UNIVERSITÉ LILLE III - CHARLES-DE-GAULLE, VILLENEUVE D'ASCQ, FRANCE.

UNIVERSITÉ GASTON BERGER, SAINT-LOUIS, SÉNÉGAL.

École Doctorale SESAM
École Doctorale des Sciences et des Technologies

Thèse de doctorat

Mention : Mathématiques et Applications

présentée par

Aladji BASSÈNE

Contribution à la modélisation spatiale des événements extrêmes.

---

dirigée par Sophie Dabo-Niang et Aliou Diop
co-dirigée par Baba THIAM.

Soutenue le 06 mai 2016 devant le jury composé de :

M. Célestin C. Kokonendji	Rapporteur
M. Djamel Meraghni	Rapporteur
Mme Laurence Broze	Examinatrice
M. Mohamed El Machkouri	Examinateur
M. Abdou Ka Diongue	Examinateur
Mme Sophie Dabo	Directrice
M. Aliou Diop	Directeur
M. Baba Thiam	Co-encadrant

Résumé

Résumé en français

Dans cette de thèse, nous nous intéressons à la modélisation non paramétrique de données extrêmes spatiales. Nos résultats sont basés sur un cadre principal de la théorie des valeurs extrêmes, permettant ainsi d'englober les lois de type Pareto. Ce cadre permet aujourd'hui d'étendre l'étude des événements extrêmes au cas spatial à condition que les propriétés asymptotiques des estimateurs étudiés vérifient les conditions classiques de la **Théorie des Valeurs Extrêmes (TVE)** en plus des conditions locales sur la structure des données proprement dites. Dans la littérature, il existe un vaste panorama de modèles d'estimation d'événements extrêmes adaptés aux structures des données pour lesquelles on s'intéresse. Néanmoins, dans le cas de données extrêmes spatiales, hormis les modèles max-stables, il n'en existe que peu ou presque pas de modèles qui s'intéressent à l'estimation fonctionnelle de l'indice de queue ou de quantiles extrêmes. Par conséquent, nous étendons les travaux existants sur l'estimation de l'indice de queue et des quantiles dans le cadre de données indépendantes ou temporellement dépendantes. La spécificité des méthodes étudiées réside sur le fait que les résultats asymptotiques des estimateurs prennent en compte la structure de dépendance spatiale des données considérées, ce qui est loin d'être trivial. Cette thèse s'inscrit donc dans le contexte de la statistique spatiale des valeurs extrêmes. Elle y apporte trois contributions principales.

• Dans la première contribution de cette thèse permettant d'appréhender l'étude de variables réelles spatiales au cadre des valeurs extrêmes, nous proposons une estimation de l'indice de queue d'une distribution à queue lourde. Notre approche repose sur l'estimateur de Hill (1975). Les propriétés asymptotiques de l'estimateur introduit sont établies lorsque le processus spatial est adéquatement approximé par un processus $M-$dépendant, linéaire causal ou lorsqu'il satisfait une condition de mélange fort ($\alpha-$mélange).

• Dans la pratique, il est souvent utile de lier la variable d'intérêt Y avec une co-variable X. Dans cette situation, l'indice de queue dépend de la valeur observée x de la co-variable X et sera appelé indice de queue conditionnelle. Dans la plupart des applications, l'indice de queue des valeurs extrêmes n'est pas l'intérêt principal et est utilisé pour estimer par exemple des quantiles extrêmes. La contribution de ce chapitre consiste à adapter l'estimateur de l'indice de queue introduit dans la première partie au cadre conditionnel et d'utiliser ce dernier afin de proposer un estimateur des quantiles conditionnels extrêmes. Nous examinons les modèles dits "à plan fixe" ou "fixed design" qui correspondent à la situation où la variable explicative est déterministe et nous utlisons l'approche de la fenêtre mobile ou "window moving approach" pour capter la co-variable. Nous étudions le comportement asymptotique des estimateurs proposés et donnons des résultats numériques basés sur des données simulées avec le logiciel "R".

• Dans la troisième partie de cette thèse, nous étendons les travaux de la deuxième partie au cadre des modèles dits "à plan aléatoire" ou "random design" pour lesquels les données sont des observations spatiales d'un couple (Y, X) de variables aléatoires réelles. Pour ce dernier modèle, nous proposons un estimateur de l'indice de queue lourde en utilisant la méthode des noyaux pour capter la co-variable. Nous utilisons un estimateur de l'indice de queue conditionnelle appartenant à la famille de l'estimateur introduit par Goegebeur et al. (2014b).

Mots-clefs

Statistique spatiale ; Données extrêmes ; Données M-dépendantes ; Processus linéaire causal ; Processus $\alpha-$mélangeant ; Estimation non paramétrique ; Estimateur à noyau ; Estimation de l'indice de queue ; Estimation de quantiles extrêmes ; Estimateur de Hill ; Consistance ; Normalité asymptotique.

Contributions to modeling spatial extremal events and applications

Abstract

In this thesis, we investigate nonparametric modeling of spatial extremes. Our results are based on the main result of the theory of extreme values, thereby encompass Pareto laws. This framework allows today to extend the study of extreme events in the spatial case provided if the asymptotic properties of the proposed estimators satisfy the standard conditions of the Extreme Value Theory (EVT) in addition to the local conditions on the data structure themselves. In the literature, there exists a vast panorama of extreme events models, which are adapted to the structures of the data of interest. However, in the case of extreme spatial data, except max-stables models, little or almost no models are interested in non-parametric estimation of the tail index and/or extreme quantiles. Therefore, we extend existing works on estimating the tail index and quantile under independent or time-dependent data. The specificity of the methods studied resides in the fact that the asymptotic results of the proposed estimators take into account the spatial dependence structure of the relevant data, which is far from trivial. This thesis is then written in the context of spatial statistics of extremes. She makes three main contributions.

• In the first contribution of this thesis, we propose a new approach of the estimator of the tail index of a heavy-tailed distribution within the framework of spatial data. This approach relies on the estimator of Hill (1975). The asymptotic properties of the estimator introduced are established when the spatial process is adequately approximated by a spatial $M-$dependent process, spatial linear causal process or when the process satisfies a strong mixing condition.

• In practice, it is often useful to link the variable of interest Y with covariate X. In this situation, the tail index depends on the observed value x of the covariate X and the unknown fonction $\gamma(.)$ will be called conditional tail index. In most applications, the tail-index of an extreme value is not the main attraction, but it is used to estimate for instance extreme quantiles. The contribution of this chapter is to adapt the estimator of the tail index introduced in the first part in the conditional framework and use it to propose an estimator of conditional extreme quantiles. We examine the models called "fixed design" which corresponds to the situation where the explanatory variable is deterministic. To tackle the covariate, since it is deterministic, we use the window moving approach. We study the asymptotic behavior of the estimators proposed and some numerical results using simulated data with the software "R".

- In the third part of this thesis, we extend the work of the second part of the frame models called "random design" for which the data are spatial observations of a pair (Y, X) of real random variables . In this last model, we propose an estimator of heavy tail-index using the kernel method to tackle the covariate. We use an estimator of the conditional tail index belonging to the family of the estimators introduced by Goegebeur et al. (2014b).

Keywords

Spatial statistics ; Extreme values ; Spatial $M-$dependent processes ; Spatial linear causal processes ; $\alpha-$ mixing processes ; Nonparametric estimation ; Kernel estimator ; Heavy tail index estimate ; Extreme quantiles estimate ; Hill's estimator ; Consistency ; Asymptotic normality.

Table des matières

Remerciements

Cette thèse n'aurait jamais pu voir le jour sans le soutien d'un grand nombre de personnes que je tiens à remercier pour leurs encouragements, leur générosité et leur gentillesse.

En premier lieu, je tiens à exprimer ma profonde gratitude à l'égard de ma directrice de thèse Sophie DABO-NIANG et la confiance qu'elle m'a accordé en acceptant d'encadrer ma thèse et de l'avoir si bien fait. J'aimerai également lui dire à quel point j'ai apprécié ses compétences, sa rigueur et sa clairvoyance. J'ai été extrêmement sensible à ses qualités humaines d'écoute, de disponibilité et de bienveillance, sans compter son humour et sa simplicité. J'ai pris un trés grand plaisir à travailler à ses côtés ; Sophie, merci de m'avoir proposé cette thèse à la fin du master. Merci aussi pour ta grande maîtrise de la statistique spatiale. Et enfin, merci pour ton soutien dans les moments difficiles. Il y en a eu quelques uns, voir même beaucoup, pendant cette thèse, et elle a su, à sa manière, trouver chaque fois les mots justes pour me remotiver. Par ailleurs, j'admets (seulement maintenant !) être sûrement quelqu'un de difficile à encadrer ; j'ai donc dû lui causer des soucis plus d'une fois. J'ai pourtant essayé de faire du mieux que je pouvais... En tout cas, merci pour tout.

Je souhaite également remercier profondément mon co-directeur de thèse Aliou DIOP qui malgré la distance a toujours été présent. Par ses conseils avisés, ses multiples relectures et remarques pertinentes, il m'a offert le meilleur encadrement possible afin de bien mener mes travaux de recherche. J'ai beaucoup apprécié sa gentillesse, sa disponibilité et sa patience. Merci aussi pour ta grande maîtrise de la théorie des valeurs extrêmes. Aliou, Je ne saurais jamais te remercier assez tant sur le plan humain que sur le plan professionnel.

Je remercie aussi mon co-encadrant Baba THIAM, pour ses conseils, ses multiples relectures, remarques pertinentes et ses conseils pour l'après thèse. Baba, aujourd'hui je ne saurais te remercier assez pour tout ce que tu as fait pour moi et sache que je t'en suis très reconnaissant.

Merci à tous les trois de m'avoir tant apporté aussi bien humainement que scientifiquement. Merci pour nos discussions si enrichissantes et d'avoir toujours été disponibles pour moi. J'espère avoir été digne de la confiance que vous m'avez accordée et que nous continuerons à travailler ensemble. J'ai beaucoup appris à vos côtés et je suis très honoré de vous avoir eu pour encadreurs.

Je remercie vivement les rapporteurs Célestin C. KOKONENDJI et Djamel MERAGHNI d'avoir eu la gentillesse et la patience de lire mes travaux. Merci également à Laurence BROZE , Mohamed EL MACHKOURI et Abdou KA DIONGUE de me faire l'honneur de faire partie de mon jury.

Mes remerciements vont également aux autres membres de l'équipe et plus particulièrement à celles et ceux avec qui j'ai partagé mon bureau : Camille, Emad, Floriant, Med

Salem et Sara pour nos discussions qui portent sur les *mathématiques*, nos travaux futurs et surtout sur nos différentes cultures. J'ai eu le plaisir et la chance de partager le bureau C06 avec vous. Il n'est pas toujours évident de partager un bureau avec des personnes aussi gentilles et compréhensibles que vous.

Je me dois également remercier l'ensemble des personnes présentes à l'UFR MIME pour leur gentillesse, leur efficacité et surtout le travail fourni pour permettre le bon déroulement de chaque année universitaire et l'organisation des missions.

J'en profite également pour remercier toute ma famille et mes ami(e)s pour ce que nous partageons ensemble : discussions, rires, sorties, expériences culinaires, dance, repas, et bien d'autres. Merci à mon cousin, Jean Goubibé BASSÈNE, Ali Yassin SANGHARÉ et à Pape Ibrahima KAMARA (PIK) de m'avoir gentiment aidé et à relire mes fichiers. PIK, merci de m'avoir toujours soutenu dans les moments difficiles avec ton franc parler, d'avoir accepté de relire mes articles et surtout de m'avoir encouragé jusqu'au bout de cette thèse.

Pour terminer, je remercie très sincèrement ma famille qui m'a toujours encouragé et soutenu. Je n'aurais jamais pu réaliser tout ce parcours sans leur grande présence ! Et Merci aussi à mon défunt oncle Jean Pierre KANFOM, qui est pour moi un papa. Oncle, aujourd'hui, je comprends mieux pourquoi tu insistais pour que je fasse une thèse et de me répéter tout le temps *"Le but d'une encyclopédie est de rassembler les connaissances éparses..."*

Jean Pierre, je te dédie cette thèse !

"Le but d'une encyclopédie est de rassembler les connaissances éparses..."

Lexique

Notations (Français)

$\mathbb{R}$	ensemble des nombres réels : $]-\infty, +\infty[$
$\mathbb{R}^d$	espace euclidien réel de dimension d
$\mathbb{Q}$	ensemble des nombres rationnels : $\frac{k}{p}$
$\mathbb{Z}$	ensemble des entiers relatifs :$\cdots, -2, -1, 0, 1, 2, \cdots$
$\mathbb{N}$	ensemble des entiers naturels : $0, 1, 2, \cdots$
$\mathbb{N}^*$	ensemble des entiers naturels non nuls : $1, 2, \cdots$
$\{a_i : i \in I\}$	ensemble des points a_i indexés par un ensemble I
$\overline{A}$ (ou A^c)	complémentaire de A
$A \cup B$	union de A et B
$A \cap B$	intersection de A et B
$A \subset B$	A est strictement inclus dans B
$A \subseteq B$	A est inclus ou égal à B
$\mathrm{Card}(A)$	cardinal de A
$\Vert . \Vert$	norme
	Soit un vecteur $X = (x_1, \cdots, x_p)$
$X^\top$	transposée de X
$\Vert X \Vert_1$	norme 1 : $\Vert X \Vert_1 = \sqrt{\lvert x_1 \rvert + \cdots + \lvert x_d \rvert}$
$\Vert X \Vert_2$	norme euclidienne : $\Vert X \Vert_2 = \sqrt{x_1^2 + \cdots + x_d^2}$
$\Vert X \Vert_\infty$	norme infinie : $\Vert X \Vert_\infty = \max\left(\lvert x_1 \rvert, \cdots, \lvert x_p \rvert\right)$
	remarque : $\Vert X \Vert_\infty \leq \Vert X \Vert_2 \leq \Vert X \Vert_1$
$\lfloor a \rfloor$	partie entière de a
$*$	produit scalaire
■	fin d'une preuve
$X_{\mathbf{i}}$	variable X observée au site **i**
i.i.d	indépendantes et identiquement distribuées
$\sigma(...)$	tribu engendrée par (...)

$(\Omega, \mathcal{F}, P)$	espace de probabilité Ω : ensemble non vide $\mathcal{F}$: $\sigma-$Algèbre de sous-ensemble de Ω P : mesure de probabilité sur $\mathcal{F}$.
$\xrightarrow{\mathbb{P}}$	covergence en probabilité
$\xrightarrow{\mathcal{D}}$	convergence en loi
$a.s.$	presque sûrement
$B(t, r)$	boule ouverte de centre t et de rayon r
$\mathrm{d}\,(.,.)$	métrique sur un espace métrique E, $\mathrm{d} = d$
(E, d)	espace métrique associé à la métrique d
$\mathcal{L}_r\,(E, \mathcal{F}, \mu)$	(ou $\mathcal{L}_r\,(E)$ ou $\mathcal{L}_r\,(\mathcal{F})$ ou $\mathcal{L}_r\,(\mathcal{B}(\mathbf{I}))$ avec $\mathcal{B}(\mathbf{I}) \subseteq \mathcal{F}$) espace de classes des fonctions mesurables f telles que $\Vert f\Vert_r = \left(\int_E \lvert f\rvert^r \, d\mu\right)^{1/r} < +\infty, \quad 1 \leq r < +\infty$ $\Vert f\Vert_\infty = \inf\left\{a : \mu\left\{f > a\right\} = 0\right\} < +\infty, \quad r = +\infty$
$a_n = O\left(b_n\right)$	il existe une constante c telle que $a_n \leq c b_n$
$a_n = o\left(b_n\right)$	$a_n/b_n \longrightarrow 0$
$a_n = O_{\mathbb{P}}\left(b_n\right)$	si $\forall \epsilon > 0,\ \exists M_\epsilon,\ \mathbb{P}\left(\lvert a_n\rvert \geq M_\epsilon b_n\right) \leq \epsilon$
$a_n = o_{\mathbb{P}}\left(b_n\right)$	$a_n/b_n \xrightarrow{\mathbb{P}} 0$

Notations (English)

$\mathbb{R}$	set of real numbers : $]-\infty, +\infty[$
$\mathbb{R}^d$	Euclidean space of dimension d
$\mathbb{Q}$	set of rational numbers : $\frac{k}{p}$
$\mathbb{Z}$	set of integers :$\cdots, -2, -1, 0, 1, 2, \cdots$
$\mathbb{N}$	set of natural numbers : $0, 1, 2, \cdots$
$\mathbb{N}^*$	set of non-zero natural numbers : $1, 2, \cdots$
$\{a_i : i \in I\}$	set of points a_i indexed by a set I
$\overline{A}$ (or A^c)	complement of A
$A \cup B$	union of A and B
$A \cap B$	intersection of A and B
$A \subset B$	A is strictly included in B
$A \subseteq B$	A is included or equal to B
$\mathrm{Card}(A)$	cardinality of A
$\lVert . \rVert$	norm
	Let a vector $X = (x_1, \cdots, x_p)$
$X^\top$	transpose of X
$\lVert X \rVert_1$	1-norm : $\lVert X \rVert_1 = \sqrt{\lvert x_1 \rvert + \cdots + \lvert x_d \rvert}$
$\lVert X \rVert_2$	euclidean norm : $\lVert X \rVert_2 = \sqrt{x_1^2 + \cdots + x_d^2}$
$\lVert X \rVert_\infty$	infinity norm : $\lVert X \rVert_\infty = \max(\lvert x_1 \rvert, \cdots, \lvert x_p \rvert)$
	remark : $\lVert X \rVert_\infty \leq \lVert X \rVert_2 \leq \lVert X \rVert_1$
$\lfloor a \rfloor$	integer part of a
$*$	scalar product
$\blacksquare$	end of a proof
$X_{\mathbf{i}}$	variable X observed at site **i**
i.i.d	independent and identically distributed
$\sigma(...)$	$\sigma-$Algebra generated by (...)
$(\Omega, \mathcal{F}, P)$	probability space
	Ω : nonempty set
	$\mathcal{F}$: $\sigma-$algebra of subset of Ω
	P : probability measure on $\mathcal{F}$.
$\xrightarrow{\mathbb{P}}$	convergence in probability
$\xrightarrow{\mathcal{D}}$	convergence in distribution
$a.s.$	almost surely

$B(t,r)$	open ball of centre t and radius r
$\mathrm{d}(.,.)$	metric on a metric space E, $\mathrm{d} = d$
(E,d)	metric space and its metric d
$\mathcal{L}_r(E,\mathcal{F},\mu)$	(ou $\mathcal{L}_r(E)$ ou $\mathcal{L}_r(\mathcal{F})$ ou $\mathcal{L}_r(\mathcal{B}(\mathbf{I}))$ avec $\mathcal{B}(\mathbf{I}) \subseteq \mathcal{F}$) space of classes of mesurable functions f such that $\lVert f\rVert_r = \left(\int_E \lvert f\rvert^r \, d\mu\right)^{1/r} < +\infty, \quad 1 \le r < +\infty$ $\lVert f\rVert_\infty = \inf\{a : \mu\{f > a\} = 0\} < +\infty, \quad r = +\infty$
$a_n = O(b_n)$	there exists a constant c such that $a_n \le c b_n$
$a_n = o(b_n)$	$a_n/b_n \longrightarrow 0$
$a_n = O_{\mathbb{P}}(b_n)$	if $\forall \epsilon > 0,\ \exists M_\epsilon,\ \mathbb{P}(\lvert a_n\rvert \ge M_\epsilon b_n) \le \epsilon$
$a_n = o_{\mathbb{P}}(b_n)$	$a_n/b_n \xrightarrow{\mathbb{P}} 0$

Chapitre 1

Introduction générale

Sommaire

"Il est impossible que l'improbable n'arrive jamais."

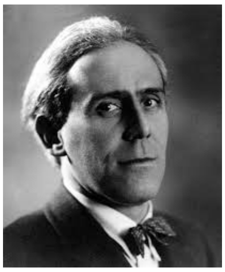

Figure 1.1 – Emile Julius Gumbel, 1891-1996.

La théorie des valeurs extrêmes a longtemps reçu beaucoup d'attention sur le plan théorique. Depuis quelques décennies, cette théorie est devenue incontournable dans la modélisation des événements extrêmes. Le développement des outils informatiques et les entrepôts de données bien fournis permettent de nos jours de diversifier les domaines d'applications de la théorie des valeurs extrêmes. Ainsi, les domaines d'applications de cette théorie sont très variés : hydrologie, biologie, ingénierie, gestion de l'environnement, météorologie, finance, assurance, sciences sociales, etc. Cette théorie est associée aux phénomènes dits rares voire même improbables.

1.1 Motivations

Les phénomènes rares et/ou catastrophiques dominent l'actualité quotidienne par leur caractère imprévisible. Ils sont variés et souvent de caractères physiques en particulier les catastrophes naturelles : les séismes, les éruptions volcaniques, les tsunamis, les mouvements de terrain, les inondations, les tempêtes, les cyclones, les orages etc faisant ainsi des ravages sur leur passage. Les exemples ci-dessous illustrent les dégâts matériels et humains que peut engendrer un événement catastrophique rare.

FIGURE 1.2 – Le 1er Février 1953, lors d'une forte tempête, la mer passe par-dessus plusieurs digues aux Pays-Bas (à gauche) et le village Wieringerwerf (avant et après).

FIGURE 1.3 – Tempête Xynthia, La Faute-sur-Mer, 1 Mars 2010 : Hauteur de vagues exceptionnelles et beaucoup de dommages enregistrés.

À l'échelle mondiale, on recense annuellement environ un millier de grandes catastrophes « naturelles » en majeure partie provoquées par les crues, événements naturels les plus fréquents et les plus destructeurs ; leurs causes initiales sont toujours météorologiques : moussons, cyclones, tempêtes (Source Wikipédia (2016)).

En confondant les effets désastreux de certains de ces événements avec les causes des catastrophes qui en résultent, il a été longtemps considéré que les causes étaient des punitions et que les effets étaient inéluctables, fatals, prescrits... La science et la technique permettent maintenant de caractériser les événements, de prévoir leurs effets, d'établir et distinguer les causes naturelles d'avec les causes humaines des catastrophes pour améliorer la prévention et la gestion des secours.

1.1.1 Approche probabiliste

L'approche standard en théorie des probabilités place l'accent sur le comportement en moyenne et la variabilité autour de la moyenne, par le biais d'outils probabilistes comme par exemple la loi des grands nombres, le théorème central limite ou encore l'analyse de la variance. Cette approche ne fournit pas d'informations fiables sur les évènements extrêmes c'est-à-dire sur les queues de distributions des événements. Pour caractériser et quantifier le comportement de ces événements extrêmes, une nouvelle théorie est nécessaire, la **Théorie des Valeurs Extrêmes (TVE).** Cette théorie englobe des modèles stochastiques extrêmes adéquats pour modéliser et décrire la survenue et l'intensité d'évènements dits rares c'est-à-dire qui présentent des variations à très grandes amplitudes ou à très faibles amplitudes et ayant une très faible probabilité d'apparition. L'étude des lois de ces événements extrêmes n'est possible que si le comportement de ces derniers est dû au hasard (notion de probabilité). Ils sont dits extrêmes quand il s'agit de valeurs beaucoup plus grandes ou plus petites que celles observées habituellement. La modélisation de tels événements est de nos jours un champ de recherches particulièrement actif due notamment à l'importance de leurs impacts socio-économiques et à la longue collecte de données enregistrant de tels événements. L'analyse des valeurs extrêmes requiert l'estimation d'un indice de queue qui donne une indication essentielle sur la forme de la queue de distribution des événements.

Cette théorie développée par Fisher et al. (1928) sur les lois limites possibles du maximum d'un échantillon a montré que la théorie des valeurs extrêmes était quelque chose de spécial et pas comme la théorie classique de la limite centrale. Depuis, ce résultat de Fisher et al. (1928) a été étudié par Gnedenko (1943) qui obtient rigoureusement la convergence, dont la preuve fut simplifiée par Haan (1976). L'unification de ce résultat est due aux travaux de Von Mises (1936) et Jenkinson (1955). Cette analyse repose principalement sur des distributions limites des extrêmes et leurs domaines d'attraction. Cependant, on y retrouve deux modèles :

- La loi généralisée des valeurs extrêmes **(GEV : «Generalized Extreme Value»)**,
- La loi de Paréto généralisée **(GPD : «Generalized Pareto Distribution»)**.

1.1.2 Limites de cette approche

Par définition, les événements extrêmes sont peu nombreux de par leurs fréquences d'apparitions rares voir même inexistants rendant ainsi leurs modélisations difficiles. L'information la plus précise est celle contenue dans les valeurs observées les plus extrêmes. De ce fait, à partir de peu de données, on doit construire des modèles nous permettant d'extrapoler et de prédire un événement sans commune mesure ce qui conduit à deux problèmes en pratique.

- Le premier est lié à la taille de l'échantillon qui est souvent faible remettant en question l'applicabilité des résultats asymptotiques (La taille minimale $n = 50$ a été recommandée par Stedinger (2000) pour avoir des estimations robustes). Néanmoins, cette taille ne peut pas fournir assez d'informations si on s'intéresse à la prédiction d'événements extrêmes sur de longues périodes.
- Le second est dû au fait qu'une loi de probabilité ne donne pas toujours un bon ajustement dans toutes les applications (cf. Bobée & Rasmussen (1995)). Dans ce cas, il est nécessaire d'effectuer un classement des distributions en fonction du comportement de

leurs queues et à partir de considérations physiques ou statistiques, d'établir des critères de discrimination entre les différentes classes dans le cas d'un échantillon de faible taille.

1.1.3 Domaines d'application

La théorie des valeurs extrêmes (TVE) fournit une base mathématique probabiliste rigoureuse sur laquelle il est possible de construire des modèles statistiques permettant de prévoir l'intensité et la fréquence de ces événements extrêmes. La TVE est très répandue ces dernières décennies dans la littérature car elle permet d'apporter des réponses à de nombreux problèmes pratiques et c'est dans ce contexte que la théorie des valeurs extrêmes développée par Fisher et al. (1928) a trouvé toute sa place. Les domaines d'applications utilisant les modèles de la TVE n'ont cessé de se développer ces dernières années. En hydrologie, le domaine d'application historique dû notamment aux travaux de Gumbel & Lieblein (1954), domaine dans lequel la prévision des crues par exemple est particulièrement importante (cf. Davison & Smith (1990) ; Katz et al. (2002)). En climatologie avec l'étude et la prédiction des évènements climatiques extrêmes comme les précipitations extrêmes, les canicules, les chutes de neige, les avalanches (cf. Rootzén & Tajvidi (2001) ; Heneka et al. (2006) ; Brodin & Rootzén (2009)). En météorologie où l'étude de la vitesse du vent, par exemple, permet d'évaluer le degré de résistance des matériaux face à la pression exercée par le vent (au cours d'une tempête par exemple) sur les bâtiments ou les structures de génie civil (cf. Coles & Walshaw (1994) ; Smith (2001) ; Klajnmic (2004) ; Khaliq et al. (2006) ; Davison et al. (2012)). En environnement, avec la modélisation de grands feux de forêts comme des événements extrêmes (cf. Alvarado et al. (1998) voir aussi Ferrez et al. (2011)). Dans les domaines des sciences humaines et sociales, plus particulièrement dans le domaine de la démographie, tout un débat qui a été initié par Gumbel (1937), auquel Fréchet a pris une part active, sur la notion de "durée extrême de la vie humaine" et sur sa mesure. Aarssen & Haan (1994) ; Han (2005) proposèrent des résultats afin de calculer l'âge limite possible de l'être humain. En assurance, avec l'étude de la survenue des sinistres d'intensité exceptionnelle qui peuvent avoir des conséquences négatives sur les résultats et la solvabilité des organismes d'assurance (cf. Barrois (1834) ; McNeil & Saladin (1997) ; Rootzén & Tajvidi (1997)). En finance, elle apporte une réponse immédiate à la remise en cause de l'hypothèse de normalité surtout avec les observations à hautes amplitudes (cf. Embrechts et al. (1997) ; Danielsson & de Vries (1997) ; Embrechts et al. (1999) ; McNeil & Frey (2000) ; Longin (2000) ; Gencay & Selcuk (2004)). Cependant, Bouleau (1991) met en garde sur les mauvaises utilisations de la théorie des valeurs extrêmes.

1.2 Présentation des travaux de recherche

Le thème principal de ce mémoire de thèse repose sur l'étude des valeurs extrêmes et s'articule autour de trois points :

- Estimation des indices de queues des valeurs extrêmes pour certaines classes de processus spatialement dépendants.
- Estimation conditionnelle des indices de queues de distribution et des quantiles extrêmes correspondants dans un cadre de **"plan fixe"** ou **"fixed design"** dans le cadre spatial.
- Estimation conditionnelle des indices de queues de distribution et des quantiles en présence d'une co-variable **aléatoire** dans le cadre spatial.

Notre contribution

Cette thèse s'inscrit dans sa globalité comme étant une contribution à la théorie des valeurs extrêmes dans le cas spatial et ses applications statistiques.

Contribution du Chapitre 3

Le chapitre 3 présente une approche générale. Il nous permet de définir le cadre spatial dans lequel on se place et les notions nécessaires permettant de mener à bien ce travail de thèse. Nous mettons en place notre modèle spatial et la méthode d'identification des sites avant de définir notre estimateur d'indice $\gamma_{\mathbf{n}}$ des valeurs extrêmes en adaptant l'estimateur de Hill (1975) au cadre spatial. Cet estimateur s'avère consistant et asymptotiquement normal sous certaines conditions, et ceci pour des données non spatiales. La contribution majeure de ce chapitre est la mise en place d'un estimateur consistant des valeurs extrêmes pour certaines classes de processus spatialement dépendantes et sous certaines conditions en plus des conditions usuelles dans le cas non spatial. Cet estimateur est asymptotiquement normal quand on l'applique à un processus spatial qui satisfait la condition de mélange forte. Cette normalité asymptotique peut être obtenue pour les autres classes de processus étudiées dans ce chapitre. Le travail effectué étend les travaux de Hsing (1991), Cline (1983), Resnick & Stărică (1993) et Basrak & Tafro (2014) au cadre où le processus est indexé dans un espace de dimension plus grande ou égale à 2.

Contribution du Chapitre 4

Dans la pratique, il est souvent utile de lier la variable d'intérêt Y avec une covariable X. Le chapitre 4 est consacré à l'étude conditionnelle des événements extrêmes où la covariable est **déterministe** en utilisant la classe de processus qui satisfont la condition de mélange fort. Dans cette situation, l'indice de queue $\gamma(.)$ est une fonction inconnue qui dépend de la valeur observée x de la covariable X et sera appelé **indice de queue conditionnelle**. Pour capter la covariable, comme dans Gardes & Girard (2008), la technique de la "fenêtre mobile" ou "windows moving approach" a été utilisée. Plus précisément, nous utilisons une boule $B\left(r,t\right)$ de centre t et de rayon r pour capter la covariable et ensuite utiliser les observations d'intérêt pour lesquelles les co-variables correspondantes sont dans $B\left(r,t\right)$ pour estimer la fonction de queue $\gamma(x)$. A la différence de l'estimateur proposé au chapitre 3, l'estimateur proposé ici tient compte d'une co-variable enregistrée dans une certaine boule. Nous avons également utilisé l'estimateur d'indices de queue pour proposer un estimateur de son *quantile extrême* correspondant. Cet estimateur de queue est consistant pour la classe de champs spatiaux étudiés. Il est également asymptotiquement normal pour un processus satisfaisant la condition de mélange fort. La normalité asymptotique de son quantile extrême correspondant a été établie. Le travail effectué dans ce chapitre étend les travaux de Goegebeur et al. (2014b), Cline (1983), Daouia et al. (2011) et Gardes & Girard (2008) au cadre spatial.

Contribution du Chapitre 5

Le chapitre 5 étend les travaux effectués au chapitre 4 au cas où la covariable considérée est aléatoire. Dans cette situation, nous avons utilisé un estimateur d'indice de queue appartenant à la famille d'estimateurs non paramétrique proposée par Goegebeur et al. (2014b) dans le cadre non spatial. Cette famille généralise l'estimateur de Hill (1975). Il faut noter que cette famille d'estimateurs est conçue à l'aide des méthodes des noyaux. Un estimateur similaire est proposé par Csörgó et al. (1985) dans le cas univarié. Il a été établi, sous certaines conditions en plus des conditions usuelles de la théorie des valeurs, que cet estimateur est asymptotique normal ainsi que son quantile extrême correspondant. Le travail effectué dans ce chapitre étend les travaux de Goegebeur et al. (2014b) et Daouia et al. (2011) au cadre spatial.

1.3 Organisation de la thèse

Le chapitre 2 présente un état de l'art de la théorie des valeurs extrêmes d'une part et de la statistique spatiale d'autre part. La section 2.2 rappelle les principaux résultats et définitions de la théorie des valeurs extrêmes utiles dans nos travaux, plus particulièrement, l'estimation de l'indice de queue lourde et de son quantile extrême correspondant qui constituent la problématique de cette thèse. Dans cette partie, nous partons du résultat principal de la TVE qui montre qu'à l'exception de certaines lois pathologiques, on peut regrouper les lois usuelles en des groupes appelés domaines d'attractions. Nous exposons les critères pour qu'une loi appartienne à l'un de ces groupes. Enfin, nous étudierons des estimateurs de l'indice de queue et de son quantile extrême correspondant. D'autre part, dans la section 2.3 nous rappelons quelques concepts de la statistique spatiale et donnons quelques références sur les méthodes d'estimation en statistique spatiale, particulièrement non-paramétriques. Les chapitres 3, 4, 5 donnent les contributions de la thèse énoncées ci-dessus. Le document se termine par un chapitre de conclusion et de perspectives.

Communications écrites et orales

Travaux et Publications

- Conditional tail index estimation for random fields. (une Note de CRAS correspondante est en révision).
- Tail index estimation for random fields and application to infinite order spatial moving average processes. (A soumettre)
- Fixed-design conditional tail index and quantile estimation for random fields. (soumis)
- On nonparametric conditional tail and extreme quantile estimation with random covariate in a spatial context. (A soumettre)

Séminaires et Conférences

- École CIMPA-SÉNÉGAL : Méthodes statistiques pour l'évaluation des risques extrêmes. "On nonparametric conditional tail and extreme quantile estimation with random covariate in a spatial context". Avril 2016.
- 15è*mes* Forum des jeunes mathématicien-ne-s, Lille. "On tail index estimation for random fields and application to infinite order spatial moving average processes". Novembre 2015.
- 47è*mes* Journées de Statistique de la SFdS, Lille. "On fixed-design conditional tail index and quantile estimation for random fields". Juin 2015.
- Séminaire Inter Universitaire de Théorie Economique, laboratoire LEM, Lille. "Conditional tail index estimation for random fields". Mars 2015.
- Rencontre internationnale des jeunes chercheurs africains en France à l'Institut Henri Poincaré, Paris. "On fixed-design conditional tail index and quantile estimation for random fields". Novembre 2014.

Participation à des conférences, séminaires et journées d'études

- Les séminaires du "Jeudi" d'Économétrie et de Statistique, laboratoire LEM, Lille, 2011-2015.
- Workshop "Kernel methods for big data", Université de Lille 1, Avril 2014.
- Workshop on "Extreme Value Theory, with an emphasis on spatial and temporal aspects", Besançon, November 3-5, 2014.
- 46è*mes* Journées de Statistique de la SFdS, Toulouse. Juin 2014.
- Journées Internationales Analyse Statistique : Théorie et Applications, Oujda (Maroc), Juin 2012.
- Journées de Statistique Spatiale organisée par Liliane Bel, Agroparistech, Paris, Mars 2012.
- European Statistical Meeting : Advances in the Treatment of Missing Data, Bruxelles, Novembre 2011.

Chapitre 2

Concepts fondamentaux

Sommaire

Le but de ce chapitre est de décrire et de présenter les principaux résultats classiques sur la théorie des valeurs extrêmes et sur la statistique spatiale qui permettent ainsi de faciliter la lecture de la thèse. Certains de ces résultats seront utilisés dans les autres chapitres.

2.1 Introduction

Les valeurs extrêmes pour une série d'observation sont des valeurs beaucoup plus fortes (ou beaucoup plus faibles) que celles observées autour de la moyenne. Etudier ces valeurs, revient à analyser la plus grande ou plus petite observation d'un échantillon. Ainsi, la théorie des valeurs extrêmes (notée TVE dans la suite) qui s'intéresse aux observations extrêmes peut être vue comme l'analogue de la théorie statistique classique qui est principalement basée sur l'étude de la moyenne d'un échantillon. Elle est développée pour l'estimation de la probabilité d'occurrence d'événements extrêmes notamment celles de dépassement de seuils élevés c'est-à-dire pour la modélisation statistique des événements rares. La TVE a pour objectif d'étudier et de caractériser le comportement des valeurs extrêmes d'un échantillon de variables aléatoires (les températures extrêmes, les crues de mer par exemple). On souhaite estimer des petites probabilités ou des quantités dont la probabilité d'observation est très faible, c'est-à-dire proche de zéro. Ces quantités sont appelées quantiles extrêmes et se situent dans les queues de distributions des lois de probabilité. On parle de quantile extrême lorsque la probabilité d'observation du quantile converge vers zéro quand la taille de l'échantillon tend vers l'infini. Le principal objectif de ce travail est de donner les modèles probabilistes de la TVE qui permettent d'extrapoler le comportement de la queue de distribution des données à partir des plus grandes observations. La TVE est principalement basée sur le théorème de Fisher et al. (1928).

Dans ce chapitre, tout d'abord, nous donnons un rappel général sur la TVE dans un cadre univarié. Ensuite, nous exposerons les domaines d'attractions selon le signe de l'indice de queue de la fonction de distribution, puis les fonctions à variations régulières. Enfin, nous exposons plus en détails la caractérisation du domaine d'attraction de Fréchet.

2.2 Généralités sur la theorie des valeurs extrêmes

Lois des valeurs extrêmes

L'estimation des quantiles requiert l'estimation du paramètre de la queue de distribution. Le théorème fondamental de la TVE (connu sous le nom de Théorème de Fisher et al. (1928)) donne les lois limites possibles du maximum de l'échantillon et permet ainsi d'avoir une certaine connaissance sur la queue de distribution.

Formulation.

Concrètement, on considère n variables aléatoires réelles $X_1, X_2, \cdots, X_n$ indépendantes et identiquement distribuées (i.i.d) de fonction de répartition $F(x) = \mathbb{P}(X_1 \leq x)$. Soit $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ la statistique d'ordre associée à ces n variables aléatoires. Pour tout $i = 1, \cdots, n$, la variable aléatoire $X_{(i)}$ s'appelle la $(n-i+1)ème$ statistique d'ordre de l'échantillon. Il existe deux statistiques d'ordre qui sont particulièrement intéressantes pour l'étude des événements extrêmes. Ce sont les statistiques d'ordre extrême qui correspondent à la plus petite statistique d'ordre $X_{(1)}$ (ou statistique du minimum).

$$X_{(1)} = \min(X_1; \cdots; X_n)$$

et à la plus grande statistique d'ordre $X_{(n)}$ (ou statistique du maximum).

$$X_{(n)} = \max(X_1; \cdots; X_n).$$

Le comportement asymptotique du maximum $X_{(n)}$ (resp. minimum $X_{(1)}$) permet de rendre compte sur la fin de la distribution. Ces deux statistiques d'ordre sont liées l'une à l'autre à l'aide de la relation

$$X_{(1)} = -\max(-X_1, \cdots, -X_n).$$

Étant donné que les variables aléatoires sont i.i.d alors la fonction de répartition F_n du maximum est donné par

$$\begin{aligned} F_n(x) &= \mathbb{P}\left(X_{(n)} \leq x\right) = \mathbb{P}(X_1 \leq x, \cdots, X_n \leq x) \\ &= \prod_{i=1}^{n} \mathbb{P}(X_i \leq x) = [F(x)]^n . \end{aligned} \tag{2.1}$$

La fonction de répartition F étant souvent inconnue, il n'est généralement pas possible de déterminer la distribution du maximum à partir de ce résultat puisque les valeurs extrêmes se trouvent, à droite et à la fin du support de la distribution.

Le principal but de la TVE est l'estimation d'un événement rare, ce qui revient à estimer la probabilité d'observer un phénomène extrême ayant une valeur plus grande que le maximum de l'échantillon. Autrement dit, on s'intéresse à la probabilité que le quantile extrême soit plus grand que le maximum de l'échantillon. Il est donc essentiel de définir explicitement la notion de quantile.

Définition 2.2.1. *Le quantile d'ordre α de la fonction de distribution F est défini par*

$$q(\alpha) := F^{-1}(\alpha) = \inf\{y : F(y) \geq \alpha\}, \quad \textit{avec} \;\; \alpha \in]0;1[,$$

où F^{-1} est l'inverse généralisée de F. Par convention $\inf \varnothing = \infty$. Notons que l'inverse généralisé d'une fonction coïncide avec l'inverse classique lorsque la fonction est continue.

Ainsi, un quantile sera dit extrême si l'on remplace son ordre α par une suite $\alpha_n \to 0$ quand $n \to \infty$. Concrètement, un quantile extrême d'ordre $1 - \alpha_n$ de la fonction de distribution F est définie par :

$$q(\alpha_n) := \overline{F}^{-1}(\alpha_n) \;\; \textit{avec} \;\; \alpha_n \to 0 \;\; \textit{quand} \;\; n \to \infty,$$

où $\overline{F} := 1 - F$ est la fonction de survie de F.

Il faut dire que le fait que l'ordre $\alpha_n \to 0$ quand $n \to \infty$ indique que l'information la plus importante pour estimer des quantiles extrêmes est contenue dans la queue de distribution.

Dans la Figure 1.2, il est plus intéressant de considérer la différence entre la hauteur réelle et celle prévue de la marée (les surcotes), que les hauteurs des marées. L'étude statistique de ces différences de hauteur a pour but d'apporter des réponses aux questions suivantes :

- trouver h (une hauteur de marée) tel que la probabilité pour que la différence entre la hauteur réelle et la hauteur prévue de la marée soit supérieure à h soit α avec $\alpha \in]0;1[$.
- trouver h tel que la probabilité pour que la plus haute différence entre la hauteur réelle et la hauteur prévue de la marée annuelle soit supérieure à h soit α avec $\alpha \in]0;1[$ (typiquement de l'ordre de 10^{-3} ou 10^{-4}).

Si l'on considère F comme étant la fonction de répartition de la loi de ces différences de hauteur, alors dans la première question, h est le **quantile d'ordre** $1 - \alpha \in]0;1[$ de la loi de ces différences de hauteur :

$$h = q(1 - \alpha) := \inf\{x \in \mathbb{R} : F(x) \geq 1 - \alpha\}.$$

Pour simplifier, on notera $q(\alpha)$ le quantile d'ordre $1-\alpha$. Il faut noter que la loi des surcotes et donc les quantiles associés sont inconnus.

Par soucis de simplification, on supposera F continue. Nous voulons déterminer, à partir de l'échantillon, le quantile extrême d'ordre $\alpha_n \to 0$ lorsque $n \to \infty$. Si l'on remplace x par le quantile extrême $q(\alpha_n)$ dans l'equation (2.1) et si la taille de l'échantillon tend vers l'infini, nous avons

$$\begin{aligned} F_n(q(\alpha_n)) &= [F(q(\alpha_n))]^n \\ &= (1 - \alpha_n)^n \\ &= \exp(n \log(1 - \alpha_n)) \\ &= \exp(-n\alpha_n(1 + o(1))) \quad \text{quand} \;\; n \to \infty. \end{aligned}$$

Cette probabilité dépend donc du comportement asymptotique de $n\alpha_n$. Estimer les quantiles extrêmes revient à étudier la limite de $n\alpha_n$ lorsque n tend vers l'infini. On distingue trois situations selon la vitesse de convergence de α_n vers 0 :

- **Premier cas** : Si $n\alpha_n \to 0$, alors $\mathbb{P}\left(X_{(n)} \leq q(\alpha_n)\right) \longrightarrow 0$.

Dans ce cas, α_n converge vers 0 plus lentement que $1/n$, i.e. α_n est très petit par rapport au nombre d'observations, une estimation empirique est impossible. Un modèle semi-paramétrique est nécessaire. Ceci signifie que l'on va modéliser la queue de distribution par une loi de valeurs extrêmes, tandis que les probabilités non extrêmes seront estimées empiriquement. L'estimation du quantile extrême repose donc sur une interpolation à l'intérieur de l'échantillon. Autrement dit, le quantile à estimer se trouve presque sûrement dans l'échantillon disponible et presque sûrement plus petit que l'observation maximale $X_{(n)}$. Dans une telle situation, l'estimateur du quantile extrême est donné par $X_{(n-\lfloor n\alpha\rfloor+1)}$.

- **Deuxième cas** : Si $n\alpha_n \to c$ avec $c \in [1;\infty[$, alors $\mathbb{P}\left(X_{(n)} \leq q(\alpha_n)\right) \longrightarrow e^{-c}$.

Dans cette situation, étant donné que pour n suffisamment grand, $n\alpha_n$ converge vers $c > 0$ et donc l'estimation repose sur une valeur extrême de l'échantillon c'est-à-dire contrairement au cas précédent, le quantile extrême se trouve avec une faible probabilité à l'intérieur de l'échantillon. Par conséquent, l'estimateur proposé est le même que celui de la première situation.

- **Troisième cas** : Si $n\alpha_n \to c$ avec $c \in [0;1[$, alors $\mathbb{P}\left(X_{(n)} \leq q(\alpha_n)\right) \longrightarrow e^{-c}$.

Dans cette situation, α_n converge vers 0 avec la même vitesse ou plus rapide que $1/n$ et donc le quantile extrême est plus grand que $X_{(n)}$ avec probabilité positive $e^{-c} \geq e^{-1}$. Par conséquent, son estimation est plus difficile car il nécessite une estimation en dehors de l'échantillon. Dans ce cas, l'estimateur du quantile ne peut être obtenu en inversant simplement la fonction de répartition empirique.

L'utilisation des lois de valeurs extrêmes repose sur les propriétés des statistiques d'ordre et sur des méthodes d'extrapolation. Plus précisément, elle repose sur des convergences en loi du maximum de variables aléatoires convenablement renormalisées. Les lois limites possibles sont connues. Elles sont appelées les lois de valeurs extrêmes. L'estimation des quantiles extrêmes se déroulera en deux étapes :

- Identification de la loi des valeurs extrêmes associée aux données.
- Estimation des paramètres de renormalisation.

Identification de la loi des valeurs extrêmes

Si l'on reprend l'equation (2.1), le résultat central de la théorie des valeurs extrêmes concerne la distribution asymptotique $\mathcal{H}$ du maximum. Étant donné que les variables aléatoires sont i.i.d, on obtient alors la distribution asymptotique du maximum en faisant tendre n vers l'infini,

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} [F(x)]^n = \begin{cases} 1 & \text{si } x \geq x_F \\ 0 & \text{si } x < x_F, \end{cases} \tag{2.2}$$

où $x_F = \sup\{x : F(x) < 1\}$ est *le point terminal à droite (right-end point)* de la distribution F (voir Embrechts et al. (1997) pour plus de précisions sur x_F). On s'intéresse par conséquent à une loi non dégénérée pour le maximum, la théorie des valeurs extrêmes permet de donner une réponse à cette problématique. Les premiers résultats sur la caractérisation du comportement asymptotique des maxima $X_{(n)}$ convenablement normalisés ont été obtenus par Fisher et al. (1928). Le théorème suivant explicite ces résultats.

Theorem 2.2.2. *(Embrechts et al., 1997, Th. 3.2.3)*

Soit $(X_n)_{n\geq 1}$ *une suite de variables aléatoires i.i.d. de même loi* F *avec* $F(x) = \mathbb{P}(X_1 \leq x)$. *S'il existe deux suites normalisantes réelles* $(a_n > 0, n \geq 1)$ *et* $(b_n \in \mathbb{R}, n \geq 1)$ *et une loi non-dégénérée* $\mathcal{H}$ *telles que*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_{(n)} - b_n}{a_n} \leq x\right) = \mathcal{H}(x),\ \forall x \in \mathbb{R},$$

alors $\mathcal{H}$ *appartient à l'un des trois types de distribution suivants :*

$$\textit{Type I (Gumbel) :}\quad \Lambda(x) = \exp\left(-e^{-x}\right),\quad \forall x \in \mathbb{R},$$

$$\textit{Type II (Fréchet) :}\quad \Phi_\alpha(x) = \begin{cases} 0 & \textit{si } x \leq 0 \\ \exp\left(-x^{-\alpha}\right) & \textit{si } x > 0 \end{cases}$$

$$\textit{Type III (Weibull) :}\quad \Psi_\alpha(x) = \begin{cases} \exp\left(-(-x)^{\alpha}\right) & \textit{si } x \leq 0 \\ 1 & \textit{si } x > 0 \end{cases}$$

L'ensemble des lois limites s'obtient donc en considérant les lois de $cW + d$, où W suit une loi de Weibull, de Gumbel ou de Fréchet.

Définition 2.2.3. *Les trois lois de probabilité ci-dessus sont appelées distributions de valeurs extrêmes. Une distribution des valeurs extrêmes du maximum (resp. minimum) est donc une distribution* $\mathcal{H}$ *pouvant s'obtenir comme limite d'une suite convenablement normalisée de maxima (resp. minima) de variables aleatoires i.i.d.*

Le résultat du Théorème 2.2.2 est d'un intérêt capital, car si l'ensemble des distributions est "grand", l'ensemble des distributions de valeurs extrêmes est lui très restreint. Il présente ainsi dans l'esprit des similarités avec le Théorème Central Limite (TCL) qui définit quant à lui le comportement limite de la somme ou de la moyenne d'une suite de variables aléatoires i.i.d. Stuart Coles (Coles, 2001, page 47) fait le parallèle suivant entre le théorème de limite centrale et celui de Fisher et al. (1928) :

> *"The remarkable feature of this result is that the three types of extreme value distributions are the only possible limits for the distribution of the* $\frac{X_{(n)}-b_n}{a_n}$, *regardless of the distribution F for the population. It is in this sense that the theorem provides an extreme value analog of the central limit theorem"*

Il faut noter que les suites $\{a_n \in \mathbb{R}_+^*, n \geq 1\}$ et $\{b_n \in \mathbb{R}, n \geq 1\}$ dependent de la loi de X. C'est ainsi que plusieurs auteurs se sont intéressés à l'estimation de ces paramètres. C'est le cas de (Embrechts et al., 1997, p.145) qui utilisent les suites de renormalisation théoriques associées à la loi normale standard pour illustrer la convergence en loi de la suite de variables aléatoires $\left\{a_n^{-1}\left(X_{(n)} - b_n\right),\ n \geq 1\right\}$ vers une limite non dégénérée Λ. Ainsi les suites de renormalisation théoriques associées à la loi normale standard dans (Embrechts et al., 1997, p.145) sont alors définies par :

$$a_n = (2\log n)^{-1/2} \quad \text{et} \quad b_n = (2\log n)^{1/2} - \frac{\log\log n + \log 4\pi}{2\,(\log n)^{1/2}} + o\left((\log n)^{-1/2}\right)$$

où $u_n = o(v_n)$ signifie que $\lim_{n\to\infty} u_n/v_n = 0$.

Lois de valeurs extrêmes généralisées (Generalized Extreme Value GEV)

Le Théorème 2.2.2 donne un résultat très intéressant. Quelle que soit la loi limite de la variable parente, la loi limite des extrêmes a toujours la même forme. Même si les comportements de ces lois sont complètement différents, elles peuvent être combinées en une seule paramétrisation contenant un unique paramètre[1] noté γ qui contrôle la "lourdeur" de la queue de distribution appelé indice des valeurs extrêmes (cf. Fisher et al. (1928), Gnedenko (1943), Von Mises (1936), Jenkinson (1955)).

$$\mathcal{H}_\gamma(x) = \begin{cases} \exp\left(-(1+\gamma x)^{-\frac{1}{\gamma}}\right), & \gamma \neq 0 \quad 1+\gamma x > 0 \\ \exp(-\exp(-x)), & \gamma = 0 \quad -\infty \leq x \leq +\infty. \end{cases} \tag{2.3}$$

où $\mathcal{H}$ est une fonction de répartition non-dégénérée. Le résultat suivant est appelé Théorème d'unification des trois types.

Theorem 2.2.4. *S'il existe des constantes $a_n > 0$ et $b_n \in \mathbb{R}$ et une distribution limite non dégénérée $\mathcal{H}$ telle que*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_{(n)} - b_n}{a_n} \leq x\right) = \mathcal{H}(x),\ \forall x \in \mathbb{R},$$

alors $\mathcal{H}$ est de la forme

$$\mathcal{H}_\gamma(x) = \begin{cases} \exp\left(-\,(1+\gamma x)_+^{-1/\gamma}\right) & si\ \gamma \neq 0 \\ \exp\left(-\exp\left(-x\right)\right) & si\ \gamma = 0 \end{cases} \tag{2.4}$$

définie sur un support $S = \{x : 1+\gamma x > 0\}$ que nous notons $\mathcal{GEV}_\gamma$. La fonction densité est donnée par

$$g(x) = (1+\gamma x)_+^{-\frac{1+\gamma}{\gamma}} \exp\left(-\,(1+\gamma x)_+^{-1/\gamma}\right),$$

avec $x_+ = \max(0, x)$.

Cas particuliers

- red Si $\gamma < 0$ alors $\mathcal{H}_\gamma(x)$ est du même type que $\Psi_\gamma(x) = \exp\left(-\,(-x)^{-1/\gamma}\right)$. *(Distribution négative de Weibull de paramètre $-\frac{1}{\gamma}$).*
- Si $\gamma = 0$ alors $\mathcal{H}_\gamma(x)$ est du même type que $\Lambda(x) = \exp\left(-\exp\left(-x\right)\right)$. *(Distribution de Gumbel).*
- Si $\gamma > 0$ alors $\mathcal{H}_\gamma(x)$ est du même type que $\Phi_\gamma(x) = \exp\left(-\,(1+\gamma x)_+^{-1/\gamma}\right)$. *(Distribution de Fréchet de parametre $\frac{1}{\gamma}$).*

Ainsi, à travers ses différentes valeurs possibles, le paramètre de forme γ donne une grande flexibilité à la distribution GEV de sorte à prendre en compte les trois types de comportement asymptotiques représentés par les distributions extrêmes ci-dessus.

---

[1] La connaissance de γ permet à elle seule de caractériser à un changement d'échelle près le comportement du maximum normalisé.

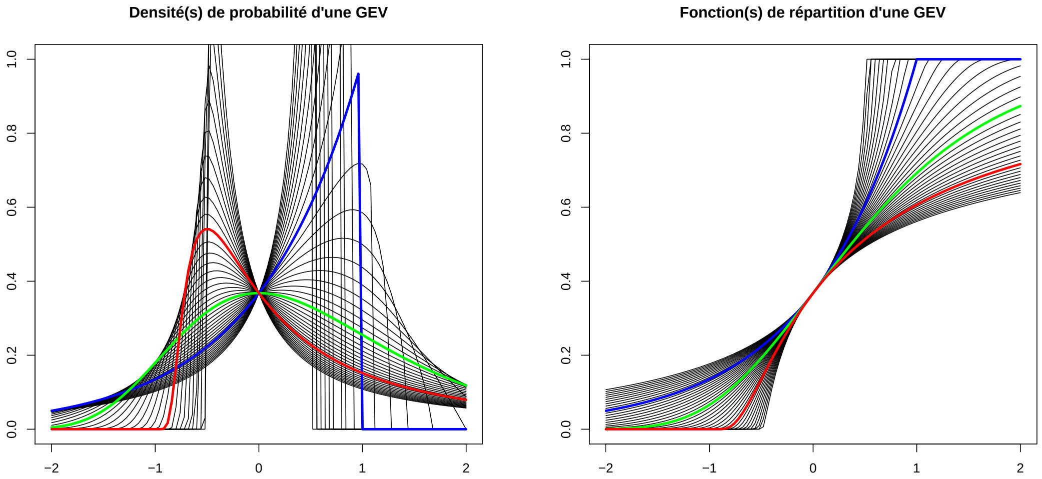

Figure 2.1 – A gauche, densités de la distribution GEV et à droite, fonctions de répartition d'une distribution GEV pour $\gamma \in [-2, 2]$. En particulier, vert : $\gamma = 0$, rouge : $\gamma = 1$ et bleue : $\gamma = -1$

La loi des excès

L'approche basée sur les distributions GEV peut être réductrice du fait que l'utilisation d'un seul maxima conduit à une perte d'information continue dans les autres grandes valeurs de l'échantillon. D'où l'utilisation d'une nouvelle approche, la méthode des excès au-delà d'un seuil (ou *Peak Over Threshold, P.O.T.*). Elle a été développée par Pickands (1975) et repose sur le comportement des valeurs observées au-delà d'un seuil déterministe. L'idée de base de cette approche consiste à choisir un seuil suffisamment élevé et à étudier les excès au-delà de ce seuil. Autrement dit, elle consiste à observer non pas le maximum ou les plus grandes valeurs mais toutes les valeurs des réalisations qui excèdent un certain seuil élevé.

Plus précisément, soit $u < x_F$ et F_u la fonction de répartition des excès définie par

$$F_u(y) = \mathbb{P}\left(X - u < y \mid X > u\right) = \frac{F(u+y) - F(u)}{1 - F(u)} \text{ pour } 0 \leq y \leq x_F - u.$$

Ce qui équivaut à

$$F_u(y) = \mathbb{P}\left(X < x \mid X > u\right) = \frac{F(x) - F(u)}{1 - F(u)} \text{ pour } x \geq u.$$

Cette méthode permet de déterminer par quelle loi de probabilité on peut approcher la distribution conditionnelle des excès F_u lorsque le seuil tend vers le point terminal x_F. On a alors le résultat suivant :

Theorem 2.2.5. *Balkema & de Haan (1974) et Pickands (1975)*

Si F appartient à l'un des trois domaines d'attraction de la loi des valeurs extrêmes (Gumbel, Fréchet ou Weibull), alors il existe une fonction $\sigma(u)$ strictement positive et $\gamma \in \mathbb{R}$ tels que

$$\lim_{u \uparrow x_F} \sup_{0 \leq y \leq x_F - u} \left| F_u(y) - \mathcal{G}_{\gamma,\sigma}(y) \right| = 0, \tag{2.5}$$

où $\mathcal{G}_{\gamma,\sigma}$ est la fonction de répartition de la loi de Paréto Généralisée définie par

$$\mathcal{G}_{\gamma,\sigma}(y) = \begin{cases} 1 - (1 + \gamma y/\sigma)^{-\frac{1}{\gamma}} & si\ \gamma \neq 0,\ \sigma > 0 \\ 1 - \exp(-y/\sigma) & si\ \gamma = 0,\ \sigma > 0, \end{cases} \tag{2.6}$$

où $y \in [0, x_F]$ *si* $\gamma \geq 0$ *et* $y \in [0, \min(-\sigma/\gamma; x_F)]$ *si* $\gamma < 0$.

La distribution conditionnelle converge donc vers la fonction $\mathcal{G}_{\gamma,\sigma}(y)$ qui correspond à la fonction de répartition de la loi Pareto généralisée, notée GPD (Generalized Pareto Distribution). Noter que le cas $\gamma = 0$ correspond à la distribution exponentielle de moyenne σ et que le cas $\gamma = -1$ correspond à la loi uniforme sur $[0, \sigma]$. La figure ci-dessus illustre le comportement de différentes densités et distributions d'une GPD standard pour $\gamma \in [-2, 2]$.

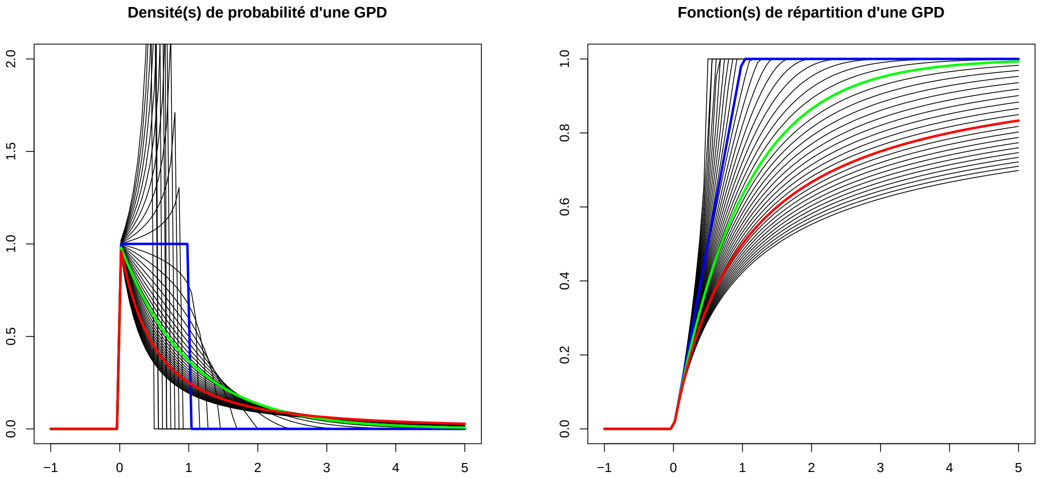

FIGURE 2.2 – A gauche, densités d'une distribution GPD et à droite, fonctions de répartition d'une distribution GPD pour $\gamma \in [-2, 2]$. En particulier, $\gamma = 0$, $\gamma = 1$ et $\gamma = -1$

Les trois domaines d'attraction

Dans le cas univarié, le problème consiste à répondre à la question suivante : étant donné une loi $\mathcal{H}$ de type extrême (donc appartenant à l'une des trois familles de Fréchet, de Gumbel ou de Weibull) :
Quels sont les critères vérifiés par une distribution quelconque F *pour que la loi du maximum normalisé de la suite de variables aléatoires i.i.d. de loi* F *converge vers* $\mathcal{H}$ *?*

Différentes caractérisations de ces trois domaines d'attraction ont été proposées dans Haan (1985) ; Haan & Ferreira (2006) ; Embrechts et al. (1997) ; Resnick (1987). Selon le signe de γ, on distingue trois cas de domaines d'attraction que l'on note $\mathcal{DA}$:

• Si $\gamma < 0$, F appartient au $\mathcal{DA}$ de **Weibull** que l'on note $F \in \mathcal{DA}(\Psi_\gamma)$. Dans ce domaine, nous trouvons des distributions à support fini, ce qui implique que le support du maximum est borné à droite. Les distributions de type Weibull sont souvent utilisées pour décrire la résistance mécanique d'un matériau ou encore le temps de fonctionnement d'un appareil électronique ou mécanique. Elles peuvent être utilisées pour estimer le potentiel éolien d'un site, de modéliser la probabilité qu'un vent souffle à telle vitesse sur ce site.

• Si $\gamma = 0$, F appartient au $\mathcal{DA}$ de **Gumbel** que l'on note $F \in \mathcal{DA}(\Lambda)$. Nous trouvons des distributions avec des queues épaisses (log-normale) et des distributions à queues légères (exponentielle, normale), c'est-à-dire, les lois dont la fonction de survie décroit vers zéro à une vitesse exponentielle. Les distributions de type Gumbel peuvent, par exemple, servir à prévoir le niveau des crues d'un fleuve si l'on possède le relevé des débits sur dix ans.

Elles peuvent aussi servir à prédire la probabilité d'un événement critique, comme un tremblement de terre par exemple.

• Si $\gamma > 0$, F appartient au $\mathcal{DA}$ de **Fréchet** que l'on note $F \in \mathcal{DA}(\Phi_\gamma)$, nous y trouvons des distributions qui ont des queues épaisses. Autrement dit, ce domaine contient toutes les lois dont la fonction de survie décroit comme une fonction puissance. Les distributions de type Fréchet sont souvent utilisées en climatologie, en hydrologie, en finance (gestion de la qualité, gestion du risque), en réassurance.

Dans le tableau 2.1, nous avons une liste non exhaustive de lois et leur domaine d'attraction comme dans les tableaux 3.4.2 à 3.4.4 dans Embrechts et al. (1997).

Domaines d'attraction	**Weibull** $\gamma < 0$	**Gumbel** $\gamma = 0$	**Fréchet** $\gamma > 0$
Lois	Beta Uniforme ReverseBurr	Gumbel Weibull Gamma Normale Logistique Log-normale Exponentielle Benktander-type-I Benktander-type-II	Burr Pareto Cauchy Student Log-gamma Log-logistique α Stable avec $\alpha < 2$ Fréchet

TABLE 2.1 – Exemples de lois classées selon leurs domaines d'attraction.

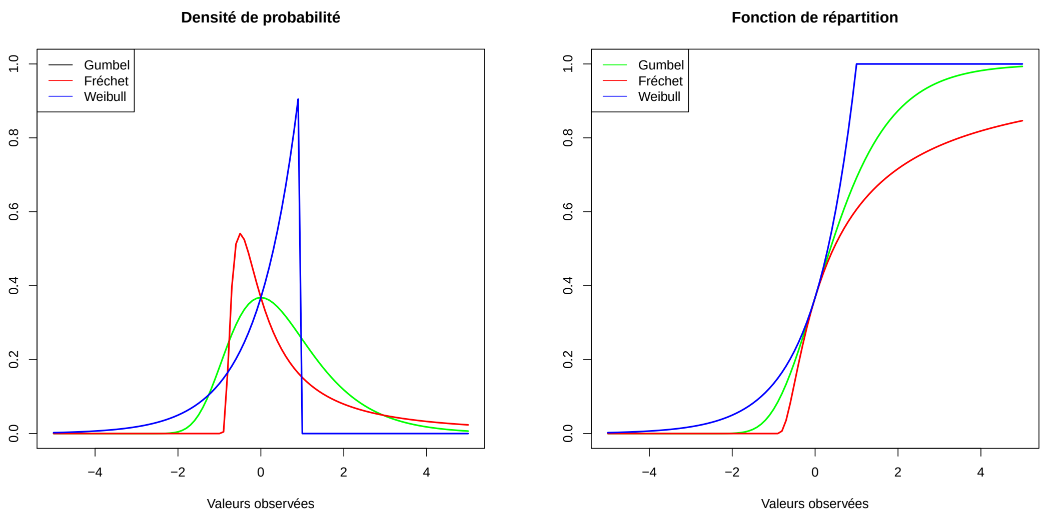

FIGURE 2.3 – Densités des lois de Gumbel ($\gamma = 0$), Fréchet ($\gamma = 1$) et Weibull ($\gamma = -1$) à gauche et leurs fonctions de répartitions associées à droite.

A titre d'illustration, nous montrons l'exemple d'une variable aléatoire X suivant une loi exponentielle de paramètre 1. Le maximum $X_{(n)}$ de cette distribution converge vers la distribution de Gumbel. La fonction de répartition de la loi exponentielle est donnée par

$F(x) = 1 - \exp(-x)$. Et que

$$\begin{aligned}\mathbb{P}\left(\frac{X_{(n)} - b_n}{a_n} \leq x\right) &= \mathbb{P}\left(X_{(n)} \leq a_n x + b_n\right) \\ &= F^n\left(a_n x + b_n\right) \\ &= \left[1 - \exp(-a_n x - b_n)\right]^n .\end{aligned}$$

Si nous posons $a_n = 1$ et $b_n = \log n$, nous avons

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left(\frac{X_{(n)} - b_n}{a_n} \leq x\right) &= \lim_{n\to\infty}\left(1 - \frac{1}{n}\exp(-x)\right)^n \\ &= \exp(-\exp(-x)) \\ \text{puisque} \quad \exp(x) = 1 + x + \frac{x^2}{2!} + \cdots &= \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n .\end{aligned}$$

Pour décrire plus en détail les domaines d'attraction, il est nécessaire de décrire précisément le comportement de $\overline{F}$ quand x converge vers x_F. Cette description fait appel aux classes de fonctions à variations régulières.

Notions de fonctions à variations régulières

La notion de fonctions à variations régulières joue un rôle essentiel dans la théorie des valeurs extrêmes. Nous donnons ici un aperçu des principaux résultats sur ces fonctions. Pour plus de détails, se référer à Bingham et al. (1987), Resnick (1987), Bingham et al. (1989) ou encore Geluk & Haan (1987) parmi tant d'autres.

Définition 2.2.6. *Soit r une fonction mesurable au sens de Lebesgue. On dit que $r : \mathbb{R}_+ \longrightarrow \mathbb{R}$ est à variations régulières à l'infini si et seulement si, il existe une fonction g telle que : $\forall t > 0$, $\forall x > 0$, on a :*

$$\lim_{t\to\infty} \frac{r(tx)}{r(t)} = g(x). \tag{2.7}$$

On a

$$\begin{aligned}g(xy) = \lim_{t\to\infty}\frac{r(txy)}{r(t)} &= \lim_{t\to\infty}\frac{r(txy)}{r(tx)} \times \frac{r(tx)}{r(t)} \\ &= \lim_{t\to\infty}\frac{r(txy)}{r(tx)} \lim_{t\to\infty}\frac{r(tx)}{r(t)} = g(y)g(x).\end{aligned}$$

Autrement dit, la fonction g vérifie : $\forall x,\, y > 0$, $g(xy) = g(x)g(y)$, la version multiplicative de l'équation fonctionnelle de Cauchy dont la solution générale est de la forme $g(x) = x^{-\delta}$, $\delta \in \mathbb{R}$. On a alors la définition suivante.

Définition 2.2.7. *On dit que F est une fonction à variations régulières d'indice α et on note $F \in \mathcal{RV}_\alpha$ si $\forall x > 0$;*

$$\lim_{t\to\infty} F(tx)/F(t) = x^\alpha .$$

En particulier,

- *Si $\alpha = 0$, F est dite fonction à **variations lentes** à l'infini et on note $F \in \mathcal{RV}_0$.*
- *Si $\alpha = +\infty$, F est dite fonction à **variations rapides** à l'infini.*

La proposition suivante donne les propriétés principales des fonctions à variations régulières voir Resnick (1987).

Proposition 2.2.8. *a. Si $F \in \mathcal{RV}_\alpha$ alors il existe une fonction à variations lentes $L(x)$ telle que*

$$F(x) = x^\alpha L(x).$$

b. La notion de variations régulières est stable par somme, multiplication, composition, intégration et différentiation. Plus concrètement,

- si $F \in \mathcal{RV}_\alpha$ et $G \in \mathcal{RV}_\beta$ alors $F + G \in \mathcal{RV}_{\max(\alpha,\beta)}$ et $F \circ G \in \mathcal{RV}_{\alpha\beta}$,

- si $F \in \mathcal{RV}_\alpha$ alors $F^k \in \mathcal{RV}_{k\alpha}$,

- si $f \in \mathcal{RV}_\alpha$ avec $\alpha \in \mathbb{R}$ alors pour toute primitive F de f on a : $F \in \mathcal{RV}_{\alpha+1}$. Et plus spécifiquement :

$\ast$ si $\alpha \geq -1$ alors $f \in \mathcal{RV}_\alpha$ implique que $F \in \mathcal{RV}_{\alpha+1}$ et on a :

$$\lim_{x\to\infty} \frac{xf(x)}{F(x) - F(0)} = \alpha + 1.$$

$\ast$ Si $\alpha < -1$ (ou si $\alpha = -1$ et $\lim_{t\to\infty} F(t) - F(x) < \infty$) alors $f \in \mathcal{RV}_\alpha$ implique que $\lim_{t\to\infty} F(t) - F(x) < \infty$ et $F \in \mathcal{RV}_{\alpha+1}$. De plus on a :

$$\lim_{x\to\infty} \frac{xf(x)}{F(x) - F(0)} = -\alpha - 1.$$

- si $F \in \mathcal{RV}_\alpha$ et si sa dérivée f existe et est monotone pour tout $x > 0$ alors $f \in \mathcal{RV}_{\alpha-1}$.

Le lemme suivant donne des informations sur l'inverse d'une fonction à variations régulières. La preuve est donnée dans l'ouvrage de (Bingham et al., 1987, Th. 1.5.12) ou dans le livre de (Resnick, 1987, Pro. 2.6).

Lemma 2.2.9. *- Si F est à variations régulières d'indice $\alpha > 0$, alors F^{-1} est à variations régulières d'indice $1/\alpha$.*

- Si F est à variations régulières d'indice $\alpha < 0$, alors F^{-1} est à variations régulières d'indice $-1/\alpha$.

La première partie de la Proposition 2.2.8 montre que l'étude des fonctions à variations régulières à l'infini se ramène à celle des fonctions à variations lentes. Parmi les fonctions à variations lentes, on peut citer :

- les fonctions possédant une limite strictement positive à l'infini ;
- les fonctions de la forme $L(x) = |\log x|^\beta$, $\beta \in \mathbb{R}$;
- les fonctions L telles que

$$\exists M > 0,\ \forall x \geq M,\ L(x) = c + dx^{-\beta}(1 + o(1)).$$

L'ensemble de ces fonctions L est appelé la **classe de Hall**. On peut caractériser les fonctions à variations lentes de manière plus précise à l'aide du théorème de Karamata.

Theorem 2.2.10. *(Représentation de Karamata, Resnick (1987))*
Toute fonction à variations lentes à l'infini L peut s'écrire sous la forme :

$$L(x) = c(x) \exp\left\{\int_1^x \frac{\varepsilon(u)}{u} du\right\}, \tag{2.8}$$

où $c > 0$ et ε sont deux fonctions mesurables telles que

$$\lim_{x\to\infty} c(x) = c > 0 \quad et \quad \lim_{x\to\infty} \varepsilon(x) = 0.$$

De plus, si la fonction c est constante, la fonction L est dite **normalisée**. Dans ce cas, l'équation (2.8) est dérivable de dérivée L' définie pour tout $x > 0$ par

$$L'(x) = \frac{\varepsilon(x)L(x)}{x}.$$

En particulier, on a

$$\lim_{x\to\infty} \frac{xL'(x)}{L(x)} = 0.$$

Proposition 2.2.11. *Pour toute fonction à variations lentes L à l'infini, on a :*

$$\lim_{x\to\infty} \frac{\log(L(x))}{\log(x)} = 0. \tag{2.9}$$

Pour plus d'information sur la théorie des fonctions à variations régulières, on renvoie le lecteur à Bingham et al. (1987) ; Geluk & Haan (1987).

Sachant une distribution F, *comment déterminer le domaine d'attraction auquel elle appartient et trouver les constantes de normalisation associées ?* La réponse à cette question est relativement complexe et dépend de la forme de la fonction F. Les résultats suivants donnent les critères les plus utilisés.

Caractérisation du domaine d'attraction de Fréchet $\mathcal{DA}(\Phi_\gamma)$

La distribution de type Fréchet Φ_γ de paramètre γ est telle que :
$\forall x > 0,\ \overline{\Phi}(x) = \exp\left(-x^{1/\gamma}\right),\ \forall\gamma > 0$, donc $\overline{\Phi}(x) \sim x^{-1/\gamma}$ au voisinage de l'infini. Le résultat ci-dessous énoncé dans Gnedenko (1943) donne une caractérisation du $\mathcal{DA}(\Phi_\gamma)$. Une démonstration de ce résultat est donnée dans Resnick (1987).

Theorem 2.2.12. *Caractérisation du $\mathcal{DA}(\Phi_\gamma)$*

Une distribution F ayant pour point terminal x_F appartient au domaine d'attraction de Fréchet $\mathcal{DA}(\Phi_\gamma)$, $\gamma > 0$, ($\overline{F} = 1 - F \in \mathcal{RV}$) si et seulement si $x_F = +\infty$ et $\overline{F}$ est à variations régulières d'indice $-1/\gamma$ à l'infini i.e.

$$\lim_{t\to+\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma}. \tag{2.10}$$

Dans ce cas, les suites de normalisation (a_n) et (b_n) sont données par :

$$a_n = F^{-1}\left(1 - \frac{1}{n}\right) \quad \text{et} \quad b_n = 0,\ \forall n > 0.$$

Remarque 2.2.13. *Si $F \in \mathcal{DA}(\Phi_\gamma)$, alors il existe une fonction L à variations lentes telle que :*

$$\overline{F}(x) = x^{-1/\gamma}L(x). \tag{2.11}$$

*Les distributions appartenant au $\mathcal{DA}(\Phi_\gamma)$ sont dites **à queue épaisse** puisque leurs moments d'ordre supérieur à γ n'existe pas i.e. $\mathbf{E}\left[X_+\right]^\theta = +\infty$, $\forall\theta > \gamma$, où $X_+ = \max\{0, X\}$.*

Caractérisations générales

On a l'ensemble des lois limites qui s'obtient en considérant les lois de $cW+d$, où W suit une loi de Weibull, de Gumbel ou de Fréchet. De plus, ces trois lois peuvent être combinées en une seule paramétrisation contenant un unique paramètre. En conséquence, une propriété globale de caractérisation peut être développée et qui englobe toutes les trois lois limites pour les maxima. Autrement dit, on cherche à caractériser $\mathcal{H}_\gamma$ sans se préoccuper du signe de γ qui spécifie la loi limite approchée par $\mathcal{H}_\gamma$. Pour cela, soit U une fonction définie par

$$U = F^{-1}\left(1-t^{-1}\right),\ \forall t>1,$$

où F^{-1} est l'inverse généralisé de F. Les calculs dans la démonstration du Théorème 2.2 (voir Haan & Ferreira (2006)) suggèrent que si $F \in \mathcal{DA}\left(\mathcal{H}_\gamma\right)$, alors on a, à un changement d'échelle près, l'existence d'une fonction $a(\cdot)$ telle que pour tout $x>0$,

$$\lim_{x\to\infty} \frac{U(tx)-U(x)}{a(x)} = \begin{cases} \frac{t^\gamma-1}{\gamma} & \text{si } \gamma \neq 0 \\ \log t & \text{si } \gamma = 0. \end{cases}$$

En particulier, si $x, y>0$ et $y \neq 1$, on a

$$\lim_{z\to\infty} \frac{U(zx)-U(z)}{U(zy)-U(z)} = \lim_{z\to\infty} \frac{U(zx)-U(z)}{a(z)} . \frac{a(z)}{U(zy)-U(z)} = \begin{cases} \frac{x^\gamma-1}{y^\gamma-1} & \text{si } \gamma \neq 0 \\ \frac{\log x}{\log y} & \text{si } \gamma = 0. \end{cases} \tag{2.12}$$

La proposition suivante assure que cette condition est suffisante pour que $F \in \mathcal{DA}\left(\mathcal{H}_\gamma\right)$

Proposition 2.2.14. *Pour tout $\gamma \in \mathbb{R}$, il y a équivalence entre $F \in \mathcal{DA}\left(\mathcal{H}_\gamma\right)$ et quelques soient $x>0$, $y>0$ et $y \neq 1$,*

$$\lim_{z\to\infty} \frac{U(zx)-U(z)}{U(zy)-U(z)} = \begin{cases} \frac{x^\gamma-1}{y^\gamma-1} & si\ \gamma \neq 0 \\ \frac{\log x}{\log y} & si\ \gamma = 0. \end{cases} \tag{2.13}$$

Pour la preuve de ce résultat nous vous référons au livre de (Embrechts et al., 1997, Th. 3.4.5). Il faut noter que ce résultat est souvent utilisé pour construire l'estimateur de Pickand de γ.

La convergence en loi du maximum renormalisé exige la recherche de suites $(a_n>0;\ n\geq 0)$ et $(b_n \in \mathbb{R},\ n \geq 1)$ telles que

$$\mathbb{P}\left(\frac{X_{(n)}-b_n}{a_n} \leq x\right) = F^n\left(a_n x + b_n\right) = \left(1-\overline{F}(a_n x+b_n)\right)^n$$

converge vers une limite non triviale.

Proposition 2.2.15. *On a $F \in \mathcal{DA}\left(\mathcal{H}_\gamma\right)$ si et seulement si*

$$n\overline{F}(a_n x+b_n) \underset{n\longrightarrow\infty}{\longrightarrow} -\log \mathcal{H}_\gamma(x), \tag{2.14}$$

pour certaines suites $a_n>0$ et $b_n \in \mathbb{R}$ pour tout $n \geq 1$. On a alors la convergence en loi de $\left\{a_n^{-1}\left(X_{(n)}-b_n\right),\ n\geq 1\right\}$ vers une variable aléatoire de fonction de répartition $\mathcal{H}_\gamma$.

Nous avons un résultat plus général.

Proposition 2.2.16. *Pour tout $\gamma \in \mathbb{R}$, il y a équivalence entre $F \in \mathcal{DA}(\mathcal{H}_\gamma)$ et il existe une fonction mesurable f telle que pour tout $1 + \gamma x > 0$, on a*

$$\lim_{x \uparrow x_F} \frac{\overline{F}(u + xf(u))}{\overline{F}(u)} = \begin{cases} (1+\gamma x)^{-1/\gamma} & si\ \gamma \neq 0 \\ e^{-x} & si\ \gamma = 0. \end{cases} \tag{2.15}$$

Dans les sections précédentes, on a vu que les lois limites approchées par $\mathcal{H}_\gamma$ sont toutes indexées par un paramètre γ. Il est donc nécessaire d'estimer ce paramètre lorsqu'on s'intéresse à la queue de distribution.

Estimation des paramètres

Nous avons vu que la loi asymptotique du maximum normalisé de variables aléatoires continues est de trois types possibles (Fréchet, Weibull et Gumbel) indexée par un paramètre noté γ. De même, la loi asymptotique des excès au delà d'un seuil donné, est une loi de Pareto généralisée indexée par le paramètre γ et un deuxième paramètre σ. Ces paramètres, appelés respectivement indice de valeurs extrêmes et paramètre d'échelle, apportent une information sur la forme de la queue de distribution de F (notamment, selon que $\gamma > 0$, $\gamma < 0$ ou $\gamma = 0$). C'est pourquoi il est nécessaire d'estimer ces indices. De nombreux estimateurs de ce paramètre et de son quantile extrême correspondant ont été proposés dans la littérature. Parmi ces estimateurs, on peut citer l'estimateur de Hill (1975), l'estimateur de Pickands (1975), l'estimateur des moments de Dekkers et al. (1989), les estimateurs basés sur les méthodes des moments, des moments pondérés et du maximum de vraisemblance Hosking & Wallis (1987), l'estimateur du rapport des moments de Danielsson et al. (1996), l'estimateur de Peng (1998), l'estimateur basé sur le QQ-plot, l'estimateur basé sur le graphique de la moyenne des excès Beirlant et al. (1996), l'estimateur construit par des méthodes de régression de Beirlant et al. (2002). Notons aussi qu'il existe d'autres estimateurs du paramètre γ (cf. Csörgó et al. (1985), Kratz & Resnick (1996), Schultze & Steinebach (1996), Beirlant et al. (2006), Haan & Ferreira (2006), Lô (1986), Embrechts et al. (1997), Beirlant et al. (1999), Beirlant et al. (2002)) parmi tant d'autres. Cependant toutes ces méthodes présentent de grandes différences lorsqu'on les applique à des données simulées ou réelles même si théoriquement elles partagent les mêmes propriétés de consistance et de normalité asymptotique. Le fait qu'en pratique ces méthodes semblent très différentes est souvent associé à la taille de l'échantillon qui n'est pas souvent très grande dans la région d'intérêt. Dans ce cas, il n'existe pas de solution miracle pour estimer le paramètre γ. Parmi tous ces estimateurs, les plus utilisés sont l'estimateur de Hill (1975), l'estimateur de Pickands (1975) et l'estimateur des moments de Dekkers et al. (1989).

Dans la suite, nous donnons une étude un peu plus détaillée de l'estimateur de Hill (1975) puisque c'est cet estimateur qui est étudié dans cette thèse. Cette étude nous permet de mettre en évidence, dans le cas univarié, les propriétés necessaires pour l'étude de la consistance et de la normalité asymptotique de l'estimateur de Hill (1975). Notons que nous ferons appel à ces propriétés dans les chapitres suivants.

Estimateur de Hill (1975) et son quantile extrême correspondant

Dans tout ce paragraphe, on suppose que $\gamma > 0$ cela sous entend que la fonction de répartition doit appartenir au domaine d'attraction de Fréchet. La définition de la fonction de répartition F de type Paréto peut être écrite comme suit :

$$\frac{1 - F(xu)}{1 - F(u)} \longrightarrow x^{-\frac{1}{\gamma}}, \text{ quand } u \to \infty \text{ et } \forall x > 1.$$

Cette définition est générale et elle est obtenue en approximant la quantité de gauche par :

$$\mathbb{P}\left(\frac{X}{u} > x \,|\, X > u\right) \sim x^{-\frac{1}{\gamma}}, \text{ pour } u \text{ assez grand et } x > 1.$$

Ou encore à l'aide d'une fonction à variation lente

$$\mathbb{P}\left(\frac{X}{u} > x \,|\, X > u\right) = x^{-\frac{1}{\gamma}} L_F(x), \ \text{ avec } \ L_F(tx)/L_F(t) \underset{t\to\infty}{\longrightarrow} 1.$$

La construction de cet estimateur est basée sur la méthode du maximum de vraisemblance où on se sert des statistiques d'ordre supérieures à un certain seuil u, permettant ainsi de garder que les observations les plus grandes, de façon à ce que qu'elles suivent approximativement une distribution de Paréto. À cette effet, soit $T_i(u) = \frac{X_i}{u}$ les excès relatifs au-delà de u avec $X_i > u$ tel que i soit le $j^{\text{ème}}$ excès et $j = 1, \ldots, N_u$ où N_u est la taille de l'échantillon au-delà de u. Ainsi, log vraisemblance s'écrit :

$$\log L\left(T_1(u), \ldots, T_{N_u}(u)\right) = -N_u \log \gamma - (1 + \frac{1}{\gamma}) \sum_{j=1}^{N_u} \log T_j(u).$$

On en déduit alors,

$$\frac{d \log L}{d\gamma} = -\frac{N_u}{\gamma} - \frac{1}{\gamma^2} \sum_{j=1}^{N_u} \log T_j(u) = 0 \Leftrightarrow \gamma = \frac{1}{N_u} \sum_{j=1}^{N_u} \log T_j(u). \tag{2.16}$$

Fixons le niveau d'excès u, la valeur de la statistique d'ordre $X_{(n-k_n)}$ où k_n remplace N_u est telle que $1 \leq N_u = k_n = o(n)$. On obtient ainsi l'estimateur de Hill (1975) défini par :

$$\hat{\gamma}_{k_n,n}^{Hill} = \frac{1}{k_n} \sum_{i=1}^{k_n} \log X_{(n-i+1)} - \log X_{(n-k_n)} = \frac{1}{k_n} \sum_{i=1}^{k_n} i \left(\log X_{(n-i+1)} - \log X_{(n-i)} \right). \tag{2.17}$$

Cependant, il existe des variantes de l'estimateur de Hill (1975) qui estiment γ (cf. Haan & Ferreira (2006), Beirlant et al. (2006)). La consistance de cet estimateur a été largement étudiée (cf. Mason (1982) pour la consistance faible, Deheuvels et al. (1988) pour la consistance forte) tandis que la normalité asymptotique a été examinée par Davis & Resnick (1984), Haeusler & Teugels (1985), Csörgó et al. (1985), Hsing (1991), Smith (1987) parmi tant d'autres. Si l'étude de la convergence en probabilité ne dépend que de k_n et de la taille de l'échantillon n qui doit être telle que $1 \leq k_n \leq n$ et $k_n/n \to 0$, l'étude de la normalité asymptotique quand à elle réside, en plus des conditions de consistance, sur des conditions de régularité sur la fonction de répartition F. Ces conditions de régularités sont appelées **les conditions de variations régulières de second ordre**, ce qui nous amène à la définition ci-dessous.

Définition 2.2.17. *(Conditions de variations régulières du second ordre)*
On dit que $F \in \mathcal{DA}(\Phi_\gamma)$ avec $\gamma > 0$ est à variations régulières du second ordre à l'infini si elle satisfait l'une des conditions ci-dessous.

- *Il existe un réel $\rho \leq 0$ et une fonction non nulle de signe constant, d'indice ρ que l'on note $A(t) \longrightarrow 0$ quand $t \to \infty$ tels que pour tout $x > 0$,*

$$\lim_{t\to\infty} \frac{x^{\frac{1}{\gamma}} \overline{F}(tx)/\overline{F}(t) \ - \ 1}{A(t)} = \begin{cases} \frac{x^\rho - 1}{\rho} & si \ \rho < 0 \\ \log x & si \ \rho = 0, \end{cases} \tag{2.18}$$

ou encore à l'aide de la fonction logarithme

$$\lim_{t\to\infty} \frac{\frac{1}{\gamma} \log x + \log \overline{F}(tx) - \log \overline{F}(t)}{A(t)} = \begin{cases} \frac{x^\rho - 1}{\rho} & si \ \rho < 0 \\ \log x & si \ \rho = 0. \end{cases} \tag{2.19}$$

- *Il existe un réel* $\rho \leq 0$ *et une fonction non nulle de signe constant, d'indice* ρ *que l'on note* $B(t) \longrightarrow 0$ *quand* $t \to \infty$ *tels que pour tout* $x > 0$,

$$\lim_{t\to\infty} \frac{x^{-\gamma}U(tx)/U(t) \; - \; 1}{B(t)} = \begin{cases} \frac{x^{\rho}-1}{\rho} & si \ \rho < 0 \\ \log x & si \ \rho = 0, \end{cases} \tag{2.20}$$

ou encore à l'aide de la fonction logarithme

$$\lim_{t\to\infty} \frac{\log U(tx) - \log U(t) - \gamma \log x}{B(t)} = \begin{cases} \frac{x^{\rho}-1}{\rho} & si \ \rho < 0 \\ \log x & si \ \rho = 0, \end{cases} \tag{2.21}$$

avec $U(t) = F^{-1}\left(1 - t^{-1}\right)$, $\forall t > 1$.

Les relations (2.18)-(2.21) impliquent que :

- la fonction $|A|$ est à variations régulières à l'infini d'indice γ/ρ et on note $|A| \in \mathcal{RV}_{\gamma/\rho}$,
- la fonction $|B|$ est à variations régulières à l'infini d'indice ρ et on note $|B| \in \mathcal{RV}_{\rho}$.

Le paramètre ρ contrôle la vitesse de convergence de $L_F(tx)/L_F(t)$ vers 1 quand $t \to \infty$ ou de manière équivalente $L_U(tx)/L_U(t)$ vers 1 quand $t \to \infty$, pour plus de détails, voir Geluk & Haan (1987). Plus ρ est proche de 0, plus la convergence sera lente et donc l'estimation de γ sera difficile. Beaucoup d'auteurs se sont intéressés à l'estimation de ce paramètre ρ afin de réduire le biais asymptotique qui apparaît lors de l'estimation de γ (cf. Alves et al. (2003a,b), Gomes & Martins (2001), Gomes et al. (2000), Peng (1998)). Ces conditions permettent d'énoncer les propriétés asymptotiques de l'estimateur de Hill (1975).

Theorem 2.2.18. *(Propriétés asymptotiques de* $\hat{\gamma}^{Hill}_{k_n,n}$*)*
Soit $(k_n)_{n\geq 1}$ *une suite d'entiers telle que* $1 \leq k_n \leq n$ *avec* $k_n \to \infty$ *et* $k_n/n \to 0$ *quand* $n \to \infty$. *On obtient les résultats suivants :*

- $$\hat{\gamma}^{Hill}_{k_n,n} \xrightarrow{\mathbb{P}} \gamma.$$
- *Si de plus,* $k_n/\log\log n \to \infty$ *quand* $n \to \infty$, *alors*
$$\hat{\gamma}^{Hill}_{k_n,n} \xrightarrow{a.s.} \gamma.$$
- *Si la condition* (2.18) *est satisfaite avec* $\sqrt{k_n}A\left(\frac{n}{k_n}\right) \longrightarrow \lambda \in \mathbb{R}$ *quand* $n \to \infty$, *alors*
$$\sqrt{k_n}\left(\hat{\gamma}^{Hill}_{k_n,n} - \gamma\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right).$$

Remarque 2.2.19. *Si* $\sqrt{k_n}A\left(\frac{n}{k_n}\right) \longrightarrow 0$ *quand* $n \to \infty$, *alors on a*

$$\sqrt{k_n}\left(\frac{\hat{\gamma}^{Hill}_{k_n,n}}{\gamma} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Ce qui permet d'obtenir un intervalle de confiance de niveau $(1-\alpha)$, $0 < \alpha < 1$ *pour l'estimateur* $\hat{\gamma}^{Hill}_{k_n,n}$ *dont les bornes sont données par*

$$\hat{\gamma}^{Hill}_{k_n,n}\left(1 + k_n^{1/2}Z_{1-\alpha/2}\right)^{-1} \quad et \ \hat{\gamma}^{Hill}_{k_n,n}\left(1 - k_n^{1/2}Z_{1-\alpha/2}\right)^{-1},$$

où Z_α *est le quantile d'ordre* α *de la loi normale centre réduite.*

Dans l'analyse des extrêmes, l'estimation du paramètre γ n'est pas l'objectif première, c'est plutôt l'estimation de petites quantités de probabilités. Ces quantités sont appelées quantiles extrêmes car l'ordre de ces quantiles tend vers zéro lorsque la taille n de l'échantillon tend vers l'infini. Puisque l'estimateur de Hill (1975) peut être interprété comme un estimateur de la droite affine du "Paréto Quantile Plot", il apparaît donc naturel d'estimer son quantile extrême correspondant. Ainsi, en extrapolant le long de la droite du "Paréto quantile Plot" d'équation

$$y = \log X_{(n-k_n)} + \hat{\gamma}_{k_n,n}^{Hill}\left(x + \log \frac{k_n+1}{n+1}\right),$$

on obtient l'estimateur du quantile extrême $q(\alpha_n) = F^{-1}(1-\alpha_n)$. Plus particulièrement, si l'on pose $x = \log \alpha_n^{-1}$, alors on obtient l'estimateur le plus célèbre introduit par Weissman (1978) défini comme suit :

$$\hat{q}_{k_n,n}^{Hill}(\alpha_n) = X_{(n-k_n)}\left(\frac{k_n+1}{(n+1)\alpha_n}\right)^{\hat{\gamma}_{k_n,n}^{Hill}}. \tag{2.22}$$

À l'aide du Lemme 2.2.9, cette écriture de la fonction de survie $\overline{F}$ implique :

$$\begin{aligned}\overline{F}^{-1}(1-\alpha_n) &= U(\alpha_n^{-1}) = \alpha_n^{-\gamma} L_U\left(\alpha_n^{-1}\right), \quad \text{avec} \quad \alpha_n \leq 1/n\\ \overline{F}^{-1}(1-\beta_n) &= U(\beta_n^{-1}) = \beta_n^{-\gamma} L_U\left(\beta_n^{-1}\right), \quad \text{avec} \quad \beta_n \geq 1/n\end{aligned}$$

où L_U est une finction à variations lentes à l'infini. Alors, pour β_n suffisamment petit lorsque $n \to \infty$, on a

$$\overline{F}^{-1}(\alpha_n) \simeq \overline{F}^{-1}(\beta_n)\left(\frac{\beta_n}{\alpha_n}\right)^{\gamma}. \tag{2.23}$$

En remplaçant $\overline{F}^{-1}(\beta_n)$ par son estimateur naturel $X_{(n-\lfloor n\beta_n \rfloor)}$ et γ par un estimateur quelconque d'indices positifs et plus particulièrement par l'estimateur $\hat{\gamma}_{k_n,n}^{Hill}$ de Hill (1975), on obtient ainsi l'estimateur général de type Weissman (1978) défini par :

$$\hat{q}_{k_n}^{W}(\alpha_n) = X_{(n-\lfloor n\beta_n \rfloor)}\left(\frac{\beta_n}{\alpha_n}\right)^{\hat{\gamma}_{k_n,n}^{Hill}}. \tag{2.24}$$

dont les propriétés asymptotiques sont discutées sous certaines conditions de la fonction de distribution F, la séquence intermédiaire k_n et l'ordre α_n du quantile permettant ainsi de construire un intervalle de confiance (cf Embrechts et al. (1997), Matthys & Beirlant (2003), Ferreira et al. (2003) et Markovich (2005)).

2.3 Généralités sur la statistique spatiale

De nos jours, avec le développement des nouvelles technologies, de nombreuses données sont recueillies dans différentes localisations avec leurs positions géographiques. C'est le cas en géologie, séismologie, géographie, épidémiologie, agronomie, sciences de l'environnement et de la terre, météorologie, économie, dans le traitement d'images, dans l'industrie pétrolière et bien d'autres. L'étude de ces types de données ne peut se faire sans tenir compte de leurs positions géographiques. Ce qui conduit à un déploiement des méthodes paramétriques et non paramétriques usuelles (estimation de densité, prédiction, régression, tests, etc.) pour l'analyse de telles données spatiales. L'analyse spatiale est un terme général pour décrire une technique qui utilise les informations spatiales afin de mieux comprendre les processus générant l'attribut observé sur un domaine bien défini. La statistique spatiale étudie ainsi des phénomènes observés sur un ensemble spatial $\mathcal{S} \in \mathbb{R}^N$, $N \geq 2$. Il existe une forte dynamique autour des outils et méthodes statistiques prenant en compte l'information spatiale contenue dans certaines données, conduisant ainsi à la caractérisation des données géoréférencées.

Les trois types de données géoréférencées.

Il existe trois grands types de données spatiales : les données géostatistiques, les données latticielles et les données ponctuelles. Plus spécifiquement, on note $s \in \mathcal{S}$ la localisation d'un site de mesure et on considère un phénomène $Z = \left\{ Z_s,\ s \in \mathcal{S} \in \mathbb{R}^N \right\}$: température, densité de populations, cumul de précipitations sur un réseau météorologique etc. Z est une famille de variables aléatoires indexée par $\mathcal{S}$, un ensemble spatial. Pour plus d'information sur ces types de données, nous nous référons aux travaux de Ripley (1981), Cressie (1992), Guyon (1995), Gaetan & Guyon (2008) et Cressie & Wikle (2011).

• **Les données géostatistiques** : Si les données sont mesurées en tout point d'un domaine continu, on se place dans le cadre de la géostatistique. Dans ce cas, la localisation s est généralement de nature géographique. On peut citer par exemple les stations de mesure du cumul de précipitations sur une région $\mathcal{S} \subset \mathbb{R}^2$, points de prélèvement d'un échantillon du sous-sol terrestre pour la mesure de densité en minerai de fer d'une zone de prospection $\mathcal{S} \subset \mathbb{R}^2$ etc. Son domaine historique d'utilisation est l'estimation des gisements miniers, mais son domaine d'application actuel est beaucoup plus large notamment en géologie, biologie, télécommunication, santé, climatologie... Ci-dessous (Figure 2.4), deux exemples d'illustration de données géostatistiques issues du package **geoR** du logiciel R.

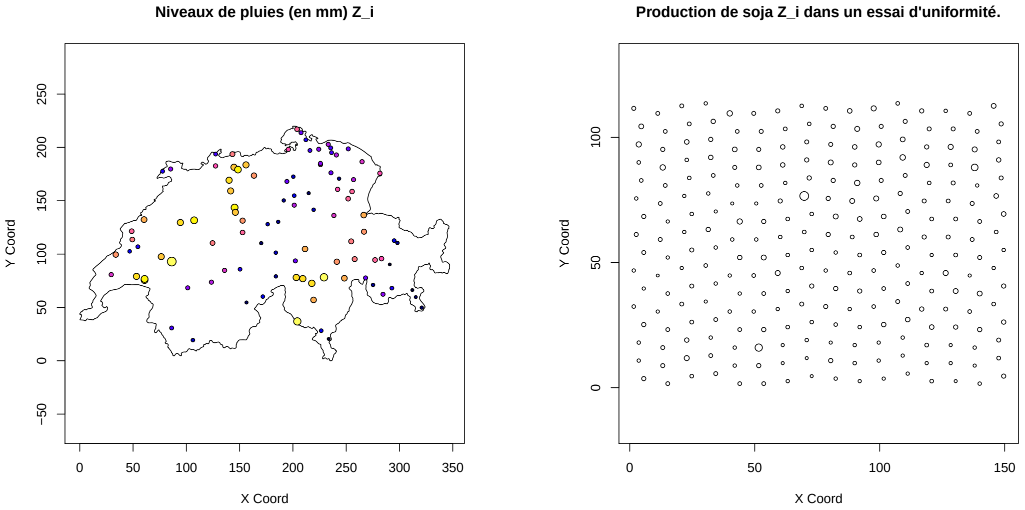

FIGURE 2.4 – A gauche, Cumul de pluies dans 100 stations météo suisses, choisies au hasard parmi 367 stations, le jour du passage du nuage de Tchernobyl (Jeu sic.100 du package geoR de R). A droite, données sur la production de soja dans un essai d'uniformité mesurées dans des parcelles de taille 5 x 5 mètres et d'autres propriétés du sol mesurées en des points donnés par les coordonnées géographiques (Jeu soja98 du package geoR de R).

• **Les données latticielles** : Si les données sont liées à un réseau discret, on parle de données latticielles, plus précisément, elles constituent un ensemble partiellement ordonné dans lequel chaque site a des voisins. Dans ce cas, l'ensemble spatial $\mathcal{S}$ est discret et fixé.

Les sites représentent en général des unités géographiques, repérées par un graphe de voisinage. On peut citer par exemple des zones liées par un réseau de transport, des sites d'arbres fruitiers dans un verger $\mathcal{S} \subset \mathbb{Z}^2$, des cantons et le nombre de personnes affectées par une maladie sur chaque canton, des pixels d'une image, etc. Ci-dessous (Figure 2.5), un exemple d'illustration de données latticielles issues du package **spdep** du logiciel R.

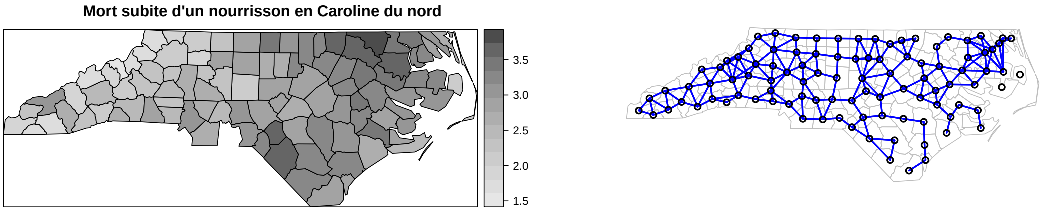

Figure 2.5 – À gauche, le nombre de cas enregistrés dans 100 Comtés de la Caroline du nord entre 1974-1978. À droite, un graphe de voisinages correspondant (Jeu sids du package spdep de R).

• **Les données ponctuelles** : Les processus ponctuels surviennent lorsque les sites où ont lieu les observations sont aléatoires. Ce type de processus sont des extensions des processus ponctuels indexés dans $\mathbb{R}$ au cadre $\mathbb{R}^N$. La question que l'on se pose dans l'étude de ces types de données est : la localisation des sites est-elle homogénéité ? Est-elle plutôt régulière ? Ou présente-t-elle des agrégats ? Par exemple, on retrouve ce type de données dans l'étude de la répartition spatiale d'une certaine espèce. Ci-dessous (Figure 2.6), des exemples d'illustration de données latticielles issues des package **spatstat** du logiciel R.

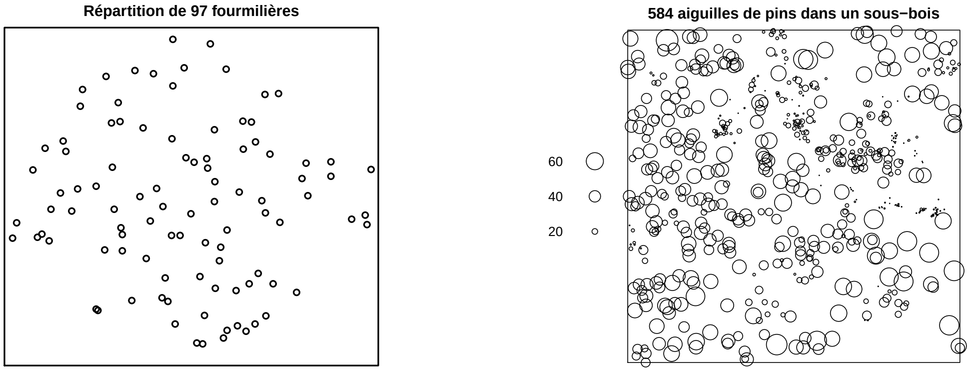

Figure 2.6 – A gauche, 97 fourmilières (jeu ants du package spatstat). A droite, positions et tailles de 584 aiguilles de pins dans un sous-bois (jeu longleaf du package spatstat de R).

L'analyse de ces trois catégories de données utilisées en statistique spatiale nécessite ainsi des méthodes de statistique spatiale variées adaptées à chaque type de données. Ces méthodes vont de la visualisation, de l'exploration de la dépendance spatiale à la modélisation (par exemple la prédiction d'un phénomène en des sites non-observés).

Un des problèmes les plus importants en statistique spatiale est la prédiction. Dans ce cadre, à la différence de la statistique "classique", la statistique spatiale a la capacité de prédire (en tenant compte de la dépendance des données d'un phénomène) en un lieu non observé à l'aide des observations effectuées dans d'autres sites. La prédiction spatiale repose ainsi sur l'utilisation des observations disponibles aux sites voisins, nécessitant ainsi l'étude de la structure de dépendance entre les lieux voisins. Cette structure de dépendance est le point commun entre les trois types de données géo-réferencées. Il faut noter également qu'il existe des différences majeures entre les approches spatiale et temporelle. La plus grande différence est la non-existence des notions de passé et futur et d'un ordre naturel dans le cadre spatial contrairement au cas temporel. Ainsi les modèles des séries temporelles ne peuvent pas être directement appliqués aux données spatiales. Pour en savoir plus sur les différences et les similarités existantes entre les séries spatiales et chronologiques, nous référons au lecteur l'article de Tjøstheim (1987).

L'estimation spatiale

Soient $\mathcal{S}$ un ensemble spatial et $Z = \{Z_s,\ s \in \mathcal{S}\}$ un champ aléatoire spatial indexé dans $\mathcal{S}$ où s est considéré comme un site. L'ensemble $\mathcal{S}$ joue un rôle essentiel dans l'analyse spatiale car une fois défini, il est le premier indicateur du type de données dont on dispose (géostatistiques, latticielles ou ponctuelles). Les premiers outils de la statistique spatiale furent paramétriques. Il existe une vaste littérature sur les modèles paramétriques en statistique spatiale allant de l'estimation de la variabilité spatiale et des tests associés à la prédiction (cf. Cressie (1992), Guyon (1995), Stein (1999), Cressie & Wikle (2011), etc). Les méthodes non paramétriques spatiales sont plus récentes mais il existe dynamique assez forte dans ce cadre. On peut citer entre autres Journel (1983), Tran (1990), Tran & Yakowitz (1993), Carbon et al. (1996), Biau & Cadre (2004), Menezes et al. (2010), Ould-Abdi et al. (2010), Dabo-Niang & Thiam (2010), El Machkouri (2011), García-Soidán & Menezes (2012), Dabo-Niang & Yao (2013), Ternynck (2014), Basrak & Tafro (2014),...dans le cadre de l'estimation d'une densité, de la régression ou la prédiction spatiale non-paramétrique.

Dans le cadre des données géostatistiques, principalement considérées dans cette thèse, on considère $\mathcal{S}$ comme un sous-ensemble fixé de $\mathbb{R}^N$ avec $N > 1$. On dénote par $(s_1, s_2, \cdots, s_n)$ des sites localisés dans $\mathcal{S} \subseteq \mathbb{R}^N$ et on suppose qu'on dispose d'observations $Z_{s_1}, ..., Z_{s_n}$. Dans les estimations paramétriques et non-paramétriques, lorsque qu'on s'intéresse aux propriétés asymptotiques des estimateurs, on étudie généralement leurs comportements lorsque la taille de l'échantillon converge vers l'infini. Deux structures sont généralement utilisées pour l'étude asymptotique (cf. Gaetan & Guyon (2008)) dans un cadre spatial :

- L'asymptotique extensive (*increasing domain asymptotics*) : c'est la situation où le nombre d'observations disponible croît avec le domaine d'observation $\mathcal{S}$.
- L'asymptotique intensive (*infill asymptotics*) : c'est la situation où les observations augmentent dans un domaine $\mathcal{S}$ fixé et borné (on a de plus en plus de sites d'observation tandis que la région reste fixe).

Dans nos travaux, l'asymptotique extensive est considérée.

Dans la suite, nous commençons par donner quelques définitions, puis nous donnons quelques méthodes d'estimation paramétriques et non-paramétriques utiles dans le cadre de nos travaux.

Modélisation paramétrique de données géostatistiques

Définition 2.3.1. ***Processus stationnaires au second ordre***
Z est un processus stationnaire au second ordre sur $\mathcal{S}$ si Z est de moyenne constante et de covariance invariante par translation :

$$\forall s, t \in \mathcal{S} : \ \mathbf{E}\left[Z_s\right] = m \ \ et \ \ c(s,t) = \operatorname{Cov}\left(Z_s, Z_t\right) = c\left(s-t\right)$$

avec $c(.)$ une fonction de covariance stationnaire.

Définition 2.3.2. ***Isotropie***
Un champ Z, stationnaire au second ordre, est dit isotropique, si la covariance entre Z_s et Z_t ne dépend que de la distance $\|s-t\|$ entre s et t :

$$\forall s,t \in \mathcal{S}, \quad c(s-t) = r\left(\|s-t\|\right),$$

avec $r(\cdot)$ la fonction de covariance isotropique.

Une façon de s'affranchir de l'hypothèse de stationnarité au second ordre est de considérer le processus des $h-$accroissements de Z, à h fixé, on parle de **processus intrinsèque**.

Définition 2.3.3. ***Processus intrinsèque***
On dit que Z est un processus intrinsèque si, pour tout $h \in \mathcal{S}$:

- *Z est de moyenne constante, $\mathbf{E}\left[Z_s\right] = m$, $\forall s \in \mathcal{S}$,*
- *le processus $I^{(h)} = \left\{I_s^{(h)} = Z_{s+h} - Z_s,\ s \in \mathcal{S}\right\}$ est stationnaire au second ordre :*

$$\mathbf{E}\left[Z_{s+h} - Z_s\right] = 0, \quad et \quad \mathrm{Var}\left(Z_{s+h} - Z_s\right) = 2\gamma(h)$$

*où $2\gamma(\cdot)$ est appelé le **variogramme** de Z et $\gamma(\cdot)$ est le **semi-variogramme**.*

Le variogramme est donc un outil géostatistique pour la modélisation de la structure de covariance. Voici quelques modèles fréquemment utilisés pour un variogramme isotropique :

- le modèle exponential

$$\gamma(h;a,\sigma^2) = \sigma^2\left(1 - \exp\left(-\frac{\|h\|}{a}\right)\right),$$

- le modèle sphérique

$$\gamma(h;a,\sigma^2) = \begin{cases} \sigma^2\left(\frac{3}{2}\frac{\|h\|}{a} - \frac{1}{2}\frac{\|h\|^3}{a^3}\right) & \text{si } \|h\| \le a \\ \sigma^2 & \text{sinon} \end{cases},$$

- le modèle gaussien

$$\gamma(h;\sigma^2,v) = \sigma^2\left\{1 - \exp\left(-\left(\frac{\|h\|}{a}\right)^2\right)\right\},$$

- le modèle de Matern

$$\gamma(h;a,\sigma^2,v) = \sigma^2\left\{1 - 2^{1-v}\left(\frac{\|h\|}{a}\right)^v \mathcal{K}_v\left(\frac{\|h\|}{a}\right)\right\}/\Gamma(v),$$

$\mathcal{K}_v(\cdot)$ est la fonction de Bessel modifiée de 2-ème espèce de paramètre $v > 0$.

Si Z est stationnaire de covariance $c(.)$, alors Z est intrinsèque de variogramme : $2\gamma(h) = 2\left(c(0) - c(h)\right)$. Notons que la réciproque est fausse.

Anisotropie et variogramme : L'anisotropie est une notion contraire à celle d'isotropie. Si $\overrightarrow{e}$ est une direction de $\mathbb{R}^N$, $\|\overrightarrow{e}\| = 1$ alors le *variogramme directionnel* d'un champ intrinsèque dans la direction $\overrightarrow{e}$ est défini par :

$$2\gamma(h) = \mathrm{Var}\left(Z_{s+h\overrightarrow{e}} - Z_s\right), \ \forall h \in \mathbb{R}.$$

On dit qu'il y a anisotropie du processus si deux variogrammes directionnels au moins diffèrent.

On distingue deux types d'anisotropie :

- *l'anisotropie géométrique* qui est associée à la déformation linéaire d'un modèle isotropique
- *l'anisotropie de support* qui correspond à une stratification du variogramme sur plusieurs sous-espaces de $\mathbb{R}^N$.

L'estimation du variogramme : L'estimateur empirique du variogramme est :

$$2\hat{\gamma}_n(h) = \frac{1}{\mathrm{Card}(N(h))} \sum_{s_i, s_j \in N(h)} \left(Z_{s_i} - Z_{s_j}\right)^2, \ h \in \mathbb{R}^N$$

où $N(h) = \{(s_i, s_j) : h - \Delta \leq s_i - s_j \leq h + \Delta;\ i, j = 1, \cdots, n\}$ est une classe approximante à tolérance Δ de $h \in \mathbb{R}^N$.
En pratique, on choisit Δ de telle sorte que chaque classe comporte au moins 30 couples de points. Si Z est stationnaire au second ordre, la covariance est estimée empiriquement par :

$$\hat{c}_n(h) = \frac{1}{\mathrm{Card}(N(h))} \sum_{s_i, s_j \in N(h)} \left(Z_{s_i} - \bar{Z}\right), \ h \in \mathbb{R}^N,$$

où $\bar{Z} = n^{-1} \sum\limits_{i=1}^{n} Z_{s_i}$.
L'avantage d'utiliser $\hat{\gamma}_n(h)$ en comparaison de $\hat{c}_n(h)$ est de ne pas demander au préalable l'estimation de la moyenne m. Cependant, le variogramme empirique n'est pas admissible, il ne vérifie pas les propriétés d'une fonction de variogramme.
Une alternative est d'utiliser (dans le cadre isotropique) un variogramme paramétrique se rapprochant le plus de l'estimation empirique. Il existe plusieurs méthodes d'estimation paramétrique d'un modèle de variogramme paramétrique $\gamma(\cdot|\theta)$: par moindres carrés ordinaires ou généralisés, ou pondérés, par maximum de vraisemblance.
L'estimation du variogramme est une étape importante dans la prédiction spatiale paramétrique.

Le Krigeage et le Cokrigeage : Comme mentionné ci-dessus, l'objectif principale en statistique spatiale est la prédiction, appelée Krigeage dans le cadre paramétrique. Il s'agit de répondre au besoin de nombreuses disciplines scientifiques, qui est de prédire ce qui se passe en un site non-observé, à partir d'observations en d'autres sites. Le Krigeage est une méthode d'interpolation spatiale issue de la géostatistique, il s'agit du meilleur prédicteur linéaire non-biaisé.
Supposons que nous disposons de la fonction de covariance spatiale $c(h)$ ou du variogramme $2\gamma(h)$ du processus Z supposé stationnaire au second ordre et observé en $(s_1, ..., s_n)$. En pratique le variogramme γ n'est pas connu, il faut l'estimer comme mentionné ci-dessus.
Avec une interpolation par Krigeage, il s'agira de donner le meilleur prédicteur linéaire de Z_{s_0} en un site s_0 non observé. Le problème consiste à déterminer la pondération W_i telle que :

$$\hat{Z}_{s_0} = \sum_{i=1}^{n} W_i Z_{s_i}$$

soit le meilleur prédicteur sans biais de Z_{s_0}. Ces poids dépendent ainsi du variogramme γ.
Il existe trois types de Krigeage :

- le *Krigeage simple* nécessite que le processus considéré soit stationnaire, de moyenne et variance connues,
- le *Krigeage ordinaire* suppose que le processus considéré soit stationnaire, de moyenne inconnue
- le *Krigeage universel* suppose que le processus considéré soit non-stationnaire.

Le Cokrigeage est une extension du Krigeage lorsqu'on dispose de plusieurs co-variables (plusieurs processus $Z^1, ..., Z^k$). Avec une interpolation par Cokrigeage, il s'agit de prédire $Z^l_{s_0}$, à partir des données des k variables mesurées sur n sites et pas seulement à partir des données du processus Z^l.

Les modèles paramétriques de Krigeage et Cokrigeage supposent généralement que le processus étudié est gaussien ce qui n'est pas toujours raisonnable. Le cadre non-paramétrique permet de s'affranchir de cette hypothèse et est donc une alternative (depuis les année 90) aux méthodes paramétriques.

L'estimation non paramétrique spatiale

Dans ce mémoire de thèse, nous nous intéressons particulièrement à l'estimation de queues de distribution des valeurs extrêmes et de leurs quantiles extrêmes correspondants par des méthodes non paramétriques spatiales. Les premières méthodes non-paramétriques spatiales concernent l'estimation de la densité et de la régression. La plupart de ces méthodes font appel à l'estimateur à noyau.
Supposons que l'ensemble spatial $\mathcal{S}$ est une région rectangulaire définie par

$$\mathcal{I}_{\mathbf{n}} = \{s = (s_1, \ldots, s_N), 1 \leq s_k \leq n_k, k = 1, \ldots, N\}.$$

Dans la suite, les champs aléatoires considérés sont également définis sur $\mathcal{I}_{\mathbf{n}}$ et on utilise la notation $\widehat{\mathbf{n}} = n_1 \times n_2 \times \cdots \times n_N$ pour faire référence à la taille de l'échantillon.

Estimation par la méthode à noyau d'une densité de probabilité. L'estimation non paramétrique d'une densité permet généralement d'étudier le comportement de la distribution d'une certaine variable. Soit $Z = \left\{Z_s,\ s \in \mathbb{R}^N\right\}$ un champ aléatoire spatial strictement stationnaire de densité marginale f inconnue avec $Z_s \in \mathbb{R}^d$, observé dans un domaine rectangulaire $\mathcal{S} = \mathcal{I}_{\mathbf{n}}$. L'estimateur à noyau de la fonction de densité en un point $z \in \mathbb{R}^d$ est donné par :

$$f_{\mathbf{n}}(z) = \frac{1}{\widehat{\mathbf{n}} b_{\mathbf{n}}^d} \sum_{s \in \mathcal{I}_{\mathbf{n}}} K\left(\frac{z - Z_s}{b_{\mathbf{n}}}\right)$$

où $K(\cdot)$ est une fonction noyau et $(b_{\mathbf{n}})$ est une suite de nombres réels positifs tendant vers zéro quand $\mathbf{n}$ tend vers l'infini. Cet estimateur a été proposé et étudié par Tran (1990). Les résultats de ce dernier ont été étendus à divers cadres par Tran & Yakowitz (1993), Carbon et al. (1996) Carbon et al. (1997), Hallin et al. (2001), Biau (2002), El Machkouri (2011), Biau (2003) et Fazekas & Chuprunov (2006),...

Estimation par la méthode à noyau d'une régression spatiale. L'estimation de la fonction de régression peut être utilisée à des fins de prévision spatiale. Soit $Z_s = (X_s, Y_s)_{s \in \mathcal{I}_{\mathbf{n}}}$ un échantillon d'un champ aléatoire $Z = \left\{(X_s, Y_s),\ s \in \mathbb{R}^N\right\}$ strictement stationnaire obéissant au modèle de régression $r(x) = \mathbb{E}[Y_s | X_s = x]$, c'est-à-dire basé sur l'espérance conditionnelle de Y_s sachant X_s où $X_s \in \mathbb{R}^d$. L'estimateur à noyau de la fonction de régression $r(\cdot)$ est donné par :

$$r_{\mathbf{n}}(x) = \frac{\frac{1}{\widehat{\mathbf{n}} b_{\mathbf{n}}^d} \sum_{s \in \mathcal{I}_{\mathbf{n}}} Y_s K\left(\frac{x - X_s}{b_{\mathbf{n}}}\right)}{f_{\mathbf{n}}(x)}.$$

Cet estimateur est au cœur d'une dynamique de recherche depuis le début des années 2000 (cf. Lu & Chen (2002), Hallin et al. (2004), Carbon et al. (2007), Biau & Cadre (2004), Carbon et al. (2007), Dabo-Niang & Yao (2007), Li & Tran (2009), Gheriballah et al. (2010), Menezes et al. (2010), Karácsony & Filzmoser (2010), Robinson (2011) dans le cas de données réelles). Dans le cadre du "fixed-design setting", le modèle de régression suivant

$$Y_{\mathbf{i}} = r\left(\frac{\mathbf{i}}{n}\right) + \epsilon_{\mathbf{i}}$$

où $(\epsilon_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$ est un champ aléatoire stationnaire réel de moyenne nulle avec $\mathbf{i} \in \{1, \ldots, n\}^N$ et $n \in \mathbb{N}^*$ a été considéré dans la littérature, voir par exemple El Machkouri (2007) et El Machkouri & Stoica (2010).

D'autres formes de régressions basées sur d'autres caractéristiques conditionnelles (quantile et mode) qui utilisent la méthode par noyau sont proposées dans la littérature (cf. Hallin et al. (2009), Dabo-Niang & Thiam (2010), Ould-Abdi et al. (2010), Dabo-Niang et al. (2014),...).

Estimation de l'indice des queues dans un cadre spatial

A notre connaissance, le seul travail qui existe dans le cadre de l'estimation de γ pour des données spatiales est celui de Basrak & Tafro (2014). Ce dernier considère l'estimation de l'indice de queue extrême γ d'un processus moyenne mobile $(X_{i,j}, i, j \in \mathbb{Z})$ observé sur une grille régulière, et montre la convergence en probabilité d'un estimateur γ_n dans le cas où les sites d'observations sont dans $\{(i,j),\ -n \leq i, j \leq n\}$. Il étend ainsi les travaux de Davis & Resnick (1985), Hsing (1991) concernant les processus temporels moyennes mobiles au cadre bivarié.
Dans ce travail, on étend les travaux de Davis & Resnick (1985), Hsing (1991), Resnick & Stărică (1998) et Basrak & Tafro (2014) au cadre de l'estimation de l'indice γ pour des données spatiales observées dans une région multidimensionnelle $\left(X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right)$ où $N \geq 2$ contrairement à Basrak & Tafro (2014) qui se limite au cas $N = 2$. La différence majeure entre notre travail et ces derniers n'est pas au niveau de l'écriture de l'estimateur. Elle est essentiellement technique, elle se trouve au niveau de l'extension des hypothèses utilisées dans le cadre de processus indexés de manière univarié au cadre multivarié et concerne la gestion de la dépendance spatiale des observations dans la région rectangulaire $\mathcal{I}_{\mathbf{n}}$. Ceci est particulièrement visible au niveau des preuves des résultats asymptotiques. De plus, nous considérons la normalité asymptotique de notre estimateur et nous avons étendu nos travaux aux cas de présence d'une covariable (déterministe ou aléatoire) et à l'estimation de quantiles conditionnels contrairement à Basrak & Tafro (2014).

Chapitre 3

On tail index estimation for random fields and application to infinite order spatial moving average processes

Sommaire

3.1 Résumé en français

Considérons un processus spatial $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$ indexé par $\mathbb{Z}^N$ avec $N > 1$ à valeurs réelles de fonction de répartition F satisfaisant,

$$F(x) = 1 - x^{-\frac{1}{\gamma}} L\left(x\right),\ \ x > 0,$$

où γ est l'unique paramètre positif inconnu appelé index des valeurs extrêmes, et L est une fonction à variations lentes

$$\forall \lambda > 0,\ \lim_{x\to\infty} \frac{L(\lambda x)}{L(x)} = 1.$$

Dans ce chapitre, nous nous intéressons à l'estimateur de Hill de l'indice (ou index) de queue γ.
Nous supposons que le processus est observé sur un domaine rectangulaire

$$\mathcal{I}_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{N}^N, 1 \le i_k \le n, k = 1, ..., N\},$$

où un point $\mathbf{i} = (i_1, \cdots, i_N) \in \mathbb{N}^N$ sera désigné comme étant un site et $\mathbf{n} = (n, \ldots, n) \in \mathbb{N}^N$ par soucis de simplicité.
Dans le cas de données dépendantes où $N = 1$ (particulièrement dans le cadre temporel), l'estimateur de l'indice des valeurs extrêmes a été considéré entre autres dans Hsing (1991), Resnick & Stărică (1995, 1998), qui généralisent ainsi les travaux existants dans le cadre de données indépendantes (e.g. Hill (1975), Resnick & Stărică (1993, 1997), Haan & Ferreira (2006)). L'estimateur étudié dans ce cadre non-spatial par les auteurs cités ci-dessus est construit comme suit.

- On définit d'abord une statistique d'ordre associée à n observations $X_1, \ldots, X_n$ de la sorte $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(n)}$.
- Ensuite on choisit une séquence intermédiaire k_n telle que $1 \leq k_n \leq n$ et $k_n = o(n)$ quand $n \to \infty$.
- Finalement, on définit l'estimateur de Hill suivant

$$\gamma_n = \frac{1}{k_n} \sum_{i=1}^{k_n} \log \left(\frac{X_{(i)}}{X_{(k_n+1)}} \right). \tag{3.1}$$

Notre objectif est de proposer une version de l'estimateur (3.1) dans le cas où les variables étudiées sont de nature spatiale. Cette extension n'est pas triviale du fait de la non-existence d'un ordre naturel dans $\mathbb{Z}^N$, $N \geq 2$ contrairement au cadre où $N = 1$ considéré dans Hsing (1991), Resnick & Stărică (1998, 1995).
Avant de définir l'estimateur de Hill dans ce cadre spatial, il est judicieux de considérer quelques transformations.
Nous considérons l'échantillon spatial comme un tableau triangulaire, c'est-à-dire les $\{X_\mathbf{i}\}_{\mathbf{i}\in\mathcal{I}_\mathbf{n}}$ seront ré-écrits $\{X_i\}_{1\leq i\leq n^N=\hat{\mathbf{n}}}$ (voir Robinson (2011)). Nous identifions ainsi chacun des indices $i = 1, \cdots, n^N$ avec un emplacement $\mathbf{i}$ dans l'ensemble $\mathcal{I}_\mathbf{n}$.

Plus généralement, soit g une fonction bijective et continue telle que

$$\begin{aligned} g : \mathbb{N}^N &\longrightarrow \mathbb{N} \\ (i_1, ..., i_N) &\longrightarrow i. \end{aligned}$$

Cette fonction transforme les indices spatiaux en indices univariés. Par exemple, lorsque $N = 2$, un site $\mathbf{i} = (i_1, i_2) \in \mathcal{I}_\mathbf{n}$, peut être indexé par $i = n(i_1 - 1) + i_2$.
L'ensemble $\{X_\mathbf{i}, \mathbf{i} \in \mathcal{I}_\mathbf{n}\}$ est alors ré-écrit comme suit :

$$\{X_i, i \in J_\mathbf{n} = g(\mathcal{I}_\mathbf{n})\}.$$

Notons $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(\hat{\mathbf{n}})}$ la statistique d'ordre associée aux $n^N = \hat{\mathbf{n}}$ variables X_i de $J_\mathbf{n}$.
Soit $\mathbf{k_n} = (k_\mathbf{n}, ..., k_\mathbf{n})$ une suite intermédiaire d'éléments de $\mathbb{N}^N$, telle que $k_\mathbf{n}$ est plus petit que n c'est-à-dire $k_\mathbf{n} = o(n)$. Nous définissons l'estimateur de l'indice de queue dans notre cadre spatial par

$$\gamma_\mathbf{n} = \frac{1}{\hat{k}_\mathbf{n}} \sum_{i=1}^{\hat{k}_\mathbf{n}} \log \left(\frac{X_{(i)}}{X_{(\hat{k}_\mathbf{n}+1)}} \right)$$

où $\hat{k}_\mathbf{n} = k_\mathbf{n}^N$.
Sous des conditions classiques des valeurs extrêmes et la convergence de la mesure empirique de queues lourdes, nous déduisons la consistance de l'estimateur proposé, c'est-à-dire

$$\gamma_\mathbf{n} \xrightarrow{\mathbb{P}} \gamma.$$

Cette consistance est prouvée dans le cas où le processus spatial est linéaire causale et dans le cas où il satisfait la condition de mélange fort ($\alpha-$mélange).

A notre connaissance, le seul travail qui existe dans le cadre de l'estimation de γ pour des données spatiales est celui de Basrak & Tafro (2014). Ce dernier considère l'estimation de l'indice de queue extrême d'un processus moyenne mobile $(X_{i,j}, i, j \in \mathbb{Z})$ observé sur une grille régulière, et montre la convergence en probabilité d'un estimateur identique à $\gamma_{\mathbf{n}}$ dans le cas où les sites d'observations sont dans $\{(i,j),\ -n \leq i, j \leq n\}$. Il étend ainsi les travaux de Davis & Resnick (1985), Hsing (1991) concernant les processus temporels moyennes mobiles au cadre bivarié.
Dans ce travail, on étend les travaux de Davis & Resnick (1985), Hsing (1991), Resnick & Stărică (1998) et Basrak & Tafro (2014) au cadre de l'estimation de l'indice γ pour des données spatiales dans une région multidimensionnelle $\left(X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right)$ où $N \geq 2$ contrairement à Basrak & Tafro (2014) qui se limitent au cas $N = 2$.

3.2 Introduction

Nowadays, in many fields, data are now collected with geographical positions such as oceanography, epidemiology, forestry survey, economy and many others. The study of these kinds of data or any characteristic of such data cannot be done without taking into account their respective geographical positions. It leads to a dynamic deployment of the known parametric and nonparametric methods (density estimation, prediction, regression, test, etc.) to spatial analysis. Spatial analysis is a general term to describe a technique that uses the spatial information in order to better understand the processes generating the observed attribute.
Several efficient spatial statistical tools are more and more developed to model spatial data. For some backgrounds in parametric spatial statistic modeling, we refer to Ripley (1981, 1991), Cressie (1992), Guyon (1995), Stein (1999), Cressie & Wikle (2011) and the references therein. More recently, nonparametric models are developed to reveal structure in data that might be missed by classical parametric ones. Some works have been done to study nonparametric variogram, quantile, density or regression problems for spatial data. We refer, for example, to Journel (1983), Tran (1990), Tran & Yakowitz (1993), Carbon et al. (1996), Biau & Cadre (2004), Menezes et al. (2010), Ould-Abdi et al. (2010), Dabo-Niang & Thiam (2010), El Machkouri (2011), García-Soidán & Menezes (2012), Dabo-Niang & Yao (2013) to name a few.

The study of extreme events (for example hurricanes, floods or earthquakes which can cause significant damage to structures such as bridges, towers and buildings) is increasingly extended to many other areas. Taking in consideration, the interest focused on the study of extreme values is to develop sophistic statistical tools for modeling extreme events. This chapter deals with extreme values estimation for spatial data.

It is well known that adequate stochastic models for the characterization and quantification of the behaviour of extremal events are needed, leading to one of the three types of extreme values limit distributions, firstly identified by Fisher et al. (1928). They play a fundamental role in the analysis of extreme events. This result obtained in 1928 on the possible limit laws of the sample maximum has created the idea that extreme values theory was something rather special and not like the classical central limit theory. Extreme values theory of stationary (non-spatial) random sequences in the univariate case has been extensively studied (e.g, Leadbetter (1974), Leadbetter et al. (1983), Hsing et al. (1988), Davis & Resnick (1985, 1988), Leadbetter & Hsing (1990), Davis & Resnick (1991)). For more details on extreme values theory we refer the reader to the book of Embrechts et al. (1997) which reminiscent the main theoretical results on extreme values theory. See also Reiss & Thomas (2001) which offers some practical examples in finance, insurance and environmental sciences. Several papers pay attention on the possible application of the extreme values theory in different areas such as in hydrology (Davison & Smith (1990), Katz et al. (2002)), in meteorology (Coles & Walshaw (1994), Smith (2001), Klajnmic (2004)) and in demography field (Gumbel (1937)) among others. The extension of the univariate to the multivariate extreme values distribution is considered for instance in Haan & Resnick (1977) and Beirlant et al. (2006). In the simplest case, the process underlying the extremes is assumed to be independent or at least stationary and satisfying a mixing condition (see Leadbetter (1974)).

However, all this theory does not take into account the spatial dependence. As so far, max-stable processes have mostly been used for the statistical modeling of spatial data (see for example, Coles (2001) and Coles & Tawn (1996) who modeled extremal rainfall fields). Padoan et al. (2010) described a practicable pairwise likelihood estimation procedure applied to the rainfall data using max-stable processes.

An interesting application to wind gusts is shown in Coles & Walshaw (1994), who used max-stable processes to model the angular dependence for wind speed directions.

In this chapter, we are interested in nonparametric tail index estimation of real-valued stationary spatial process $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$, $N \geq 2$ with same marginal distribution as a real random variable X. To the best of our knowledge, the only existing work on tail index estimation for spatial process is in Basrak & Tafro (2014). This article considered the tail estimation for moving averages and moving maxima of $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^2}$ observed in bivariate regular lattice, and shows the consistency (convergence in probability) of the tail estimate extending the work of Hsing (1991).

In this work, we extend the works of Hsing (1991) and Basrak & Tafro (2014) by dealing with the estimation of the tail index of a heavy-tailed distribution when it is applied to certain classes of heavy-tailed stationary multi-dimensionally indexed spatial processes $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$ where $N \geq 2$. We also consider, quantile estimation based on that of the tail index that drives the tail heaviness of the distribution of X and thus plays a central role in the analysis of extremes, making its estimation a crucial issue. Then, we consider the problem of estimating the extreme values index γ of the distribution F of X. We propose here a tail index estimator for spatial data based on a class of functions satisfying some mild conditions. To the best of our knowledge, such estimator has not been investigated before. Estimating an extreme spatial event, particularly, the extreme index of a distribution function F heavy-tailed is far from being trivial. This parameter controls the behavior of F at infinity, which implies that its estimate is necessary, especially when one wants to estimate extreme quantiles. There are several methods for estimating tail index for different types of process, the most widely used estimator of the extreme values index was proposed by Hill (1975). Instead of parametric estimation, we consider in the following, that the distribution of the random variables of interest is heavy-tailed and depends on a positive parameter γ. The consistency of our estimator is studied in the case where the spatial process is causal linear and when it is $\alpha-$ mixing. The proposed method is based on a function that identifies the sites and the use of the big blocks to select the observations that will be used in the estimation of the extreme values index γ thanks to the blocking technique used in Carbon et al. (1997) which is reminiscent of the blocking scheme in Tran (1990).

The chapter is organized as follows. We define the tail estimator in Section 3.3 and establish its weak consistency in the infinite order moving average case, while its asymptotic normality and also its weak consistency is given in the mixing random fields case. Section 3.4 is devoted to some preliminary results and proofs of our results.

3.3 Tail estimation

Let $\mathbb{Z}^N$; $N \geq 1$, denotes the integer lattice points in the N-dimensional Euclidean space and $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$ be an $\mathbb{R}-$valued measurable spatial process, where $X_{\mathbf{i}}$ has same distribution as X defined on the probability space (Ω, A, P). Let $\mathbb{R}$ be associated to a metric $d(.,.)$. We assume that the conditions of regularly varying tail probabilities of X is given by, for all $x > 0$,

$$\mathbb{P}\left(X > \ x\right) = x^{-\frac{1}{\gamma}} L\left(x\right), \tag{3.2}$$

where $\gamma > 0$ is unknown and $L(.)$ is a slowly varying function at infinity that is, for $\lambda > 0$

$$\lim_{x\to\infty} \frac{L\left(\lambda x\right)}{L(x)} = 1. \tag{3.3}$$

The unknown positive parameter γ is referred to the tail index. The regular variation condition in (3.2) leads to the consideration of state space $]0, +\infty]$.

In the following, we are interested in nonparametric estimation of the positive tail-index γ for spatial data. First, a point $\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N$ will be referred to as a site and define the following rectangular domain

$$\mathcal{I}_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{N}^N, 1 \leq i_k \leq n_k, k = 1, ..., N\}, \text{ with } \mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N.$$

Given a sample $(X_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}$ of dependent observations from (3.2), the aim is to build an estimator of γ. For simplicity, let $n_i = n$, $\forall i = 1, ..., N$. In the sequel, all the limits are considered when $n \to \infty$. For $\mathbf{n} = (n, ..., n) \in \mathbb{N}^N$, we set $\widehat{\mathbf{n}} = n^N$.

For convenient, we sometimes treat the spatial sample as triangular arrays, that is $(X_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}$ are written $(X_i)_{1\leq i\leq n^N}$, $n^N = \hat{\mathbf{n}}$ is the sample size (see Robinson (2011)). We can identify each of the indices $\widetilde{i} = 1, ..., \hat{\mathbf{n}}$ with a location $\mathbf{i}$ in the space $I_{\mathbf{n}}$. More generally, let g be a continuous bijective function such that :

$$\begin{aligned} g : \mathbb{N}^N &\longrightarrow \mathbb{N} \\ (i_1, ..., i_N) &\longrightarrow \widetilde{i}. \end{aligned}$$

For instance, when we have a 2-dimensional regularly-spaced lattice ($N = 2$), where both the number n_1 of rows and the number n_2 of columns increase with $\hat{\mathbf{n}} = n_1 * n_2$, the spatial points $\mathbf{i} = (i_1, i_2) \in \mathcal{I}_{\mathbf{n}}$ can be indexed by $\widetilde{i} = n_2(i_1 - 1) + i_2$. The set $\{X_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}}\}$ can be rewritten as the following array :

$$\left\{X_i, i \in J_{\mathbf{n}} = g\left(\mathcal{I}_{\mathbf{n}}\right)\right\}.$$

Let us denote by $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(\widehat{\mathbf{n}})}$ the order statistics associated to the $n^N = \hat{\mathbf{n}}$ variables $X_{i,}$ $i \in J_{\mathbf{n}}$.

Let $(\mathbf{k_n})_{\mathbf{n}}$ be a sequence of elements in $\mathbb{N}^N$ whose components are equal to $k_{\mathbf{n}} = k_n \in \mathbb{N}$ such that $1 \leq k_n \leq \hat{\mathbf{n}}$. As in the classical non-spatial data case, we shall assume that $(k_{\mathbf{n}})$ is an intermediate sequence of integers such that :

$$k_{\mathbf{n}} \to \infty \text{ and } k_{\mathbf{n}} = o\left(n\right) \ \text{ as } n \to \infty. \tag{3.4}$$

We can formally write

$$\mathbf{k_n} \to \infty \text{ and } \mathbf{k_n} = o\left(\mathbf{n}\right) \ \text{ as } \mathbf{n} \to \infty.$$

Then, we extend the Hill estimator to our spatial context. We base inference on the $\hat{k}_{\mathbf{n}}$ top-order statistics, and as in semi-parametric estimation of parameters of extreme events, see for example Resnick & Stărică (1998, 1995) :

$$\gamma_{\mathbf{n}} = \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{i=1}^{\hat{k}_{\mathbf{n}}} \log\left(\frac{X_{(i)}}{X_{(\hat{k}_{\mathbf{n}}+1)}}\right), \tag{3.5}$$

where $\hat{k}_{\mathbf{n}} = k_{\mathbf{n}}^N$. The choice of the intermediate sequence $(k_{\mathbf{n}})$ is a compromise between bias and variance of the estimator. In practice, one can try to choose an optimal $k_{\mathbf{n}}$ which minimizes asymptotic mean square error (see Danielsson et al. (2001) and Drees & Kaufmann (1998)) for more details.

To our knowledge, the only work that exists as part of the estimate of γ for spatial data is that of Basrak & Tafro (2014). The authors consider the tail index estimate of a moving average process $(X_{i,j})_{,i,j\in\mathbb{Z}}$ observed on a regular grid $\{(i,j),\ -n \le i,j \le n\}$, and show the convergence in probability of an estimate similar to $\gamma_{\mathbf{n}}$. Thus, they extend the work of Davis & Resnick (1985) and Hsing (1991) for moving averages temporal processes in the bivariate framework.

In this work, we then extend the work of Davis & Resnick (1985), Hsing (1991), Resnick & Stărică (1998) and Basrak & Tafro (2014) to the context of estimating the index γ for spatial data observed in a multidimensional region $\left(X_{\mathbf{i}},\ \mathbf{i} \in \mathbb{Z}^N\right)$ where $N \geq 2$ unlike Basrak & Tafro (2014) which is limited to the case $N = 2$. The major difference between our work and the latter is not at the writing of the estimator $\gamma_{\mathbf{n}}$. It is essentially technical, it concerns the extension of the assumptions used in the indexed process under unidimensional case to multidimensional framework in the management of spatial dependence observations in the region $\mathcal{I}_{\mathbf{n}}$. This is particularly seen in the proofs of the asymptotic results. Furthermore, we consider the asymptotic normality of our estimator and extend (in the following chapters) this present work to fixed and random covariate and also conditional quantile estimation unlike Basrak & Tafro (2014).

3.3.1 Infinite order spatial moving average framework

Let $\mathbf{i} = (i_1, \ldots, i_N)$ and $\mathbf{j} = (j_1, \ldots, j_N)$ be two N-tuples in $\mathbb{N}^N$. We said that $\mathbf{i} \geq_g \mathbf{j}$ if $i_k \geq j_k\ \forall k = 1, \ldots, N$.

In this paragraph, we are particularly interested to the causal linear spatial processes $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$ defined by

$$X_{\mathbf{i}} = \sum_{\mathbf{k} \geq_g \mathbf{0}} a_{\mathbf{k}} \epsilon_{\mathbf{i}-\mathbf{k}},\quad \mathbf{i} \in \mathbb{Z}^N, \tag{3.6}$$

where $\sum_{\mathbf{k}\geq_g \mathbf{0}} a_{\mathbf{k}}^2 < \infty$, one at least of the $a_{\mathbf{k}}$ is positive and $\{\epsilon_{\mathbf{k}}\}_{\mathbf{k}\in\mathbb{Z}^N}$ are i.i.d (supposed positive for simplicity) with regularly varying tail probabilities *i.e.* for $y > 0$,

$$P(\epsilon_{\mathbf{1}} > y) = y^{-\frac{1}{\gamma}} L(y), \tag{3.7}$$

with L a slowly varying function at infinity. All the previous conditions on the tail distribution given in the beginning of Section 3.3 will be replaced by that given in this specific subsection.

Although the causal model may not seem as natural in two dimensions as it does in one, the spatial causal linear model provides an appropriate representation of many general patterns of a stationary random field. We refer the reader to Basu & Reinsel (1993), Ivanoff et al. (2010), Doreian (1980) and the references therein. Our motivation in this section is to extend some models of extreme values theory of causal time-series or bivariate indexed processes to causal linear spatial processes defined in (3.6). In the extreme non spatial case, Barbe & McCormick (2009) established some asymptotic expansions for infinite weighted convolution of distributions having regular varying tails and developed various applications to statistics and probability models included causal linear models and ARMA models, see also Cline (1983).

Many spatial processes including causal ARMA can be represented in the form (3.6) see for example Christensen (1991) or Lozano et al. (2009). Parameters estimation in the non extreme spatial case of (3.6) has been studied by a number of authors including Tjøstheim (1978), Tjøstheim (1983), Kashyap (1984), Huang & Anh (1992), Basu & Reinsel (1993), Fazekas (2003), Illig & Truong-Van (2004) to name a few.

The tail estimation for some class of processes including autoregressive, linear, bilinear processes, dependent data has been investigated in non-spatial case ($N = 1$) by Resnick & Stărică (1998), Resnick & Stărică (1995), Hsing (1991), Datta & McCormick (1998), Borkovec & Klüppelberg (2001), Drees (2008), Brockwell & Lindner (2010), Davis & Resnick (1996), Novak (2011) among others. In the bivariate spatial case (N=2), Basrak & Tafro (2014) considered tail estimation for autoregressive processes and prove the consistency in probability of the tail index estimator.

The present extension of these authors results to high dimension ($N \geq 2$) is not trivial.

Note that when $N = 1$, Cline (1983) proved that $\sum_{\mathbf{k}\geq_g \mathbf{0}} a_{\mathbf{k}}\epsilon_{\mathbf{k}}$ has also regularly varying tail probabilities. This can be easily extended to high dimension ($N > 1$), see the following proposition (its proof is omitted since it is very similar to that of Theorem 2.3 of Cline (1983)).

Proposition 3.3.1. *Let* $\{X_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{Z}^N}$ *be a causal linear process as defined in (3.6)-(3.7), such that* $\sum_{\mathbf{k}\geq_g \mathbf{0}} |a_{\mathbf{k}}|^{\delta} < \infty,\ 0 < \delta < \min(\gamma, 1)$. *Then,*

$$\lim_{y\to\infty} \frac{P\left(\left|\sum_{\mathbf{k}\geq_g \mathbf{0}} a_{\mathbf{k}}\epsilon_{\mathbf{k}}\right| > y\right)}{P(\epsilon_{\mathbf{1}} > y)} = \sum_{\mathbf{k}\in\mathbb{Z}^N,\mathbf{k}\geq_g \mathbf{0}} |a_{\mathbf{k}}|^{1/\gamma}. \tag{3.8}$$

The following theorem gives the main consistency result of the conditional tail index estimator $\gamma_{\mathbf{n}}$.

Theorem 3.3.2. *Let* $\{X_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{Z}^N}$ *observed on* $\mathcal{I}_{\mathbf{n}}$ *be a causal spatial linear process defined in (3.6)-(3.7) such that* $\sum_{\mathbf{k}\geq_g \mathbf{0}} |a_{\mathbf{k}}|^{\delta} < \infty,\ 0 < \delta < \min(\gamma, 1)$. *Suppose also that (3.4) holds, then as* $\mathbf{n}\to\infty$,

$$\gamma_{\mathbf{n}} \xrightarrow{\mathbb{P}} \gamma, \tag{3.9}$$

where $\xrightarrow{\mathbb{P}}$ *denotes the convergence in probability.*

3.3.2 Mixing random fields framework

We assume in this section that the spatial dependence of the process $(X_{\mathbf{i}}, \mathbf{i}\in\mathbb{Z}^N)$ is measured by means of $\alpha - mixing$. Then, we consider the $\alpha - mixing$ coefficients of the field defined by :

Mixing **condition :** Let E and $E^{'}$ be two sets of sites. Let $\mathcal{B}(E) = \mathcal{B}(X_{\mathbf{i}}, \mathbf{i}\in E)$ and $\mathcal{B}\left(E^{'}\right) = \mathcal{B}\left(X_{\mathbf{i}}, \mathbf{i}\in E^{'}\right)$ be $\sigma-$fields generated by the random variables $(X_{\mathbf{i}})_{\mathbf{i}}$ with $\mathbf{i}$ being elements of E and $E^{'}$, respectively. There exists a function $\varphi(t) \downarrow 0$ as $t\to\infty$, such that whenever $E,\ E^{'}$ subsets of $\mathbb{Z}^N$ with finite cardinals,

$$\begin{aligned}\alpha\left(\mathcal{B}\left(E\right),\mathcal{B}\left(E^{'}\right)\right) &= \sup_{B\in\mathcal{B}(E),\,C\in\mathcal{B}\left(E^{'}\right)} \left|\mathbb{P}\left(B\cap C\right)-\mathbb{P}\left(B\right)\mathbb{P}\left(C\right)\right| \\ &\leq \psi\left(\text{Card}\left(E\right),\text{Card}\left(E^{'}\right)\right)\varphi\left(\text{dist}\left(E,E^{'}\right)\right), \qquad (3.10)\end{aligned}$$

where Card(E) (*resp.* Card$\left(E^{'}\right)$) denotes the cardinality of E (*resp.* $E^{'}$), dist$\left(E,E^{'}\right)$ the Euclidean distance between E and $E^{'}$ in $\mathbb{Z}^N$ defined by :

$$d\left(E,E^{'}\right)=\min\left\{\sqrt{|i_1-i'_1|^2+\cdots+|i_N-i'_N|^2}:(i_1,\cdots,i_N)\in E,\ \left(i'_1,\cdots,i'_N\right)\in E^{'}\right\}$$

and $\psi:\mathbb{N}^2\to\mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable. Throughout the chapter, it will be assumed that ψ satisfies

$$\psi\left(n,m\right)\leq C\min\left(n,m\right),\qquad \forall n,m\in\mathbb{N} \tag{3.11}$$

for some $C>0$.

We also assume that the process satisfies a polynomial mixing condition :

$$\varphi(t)\leq Ct^{-\theta},\quad \theta>0,\quad t\in\mathbb{R}_+^*. \tag{3.12}$$

If $\psi\equiv 1$, then $X_{\mathbf{i}}$ is called strongly mixing. Many stochastic processes, among them various useful time series models satisfy strong mixing properties, which are relatively easy to check. Conditions (3.11)-(3.12) are used in Tran (1990), Carbon et al. (1996, 1997). See Doukhan (1994) for discussion on mixing and examples. Here, we assume that $\mathbb{Z}^N$ is endowed with the lexicographic order.

The consistency result of Theorem 3.3.2 is extended to mixing case as follows using the additionional condition

$$\hat{\mathbf{n}}\hat{k}_{\mathbf{n}}^{-\theta_1}(\log(\hat{k}_{\mathbf{n}}))^{\frac{\theta-2N}{2N}}\to 0,\ \text{with}\quad \theta_1=\min\left(2,\theta/2N\right). \tag{3.13}$$

This condition plays an essential role in the study of the consistency of our Hill estimator in this case because it allows us to choose an intermediate sequence converging to infinity more slowly than the sample size as is done in theory extreme values, while taking account the spatial dependency.

Theorem 3.3.3. *Let $\{X_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{Z}^N}$ observed on $\mathcal{I}_{\mathbf{n}}$ be a mixing spatial process satisfying (3.11)-(3.12) with $\theta>2N$. Suppose also that (3.2)-(3.4) and (3.13) hold, then as $\mathbf{n}\to\infty$,*

$$\gamma_{\mathbf{n}}\xrightarrow{\mathbb{P}}\gamma, \tag{3.14}$$

where $\xrightarrow{\mathbb{P}}$ denotes the convergence in probability.

To prove asymptotic normality of $\gamma_{\mathbf{n}}$, we need the following additional conditions.

(C1) The cumulative distribution function F is absolutely continuous in $\mathbb{R}$.

(C2) $\theta>2(N+1)$, where θ is defined in (3.12).

($\mathcal{R}_2$) There exist a constant $\rho<0$ and a rate function $\mathcal{A}(.)$ satisfying $\mathcal{A}(x)\to 0$ as $x\to\infty$, with shape ρ such that : for $t>1$,

$$\lim_{x\to+\infty}\frac{L\left(tx\right)/L\left(x\right)-1}{\mathcal{A}(x)}=\mathcal{K}(t),$$

where $\mathcal{K}(\cdot)$ is a function such that $\mathcal{K}(t)=\int_1^t\delta^{\rho-1}d\delta$.

Set

$$b(t) = F^{-1}\left(1 - t^{-1}\right), \qquad t > 1 \tag{3.15}$$

$$F^{-1}(y) = \inf\left\{x : F(x) \geq y\right\}, \qquad 0 < y < 1,$$

F is the distribution function of X. Then regular variation implies

$$\overline{F}\left(b\left(t\right)\right) \sim t^{-1},\ \overline{F} := 1 - F. \tag{3.16}$$

The following result gives the asymptotic normality of $\gamma_{\mathbf{n}}$.

Theorem 3.3.4. *Let $\{X_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{Z}^N}$ observed on $\mathcal{I}_{\mathbf{n}}$ be a mixing spatial process satisfying (3.2)-(3.4), (3.11)-(3.13). Assume that conditions (C1)-(C2) and $(\mathcal{R}_2)$ hold. In addition, suppose that there exists a constant $\lambda \in \mathbb{R}$ such that*

$$\sqrt{\hat{k}_{\mathbf{n}}}\mathcal{A}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) \longrightarrow \lambda < \infty \ \ \text{as} \ \ \mathbf{n} \to \infty,$$

with $b(.)$ defined in (3.15). Then

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}} - \gamma\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda\gamma}{\gamma^{-1} - \rho}, \gamma^2\right). \tag{3.17}$$

It appears that the estimator is asymptotically Gaussian with asymptotic variance proportional to $\gamma^2/\hat{k}_{\mathbf{n}}$ for mixing random fields. This result is similar to the one established by Hsing (1991) in the non-spatial case.
A similar asymptotic normality result can also be obtained for spatial moving average of Section 3.3.1.

Remark 3.3.5. *The conditions used in the previous results are similar to that used in Hsing (1991), for more discussions on these conditions specific to tail index estimation see for instance Hsing (1991) and the references therein. For instance, assumption $(\mathcal{R}_2)$ is the so-called second order condition classically used to establish the asymptotic normality of tail index estimators. Note that the parameter ρ controls the rate of convergence of $L(tx)/L(x)$ to 1 (see Bingham et al. (1987) and Haan & Ferreira (2006) for further details). A particular condition, specific to our context is (3.13). This kind of assumption can be found in nonparametric inference for spatial data, see for instance Carbon et al. (1996, 1997). That is an extension of conditions in Theorem 3.3 of Hsing (1991) in non-spatial dependent case, permitting to control several covariance terms (see also condition D^* in Appendix A1) involved in the proof of the consistency results of the estimate. In independent data case, (3.13) is not needed, $\hat{k} = o(\widehat{\mathbf{n}})$ suffices.*

3.3.3 Conclusion and forthcoming studies

In this chapter, we propose an estimator of the tail index of a heavy-tailed distribution for multidimensionally indexed spatial data. We establish its main consistency properties under very general conditions by showing the convergence in probability of Hill's estimator, deduced from a consistency result of the tail empirical measure. This convergence in probability is obtained by considering some classes of stationary processes including spatial causal linear processes and $\alpha - mixing$ processes which satisfy certain condition. We first employed the blocking scheme as in Tran (1990) which specifies the locations where the useful data are concentrated. Additionally, we adapt the classical assumptions used in tail index estimation to the spatial case. Combining all these arguments with the properties of our considering classes of processes, we prove consistency of the extended version of Hill's estimator.

3.4 Appendix A1 : Tail empirical measure and technical lemmas

3.4.1 Tail empirical measure

The tail empirical measure plays a central role for the consistency of the Hill's estimator. Before going further, let us define this measure.

Let $\mathbb{E} =:]0, \infty[$ and $\mathcal{E}$ be the Borel σ-algebra generated by open sets of $\mathbb{E}$. Let $\mathbf{C}_{\mathbf{K}}^{+}(\mathbb{E})$ be the space of continuous nonnegative functions on $\mathbb{E}$ with compact support $\mathbf{K}$ and $\mathbf{M}_{+}(\mathbb{E})$ be the space of positive Radon measures on $\mathbb{E}$ endowed with the vague topology (see Resnick (1987) and Kallenberg (1983)).
Define the measure μ

$$\mu : \mathcal{E} \to \mathbb{R}_{+},$$

with density $\mu(dx) = \gamma^{-1} x^{-\gamma^{-1}-1} dx$.

For $x \in \mathbb{E}$ and $A \in \mathcal{E}$, define the measure ε_x by

$$\varepsilon_x(A) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A. \end{cases} \tag{3.18}$$

Then, we define the tail empirical measure by

$$\widehat{\mu}_{\widehat{\mathbf{n}}}(.) = \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \varepsilon_{\frac{X_{\mathbf{i}}}{b\left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)}}(.)\,. \tag{3.19}$$

The following result gives the convergence of the tail empirical measure estimate corresponding to a spatial process $\{X_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{Z}^N}$ observed on $\mathcal{I}_{\mathbf{n}}$. For that we need the following assumption :

Condition D^* **:** *For any disjoints collections of sites* $\mathbf{I}(\mathbf{j}) = \{\mathbf{i},\, j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\, k=1,...,N\}$, *each of size* p^N $(p^N = o(\hat{\mathbf{n}})$, $p = p_{\mathbf{n}} \to \infty$, $q = q_{\mathbf{n}} \to \infty$, $q/p \to 0)$ *and separated at least by* q, *we have for* $f \in \mathbf{C}_{\mathbf{K}}^{+}(\mathbb{E})$

$$\mathbf{E}\left(\prod_{\substack{j_k=0\\k=1,...,N}}^{r_k-1} \exp\left(-\frac{1}{\hat{k}_{\mathbf{n}}}\sum_{\mathbf{i}\in\mathbf{I}(\mathbf{j})} f\left(X_{\mathbf{i}}^{*}\right)\right)\right) - \prod_{\substack{j_k=0\\k=1,...,N}}^{r_k-1} \mathbf{E}\left(\exp\left(-\frac{1}{\hat{k}_{\mathbf{n}}}\sum_{\mathbf{i}\in\mathbf{I}(\mathbf{j})} f\left(X_{\mathbf{i}}^{*}\right)\right)\right) \to 0,$$

where $X_{\mathbf{i}}^{*} = \frac{X_{\mathbf{i}}}{b(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}})}$, $\mathbf{r} = (r_1, ..., r_N) \in \mathbb{N}^N$.

Remark 3.4.1. *Condition* D^* *is an extension of condition* D *in Adler (1978), see also relation (2.4) of Resnick & Stărică (1998). It consists of defining blocks* $\mathbf{I}(\mathbf{j})$ *such that the variables* $\sum_{\mathbf{i}\in\mathbf{I}(\mathbf{j})} f\left(X_{\mathbf{i}}^{*}\right)$ *are asymptotically independent. The sets defined in for instance Resnick & Stărică (1998) are of form*
$I(1) = [1, \hat{k}_n - l_n]$, $I(2) = [\hat{k}_n + 1, 2\hat{k}_n - l_n]$, ..., $I([n/\hat{k}_n]) = [([n/\hat{k}_n]-1)\hat{k}_n + 1, [n/\hat{k}_n]\hat{k}_n - l_n]$ *where* $l_n \to \infty$, $l_n = o(\hat{k}_n)$. *These sets are each of size* $\hat{k}_n - l_n$ *and separated by* l_n.
In our spatial setting, we extend sets $I(j)$ *to spatial sets* $\mathbf{I}(\mathbf{j})$ *each of size* p^N *separated by* $q_{\mathbf{n}}$ *using the blocking scheme in Tran (1990). We show in the following that the considered* $M-$*dependent and mixing processes satisfy* D^* *with an appropriate choice of* M *or* p, q.

Proposition 3.4.2. *Consider a spatial process* $\{X_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{Z}^N}$ *observed on* $\mathcal{I}_{\mathbf{n}}$ *where* $X_{\mathbf{i}}$ *has same law as* X*. Suppose also that* D^* *holds. If in addition the following two conditions hold :*

(i) $\forall f \in \mathbf{C}_{\mathbf{K}}^{+}(\mathbb{E})$, $\lim\limits_{\mathbf{n}\to\infty} \frac{\widehat{\mathbf{r}}}{\hat{k}_{\mathbf{n}}^2} \left\{ \sum\limits_{\substack{l_u,\, i_u=1,\\ u=1,\cdots,N,\, \mathbf{i}\neq\mathbf{l}}}^{p+q} \mathbf{E}\left[f(X_{\mathbf{i}}^*)f(X_{\mathbf{l}}^*)\right] \right\} = 0$, $\widehat{\mathbf{r}} = r_1 \times \ldots \times r_N$.

(ii) $\frac{\widehat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}} \mathbb{P}\left(X_{\mathbf{i}}^* \in \cdot\ \right) \xrightarrow{v} \mu(\cdot)$. [2]

Then

$$\hat{\mu}_{\widehat{\mathbf{n}}}(\cdot) \Longrightarrow \mu(\cdot)\,\text{[3]} \quad \textit{in } \mathbf{M}^{+}(\mathbb{E}). \tag{3.20}$$

Recall that $n^N = \widehat{\mathbf{n}}$. The following proposition extends the previous one to $M-$dependent process.

Proposition 3.4.3. *Consider an* $M-$*dependent spatial process* $(X_{\mathbf{i}}^M)_{\mathbf{i}\in\mathbb{Z}^N}$ *in the sense that* $X_{\mathbf{i}}^M$ *and* $X_{\mathbf{j}}^M$ *are independent if* $\|\mathbf{i}-\mathbf{j}\|_\infty \geq M$*. Let* $(X_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}$ *be a spatial process where* $X_{\mathbf{i}}$ *has same law as* X *and is observed on* $\mathcal{I}_{\mathbf{n}}$*. Suppose that there exists Radon measures* $\mu^{(M)}$ *on* $\mathbb{E}$ *such that for any* $M \geq 1$

$$\frac{\widehat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}} P\left(X_{\mathbf{i}}^{*(M)} \in .\right) \xrightarrow{v} \mu^{(M)}(\cdot). \tag{3.21}$$

If in addition the following two conditions hold :

(i) $\mu^{(M)} \xrightarrow{v} \mu$ *as* $M \to \infty$.

(ii) $\lim\limits_{M\to\infty} \limsup\limits_{\mathbf{n}\to\infty} \frac{\widehat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}} P\left(|X_{\mathbf{i}}^{*(M)} - X_{\mathbf{i}}^*| > \varepsilon\right) = 0$ *for all* $\varepsilon > 0$.

Then

$$\hat{\mu}_{\widehat{\mathbf{n}}}(\cdot) \Longrightarrow \mu(\cdot) \ \textit{ in } \mathbf{M}^{+}(\mathbb{E}). \tag{3.22}$$

3.4.2 Technical lemmas

Lemma 3.4.4 (Tran (1990))**.** ***(i)*** *Suppose that* (3.10) *holds. Denote by* $\mathcal{L}_r(\mathcal{F})$ *the class of* $\mathcal{F}$*-measurable r.v's* X *satisfying* $\|X\|_r = (E\,|X|^r)^{\frac{1}{r}} < \infty$*. Suppose* $X_1 \in \mathcal{L}_r(\mathcal{B}(\mathbf{I}_1))$ *and* $X_2 \in \mathcal{L}_r(\mathcal{B}(\mathbf{I}_2))$*, where* $\mathbf{I}_1$*,* $\mathbf{I}_2$ *are two sets of spatial sites. Assume also that* $1 \leq r, s, t < \infty$ *and* $r^{-1} + s^{-1} + t^{-1} = 1$*. Then*

$$|\mathbf{E}X_1X_2 - \mathbf{E}X_1\mathbf{E}X_2| \leq C\,\|X_1\|_r\,\|X_2\|_s\,\{\psi\left(\mathit{Card}(\mathbf{I}_1), \mathit{Card}(\mathbf{I}_2)\right)\varphi\left(\mathrm{d}(\mathbf{I}_1,\mathbf{I}_2)\right)\}^{\frac{1}{t}}.$$

(ii) *For r.v's bounded with probability* 1 *the right-hand side of the last inequality can be replaced by*

$$C\psi\left(\mathit{Card}(\mathbf{I}_1), \mathit{Card}(\mathbf{I}_2)\right)\varphi\left(\mathrm{d}(\mathbf{I}_1,\mathbf{I}_2)\right).$$

Lemma 3.4.5 (Tran (1990))**.** *Let* $(\xi_1, \cdots, \xi_n)$ *be a random vector such that* $\left|\mathbf{E}\left[\prod\limits_{s=i}^{n}\xi_i\right]\right| < \infty$*, with* $|C\xi_i| \leq 1$*,* $i = 1, \cdots, n$*. Then*

$$\left|\mathbf{E}\left[\prod_{s=1}^{n}\xi_s\right] - \prod_{s=1}^{n}\mathbf{E}\left[\xi_s\right]\right| \leq \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\Big|\mathbf{E}\left[(\xi_i-1)(\xi_j-1)\right] \times \prod_{s=j+1}^{n}\xi_s$$
$$- \mathbf{E}\left[(\xi_i-1)\right]\mathbf{E}\left[(\xi_j-1)\right]\prod_{s=j+1}^{n}\xi_s\Big|.$$

---

[2] A sequence $(\mu_n) \in \mathbf{M}_{+}(\mathbb{E})$ converges vaguely (see Resnick (1987)) to μ (written $\mu_n \xrightarrow{v} \mu$) if $\mu_n(f) \to \mu(f)$ for all $f \in \mathbf{C}_{\mathbf{K}}^{+}(\mathbb{E})$

[3] $\Longrightarrow$ denotes the weak convergence (see Davis & Resnick (1988))

Lemma 3.4.6 (Carbon et al. (1996))**.** *Suppose $S_1, S_2, \cdots, S_r$ be sets containing m sites each with dist $(S_i, S_j) \geq \delta, \ \ \delta > 0$ for all $i \neq j$ where $1 \leq i, j \leq r$. Suppose $Y_1, \cdots, Y_r$ is a sequence of real-valued random variables measurable with respect to $\mathcal{B}(S_1), \cdots, \mathcal{B}(S_r)$ respectively, and Y_i takes values in $[a; b]$. Then there exists a sequence of independent random variables $Y_1^*, \cdots, Y_r^*$ independent of $Y_1, \cdots, Y_r$ such that Y_i^* has the same distribution as Y_i and satisfies :*

$$\sum_{i=1}^{r} \mathbb{E}\left|Y_i - Y_i^*\right| \leq 2r\,(b-a)\,\psi\left((r-1)m, m\right)\varphi\left(\delta\right). \tag{3.23}$$

3.5 Appendix A2 : Intermediate lemmas and proofs of main results

In this section, we establish the proofs of our main results.

We will employing the blocking technique used in Carbon et al. (1997) which is reminiscent of the blocking scheme in Tran (1990). Without loss of generality assume that $\widehat{\mathbf{n}} = \widehat{\mathbf{r}}(p+q)^N$ where $\widehat{\mathbf{r}} = r_1 \times ... \times r_N$, $q \to \infty$, $q/p \to 0$. Denote

$$\begin{aligned}
\mathbf{I}(1,\mathbf{j}) &= \{\mathbf{i},\, j_k(p+q)+1 \leq i_k \leq j_k(p+q)+p,\, k=1,...,N\} \\
\mathbf{I}(2,\mathbf{j}) &= \{\mathbf{i},\, j_k(p+q)+1 \leq i_k \leq j_k(p+q)+p,\, k=1,...,N-1, \\
&\quad j_N(p+q)+p+1 \leq i_N \leq (j_N+1)(p+q)\}, \\
\mathbf{I}(3,\mathbf{j}) &= \{\mathbf{i},\, j_k(p+q)+1 \leq i_k \leq j_k(p+q)+p,\ k=1,...,N-2, \\
&\quad j_{N-1}(p+q)+p+1 \leq i_{N-1} \leq (j_{N-1}+1)(p+q), \\
&\quad j_N(p+q)+1 \leq i_N \leq j_N(p+q)+p\}, \\
\mathbf{I}(4,\mathbf{j}) &= \{\mathbf{i},\, j_k(p+q)+1 \leq i_k \leq j_k(p+q)+p,\, k=1,...,N-2, \\
&\quad j_{N-1}(p+q)+p+1 \leq i_{N-1} \leq (j_{N-1}+1)(p+q), \\
&\quad j_N(p+q)+p+1 \leq i_N \leq (j_N+1)(p+q)\}
\end{aligned} \tag{3.24}$$

and so on. The last two terms are

$$\begin{aligned}
\mathbf{I}(2^{N-1},\mathbf{j}) &= \{\mathbf{i},\, j_k(p+q)+p+1 \leq i_k \leq (j_k+1)(p+q), k=1,...,N-1, \\
&\quad j_N(p+q)+1 \leq i_N \leq j_N(p+q)+p\}
\end{aligned}$$

and

$$\mathbf{I}(2^N,\mathbf{j}) = \{\mathbf{i},\, j_k(p+q)+p+1 \leq i_k \leq (j_k+1)(p+q),\, k=1,...,N,\}$$

For each integer $1 \leq i \leq 2^N$, define

$$D(i) = \{\mathbf{I}(i,\mathbf{j}),\, 0 \leq j_k \leq r_k - 1,\, k=1,...,N\} \tag{3.25}$$

Note that the set $\mathcal{I}_{\mathbf{n}}$ is then decomposed into these 2^N small and large blocks in sets $D(i)$. If it is not the case that $\widehat{\mathbf{n}} = \widehat{\mathbf{r}}(p+q)^N$ then an additional set $D(2^N+1)$ containing all the sites of $\mathcal{I}_{\mathbf{n}}$ not in these blocks $\mathbf{I}(i,\mathbf{j})$ can be added, this will not change the proof a lot.

For simplicity let $\hat{k} = \hat{k}_{\mathbf{n}}$.

3.5.1 Appendix A21 : Convergence in probability

Proof of Proposition 3.4.2

We follow the same lines as Resnick & Stărică (1998) by adapting them in our spatial context. It suffices to prove that for any function $f \in \mathbf{C}_{\mathbf{K}}^{+}(\mathbb{E})$

$$\lim_{\mathbf{n}\to\infty} \mathbb{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} f\left(X_{\mathbf{i}}^{*}\right)\right)\right\} = \exp\left(-\mu(f)\right).$$

Let us use the blocks $\mathbf{I}(i,\mathbf{j})$ defined above. For simplicity, let $f_{\mathbf{i}} = f\left(X_{\mathbf{i}}^{*}\right)$, and

$$\Psi_p = \mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}_{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} f_{\mathbf{i}}\right)\right\} - \exp\left(-\mu(f)\right).$$

Now we employ the previous blocking scheme in Tran (1990). Denote

$$\begin{aligned}
\mathbf{U}(1,\mathbf{j}) &= \sum_{\mathbf{i}\in I(1,\mathbf{j})} f_{\mathbf{i}},\ \mathbf{U}(2,\mathbf{j}) = \sum_{I(2,\mathbf{j})} f_{\mathbf{i}}\\
\mathbf{U}(3,\mathbf{j}) &= \sum_{I(3,\mathbf{j})} f_{\mathbf{i}},\ \mathbf{U}(4,\mathbf{j}) = \sum_{I(4,\mathbf{j})} f_{\mathbf{i}}
\end{aligned}$$

and so on. The last two terms are

$$\mathbf{U}(2^{N-1},\mathbf{j}) = \sum_{I(2^{N-1},\mathbf{j})} f_{\mathbf{i}},\ \mathbf{U}(2^{N},\mathbf{j}) = \sum_{I(2^{N},\mathbf{j})} f_{\mathbf{i}}.$$

For each integer $1 \le i \le 2^N$, define

$$T(i) = \sum_{\mathbf{j}\in D(i)} \mathbf{U}(i,\mathbf{j}). \tag{3.26}$$

Then, we set

$$\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} f_{\mathbf{i}} = \sum_{i=1}^{2^N} T(i).$$

We have :

$$\begin{aligned}
|\Psi_p| \le & \left|\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} f_{\mathbf{i}}\right)\right\} - \mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}T(1)\right)\right\}\right|\\
& + \left|\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}T(1)\right)\right\} - \left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{j})\right)\right\}\right)^{\widehat{\mathbf{r}}}\right|\\
& + \left|\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{j})\right)\right\}\right)^{\widehat{\mathbf{r}}} - \left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\ k=1,\dots,N}}^{p+q} f_{\mathbf{i}}\right)\right\}\right)^{\widehat{\mathbf{r}}}\right|\\
& + \left|\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\ k=1,\dots,N}}^{p+q} f_{\mathbf{i}}\right)\right\}\right)^{\widehat{\mathbf{r}}} - \exp\left\{-\mu(f)\right\}\right|\\
= & \ \mathbf{I} + \mathbf{II} + \mathbf{III} + \mathbf{IV}.
\end{aligned}$$

Note that

$$\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{j})\right)\right\}\right)^{\widehat{\mathbf{r}}} = \prod_{\substack{t_k=0\\k=1,\ldots,N}}^{r_k-1}\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{j})\right)\right\}\right)$$

Now we will examine the individual terms in turn :

Step 1.

$$\begin{aligned}
\mathbf{I} &= \left|\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} f_\mathbf{i}\right)\right\} - \mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}T(1)\right)\right\}\right| \\
&= \left|\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\left(T(1)+\sum_{j=2}^{2^N}T(j)\right)\right)\right\} - \mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}T(1)\right)\right\}\right| \\
&\leqslant \mathbf{E}\left|1-\exp\left(-\frac{1}{\hat{k}}\sum_{j=2}^{2^N}T(j)\right)\right| \\
&\leq \mathbf{E}\left|\frac{1}{\hat{k}}\sum_{j=2}^{2^N}T(j)\right| \\
&\leqslant \frac{1}{\hat{k}}\sum_{j=2}^{2^N}\mathbf{E}(T(j)) \leqslant \sum_{j=2}^{2^N}\widehat{\mathbf{r}}p^{N-j+1}q^{j-1}\frac{1}{\hat{k}}\mathbf{E}\left(f_\mathbf{1}\right) \leq \sum_{j=2}^{2^N}\widehat{\mathbf{n}}(p+q)^{-N}p^{N-j+1}q^{j-1}\frac{1}{\hat{k}}\mathbf{E}\left(f_\mathbf{1}\right).
\end{aligned}$$

since each $U(i,\mathbf{j})$ contains $p^{N-i+1}q^{i-1}$ sites. Notice that for $j \geq 2$,

$$\frac{p^{N-j+1}q^{j-1}}{(p+q)^N} = \frac{p^{N-j+1}}{(p+q)^{N-j+1}}\frac{q^{j-1}}{(p+q)^{j-1}} \to 0$$

since $q/p \to 0$ implies that $\frac{p+q}{p} \to 1$ and $\frac{q}{p+q} \to 0$.
Since $f \in \mathbf{C}_\mathbf{K}^+(\mathbb{E})$, we set $[c,\infty]$ for the support of f, where c is a positive constant and set $\|f\| = \sup_{x\in\mathbb{E}} f(x)$, then

$$f \leq \|f\|\,\mathbb{I}_{[c;\infty)}.$$

Therefore, I tends to zero, according condition $ii)$.

Step 2. **II** goes to zero, by D^*.

Step 3. It is treated in a similar way as **I** :

$$
\begin{aligned}
\mathbf{III} &= \left| \left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{j})\right)\right\}\right)^{\widehat{\mathbf{r}}} - \left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\k=1,\dots,N}}^{p+q} f_{\mathbf{i}}\right)\right\}\right)^{\widehat{\mathbf{r}}}\right| \\
&\leq \widehat{\mathbf{r}}\left| \left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{0})\right)\right\}\right) - \left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\k=1,\dots,N}}^{p+q} f_{\mathbf{i}}\right)\right\}\right)\right| \\
&\leq \widehat{\mathbf{r}}\mathbf{E}\left| 1 - \left(\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{p+q}}/\mathcal{I}_{\mathbf{p}}} f_{\mathbf{i}}\right)\right\}\right)\right| \\
&\leq \widehat{\mathbf{r}}\mathbf{E}\left| \frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{p+q}}/\mathcal{I}_{\mathbf{p}}} f_{\mathbf{i}}\right| \\
&\leq \frac{\widehat{\mathbf{r}}}{\hat{k}}\mathbf{E}\left(\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{p+q}}/\mathcal{I}_{\mathbf{p}}} f_{\mathbf{i}}\right) \leq \widehat{\mathbf{r}}((p+q)^N - p^N)\frac{1}{\hat{k}}\mathbf{E}\left(f_{\mathbf{1}}\right) \\
&\leq \frac{((p+q)^N - p^N)}{(p+q)^N}\frac{\widehat{\mathbf{n}}}{\hat{k}}\mathbf{E}\left(f_{\mathbf{1}}\right).
\end{aligned}
$$

Since $\frac{((p+q)^N - p^N)}{(p+q)^N} \to 0$, using again condition *ii*), **III** tends to 0 as $n \to \infty$.

Step 4. From term **IV**, we set $y_{\mathbf{i}} = 1 - \exp\left(-\frac{1}{\hat{k}}f_{\mathbf{i}}\right)$ and apply the following inequality :

$$
\begin{aligned}
1 - \sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} y_{\mathbf{i}} \leq \prod_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} (1 - y_{\mathbf{i}}) \leq 1 - \sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} y_{\mathbf{i}} \\
+ \sum_{\substack{i_k=1\\k=1,\dots,N\\ \mathbf{i}\neq\mathbf{l}}}^{(p+q)} \sum_{\substack{l_k=1\\k=1,\dots,N}}^{(p+q)} y_{\mathbf{i}}y_{\mathbf{l}}, \qquad 0 \leq y_{\mathbf{i}} \leq 1, \forall \mathbf{i}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} f_{\mathbf{i}}\right)\right\} &= \mathbf{E}\left\{\prod_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} (1 - y_{\mathbf{i}})\right\} \\
&\leq 1 - \mathbf{E}\left\{\sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} y_{\mathbf{i}}\right\} + \mathbf{E}\left\{\sum_{\substack{i_k=1\\k=1,\dots,N\\ \mathbf{i}\neq\mathbf{l}}}^{(p+q)} \sum_{\substack{l_k=1\\k=1,\dots,N}}^{(p+q)} y_{\mathbf{i}}y_{\mathbf{l}}\right\} \\
&\leq 1 - (p+q)^N\mathbf{E}\left\{y_{\mathbf{1}}\right\} + \mathbf{E}\left\{\sum_{\substack{l_k,\, i_k=1,\ k=1,\dots,N\\ \mathbf{i}\neq\mathbf{l}}}^{(p+q)} y_{\mathbf{i}}y_{\mathbf{l}}\right\}
\end{aligned}
$$

$$\leq 1-\mathbf{E}\left\{\sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} y_{\mathbf{i}}\right\}+\mathbf{E}\left\{\sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)}\sum_{\substack{l_k=1\\k=1,\dots,N\\\mathbf{i}\neq\mathbf{l}}}^{(p+q)} y_{\mathbf{i}}y_{\mathbf{l}}\right\}$$

$$\leq 1-(p+q)^N\mathbf{E}\left\{y_{\mathbf{1}}\right\}+\mathbf{E}\left\{\sum_{\substack{l_k,\,i_k=1,\;k=1,\dots,N\\\mathbf{i}\neq\mathbf{l}}}^{(p+q)} y_{\mathbf{i}}y_{\mathbf{l}}\right\}.$$

We have

$$\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} f_{\mathbf{i}}\right)\right\}\right)^{\widehat{\mathbf{r}}} \tag{3.27}$$

$$\leq\left(1-\frac{(p+q)^N\widehat{\mathbf{r}}\left(\mathbf{E}\left(y_{\mathbf{1}}\right)-(p+q)^{-N}\left\{\mathbf{E}\left\{\sum_{\substack{l_k,\,i_k=1,\;k=1,\dots,N\\\mathbf{i}\neq\mathbf{l}}}^{(p+q)} y_{\mathbf{i}}y_{\mathbf{l}}\right\}\right\}\right)}{\widehat{\mathbf{r}}}\right)^{\widehat{\mathbf{r}}}. \tag{3.28}$$

Notice that

$$\widehat{\mathbf{r}}\mathbf{E}\left\{\sum_{\substack{l_u,\,i_u=1,\,u=1,\dots,N\\\mathbf{i}\neq\mathbf{l}}}^{(p+q)} y_{\mathbf{i}}y_{\mathbf{l}}\right\}\leq C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\left\{\sum_{\substack{l_u,\,i_u=1,\,u=1,\dots,N\\\mathbf{i}\neq\mathbf{l}}}^{(p+q)}\mathbf{E}\left(f_{\mathbf{i}}f_{\mathbf{l}}\right)\right\}.$$

By assumption, we have :

$$\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\left\{\sum_{\substack{l_u,\,i_u=1,\,u=1,\dots,N\\\mathbf{i}\neq\mathbf{l}}}^{(p+q)}\mathbf{E}\left(f_{\mathbf{i}}f_{\mathbf{l}}\right)\right\}\to 0.$$

We also have

$$(p+q)^N\widehat{\mathbf{r}}\mathbf{E}\left(y_{\mathbf{1}}\right)=\widehat{\mathbf{n}}\mathbf{E}\left(y_{\mathbf{1}}\right)\to\mu(f).$$

Hence the right hand side of (3.28), goes to $\exp\{-\mu(f)\}$. Thus we conclude

$$\lim_{n\to\infty}\sup\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} f_{\mathbf{i}}\right)\right\}\right)^{\widehat{\mathbf{r}}}\leqslant\exp\left\{-\mu\left(f\right)\right\}.$$

A slightly simpler argument gives :

$$\lim_{n\to\infty}\inf\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\sum_{\substack{i_k=1\\k=1,\dots,N}}^{(p+q)} f_{\mathbf{i}}\right)\right\}\right)^{\widehat{\mathbf{r}}}\geqslant\exp\left\{-\mu\left(f\right)\right\}.$$

This ends the proof.■

Proof of Proposition 3.4.3

Let us show that

$$\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\varepsilon_{X_{\mathbf{i}}^{*(M)}}\to\mu^{(M)}$$

by checking hypotheses of Proposition 3.4.2 for $\{X_{\mathbf{i}}^{(M)}\}_{\mathbf{i}\in\mathbb{Z}^N}$. Since (3.21) holds it suffices to show that i) (of Proposition 3.4.2) and D^* hold. Condition D^* holds since by taking $q>M$, $\{\sum\limits_{\mathbf{i}\in I(\mathbf{j})} f(X_{\mathbf{i}}^{*(M)}),\, \mathbf{j}\in\{\mathbf{0},...,\hat{\mathbf{r}}\}\}$ are independent random variables, this is due to the fact that $\{X_{\mathbf{i}}^{(M)}\}_{\mathbf{i}\in\mathbb{Z}^N}$ is $M-$dependent.

Proof of $i)$ **:**

We know that $f\in\mathbf{C}^{+}_{\mathbf{K}}(\mathbb{E})$, as above set $[c,\infty]$ for the support of f, where c is a positive constant and $\|f\|=\sup_{x\in\mathbb{E}}f(x)$, then

$$f\leq\|f\|\,\mathbb{I}_{[c;\infty)},$$

and

$$\mathbf{E}\left\{f\left(X_{\mathbf{i}}^{*(M)}\right)f\left(X_{\mathbf{l}}^{*(M)}\right)\right\}\leq\|f\|^2\,\mathbb{P}\left(X_{\mathbf{i}}^{*(M)}>c, X_{\mathbf{l}}^{*(M)}>c\right).$$

We will study (i) for all $\mathbf{i}\neq\mathbf{l}$ in two sets defined as :

let

$$\begin{aligned}S_1=\{&j_u(p+q)+1\leq l_u,\, i_u\leq(j_u+1)(p+q),\ u=1,...,N,\\ &\mathbf{i}\neq\mathbf{l}:\ 1<\|\mathbf{i}-\mathbf{l}\|\leq M,\, \mathbf{j}=(j_1,...,j_N)\}\ \text{ and}\\ S_2=\{&j_u(p+q)+1\leq l_u,\, i_u\leq(j_u+1)(p+q),\ u=1,...,N,\\ &\mathbf{i}\neq\mathbf{l}:\ M+1\leq\|\mathbf{i}-\mathbf{l}\|<(p+q),\, \mathbf{j}=(j_1,...,j_N)\}.\end{aligned}$$

We have that for $f\in\mathbf{C}^{+}_{\mathbf{K}}(\mathbb{E})$,

$$\begin{aligned}\lim_{m\to\infty}\quad & \frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\substack{l_u,\,i_u=j_u(p+q)+1\\ u=1,...,N,\ \mathbf{i}\neq\mathbf{l}}}^{(j_u+1)(p+q)}\mathbf{E}\left[f\left(X_{\mathbf{i}}^{*(M)}\right)f\left(X_{\mathbf{l}}^{*(M)}\right)\right]\\ \leq\ & C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_1}\mathbb{P}\left(X_{\mathbf{i}}^{*(M)}>c,\, X_{\mathbf{l}}^{*(M)}>c\right)+C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_2}\mathbb{P}\left(X_{\mathbf{i}}^{*(M)}>c,\, X_{\mathbf{l}}^{*(M)}>c\right)\\ \leq\ & C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_1}\mathbb{P}\left(X_{\mathbf{i}}^{*(M)}>c\right)+C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_2}\mathbb{P}\left(X_{\mathbf{i}}^{*(M)}>c\right)\mathbb{P}\left(X_{\mathbf{l}}^{*(M)}>c\right)\\ \leq\ & C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}(p+q)^N M^N\mathbb{P}\left(X_{\mathbf{1}}^{*(M)}>c\right)+C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}(p+q)^{2N}\mathbb{P}\left(X_{\mathbf{1}}^{*(M)}>c\right)\mathbb{P}\left(X_{\mathbf{1}}^{*(M)}>c\right)\\ \leq\ & C\frac{M^N}{\hat{k}}\frac{\widehat{\mathbf{n}}}{\hat{k}}\mathbb{P}\left(X_{\mathbf{1}}^{*(M)}>c\right)+C\frac{(p+q)^N}{\widehat{\mathbf{n}}}\frac{\widehat{\mathbf{n}}^2}{\hat{k}^2}\mathbb{P}\left(X_{\mathbf{1}}^{*(M)}>c\right)\mathbb{P}\left(X_{\mathbf{1}}^{*(M)}>c\right).\end{aligned}$$

Since $\frac{\widehat{\mathbf{n}}}{\hat{k}}\mathbb{P}\left(X_{\mathbf{i}}^{*(M)}>c\right)\to\mu\left(c,\infty\right]$, $(p+q)^N=o(\widehat{\mathbf{n}})$ and $M=M_{\mathbf{n}}\to+\infty$, these two terms go to zero by choosing M such that $M^N=o(\hat{k})$. The proof of (i) is completed. The proof of the Proposition then follows from Proposition 3.3 of Resnick & Stărică (1995). ■

Proof of Theorem 3.3.2. Recall that

$$X_{\mathbf{i}} = \sum_{\mathbf{k} \geq_g \mathbf{0}} a_{\mathbf{k}} \epsilon_{\mathbf{i}-\mathbf{k}}, \quad \mathbf{i} \in \mathbb{Z}^N \tag{3.29}$$

This can be rewritten (by renumbering the sites) :

$$\tilde{X}_i = \sum_{k \geq 0} \tilde{a}_k \tilde{\epsilon}_{i-k}, \quad i \in \mathbb{Z} \tag{3.30}$$

Let

$$X_{\mathbf{i}}^{(M)} = \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{M}} a_{\mathbf{k}} \epsilon_{\mathbf{i}-\mathbf{k}}, \quad \mathbf{i} \in \mathbb{Z}^N \tag{3.31}$$

rewritten as

$$\tilde{X}_i^{(M)} = \sum_{k=0}^{M^N} \tilde{a}_k \tilde{\epsilon}_{i-k}, \quad i \in \mathbb{Z}. \tag{3.32}$$

Here, let $M^N = o(\hat{k})$. Let us apply Proposition 3.4.3. First, we prove the existence of the measures $\mu^{(M)}$ and (i). By (3.8) and Proposition 3.2 of Resnick & Stărică (1995), we have for $x > 0$

$$\frac{\widehat{\mathbf{n}}}{\hat{k}} P\left(\frac{\sum_{k=0}^{M^N} \tilde{a}_k \tilde{\epsilon}_{i-k}}{b\left(\frac{\mathbf{n}}{\hat{k}}\right)} > x \right) \to \frac{\sum_{k=0,\, \tilde{a}_k>0}^{M^N} \tilde{a}_k^{\gamma}}{\sum_{k=0,\, \tilde{a}_k>0}^{\infty} \tilde{a}_k^{\gamma}} x^{-1/\gamma}, \tag{3.33}$$

this permits to define the Radon measures $\mu^{(M)}$, as

$$\mu^{(M)}((x,\infty]) = \frac{\sum_{k=0,\, \tilde{a}_k>0}^{M^N} \tilde{a}_k^{\gamma}}{\sum_{k=0,\, \tilde{a}_k>0}^{\infty} \tilde{a}_k^{\gamma}} x^{-1/\gamma}, \tag{3.34}$$

and note that $\mu^{(M)} \xrightarrow{v} \mu$ where $\mu((x,\infty]) = x^{-1/\gamma}$. Finally, let us prove (ii). We have (using similar arguments are that of Proposition 3.2 of Resnick & Stărică (1995))

$$\lim_{M\to\infty} \limsup_{\mathbf{n}\to\infty} \frac{\widehat{\mathbf{n}}}{\hat{k}} P\left(|X_{\mathbf{i}}^{*(M)} - X_{\mathbf{i}}^{*}| > \varepsilon \right) = \lim_{M\to\infty} \frac{\sum_{k=M^N+1,\, \tilde{a}_k>0}^{\infty} \tilde{a}_k^{\gamma}}{\sum_{k=0,\, \tilde{a}_k>0}^{\infty} \tilde{a}_k^{\gamma}} \varepsilon^{-1/\gamma} = 0,$$

condition(ii) of Proposition 3.4.3 is then verified.
Then, we make use of Proposition 2.4 of Resnick & Stărică (1995) which shows that the convergence of the tail measure implies the consistency of Hill's estimator.■

Proof of Theorem 3.3.3. It suffices to show that D^*, i) and ii) (of Proposition 3.4.2) hold.
(ii) : For any $x > 0$, it suffices to show that

$$\frac{\widehat{\mathbf{n}}}{\hat{k}} \mathbb{P}\left(X_{\mathbf{i}} / b\left(\frac{\widehat{\mathbf{n}}}{\hat{k}}\right) > x \right) \to \mu(x; +\infty]. \tag{3.35}$$

By the regular variation (3.2) and the definition of the function $b(.)$, it is easy to see that

$$\begin{aligned}
\lim_{\mathbf{n}\to\infty} \frac{\widehat{\mathbf{n}}}{\hat{k}} \mathbb{P}\left(X_{\mathbf{i}} / b\left(\frac{\widehat{\mathbf{n}}}{\hat{k}}\right) > x \right) &= \lim_{\mathbf{n}\to\infty} \mathbb{P}\left(X_{\mathbf{i}} > x b\left(\frac{\widehat{\mathbf{n}}}{\hat{k}}\right) \right) / \overline{F}\left(b\left(\frac{\widehat{\mathbf{n}}}{\hat{k}}\right) \right) \\
&= \lim_{\mathbf{n}\to\infty} \frac{\mathbb{P}\left(X_{\mathbf{i}} > x b\left(\frac{\widehat{\mathbf{n}}}{\hat{k}}\right) \right)}{\mathbb{P}\left(X_{\mathbf{1}} > b\left(\frac{\widehat{\mathbf{n}}}{\hat{k}}\right) \right)} \\
&= x^{-\frac{1}{\gamma}} = \mu(x; +\infty].
\end{aligned} \tag{3.36}$$

Then (3.35) permits to prove (ii).
Proof of (i) **:**
We have to prove that :

$$\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\left\{\sum_{\substack{l_u,\, i_u=1,\ u=1,\ldots,N\\ \mathbf{i}\neq\mathbf{l}}}^{(p+q)} \mathbf{E}\left(f_{\mathbf{i}}f_{\mathbf{l}}\right)\right\} \to 0,$$

where $f_{\mathbf{i}} = f\left(X_{\mathbf{i}}/b\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right)$. Let

$$S_1 = \{1 \le l_u,\, i_u \le p+q,\ u=1,\ldots,N,\, \mathbf{i}\neq\mathbf{l} :\ 1 < \|\mathbf{i}-\mathbf{l}\| \le c_{\mathbf{n}}\}$$

and

$$S_2 = \{1 \le l_u,\, i_u \le p+q,\ u=1,\ldots,N,\, \mathbf{i}\neq\mathbf{l} :\ c_{\mathbf{n}} < \|\mathbf{i}-\mathbf{l}\|\}\,.$$

For $f \in \mathbf{C}^{+}_{\mathbf{K}}(\mathbb{E})$, note $[c,\infty)$ the support of f, where c is a positive constant and set $\|f\| = \sup_{x\in\mathbb{E}} f(x)$, then

$$f \le \|f\|\,\mathbb{I}_{[c;\infty)}$$

and

$$\begin{aligned}
\lim_{m\to\infty}\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\substack{l_u,\, i_u=1,\\ u=1,\ldots,N\\ \mathbf{i}\neq\mathbf{l}}}^{(p+q)} & \mathbf{E}\left[f_{\mathbf{i}}f_{\mathbf{l}}\right] \\
\le\ & C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_1}\mathbb{P}\left(X_{\mathbf{i}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right),\, X_{\mathbf{l}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) \\
+\ & C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_2}\mathbb{P}\left(X_{\mathbf{i}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right),\, X_{\mathbf{l}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) \\
\le\ & C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_1}\mathbb{P}\left(X_{\mathbf{i}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) + C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\mathbf{i},\mathbf{l}\in S_2}\psi(1,1)\varphi(\|\mathbf{i}-\mathbf{l}\|) \\
\le\ & C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}(p+q)^{2N}\mathbb{P}\left(X_{\mathbf{1}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) + C\frac{\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\|\mathbf{i}-\mathbf{l}\|>c_n}\psi(1,1)\varphi(\|\mathbf{i}-\mathbf{l}\|) \\
\le\ & C\frac{(p+q)^N}{\hat{k}}\frac{\widehat{\mathbf{n}}}{\hat{k}}\mathbb{P}\left(X_{\mathbf{1}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) + C\frac{(p+q)^N\widehat{\mathbf{r}}}{\hat{k}^2}\sum_{\|\mathbf{i}\|>c_{\mathbf{n}}}\varphi(\|\mathbf{i}\|) \\
\le\ & C\left(\frac{(p+q)^N}{\hat{k}}\frac{\widehat{\mathbf{n}}}{\hat{k}}\mathbb{P}\left(X_{\mathbf{1}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) + \frac{(p+q)^N\widehat{\mathbf{r}}}{\hat{k}^2}c_{\mathbf{n}}^{-N}\sum_{\|\mathbf{i}\|>c_{\mathbf{n}}}\|\mathbf{i}\|^N\varphi(\|\mathbf{i}\|)\right) \\
\le\ & C\left(\frac{(p+q)^N}{\hat{k}}\frac{\widehat{\mathbf{n}}}{\hat{k}}\mathbb{P}\left(X_{\mathbf{1}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) + \frac{1}{\hat{k}}\sum_{\|\mathbf{i}\|>0}\|\mathbf{i}\|^{N-\theta}\right)
\end{aligned}$$

by taking $c_{\mathbf{n}} = \left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)^{1/N} \to +\infty$. Since $\frac{\widehat{\mathbf{n}}}{\hat{k}}\mathbb{P}\left(X_{\mathbf{1}} > cb\left(\frac{\hat{\mathbf{n}}}{\hat{k}}\right)\right) \to \mu\left(c,\infty\right]$, and $\theta > N+1$, if we let here p, q, such that $q = [\hat{k}^{1/2N}(\log(\hat{k}))^{\frac{2N-\theta}{2N\theta}}]$, $p = [\hat{k}^{1/2N}]$ then, $(p+q)^N = o(\hat{k})$, the proof of (i) is then completed.

Proof of D^* **:**

We have to prove that :

$$\mathbf{E}\left(\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathbf{I}(1,\mathbf{j})}f_{\mathbf{i}}\right)\right)-\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}\mathbf{E}\left(\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathbf{I}(1,\mathbf{j})}f_{\mathbf{i}}\right)\right)\to 0.$$

We have

$$\begin{aligned}&\mathbf{E}\left(\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathbf{I}(1,\mathbf{j})}f_{\mathbf{i}}\right)\right)-\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}\mathbf{E}\left(\exp\left(-\frac{1}{\hat{k}}\sum_{\mathbf{i}\in\mathbf{I}(1,\mathbf{j})}f_{\mathbf{i}}\right)\right),\\ &=\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}T(1)\right)\right\}-\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{j})\right)\right\}\right).\end{aligned}$$

Enumerate the variables $\mathbf{U}(1,\mathbf{j})/\hat{k}$ in an arbitrary manner and refer to them as $\tilde{U}_1,\dots,\tilde{U}_{\widehat{\mathbf{r}}}$ and remember that distinct sets of sites $\mathbf{I}(1,\mathbf{j})$ are distant of at least q. Then, Lemma 3.4.5 permits to write that :

$$\begin{aligned}Q&=\left\|\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}T(1)\right)\right\}-\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}\left(\mathbf{E}\left\{\exp\left(-\frac{1}{\hat{k}}\mathbf{U}(1,\mathbf{j})\right)\right\}\right)\right\|\\ &\le\sum_{k=1}^{\widehat{\mathbf{r}}-1}\sum_{j=k+1}^{\widehat{\mathbf{r}}}\left|\mathbf{E}\left\{\exp\left(\tilde{U}_k\right)-1\right\}\times\left\{\exp\left(\tilde{U}_j\right)-1\right\}\prod_{s=j+1}^{\widehat{\mathbf{r}}}\exp\left(\tilde{U}_s\right)\right.\\ &\qquad\qquad\left.-\mathbf{E}\left\{\exp\left(\tilde{U}_k\right)-1\right\}\mathbf{E}\left\{\exp\left(\tilde{U}_j\right)-1\right\}\prod_{s=j+1}^{\widehat{\mathbf{r}}}\exp\left(\tilde{U}_s\right)\right|\end{aligned}$$

Let $\tilde{I}(j)$ be the sets $\mathbf{I}(1,\mathbf{j})$ involved with $\tilde{U}_j$, then by Lemma 3.4.4 (ii) we have :

$$\begin{aligned}&\left|\mathbf{E}\left\{\exp\left(\tilde{U}_k\right)-1\right\}\times\left\{\exp\left(\tilde{U}_j\right)-1\right\}-\mathbf{E}\left\{\exp\left(\tilde{U}_k\right)-1\right\}\mathbf{E}\left\{\exp\left(\tilde{U}_j\right)-1\right\}\right|\\ &\qquad\qquad\le C\varphi\left(d\left(\tilde{I}(j),\tilde{I}(k)\right)\right)p^N.\end{aligned}$$

Then we have,

$$\begin{aligned}Q&\le Cp^N\sum_{k=1}^{\widehat{\mathbf{r}}-1}\sum_{j=k+1}^{\widehat{\mathbf{r}}}\varphi\left(d\left(\tilde{I}(j),\tilde{I}(k)\right)\right)\\ &\le Cp^N\widehat{\mathbf{r}}\sum_{k=2}^{\widehat{\mathbf{r}}}\varphi\left(d\left(\tilde{I}(1),\tilde{I}(k)\right)\right)\\ &\le Cp^N\widehat{\mathbf{r}}\sum_{i=1}^{\infty}\sum_{k:iq\le d\left(\tilde{I}(1),\tilde{I}(k)\right)\le(i+1)q}\varphi\left(d\left(\tilde{I}(1),\tilde{I}(k)\right)\right)\\ &\le Cp^N\widehat{\mathbf{r}}\sum_{i=1}^{\infty}i^{N-1}\varphi\left(iq\right)\le Cp^N\widehat{\mathbf{r}}\sum_{i=1}^{\infty}i^{N-1-\theta}q^{-\theta}\le C\widehat{\mathbf{n}}q^{-\theta}.\end{aligned}$$

The proof is then completed if one takes q such that $q=\left[\hat{k}^{1/2N}(\log(\hat{k}))^{\frac{2N-\theta}{2N\theta}}\right]$, $p=\left[\hat{k}^{1/2N}\right]$ then $\widehat{\mathbf{n}}q^{-\theta}\to 0$, by (3.13) with $\theta>2N$.

The proof of the proposition follows from Proposition 3.3 of of Resnick & Stărică (1995).■

3.5.2 Appendix A22 : Intermediate results and their proofs

Let us give some notations and definitions, we have

$$\gamma_{\mathbf{n}} = \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{i=1}^{\hat{k}_{\mathbf{n}}} \log\left(X_{(i)}/X_{(\hat{k}_{\mathbf{n}}+1)}\right) \text{ and define :}$$

$$\widetilde{\gamma}_{\mathbf{n}} := \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{i=1}^{\hat{k}_{\mathbf{n}}} \log\left(X_{(i)}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right),$$

$$\gamma_{\mathbf{n}}^{+} := \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_{+} = \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{i=1}^{n^N} \log\left(X_i/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_{+}, \quad \text{where } z_{+} = \max(0, z)$$

$$I_{\mathbf{n}}^{(\tau)} := \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbb{1}_{\left\{\log X_{\mathbf{i}} > \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) + \tau/\sqrt{\hat{k}}\right\}}, \quad \text{where } \tau \in \mathbb{R}.$$

We also define the following quantities

$$\begin{aligned}
\eta_{\mathbf{i},\mathbf{n}} &:= \left[\log(X_{\mathbf{i}}) - \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\right]_{+} \quad \text{and} \quad \Delta_{\mathbf{i},\mathbf{n}} := \eta_{\mathbf{i},\mathbf{n}} - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right]. \\
\eta_{\mathbf{i},\mathbf{n},\tau} &:= \mathbb{1}_{\left\{\log(X_{\mathbf{i}}) - \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) > \tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}} \quad \text{and} \quad \Delta_{\mathbf{i},\mathbf{n},\tau} := \eta_{\mathbf{i},\mathbf{n},\tau} - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}\right] \\
S_{\mathbf{n},\gamma} &:= \hat{k}_{\mathbf{n}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n}} \quad \text{and} \quad S_{\mathbf{n},\tau} = \hat{k}_{\mathbf{n}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n},\tau}.
\end{aligned}$$

In addition, let

$$\mathbf{H}_{\mathbf{n}}^{+} = \sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}^{+} - \mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right) \quad \text{and} \quad \mathbf{H}_{\mathbf{n}}^{(\tau)} = \sqrt{\hat{k}_{\mathbf{n}}}\left(I_{\mathbf{n}}^{(\tau)} - \mathbf{E}\left[I_{\mathbf{n}}^{(\tau)}\right]\right).$$

First, by the dominated convergence and Potter's Theorem (see also, Hsing (1991) p. 1548), the k^{th} moment of $\left[\log(X_{\mathbf{1}}) - \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\right]_{+}$ is approximated as

$$\begin{aligned}
\mathbf{E}\left[\log(X_{\mathbf{1}}) - \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\right]_{+}^{k} &= \int_0^\infty \mathbb{P}\left[\left(\log\left(X_{\mathbf{1}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\right)^k > u\right] du \\
&= \int_0^\infty \mathbb{P}\left[X_{\mathbf{1}} > b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right) e^{u^{1/k}}\right] du \\
&= \int_0^\infty \overline{F}\left(e^{u^{1/k}} b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) du \\
&= \overline{F}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) \int_0^\infty \overline{F}\left(e^{u^{1/k}} b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) / \overline{F}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) du \\
&\sim \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}} \int_0^\infty e^{-\gamma^{-1} u^{1/k}} du = \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}} k! \gamma^k.
\end{aligned}$$

Then,

$$\mathbf{E}\left[\log(X_{\mathbf{1}}) - \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\right]_{+}^{k} \sim \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}} k! \gamma^k. \tag{3.37}$$

The following lemma gives results on the variances of $S_{\mathbf{n},\gamma}$ and $S_{\mathbf{n},\tau}$.

Lemma 3.5.1. *Under model (3.2) and the conditions in Theorem 3.3.4, we have*

1. $\hat{k}_{\mathbf{n}} \mathrm{Var}\left(S_{\mathbf{n},\gamma}\right) = \gamma^2 \left(1 + o(1)\right)$.
2. $\hat{k}_{\mathbf{n}} \mathrm{Var}\left(S_{\mathbf{n},\tau}\right) = 1 + o(1)$.

Proof of Lemma 3.5.1

1. Before going further, let us denote

$$I_{\mathbf{n},\gamma} := \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathrm{Var}\left(\eta_{\mathbf{i},\mathbf{n}}\right) \quad \text{and} \quad R_{\mathbf{n},\gamma} = \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} \left|\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}}, \eta_{\mathbf{j},\mathbf{n}}\right)\right|.$$

In this part, we compute the variance of $S_{\mathbf{n},\gamma}$ by using the method proposed by Tran (1990). Recall that $\eta_{\mathbf{i},\mathbf{n}} := \left[\log\left(X_{\mathbf{i}}\right) - \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\right]_+$ and $\Delta_{\mathbf{i},\mathbf{n}} := \eta_{\mathbf{i},\mathbf{n}} - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right]$. It is easily seen that

$$\mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n}}\right) = \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathrm{Var}\left(\eta_{\mathbf{i},\mathbf{n}}\right) + \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} \mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}}, \eta_{\mathbf{j},\mathbf{n}}\right).$$

Clearly, $\mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n}}\right) \leq I_{\mathbf{n},\gamma} + R_{\mathbf{n},\gamma}$, and

$$I_{\mathbf{n},\gamma} = \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^2} \left(\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}^2\right] - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right]^2\right).$$

It follows that

$$I_{\mathbf{n},\gamma} = \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^2} \left(\mathbf{E}\left[\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+^2\right] - \mathbf{E}\left[\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+\right]^2\right).$$

From (3.37), we have,

$$\mathbf{E}\left[\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+\right] \sim \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma \quad \text{and} \quad \mathbf{E}\left[\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+^2\right] \sim \frac{2\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma^2,$$

then

$$I_{\mathbf{n},\gamma} \sim \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^2} \left(2\gamma^2 \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}} - \frac{\hat{k}_{\mathbf{n}}^2}{\hat{\mathbf{n}}^2}\gamma^2\right) = 2\gamma^2 \hat{k}_{\mathbf{n}}^{-1} - \hat{\mathbf{n}}^{-1}\gamma^2. \tag{3.38}$$

Let us treat the term $R_{\mathbf{n},\gamma}$. We want to show that there exists a constant C not depending on $\mathbf{n}$ such that $R_{\mathbf{n},\gamma} \leq C\hat{k}_{\mathbf{n}}^{-1}$ for $\mathbf{n}$ large enough. Notice that

$$\begin{aligned}\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}}, \eta_{\mathbf{j},\mathbf{n}}\right) = {} & \hat{k}_{\mathbf{n}}^{-2}\mathbf{E}\left[\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+ \log\left(X_{\mathbf{j}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+\right] \\ & - \hat{k}_{\mathbf{n}}^{-2}\mathbf{E}\left[\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+\right] \mathbf{E}\left[\log\left(X_{\mathbf{j}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+\right]\end{aligned}$$

We will study the covariance between $\eta_{\mathbf{i},\mathbf{n}}$ and $\eta_{\mathbf{j},\mathbf{n}}$ for all $\mathbf{i} \neq \mathbf{j}$ in two sets defined as follows, we set : $\mathcal{S}_{\Omega_{\mathbf{n}}} = \left\{\mathbf{i} \neq \mathbf{j} \in \mathcal{I}_{\mathbf{n}} : 0 < \mathrm{d}\left(\mathbf{i}, \mathbf{j}\right) < \Omega_{\mathbf{n}}\right\}$ and denote by $\mathcal{S}_{\Omega_{\mathbf{n}}}^c$ the complement of $\mathcal{S}_{\Omega_{\mathbf{n}}}$ where $\Omega_{\mathbf{n}}$ is a sequence of real numbers which tends to ∞ as $\mathbf{n} \to \infty$. Define

$$R_{1,\mathbf{n},\gamma} := \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}} \left|\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}}, \eta_{\mathbf{j},\mathbf{n}}\right)\right| \quad \text{and} \quad R_{2,\mathbf{n},\gamma} := \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}^c} \left|\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}}, \eta_{\mathbf{j},\mathbf{n}}\right)\right|.$$

Clearly, $R_{\mathbf{n},\gamma} \leq R_{1,\mathbf{n},\gamma} + R_{2,\mathbf{n},\gamma}$. Applying the first part of Lemma 3.4.4, on the $\eta_{\mathbf{i},\mathbf{n}}$, we get,

$$\left|\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}}, \eta_{\mathbf{j},\mathbf{n}}\right)\right| \leq \|\eta_{\mathbf{i},\mathbf{n}}\|_s \|\eta_{\mathbf{j},\mathbf{n}}\|_u \left(\psi(1,1)\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)\right)^{1/t},$$

with $1/t + 1/s + 1/u = 1$, we have $\|\eta_{\mathbf{i},\mathbf{n}}\|_s^s = \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}} s!\gamma^s$. It follows that

$$\left|\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}}, \eta_{\mathbf{j},\mathbf{n}}\right)\right| \leq \frac{\hat{k}_{\mathbf{n}}^{1/s+1/u}}{\hat{\mathbf{n}}^{1/s+1/u}} (s!)^{1/s}(u!)^{1/u}\gamma^2 \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}.$$

- ▶ First, concerning $R_{1,\mathbf{n},\gamma}$, we have by setting $u = s = 4$, $t = 2$

$$\begin{aligned} R_{1,\mathbf{n},\gamma} &\leq (s!)^{1/s}(u!)^{1/u}\gamma^2 \frac{\hat{k}_{\mathbf{n}}^{\frac{1}{s}+\frac{1}{u}-2}}{\hat{\mathbf{n}}^{\frac{1}{s}+\frac{1}{u}}} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{S}_{\Omega_{\mathbf{n}}}} \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t} \\ &\leq (s!)^{1/s}(u!)^{1/u}\gamma^2 \frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{t}-1}}{\hat{\mathbf{n}}^{1-\frac{1}{t}}} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{S}_{\Omega_{\mathbf{n}}}} \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t} \\ \hat{k}_{\mathbf{n}} R_{1,\mathbf{n},\gamma} &\leq (s!)^{1/s}(u!)^{1/u}\gamma^2 \frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{2}}}{\hat{\mathbf{n}}^{\frac{1}{2}}} \Omega_{\mathbf{n}}^N \sum_{\|\mathbf{i}\| < \Omega_{\mathbf{n}}} \|\mathbf{i}\|^{-\theta/2}. \end{aligned} \tag{3.39}$$

- ▶ For the term $R_{2,\mathbf{n},\gamma}$, we have

$$\begin{aligned} R_{2,\mathbf{n},\gamma} &\leq (s!)^{1/s}(u!)^{1/u}\gamma^2 \frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{t}-1}}{\hat{\mathbf{n}}^{1-\frac{1}{t}}} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{S}^c_{\Omega_{\mathbf{n}}}} \sum \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t} \\ \hat{k}_{\mathbf{n}} R_{2,\mathbf{n},\gamma} &\leq (s!)^{1/s}(u!)^{1/u}\gamma^2 \frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{2}}}{\hat{\mathbf{n}}^{-\frac{1}{2}}} \Omega_{\mathbf{n}}^{-N} \sum_{\|\mathbf{i}\| > \Omega_{\mathbf{n}}} \|\mathbf{i}\|^{\frac{2N-\theta}{2}}. \end{aligned} \tag{3.40}$$

Since $\theta > 2(N+1)$ (then $\sum_{\mathbf{i}>0} \|\mathbf{i}\|^{\frac{2N-\theta}{2}} < \infty$), letting $\Omega_{\mathbf{n}} = \left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)^{1/(2N)}$ and combining (3.39) and (3.40), gives

$$R_{\mathbf{n},\gamma} = O\left(\hat{k}_{\mathbf{n}}^{-1}\right). \tag{3.41}$$

The proof of statement **1** follows from (3.38), (3.41) and by the fact that

$$\hat{k}_{\mathbf{n}} \mathrm{Var}\left(S_{\mathbf{n},\gamma}\right) = \hat{k}_{\mathbf{n}} \mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n}}\right) = 2\gamma^2 - \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma^2 + o(1).$$

2. We will compute in the same way as statement **1** the variance of $S_{\mathbf{n},\tau}$. First, define

$$I_{\mathbf{n},\tau} := \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \mathrm{Var}\left(\eta_{\mathbf{i},\mathbf{n},\tau}\right) \quad \text{and} \quad R_{\mathbf{n},\tau} = \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i} \neq \mathbf{j} \in \mathcal{I}_{\mathbf{n}}} \left|\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau}, \eta_{\mathbf{j},\mathbf{n},\tau}\right)\right|.$$

Since $\eta_{\mathbf{i},\mathbf{n},\tau} := \mathbb{1}_{\left\{\log(X_{\mathbf{i}}) - \log\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) > \tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}$ and $\Delta_{\mathbf{i},\mathbf{n},\tau} := \eta_{\mathbf{i},\mathbf{n},\tau} - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}\right]$, it is easily seen that

$$\mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n},\tau}\right) = \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \mathrm{Var}\left(\eta_{\mathbf{i},\mathbf{n},\tau}\right) + \hat{k}_{\mathbf{n}}^{-2} \sum_{\mathbf{i} \neq \mathbf{j} \in \mathcal{I}_{\mathbf{n}}} \mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau}, \eta_{\mathbf{j},\mathbf{n},\tau}\right).$$

Clearly, $\mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n},\tau}\right) \leq I_{\mathbf{n},\tau} + R_{\mathbf{n},\tau}$, we deduce that,

$$I_{\mathbf{n},\tau} = \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^2} \left(\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}^2\right] - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}\right]^2\right).$$

From (3.37), we have,

$$\mathbf{E}\left[\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}\right]\sim\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\text{ and }\mathbf{E}\left[\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}\right]^2\sim\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}.$$

then

$$I_{\mathbf{n},\tau}\sim\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^2}\left(\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}-\frac{\hat{k}_{\mathbf{n}}^2}{\hat{\mathbf{n}}^2}\right)=\hat{k}_{\mathbf{n}}^{-1}-\hat{\mathbf{n}}^{-1}. \tag{3.42}$$

We now treat the term $R_{\mathbf{n},\tau}$, we want to show that there exists a constant C not depending on $\mathbf{n}$ such that $R_{\mathbf{n},\tau}\leq C\hat{k}_{\mathbf{n}}^{-1}$ for $\mathbf{n}$ large enough.

$$\begin{aligned}\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau},\eta_{\mathbf{j},\mathbf{n},\tau}\right)&=\hat{k}_{\mathbf{n}}^{-2}\mathbf{E}\left[\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}\mathbb{1}_{\left\{\log\left(X_{\mathbf{j}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}\right]\\&\quad-\hat{k}_{\mathbf{n}}^{-2}\mathbf{E}\left[\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}\right]\mathbf{E}\left[\mathbb{1}_{\left\{\log\left(X_{\mathbf{j}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}\right]\end{aligned}$$

In this part, to study the covariance between $\eta_{\mathbf{i},\mathbf{n},\tau}$ and $\eta_{\mathbf{j},\mathbf{n},\tau}$ for all $\mathbf{i}\neq\mathbf{j}$, we use the previously defined two sets which are $\mathcal{S}_{\Omega_{\mathbf{n}}}=\{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}:0<\mathrm{d}\left(\mathbf{i},\mathbf{j}\right)<\Omega_{\mathbf{n}}\}$ and $\mathcal{S}_{\Omega_{\mathbf{n}}}^c$. Denote

$$R_{1,\mathbf{n},\tau}:=\hat{k}_{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau},\eta_{\mathbf{j},\mathbf{n},\tau}\right)\right|\text{ and }R_{2,\mathbf{n},\tau}:=\hat{k}_{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}^c}\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau},\eta_{\mathbf{j},\mathbf{n},\tau}\right)\right|.$$

Clearly, $R_{\mathbf{n},\tau}\leq R_{1,\mathbf{n},\tau}+R_{2,\mathbf{n},\tau}$. Applying again the first part of Lemma 3.4.4, on the $\eta_{\mathbf{i},\mathbf{n},\tau}$, we get,

$$\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau},\eta_{\mathbf{j},\mathbf{n},\tau}\right)\right|\leq\|\eta_{\mathbf{i},\mathbf{n},\tau}\|_s\|\eta_{\mathbf{j},\mathbf{n},\tau}\|_u(\psi(1,1)\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right))^{1/t},$$

with $1/t+1/s+1/u=1$, we have $\|\eta_{\mathbf{i},\mathbf{n}}\|_s^s=\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}$. It follows that

$$\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau},\eta_{\mathbf{j},\mathbf{n},\tau}\right)\right|\leq\frac{\hat{k}_{\mathbf{n}}^{1/s+1/u}}{\hat{\mathbf{n}}^{1/s+1/u}}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}.$$

- Let $u=s=4$, $t=2$, first, concerning $R_{1,\mathbf{n},\tau}$, we have

$$\begin{aligned}R_{1,\mathbf{n},\tau}&\leq C\frac{\hat{k}_{\mathbf{n}}^{\frac{1}{s}+\frac{1}{u}-2}}{\hat{\mathbf{n}}^{\frac{1}{s}+\frac{1}{u}}}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}=C\frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{t}-1}}{\hat{\mathbf{n}}^{1-\frac{1}{t}}}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t},\\\hat{k}_{\mathbf{n}}R_{1,\mathbf{n},\tau}&\leq C\frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{2}}}{\hat{\mathbf{n}}^{\frac{1}{2}}}\Omega_{\mathbf{n}}^N\sum_{\|\mathbf{i}\|<\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{-\theta/2}.\end{aligned} \tag{3.43}$$

- For the term $R_{2,\mathbf{n},\tau}$, we have

$$\begin{aligned}R_{2,\mathbf{n},\tau}&\leq C\frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{t}-1}}{\hat{\mathbf{n}}^{1-\frac{1}{t}}}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}^c}\sum\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}\\\hat{k}_{\mathbf{n}}R_{2,\mathbf{n},\tau}&\leq C\frac{\hat{k}_{\mathbf{n}}^{-\frac{1}{2}}}{\hat{\mathbf{n}}^{-\frac{1}{2}}}\Omega_{\mathbf{n}}^{-N/2}\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{\frac{2N-\theta}{2}}.\end{aligned} \tag{3.44}$$

Since $\theta > 2(N+1)$ then $\sum\limits_{\mathbf{i}>0} \|\mathbf{i}\|^{\frac{2N-\theta}{2}} < \infty$, letting $\Omega_{\mathbf{n}} = \left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)^{1/(2N)}$, then (3.43) and (3.44) give

$$R_{\mathbf{n},\tau} = O\left(\hat{k}_{\mathbf{n}}^{-1}\right). \tag{3.45}$$

The proof of statement **2** follows from (3.42), (3.45) and by the fact that

$$\hat{k}_{\mathbf{n}}\mathrm{Var}\left(S_{\mathbf{n},\tau}\right) = \hat{k}_{\mathbf{n}}\mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n},\tau}\right) = 1 - \frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}} + o(1).\blacksquare$$

Lemma 3.5.2. *Under (3.2) and the conditions in Theorem 3.3.4, we have*

1. $\mathbf{H}_{\mathbf{n}}^{+} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\ 2\gamma^2\right)$ *and* $\mathbf{H}_{\mathbf{n}}^{(\tau)} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\ 1\right)$.
2. *Moreover as* $\mathbf{n}\to\infty$, $\mathbf{H}_{\mathbf{n}} := \left(\mathbf{H}_{\mathbf{n}}^{(\tau)}, \mathbf{H}_{\mathbf{n}}^{+}\right)^{\top}$ *converges in distribution to a bivariate Gaussian vector* $\mathcal{N}\left(0,\Sigma\right)$, *where*

$$\Sigma := \begin{pmatrix} 1 & \gamma \\ \gamma & 2\gamma^2 \end{pmatrix}$$

and $\top$ *denotes the transpose.*

Proof of Lemma 3.5.2

1. ▶ We first prove asymptotic normality of $\mathbf{H}_{\mathbf{n}}^{+}$. With the notations that are given at the beginning of this section, it follows that

$$\mathbf{H}_{\mathbf{n}}^{+} = \hat{k}_{\mathbf{n}}^{1/2}S_{\mathbf{n},\gamma} \quad \text{where} \qquad S_{\mathbf{n},\gamma} := \hat{k}_{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n}},$$
$$\Delta_{\mathbf{i},\mathbf{n}} := \eta_{\mathbf{i},\mathbf{n}} - \mathbb{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right] \ \text{ and } \ \eta_{\mathbf{i},\mathbf{n}} := \left[\log\left(X_{\mathbf{i}}\right)/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right]_{+}.$$

First, using (3.13), take sequences of positive integers $p_{\mathbf{n}} = p$, $q_{\mathbf{n}} = q$ tending to infinity such that $qp^{-1} = o(1)$. Therefore, $q = q_{\mathbf{n}} = o(p)$. For some integers $r_1, \cdots, r_N$, we assume that $n_1 = r_1(q+p), \cdots, n_N = r_N(q+p)$.

Then, the random variables $\Delta_{\mathbf{i},\mathbf{n}}$'s are set as above into large blocks and small blocks, that is

$$\begin{aligned} U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right) &= \sum_{\mathbf{i}\in I(1,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}},\ U_{\mathbf{n}}\left(2,\gamma,\mathbf{j}\right) = \sum_{I(2,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}} \\ U_{\mathbf{n}}\left(3,\gamma,\mathbf{j}\right) &= \sum_{I(3,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}},\ U_{\mathbf{n}}\left(4,\gamma,\mathbf{j}\right) = \sum_{I(4,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}} \end{aligned}$$

and so on. The last two terms are

$$U_{\mathbf{n}}\left(2^{N-1},\gamma,\mathbf{j}\right) = \sum_{I(2^{N-1},\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}},\ U_{\mathbf{n}}\left(2^{N},\gamma,\mathbf{j}\right) = \sum_{I(2^{N},\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}}$$

For each integer $1 \le i \le 2^N$, define

$$T_{\mathbf{n},\gamma,i} = \textstyle\sum_{\mathbf{j}\in D(i)}U_{\mathbf{n}}\left(i,\gamma,\mathbf{j}\right). \tag{3.46}$$

Let $K_{\mathbf{n},\gamma} = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n}}$, clearly $K_{\mathbf{n},\gamma} = \sum\limits_{i=1}^{2^N}T_{\mathbf{n},\gamma,i}$. With this notation, $T_{\mathbf{n},\gamma,1}$ is the sum of the random variables $\Delta_{\mathbf{i},\mathbf{n}}$ in large blocks. For $2 \le i \le 2^N$, the $T_{\mathbf{n},\gamma,i}$ are sums of random variables in small blocks. Using Lemma 3.5.1 (see Tran (1990)), it suffices to show that as $\mathbf{n} \to \infty$,

$$Q_{1,\mathbf{n},\gamma} \equiv \left| \mathbf{E}\left[\exp\left(izT_{\mathbf{n},\gamma,1}\right)\right] - \prod_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[\exp\left(izU_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right)\right)\right] \right| \to 0, \tag{3.47}$$

$$Q_{2,\mathbf{n},\gamma} \equiv \hat{k}_{\mathbf{n}}^{-1}\mathbf{E}\left[\left(\sum_{i=2}^{2^N} T_{\mathbf{n},\gamma,i}\right)^2\right] \to 0, \tag{3.48}$$

$$Q_{3,\mathbf{n},\gamma} \equiv \hat{k}_{\mathbf{n}}^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right)^2\right] \to \overline{\sigma}^2, \quad \text{where } \overline{\sigma} = \gamma^2 \tag{3.49}$$

$$Q_{4,\mathbf{n},\gamma} \equiv \hat{k}_{\mathbf{n}}^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right)^2 \mathbb{1}_{\left\{|U_{\mathbf{n}}(1,\gamma,\mathbf{j})| \geq \varepsilon\overline{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right\}}\right] \to 0 \ \ \forall\varepsilon > 0. \tag{3.50}$$

We have

$$\frac{H_{\mathbf{n}}^{+}}{\overline{\sigma}} = \frac{K_{\mathbf{n},\gamma}}{\overline{\sigma}\sqrt{\hat{k}_{\mathbf{n}}}} = \frac{T_{\mathbf{n},\gamma,1}}{\overline{\sigma}\ \sqrt{\hat{k}_{\mathbf{n}}}} + \sum_{i=2}^{2^N} \frac{T_{\mathbf{n},\gamma,i}}{\overline{\sigma}\sqrt{\hat{k}_{\mathbf{n}}}}.$$

In summary, equation (3.47) ensures that the random variables $U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right)$ are asymptotically independent. The asymptotic normality of $T_{\mathbf{n},\gamma,1}/\left(\overline{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right)$ follows from (3.49) and the Lindeberg-Feller condition (3.50). Equation (3.48) shows that the term $\sum_{i=2}^{2^N} T_{\mathbf{n},\gamma,i}/\left(\overline{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible. Now, let us prove (3.47)-(3.50).

Let us begin by (3.47).
Enumerate the random variables $U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right)$ in an arbitrary order say, $U_1^*,\cdots,U_{\hat{r}}^*$, where $\hat{r} = \prod_{k=1}^{N} r_k = \hat{\mathbf{n}}(p+q)^{-N} \leq \hat{\mathbf{n}}p^N$. Let

$$\mathcal{S}_{\mathbf{n}}(1,\gamma,\mathbf{j}) = \{i : j_k(q+p)+1 \leq i_k \leq j_k(q+p)+p,\ \ k=1,\cdots,N\}$$

be the sets of distinct sites that are distant from a distance at least q. Clearly $\mathcal{S}_{\mathbf{n}}(1,\gamma,\mathbf{j})$ contains p^N sites, where $\mathcal{S}_{\mathbf{n}}(1,\gamma,\mathbf{j})$ is the set of sites involved with $U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right)$. By Lemma 3.4.5, we have that :

$$\begin{aligned} Q_{1,\mathbf{n},\gamma} \leq \sum_{k=1}^{\hat{r}-1}\sum_{j=k+1}^{\hat{r}} &\left|\mathbf{E}\left[(\exp\left\{iuU_k^*\right\}-1)(\exp\left\{iuU_j^*\right\}-1)\right]\prod_{s=j+1}^{\hat{r}}\exp\left\{iuU_s^*\right\}\right. \\ &\left. -\,\mathbf{E}\left[(\exp\left\{iuU_k^*\right\}-1)\right]\mathbf{E}\left[(\exp\left\{iuU_j^*\right\}-1)\right]\prod_{s=j+1}^{\hat{r}}\exp\left\{iuU_s^*\right\}\right|. \end{aligned}$$

Let $\widetilde{\mathcal{S}}_j$ be the sets of sites involved with U_j^*. Using Lemma 3.4.5, it follows that

$$\begin{aligned}&\left|\mathbf{E}\left[(\exp\{iuU_k^*\}-1)(\exp\left\{iuU_j^*\right\}-1)\right]-\mathbf{E}\left[(\exp\{iuU_k^*\}-1)\right]\mathbf{E}\left[(\exp\left\{iuU_j^*\right\}-1)\right]\right|\\ &\quad\leq Cp^N\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_j,\widetilde{\mathcal{S}}_k)\right)\end{aligned}$$

thus

$$\begin{aligned}Q_{1,\mathbf{n},\gamma}&\leq Cp^N\sum_{k=1}^{\hat{r}-1}\sum_{j=k+1}^{\hat{r}}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_j,\widetilde{\mathcal{S}}_k)\right)\\ &\leq Cp^N\hat{r}\sum_{k=2}^{\hat{r}}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)\right)\\ &\leq Cp^N\hat{r}\sum_{i=1}^{\infty}\sum_{k:iq\leq \mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)<(i+1)q}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)\right)\\ &\leq Cp^N\hat{r}\sum_{i=1}^{\infty}i^{N-1}\varphi(iq)\leq C\hat{\mathbf{n}}\sum_{i=}^{\infty}i^{N-1}\varphi(iq)\\ &\leq C\hat{\mathbf{n}}\sum_{i=}^{\infty}i^{N-1-\theta}q^{-\theta}\leq C\hat{\mathbf{n}}q^{-\theta},\end{aligned}$$

this tends to zero if one take q such that $\hat{\mathbf{n}}q^{-\theta}\to 0$.■
To prove (3.48) it suffices to show that :

$$\hat{k}_{\mathbf{n}}^{-1}\mathbf{E}\left[T_{\mathbf{n},\gamma,i}\right]^2\longrightarrow 0\quad\text{for each}\quad 2\leq i\leq 2^N.$$

Without loss of generality, we only consider the case where $i=2$. As in the proof of (3.47), enumerate the random varible's $U_{\mathbf{n}}(2,\gamma,\mathbf{j})$ in an arbitrary way and refer them as $U_1^+,\cdots,U_{\hat{Q}}^+$. Then

$$\begin{aligned}\mathbf{E}\left[T_{\mathbf{n},\gamma,2}\right]^2&=\sum_{i=0}^{\hat{r}}\operatorname{Var}\left(U_i^+\right)+2\sum_{j=0}^{\hat{r}-1}\sum_{i=j+1}^{\hat{r}}\operatorname{Cov}\left(U_i^+,U_j^+\right)\\ &:=\quad L_{1,\gamma}\quad+\quad L_{2,\gamma}\end{aligned}\tag{3.51}$$

Since the $X_{\mathbf{i}}$ are of same law as X

$$\begin{aligned}\operatorname{Var}\left(U_i^+\right)&=\operatorname{Var}\left(\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}\Delta_{\mathbf{i},\mathbf{n}}\right)^2\\ &:=p^{N-1}q\operatorname{Var}(\Delta_{\mathbf{i},\mathbf{n}})+\underset{i_k\neq j_k\text{ for some }1\leq k\leq N}{\sum_{\substack{j_k=1\\k=1,\cdots,N-1}}^{p}\sum_{j_N=1}^{q}\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}}\mathbf{E}\left[\Delta_{\mathbf{i},\mathbf{n}}\Delta_{\mathbf{j},\mathbf{n}}\right]\end{aligned}\tag{3.52}$$

Recall that $\Delta_{\mathbf{i},\mathbf{n}}=\eta_{\mathbf{i},\mathbf{n}}-\mathbb{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right]$ and $\eta_{\mathbf{i},\mathbf{n}}=\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_+$, and combining with (3.37), it follows that

$$\operatorname{Var}(\Delta_{\mathbf{i},\mathbf{n}})=\mathbf{E}\left[\Delta_{\mathbf{i},\mathbf{n}}^2\right]\leq\gamma^2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\left(2-\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\right)\leq C.\tag{3.53}$$

Let $\chi = 2(1-\varsigma)/\varsigma$, $0<\varsigma<1$, now using the second statement of Lemma 3.4.4 and (3.37) with the Lebesgue theorem, we have that

$$\begin{aligned}\mathbf{E}\left|\Delta_{\mathbf{i},\mathbf{n}}\Delta_{\mathbf{j},\mathbf{n}}\right| &\leq C\left(\int_{\mathbb{R}}\left|\log\left(u/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_{+}\right|^{2+\chi} f(u)du\right)^{\varsigma}\varphi\left(\|\mathbf{i}\|\right)^{1-\varsigma}\\ &\leq C\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\gamma^{2}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1-\varsigma}. \qquad (3.54)\end{aligned}$$

Combining (3.52), (3.53) and (3.54), it follows that

$$\begin{aligned}\operatorname{Var}\left(U_{i}^{+}\right) &\leq Cp^{N-1}q\left(1+\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\gamma^{2}\sum_{\substack{i_k=1\\ k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}\varphi\left(\|\mathbf{i}\|\right)^{1-\varsigma}\right)\\ &\leq Cp^{N-1}q\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\gamma^{2}\sum_{\substack{i_k=1\\ k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1-\varsigma}\\ &\leq Cp^{N-1}q\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\gamma^{2}\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}. \qquad (3.55)\end{aligned}$$

By (3.51)-(3.55), it follows that :

$$L_{1,\gamma} \leq C\hat{r}p^{N-1}q\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\gamma^{2}\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}. \qquad (3.56)$$

Let

$$\begin{aligned}\mathcal{S}_{\mathbf{n}}(2,\gamma,\mathbf{j}) \quad = \quad &\{\mathbf{i}: j_k(p+q+1)\leq i_k\leq j_k(p+q)+p,\ 1\leq k<N\\ &j_k(p+q)+p+1\leq i_N\leq (j_N+1)(p+q)\}\end{aligned}$$

Then $U_{\mathbf{n}}(2,\gamma,\mathbf{j})$ is the sum of $\Delta_{\mathbf{i},\mathbf{n}}$ with sites in $\mathcal{S}_{\mathbf{n}}(2,\gamma,\mathbf{j})$. Since $p>q$, if $\mathbf{j}$ and $\mathbf{j}'$ belong to two distinct sets $\mathcal{S}_{\mathbf{n}}(2,\gamma,\mathbf{j})$ and $\mathcal{S}_{\mathbf{n}}(2,\gamma,\mathbf{j}')$, then $j_k \neq j_k^{'}$ for some $1\leq k\leq N$ and $\|\mathbf{j}-\mathbf{j}'\|>q$. Then by (3.54), we have

$$\begin{aligned}L_{2,\gamma} &\leq C\sum_{\substack{j_k=1\\ k=1,\cdots,N}}^{n_k}\sum_{\substack{i_k=1\\ k=1,\cdots,N\\ \|\mathbf{i}-\mathbf{j}\|>q}}^{n_k}\mathbf{E}\left[\Delta_{\mathbf{i},\mathbf{n}}\Delta_{\mathbf{j},\mathbf{n}}\right]\\ &\leq C\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\gamma^{2}\hat{\mathbf{n}}\sum_{\substack{i_k=1\\ k=1,\cdots,N;\ \|\mathbf{i}\|>q}}^{n_k}\varphi(\|\mathbf{i}\|)^{1-\varsigma}\\ &\leq C\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}}\gamma^{2}\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}. \qquad (3.57)\end{aligned}$$

From (3.51), (3.56) and (3.57), we get

$$\begin{aligned}
\hat{k}_{\mathbf{n}}^{-1}\mathbf{E}\left[T_{\mathbf{n},\gamma,2}\right]^2 &\leq C\hat{r}p^{N-1}q\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}^2}\gamma^2\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}+C\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma^2\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}\\
&\leq C\gamma^2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\left(\hat{\mathbf{n}}^{-1}\hat{r}p^{N-1}q\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}+\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}\right)\\
&\leq C\gamma^2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\left(\frac{q}{p}\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}+\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}\right)\\
&\leq C\gamma^2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\left(\frac{q}{p}\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}+\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}\right)\\
&\leq C\gamma^2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\left(\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}+\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}\right)\leq C\gamma^2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\sum_{i=1}^{\infty}i^{N-1-\theta(1-\varsigma)}.
\end{aligned}$$

This tends to zero it suffices to take $0<\varsigma<\frac{\theta-N}{\theta}$. ■
To prove (3.49), let us denote by

$$K_{\mathbf{n},\gamma,1}=T_{\mathbf{n},\gamma,1}\quad\text{and}\quad K_{\mathbf{n},\gamma,2}=\sum_{i=2}^{2^N}T_{\mathbf{n},\gamma,i}. \tag{3.58}$$

$K_{\mathbf{n},\gamma,1}$ is the sum of random variable's $\Delta_{\mathbf{j},\mathbf{n}}$ in large blocks and $K_{\mathbf{n},\gamma,2}$ is that of random variable's in small blocks. Lemma 3.5.1, statement 1 implies that $\hat{k}_{\mathbf{n}}^{-1}\mathbf{E}\left[S_{\mathbf{n},\gamma}^2\right]\to\overline{\sigma}^2=\gamma^2$. This combining with (3.48) shows that $\hat{k}_{\mathbf{n}}^{-1}\mathbf{E}\left[K_{\mathbf{n},\gamma,1}^2\right]\longrightarrow\overline{\sigma}^2$. Now

$$\begin{aligned}
\hat{k}_{\mathbf{n}}^{-1}\mathbf{E}\left[K_{\mathbf{n},\gamma,1}^2\right]&=\hat{k}_{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbf{E}\left[U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right)\right]^2\\
&+\hat{k}_{\mathbf{n}}^{-1}\sum_{\substack{j_k=1\\k=1,\cdots,N\\i_k\neq j_k\text{ for some }1\leq k\leq N}}^{r_k-1}\sum_{\substack{i_k=1\\k=1,\cdots,N}}^{r_k-1}\operatorname{Cov}\left(U_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right),U_{\mathbf{n}}\left(1,\gamma,\mathbf{i}\right)\right).
\end{aligned}\tag{3.59}$$

Observe that (3.49) follows from (3.59) if the last term of (3.59) tends to zero as $\mathbf{n}\to\infty$. By the same argument used to obtain a bound for $L_{2,\gamma}$ (see (3.51)), the last term of (3.59) is bounded by

$$C\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma^2\sum_{\substack{i_k=1\\k=1,\cdots,N;\ \|\mathbf{i}\|>q}}^{n_k}\varphi(\|\mathbf{i}\|)^{1-\varsigma}\leq C\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma^2\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}. \tag{3.60}$$

this last tends to zero as above.■
Prove now (3.50).

First, we have that : $\mathbf{E}\left[\Delta_{\mathbf{i},\mathbf{n}}^2\right]\leq2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma^2$, $\mathbf{E}\left[\Delta_{\mathbf{i},\mathbf{n}}\right]\leq\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma$. Note that by Markov, inequality we have

$$\mathbb{P}\left[\left|U_{\mathbf{n}}\left(1,x,\mathbf{j}\right)\right|>p^N\left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)^{1/2}\right]\leq\frac{p^N\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}\gamma}{p^N\left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)^{1/2}}=\gamma\left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)^{-3/2}\to0.$$

Therefore for $\mathbf{n}$ large enough, $U_{\mathbf{n}}(1,x,\mathbf{j}) < p^N \left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)^{1/2}$, and thus

$$\mathbf{E}\left[U_{\mathbf{n}}(1,x,\mathbf{j})^2 \mathbb{1}_{\left\{|U_{\mathbf{n}}(1,x,\mathbf{j})|>\varepsilon\overline{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right\}}\right] \leq p^{2N}\left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\right)\mathbb{P}\left[|U_{\mathbf{n}}(1,x,\mathbf{j})| > \varepsilon\overline{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right].$$

By (3.48), we have $E\left(U_{\mathbf{n}}(1,x,\mathbf{j})^2\right) \sim \overline{\sigma}^2\frac{\hat{k}_{\mathbf{n}}}{\hat{r}}$ for $\mathbf{n}$ large enough, then using Tchebychev inequality

$$\mathbb{P}\left[|U_{\mathbf{n}}(1,x,\mathbf{j})| > \varepsilon\overline{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right] \leq \frac{\overline{\sigma}^2\frac{\hat{k}_{\mathbf{n}}}{\hat{r}}}{\overline{\sigma}^2\hat{k}_{\mathbf{n}}\varepsilon^2} = \hat{r}_{\mathbf{n}}^{-1}\varepsilon^{-2}.$$

Therefore

$$Q_{4,\mathbf{n},\gamma} \leq C\frac{p^{2N}\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^2\varepsilon^2} \longrightarrow 0 \text{ by (3.13).}\blacksquare \tag{3.61}$$

► Now we are going to give the asymptotic normality of the vector $\mathbf{H}_{\mathbf{n}}^{(\tau)}$. The variance of $\mathbf{H}_{\mathbf{n}}^{(\tau)}$ is obtained in the same manner as the variance of $\mathbf{H}_{\mathbf{n}}^{+}$, it suffices to note that $\mathbf{H}_{\mathbf{n}}^{+}$ has the same writing as $\mathbf{H}_{\mathbf{n}}^{(\tau)}$. Recall the notations given at the beginning of this section, and note that

$$\mathbf{H}_{\mathbf{n}}^{(\tau)} = \hat{k}_{\mathbf{n}}^{1/2}S_{\mathbf{n},\tau} \text{ where } S_{\mathbf{n},\tau} := \hat{k}_{\hat{\mathbf{n}}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n},\tau},$$
$$\Delta_{\mathbf{i},\mathbf{n},\tau} = \eta_{\mathbf{i},\mathbf{n},\tau} - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}\right] \text{ and } \eta_{\mathbf{i},\mathbf{n},\tau} = \mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}.$$

Using again (3.13), one can find a sequence of positive integers $p_{\mathbf{n}}$ such that $p_{\mathbf{n}} = p$, $qp^{-1} = o(1)$ which tends to zero as $\mathbf{n} \to \infty$. As above, let $r_1, \cdots, r_N$, and $n_1 = r_1(q+p), \cdots, n_N = r_N(q+p)$.
Then, the random variables $\Delta_{\mathbf{i},\mathbf{n},\tau}$'s are now set into large blocks and small blocks. From now, the asymptotic normality of $\mathbf{H}_{\mathbf{n}}^{(\tau)}$ follows the same lines as the proof of the asymptotic normality of $\mathbf{H}_{\mathbf{n}}^{+}$, that is

$$\begin{aligned} U_{\mathbf{n}}(1,\tau,\mathbf{j}) &= \sum_{\mathbf{i}\in I(1,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n},\tau},\ U_{\mathbf{n}}(2,\tau,\mathbf{j}) = \sum_{I(2,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n},\tau} \\ U_{\mathbf{n}}(3,\tau,\mathbf{j}) &= \sum_{I(3,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n},\tau},\ U_{\mathbf{n}}(4,\tau,\mathbf{j}) = \sum_{I(4,\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n},\tau} \end{aligned}$$

and so on, with

$$U_{\mathbf{n}}\left(2^{N-1},\tau,\mathbf{j}\right) = \sum_{I(2^{N-1},\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n},\tau},\ U_{\mathbf{n}}\left(2^{N},\tau,\mathbf{j}\right) = \sum_{I(2^{N},\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n},\tau}.$$

For each integer $1 \leq i \leq 2^N$, define

$$T_{\mathbf{n},\tau,i} = \textstyle\sum_{\mathbf{j}\in D(i)}U_{\mathbf{n}}(i,\tau,\mathbf{j}). \tag{3.62}$$

Clearly $K_{\mathbf{n},\tau} = \sum_{i=1}^{2^N}T_{\mathbf{n},\tau,i}$. Using again Lemma 3.5.1, the general approach is to show that as $\mathbf{n} \to \infty$,

$$Q_{1,\mathbf{n},\tau} \equiv \left| \mathbf{E}\left[\exp\left(izT_{\mathbf{n},\tau,1}\right)\right] - \prod_{\substack{j_k=0 \\ k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[\exp\left(izU_{\mathbf{n}}\left(1,\tau,\mathbf{j}\right)\right)\right] \right| \to 0, \tag{3.63}$$

$$Q_{2,\mathbf{n},\tau} \equiv \hat{k}_{\mathbf{n}}^{-1}\mathbb{E}\left[\left(\sum_{i=2}^{2^N} T_{\mathbf{n},\tau,i}\right)^2\right] \to 0, \tag{3.64}$$

$$Q_{3,\mathbf{n},\tau} \equiv \hat{k}_{\mathbf{n}}^{-1} \sum_{\substack{j_k=0 \\ k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[U_{\mathbf{n}}\left(1,\tau,\mathbf{j}\right)^2\right] \to \widetilde{\sigma}^2, \quad \text{where} \quad \widetilde{\sigma}^2 = 1 + o(1) \tag{3.65}$$

$$Q_{4,\mathbf{n},\tau} \equiv \hat{k}_{\mathbf{n}}^{-1} \sum_{\substack{j_k=0 \\ k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[U_{\mathbf{n}}\left(1,\tau,\mathbf{j}\right)^2 \mathbb{1}_{\left\{\left|U_{\mathbf{n}}(1,\tau,\mathbf{j})\right| \geq \varepsilon\widetilde{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right\}}\right] \to 0 \quad \text{for every } \varepsilon > 0. \tag{3.66}$$

Note that

$$\frac{S_{\mathbf{n},\tau}}{\widetilde{\sigma}\sqrt{\hat{k}_{\mathbf{n}}}} = \frac{T_{\mathbf{n},\tau,1}}{\widetilde{\sigma}\sqrt{\hat{\mathbf{n}}}} + \sum_{i=2}^{2^N} \frac{T_{\mathbf{n},\tau,i}}{\widetilde{\sigma}\sqrt{\hat{k}_{\mathbf{n}}}}.$$

In summary, equation (3.63) ensures that the random variables $U_{\mathbf{n}}\left(1,\tau,\mathbf{j}\right)$ are asymptotically independent. The asymptotic normality of $T_{\mathbf{n},\tau,1}/\left(\widetilde{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right)$ follows from (3.65) and the Lindeberg-Feller condition (3.66). Equation (3.64) shows that the term $\sum_{i=2}^{2^N} T_{\mathbf{n},\tau,i}/\left(\widetilde{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible. The proofs of (3.63)-(3.66) follow the same lines as above and therefore are omitted.

► Let us prove that $\mathbf{H}_{\mathbf{n}} := \left(\mathbf{H}_{\mathbf{n}}^{(\tau)}, \mathbf{H}_{\mathbf{n}}^{+}\right)^{\top}$ converges in distribution to $\mathcal{N}\left(0,\Sigma\right)$. According to Cramèr-Wold's device (see Van der Vaart (1998)), it is sufficient to prove that $l^{\top}\mathbf{H}_{\mathbf{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, l^{\top}\Sigma l\right)$ for all $l = (l_1,l_2)^{\top} \in \mathbb{R}^2,\ l \neq 0$. Some simple calculation yields :

$$\begin{aligned}
l^{\top}\mathbf{H}_{\mathbf{n}} &:= \hat{k}_{\mathbf{n}}^{1/2}\,\frac{1}{\hat{k}_{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n}}^{\dagger}, \\
\Delta_{\mathbf{i},\mathbf{n}}^{\dagger} &:= l_1\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) > \tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}} + l_2\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_{+} \\
&\qquad - \mathbf{E}\left[l_1\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right) > \tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}} + l_2\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_{+}\right] \\
&:= A_{\mathbf{i},\mathbf{n},\gamma,\tau} - \mathbf{E}\left[A_{\mathbf{i},\mathbf{n},\gamma,\tau}\right].
\end{aligned}$$

Thus, to establish the asymptotic normality of $\mathbf{H}_{\mathbf{n}}$, we verify the Lindeberg-Feller condition for triangular arrays. We first calculate the variance of $\mathbf{H}_{\mathbf{n}}$. We have

$$\operatorname{Var}\left(\hat{k}_{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n}}^{\dagger}\right) = \hat{k}_{\mathbf{n}}^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\operatorname{Var}\left(\Delta_{\mathbf{i},\mathbf{n}}^{\dagger}\right) + 2\hat{k}_{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\sum_{\mathbf{i}\neq\mathbf{j}}\operatorname{Cov}\left(\Delta_{\mathbf{i},\mathbf{n}}^{\dagger},\Delta_{\mathbf{j},\mathbf{n}}^{\dagger}\right).$$

Denote by $I_{\mathbf{n}}^{\dagger} := \hat{k}_{\mathbf{n}}^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\operatorname{Var}\left(A_{\mathbf{i},\mathbf{n},\gamma,\tau}\right)$ and by $R_{\mathbf{n}}^{\dagger} := \hat{k}_{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\sum_{\mathbf{i}\neq\mathbf{j}}\left|\operatorname{Cov}\left(A_{\mathbf{i},\mathbf{n},\gamma,\tau},A_{\mathbf{j},\mathbf{n},\gamma,\tau}\right)\right|$.

Clearly, $\mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\right)\le I^{\dagger}_{\mathbf{n}}+R^{\dagger}_{\mathbf{n}}$, with

$$\begin{aligned}
I^{\dagger}_{\mathbf{n}} &= \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^{2}}\left(\mathbf{E}\left[A^{2}_{\mathbf{i},\mathbf{n},\gamma,\tau}\right]-\mathbf{E}\left[A_{\mathbf{in},\gamma,\tau}\right]^{2}\right)\\
&= \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^{2}}\mathbf{E}\left[\left(l_1\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}+l_2\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_{+}\right)^{2}\right]\\
&\quad -\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^{2}}\mathbf{E}\left[l_1\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\tau/\sqrt{\hat{k}_{\mathbf{n}}}\right\}}+l_2\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)_{+}\right]^{2}.\\
&= \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^{2}}l_1^{2}\left(\mathbf{E}\left[\eta^{2}_{\mathbf{i},\mathbf{n},\tau}\right]-\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}\right]^{2}\right)+l_2^{2}\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^{2}}\left(\mathbf{E}\left[\eta^{2}_{\mathbf{i},\mathbf{n}}\right]-\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right]^{2}\right)\\
&\quad +2l_1l_2\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^{2}}\left(\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}\eta_{\mathbf{i},\mathbf{n}}\right]-\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},\tau}\right]\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right]\right).
\end{aligned}$$

From (3.37) and Lemma 3.5.1, some calculations yields

$$\begin{aligned}
I^{\dagger}_{\mathbf{n}} &= \frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}^{2}}\left\{l_1^{2}\left(\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}-\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\right)+2l_1l_2\gamma\left(\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}-\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\right)+l_2^{2}\gamma\left(2\frac{\hat{k}_{\mathbf{n}}}{\hat{\mathbf{n}}}-\frac{\hat{k}_{\mathbf{n}}^{2}}{\hat{\mathbf{n}}^{2}}\right)\right\}\\
&= \left\{l_1^{2}\left(\hat{k}_{\mathbf{n}}^{-1}-\hat{\mathbf{n}}^{-1}\right)+2l_1l_2\gamma\left(\hat{k}_{\mathbf{n}}^{-1}-\hat{\mathbf{n}}^{-1}\right)+l_2^{2}\gamma^{2}\left(2\hat{k}_{\mathbf{n}}^{-1}-\hat{\mathbf{n}}^{-1}\right)\right\} \qquad (3.67)
\end{aligned}$$

Now let us examine $R^{\dagger}_{\mathbf{n}}$, simple calculations yields

$$\begin{aligned}
\mathrm{Cov}\left(A_{\mathbf{i},\mathbf{n},\tau,\gamma},A_{\mathbf{j},\mathbf{n},\tau,\gamma}\right) &= l_1^{2}\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau},\eta_{\mathbf{j},\mathbf{n},\tau}\right)+l_2^{2}\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)\\
&\quad +l_1l_2\left\{\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n},\tau}\right)+\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n},\tau},\eta_{\mathbf{j},\mathbf{n}}\right)\right\}.
\end{aligned}$$

Using the same method as in Lemma 3.5.1, it follows

$$R^{\dagger}_{\mathbf{n}}=O\left(\hat{k}_{\mathbf{n}}^{-1}\right). \qquad (3.68)$$

Then combining (3.67) and (3.68), we get

$$\begin{aligned}
\lim_{\mathbf{n}\to+\infty}\mathrm{Var}\left(l^{\top}\mathbf{H}_{\mathbf{n}}\right) &= \lim_{\mathbf{n}\to+\infty}\hat{k}_{\mathbf{n}}\mathrm{Var}\left(\hat{k}_{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\right)\\
&= \lim_{\mathbf{n}\to+\infty}\hat{k}_{\mathbf{n}}\ I^{\dagger}_{\mathbf{n}}(x)=l_1^{2}+2l_1l_2\gamma+2l_2^{2}\gamma^{2}\\
&= l^{\top}\Sigma l.
\end{aligned}$$

To complete the proof of the asymptotic normality of $l^{\top}\mathbf{H}_{\mathbf{n}}$ let us use same notations as above. Denote

$$S^{\dagger}_{\mathbf{n}}:=\hat{k}_{\mathbf{n}}^{-1}\sum_{i\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\quad\text{then}\quad l^{\top}\mathbf{H}_{\mathbf{n}}=\hat{k}_{\mathbf{n}}^{1/2}S^{\dagger}_{\mathbf{n}}.$$

Using the same notations as in Lemma 3.5.2, the random variables $\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}$'s are set into large blocks and small blocks, that is

$$\begin{aligned}
U^{\dagger}_{\mathbf{n}}\left(1,\gamma,\mathbf{j}\right) &= \sum_{\mathbf{i}\in I(1,\mathbf{j})}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}},\ U^{\dagger}_{\mathbf{n}}\left(2,\gamma,\mathbf{j}\right)=\sum_{I(2,\mathbf{j})}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\\
U^{\dagger}_{\mathbf{n}}\left(3,\gamma,\mathbf{j}\right) &= \sum_{I(3,\mathbf{j})}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}},\ U^{\dagger}_{\mathbf{n}}\left(4,\gamma,\mathbf{j}\right)=\sum_{I(4,\mathbf{j})}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}
\end{aligned}$$

and so on,

$$U_{\mathbf{n}}^{\dagger}\left(2^{N-1},\gamma,\mathbf{j}\right) = \sum_{I(2^{N-1},\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}}^{\dagger},\ U_{\mathbf{n}}^{\dagger}\left(2^{N},\gamma,\mathbf{j}\right)=\sum_{I(2^{N},\mathbf{j})}\Delta_{\mathbf{i},\mathbf{n}}^{\dagger}$$

For each integer $1 \leq i \leq 2^N$, define

$$T_{\mathbf{n},\gamma,i}^{\dagger} = \textstyle\sum_{\mathbf{j}\in D(i)} U_{\mathbf{n}}^{\dagger}\left(i,\gamma,\mathbf{j}\right). \tag{3.69}$$

We have $S_{\mathbf{n}}^{\dagger} = \sum_{i=1}^{2^N} T_{\mathbf{n},\gamma,i}^{\dagger}$. Using Lemma 3.5.1, it suffices to show that as $\mathbf{n}\to\infty$,

$$Q_{1,\mathbf{n},\gamma}^{\dagger} \equiv \left|\mathbb{E}\left[\exp\left(izT_{\mathbf{n},\gamma,1}^{\dagger}\right)\right]-\prod_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbb{E}\left[\exp\left(izU_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j}\right)\right)\right]\right|\to 0, \tag{3.70}$$

$$Q_{2,\mathbf{n},\gamma}^{\dagger} \equiv \hat{k}_{\mathbf{n}}^{-1}\mathbb{E}\left[\left(\sum_{i=2}^{2^N}T_{\mathbf{n},\gamma,i}^{\dagger}\right)^2\right]\to 0, \tag{3.71}$$

$$Q_{3,\mathbf{n},\gamma}^{\dagger} \equiv \hat{k}_{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbb{E}\left[U_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j}\right)^2\right]\to\widehat{\sigma}^2, \tag{3.72}$$

$$Q_{4,\mathbf{n},\gamma}^{\dagger} \equiv \hat{k}_{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbb{E}\left[U_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j}\right)^2\mathbb{1}_{\left\{\left|U_{\mathbf{n}}^{\dagger}(1,\gamma,\mathbf{j})\right|\geq\varepsilon\widehat{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right\}}\right]\to 0 \quad \text{for every } \varepsilon>0 \tag{3.73}$$

Note that

$$\frac{S_{\mathbf{n}}^{\dagger}}{\widehat{\sigma}\sqrt{\hat{k}_{\mathbf{n}}}}=\frac{T_{\mathbf{n},\gamma,1}^{\dagger}}{\widehat{\sigma}\sqrt{\hat{k}_{\mathbf{n}}}}+\sum_{i=2}^{2^N}\frac{T_{\mathbf{n},\gamma,i}^{\dagger}}{\widehat{\sigma}\sqrt{\hat{k}_{\mathbf{n}}}},$$

where $\hat{\sigma}^2 = l_1^2 + 2l_1l_2\gamma + 2l_2^2\gamma^2$.
In summary, equation (3.70) ensures that the random variables $U_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j}\right)$ are asymptotically independent. The asymptotic normality of $T_{\mathbf{n},\gamma,1}^{\dagger}/\left(\widehat{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right)$ follows from (3.72) and the Lindeberg-Feller condition (3.73). Equation (3.71) shows that the term $\sum_{i=2}^{2^N}T_{\mathbf{n},\gamma,i}^{\dagger}\left(x,i\right)/\left(\widehat{\sigma}\hat{k}_{\mathbf{n}}^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible. The proofs of (3.70)-(3.73) follow the same lines as above and are therefore omitted. ■

3.5.3 Appendix A23 : Asymptotic normality

Proof of Theorem 3.3.4.

First observe that

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}-\gamma\right)=\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)+\sqrt{\hat{k}_{\mathbf{n}}}\left(\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]-\gamma\right)$$

$$\text{and } \sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)=\sqrt{\hat{k}_{\mathbf{n}}}\left(\widetilde{\gamma}_{\mathbf{n}}-\log\left(X_{(\hat{k}_{\mathbf{n}}+1),n}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right),$$

due to the fact that $\gamma_{\mathbf{n}}=\widetilde{\gamma}_{\mathbf{n}}-\log\left(X_{(\hat{k}_{\mathbf{n}}+1),n}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)$, see Appendix A22. Then, to prove the asymptotic normality of $\gamma_{\mathbf{n}}$, it suffices (see Hsing (1991)) to prove that there exists a random vector (V_1, V_2) such that

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma\left(I_{\mathbf{n}}^{(\tau)}-EI_{\mathbf{n}}^{(\tau)}\right),\left(\gamma_{\mathbf{n}}^{+}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\right)^{\top}\to(V_1,V_2) \tag{3.74}$$

and

$$\frac{1}{\hat{k}_{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\left(Y_{\mathbf{i},\mathbf{n}}-\mathbf{E}\left[Y_{\mathbf{i},\mathbf{n}}\right]\right)\xrightarrow{\mathbb{P}}0. \tag{3.75}$$

We have (see Hsing (1991))

$$\begin{aligned}
&\lim_{\mathbf{n}\to\infty}P\left[\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}^{+}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\leq x,\ \sqrt{\hat{k}_{\mathbf{n}}}\left(\log\left(X_{(\hat{k}_{\mathbf{n}}+1),n}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\right)\leq\tau\right] && (3.76)\\
&=\lim_{\mathbf{n}\to\infty}P\left[\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}^{+}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\leq x,\ \gamma\sqrt{\hat{k}_{\mathbf{n}}}\left(\left(I_{\mathbf{n}}^{(\tau)}-EI_{\mathbf{n}}^{(\tau)}\right)\right)\leq\tau+o(1)\right] && (3.77)\\
& && (3.78)
\end{aligned}$$

Then, if (3.74) and (3.75) hold, the last limit implies that

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\log\left(X_{(\hat{k}_{\mathbf{n}}+1),n}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right),\left(\gamma_{\mathbf{n}}^{+}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\right)^{\top}\to(V_1,V_2). \tag{3.79}$$

By Lemma 3.5.3 (see Lemma 2.1 and Theorem 3.3 of Hsing (1991)), we have

$$\begin{aligned}
\left|\gamma_{\mathbf{n}}-\tilde{\gamma}_{\mathbf{n}}\right|+\left|\gamma_{\mathbf{n}}-\gamma_{\mathbf{n}}^{+}\right|+\left|\tilde{\gamma}_{\mathbf{n}}-\gamma_{\mathbf{n}}^{+}\right|&\xrightarrow{\mathbb{P}}0, && (3.80)\\
\sqrt{\hat{k}_{\mathbf{n}}}\left(\tilde{\gamma}_{\mathbf{n}}-\gamma_{\mathbf{n}}^{+}\right)&\xrightarrow{\mathbb{P}}0 && (3.81)
\end{aligned}$$

Thus, if (3.74) and (3.75) hold, the last previous results imply that

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\log\left(X_{(\hat{k}_{\mathbf{n}}+1),n}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right),\left(\tilde{\gamma}_{\mathbf{n}}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\right)^{\top}\to(V_1,V_2). \tag{3.82}$$

and by continuous mapping theorem

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\to(V_2-V_1). \tag{3.83}$$

Note that (3.74) holds since by Lemma 3.5.2

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma\left(I_{\mathbf{n}}^{(\tau)}-EI_{\mathbf{n}}^{(\tau)}\right),\left(\gamma_{\mathbf{n}}^{+}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\right)^{\top}\to\mathcal{N}\left((0,0)^{\top},\Sigma\right) \tag{3.84}$$

where

$$\Sigma:=\begin{pmatrix}1&\gamma\\ \gamma&2\gamma^2\end{pmatrix}.$$

Thus by a straightforward application of the Delta-method, we have

$$\sqrt{\hat{k}_{\mathbf{n}}}\left(\gamma_{\mathbf{n}}-\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]\right)\to\mathcal{N}\left(0,\gamma^2\right). \tag{3.85}$$

Now observe that

$$\begin{aligned}
\sqrt{\hat{k}_{\mathbf{n}}}\left(\mathbf{E}\left[\gamma_{\mathbf{n}}^{+}\right]-\gamma\right) &= \sqrt{\hat{k}_{\mathbf{n}}}\left(\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}\overline{F}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\int_{0}^{\infty}\frac{\overline{F}\left(yb\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)}{\overline{F}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)}\frac{dy}{y}-\gamma\right)\\
&= \sqrt{\hat{k}_{\mathbf{n}}}\left(\int_{1}^{\infty} y^{-\gamma^{-1}-1}\left(\frac{L\left(yb\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)}{L\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)}-1+1\right)dy-\gamma\right)\\
&= \sqrt{\hat{k}}\left(\int_{1}^{\infty} y^{-\gamma^{-1}-1}\left(\mathcal{A}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\frac{\frac{L\left(yb\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)}{L\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)}-1}{\mathcal{A}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)}+1\right)dy-\gamma\right)\\
&= \sqrt{\hat{k}_{\mathbf{n}}}\left(\int_{1}^{\infty}\left(\mathcal{A}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)y^{-\gamma^{-1}-1}\frac{y^{\rho}-1}{\rho}-y^{-\gamma^{-1}-1}\right)dy-\gamma\right)\\
&= \rho^{-1}\sqrt{\hat{k}_{\mathbf{n}}}\mathcal{A}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\int_{1}^{\infty}\left(y^{-\gamma^{-1}+\rho-1}-y^{-\gamma^{-1}-1}\right)dy\\
&\quad+\sqrt{\hat{k}_{\mathbf{n}}}\left(\int_{1}^{\infty} y^{-\gamma^{-1}-1}dy-\gamma\right)\\
&= \rho^{-1}\sqrt{\hat{k}_{\mathbf{n}}}\mathcal{A}\left(b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)\left(\frac{1}{\gamma^{-1}-\rho}-\frac{1}{\gamma^{-1}}\right)\longrightarrow\frac{\lambda\gamma}{\gamma^{-1}-\rho}\quad\text{as}\quad\mathbf{n}\to\infty.
\end{aligned}$$

To achieve the proof, it suffices to note that $I_{\mathbf{n}}^{(\tau)}-\mathbf{E}\left[I_{\mathbf{n}}^{(\tau)}\right]\xrightarrow{\mathbb{P}}0$ which amounts to prove (3.75). This last is provided from Lemma 3.5.3. ■

Lemma 3.5.3. *Under conditions of Theorem 3.3.4, we have*

$$\frac{1}{\hat{k}_{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\left(\mathrm{Y}_{\mathbf{i},\mathbf{n}}-\mathbf{E}\left[\mathrm{Y}_{\mathbf{i},\mathbf{n}}\right]\right)\xrightarrow{\mathbb{P}}0,$$

where $\mathrm{Y}_{\mathbf{i},\mathbf{n}}=\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\varepsilon\right\}}$ *for all real* ε.

Proof of Lemma 3.5.3

Let

$$I_{\mathbf{i},\mathbf{n}}(\varepsilon)=\hat{k}_{\mathbf{n}}^{-1}\mathbb{1}_{\left\{\log\left(X_{\mathbf{i}}/b\left(\hat{\mathbf{n}}/\hat{k}_{\mathbf{n}}\right)\right)>\varepsilon\right\}},\qquad\text{and}\;\;\Delta_{\mathbf{i},\mathbf{n}}(\varepsilon)=I_{\mathbf{i},\mathbf{n}}(\varepsilon)-\mathbf{E}\left[I_{\mathbf{i},\mathbf{n}}(\varepsilon)\right].$$

Then

$$S_{\mathbf{n}}(\varepsilon)=\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n}}(\varepsilon)=\sum_{\substack{i_k=1\\k=1,\cdots,N}}^{n_k}\Delta_{\mathbf{i},\mathbf{n}}(\varepsilon).$$

Let us use a simpler spatial block decomposition (Tran (1990), Carbon et al. (1996) and Dabo-Niang & Thiam (2010)) slightly different from the previous ones. Without loss of generality, assume that $n_k=2vq_k$ for all $k=1,\cdots,N$ and $\hat{Q}=q_1\times\cdots\times q_N$. The random variables $\Delta_{\mathbf{i},\mathbf{n}}(\varepsilon)$ can be grouped into $2^N q_1 \times q_2 \times \cdots \times q_N$ cubic blocks of side υ. Denote

$$U_{\mathbf{n}}(1,\varepsilon,\mathbf{j}) := \sum_{\substack{i_k=2j_k\upsilon+1\\k=1,\cdots,N}}^{(2j_k+1)\upsilon} \Delta_{\mathbf{i},\mathbf{n}}(\varepsilon), \quad U_{\mathbf{n}}(2,\varepsilon,\mathbf{j}) := \sum_{\substack{i_k=2j_k\upsilon+1\\k=1,\cdots,N-1}}^{(2j_k+1)\upsilon} \sum_{i_N=(2j_N+1)\upsilon+1}^{2(j_{N-1}+1)\upsilon} \Delta_{\mathbf{i},\mathbf{n}}(\varepsilon)$$

$$U_{\mathbf{n}}(3,\varepsilon,\mathbf{j}) := \sum_{\substack{i_k=2j_k\upsilon+1\\k=1,\cdots,N-2}}^{(2j_k+1)\upsilon} \sum_{i_{N-1}=(2j_{N-1}+1)\upsilon+1}^{2(j_{N-1}+1)\upsilon} \sum_{i_N=2j_N\upsilon+1}^{(2j_N+1)\upsilon} \Delta_{\mathbf{i},\mathbf{n}}(\varepsilon)$$

$$U_{\mathbf{n}}(4,\varepsilon,\mathbf{j}) := \sum_{\substack{i_k=2j_k\upsilon+1\\k=1,\cdots,N-2}}^{(2j_k+1)\upsilon} \sum_{i_{N-1}=(2j_{N-1}+1)\upsilon+1}^{2(j_{N-1}+1)\upsilon} \sum_{i_N=(2j_N+1)\upsilon+1}^{2(j_N+1)\upsilon} \Delta_{\mathbf{i},\mathbf{n}}(\varepsilon)$$

and so on. Not that :

$$U_{\mathbf{n}}\left(2^{N-1},\varepsilon,\mathbf{j}\right) := \sum_{\substack{i_k=(2j_k+1)\upsilon+1\\k=1,\cdots,N-1}}^{2(j_k+1)\upsilon} \sum_{i_N=2j_N\upsilon+1}^{(2j_N+1)\upsilon} \Delta_{\mathbf{i},\mathbf{n}}(\varepsilon) \text{ and}$$

$$U_{\mathbf{n}}\left(2^{N},\varepsilon,\mathbf{j}\right) := \sum_{\substack{i_k=(2j_k+1)\upsilon+1\\k=1,\cdots,N}}^{2(j_k+1)\upsilon} \Delta_{\mathbf{i},\mathbf{n}}(\varepsilon).$$

For each integer $1 \le i \le 2^N$, let

$$T_{\mathbf{n},\varepsilon,i} = \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{q_k-1} U_{\mathbf{n}}(i,\varepsilon,\mathbf{j}), \quad S_{\mathbf{n},\varepsilon} = \sum_{i=1}^{2^N} T_{\mathbf{n},\varepsilon,i}.$$

Observe that, for $\eta > 0$

$$\mathbb{P}(|S_{\mathbf{n},\varepsilon}| > \eta) = \mathbb{P}\left(\left|\sum_{i=1}^{2^N} T_{\mathbf{n},\varepsilon,i}\right| > \eta\right) = 2^N \mathbb{P}\left(|T_{\mathbf{n},\varepsilon,i}| > \eta/2^N\right).$$

Without loss of generality, we only consider the case where $i = 1$. It is easily seen that $T_{\mathbf{n},\varepsilon,i}$ is the sum of $\hat{Q} = q_1 \cdots q_N$ of the $U_{\mathbf{n}}(1,\varepsilon,\mathbf{j})$'s. Note that $U_{\mathbf{n}}(1,\varepsilon,\mathbf{j})$ is measurable with respect to the $\sigma-$ field generated by $X_{\mathbf{i}}$ belongs to the set of sites :

$$\{\mathbf{i} : \ 2j_k\upsilon + 1 \le i_k \le (2j_k+1)\upsilon, \quad k = 1,\cdots,N\}.$$

This sets of sites are separated by a distance at least υ. According to Lemma 3.4.6, enumerate the random variables $U_{\mathbf{n}}(1,\varepsilon,\mathbf{j})$ with $0 \le j_k \le q_k - 1, \ k = 1,\cdots,N$ in an arbitrary order, say $Y_1,\cdots,Y_{\hat{Q}}$ with $|Y_i| \le C\upsilon^N \hat{k}_{\mathbf{n}}^{-1}$ and approximate them by the independent random variables $Y_1^*,\cdots,Y_{\hat{Q}}^*$ such that : for all $i = 1,\cdots,\hat{Q}$

$$\mathbb{E}|Y_i - Y_i^*| \le C\hat{k}_{\mathbf{n}}^{-1}\upsilon^N \psi\left(\hat{\mathbf{n}},\upsilon^N\right)\varphi(\upsilon).$$

Markov's inequality leads to

$$\mathbb{P}\left(\sum_{i=1}^{\hat{Q}} |Y_i - Y_i^*| > \eta/2^{N+1}\right) \le C2^{N+1}\hat{Q}\hat{k}_{\mathbf{n}}^{-1}\upsilon^N\psi\left(\hat{\mathbf{n}},\upsilon^N\right)\varphi(\upsilon)\eta^{-1}. \tag{3.86}$$

By Bernstein's inequality, we get

$$\mathbb{P}\left(\left|\sum_{i=1}^{\hat{Q}} Y_i^*\right| > \eta/2^{N+1}\right) \le 2\exp\left(-\frac{\eta^2/\left(2^{N+1}\right)^2}{4\sum\limits_{i=1}^{\hat{Q}} \mathbf{E}\left[Y_i^{*2}\right] + 2C\hat{k}_{\mathbf{n}}^{-1}\upsilon^N\eta/2^{N+1}}\right). \tag{3.87}$$

Combining (3.86) and (3.87), we have

$$\begin{aligned}\mathbb{P}(\,|S_{\mathbf{n},\varepsilon}| > \eta) &\le 2^N\mathbb{P}\left(\sum_{i=1}^{\hat{Q}} |Y_i - Y_i^*| > \eta/2^{N+1}\right) + 2^N\mathbb{P}\left(\left|\sum_{i=1}^{\hat{Q}} Y_i^*\right| > \eta/2^{N+1}\right)\\ &\le C2^{N+1}\hat{Q}\hat{k}_{\mathbf{n}}^{-1}\upsilon^N\psi\left(\hat{\mathbf{n}},\upsilon^N\right)\varphi\left(\upsilon\right)\eta^{-1}\\ &\quad + 2\exp\left(-\frac{\eta^2/\left(2^{N+1}\right)^2}{4\sum\limits_{i=1}^{\hat{Q}} \mathbf{E}\left[Y_i^{*2}\right] + 2^{N+1}C\hat{k}_{\mathbf{n}}^{-1}\upsilon^N\eta/2^{N+1}}\right).\end{aligned}$$

From (3.37), one has $\sum\limits_{i=1}^{\hat{Q}} \mathbf{E}\left[Y_i^{*2}\right] = O(\hat{k}_{\mathbf{n}}^{-1})$. Since $\hat{\mathbf{n}} = 2^N\hat{Q}\upsilon^N$, we have

$$\begin{aligned}\mathbb{P}(\,|S_{\mathbf{n},\varepsilon}| > \eta) &\le C2^{N+1}\hat{Q}\upsilon^N\hat{k}_{\mathbf{n}}^{-1}\psi\left(\hat{\mathbf{n}},\upsilon^N\right)\varphi\left(\upsilon\right)\eta^{-1} + 2^{N+1}\exp\left(-\frac{\eta^2\hat{k}_{\mathbf{n}}}{2^{2N+4}C + 2^{N+2}C\upsilon^N\eta}\right).\\ &\le 2C\hat{\mathbf{n}}\hat{k}_{\mathbf{n}}^{-1}\psi\left(\hat{\mathbf{n}},\upsilon^N\right)\varphi\left(\upsilon\right)\eta^{-1} + 2^{N+1}\exp\left(-\frac{\eta^2\hat{k}_{\mathbf{n}}}{2^{2N+4}C + 2^{N+2}C\upsilon^N\eta}\right).\end{aligned}$$

Let $\tilde{\lambda} > 0$ and set $\eta = \tilde{\lambda}\sqrt{\frac{\log\hat{k}_{\mathbf{n}}}{\hat{k}_{\mathbf{n}}}}$, $\upsilon = \left(\frac{\hat{k}_{\mathbf{n}}}{\log\hat{k}_{\mathbf{n}}}\right)^{1/(2N)}$ and a simple computation show that for sufficiently large $\mathbf{n}$ there exists $a > 1$, such that

$$\begin{aligned}\mathbb{P}(\,|S_{\mathbf{n},\varepsilon}| > \eta) &\le 2C\tilde{\lambda}\upsilon^{2N-\theta}\frac{\hat{\mathbf{n}}}{\hat{k}_{\mathbf{n}}} + 2^{N+1}\exp\left(\frac{-\lambda^2\log\hat{k}_{\mathbf{n}}}{2^{2N+4}C + 2^{N+2}C\lambda}\right)\\ &\le C\hat{\mathbf{n}}\hat{k}_{\mathbf{n}}^{\frac{-\theta}{2N}}\left(\log\hat{k}_{\mathbf{n}}\right)^{-1+\theta/2N} + C\hat{k}_{\mathbf{n}}^{-a},\end{aligned}$$

which tends to zero since by (3.13), $\hat{\mathbf{n}}\hat{k}_{\mathbf{n}}^{\frac{-\theta}{2N}}\left(\log\hat{k}_{\mathbf{n}}\right)^{-1+\theta/2N} \to 0$. The proof of the Lemma 3.5.3 is completed. ■

Chapitre 4

On fixed-design conditional tail index and quantile estimation for random fields

Sommaire

4.1 Résumé en français

Dans le chapitre 3, nous nous sommes intéressés à l'estimation non paramétrique de l'indice de queue et des quantiles d'un processus spatial $(Y_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$. Nous avons proposé une nouvelle approche de l'estimateur de l'indice de queue d'une distribution à queue lourde qui permet de tenir compte à la fois de la distance entre les observations et de celle entre les sites. Dans ce chapitre, nous proposons une adaptation de l'estimateur présenté au chapitre précédent au cadre conditionnel c'est-à-dire le cas où une co-variable déterministe est enregistrée simultanément avec la variable d'intérêt Y. On parle de modèle à "plan fixe" où "fixed design" en anglais. On souhaite estimer l'indice des queues ainsi que les quantiles associés de la variable aléatoire Y sachant une co-variable.

On considère un processus $(Y_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$, $N \geq 2$, à valeurs dans $\mathbb{R}$ et une variable déterministe $x_{\mathbf{i}}$ enregistrée au site $\mathbf{i}$.

A la place du couple $(Y_{\mathbf{i}}, x_{\mathbf{i}})$, nous noterons dans la suite $(Y_{x_{\mathbf{i}}})$, où $x_{\mathbf{i}} \in \mathbb{R}^d$ est l'information déterministe qui peut inclure $\mathbf{i}$.
Nous supposons que la distribution de Y sachant $x \in \mathbb{R}^d$ vérifie : $\forall x \in \mathbb{R}^d$, $\forall y > 0$,

$$F\left(y, x\right) = 1 - y^{-\frac{1}{\gamma(x)}} L\left(y, x\right), \tag{4.1}$$

Nous considérons les observations dépendantes $Y_{x_{\mathbf{i}}}$ dans un domaine rectangulaire

$$\mathcal{I}_{\mathbf{n}} = \{\mathbf{i} = (i_1, \cdots, i_N) \in \mathbb{Z}^N, 1 \leq i_k \leq n_k, k = 1, \cdots, N\}, \mathbf{n} = (n_1, \cdots, n_N) \in \mathbb{N}^N,$$

suivant le modèle (4.1). Nous cherchons à estimer l'indice $\gamma(x)$, $x \in \mathbb{R}^d$ ainsi que les quantiles associés à l'aide des observations $\{Y_{x_{\mathbf{i}}}\}_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}}$. Tout d'abord, avant de présenter notre estimateur, nous allons rappeler l'approche de la fenêtre mobile définie dans Gardes & Girard (2008). Pour $r > 0$, soit

$$B(x, r) = \{\omega \in \mathbb{R}^d, d(\omega, x) \leq r\}$$

la boule de centre x et de rayon r et soit $r_{\mathbf{n},x}$ une suite de nombres réels positifs tendant vers 0 lorsque $\mathbf{n} \to \infty$ (on dira que $\mathbf{n} \to \infty$ lorsque $\min\{n_k\} \to \infty$). Puisque l'estimateur est basé sur les variables $Y_{x_{\mathbf{i}}}$ pour lesquelles les co-variables $x_{\mathbf{i}}$ appartiennent à la boule $B(x, r_{\mathbf{n},x})$, la proportion des points dans cette boule est donc donnée par

$$\phi(r_{\mathbf{n},x}) = \frac{1}{\hat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \mathbb{1}_{\{x_{\mathbf{i}} \in B(x, r_{\mathbf{n},x})\}},$$

où $\hat{\mathbf{n}} = n_1 \times \cdots \times n_N$ est la taille de l'échantillon. Ainsi, le nombre d'observations non aléatoire $(Y_{x_{\mathbf{i}}})$ où $x_{\mathbf{i}} \in B(x, r_{\mathbf{n},x})$ est donné par $m_{\mathbf{n},x} = \hat{\mathbf{n}}\phi(r_{\mathbf{n},x})$.
Dans le cas conditionnel mais non spatial ($N = 1$, $\hat{\mathbf{n}} = n$), l'estimateur de l'indice des valeurs extrêmes peut être construit de la façon suivante :

- On définit d'abord une statistique d'ordre associée aux $m_{n,x}$ variables dont les co-variables sont dans la boule $B(x, r_{\mathbf{n},x})$ de la sorte : $Y_{(1),m_{n,x}} \geq Y_{(2),m_{n,x}} \geq \cdots \geq Y_{(m_{n,x}),m_{n,x}}$.
- Ensuite on choisit une suite intermédiaire k_x telle que $1 \leq k_x \leq m_{n,x}$ et $k_x = o(m_{n,x})$ quand $n \to \infty$.
- L'estimateur conditionnel de Hill suivant a été considéré dans la littérature (Gardes & Girard (2008))

$$\gamma_n(x) = \frac{1}{k_x} \sum_{i=1}^{k_x} \log\left(\frac{Y_{(i),m_{n,x}}}{Y_{(k_x+1),m_{n,x}}}\right) \tag{4.2}$$

Notre objectif est de proposer une nouvelle version de l'estimateur (4.2) dans un cadre spatial. Une extension directe de (4.2) dans notre cadre peut être construite de la manière suivante. Soit $\left\{\widetilde{Z}_{\mathbf{i}}(x), x_{\mathbf{i}} \in B(x, r_{\mathbf{n},x})\right\}$ l'ensemble des $Y_{x_{\mathbf{i}}}$, $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$ pour lesquelles les co-variables associées $x_{\mathbf{i}}$ appartiennent à la boule $B(x, r_{\mathbf{n},x})$. Alors, l'ensemble $\left\{\widetilde{Z}_{\mathbf{i}}(x)\right\}$ peut être énuméré dans l'ordre suivant : $\widetilde{Z}_{(1),m_{\mathbf{n},x}} \geq \widetilde{Z}_{(2),m_{\mathbf{n},x}} \geq \cdots \geq \widetilde{Z}_{(m_{\mathbf{n},x}),m_{\mathbf{n},x}}$. Soit $\hat{k}_{\mathbf{n},x}$ une suite intermédiaire d'éléments de $\mathbb{N}$ telle que $\hat{k}_{\mathbf{n},x} \leq m_{\mathbf{n},x}$. Nous supposerons que

$$\hat{k} = \hat{k}_{\mathbf{n},x} \to \infty\,; \, \hat{k}_{\mathbf{n},x} = o\left(m_{\mathbf{n},x}\right) \text{ quand } \mathbf{n} \to \infty.$$

On pourrait ainsi étendre (4.2) ci-dessus au cadre spatial en considérant l'estimateur :

$$\gamma_{\mathbf{n}}(x) = \frac{1}{\hat{k}_{\mathbf{n},x}} \sum_{i=1}^{\hat{k}_{\mathbf{n},x}} \log \left(\frac{\widetilde{Z}_{(i),m_{\mathbf{n},x}}}{\widetilde{Z}_{(\hat{k}_{\mathbf{n},x}+1),m_{\mathbf{n},x}}} \right) \tag{4.3}$$

Une autre alternative (voir Goegebeur et al. (2014b)) que nous étudierons dans la suite, permettant de choisir autrement les plus grandes observations $\widetilde{Z}_{\mathbf{i}}(x)$, est :

$$\widehat{\gamma}_{v_{\mathbf{n}}}(x) = \frac{\sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\left(\log \frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}}{\sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}}, \ x \in \mathbb{R}^d \tag{4.4}$$

où $r_{\mathbf{n}}$ est une suite de réels positifs tendant vers 0 si $\mathbf{n} \to \infty$ appelé fenêtre, $K(x) = \mathbb{1}_{\{\|x\|\leq 1\}}$ et $\mathbb{1}_{\{.\}}$ est la fonction indicatrice $(v_{\mathbf{n}})$ est une suite non aléatoire tendant vers ∞ et qui peut dépendre de la co-variable x. La suite $(v_{\mathbf{n}})$ est introduite pour sélectionner les plus grandes observations $Y_{\mathbf{i}}$ telles que $x_{\mathbf{i}} \in B(x, r_{\mathbf{n}})$. Cet estimateur est une extension au cadre spatial de l'estimateur proposé par Goegebeur et al. (2014b) dans un cadre de co-variable aléatoire et d'observations indépendantes.
Sous des conditions classiques des valeurs extrêmes, des propriètés de dépendance du processus, nous obtenons la consistance de l'estimateur $\widehat{\gamma}_{\mathbf{n}}(x)$. Nous en déduisons ensuite une estimation des quantiles extrêmes et en étudions leur consistance.

4.2 Introduction

In many fields such as oceanography, epidemiology, forestry survey, economy and many others, with the development of new technologies, data are now collected with geographical positions. The study of these kind of data or any characteristic of such data cannot be done without taking into account their respective geographical positions. It leads to a dynamic deployment of the known parametric and nonparametric methods (density estimation, prediction, regression, test, etc.) to spatial analysis. It is a general term to describe a technique that uses the spatial information in order to better understand the processes generating the observed attribute.
Several efficient spatial statistical tools are more and more developed to model such data. For some backgrounds in parametric spatial statistic modeling, we refer to Ripley (1981, 1991), Cressie (1992), Guyon (1995), Stein (1999), Cressie & Wikle (2011) and the references therein. More recently, nonparametric models are developed to reveal structure in data that might be missed by classical parametric ones. Some works have been done to study nonparametric variogram, quantile, density or regression problems for spatial data. We refer to Journel (1983), Tran (1990), Tran & Yakowitz (1993), Carbon et al. (1997, 1996), Biau & Cadre (2004), Menezes et al. (2010), Ould-Abdi et al. (2010), Dabo-Niang & Thiam (2010), El Machkouri (2011), Dabo-Niang & Yao (2013) and the references therein.
The study of extreme events (for example hurricanes, floods or earthquakes which can cause significant damage to structures such as bridges, towers and buildings) is increasingly extended to many other areas. Taking in consideration, the interest focused on the study of extreme values, that is, to develop sophisticated statistical tools for modeling extreme events. This chapter deals with extreme values estimation for spatial data.
The result obtained on the possible limit laws of the sample maximum has created

the idea that extreme values theory was something rather special and not like the classical central limit theory. Extreme values theory of stationary (non-spatial) random sequences in the univariate case has been extensively studied (e.g, Leadbetter (1974), Leadbetter et al. (1983), Hsing et al. (1988), Davis & Resnick (1985, 1988), Leadbetter & Hsing (1990), Davis & Resnick (1991)). Several papers pay attention on the possible application of the extreme values theory in different areas such as in hydrology (Davison & Smith (1990), Katz et al. (2002)), in meteorology (Coles & Walshaw (1994), Smith (2001), Klajnmic (2004)) among others. The extension of the univariate to the multivariate extreme values distribution is considered for instance in Haan & Resnick (1977) and Beirlant et al. (2006). In the simplest case, the process underlying the extremes is assumed to be independent or at least stationary and satisfying a mixing condition (see Leadbetter (1974)).

However, all this theory does not take into account the spatial dependence. As so far, max-stable processes have mostly been used for the statistical modeling of spatial data, see for example, Coles (2001) and Coles & Tawn (1996) who modeled extremal rainfall fields. Padoan et al. (2010) described a practicable pairwise likelihood estimation procedure applied to the rainfall data using max-stable processes. An interesting application to wind gusts is shown in Coles & Walshaw (1994) who used max-stable processes to model the angular dependence for wind speed directions.

In this chapter, we are interested in nonparametric conditional tail index estimation for spatial process. When some covariate is recorded simultaneously with the quantity of interest Y, the heavy-tail index can depend on the covariate and is referred in the sequel to as the conditional heavy-tail index. It involves the estimation of the so-called extreme values index or tail index. This parameter drives the tail heaviness of the distribution of Y and thus plays a central role in the analysis of extremes, making its estimation a crucial issue. Then, we consider the situation where some covariate information X is available to the investigator, and the distribution of Y depends on X. For every $X = x$, we consider the problem of estimating the conditional extreme values index $\gamma(x)$ of the distribution $F(y, x)$ of Y given $X = x$. Several papers already address the estimation of the conditional extreme values index (cf. Davison & Smith (1990), Smith (1989), Beirlant et al. (2006), Hall & Tajvidi (2000), Davison & Ramesh (2000)). More recently, a family of nonparametric estimators for the conditional tail index of Pareto type distribution when the covariate information is available can be found in Goegebeur et al. (2014b).

Here, we propose here a conditional tail index estimator for spatial data with fixed design covariate, based on a class of functions satisfying some mild conditions. To the best of our knowledge, such estimator has not been investigated before. The only work on tail index estimation for spatial process is that of Basrak & Tafro (2014). That article consider the tail estimation (with no covariate) for moving averages and moving maxima of $(X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^2)$ observed in bivariate regular lattice, and shows the consistency (convergence in probability) of the tail estimate.

However, some works exist on the non-parametric estimation of conditional quantiles in a spatial non-extreme case. For instance, the reader can refer to Hallin et al. (2009), Ould-Abdi et al. (2010) and Dabo-Niang & Thiam (2010).

In the case of spatial framework, Carreau & Girard (2011) propose to estimate spatial extreme quantiles by a weighted log-likelihood approach.

Estimating an extreme spatial event, particularly, the extreme index of a distribution function F heavy-tailed is far from being trivial. This parameter controls the behavior of F at infinity, which implies that its estimate is necessary, especially when one wants to estimate extreme quantiles. There are several methods for estimating tail index for different types of process, the most widely used estimator of the extreme values index was proposed by Hill (1975). Instead of parametric estimation, we consider in the following, that the conditional distribution of the random variable of interest given a covariate is heavy-tailed and depends on a nonlinear function of the covariate. Indeed, the conditional distribution is of polynomial type, with a rate of convergence driven by the conditional tail index. To estimate the unknown conditional tail index, we adapt nonparametric time-dependent extreme values smoothing methods to spatial data using some ordering method. The proposed methodology is based on a selection of the observations used in the estimator of the extreme values index, thanks to a moving window approach.

4.3 Model

Let us denote by $\left(Z_{\mathbf{i}} = (Y_{\mathbf{i}}, x_{\mathbf{i}}), \mathbf{i} \in \mathbb{Z}^N\right)$ the $\mathbb{R} \times \mathbb{R}^d-$valued measurable spatial process, with fixed non-random $x_{\mathbf{i}}$. For $y > 0$ and x, denote by $F(y, x)$ the conditional distribution function of random variable Y given x. We assume that for all x, the conditional cumulative distribution function of Y is heavy-tail. More specifically, we have

$$\overline{F}(y,x) := 1 - F(y,x) \sim y^{-\frac{1}{\gamma(x)}} L_F(y,x), \quad (\gamma(.) > 0)\,. \tag{4.5}$$

The conditional Pareto-type model can also be stated in an equivalent way in terms of the conditional tail quantile function $q(\alpha, x) := F^{-1}(1-\alpha, x)$ for $0 < \alpha < 1$, defined as follows

$$q(\alpha,x) = \alpha^{-\gamma(x)} L_q(\alpha^{-1}, x), \tag{4.6}$$

where $F^{-1}(., x)$ is the inverse function of the conditional distribution function. For x fixed, the unknown functions $L_F(., x)$ and $L_q(., x)$ are slowly varying at infinity, that is for all $t > 0$:

$$\lim_{y\to\infty} \frac{L_q(ty,x)}{L_q(y,x)} = \lim_{y\to\infty} \frac{L_F(ty,x)}{L_F(y,x)} = 1. \tag{4.7}$$

The unkown function $\gamma(x)$ is called the conditional tail index function, it describes the tail heaviness of the conditional response distribution, and is a function that needs to be adequately estimated from the available data. It is clear that for every $x \in \mathbb{R}^d$, $F(., x)$ belongs to the domain of attraction of Fréchet distribution with shape $\gamma(x)$.

For $\alpha \in (0, 1)$, the conditional quantile of Y given x of order α denoted $q(\alpha, x)$ is a solution of the equation :

$$F(q(\alpha,x),x) = \alpha.$$

We address the problem of estimating conditional high quantiles, defined as the graph of the functions $x \longmapsto q(\alpha_{\mathbf{n}}, x)$, $x \in \mathbb{R}^d$ where $\alpha_{\mathbf{n}} \to 0$ as $\mathbf{n} \to \infty$. A natural estimator $\widehat{q}(\alpha, x)$ of $q(\alpha, x)$ is defined such that :

$$\widehat{F}(\widehat{q}(\alpha,x),x) = \alpha.$$

Our purpose is to evaluate the asymptotic normality of the conditional tail index $\gamma(x)$ and the conditional extreme quantiles $q(\alpha, x)$ from a sample $(Z_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$ of random fields distributed according to (4.5).

Given a sample $Z_{\mathbf{i}} = (Y_{\mathbf{i}}, x_{\mathbf{i}})$ of observations from (4.5) over a rectangular domain

$$\mathcal{I}_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N, 1 \leq i_k \leq n, k = 1, \cdots, N\},$$

where $\mathbf{n} = (n_1, \cdots, n_N) \in \mathbb{N}^N$. Our aim is to build and evaluate a point-wise estimator of the function $\gamma(x)$ and the quantile $q(\alpha, x)$. More precisely, for a given $x \in \mathbb{R}^d$, we want to build and evaluate the asymptotic normality of $\gamma(x)$ and the quantile function $q(\alpha, x)$, focusing on the case where the design points $(x_{\mathbf{i}})$ are nonrandom. A point $\mathbf{i} = (i_1, \cdots, i_N) \in \mathbb{N}^N$ will be referred to as a site. We will write $\mathbf{n} \to \infty$ if $n \to \infty$. In the sequel, all the limits are considered when $\mathbf{n} \to \infty$. For $\mathbf{n} = (n_1, \cdots, n_N) \in \mathbb{N}^N$, we set $\widehat{\mathbf{n}} = n_1 \cdots n_N$. We assume that the spatial dependence of the process is measured by means of $\alpha - mixing$. Then, we consider the $\alpha - mixing$ coefficients of the field $(Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N)$, defined by :

***Mixing* condition :** Let E and E' be two sets of sites. Let $\mathcal{B}(E) = \mathcal{B}(Y_{\mathbf{i}}, \mathbf{i} \in E)$ and $\mathcal{B}\left(E'\right) = \mathcal{B}\left(Y_{\mathbf{i}}, \mathbf{i} \in E'\right)$ be $\sigma-$fields generated by the random variables $(Y_{\mathbf{i}})_{\mathbf{i}}$ with $\mathbf{i}$ being elements of E and E', respectively. There exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that whenever E, E' subsets of $\mathbb{Z}^N$ with finite cardinals,

$$\begin{aligned} \alpha\left(\mathcal{B}(E), \mathcal{B}\left(E'\right)\right) &= \sup_{B\in\mathcal{B}(E),\, C\in\mathcal{B}\left(E'\right)} |P(B \cap C) - P(B) P(C)| \\ &\leq \psi\left(\mathrm{Card}(E), \mathrm{Card}\left(E'\right)\right) \varphi\left(\mathrm{d}\left(E, E'\right)\right) \end{aligned} \tag{4.8}$$

where $\mathrm{Card}(E)$ (*resp.* $\mathrm{Card}\left(E'\right)$) denotes the cardinality of E (*resp.* E'), $\mathrm{d}\left(E, E'\right)$ the Euclidean distance between E and E' in $\mathbb{Z}^N$ defined by :

$$\mathrm{d}\left(E, E'\right) = \min\left\{\left(\sum_{k=1}^{N} |i_k - i'_k|^2\right)^{1/2} : (i_1, \cdots, i_N) \in E,\ \left(i'_1, \cdots, i'_N\right) \in E'\right\}$$

and $\psi : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable. Throughout the chapter, it will be assumed that ψ satisfies either

$$\psi(n, m) \leq C \min(n, m) \quad \forall n, m \in \mathbb{N} \tag{4.9}$$

or

$$\psi(n, m) \leq C(n + m + 1)^{\widetilde{\beta}}, \quad \forall n, m \in \mathbb{N} \tag{4.10}$$

for some $\widetilde{\beta} \geq 1$ and some $C > 0$. In the remainder of this chapter, we consider the first function defined in (4.9). Such function $\psi(m, n)$ can be found, for instance, in Tran (1990) ; Carbon et al. (1996, 1997), Biau & Cadre (2004) ; Dabo-Niang & Thiam (2010) ; Ould-Abdi et al. (2010) among others.

We also assume that one of both following conditions on $\varphi(t)$ is verified. These conditions are defined by

$$\varphi(t) = O(t^{-\theta}), \quad \text{for some } \theta > 0 \tag{4.11}$$

i.e that $\varphi(t)$ tends to zero at a polynomial rate, or

$$\varphi(t) = O(e^{-st}), \quad \text{for some } s > 0, \tag{4.12}$$

i.e that $\varphi(t)$ tends to zero at an exponential rate. We will consider, for simplicity, the function φ tending to zero at a polynomial rate. However, similar result could be obtained with $\varphi(t)$ tending to zero at an exponential rate which implies the polynomial case.

If $\psi \equiv 1$, then $Z_{\mathbf{i}}$ is called strongly mixing. Many stochastic processes, among them various useful time series models satisfy strong mixing properties, which are relatively easy to check. Conditions (4.8)-(4.12) are used in Tran (1990), Carbon et al. (1996, 1997). See Doukhan (1994) for discussion on mixing and examples.

4.4 Defining the estimator

Many estimators of the extreme values index and extreme quantiles have been proposed when some covariate information is available (e.g, Beirlant et al. (2006), Gardes & Girard (2008), Gardes & Girard (2010), Ndao et al. (2014)). When both spatial dependence and covariates are present, we propose to estimate these quantities by using a moving window approach. In this context, it is indeed natural to use the information provided by covariate for estimating the parameters of interest, which means that, we need to work locally in a neighborhood of the point studied.

Let $r > 0$ and $t \in \mathbb{R}^d$, let $B(t,r)$ in $\mathbb{R}^d$, the ball centered at point t with radius r that is

$$B(t,r) = \left\{w \in \mathbb{R}^d, d(w,t) \leq r\right\}$$

and let $r_{\mathbf{n},t}$ be a positive sequence tending to 0 as $\mathbf{n}$ goes to infinity. The proposed estimator uses the moving window approach as in Gardes & Girard (2008) since it is based on the response variables $Y_{\mathbf{i}}$'s for which the associated covariates $x_{\mathbf{i}}$'s belong to ball $B(t, r_{\mathbf{n},t})$. The proportion of such design points is thus defined by :

$$\phi\left(r_{\mathbf{n},t}\right) = \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbb{1}_{\{x_{\mathbf{i}}\in B(t,r_{\mathbf{n},t})\}} \text{ where } \widehat{\mathbf{n}} = \prod_{i=1}^{N} n_i.$$

It describes how the design points concentrates in the neighborhood of t when $r_{\mathbf{n},t}$ goes to 0, similary to the small ball probability does. Let $m_{\mathbf{n},t}$ be a sequence of integers such that $\mathbf{1} < m_{\mathbf{n},t} < \widehat{\mathbf{n}}$ and let $\left\{x_{\mathbf{1}}, \cdots, x_{m_{\mathbf{n},t}}\right\}$ the $m_{\mathbf{n},t}$ covariates recorded in the ball $B(t, r_{\mathbf{n},t})$. And let us denote by $\left\{\mathcal{Y}_{\mathbf{1}}, \cdots, \mathcal{Y}_{m_{\mathbf{n},t}}\right\}$ the associated observations. The nonrandom number of observations $\left(\widetilde{Z}_{\mathbf{i}}(t)\right)_{\mathbf{i}} := (\mathcal{Y}_{\mathbf{i}}, x_{\mathbf{i}})$ in $\mathbb{R} \times B(t, r_{\mathbf{n},t})$ are given by $m_{\mathbf{n},t} = \widehat{\mathbf{n}}\phi(r_{\mathbf{n},t})$.

In the non-spatial conditional case ($N = 1$, $\widehat{\mathbf{n}} = n$), the estimator of the tail index of extreme values can be constructed as follows :

- We first define an order statistic associated to the $m_{n,x}$ variables whose covariates are in the ball $B(x, r_{n,x})$ as follows : $\mathcal{Y}_{(1),m_{n,x}} \geq \mathcal{Y}_{(2),m_{n,x}} \geq \cdots \geq \mathcal{Y}_{(m_{n,x}),m_{n,x}}$.
- After that, we choose an intermediate sequence $\hat{k}_x$ such that $1 \leq \hat{k}_x \leq m_{n,x}$ and $\hat{k}_x = o(m_{n,x})$ as $n \to \infty$.
- The following conditional Hill estimator was considered in the literature (Gardes & Girard (2008))

$$\widehat{\gamma}_n(x) = \frac{1}{\hat{k}_x} \sum_{i=1}^{\hat{k}_x} \log\left(\frac{\mathcal{Y}_{(i),m_{n,x}}}{\mathcal{Y}_{(\hat{k}_x+1),m_{n,x}}}\right). \tag{4.13}$$

Our goal is to propose a new version of the estimator (4.13) in the spatial case. To this end, an alternative of the estimator (4.13) allowing to choose the biggest observations $Y_{\mathbf{i}}$, that we will study in the sequel, is

$$\widehat{\gamma}_{v_{\mathbf{n}}}(x) = \frac{\sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}}{\sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}},\ x\in\mathbb{R}^d \tag{4.14}$$

where $r_{\mathbf{n}} = r_{\mathbf{n},x}$ is a sequence of positive real numbers tending to zero as $\mathbf{n}\to\infty$ which is the bandwidth, $K(x) = \mathbb{1}_{\{\|x\|\le 1\}}$, $(v_{\mathbf{n}})$ is a nonrandom sequence tending to ∞ as $\mathbf{n}\to\infty$. The nonrandom $(v_{\mathbf{n}})$ is introduced to select the largest values of $Y_{\mathbf{i}}$ for which the associated covariate $x_{\mathbf{i}}\in B(x, r_{\mathbf{n}})$ to estimate $\gamma(x)$. Its choice is thus very important in practice. Let us notice that in (4.14), the term $\pi_{\mathbf{n}}(v_{\mathbf{n}}, x) := \sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}$ is the number of observations used to estimate $\widehat{\gamma}_{v_{\mathbf{n}}}(x)$. Clearly, its choice is very crucial since it is a compromise between bias and variance of the estimator, thus, for most values of $\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)$, the statistic $\widehat{\gamma}_{v_{\mathbf{n}}}(x)$ is equal to zero. Note that, the estimator (4.14) we consider here is version of Hill's estimator proposed by Goegebeur et al. (2014b) in the random covariate and non-spatial case.

Remarque 4.4.1. *For example, to specify the design points* $x_{\mathbf{i}}$*, one can choose* $x_{\mathbf{i}}$ *such that* $x_{\mathbf{i}} = \frac{\mathbf{i}}{\mathbf{n}} = (\frac{i_1}{n}, \frac{i_2}{n}, \cdots, \frac{i_N}{n})$ *with* $\mathbf{n} = (n, ..., n)$*, i.e.* $n_j = n$*,* $j = 1, \cdots, n$*. We have* $x_{\mathbf{i}}\in[0,1]^N$ *and* $x\in[0,1]^N$*. Then equation (4.14) becomes in this framework : for each fixed* $x\in[0,1]^N$*, the estimator of* $\gamma(x)$ *can then be written as follows*

$$\widehat{\gamma}_{v_{\mathbf{n}}}(x) = \frac{\sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-\frac{\mathbf{i}}{\mathbf{n}}}{r_{\mathbf{n}}}\right)\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}}{\sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-\frac{\mathbf{i}}{\mathbf{n}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}},\ x\in\mathbb{R}^d \tag{4.15}$$

In this case, one can assume that K *is a symmetric, non-negative density with supported by* $[-1,1]^N$ *and* $\int K^2(u)du < \infty$*. Assume also that it satisfies a Lipschitz condition* $|K(x) - K(y)| \le R\|x-y\|$ *for any* $x, y\in[-1,1]^N$*,* $R > 0$ *and there exist two constants* $C_1, C_2 > 0$ *such that* $C_1 \le K(x) \le C_2$ *for any* $x, y\in[-1,1]^N$*. Under the conditions on* K*, El Machkouri & Stoica (2010), Lemma 1 show that if* $\widehat{\mathbf{n}}r_{\mathbf{n}}^{N+1}\to\infty$ *then* $\frac{m_{\mathbf{n},x}}{\widehat{\mathbf{n}}r_{\mathbf{n}}^N}\to 1$ *for* $x\in[0,1]^N$*. In this case, one can obtain the results of Theorems 4.6.1 by replacing* $m_{\mathbf{n},x}$ *by* $\widehat{\mathbf{n}}r_{\mathbf{n}}^N$*.*

4.5 Estimating the conditional extreme quantiles

In most applications, the extreme value index is not the primary object of interest, but for instance exceedance probabilities or extreme quantiles are to be estimated. In this part, we address the estimation of the conditional extreme quantiles $q(\alpha, x)$ of order $\alpha_{\mathbf{n}}$ of the distribution of Y given x. Those quantiles verify the following equation :

$$\overline{F}\left(q(\alpha_{\mathbf{n}}, x), x\right) = \alpha_{\mathbf{n}},$$

where $\alpha_{\mathbf{n}}$ goes to 0 as $\mathbf{n}\to+\infty$ and $\overline{F}(y, x)$ denotes the survival function of Y given x.

To estimate the conditional survival function of Y given x, we propose to use the following empirical estimator used by Gardes et al. (2012), that is

$$\begin{aligned}\widehat{\overline{F}}_{\mathbf{n}}(y,x) &= \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} R_{\mathbf{ni}}(x)\mathbb{1}_{\{Y_{\mathbf{i}}>y\}}, \quad \text{with} \\ R_{\mathbf{ni}}(x) &= K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right) / \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right),\end{aligned} \tag{4.16}$$

where $K(\cdot)$, $\mathbb{1}_{\{\cdot\}}$ and $r_{\mathbf{n}}$ are defined in the previous section. Let us highlight that, in the last equality, the term $m_{\mathbf{n},x} = \sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)$ is the number of covariates $x_{\mathbf{i}}$ which lie in the ball $B(x, r_{\mathbf{n}})$ centered in x with radius $r_{\mathbf{n}}$.

Then, for fixed $\alpha \in (0,1)$, the associated conditional quantile function $\widehat{\overline{F}}_{\mathbf{n}}^{-1}(\cdot,x)$ can be defined by the generalized inverse function of $\widehat{\overline{F}}_{\mathbf{n}}(\cdot,x)$ as follows :

$$\widehat{\overline{F}}_{\mathbf{n}}^{-1}(\alpha,x) = \inf\left\{y\in\mathbb{R} : \widehat{\overline{F}}_{\mathbf{n}}(y,x) \le \alpha\right\}. \tag{4.17}$$

According to equation (4.6), if $\alpha_{\mathbf{n}}$ is small enough and $\beta_{\mathbf{n}} < \alpha_{\mathbf{n}}$, one can obtain the following approximation for all $\gamma(x) > 0$,

$$q(\alpha_{\mathbf{n}},x) = \alpha_{\mathbf{n}}^{-\gamma(x)} L_q\left(\alpha_{\mathbf{n}}^{-1},x\right) \tag{4.18}$$

$$q(\beta_{\mathbf{n}},x) = \beta_{\mathbf{n}}^{-\gamma(x)} L_q\left(\beta_{\mathbf{n}}^{-1},x\right). \tag{4.19}$$

Then dividing (4.18) by (4.19) and by the slowly varying function, it follows that

$$q(\alpha_{\mathbf{n}},x) \approx q(\beta_{\mathbf{n}},x)\left(\frac{\alpha_{\mathbf{n}}}{\beta_{\mathbf{n}}}\right)^{-\gamma(x)} \tag{4.20}$$

where $q(\alpha_{\mathbf{n}},x)$ is a choosen quantile in the sample which is easily obtained from (4.17) and $q(\beta_{\mathbf{n}},x)$ a chosen quantile outside the sample. So, we extrapolate $(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}})^{-\gamma(x)}$ to estimate an extreme quantile of order $\beta_{\mathbf{n}}$ arbitrary small and replacing $q(\alpha_{\mathbf{n}},x)$ by the estimator $\widehat{q}(\alpha_{\mathbf{n}},x) := \widehat{\overline{F}}_{\mathbf{n}}^{-1}(\alpha_{\mathbf{n}},x)$ given in (4.17) and $\gamma(x)$ by the above estimator $\widehat{\gamma}_{v_{\mathbf{n}}}(x)$. This allows us to obtain the estimator of Weissman (1978),

$$\widehat{q}_{\mathbf{n}}^{W}(\beta_{\mathbf{n}},x) := \widehat{q}(\alpha_{\mathbf{n}},x)\,(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}})^{\widehat{\gamma}_{v_{\mathbf{n}}}(x)}. \tag{4.21}$$

Note that, the best known quantile estimator of Weissman (1978) is based on a semi-parametric approach and on the estimation of $\gamma(x) > 0$.

This estimator extends the conditional extreme quantile estimator proposed in Gardes & Girard (2010) and used recently in Ndao et al. (2014) when the data are not spatially dependent.

4.6 Assumptions and main results

To prove asymptotic normality of $\widehat{\gamma}_{v_{\mathbf{n}}}(x)$ and $\widehat{q}_{\mathbf{n}}^{W}(\beta_{\mathbf{n}},x)$, we now give some notations and assumptions. We first denote $m_x := m_{\mathbf{n},x}$ and for a given x, $\overline{F}(\cdot,x) :=$

$1 - F(\cdot, x)$ is regularly varying at infinity, i.e, there exists an unknown positive continuous function $\gamma(\cdot)$ such that, $\forall x \in \mathbb{R}^d$,

$$\lim_{y\to\infty} \frac{\overline{F}\,(ty,x)}{\overline{F}\,(y,x)} = t^{-1/\gamma(x)}, \quad \forall x > 0. \tag{4.22}$$

We are interested in estimating $\gamma(x) > 0$ based on largest values of $Y_{\mathbf{i}}$ for which the associated covariates fall in the ball $B(x, r_{\mathbf{n}})$ and we need the following conditions :

($\mathcal{F}$) The function γ is positive and continuous on $\mathbb{R}^d$ with support Ω included in the unit ball on $\mathbb{R}^d$. Suppose also that, the cumulative distribution function $F(\cdot, x)$ is absolutely continuous for any $x \in \mathbb{R}^d$ and there exist constants $c_{\overline{F}} > 0$ and $y_0 > 1$ such that

$$\sup_{y\geq y_0} \left| \frac{\log \overline{F}(y,x)}{\log \overline{F}(y,z)} - 1 \right| \leq c_{\overline{F}} d(x,z).$$

($\mathcal{G}$) f is continuous in $\mathbb{R}^{d+1}$ and for any $\mathbf{i} \neq \mathbf{j}$, the pairs $(X_{\mathbf{i}}, X_{\mathbf{j}})$ admits uniformly joint density $f_{\mathbf{i},\mathbf{j}}$ such that there exists positive constants c satisfying

$$\sup |f_{\mathbf{i},\mathbf{j}}(x,y) - f_{\mathbf{i}}(x) f_{\mathbf{j}}(y)| \leq c, \quad \forall x,\ y \in \mathbb{R}^{d+1}.$$

Since the slowly-varying functions are of interest only asymptotically, without loss of generality, the Karamata representation (see for instance Theorem 1.3.1 in Bingham et al. (1987) or Daouia et al. (2011)) of the slowly-varying function can be written as follows

$$L_q\,(y,x) = c(x) \exp\left(\int_1^y \frac{\mathcal{A}(u,x)}{u} du \right), \tag{4.23}$$

where $c(x)$ is a positive function and $\mathcal{A}(y,x) \to 0$ as $y \to \infty$.

($\mathcal{Q}$) The function $|\mathcal{A}(\cdot, x)|$ is ultimately decreasing and regularly varying with index $\rho(x)/\gamma(x) < 0$, that is for all $t > 0$,

$$|\mathcal{A}(tu,x)| \,/\, |\mathcal{A}(u,x)| \longrightarrow t^{\rho(x)/\gamma(x)} \text{ as } u \to \infty.$$

Moreover, $\mathcal{A}(\cdot, x)$ is assumed to have a constant sign at infinity.

Remarque that conditions (4.23) and ($\mathcal{Q}$) entail that

($\mathcal{R}_2$) for all $t > 0$,

$$\lim_{y\to\infty} \frac{\overline{F}(ty,x)/\overline{F}(y,x) - t^{-1/\gamma(x)}}{\mathcal{A}(y,x)} = t^{-1/\gamma(x)} \frac{t^{\rho(x)/\gamma(x)} - 1}{\gamma(x)\rho(x)}.$$

($\mathcal{W}$) The sequence satisfy the following conditions :

(i) $m_x \to \infty$ as $\mathbf{n} \to \infty$ and there exists a sequence of integer $q = q_{\mathbf{n}} \to \infty$ such that

$$q = o\left[m_x \left(\overline{F}(v_{\mathbf{n}}, x) \right)^{(1+(1-\varsigma))/2N} \right]^{2N} m_x q^{-\theta} \longrightarrow 0, \tag{4.24}$$

with $0 < \varsigma < (\theta - N)/\theta$, $\theta > 2(N+1)$, where θ is defined in (4.11).

(ii) $\left(\overline{F}(v_{\mathbf{n}}, x) \right)^{-(1-\varsigma)} q^{N-\theta(1-\varsigma)} \longrightarrow 0$ as $\mathbf{n} \to \infty$.

Comments Note that Assumption $(\mathcal{R}_2)$ is the so-called second order condition classically used to establish the asymptotic normality of tail index estimators. Note that the second order parameter $\rho(x)$ controls the rate of convergence of $L(\lambda y, x)/L(y, x)$ to 1 (see Bingham et al. (1987) and Haan & Ferreira (2006) for further details). In particular, if $\rho(x)$ is close to 0, the convergence is slow and thus, the estimation of the conditional tail index is difficult. Note that this condition is also used in Goegebeur et al. (2014b).Finally, $(\mathcal{W})$ are technical conditions required for the bandwidth in order to get convergence results of our proposed estimator, they are related to the spatial context. $(\mathcal{G})$ is a local dependency condition classical in non-parametric inference for spatial data, see for instance Carbon et al. (1996, 1997), Ould-Abdi et al. (2010).

The following theorem gives the asymptotic convergence of the conditional tail index estimator $\widehat{\gamma}_{v_{\mathbf{n}}}(x)$.

Theorem 4.6.1. *Let $\{(Y_{\mathbf{i}}, x_{\mathbf{i}})\}_{\mathbf{i}\in\mathbb{Z}^N}$ (observed on $\mathcal{I}_{\mathbf{n}}$) where $\{Y_{\mathbf{i}}\}$ is a mixing spatial process satisfying (4.8)-(4.10). Assume that the conditions $(\mathcal{F})$, $(\mathcal{W})$, $(\mathcal{Q})$, (4.23) and (4.22) hold. Suppose also that there exists a bounded function $\lambda(\cdot) \in \mathbb{R}$ and a sequence $(v_{\mathbf{n}})$ such that :*

- ▷ *As $\mathbf{n} \to \infty$, $v_{\mathbf{n}} \to \infty$, $m_x \to \infty$ and $m_x\overline{F}(v_{\mathbf{n}}, x) \longrightarrow \infty$.*
- ▷ *As $\mathbf{n} \to \infty$, $\sqrt{m_x\overline{F}(v_{\mathbf{n}}, x)}\mathcal{A}(v_{\mathbf{n}}, x) \longrightarrow \lambda(x) < \infty$.*

Then for all $x \in \mathbb{R}^d$,

$$\sqrt{m_x\overline{F}(v_{\mathbf{n}}, x)}\left(\widehat{\gamma}_{v_{\mathbf{n}}}(x) - \gamma(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)}, \gamma^2(x)\right). \tag{4.25}$$

It appears that the estimator is asymptotically Gaussian with asymptotic variance proportional to $\gamma^2(x)/m_x\overline{F}(v_{\mathbf{n}}, x)$ for mixing random fields. This result is similar to the one established by Goegebeur et al. (2014b) in the non-spatial conditional random variable case.

We now examine the asymptotic properties of the estimator of the extreme quantile given in (4.21).

The following result gives the asymptotic normality of the extreme conditional quantile (4.21) satisfying the above conditions.

Theorem 4.6.2. *Assume that conditions of Theorem 4.6.1 holds. In addition, let $(\alpha_{m_x})_{m_x\geq 1}$ and $(\beta_{m_x})_{m_x\geq 1}$ be positive sequences such that*

- ▷ *As $\mathbf{n} \to \infty$, $\alpha_{m_x} \to 0$, $\beta_{m_x}/\alpha_{m_x} \to 0$, $\left|\frac{\hat{q}(\alpha_{m_x},x)}{q(\alpha_{m_x},x)} - 1\right| \xrightarrow{\mathbb{P}} 0$,*
- ▷ *As $\mathbf{n} \to \infty$, $\sqrt{m_x\overline{F}(v_{\mathbf{n}}, x)}\ \mathcal{A}(q(\alpha_{m_x}, x), x) \to \lambda(x)$.*

Then, for any $x \in \mathbb{R}^d$,

$$\frac{\sqrt{m_x\overline{F}(v_{\mathbf{n}}, x)}}{\log\left(\alpha_{m_x}/\beta_{m_x}\right)}\left(\frac{\hat{q}^W_{\mathbf{n}}(\beta_{m_x}, x)}{q(\beta_{m_x}, x)} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)},\ \gamma^2(x)\right). \tag{4.26}$$

4.7 A small simulation study

In this section, we conduct a comprehensive simulation study to assess the performance of the estimators (4.14) and (4.21).

4.7.1 Models

In this section, we illustrate our results by a small simulation study. For simplicity, we take $N = 2$ and consider a random process $\{(Y_{\mathbf{i}}, x_{\mathbf{i}}), \mathbf{i} \in \mathbb{Z}^2\}$ with values in $\mathbb{R} \times \mathbb{R}^d$ and $x_{\mathbf{i}}$ is non random. We assume that the process is observed over a subset

$$\mathcal{I}_{\mathbf{n}} = \left\{ \mathbf{i} = (i,j) \in \mathbb{N}^{*2};\, 1 \leq i \leq n_1;\, 1 \leq j \leq n_2 \right\}, \quad \mathbf{n} = (n_1, n_2) \in \mathbb{N}^2.$$

The aim of this section is to evaluate the performance of the conditional tail index estimator (4.14) towards simulation. Without loss of generality, we take $n_1 = n_2$ and simulate $M = 100$ samples $(Y_{\mathbf{i}}, x_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}}$ of different sizes $\hat{\mathbf{n}} = n_1 \times n_2$ with $n_1 \in \{20,\, 30\}$ and $x_{\mathbf{i}} \in [0,1]$. We consider two models where the conditional distribution of $Y_{\mathbf{i}}$ given $x_{\mathbf{i}}$ is Paréto and Fréchet distributions :

$$\textbf{Model A :} \quad Y_{\mathbf{i}} = Y_{\mathbf{i}}(x_{\mathbf{i}}) = U_{\mathbf{i}} \left(1 - V_{\mathbf{i}}\right)^{-1/\gamma(x_{\mathbf{i}})}, \quad \mathbf{i} \in \mathbb{N}^{*2}, \tag{4.27}$$

$$\textbf{Model B :} \quad Y_{\mathbf{i}} = Y_{\mathbf{i}}(x_{\mathbf{i}}) = U_{\mathbf{i}} \left(- \log V_{\mathbf{i}}\right)^{-1/\gamma(x_{\mathbf{i}})}, \quad \mathbf{i} \in \mathbb{N}^{*2}, \tag{4.28}$$

where $V = (V_{\mathbf{i}},\, \mathbf{i} \in \mathbb{N}^{*2})$ is a uniform random variable and $U_{\mathbf{i}}$ is introduced to control the spatial local dependency. We choose :

- First case : $U_{\mathbf{i}} = \frac{1}{\hat{\mathbf{n}}} \sum_{\mathbf{i},\mathbf{k} \in I_{\mathbf{n}}} \exp\left(-\frac{\|\mathbf{i}-\mathbf{k}\|}{\rho}\right)$.
- Second case : $U_{\mathbf{i}} = \left| \sin\left(2i\pi/T\right) + \sin\left(2j\pi/T\right) \right|$ where $\mathbf{i} = (i,j) \in \mathbb{N}^{*2}$.
- Third case : $U_{\mathbf{i}} = a^{i+j}$ where $\mathbf{i} = (i,j)$ and $a \in \mathbb{R}^*$.

In the two first cases, the spatial dependence is controlled by the value of ρ and T respectively. In fact, the greater ρ is, the weaker the spatial dependency is. According to this fact, we provide simulation results obtained with different values of ρ and T ($\rho = T = 5,\, 50$) and $a = .98$.

A sample of size $\hat{\mathbf{n}} = 900$ obtained, from **cases 1, 2** with $\rho = T = 5,\, 50$ and **case 3** with parameter $a = .98$ are represented respectively in Figures 4.1 and 4.2. We denote by $\left\{\widetilde{Z}_{\mathbf{i}}(x_{\mathbf{i}})\right\}_{\mathbf{i} \in \mathbb{N}^{*2}}$ the selected variables $Y_{\mathbf{i}}$'s for which the associate covariates $x_{\mathbf{i}}$'s belong to the ball $B(x_0, r_{\mathbf{n}})$ where x_0 is the point where we want to estimate the function $\gamma(\cdot)$. The conditional distribution of $\left\{\widetilde{Z}_{\mathbf{i}}(x_{\mathbf{i}})\right\}_{\mathbf{i} \in \mathbb{N}^{*2}}$ according to (4.27) and (4.28) is Paréto and Fréchet distribution respectively. The following conditional tail-index has been chosen as in Daouia et al. (2011), see also Ndao et al. (2014) and Ndao et al. (2016)

$$\gamma(x) = 0.5 \left(0.1 + \sin\left(\pi x\right)\right) \left(1.1 + 0.5 \exp\left\{-64 \left(x - 0.5\right)^2\right\}\right). \tag{4.29}$$

The pattern of $\gamma(\cdot)$ is given in Figure 4.3.
For simplicity, let $m_{\mathbf{n},x_0} = m_{x_0}$, $\hat{k}_{\mathbf{n},x_0} = \hat{k}$, $\widehat{\gamma}_{v_{\mathbf{n}}}(x_0) = \widehat{\gamma}_{v_{\mathbf{n}},\hat{k}}(x_0)$ and $r_{\mathbf{n}} = r$, where $(\hat{k}+1)$ is the number of largest observations of interest. For each sample, we estimate

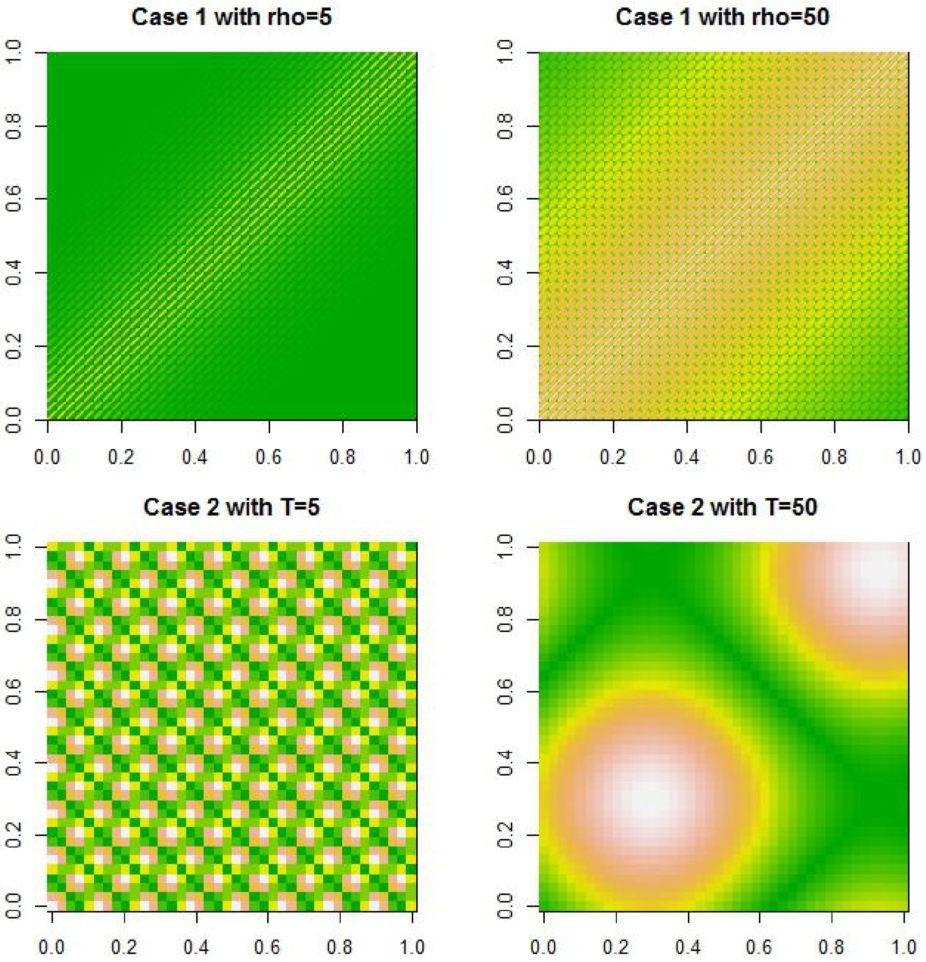

FIGURE 4.1 – Representation of the field $\{U_{\mathbf{i}}\}_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}$ with $\hat{\mathbf{n}} = 900$.

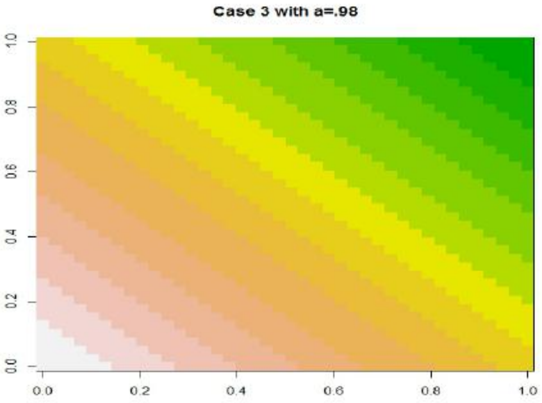

FIGURE 4.2 – Representation of the field $\{U_{\mathbf{i}}\}_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}$ with $\hat{\mathbf{n}} = 900$.

$\gamma(\cdot)$ at point $x_0 = 0.5$ $(\gamma(0.5) = 0.33)$ by the Hill estimator proposed in Section 4.4. The moving window approach is used with the ball $B(x_0, r)$.

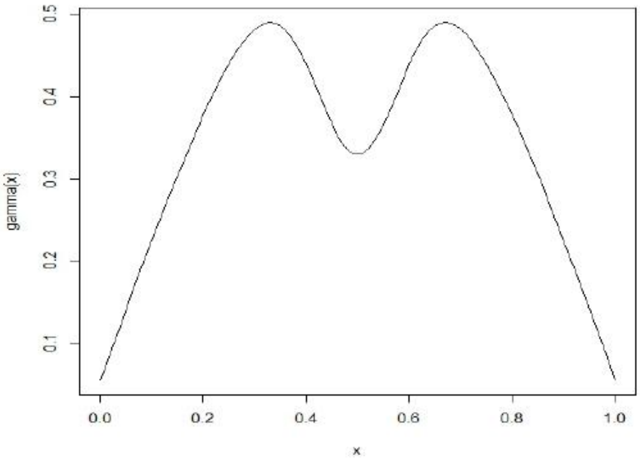

Figure 4.3 – Representation of $\gamma(\cdot)$ on $[0,1]$.

Thus, keeping in mind that this method should be modified in practice since $\gamma(x)$ is unknown. In order to calculate our estimators, we need to choose the bandwidth r and threshold $v_{\mathbf{n}}$. We take $v_{\mathbf{n}}$ as the $(m_x-\hat{k})$ th order statistic $Y_{(m_x-\hat{k})}$, as is classical in extreme values statistics.

Now, we propose a procedure for choosing the pair $(r,\hat{k})$. This procedure is adapted from Goegebeur et al. (2014b). First, we select the bandwidth r using the following cross-validation criterion, introduced by Yao (1999) and also implemented by Daouia et al. (2011) and used recently by Ndao et al. (2016)

$$r^* := \arg\min_{r\in\mathcal{R}_0} \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \sum_{\mathbf{j}\in\mathcal{I}_\mathbf{n}} \left(\mathbb{1}_{\{Y_\mathbf{i}>Y_\mathbf{j}\}} - \widehat{\overline{F}}_{\mathbf{n},-\mathbf{i}}\left(Y_\mathbf{j}, X_\mathbf{j}\right)\right)^2,$$

where $\mathcal{R}_0$ is a grid of r $(\mathcal{R}_0 = \{r_1 = 1/5\log(m_x) < r_2 < \cdots < r_{max} = .5\})$ where the points $r_2,\cdots,r_{max}$ are regularly distributed between r_1 and r_{max-1} and $\widehat{\overline{F}}_{\mathbf{n},-\mathbf{i}}$ is the conditional distribution estimator (4.16) calculated on the sample $(Y_\mathbf{i}, x_\mathbf{i})$, $\mathbf{i} = \mathbf{1},\cdots,m_x-1$.

The procedure that is used is as follows : for all $x\in B(x,r)$,

1. we first compute the estimate $\widehat{\gamma}_{Y_{(m_x-\hat{k})}}(x)$ for $\gamma(x)$ with $\hat{k} = 1,2,\cdots,m_x-1$,

2. second, we divided into several successive blocks of same size the estimate $\widehat{\gamma}_{Y_{(m_x-\hat{k})}}(x)$, that is, one block for $\hat{k}\in\{1,2,\cdots,15\}$, second block for $\hat{k}\in\{16,\cdots,30\}$, and so on,

3. third, we evaluate standard deviation of the estimate within blocks,

4. at the end, we choose the $\hat{k}-$value denoted $\hat{k}^*$, to be used from the block with minimal standard deviation. Specifically, we opt to take the median value of the $\hat{k}-$value in the block.

Conclusion, we evaluate the estimator (4.14) of $\gamma(x)$ with $\left(r,\,\hat{k}\right) = \left(r^*,\,\hat{k}^*\right)$.

4.7.2 Estimates of the conditional tail index

For each configuration, we calculate the averaged estimate of $\gamma(x)$ and the averaged empirical mean square error (AMSE) and averaged mean absolute error (AMAE) over the M estimates. From the AMSE and AMAE, we calculate the corresponding standard deviation. Table 4.7.2 contains the numerical results for the estimates of $\gamma(x_0)$, the AMSE and the AMAE of the estimator $\widehat{\gamma}_{v_{\mathbf{n}},\hat{k}}(x_0)$ of the tail index obtained with models **A** and **B** for $\hat{\mathbf{n}} = 400$, 900 and **case 3** with $a = .98$.

$\hat{\mathbf{n}}$	Models	**case**	ρ	$\hat{k}^*$	$\widehat{\gamma}_{v_{\mathbf{n}}}(.)$	AMSE	AMAE
20×20	**A**	**1**	5	124	0.3175	0.0014 (0.018)	0.029 (0.065)
			50	389	0.3309	0.0004 (0.025)	0.016 (0.073)
		2	5	138	0.3219	0.0014 (0.019)	0.029 (0.069)
			50	318	0.3295	0.0006 (0.025)	0.020 (0.078)
		3		369	0.3313	0.0005 (0.022)	0.018 (0.070)
	B	**1**	5	119	0.3248	0.0017 (0.018)	0.032 (0.065)
			50	334	0.3303	0.0006 (0.027)	0.020 (0.082)
		2	5	131	0.3311	0.0011 (0.018)	0.027 (0.069)
			50	282	0.3250	0.0005 (0.026)	0.019 (0.081)
		3		324	0.3325	0.0009 (0.021)	0.024 (0.072)
30×30	**A**	**1**	5	208	0.3223	0.0013 (0.019)	0.028 (0.071)
			50	841	0.3307	0.0002 (0.028)	0.012 (0.079)
		2	5	259	0.3186	0.0006 (0.023)	0.021 (0.077)
			50	426	0.3307	0.0004 (0.025)	0.015 (0.082)
		3		734	0.3301	0.0003 (0.019)	0.014 (0.069)
	B	**1**	5	199	0.3345	0.0012 (0.018)	0.026 (0.068)
			50	739	0.3306	0.0002 (0.028)	0.012 (0.085)
		2	5	246	0.3286	0.0004 (0.023)	0.017 (0.080)
			50	387	0.3283	0.0004 (0.025)	0.016 (0.082)
		3		660	0.3305	0.0005 (0.019)	0.017 (0.069)

Table 4.1 – Simulation results of the $M = 100$ estimate of $\gamma(x_0)$ according to the models **A** and **B**, the cases **1**, **2** and **3** and the value of $\rho = 5, 50$ and $a = .98$: for each simulation scenario, from left to right, we have respectively, $\hat{k}^*$ the intermediate sequence, the estimate $\widehat{\gamma}_{v_{\mathbf{n}}}(.)$, the average mean squared errors (AMSE), the average mean absolute errors (AMAE) and in brackets the corresponding standard deviation.

From Table 4.7.2, based on $(r^*, \hat{k}^*)$, we calculate the empirical average mean squared errors (AMSE) and average mean absolute errors (AMAE) of the estimate of $\gamma(x_0)$. These illustrations show the good performance of the estimate. As we can see, in the two models, the AMSE decreases for a large range of values of $\hat{k}^*$ and for large values of $\hat{\mathbf{n}}$. This brings up the important question of how to choose the adequate $\hat{k}^*$ that means in this case, how to choose the adequate radius r^* of the ball for a given data set of size $\hat{\mathbf{n}}$. It is well known that the choice of the most appropriate value for $\hat{k}^*$ is a difficult strategy, we refer the reader to Drees & Kaufmann (1998), Haan & Peng (1998), Danielsson et al. (2001), Yao (1999), Gardes & Girard (2008) and Lekina (2010) for more details.
Regardless the models themselves, AMSE is less important when ρ (the parameter that controls the spatial dependency) increases. This can be explained by the fact that when ρ decreases, the procedure use few points in the ball to estimate $\gamma(x_0)$. When the spatial dependence parameter ρ increases, the process is closed to the independence case, then the distribution does not change much. Therefore, in summary, the higher the values of ρ (which means less spatial dependence) is, the smaller the absolute mean square error of the estimate is. If we consider **case 3**, we can also see that the AMSE decreases with the size of the sample size.

4.7.3 Estimates of the conditional extreme quantiles

In this section, we evaluate the performance of the estimator (4.21) of the conditional extreme quantile $q(1/1000, 0.5)$ of order $1-1/1000$ of the conditional distribution of Y given $x = 0.5$ with $q(1/1000, 0.5) = 9.77$. For each configuration scenario of the simulation design parameters, we calculate the conditional estimate (4.21), based on the Hill estimator of the conditional tail index. Then, for each sample size $\mathbf{n}$ and the spatial dependence controllers (ρ, T and a), we obtain the averaged value of the $\widehat{q}_{\mathbf{n}}^{W}(1/1000, 0.5)$, the AMSE and AMAE with $\left(r^*, \hat{k}^*\right)$. The results are given in Table 4.7.3.

From table 4.7.3, we observe the performance of the proposed estimator. Based on the AMSE and the AMAE, in one hand, when the spatial dependence is very strong ($\rho = T = 5$) and the sample size decreases, we observe that the performance of the proposed estimator deteriorates for models **A** and **B** (cases **1** and **2**), the bias can be quite large. In other hand, when the spatial dependence is low ($\rho = T = 50$), the bias is quite limited. At the end, for the case **3** and for both models, the bias of the proposed estimator stays limited, the estimator performs better. When the sample size is sufficiently large, the bias of the proposed estimator stays small in all simulation study.

$\hat{\mathbf{n}}$	Models	**case**	ρ	$\widehat{q}^{W}_{\mathbf{n}}(.)$	AMSE	AMAE
20×20	**A**	**1**	5	9.37	15.68 (8.18)	3.12 (0.99)
			50	9.64	2.01 (19.57)	1.19 (2.03)
		2	5	8.93	6.03 (5.16)	1.95 (0.83)
			50	9.22	2.47 (11.04)	1.31 (1.51)
		3		9.41	5.52 (17.48)	1.84 (1.81)
	B	**1**	5	7.88	15.46 (10.46)	3.46 (1.08)
			50	8.99	4.88 (19.35)	1.83 (1.93)
		2	5	8.31	9.33 (2.52)	2.59 (0.37)
			50	9.20	3.57 (11.05)	1.58 (1.43)
		3		8.93	7.47 (16.26)	2.16 (1.70)
30×30	**A**	**1**	5	9.67	13.38 (7.83)	2.97 (1.13)
			50	9.89	1.57 (20.47)	0.97 (2.12)
		2	5	9.92	5.02 (6.06)	1.79 (2.21)
			50	9.53	2.64 (10.54)	1.33 (1.48)
		3		9.08	3.78 (16.53)	1.58 (1.86)
	B	**1**	5	8.36	12.26 (11.56)	2.95 (1.27)
			50	9.85	2.61 (21.00)	1.29 (2.12)
		2	5	8.99	6.91 (14.98)	2.01 (1.68)
			50	9.21	3.05 (11.32)	1.52 (1.49)
		3		9.45	4.19 (15.64)	1.62 (1.81)

TABLE 4.2 – Simulation results of the $M = 100$ estimate of $\gamma(x_0)$ according to the models **A** and **B**, the cases **1**, **2** and **3** and the value of $\rho = 5$, 50 and $a = .98$: for each simulation scenario, from left to right, we have respectively, the estimate $\widehat{q}^{W}_{\mathbf{n}}(.)$, the average mean squared errors (AMSE), the average mean absolute errors (AMAE) and in brackets the corresponding standard deviation.

4.8 Conclusion and forthcoming studies

In this chapter, we propose an estimator of the tail index and extreme quantile of a heavy-tailed distribution for spatial data in the presence of covariates. We establish main consistency properties under general conditions by showing the convergence in probability of the Hill estimator deduced from a consistency result of the tail empirical measure. This convergence in probability is obtained by considering some $\alpha-$mixing condition on the underlying process in addition to the usual classical ones used in tail index estimation. Asymptotic normality of the tail index is also given. This last permits to derive asymptotic normality of the Weissman-type extreme quantile estimate. Numerical illustrations show that the Hill estimator works well when we are in the presence of spatial dependence. This numerical study will be extended to extreme quantile estimates. Other types (moment-type, UH-type,...) of tail estimator and corresponding quantile should be considered. As we consider here, fixed design covariates, a natural extension to the random covariate case will be a subject of our forthcoming studies.

4.9 Appendix : Technical lemmas and proofs of main results

4.9.1 Appendix A1 : Technical lemmas

In this section, we establish for simplicity the proofs of Theorem 4.6.1 and Theorem 4.6.2.

We need the three following lemmas whose proofs are omitted :

Lemma 4.9.1 (Tran (1990)). ***(i)*** *Suppose that* (4.8) *holds. Denote by* $\mathcal{L}_r(\mathcal{F})$ *the class of* $\mathcal{F}$*-measurable r.v's* X *satisfying* $\|X\|_r = (E|X|^r)^{\frac{1}{r}} < \infty$. *Suppose* $X_1 \in \mathcal{L}_r(\mathcal{B}(\mathbf{I}_1))$ *and* $X_2 \in \mathcal{L}_r(\mathcal{B}(\mathbf{I}_2))$, *where* $\mathbf{I}_1$, $\mathbf{I}_2$ *are two sets of spatial sites. Assume also that* $1 \le r, s, t < \infty$ *and* $r^{-1} + s^{-1} + t^{-1} = 1$. *Then*

$$|\mathbf{E}X_1X_2 - \mathbf{E}X_1\mathbf{E}X_2| \le C\|X_1\|_r\|X_2\|_s\left\{\psi\left(Card(\mathbf{I}_1), Card(\mathbf{I}_2)\right)\varphi\left(\mathrm{d}(\mathbf{I}_1,\mathbf{I}_2)\right)\right\}^{\frac{1}{t}}.$$

(ii) *For r.v's bounded with probability* 1 *the right-hand side of the last inequality can be replaced by*

$$C\psi\left(Card(\mathbf{I}_1), Card(\mathbf{I}_2)\right)\varphi\left(\mathrm{d}(\mathbf{I}_1,\mathbf{I}_2)\right).$$

Lemma 4.9.2 (Tran (1990)). *Let* $(\xi_1, \cdots, \xi_n)$ *be a random vector such that* $\left|\mathbf{E}\left[\prod_{s=i}^{n} \xi_i\right]\right| < \infty$, $i = 1, \cdots, n$, $|C\xi_i| \le 1$, $i = 1, \cdots, n$. *Then*

$$\left|\mathbf{E}\left[\prod_{s=1}^{n} \xi_s\right] - \prod_{s=1}^{n}\mathbf{E}\left[\xi_s\right]\right| \le \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left|\mathbf{E}\left[(\xi_i-1)(\xi_j-1)\right] \times \prod_{s=j+1}^{n}\xi_s - \mathbf{E}\left[(\xi_i-1)\right]\mathbf{E}\left[(\xi_j-1)\right]\prod_{s=j+1}^{n}\xi_s\right|.$$

Lemma 4.9.3 (Carbon et al. (1996)). *Suppose* $S_1, S_2, \cdots, S_r$ *be sets containing* m *sites each with* $dist(S_i, S_j) \ge \delta$, $\delta > 0$ *for all* $i \ne j$ *where* $1 \le i, j \le r$. *Suppose* $Y_1, \cdots, Y_r$ *is a sequence of real-valued random variables measurable with respect to* $\mathcal{B}(S_1), \cdots, \mathcal{B}(S_r)$ *respectively, and* Y_i *takes values in* $[a; b]$. *Then there exists a sequence of independent random variables* $Y_1^*, \cdots, Y_r^*$ *independent of* $Y_1, \cdots, Y_r$ *such that* Y_i^* *has the same distribution as* Y_i *and satisfies :*

$$\sum_{i=1}^{r}\mathbb{E}\left|Y_i - Y_i^*\right| \le 2r\,(b-a)\,\Phi\left((r-1)m, m\right)\varphi\left(\delta\right). \tag{4.30}$$

4.9.2 Appendix A2 : Definitions and notations

In this section, we establish the proofs of our main results by employing the blocking technique used in Carbon et al. (1997) and Tran (1990). Without loss of generality assume that $\sum\limits_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(\frac{x-x_\mathbf{i}}{r_\mathbf{n}}\right) := m_x = \widehat{\mathbf{r}}(p+q)^N$ where $\widehat{\mathbf{r}} = r_1 \times ... \times r_N$. Denote

$$\begin{aligned}
\mathbf{I}(1,x,\mathbf{j}) &= \{\mathbf{i},\ j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\ k=1,\cdots,N\} \qquad (4.31)\\
\mathbf{I}(2,x,\mathbf{j}) &= \Big\{\mathbf{i},\ j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\ k=1,\cdots,N-1,\\
&\quad j_N(p+q)+p+1 \le i_N \le (j_N+1)(p+q)\Big\},\\
\mathbf{I}(3,x,\mathbf{j}) &= \Big\{\mathbf{i},\ j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\ k=1,\cdots,N-2,\\
&\quad j_{N-1}(p+q)+p+1 \le i_{N-1} \le (j_{N-1}+1)(p+q),\\
&\quad j_N(p+q)+1 \le i_N \le j_N(p+q)+p\Big\},\\
\mathbf{I}(4,x,\mathbf{j}) &= \Big\{\mathbf{i},\ j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\ k=1,\cdots,N-2,\\
&\quad j_{N-1}(p+q)+p+1 \le i_{N-1} \le (j_{N-1}+1)(p+q),\\
&\quad j_N(p+q)+p+1 \le i_N \le (j_N+1)(p+q)\Big\}
\end{aligned}$$

and so on. The last two terms are

$$\begin{aligned}
\mathbf{I}(2^{N-1},x,\mathbf{j}) &= \Big\{\mathbf{i},\ j_k(p+q)+p+1 \le i_k \le (j_k+1)(p+q), k=1,\cdots,N-1,\\
&\quad j_N(p+q)+1 \le i_N \le j_N(p+q)+p\Big\} \qquad (4.32)
\end{aligned}$$

and

$$\mathbf{I}(2^N,x,\mathbf{j}) = \{\mathbf{i},\ j_k(p+q)+p+1 \le i_k \le (j_k+1)(p+q),\ k=1,\cdots,N,\} \qquad (4.33)$$

For each integer $1 \le i \le 2^N$, define

$$D(i,x) = \{\mathbf{I}(i,x,\mathbf{j}),\ 0 \le j_k \le r_k - 1,\ k=1,\cdots,N\} \qquad (4.34)$$

Note that the set $\mathcal{I}_\mathbf{n}$ is then decomposed into these 2^N small and large blocks in sets $D(i)$. If it is not the case that $m_x = \widehat{\mathbf{r}}(p+q)^N$ then an additional set $D(2^N+1,x)$ containing all the sites of $\mathcal{I}_\mathbf{n}$ not in these blocks $\mathbf{I}(i,x,\mathbf{j})$.

In all the proof we use the decomposition (4.31) additionally to (4.32), (4.33) and (4.34).

4.9.3 Appendix A3 : Intermediate results and their proofs

In this section we will give some intermediate results and their proofs which will helps us to prove the asymptotic normality of the Hill's estimator $\widehat{\gamma}_{v_\mathbf{n}}(x)$ for $\alpha-$mixing process. For simplicity, we begin with some notations and definitions. First, we rewrite $\widehat{\overline{F}}_\mathbf{n}(\cdot,x)$ the estimator of $\overline{F}(\cdot,x)$ as follows :

$$\widehat{\overline{F}}_\mathbf{n}(t,x) = \pi_\mathbf{n}(t,x)/g_\mathbf{n}(x,r_\mathbf{n}),$$

where $\pi_\mathbf{n}(t,x) := \sum\limits_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(\frac{x-x_\mathbf{i}}{r_\mathbf{n}}\right) \mathbb{1}_{\{y_\mathbf{i}>t\}}$ and $g_\mathbf{n}(x,r_\mathbf{n}) := m_x = \sum\limits_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(\frac{x-x_\mathbf{i}}{r_\mathbf{n}}\right)$.
Let $\pi(t,x) = m_x\overline{F}(t,x)$ and $\widehat{\phi}_{v_\mathbf{n}}(x,r_\mathbf{n}) := m_x^{-1} \sum\limits_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K\left(\frac{x-x_\mathbf{i}}{r_\mathbf{n}}\right) \left(\log\frac{Y_\mathbf{i}}{v_\mathbf{n}}\right) \mathbb{1}_{\{Y_\mathbf{i}>v_\mathbf{n}\}}$.

In addition, we define the following quantities :

$$H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}}) := \sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\left(\frac{\pi_{\mathbf{n}}(v_{\mathbf{n}},x)-\mathbf{E}\left[\pi_{\mathbf{n}}(v_{\mathbf{n}},x)\right]}{\pi(v_{\mathbf{n}},x)}\right)$$
$$H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}}) := \sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\left(\frac{\widehat{\phi}_{v_{\mathbf{n}}}(x,r_{\mathbf{n}})-\mathbf{E}\left[\widehat{\phi}_{v_{\mathbf{n}}}(x,r_{\mathbf{n}})\right]}{\pi(v_{\mathbf{n}},x)}\right).$$

We also introduce the following notations :

$$\widetilde{\eta}_{\mathbf{i},\mathbf{n},x} := \frac{1}{\overline{F}(v_{\mathbf{n}},x)}K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\left[\log\left(Y_{\mathbf{i}}/v_{\mathbf{n}}\right)\right]\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}},\ \widetilde{\Delta}_{\mathbf{i},\mathbf{n},x} := \widetilde{\eta}_{\mathbf{i},\mathbf{n},x}-\mathbf{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right].$$
$$\eta_{\mathbf{i},\mathbf{n},x} := \frac{1}{\overline{F}(v_{\mathbf{n}},x)}K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}},\ \Delta_{\mathbf{i},\mathbf{n},x} := \eta_{\mathbf{i},\mathbf{n},x}-\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},x}\right]$$
$$\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}}) = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\frac{\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}}{m_x}\ \text{ and }\ S_{\mathbf{n}}(x,r_{\mathbf{n}}) = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\frac{\Delta_{\mathbf{i},\mathbf{n},x}}{m_x}.$$

Then we have :

$$H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}}) = m_x^{1/2}S_{\mathbf{n}}(x,r_{\mathbf{n}})\ \text{ and }\ H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}}) = m_x^{1/2}\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}}).$$

The following lemma gives the variance of $\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}})$ and $S_{\mathbf{n}}(x,r_{\mathbf{n}})$.

Lemma 4.9.4. *Under model (4.5), (4.8) and the conditions in Theorem 4.6.1. We have*

1. $m_x\mathrm{Var}\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}}) = 2\gamma^2(x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}$.
2. $m_x\mathrm{Var}S_{\mathbf{n}}(x,r_{\mathbf{n}}) = 1+o(1)$.

Proof of Lemma 4.9.4

1. In this statement, before going further, let us denote

$$\widetilde{I}_{\mathbf{n},\gamma,x} := m_x^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\mathrm{Var}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right)\quad\text{and}$$
$$\widetilde{R}_{\mathbf{n},\gamma,x} := m_x^{-2}\sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right)\right|.$$

In this part, we give the variance of $\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}})$. Recall that $\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x} := \widetilde{\eta}_{\mathbf{i},\mathbf{n},x}-\mathbf{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right]$, where $\widetilde{\eta}_{\mathbf{i},\mathbf{n},x} := \frac{1}{\overline{F}(v_{\mathbf{n}},x)}K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\log\left(Y_{\mathbf{i}}/v_{\mathbf{n}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}$. It is easily seen that

$$\mathrm{Var}\left(m_x^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}\right) = m_x^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\mathrm{Var}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right)+m_x^{-2}\sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right)$$

Clearly, $\mathrm{Var}\left(m_x^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}\right) \leq \widetilde{I}_{\mathbf{n},\gamma,x} + \widetilde{R}_{\mathbf{n},\gamma,x}$. Observe that, for all the covariates recorded in the ball $B\left(x,r_{\mathbf{n}}\right)$,

$$\widetilde{I}_{\mathbf{n},\gamma,x} = \frac{1}{m_x\left[\overline{F}(v_{\mathbf{n}},x)\right]^2}\left(\mathbf{E}\left[\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)^2\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]-\mathbf{E}\left[\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]^2\right).$$

From Lemma 1 of Goegebeur et al. (2014b) and the second order condition $(\mathcal{R}_2)$ together with all the covariates in the ball, we have

$$\mathbf{E}\left[\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)^2 \mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] = 2\gamma^2(x)\overline{F}(v_{\mathbf{n}},x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\},$$

and the first statement of Lemma 4.9.5 below, we get

$$\mathbf{E}\left[\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right) \mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] = \gamma(x)\overline{F}(v_{\mathbf{n}},x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\},$$

It follows that

$$\begin{aligned}\widetilde{I}_{\mathbf{n},\gamma,x} &= 2\gamma^2(x)\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1} m_x^{-1}\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}\\ &\quad-\gamma^2(x)m_x^{-1}\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\}^2.\end{aligned} \tag{4.35}$$

Now we will concentrate on the term say $\widetilde{R}_{\mathbf{n},\gamma,x}$. We seek to prove that, for all the covariates recorded in the ball $B(x,r_{\mathbf{n}})$, there exists a constant C not depending on $\hat{\mathbf{n}}$ such that $\widetilde{R}_{\mathbf{n},\gamma,x} \leq Cm_x^{-1}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1}$ for $\hat{\mathbf{n}}$ large enough. Recall that the covariance between $\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}$ and $\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}$ is written as

$$\begin{aligned}\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right) &= m_x^{-2}\mathbf{E}\left[\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\left(\log\frac{Y_{\mathbf{j}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\left\{Y_{\mathbf{j}}>v_{\mathbf{n}}\right\}}\right]\\ &\quad- m_x^{-2}\mathbf{E}\left[\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\mathbf{E}\left[\left(\log\frac{Y_{\mathbf{j}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\left\{Y_{\mathbf{j}}>v_{\mathbf{n}}\right\}}\right].\end{aligned}$$

We will study the covariance between $\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}$ and $\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}$ for all $\mathbf{i}\neq\mathbf{j}$ in two sets defined as follows, we set : $\mathcal{S}_{\Omega_{\mathbf{n}}} = \{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}} : 0<\mathrm{d}(\mathbf{i},\mathbf{j})<\Omega_{\mathbf{n}}\}$ and denote by $\mathcal{S}^c_{\Omega_{\mathbf{n}}}$ the complement of $\mathcal{S}_{\Omega_{\mathbf{n}}}$ where $\Omega_{\mathbf{n}}$ is a sequence of real numbers which tends to ∞ as $\mathbf{n}\to\infty$. Define also the two following quantities

$$\begin{aligned}\widetilde{R}_{1,\mathbf{n},\gamma,x} &:= m_x^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\eta_{\mathbf{j},\mathbf{n},x}\right)\right| \quad \text{and,}\\ \widetilde{R}_{2,\mathbf{n},\gamma,x} &:= m_x^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^c_{\Omega_{\mathbf{n}}}}\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right)\right|.\end{aligned}$$

Clearly, $\widetilde{R}_{\mathbf{n},\gamma,x} \leq \widetilde{R}_{1,\mathbf{n},\gamma,x} + \widetilde{R}_{2,\mathbf{n},\gamma,x}$.
Note that the key to this proof lies in the Lemma 4.9.1. Now applying the first part of Lemma 4.9.1, on the $\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}$, we get,

$$\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right)\right| \leq \|\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\|_s\|\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\|_u(\psi(1,1)\varphi(\|\mathbf{i}-\mathbf{j}\|))^{1/t},$$

with $t^{-1}+s^{-1}+u^{-1}=1$, we have

$$\begin{aligned}\|\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\|_s^s &= \left[\overline{F}(v_{\mathbf{n}},x)\right]^{-s}\mathbf{E}\left[\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)^s\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\\ &= s!\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1-s}\gamma^s(x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^s}-1\right](1+o(1))\right\}.\end{aligned}$$

It follows that

$$\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}, \widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right)\right| \leq \left[\overline{F}\left(v_{\mathbf{n}}, x\right)\right]^{-\frac{1}{t}-1} \gamma^2(x)\widetilde{\varepsilon}_{s,u}(x)\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}, \tag{4.36}$$

where

$$\begin{aligned}\widetilde{\varepsilon}_{s,u}(x) := (s!)^{1/s}&\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^s}-1\right](1+o(1))\right\}^{\frac{1}{s}} \\ &\times (u!)^{1/u}\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^u}-1\right](1+o(1))\right\}^{\frac{1}{u}}.\end{aligned}$$

- First, concerning $\widetilde{R}_{1,\mathbf{n},\gamma,x}$, we have by setting $u=s=4$, $t=2$

$$\begin{aligned}\widetilde{R}_{1,\mathbf{n},\gamma,x} &\leq m_x^{-2}\|\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\|_4\|\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\|_4\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/2} \\ &\leq m_x^{-2}\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-\frac{3}{2}}\gamma^2(x)\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^N\sum_{\|\mathbf{i}\|\leq\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/2}. \\ &\leq m_x^{-2}\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-\frac{3}{2}}\gamma^2(x)\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^N\sum_{\|\mathbf{i}\|\leq\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/2}.\end{aligned}$$

$$m_x\overline{F}\left(v_{\mathbf{n}},x\right)\widetilde{R}_{1,\mathbf{n},\gamma,x} \leq m_x^{-1}\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-\frac{1}{2}}\gamma^2(x)\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^N\sum_{\mathbf{i}>0}\|\mathbf{i}\|^{-\theta/2}. \tag{4.37}$$

- For the term $R_{2,\mathbf{n},\gamma,x}$, we have

$$\begin{aligned}\widetilde{R}_{2,\mathbf{n},\gamma,x} &\leq m_x^{-2}\|\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\|_4\|\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\|_4\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^c_{\Omega_{\mathbf{n}}}}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/2} \\ &\leq m_x^{-1}\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-\frac{3}{2}}\gamma^2(x)\widetilde{\varepsilon}_{s,u}(x)\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/2}.\end{aligned}$$

$$\begin{aligned}m_x\overline{F}\left(v_{\mathbf{n}},x\right)\widetilde{R}_{2,\mathbf{n},\gamma,x} &\leq \left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-\frac{1}{2}}\gamma^2(x)\widetilde{\varepsilon}_{s,u}(x)\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{-N/2}\|\mathbf{i}\|^{N/2}\varphi\left(\|\mathbf{i}\|\right)^{1/2} \\ &\leq C\gamma^2(x)\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-\frac{1}{2}}\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^{-N/2}\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{\frac{N-\theta}{2}}.\end{aligned} \tag{4.38}$$

By assumptions of Theorem 4.6.1, we have $\widetilde{\varepsilon}_{s,u}(x)\longrightarrow c\in\mathbb{R}_+^*$ as $\mathbf{n}\to\infty$. Then combining (4.37) and (4.38), since $\theta>2(N+1)$ and $\Omega_{\mathbf{n}}=\left(m_x\overline{F}(v_{\mathbf{n}},x)^{1/2}\right)^{1/N}$, it follows that

$$\widetilde{R}_{\mathbf{n},\gamma,x}=O\left(\frac{1}{m_x\overline{F}\left(v_{\mathbf{n}},x\right)}\right). \tag{4.39}$$

The proof of statement **1** follows from (4.35), (4.39) and by the fact that

$$\begin{aligned}m_x\mathrm{Var}\left(\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}})\right) &= m_x\overline{F}\left(v_{\mathbf{n}},x\right)\mathrm{Var}\left(m_x^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}\right) \\ &= 2\gamma^2(x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left(\frac{1}{(1-\rho(x))^2}-1\right)(1+o(1))\right\} \\ &\quad -\gamma^2(x)\overline{F}\left(v_{\mathbf{n}},x\right)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left(\frac{1}{(1-\rho(x))}-1\right)(1+o(1))\right\}^2\end{aligned}$$

2. We will compute in the same way as statement **1** the variance of $S_{\mathbf{n}}(x, r_{\mathbf{n}})$. Notice that, in the proof we consider all the covariates recorded in the ball $B(x, r_{\mathbf{n}})$ as previously. First, define

$$I_{\mathbf{n},x} := m_x^{-2} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \operatorname{Var}(\eta_{\mathbf{i},\mathbf{n},x}) \qquad \text{and} \qquad R_{\mathbf{n},x} := m_x^{-2} \sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} \left|\operatorname{Cov}(\eta_{\mathbf{i},\mathbf{n},x}, \eta_{\mathbf{j},\mathbf{n},x})\right|.$$

Carrying over the notation from before, $\eta_{\mathbf{i},\mathbf{n},x} := \left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1} \mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}$ and $\Delta_{\mathbf{i},\mathbf{n},x} := \eta_{\mathbf{i},\mathbf{n},x} - \mathbb{E}\left[\eta_{\mathbf{i},\mathbf{n},x}\right]$, it is easily seen that

$$\operatorname{Var}\left(m_x^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n},x}\right) = m_x^{-2} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \operatorname{Var}(\eta_{\mathbf{i},\mathbf{n},x}) + m_x^{-2} \sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} \sum_{\mathbf{i}\neq\mathbf{j}} \operatorname{Cov}(\eta_{\mathbf{i},\mathbf{n},x}, \eta_{\mathbf{j},\mathbf{n},x}).$$

Clearly, $\operatorname{Var}\left(m_x^{-1} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{n},x}\right) \leq I_{\mathbf{n},x} + R_{\mathbf{n},x}$.

First, we have

$$I_{\mathbf{n},x} = \frac{\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-2}}{m_x} \left(\mathbf{E}\left[\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] - \mathbf{E}\left[\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]^2\right).$$

We get also

$$\mathbf{E}\left[K^2\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] = K^2\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\overline{F}(v_{\mathbf{n}}, x) = \overline{F}(v_{\mathbf{n}}, x)$$

and

$$\mathbf{E}\left[K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] = K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\overline{F}(v_{\mathbf{n}}, x) = \overline{F}(v_{\mathbf{n}}, x). \tag{4.40}$$

It follows that

$$I_{\mathbf{n},x} = \frac{1}{m_x \overline{F}(v_{\mathbf{n}}, x)} \left(1 - \overline{F}(v_{\mathbf{n}}, x)\right). \tag{4.41}$$

We now treat the term $R_{\mathbf{n},x}$, we want to show that there exists a constant C not depending on $\hat{\mathbf{n}}$ such that $R_{\mathbf{n},x} \leq C m_x^{-1} \left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1}$ for $\hat{\mathbf{n}}$ large enough.

$$\begin{aligned}\operatorname{Cov}(\eta_{\mathbf{i},\mathbf{n},x}, \eta_{\mathbf{j},\mathbf{n},x}) = {} & m_x^{-2}\mathbf{E}\left[\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1}\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1}\mathbb{1}_{\{Y_{\mathbf{j}}>v_{\mathbf{n}}\}}\right] \\ & - m_x^{-2}\mathbf{E}\left[\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1}\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\mathbf{E}\left[\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1}\mathbb{1}_{\{Y_{\mathbf{j}}>v_{\mathbf{n}}\}}\right]\end{aligned}$$

We will study the covariance between $\eta_{\mathbf{i},\mathbf{n},x}$ and $\eta_{\mathbf{j},\mathbf{n},x}$ for all $\mathbf{i}\neq\mathbf{j}$ in two sets $\mathcal{S}_{\Omega_{\mathbf{n}}}$ and $\mathcal{S}^c_{\Omega_{\mathbf{n}}}$ defined in the previous calculation of the covariance. Define

$$R_{1,\mathbf{n},x} := m_x^{-2} \sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}} \left|\operatorname{Cov}(\eta_{\mathbf{i},\mathbf{n},x}, \eta_{\mathbf{j},\mathbf{n},x})\right| \text{ and } R_{2,\mathbf{n},x} := m_x^{-2} \sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^c_{\Omega_{\mathbf{n}}}} \left|\operatorname{Cov}(\eta_{\mathbf{i},\mathbf{n},x}, \eta_{\mathbf{j},\mathbf{n},x})\right|.$$

Clearly, $R_{\mathbf{n},x} \leq R_{1,\mathbf{n},x} + R_{2,\mathbf{n},x}$. Applying the first part of Lemma 4.9.1, on the $\eta_{\mathbf{i},\mathbf{n},x}$, we get,

$$\left|\operatorname{Cov}(\eta_{\mathbf{i},\mathbf{n},x}, \eta_{\mathbf{j},\mathbf{n},x})\right| \leq \|\eta_{\mathbf{i},\mathbf{n},x}\|_s \|\eta_{\mathbf{j},\mathbf{n},x}\|_u (\psi(1,1)\varphi(\|\mathbf{i}-\mathbf{j}\|))^{1/t},$$

with $t^{-1}+s^{-1}+u^{-1}=1$, in particular, we choose $s=u=4$ and $t=2$. We have

$$\|\eta_{\mathbf{i},\mathbf{n},x}\|_s^s=\mathbf{E}\left[\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-1}\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]=\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{1-s}.$$

It follows that

$$\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},x},\eta_{\mathbf{j},\mathbf{n},x}\right)\right|\leq\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-2}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/s+1/u}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}.$$

- ▶ Let $u=s=4$, $t=2$, first, concerning $R_{1,\mathbf{n},x}$, we have

$$\begin{aligned}
R_{1,\mathbf{n},x}&\leq Cm_x^{-2}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1-\frac{1}{2}}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/2}\\
&\leq Cm_x^{-2}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1-\frac{1}{2}}\Omega_{\mathbf{n}}^{N}\sum_{\|\mathbf{i}\|<\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/2}\\
m_x\overline{F}(v_{\mathbf{n}},x)R_{1,\mathbf{n},x}&\leq Cm_x^{-1}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-\frac{1}{2}}\Omega_{\mathbf{n}}^{N}\sum_{\mathbf{i}>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{-\theta/2}.
\end{aligned}\tag{4.42}$$

- ▶ For the term $R_{2,\mathbf{n},x}$, we have

$$\begin{aligned}
R_{2,\mathbf{n},x}&\leq Cm_x^{-2}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1-\frac{1}{2}}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^c_{\Omega_{\mathbf{n}}}}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/2}\\
&\leq Cm_x^{-1}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1-\frac{1}{2}}\Omega_{\mathbf{n}}^{-N/2}\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{\frac{N-\theta}{2}}\\
m_x\overline{F}(v_{\mathbf{n}},x)R_{2,\mathbf{n},x}&\leq C\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-\frac{1}{2}}\Omega_{\mathbf{n}}^{-N/2}\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{\frac{N-\theta}{2}}.
\end{aligned}\tag{4.43}$$

By assumptions of Theorem 4.6.1 and combining (4.42) and (4.43), since $\theta>2(N+1)$ and $\Omega_{\mathbf{n}}=\left(m_x\overline{F}(v_{\mathbf{n}},x)^{1/2}\right)^{1/N}$, it follows that

$$\widetilde{R}_{\mathbf{n},x}=O\left(\frac{1}{m_x\overline{F}\left(v_{\mathbf{n}},x\right)}\right).\tag{4.44}$$

The proof of statement **2** follows from (4.41), (4.44) and by the fact that

$$m_x\operatorname{Var}\left(S_{\mathbf{n}}(x,r_{\mathbf{n}})\right)=m_x\overline{F}(v_{\mathbf{n}},x)\operatorname{Var}\left(m_x^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n},\tau}\right)=1-\frac{\overline{F}(v_{\mathbf{n}},x)}{m_x}+o(1).\blacksquare$$

Lemma 4.9.5. *Under model (4.5), (4.8) and the conditions in Theorem 4.6.1, we have*

1. $H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}})\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\ 2\gamma^2(x)\right)$ *and* $H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}})\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\ 1\right)$.
2. *Moreover as* $\mathbf{n}\to\infty$, $H_{\mathbf{n}}(\widehat{\gamma}_{v_{\mathbf{n}}},x,r_{\mathbf{n}}):=\left(H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}}),H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}})\right)^{\top}$ *converges in distribution to a bivariate Gaussian vector* $\mathcal{N}\left(0,\Sigma\right)$, *where*

$$\Sigma:=\begin{pmatrix}1&\gamma(x)\\\gamma(x)&2\gamma^2(x)\end{pmatrix}$$

and $\top$ *denotes the transpose.*

Proof of Lemma 4.9.5

1.

▶ Concerning the asymptotic normality of $H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}})$, with the notations that are given at the beginning of this section, it follows that

$$H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}})=m_x^{1/2}\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}})\ \text{ where }\quad \widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}}):=m_x^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x},$$

$$\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}:=\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}-\mathbb{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right]\ \text{ and }\ \widetilde{\eta}_{\mathbf{i},\mathbf{n},x}:=\frac{1}{\overline{F}(v_{\mathbf{n}},x)}K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}.$$

First, using (4.24), one can find sequences of positive integers $p_{\mathbf{n}}=p$, $q_{\mathbf{n}}=q$ tending to infinity such that $qp^{-1}=o(1)$. For some integers $r_1,\cdots,r_N$, we assume that $n_1=r_1(q+p),\cdots,n_N=r_N(q+p)$.

Then, the random variables $\widetilde{A}_{\mathbf{i},\mathbf{n},x}:=\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}$'s are now set into large blocks and small blocks, that is

$$\begin{aligned}\widetilde{U}_{\mathbf{n}}(1,\gamma,\mathbf{j},x)&=\sum_{\mathbf{i}\in I(1,x,\mathbf{j})}\widetilde{A}_{\mathbf{i},\mathbf{n},x},\quad \widetilde{U}_{\mathbf{n}}(2,\gamma,\mathbf{j},x)=\sum_{I(2,x,\mathbf{j})}\widetilde{A}_{\mathbf{i},\mathbf{n},x}\\ \widetilde{U}_{\mathbf{n}}(3,\gamma,\mathbf{j},x)&=\sum_{I(3,x,\mathbf{j})}\widetilde{A}_{\mathbf{i},\mathbf{n},x},\quad \widetilde{U}_{\mathbf{n}}(4,\gamma,\mathbf{j},x)=\sum_{I(4,x,\mathbf{j})}\widetilde{A}_{\mathbf{i},\mathbf{n},x}\end{aligned}$$

and so on. The last two terms are

$$\widetilde{U}_{\mathbf{n}}\left(2^{N-1},\gamma,\mathbf{j},x\right)=\sum_{I(2^{N-1},x,\mathbf{j})}\widetilde{A}_{\mathbf{i},\mathbf{n},x},\ \widetilde{U}_{\mathbf{n}}\left(2^{N},\gamma,\mathbf{j},x\right)=\sum_{I(2^{N},x,\mathbf{j})}\widetilde{A}_{\mathbf{i},\mathbf{n},x}$$

For each integer $1\le i\le 2^N$, define

$$\widetilde{T}_{\mathbf{n},\gamma,x,i}=\sum_{\mathbf{j}\in D(i,x)}\widetilde{U}_{\mathbf{n}}(i,\gamma,\mathbf{j},x).\tag{4.45}$$

Denote $\widetilde{K}_{\mathbf{n},\gamma,x}:=\sum_{i=1}^{2^N}\widetilde{T}_{\mathbf{n},\gamma,x,i}$. With this notation, $\widetilde{T}_{\mathbf{n},\gamma,x,1}$ is the sum of the random variables $\widetilde{\Delta}_{\mathbf{i},\mathbf{n}}$ in large blocks. For $2\le i\le 2^N$, the $\widetilde{T}_{\mathbf{n},\gamma,x,i}$ are sums of random variables in small blocks. If it is not the case that $n_1=r_1(p+q),\cdots,n_N=r_N(p+q)$ for some integers $r_1,\cdots,r_N$, then a term, say, $\widetilde{T}_{\mathbf{n},\gamma,x,2^N+1}$ containing all the $\widetilde{A}_{\mathbf{i},\mathbf{n},x}$'s at the end not included in the big or small blocks can be added. This term will not change the proof much. Using Lemma 4.9.4, the general approach is to show that as $\mathbf{n}\to\infty$,

$$\widetilde{Q}_{1,\mathbf{n},\gamma,x}\equiv\left|\mathbf{E}\left[\exp\left(iz\widetilde{T}_{\mathbf{n},\gamma,x,1}\right)\right]-\prod_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbf{E}\left[\exp\left(iz\widetilde{U}_{\mathbf{n}}(1,\gamma,\mathbf{j},x)\right)\right]\right|\to 0,\tag{4.46}$$

$$\widetilde{Q}_{2,\mathbf{n},\gamma,x}\equiv m_x^{-1}\mathbf{E}\left[\left(\sum_{i=2}^{2^N}\widetilde{T}_{\mathbf{n},\gamma,x,i}\right)^2\right]\to 0,\tag{4.47}$$

$$\widetilde{Q}_{3,\mathbf{n},\gamma,x}\equiv m_x^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbf{E}\left[\widetilde{U}_{\mathbf{n}}(1,\gamma,\mathbf{j},x)^2\right]\to\widetilde{\sigma}^2(x),\tag{4.48}$$

$$\widetilde{Q}_{4,\mathbf{n},\gamma,x}\equiv m_x^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbf{E}\left[\widetilde{U}_{\mathbf{n}}(1,\gamma,\mathbf{j},x)^2\,\mathbb{1}_{\left\{\left|\widetilde{U}_{\mathbf{n}}(1,\gamma,\mathbf{j},x)\right|\ge\varepsilon\widetilde{\sigma}(x)m_x^{1/2}\right\}}\right]\to 0,\tag{4.49}$$

where $\widetilde{\sigma}(x) = 2\gamma^2(x)$ and for every $\varepsilon > 0$. We have

$$\frac{H_{v_{\mathbf{n}},\phi}(x, r_{\mathbf{n}})}{\widetilde{\sigma}(x)} = \frac{\widetilde{K}_{\mathbf{n},\gamma,x}}{\widetilde{\sigma}(x)\sqrt{m_x}} = \frac{\widetilde{T}_{\mathbf{n},\gamma,x,1}}{\widetilde{\sigma}(x)\ \sqrt{m_x}} + \sum_{i=2}^{2^N} \frac{\widetilde{T}_{\mathbf{n},\gamma,x,i}}{\widetilde{\sigma}(x)\sqrt{m_x}}.$$

To complete the proof of this lemma, we need to prove the following two lemmas.

Lemma 4.9.6. *Under conditions of Lemma 4.9.5, we have that as* $\mathbf{n} \to \infty$*,*

$$\lim_{\mathbf{n}\to\infty} \widetilde{Q}_{1,\mathbf{n},\gamma,x} = 0 \ \ \textit{and} \ \ \lim_{\mathbf{n}\to\infty} \widetilde{Q}_{2,\mathbf{n},\gamma,x} = 0.$$

This lemma shows that as $\mathbf{n} \to \infty$, the random variables $\widetilde{U}_{\mathbf{n}}\left(1, \gamma, x, \mathbf{j}\right)$ are asymptotically independent and the term $\sum_{i=2}^{2^N} \widetilde{T}_{\mathbf{n},\gamma,x,i} / \left(\widetilde{\sigma}(x) m_x^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible.

Lemma 4.9.7. *Under conditions of Lemma 4.9.5, we have that as* $\mathbf{n} \to \infty$*,*

$$\lim_{\mathbf{n}\to\infty} \widetilde{Q}_{3,\mathbf{n},\gamma,x} = \widetilde{\sigma}^2(x) \ \ \textit{and} \ \ \lim_{\mathbf{n}\to\infty} \widetilde{Q}_{4,\mathbf{n},\gamma,x} = 0.$$

where $\widetilde{\sigma}^2(x) = 2\gamma^2(x)$*.*

This last result permit to deduce the asymptotic normality of $\widetilde{T}_{\mathbf{n},\gamma,x,1} / \left(\widetilde{\sigma}(x) m_x^{1/2}\right)$, since (4.48) goes to zero and the Lindeberg-Feller condition is satisfied.

Proof of Lemma 4.9.6 First, we use Lemma 4.9.2 to enumerate the random variables $\widetilde{U}_{\mathbf{n}}\left(1, \gamma, \mathbf{j}, x\right)$ in an arbitrary order say, $U_1^*, \cdots, U_{\hat{Q}}^*$, where $\hat{Q} = \prod_{k=1}^{N} r_k = m_x(p+q)^{-N} \le m_x p^N$. Define the following sets :

$$\mathcal{S}_{\mathbf{n}}(1, \gamma, \mathbf{j}, x) = \left\{ i : j_k(q+p) + 1 \le i_k \le j_k(q+p) + p,\ \ k = 1, \cdots, N \right\}.$$

Note that the sets of distinct sites are distant from a distance at least q. Remark that the set $\mathcal{S}_{\mathbf{n}}(1, \gamma, \mathbf{j}, x)$ contains p^N sites, where $\mathcal{S}_{\mathbf{n}}(1, \gamma, \mathbf{j}, x)$ is the set of sites involved with $\widetilde{U}_{\mathbf{n}}\left(1, \gamma, \mathbf{j}, x\right)$. Now observe that, by Lemma 4.9.2, we have that :

$$\begin{aligned}\widetilde{Q}_{1,\mathbf{n},\gamma,x} \le \sum_{k=1}^{\hat{Q}-1} \sum_{j=k+1}^{\hat{Q}} \Bigg| \mathbf{E}\left[(\exp\left\{iuU_k^*\right\} - 1)(\exp\left\{iuU_j^*\right\} - 1) \right] \prod_{s=j+1}^{\hat{Q}} \exp\left\{iuU_s^*\right\} \\ - \mathbf{E}\left[(\exp\left\{iuU_k^*\right\} - 1) \right] \mathbf{E}\left[(\exp\left\{iuU_j^*\right\} - 1) \right] \prod_{s=j+1}^{\hat{Q}} \exp\left\{iuU_s^*\right\} \Bigg|.\end{aligned}$$

Let $\widetilde{\mathcal{S}}_j$ be the sets of sites involved with U_j^*. Using again Lemma 4.9.2, it follows that

$$\begin{aligned}&\Big|\mathbf{E}\left[(\exp\{iuU_k^*\}-1)(\exp\left\{iuU_j^*\right\}-1)\right]-\mathbf{E}\left[(\exp\{iuU_k^*\}-1)\right]\mathbf{E}\left[(\exp\left\{iuU_j^*\right\}-1)\right]\Big|\\ &\quad\le Cp^N\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_j,\widetilde{\mathcal{S}}_k)\right)\end{aligned}$$

thus

$$\begin{aligned}\widetilde{Q}_{1,\mathbf{n},\gamma,x} &\le Cp^N\sum_{k=1}^{\hat{Q}-1}\sum_{j=k+1}^{\hat{Q}}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_j,\widetilde{\mathcal{S}}_k)\right)\\ &\le Cp^N\hat{Q}\sum_{k=2}^{\hat{Q}}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)\right)\\ &\le Cp^N\hat{Q}\sum_{i=1}^{\infty}\sum_{k:iq\le \mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)<(i+1)q}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)\right)\\ &\le Cp^N\hat{Q}\sum_{i=1}^{\infty}i^{N-1}\varphi(iq)\le Cm_x\sum_{i=}^{\infty}i^{N-1}\varphi(iq)\\ &\le Cm_x\sum_{i=}^{\infty}i^{N-1-\theta}q^{-\theta}\le Cm_xq^{-\theta},\end{aligned}$$

this term tends to zero if one take q as stated in $(\mathcal{W})$ and $p=\left[m_x\overline{F}(v_{\mathbf{n}},x)\right]^{\frac{1}{2N}}$.

For the term (4.47), we want to show that :

$$m_x^{-1}\mathbf{E}\left[\widetilde{T}_{\mathbf{n},\gamma,x,i}\right]^2\longrightarrow 0\quad\text{for each}\quad 2\le i\le 2^N.$$

Without loss of generality, for simplicity, we only consider the case where $i=2$. Again, we enumerate as in proof of (4.46), the random varible's $\widetilde{U}_{\mathbf{n}}(2,\gamma,\mathbf{j},x)$ in an arbitrary way and refer them as $U_1^+,\cdots,U_{\hat{Q}}^+$. Then

$$\begin{aligned}\mathbf{E}\left[\widetilde{T}_{\mathbf{n},\gamma,x,2}\right]^2 &=\sum_{i=0}^{\hat{Q}}\mathrm{Var}\left(U_i^+\right)+2\sum_{j=0}^{\hat{Q}-1}\sum_{i=j+1}^{\hat{Q}}\mathrm{Cov}\left(U_i^+,U_j^+\right)\\ &:=\quad\Gamma_{1,\gamma,x}\quad+\quad\Gamma_{2,\gamma,x}\end{aligned}\tag{4.50}$$

Let us treat first the variance term, we have

$$\begin{aligned}\mathrm{Var}\left(U_i^+\right) &=\mathrm{Var}\left(\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}\widetilde{A}_{\mathbf{i},\mathbf{n},x}\right)^2\\ &:=\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}\mathrm{Var}(\widetilde{A}_{\mathbf{i},\mathbf{n},x})+\underset{i_k\neq j_k\text{ for some }1\le k\le N}{\sum_{\substack{j_k=1\\k=1,\cdots,N-1}}^{p}\sum_{j_N=1}^{q}\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}}\mathbf{E}\left[\widetilde{A}_{\mathbf{i},\mathbf{n},x}\widetilde{A}_{\mathbf{j},\mathbf{n},x}\right]\end{aligned}\tag{4.51}$$

Recall that $\widetilde{A}_{\mathbf{i},\mathbf{n},x}=\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}$, $\widetilde{\Delta}_{\mathbf{i},\mathbf{n},x}=\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}-\mathbb{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right]$ and $\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}=$

$\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1} K\left(\frac{x-x_{\mathbf{i}}^*}{r_{\mathbf{n}}}\right)\left[\log\left(Y_{\mathbf{i}}/v_{\mathbf{n}}\right)\right]_+$, and combining with (4.35), it follows that

$$\begin{aligned}\operatorname{Var}(\widetilde{A}_{\mathbf{i},\mathbf{n},x}) &= \overline{F}(v_{\mathbf{n}},x)\mathbf{E}\left[\widetilde{\Delta}^2_{\mathbf{i},\mathbf{n},x}\right]\\ &\le 2\gamma^2(x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}\le C. \end{aligned}\tag{4.52}$$

Let $\chi = 2(1-\varsigma)/\varsigma$, $0<\varsigma<1$, now by Lemma 4.9.1 second statement and Lemma 1 of Goegebeur et al. (2014b), we have that

$$\begin{aligned}\mathbf{E}\left|\widetilde{A}_{\mathbf{i},x}\widetilde{A}_{\mathbf{j},x}\right| &\le C\left(\int_{\mathbb{R}}\left|\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-1/2}\log\left(\frac{t}{v_{\mathbf{n}}}\right)_+\right|^{2+\chi} f(t)dt\right)^{\varsigma}\varphi(\|\mathbf{i}-\mathbf{j}\|)^{1-\varsigma}\\ &\le 4C\gamma^4(x)\varphi(\|\mathbf{i}-\mathbf{j}\|)^{1-\varsigma}\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\},\end{aligned}\tag{4.53}$$

Combining (4.51), (4.52) and (4.53), it follows that

$$\begin{aligned}\operatorname{Var}\left(U_i^+\right) &\le Cp^{N-1}q\left(1+4\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}\varphi(\|\mathbf{i}\|)^{1-\varsigma}\right)\\ &\le 4Cp^{N-1}q\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{p}\sum_{i_N=1}^{q}\varphi(\|\mathbf{i}\|)\\ &\le 4Cp^{N-1}q\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma},\end{aligned}\tag{4.54}$$

where $\widetilde{\varepsilon}_1(x) := 1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))$. By (4.50)-(4.54), it follows that :

$$\Gamma_{1,\gamma,x}\le 4C\hat{Q}p^{N-1}q\gamma^4(x)\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}.\tag{4.55}$$

Let

$$\begin{aligned}\widetilde{\mathcal{S}}_{\mathbf{n}}(2,x,\mathbf{j}) \;=\; \Big\{\mathbf{i} : {} & j_k(p+q+1)\le i_k\le j_k(p+q)+p,\ 1\le k<N\\ & j_k(p+q)+p+1\le i_N\le (j_N+1)(p+q)\Big\}\end{aligned}$$

Clearly, $\widetilde{U}_{\mathbf{n}}(2,\gamma,\mathbf{j},x)$ is the sum of $\widetilde{A}_{\mathbf{i},\mathbf{n},x}$ with sites in $\mathcal{S}_{\mathbf{n}}(2,\gamma,\mathbf{j},x)$. Since $p>q$, if $\mathbf{j}$ and $\mathbf{j}'$ belong to two distinct sets $\mathcal{S}_{\mathbf{n}}(2,\gamma,\mathbf{j},x)$ and $\mathcal{S}_{\mathbf{n}}(2,\gamma,\mathbf{j}',x)$, then $j_k\neq j_k^{'}$ for some $1\le k\le N$ and $\|\mathbf{j}-\mathbf{j}'\|>q$. Then by (4.53), we have

$$\begin{aligned}\Gamma_{2,\gamma,x} &\le C\sum_{\substack{j_k=1\\k=1,\cdots,N}}^{n_k}\sum_{\substack{i_k=1\\k=1,\cdots,N\\ \|\mathbf{i}-\mathbf{j}\|>q}}^{n_k}\mathbf{E}\left[\widetilde{A}_{\mathbf{i},\mathbf{n},x}\widetilde{A}_{\mathbf{j},\mathbf{n},x}\right]\\ &\le 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)m_x\sum_{\substack{i_k=1\\k=1,\cdots,N;\ \|\mathbf{i}\|>q}}^{n_k}\varphi(\|\mathbf{i}\|)^{1-\varsigma}\\ &\le 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)m_x\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}.\end{aligned}\tag{4.56}$$

From (4.50), (4.55) and (4.56), we get

$$\begin{aligned}
m_x^{-1}\mathbf{E}\left[\widetilde{T}_{\mathbf{n},\gamma,x,2}\right]^2 &\leq 4C\hat{Q}p^{N-1}qm_x^{-1}\gamma^4(x)\sum_{i=1}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma} + 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\sum_{i=q}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma}\\
&\leq 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\left(\hat{Q}p^{N-1}qm_x^{-1}\sum_{i=1}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma} + \sum_{i=q}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma}\right)\\
&\leq 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\left(q/p\sum_{i=1}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma} + \sum_{i=q}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma}\right)\\
&\leq 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\left(q/p+1\right)\sum_{i=1}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma}\\
&\leq 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\left(q/p+1\right)\sum_{i=1}^{\infty} i^{N-1-\theta(1-\varsigma)}.
\end{aligned}$$

Which tends to zero it suffices to take $0 < \varsigma < \frac{\theta - N}{\theta}$.■

Proof of Lemma 4.9.7 To prove that $\widetilde{Q}_{3,\mathbf{n},\gamma,x} \to \widetilde{\sigma}^2(x)$ as $\mathbf{n} \to \infty$, let us denote by

$$\widetilde{K}_{\mathbf{n},\gamma,x,1} = \widetilde{T}_{\mathbf{n},\gamma,x,1} \text{ and } \widetilde{K}_{\mathbf{n},\gamma,x,2} = \sum_{i=2}^{2^N} \widetilde{T}_{\mathbf{n},\gamma,x,i}. \tag{4.57}$$

The sum of the random variables $\widetilde{A}_{\mathbf{1},\mathbf{n},x}$ in large blocks is assigned to $\widetilde{K}_{\mathbf{n},\gamma,x,1}$ and $\widetilde{K}_{\mathbf{n},\gamma,x,2}$ is the sum of random variables $\widetilde{A}_{\mathbf{i},\mathbf{n},x}$ in small blocks. Lemma 4.9.4 statement 1 implies that $m_x^{-1}\mathbf{E}\left[\widetilde{S}_{\mathbf{n}}(x,r_{\mathbf{n}})^2\right] \longrightarrow \widetilde{\sigma}^2(x)$. This combining with (4.47) shows that $m_x^{-1}\mathbf{E}\left[\widetilde{K}^2_{\mathbf{n},\gamma,x,1}\right] \longrightarrow \widetilde{\sigma}^2(x)$. Now

$$\begin{aligned}
m_x^{-1}\mathbf{E}\left[\widetilde{K}^2_{\mathbf{n},\gamma,x,1}\right] &= m_x^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbf{E}\left[\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)\right]^2\\
&+\; m_x^{-1}\sum_{\substack{j_k=1\\k=1,\cdots,N}}^{r_k-1}\;\sum_{\substack{i_k=1\\k=1,\cdots,N\\ i_k\neq j_k \text{ for some } 1\leq k\leq N}}^{r_k-1}\operatorname{Cov}\left(\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right),\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{i},x\right)\right).
\end{aligned} \tag{4.58}$$

To prove (4.48), it suffices to show that the second term in the right hand side of (4.58) tends to zero as $\mathbf{n} \to \infty$. For this end, we use the same arguments used to obtain a bound for $\Gamma_{2,\gamma,x}$ defined in (4.50), the last term of (4.58) is bounded by

$$4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\sum_{\substack{i_k=1\\k=1,\cdots,N;\ \|\mathbf{i}\|>q}}^{n_k}\varphi(\|\mathbf{i}\|)^{1-\varsigma} \leq 4C\gamma^4(x)\widetilde{\varepsilon}_1^2(x)\sum_{i=q}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma}. \tag{4.59}$$

this last tends to zero as above.

To complete the proof, let us show from (4.49) that $\widetilde{Q}_{4,\mathbf{n},\gamma,x} \to 0$ as $\mathbf{n} \to \infty$ for every $\varepsilon > 0$. First, we have that : $\mathbf{E}\left[\widetilde{A}^2_{\mathbf{i},\mathbf{n},x}\right] \leq \overline{F}(v_{\mathbf{n}},x)\mathbf{E}\left[\widetilde{\Delta}^2_{\mathbf{i},\mathbf{n},x}\right] \leq 2\gamma^2(x)\widetilde{\varepsilon}_1(x)$, then

$\mathbf{E}\left[\widetilde{A}_{\mathbf{i},\mathbf{n},x}\right] \leq 2\gamma(x)\widetilde{\varepsilon}_1(x)\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}$. Note that by Markov, inequality we have

$$\mathbb{P}\left[\left|\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)\right| > p^N\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1/2}\right] \leq \frac{p^N\overline{F}(v_{\mathbf{n}},x)\gamma(x)}{p^N\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1/2}} = \gamma(x)\left[\overline{F}(v_{\mathbf{n}},x)\right]^{3/2} \to 0.$$

Therefore for $\mathbf{n}$ large enough, $\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right) < p^N\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1/2}$, and thus

$$\mathbf{E}\left[\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)^2 \mathbb{1}_{\left\{|U_{\mathbf{n}}(1,\gamma,\mathbf{j},x)|>\varepsilon\widetilde{\sigma}(x)m_x^{1/2}\right\}}\right] \leq \frac{p^{2N}}{\overline{F}(v_{\mathbf{n}},x)}\mathbb{P}\left[\left|\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)\right| > \varepsilon\widetilde{\sigma}(x)m_x^{1/2}\right].$$

By (4.47), we have $E\left[\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)^2\right] \sim \frac{\widetilde{\sigma}^2(x)}{\hat{Q}}$ for $\mathbf{n}$ large enough, then using Tchebychev inequality

$$\mathbb{P}\left[\left|\widetilde{U}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)\right| > \varepsilon\widetilde{\sigma}(x)m_x^{1/2}\right] \leq \frac{\frac{\widetilde{\sigma}^2(x)}{\hat{Q}}}{\widetilde{\sigma}^2(x)m_x\varepsilon^2} = \hat{Q}^{-1}m_x^{-1}\varepsilon^{-2}.$$

Therefore

$$\begin{aligned}\widetilde{Q}_{4,\mathbf{n},x} &\leq Cp^{2N}\left[\overline{F}\left(v_{\mathbf{n}},x\right)\right]^{-1}\hat{Q}^{-1}m_x^{-1}\varepsilon^{-2}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbb{P}\left[\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right) > \varepsilon\widetilde{\sigma}(x)m_x^{1/2}\right] && (4.60)\\ &\leq C\frac{p^{2N}}{m_x\overline{F}\left(v_{\mathbf{n}},x\right)\varepsilon^2}. && (4.61)\end{aligned}$$

This tends to zero by using the choice of q in (4.24) and $p = \left[m_x\overline{F}(v_{\mathbf{n}},x)\right]^{1/(2N)}$. This complete the proof of Lemma 4.9.7.■

► Now we are going to give the asymptotic normality of the quantity $H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}})$. The variance of $H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}})$ is obtained in the same manner as the variance of $H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}})$, it suffices to note that $H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}})$ has the same writing as $H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}})$. Recall the notations given at the beginning of this section, we have

$$H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}}) = m_x^{1/2}S_{\mathbf{n}}(x,r_{\mathbf{n}}) \text{ where } S_{\mathbf{n}}(x,r_{\mathbf{n}}) := m_x^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\Delta_{\mathbf{i},\mathbf{n},x},$$

$$\Delta_{\mathbf{i},\mathbf{n},x} = \eta_{\mathbf{i},\mathbf{n},x} - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},x}\right] \text{ and } \eta_{\mathbf{i},\mathbf{n},x} = \left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1}K\left(\frac{x-x_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}.$$

Using again (4.24), one can find a sequence of positive integers $p_{\mathbf{n}}$ such that $p_{\mathbf{n}} = p$, $qp^{-1} = o(1)$ which tends to zero as $\mathbf{n} \to \infty$. As above, let $r_1,\cdots,r_N$, we assume that $n_1 = r_1(q+p),\cdots,n_N = r_N(q+p)$.

Then, the random variables $A_{\mathbf{i},\mathbf{n},x} = \left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\Delta_{\mathbf{i},\mathbf{n},x}$'s are now set into large blocks and small blocks. From now, the asymptotic normality of $H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}})$ follows the same lines as the proof of the asymptotic normality of $H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}})$, that is

$$\begin{aligned}U_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right) &= \sum_{\mathbf{i}\in I(1,x,\mathbf{j})}A_{\mathbf{i},\mathbf{n},x},\ U_{\mathbf{n}}\left(2,\gamma,\mathbf{j},x\right) = \sum_{I(2,x,\mathbf{j})}A_{\mathbf{i},\mathbf{n},x}\\ U_{\mathbf{n}}\left(3,\gamma,\mathbf{j},x\right) &= \sum_{I(3,x,\mathbf{j})}A_{\mathbf{i},\mathbf{n},x},\ U_{\mathbf{n}}\left(4,\gamma,\mathbf{j},x\right) = \sum_{I(4,x,\mathbf{j})}A_{\mathbf{i},\mathbf{n},x}\end{aligned}$$

and so on. The last two terms are

$$U_{\mathbf{n}}\left(2^{N-1},\gamma,\mathbf{j},x\right) = \sum_{I(2^{N-1},x,\mathbf{j})} A_{\mathbf{i},\mathbf{n},x}, \quad U_{\mathbf{n}}\left(2^{N},\gamma,\mathbf{j},x\right) = \sum_{I(2^{N},x,\mathbf{j})} A_{\mathbf{i},\mathbf{n},x}$$

For each integer $1 \leq i \leq 2^N$, define

$$T_{\mathbf{n},\gamma,x,i} = \textstyle\sum_{\mathbf{j}\in D(i,x)} U_{\mathbf{n}}\left(i,\gamma,\mathbf{j},x\right). \tag{4.62}$$

Set $K_{\mathbf{n},\gamma,x} := \sum_{i=1}^{2^N} T_{\mathbf{n},\gamma,x,i}$. With this notation, $T_{\mathbf{n},\gamma,x,1}$ is the sum of the random variables $A_{\mathbf{i},\mathbf{n},x}$ in large blocks. For $2 \leq i \leq 2^N$, the $T_{\mathbf{n},\gamma,x,i}$ are sums of random variables in small blocks. If it is not the case that $n_1 = r_1(p+q), \cdots, n_N = r_N(p+q)$ for some integers $r_1, \cdots, r_N$, then a term, say, $T_{\mathbf{n},\gamma,x,2^N+1}$, containing all the $A_{\mathbf{i},\mathbf{n},x}$'s at the end not included in the big or small blocks can be added. This term will not change the proof much. Using again Lemma 4.9.4, the general approach is to show that as $\mathbf{n} \to \infty$,

$$Q_{1,\mathbf{n},\gamma,x} \equiv \left| \mathbf{E}\left[\exp\left(izT_{\mathbf{n},\gamma,x,1}\right)\right] - \prod_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[\exp\left(izU_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)\right)\right] \right| \to 0, \tag{4.63}$$

$$Q_{2,\mathbf{n},\gamma,x} \equiv m_x^{-1}\mathbf{E}\left[\left(\sum_{i=2}^{2^N} T_{\mathbf{n},\gamma,x,i}\right)^2\right] \to 0, \tag{4.64}$$

$$Q_{3,\mathbf{n},\gamma,x} \equiv m_x^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[U_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)^2\right] \to \overline{\sigma}^2(x), \quad \text{where} \quad \overline{\sigma}^2(x) = 1 \tag{4.65}$$

$$Q_{4,\mathbf{n},\gamma,x} \equiv m_x^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[U_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)^2 \mathbb{1}_{\left\{|U_{\mathbf{n}}(1,\gamma,\mathbf{j},x)| \geq \varepsilon\overline{\sigma}(x)m_x^{1/2}\right\}}\right] \to 0 \quad \text{for every } \varepsilon > 0. \tag{4.66}$$

Note that

$$\frac{S_{\mathbf{n}}(x,r_{\mathbf{n}})}{\overline{\sigma}(x)\sqrt{m_x}} = \frac{T_{\mathbf{n},\gamma,x,1}}{\overline{\sigma}(x)\sqrt{m_x}} + \sum_{i=2}^{2^N} \frac{T_{\mathbf{n},\gamma,x,i}}{\overline{\sigma}(x)\sqrt{m_x}}.$$

In summary, equation (4.63) ensures that the random variables $U_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right)$ are asymptotically independent. The asymptotic normality of $T_{\mathbf{n},\gamma,x,1}/\left(\overline{\sigma}(x)m_x^{1/2}\right)$ follows from (4.65) and the Lindeberg-Feller condition (4.66). Equation (4.64) shows that the term $\sum_{i=2}^{2^N} T_{\mathbf{n},\gamma,x,i}/\left(\overline{\sigma}(x)m_x^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible. The proofs of (4.63)-(4.66) follows the same lines of proofs as in Tran (1990) and are therefore omitted in this chapter.

► Let us prove that $H_{v_{\mathbf{n}}}(\widehat{\gamma}_{v_{\mathbf{n}}}, x, r_{\mathbf{n}}) := (H_{v_{\mathbf{n}},\pi}(x,r_{\mathbf{n}}), H_{v_{\mathbf{n}},\phi}(x,r_{\mathbf{n}}))^\top$ converges in distribution to $\mathcal{N}\left(0,\Sigma\right)$. According to Cramèr-Wold's device (see Van der Vaart (1998)), it is sufficient to prove that $l^\top H_{\mathbf{n}}(\widehat{\gamma}_{v_{\mathbf{n}}}, x, r_{\mathbf{n}}) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, l^\top \Sigma l\right)$ for all $l = (l_1, l_2)^\top \in$

$\mathbb{R}^2$, $l \neq 0$. Notice that, in the proof we consider all the covariates recorded in the ball $B(x, r_{\mathbf{n}})$ as previously. Some simple calculation yields :

$$
\begin{aligned}
l^{\top} H_{\mathbf{n}}(\widehat{\gamma}_{v_{\mathbf{n}}}, x, r_{\mathbf{n}}) &:= m_x^{1/2} \frac{\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{1/2}}{m_x} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta^{\dagger}_{\mathbf{i},\mathbf{n},x}, \\
\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x} &:= \left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1} \left(l_1 \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} + l_2 \log\left(Y_{\mathbf{i}}/v_{\mathbf{n}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} \right) \\
&\quad - \left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1} \mathbf{E}\left[l_1 \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} + l_2 \log\left(Y_{\mathbf{i}}/v_{\mathbf{n}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} \right] \\
&:= A^{\dagger}_{\mathbf{i},\mathbf{n},x} - \mathbf{E}\left[A^{\dagger}_{\mathbf{i},\mathbf{n},x}\right].
\end{aligned}
$$

That is a row sum of the triangular array of random variables $\left\{\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x}\right\}_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}}$ where for fixed $\hat{\mathbf{n}}$, $\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x}$, $\mathbf{i} = \mathbf{1}, \cdots, m_x$ are zero mean random variables. Thus, to establish the asymptotic normality of $H_{\mathbf{n}}(\widehat{\gamma}_{v_{\mathbf{n}}}, x, r_{\mathbf{n}})$, we verify the Lindeberg-Feller condition for triangular arrays. We first calculate the variance of $H_{\mathbf{n}}(\widehat{\gamma}_{v_{\mathbf{n}}}, x, r_{\mathbf{n}})$. We have

$$
\operatorname{Var}\left(m_x^{-1} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta^{\dagger}_{\mathbf{i},\mathbf{n},x} \right) = m_x^{-2} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \operatorname{Var}\left(\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x}\right) + 2 m_x^{-2} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{I}_{\mathbf{n}}} \sum_{\mathbf{i} \neq \mathbf{j}} \operatorname{Cov}\left(\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x}, \Delta^{\dagger}_{\mathbf{j},\mathbf{n},x}\right).
$$

Denote by $I^{\dagger}_{\mathbf{n},x} := m_x^{-2} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \operatorname{Var}\left(A^{\dagger}_{\mathbf{i},\mathbf{n},x}\right)$ and by $R^{\dagger}_{\mathbf{n},x} := m_x^{-2} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{I}_{\mathbf{n}}} \sum_{\mathbf{i} \neq \mathbf{j}} \left| \operatorname{Cov}\left(A^{\dagger}_{\mathbf{i},\mathbf{n},x}, A^{\dagger}_{\mathbf{j},\mathbf{n},x}\right) \right|$. Clearly, $\operatorname{Var}\left(m_x^{-1} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta^{\dagger}_{\mathbf{i},\mathbf{n},x} \right) \leq I^{\dagger}_{\mathbf{n},x} + R^{\dagger}_{\mathbf{n},x}$, with

$$
\begin{aligned}
I^{\dagger}_{\mathbf{n},x} &= \frac{m_x}{m_x^2} \left(\mathbf{E}\left[A^{\dagger 2}_{\mathbf{i},\mathbf{n},x}\right] - \mathbf{E}\left[A^{\dagger}_{\mathbf{i},\mathbf{n},x}\right]^2 \right) \\
&= \frac{\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-2}}{m_x} \mathbf{E}\left[\left(l_1 \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} + l_2 \log\left(Y_{\mathbf{i}}/v_{\mathbf{n}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} \right)^2 \right] \\
&\quad - \frac{\left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-2}}{m_x} \mathbf{E}\left[l_1 \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} + l_2 \log\left(Y_{\mathbf{i}}/v_{\mathbf{n}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}} \right]^2. \\
&= \frac{l_1^2}{m_x} \left(\mathbf{E}\left[\eta^2_{\mathbf{i},\mathbf{n},x}\right] - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},x}\right]^2 \right) + \frac{l_2^2}{m_x} \left(\mathbf{E}\left[\widetilde{\eta}^2_{\mathbf{i},\mathbf{n},x}\right] - \mathbf{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right]^2 \right) \\
&\quad + \frac{2 l_1 l_2}{m_x} \left(\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},x} \widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right] - \mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n},x}\right] \mathbf{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n},x}\right] \right) \\
&:= I^{\dagger}_{\mathbf{n},x,1} \quad + \quad I^{\dagger}_{\mathbf{n},x,2}.
\end{aligned}
$$

From (4.35) and (4.41), it follows

$$
\begin{aligned}
I^{\dagger}_{\mathbf{n},x,1} &= l_1^2 m_x^{-1} \left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1} \left(1 - \overline{F}(v_{\mathbf{n}}, x)\right) \\
&\quad + 2 l_2^2 m_x^{-1} \gamma^2(x) \left[\overline{F}(v_{\mathbf{n}}, x)\right]^{-1} \left\{ 1 + \frac{\mathcal{A}(v_{\mathbf{n}}, x)}{\gamma(x)\rho(x)} \left[\frac{1}{(1-\rho(x))^2} - 1 \right] (1 + o(1)) \right\} \\
&\quad - l_2^2 m_x^{-1} \gamma^2(x) \left\{ 1 + \frac{\mathcal{A}(v_{\mathbf{n}}, x)}{\gamma(x)\rho(x)} \left[\frac{1}{(1-\rho(x))} - 1 \right] (1 + o(1)) \right\}^2. \qquad (4.67)
\end{aligned}
$$

From (4.40) and Lemma 1 of Goegebeur et al. (2014b) together with the second order condition $(\mathcal{R}_2)$, we get

$$\begin{aligned} I^{\dagger}_{\mathbf{n},x,2} &= 2l_1 l_2 \gamma(x) m_x^{-1}\left(\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1}-1\right) \\ &\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}. \end{aligned} \tag{4.68}$$

Combining (4.67) and (4.68), we have

$$\begin{aligned} I^{\dagger}_{\mathbf{n},x} &= l_1^2 m_x^{-1}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1}\left(1-\overline{F}\left(v_{\mathbf{n}},x\right)\right) \\ &+ 2l_2^2 m_x^{-1}\gamma^2(x)\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1}\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\} \\ &- l_2^2 m_x^{-1}\gamma^2(x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\}^2 \\ &+ 2l_1 l_2\gamma(x) m_x^{-1}\left(\left[\overline{F}(v_{\mathbf{n}},x)\right]^{-1}-1\right) \\ &\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}. \end{aligned} \tag{4.69}$$

Now let us examine $R^{\dagger}_{\mathbf{n},x}$, simple calculations yields

$$\begin{aligned} \operatorname{Cov}\left(A^{\dagger}_{\mathbf{i},\mathbf{n},x}, A^{\dagger}_{\mathbf{j},\mathbf{n},x}\right) &= l_1^2 \operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},x},\eta_{\mathbf{j},\mathbf{n},x}\right) + l_2^2 \operatorname{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right) \\ &+ l_1 l_2\left\{\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n},x},\widetilde{\eta}_{\mathbf{j},\mathbf{n},x}\right)+\operatorname{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n},x},\eta_{\mathbf{j},\mathbf{n},x}\right)\right\}. \end{aligned}$$

Using the same method as in Lemma 4.9.4, it follows

$$R^{\dagger}_{\mathbf{n},x} = O\left(\frac{1}{m_x\overline{F}(v_{\mathbf{n}},x)}\right). \tag{4.70}$$

Then combining (4.69) and (4.70), we get

$$\begin{aligned} \lim_{\mathbf{n}\to+\infty}\operatorname{Var}\left(l^{\top}H_{\mathbf{n}}(\widehat{\gamma}_{v_{\mathbf{n}}},x,r_{\mathbf{n}})\right) &= \lim_{\mathbf{n}\to+\infty} m_x\overline{F}(v_{\mathbf{n}},x)\operatorname{Var}\left(m_x^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x}\right) \\ &= \lim_{\mathbf{n}\to+\infty} m_x\overline{F}(v_{\mathbf{n}},x)\ I^{\dagger}_{\mathbf{n},x} = l_1^2+2l_1l_2\gamma(x)+2l_2^2\gamma^2(x) \\ &= l^{\top}\Sigma l. \end{aligned}$$

To complete the proof of the asymptotic normality of $l^{\top}H_{\mathbf{n}}(\widehat{\gamma}_{v_{\mathbf{n}}},x,r_{\mathbf{n}})$ let us use same notations as above. Denote

$$S^{\dagger}_{\mathbf{n}}(x,r_{\mathbf{n}}) := m_x^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x} \quad \text{then} \quad l^{\top}\mathcal{H}_{v_{\mathbf{n}}}(x,r_{\mathbf{n}}) = m_x^{1/2}S^{\dagger}_{\mathbf{n}}(x,r_{\mathbf{n}}).$$

Using the same notations as in Lemma 4.9.5. Then the random variables $\Psi^{\dagger}_{\mathbf{i},\mathbf{n},x} := \left[\overline{F}(v_{\mathbf{n}},x)\right]^{1/2}\Delta^{\dagger}_{\mathbf{i},\mathbf{n},x}$'s are set into large blocks and small blocks, that is

$$\begin{aligned} U^{\dagger}_{\mathbf{n}}\left(1,\gamma,\mathbf{j},x\right) &= \sum_{\mathbf{i}\in I(1,x,\mathbf{j})}\Psi^{\dagger}_{\mathbf{i},\mathbf{n},x}, \quad U^{\dagger}_{\mathbf{n}}\left(2,\gamma,\mathbf{j},x\right) = \sum_{I(2,x,\mathbf{j})}\Psi^{\dagger}_{\mathbf{i},\mathbf{n},x} \\ U^{\dagger}_{\mathbf{n}}\left(3,\gamma,\mathbf{j},x\right) &= \sum_{I(3,x,\mathbf{j})}\Psi^{\dagger}_{\mathbf{i},\mathbf{n},x}, \quad U^{\dagger}_{\mathbf{n}}\left(4,\gamma,\mathbf{j},x\right) = \sum_{I(4,x,\mathbf{j})}\Psi^{\dagger}_{\mathbf{i},\mathbf{n},x} \end{aligned}$$

and so on. The last two terms are

$$U_{\mathbf{n}}^{\dagger}\left(2^{N-1},\gamma,\mathbf{j},x\right) = \sum_{I(2^{N-1},x,\mathbf{j})} \Psi_{\mathbf{i},\mathbf{n},x}^{\dagger}, \quad U_{\mathbf{n}}^{\dagger}\left(2^{N},\gamma,\mathbf{j},x\right) = \sum_{I(2^{N},x,\mathbf{j})} \Psi_{\mathbf{i},\mathbf{n},x}^{\dagger}$$

For each integer $1 \leq i \leq 2^N$, define

$$T_{\mathbf{n},\gamma,x,i}^{\dagger} = \textstyle\sum_{\mathbf{j}\in D(i,x)} U_{\mathbf{n}}^{\dagger}\left(i,\gamma,\mathbf{j},x\right).$$

Clearly $S_{\mathbf{n}}^{\dagger}(x,r_{\mathbf{n}}) = \sum\limits_{i=1}^{2^N} T_{\mathbf{n},\gamma,x,i}^{\dagger}$. With this notation, $T_{\mathbf{n},\gamma,x,1}^{\dagger}$ is the sum of the random variables $\Psi_{\mathbf{i},\mathbf{n},x}^{\dagger}$ in large blocks. For $2 \leq i \leq 2^N$, the $T_{\mathbf{n},\gamma,x,i}^{\dagger}$ are sums of random variables in small blocks. Using Lemma 4.9.4, the general approach is to show that as $\mathbf{n} \to \infty$,

$$Q_{1,\mathbf{n},\gamma,x}^{\dagger} \equiv \left|\mathbb{E}\left[\exp\left(izT_{\mathbf{n},\gamma,x,1}^{\dagger}\right)\right] - \prod_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[\exp\left(izU_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j},x\right)\right)\right]\right| \to 0, \tag{4.71}$$

$$Q_{2,\mathbf{n},\gamma,x}^{\dagger} \equiv m_x^{-1}\mathbb{E}\left[\left(\sum_{i=2}^{2^N} T_{\mathbf{n},\gamma,x,i}^{\dagger}\right)^2\right] \to 0, \tag{4.72}$$

$$Q_{3,\mathbf{n},\gamma,x}^{\dagger} \equiv m_x^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[U_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j},x\right)^2\right] \to \widehat{\sigma}^2(x), \tag{4.73}$$

$$Q_{4,\mathbf{n},\gamma,x}^{\dagger} \equiv m_x^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[U_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j},x\right)^2 \mathbb{1}_{\left\{\left|U_{\mathbf{n}}^{\dagger}(1,\gamma,\mathbf{j},x)\right| \geq \varepsilon\widehat{\sigma}(x)m_x^{1/2}\right\}}\right] \to 0, \tag{4.74}$$

for every $\varepsilon > 0$. Note that

$$\frac{S_{\mathbf{n}}^{\dagger}(x)}{\widehat{\sigma}(x)\sqrt{m_x}} = \frac{T_{\mathbf{n},\gamma,x,1}^{\dagger}}{\widehat{\sigma}(x)\sqrt{m_x}} + \sum_{i=2}^{2^N} \frac{T_{\mathbf{n},\gamma,x,i}^{\dagger}}{\widehat{\sigma}(x)\sqrt{m_x}},$$

where $\sigma^2(x) = l_1^2 + 2l_1l_2\gamma(x) + 2l_2^2\gamma^2(x)$.
In summary, equation (4.71) ensures that the random variables $U_{\mathbf{n}}^{\dagger}\left(1,\gamma,\mathbf{j},x\right)$ are asymptotically independent. The asymptotic normality of $T_{\mathbf{n},\gamma,x,1}^{\dagger}/\left(\widehat{\sigma}(x)m_x^{1/2}\right)$ follows from (4.73) and the Lindeberg-Feller condition (4.74). Equation (4.72) shows that the term $\sum\limits_{i=2}^{2^N} T_{\mathbf{n},\gamma,x,i}^{\dagger}/\left(\widehat{\sigma}(x)m_x^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible. The proofs of (4.71)-(4.74) follows the same lines of proofs as in Tran (1990) and are therefore omitted in this chapter. ■

4.9.4 Appendix A4 : Proofs of main results

Proof of Theorem 4.6.1. First, let us introduce the following quantity

$$\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\left(\widehat{\gamma}_{v_{\mathbf{n}}}(x)-\gamma(x)\right)=:\frac{\pi(v_{\mathbf{n}},x)}{\pi_{\mathbf{n}}(v_{\mathbf{n}},x)}\left(\Gamma_{\mathbf{n},x,1}\quad+\quad\Gamma_{\mathbf{n},x,2}\quad+\quad\Gamma_{\mathbf{n},x,3}\right)$$

$$\Gamma_{\mathbf{n},x,1}:=\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\left(\frac{\widehat{\phi}_{v_{\mathbf{n}}}(y,x)}{\pi(v_{\mathbf{n}},x)}-\mathbf{E}\left[\frac{\widehat{\phi}_{v_{\mathbf{n}}}(y,x)}{\pi(v_{\mathbf{n}},x)}\right]\right)$$

$$\Gamma_{\mathbf{n},x,2}:=-\gamma(x)\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\left(\frac{\pi_{\mathbf{n}}(v_{\mathbf{n}},x)}{\pi(v_{\mathbf{n}},x)}-\mathbf{E}\left[\frac{\pi_{\mathbf{n}}(v_{\mathbf{n}},x)}{\pi(v_{\mathbf{n}},x)}\right]\right)$$

$$\Gamma_{\mathbf{n},x,3}:=\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\left(\frac{\mathbf{E}\left[\widehat{\phi}_{v_{\mathbf{n}}}(y,x)\right]-\gamma(x)\mathbf{E}\left[\pi_{\mathbf{n}}(v_{\mathbf{n}},x)\right]}{\pi(v_{\mathbf{n}},x)}\right).$$

By Lemma 4.9.5, it follows that

$$\Gamma_{\mathbf{n},x,1}+\Gamma_{\mathbf{n},x,2}\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\gamma^2(x)\right). \tag{4.75}$$

For the term $\Gamma_{\mathbf{n},x,3}$, from Lemma 4.9.5, Lemma 1 of Goegebeur et al. (2014b) and the second order condition $(\mathcal{R}_2)$ together with all the covariates in the ball, it follows easily that

$$\frac{\mathbf{E}\left[\widehat{\phi}_{v_{\mathbf{n}}}(y,x)\right]}{\pi(v_{\mathbf{n}},x)}=\gamma(x)\overline{F}(v_{\mathbf{n}},x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\},$$

and

$$\frac{\mathbf{E}\left[\pi_{\mathbf{n}}(v_{\mathbf{n}},x)\right]}{\pi(v_{\mathbf{n}},x)}=\overline{F}(v_{\mathbf{n}},x).$$

Then we have

$$\Gamma_{\mathbf{n},x,3}=\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\frac{\mathcal{A}(v_{\mathbf{n}},x)}{1-\rho(x)}+o(1)\xrightarrow{\mathbb{P}}\frac{\lambda(x)}{1-\rho(x)}. \tag{4.76}$$

In other hand, from Lemma 4.9.5

$$\frac{\sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}K\left(\frac{x-X_{\mathbf{i}}}{r_{\mathbf{n}}}\right)\mathbb{1}_{Y_{\mathbf{i}}>v_{\mathbf{n}}}}{\pi(v_{\mathbf{n}},x)}=1+O_{\mathbb{P}}(1).$$

Combining (4.75) and (4.76) with (4.69), it follows

$$\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\left(\widehat{\gamma}_{v_{\mathbf{n}}}(x)\ -\ \gamma(x)\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)},\gamma^2(x)\right).$$

Theorem 4.6.1 follows by a straightforward application of the Delta-method. ■

Proof of Theorem 4.6.2 The proof is based on the following transformation provided from (4.21), that is :

$$\frac{\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}\log\frac{\hat{q}^W_{\mathbf{n}}(\beta_{\mathbf{n}},x)}{q(\beta_{\mathbf{n}},x)}=:A_{1,v_{\mathbf{n}}}(x,r_{\mathbf{n}})+A_{2,v_{\mathbf{n}}}(x,r_{\mathbf{n}})+A_{3,v_{\mathbf{n}}}(x,r_{\mathbf{n}}),$$

with

$$\begin{aligned}
A_{1,v_{\mathbf{n}}}(x,r_{\mathbf{n}}) &:= \sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\,(\widehat{\gamma}_{v_{\mathbf{n}}}(x)-\gamma(x))\\
A_{2,v_{\mathbf{n}}}(x,r_{\mathbf{n}}) &:= \frac{\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}\,(\log\hat{q}(\alpha_{\mathbf{n}},x)-\log q(\alpha_{\mathbf{n}},x))\\
A_{3,v_{\mathbf{n}}}(x,r_{\mathbf{n}}) &:= \frac{\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}\left(\log\frac{q(\alpha_{\mathbf{n}},x)}{q(\beta_{\mathbf{n}},x)}+\gamma(x)\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)\right).
\end{aligned}$$

In one hand, from Theorem 4.6.1,

$$A_{1,v_{\mathbf{n}}}(x,r_{\mathbf{n}}) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)},\gamma^2(x)\right), \tag{4.77}$$

as a straightforward consequence of Theorem 4.6.1.

Concerning the term $A_{2,v_{\mathbf{n}}}(x,r_{\mathbf{n}})$, since $\widehat{q}(\alpha_{\mathbf{n}},x)/q(\alpha_{\mathbf{n}},x) \xrightarrow{\mathbb{P}} 1$ by assumption, one can obtain easily the asymptotic normality of $\widehat{q}(\alpha_{\mathbf{n}},x)$ with rate $\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}$, and we get

$$A_{2,v_{\mathbf{n}}}(x,r_{\mathbf{n}}) = \frac{\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}\left(\widehat{q}(\alpha_{\mathbf{n}},x)/q(\alpha_{\mathbf{n}},x)-1\right)(1+o(1)) = \frac{O_{\mathbb{P}}(1)}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}. \tag{4.78}$$

For the term saying $A_{3,v_{\mathbf{n}}}(x,r_{\mathbf{n}})$, from Lemma 2 of Daouia et al. (2011), it follows immediately that

$$A_{3,v_{\mathbf{n}}}(x,r_{\mathbf{n}}) = O\left(\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}\Lambda\left(q(\alpha_{\mathbf{n}},x),x\right)\right). \tag{4.79}$$

which goes to 0 by assumption. Finally, combining (4.77)-(4.79), it follows that

$$\frac{\sqrt{m_x\overline{F}(v_{\mathbf{n}},x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}\log\frac{\widehat{q}^W_{\mathbf{n}}(\beta_{\mathbf{n}},x)}{q(\beta_{\mathbf{n}},x)} =: A_{1,v_{\mathbf{n}}}(x,r_{\mathbf{n}})+A_{2,v_{\mathbf{n}}}(x,r_{\mathbf{n}})+A_{3,v_{\mathbf{n}}}(x,r_{\mathbf{n}})$$

converges in distribution to $\mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)},\gamma^2(x)\right)$.$\blacksquare$

Chapitre 5

On nonparametric conditional tail and extreme quantile estimation with random covariate in a spatial context

Sommaire

5.1 Résumé en français

Dans le chapitre 4 qui précède, nous nous sommes penchés sur l'estimation non paramétrique de l'indice de queue conditionnel et des quantiles extrêmes associés d'un processus spatial en présence d'une co-variable déterministe. Dans ce chapitre, nous considérons la situation où la variable explicative X est aléatoire, c'est le modéle dit "design aléatoire". On se propose d'étudier un champ aléatoire $(Z_{\mathbf{i}} = (Y_{\mathbf{i}}, X_{\mathbf{i}}))_{\mathbf{i}\in\mathbb{Z}^N}$ où $N \geq 2$ à valeurs dans $\mathbb{R} \times \mathbb{R}^d$ dont les éléments $Z_{\mathbf{i}}$ ont la même distribution que le couple (Y, X), défini sur un espace de probabilité (Ω, A, P). Nous supposons également que le processus étudié satisfait la condition de mélange fort ($\alpha-$mélange). Nous étudions une estimation de l'indice de queues lourdes et les quantiles extrêmes de la fonction de distribution conditionnelle de la variable réponse spatiale $Y_{\mathbf{i}}$ étant donnée la variable explicative $X_{\mathbf{i}}$. Soit $\left(\mathbb{R}^d, d_1\right)$ muni de la métrique $d_1(.,.)$. Nous supposons que la condition de variation régulière à l'infini de la probabilité conditionnelle de Y sachant $X = x$ est donnée par : $\forall x \in \mathbb{R}^d$ et pour tout y,

$$\mathbb{P}\left(Y > \ y \mid X = x\right) = y^{-\frac{1}{\gamma(x)}} L\left(y, x\right), \tag{5.1}$$

caractérisée par $\gamma(x) > 0$, l'unique fonction positive inconnue de la co-variable x, appelée index des valeurs extrêmes et pour un x fixé, $L(., x)$ est une fonction à variations lentes à l'infini. Un point $\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N$ sera désigné comme un site. Nous considérons que les données sont observées dans un domaine rectangulaire

$$\mathcal{I}_{\mathbf{n}} = \{\mathbf{i} = (i_1, ..., i_N) \in \mathbb{Z}^N, 1 \leq i_k \leq n_k, k = 1, ..., N\}, \mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N.$$

Ainsi, nous abordons l'estimation de la fonction inconnue de l'indice de queue conditionnel des valeurs extrêmes $\gamma(\cdot)$ qui sera ensuite utilisé pour estimer les quantiles extrêmes conditionnels. Pour cela, nous considérons une version de l'estimateur de Hill (1975) appartenant à la famille d'estimateurs à noyau proposée dans la littérature (voir Goegebeur et al. (2014b)), que nous étendons au cas spatial. Cet estimateur est défini par

$$\widehat{\gamma}_{\mathbf{n}}(x) = \frac{\sum\limits_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right) \left(\log \frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}}}{\sum\limits_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}}} \tag{5.2}$$

où $K_b(u) := b^{-d} K\left(u/b\right)$ est un noyau dans $\mathbb{R}^d$, $(b := b_{\mathbf{n}})$ est une suite déterministe tendant vers zéro quand $\mathbf{n} \to \infty$, $\mathbb{1}_{\{\cdot\}}$ est la fonction indicatrice et $(v_{\mathbf{n}})$ est une séquence non aléatoire qui tend vers ∞. La suite $(v_{\mathbf{n}})$ est introduite pour sélectionner les grandes observations ; son choix est très important en pratique. Sous certaines conditions liées à la dépendance spatiale des observations, à la fonction de survie, à $b_{\mathbf{n}}$, $v_{\mathbf{n}}$, en plus des conditions usuelles en théorie des valeurs extrêmes, on montre que l'estimateur est asymptotiquement normal. Cet estimateur est ensuite utilisé pour donner une estimation de quantiles extrêmes conditionnels dont la normalité asymptotique est aussi obtenue.

5.2 Introduction

The problem of studying extreme events arises in many fields of statistical applications. For instance, in actuarial science, it is of primary interest for a company to estimate the small probabilities that a claim that represents a threat to its solvency is filled ; in hydrology, one could be interested in forecasting the maximum level reached by the seawater along a coast over a given period, or studying extreme rainfall at a given location. The aim of the extreme values theory is to estimate small probabilities. For example, when interest is in the estimation of extreme quantiles, *i.e.* quantiles of order α_n with $\alpha_n > 1 - 1/n$, n being the sample size, one has to extrapolate beyond the available data by means of an extreme values model depending on the tail index. Let Y be a random variable with cumulative distribution function F and let $Y_1, \ldots, Y_n$, be copies of Y. Extreme quantiles of F are defined as quantities of the form

$$\overline{F}^{-1}(\alpha) = \inf\left\{x : \overline{F}(x) \leq \alpha\right\} \quad \text{for} \quad 0 < \alpha < 1,$$

where α is so small that this quantile falls beyond the range of the observed data $Y_1, \cdots, Y_n$. This problem is closely related to the estimation of the extreme values or tail index of the distribution function of Y. The extreme values index drives the behavior of F in its right tail and thus plays a central role in the analysis of extremes. Many authors pay attention on the study of extreme values theory and in particular on estimation of the extreme values index and extreme quantiles see Embrechts et al. (1997), Gomes et al. (2005), Beirlant et al. (2006), Silverberg & Verspagen (2007), Beirlant et al. (2008), Müller & Rufibach (2009), Novak (2011), Davison et al. (2012), Fawcett & Walshaw (2014) and the references therein. In a lot of applications, some covariate information is recorded simultaneously with the quantity of interest. But all this theory does not usually take into account the spatial dependence, despite its practical interest. For instance, in climatology one may be interested in the estimation of return periods associated to extreme rainfall as a function of the geographical location. As so far, max-stable processes have mostly been used for the statistical modeling of spatial data see for example, Coles (2001) and Coles & Tawn (1996) who modeled extremal rainfall fields. Padoan et al. (2010) described a practicable pairwise likelihood estimation procedure applied to the rainfall data using max-stable processes. An interesting application to wind gusts is shown in Coles & Walshaw (1994), who used max-stable processes to model the angular dependence for wind speed directions.

In this chapter, we will consider tail index estimation for heavy tailed models when a random covariate X is recorded simultaneously with the variable of interest Y and observations are available in space. When some random covariate information X is available and the distribution of Y depends on X, the problem is to estimate the conditional extreme values index and conditional extreme quantiles of the distribution $F(., x)$ of Y given $X = x$. Several works has been done recently in the case of fixed covariates, include Davison & Smith (1990), Smith (2001), Hall & Tajvidi (2000), Beirlant et al. (2006), Gardes & Girard (2008), Gardes & Girard (2010), Ferrez et al. (2011), Gardes & Girard (2012), Ndao et al. (2014) among others. More recently, some papers already address the estimation of the conditional extreme values index and the conditional extreme quantiles in the random covariates case. Daouia et al. (2011) proposed a kernel-based estimator of conditional extreme quantiles with random covariates. Gardes & Stupfler (2014) and Goegebeur et al. (2014b) adapt Hill's estimator of the extreme values index of a heavy-tailed distribution to the presence of a random covariate. Goegebeur et al. (2014a) adapt the moment estimator introduced by Dekkers et al. (1989) to the presence of random covariates.

We propose a conditional tail index and conditional extreme quantiles estimators for spatial data based on a class of functions satisfying some mild conditions. To the best of our knowledge, such estimator has not been investigated before. The only work on tail index estimation for spatial process is that of Basrak & Tafro (2014). This last consider the tail estimation (with no covariate) for moving averages and moving maxima of $(X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^2)$ observed on bivariate regular lattice, and shows the consistency (convergence in probability) of the tail estimate.

However, there exist few results on nonparametric estimation of conditional quantiles in a spatial non-extreme case. Cui et al. (2004) studied the asymptotic distribution for a class of M-estimators including the median estimator in the framework of a linear model with spatially correlated data. Hallin et al. (2009) give a Bahadur representation and asymptotic normality results of the local linear quantile estimator.

Laksaci & Maref (2009) consider the case where the regressors take their values in a semi-metric space and show the strong and weak consistency of the conditional quantile. Ould-Abdi et al. (2010) considered the pointwise p-mean and almost-complete consistencies of a double kernel quantile estimator for real-valued random fields. Dabo-Niang & Thiam (2010) established the L_1 consistency and the asymptotic normality of the kernel conditional quantile estimator in the case of random fields. In the case of spatial extreme quantiles, Carreau & Girard (2011) propose to estimate spatial extreme quantiles by a weighted log-likelihood approach.

Estimating an extreme spatial event, particularly, the extreme index of a conditional distribution function F heavy-tailed is far from being trivial. This tail parameter controls the behavior of F at infinity, which implies that its estimate is necessary, especially when one wants to estimate extreme quantiles. There are several methods for estimating tail index for different types of processes. Instead of parametric estimation, we consider in the following, that the conditional distribution of the random variable of interest given a covariate is heavy-tailed and depends on a nonlinear function of the covariate. To estimate the unknown conditional tail index and conditional extreme quantiles, we adapt nonparametric time-dependent extreme values smoothing methods and the kernel estimator of extreme values index of a heavy-tailed distribution proposed by Goegebeur et al. (2014b) to spatial data. The major difference between our work and the latter is not at the writing of the estimator of γ. The difference is essentially technical, it concerns the extension of the assumptions used in the indexed process under unidimensional case to multidimensional framework in the management of spatial dependence observations in the region $\mathcal{I}_{\mathbf{n}}$. This is particularly seen in the proofs of the asymptotic results.

The following section describes our basic model and settings. Detailed regularity conditions and asymptotic results are presented in Section 5.4. Conclusion and some discussions are given in Section 5.5, while Section 5.6 is devoted to the proofs.

5.3 Model and definitions

Let $\left(Z_{\mathbf{i}} = (Y_{\mathbf{i}}, X_{\mathbf{i}}), \mathbf{i} \in \mathbb{Z}^N\right)$ be a $\mathbb{R} \times \mathbb{R}^d-$valued measurable spatial process, where $Z_{\mathbf{i}}$ has the same distribution as $(Y, X) \in \mathbb{R} \times \mathbb{R}^d$. Here, Y is the scalar response variable and X is some random covariate valued in $\mathbb{R}^d$ with density g and $\mathbb{R}^d$ is a metric space associated to a metric $d_1(.,.)$. We assume also that (Y, X) is of density $f(\cdot,\cdot)$.

For $y > 0$ and $x \in \mathbb{R}^d$, denote by $F(y, x)$ the conditional distribution function of Y given x. We assume that for all $x \in \mathbb{R}^d$, $F(y, x)$ is heavy-tail. More specifically, we have

$$\overline{F}(y, x) := 1 - F(y, x) = y^{-\frac{1}{\gamma(x)}} L(y, x), \tag{5.3}$$

where $\gamma(\cdot)$ is unknown positive continuous function of the covariate x.

The positive function $\gamma(x)$ describes the tail heaviness of the conditional response distribution, and is a function that needs to be adequately estimated from the available data. It is clear that for every $x \in \mathbb{R}^d$, $F(\cdot, x)$ belongs to the domain of attraction of Frèchet distribution with shape $\gamma(x)$.

The conditional Pareto-type model can also be stated in an equivalent way in terms of the conditional tail quantile function $q(\alpha, x) := F^{-1}(1-\alpha, x)$ for $\alpha \in]0,1[$ defined as follows

$$q(\alpha, x) = \alpha^{-\gamma(x)} L_q(\alpha^{-1}, x). \tag{5.4}$$

Where $F^{-1}(\cdot, x)$ is the inverse function of the conditional distribution function. For x fixed, the unknown functions $L_F(., x)$ and $L_q(., x)$ are slowly varying at infinity, that is for all $t > 0$:

$$\lim_{y\to\infty} \frac{L_q(ty, x)}{L_q(y, x)} = \lim_{y\to\infty} \frac{L_F(ty, x)}{L_F(y, x)} = 1. \tag{5.5}$$

For $\alpha \in]0,1[$, the α-th conditional quantile of Y given x denoted by $q(\alpha, x)$ is a solution of the equation :

$$F(q(\alpha, x), x) = \alpha.$$

Let us now give some notations which will be used throughout this chapter. For $N \geq 1$, $\mathbb{Z}^N$ denotes the integer lattice points in the $N-$dimensional Euclidean space. A bold point $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{Z}^N$ will be refered to a site and $\mathbf{n} = (n_1, \ldots, n_N) \in \mathbb{N}^N$. We will write $\mathbf{n} \to \infty$ if $\min\{n_k\} \to \infty$. All the limits are considered when $\mathbf{n} \to \infty$. For $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N$, we set $\widehat{\mathbf{n}} = n_1 \cdots n_N$. Finally, we will denote by $\mathcal{I}_\mathbf{n}$ a rectangular region defined in the following way :

$$\mathcal{I}_\mathbf{n} = \{\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{Z}^N, 1 \leq i_k \leq n_k, k = 1, \ldots, N\}.$$

Now, assume that the sample $Z_\mathbf{i} = (Y_\mathbf{i}, X_\mathbf{i})$ is observed over the rectangular domain $\mathcal{I}_n$ with conditional distribution function of $Y_\mathbf{i}$ given $X_\mathbf{i}$ defined in (5.3). The aim of this chapter is to provide and study a pointwise estimator of the conditional function $\gamma(x)$ and the conditional high quantiles $q(\alpha_\mathbf{n}, x)$, where $\alpha_\mathbf{n} \to 0$ as $\mathbf{n} \to \infty$. More precisely, for a given $x \in \mathbb{R}^d$, we give asymptotic normality for the estimator of $\gamma(x)$ and the conditional high quantiles $q(\alpha_n, x)$, where $\alpha_\mathbf{n} \to 0$ as $\mathbf{n} \to \infty$.

We will assume that the spatial dependence of the process is measured by means of $\alpha-$*mixing*. Then, we consider the $\alpha-$*mixing* coefficients of the field $(Z_\mathbf{i}, \mathbf{i} \in \mathbb{Z}^N)$, defined by :

***Mixing* condition :** Let E and E' be two sets of sites. Let $\mathcal{B}(E) = \mathcal{B}(Z_\mathbf{i}, \mathbf{i} \in E)$ and $\mathcal{B}(E') = \mathcal{B}(Z_\mathbf{i}, \mathbf{i} \in E')$ be $\sigma-$fields generated by the random variables $(Z_\mathbf{i})_\mathbf{i}$ with $\mathbf{i}$ being elements of E and E', respectively. There exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that whenever E, E' subsets of $\mathbb{Z}^N$ with finite cardinals,

$$\begin{aligned} \alpha(\mathcal{B}(E), \mathcal{B}(E')) &= \sup_{B\in\mathcal{B}(E),\, C\in\mathcal{B}(E')} |P(B\cap C) - P(B)P(C)| \\ &\leq \psi(\mathrm{Card}(E), \mathrm{Card}(E'))\, \varphi(\mathrm{d}(E, E')) \end{aligned} \tag{5.6}$$

where $\mathrm{Card}(E)$ (*resp.* $\mathrm{Card}(E')$) denotes the cardinality of E (*resp.* E'), $\mathrm{d}(E, E')$ the Euclidean distance between E and E' in $\mathbb{Z}^N$ defined by :

$$\mathrm{d}(E, E') = \min\left\{\left((i_1 - i'_1)^2 + \cdots + (i_N - i'_N)^2\right)^{1/2} : (i_1, \cdots, i_N) \in E,\ (i'_1, \cdots, i'_N) \in E'\right\},$$

and $\psi : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable. Usually, it will be assumed that ψ satisfies either

$$\psi(n, m) \leq C \min(n, m) \quad \forall n, m \in \mathbb{N} \tag{5.7}$$

or

$$\psi(n, m) \leq C(n + m + 1)^{\widetilde{\beta}}, \quad \forall n, m \in \mathbb{N} \tag{5.8}$$

for some $C > 0$ and some $\widetilde{\beta} \geq 1$. In the remainder of this chapter, we consider the first function defined in (5.7). Such function $\psi(m, n)$ can be found, for instance, in Tran (1990), Carbon et al. (1996, 1997), Biau & Cadre (2004), Dabo-Niang & Thiam (2010), Ould-Abdi et al. (2010) among others.

The function $\varphi(t)$ appearing in (5.6) is also usually assumed to satisfied one of the two following conditions :

$$\varphi(t) = O(t^{-\vartheta}) \quad \text{for some } \vartheta > 0 \tag{5.9}$$

or

$$\varphi(t) = O(e^{-st}) \quad \text{for some } s > 0. \tag{5.10}$$

However, in this work, for the sake of simplicity, we will suppose that φ satisfies (5.9) which means that φ decreases to zero at a polynomial rate. However, similar results of this chapter could be obtained when $\varphi(t)$ satisfies (5.10), that is φ tend to zero at an exponential rate which implies the polynomial case.

Note that, when $\psi \equiv 1$, then the process $(Z_{\mathbf{i}})_{\mathbf{i} \in \mathbb{Z}^N}$ is called strongly mixing. Many stochastic processes, among them various useful time series models satisfy strong mixing properties which are relatively easy to check. Conditions (5.6)-(5.10) are used in Tran (1990), Carbon et al. (1996, 1997). For discussions on mixing and examples, see Doukhan (1994).

5.3.1 Estimating the conditional tail index

In this subsection, we address the estimation of the unknown conditional extreme values index function $\gamma(\cdot)$ which will be used to estimate the conditional extreme quantiles. For this end, we consider the following kernel version of Hill's estimator proposed by Goegebeur et al. (2014b) in the non-spatial case, that is

$$\widehat{\gamma}_{\mathbf{n}}(x) = \frac{\sum\limits_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right)\left(\log \frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}}}{\sum\limits_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}}} := \frac{\widehat{\phi}_{\mathbf{n}}\left(v_{\mathbf{n}}, x\right)}{\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)}, \tag{5.11}$$

where $K_b(u) := b^{-d} K\left(u/b\right)$ is a probability density function on $\mathbb{R}^d$, $b := b_{\mathbf{n}}$ is a positive nonrandom sequence tending to 0 as $\mathbf{n} \to \infty$ called the bandwidth sequence, $\mathbb{1}_{\{\cdot\}}$ is the indicator function and $(v_{\mathbf{n}})$ is a nonrandom sequence tending to ∞.

The sequence $(v_{\mathbf{n}})$ is introduced in order to select only the largest observations to estimate $\gamma(x)$. Its choice is thus very important in practice. Goegebeur et al. (2014b) prove that, under appropriate regularity assumptions, estimator (5.11) will converge to the extreme values index $\gamma(x)$ of the conditional distribution of Y given $X = x$. Notice that the estimator (5.11) is similar to the one proposed by Csörgó et al. (1985) in the unconditional non spatial case.

5.3.2 Estimating the conditional extreme quantiles

In most applications, the extreme values index is not the primary object of interest, but for instance exceedance probabilities or extreme quantiles are to be estimated. Here, we address the estimation of the conditional extreme quantiles $q(\alpha_{\mathbf{n}}, x)$ of order $\alpha_{\mathbf{n}}$ of the distribution of Y given $X = x$. Those quantiles verify the following equation : $\overline{F}\left(q(\alpha_{\mathbf{n}}, x), x\right) = \alpha_{\mathbf{n}}$ where $\alpha_{\mathbf{n}}$ goes to 0 as $\mathbf{n} \to +\infty$ and $\overline{F}\left(y, x\right)$ denotes the distribution function of Y given $X = x$. To estimate the conditional distribution function of Y given $X = x$, we propose the kernel estimator introduced by Collomb (1980) and used recently in Daouia et al. (2011). This estimator is defined as follows, for $(y, x) \in \mathbb{R} \times \mathbb{R}^d$,

$$\widehat{\overline{F}}\left(y, x\right) = \frac{\sum\limits_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > y\}}}{\sum\limits_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right)}. \tag{5.12}$$

The associated estimator of the conditional quantile function $\widehat{\overline{F}}^{-1}\left(\cdot, x\right)$ for fixed $\alpha \in (0;1)$ can be defined via the generalized inverse of $\widehat{\overline{F}}\left(\cdot, x\right)$ as follows

$$\widehat{\overline{F}}^{-1}\left(\alpha, x\right) = \inf\left\{y \in \mathbb{R} \mid \widehat{\overline{F}}\left(y, x\right) \leq \alpha\right\}. \tag{5.13}$$

The best known extreme quantile estimators based on a semi-parametric approach is the estimator of Weissman (1978). This estimator is based on the estimation of $\gamma(x) > 0$. If $\alpha_{\mathbf{n}}$ is small enough and $\beta_{\mathbf{n}} < \alpha_{\mathbf{n}}$ we obtain the following quantities for all $\gamma(x) > 0$,

$$q(\alpha_{\mathbf{n}}, x) = \alpha_{\mathbf{n}}^{-\gamma(x)} L_q(\alpha_{\mathbf{n}}^{-1}, x), \tag{5.14}$$

$$q(\beta_{\mathbf{n}}, x) = \beta_{\mathbf{n}}^{-\gamma(x)} L_q(\beta_{\mathbf{n}}^{-1}, x). \tag{5.15}$$

Then dividing (5.14) by (5.15) and by the slowly varying function, we get the following approximation

$$q(\alpha_{\mathbf{n}}, x) \approx q(\beta_{\mathbf{n}}, x) \left(\frac{\alpha_{\mathbf{n}}}{\beta_{\mathbf{n}}}\right)^{-\gamma(x)}, \tag{5.16}$$

where $q(\beta_{\mathbf{n}}, x)$ is a chosen quantile in the sample which is easily estimated by inverting the empirical distribution function and $q(\alpha_{\mathbf{n}}, x)$ a chosen quantile outside the sample. So we estimate extremal quantiles of order $\beta_{\mathbf{n}}$ arbitrary small by extrapolating $\left(\frac{\alpha_{\mathbf{n}}}{\beta_{\mathbf{n}}}\right)^{-\gamma(x)}$ and replacing $q(\alpha_{\mathbf{n}}, x)$ by the estimator $\hat{q}(\alpha_{\mathbf{n}}, x) := \widehat{\overline{F}}^{-1}\left(\alpha_{\mathbf{n}}, x\right)$ given in (5.13) and $\gamma(x)$ by the above estimator $\widehat{\gamma}_{\mathbf{n}}(x)$. This allows us to obtain the estimator of Weissman (1978),

$$\hat{q}_{\mathbf{n}}^{W}\left(\beta_{\mathbf{n}}, x\right) := \hat{q}(\alpha_{\mathbf{n}}, x) \left(\frac{\alpha_{\mathbf{n}}}{\beta_{\mathbf{n}}}\right)^{\gamma_{\mathbf{n}}(x)}. \tag{5.17}$$

More details about the properties of this estimator can be found in Weissman (1978), Embrechts et al. (1997) and Haan & Ferreira (2006).

5.4 Assumptions and main results

Recall that, throughout this chapter, $\mathbb{R}^d$ is a metric space associated with a metric d_1. Moreover, we will assume that the following assumptions hold :

($\mathcal{F}$) The function $\gamma(x)$ is positive and continuous on $\mathbb{R}^d$ with support Λ included in the unit ball on $\mathbb{R}^d$ and suppose that there exists a constant $c_{\overline{F}} > 0$ and $y_0 > 1$ such that

$$\sup_{y \geq y_0} \left| \frac{\log \overline{F}(y, x)}{\log \overline{F}(y, z)} - 1 \right| \leq c_{\overline{F}} d_1(x, z).$$

The Karamata representation theorem of the slowly-varying function can be written as

$$L_q\,(y,x) = c(y,x) \exp\left(\int_1^y \frac{\mathcal{A}(u,x)}{u} du \right), \tag{5.18}$$

where $\lim\limits_{y\to\infty} c(y,x) = c(x) > 0$ and $\lim\limits_{y\to\infty} \mathcal{A}(y,x) = 0$.

($\mathcal{Q}$) The function $|\mathcal{A}(\cdot,x)|$ is ultimately decreasing and regularly varying with index $\rho(x)/\gamma(x) < 0$, that is for all $t > 0$,

$$\lim_{u\to\infty} \frac{|\mathcal{A}(tu,x)|}{|\mathcal{A}(u,x)|} = t^{\rho(x)/\gamma(x)}.$$

Moreover, $\mathcal{A}(\cdot,x)$ is assumed to have a constant sign at infinity.

Note that conditions (5.18) and ($\mathcal{Q}$) imply that

($\mathcal{R}_2$) for all $t > 0$,

$$\lim_{y\to\infty} \frac{\overline{F}(ty,x)/\overline{F}(y,x) - t^{-1/\gamma(x)}}{\mathcal{A}(y,x)} = t^{-1/\gamma(x)}\, \frac{t^{\rho(x)/\gamma(x)} - 1}{\gamma(x)\rho(x)}.$$

($\mathcal{K}$) The Kernel K is a density function with support Λ included in the unit ball on $\mathbb{R}^d$.

($\mathcal{W}$) The bandwidth b is such that

(i) $\hat{\mathbf{n}}\, b^d \to \infty$ as $\mathbf{n} \to \infty$ and there exists a sequence of integers $p = p_{\mathbf{n}} \to \infty$, such that

$$p = o\left[\hat{\mathbf{n}} \left(b^d \overline{F}(v_{\mathbf{n}},x)\right)^{(1+(1-\varsigma))2N}\right]^{2N} \quad \hat{\mathbf{n}} p^{-\vartheta} \to 0 \text{ with } \quad 0 < \varsigma < (\vartheta - N)/\vartheta, \tag{5.19}$$

where $\vartheta > 2(N+1)$, ϑ is defined in (5.9).

(ii) $\left(b^d \overline{F}(v_{\mathbf{n}},x)\right)^{-(1-\varsigma)} p^{N-\vartheta(1-\varsigma)} \to 0$ as $\mathbf{n} \to \infty$.

($\mathcal{G}$) (i) For any x, $x' \in \mathbb{R}^d$, there exists $c_g > 0$ such that $|g(x) - g(x')| \le c_g d_1\,(x,x')$. Moreover, f is continuous in $\mathbb{R}^{d+1}$.

(ii) For any $\mathbf{i} \neq \mathbf{j}$, the pairs $(X_{\mathbf{i}}, X_{\mathbf{j}})$ $((Y_{\mathbf{i}}, X_{\mathbf{i}}), (Y_{\mathbf{j}}, X_{\mathbf{j}}))$ admit respectively uniformly joint densities $g_{\mathbf{i},\mathbf{j}}$ and $f_{\mathbf{i},\mathbf{j}}$ such that there exists positive constants C and c satisfying

$$\sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} |g_{\mathbf{i},\mathbf{j}}(x,y) - g(x)g(y)| \le C.$$

$$\sup_{(u,v)\in\mathbb{R}^{d+1}\times\mathbb{R}^{d+1}} |f_{\mathbf{i},\mathbf{j}}(u,v) - f(u)f(v)| \le c.$$

Comments 1 : Note that Assumption ($\mathcal{F}$) controls the oscillation of $\log \overline{F}(y;x)$ with respect to its second argument x. Assumption ($\mathcal{R}_2$) is the so-called second order condition classically used to establish the asymptotic normality of tail index estimators. Note that the second order parameter $\rho(x)$ controls the rate of convergence of $L(\lambda y, x)/L(y,x)$ to 1 (see Bingham et al. (1987) and Haan & Ferreira (2006) for further details). In particular, if $\rho(x)$ is close to 0, the convergence is slow and thus, the estimation of the conditional tail index is difficult. Note that this condition is also used in Goegebeur et al. (2014b). $(\mathcal{K})$ is classically used in nonparametric estimation. To provide a local Hill-type estimator (5.11), one can choose for instance K such that $K = \mathbb{1}_{\{S\}}/|S|$, where S is the $d-$dimensional unit sphere with volume $|S|$. Finally, $(\mathcal{W})$ are technical conditions required for the bandwidth in order to get convergence results of our proposed estimator, they are related to the spatial context. $(\mathcal{G})$ is a local dependency condition classical in non-parametric inference for spatial data, see for instance Carbon et al. (1996, 1997), Ould-Abdi et al. (2010).

The following theorem gives the asymptotic convergence of the conditional tail index estimator $\gamma_{\mathbf{n}}(x)$.

Theorem 5.4.1. *Suppose the spatial process* $\left\{Z_{\mathbf{i}} = (Y_{\mathbf{i}}, X_{\mathbf{i}})\,;\ \mathbf{i} \in \mathbb{Z}^N\right\}$ *satisfies the above mixing condition* (5.6)*. If* $(\mathcal{F})$, $(\mathcal{K})$, $(\mathcal{Q})$, $(\mathcal{W})$ *and* (5.18) *hold and let* $(v_{\mathbf{n}})$ *be a nonrandom sequence such that*

▷ *As* $\mathbf{n} \to \infty$, $v_{\mathbf{n}} \to \infty$, $\hat{\mathbf{n}} b_{\mathbf{n}}^d \overline{F}\left(v_{\mathbf{n}}, x\right) \to \infty$ *and* $\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^d \overline{F}(v_{\mathbf{n}}, x)}\, b_{\mathbf{n}} \ln v_{\mathbf{n}} \to 0.$

▷ *As* $\mathbf{n} \to \infty$, $\sqrt{\hat{\mathbf{n}} b^d \overline{F}(v_{\mathbf{n}}, x)}\,\mathcal{A}(v_{\mathbf{n}}, x) \to \lambda(x) < \infty.$

Then, for all $x \in \mathbb{R}^d$ *such that* $g(x) > 0$,

$$\sqrt{\hat{\mathbf{n}} b^d \overline{F}_{\mathbf{n}}(v_{\mathbf{n}}, x)}\,\left(\widehat{\gamma}_{\mathbf{n}}(x) \ - \ \gamma(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)}, \gamma^2(x)\frac{\|K\|_2^2}{g(x)}\right). \tag{5.20}$$

Note that the asymptotic variance in (5.20) is the same as the asymptotic variance obtained by Csörgó et al. (1985) in the unconditional non spatial case up to the scale factor $\frac{\|K\|_2^2}{g(x)}$.

We now examine the asymptotic properties of the estimator of the extreme quantile given in (5.17). First, let us give some comments on the auxiliary function $\mathcal{A}(\cdot, x)$.

Comments 2 : The function $\mathcal{A}(\cdot, x)$ in (5.18) plays an important role in extreme values theory since it drives the speed of convergence in (5.5) and more generally the bias of extreme values estimators. Considering the estimation of extreme quantiles $q(\alpha_{\mathbf{n}}, x)$ with $\alpha_{\mathbf{n}} \to 0$ as $\mathbf{n} \to \infty$; as mentioned in Daouia et al. (2011), this kernel estimator of extreme quantiles requires a stringent condition on the order $\alpha_{\mathbf{n}}$ of the quantile, since by construction it cannot extrapolate beyond the maximum observation.

The following result gives the asymptotic normality of the extreme conditional quantile (5.17) satisfying the above conditions.

Theorem 5.4.2. *Assume that conditions of Theorem 5.4.1 holds. In addition, let* $(\alpha_{\mathbf{n}})_{\mathbf{n}\geq 1}$ *and* $(\beta_{\mathbf{n}})_{\mathbf{n}\geq 1}$ *be positive sequences such that*

▷ *As* $\mathbf{n} \to \infty$, $\alpha_{\mathbf{n}} \to 0$, $\beta_{\mathbf{n}}/\alpha_{\mathbf{n}} \to 0$, $\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^d \overline{F}(v_{\mathbf{n}}, x)}\,\mathcal{A}\left(q(\alpha_{\mathbf{n}}, x), x\right) \to \lambda(x).$

▷ *As* $\mathbf{n} \to \infty$, $\left|\frac{\hat{q}(\alpha_{m_x}, x)}{q(\alpha_{m_x}, x)} - 1\right| \xrightarrow{\mathbb{P}} 0.$

Then, for any $x \in \mathbb{R}^d$ *such that* $g(x) > 0$,

$$\frac{\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^d \overline{F}(v_{\mathbf{n}}, x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}\left(\frac{\hat{q}_{\mathbf{n}}^W(\beta_{\mathbf{n}}, x)}{q(\beta_{\mathbf{n}}, x)} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)},\ \gamma^2(x)\frac{\|K\|_2^2}{g(x)}\right). \tag{5.21}$$

5.5 Conclusion

In this chapter, we propose kernel estimators of the tail index and quantile of a heavy-tailed distribution for spatial data in the presence of random covariates. We establish consistency properties under general conditions by showing the asymptotic normality of a tail index kernel estimator. This convergence is obtained by considering some $\alpha-$mixing condition on the underlying process in addition to the usual classical ones used in tail index estimation. The consistency result of the tail estimate permits to derive an asymptotic normality result of a Weissman-type extreme quantile. Other types (moment-type, UH-type, ...) of tail index estimator and corresponding quantiles will be considered. As we only consider here consistency results, numerical illustrations are a subject of futur investigations.

5.6 Appendix : Technical lemmas et proofs of main results

5.6.1 Appendix A1 : Technical lemmas

In this section, we establish the main results and give the necessary technical lemma. To prove our results, we need the three following lemmas. There proofs are omitted.

Lemma 5.6.1 (Tran (1990)). ***(i)*** *Suppose that* (5.6) *holds. Denote by* $\mathcal{L}_r(\mathcal{F})$ *the class of* $\mathcal{F}$*-measurable r.v's* X *satisfying* $\|X\|_r = (\mathbf{E}\,|X|^r)^{\frac{1}{r}} < \infty$. *Suppose* $X_1 \in \mathcal{L}_r(\mathcal{B}(\mathbf{I}_1))$ *and* $X_2 \in \mathcal{L}_r(\mathcal{B}(\mathbf{I}_2))$. *Assume also that* $1 \leq r, s, t < \infty$ *and* $r^{-1}+s^{-1}+t^{-1} = 1$. *Then*

$$|\mathbf{E}(X_1X_2) - \mathbf{E}(X_1)\mathbf{E}(X_2)| \leq C\,\|X_1\|_r\,\|X_2\|_s \left\{\psi\left(Card(\mathbf{I}_1), Card(\mathbf{I}_2)\right)\varphi\left(\mathrm{d}(\mathbf{I}_1,\mathbf{I}_2)\right)\right\}^{\frac{1}{t}}.$$

(ii) *For r.v's bounded with probability* 1, *the right-hand side of the last inequality can be replaced by*

$$C\psi\left(Card(\mathbf{I}_1), Card(\mathbf{I}_2)\right)\varphi\left(\mathrm{d}(\mathbf{I}_1,\mathbf{I}_2)\right).$$

Lemma 5.6.2 (Tran (1990)). *Let* $(\xi_1, \cdots, \xi_n)$ *be a random vector such that* $\left|\mathbf{E}\left[\prod_{s=i}^{n}\xi_i\right]\right| < \infty$, $i = 1, \cdots, n$, $|C\xi_i| \leq 1$, $i = 1, \cdots, n$. *Then*

$$\left|\mathbf{E}\left[\prod_{s=1}^{n}\xi_s\right] - \prod_{s=1}^{n}\mathbf{E}\left[\xi_s\right]\right| \leq \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left|\mathbf{E}\left[\left(\xi_i-1\right)\left(\xi_j-1\right)\right]\prod_{s=j+1}^{n}\xi_s\right.$$
$$\left.\times -\mathbf{E}\left[\left(\xi_i-1\right)\right]\mathbf{E}\left[\left(\xi_j-1\right)\right]\prod_{s=j+1}^{n}\xi_s\right|.$$

Lemma 5.6.3 (Carbon et al. (1996)). *Suppose* $S_1, S_2, \cdots, S_r$ *be sets containing* m *sites each with* $dist(S_i, S_j) \geq \delta$, $\delta > 0$ *for all* $i \neq j$ *where* $1 \leq i, j \leq r$. *Suppose* $Y_1, \cdots, Y_r$ *is a sequence of real-valued random variables measurable with respect to* $\mathcal{B}\left(S_1\right), \cdots, \mathcal{B}\left(S_r\right)$ *respectively, and* Y_i *takes values in* $[a; b]$. *Then there exists a sequence of independent random variables* $Y_1^*, \cdots, Y_r^*$ *independent of* $Y_1, \cdots, Y_r$ *such that* Y_i^* *has the same distribution as* Y_i *and satisfies :*

$$\sum_{i=1}^{r}\mathbf{E}\,|Y_i - Y_i^*| \leq 2r\,(b-a)\,\Phi\left((r-1)m, m\right)\varphi\left(\delta\right). \tag{5.22}$$

5.6.2 Appendix A2 : Notations and definitions

First, we establish the proofs of our main results by employing the blocking technique used in Carbon et al. (1997) and Tran (1990). Without loss of generality assume that $\hat{\mathbf{n}} = \widehat{\mathbf{r}}(p+q)^N$ where $\widehat{\mathbf{r}} = r_1 \times ... \times r_N$. Denote

$$\begin{aligned}
\mathbf{I}(1,x,\mathbf{j}) &= \{\mathbf{i},\, j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\, k=1,...,N\} \qquad (5.23)\\
\mathbf{I}(2,x,\mathbf{j}) &= \Big\{\mathbf{i},\, j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\, k=1,...,N-1,\\
&\quad j_N(p+q)+p+1 \le i_N \le (j_N+1)(p+q)\Big\},\\
\mathbf{I}(3,x,\mathbf{j}) &= \Big\{\mathbf{i},\, j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\, k=1,...,N-2,\\
&\quad j_{N-1}(p+q)+p+1 \le i_{N-1} \le (j_{N-1}+1)(p+q),\\
&\quad j_N(p+q)+1 \le i_N \le j_N(p+q)+p\Big\},\\
\mathbf{I}(4,x,\mathbf{j}) &= \Big\{\mathbf{i},\, j_k(p+q)+1 \le i_k \le j_k(p+q)+p,\, k=1,...,N-2,\\
&\quad j_{N-1}(p+q)+p+1 \le i_{N-1} \le (j_{N-1}+1)(p+q),\\
&\quad j_N(p+q)+p+1 \le i_N \le (j_N+1)(p+q)\Big\}
\end{aligned}$$

and so on. The last two terms are

$$\begin{aligned}
\mathbf{I}(2^{N-1},x,\mathbf{j}) &= \Big\{\mathbf{i},\, j_k(p+q)+p+1 \le i_k \le (j_k+1)(p+q), k=1,...,N-1,\\
&\quad j_N(p+q)+1 \le i_N \le j_N(p+q)+p\Big\} \qquad (5.24)
\end{aligned}$$

and

$$\mathbf{I}(2^N,x,\mathbf{j}) = \{\mathbf{i},\, j_k(p+q)+p+1 \le i_k \le (j_k+1)(p+q),\, k=1,...,N,\} \qquad (5.25)$$

For each integer $1 \le i \le 2^N$, define

$$D(i,x) = \{\mathbf{I}(i,x,\mathbf{j}),\, 0 \le j_k \le r_k - 1,\, k=1,...,N\} \qquad (5.26)$$

Note that the set $\mathcal{I}_{\mathbf{n}}$ is then decomposed into these 2^N small and large blocks in sets $D(i,x)$. If it is not the case that $\hat{\mathbf{n}} = \widehat{\mathbf{r}}(p+q)^N$ then use an additional set $D(2^N+1,x)$ containing all the sites of $\mathcal{I}_{\mathbf{n}}$ not in these blocks $\mathbf{I}(i,x,\mathbf{j})$.

Now we will give some definitions and notations to simplify the formulas. First we rewrite $\widehat{\overline{F}}(.,x)$ the estimator of $\overline{F}(.,x)$ as follows

$$\widehat{\overline{F}}(t,x) = \pi_{\mathbf{n}}(t,x)/g_{\mathbf{n}}(x)$$

where $\pi_{\mathbf{n}}(t,x) := \frac{1}{\hat{\mathbf{n}}} \sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right) \mathbb{1}_{\{Y_{\mathbf{i}}>t\}}$ and $g_{\mathbf{n}}(x) := \frac{1}{\hat{\mathbf{n}}} \sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right)$ is the classical kernel density estimator of g. Let $\pi(t,x) = g(x)\overline{F}(t,x)$ and $\widehat{\phi}_{v_{\mathbf{n}}}\left(y,x\right) := \frac{1}{\hat{\mathbf{n}}} \sum\limits_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_b\left(x - X_{\mathbf{i}}\right) \left(\log \frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right) \mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}$. In addition we define the following quantities

$$\begin{aligned}
\mathbb{W}_{\mathbf{n},\pi}(x) &:= \sqrt{\hat{\mathbf{n}} b^d \pi\left(v_{\mathbf{n}},x\right)} \left(\frac{\pi_{\mathbf{n}}\left(v_{\mathbf{n}},x\right) - \mathbf{E}\left[\pi_{\mathbf{n}}\left(v_{\mathbf{n}},x\right)\right]}{\pi\left(v_{\mathbf{n}},x\right)}\right)\\
\mathbb{W}_{\mathbf{n},\phi}\left(x,y\right) &:= \sqrt{\hat{\mathbf{n}} b^d \pi(v_{\mathbf{n}},x)} \left(\frac{\widehat{\phi}_{v_{\mathbf{n}}}\left(y,x\right) - \mathbf{E}\left[\widehat{\phi}_{v_{\mathbf{n}}}\left(y,x\right)\right]}{\pi(v_{\mathbf{n}},x)}\right).
\end{aligned}$$

To simplify the formulas for simple reading, we introduce the following notations. Set

$$\eta_{\mathbf{j},\mathbf{n}} := \frac{1}{\pi(v_{\mathbf{n}},x)} K_b\Big(x - X_{\mathbf{j}}\Big) \mathbb{1}_{\big\{Y_{\mathbf{j}}>v_{\mathbf{n}}\big\}}, \qquad \Delta_{\mathbf{j},\mathbf{n}} := \eta_{\mathbf{j},\mathbf{n}} - \mathbf{E}\left[\eta_{\mathbf{j},\mathbf{n}}\right],$$
$$\widetilde{\eta}_{\mathbf{i},\mathbf{n}} := \frac{1}{\pi(v_{\mathbf{n}},x)} K_b\left(x - X_{\mathbf{i}}\right) \left(\log \frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right) \mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}, \quad \widetilde{\Delta}_{\mathbf{i},\mathbf{n}} := \widetilde{\eta}_{\mathbf{i},\mathbf{n}} - \mathbf{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\right].$$

Finally, define

$$S_{\mathbf{n}}(x) := \sum_{\mathbf{i}\in I_{\mathbf{n}}} \Big[b^d \pi\left(v_{\mathbf{n}},x\right)\Big]^{1/2} \frac{\Delta_{\mathbf{i},\mathbf{n}}}{\hat{\mathbf{n}}} \text{ and } \widetilde{S}_{\mathbf{n}}(x,y) := \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Big[b^d \pi(v_{\mathbf{n}},x)\Big]^{1/2} \frac{\widetilde{\Delta}_{\mathbf{i},\mathbf{n}}}{\hat{\mathbf{n}}}. \quad (5.27)$$

Then we can write

$$\mathbb{W}_{\mathbf{n},\pi}(x) = \hat{\mathbf{n}}^{1/2} S_{\mathbf{n}}(x), \quad \mathbb{W}_{\mathbf{n},\phi}\left(x,y\right) = \hat{\mathbf{n}}^{1/2} \widetilde{S}_{\mathbf{n}}(x,y). \qquad (5.28)$$

It seems that what we need is some way of dealing with a sequence of sequences, say, $\{\Delta_{\mathbf{i},\mathbf{n}}\}_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}$. This is exactly the idea of a triangular array of random variables.

5.6.3 Appendix A3 : Intermediate lemmas

The following lemma gives the variances of $S_{\mathbf{n}}(x)$ and $\widetilde{S}_{\mathbf{n}}(x,y)$.

Lemma 5.6.4. *Suppose* (5.3) *and* (5.6) *and conditions of Theorem 5.4.1 hold, then we have*

1. $\hat{\mathbf{n}}\mathrm{Var}S_{\mathbf{n}}(x) = \overline{\sigma}(x)^2\left(1+o(1)\right)$, *where* $\overline{\sigma}(x)^2 = \|K\|_2^2$.
2. $\hat{\mathbf{n}}\mathrm{Var}\widetilde{S}_{\mathbf{n}}(x,y) = \widetilde{\sigma}(x)^2\left(1+o(1)\right)$, *where*

$$\begin{aligned}\widetilde{\sigma}^2 &= 2\gamma^2(x)\,\|K\|_2^2\left(1+O(b\log v_{\mathbf{n}})\right)\\ &\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}.\end{aligned}$$

Proof of Lemma 5.6.4

1. Before going further, let us denote

$$I_{\mathbf{n},x} := \hat{\mathbf{n}}^{-2} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathrm{Var}\left(\eta_{\mathbf{i},\mathbf{n}}\right) \qquad \text{and} \qquad R_{\mathbf{n},x} := \hat{\mathbf{n}}^{-2} \sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} \left|\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)\right|.$$

In this part, we give the variance of $S_{\mathbf{n}}(x)$ by using the method proposed by Tran (1990). Carrying over the above, it is easily seen that

$$\mathrm{Var}\left(\hat{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n}}\right) = \hat{\mathbf{n}}^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\mathrm{Var}\left(\eta_{\mathbf{i},\mathbf{n}}\right) \;+\; \hat{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\sum_{\mathbf{i}\neq\mathbf{j}}\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right).$$

Clearly, $\mathrm{Var}\left(\hat{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i},\mathbf{n}}\right) \leq I_{\mathbf{n},x} \;+\; R_{\mathbf{n},x}$. We deduce that,

$$I_{\mathbf{n},x} = \hat{\mathbf{n}}^{-1}\mathbf{E}\left[\Delta_{\mathbf{i},\mathbf{n}}^2\right] = \hat{\mathbf{n}}^{-1}\left(\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}^2\right]-\mathbf{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right]^2\right).$$

It follows

$$I_{\mathbf{n},x} = \frac{1}{\hat{\mathbf{n}}\pi^2(v_{\mathbf{n}},x)}\left(\mathbf{E}\left[K_b^2(x-X_{\mathbf{i}})\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]-\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]^2\right).$$

From Lemma 2 of Goegebeur et al. (2014b),

$$\mathbf{E}\left[K_b^2(x-X_{\mathbf{i}})\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] = \pi\left(v_{\mathbf{n}},x\right)b^{-d}\left\|K\right\|_2^2\left(1+O(b\log v_{\mathbf{n}})\right).$$

And from the first statement of Lemma 5.6.6 below,

$$\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] = \pi\left(v_{\mathbf{n}},x\right)\left(1+O(b\log v_{\mathbf{n}})\right).$$

It follows that

$$I_{\mathbf{n},x} = \hat{\mathbf{n}}^{-1}b^{-d}\pi^{-1}(v_{\mathbf{n}},x)\left\|K\right\|_2^2\left(1+O(b\log v_{\mathbf{n}})\right) - \hat{\mathbf{n}}^{-1}\left(1+O(b\log v_{\mathbf{n}})\right)^2. \qquad (5.29)$$

We now treat the term $R_{\mathbf{n},x}$. We want to show that there exists a constant C not depending on $\hat{\mathbf{n}}$ such that $R_{\mathbf{n},x} \leq C\hat{\mathbf{n}}^{-1}b^{-d}\pi^{-1}\left(v_{\mathbf{n}},x\right)$ for $\hat{\mathbf{n}}$ large enough.

$$\begin{aligned}\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right) &= \pi^{-2}\left(v_{\mathbf{n}},x\right)\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}K_b(x-X_{\mathbf{j}})\mathbb{1}_{\left\{Y_{\mathbf{j}}>v_{\mathbf{n}}\right\}}\right]\\ &\quad - \pi^{-2}\left(v_{\mathbf{n}},x\right)\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\mathbf{E}\left[K_b(x-X_{\mathbf{j}})\mathbb{1}_{\left\{Y_{\mathbf{j}}>v_{\mathbf{n}}\right\}}\right].\end{aligned}$$

We will study the covariance between $\eta_{\mathbf{i},\mathbf{n}}$ and $\eta_{\mathbf{j},\mathbf{n}}$ for all $\mathbf{i}\neq\mathbf{j}$ in two sets defined as follows, let : $\mathcal{S}_{\Omega_{\mathbf{n}}} = \left\{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}} : 0<\mathrm{d}\left(\mathbf{i},\mathbf{j}\right)<\Omega_{\mathbf{n}}\right\}$ and denote by $\mathcal{S}^c_{\Omega_{\mathbf{n}}}$ the complement of $\mathcal{S}_{\Omega_{\mathbf{n}}}$ where $\Omega_{\mathbf{n}}$ is a sequence of real numbers which tends to ∞ as $\mathbf{n}\to\infty$. Define

$$R_{1,\mathbf{n},x} := \hat{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)\right| \quad\text{and}\quad R_{2,\mathbf{n},x} := \hat{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^c_{\Omega_{\mathbf{n}}}}\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)\right|.$$

Clearly, $R_{\mathbf{n},x} \leq R_{1,\mathbf{n},x} + R_{2,\mathbf{n},x}$. Applying the first part of Lemma 5.6.1, we get,

$$\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)\right| \leq \|\eta_{\mathbf{i},\mathbf{n}}\|_s\|\eta_{\mathbf{j},\mathbf{n}}\|_u(\psi(1,1)\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right))^{1/t},$$

with $1/t+1/s+1/u=1$, in particular, we choose $s=u=4$ and $t=2$. We have

$$\begin{aligned}\|\eta_{\mathbf{i},\mathbf{n}}\|_s^s &= \pi^{-s}\left(v_{\mathbf{n}},x\right)\mathbf{E}\left[K_b^s(x-X_{\mathbf{i}})\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\\ &= \pi^{1-s}\left(v_{\mathbf{n}},x\right)b^{-(s-1)d}\left\|K\right\|_s^s\left(1+O(b\log v_{\mathbf{n}})\right).\end{aligned}$$

It follows that

$$\left|\operatorname{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)\right| \leq \pi^{-\frac{1}{t}-1}\left(v_{\mathbf{n}},x\right)b^{-(1+\frac{1}{t})d}\left\|K\right\|_s\left\|K\right\|_u\varepsilon_{s,u}(x)\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t},$$

with $\varepsilon_{s,u}(x) := \left(1+O(b\log v_{\mathbf{n}})\right)^{1-\frac{1}{t}}$.

- ▶ For the term $R_{1,\mathbf{n},x}$, it follows that

$$\begin{aligned}R_{1,\mathbf{n},x} &\leq \hat{\mathbf{n}}^{-2}\|\eta_{\mathbf{i},\mathbf{n}}\|_s\|\eta_{\mathbf{j},\mathbf{n}}\|_u\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\sum\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}\\ &\leq \hat{\mathbf{n}}^{-2}\pi^{-\frac{1}{t}-1}\left(v_{\mathbf{n}},x\right)b^{-(1+\frac{1}{t})d}\left\|K\right\|_s\left\|K\right\|_u\varepsilon_{s,u}(x)\Omega_{\mathbf{n}}^N\sum_{\|\mathbf{i}\|\leq\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/t},\end{aligned}$$

$$\hat{\mathbf{n}}b^d\pi\left(v_{\mathbf{n}},x\right)R_{1,\mathbf{n},x} \leq \hat{\mathbf{n}}^{-1}\pi^{-\frac{1}{t}}\left(v_{\mathbf{n}},x\right)b^{-\frac{d}{t}}\Omega_{\mathbf{n}}^N\left\|K\right\|_s\left\|K\right\|_u\varepsilon_{s,u}(x)\sum_{\mathbf{i}}\|\mathbf{i}\|^{-\vartheta/t}. \qquad (5.30)$$

- For the term $R_{2,\mathbf{n},x}$, we get

$$R_{2,\mathbf{n},x} \leq \hat{\mathbf{n}}^{-2} \|\eta_{\mathbf{i},\mathbf{n}}\|_s \|\eta_{\mathbf{j},\mathbf{n}}\|_u \sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^c_{\Omega_\mathbf{n}}} \sum \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t} \tag{5.31}$$

$$\leq \hat{\mathbf{n}}^{-1} \pi^{-\frac{1}{t}-1}\left(v_\mathbf{n}, x\right) b^{-(1+\frac{1}{t})d} \|K\|_s \|K\|_u \varepsilon_{s,u}(x) \sum_{\|\mathbf{i}\|>\Omega_\mathbf{n}} \|\mathbf{i}\|^N \|\mathbf{i}\|^{-N} \varphi\left(\|\mathbf{i}\|\right)^{1/t}.$$

$$\hat{\mathbf{n}} b^d \pi\left(v_\mathbf{n}, x\right) R_{2,\mathbf{n},x} \leq \pi^{-\frac{1}{t}}\left(v_\mathbf{n}, x\right) b^{-\frac{d}{t}} \Omega_\mathbf{n}^{-N} \|K\|_s \|K\|_u \varepsilon_{s,u}(x) \sum_{\mathbf{i}\geq 0}^{\infty} \|\mathbf{i}\|^{\frac{tN-\vartheta}{t}}. \tag{5.32}$$

Since $\vartheta > 2(N+1)$ (then $\sum_{\mathbf{i}>0} \|\mathbf{i}\|^{\frac{N-\vartheta}{t}} < \infty$) and by the fact that $\varepsilon_{s,u}(x) \longrightarrow 1$ as $\mathbf{n} \to \infty$, letting $\Omega_\mathbf{n} = \left(\hat{\mathbf{n}} b^{d/2} \pi(v_\mathbf{n}, x)^{1/2}\right)^{1/N}$ and combining (5.30) and (5.32), gives

$$R_{\mathbf{n},x} = O\left(b^{-d} \hat{\mathbf{n}}^{-1} \pi^{-1}\left(v_\mathbf{n}, x\right)\right). \tag{5.33}$$

The proof of statement 1 follows from (5.29), (5.33) and by the fact that

$$\begin{aligned}\hat{\mathbf{n}} \mathrm{Var}\left(S_\mathbf{n}(x)\right) &= \hat{\mathbf{n}} b^d \pi\left(v_\mathbf{n}, x\right) \mathrm{Var}\left(\hat{\mathbf{n}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \Delta_{\mathbf{i},\mathbf{n}}\right) \\ &= \|K\|_2^2 \left(1 + O(b \log v_\mathbf{n})\right) - \hat{\mathbf{n}}^{-1} \left(1 + O(b \log v_\mathbf{n})\right)^2 + o(1).\end{aligned}$$

2. The variance of $\widetilde{S}_\mathbf{n}(x,y)$ will be treated exactly as in the first statement.
Let us define

$$\widetilde{I}_{\mathbf{n},x,y} := \hat{\mathbf{n}}^{-2} \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \mathrm{Var}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\right), \quad \text{and} \quad \widetilde{R}_{\mathbf{n},x,y} := \hat{\mathbf{n}}^{-2} \sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_\mathbf{n}} \sum_{\mathbf{i}\neq\mathbf{j}} \left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}}, \widetilde{\eta}_{\mathbf{j},\mathbf{n}}\right)\right|.$$

Then, we get

$$\mathrm{Var}\left(\hat{\mathbf{n}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \widetilde{\Delta}_{\mathbf{i},\mathbf{n}}\right) = \hat{\mathbf{n}}^{-2} \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \mathrm{Var}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\right) + \hat{\mathbf{n}}^{-2} \sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_\mathbf{n}} \sum_{\mathbf{i}\neq\mathbf{j}} \mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}}, \widetilde{\eta}_{\mathbf{j},\mathbf{n}}\right)$$

Clearly, $\mathrm{Var}\left(\hat{\mathbf{n}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} \widetilde{\Delta}_{\mathbf{i},\mathbf{n}}\right) \leq \widetilde{I}_{\mathbf{n},x,y} + \widetilde{R}_{\mathbf{n},x,y}$. We deduce that,

$$\widetilde{I}_{\mathbf{n},x,y} = \hat{\mathbf{n}}^{-1} \mathbf{E}\left[\widetilde{\Delta}^2_{\mathbf{i},\mathbf{n}}\right] = \hat{\mathbf{n}}^{-1} \left(\mathbf{E}\left[\widetilde{\eta}^2_{\mathbf{i},\mathbf{n}}\right] - \mathbf{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\right]^2\right).$$

It follows that

$$\widetilde{I}_{\mathbf{n},x,y} = \frac{1}{\hat{\mathbf{n}} \pi^2(v_\mathbf{n}, x)} \left(\mathbf{E}\left[K_b^2(x - X_\mathbf{i}) \left(\log \frac{Y_\mathbf{i}}{v_\mathbf{n}}\right)^2 \mathbb{1}_{\{Y_\mathbf{i} > v_\mathbf{n}\}}\right] - \mathbf{E}\left[K_b(x - X_\mathbf{i}) \left(\log \frac{Y_\mathbf{i}}{v_\mathbf{n}}\right) \mathbb{1}_{\{Y_\mathbf{i} > v_\mathbf{n}\}}\right]^2\right).$$

From Lemma 2 of Goegebeur et al. (2014b),

$$\begin{aligned}\mathbf{E}\left[K_b^2(x - X_\mathbf{i}) \left(\log \frac{Y_\mathbf{i}}{v_\mathbf{n}}\right)^2 \mathbb{1}_{\{Y_\mathbf{i} > v_\mathbf{n}\}}\right] &= 2\gamma^2(x) \pi\left(v_\mathbf{n}, x\right) b^{-d} \|K\|_2^2 \left(1 + O(b \log v_\mathbf{n})\right) \\ &\quad \times \left\{1 + \frac{\mathcal{A}(v_\mathbf{n}, x)}{\gamma(x)\rho(x)} \left[\frac{1}{(1-\rho(x))^2} - 1\right] (1 + o(1))\right\}.\end{aligned}$$

And from the first statement of Lemma 5.6.6 below, see also Lemma 2 of Goegebeur et al. (2014b),

$$\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]=\gamma(x)\pi\left(v_{\mathbf{n}},x\right)\left(1+O(b\log v_{\mathbf{n}})\right)$$
$$\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\}.$$

It follows that

$$\begin{aligned}\widetilde{I}_{\mathbf{n},x,y}&=2\gamma^2(x)\pi^{-1}\left(v_{\mathbf{n}},x\right)\hat{\mathbf{n}}^{-1}b^{-d}\left\|K\right\|_2^2\left(1+O(b\log v_{\mathbf{n}})\right)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}\\&\quad-\gamma^2(x)\hat{\mathbf{n}}^{-1}\left(1+O(b\log v_{\mathbf{n}})\right)^2\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\}^2.\end{aligned}\tag{5.34}$$

Now let us examine the term $\widetilde{R}_{\mathbf{n},x,y}$. We want to show that there exists a constant not depending on $\mathbf{n}$ such that $\widetilde{R}_{\mathbf{n},x,y}\leq C\hat{\mathbf{n}}^{-1}b^{-d}\pi^{-1}\left(v_{\mathbf{n}},x\right)$ for $\mathbf{n}$ large enough. We have

$$\begin{aligned}\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}},\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\right)&=\pi^{-2}\left(v_{\mathbf{n}},x\right)\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}K_b(x-X_{\mathbf{j}})\left(\log\frac{Y_{\mathbf{j}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\left\{Y_{\mathbf{j}}>v_{\mathbf{n}}\right\}}\right]\\&\quad-\pi^{-2}\left(v_{\mathbf{n}},x\right)\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\mathbf{E}\left[K_b(x-X_{\mathbf{j}})\left(\log\frac{Y_{\mathbf{j}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\left\{Y_{\mathbf{j}}>v_{\mathbf{n}}\right\}}\right]\end{aligned}$$

In this part, to study the covariance between $\widetilde{\eta}_{\mathbf{i},\mathbf{n}}$ and $\widetilde{\eta}_{\mathbf{j},\mathbf{n}}$ for all $\mathbf{i}\neq\mathbf{j}$, we use the previously defined two sets which are $\mathcal{S}_{\Omega_{\mathbf{n}}}=\left\{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}:\,0<\mathrm{d}\left(\mathbf{i},\mathbf{j}\right)<\Omega_{\mathbf{n}}\right\}$ and $\mathcal{S}^c_{\Omega_{\mathbf{n}}}$. Denote

$$\widetilde{R}_{1,\mathbf{n},x,y}:=\hat{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}},\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\right)\right|\quad\text{and}\quad\widetilde{R}_{2,\mathbf{n},x,y}:=\hat{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^c_{\Omega_{\mathbf{n}}}}\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}},\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\right)\right|.$$

Clearly we have that : $\widetilde{R}_{\mathbf{n},x,y}\leq\widetilde{R}_{1,\mathbf{n},x,y}\ +\ \widetilde{R}_{2,\mathbf{n},x,y}$. Applying again the first part of Lemma 5.6.1, on the $\widetilde{\eta}_{\mathbf{i},\mathbf{n}}$, permits to get

$$\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}},\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\right)\right|\leq\|\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\|_s\|\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\|_u(\psi(1,1)\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right))^{1/t},$$

with $1/t+1/s+1/u=1,\ u=s=4,\ t=2$. We have

$$\begin{aligned}\|\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\|_s^s&=\pi^{-s}\left(v_{\mathbf{n}},x\right)\mathbf{E}\left[K_b^s(x-X_{\mathbf{i}})\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)^s\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\\&=s!\pi^{1-s}\left(v_{\mathbf{n}},x\right)\gamma^s(x)b^{-(s-1)d}\left\|K\right\|_s^s\\&\quad\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^s}-1\right](1+o(1))\right\}\left(1+O(b\log v_{\mathbf{n}})\right).\end{aligned}$$

It follows that

$$\left|\mathrm{Cov}\left(\widetilde{\eta}_{\mathbf{i},\mathbf{n}},\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\right)\right|\leq\pi^{-\frac{1}{t}-1}\left(v_{\mathbf{n}},x\right)\gamma^2(x)b^{-(1+\frac{1}{t})d}\left\|K\right\|_s\left\|K\right\|_u\widetilde{\varepsilon}_{s,u}(x)\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t},\tag{5.35}$$

with

$$\widetilde{\varepsilon}_{s,u}(x) := (s!)^{1/s}\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^{s}}-1\right](1+o(1))\right\}^{\frac{1}{s}}(1+O(b\log v_{\mathbf{n}}))^{\frac{1}{s}}$$
$$\times\,(u!)^{1/u}\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^{u}}-1\right](1+o(1))\right\}^{\frac{1}{u}}(1+O(b\log v_{\mathbf{n}}))^{\frac{1}{u}}.$$

Lets us now examine the terms $\widetilde{R}_{1,\mathbf{n},x,y}$ and $\widetilde{R}_{2,\mathbf{n},x,y}$ separately.

- First, we get

$$\begin{aligned}\widetilde{R}_{1,\mathbf{n},x,y} &\leq \hat{\mathbf{n}}^{-2}\|\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\|_s\|\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\|_u\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}_{\Omega_{\mathbf{n}}}}\sum\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}\\
&\leq \hat{\mathbf{n}}^{-2}\pi^{-\frac{1}{t}-1}\left(v_{\mathbf{n}},x\right)\gamma^2(x)b^{-(1+\frac{1}{t})d}\left\|K\right\|_s\left\|K\right\|_u\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^{N}\sum_{\|\mathbf{i}\|\leq\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/t}.\\
&\leq \hat{\mathbf{n}}^{-2}\pi^{-\frac{1}{t}-1}\left(v_{\mathbf{n}},x\right)\gamma^2(x)b^{-(1+\frac{1}{t})d}\left\|K\right\|_s\left\|K\right\|_u\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^{N}\sum_{\|\mathbf{i}\|\leq\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/t}.\\
\hat{\mathbf{n}}b^{d}\pi\left(v_{\mathbf{n}},x\right)\widetilde{R}_{1,\mathbf{n},x,y} &\leq \hat{\mathbf{n}}^{-1}\pi^{-\frac{1}{t}}\left(v_{\mathbf{n}},x\right)\gamma^2(x)b^{-\frac{d}{t}}\left\|K\right\|_s\left\|K\right\|_u\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^{N}\sum_{\mathbf{i}>0}\|\mathbf{i}\|^{-\vartheta/t}.\end{aligned}\tag{5.36}$$

- Second, we have

$$\begin{aligned}\widetilde{R}_{2,\mathbf{n},x,y} &\leq \hat{\mathbf{n}}^{-2}\|\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\|_s\|\widetilde{\eta}_{\mathbf{j},\mathbf{n}}\|_u\sum_{\mathbf{i},\mathbf{j}\in\mathcal{S}^{c}_{\Omega_{\mathbf{n}}}}\sum\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)^{1/t}\\
&\leq \hat{\mathbf{n}}^{-1}\pi^{-\frac{1}{t}-1}\left(v_{\mathbf{n}},x\right)\gamma^2(x)b^{-(1+\frac{1}{t})d}\left\|K\right\|_s\left\|K\right\|_u\widetilde{\varepsilon}_{s,u}(x)\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right)^{1/t}.\\
\hat{\mathbf{n}}b^{d}\pi\left(v_{\mathbf{n}},x\right)\widetilde{R}_{2,\mathbf{n},x,y} &\leq \pi^{-\frac{1}{t}}\left(v_{\mathbf{n}},x\right)\gamma^2(x)b^{-\frac{d}{t}}\left\|K\right\|_s\left\|K\right\|_u\widetilde{\varepsilon}_{s,u}(x)\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{-N/t}\|\mathbf{i}\|^{N/t}\varphi\left(\|\mathbf{i}\|\right)^{1/t}\\
&\leq C\gamma^2(x)\pi^{-\frac{1}{t}}\left(v_{\mathbf{n}},x\right)b^{-\frac{d}{t}}\left\|K\right\|_s\left\|K\right\|_u\widetilde{\varepsilon}_{s,u}(x)\Omega_{\mathbf{n}}^{-N/t}\sum_{\|\mathbf{i}\|>\Omega_{\mathbf{n}}}\|\mathbf{i}\|^{\frac{N-\vartheta}{t}}.\end{aligned}\tag{5.37}$$

By assumptions of Theorem 5.4.1, we have $\widetilde{\varepsilon}_{s,u}(x)\longrightarrow 1$ as $\mathbf{n}\to\infty$. Then combining (5.36) and (5.37), since $\vartheta>2(N+1)$ and $\Omega_{\mathbf{n}}=\left(\hat{\mathbf{n}}b^{d/2}\pi(v_{\mathbf{n}},x)^{1/2}\right)^{1/N}$, it follows that

$$\widetilde{R}_{\mathbf{n},x,y}=O\left(\frac{1}{\hat{\mathbf{n}}b^{d}\pi\left(v_{\mathbf{n}},x\right)}\right).\tag{5.38}$$

The proof of the lemma follows from (5.34), (5.38) and by the fact that

$$\begin{aligned}\hat{\mathbf{n}}\mathrm{Var}\left(\widetilde{S}_{\mathbf{n}}(x,y)\right) &= \hat{\mathbf{n}}b^{d}\pi\left(v_{\mathbf{n}},x\right)\mathrm{Var}\left(\hat{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\widetilde{\Delta}_{\mathbf{i},\mathbf{n}}\right)\\
&= 2\gamma^2(x)\left\|K\right\|_2^2\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^{2}}-1\right](1+o(1))\right\}(1+O(b\log v_{\mathbf{n}}))\\
&\quad - b^{d}\pi\left(v_{\mathbf{n}},x\right)\gamma^2(x)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\}^{2}(1+O(b\log v_{\mathbf{n}}))^{2}\end{aligned}$$

5.6.4 Appendix A4 : Intermediate lemmas

Lemma 5.6.5. *. Suppose* (5.3)*,* ($\mathcal{Q}$) *and* ($\mathcal{F}$) *hold. If* $\upsilon_{\mathbf{n}} \to \infty$ *and* $b \log \upsilon_{\mathbf{n}} \to 0$ *as* $\mathbf{n} \to \infty$*, then*

$$\sup_{\mathrm{d}(x,x') \leq b} \left| \frac{\overline{F}(\upsilon_{\mathbf{n}}, x)}{\overline{F}(\upsilon_{\mathbf{n}}, x')} - 1 \right| = O\left(b \log \upsilon_{\mathbf{n}}\right).$$

Proof of Lemma 5.6.5 We refer the reader to the proof of Lemma 1 in Daouia et al. (2011).■

Lemma 5.6.6. *Suppose* (5.3) *and* (5.6) *and the conditions of Theorem 5.4.1 hold, then for all* $x \in \mathbb{R}^d$ *such that* $g(x) > 0$

a. $\mathbf{E}\left[\pi_{\mathbf{n}}\left(\upsilon_{\mathbf{n}}, x\right)\right] = \pi\left(\upsilon_{\mathbf{n}}, x\right)\left(1 + O(b \log \upsilon_{\mathbf{n}})\right)$ *and*

$$\mathbf{E}\left[\widehat{\phi}_{\upsilon_{\mathbf{n}}}\left(y, x\right)\right] = \gamma(x)\pi\left(\upsilon_{\mathbf{n}}, x\right) \times \left\{1 + \frac{\mathcal{A}(\upsilon_{\mathbf{n}}, x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))} - 1\right](1 + o(1))\right\}(1 + O(b \log \upsilon_{\mathbf{n}})).$$

.

b. $\mathbb{W}_{\mathbf{n},\pi}(x) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \|K\|_2^2\right)$ *and* $\mathbb{W}_{\mathbf{n},\phi}(x, y) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\gamma^2(x)\|K\|_2^2\right)$ *as* $\mathbf{n} \to \infty$*. Moreover as* $\mathbf{n} \to \infty$*,*
$\mathbb{W}_{\mathbf{n}}(x) := \left(\mathbb{W}_{\mathbf{n},\pi}(x); \mathbb{W}_{\mathbf{n},\phi}(x, y)\right)^{\top}$ *converges in distribution to a bivariate Gaussian vector* $\mathcal{N}\left(0, \Sigma\right)$ *; where*

$$\Sigma := \begin{pmatrix} \|K\|_2^2 & \gamma(x)\|K\|_2^2 \\ \gamma(x)\|K\|_2^2 & 2\gamma^2(x)\|K\|_2^2 \end{pmatrix}$$

and $\top$ *denotes the transpose.*

Proof of Lemma 5.6.6
a. It is directly derived from Lemma 2 of Goegebeur et al. (2014b)

b.
▶ First, we prove asymptotic normality of $\mathbb{W}_{\mathbf{n},\phi}(x, y)$. With the notations that are given at the beginning of this section, it follows that

$$\mathbb{W}_{\mathbf{n},\phi}(x, y) = \hat{\mathbf{n}}^{1/2}\widetilde{S}_{\mathbf{n}}(x, y) \quad \text{where} \quad \widetilde{S}_{\mathbf{n}}(x, y) := \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \left[b^d \pi\left(\upsilon_{\mathbf{n}}, x\right)\right]^{1/2} \frac{\widetilde{\Delta}_{\mathbf{i},\mathbf{n}}}{\hat{\mathbf{n}}},$$

$$\widetilde{\Delta}_{\mathbf{i},\mathbf{n}} := \widetilde{\eta}_{\mathbf{i},\mathbf{n}} - \mathbb{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\right] \quad \text{and} \quad \widetilde{\eta}_{\mathbf{i},\mathbf{n}} := \frac{1}{\pi(\upsilon_{\mathbf{n}}, x)} K_b\left(x - X_{\mathbf{i}}\right)\left(\log \frac{Y_{\mathbf{i}}}{\upsilon_{\mathbf{n}}}\right) \mathbb{1}_{\{Y_{\mathbf{i}} > \upsilon_{\mathbf{n}}\}}$$

First, using (5.19), one can find sequences of positive integers $w_{\mathbf{n}} = w$, $q_{\mathbf{n}} = q$ tending to infinity such that $qw^{-1} = o(1)$. Therefore, $q = q_{\mathbf{n}} = o(w)$. For some integers $r_1, \cdots, r_N$, we assume that $n_1 = r_1(q + w), \cdots, n_N = r_N(q + w)$. Define $\widetilde{A}_{\mathbf{i},x} = \left[b^d \pi\left(\upsilon_{\mathbf{n}}, x\right)\right]^{1/2} \widetilde{\Delta}_{\mathbf{i},\mathbf{n}}$, then, the random variables $\widetilde{A}_{\mathbf{i},x}$'s are now set into large blocks and small blocks, that is

$$\begin{aligned} \widetilde{U}_{\mathbf{n}}\left(1, x, \mathbf{j}\right) &= \sum_{\mathbf{i} \in I(1,\mathbf{j},x)} \widetilde{A}_{\mathbf{i},x}, \quad \widetilde{U}_{\mathbf{n}}\left(2, x, \mathbf{j}\right) = \sum_{\mathbf{i} \in I(2,\mathbf{j},x)} \widetilde{A}_{\mathbf{i},x} \\ \widetilde{U}_{\mathbf{n}}\left(3, x, \mathbf{j}\right) &= \sum_{\mathbf{i} \in I(3,\mathbf{j},x)} \widetilde{A}_{\mathbf{i},x}, \quad \widetilde{U}_{\mathbf{n}}\left(4, x, \mathbf{j}\right) = \sum_{\mathbf{i} \in I(4,\mathbf{j},x)} \widetilde{A}_{\mathbf{i},x}. \end{aligned}$$

and so on. The last two terms are

$$\widetilde{U}_{\mathbf{n}}\left(2^{N-1},\mathbf{j},x\right) \;=\sum_{I(2^{N-1},\mathbf{j},x)}\widetilde{A}_{\mathbf{i},x},\;\; \widetilde{U}_{\mathbf{n}}\left(2^{N},\mathbf{j},x\right)=\sum_{I(2^{N},\mathbf{j},x)}\widetilde{A}_{\mathbf{i},x}$$

For each integer $1\leq i\leq 2^N$, define

$$\widetilde{T}_{\mathbf{n}}\left(x,i\right) \;=\sum_{\mathbf{j}\in D(i,x)}\widetilde{U}_{\mathbf{n}}\left(i,\mathbf{j},x\right). \tag{5.39}$$

Clearly $\widetilde{S}_{\mathbf{n}}(y,x)=\sum_{i=1}^{2^N}\widetilde{T}_{\mathbf{n}}\left(x,i\right)$. With this notation, $\widetilde{T}_{\mathbf{n}}\left(x,1\right)$ is the sum of the random variables $\widetilde{A}_{\mathbf{i},x}$ in large blocks. For $2\leq i\leq 2^N$, the $\widetilde{T}_{\mathbf{n}}\left(x,i\right)$ are sums of random variables in small blocks. Denote

$$\widetilde{Q}_{1,\mathbf{n},x}:=\left|\mathbb{E}\left[\exp\left(iz\widetilde{T}_{\mathbf{n}}\left(x,1\right)\right)\right]-\prod_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbb{E}\left[\exp\left(iz\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)\right)\right]\right|, \tag{5.40}$$

$$\widetilde{Q}_{2,\mathbf{n},x}:=\hat{\mathbf{n}}^{-1}\mathbb{E}\left[\left(\sum_{i=2}^{2^N}\widetilde{T}_{\mathbf{n}}\left(x,i\right)\right)^2\right], \tag{5.41}$$

$$\widetilde{Q}_{3,\mathbf{n},x}:=\hat{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbb{E}\left[\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)^2\right], \tag{5.42}$$

$$\widetilde{Q}_{4,\mathbf{n},x}:=\hat{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1}\mathbb{E}\left[\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)^2\mathbb{1}_{\left\{\left|\widetilde{U}_{\mathbf{n}}(1,x,\mathbf{j})\right|\geq\varepsilon\widetilde{\sigma}(x)\hat{\mathbf{n}}^{1/2}\right\}}\right]\quad\text{for every }\varepsilon>0. \tag{5.43}$$

Note that

$$\frac{\widetilde{S}_{\mathbf{n}}(y,x)}{\widetilde{\sigma}(x)\sqrt{\hat{\mathbf{n}}}}=\frac{\widetilde{T}_{\mathbf{n}}\left(x,1\right)}{\widetilde{\sigma}(x)\sqrt{\hat{\mathbf{n}}}}+\sum_{i=2}^{2^N}\frac{\widetilde{T}_{\mathbf{n}}\left(x,i\right)}{\widetilde{\sigma}(x)\sqrt{\hat{\mathbf{n}}}},$$

where in this proof $\widetilde{\sigma}(x)^2=2\gamma(x)^2\left\|K\right\|_2^2$. To complete the proof of this lemma, we need to prove the following two corollaries.

Corollary 5.6.7. *Under conditions of Lemma 5.6.6, we have that as* $\mathbf{n}\to\infty$*,*

$$\lim_{\mathbf{n}\to\infty}\widetilde{Q}_{1,\mathbf{n},x}=0\;\;\textit{and}\;\;\lim_{\mathbf{n}\to\infty}\widetilde{Q}_{2,\mathbf{n},x}=0.$$

This corollary shows that as $\mathbf{n}\to\infty$, the random variables $\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)$ are asymptotically independent and the term $\sum_{i=2}^{2^N}\widetilde{T}_{\mathbf{n}}\left(x,i\right)/\left(\widetilde{\sigma}\hat{\mathbf{n}}^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible.

Corollary 5.6.8. *Under conditions of Lemma 5.6.6, we have that as* $\mathbf{n}\to\infty$*,*

$$\lim_{\mathbf{n}\to\infty}\widetilde{Q}_{3,\mathbf{n},x}=\widetilde{\sigma}(x)^2\;\;\textit{and}\;\;\lim_{\mathbf{n}\to\infty}\widetilde{Q}_{4,\mathbf{n},x}=0.$$

where $\widetilde{\sigma}(x)^2=2\gamma(x)^2\left\|K\right\|_2^2$.

This last result permits to deduce the asymptotic normality of $\widetilde{T}_{\mathbf{n}}(x,1)\,/\,\left(\widetilde{\sigma}(x)\hat{\mathbf{n}}^{1/2}\right)$, since (5.42) goes to zero and the Lindeberg-Feller condition (5.43) $\to 0$ is satisfied.

Proof of Corollary 5.6.7

To prove that $\widetilde{Q}_{1,\mathbf{n},x} \to 0$ as $\mathbf{n} \to \infty$, we first use Lemma 5.6.2 to enumerate the random variables $\widetilde{U}_{\mathbf{n}}(1,x,\mathbf{j})$ in an arbitrary order say, $U_1^*,\cdots,U_M^*$, where $M = \prod_{k=1}^{N} r_k = \hat{\mathbf{n}}(w+q)^{-N} \leq \hat{\mathbf{n}}w^N$. Define the following sets :

$$\mathcal{S}_{\mathbf{n}}(1,x,\mathbf{j}) = \left\{i : j_k(q+w)+1 \leq i_k \leq j_k(q+w)+w,\ k=1,\cdots,N\right\}.$$

Note that the sets of distinct sites are distant from a distance at least q. Remark that the set $\mathcal{S}_{\mathbf{n}}(1,x,\mathbf{j})$ contains w^N sites, and is the set of sites involved with $\widetilde{U}_{\mathbf{n}}(1,x,\mathbf{j})$. We now use Lemma 5.6.2 to get

$$\begin{aligned}\widetilde{Q}_{1,\mathbf{n},x} \leq \sum_{k=1}^{M-1}\sum_{j=k+1}^{M} \Bigg| &\mathbf{E}\left[(\exp\{iuU_k^*\}-1)(\exp\left\{iuU_j^*\right\}-1)\right]\prod_{s=j+1}^{M}\exp\{iuU_s^*\} \\ &- \mathbf{E}\left[(\exp\{iuU_k^*\}-1)\right]\mathbf{E}\left[(\exp\left\{iuU_j^*\right\}-1)\right]\prod_{s=j+1}^{M}\exp\{iuU_s^*\}\Bigg|.\end{aligned}$$

Let $\widetilde{\mathcal{S}}_j$ be the sets of sites involved with U_j^*. Using again Lemma 5.6.2, it follows that

$$\begin{aligned}&\left|\mathbf{E}\left[(\exp\{iuU_k^*\}-1)(\exp\left\{iuU_j^*\right\}-1)\right] - \mathbf{E}\left[(\exp\{iuU_k^*\}-1)\right]\mathbf{E}\left[(\exp\left\{iuU_j^*\right\}-1)\right]\right| \\ &\quad\leq C\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_j,\widetilde{\mathcal{S}}_k)\right)\end{aligned}$$

thus

$$\begin{aligned}\widetilde{Q}_{1,\mathbf{n},x} &\leq Cw^N\sum_{k=1}^{M-1}\sum_{j=k+1}^{M}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_j,\widetilde{\mathcal{S}}_k)\right) \\ &\leq Cw^N M\sum_{k=2}^{M}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)\right) \\ &\leq Cw^N M\sum_{i=1}^{\infty}\sum_{k:iq\leq \mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)<(i+1)q}\varphi\left(\mathrm{d}(\widetilde{\mathcal{S}}_1,\widetilde{\mathcal{S}}_k)\right) \\ &\leq Cw^N M\sum_{i=1}^{\infty} i^{N-1}\varphi(iq) \leq C\hat{\mathbf{n}}\sum_{i=1}^{\infty} i^{N-1}\varphi(iq) \\ &\leq C\hat{\mathbf{n}}\sum_{i=1}^{\infty} i^{N-1-\vartheta}q^{-\vartheta} \leq C\hat{\mathbf{n}}q^{-\vartheta},\end{aligned}$$

this term tends to zero if one takes q as stated in $\mathcal{W}$ and $w = \left[\hat{\mathbf{n}}b^d\right]^{\frac{1}{2N}}$.

From (5.41), we want to show that $\widetilde{Q}_{2,\mathbf{n},x} \to 0$ as $\mathbf{n} \to \infty$. It is enough to show, as $\mathbf{n} \to \infty$, that

$$\hat{\mathbf{n}}^{-1}\mathbf{E}\left[\widetilde{T}_{\mathbf{n}}(x,i)\right]^2 \longrightarrow 0 \quad \text{for each} \quad 2 \leq i \leq 2^N.$$

Without loss of generality, for simplicity, we only consider the case where $i=2$. Again, we enumerate as in the previous proof, the random variables $\widetilde{U}_{\mathbf{n}}(2,x,\mathbf{j})$ in an arbitrary way and refer them as $U_1^+, \cdots, U_M^+$. Then

$$\begin{aligned}\mathbf{E}\left[\widetilde{T}_{\mathbf{n}}(x,2)\right]^2 &= \sum_{i=0}^{M}\mathrm{Var}\left(U_i^+\right) + 2\sum_{j=0}^{M-1}\sum_{i=j+1}^{M}\mathrm{Cov}\left(U_i^+,U_j^+\right) \\ &:= \quad L_{1,x} \quad + \quad L_{2,x}\end{aligned} \tag{5.44}$$

Let us treat first the variance term, we have

$$\begin{aligned}\mathrm{Var}\left(U_i^+\right) &= \mathrm{Var}\left(\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{w}\ \sum_{i_N=1}^{q}\widetilde{A}_{\mathbf{i},x}\right) \\ &:= w^{N-1}q\mathrm{Var}(\widetilde{A}_{\mathbf{i},x}) + \sum_{\substack{j_k=1\\k=1,\cdots,N-1}}^{w}\ \sum_{j_N=1}^{q}\ \sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{w}\ \sum_{\substack{i_N=1\\ \\ i_k\neq j_k \text{ for some } 1\leq k\leq N}}^{q}\mathbf{E}\left[\widetilde{A}_{\mathbf{i},x}\widetilde{A}_{\mathbf{j},x}\right]\end{aligned} \tag{5.45}$$

Recall that $\widetilde{A}_{\mathbf{i},x} = \left[b^d\pi\left(v_{\mathbf{n}},x\right)\right]^{1/2}\widetilde{\Delta}_{\mathbf{i},\mathbf{n}}$, $\widetilde{\Delta}_{\mathbf{i},\mathbf{n}} = \widetilde{\eta}_{\mathbf{i},\mathbf{n}} - \mathbb{E}\left[\widetilde{\eta}_{\mathbf{i},\mathbf{n}}\right]$,
$\widetilde{\eta}_{\mathbf{i},\mathbf{n}} := \frac{1}{\pi(v_{\mathbf{n}},x)}K_b\Big(x - X_{\mathbf{i}}\Big)\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}$, and combining with (5.34), it follows that

$$\begin{aligned}\mathrm{Var}(\widetilde{A}_{\mathbf{i},x}) &\leq 2\gamma^2(x)\left\|K\right\|_2^2\left(1+O(b\log v_{\mathbf{n}})\right)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\} \\ &\leq C2\gamma^2(x)\left\|K\right\|_2^2\widetilde{\varepsilon}_1(x) \leq C.\end{aligned} \tag{5.46}$$

where $\widetilde{\varepsilon}_1(x) := \left(1+O(b\log v_{\mathbf{n}})\right)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}$.

Now, for the second term on the right hand of (5.45), let $\chi = 2(1-\varsigma)/\varsigma,\ 0 < \varsigma < 1$, by Lemma 5.6.1 second statement together with Lemma 2 of Goegebeur et al. (2014b), we have that

$$\begin{aligned}\mathbf{E}\left|\widetilde{A}_{\mathbf{i},x}\widetilde{A}_{\mathbf{j},x}\right| &\leq C\left(\int_{\mathbb{R}^d\times\mathbb{R}}\left|\left[b^d\pi^{-1}\left(v_{\mathbf{n}},x\right)\right]^{1/2}K_b(x-u)\log\left(\frac{t}{v_{\mathbf{n}}}\right)_+\right|^{2+\chi}f(u)g(t)dudt\right)^{\varsigma}\varphi(\|\mathbf{i}-\mathbf{j}\|)^{1-\varsigma} \\ &\leq 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\varphi(\|\mathbf{i}-\mathbf{j}\|)^{1-\varsigma},\end{aligned} \tag{5.47}$$

Combining (5.45), (5.46) and (5.47), it follows that

$$\begin{aligned}\mathrm{Var}\left(U_i^+\right) &\leq Cw^{N-1}q\left(1+4b^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{w}\ \sum_{i_N=1}^{q}\varphi(\|\mathbf{i}\|)^{1-\varsigma}\right) \\ &\leq 4Cw^{N-1}qb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\sum_{\substack{i_k=1\\k=1,\cdots,N-1}}^{w}\ \sum_{i_N=1}^{q}\varphi(\|\mathbf{i}\|) \\ &\leq 4Cw^{N-1}qb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}.\end{aligned} \tag{5.48}$$

By (5.45)-(5.48), it follows that :

$$L_{1,x} \leq 4CMw^{N-1}qb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}. \tag{5.49}$$

Let

$$\begin{aligned}\widetilde{\mathcal{S}}_{\mathbf{n}}(2,x,\mathbf{j}) \quad = \quad & \Big\{\mathbf{i} : j_k(w+q+1) \leq i_k \leq j_k(w+q)+w,\ 1 \leq k < N \\ & j_k(w+q)+w+1 \leq i_N \leq (j_N+1)(w+q)\Big\}.\end{aligned}$$

Clearly, $\widetilde{U}_{\mathbf{n}}(2,x,\mathbf{j})$ is the sum of $\widetilde{A}_{\mathbf{i},x}$ with sites in $\widetilde{\mathcal{S}}_{\mathbf{n}}(2,x,\mathbf{j})$. Since $w > q$, if $\mathbf{j}$ and $\mathbf{j}'$ belong to two distinct sets $\widetilde{\mathcal{S}}_{\mathbf{n}}(2,x,\mathbf{j})$ and $\widetilde{\mathcal{S}}_{\mathbf{n}}(2,x,\mathbf{j}')$, then $j_k \neq j_k^{'}$ for some $1 \leq k \leq N$ and $\|\mathbf{j}-\mathbf{j}'\| > q$. With (5.47), we obtain

$$\begin{aligned}L_{2,x} &\leq C \sum_{\substack{j_k=1\\k=1,\cdots,N}}^{n_k} \sum_{\substack{i_k=1\\k=1,\cdots,N\\ \|\mathbf{i}-\mathbf{j}\|>q}}^{n_k} \mathbf{E}\left[\widetilde{A}_{\mathbf{i},x}\widetilde{A}_{\mathbf{j},x}\right] \\ &\leq 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\hat{\mathbf{n}} \sum_{\substack{i_k=1\\k=1,\cdots,N;\ \|\mathbf{i}\|>q}}^{n_k} \varphi(\|\mathbf{i}\|)^{1-\varsigma} \\ &\leq 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\hat{\mathbf{n}}\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}. \end{aligned} \tag{5.50}$$

From (5.39), (5.49) and (5.50), we get
$\hat{\mathbf{n}}^{-1}\mathbf{E}\left[\widetilde{T}_{\mathbf{n}}(x,2)\right]^2$

$$\begin{aligned}&\leq 4CMw^{N-1}q\hat{\mathbf{n}}^{-1}b^{-d}\pi\left(v_{\mathbf{n}},x\right)\gamma^4(x)\left\|K\right\|_2^4\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma} \\ &+ 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma} \\ &\leq 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\left(Mw^{N-1}q\hat{\mathbf{n}}^{-1}\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}+\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}\right) \\ &\leq 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\left(q/w\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}+\sum_{i=q}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma}\right) \\ &\leq 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\left(q/w+1\right)\sum_{i=1}^{\infty}i^{N-1}\varphi(i)^{1-\varsigma} \\ &\leq 4Cb^{-d(1-\varsigma)}\pi\left(v_{\mathbf{n}},x\right)^{-(1-\varsigma)}\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\left(q/w+1\right)\sum_{i=1}^{\infty}i^{N-1-\vartheta(1-\varsigma)}. \end{aligned} \tag{5.51}$$

This tends to zero it suffices to take $0 < \varsigma < \frac{\vartheta-N}{\vartheta}$ with the previous choices of q, w and $(\mathcal{W})$. This complete the proof of Corollary 5.6.7.■

Proof of Corollary 5.6.8
To prove that $\widetilde{Q}_{3,\mathbf{n},x} \to \widetilde{\sigma}(x)^2$ as $\mathbf{n} \to \infty$, let us denote by

$$K_{\mathbf{n},x,1} = \widetilde{T}_{\mathbf{n}}(x,1) \quad \text{and} \quad K_{\mathbf{n},x,2} = \sum_{i=2}^{2^N}\widetilde{T}_{\mathbf{n}}(x,i). \tag{5.52}$$

The sum of the random variables $\widetilde{A}_{\mathbf{1},x}$ in large blocks is assigned to $K_{\mathbf{n},x,1}$ and $K_{\mathbf{n},x,2}$ is the sum of random variables $\widetilde{A}_{\mathbf{i},x}$ in small blocks. Lemma 5.6.4 statement 2 implies that $\hat{\mathbf{n}}^{-1}\mathbf{E}\left[\widetilde{S}_{\mathbf{n}}(x,y)^2\right] \longrightarrow \widetilde{\sigma}(x)^2$. This combining with (5.41) shows that $\hat{\mathbf{n}}^{-1}\mathbf{E}\left[K^2_{\mathbf{n},x,1}\right] \longrightarrow \widetilde{\sigma}(x)^2$. Now

$$\hat{\mathbf{n}}^{-1}\mathbf{E}\left[K^2_{\mathbf{n},x,1}\right] = \hat{\mathbf{n}}^{-1}\sum_{\substack{j_k=0\\ k=1,\cdots,N}}^{r_k-1} \mathbf{E}\left[U_{\mathbf{n}}\left(1,x,\mathbf{j}\right)\right]^2 \;+\; \hat{\mathbf{n}}^{-1} \sum_{\substack{j_k=1\\ k=1,\cdots,N\\ i_k\neq j_k \text{ for some } 1\leq k\leq N}}^{r_k-1} \sum_{\substack{i_k=1\\ k=1,\cdots,N}}^{r_k-1} \operatorname{Cov}\left(U_{\mathbf{n}}\left(1,x,\mathbf{j}\right), U_{\mathbf{n}}\left(1,x,\mathbf{i}\right)\right). \tag{5.53}$$

To prove (5.42), it suffices to show that the second term in the right hand side of (5.53) tends to zero as $\mathbf{n}\to\infty$. For this end, we use the same arguments used to obtain a bound for $L_{2,x}$ defined in (5.50), to have that the last term of (5.53) is bounded by

$$\begin{aligned} &4Cb^{-d}\pi\left(v_{\mathbf{n}},x\right)\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x) \sum_{\substack{i_k=1\\ k=1,\cdots,N;\ \|\mathbf{i}\|>q}}^{n_k} \varphi(\|\mathbf{i}\|)^{1-\varsigma} \\ &\qquad\leq 4Cb^{-d}\pi\left(v_{\mathbf{n}},x\right)\gamma^4(x)\left\|K\right\|_2^4\widetilde{\varepsilon}_1^2(x)\sum_{i=q}^{\infty} i^{N-1}\varphi(i)^{1-\varsigma}. \end{aligned} \tag{5.54}$$

This last tends to zero as above.

To complete the proof, let us show that from (5.43), $\widetilde{Q}_{4,\mathbf{n},x}\to 0$ as $\mathbf{n}\to\infty$ for every $\varepsilon > 0$. First, we have that

$$\mathbf{E}\left[\widetilde{A}^2_{\mathbf{i},x}\right] \leq 2C\gamma^2(x)\left\|K\right\|_2^2\widetilde{\varepsilon}_1(x) \leq C \;\text{ and }\; \mathbf{E}\left[\widetilde{A}_{\mathbf{i},x}\right] \leq \left[b^d\pi\left(v_{\mathbf{n}},x\right)\right]^{1/2}\gamma(x)\left\|K\right\|_2^2\widetilde{\varepsilon}_1(x).$$

Note that by Markov inequality we have

$$\begin{aligned} \mathbb{P}\left[\left|\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)\right| > w^N\left(b^{-d}\pi\left(v_{\mathbf{n}},x\right)^{-1}\right)^{1/2}\right] &\leq \frac{w^N\left(b^d\pi\left(v_{\mathbf{n}},x\right)\right)^{1/2}}{w^N\left(b^{-d}\pi\left(v_{\mathbf{n}},x\right)^{-1}\right)^{1/2}}\gamma(x)\left\|K\right\|_2^2 \\ &= \gamma(x)\left\|K\right\|_2^2 b^d\pi\left(v_{\mathbf{n}},x\right) \to 0. \end{aligned}$$

Therefore for $\mathbf{n}$ large enough, $\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right) < w^N\left(b^{-d}\pi\left(v_{\mathbf{n}},x\right)^{-1}\right)^{1/2}$, and thus

$$\mathbf{E}\left[\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)^2\mathbb{1}_{\left\{\left|U_{\mathbf{n}}(1,x,\mathbf{j})\right|>\varepsilon\overline{\sigma}(x)\hat{\mathbf{n}}^{1/2}\right\}}\right] \leq w^{2N}\left(b^{-d}\pi\left(v_{\mathbf{n}},x\right)^{-1}\right)\mathbb{P}\left[\left|\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)\right| > \varepsilon\overline{\sigma}(x)\hat{\mathbf{n}}^{1/2}\right].$$

By (5.41), we have $E\left[\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)^2\right] \sim \frac{\widetilde{\sigma}(x)^2}{M}$ for $\mathbf{n}$ large enough, then using Tchebychev inequality

$$\mathbb{P}\left[\left|\widetilde{U}_{\mathbf{n}}\left(1,x,\mathbf{j}\right)\right| > \varepsilon\widetilde{\sigma}(x)\hat{\mathbf{n}}^{1/2}\right] \leq \frac{\frac{\widetilde{\sigma}(x)^2}{M}}{\widetilde{\sigma}(x)^2\hat{\mathbf{n}}\varepsilon^2} = \left(M\hat{\mathbf{n}}\right)^{-1}\varepsilon^{-2}.$$

Therefore

$$\begin{aligned} \widetilde{Q}_{4,\mathbf{n},x} &\leq Cw^{2N}\left(b^{-d}\pi\left(v_{\mathbf{n}},x\right)^{-1}\right)\left(M\hat{\mathbf{n}}\right)^{-1}\varepsilon^{-2}\sum_{\substack{j_k=0\\ k=1,\cdots,N}}^{r_k-1}\mathbb{P}\left[U_{\mathbf{n}}\left(1,x,\mathbf{j}\right) > \varepsilon\overline{\sigma}(x)\hat{k}_{\mathbf{n}}^{1/2}\right] \\ &\leq C\frac{w^{2N}}{\hat{\mathbf{n}}b^d\pi\left(v_{\mathbf{n}},x\right)\varepsilon^2}. \end{aligned}$$

This tends to zero by using the choice of q in $\mathcal{W}$ and $w = \left[\hat{\mathbf{n}} b^d\right]^{1/(2N)}$. This complete the proof of Corollary 5.6.8.■

Now, to complete the proof of asymptotic normality of $\mathbb{W}_{\mathbf{n},\phi}(x,y)$, as $\mathbf{n} \to \infty$, in one hand from Corollary 5.6.7, $\widetilde{Q}_{1,\mathbf{n},x}$ shows that the random variable's $\widetilde{U}_{\mathbf{n}}(1,x,\mathbf{j})$ are asymptotically independent and $\widetilde{Q}_{2,\mathbf{n},x}$ shows that the random variable's $\sum_{i=2}^{2^N} \widetilde{T}_{\mathbf{n}}(x,i) \,/\, \left(\widetilde{\sigma}\hat{\mathbf{n}}^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible. In other hand, from Corollary 5.6.8, as $\mathbf{n} \to \infty$, $\widetilde{Q}_{3,\mathbf{n},x}$ shows that the random variable's $\widetilde{T}_{\mathbf{n}}(x,1) \,/\, \left(\widetilde{\sigma}\hat{\mathbf{n}}^{1/2}\right)$ which is in large blocks is asymptotically normal with variance equal to 1 and $\widetilde{Q}_{4,\mathbf{n},x}$ shows that the random variables in large blocks verify the Lindeberg-Feller condition. Then, with these arguments, as $\mathbf{n} \to \infty$, we have

$$\mathbb{W}_{\mathbf{n},\phi}(x,y) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \widetilde{\sigma}(x)^2\right).$$

► Now we are going to give the asymptotic normality of the quantity $\mathbb{W}_{\mathbf{n},\pi}(x)$. The variance of $\mathbb{W}_{\mathbf{n},\pi}(x)$ is obtained in the same manner as the variance of $\mathbb{W}_{\mathbf{n},\phi}(x,y)$, it suffices to note that $\mathbb{W}_{\mathbf{n},\pi}(x)$ has the same writing as $\mathbb{W}_{\mathbf{n},\phi}(x,y)$. Recall the notations given at the beginning of this section

$$\mathbb{W}_{\mathbf{n},\pi}(x) = \hat{\mathbf{n}}^{1/2} S_{\mathbf{n}}(x) \quad \text{where} \quad S_{\mathbf{n}}(x) := \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \left[b^d \pi\left(v_{\mathbf{n}}, x\right)\right]^{1/2} \frac{\Delta_{\mathbf{i},\mathbf{n}}}{\hat{\mathbf{n}}},$$
$$\Delta_{\mathbf{i},\mathbf{n}} := \eta_{\mathbf{i},\mathbf{n}} - \mathbb{E}\left[\eta_{\mathbf{i},\mathbf{n}}\right] \quad \text{and} \quad \eta_{\mathbf{i},\mathbf{n}} := \frac{1}{\pi(v_{\mathbf{n}},x)} K_b\Big(x - X_{\mathbf{j}}\Big) \mathbb{1}_{\{Y_{\mathbf{i}} > v_{\mathbf{n}}\}}.$$

First, using again (5.19), one can find sequences of positive integers $w_{\mathbf{n}} = w$, $q_{\mathbf{n}} = q$ tending to infinity such that $qw^{-1} = o(1)$. Therefore, $q = q_{\mathbf{n}} = o(w)$. For some integers $r_1, \cdots, r_N$, we assume that $n_1 = r_1(q+w), \cdots, n_N = r_N(q+w)$. Define $A_{\mathbf{i},x} = \left[b^d \pi\left(v_{\mathbf{n}}, x\right)\right]^{1/2} \Delta_{\mathbf{i},\mathbf{n}}$, then, the random variables $A_{\mathbf{i},x}$'s are now set into large blocks and small blocks, that is

$$\begin{aligned} U_{\mathbf{n}}(1,x,\mathbf{j}) &= \sum_{\mathbf{i}\in I(1,\mathbf{j},x)} A_{\mathbf{i},x}, \; U_{\mathbf{n}}(2,x,\mathbf{j}) = \sum_{\mathbf{i}\in I(2,\mathbf{j},x)} A_{\mathbf{i},x} \\ U_{\mathbf{n}}(3,x,\mathbf{j}) &= \sum_{\mathbf{i}\in I(3,\mathbf{j},x)} A_{\mathbf{i},x}, \; U_{\mathbf{n}}(4,x,\mathbf{j}) = \sum_{\mathbf{i}\in I(4,\mathbf{j},x)} A_{\mathbf{i},x} \end{aligned}$$

and so on. The last two terms are

$$U_{\mathbf{n}}\left(2^{N-1},x,\mathbf{j}\right) = \sum_{\mathbf{i}\in I(2^{N-1},\mathbf{j},x)} A_{\mathbf{i},x}, \; U_{\mathbf{n}}\left(2^{N},x,\mathbf{j}\right) = \sum_{\mathbf{i}\in I(2^{N},\mathbf{j},x)} A_{\mathbf{i},x}$$

For each integer $1 \le i \le 2^N$, define

$$T_{\mathbf{n}}(x,i) = \sum_{\mathbf{j}\in D(i,x)} U_{\mathbf{n}}(i,x,\mathbf{j}). \tag{5.55}$$

With this notation, $T_{\mathbf{n}}(x,1)$ is the sum of the random variables $A_{\mathbf{i},x}$ in large blocks. For $2 \le i \le 2^N$, the $T_{\mathbf{n}}(x,i)$ are sums of random variables in small blocks. To achieve the proof of asymptotic normality of $\mathbb{W}_{\mathbf{n},\pi}(x)$, we first denote

$$Q_{1,\mathbf{n},x} := \left| \mathbb{E}\left[\exp\left(izT_{\mathbf{n}}\left(x,1\right)\right)\right] - \prod_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[\exp\left(izU_{\mathbf{n}}\left(1,x,\mathbf{j}\right)\right)\right] \right|, \tag{5.56}$$

$$Q_{2,\mathbf{n},x} := \hat{\mathbf{n}}^{-1}\mathbb{E}\left[\left(\sum_{i=2}^{2^N} T_{\mathbf{n}}\left(x,i\right)\right)^2\right], \tag{5.57}$$

$$Q_{3,\mathbf{n},x} := \hat{\mathbf{n}}^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[U_{\mathbf{n}}\left(1,x,\mathbf{j}\right)^2\right], \tag{5.58}$$

$$Q_{4,\mathbf{n},x} := \hat{\mathbf{n}}^{-1} \sum_{\substack{j_k=0\\k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[U_{\mathbf{n}}\left(1,x,\mathbf{j}\right)^2 \mathbb{1}_{\left\{\left|U_{\mathbf{n}}(1,x,\mathbf{j})\right| \geq \varepsilon\overline{\sigma}(x)\hat{\mathbf{n}}^{1/2}\right\}}\right], \quad \varepsilon > 0. \tag{5.59}$$

Note that

$$\frac{S_{\mathbf{n}}(x)}{\overline{\sigma}\sqrt{\hat{\mathbf{n}}}} = \frac{T_{\mathbf{n}}\left(x,1\right)}{\overline{\sigma}\sqrt{\hat{\mathbf{n}}}} + \sum_{i=2}^{2^N} \frac{T_{\mathbf{n}}\left(x,i\right)}{\overline{\sigma}(x)\sqrt{\hat{\mathbf{n}}}},$$

wher $\overline{\sigma}(x)^2 = \|K\|_2^2$. From now, the asymptotic normality of $\mathbb{W}_{\mathbf{n},\pi}(x)$ follows the same lines as the proof of the asymptotic normality of $\mathbb{W}_{\mathbf{n},\phi}(x,y)$. A straightforward application of Corollaries 5.6.7 and 5.6.8 where we replace $\left(\log \frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right) \mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}$ by 1, permits to have that as $\mathbf{n} \to \infty$, equation (5.56) ensures that the random variables $U_{\mathbf{n}}\left(1,x,\mathbf{j}\right)$ are asymptotically independent. The asymptotic normality of $T_{\mathbf{n}}\left(x,1\right)/\left(\overline{\sigma}\hat{\mathbf{n}}^{1/2}\right)$ follows from (5.58) and the Lindeberg-Feller condition (5.59). Equation (5.57) shows that the term $\sum_{i=2}^{2^N} T_{\mathbf{n}}\left(x,i\right)/\left(\overline{\sigma}\hat{\mathbf{n}}^{1/2}\right)$ which is the sum of random variables in small blocks is asymptotically negligible. In summary, we have that

$$\begin{aligned} &\lim_{\mathbf{n}\to\infty} Q_{1,\mathbf{n},x} = 0, \quad \lim_{\mathbf{n}\to\infty} Q_{2,\mathbf{n},x} = 0, \\ &\lim_{\mathbf{n}\to\infty} Q_{3,\mathbf{n},x} = \overline{\sigma}(x)^2 \quad \text{and} \quad \lim_{\mathbf{n}\to\infty} Q_{4,\mathbf{n},x} = 0. \end{aligned}$$

Then

$$\mathbb{W}_{\mathbf{n},\pi}(x) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\overline{\sigma}(x)^2\right).$$

completes the proof.

► We now prove that $\mathbb{W}_{\mathbf{n}}\left(x,y\right) := \left(\mathbb{W}_{\mathbf{n},\pi}\left(x\right);\mathbb{W}_{\mathbf{n},\phi}\left(x,y\right)\right)^\top$ converges in distribution to $\mathcal{N}\left(0,\Sigma\right)$. According to Cramèr-Wold's device (see Van der Vaart (1998)), it is sufficient to prove that $l^\top\mathbb{W}_{\mathbf{n}}\left(x,y\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,l^\top\Sigma l\right)$ for all $l = (l_1,l_2)^\top \in \mathbb{R}^2,\ l \neq 0$.

Some simple calculation yields :

$$
\begin{aligned}
l^{\top}\mathbb{W}_{\mathbf{n}}(x,y) &:= \frac{1}{\Psi_{\mathbf{n}}}\frac{1}{\hat{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}} \quad \text{where} \quad \Psi_{\mathbf{n}} := \left[\hat{\mathbf{n}}b^{d}\pi(v_{\mathbf{n}},x)\right]^{-1/2},\\
\Delta^{\dagger}_{\mathbf{i},\mathbf{n}} &:= \frac{1}{\pi(v_{\mathbf{n}},x)}K_b(x-X_{\mathbf{i}})\left\{l_1\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}} + l_2\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right\}\\
&\quad - \mathbf{E}\left[\frac{1}{\pi(v_{\mathbf{n}},x)}K_b(x-X_{\mathbf{i}})\left\{l_1\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}} + l_2\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right\}\right]\\
&:= A_{\mathbf{i},x,y} - \mathbf{E}\left[A_{\mathbf{i},x,y}\right].
\end{aligned}
$$

Thus, to establish the asymptotic normality of $\mathbb{W}_{\mathbf{n}}(x,y)$, we do as above by verifying the Lindeberg-Feller condition for triangular arrays. We first calculate the variance of $\mathbb{W}_{\mathbf{n}}(x,y)$. We have

$$
\operatorname{Var}\left(\hat{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\right) = \hat{\mathbf{n}}^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\operatorname{Var}\left(\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\right) + 2\hat{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\sum_{\mathbf{i}\neq\mathbf{j}}\operatorname{Cov}\left(\Delta^{\dagger}_{\mathbf{i},\mathbf{n}},\Delta^{\dagger}_{\mathbf{j},\mathbf{n}}\right).
$$

Denote by $I^{\dagger}_{\mathbf{n}}(x) := \hat{\mathbf{n}}^{-2}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\operatorname{Var}\left(A_{\mathbf{i},x,y}\right)$ and by $R^{\dagger}_{\mathbf{n}}(x) := \hat{\mathbf{n}}^{-2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\sum_{\mathbf{i}\neq\mathbf{j}}\operatorname{Cov}\left(A_{\mathbf{i},x,y},A_{\mathbf{j},x,y}\right)$. Clearly, $\operatorname{Var}\left(\hat{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\right) \leq I^{\dagger}_{\mathbf{n}}(x) + R^{\dagger}_{\mathbf{n}}(x)$, with

$$
\begin{aligned}
I^{\dagger}_{\mathbf{n}}(x) &= \hat{\mathbf{n}}^{-1}\left(\mathbf{E}\left[A^2_{\mathbf{i},x,y}\right] - \mathbf{E}\left[A_{\mathbf{i},x,y}\right]^2\right)\\
&= \hat{\mathbf{n}}^{-1}\mathbf{E}\left[K_b^2(x-X_{\mathbf{i}})\left\{l_1\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}} + l_2\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right\}^2\right]\\
&\quad - \hat{\mathbf{n}}^{-1}\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\left\{l_1\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}} + l_2\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right\}\right]^2.
\end{aligned}
$$

From Lemma 2 of Goegebeur et al. (2014b), we have that

$$
\begin{aligned}
\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] &= \gamma(x)\pi(v_{\mathbf{n}},x)\left(1+O(b\log v_{\mathbf{n}})\right)\\
&\quad\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\},\\
\mathbf{E}\left[K_b^2(x-X_{\mathbf{i}})\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] &= \gamma(x)\pi(v_{\mathbf{n}},x)b^{-d}\left\|K\right\|_2^2\left(1+O(b\log v_{\mathbf{n}})\right)\\
&\quad\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\},\\
\text{and}\quad \mathbf{E}\left[K_b^2(x-X_{\mathbf{i}})\left(\log\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)^2\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right] &= 2\gamma^2(x)\pi\left(v_{\mathbf{n}},x\right)b^{-d}\left\|K\right\|_2^2\left(1+O(b\log v_{\mathbf{n}})\right)\\
&\quad\times\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\}.
\end{aligned}
$$

Then we have

$$\begin{aligned}I^{\dagger}_{\mathbf{n}}(x) &= l_1^2\hat{\mathbf{n}}^{-1}\pi(v_{\mathbf{n}},x)^{-2}\left\{\mathbf{E}\left[K_b^2\left(x-X_{\mathbf{i}}\right)\mathbb{1}_{Y>v_{\mathbf{n}}}\right]-\mathbf{E}\left[K_b\left(x-X_{\mathbf{i}}\right)\mathbb{1}_{Y>v_{\mathbf{n}}}\right]^2\right\}\\
&\quad + 2l_1l_2\hat{\mathbf{n}}^{-1}\pi(v_{\mathbf{n}},x)^{-2}\left\{\mathbf{E}\left[K_b^2(x-X_{\mathbf{i}})\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]-\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]\right\}\\
&\quad + l_2^2\hat{\mathbf{n}}^{-1}\pi(v_{\mathbf{n}},x)^{-2}\left\{\mathbf{E}\left[K_b^2(x-X_{\mathbf{i}})\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)^2\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]-\mathbf{E}\left[K_b(x-X_{\mathbf{i}})\log\left(\frac{Y_{\mathbf{i}}}{v_{\mathbf{n}}}\right)\mathbb{1}_{\{Y_{\mathbf{i}}>v_{\mathbf{n}}\}}\right]^2\right\}\\
&=: I^{\dagger}_{\mathbf{n},1}(x)+I^{\dagger}_{\mathbf{n},2}(x)+I^{\dagger}_{\mathbf{n},3}(x). \qquad (5.60)\end{aligned}$$

Using this lengthy but simple calculation, we get

$$I^{\dagger}_{\mathbf{n},1}(x) = l_1^2\hat{\mathbf{n}}^{-1}\pi(v_{\mathbf{n}},x)^{-1}b^{-d}\left(1+O(b\log v_{\mathbf{n}})\right)\left\{\|K\|_2^2-b^d\pi(v_{\mathbf{n}},x)\left(1+O(b\log v_{\mathbf{n}})\right)\right\}, \qquad (5.61)$$

It follows that,

$$\begin{aligned}I^{\dagger}_{\mathbf{n},2}(x) &= 2l_1l_2\hat{\mathbf{n}}^{-1}b^{-d}\gamma(x)\pi(v_{\mathbf{n}},x)^{-1}\left(1+O(b\log v_{\mathbf{n}})\right)\\
&\quad\times\left(\|K\|_2^2-b^d\right)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\}, \qquad (5.62)\end{aligned}$$

thus

$$\begin{aligned}I^{\dagger}_{\mathbf{n},3}(x) &= l_2^2\hat{\mathbf{n}}^{-1}\gamma^2(x)\pi(v_{\mathbf{n}},x)^{-1}b^{-d}\left(1+O\left(b\log v_{\mathbf{n}}\right)\right)\left(I^{\dagger}_{\mathbf{n},4}(x)+I^{\dagger}_{\mathbf{n},5}(x)\right), \qquad (5.63)\\
I^{\dagger}_{\mathbf{n},4}(x) &:= 2\|K\|_2^2\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))^2}-1\right](1+o(1))\right\},\\
I^{\dagger}_{\mathbf{n},5}(x) &:= \pi(v_{\mathbf{n}},x)b^d\left(1+O\left(b\log v_{\mathbf{n}}\right)\right)\left\{1+\frac{\mathcal{A}(v_{\mathbf{n}},x)}{\gamma(x)\rho(x)}\left[\frac{1}{(1-\rho(x))}-1\right](1+o(1))\right\}^2.\end{aligned}$$

Now let us examine $R^{\dagger}_{\mathbf{n}}(x)$, simple calculations yield

$$\begin{aligned}\mathrm{Cov}\left(A_{\mathbf{i},x,y},A_{\mathbf{j},x,y}\right) &= l_1^2\pi^{-2}(v_{\mathbf{n}},x)\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)+l_2^2\pi^{-2}(v_{\mathbf{n}},x)\mathrm{Cov}\left(\tilde{\eta}_{\mathbf{i},\mathbf{n}},\tilde{\eta}_{\mathbf{j},\mathbf{n}}\right)\\
&\quad + l_1l_2\pi^{-2}(v_{\mathbf{n}},x)\left\{\mathrm{Cov}\left(\eta_{\mathbf{i},\mathbf{n}},\tilde{\eta}_{\mathbf{j},\mathbf{n}}\right)+\mathrm{Cov}\left(\tilde{\eta}_{\mathbf{i},\mathbf{n}},\eta_{\mathbf{j},\mathbf{n}}\right)\right\}.\end{aligned}$$

Using the same method as in Lemma 5.6.4, it follows

$$R^{\dagger}_{\mathbf{n}}(x) = O\left(\frac{1}{\hat{\mathbf{n}}b^d\pi(v_{\mathbf{n}},x)}\right). \qquad (5.64)$$

Then combining (5.61), (5.62), (5.63) and (5.64), we get

$$\begin{aligned}\lim_{\mathbf{n}\to+\infty}\mathrm{Var}\left(l^{\top}\mathbb{W}_{\mathbf{n}}(x,y)\right) &= \lim_{\mathbf{n}\to+\infty}\hat{\mathbf{n}}b^d\pi(v_{\mathbf{n}},x)\mathrm{Var}\left(\hat{\mathbf{n}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}\right)\\
&= \lim_{\mathbf{n}\to+\infty}\frac{\hat{\mathbf{n}}b^d}{\pi(v_{\mathbf{n}},x)}I^{\dagger}_{\mathbf{n}}(x) = l_1^2\|K\|_2^2+2l_1l_2\gamma(x)\|K\|_2^2+2l_2^2\gamma^2(x)\|K\|_2^2\\
&= l^{\top}\Sigma l.\end{aligned}$$

To complete the proof of the asymptotic normality of $l^{\top}\mathbb{W}_{\mathbf{n}}(x,y)$, let us introduce some notation. Denote

$$S^{\dagger}_{\mathbf{n}}(x,y) := \sum_{i\in\mathcal{I}_{\mathbf{n}}}\left[b^d\pi(v_{\mathbf{n}},x)\right]^{1/2}\frac{\Delta^{\dagger}_{\mathbf{i},\mathbf{n}}}{\hat{\mathbf{n}}} \quad \text{then} \quad l^{\top}\mathbb{W}_{\mathbf{n}}(x,y) = \hat{\mathbf{n}}^{1/2}S^{\dagger}_{\mathbf{n}}(x,y).$$

Again using (5.19), one can find sequences of positive integers $w_{\mathbf{n}} = w$, $q_{\mathbf{n}} = q$ tending to infinity such that $qw^{-1} = o(1)$. Therefore, $q = q_{\mathbf{n}} = o(w)$. For some integers $r_1, \cdots, r_N$, we assume that $n_1 = r_1(q+w), \cdots, n_N = r_N(q+w)$. With $A^{\dagger}_{\mathbf{i},x} := \sum_{i \in \mathcal{I}_{\mathbf{n}}} \left[b^d \pi(v_{\mathbf{n}}, x)\right]^{1/2} \Delta^{\dagger}_{\mathbf{i},\mathbf{n}}$. Then the random variables $A^{\dagger}_{\mathbf{i},x}$'s are set into large blocks and small blocks, that is

$$\begin{aligned} U^{\dagger}_{\mathbf{n}}(1, x, \mathbf{j}) &= \sum_{\mathbf{i} \in I(1,\mathbf{j},x)} A^{\dagger}_{\mathbf{i},x}, \quad U^{\dagger}_{\mathbf{n}}(2, x, \mathbf{j}) = \sum_{\mathbf{i} \in I(2,\mathbf{j},x)} A^{\dagger}_{\mathbf{i},x} \\ U^{\dagger}_{\mathbf{n}}(3, x, \mathbf{j}) &= \sum_{\mathbf{i} \in I(3,\mathbf{j},x)} A^{\dagger}_{\mathbf{i},x}, \quad U^{\dagger}_{\mathbf{n}}(4, x, \mathbf{j}) = \sum_{\mathbf{i} \in I(4,\mathbf{j},x)} A^{\dagger}_{\mathbf{i},x}, \end{aligned}$$

and so on. The last two terms are

$$U^{\dagger}_{\mathbf{n}}\left(2^{N-1}, x, \mathbf{j}\right) = \sum_{\mathbf{i} \in I(2^{N-1},\mathbf{j},x)} A^{\dagger}_{\mathbf{i},x} \quad U^{\dagger}_{\mathbf{n}}\left(2^{N}, x, \mathbf{j}\right) = \sum_{\mathbf{i} \in I(2^{N},\mathbf{j},x)} A_{\mathbf{i},x} \ A^{\dagger}_{\mathbf{i},x}.$$

For each integer $1 \leq i \leq 2^N$, define

$$T^{\dagger}_{\mathbf{n}}(x, i) = \sum_{\mathbf{j} \in D(i,x)} U^{\dagger}_{\mathbf{n}}(i, x, \mathbf{j}).$$

Using Lemma 5.6.4, as previously let us denote

$$Q^{\dagger}_{1,\mathbf{n},x} := \left| \mathbb{E}\left[\exp\left(izT^{\dagger}_{\mathbf{n}}(x,1)\right)\right] - \prod_{\substack{j_k=0 \\ k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[\exp\left(izU^{\dagger}_{\mathbf{n}}(1,x,\mathbf{j})\right)\right] \right|, \tag{5.65}$$

$$Q^{\dagger}_{2,\mathbf{n},x} := \hat{\mathbf{n}}^{-1} \mathbb{E}\left[\left(\sum_{i=2}^{2^N} T^{\dagger}_{\mathbf{n}}(x,i)\right)^2\right], \tag{5.66}$$

$$Q^{\dagger}_{3,\mathbf{n},x} := \hat{\mathbf{n}}^{-1} \sum_{\substack{j_k=0 \\ k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[U^{\dagger}_{\mathbf{n}}(1,x,\mathbf{j})^2\right], \tag{5.67}$$

$$Q^{\dagger}_{4,\mathbf{n},x} := \hat{\mathbf{n}}^{-1} \sum_{\substack{j_k=0 \\ k=1,\cdots,N}}^{r_k-1} \mathbb{E}\left[U^{\dagger}_{\mathbf{n}}(1,x,\mathbf{j})^2 \, \mathbb{1}_{\left\{\left|U^{\dagger}_{\mathbf{n}}(1,x,\mathbf{j})\right| \geq \varepsilon \widehat{\sigma}(x) \hat{\mathbf{n}}^{1/2}\right\}}\right], \quad \varepsilon > 0. \tag{5.68}$$

Note that

$$\frac{S^{\dagger}_{\mathbf{n}}(x,y)}{\widehat{\sigma}(x)\sqrt{\hat{\mathbf{n}}}} = \frac{T^{\dagger}_{\mathbf{n}}(x,1)}{\widehat{\sigma}(x)\sqrt{\hat{\mathbf{n}}}} + \sum_{i=2}^{2^N} \frac{T^{\dagger}_{\mathbf{n}}(x,i)}{\widehat{\sigma}(x)\sqrt{\hat{\mathbf{n}}}},$$

where $\widehat{\sigma}(x)^2 = l_1^2 \|K\|_2^2 + 2l_1 l_2 \gamma(x) \|K\|_2^2 + 2l_2^2 \gamma^2(x) \|K\|_2^2$. A straightforward application of Corollaries 5.6.7 and 5.6.8, permits to have

$$\begin{aligned} &\lim_{\mathbf{n}\to\infty} Q^{\dagger}_{1,\mathbf{n},x} = 0, \quad \lim_{\mathbf{n}\to\infty} Q^{\dagger}_{2,\mathbf{n},x} = 0 \\ &\lim_{\mathbf{n}\to\infty} Q^{\dagger}_{3,\mathbf{n},x} = \widehat{\sigma}(x)^2 \quad \text{and} \quad \lim_{\mathbf{n}\to\infty} Q^{\dagger}_{4,\mathbf{n},x} = 0. \end{aligned}$$

Then, as above

$$l^{\top} \mathbb{W}_{\mathbf{n}}(x,y) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \widehat{\sigma}(x)^2\right).$$

This ends the proof of the Lemma.■

5.6.5 Appendix A5 : Proofs of main results

Proof of Theorem 5.4.1. Let us decompose $\sqrt{\hat{\mathbf{n}} b^d \overline{F}(v_{\mathbf{n}}, x)}\,(\hat{\gamma}_{\mathbf{n}}(x) - \gamma(x))$ in the following way :

$$\begin{aligned}
\sqrt{\hat{\mathbf{n}} b^d \overline{F}(v_{\mathbf{n}}, x)}\,(\hat{\gamma}_{\mathbf{n}}(x) - \gamma(x)) &=: \frac{\pi(v_{\mathbf{n}}, x)}{\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)} \{V_1 + V_2 + V_3\} \\
V_1 &:= g^{-1/2}(x) \sqrt{\hat{\mathbf{n}} b^d \pi(v_{\mathbf{n}}, x)} \left(\frac{\widehat{\phi}_{v_{\mathbf{n}}}(y,x)}{\pi(v_{\mathbf{n}}, x)} - \mathbf{E}\left[\frac{\widehat{\phi}_{v_{\mathbf{n}}}(y,x)}{\pi(v_{\mathbf{n}}, x)} \right] \right) \\
V_2 &:= -\gamma(x) g^{-1/2}(x) \sqrt{\hat{\mathbf{n}} b^d \pi(v_{\mathbf{n}}, x)} \left(\frac{\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)}{\pi(v_{\mathbf{n}}, x)} - \mathbf{E}\left[\frac{\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)}{\pi(v_{\mathbf{n}}, x)} \right] \right) \\
V_3 &:= \sqrt{\hat{\mathbf{n}} b^d \overline{F}(v_{\mathbf{n}}, x)} \left(\frac{\mathbf{E}\left[\phi_{\mathbf{n}}(y,x)\right] - \gamma(x) \mathbf{E}\left[\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)\right]}{\pi(v_{\mathbf{n}}, x)} \right).
\end{aligned}$$

By Lemma 5.6.6, it follows that

$$\begin{aligned}
V_1 &\xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\gamma^2(x) \frac{\|K\|_2^2}{g(x)}\right) \\
-V_2 &\xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \gamma^2(x) \frac{\|K\|_2^2}{g(x)}\right) \\
V_1 + V_2 &\xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \gamma^2(x) \frac{\|K\|_2^2}{g(x)}\right).
\end{aligned} \tag{5.69}$$

For the term V_3, by assumptions and from Lemma 1 and 2 of Goegebeur et al. (2014b), we have

$$\begin{aligned}
V_3 &= \sqrt{\hat{\mathbf{n}} b^d \overline{F}(v_{\mathbf{n}}, x)} \left(\frac{\mathbf{E}\left[\widehat{\phi}_{v_{\mathbf{n}}}(y,x)\right] - \gamma(x) \mathbf{E}\left[\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)\right]}{\pi(v_{\mathbf{n}}, x)} \right) \\
&= g^{-1/2}(x) \sqrt{\hat{\mathbf{n}} b^d \overline{F}(v_{\mathbf{n}}, x) g(x)} \left(\frac{\mathbf{E}\left[\widehat{\phi}_{v_{\mathbf{n}}}(y,x)\right] - \gamma(x) \mathbf{E}\left[\pi_{\mathbf{n}}(v_{\mathbf{n}}, x)\right]}{\pi(v_{\mathbf{n}}, x)} \right) \\
&= g^{-1/2}(x) \sqrt{\hat{\mathbf{n}} b^d \overline{F}(v_{\mathbf{n}}, x) g(x)} \left\{ \frac{\mathcal{A}(v_{\mathbf{n}}, x)}{\gamma(x)\rho(x)} \left[\frac{1}{(1-\rho(x))^2} - 1 \right] (1 + o(1)) + O(b \log v_{\mathbf{n}}) \right\} \\
V_3 &\xrightarrow{\mathbb{P}} \frac{\lambda(x)}{1 - \rho(x)}.
\end{aligned} \tag{5.70}$$

In other hand, from Lemma 5.6.6

$$\frac{\sum\limits_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_b(x - X_{\mathbf{i}}) \mathbb{1}_{Y > v_{\mathbf{n}}}}{\pi(v_{\mathbf{n}}, x)} = 1 + O_{\mathbb{P}}(1).$$

Combining (5.69) and (5.70) with (5.60), it follows

$$\sqrt{\hat{\mathbf{n}} b^d_{\mathbf{n}} \overline{F}(v_{\mathbf{n}}, x)}\,(\widehat{\gamma}_{v_{\mathbf{n}}}(x) \;-\; \gamma(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda(x)}{1 - \rho(x)} \,; \gamma^2(x) \frac{\|K\|_2^2}{g(x)} \right)$$

Theorem 5.4.1 follows by a straightforward application of the Delta-method. ■

Proof of Theorem 5.4.2 The proof is based on the following transformation provided from (5.17), that is :

$$\frac{\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \overline{F}(\upsilon_{\mathbf{n}}, x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)} \left(\log \hat{q}_{\mathbf{n}}^{W}(\beta_{\mathbf{n}}, x) - \log q(\beta_{\mathbf{n}}, x)\right) =: A_{1,\mathbf{n}}(x) + A_{2,\mathbf{n}}(x) + A_{3,\mathbf{n}}(x),$$

with

$$\begin{aligned}
A_{1,\mathbf{n}}(x) &:= \sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \overline{F}(\upsilon_{\mathbf{n}}, x)}\left(\hat{\gamma}_{\mathbf{n}}(x) - \gamma(x)\right) \\
A_{2,\mathbf{n}}(x) &:= \frac{\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \overline{F}(\upsilon_{\mathbf{n}}, x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)} \left(\log\left(\hat{q}(\alpha_{\mathbf{n}}, x)\right) - \log\left(q(\alpha_{\mathbf{n}}, x)\right)\right) \\
A_{3,\mathbf{n}}(x) &:= \frac{\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \overline{F}(\upsilon_{\mathbf{n}}, x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)} \left(\log\left(q(\alpha_{\mathbf{n}}, x)\right) - \log\left(q(\beta_{\mathbf{n}}, x)\right) + \gamma(x) \log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)\right).
\end{aligned}$$

In one hand, from Theorem 5.4.1,

$$A_{1,\mathbf{n}}(x) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)}, \gamma^{2}(x) \frac{\|K\|_{2}^{2}}{g(x)}\right), \tag{5.71}$$

as a straightforward consequence of Theorem 5.4.1. In the other hand, by assumption, it follows that $\hat{q}(\alpha_{\mathbf{n}}, x)/q(\alpha_{\mathbf{n}}, x) \xrightarrow{\mathbb{P}} 1$, and by Theorem 2 (ii) of Dabo-Niang & Thiam (2010) one can obtain easily (with the help of the assumptions of the theorem) the asymptotic normality of $\hat{q}(\alpha_{\mathbf{n}}, x)$ with rate $\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \overline{F}(\upsilon_{\mathbf{n}}, x)}$), and then we have

$$A_{2,\mathbf{n}}(x = \frac{\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \overline{F}(\upsilon_{\mathbf{n}}, x)}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)} \left(\hat{q}(\alpha_{\mathbf{n}}, x)/q(\alpha_{\mathbf{n}}, x) - 1\right)\left(1 + o(1)\right) = \frac{O_{\mathbb{P}}(1)}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)}.$$

For the last term $A_{3,\mathbf{n}}(x)$, we have from Lemma 2 of Daouia et al. (2011),

$$A_{3,\mathbf{n}}(x) = O\left(\sqrt{\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \overline{F}(\upsilon_{\mathbf{n}}, x)} \delta\left(q(\alpha_{\mathbf{n}}, x), x\right)\right)$$

which converges to 0 by assumption. Finally,

$$\frac{\sqrt{\hat{\mathbf{n}} b^{d} \alpha_{\mathbf{n}}}}{\log\left(\alpha_{\mathbf{n}}/\beta_{\mathbf{n}}\right)} \left(\log \hat{q}_{\mathbf{n}}^{W}(\beta_{\mathbf{n}}, x) - \log q(\beta_{\mathbf{n}}, x)\right) =: A_{1,\mathbf{n}}(x) + A_{2,\mathbf{n}}(x) + A_{3,\mathbf{n}}(x)$$

converges in distribution to $\mathcal{N}\left(\frac{\lambda(x)}{1-\rho(x)}, \gamma^{2}(x) \frac{\|K\|_{2}^{2}}{g(x)}\right)$.

Chapitre 6

Conclusion générale et perspectives

6.1 Conclusion générale

Dans ce mémoire de thèse, nous avons principalement considéré la modélisation non paramétrique dans le cas de données extrêmes spatiales. Plus particulièrement, nous avons modélisé deux classes de données, à savoir les processus spatialement linéairement dépendants et les processus spatiaux satisfaisant une condition de mélange fort ($\alpha-$mélangeant). Dans le premier chapitre de cette thèse, nous avons proposé un estimateur de queue lourde pour une distribution spatiale et ceci par le biais de l'estimateur de Hill (1975). Dans les deux derniers chapitres, nous avons étendu nos travaux aux cas d'une présence de co-variable fixe et aléatoire. En effet, il est intéressant de lier la variable d'intérêt Y à une co-variable X, permettant ainsi d'évaluer par exemple les conséquences que peut occasionner une variable d'intérêt (températures extrêmes par exemple) sur la population (taux de mortalité élevé sur les périodes de canicule). Nous avons utilisé la technique de la "fenêtre mobile" ou "noyau" pour capter la co-variable exigeant ainsi que la taille de l'échantillon soit assez importante pour pouvoir disposer de suffisamment de données et prétendre effectuer une estimation de la fonction inconnue de l'indice de queue conditionnelle. Ensuite, nous avons utilisé cet estimateur de queues lourdes pour proposer un estimateur de son quantile extrême.

Sous certaines hypothèses sur la structure de dépendance spatiale des données, en plus des hypothèses usuelles en théorie des valeurs extrêmes, les convergences en probabilité et en loi, avec vitesse sont obtenues.

Les illustrations numériques avec le "fixed design" montrent que l'estimateur est performant lorsqu'on l'applique à des données spatiales extrêmes issues d'une distribution de type Paréto et une distribution de type Fréchet.

À travers ce travail de thèse, nous contribuons à la statistique spatiale des valeurs extrêmes de manière assez théorique. Notre apport réside sur l'introduction d'estimateurs de queues lourdes pour des données spatiales extrêmes et de leurs quantiles extrêmes correspondants dans le cas non-conditionnel comme dans le cas conditionnel. Tout au long de ce travail, des questions et remarques sont apparues laissant place à quelques perspectives de recherche que nous développons ci-après.

6.2 Perspectives

Tout d'abord, une étude plus poussée de simulations sera faite pour étudier les performances numériques des estimateurs proposés. Les premiers résultats numériques obtenus peuvent être améliorés en effectuant un choix adéquat du nombre d'observations intermédiaires à utiliser pour l'estimation. En effet, tout comme en statistique univarié (non-spatial), le choix du nombre intermédiaire $\hat{k}$ (dans le cas du "fixed design") est assez difficile à effectuer, ce choix s'avère encore plus compliqué puisque les observations ne sont pas régulièrement distribuées dans l'ensemble spatial considéré, il faut donc tenir compte de la dépendance spatiale locale entre les sites.
Du fait de la rareté d'une littérature assez conséquente qui traite le problème d'estimation des indices de queues et des quantiles extrêmes spatiaux, beaucoup de questions sont soulevées conduisant ainsi à un début de chantier de recherche sur ce domaine. Dans la continuité de ce travail de thèse, nous envisageons d'étudier un vecteur $X_{\mathbf{i}} = \left(X_{\mathbf{i}}^{(1)}, \cdots, X_{\mathbf{i}}^{(d)}\right)_{\mathbf{i}\in\mathbb{Z}^N}$, de fonction de distribution multivariée F. Il apparaît immédiatement la question : *comment étudier un tel vecteur multivarié lorsqu'il suit une distribution spatiale des extrêmes* $F = (F_1, \cdots, F_d)$ *?* Plus précisément, on s'intéressera au cas où F appartient à un des trois domaines d'attraction usuels. Nous pensons utiliser la théorie des copules pour aborder ce problème.
Dans le chapitre 3, nous avons considéré deux classes de processus stationnaires. La question que l'on se pose est comment généraliser l'estimateur proposé ? C'est-à-dire, quand on a un processus spatial non stationnaire. Ce cas a été abordé par Diop & Diouf (2010) où ils considérèrent une classe de processus non stationnaires dans un cadre non spatial. La seconde question est liée directement au fait que l'estimateur proposé dans ce chapitre présente l'inconvénient de ne pas être à mesure d'effectuer des estimations quand on est dans les domaines d'attraction de Gumbel et Weibull, ce qui nécessite de nouveaux estimateurs adaptés à ces cas, à savoir l'estimateur de Pickands (1975) ou de Dekkers & Haan (1989).
Dans le chapitre 5, nous avons seulement considéré un processus satisfaisant la condition de mélange fort pour étudier le comportement asymptotique de notre estimateur. Il serait intéressant d'élargir ce travail au cas où l'on a un processus linéaire causal, lié à une co-variable aléatoire.
Dans tout ce travail, nous ne nous sommes pas penchés sur des processus ARCH et GARCH, qui ont été étudiés dans le cas non spatial entres autres par Mikosch & Stărică (2000), Hall & Yao (2003), Davis & Mikosch (2009), Resnick et al. (1997).

Bibliographie

Aarssen, K., & Haan, L. d. (1994). On the maximal life span of humans. *Mathematical Population Studies*, *4*, 259–281.

Adler, R. J. (1978). Weak convergence results for extremal processes generated by dependent random variables. *The Annals of Probability*, *6*, 660–667.

Alvarado, E., Sandberg, D. V., & Pickford, S. G. (1998). *Modeling large forest fires as extreme events.* National Emergency Training Center.

Alves, M. F., Gomes, M. I., & Haan, L. d. (2003a). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, *60*, 193–214.

Alves, M. F., Haan, L. d., & Lin, T. (2003b). Estimation of the parameter controlling the speed of convergence in extreme value theory. *Mathematical Methods of Statistics*, *12*, 155–176.

Balkema, A. A., & de Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, *2*, 792–804.

Barbe, P., & McCormick, W. P. (2009). *Asymptotic expansions for infinite weighted convolutions of heavy tail distributions and applications.* 919-924. American Mathematical Soc.

Barrois, M. T. (1834). *Essai sur l'application du calcul des probabilités aux assurances contre l'incendie.*

Basrak, B., & Tafro, A. (2014). Extremes of moving averages and moving maxima on a regular lattice. *Probability and Mathematical Statistics-Poland*, *34*, 61–79.

Basu, S., & Reinsel, G. (1993). Properties of the spatial unilateral first-order arma model. *Advances in Applied Probability*, *25*, 631–648.

Beirlant, J., Dierckx, G., Goegebeur, Y., & Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes*, *2*, 177–200.

Beirlant, J., Dierckx, G., Guillou, A., & Stărică, C. (2002). On exponential representations of log-spacings of extreme order statistics. *Extremes*, *5*, 157–180.

Beirlant, J., Figueiredo, F., Gomes, M. I., & Vandewalle, B. (2008). Improved reduced-bias tail index and quantile estimators. *Journal of Statistical Planning and Inference*, *138*, 1851–1870.

Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. (2006). *Statistics of Extremes : Theory and Applications*. John Wiley & Sons.

Beirlant, J., Vynckier, P., & Teugels, J. L. (1996). Excess functions and estimation of the extreme-value index. *Bernoulli*, *2*, 293–318.

Biau, G. (2002). Optimal asymptotic quadratic errors of density estimators on random fields. *Statistics & Probability Letters*, *60*, 297–307.

Biau, G. (2003). Spatial kernel density estimation. *Mathematical Methods of Statistics*, *12*, 371–390.

Biau, G., & Cadre, B. (2004). Nonparametric spatial prediction. *Statistical Inference for Stochastic Processes*, *7*, 327–349.

Bingham, N. H., Goldie, C. M., & Teugels, J. L. (1987). *Regular variation* volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

Bingham, N. H., Goldie, C. M., & Teugels, J. L. (1989). *Regular variation* volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

Bobée, B., & Rasmussen, P. F. (1995). Recent advances in flood frequency analysis. *Reviews of Geophysics*, *33*, 1111–1116.

Borkovec, M., & Klüppelberg, C. (2001). The tail of the stationary distribution of an autoregressive process with arch (1) errors. *The Annals of Applied Probability*, *11*, 1220–1241.

Bouleau, N. (1991). Splendeurs et misères des lois de valeurs extêmes. *Revue Risques*, *4*, 85–92.

Brockwell, P. J., & Lindner, A. (2010). Strictly stationary solutions of autoregressive moving average equations. *Biometrika*, *97*, 765–772.

Brodin, E., & Rootzén, H. (2009). Univariate and bivariate gpd methods for predicting extreme wind storm losses. *Insurance : Mathematics and Economics*, *44*, 345–356.

Carbon, M., Francq, C., & Tran, L. T. (2007). Kernel regression estimation for random fields. *Journal of Statistical Planning and Inference*, *137*, 778–798.

Carbon, M., Hallin, M., & Tat Tran, L. (1996). Kernel density estimation for random fields : the l 1 theory. *Journal of Nonparametric Statistics*, *6*, 157–170.

Carbon, M., Tran, L. T., & Wu, B. (1997). Kernel density estimation for random fields (density estimation for random fields). *Statistics & Probability Letters*, *36*, 115–125.

Carreau, J., & Girard, S. (2011). Spatial extreme quantile estimation using a weighted log-likelihood approach. *Journal de la Société Française de Statistique*, *152*, 66–83.

Christensen, R. (1991). *Linear models for multivariate, time series, and spatial data.* Springer Science & Business Media.

Cline, D. (1983). Infinite series of random variables with regularly varying tails. *Tech. Rpt., Inst. Appl. Math. Statist., Univ. British Columbia*, .

Coles, S. (2001). *An introduction to statistical modeling of extreme values* volume 208. Springer.

Coles, S., & Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, *58*, 329–347.

Coles, S. G., & Walshaw, D. (1994). Directional modelling of extreme wind speeds. *Applied Statistics*, *43*, 139–157.

Collomb, G. (1980). Estimation non paramétrique de probabilités conditionnelles. *Comptes Rendus de l'Académie des Sciences*, *291*, 427–430.

Cressie, N. (1992). Statistics for spatial data. *Terra Nova*, *4*, 613–617.

Cressie, N., & Wikle, C. K. (2011). *Statistics for spatio-temporal data.* John Wiley & Sons.

Csörgó, S., Deheuvels, P., & Mason, D. (1985). Kernel estimates of the tail index of a distribution. *The Annals of Statistics*, *13*, 1050–1077.

Cui, H., He, X., & Ng, K. W. (2004). M-estimation for linear models with spatially-correlated errors. *Statistics & Probability Letters*, *66*, 383–393.

Dabo-Niang, S., Ould-Abdi, S. A., Ould-Abdi, A., & Diop, A. (2014). Consistency of a nonparametric conditional mode estimator for random fields. *Statistical Methods & Applications*, *23*, 1–39.

Dabo-Niang, S., & Thiam, B. (2010). Robust quantile estimation and prediction for spatial processes. *Statistics & Probability Letters*, *80*, 1447–1458.

Dabo-Niang, S., & Yao, A.-F. (2007). Kernel regression estimation for continuous spatial processes. *Mathematical Methods of Statistics*, *16*, 298–317.

Dabo-Niang, S., & Yao, A.-F. (2013). Kernel spatial density estimation in infinite dimension space. *Metrika*, *76*, 19–52.

Danielsson, J., Haan, L. d., Peng, L., & de Vries, C. G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, *76*, 226–248.

Danielsson, J., Jansen, D. W., & De vries, C. G. (1996). The method of moments ratio estimator for the tail shape parameter. *Communications in Statistics-Theory and Methods*, *25*, 711–720.

Danielsson, J., & de Vries, C. G. (1997). Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance*, *4*, 241–257.

Daouia, A., Gardes, L., Girard, S., & Lekina, A. (2011). Kernel estimators of extreme level curves. *Test*, *20*, 311–333.

Datta, S., & McCormick, W. P. (1998). Inference for the tail parameters of a linear process with heavy tail innovations. *Annals of the Institute of Statistical Mathematics*, *50*, 337–359.

Davis, R., & Resnick, S. (1984). Tail estimates motivated by extreme value theory. *The Annals of Statistics*, *12*, 1467–1487.

Davis, R., & Resnick, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *The Annals of probability*, *13*, 179–195.

Davis, R., & Resnick, S. (1988). Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stochastic Processes and their Applications*, *30*, 41–68.

Davis, R. A., & Mikosch, T. (2009). Extreme value theory for garch processes. In *Handbook of Financial Time Series* (pp. 187–200). Springer.

Davis, R. A., & Resnick, S. I. (1991). Extremes of moving averages of random variables with finite endpoint. *The Annals of Probability*, *19*, 312–328.

Davis, R. A., & Resnick, S. I. (1996). Limit theory for bilinear processes with heavy-tailed noise. *The Annals of Applied Probability*, *6*, 1191–1210.

Davison, A., & Smith, R. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B. Methodological*, *52*, 393–442.

Davison, A. C., Padoan, S., & Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical Science*, *27*, 161–186.

Davison, A. C., & Ramesh, N. (2000). Local likelihood smoothing of sample extremes. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, *62*, 191–208.

Deheuvels, P., Haeusler, E., & Mason, D. M. (1988). Almost sure convergence of the hill estimator. In *Mathematical Proceedings of the Cambridge Philosophical Society* (pp. 371–381). Cambridge Univ Press volume 104.

Dekkers, A. L., Einmahl, J. H., & De Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, *17*, 1833–1855.

Dekkers, A. L., & Haan, L. d. (1989). On the estimation of the extreme-value index and large quantile estimation. *The Annals of Statistics*, *17*, 1795–1832.

Diop, A., & Diouf, S. (2010). Extreme value theory for nonstationary random coefficients time series with regularly varying tails. *Journal Afrika Statistika*, *5*, 268–278.

Doreian, P. (1980). Linear models with spatially distributed data spatial disturbances or spatial effects ? *Sociological Methods & Research*, *9*, 29–60.

Doukhan, P. (1994). *Mixing* volume 85 of *Lecture Notes in Statistics*. Springer New York.

Drees, H. (2008). Some aspects of extreme value statistics under serial dependence. *Extremes*, *11*, 35–53.

Drees, H., & Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and their Applications*, *75*, 149–172.

El Machkouri, M. (2007). Nonparametric regression estimation for random fields in a fixed-design. *Statistical Inference for Stochastic Processes*, *10*, 29–47.

El Machkouri, M. (2011). Asymptotic normality of the parzen–rosenblatt density estimator for strongly mixing random fields. *Statistical Inference for Stochastic Processes*, *14*, 73–84.

El Machkouri, M., & Stoica, R. (2010). Asymptotic normality of kernel estimates in a regression model for random fields. *Journal of Nonparametric Statistics*, *22*, 955–971.

Embrechts, P., Mikosch, T., & Klüppelberg, C. (1997). *Modelling Extremal Events : For Insurance and Finance.* London, UK : Springer-Verlag.

Embrechts, P., Resnick, S. I., & Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. *North American Actuarial Journal*, *3*, 30–41.

Fawcett, L., & Walshaw, D. (2014). Estimating the probability of simultaneous rainfall extremes within a region : a spatial approach. *Journal of Applied Statistics*, *41*, 959–976.

Fazekas, I. (2003). Limit theorems for the empirical distribution function in the spatial case. *Statistics & Probability Letters*, *62*, 251–262.

Fazekas, I., & Chuprunov, A. (2006). Asymptotic normality of kernel type density estimators for random fields. *Statistical Inference for Stochastic Processes*, *9*, 161–178.

Ferreira, A., Haan, L. d., & Peng, L. (2003). On optimising the estimation of high quantiles of a probability distribution. *Statistics*, *37*, 401–434.

Ferrez, J., Davison, A., & Rebetez, M. (2011). Extreme temperature analysis under forest cover compared to an open field. *Agricultural and Forest Meteorology*, *151*, 992–1001.

Fisher, R. A., Tippett, L. et al. (1928). *Limiting Forms of the Frequency Distribution of the Largest of Smallest Member of a Sample.* volume 24.

Gaetan, C., & Guyon, X. (2008). *Modélisation et statistique spatiales.* New York : Springer.

García-Soidán, P., & Menezes, R. (2012). Estimation of the spatial distribution through the kernel indicator variogram. *Environmetrics*, *23*, 535–548.

Gardes, L., & Girard, S. (2008). A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, *99*, 2368–2388.

Gardes, L., & Girard, S. (2010). Conditional extremes from heavy-tailed distributions : An application to the estimation of extreme rainfall return levels. *Extremes*, *13*, 177–204.

Gardes, L., & Girard, S. (2012). Functional kernel estimators of large conditional quantiles. *Electronic Journal of Statistics*, *6*, 1715–1744.

Gardes, L., Guillou, A., & Schorgen, A. (2012). Estimating the conditional tail index by integrating a kernel conditional quantile estimator. *Journal of Statistical Planning and Inference*, *142*, 1586–1598.

Gardes, L., & Stupfler, G. (2014). Estimation of the conditional tail index using a smoothed local hill estimator. *Extremes*, *17*, 45–75.

Geluk, J. L., & Haan, L. d. (1987). Regular variation, extensions and tauberian theorems. *CWI tracts*, *40*, 1–131.

Gencay, R., & Selcuk, F. (2004). Extreme value theory and value-at-risk : Relative performance in emerging markets. *International Journal of Forecasting*, *20*, 287–303.

Gheriballah, A., Laksaci, A., & Rouane, R. (2010). Robust nonparametric estimation for spatial regression. *Journal of Statistical Planning and Inference*, *140*, 1656–1670.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, *44*, 423–453.

Goegebeur, Y., Guillou, A., & Osmann, M. (2014a). A local moment type estimator for the extreme value index in regression with random covariates. *Canadian Journal of Statistics*, *42*, 487–507.

Goegebeur, Y., Guillou, A., & Schorgen, A. (2014b). Nonparametric regression estimation of conditional tails : the random covariate case. *Statistics*, *48*, 732–755.

Gomes, M. I., Figueiredo, F., & Mendonça, S. (2005). Asymptotically best linear unbiased tail estimators under a second-order regular variation condition. *Journal of Statistical Planning and Inference*, *134*, 409–433.

Gomes, M. I., & Martins, M. J. (2001). Generalizations of the hill estimator–asymptotic versus finite sample behaviour. *Journal of Statistical Planning and Inference*, *93*, 161–180.

Gomes, M. I., Martins, M. J., & Neves, M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events-the jackknife methodology. *Extremes*, *3*, 207–229.

Gumbel, E. J. (1937). *La durée extrême de la vie humaine*. Hermann et Cie. Paris.

Gumbel, E. J., & Lieblein, J. (1954). *Statistical theory of extreme values and some practical applications : a series of lectures* volume 33. US Government Printing Office Washington.

Guyon, X. (1995). *Random fields on a network : modeling, statistics, and applications*. Springer Science & Business Media (Probability and its Applications).

Haan, L. d. (1976). Sample extremes : an elementary introduction. *Statistica Neerlandica*, *30*, 161–172.

Haan, L. d., & Ferreira, A. (2006). *Extreme Value Theory : An Introduction*. Springer Series in Operations Research and Financial Engineering. New York : Springer.

Haan, L. d., & Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica*, *52*, 60–70.

Haan, L. d., & Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, *40*, 317–337.

Haan, L. F. M. d. (1985). *Extremes in higher dimensions : the model and some statistics*. Erasmus University.

Haeusler, E., & Teugels, J. L. (1985). On asymptotic normality of hill's estimator for the exponent of regular variation. *The Annals of Statistics*, *13*, 743–756.

Hall, P., & Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, *15*, 153–167.

Hall, P., & Yao, Q. (2003). Inference in arch and garch models with heavy-tailed errors. *Econometrica*, *71*, 285–317.

Hallin, M., Lu, Z., & Tran, L. T. (2001). Density estimation for spatial linear processes. *Bernoulli*, *7*, 657–668.

Hallin, M., Lu, Z., & Tran, L. T. (2004). Local linear spatial regression. *The Annals of Statistics*, *32*, 2469–2500.

Hallin, M., Lu, Z., & Yu, K. (2009). Local linear spatial quantile regression. *Bernoulli*, *15*, 659–686.

Han, Z. J. (2005). Living to 100 and beyond : An extreme value study. In *Living to 100 and Beyond Symposium Monograph*.

Heneka, P., Hofherr, T., Ruck, B., & Kottmeier, C. (2006). Winter storm risk of residential structures–model development and application to the german state of baden-württemberg. *Natural Hazards and Earth System Science*, *6*, 721–733.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, *3*, 1163–1174.

Hosking, J. R., & Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, *29*, 339–349.

Hsing, T. (1991). On tail index estimation using dependent data. *The Annals of Statistics*, *19*, 1547–1569.

Hsing, T., Hüsler, J., & Leadbetter, M. (1988). On the exceedance point process for a stationary sequence. *Probability Theory and Related Fields*, *78*, 97–112.

Huang, D., & Anh, V. (1992). Estimation of spatial arma models. *Australian Journal of Statistics*, *34*, 513–530.

Illig, A., & Truong-Van, B. (2004). *Asymptotic results for spatial ARMA models.* Centre National de la Recherche Scientifique (CNRS) et Université de Toulouse (France). Laboratoire de Statistique et Probabilités [LSP], préprint.

Ivanoff, B. G., Weber, N. et al. (2010). Asymptotic results for spatial causal arma models. *Electronic Journal of Statistics*, *4*, 15–35.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, *81*, 158–171.

Journel, A. G. (1983). Nonparametric estimation of spatial distributions. *Journal of the International Association for Mathematical Geology*, *15*, 445–468.

Kallenberg, O. (1983). *Random Measures.* 3rd ed. Akademie, Berlin.

Karácsony, Z., & Filzmoser, P. (2010). Asymptotic normality of kernel type regression estimators for random fields. *Journal of Statistical Planning and Inference*, *140*, 872–886.

Kashyap, R. L. (1984). Characterization and estimation of two-dimensional arma models. *Information Theory, IEEE Transactions on*, *30*, 736–745.

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*, 1287–1304.

Khaliq, M., Ouarda, T., Ondo, J.-C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations : A review. *Journal of hydrology*, *329*, 534–552.

Klajnmic, H. (2004). Estimation et comparaison des niveaux de retour des vitesses extrêmes des vents. *XXXVIemes Journées de Statistiques, Montpellier, France, XXXVIemes.*

Kratz, M., & Resnick, S. I. (1996). The qq-estimator and heavy tails. *Stochastic Models*, *12*, 699–724.

Laksaci, A., & Maref, F. (2009). Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes. *Comptes Rendus Mathematique*, *347*, 1075–1080.

Leadbetter, M., Lindgren, G., & Rootzén, H. (1983). Extremes and related properties of random sequences and processes. *Springer Series in Statistics. Springer-Verlag, New York 1983*, .

Leadbetter, M. R. (1974). On extreme values in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, *28*, 289–303.

Leadbetter, M. R., & Hsing, T. (1990). Limit theorems for strongly mixing stationary random measures. *Stochastic processes and their applications*, *36*, 231–243.

Lekina, A. (2010). *Estimation non-paramétrique des quantiles extrêmes conditionnels*. Ph.D. thesis Université Joseph-Fourier-Grenoble I.

Li, J., & Tran, L. T. (2009). Nonparametric estimation of conditional expectation. *Journal of Statistical Planning and Inference*, *139*, 164–175.

Lô, G. S. (1986). *Sur quelques estimateurs de l'index d'une loi de Pareto : estimateur de Hill, de S. Csörgó-Deheuvels-Mason, de Haan/Resnick et lois limites de sommes de valeurs extrêmes pour une variable aléatoire dans le domaine d'attraction de Gumbel*. Ph.D. thesis.

Longin, F. M. (2000). From value at risk to stress testing : The extreme value approach. *Journal of Banking & Finance*, *24*, 1097–1130.

Lozano, A. C., Li, H., Niculescu-Mizil, A., Liu, Y., Perlich, C., Hosking, J., & Abe, N. (2009). *Spatial-temporal causal modeling for climate change attribution*.

Lu, Z.-d., & Chen, X. (2002). Spatial nonparametric regression estimation : Non-isotropic case. *Acta Mathematicae Applicatae Sinica (English Series)*, *18*, 641–656.

Markovich, N. M. (2005). High quantile estimation for heavy-tailed distributions. *Performance Evaluation*, *62*, 178–192.

Mason, D. M. (1982). Laws of large numbers for sums of extreme values. *The Annals of Probability*, *10*, 754–764.

Matthys, G., & Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica*, *13*, 853–880.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series : an extreme value approach. *Journal of empirical finance*, *7*, 271–300.

McNeil, A. J., & Saladin, T. (1997). *The peaks over thresholds method for estimating high quantiles of loss distributions*.

Menezes, R., Garcia-Soidan, P., & Ferreira, C. (2010). Nonparametric spatial prediction under stochastic sampling design. *Journal of Nonparametric Statistics*, *22*, 363–377.

Mikosch, T., & Stărică, C. (2000). Limit theory for the sample autocorrelations and extremes of a garch (1, 1) process. *The Annals of Statistics*, *28*, 1427–1451.

Müller, S., & Rufibach, K. (2009). Smooth tail-index estimation. *Journal of Statistical Computation and Simulation*, *79*, 1155–1167.

Ndao, P., Diop, A., & Dupuy, J. (2016). Nonparametric estimation of the conditional extreme-value index with random covariates and censoring. *Journal of Statistical Planning and Inference*, *168*, 20 – 37.

Ndao, P., Diop, A., & Dupuy, J.-F. (2014). Nonparametric estimation of the conditional tail index and extreme quantiles under random censoring. *Computational Statistics & Data Analysis*, *79*, 63–79.

Novak, S. Y. (2011). *Extreme value methods with applications to finance.* CRC Press.

Ould-Abdi, S., Dabo-Niang, S., Diop, A., & Ould-Abdi, A. O. (2010). Consistency of a nonparametric conditional quantile estimator for random fields. *Mathematical Methods of Statistics*, *19*, 1–21.

Padoan, S. A., Ribatet, M., & Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, *105*, 263–277.

Peng, L. (1998). Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters*, *38*, 107–115.

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, *3*, 119–131.

Reiss, R.-D., & Thomas, M. (2001). *Statistical analysis of extreme values.* Springer.

Resnick, S., & Stărică, C. (1993). *Consistency of Hill's Estimator for Dependent Data.* Technical Report Cornell University Operations Research and Industrial Engineering.

Resnick, S., & Stărică, C. (1995). Consistency of hill's estimator for dependent data. *Journal of Applied probability*, *32*, 139–167.

Resnick, S., & Stărică, C. (1997). Asymptotic behavior of hill's estimator for autoregressive data. *Communications in Statistics. Stochastic Models*, *13*, 703–721.

Resnick, S., & Stărică, C. (1998). Tail index estimation for dependent data. *The Annals of Applied Probability*, *8*, 1156–1183.

Resnick, S. I. (1987). *Extreme values, regular variation, and point processes* volume 4. Springer Verlag.

Resnick, S. I. et al. (1997). Heavy tail modeling and teletraffic data : special invited paper. *The Annals of Statistics*, *25*, 1805–1869.

Ripley, B. (1981). *Spatial statistics.* New York : Wiley.

Ripley, B. D. (1991). *Statistical inference for spatial processes.* Cambridge university press.

Robinson, P. M. (2011). Asymptotic theory for nonparametric regression with spatial data. *Journal of Econometrics*, *165*, 5–19.

Rootzén, H., & Tajvidi, N. (1997). Extreme value statistics and wind storm losses : a case study. *Scandinavian Actuarial Journal*, *1997*, 70–94.

Rootzén, H., & Tajvidi, N. (2001). Can losses caused by wind storms be predicted from meteorological observations ? *Scandinavian Actuarial Journal*, *2001*, 162–175.

Schultze, J., & Steinebach, J. (1996). On least squares estimates of an exponential tail coefficient. *Statistics & Risk Modeling*, *14*, 353–372.

Silverberg, G., & Verspagen, B. (2007). The size distribution of innovations revisited : an application of extreme value statistics to citation and value measures of patent significance. *Journal of Econometrics*, *139*, 318–339.

Smith, R. L. (1987). Estimating tails of probability distributions. *The annals of Statistics*, *15*, 1174–1207.

Smith, R. L. (1989). Extreme value analysis of environmental time series : an application to trend detection in ground-level ozone. *Statistical Science*, *4*, 367–377.

Smith, R. L. (2001). Extreme Value Statistics in Meteorology and the Environment. Chapter 8, NSF-CBMS conference notes, pp. 300-357. URL : `http://www.stat.unc.edu/postscript/rs/envnotes.pdf`.

Stedinger, J. R. (2000). *Flood frequency analysis and statistical estimation of flood risk*.

Stein, M. L. (1999). *Interpolation of spatial data : some theory for kriging*. Springer.

Ternynck, C. (2014). Spatial regression estimation for functional data with spatial dependency. *Journal de la Société Française de Statistique*, *155*, 138–160.

Tjøstheim, D. (1978). Statistical spatial series modelling. *Advances in Applied Probability*, *10*, 130–154.

Tjøstheim, D. (1983). Statistical spatial series modelling ii : some further results on unilateral lattice processes. *Advances in Applied Probability*, *15*, 562–584.

Tjøstheim, D. (1987). *Spatial series and time series : similarities and differences*.

Tran, L. T. (1990). Kernel density estimation on random fields. *Journal of Multivariate Analysis*, *34*, 37–53.

Tran, L. T., & Yakowitz, S. (1993). Nearest neighbor estimators for random fields. *Journal of Multivariate Analysis*, *44*, 23–46.

Van der Vaart, A. (1998). *Asymptotic Statistics* volume 3. Cambridge University Press.

Von Mises, R. (1936). La distribution de la plus grande de n valeurs. *Rev. math. Union interbalcanique*, *1*.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, *73*, 812–815.

Wikipédia (2016). Catastrophe naturelle. URL : `https://fr.wikipedia.org/wiki/Catastrophe_naturelle`.

Yao, Q. (1999). Conditional predictive regions for stochastic processes. *Rapport technique, University of Kent at Canterbury*, *99*.