|w'\|^3\|h^{(3)}\|_\infty\right) \\ &\leqslant \frac{1}{6^\delta}\|h^{(2)}\|_\infty^{1-\delta}\|h^{(3)}\|_\infty^\delta \left(\|w\|^{2+\delta} + \|w\|^{2(1-\delta)}\|w'\|^{3\delta} + \|w\|^{3\delta}\|w'\|^{2(1-\delta)} + \|w'\|^{2+\delta}\right), \end{aligned} \tag{5.23}$$

where the last inequality (5.23) is given by using the inequality $1 \wedge a \leqslant a^\delta$, with $a \geqslant 0$ and $\delta \in [0,1]$.

Substituting $h_{ij}, V_{\mathbf{n},ij}, W_{\mathbf{n},ij}$ and $W'_{\mathbf{n},ij}$ to h, v, w and w' in the preceding inequality (5.23) and taking expectations, a bound is given to $\Delta_{\mathbf{n},ij}$. Indeed,

$$\begin{aligned}\mathbb{E}\left(h_{ij}(V_{\mathbf{n},ij}+W_{\mathbf{n},ij})-h_{ij}(V_{\mathbf{n},ij}+W'_{\mathbf{n},ij})\right) &= \mathbb{E}\left(h_{ij}(V_{\mathbf{n},ij}+W_{\mathbf{n},ij})-h_{ij}(V_{\mathbf{n},ij}+W'_{\mathbf{n},ij})\right)+0\\ &= \mathbb{E}\left(h_{ij}(V_{\mathbf{n},ij}+W_{\mathbf{n},ij})-h_{ij}(V_{\mathbf{n},ij}+W'_{\mathbf{n},ij})\right)-\mathbb{E}\left(h^{(1)}_{ij}(V_{\mathbf{n},ij})(W_{\mathbf{n},ij}-W'_{\mathbf{n},ij})\right)\\ &\quad -\frac{1}{2}\mathbb{E}\left(h^{(2)}_{ij}(V_{\mathbf{n},ij})(W_{\mathbf{n},ij},W_{\mathbf{n},ij})-h^{(2)}_{ij}(V_{\mathbf{n},ij})(W'_{\mathbf{n},ij},W'_{\mathbf{n},ij})\right),\end{aligned}$$

because $V_{\mathbf{n},ij}$ is independent from $W_{\mathbf{n},ij}$ and $W'_{\mathbf{n},ij}$ and because $\mathbb{E}W_{\mathbf{n},ij} = \mathbb{E}W'_{\mathbf{n},ij} = 0$, and $\operatorname{Cov}(W_{\mathbf{n},ij}) = \operatorname{Cov}(W'_{\mathbf{n},ij})$ for all $(i,j) \in D_{\mathbf{n},2}$.

On the other hand, using Jensen's inequality and the independence of the random variables $(W_{\mathbf{n},ij})_{(i,j)\in D_{\mathbf{n},2}}$ and $(W'_{\mathbf{n},ij})_{(i,j)\in D_{\mathbf{n},2}}$, we derive

$$\mathbb{E}\|W'_{\mathbf{n},ij}\|^{2+\delta} \leqslant \left(\mathbb{E}\|W'_{\mathbf{n},ij}\|^4\right)^{\frac{1}{2}+\frac{\delta}{4}} \quad \text{and} \quad \mathbb{E}\|W'_{\mathbf{n},ij}\|^4 \leqslant 3\cdot\left(\mathbb{E}\|W_{\mathbf{n},ij}\|^2\right)^2,$$

because $W'_{\mathbf{n},ij}$ is a Gaussian random variable with the same covariance that $W_{\mathbf{n},ij}$. Therefore,

$$\mathbb{E}\|W'_{\mathbf{n},ij}\|^{2+\delta} \leqslant \left(3\cdot\left(\mathbb{E}\|W_{\mathbf{n},ij}\|^2\right)^2\right)^{\frac{1}{2}+\frac{\delta}{4}} = 3^{\frac{1}{2}+\frac{\delta}{4}}\left(\mathbb{E}\|W_{\mathbf{n},ij}\|^2\right)^{1+\frac{\delta}{2}} \leqslant 3^{\frac{1}{2}+\frac{\delta}{4}}\mathbb{E}\|W_{\mathbf{n},ij}\|^{2+\delta} \tag{5.24}$$

$$\begin{aligned}\mathbb{E}\|W'_{\mathbf{n},ij}\|^{2(1-\delta)}\mathbb{E}\|W_{\mathbf{n},ij}\|^{3\delta} &\leqslant \left(\mathbb{E}\|W'_{\mathbf{n},ij}\|^2\right)^{1-\delta}\mathbb{E}\|W_{\mathbf{n},ij}\|^{3\delta}\\ &\leqslant \left(\mathbb{E}\|W_{\mathbf{n},ij}\|^2\right)^{1-\delta}\mathbb{E}\|W_{\mathbf{n},ij}\|^{3\delta} \leqslant \mathbb{E}\|W_{\mathbf{n},ij}\|^{2+\delta}\end{aligned} \tag{5.25}$$

Besides, for $3\delta < 2$,

$$\mathbb{E}\|W_{\mathbf{n},ij}\|^{2(1-\delta)}\mathbb{E}\|W'_{\mathbf{n},ij}\|^{3\delta} \leqslant \mathbb{E}\|W_{\mathbf{n},ij}\|^{2(1-\delta)}\left(\mathbb{E}\|W'_{\mathbf{n},ij}\|^2\right)^{\frac{3\delta}{2}} \leqslant \mathbb{E}\|W_{\mathbf{n},ij}\|^{2+\delta}, \tag{5.26}$$

else

$$\begin{aligned}\mathbb{E}\|W_{\mathbf{n},ij}\|^{2(1-\delta)}\mathbb{E}\|W'_{\mathbf{n},ij}\|^{3\delta} &\leqslant \mathbb{E}\|W_{\mathbf{n},ij}\|^{2(1-\delta)}\left(\mathbb{E}\|W'_{\mathbf{n},ij}\|^4\right)^{\frac{3\delta}{4}}, \quad \text{because } 3\delta \leqslant 4\\ &\leqslant 3^{\frac{3\delta}{4}}\mathbb{E}\|W_{\mathbf{n},ij}\|^{2(1-\delta)}\left(\mathbb{E}\|W_{\mathbf{n},ij}\|^2\right)^{\frac{3\delta}{2}} \leqslant 3^{\frac{1}{2}+\frac{\delta}{4}}\mathbb{E}\|W_{\mathbf{n},ij}\|^{2+\delta}.\end{aligned} \tag{5.27}$$

The inequalities (5.24)-(5.27) allow to consider the terms between parentheses in the last inequality in (5.23). Recall that $\|h^{(k)}_{ij}\|_\infty \leqslant \|h^{(k)}\|_\infty$ for all $(i,j) \in D_{\mathbf{n},2}$ and $0 \leqslant k \leqslant 3$. Therefore, we obtain that

$$\Delta_{\mathbf{n},ij} \leqslant \frac{2(1+3^{\frac{1}{2}+\frac{\delta}{4}})}{6^\delta}\|h^{(2)}\|_\infty^{1-\delta}\|h^{(3)}\|_\infty^{\delta}\mathbb{E}\|W_{\mathbf{n},ij}\|^{2+\delta} \leqslant 6\|h^{(2)}\|_\infty^{1-\delta}\|h^{(3)}\|_\infty^{\delta}\mathbb{E}\|W_{\mathbf{n},ij}\|^{2+\delta},$$

because, for all $\delta \in [0,1]$, $C(\delta) = \frac{2(1+3^{\frac{1}{2}+\frac{\delta}{4}})}{6^{\delta}} \leqslant C(0) = 2(1+\sqrt{3}) < 6$.

As a consequence, from Assumption (C.1'),

$$\Delta_{\mathbf{n}} \leqslant 6\ \|h^{(2)}\|_{\infty}^{1-\delta}\ \|h^{(3)}\|_{\infty}^{\delta}\ A_{\mathbf{n}}.$$

□

Proof of Lemma 5.2. Consider $(W^*_{\mathbf{n},j_1j_2})_{(j_1,j_2)\in D_{\mathbf{n},2}}$ an array of random variables satisfying Assumption (C.1') and such that $(W^*_{\mathbf{n},j_1j_2})_{(j_1,j_2)\in D_{\mathbf{n},2}}$ is independent of $(W_{\mathbf{n},j_1j_2})_{(j_1,j_2)\in D_{\mathbf{n},2}}$ and $(W'_{\mathbf{n},j_1j_2})_{(j_1,j_2)\in D_{\mathbf{n},2}}$. Moreover, assume that $W^*_{\mathbf{n},j_1j_2}$ has the same distribution as $W_{\mathbf{n},j_1j_2}$ for $(j_1,j_2) \in D_{\mathbf{n},2}$.

Then, using the same decomposition (5.22) in the proof of the previous lemma, one can also write,

$$\begin{aligned}\Delta_{\mathbf{n},j_1j_2} \leqslant \Big|\mathbb{E}\left(h_{j_1j_2}(V_{\mathbf{n},j_1j_2}+W_{\mathbf{n},j_1j_2}) - h_{j_1j_2}(V_{\mathbf{n},j_1j_2}+W^*_{\mathbf{n},j_1j_2})\right)\Big| \\ + \Big|\mathbb{E}\left(h_{j_1j_2}(V_{\mathbf{n},j_1j_2}+W^*_{\mathbf{n},j_1j_2}) - h_{j_1j_2}(V_{\mathbf{n},j_1j_2}+W'_{\mathbf{n},j_1j_2})\right)\Big|. \quad (5.28)\end{aligned}$$

Consider now the special case of complex exponential functions $h(x) = \exp(i\langle \mathbf{t}, x\rangle)$ for $\mathbf{t} \in \mathbb{R}^k$, where $\langle a, b\rangle$ denotes the scalar product in $\mathbb{R}^k$. Then, from the previous lemma, the second term of the RHS of the inequality (5.28) is bounded by

$$6\ \|h^{(2)}\|_{\infty}^{1-\delta}\ \|h^{(3)}\|_{\infty}^{\delta}\ \mathbb{E}\|W_{\mathbf{n},j_1j_2}\|^{2+\delta} \leqslant 6\ \|\mathbf{t}\|^{2+\delta}\ \ \mathbb{E}\|W_{\mathbf{n},j_1j_2}\|^{2+\delta}.$$

For the first term of the RHS of the inequality (5.28), first notice that for a $\mathbb{R}^k-$valued random vector X independent from $(W'_{\mathbf{n},j_1j_2})_{(j_1,j_2)\in D_{\mathbf{n},2}}$,

$$\begin{aligned}\mathbb{E}h_{j_1j_2}(X) &= \mathbb{E}\left(h\left(X + \sum_{u=0}^{j_1-1}\sum_{v=1}^{m_2} W'_{\mathbf{n},uv} + \sum_{v=0}^{j_2-1} W'_{\mathbf{n},j_1v}\right)\right) \\ &= \exp\left(-\frac{1}{2}\mathbf{t}^T\left(\sum_{u=0}^{j_1-1}\sum_{v=1}^{m_2} C_{\mathbf{n},uv} + \sum_{v=0}^{j_2-1} C_{\mathbf{n},j_1v}\right)\mathbf{t}\right)\cdot \mathbb{E}\left(\exp\left(i\langle \mathbf{t}, X\rangle\right)\right),\end{aligned}$$

because $W'_{\mathbf{n},j_1j_2} \sim \mathcal{N}_k\left(\mathbf{0}, C_{\mathbf{n},j_1j_2}\right)$, where $C_{\mathbf{n},j_1j_2} := \mathrm{Cov}(W_{\mathbf{n},j_1j_2})$, is the covariance matrix of the vector $W_{\mathbf{n},j_1j_2}$, for $(j_1,j_2) \in D_{\mathbf{n},2}$. For $j_1 = 0$ or $j_2 = 0$, recall that $W_{\mathbf{n},j_1j_2} = 0$. In this case, we also set $C_{\mathbf{n},j_1j_2} = 0$.

Thus,

$$
\begin{aligned}
&\left|\mathbb{E}\left(h_{j_1j_2}(V_{\mathbf{n},j_1j_2}+W_{\mathbf{n},j_1j_2})-h_{j_1j_2}(V_{\mathbf{n},j_1j_2}+W^*_{\mathbf{n},j_1j_2})\right)\right| \\
&\quad=\left|\exp\left(-\frac{1}{2}\mathbf{t}^T\left(\sum_{u=0}^{j_1-1}\sum_{v=1}^{m_2}C_{\mathbf{n},uv}+\sum_{v=0}^{j_2-1}C_{\mathbf{n},j_1v}\right)\mathbf{t}\right)\right. \\
&\qquad\qquad\left.\times\,\mathbb{E}\left(\exp\left(i\langle\mathbf{t},V_{\mathbf{n},j_1j_2}\rangle\right)\cdot\left(\exp\left(i\langle\mathbf{t},W_{\mathbf{n},j_1j_2}\rangle\right)-\exp\left(i\langle\mathbf{t},W^*_{\mathbf{n},j_1j_2}\rangle\right)\right)\right)\right| \\
&\quad=\left|\exp\left(-\frac{1}{2}\mathbf{t}^T\left(\sum_{u=0}^{j_1-1}\sum_{v=1}^{m_2}C_{\mathbf{n},uv}+\sum_{v=0}^{j_2-1}C_{\mathbf{n},j_1v}\right)\mathbf{t}\right)\right| \\
&\qquad\qquad\qquad\times\left|\mathrm{Cov}\left(\exp\left(i\langle\mathbf{t},V_{\mathbf{n},j_1j_2}\rangle\right),\exp\left(i\langle\mathbf{t},W_{\mathbf{n},j_1j_2}\rangle\right)\right)\right| \\
&\qquad\qquad\qquad\leqslant\left|\mathrm{Cov}\left(\exp\left(i\langle\mathbf{t},V_{\mathbf{n},j_1j_2}\rangle\right),\exp\left(i\langle\mathbf{t},W_{\mathbf{n},j_1j_2}\rangle\right)\right)\right|.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\Delta_{\mathbf{n}}&=\sum_{(j_1,j_2)\in D_{\mathbf{n},2}}\Delta_{\mathbf{n},j_1j_2} \\
&\leqslant\sum_{(j_1,j_2)\in D_{\mathbf{n},2}}\left(\left|\mathrm{Cov}\left(\exp\left(i\langle\mathbf{t},V_{\mathbf{n},j_1j_2}\rangle\right),\exp\left(i\langle\mathbf{t},W_{\mathbf{n},j_1j_2}\rangle\right)\right)\right|+6\|\mathbf{t}\|^{2+\delta}\mathbb{E}\|W_{\mathbf{n},j_1j_2}\|^{2+\delta}\right) \\
&=T_{\mathbf{n},\mathbf{t}}(\mathbf{f}_k)+6\|\mathbf{t}\|^{2+\delta}A_{\mathbf{n}}.
\end{aligned}
$$

□

Proof of Theorem 5.1. The assumptions (C.1) and (C.2) imply that $B_{\mathbf{n}}(\epsilon)\underset{\mathbf{n}\to\infty}{\longrightarrow}0$ and that $a_{\mathbf{n}}=\sum_{i=1}^k c(f_i,f_i)<\infty$, respectively. Therefore, taking into account Remark 5.2, we obtain from Lemma 5.2 that, for each $k\in\mathbb{N}$,

$$
\Delta_{\mathbf{n}}=\left|\mathbb{E}\left(h\left(\sum_{(i,j)\in D_{\mathbf{n},2}}W_{\mathbf{n},ij}\right)-h\left(\sum_{(i,j)\in D_{\mathbf{n},2}}W'_{\mathbf{n},ij}\right)\right)\right|\underset{\mathbf{n}\to\infty}{\longrightarrow}0,
$$

for all $\mathbf{t}\in\mathbb{R}^k$ and $h(x)=\exp(i\langle\mathbf{t},x\rangle)$, because by hypotheses, $T_{\mathbf{n},\mathbf{t}}(\mathbf{f}_k)\underset{\mathbf{n}\to\infty}{\longrightarrow}0$ for all $\mathbf{t}\in\mathbb{R}^k$ and all $\mathbf{f}_k=(f_1,\ldots,f_k)\in\mathcal{F}^k$. Notice that $W'_{\mathbf{n}}:=\sum_{(i,j)\in D_{\mathbf{n},2}}W'_{\mathbf{n},ij}\sim\mathcal{N}_k(\mathbf{0},\mathrm{Cov}(W_{\mathbf{n},11}))$ and that $|\mathbb{E}\left(h(W'_{\mathbf{n}})-h(W)\right)|\underset{\mathbf{n}\to\infty}{\longrightarrow}0$, where $W\sim\mathcal{N}_k(\mathbf{0},\Sigma_k)$, with $\Sigma_k=(c(f_i,f_j))_{(i,j)\in[k]^2}$. Using triangular inequality, we deduce that

$$
\left|\mathbb{E}\left(h\left(\sum_{(i,j)\in D_{\mathbf{n},2}}W_{\mathbf{n},ij}\right)-h(W)\right)\right|\underset{\mathbf{n}\to\infty}{\longrightarrow}0,
$$

and therefore $(Z_{\mathbf{n}}(f_1), \ldots, Z_{\mathbf{n}}(f_k)) = \sum_{(i,j)\in D_{\mathbf{n},2}} W_{\mathbf{n},ij} \xrightarrow[\mathbf{n}\to\infty]{\mathcal{D}} W$.

□

Proof of Proposition 5.1. Consider the expression (5.17) of the iso-extremogram estimator. Then, for $(h_s, h_t) \in [0 : L_s] \times [0, L_t]$, we obtain that

$$\frac{\sqrt{nv_n}}{r_1 r_2}\left(\widehat{\rho^*}_{A,B}(h_s,h_t) - \rho^*_{A,B,n}(h_s,h_t)\right) = \frac{Z_{\mathbf{n}}(f_{A,B,h_s,h_t}) - \left(\frac{mh_t v_n}{\sqrt{nv_n}} + Z_{\mathbf{n}}(f_{A,A,0,0})\right)\rho^*_{A,B,n}(h_s,h_t) + \frac{\delta_{\mathbf{n}}}{\sqrt{nv_n}} + R}{\frac{r_1r_2}{\sqrt{nv_n}} Z_{\mathbf{n}}(f_{A,A,0,0}) + 1 + \frac{r_1 r_2 R_{A,A,0,0}}{nv_n}}, \tag{5.29}$$

where $Z_{\mathbf{n}}(\cdot)$ denotes the empirical process of cluster functionals (5.4). Besides, here $R = \frac{1}{\sqrt{nv_n}}\left(R_{A,B,h_s,h_t} - \rho^*_{A,B,n} R_{A,A,0,0}\right)$ and $m = m_1 m_2 m_3$.

Now, notice that Chebyshev's inequality applied on the random variables R and $r_1 r_2 (nv_n)^{-1} R_{A,A,0,0}$ implies that these variables converge to zero in probability as $\mathbf{n} \to \infty$. Similarly, applying Chebyshev's inequality together with the condition $\sqrt{nv_n} = o(r)$, we prove that $(nv_n)^{-1/2}\delta_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{P} 0$. This last condition ($\sqrt{nv_n} = o(r)$) also guarantees that $mh_t v_n (nv_n)^{-1/2} \xrightarrow[\mathbf{n}\to\infty]{} 0$. Again, Chebyshev's inequality on the random variable $\frac{r_1 r_2}{\sqrt{nv_n}} Z_{\mathbf{n}}(f_{A,A,0,0})$, followed by the condition (C.3) and $r = o(nv_n r_3)$, implies that this converges to zero in probability as $\mathbf{n} \to \infty$.

Thus,

$$\frac{\sqrt{nv_n}}{r_1 r_2}\left(\widehat{\rho^*}_{A,B}(h_s,h_t) - \rho^*_{A,B,n}(h_s,h_t)\right) = Z_{\mathbf{n}}(f_{A,B,h_s,h_t}) - \rho^*_{A,B,n}(h_s,h_t) Z_{\mathbf{n}}(f_{A,A,0,0}) + o(1).$$

The assumption 3 of the proposition and the existence of the covariance function $\sigma_{A,B}$ of (C.2′) implies, from Theorem 5.1, that $(Z_n(f_{A,B,h_s,h_t}))_{(h_s,h_t)\in[0:L_s]\times[0:L_t]}$ converges to a centred gaussian variable with covariance matrix $(\sigma_{A,B}(\mathbf{h},\mathbf{h}'))_{\mathbf{h},\mathbf{h}'\in[0:L_s]\times[0:L_t]}$, for each $(L_s, L_t) \in \mathbb{N}_0^2$. Using the same argument, we prove that $Z_n(f_{A,A,0,0})$ converges to a centred gaussian variable with variance $\sigma'_{A,A}(0)$.

Finally, considering the existence of the covariance function $\sigma'_{A,B}$ in (C.2′), we obtain the result. □

Bibliography

Adler, R. A. *The geometry of random fields*, (1981) John Wiley, New York. 77

Andrews, D. *Non strong mixing autoregressive processes*. J. Appl. Prob. 21 (1984): 930 – 934. 4, 12, 85

Ango Nze P. & Doukhan P. *Weak dependence and applications to econometrics*. Econom Theory 20 (2004) : 995 – 1045. 4, 12

Bardet, J.M., Doukhan, P., Lang, G. & Ragache, N. *Dependent Lindeberg Central Limit Theorem and Some Applications*. ESAIM: Probability and Statistics 12 (2007) : 154 – 172. 6, 10, 12, 13, 18, 36, 99, 101, 110

Basrak, B. & Segers, J. *Regularly varying multivariate time series*. Stoch. Proc. Appl. 119 (2009) : 1055 – 1080. 29

Bradley, R. C. *On the strong mixing and weak Bernoulli conditions*. Z. Wahrsch. Verw. Gebiete 51 (1980): 49–54.

Bradley, R. C. *Central limit theorems under weak dependence*. J. Multivar. Anal. 11 (1981) : 1–16.

Davis R.A. & Mikosch, T. *Extreme value theory for space-time processes with heavy-tailed distributions*. Stochastic Processes and their Applications 118 (2008) : 560 – 584. 104

Davis R.A. & Mikosch, T. *The extremogram: a correlogram for extreme events*. Bernoulli 15 (2009) : 977 – 1009. v, vi, 1, 11, 13, 27, 29, 48

Davis R.A. & Mikosch, T. *Towards estimating extremal serial dependence via the bootstrapped extremogram*. Journal of Econometrics 170 (2012) : 142 – 152. 34

Dedecker, J. & Prieur, C. *Coupling for $\tau-$dependent sequences and applications*. Journal of Theor. Prob. 17 – 4 (2004) : 861 – 885. v, vi, 7, 8, 47, 49, 50, 51, 84

Dedecker, J. & Prieur, C. *Couplage pour la distance minimale*. C. R. Acad. Sci. Paris Serie 1 338-10 (2004) : 805 – 808. 8, 49, 50, 63

Dedecker, J. & Prieur, C. *New dependence coefficients. Examples and applications to statistics*. Prob. Theor. and Rel. Fields 132 (2005) : 203 – 236. 8

Dedecker, J., Doukhan, P., Lang, G., León, J.R., Louhichi,S. & Prieur, C. *Weak dependence: With Examples and Applications*. Lecture Notes in Statistics 190 (2007), Springer-Verlag. 6, 8, 19, 21, 22, 23, 81, 84

Doukhan, P. *Mixing: properties and examples*. Lecture Notes in Statistics 85 (1994), Springer-Verlag. 5, 12, 78

Doukhan, P., Massart, P., and Rio, E. *Invariance principles for absolutely regular empirical processes*. Annales de l'institut Henri Poincaré., section B, Probabilités et Statistiques, tome 31, no. 2 (1995) : 393–427. 7

Doukhan, P. *Stochastic Models for Time Series*. Mathématiques et applications (2007), Springer-Verlag.

Doukhan, P. & Louhichi, S. *A new weak dependence condition and applications to moment inequalities*. Stoch. Proc. Appl. 84 (1999) : 313 – 342. v, vi, 1, 4, 5, 6, 11, 12, 18, 19, 20, 77, 78, 79, 81, 84, 103

Doukhan, P. & Lang, G. *Rates in the empirical central limit theorem for stationary weakly dependent random fields*. Stat. Inference Stoch. Process. 5 (2002) : 199 – 228. 81, 84, 85

Doukhan, P., Teyssiere, G. and Winant, P. *A LARCH(∞) Vector Valued Process*. In: P. Bertail, P. Doukhan and P. Soulier (eds). Lecture Notes in Statistics, Dependence in Probability and Statistics 187 (2006): 245 –258. 83

Doukhan, P. & Truquet, L. *A fixed point approach to model random fields*. Alea 3 (2007) : 111-132. 84

Doukhan, P. & Gómez, J.G. *Limit theorems for empirical processes of cluster functionals without mixing*. Submitted to Extremes (April 2017).

Drees, H. *Weighted Approximations of Tail Processes for $\beta-$Mixing Random Variables*. Ann. Appl. Probab. 10 (2000) : 1274-1301. 2

Drees, H. *Tail empirical processes under mixing conditions*. In: H.G. Dehling, T. Mikosch and M. Sorensen (eds.), Empirical Process Techniques for Dependent Data (2002) : 325-342. 2

DREES, H. *Extreme Quantile Estimation for Dependent Data with Applications to Finance*. Bernoulli 9 (2003): 617-657.

DREES, H. *Bias correction for blocks estimators of the extremal index*. Preprint `https://arxiv.org/pdf/1107.0935.pdf` (2011). v, vi, 1, 48, 57, 58, 72

DREES, H & ROOTZÉN, H. *Limit Theorems for Empirical Processes of Cluster Functionals*. Ann. Stat. 4 (2010) : 2145 – 2186. v, vi, 3, 4, 11, 12, 13, 17, 18, 40, 47, 48, 61, 70, 71, 76

EBERLEIN, E. *Weak convergence of partial sums of absolutely regular sequences*. Statist. Prob. Letters 2 (1984) : 291 – 293. 18

FEDERER, H. *Geometric Measure Theory* (1969), Springer, Berlin. 66

FERREIRA, H. & PEREIRA, L. *How to compute the extremal index of stationary random fields*. Statistics & Probability Letters 78(11), (2008) : 1301 – 1304. 78

GÓMEZ, J.G. *Dependent Lindeberg CLT for the fidis of empirical processes of cluster functionals.* Accepted in Statistics: A Journal of Theoretical and Applied Statistics (2017). 48, 49

Gómez, J.G. *On extreme values in stationary weakly dependent random fields*. In preparation.

Gómez, J.G. *Extreme cluster functionals in stationary weakly dependent random fields and space-time processes*. In preparation. 103

HENZE, N. & ZIRKLER, B. *A class of invariant consistent tests for multivariate normality*. Communications in Statistics - Theory and Methods 19 - 10 (1990) : 3595 – 3617. 32

KORKMAZ, S., GOKSULUK, D. AND ZARASIZ, G. *An R package for assessing multivariate normality*. The R Journal 6/2 (2014) : 151 – 162. 32

LEADBETTER, M.R., LINDGREN, G. AND ROOTZÉN, H. *Extremes and Related Properties of Random Sequences and Processes* (1983), Springer - Verlag, New York. 9, 77, 88

LEADBETTER, M.R. & ROOTZÉN, H. *On Extreme Values in Stationary Random Fields*. Stochastic Processes and Related Topics, Trends Math. Birkhauser Boston (1998) : 275 – 285. 9, 77, 78, 79, 85, 88, 95

MARDIA, K. V. *Measures of multivariate skewness and kurtosis with applications*. Biometrika 57 - 3 (1970) : 519 – 530. 32

Newman, C.M. *Asymptotic independence and limit theorems for positively and negatively dependent random variables*. In: Tong, Y.L. (Ed.), Inequalities in Statistics and Probability, IMS Lecture Notes-Monograph Series; Vol. 5 (1984) : 127 – 140. 12

Perfekt, R. *Extremal behaviour of stationary Markov chains with applications*. Ann. of Appl. Prob. 4 (1994) : 529 – 548.

Petrov, V. V. *Sums of Independent Random Variables* (1975), Springer, Berlin. 67

Pickands, J. *Statistical inference using extreme order statistics*. Ann. Statist. 3 (1975) : 119 – 131. 15

Rootzén, H. *The tail empirical process for stationary sequences*. Technical report, Department of Mathematics, Chalmers University, Sweden; (1995). 2

Rootzén, H., Leadbetter, M. R. and Haan, L. *On the distribution of tail array sums for strongly mixing stationary sequences*. Ann. Appl. Prob. 8 (1998) : 868 – 885. 3

Rootzén, H. *Weak convergence of the tail empirical function for dependent sequences*. Stoch. Proc. Appl. 119 (2009) : 468–490. 2

Rosenblatt, M. *A central limit theorem and a strong mixing condition*. Proc. Nat. Ac. Sc. U.S.A. 42 (1956) : 43 – 47. 12, 78

Royston, P. *Approximating the Shapiro-Will W test for non-normality*. Statistics and Computing 2 - 3 (1992) : 117 – 119. 32

Segers, J. *Functionals of clusters of extremes*. Adv. Appl. Prob. 35 (2003) : 1028 – 1045. 3, 25, 40, 53

Yun, S. *The extremal index of a higher-order stationary Markov chain*. Ann. Appl. Prob. 8 (1998) : 408 – 437.

Yun, S. *The distribution of cluster functionals of extreme events in a dth-order Markov chain*. J. Appl. Prob. 37 (2000) : 29 – 44. 3

Van Der Vaart, A.W. and Wellner, J. A. *Weak Convergence and Empirical Processes. With Applications to Statistics*. Springer Series in Statistics (1996), Springer-Verlag. 8, 18, 48, 49, 54, 55, 56, 71