Some contributions in probability and statistics of extremes.

Marie Kratz

HAL Id: tel-00239329
https://theses.hal.science/tel-00239329

Submitted on 5 Feb 2008

Document de synthèse présenté pour l'Habilitation à Diriger des Recherches

Spécialité Mathématiques

par

Marie KRATZ

SAMOS-MATISSE, UMR 8575, Université Paris I
et MAP5, UMR 8145, Université Paris V

Soutenue le 15 novembre 2005 devant le jury composé de

Jean-Marc Azaïs
Yuri Davydov
Anne Estrade
Xavier Guyon
Annie Millet
Philippe Soulier, rapporteur
Mario Wschebor, rapporteur

Remerciements

A mes co-auteurs, Jürg Hüsler et Holger Rootzén rencontrés au Center for Stochastic Processes à Chapel Hill lors de ma thèse, Sid Resnick qui m'a prise en post-doctorat à Cornell et que je remercie encore, José León dit Chichi, Pierre Picco, Miguel Atencia et Gonzalo Joya, avec lesquels ce fut et c'est un réel plaisir de travailler et de partager toutes les joies et les peines liées à l'activité de recherche.

A Chichi et Bernard Bru, que j'estime tant pour leurs qualités humaines que scientifiques, et dont l'exemple est pour moi un formidable moteur. Encore un grand merci !

Aux membres du jury : Philippe Soulier et Mario Wschebor, rapporteurs choisis par la commission parisienne des thèses, Xavier Guyon, qui a présenté mon dossier à Paris 1, Jean-Marc Azaïs, Yuri Davydov, Anne Estrade et Annie Millet, que je remercie chaleureusement de m'avoir fait l'honneur de participer à mon jury.

A Jean-Marc Azaïs, Xavier Guyon, Chichi León, Philippe Soulier et Mario Wschebor, pour leur lecture détaillée de mon travail et leurs remarques constructives.

Aux collègues de Paris V, en particulier à Mamadou Mboup, à Georges Koepfler, pour leur disponibilité et leur gentillesse, à Anne Estrade pour son dynamisme, à Martine Aubel pour son dévouement et sa bonne humeur, et aussi à Renaud Rioboo, collègue de Paris VI, pour son aide précieuse.

A tous les Samosiens, en particulier à Marie Cotrell débordante d'énergie et de bonne humeur, qui font du Samos un laboratoire hors du commun, lieu également privilégié pour rencontrer et travailler avec de nombreux invités, en ce qui me concerne, Chichi León, Mario Wschebor, Wilfredo Urbina et Gonzalo Joya.

Table des matières

Présentation générale

Les problèmes posés par l'hydrologie, la finance, l'assurance, et plus récemment par l'écologie/les études environnementales (évaluation du réchauffement climatique, détermination des pics de pollution), les grands réseaux de télécommunication (évaluation des risques de congestion des réseaux), la théorie du signal, le traitement d'image, la fiabilité des systèmes complexes, ont renouvelé l'intérêt ces dernières années pour l'étude des comportements extrêmes des processus stochastiques, stationnaires ou non, gaussiens, α-stables, etc...
C'est principalement dans ce cadre que se déroulent mes activités de recherche depuis une dizaine d'années, autour de plusieurs thèmes regroupés en deux axes :

• le premier axe est constitué de l'étude des franchissements de niveau par des processus gaussiens, ou de façon plus générale, l'étude de fonctionnelles non linéaires de processus gaussiens

• le second axe traite de problèmes de statistique et de mécanique statistique en temps discret tels :
⋆ l'étude du processus des excédences
⋆ l'estimation de paramètres et leur validité lors de la modélisation par des processus de distribution à queue épaisse (heavy tailed),
⋆ l'étude de fonctions liées aux statistiques d'ordre de processus gaussiens en vue d'application à la mécanique statistique, ...

Cette présentation générale est une introduction à la synthèse de mes travaux constituée de deux chapitres (correspondant aux deux axes mentionnés ci-dessus).

Pour plus de clarté, les références à mes articles seront regroupées en fin de cette présentation générale (indépendamment des bibliographies de la synthèse), suite à l'exposition de mes perspectives de recherche ; elles figureront en caractères gras dans cette introduction. Pour les autres références, il convient de se reporter aux bibliographies des chapitres et sections correspondants.

Chapitre 1 - Etude de fonctionnelles non linéaires de processus gaussiens.

J'ai profité de ce travail pour présenter, outre ma modeste contribution dans ce domaine, un synopsis des études faites sur le nombre de dépassements ou de franchissements de niveau par un processus gaussien $X = (X_t, t \in \mathbb{R})$, rédigé en anglais pour le rendre plus facilement accessible aux chercheurs intéressés par ce sujet et qui sera soumis à publication. Un important travail de bibliographie y a été joint.
L'étude des franchissements de niveau est un domaine riche en potentiel de champs d'applications. D'autre part, il m'a donné accès à de vastes domaines de probabilité grâce à la variété des approches choisies pour son étude.

Alors qu'un des chapitres de ma thèse ([**1**]) et deux articles écrits en collaboration avec J. Hüsler ([**3**] et [**4**]) ont été consacrés à l'approximation poissonnienne du processus des excédences $(\mathbb{1}_{(X_1 > u_n)}, \cdots, \mathbb{1}_{(X_n > u_n)}, \cdots)$ sous diverses hypothèses portant sur la suite de v.a. $(X_i, i \in \mathbb{N}^*)$ de loi commune ou sur la suite numérique des niveaux $(u_n)_n$, l'attention sera portée ici au cas de processus stochastique ou champ aléatoire à temps continu, la notion d'excédences faisant alors place à celle de dépassements ou de franchissements de niveau.

Soit $X = (X_t, t \in \mathbb{R}^d)$ processus stochastique réel. Nous nous intéressons à la mesure de l'ensemble aléatoire $C_x^X := \{t : X_t = x\}$ de niveau x (pouvant dépendre du temps : $x = x(t)$).

- **Le cas gaussien unidimensionnel** $(d = 1)$ est traité non seulement en citant les principaux résultats, mais en donnant également les idées clefs des méthodes utilisées, spécifiques ou non aux processus gaussiens, en particulier de celles pouvant être facilement adaptées à une dimension supérieure.

  Interviennent dans cette partie les contributions [**7**], [**8**], [**9**], [**10**], [**14**] et [**13**].

  Soit $X = (X_s, s \geq 0)$ un processus réel centré stationnaire gaussien, de variance 1, à trajectoires différentiables, de fonction de corrélation $r(t) = \mathbb{E}[X_0 X_t]$, donnée aussi en fonction de la distribution spectrale F par $r(t) = \int_0^\infty \cos(\lambda t) dF(\lambda)$.

  λ_2 désignera le second moment spectral (quand il existe), i.e. $\lambda_2 = \int_0^\infty \lambda^2 dF(\lambda)$.

  Le nombre de franchissements par X sur un intervalle $[0, t]$, $t > 0$, d'un niveau donné x ou d'une courbe $\psi(.)$ sera noté respectivement par $N_t(x)$ ou $N_t(\psi)$, et $N_t^+(x)$ désignera le nombre de dépassements par X. Nous rappelons que X_s est dit avoir un dépassement de x en $s_0 > 0$ s'il existe $\varepsilon > 0$ tel que $X_s \leq x$ lorsque $s \in (s_0 - \varepsilon, s_0)$ et $X_s \geq x$ pour $s \in (s_0, s_0 + \varepsilon)$.

  Nous nous intéresserons aux moments (factoriels ou non) du nombre de franchissements ou de dépassements, en vue des résultats en loi de ces variables aléatoires ; nous chercherons en particulier des conditions de finitude des moments, nécessaires aux applications. Puis nous ferons un parallèle entre nombre de franchissements et temps local, et parlerons de l'approximation du temps local par le nombre de franchissements.

  Finalement, nous étudierons les comportements asymptotiques du nombre de franchissements ou de dépassements.

  ▷ **Moments et moments factoriels.**

  Remarquons tout d'abord que les conditions de finitude des moments sont locales. L'inégalité de Hölder implique $(\mathbb{E}[N_{2t}(x)])^k \leq 2^k (\mathbb{E}[N_t(x)])^k$, ainsi $(\mathbb{E}[N_t(x)])^k < \infty$ pour un certain t, signifie que c'est vrai pour tout t.

  Commençons par l'un des résultats les plus connus, établi par Rice ([143]) en 1945 par des méthodes intuitives et motivé par la compression d'un signal en ne conservant que les séjours au dessus ou au dessous d'un niveau donné :

  $$\mathbb{E}[N_t(x)] = te^{-x^2/2} \sqrt{-r''(0)} / \pi$$

  signifie que le nombre de franchissements est en moyenne le plus important en 0 et qu'il décroît exponentiellement avec le niveau.

  Cette égalité a été démontrée 20 ans plus tard par Itô ([80]) et Ylvisaker ([182]), permettant d'obtenir ainsi une condition nécessaire et suffisante de finitude de la moyenne du

nombre de franchissements de tout niveau x fixé, soit :

$$IE[N_t(x)] < \infty \quad \Leftrightarrow \quad -r''(0) < \infty \quad \Leftrightarrow \quad \lambda_2 < \infty.$$

Ce résultat a été généralisée par Cramér et Leadbetter ([40]) et par Ylvisaker ([183]) au nombre de franchissements d'une courbe continue différentiable.
Cramér et Leadbetter ([40]), Belayev ([20]), Ylvisaker ([183]) donnèrent les expressions explicites des moments (factoriels) d'ordre 2 dans le cas gaussien, par exemple le second moment

$$IE[N_t^2(x)] = IE[N_t(x)] + 2\int_0^t (t-u) \int_{I\!R^2} |\dot{x}||\dot{y}|\phi_u(x, \dot{x}, x, \dot{y})d\dot{x}d\dot{y}du,$$

où ϕ_u est la densité gaussienne de $(X_0, \dot{X}_0, X_u, \dot{X}_u)$.
En 1972, Geman ([59]) démontra que la condition suffisante, proposée en 1967 par Cramér et Leadbetter ([40]), pour avoir une variance du nombre de franchissements en 0 finie, était aussi nécessaire. Cette condition nécessaire et suffisante, désormais connue dans la littérature sous le nom de condition de Geman, s'énonce ainsi :

$$\exists\, \delta > 0 \ , \ \ L(t) := \frac{r''(t) - r''(0)}{t} \ \in L^1([0, \delta], dx) \quad \Leftrightarrow \quad E[N_t^2(0)] < \infty.$$

Restait la généralisation de ce résultat à n'importe quel niveau. Cuzick ([42], [45] et [46]) proposa, quelques années plus tard, des conditions suffisantes pour généraliser ce résultat à d'autres niveaux et aux moments d'ordres supérieurs, mais il fallut attendre encore plus de quinze ans pour obtenir un résultat général : la condition de Geman est nécessaire et suffisante pour tout niveau et pour les franchissements d'une fonction $\mathcal{C}^1$, à dérivée modérée mais non nécessairement $\mathcal{C}^1$.

Résultat principal. ([**13**], avec J. León).
1) Pour tout niveau x donné,

$$IE[N_t^2(x)] < \infty \quad \Leftrightarrow \quad L(t) = \frac{r''(t) - r''(0)}{t} \ \in L^1([0, \delta], dx) \quad \text{(condition de Geman)}.$$

2) Soit ψ fonction réelle continue dérivable telle que,
si γ désigne le module de continuité de $\dot{\psi}$, $\int_0^\delta \frac{\gamma(s)}{s}ds < \infty$, avec $\delta > 0$.
Alors
$$IE[N_t^2(\psi)] < \infty \quad \Leftrightarrow \quad L(t) \ \in L^1([0, \delta], dx).$$

Notons que Wschebor (1985, [180]), considérant deux niveaux distincts x et y ($x \neq y$), obtint sous la condition de Geman l'expression explicite :

$$IE[N_t(x)N_t(y)] = \int_0^t (t-u) \int_{I\!R^2} |\dot{x}||\dot{y}| \left(\phi_u(x, \dot{x}, y, \dot{y}) + \phi_u(y, \dot{x}, x, \dot{y})\right) d\dot{x}d\dot{y}du,$$

ce qui montre que la fonction $IE[N_t(x)N_t(y)]$ n'est pas continue sur la diagonale puisque à la limite lorsque $x \to y$, cette expression diffère de celle de Cramér et al.

Parmi les auteurs ayant étudié les moments d'ordres supérieurs à 2, nous pouvons citer, entre autres, Cramér et Leadbetter (1965, [40]), Ylvisaker (1966, [183]), Belyaev (1967, [20]), Cuzick (1975-78, [42], [45] et [46]), Marcus (1977, [111]), Nualart et Wschebor (1991, [127]), Azaïs et Wschebor (2001, [13]). Ont été proposées, d'une part des conditions nécessaires et suffisantes peu manipulables pour obtenir la finitude du moment factoriel d'ordre k du nombre de zéros du processus, d'autre part des conditions suffisantes plus explicites et adaptées au cas plus général de franchissements de niveau quelconque ou de courbe spécifique.

Mentionnons par exemple le résultat de Nualart et Wschebor (1991, [127]) :

Si $X = (X_t)_{t \in I \subset \mathbb{R}}$ est un processus gaussien à trajectoires C^∞, tel que var$(X_t) \geq a > 0$, $t \in I$, alors $\mathbb{E}[N_t(u)(N_t(u) - 1)\cdots(N_t(u) - k + 1)] < \infty$ pour chaque niveau $u \in \mathbb{R}$ et chaque $k \in \mathbb{N}^*$.

Il existe deux outils fondamentaux dans l'étude générale des processus gaussiens : le lemme de Slepian et le lemme de concentration de la mesure, comme le fait remarquer Talagrand dans son dernier livre ([167]). Pour l'étude des franchissements d'un processus gaussien, les techniques utilisées pour démontrer les divers résultats ci-dessus reposent certes sur le lemme de Slepian (ou lemme de comparaison normale), mais aussi sur une autre méthode importante, la méthode des moments, dite encore méthode de Rice, qui peut être généralisée au cas non gaussien. Ces deux méthodes sont exposées dans le premier chapitre de la synthèse. L'idée du lemme de comparaison normale, outil spécifiquement gaussien, dont la première version est dûe à PLackett (1954, [139]) et Slepian (1962, [160]) et l'une des dernières versions à Li & Shao (2002, [98]), est de comparer les distributions respectives du supremum de deux processus gaussiens standards en comparant leurs fonctions de covariance respectives (les extensions à cette méthode ayant pour but de définir les bornes de la différence entre ces deux distributions). Quant à la méthode de Rice, qui peut être généralisée au cas non gaussien, elle consiste à utiliser les deux premiers moments du nombre de franchissements par un processus pour estimer la probabilité de dépassement d'un niveau donné par une trajectoire du processus.

▷ **Franchissements et temps local.**

⋆ Nous nous sommes intéressés dans un premier temps à la représentation dans le chaos de Wiener du nombre de franchissements de x par X, afin de rendre plus aisée l'obtention des théorèmes limites.

Rappelons que, si W désigne le mouvement brownien ou processus de Wiener standard, H_n le polynôme d'Hermite d'ordre n, et $H(X)$ l'espace des fonctionnelles réelles de X de carré intégrable, alors $H(X) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$, où $\mathcal{H}_n$ est le sous espace linéaire fermé de $L^2(\Omega, \mathcal{F}, P)$ généré par les variables aléatoires $\{H_n(W(h)), h \in L^2(\mathbb{R}, dx), ||h|| = 1\}$, appelé chaos de Wiener, où $W(h)$ désigne l'intégrale stochastique de h par rapport à W et $\mathcal{H}_0$ est l'ensemble des fonctions constantes réelles. Les intégrales multiples de Wiener-Itô (notées MWI) peuvent être également utilisées et elles sont liées aux polynômes d'Hermite par la relation

$$H_n(W(h)) = I_n(h^{\otimes n}) \qquad \forall h \in L^2(\mathbb{R}, dx), \; ||h|| = 1.$$

Polynômes d'Hermite ou MWI constituent un outil fort utile pour l'étude de fonctionnelles non linéaires des processus gaussiens stationnaires.

Si X possède une mesure spectrale absolument continue et une fonction de corrélation r deux fois dérivable et si $var\,(N_t(\psi)) < \infty$, en utilisant le calcul stochastique de Wiener-Itô, Slud a montré en 1991 ([161]) et en 1994 ([162] ou [163]) que

$$N_t(\psi(y)) = I\!E[N_t(\psi(y))] + \sum_{n=1}^{\infty} I_n(F_n) \quad dans \ \ L^2(\Omega), \quad où$$

— *dans le cas d'un niveau donné x (i.e. $\psi(y) = x, \ \forall y$), la moyenne de $N_t(x)$ est donnée par la formule de Rice et*

$$F_n(\underline{\lambda}) = \frac{e^{-\frac{x^2}{2}}}{\pi} \int_0^t e^{is(\lambda_1 + \cdots + \lambda_n)} \times$$

$$\sum_{l=0}^{[\frac{n}{2}]} (-r''(0))^{\frac{1}{2}-l} H_{n-2l}(x) \frac{(-1)^{l+1} H_{2l}(0)}{(2l)!(2l-1)} \sum_{1 \le m_1 < \cdots < m_{2l} \le n} \lambda_{m_1} \cdots \lambda_{m_{2l}} ds,$$

— *et dans le cas d'une courbe $\mathcal{C}^1$, la moyenne de $N_t(\psi)$ est donnée par la formule de Rice généralisée, et*

$$F_n(\underline{\lambda}) = \int_0^t e^{is(\lambda_1 + \cdots + \lambda_n)} \frac{e^{-\frac{u^2}{2}}}{\pi} \left(\sqrt{-r''(0)} H_n(u) - \right.$$

$$\left. \sum_{l=1}^{n} \frac{i^l}{l!} H_{n-l}(u) \sum_{1 \le m_1 < \cdots < m_l \le n} \lambda_{m_1} \cdots \lambda_{m_l} \int_0^{(-r''(0))^{-\frac{1}{2}}} e^{-\frac{z^2 y^2}{2}} H_l(-zy) y^{l-2} dy \right)_{u=\psi(s), z=\psi'(s)} ds,$$

où $\underline{\lambda}$ est un vecteur de composantes $(\lambda_i, 1 \le i \le n)$.

Nous avons introduit avec J. León ([**8**]) une nouvelle méthode ayant l'avantage d'être directe et beaucoup plus simple que celle développée par Slud, pour obtenir une représentation, elle même plus simple, du nombre de franchissements de niveau donné en polynômes d'Hermite.

Le point de départ de cette méthode est la formule heuristique du nombre de franchissements :

$$N_t(x) = \int_0^t \delta_x(X_s)|\dot{X}_s|ds$$

rendue exacte par approximation de la 'fonction de Dirac' $\delta_x(u) = \infty$ si $u = x$ et 0 sinon. L'astuce ensuite a été d'utiliser l'indépendance de X_s et de $\dot{X}_s$ à s fixé et de développer $|\dot{X}_s|$ en polynômes d'Hermite en $\dot{X}_s$ plutôt qu'en X_s, comme l'avait fait Slud.

Mentionnons alors l'expression obtenue sous la condition de Geman et la condition additionnelle $|r''(t+h) - r''(t)| \le |h|L_1(h)$, $L_1(h)$ étant une fonction paire appartenant à $L^1([0, \delta], dx)$.

Résultat principal.

$$\frac{N_t(x)}{\sqrt{-r''(0)}} = \sum_{q=0}^{\infty} \sum_{l=0}^{[\frac{q}{2}]} b_{q-2l}(x) a_{2l} \int_0^t H_{q-2l}(X_s) H_{2l}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right) ds,$$

où $(a_k)_{k \geq 0}$ et $(b_k)_{k \geq 0}$ sont les coefficients d'Hermite respectivement de la fonction $|.|$ et de la densité gaussienne standard.

$\star$ Revenons maintenant à la formule heuristique du nombre de franchissements.

Supposons qu'il y ait sur $[0, t]$ un nombre fini de points $0 = \alpha_0 < \alpha_1 < .... < \alpha_{k-1} < \alpha_k = t$ où la dérivée de X change de signe.

Alors dire que X franchit le niveau x se traduit par $|\mathbb{1}_{(X_{\alpha_{i+1}} > x)} - \mathbb{1}_{(X_{\alpha_i} > x)}| = 1$,

ce qui implique $N_t(x) = \sum_{i=0}^{k-1} |\mathbb{1}_{(X_{\alpha_{i+1}} > x)} - \mathbb{1}_{(X_{\alpha_i} > x)}| = \sum_{i=0}^{k-1} |Y_x(X_{\alpha_{i+1}}) - Y_x(X_{\alpha_i})|$,

où $Y_x(u) = \mathbb{1}_{[x,\infty)}(u)$ est la fonction Heaviside dont la dérivée généralisée est la 'fonction de Dirac' .

Ainsi, formellement $N_t(x) = \sum_{i=0}^{k-1} \int_{\alpha_i}^{\alpha_{i+1}} \delta_x(X_s)|\dot{X}_s|ds$, soit $N_t(x) = \int_0^t \delta_x(X_s)|\dot{X}_s|ds$.

Notons que cette formule heuristique permet de retrouver directement des résultats classiques, tels

- la formule de Rice : $\mathbb{E}[N_t(x)] = \int_0^t \mathbb{E}[\delta_x(X_s)]\mathbb{E} \mid \dot{X}_s \mid ds = \; te^{-x^2/2} \; \sqrt{-r''(0)} \; /\pi$;

- plus généralement, si g est une fonction positive sur $\mathbb{R}$, et G l'une de ses primitives, alors, comme remarqué par Cabaña (1985, [34]) :

$$\int_{\mathbb{R}} N_t(x)g(x)dx = \int_0^t \left[\int_{\mathbb{R}} \delta_{X_s}(x)g(x)dx \right] \mid \dot{X}_s \mid ds \;\; = \;\; \int_0^t g(X_s) \mid \dot{X}_s \mid ds$$
$$= \;\; \sum_{i=0}^{k-1} \mid G(X_{\alpha_{i+1}}) - G(X_{\alpha_i}) \mid .$$

- la formule de Kac : puisque $\hat{\delta}_0(t) = 1$, alors $\delta_0(u) = \frac{1}{2\pi} \int_{\mathbb{R}} cos(tu)dt$ en appliquant de façon formelle l'inversion de Fourier, d'où

$$N_t(0) = \int_0^t \delta_0(X_s) \mid \dot{X}_s \mid ds = \frac{1}{2\pi} \int_{\mathbb{R}} \int_0^t cos(\xi X_s) \mid \dot{X}_s \mid d\xi ds.$$

Nous avons cherché à valider mathématiquement cette formule heuristique du nombre de franchissements (cf. [**14**]), en nous inspirant de l'étude menée pour le cas plus simple que constitue le temps local. Il existe, en effet, des analogies entre ces deux formules dans leurs présentations heuristiques et non heuristiques, comme le montre le tableau suivant :

Temps local L_t^x pour $\check{X} = (\check{X}_s, s \geq 0)$ (pour lequel L_t^x existe)	Nombre de franchissements N_t^x pour $X = (X_s, s \geq 0)$ (pour lequel N_t^x existe)
Temps de séjour de $\check{X}$ au niveau x dans $[0,t]$: $$S_x(t) = \int_0^t \mathbb{1}_{(\check{X}_s \geq x)} ds := \int_0^t Y_x(\check{X}_s) ds$$	
Temps local = densité de $S_x(t)$	Nombre de franchissements
Formellement	
$$L_t^x = \int_0^t \delta_x(\check{X}_s) ds$$	$$N_t^x = \int_0^t \delta_x(X_s)|\dot{X}_s| ds$$
Mathématiquement la fonction Heaviside Y_x et la "fonction de Dirac" δ_x sont approchées respectivement par la distribution et la densité d'une v.a. gaussienne de moyenne x et de variance σ^2, avec $\sigma \to 0$.	
De plus, $\forall f \in \mathcal{C}^\infty$,	
$$\int_0^t f(\check{X}_s) ds = \int_{-\infty}^\infty f(x) L_t^x dx$$	$$\int_0^t f(X_s)|\dot{X}_s| ds = \int_{-\infty}^\infty f(x) N_t(x) dx$$ formule de Banach-Kac (1925-1943)

Introduisons alors l'espace de Sobolev $I\!D^{2,\alpha}$, $\alpha \in I\!R$, à savoir :

$f \in H(X)$, telle que $f = I\!E[f] + \sum_{n=1}^\infty I_n(f_n) = \sum_{n=0}^\infty I_n(f_n)$, appartient à $I\!D^{2,\alpha}$ si et seulement

si $||f||_{2,\alpha}^2 = \sum_{n=1}^\infty (1+n)^\alpha ||f_n||_2^2 = ||(I-L)^{\alpha/2} f||_2^2 < \infty$, où L est l'opérateur défini sur $H(X)$

par $Lf = \sum_{n=0}^\infty -n I_n(f_n)$.

L coincide avec le générateur infinitésimal du semi-groupe d'Ornstein-Uhlenbeck $\{T_u, u \geq 0\}$ des opérateurs de contraction sur $H(X)$ définis par $T_u(f) = \sum_{n=1}^\infty e^{-nu} I_n(f_n)$, pour tout

$f = \sum_{n=0}^\infty I_n(f_n)$.

En particulier, nous avons pour $\beta > 0$, $(I-L)^{-\beta} = \dfrac{1}{\Gamma(\beta)} \int_0^\infty e^{-u} u^{\beta-1} T_u du$.

Il a été démontré que le temps local du mouvement brownien appartient à l'espace de Sobolev $I\!D^{p,\alpha}$, $\forall p \geq 2$ et $0 < \alpha < 1/2$ (Nualart & Vives (1992, [125] et [126]), Imkeller et al. (1995, [78])).

Grâce au calcul de Malliavin, nous avons obtenu le résultat suivant ([**14**]) :

Résultat. $\delta_x(X_s) = \lim\limits_{\varepsilon \to 0} p_\varepsilon(X_s - x)$ *existe dans* $\mathbb{D}^{2,\alpha}$ *pour tout* $\alpha < -1/2$ (p_ε *désignant la densité gaussienne centrée de variance* ε) *et s'exprime par* :

$$\delta_x(X_s) = \sum_{n=0}^{\infty} c_n(x) H_n(X_s) = \sum_{n=0}^{\infty} c_n(x) I_n(h_s^{\otimes n}), \ avec \ c_n(x) = \frac{1}{n!} H_n(x) p_1(x), \forall n \geq 0.$$

$|\ \dot{X}_s\ |$ *existe dans* $\mathbb{D}^{2,\alpha}$ *pour tout* $\alpha \in \mathbb{R}$ *et*

$$|\ \dot{X}_s\ | = \sum_{n=0}^{\infty} a_{2n} H_{2n}(\dot{X}_s) = \sum_{n=0}^{\infty} a_{2n} I_{2n}(\dot{h}_s^{\otimes 2n}), \ avec \ a_{2n} = \sqrt{\frac{2}{\pi}} \frac{(-1)^{n+1}}{2^n n! (2n-1)}, \forall n \geq 0.$$

Supposons F absolument continue, auquel cas $r(t) = \int_{-\infty}^{\infty} b(t+s) b(s) ds$ *avec* $b \in L^2(\mathbb{R})$. *Soit, pour δ appartenant au voisinage de 0, le moment d'ordre 2, indépendant en α, défini par* $I_{\delta,\varepsilon} := \mathbb{E}\left(\int_0^{\delta} (t-s) r(s) \mid \dot{X}_s \mid \frac{x - X_s}{\varepsilon} p_{\varepsilon,x}(X_s)\ ds \right)^2$.
Si $I_{\delta,\varepsilon}$ converge uniformément en $\varepsilon > 0$, alors, sous la condition de Geman,

$$
\begin{aligned}
N_t(x) &= \sum_{n=0}^{\infty} \sum_{m=0}^{[n/2]} a_{2m} c_{n-2m}(x) \int_0^t H_{n-2m}(X_s) H_{2m}(\dot{X}_s) ds \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{[n/2]} a_{2m} c_{n-2m}(x) \int_0^t I_{n-2m}(h_s^{\otimes n-2m}) I_{2m}(\dot{h}_s^{\otimes 2m}) ds,
\end{aligned}
$$

appartient à l'espace de Sobolev $\mathbb{D}^{2,\alpha}$ *pour tout* $0 < \alpha < 1/4$.

Comme pour le temps local, l'intégrale intervenant dans la définition du nombre de franchissements aurait une conséquence sur l'ordre de l'espace de Sobolev.
La démonstration revient à prouver que $(I-L)^{\frac{\alpha-1}{2}} D_\theta N_t^\varepsilon(x)$ est borné dans $L^2(\Omega, L^2(\mathbb{R}, dx))$, uniformément en $\varepsilon > 0$, où D_θ est l'opérateur dérivé, N_t^ε est l'approximation du nombre de franchissements obtenue en remplaçant la fonction de Dirac par la densité gaussienne appropriée, et $\theta \in \mathbb{R}$.
Nous avons montré que $||(I-L)^{\frac{\alpha-1}{2}} D_\theta N_t^\varepsilon(x)||_{L^2(\Omega, L^2(\mathbb{R}, dx))}^2 \leq \int_{\mathbb{R}} J_{\alpha,\varepsilon}(\theta) d\theta + I_{t,\alpha,\varepsilon} + I_{\delta,\varepsilon}$,

avec $\delta > 0$, $\int_{\mathbb{R}} J_{\alpha,\varepsilon}(\theta) d\theta < \infty$ uniformément en $\varepsilon > 0$, ceci pour tout $\alpha < 1/2$, et $I_{t,\alpha,\varepsilon} < \infty$ uniformément en $\varepsilon > 0$, ceci pour tout $\alpha < 1/4$.
Il reste à chercher sous quelles conditions raisonnables la convergence uniforme en ε de $I_{\delta,\varepsilon}$ peut avoir lieu.
Notons que nous avons appliqué l'inégalité de Cauchy-Schwarz pour simplifier l'étude de $I_{\delta,\varepsilon}$ (cf. chapitre 1). Aussi, si cette étude mène à une impasse, nous serons amenés à revenir au pas antérieur, i.e. à discuter selon α de la convergence uniforme en ε de $\int_0^{\delta} (t-s) r(s) \mathbb{E}\left[g_\varepsilon(s) (I-L)^{\alpha-1}(g_\varepsilon(0)) \right] ds$, où $g_\varepsilon(s) := \mid \dot{X}_s \mid \frac{x - X_s}{\varepsilon} p_{\varepsilon,x}(X_s)$.

⋆ D'autres travaux abordent le problème de l'approximation du temps local d'un processus irrégulier X par le nombre de franchissements de la régularisation de ce processus. Wschebor (1984 [179], 1992 [181]) a montré que le nombre de franchissements d'un niveau x sur $[0, t]$ par le mouvement brownien régularisé par convolution converge, à un facteur d'échelle près, dans L^p vers le temps local au niveau x du mouvement brownien sur cet intervalle, ceci pour tout $p \in I\!N$. Citons aussi les travaux de Azaïs, Florens-Zmirou (1987 [8], 1990 [7]), Berzin, León, Ortega (1992 [97], 1998 [25]), Azaís et Wschebor (1996, [11]).

▷ **Comportements asymptotiques.**

Nous savons caractériser deux ou trois types de comportements asymptotiques du nombre de franchissements ou de dépassements selon les hypothèses faîtes sur X.

⋆ Sous des propriétés de mélangeance assurant que les dépassements sur des intervalles suffisamment séparés peuvent être considérés comme asymptotiquement indépendants, le processus du nombre de dépassements d'un niveau a un **comportement poissonnien**. Donnons une version de ce résultat proposée par Leadbetter, Lindgren et Rootzen (1983, [94]) :

Supposons que la corrélation r du processus stationnaire gaussien X satisfait au voisinage de 0 : $r(s) = 1 - \dfrac{\lambda_2}{2}s^2 + o(s^2)$ et en l'infini la condition de Berman : $r(s) \log s \underset{s \to \infty}{\to} 0$. Supposons que le niveau $x = x(t)$ et t tendent vers l'infini tels que $I\!EN_x^+(t) = t\mu(x) \underset{t \to \infty}{\to} \tau > 0$. Alors le processus ponctuel des dépassements normalisé (i.e. ayant des points en s/t lorsque X_s dépasse x en s) converge quand $t \to \infty$ vers un processus de Poisson d'intensité τ.

Il est utile d'avoir une idée de la vitesse de convergence en vue des applications. Pickersgill et Piterbarg ([134]) établirent en 1987 une vitesse de convergence de l'ordre de $t^{-\nu}$, mais sans aucune information sur la taille de ν. Dix ans plus tard, Kratz et Rootzén ([**7**]) calculèrent des bornes des moments du nombre de dépassements et identifièrent la vitesse de convergence de l'ordre de $t^{-\delta}$, avec $\delta = \dfrac{1}{2} \wedge \inf_{s \geq 0} \rho(s)$ où $\rho(s) = \dfrac{(1 - r(s))^2}{1 - r^2(s) + r'(s)|r'(s)|}$, en procédant par discrétisation et en combinant la méthode de comparaison normale avec la méthode de Stein-Chen. La méthode de comparaison normale donne des bornes de la distance entre deux lois gaussiennes, en fonction de leurs covariances respectives, et la méthode de Stein-Chen permet d'approcher la loi d'une v.a.r. W par le calcul de $I\!E[\mathbb{1}_{W \leq y}]$ et utilise une caractérisation de la loi de Poisson, à savoir, une v.a. Z à valeurs entières suit une loi de Poisson de paramètre λ si et seulement si pour toute fonction bornée f, $I\!E[\lambda f(Z + 1) - Z f(Z)] = 0$; les grandes lignes de ces méthodes sont données dans le travail de synthèse.

Résultat principal. *Suposons que r satisfait $r(s) = 1 - \dfrac{\lambda_2}{2}s^2 + o(s^2)$ au voisinage de 0, la condition de Berman $r(s) \log s \underset{s \to \infty}{\to} 0$ en l'infini et que r et sa dérivée décroissent de façon polynomiale d'ordre $\alpha > 2$. Alors il existe une constante C (explicitée dans [**7**]) dépendant de $r(s)$ mais non de x ou de t telle que pour $t \geq t_0 > 1$, la distance en variation totale entre $N_x^+(t)$ et la loi de Poisson d'intensité $\mu(x)$ satisfaît*

$$d(N_x^+(t), \mathcal{P}(t\mu(x))) \leq C \frac{(\log t)^{1+1/\alpha}}{t^\delta}.$$

⋆ Dès les années 70, Malevitch ([108]), Cuzick ([43]) obtinrent un théorème limite centrale (TLC) pour le nombre de zéros de X sous certaines conditions sur la covariance de X. L'idée de Malevitch, reprise par Cuzick tout en affaiblissant les conditions, était d'approcher X par un processus m-dépendant, afin d'utiliser le TLC pour les processus m-dépendants de Hoeffding et Robbins (1948, [67]). En 1978, Piterbarg ([132]) généralisa ce résultat pour tout niveau x donné (avec des conditions de Cuzick modifiées), en combinant la méthode de comparaison normale à celle de discrétisation. Berman (1971-1992, [22]) travailla également sur les TLC pour différents types de conditions de mélangeance, non pas pour le nombre de dépassements d'un niveau par X, mais pour une notion qui lui est liée, celle du temps passé au-dessus d'un niveau, appelée encore temps de séjour de X au-dessus de x. Dans les années 90, Slud ([163]) appliqua un TLC fonctionnel obtenu précédemment avec Chambers ([36]) par la théorie des MWI, pour avoir un TLC pour le nombre de zéros de X, ceci sans les conditions suffisantes de Cuzick assurant la non-nullité de la variance asymptotique. Puis il généralisa ce TLC au cas d'un niveau quelconque et d'une courbe de dépassement périodique pour un processus de corrélation décroissant rapidement vers 0.

Kratz et León ([**10**]) ont proposé en 2001 une méthode générale ayant le grand avantage d'être applicable à des processus multidimensionnels et permettant d'obtenir un TLC pour toute fonctionnelle de niveau de processus gaussiens. Les techniques utilisées combinent deux approches, celle développée par ces auteurs en 1997 ([**8**]) utilisant le développement en polynômes d'Hermite, et la méthode de m-dépendance.

Résultat principal. *Supposons que la fonction de corrélation r du processus X stationnaire gaussien standard satisfait $r \in L^1$ et $r^{(iv)} \in L^2$. Pour s fixé, soit Z_s la v.a. indépendante de X_s et $\dot{X}_s$ telle que $\dfrac{\ddot{X}_s}{\sqrt{r^{(iv)}(0)}} = \rho_1 X_s + \rho_2 Z_s$*

avec $\rho_1 = \dfrac{r''(0)}{\sqrt{r^{(iv)}(0)}}, \rho_2 = \sqrt{1 - \rho_1^2}$.

Soit F_t^X le développement d'Hermite de $H(X)$ donnée par

$$F_t^X = \sum_{q=0}^{\infty} \sum_{0 \leq n+m \leq q} d_{qnm} \int_0^t H_n(X_s) H_{q-(n+m)} \left(\frac{\dot{X}_s}{\sqrt{-r''(0)}} \right) H_m(Z_s) ds,$$

avec d_{qnm} tel que $\forall q \geq 0, \displaystyle\sum_{0 \leq n+m \leq q} d_{qnm}^2 n! m! (q - (n+m))! < C(q)$, et $(C(q))_q$ une suite bornée. Alors

$$\frac{F_t^X - \mathbb{E}[F_t^X]}{\sqrt{t}} \xrightarrow[t \to \infty]{} \mathcal{N}(0, \sigma^2),$$

où $\mathbb{E}[F_t^X] = t d_{000}$ et $\sigma^2 = \displaystyle\sum_{q=1}^{\infty} \sigma^2(q) < \infty$, avec

$$\sigma^2(q) = \sum_{0 \leq n_1+m_1 \leq q} \sum_{0 \leq n_2+m_2 \leq q} d_{qn_1m_1} d_{qn_2m_2} \int_0^{\infty} \mathbb{E}[I_{qn_1m_1}(0) I_{qn_2m_2}(s)] ds,$$

$$et \quad I_{qnm}(s) = H_n(X_s) H_{q-(n+m)} \left(\frac{\dot{X}_s}{\sqrt{-r''(0)}} \right) H_m(Z_s).$$

10

Les conditions portant sur r et ses dérivées s'affaiblissent selon la fonctionnelle considérée et se réduisent à $r \in L^1$ et $r'' \in L^2$, si seuls X et sa dérivée interviennent, ou encore à $r \in L^1$ si seul X intervient. Ce théorème permet de retrouver en particulier le TLC de Slud pour $N_t(x)$, à savoir

$$\sqrt{t}\left(\frac{N_t(x)}{t} - \frac{\sqrt{-r''(0)}}{\pi}e^{-x^2/2}\right) \xrightarrow[t\to\infty]{} \mathcal{N}(0, -r''(0)\sigma^2),$$

avec $\sigma^2 < \infty$ donné par

$$\sigma^2 = \sum_{q=1}^{\infty} \sigma^2(q) = \sum_{q=1}^{\infty} \sum_{n_1=0}^{[q/2]} \sum_{n_2=0}^{[q/2]} b_{q-2n_1}(x) a_{2n_1} b_{q-2n_2}(x) a_{2n_2} \int_0^{\infty} E\left[H_{2n_1}(X_0)H_{q-2n_1}\left(\frac{\dot{X}_0}{\sqrt{-r''(0)}}\right)\right.$$
$$\left. H_{2n_2}(X_s)H_{q-2n_2}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right)\right] ds.$$

Il est généralement difficile de vérifier que la variance asymptotique est non nulle. Nous avons pu le faire en démontrant que le déterminant de la matrice associée à $\sigma^2(2)$ (exprimée dans le domaine des fréquences) est strictement positif.

$\star$ Remarquons pour conclure qu'il est également possible de se placer sous une condition de longue dépendance régulière : $r(s) = (1 + |s|)^{-\alpha}L(s)$, où L est une fonction à variation lente à l'infini et $0 < \alpha < 1/2$, et d'obtenir un NCLT (non central limit theorem) pour le nombre de franchissements ; c'est ce qu'a obtenu Slud (1994, [163]) par des techniques propres aux MWI, en particulier via le NCLT de Major ([106]) pour des champs gaussiens stationnaires à longue dépendance régulière.

- Nous présentons ensuite de façon plus succinte, en se basant principalement sur les contributions [8], [9] et [10], **quelques extensions** de cette étude en abordant d'autres fonctionnelles non linéaires d'un processus gaussien, ainsi que le cas bidimensionnel en s'intéressant à la longueur des courbes de niveau de champs aléatoires gaussiens.

  $\star$ L'application du TLC de toute fonctionnelle de niveau de processus gaussiens du type F_t^X énoncé ci-dessus permet d'obtenir un TLC d'une part pour des fonctionnelles de temps d'occupation du processus stationnaire gaussien X de corrélation $r \in L^1$, telles le temps de séjour de X au dessus d'un niveau x sur l'intervalle $[0, t]$, le temps local L_t^x (quand il existe) (ces TLC sont exposés dans la synthèse), d'autre part pour le nombre de maxima dans un intervalle avec, en vue, l'application à l'étude des mers aléatoires ; nous pouvons en effet obtenir le comportement asymptotique, par exemple, de l'amplitude des vagues, ou de l'amplitude associée à l'accélération, ou encore du nombre de maxima locaux positifs de vagues ([9]). Citons par exemple le résultat lié au nombre de maxima locaux positifs de vagues.

  Résultat. *Soit X processus gaussien stationnaire de corrélation r et soit $M_{[a,b]}^X$ le nombre de maxima locaux de X_s, $0 \leq s \leq t$ sur l'intervalle $[a, b]$. Sous certaines hypothèses sur r (cf. [9]), nous avons*

$$\sqrt{T}\left(\frac{M_{[\beta,+\infty)}^X}{M_{[0,+\infty)}^X} - \frac{E[M_{[\beta,+\infty)}^X]}{E[M_{[0,+\infty)}^X]}\right) \xrightarrow{d} N(0, \sigma^2) \quad quand \ T \longrightarrow \infty,$$

11

*où σ^2 est donné explicitement (cf. chapitre 1, (2.63), Corollaire 1.2.1) et pour lequel un estimateur consistant a été proposé dans [**9**].*

$\star$ La méthode mise en place pour établir le TLC général de toute fonctionnelle de niveau de processus gaussiens est aisément adaptable à une dimension supérieure ; elle a permis d'obtenir en dimension 2, le comportement asymptotique de la longueur de courbe d'un champ gaussien standard stationnaire $X := (X_{s,t}; (s,t) \in I\!\!R^2)$ sous certaines conditions. Plus précisément, soit $\mathcal{L}_{Q(T)}^X(u)$ la longueur de la courbe de niveau u, $\{(s,t) \in Q(T) \ : \ X_{s,t} = u\}$, où $Q(T)$ désigne le carré $[-T,T] \times [-T,T]$. Supposons X isotropique, de fonction de corrélation r ayant des dérivées partielles $\partial_{ij} r$, pour $1 \le i+j \le 2$, et telle que $r \in L^1$, $r^2(0,0) - r^2(s,t) \neq 0$, for $(s,t) \neq (0,0)$, et $\partial_{02} r, \partial_{20} r \in L^2$.

Rappelons le théorème de Federer([57]), à savoir, pour $g \in \mathbf{C}(I\!\!R)$,

$$\int_{-\infty}^{\infty} g(u)\mathcal{L}_{Q(T)}^X(u)du = \int\int_{Q(T)} g(X_{s,t})||\nabla X_{s,t}||dsdt.$$

Nous en avons déduit la représentation dans le chaos de Wiener de $\mathcal{L}_{Q(T)}^X(u)$, dont nous avons étudié le comportement asymptotique (cf. [**10**]).

Résultat principal (TCL).

$$\frac{\mathcal{L}_{Q(T)}^X(u) - I\!\!E[\mathcal{L}_{Q(T)}^X(u)]}{|Q(T)|^{1/2}} \xrightarrow[T\to\infty]{} \mathcal{N}(0, \sigma^2),$$

avec $I\!\!E[\mathcal{L}_{Q(T)}^X(u)] = |Q(T)|\dfrac{\sqrt{2\pi}\phi(u)}{B(1,1/2)}$ *et*

$$\sigma^2 = \sum_{q=0}^{\infty} \sum_{0 \le m_1+l_1 \le [q/2]} \sum_{0 \le m_2+l_2 \le [q/2]} \delta_{q,2m_1,2l_1}(u)\delta_{q,2m_2,2l_2}(u) \int\int_{I\!\!R^2} I\!\!E[I_{q,2m_1,2l_1}(0,0)I_{q,2m_2,2l_2}(s,t)]dsdt,$$

où $I_{q,2m,2l}(s,t) := H_{q-2(m+l)}(X_{s,t})H_{2m}(\partial_{10}X_{s,t})H_{2l}(\partial_{01}X_{s,t})$ *et*
$\delta_{q,2m,2l}(u) := d_{q-2(m+l)}(u)c_{2m,2l}$, (c_k) *et* (d_k) *étant les coefficients de Hermite des fonctions* $||\nabla X_{s,t}||$ *et* ϕ *(densité normale centrée réduite) respectivement.*

Chapitre 2 - Quelques problèmes de statistique et de mécanique statistique en temps discret.

- Avant de présenter les études figurant dans le second chapitre de la synthèse, parlons brièvement de l'étude du processus des excédences.

  Soient $(X_i)_{i \in I\!\!N^*}$ une suite de variables aléatoires de même fonction de répartition F, et $(u_n)_{n \in I\!\!N^*}$ une suite numérique.

  De nombreux auteurs se sont intéressés au comportement asymptotique du processus des excédences $(I\!\!I_{(X_1 > u_n)}, \cdots, I\!\!I_{(X_n > u_n)}, \cdots)$, en particulier à l'approximation poissonnienne du processus ponctuel $N_n = \sum_{j=1}^{n} I\!\!I_{(X_j > u_n)}$, sous diverses hypothèses portant sur $(X_i, i \in I\!\!N^*)$ ou sur $(u_n)_n$. La vitesse de convergence de (N_n) vers le processus de Poisson limite a été calculée dans de nombreux cas.

⋆ Dans le cas de variables i.i.d., une application du théorème 6.1 de Deheuvels et Pfeifer ([51], 1988) permet de mesurer la distance (en variation totale) entre la loi du processus ponctuel N_n et la loi du processus de Poisson Π_n d'intensité $nF(. \cap]u_n; \infty))$, lorsque $n \to \infty$:

$$d_{tv}\left(\mathcal{L}(N_n), \mathcal{L}(\Pi_n)\right) \ \sim \ \begin{cases} \frac{1-F(u_n)}{\sqrt{2\pi e}} & \text{si } \begin{cases} n(1 - F(u_n)) \to \infty \\ 1 - F(u_n) < 1 - \varepsilon, \text{ avec } \varepsilon > 0 \text{ fixé} \end{cases} \\ n(1 - F(u_n))^2 & \text{si } \ n(1 - F(u_n)) \to 0. \end{cases}$$

Notons également que dans le cas de v.a. iid, on peut se ramener au cas du processus empirique uniforme, sur lequel beaucoup de travaux ont été effectués.

Citons par exemple trois résultats concernant l'approximation en loi du processus empirique uniforme $(nF_n(s), 0 \le s \le a_n)$ par une suite $(\Pi_n(s), 0 \le s \le a_n)$ de processus de Poisson homogènes d'intensité $n \ge 1$, où la suite (a_n) satisfaît $0 < a_n < 1$ et $a_n \underset{n\to\infty}{\to} 0$

(et $F_n(s) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{(U_j \le s)}$, avec $(U_j)_{j\ge 1}$ v.a. iid uniformes sur $(0,1)$).

Des bornes supérieures asymptotiquement optimales (dépendant de a_n) ont été obtenues par Deheuvels et Pfeifer ([51], 1988) pour

$$\sup_A |I\!P[\{nF_n(a_n t), 0 \le t \le 1\} \in A] - I\!P[\{n\Pi_n(a_n t), 0 \le t \le 1\} \in A]|$$

(avec A partie borélienne de l'ensemble $I(0,1)$ des fonction monotones non décroissantes continues à droite définies sur $[0,1]$, muni de la topologie de la convergence uniforme).

Major ([107], 1990) a montré que, sous la condition supplémentaire $\sqrt{n} \, a_n \underset{n\to\infty}{\to} 0$, il existe un espace de probabilité sur lequel F_n et Π_n sont définis simultanément de telle manière que $I\!P[\sup_{0 \le s \le a_n} |n(F_n(s) - s) - (\Pi_n(s) - ns)| = 0] \underset{n\to\infty}{\to} 1$.

Kratz ([2], 1993) s'est intéressée à l'approximation poissonnienne relative du processus empirique et a obtenu un encadrement de $\sup_A \dfrac{I\!P[\{nF_n(a_n t), 0 \le t \le 1\} \in A]}{I\!P[\{n\Pi_n(a_n t), 0 \le t \le 1\} \in A]}$.

⋆ Dans le cas d'une suite stationnaire, certaines conditions de dépendance, moins fortes que celles de mélangeance forte, ont été définies par Leadbetter : la condition $D(u_n)$ pour rangs éloignés, et la condition locale $D'(u_n)$ qui exclut la possibilité de groupes d'excédences sur un petit intervalle (cf. Leadbetter et al., [94]). Sous ces conditions et pour $(u_n)_n$ telle que $n(1 - F(u_n)) = nI\!P[X_1 > u_n] \underset{n\to\infty}{\to} \tau \ge 0$, Leadbetter (1974, 1983 [94]) a démontré que le processus ponctuel défini sur $(0,1]$ et formé des points (j/n) tels que $X_j > u_n$, converge en loi vers un processus de Poisson Π sur $[0, \infty)$ de paramètre τ. Dans le cas gaussien, les conditions $D(u_n)$ et $D'(u_n)$ peuvent être remplacées par la condition de Berman $r_k \log k \underset{k\to\infty}{\to} 0$ et la condition $\tau_n := n(1 - \Phi(u_n))$ bornée, Φ désignant la fonction de répartition normale.

Holst et Janson (1990, [68]) ont calculé la vitesse de convergence en utilisant la méthode de Stein-Chen et ont obtenu le résultat suivant :

Soit $(X_i)_i$ suite stationnaire de va gaussiennes standard, de corrélations $(r_k)_k$ telles que $r_k \le \dfrac{A}{\log k}$, avec $k \ge 2$ et A constante. Soit $\rho := max(0, r_1, r_2, \cdots)$.

Supposons $\tau_n := n(1 - \Phi(u_n)) \leq B < \infty$. *Alors, quand* $n \to \infty$,

$$d_{tv}\left(\mathcal{L}(N_n), \mathcal{P}(\tau_n)\right) = O\left(n^{-\frac{1-\rho}{1+\rho}}(\log n)^{-\frac{\rho}{1+\rho}} + \frac{\log n}{n}\sum_{k=1}^{n}|r_k|\right),$$

où $\mathcal{P}(\tau_n)$ *désigne la loi de Poisson de paramètre* τ_n.

En particulier, on retrouve que si $r_k \log k \underset{k\to\infty}{\to} 0$ et si $\tau_n \underset{n\to\infty}{\to} \tau$, alors $N_n \overset{d}{\to} \mathcal{P}(\tau)$.

$\star$ Le cas non stationnaire a été considéré par Hüsler (1983, 1986 [74]) ; il a montré en remplaçant la condition locale de Leadbetter par une autre condition locale, que le processus N_n converge aussi en loi vers un processus de Poisson d'intensité τ telle que $\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}(1 - F_i(u_{ni})) = \tau > 0$ (où F_i est la fonction de répartition de X_i et (u_{ni}) sont les différents niveaux).

Dans le cas gaussien, la vitesse de convergence a été calculée par Hüsler et Kratz pour un niveau donné u_n ([**3**], 1994) puis généralisée à des niveaux distincts ([**4**], 1995). La vitesse de convergence dépend alors de la corrélation ayant la plus grande valeur et du niveau le plus bas.

Résultat. *Soit* $\{X_i, i \geq 1\}$ *suite gaussienne standard de corrélations* $\{r_{ij}, i, j \geq 1\}$ *telles que* $\begin{cases} \rho_n < 1, \ \forall n \geq 1 \\ \rho_n \log n \underset{n\to\infty}{\to} 0, \quad \text{(Berman's condition)}. \end{cases}$

Soient les niveaux $\{u_{ni}\}$ *tels que* $\begin{cases} \limsup\limits_{n\to\infty} \lambda_n < \infty, \ \text{avec} \ \lambda_n = \sum_{i=1}^{n} I\!\!P[X_i > u_{ni}], \\ u_{n,min} = \min\limits_{1\leq i\leq n} u_{ni} \underset{n\to\infty}{\to} \infty. \end{cases}$

Alors nous avons :

i) si $\rho > 0$, $d(\mathcal{L}(N_n), \mathcal{P}(\lambda_n)) = O\left(\dfrac{1}{u_{n,min}}\exp\left\{-\dfrac{1-\rho}{2(1+\rho)}\, u_{n,min}^2\right\} + \Delta(e^{u_{n,min}^2/\alpha})\right)$,

avec $\Delta(s) = \sup\{\rho_k \log k : k \geq s\}$ *et* $\alpha > 0$ *tel que* $\alpha > (1+\rho)/\rho$;

ii) si $\rho = 0$, $d(\mathcal{L}(N_n), \mathcal{P}(\lambda_n)) = O\left(u_{n,min}\exp\left\{-\dfrac{u_{n,min}^2}{2}\right\}\sum_{l\leq n}\rho_l\right)$.

Revenons maintenant aux deux principaux thèmes développés dans le second chapitre de la synthèse.

- **Modélisation par des lois à queue de distribution épaisse**.

  On dit qu'une variable aléatoire X a une distribution F à queue épaisse d'ordre α si $1 - F(x) = x^{-\alpha}L(x)$, quand $x \to \infty$, où L est une fonction à variation lente à l'infini,

  $$\text{i.e. } L \text{ satisfait pour tout } x \geq 1, \ \lim_{t\to\infty}\frac{L(tx)}{L(t)} = 1,$$

  ou encore si $1 - F$ est une fonction à variation régulière d'indice α (à l'infini),

  $$\text{i.e. pour tout } x > 0, \ \lim_{t\to\infty}\frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}. \tag{0.1}$$

Il s'agit d'un sujet d'actualité tant en théorie qu'en pratique (cf. [8] (section 2.1.3) et les références citées), en particulier en informatique ([22]) où de nombreux phénomènes

comme par exemple les temps d'attente entre les arrivées de paquets sur réseaux peuvent être modélisés par des lois à queue de distribution épaisse.

▷ Dans un premier temps nous nous sommes intéressés à la **détection des phénomènes dits à queue épaisse**. Pour cela, nous avons mis en place une méthode graphique permettant cette détection ainsi que l'estimation de l'épaisseur de la queue, i.e. du paramètre α intervenant dans la définition (0.1), à partir de données supposées indépendantes et de même loi. Cette étude, en collaboration avec Resnick (cf. [**6**]), avait pour but de proposer une alternative graphique à l'approche de Hill (1975, [18]) qui proposa un estimateur de l'indice de queue α défini par $\widehat{\alpha_H^{-1}} = \dfrac{1}{k_n} \sum\limits_{i=1}^{k_n} \log\left(\dfrac{X_{n+1-i,n}}{X_{n-k_n,n}}\right)$, basé sur les k_n statistiques d'ordre les plus élevées $X_{n-k_n+1,n} \leq \cdots \leq X_{n,n}$ d'un n-échantillon $(X_i)_{1\leq i \leq n}$ de distribution F. L'estimateur de Hill $\widehat{\alpha_H^{-1}}$ a de bonnes propriétés symptotiques lorsque (k_n) satisfaît les conditions

$$1 \leq k_n < n, \quad k_n \to \infty, \quad \frac{k_n}{n} \to 0, \quad \text{quand} \quad n \to \infty.$$

Mais la qualité de cet ajustement dépend fortement du choix de (k_n) pour lequel n'existe pas de critère automatique, sinon le bon sens.

Notre approche s'inspire de la méthode de l'ajustement par quantiles appelé encore QQ-plot, qui consiste à représenter les points $\{(Q(i/(n+1)), x_{i,n}), 1 \leq i \leq n\}$, où Q désigne la fonction quantile de la loi théorique G modélisant les données $(x_i, 1 \leq i \leq n)$ et $x_{i,n}$ est la $i$ème statistique d'ordre, i.e. le quantile associé de la fonction de répartition empirique (d'où le nom de QuantileQuantile-plot). Nous avons une idée visuelle de la qualité de l'ajustement en regardant si les points $(Q(i/(n+1)), x_{i,n})$ sont proches de la première bissectrice.

Dans le cas où la loi d'ajustement est de la forme $G_{\mu,\sigma}$ (μ, σ étant inconnus), par exemple $G_{\mu,\sigma} = G_{0,1}\left(\dfrac{x-\mu}{\sigma}\right)$, alors, la fonction quantile $Q_{\mu,\sigma}$ de la loi théorique $G_{\mu,\sigma}$ satisfaisant $Q_{\mu,\sigma} = \sigma Q_{0,1} + \mu$, le graphe des points $\{(Q_{0,1}(i/(n+1)), X_{i,n}), 1 \leq i \leq n\}$, doit être approximativement une droite affine de pente σ et d'ordonnée à l'origine μ, permettant de mesurer visuellement la qualité de l'ajustement et de proposer des estimateurs de σ et de μ par la méthode des moindres carrés. Dans le cas gaussien, la droite ainsi tracée est la droite de Henry.

Appliquons cette technique au cas de distributions à queues lourdes.

Intéressons-nous dans un premier temps au cas où nous proposons d'ajuster les données $(x_i)_{1\leq i \leq n}$ par une loi F_α de type Pareto de paramètre de forme $\alpha > 0$, i.e. satisfaisant $F_\alpha(x) = 1 - x^{-\alpha}$, $x \geq 1$.

Posons pour $y > 0$, $G_{0,\alpha}(y) := P[\log X_1 > y] = e^{-\alpha y}$. Alors le graphe

$$\{(Q_{0,1}(i/(n+1)), \log X_{i,n}), 1 \leq i \leq n\} = \{(-\log(1-i/(n+1)), \log X_{i,n}), 1 \leq i \leq n\}$$

devrait être approximativement une droite passant par l'origine de pente α^{-1}. Nous déduisons alors un estimateur de la pente par la méthode des moindres carrés, appelé qq-estimateur et défini par

$$\widehat{\alpha^{-1}} = \frac{\sum_{i=1}^{n} -\log(\frac{i}{n+1})\{n \log X_{n-i+1,n} - \sum_{j=1}^{n} \log X_{n-j+1,n}\}}{n \sum_{i=1}^{n}(-\log(\frac{i}{n+1}))^2 - (\sum_{i=1}^{n} -\log(\frac{i}{n+1}))^2}.$$

15

Considérons maintenant le cas plus général de lois F d'ajustement à queues lourdes telles que $1 - F$ soit à variation régulière d'ordre α.

Après avoir remarqué que pour t grand, $\dfrac{1 - F(tx)}{1 - F(t)} = I\!P\left[\dfrac{X_1}{t} > x \mid X_1 > t\right] \approx x^{-\alpha}$, alors

pour $k_n \to \infty$ tel que $k_n/n \to 0$, conditionnellement à $X_{n-k_n,n}$, $\left(\dfrac{X_{n-k_n+i,n}}{X_{n-k_n,n}}, i = 1, \ldots, k_n\right)$

se comporte comme les statistiques d'ordre d'un k_n-échantillon de loi concentrée sur $(1, \infty)$

et de queue satisfaisant $\dfrac{1 - F(X_{n-k_n,n}x)}{1 - F(X_{n-k_n,n})} \approx x^{-\alpha}$.

Nous pouvons donc définir comme précédemment un estimateur $\widehat{\alpha^{-1}}$ de la pente basé sur les k_n plus grandes statistiques d'ordre en utilisant la méthode des moindres carrés. Comment procéder en pratique ?

Tracer un QQ-plot à partir de toutes les données ; considérer k_n pour lequel le graphe apparaît linéaire ; puis calculer la pente par la méthode des moindres carrés à partir des k_n plus grandes statistiques d'ordre et les quantiles exponentiels correspondants.

De façon alternative, un graphe de $\{(k_n, \widehat{\alpha^{-1}}(k_n)), 1 \le k_n \le n\}$ peut être tracé pour ensuite repérer la vraie valeur de α^{-1} en regardant la région stable du graphe.

Nous pouvons alors démontrer les propriétés suivantes du qq-estimateur.

Résultat principal. *Soit $(X_1, \ldots, X_n)$ un n-échantillon de distribution F à variation régulière d'indice α. Alors le qq–estimateur $\widehat{\alpha^{-1}}$ défini par*

$$\widehat{\alpha^{-1}} = \frac{\displaystyle\sum_{i=1}^{k_n} -\log\left(\frac{i}{k_n+1}\right)\left(k_n \log\left(X_{n-i+1,n}\right) - \sum_{j=1}^{k_n} \log\left(X_{n-j+1,n}\right)\right)}{k_n \displaystyle\sum_{i=1}^{k_n}\left(-\log\left(\frac{i}{k_n+1}\right)\right)^2 - \left(\sum_{i=1}^{k_n} -\log\left(\frac{i}{k_n+1}\right)\right)^2}$$

est faiblement consistant pour $1/\alpha$, i.e. $\widehat{\alpha^{-1}} \overset{P}{\to} \alpha^{-1}$, quand $n \to \infty$, lorsque k_n satisfait $k_n \to \infty$ et $k_n/n \to 0$ quand $n \to \infty$.

Cet estimateur est également asymptotiquement normal, lorsqu'on ajoute sur F une hypothèse de variation régulière au second ordre et une condition sur le comportement asymptotique de k_n.

Soit U la fonction quantile de $1/(1 - F)$, i.e. la fonction positive définie sur $I\!R^{+*}$ par $U(t) = \left(\dfrac{1}{1 - F}\right)^{\leftarrow}(t) = \inf\{x \in I\!R : 1/(1 - F)(x) \ge t\}$, et soit $\gamma = \alpha^{-1}$. Supposons qu'il existe une fonction positive A tendant vers 0 en l'infini telle que pour tout $x > 1$,

$$\frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} \underset{t\to\infty}{\to} cx^\gamma\left(\frac{x^\rho - 1}{\rho}\right),$$

avec $c \in I\!R^*$ et $\rho \le 0$ (si $\rho = 0$, on pose $(x^\rho - 1)/\rho = \log x$).
Soit (k_n) une suite telle que, lorsque $n \to \infty$,

$$k_n \to \infty, \qquad k_n/n \to 0, \qquad \sqrt{k_n}\, A(n/k_n) \to 0.$$

Sous ces conditions, $\sqrt{k_n}\left(\widehat{\alpha^{-1}} - \alpha^{-1}\right) \overset{d}{\to} \mathcal{N}\left(0, 2\alpha^{-2}\right).$

Prenons l'exemple d'un n-échantillon de distribution F définie pour tout $x > 1$ par

$$1 - F(x) = x^{-\alpha} + cx^{-\beta}, \quad \text{avec } c > 0, \text{ et } \beta > \alpha > 0.$$

Alors $A(s) \underset{s \to \infty}{\sim} c\alpha^{-1}(\beta/\alpha - 1)s^{1-\beta/\alpha}$ et $\rho = 1 - \frac{\beta}{\alpha}$.

Il est naturel de comparer cet estimateur à l'estimateur de Hill. La variance asymptotique de l'estimateur de Hill, égale à α^{-2}, est meilleure que celle du qq-estimateur ; il faut cependant relativiser ce critère du fait que dans certaines circonstances l'estimateur de Hill peut présenter des biais considérables et que sur des exemples (cf. [6]) le qq-plot apparait plus robuste que le Hill-plot. Notons également que dans le cas de l'estimation par qq-plot, l'information contenue dans les résidus peut permettre de diminuer le biais des estimateurs dans le cas de queue lourde non Pareto, ce qui constitue l'un des avantages de cette méthode par rapport à celle de Hill.
Tout comme Csörgő, Deheuvels et Mason (1985, [6]) ont proposé une classe d'estimateurs à noyau contenant l'estimateur de Hill, signalons que Viharos (1999, [27]) a construit une classe d'estimateurs d'indice de queue contenant une forme asymptotique équivalente au qq-estimateur, ceci à partir de la méthode des moindres carrés pondérés.

▷ La seconde étude concerne la **modélisation de séries temporelles à innovations à queues de distribution épaisses**. Il y a essentiellement deux approches pour aborder cette étude, l'une basée sur les modèles structurels (prenant en compte les caractéristiques physiques de l'objet d'étude ; voir par exemple la discussion de Willinger et Paxon dans [22]) et l'autre, plus classique, basée sur l'analyse statistique des séries temporelles. Nous avons choisi cette seconde approche, en commençant naturellement par le cas des modèles linéaires tels les processus ARMA. Le travail que je présente s'inscrit dans un programme de recherche entrepris en 1992 par Feigin et Resnick ([10]) sur l'estimation des paramètres de processus ARMA à innovations positives et à queues épaisses (à gauche ou à droite). L'idée est de construire par des techniques de programmation linéaire des estimateurs ayant de meilleures vitesses de convergence que les estimateurs classiques de type Yule-Walker ou du maximum de vraisemblance lorsque nous considérons les hypothèses de queues épaisses. Alors que Feigin et Resnick ont commencé par traiter le cas autorégressif pur (1992, [10] ; 1994, [11]) et ont combiné cette méthode avec celle de bootstrap afin de rendre l'inférence possible (1997, [12]), nous nous sommes intéressés avec Resnick (cf. [5]) au cas moyenne mobile.
Soit le processus $X = (X_t, t \in \mathbb{Z})$ moyenne mobile d'ordre q défini par

$$X_t = Z_t + \sum_{i=1}^{q} \theta_i Z_{t-i} := \Theta(B)Z_t, \quad t \in \mathbb{Z},$$

où $\{Z_t\}$ est un bruit blanc fort, d'infimum (left endpoint) égal à 0, et de loi commune F. Les coefficients $\theta_1, \ldots, \theta_q$ sont supposés positifs ($\theta_i \geq 0, \forall 1 \leq i \leq q-1$ et $\theta_q > 0$) et satisfont la condition d'inversibilité (i.e. le polynôme moyenne mobile n'a pas de racines dans le disque unité $\{z : |z| \leq 1\}$).

Cette condition d'inversibilité permet de se ramener au cas d'un $AR(\infty)$ et de construire par programmation linéaire, au vu du cas $AR(p)$, une estimation des coefficients du polynôme moyenne mobile de X comme suit :

$$\hat{\boldsymbol{\theta}} = \arg \max_{D_n} \sum_{i=1}^{q} \eta_i$$

où l'ensemble des contraintes est donné par

$$D_n = \{\boldsymbol{\eta} \in \mathbb{R}_+^q : \sum_{k=0}^{2l}(I - \sum_{i=0}^{q} \eta_i B^i)^k X_t \geq 0, t = 2lq + 1, \cdots, n; \eta_0 = 1, \sum_{i=0}^{q} \eta_i z^i \neq 0, |z| \leq 1\}.$$

Cet estimateur a été appelé lp-estimateur. Se posait le problème du choix de l (indice de troncation de l'$AR(\infty)$) ; la théorie asymptotique nous a poussé à prendre le plus petit entier l tel que $2l \geq q$.

Nous avons également défini deux types de conditions de variation régulière et de conditions sur les moments pour la loi de la suite d'innovations, à savoir :

Condition L (left tail) : la distribution F des innovations Z_t satisfait

$$\exists \, \alpha > 0 \text{ tel que } \forall x > 0, \ \lim_{s \downarrow 0} \frac{F(sx)}{F(s)} = x^{\alpha};$$

$$\exists \, \beta > \alpha, \ E(Z_t^{\beta}) = \int_0^{\infty} u^{\beta} F(du) < \infty.$$

Cette condition est faible. Elle est satisfaite par exemple par les lois de Weibull, les lois gamma, et par toute fonction F ayant une densité f continue en 0 telle que $f(0) > 0$ (dans ce cas, $\alpha = 1$).

Condition R (right tail) : La distribution F des innovations Z_t satisfait

$$\exists \, \alpha > 0 \text{ tel que } \forall x > 0, \ \lim_{s \to \infty} \frac{1 - F(sx)}{1 - F(s)} = x^{-\alpha};$$

$$\exists \, \beta > \alpha, \ E(Z_t^{-\beta}) = \int_0^{\infty} u^{-\beta} F(du) < \infty.$$

Par exemple, les lois de Pareto et les lois stables à densité positive satisfont la condition R.

En utilisant des techniques propres aux processus ponctuels, nous avons obtenu :

Résultat principal.

• *Supposons l'ordre q du processus X égal à 1 et la distribution F des innovations absolument continue. Désignons par $\theta^{(0)} \in (0,1)$ la vraie valeur du coefficient du polynôme moyenne mobile de X, par $\{Y_{k,1}, Y_{k,2}, k \geq 1\}$ des variables aléatoires indépendantes de distribution commune F, et par $\Gamma_k = E_1 + \cdots + E_k, k \geq 1$, la somme de variables exponentielles standard i.i.d., indépendante de $\{(Y_{k,1}, Y_{k,2})\}$.*

Soient les fonctions quantiles $a(n)$ et $b(n) = b_n$ définies respectivement par

$$a_n = a(n) = F^{\leftarrow}(1/n) \text{ et } b_n = \left(\frac{1}{1-F}\right)^{\leftarrow}(n) = F^{\leftarrow}(1 - 1/n).$$

18

i) *Supposons la* **Condition L**.
La loi limite du lp-estimateur $\hat{\theta}$ défini précédemment est telle que sur $[0, \infty)$,

$$\frac{(\hat{\theta} - \theta^{(0)})}{a(\sqrt{n})} \xrightarrow{d} \frac{(\theta^{(0)})^{3/2\alpha}}{c(\alpha)^{1/2\alpha}} \bigwedge_{\substack{1 \leq k < \infty \\ Y_{k,1} > Y_{k,2}}} \frac{\Gamma_k^{1/2\alpha}}{|Y_{k,1} - (\theta^{(0)})^3 Y_{k,2}|},$$

où $c(\alpha)$ est une constante définie par l'intégrale Beta $c(\alpha) = \int_0^1 (1-s)^\alpha \alpha s^{\alpha-1} ds$.

En d'autres termes, la distribution limite de $\hat{\theta}$ est une loi de Weibull, de vitesse de convergence $1/a(\sqrt{n})$, i.e. de variation régulière d'ordre $\alpha/2$:

$$\lim_{n \to \infty} P[a(\sqrt{n})^{-1}(\hat{\theta} - \theta^{(0)}) \leq x] = 1 - \exp\{-kx^{2\alpha}\}, \quad x > 0,$$

*avec $k = (\theta^{(0)})^{-3\alpha} c(\alpha) E\left(|Y_{k,1} - (\theta^{(0)})^3 Y_{k,2}|^{2\alpha} \mathbb{1}_{(Y_{k,1} > Y_{k,2})}\right)$ constante finie sous la condition
L.*

ii) *Supposons la* **Condition R**.
La loi limite du lp-estimateur $\hat{\theta}$ défini précédemment est telle que sur $[0, \infty)$,

$$b_n(\hat{\theta} - \theta^{(0)}) \xrightarrow{d} \bigwedge_{k=1}^{\infty} \Gamma_k^{1/\alpha}(Y_{k,1} + (\theta^{(0)})^3 Y_{k,2}).$$

La distribution limite de $\hat{\theta}$ est une loi de Weibull, de vitesse de convergence b_n :

$$\lim_{n \to \infty} P[b_n(\hat{\theta} - \theta^{(0)}) \leq x] = 1 - \exp\{-cx^\alpha\}, \quad x > 0,$$

où $c = E|Y_{k,1} + (\theta^{(0)})^3 Y_{k,2}|^{-\alpha}$ constante finie sous la condition R.
• *Dans le cas général du $MA(q)$, $\theta^{(\mathbf{0})}$ désignant la vraie valeur du paramètre, la loi limite
du lp-estimateur $\hat{\theta}$ défini précédemment est telle que*

$$q_n(\hat{\theta} - \theta^{(0)}) = arg \max_{\Lambda_n} \delta' \mathbf{1}$$

où $\Lambda_n = \{\delta \in I\!R_+^q : 1 + \sum_{i=1}^{q} \left(\frac{\delta_i}{q_n} + \theta_i^{(0)}\right) z^i \neq 0, \ |z| \leq 1,$

$$\sum_{k=0}^{2l} (-1)^k \left(\sum_{i=1}^{q} \theta_i^{(0)} B^i + \frac{\delta(B)}{q_n}\right)^k X_t \geq 0, \ 2lq + 1 \leq t \leq n\},$$

et avec $q_n = a_n$ sous la condition L, respectivement $q_n = b_n$ sous la condition R.

Alors que les lp-estimateurs des coefficients autorégressifs dans le cas d'un $AR(p)$ ont une
vitesse de convergence ne dépendant pas de p, et à variation régulière d'indice α (cf. [10]
et [11]), il en est de même pour les lp-estimateurs des coefficients moyenne mobile dans
le cas d'un $MA(q)$ sous la condition R. Cela n'est pas le cas sous la condition L : leurs
vitesses de convergence se dégradent lorsque l'ordre q augmente.
Si les modèles linéaires se sont avérés insuffisamment pertinents pour le type de séries
considérées, la théorie et les outils développés pour leur étude ont servi de base pour les
recherches développées ces dernières années sur d'autres modélisations de séries tempo-
relles à queues épaisses, telles les modèles non linéaires, les modèles à régime markovien
caché, les modèles GARCH, etc ...

- **L'autre thème considéré dans ce chapitre** appartient à la mécanique statistique, plus particulièrement **à la théorie des systèmes de spins** et nous replace dans un contexte gaussien.

Le modèle inventé par Lenz, connu sous le nom de modèle d'Ising, est un modèle simplifié du système ferromagnétique, dans lequel les atomes sont placés sur les noeuds d'un maillage régulier de $\mathbb{Z}^d$ et sont représentés par des variables de spins les plus simples prenant deux valeurs ± 1, les spins n'interagissant entre eux que s'ils se trouvent être dans des noeuds voisins. Toutes ces interactions sont représentées par la fonction d'énergie ou Hamiltonien ; ainsi à une configuration donnée de spins correspond un Hamiltonien.

Une façon de simplifier le modèle d'Ising est d'une part, de remplacer la structure spatiale du maillage sur $\mathbb{Z}^d$ par un hypercube $S_N := \{-1, +1\}^N$ (avec $N \in \mathbb{N}$), où les sommets sont indexés par des entiers naturels et où tous les spins interagissent entre eux indépendamment de leurs distances, d'autre part de considérer l'Hamiltonien comme un processus gaussien indexé par l'espace des états S_N. Dans cette classe de modèles simplifiés figurent les systèmes de Derrida (datant des années 80, [6] et [7]) : modèles à énergie aléatoire (notés R.E.M.pour Random Energy Model) et leurs généralisations (notées G.R.E.M.), dont l'Hamiltonien est défini pour une configuration σ par

$$H(\sigma) = -\frac{N^{1/2}}{2^{N/2}} \sum_{\alpha \subset \{1,\dots,N\}} J_\alpha \sigma_\alpha,$$

où $\{J_\alpha, \alpha \subset \{1, \dots, N\}\}$ est une famille de v.a. gaussiennes réduites (définies sur un même espace de probabilité $(\Omega, \mathcal{F}, \mathcal{P})$) identiquement distribuées,
indépendantes dans le cas du R.E.M. et de corrélation positive dans le cas du G.R.E.M., où la somme est prise sur les 2^N sous-ensembles α de $\{1, \dots, N\}$,
et où $\sigma_\alpha = \prod_{i \in \alpha} \sigma_i$ (avec $\sigma_\emptyset = 1$), avec $\sigma_i = \pm 1, \forall i$.

Nous avons commencé à travailler dans la classe des R.E.M., modèle le plus désordonné étant donné que les interactions sont indépendantes.
Dans ce cas, $(H(\sigma), \sigma \in \{-1, +1\}^N) = (-\sqrt{N} X_i, 1 \le i \le 2^N)$, où les v.a. $(X_i)_{1 \le i \le 2^N}$ sont i.i.d. de loi normale centrée réduite (cf. [9]).
Nous avons étudié la **mesure de Gibbs** $\mu_{\mathbf{N}, \beta}$ associée au modèle, définie pour tout N comme la mesure de probabilité aléatoire sur l'espace d'états $S_N = \{-1, +1\}^N$, qui donne à une configuration σ le poids

$$\mu_{N,\beta}(\sigma) \equiv \frac{e^{-\beta H(\sigma)}}{Z_N}, \qquad \text{avec} \quad Z_N \equiv Z_N(\beta) = \sum_{\sigma} e^{-\beta H(\sigma)},$$

Z_N représentant la fonction de partition à volume fini, et β la température inverse ($\beta \ge 0$). Cette étude, developpée en collaboration avec P. Picco (cf. [**11**]), repose sur des techniques gaussiennes et des résultats que nous avons mis en place sur les statistiques d'ordre et sur les sommes d'espacements.
Pour entrer un peu plus dans ce travail, présentons quelques caractéristiques supplémentaires de ce modèle, telles :

⋆ la fonction d'énergie libre à volume fini, définie par $F_N(\beta) = -\dfrac{1}{\beta N} \log Z_N(\beta)$,

⋆ la température critique $\beta_c := \sqrt{2 \log 2}$, qui correspond au saut de la dérivée seconde de la fonction limite (non aléatoire) $F(\beta)$ de $F_N(\beta)$ quand $N \to \infty$ (la convergence ayant lieu p.s. et dans $L^p(\Omega, I\!P)$ pour $1 \le p < \infty$, pour tout $\beta \ge 0$) (cf. [16]) ;

⋆ notons également que pour $\beta > \beta_c$, la contribution la plus importante à $F_N(\beta)$ provient des termes de $Z_N(\beta)$ qui ont les énergies les plus faibles, i.e. les configurations σ pour lesquelles $H(\sigma)$ est de l'ordre de $-N\sqrt{2 \log 2} = -N\beta_c$, (voir par exemple [1]). En physique, on parle dans ce cas là de "système gelé" (alors que dans un système non gelé, les énergies les plus faibles contribuent seulement lorsque la température est nulle).

Nous avons cherché à caractériser (physiquement) la mesure (aléatoire) de Gibbs sur $(S_N, \mathcal{F}_N)$ dans le cas où la température était supérieure à la température critique ($\beta > \beta_c$) et pour N assez grand ; c'est pourquoi nous avons étudié son support, et par conséquent cherché à déterminer, pour un échantillon $(H_{rem}(\sigma), \sigma \in \{-1, +1\}^N)$ donné, sur quelles configurations σ se concentre la mesure.

Désignons par $\sigma^{(i)}$ la configuration correspondant à la $i$ème statistique d'ordre du 2^n-échantillon $(H_{rem}(\sigma))_\sigma$.

Sont entrés alors en jeu les statistiques d'ordre

$$H(\sigma^{(2^N)}) \ge H(\sigma^{(2^N-1)}) \cdots \ge H(\sigma^{(2)}) \ge H(\sigma^{(1)}),$$

ou plus exactement les espacements $H(\sigma^{(k+1)}) - H(\sigma^{(k)})$, ou encore les sommes de ces espacements. C'est pourquoi nous avons dû établir de nouveaux résultats sur les sommes des espacements de v.a. gaussiennes, ayant un intérêt propre et énoncés dans le chapitre 2, section 2.2.3, de ce travail.

Diverses questions se sont naturellement posées :

1) si l'on utilise ces sommes d'espacements, combien de termes $k = k(N)$ faut-il sommer pour considérer la mesure construite à partir de ces k termes comme une "bonne" approximation de la mesure pour N grand ?

2) la distance en variation totale d_{tv} entre la mesure de Gibbs et son approximation tend-elle vers 0, quand $N \to \infty$ p.s. ?

(on ne s'attend pas à une convergence p.s. de chacune des mesures, mais la différence peut converger p.s. : un fait similaire a été démontré pour le modèle de champ moyen de Curie-Weiss ([14])).

3) si l'on considère cette approximation, que représente cette nouvelle mesure ? est-ce la mesure uniforme sur les $k(N)$ points choisis sans remplacement parmi les 2^N points ? ou une mesure concentrée sur le minimum ? ou ...?

4) la mesure de Gibbs dépend-elle de β même si le système est gelé ? par exemple, peut-on estimer la distance en variation totale entre la mesure de Gibbs lorsque β fini est supérieur à β_c ? qu'en est-il de la mesure limite quand $\beta \to \infty$?

5) enfin, y-a-t-il un moyen simple et non coûteux de construire l'approximation de la mesure de Gibbs ?

Soit $\Psi : S_N = \{-1, +1\}^N \to I\!R$, $||\Psi||_\infty := \sup\limits_{\sigma \in S_N} |\Psi(\sigma)| < \infty$.

Désignons par $\{(\sigma)_i, i \in \{1, \ldots, 2^N\}\}$ les 2^N configurations possibles.

La mesure de Gibbs peut alors s'écrire :

$$\mu_{N,\beta}(\Psi) = \sum_{i=1}^{2^N} \Psi((\sigma)_i) \frac{e^{-\beta H((\sigma)_i)}}{Z_{\beta,N}} = \frac{\sum_{i=1}^{2^N} \Psi((\sigma)_i) e^{-\beta\sqrt{N}X_i}}{\sum_{i=1}^{2^N} e^{-\beta\sqrt{N}X_i}},$$

avec $(X_i, 1 \le i \le 2^N)$ v.a. iid gaussiennes standard associées à $\{(\sigma)_i, 1 \le i \le 2^N\}$,
ou encore,

$$\mu_{N,\beta}(\Psi) = \frac{\sum_{i=1}^{2^N} \Psi((\tilde{\sigma})_i) e^{-\frac{\beta}{\beta_c}\sqrt{2\log 2^N}X_{i,2^N}}}{\sum_{i=1}^{2^N} e^{-\frac{\beta}{\beta_c}\sqrt{2\log 2^N}X_{i,2^N}}},$$

$(\tilde{\sigma})_j$ correspondant à la configuration $(\sigma)_i$ à laquelle $X_i = X_{j,2^N}$ est associée.

Le premier problème a donc été de définir le nombre minimum k_N de termes parmi les $X_{i,2^N}$, nécessaire pour construire une bonne approximation de $\mu_{N,\beta}$.
Nous avons obtenu le résultat suivant (cf. Theorem 3.1 et Corollary 3.1 dans [**11**], rappelés dans le second chapitre de ce travail).
Résultat principal. *Soit k_N tel que*

$$k_N \uparrow \infty, \quad \frac{k_N}{N} \uparrow \infty \quad and \quad \frac{\log k_N}{N} \downarrow 0, \quad quand \quad N \to \infty.$$

Il existe $\Omega_N \subset \Omega$ et un entier N_0 tel que $\forall N > N_0$, $I\!P[\Omega_N] \ge 1 - \dfrac{4}{(N\log 2)^{1+\delta}}$ et sur Ω_N on a

$$\mu_{N,\beta}(\Psi) = \frac{\sum_{i=1}^{k_N} \Psi((\tilde{\sigma})_i) e^{-\frac{\beta}{\beta_c}\sqrt{2\log 2^N}(X_{i,2^N}-X_{1,2^N})} + B_N(\Psi)}{\sum_{i=1}^{k_N} e^{-\frac{\beta}{\beta_c}\sqrt{2\log 2^N}(X_{i,2^N}-X_{1,2^N})} + B_N(1)},$$

avec $|B_N(\Psi)| \le \dfrac{2||\Psi||_\infty}{\frac{\beta}{\beta_c}-1} \times \dfrac{1}{(k_N-1)^{\frac{1}{2}(\frac{\beta}{\beta_c}-1)}}.$
On peut alors vérifier que la mesure aléatoire $\mu_{k_N,\beta}^{(1)}$ sur $\{-1,+1\}^N$ définie par

$$\mu_{k_N,\beta}^{(1)}(\Psi) := \frac{\sum_{k=1}^{k_N} \Psi((\tilde{\sigma})_k) e^{-\frac{\beta}{\beta_c}\sqrt{2\log 2^N}(X_{k,2^N}-X_{1,2^N})}}{\sum_{k=1}^{k_N} e^{-\frac{\beta}{\beta_c}\sqrt{2\log 2^N}(X_{k,2^N}-X_{1,2^N})}},$$

est p.s. proche de la mesure de Gibbs vérifiant

$$d_{TV}(\mu_{N,\beta}, \mu_{k_N,\beta}^{(1)}) \le \frac{4}{\frac{\beta}{\beta_c}-1} \times \frac{1}{(k_N-1)^{\frac{1}{2}(\frac{\beta}{\beta_c}-1)}}.$$

Par exemple, un choix de k_N est $k_N = N\log_p(N)$.
En modifiant légèrement l'hypothèse sur k_N de sorte que :

$$k_N \uparrow \infty, \quad \frac{k_N}{N} \uparrow \infty \quad et \quad \frac{(\log k_N)(\log N)}{N} \downarrow 0, \quad \text{quand } N \to \infty,$$

on peut démontrer que pour tout k_{2^N} satisfaisant cette dernière hypothèse, sur Ω_n^* (tq $I\!P[\Omega_n^*]$ tend vers 1 quand $n := 2^N \to \infty$),

$$\forall j, \ 1 \le j \le k_n, \quad \sqrt{2\log n}(X_{j,n} - X_{1,n}) = \sum_{i=1}^{j} \frac{W_i}{i} + O\left(\frac{\log(k_n)\log\log n}{\log n}\right),$$

où (W_i) v.a. i.i.d. de loi exponentielle standard.

Nous en déduisons une autre approximation de la mesure de Gibbs (au sens $d_{TV}(\mu_{N,\beta}, \mu_{k_N,\beta}^{(2)}) \to 0$), utile pour la simulation, à savoir :

$$\mu_{k_N,\beta}^{(2)}(\Psi) := \frac{\sum_{k=1}^{k_N} \Psi((\tilde{\sigma})_k) e^{-\frac{\beta}{\beta_c} \sum_{\ell=1}^{k} \frac{W_\ell}{\ell}}}{\sum_{k=1}^{k_N} e^{-\frac{\beta}{\beta_c} \sum_{\ell=1}^{k} \frac{W_\ell}{\ell}}}.$$

Ces différents résultats ont permis de répondre aux questions que nous pouvions nous poser sur la mesure de Gibbs ; par exemple, la mesure de Gibbs n'est p.s. pas concentrée, lorsque $\beta_c < \beta < \infty$, sur le minimum ($\mu_{K_N,\beta} \neq \delta_{(\tilde{\sigma})_1}$), ni sur $\{(\tilde{\sigma})_1, \ldots, (\tilde{\sigma})_{k_0}\}$, avec k_0 fini ; il ne s'agit pas non plus d'une mesure uniformément répartie sur les k_N premiers minima $\{(\tilde{\sigma})_1, \ldots, (\tilde{\sigma})_{k_N}\}$.

Suite à cette étude, il est naturel de passer au cas plus général du G.R.E.M. en considérant non plus des v.a. gaussiennes indépendantes mais à corrélation positive.

Perspectives

• Le travail réalisé dans le cadre de l'étude de fonctionnelles non linéaires de processus gaussiens (cf. Chapitre 1) représente un investissement important qu'il conviendra de valoriser par la suite. Nombre de problèmes restent à résoudre, en particulier dans le cas non stationnaire ou multidimensionnel, en vue d'applications dans divers domaines liés à la physique, au traitement d'image, à la santé, ...

⋆ Nous nous intéressons actuellement, avec J. León et I. Iribarren (UCV, Caracas), aux problèmes de réflexion et réfraction sur des surfaces aléatoires, traités dans les années 60 par le physicien Longuet-Higggins ([102]), en utilisant entre autres les études faîtes sur les franchissements de courbes différentiables par des champs gaussiens stationnaires ([**10**], [**13**]), pour démontrer et compléter les résultats qu'il avait obtenus.

⋆ L'étude des franchissements ou dépassements de niveau par des champs gaussiens peut être utilisée également pour le traitement d'images médicales, en particulier pour l'analyse des milieux poreux.

Etant donné un champ gaussien X d'espace d'indices d-dimensionnel et un seuil, nous pouvons construire un champ gaussien seuillé Y tel que en chaque point t de l'espace d'indices, $Y(t)$ vaut 1 si $X(t)$ dépasse le seuil (ou niveau), et 0 sinon. Un milieu poreux pouvant être décrit par un champ booléen (i.e. à valeurs 0-1), nous nous intéresserons dans un premier temps aux propriétés (statistiques, morphologiques) des champs gaussiens seuillés, puis à l'opération de reconstruction d'un champ gaussien seuillé X à partir du champ booléen Y décrivant un milieu poreux, X devant avoir les mêmes propriétés que Y. Il s'agit d'un projet en collaboration avec A. Estrade (MAP5), I. Iribarren et J. León (UCV, Caracas) (dans le cadre du groupe de travail 'Milieux poreux' créé récemment au MAP5 par A. Estrade).

⋆ En ce qui concerne l'approximation du temps local par le nombre de franchissements (cf. Chapitre 1, §1.1.2), une collaboration est en cours avec W. Urbina (Univ. of Kansas et UCV, Caracas), inspirée de l'article [6] de Azaïs, dans laquelle nous étudions la convergence du nombre de franchissements vers le temps local pour des solutions d'équations différentielles stochastiques sans drift définies par rapport au mouvement brownien fractionnaire.

• Sur le thème des systèmes désordonnés se rattachent deux projets.

⋆ L'un concerne les modèles de Derrida ([6] et [7]) et cherche à géneraliser au cas des G.R.E.M. le travail [11] effectué sur des R.E.M. (avec P. Picco).

⋆ L'autre étude, en collaboration avec M. Atencia et G. Joya (Univ. Málaga), traite des modèles ou réseaux de Hopfield avec bruit aléatoire, que nous utilisons à des fins d'optimisation combinatoire. Dans sa thèse, M. Atencia a montré que, dans un cadre déterministe, les réseaux de neurones de Hopfield constituent un outil intéressant pour obtenir la solution de problèmes d'optimisation et d'ingénierie de contrôle ; il en a étudié les différentes formulations dont celle d'Abe qu'il a démontré être la plus adaptée à la résolution des problèmes d'optimisation.
Suite à ce travail, nous étudions actuellement l'influence du bruit aléatoire dans l'application des réseaux de Hopfield (formulation de Abe) à l'optimisation combinatoire. Le modèle de Abe, qui s'éxprime par une équation différentielle stochastique, se révèle stable et nous avons déterminé l'ensemble de convergence des états, en supposant l'intensité du bruit bornée. Ce résultat a fait l'objet d'un article [12] présenté à l'ESANN en avril dernier. L'étude se poursuit sur la relation entre l'ensemble de convergence du modèle déterministe et celui du modèle stochastique.

• Enfin, Y. Le Strat (Institut de Veille Sanitaire) et moi-même démarrons un projet ayant pour but de répondre à plusieurs questions relatives à la détection en temps ou spatiale d'une survenue inhabituelle de cas de maladies ou de décès, en utilisant les outils de la statistique des extrêmes. En effet, alors que de nombreuses méthodes statistiques ont été développées pour la surveillance épidémiologique, il n'existe pas encore, à notre connaissance, d'applications de la théorie des valeurs extrêmes à ce domaine.

Bibliographie

[1] M.F. Kratz, Statistique des queues de distribution et approximation Poissonnienne, *Thèse de doctorat de l'Université Pierre et Marie Curie, Paris VI* (1993).

[2] M.F. Kratz, Approximation Poissonnienne relative du processus empirique, *C.R.A.S.* **316, I**, 1221-1224 (1993).

[3] J. Hüsler and M. Kratz, On the convergence of the number of exceedances of nonstationary normal sequences, *Journal of Research of the NIST* **99**, 539-542 (1994).

[4] J. Hüsler and M. Kratz, Rate of Poisson approximation of the number of exceedances of nonstationary normal sequences, *Stoch. Proc. Applic.* **55** 301-313 (1995).

[5] P. Feigin, M. Kratz and S. Resnick, Parameter Estimation for Moving Averages with Positive Innovations, *Ann. Appl. Probab.* **6**, 1157-1190 (1996).

[6] M. Kratz and S. Resnick, The QQ-estimator and heavy tails, *Stoch. Models*, **12**, 699-724 (1996).

[7] M. Kratz and H. Rootzén, On the rate of convergence for extremes of mean square differentiable stationary normal processes, *J. Applied Probab.* **34**, 908-923 (1997).

[8] M. Kratz and J. León, Hermite polynomial expansion for non-smooth functionals of stationary Gaussian processes : crossings and extremes, *Stoch. Proc. Applic.* **66**, 237-252 (1997).

[9] M. Kratz and J. León, Central limit theorems for the number of maxima and some estimator of the second spectral moment of a stationary Gaussian process. Applications in hydroscience, *Extremes* **3 :1**, 57-86 (2000) .

[10] M. Kratz and J. León, Central limit theorems for level functionals of stationary Gaussian processes and fields, *J.Theoretical Probab.* **14** (2001) 639-672.

[11] M. Kratz and P. Picco, On a representation of Gibbs measure for R.E.M., *Ann. Applied Probab.* **14**, 651-677 (2004).

[12] M. Atencia, G. Joya and M. Kratz, Stochastic analysis of the Abe formulation of Hopfield networks, Proceedings *ESANN* (2005).

[13] M. Kratz and J. León, On the second moment of the number of crossings by a stationary Gaussian process, *Ann. Probab., to appear*

Prépublications

[14] M. Kratz, Chaos expansions and level crossings, *preprint Samos* 127 (non submitted, 2000).

[15] M. Kratz and J. León, Curve crossings and specular points, d'après Longuet-Higgins *preprint Samos-226 or Map5-18* (2005).

Rapport industriel

[16] M.F. Kratz, Modélisation et analyse statistique des longueurs de bouchons liquides dans les écoulements diphasiques dans les pipelines, *C.F.P., TOTAL* (1990).

Chapitre 1

Level crossings and other level functionals

I will present in this review the mathematical aspect of the work done on level crossings and upcrossings of a level, and will insist on the different approaches used to tackle a same problem. I will speak only briefly about examples or applications which of course were the reason of such theoretical work and which began to be much developed in the extreme value field those past years. This review of mathematical results and methods can be also very useful in mathematical statistics applications, reliability theory and other areas of applications of stochastic processes. We will try to privilege conditions on the behavior of spectral or covariance functions rather than general conditions of mixing, since they are easier to work with.

We will consider mainly continuous parameter processes, since they appear in most of the mathematical models to describe physical phenomena. Nevertheless one approach to work on level crossings or on upcrossings of a level, which are analogous to the exceedances used in the discrete case (on which a short review is given in the previous chapter *Présentation générale*), can be through discretization, using discrete parameter results, as we will see later.

Let $X = (X_t, t \in I\!R^d)$ be a real stochastic process. Our main concern is the measure of the random set of level x defined by $C_x^X := \{t : X_t = x\}$, by taking the same notation as Wschebor ([180]). The unidimensional case will be exposed in details with results and methods (especially the ones which can/could be easily adapted to a higher dimension), whereas only a partial view of the multidimensional case will be given.

Let us give the notations and specific symbols used through this chapter.
$L^p(\Omega) := \{X : I\!E[|X|^p] < \infty\}\,; \, ||.||_p : L^p$-norm.
$1\!\!1_A$: indicator function of a set A.
$\phi_{s,t}(x, \dot{x}, y, \dot{y})$: density function of $(X_s, \dot{X}_s, X_t, \dot{X}_t)\,;$
$\phi_t(x, \dot{x}, y, \dot{y}) := \phi_{0,t}(x, \dot{x}, y, \dot{y})$.
r.v. : random variable
CLT : central limit theorem
s.t. : such that
$\overset{d}{\to}$ stands for convergence in distribution.

1.1 Crossings of Gaussian processes

Theorem 1.1.1 *(Bulinskaya, 1961 [33])*
Let $X = (X_s, 0 \leq s \leq t)$ be an a.s. continuously differentiable stochastic process which one-dimensional density $p_s(x)$ is bounded in x for all s.
Then for any level u the probability of existence of a point s such that the event $X(s) = u$ and $\dot{X}(s) = 0$ occurs simultaneously is equal to 0. In particular, the probability that X_t becomes tangent to the level u, is equal to 0, in other words the probability of contingence of the level u by the process X is equal to 0.

Studies on level-crossings by stationary Gaussian processes have begun about sixty years ago. Different approachs have been proposed. Here is a survey on the literature on the number of crossings of a given level or of a differentiable curve in a fixed time interval by a continuous spectrum Gaussian process.
Besides the well-known books about Extremes, a very short survey has been proposed by Slud in 1994 (see [162] or [163]), and a more general survey about extremes including a short section about level crossings by Rootzen in 1995, to quote the most recent one (see [146],[95]), and by Piterbarg in his 1996 book (see [136], in particular for some methods described in more details than here).
My purpose here is to focalize only on crossing counts in order to be more explicit about the subject, not only recalling the main results, but giving also the main ideas about the different methods used to establish them. To make the methods easier to understand, I will consider mainly crossing counts of a given level. Note that the problem of curve (ψ) crossings by a stationary process X may also be regarded as a zero-crossing problem for the non stationary process $X^* := X - \psi$ (but stationary in the sense of the covariance), as noticed by Cramér and Leadbetter ([40]).

Let $X = (X_s, s \geq 0)$ be a real stationary Gaussian process with variance one and a.s. continuous sample functions.
Let ψ be a continuously differentiable function.
We will denote by $N_I(x)$ or $N_I(\psi))$ the number of crossings of a given level x or of a curve $\psi(.)$ respectively, by X on the interval I, and by $N_I^+(x)$ the number of upcrossings of x by X (recall that X_s is said to have an upcrossing of x at $s_0 > 0$ if for some $\varepsilon > 0$, $X_s \leq x$ in $(s_0 - \varepsilon, s_0)$ and $X_s \geq x$ in $(s_0, s_0 + \varepsilon)$). The notation will be simplified to $N_t(x)$ or $N_t(\psi)$ or $N_t^+(x)$ respectively, when $I = [0, t]$, $t > 0$.

In the case of a Gaussian process, as a consequence of the separability property and the Tsyrelson theorem (see for instance [99]), we have

Theorem 1.1.2 *For any Gaussian process X on a arbitrary parameter set T, for any x such that*

$$I\!P[\sup_{s \in T} X(s) \; < \; x] > 0,$$

there is no contingence, with probability one.

1.1.1 Moments and factorial moments

Distributional results about level or curve crossings by a Gaussian process are often obtained in terms of factorial moments for the number of crossings, that's why many authors have been working not only on moments but also on factorial moments, to find out expressions or conditions (in terms of the covariance function of the process) for their finiteness, since a certain number of applications require to know if they are finite.

Note also that the conditions governing finite crossing moments are local ones, since Hölder inequality implies that $(I\!E[N_{2t}(x)])^k \leq 2^k (I\!E[N_t(x)])^k$, and therefore when $(I\!E[N_t(x)])^k$ is finite for some t, it is finite for all t.

Introduction

● *Kac's method and formula (1943).*
We can start with Kac (cf. [82]), who studied the number N of zeros of a Gaussian random polynomial on some bounded interval of $I\!R$. To compute $I\!E[N]$, he proposed a method which used somehow, in a formal way, the approximation of the "Dirac function" $\delta_0(x)$ by the function $\frac{1}{2\varepsilon} I\!I_{[-\varepsilon,\varepsilon)}(x)$. He gave the first heuristic expression of $N_t(0)$ as a function of the process X :

Theorem 1.1.3 *(Kac formula, 1943)*

$$N_t(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^t cos(\zeta X_s)|\dot{X}_s| ds d\zeta, \quad with\ probability\ one.$$

From now on, we suppose w.l.o.g. that X is centered, with correlation function $r(t) = I\!E[X_0 X_t]$, given also in respect with the spectral function F as $r(t) = \int_0^{\infty} cos(\lambda t) dF(\lambda)$. We will denote by λ_2 the second moment (when it exists) of the spectral function, i.e. $\lambda_2 = \int_0^{\infty} \lambda^2 dF(\lambda)$.

● *Rice (1945).*
One of the most well-known first results is the one of Rice (cf. [143]) who proved by intuitive methods related to those used later by Cramér and Leadbetter (as the discretization method described below), that

Theorem 1.1.4 *(Rice formula, 1945)*

$$I\!E[N_t(x)] = te^{-x^2/2} \sqrt{-r''(0)}\ /\pi. \tag{1.1}$$

It means that the mean number of crossings is the most important for zero crossings, and decreases exponentially with the level.

● *Itô (1964), Ylvisaker (1965).*
Itô and Ylvisaker have proved (1.1) and so that the mean number of crossings is finite exactly when the second spectral moment is finite :

Theorem 1.1.5 *(Itô, 1964 ; Ylvisaker, 1965)*

$$\mathbb{E}[N_t(x)] < \infty \quad \Leftrightarrow \quad \lambda_2 < \infty \quad \Leftrightarrow \quad -r''(0) < \infty.$$

● *The discretization method.*
From the intuitive method developed by Rice (1939,1944,1945), Ivanov (1960, [79]), Bulinskaya (1961, [33]), Itô (1964, [80]) and Ylvisaker (1965, [182]) proposed rigorous proofs for zero countings. Followed the general formulation, due to Leadbetter (1966, cf. [40] p.195 or [94] p.148), known as the method of discretization, which applies also to non-normal processes.
This method is based on the approximation of the continuous process $X = (X_t, t \geq 0)$ by the sequence $(X(jq), j = 1, 2, \cdots)$ where q satisfies some conditions related to the level x of upcrossings :

$$q = q(x) \to 0 \quad \text{and} \quad xq \downarrow 0 \quad \text{as} \quad x = x(t) \to \infty \quad \text{(or equivalently as } t \to \infty). \quad (1.2)$$

Consider any sequence (q_n) s.t. $q_n \downarrow 0$ and the number $N_{n,I}$ of points $(jq_n, j = 1, 2, \cdots)$, in a fixed and bounded interval I, s.t. $\begin{cases} (j-1)q_n \in I \text{ and } jq_n \in I, \\ X((j-1)q_n) < x < X(jq_n). \end{cases}$
It can be proved (see Cramér and Leadbetter, [40] p.195, and Leadbetter et al., [94], Lemma 7.7.2) that

$$N_{n,I} \to N_I^+(x) \text{ a.s. as } n \to \infty, \quad (1.3)$$

hence that $N_I^+(x)$ is a (possibly infinite-valued) r.v., and that (see [94]) $\mathbb{E}[N_{n,I}] \underset{n \to \infty}{\to} \mathbb{E}[N_I^+(x)]$.
Following the notation of Leadbetter et al. ([94]), let

$$J_q(x) = \frac{1}{q} \mathbb{P}[X_0 < x < X_q], \quad q > 0. \quad (1.4)$$

By choosing $I = (0, t]$, we have $\nu_n := t/q_n$ points $jq_n \in I$ and by stationarity, it comes $\mathbb{E}[N_n] = (\nu_n - 1)\mathbb{P}[X_0 < x < X_{q_n}] \sim tJ_{q_n}(x)$; since the sequence (q_n) is arbitrary, it implies

$$\mathbb{E}[N_t^+(x)] = t \lim_{q \downarrow 0} J_q(x) \quad (1.5)$$

and hence that
$$\mathbb{E}[N_t^+(x)] = t\mathbb{E}[N_1^+(x)]. \quad (1.6)$$

Under mild conditions, $\lim_{q \to 0} J_q(x)$, and so $\mathbb{E}[N_1^+(x)]$, can be expressed in a simple integral form (by writting J_q as $\mathbb{P}[X_0 > x , \tilde{X}_q > q^{-1}(x - X_0)]$, with $\tilde{X}_q := q^{-1}(X_q - X_0)$, and by doing some change of variable), giving in the normal case (see [94], lemma 7.3.1) :

Lemma 1.1.1 *(Leadbetter et al.)*
Let $X = (X_t, t \geq 0)$ be a standardized stationary normal process with its second spectral moment λ_2 finite and let $(q = q(x), x)$ satisfying (1.2). Then, as $q \to 0$,

$$J_q(x) \sim \int_0^\infty zp(x, z)dz = \frac{\lambda_2^{1/2}}{2\pi}e^{-x^2/2},$$

$p(.,.)$ denoting the joint density of $(X, \dot{X})$ and J_q satisfying (1.4).

Therefore this last result, together with (1.5), provide the Rice formula for the number of upcrossings of any fixed level x per unit time by a standardized stationary normal process, namely

$$\mathbb{E}[N_1^+(x)] = e^{-x^2/2} \, \lambda_2^{1/2}/(2\pi).$$

• *Cramér and Leadbetter (1965,1967), Ylvisaker (1966).*
In the 60's, following on the work of Cramér, generalization to curve crossings and higher order moments for $N_t(.)$ were considered in a series of papers by Cramér and Leadbetter (cf. [40]) and Ylvisaker (cf. [183]).
In what concerns the curve crossings, the generalized Rice formula was obtained :

Lemma 1.1.2 *(Generalized Rice formula ; Ylvisaker, 1966 ; Cramér and Leadbetter, 1967).*
Let $X = (X_t, t \geq 0)$ be a mean zero, variance one, stationary Gaussian process, with twice differentiable covariance function r. Let ψ be a continuous differentiable real function. Then

$$\mathbb{E}[N_t(\psi)] = \sqrt{-r''(0)} \int_o^t \varphi(\psi(y)) \left[2\varphi\left(\frac{\psi'(y)}{\sqrt{-r''(0)}}\right) + \frac{\psi'(y)}{\sqrt{-r''(0)}} \left(2\Phi\left(\frac{\psi'(y)}{\sqrt{-r''(0)}}\right) - 1\right)\right] dy, (1.7)$$

where φ and Φ are respectively the standard normal density and distribution function.

Again the authors used the discretization method and approximated the continuous time number $N_t(\psi)$ of ψ-crossings by the discrete-time numbers $(N_\psi(t, q_n))$ of crossings of continuous polygonal curves agreeing with ψ at points jq_n $(j = 0, 1, \cdots, 2^n)$, with time steps $q_n = t/2^n$:

$$N_\psi(t, q_n) := \sum_{j=0}^{2^n-1} \mathbb{1}_{\left((X_{jq_n} - \psi(jq_n))(X_{(j+1)q_n} - \psi((j+1)q_n)) < 0\right)} \tag{1.8}$$

to obtain that $N_\psi(t, q_n) \uparrow N_t(\psi)$ with probability one, as $n \to \infty$. □

Moments and factorial moments of order 2

• Cramér and Leadbetter proposed a sufficient condition (known nowdays as *Geman condition*) on the correlation function of X (stationary case) in order to have the random variable $N_t(x)$ belonging to $L^2(\Omega)$, namely

Theorem 1.1.6 *(Cramér et al., 1967)*

$$\text{If} \quad \exists \delta > 0, \quad L(t) := \frac{r''(t) - r''(0)}{t} \in L^1([0, \delta], dx) \tag{1.9}$$
$$\text{then} \quad \mathbb{E}[N_t^2(0)] < \infty.$$

31

• Explicit (factorial) moment formulas for the number of crossings have been obtained by Cramér and Leadbetter (1965, 1967), Belayev (1966) and Ylvisaker (1966), based on careful computations involving joint densities of values and derivatives of the underlying Gaussian process.

In particular, the second moment of $N_t(x)$ is given by

$$IE[N_t^2(x)] = IE[N_t(x)] + 2\int_0^t (t-u)\int_{IR^2} |\dot{x}||\dot{y}|\phi_u(x,\dot{x},x,\dot{y})d\dot{x}d\dot{y}du.$$

Concerning the second factorial moment, Cramér & Leadbetter provided an explicit formula for the number of zeros of the process X (see [40], pp. 209), from which can be deduced the following formula for the second factorial moment of the number of crossings of a continuous differentiable real function ψ by X :

$$IE[N_t(\psi)(N_t(\psi)-1)] = \int_0^t \int_0^t \int_{R^2} |\dot{x}_1 - \dot{\psi}_{t_1}||\dot{x}_2 - \dot{\psi}_{t_2}|\phi_{t_1,t_2}(\psi_{t_1},\dot{x}_1,\psi_{t_2},\dot{x}_2)d\dot{x}_1 d\dot{x}_2 \ dt_1 dt_2 \ , (1.10)$$

where the density ϕ_{t_1,t_2} is supposed non-singular for all $t_1 \neq t_2$. The formula holds whether the second factorial moment is finite or not.

• *Ershov (1967).*
This author proved (see [55]) that whenever a Gaussian stationary process $X = (X_t, t \in IR)$ with covariance function r is with mixing (i.e. that $|r(|i-j|)| \leq f(|i-j|)$ with $\lim_{k\to\infty} f(k) = 0$), then the number of its x-upcrossings on all IR can not hold a finite second moment :

Theorem 1.1.7 *(Ershov, 1967)*
If $r(t) \to 0$ as $t \to \infty$,
then $var\left(N_t^+(x)\right) \ \to \ \infty, \ as \ t \to \infty.$

• *Geman (1972).*
Geman ([59]) proved in 1972 that the condition (1.9) was not only sufficient but also necessary, by showing that if $\frac{r''(t)-r''(0)}{t}$ diverges on $(0,\delta)$ then so does the integral $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |xy|\phi_s(0,x,0,y)dxdy$ appearing in the computation of the second moment. So we have

Theorem 1.1.8 *(Geman condition, 1972)*

$$Geman \ condition \ (1.9) \quad \Leftrightarrow \quad IE[N_t^2(0)] < \infty.$$

Of course, this result holds for any stationary Gaussian process when considering the number of crossings of the mean of the process.

• *Wschebor (1985)* (see [180]) provided (under the Geman condition) an explicit expression in the case of two different levels x and y, with $x \neq y$, namely :

$$IE[N_t(x)N_t(y)] = \int_0^t (t-u)\int_{IR^2} |\dot{x}||\dot{y}| \left(\phi_u(x,\dot{x},y,\dot{y}) + \phi_u(y,\dot{x},x,\dot{y})\right)d\dot{x}d\dot{y}du.$$

32

Note that this expression differs from the one of Cramér et al. when $x \to y$, which means that the function $\mathbb{E}[N_t(x)N_t(y)]$ is not continuous on the diagonal.

• *Piterbarg (1982), Kratz and Rootzén (1997).*
Let us assume that the correlation function r of X satisfies

$$r(s) = 1 - \frac{s^2}{2} + o(s^2) \qquad \text{as} \quad s \to 0, \tag{1.11}$$

$$\mathbb{E}\{(X'(s) - X'(0))^2\} = 2(r''(s) - r''(0)) \leq c|s|^{\gamma}, \qquad s \geq 0, \quad 0 < \gamma \leq 2, \tag{1.12}$$

that $r(s)$ and its first derivative decay polynomially,

$$|r(s)| \leq Cs^{-\alpha}, \qquad |r(s)| + r'(s)^2 \leq C_0 s^{-\alpha}, \qquad s \geq 0, \tag{1.13}$$

for some $\alpha > 2$ and constants c, C, C_0,
and that the range of t, x is such that

$$\varepsilon \leq \mathbb{E}[N_x^+(t)] \leq K_0 \tag{1.14}$$

for some fixed $K_0, \varepsilon > 0$; by using Piterbarg's notations, let

$$m = m(s) = \frac{r'(s)}{1 + r(s)}, \quad \sigma = \sigma(s) = \sqrt{\frac{1 - r(s)^2 - r'(s)^2}{1 - r(s)^2}}$$

and

$$m^*(s) = \sup_{v \geq s > 0} \frac{m^2(v)}{2\pi\sqrt{1 - r^2(v)}}, \quad \sigma^* = \sup_{s \geq 0} \frac{\sigma^2(s)}{2\pi\sqrt{1 - r^2(s)}}.$$

Piterbarg (cf. [135]) proposed some bounds of the second factorial moment of the number of upcrossings; Kratz and Rootzén (cf. [92]) gave a small variation of his result, but with more precise bounds, namely

Lemma 1.1.3 *(Kratz and Rootzén, 1997)*
Suppose that the previous hypotheses are satisfied (with $\gamma = 2$). Then, for $t, x \geq 1$,

$$\mathbb{E}\left[N_x^+(t)(N_x^+(t) - 1)\right] \leq K_3 t x^{2 + 2/\alpha} e^{-\frac{x^2}{2}\left(1 + \inf_{s \geq 0} \rho(s)\right)} + K_4 t^2 x^2 e^{-x^2}, \tag{1.15}$$

with $\rho(.)$ defined in (1.45) below, $K_3 = (3.1\sigma^* + 2.3m^*(0))(3C_0)^{1/\alpha}$, $K_4 = 2.6\sigma^* + 2.3m^*(0)$.

• Recently, Kratz and León (see [91]) succeded in generalizing the theorem 1.1.8 to any given level x (not only the mean of the process) and to some differentiable curve ψ.

Note that this problem of finding a simple necessary and sufficient condition for the number of crossings of any level has been subject to some investigation and nice papers, such as the ones of Cuzick ([42], [45], [46]) proposing sufficient conditions. But to get necessary conditions remained an open problem for many years. The solution of this problem is enunciated in the following theorem.

Theorem 1.1.9 *(Kratz and León, 2004)*
   1) *For any given level x, we have*

$$\mathbb{E}[N_t^2(x)] < \infty \quad \Leftrightarrow \quad \exists \delta > 0, \; L(t) = \frac{r''(t) - r''(0)}{t} \in L^1([0, \delta], dx) \quad (Geman\; condition).$$

   2) *Suppose that the continuous differentiable real function ψ satisfies for some $\delta > 0$,*

$$\int_0^\delta \frac{\gamma(s)}{s} ds < \infty, \quad where\; \gamma(\tau)\; is\; the\; modulus\; of\; continuity\; of\; \dot\psi.$$

   Then
$$\mathbb{E}[N_t^2(\psi)] < \infty \quad \Leftrightarrow \quad L(t) \in L^1([0, \delta], dx).$$

This smooth condition on ψ is satisfied by a large class of functions which includes in particular functions whose derivatives are Hölder.

The method used to prove that the Geman condition keeps being the sufficient and necessary condition to have a second moment finite in these different cases, is quite simple.
It relies mainly on the study of some functions of r and its derivatives at the neighborhood of 0, and the chaos expansion of the second moment, notion which will be explained later.

Factorial moments and moments of higher order

• Concerning moments of higher order than 2, Cramér and Leadbetter (1965) got results under very mild conditions in the stationary case, that Belayev (1966) derived in the non stationary case under slightly more restrictive conditions, weakened by Ylvisaker (1966) in the stationary case but which may also be adapted to cover nonstationary cases. Let us give for instance the kth factorial moment of the number of zero crossings by a stationary Gaussian process :

$$\begin{aligned}
M_t^k(0) \quad &:= \quad \mathbb{E}[N_t(0)(N_t(0) - 1) \cdots (N_t(0) - k + 1)] \\
&= \quad \int_0^t dt_1 \cdots \int_0^t dt_k \mathbb{E}\left[\prod_{i=1}^k |\dot X(t_j)| \mid X(t_j) = 0, \; 0 \le j \le k\right] p(0, ..., 0) \quad (1.16)
\end{aligned}$$

where $p(x_1, \cdots, x_k)$ is the joint density of the r.v. $(X_{t_1}, \cdots, X_{t_k})$.

- *Belyaev (1967).*

Concurrently to this study, Belyaev proposed in ([20]) a sufficient condition for the finiteness of the kth factorial moment $M_t^k(0)$ for the number $N_t(0)$ of zero crossings on the interval $[0, t]$ in terms of the covariance matrix Σ_k of $(X_{t_1}, \cdots, X_{t_k})$ and of

$$\sigma_i^2 := var\left(\dot{X}_{t_i} \mid X_{t_j} = 0, \ 1 \leq j \leq k\right); \tag{1.17}$$

Theorem 1.1.10 *(Belyaev, 1967)*

$$If \quad \int_0^t dt_1 \cdots \int_0^t dt_k \left(\frac{\prod_{i=1}^k \sigma_i^2}{det\Sigma_k}\right)^{1/2} < \infty \tag{1.18}$$

$$then \quad M_t^k(0) := I\!E[N_t(0)(N_t(0) - 1) \cdots (N_t(0) - k + 1)] < \infty.$$

- *Cuzick (1975-1978).*

Cuzick proved in 1975 ([42]) that Belyaev's condition (1.18) for the finiteness of the kth factorial moments for the number of zero crossings, was not only sufficient but also necessary :

Theorem 1.1.11 *(Cuzick condition, 1975)*

$$Condition \ (1.18) \quad \Leftrightarrow \quad \int_0^\varepsilon d\Delta_1 \cdots \int_0^\varepsilon d\Delta_{k-1} \left(\frac{\prod_{i=1}^k \sigma_i^2}{det\Sigma_k}\right)^{1/2} < \infty \tag{1.19}$$

$$for \ some \ \varepsilon > 0, \quad where \ \sigma_i \ satisfies \ (1.17)$$

$$\Leftrightarrow \quad M_t^k(0) = I\!E[N_t(0)(N_t(0) - 1) \cdots (N_t(0) - k + 1)] < \infty.$$

The fact that condition (1.18) is equivalent to condition (1.19) is immediate by the change of variables $\Delta_i := t_{i+1} - t_i$, $1 \leq i \leq k - 1$, after noticing the symmetry of the integrand in (1.18). Now the necessity of (1.19) comes from the fact that the lemma given below implies that

$$(1.16) \ > \ C \int_0^t dt_1 \cdots \int_0^t dt_k \left(\frac{\prod_{i=1}^k \sigma_i^2}{det\Sigma_k}\right)^{1/2}, \text{ with } C \text{ some constant.}$$

Then Cuzick tried to derive from his result (1.19), simpler sufficient conditions to have the finiteness of the kth (factorial) moments for the number of crossings.

In particular, in his 75's paper ([42]), he proved that $M_t^k(0) < \infty$ for all k, for a covariance function r having a behavior near 0 such that $r(s) = 1 - \dfrac{s^2}{2} + c\dfrac{s^3}{3!} + o(s^3)$ as $s \to 0$ (with c some constant > 0).

Later (see [45] and [46]), for X with path continuous nth derivative $X^{(n)}$ and spectral distribution function $F(\lambda)$, he proposed a series of sufficient conditions involving F and $\sigma_n^2(h) := I\!E[(X_{t+h}^{(n)} - X_t^{(n)})^2]$ for having $I\!E[(N_t(0))^k] < \infty$.

Let us give an example in terms of the spectral density of X among the sufficient conditions he proposed.

Theorem 1.1.12 *(Cuzick, 1978)*
If X has a spectral density given by $f(\lambda) := 1/\ (1 + \lambda^3 |\log \lambda|^\alpha)$, then for $k \geq 2$ and $\alpha > 3k/2 - 1$, we have that $M_t^k(0) < \infty$.

35

Those results are not restricted to zero crossings and would also apply to a large family of curves (see [41] and [45]).

However, necessary conditions are more difficult for higher monents, the main difficulty lying in obtaining sharp lower bounds for the σ_i^2 defined in (1.19).

- *Marcus (1977) : generalized Rice functions.*

Marcus (cf. [111]) generalised results of Cramér and Leadbetter and Ylvisaker by considering not only Gaussian processes and by computing quantities such as $\mathbb{E}[N_t^{j_1}(x_1) \cdots N_t^{j_k}(x_k)]$ for levels $x_1, \cdots, x_k$ and integers $j_1, \cdots, j_k$, called generalized Rice functions.

For the proofs, the author returns to the approach used by Kac ([82]), Ivanov ([79]) and Itô ([80]) to obtain the mean number of crossings at a fixed level, which consists in first finding a function that counts the level crossings of a real valued function, then in substituting X for the function (to take then the expectation).

- *Nualart and Wschebor (1991).*

We know that the general Rice formula giving the factorial moments of the number of level crossings of a stochastic process satisfying some conditions can hold wheather finite or not.

In the search of conditions for the finiteness of moments of the number of crossings, Nualart and Wschebor (see [127], 1991) proposed some sufficient conditions in the case of a general stochastic process, which in the Gaussian case, become :

Theorem 1.1.13 *(Nualart and Wschebor, 1991)*
If $X = (X_t, t \in I \subset \mathbb{R})$ is a Gaussian process having C^∞ paths and such that $var(X_t) \geq a > 0$, $t \in I$, then $M_t^k(u) < \infty$ for every level $u \in \mathbb{R}$ and every $k \in \mathbb{N}^*$.

- *Azaïs and Wschebor (2001).*

The computation of the factorial moments of the crossings of a process keeps being a subject of interest. In particular, when expressing these factorial moments by means of Rice integral formulae (of the type of (1.16) in the case of the 2nd factorial moment of the zero crossings, for instance), arises the problem of describing the behavior of the integrands (appearing in these formulae) near the diagonal; it is still an open (and difficult) problem, even though partial answers to it have been provided up to now, as for instance by Azaïs and Wschebor (see [11], [13], and references therein). Let us give an example of the type of results they got, which helps to improve the efficiency of the numerical computation of the factorial moments, in spite of the restrictive conditions.

Proposition 1.1.1 *(Azaïs and Wschebor, 2001)*
Suppose that X is a centered Gaussian process with C^{2k-1} paths (k integer ≥ 2), and that for each pairwise distinct values of the parameter $t_1, t_2, \cdots, t_k \in I$ the joint distribution of $(X_{t_h}, X'_{t_h}, \cdots, X_{t_h}^{2k-1}, h = 1, 2, \cdots, k)$ is non degenerate. Then, as $t_1, t_2, \cdots, t_k \to t^*$,

$$\int_{[0,+\infty)^k} x'_1 \cdots x'_k \, p_{X_{t_1}, \cdots, X_{t_k}, X'_{t_1}, \cdots, X'_{t_k}}(0, \cdots, 0, x'_1 \cdots x'_k) dx'_1 \cdots dx'_k \approx J_k(t^*) \prod_{1 \leq i < j \leq k} (t_j - t_i)^4,$$

where $J_k(.)$ is a continuous non-zero function.

Two reference methods in the Gaussian extreme value theory

• *The normal comparison method.*
The main tool in the Gaussian extreme value theory, and maybe one of the basic important tools of probability theory, has been certainly the so-called normal comparison technique.
This method, used in the Gaussian case, bounds the difference between two standardized normal distribution functions by a convenient function of their covariances. This idea seems intuitively reasonable since the finite dimensional distributions of a centered stationary Gaussian process is determined by its covariance function.
It has been first developed by Plackett (1954, [139]), Slepian (1962) (cf. [160]), then by Berman (1964, 1971, 1992) (cf. [23]) and by Cramér (1967) (cf. [40]) in the independent or midly dependent case. An extension of this method to the strongly dependent case has been introduced by Mittal and Ylvisaker (1975) (see [122] and also the 84's paper [121] of Mittal for a review on comparison techniques).

Theorem 1.1.14 *(Normal comparison theorem; Slepian, 1962)*
Let $\xi_1 = (\xi_1(t), t \in T)$ and $\xi_2 = (\xi_2(t), t \in T)$ (with T parameter set) be two separable normal processes (possessing continuous sample functions) having same mean and same variance functions, with respective covariance function ρ_1 and ρ_2. Suppose that $\rho_1(s,t) \leq \rho_2(s,t), \; s, t \in T.$
Then for any x, $I\!P\left[\sup_{t \in T} \xi_1(t) \leq x\right] \leq I\!P\left[\sup_{t \in T} \xi_2(t) \leq x\right].$

The proof constitutes a basis for proofs of the Berman inequality and a whole line of its generalizations, and is part of what we call the normal comparison method.
Let us illustrate it by considering a pair of Gaussian vectors of dimension n, X_1 and X_2, that we suppose to be independent, with respective distribution functions F_i, $i = 1, 2$, density functions φ_i, $i = 1, 2$, and covariance matrices $\Sigma_i = ((\sigma_i(j,k))_{j,k})$, $i = 1, 2$ such that $\sigma_1(j,j) = \sigma_2(j,j)$. Then the covariance matrix $\Sigma_h := h\Sigma_2 + (1 - h)\Sigma_1$ is positive definite. Let f_h and F_h be respectively the n-dimensional normal density and distribution function based on Σ_h.
Recall the following equation, discovered by Plackett in 1954 ([139]), recorded by Slepian (1962) ([160]) and proved later in a simplier way than the one of the author, by Berman (1987) ([23]) :

Theorem 1.1.15 *(Plackett partial differential equation, 1954)*
Let Φ_Σ be the centered normal density function with covariance matrix $\Sigma = (\sigma_{ij})_{i,j}$. Then

$$\frac{\partial \Phi_\Sigma}{\partial \sigma_{ij}} = \frac{\partial^2 \Phi_\Sigma}{\partial x_i \partial x_j}, \quad i \neq j.$$

This equation will help to compute the difference between the two normal distribution functions F_i, $i = 1, 2$. Indeed, x being in this case a real vector with coordinates $(x_i)_{1 \leq i \leq n}$, we have

$$F_2(x) - F_1(x) = \int_0^1 \frac{d}{dh} F_h(x) dh = \int_0^1 \sum_{i>j} \frac{\partial}{\partial \sigma_h(i,j)} F_h(x) \frac{d\sigma_h(i,j)}{dh} dh$$

$$= \sum_{i>j} (\sigma_2(i,j) - \sigma_1(i,j)) \int_0^1 \oint_{-\infty}^{x_k} f_h(y_1, \cdots, y_{i-1}, x_i, y_{i+1}, \cdots, y_{j-1}, x_j, y_{j+1}, \cdots, y_n) \prod_{k \neq i,j} dy_k \, dh,$$

where $\oint_{-\infty}^{x_k}$ represents the integral of order $n-2$: $\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_{i-1}} \int_{-\infty}^{x_{i+1}} \cdots \int_{-\infty}^{x_{j-1}} \int_{-\infty}^{x_{j+1}} \cdots \int_{-\infty}^{x_n}$.

Thus, if for all $i \neq j$, $\sigma_2(i,j) - \sigma_1(i,j) \geq 0$, then $F_2(x) - F_1(x) \geq 0$.

The same results hold in the case of Gaussian separable functions of arbitrary kind.

From those two last results, Berman obtained in 1964 :

Theorem 1.1.16 *(Berman inequality, 1964-1992)*
Suppose that $(X_{1i}, 1 \leq i \leq n)$ are standard normal random variables with covariance matrix $\Lambda_1 = (\Lambda_{i,j}^1)$ and $(X_{2i}, 1 \leq i \leq n)$ similarly with covariance matrix $\Lambda_2 = (\Lambda_{i,j}^2)$; let $\rho_{ij} = \max(|\Lambda_{i,j}^1|, |\Lambda_{i,j}^2|)$ and $\rho = \max_{i \neq j} \rho_{ij}$. Then, for any real numbers $x_1, \cdots, x_n$,

$$|\mathbb{P}[X_{1j} \leq x_j, 1 \leq j \leq n] - \mathbb{P}[X_{2j} \leq x_j, 1 \leq j \leq n]| \leq \qquad (1.20)$$

$$\frac{1}{2\pi} \sum_{1 \leq i < j \leq n} \frac{|\Lambda_{i,j}^1 - \Lambda_{i,j}^2|}{\sqrt{(1-\rho_{ij}^2)}} \exp\left\{-\frac{x_i^2 + x_j^2}{2(1+\rho_{ij})}\right\} \leq$$

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \sum_{1 \leq i < j \leq n} |\Lambda_{i,j}^1 - \Lambda_{i,j}^2| \exp\left\{-\frac{x_i^2 + x_j^2}{2(1+\rho)}\right\}.$$

In particular those results hold when choosing one of the two sequences with iid r.v., hence the maximum does behave like that of the associated independent sequence; it helps to prove, under some conditions, results on the maximum and on the point process of exceedances of an adequate level x_n of a stationary normal sequence with correlation function r; for instance we obtain that the point process of exceedances converges to a Poisson process under the weak dependence condition $r_n \log n \to 0$ (cf. [94], chap.4) or to a Cox process under the stronger dependence condition $r_n \log n \to \gamma > 0$, or even to a normal process if $r_n \log n \to \infty$ (cf. [94], chap.6 or [121])).

There is a discussion in Piterbarg (1988 for the russian version, 1996 for the english one [136]) about two directions in which the Berman inequality can be generalized, on one hand on arbitrary events, on the other hand for processes and fields in continuous time. Piterbarg notices that it is not possible to carry the Berman inequality (1.20) over to the processes in continuous time as elegantly as it was done for the Slepian inequality (1.1.14), but provides a solution for Gaussian stationary processes with smooth enough paths (see Theorems $C3$ and $C4$, p. 10-12 in [136]) (and also for smooth enough stationary Gaussian fields).

Finally let us mention the last refinements of Berman inequality (1.20) given by Li and Shao (cf. [98], 2002), which provide an upper bound in (1.20), cleared of the term $(1-\rho_{ij}^2)^{-1/2}$:

Theorem 1.1.17 *(Li and Shao, 2002)*
Suppose that $(X_{1i}, 1 \leq i \leq n)$ are standard normal random variables with covariance matrix $\Lambda_1 = (\Lambda_{i,j}^1)$ and $(X_{2i}, 1 \leq i \leq n)$ similarly with covariance matrix $\Lambda_2 = (\Lambda_{i,j}^2)$;
let $\rho_{ij} = \max(|\Lambda_{i,j}^1|, |\Lambda_{i,j}^2|)$ and $\rho = \max_{i \neq j} \rho_{ij}$. Then, for any real numbers $x_1, \cdots, x_n$,

$$\mathbb{P}[X_{1j} \leq x_j, 1 \leq j \leq n] - \mathbb{P}[X_{2j} \leq x_j, 1 \leq j \leq n] \leq \qquad (1.21)$$

$$\frac{1}{2\pi} \sum_{1 \leq i < j \leq n} \left(\arcsin(\Lambda_{i,j}^1) - \arcsin(\Lambda_{i,j}^2)\right)^+ \exp\left\{-\frac{x_i^2 + x_j^2}{2(1+\rho_{ij})}\right\}$$

Moreover, for $n \geq 3$, for any positive real numbers $x_1, \cdots, x_n$, and when assuming that $\Lambda_{i,j}^2 \geq \Lambda_{i,j}^1 \geq 0$ for all $1 \leq i, j \leq n$, then

$$I\!P[X_{1j} \leq x_j, 1 \leq j \leq n] \leq I\!P[X_{2j} \leq x_j, 1 \leq j \leq n] \tag{1.22}$$

$$\leq \; I\!P[X_{1j} \leq x_j, 1 \leq j \leq n] \exp\left\{ \sum_{1 \leq i < j \leq n} \ln\left(\frac{\pi - 2\arcsin(\Lambda_{i,j}^1)}{\pi - 2\arcsin(\Lambda_{i,j}^2)} \right) \exp\left\{ -\frac{x_i^2 + x_j^2}{2(1 + \Lambda_{i,j}^2)} \right\} \right\}$$

For other precise versions and extensions of this method, we can also refer to e.g. Leadbetter et al. (1983, [94]), Tong (1990, [173]), Ledoux and Talagrand (1991, [96]) and Lifshits (1995, [99]).

• *The method of moments, also called Rice method.*

This method, introduced by Rice to estimate the distribution of the maximum of a random signal, consists in using the first two moments of the number of crossings to estimate the probability of exceeding some given level by a trajectory of a (Gaussian) process, as it appears in the lemma below (see Piterbarg, 1996, [136] p.27 and chap.3). In particular it relies on the fact that the event $(X_0 < x, \max_{0 \leq s \leq t} X_s > x)$ implies the event that there is at least one upcrossing : $(N_t^+(x) \geq 1)$, knowing that the probability of more than one up/down-crossing of the level x become smaller as the level becomes larger.

This method works only for smooth processes, but can be extended to non-stationary Gaussian processes (see Rudzkis, [150] and [151], 1985) and to non-Gaussian processes.

Let $X = (X_s, s \in [0,t])$ be a.s. continuously differentiable with one-dimensional densities bounded. Then

Lemma 1.1.4

$$0 \; \leq \; I\!E[N_t^+(x)] + I\!P[X_0 \geq x] - I\!P[\max_{0 \leq s \leq t} X_s \geq x]$$

$$\leq \; \frac{1}{2}\left(I\!E[N_t^+(x)\left(N_t^+(x) - 1\right)] + I\!E[N_t^-(x)\left(N_t^-(x) - 1\right)] \right) + I\!P[X_0 \geq x, X_t \geq x],$$

where $N_t^-(x)$ denotes the number of downcrossings of level x by X on $[0,t]$ (recall that X_s is said to have a downcrossing of x at $s_0 > 0$ if for some $\varepsilon > 0$, $X_s \geq x$ in $(s_0 - \varepsilon, s_0)$ and $X_s \leq x$ in $(s_0, s_0 + \varepsilon)$).

Indeed

$$I\!P[\max_{0 \leq s \leq t} X_s \geq x] = I\!P[X_0 \geq x] + I\!P[X_0 < x, \max_{0 \leq s \leq t} X_s > x]$$

and

$$I\!P[X_0 < x, \max_{0 \leq s \leq t} X_s > x] \; = \; I\!P[X_0 < x, N_t^+(x) = 1] + I\!P[X_0 < x, N_t^+(x) \geq 2]$$

$$= \; I\!P[N_t^+(x) = 1] - I\!P[X_0 \geq x, N_t^+(x) = 1] + I\!P[X_0 < x, N_t^+(x) \geq 2]$$

$$= \; I\!P[N_t^+(x) = 1] + I\!P[N_t^+(x) \geq 2] - I\!P[X_0 \geq x, N_t^+(x) \geq 1],$$

with $I\!P[N_t^+(x) = 1] = I\!E[N_t^+(x)] - \sum_{k=2}^{\infty} k\, I\!P[N_t^+(x) = k]$

and $I\!P[X_0 \geq x, N_t^+(x) \geq 1] \leq I\!P[X_0 \geq x, X_t \geq x] + I\!P[N_t^-(x) \geq 2]$.

Now we can conclude since we also have $I\!P[N_t^-(x) \geq 2] \leq I\!E[N_t^-(x)\left(N_t^-(x)-1\right)]/2$ and

$$\sum_{k=2}^{\infty} k\, I\!P[N_t^+(x)=k] - I\!P[N_t^+(x)\geq 2] \leq I\!E[N_t^+(x)\left(N_t^+(x)-1\right)]/2. \;\square.$$

Estimates proposed previously for the (factorial) moments can then be used at this step of calculation.

Recently, Azaïs and Wschebor (2001, [14]) adapted this method to express the distribution of the maximum of a one-parameter stochastic process on a fixed interval (in particular in the Gaussian case) by means of a series (called Rice series) whose terms contain the factorial moments of the number of upcrossings, and which converges for some general classes of Gaussian processes, making the Rice method very attractive also for numerical purpose.

1.1.2 Crossings and local time

Representations of the number of crossings

Let $X = (X_s, s \geq 0)$ be a stationary Gaussian process defined on a probability space $(\Omega, \mathcal{F}, I\!P)$, with mean zero, variance one and correlation function r such that $-r''(0) = 1$. We are first interested in having a representation of the number of x-crossings of X as a sum of multiple Wiener-Itô integrals or in terms of Hermite polynomials, where the nth Hermite polynomial H_n can be defined as

$$exp(tx - \frac{t^2}{2}) = \sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!}$$

or as

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n}(e^{-x^2/2}), \qquad n \geq 0.$$

Let W be the standard Brownian Motion (or Wiener process).
Let $H(X)$ denotes the space of real square integrable functionals of the process X.
Recall that $H(X) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$, $\mathcal{H}_n$ being the Wiener Chaos, i.e. the closed linear subspace of $L^2(\Omega, \mathcal{F}, P)$ generated by the random variables $\{H_n(W(h)), h \in L^2(I\!R, dx), ||h|| = 1\}$, where $W(h)$ is the stochastic integral of h with respect to W and $\mathcal{H}_0$ is the set of real constant functions.
We can make use as well of the multiple Wiener-Itô integral I_n defined as in Major (1981, [105]), since we have

$$H_n(W(h)) = I_n(h^{\otimes n}) \qquad \forall h \in L^2(I\!R, dx), \; s.t. \; ||h|| = 1. \tag{1.23}$$

This integral operator I_n satisfies the multiplication rule, namely :
for $f_p \in L_s^2(I\!R^p, m^p)$ and $g_q \in L_s^2(I\!R^q, m^q)$, with

$$L_s^2(I\!R^n, m^n) := \big\{ f_n \in L^2(I\!R^n, \mathcal{B}(I\!R^n), m^n) : \tag{1.24}$$

$$f_n(\underline{\lambda}) = \overline{f_n(-\underline{\lambda})}, \; f_n(\underline{\lambda}) = f_n(\lambda_{p(1)}, \cdots, \lambda_{p(n)}), \forall p \in S_n \big\},$$

m^n denoting the product Borel measure on $I\!R^n$, S_n the symetric group of permutations of $\{1, \cdots, n\}$ and $\underline{\lambda} = (\lambda_1, \cdots, \lambda_n)$, then

$$I_p(f_p).I_q(g_q) = \sum_{k=0}^{p \wedge q} \frac{k!(p+q-2k)!}{p!q!} C_p^k C_q^k I_{p+q-2k}(f_p \ \hat{\otimes}_k \ g_q),$$

where $f_p \hat{\otimes}_k g_q$ denotes the average over all permutations of λ-arguments of the function

$$\int_{I\!R^k} f_p(\lambda_1, \cdots, \lambda_{p-k}, x_1, \cdots, x_k) g_q(\lambda_{p-k+1}, \cdots, \lambda_{p+q-2k}, -x_1, \cdots, -x_k) m^k(d\underline{x})$$

and $p \wedge q \equiv min(p, q)$.

We can introduce the Sobolev spaces $I\!D^{2,\alpha}$ for $\alpha \in I\!R$ as in Watanabe (1984, [176]). A functional $f \in H(X)$ with the development

$$f = I\!E[f] + \sum_{n=1}^{\infty} I_n(f_n) = \sum_{n=0}^{\infty} I_n(f_n)$$

belongs to $I\!D^{2,\alpha}$ if and only if

$$||f||_{2,\alpha}^2 = \sum_{n=1}^{\infty}(1+n)^\alpha||f_n||_2^2 = ||(I-L)^{\alpha/2}f||_2^2 < \infty,$$

where L is the operator defined on $H(X)$ by $Lf = \sum_{n=0}^{\infty} -nI_n(f_n)$.

L coincides with the infinitesimal generator of the Ornstein-Uhlenbeck semigroup $\{T_u, u \geq 0\}$ of contraction operators on $H(X)$ defined by $T_u(f) = \sum_{n=1}^{\infty} e^{-nu}I_n(f_n)$, for any $f = \sum_{n=0}^{\infty} I_n(f_n)$.

In particular we have, for $\beta > 0$,

$$(I-L)^{-\beta} = \frac{1}{\Gamma(\beta)} \int_0^\infty e^{-u}u^{\beta-1}T_u du. \tag{1.25}$$

(cf. Watanabe (1984, [176] p.24), or Nualart (1995, [124] §1.4).

• *Slud (1991, 1994) : MWI integral expansion.*
Multiple Wiener-Itô integrals (MWI) may be a tool to represent and to study nonlinear functionals of stationary Gaussian processes, as noticed by Kalliampur (1980, [83], chap.8). Slud applied the stochastic calculus of MWI integral expansions, first in 1991 (see [161]) to express the number of crossings of the mean level by a stationary Gaussian process within a fixed time interval $[0, t]$, the motivation being to obtain probabilistic limit theorems for crossings-counts, then in 1994 (see [162] or [163]) to extend his results to $\mathcal{C}^1$-curve crossings.

Theorem 1.1.18 *(Slud, 1991 and 1994)*
Let X be a mean zero, variance one, stationary Gaussian process with continuous spectral measure, and twice-differentiable correlation function r.
If $var\,(N_t(\psi)) < \infty$, then

$$N_t(\psi(y)) = \mathbb{E}[N_t(\psi(y))] + \sum_{n=1}^{\infty} I_n(F_n) \quad in \ \ L^2(\Omega), \quad where$$

$*$ *in the case of a given level x (i.e. $\psi(y) = x, \ \forall y$), the mean of $N_t(x)$ is given by the Rice formula (1.1) and*

$$F_n(\underline{\lambda}) = \frac{e^{-\frac{x^2}{2}}}{\pi} \int_0^t e^{is(\lambda_1 + \cdots + \lambda_n)} \times \tag{1.26}$$

$$\sum_{l=0}^{[\frac{n}{2}]} (-r''(0))^{\frac{1}{2}-l} H_{n-2l}(x) \frac{(-1)^{l+1} H_{2l}(0)}{(2l)!(2l-1)} \sum_{1 \leq m_1 < \cdots < m_{2l} \leq n} \lambda_{m_1} \cdots \lambda_{m_{2l}} ds,$$

$*$ *and in the case of a $\mathcal{C}^1$-curve, the mean of $N_t(\psi)$ is given by the generalized Rice formula (1.7) and*

$$F_n(\underline{\lambda}) = \int_0^t e^{is(\lambda_1 + \cdots + \lambda_n)} \frac{e^{-\frac{u^2}{2}}}{\pi} \left(\sqrt{-r''(0)} H_n(u) - \right. \tag{1.27}$$

$$\left. \sum_{l=1}^{n} \frac{i^l}{l!} H_{n-l}(u) \sum_{1 \leq m_1 < \cdots < m_l \leq n} \lambda_{m_1} \cdots \lambda_{m_l} \int_0^{(-r''(0))^{-\frac{1}{2}}} e^{-\frac{z^2 y^2}{2}} H_l(-zy) y^{l-2} dy \right)_{u=\psi(s), z=\psi'(s)} ds,$$

where $\underline{\lambda}$ is a n-vector of coordinates λ_i.

Note that the functional $N_t(\psi)$ may thus be expressed as the integral on $[0,t]$ w.r.t. Lebesgue measure, of $e^{is(\lambda_1 + \cdots + \lambda_n)}$ multiplied by the formal MWI expansion of the form $\sum_n I_n(f_n(.,s))$.

The proof is mainly based on the method of discrete approximation of Cramér and Leadbetter ([40]), already used to obtain the generalized Rice formula (1.7), by introducing the discrete-time number of crossings $N_\psi(1, 2^{-m})$ defined in (1.8), which increases to $N_\psi(1)$. Since by hypothesis in the theorem $var\,(N_\psi(1)) < \infty$, so is the limiting variance of $N_\psi(1, 2^{-m})$ (via the monotone convergence theorem); then, because of the orthogonal decomposition of $L^2(\Omega)$, the MWI integrands for $N_\psi(1)$ are the L^2 (and *a.e.*) limits of the corresponding integrands for $N_\psi(1, 2^{-m})$.
To provide the MWI integrands for $N_\psi(1, 2^{-m})$, some work on indicators of the type $\mathbb{1}_{(X_0>c)}$ or $\mathbb{1}_{(X_0>a)}\mathbb{1}_{(X_h>b)}$, is needed, after having noted that $\mathbb{1}_{((X_s-a)(X_{s+h}-b)<0)} = \mathbb{1}_{(X_s \geq a)} + \mathbb{1}_{(X_{s+h} \geq b)} - 2\mathbb{1}_{(X_s \geq a)}\mathbb{1}_{(X_{s+h} \geq b)}$.
The main new technical tools used for the study of the indicators are a generalization of the Hermite polynomial expansion for the bivariate-normal density in (1.28) and the identity (1.29) given below, because of their own interest.

Lemma 1.1.5 *(Slud, 1994)*
$\forall x, y \in I\!\!R, \ k, m, n \in I\!\!N$ *and* $|t| < 1,$

$$\sum_{j=0}^{\infty} \frac{t^j}{j!} H_{k+j}(x) H_{m+j}(y) = \frac{(-1)^{k+m} e^{\frac{x^2+y^2}{2}}}{\sqrt{1-t^2}} \frac{\partial^{k+m}}{\partial x^k \partial y^m} e^{-\frac{x^2+y^2-2xyt}{2(1-t^2)}} \tag{1.28}$$

and

$$\left(-\frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right)^n e^{-\frac{x^2+y^2-2xyt}{2(1-t^2)}} = \left(\frac{2}{1+t} \right)^{n/2} H_n \left(\frac{x+y}{\sqrt{2(1+t)}} \right) e^{-\frac{x^2+y^2-2xyt}{2(1-t^2)}}. \tag{1.29}$$

The case of a constant level is simply deduced from the general case.
Note that Slud used a different method in 1991, when considering directly a constant level, based mainly on properties of generalized hypergeometric functions; indeed he expressed $I\!\!I_{((X_s-x)(X_{s+h}-x)<0)}$ as the sum $I\!\!I_{(X_s-x<0)} I\!\!I_{(X_{s+h}-x>0)} + I\!\!I_{(X_s-x>0)} I\!\!I_{(X_{s+h}-x<0)}$, obtained the MWI expansion for the indicator $I\!\!I_{(X_s x>0)}$ by using first the Hermite polynomial expansion of this indicator, then by studying the asymptotical behavior of hypergeometrical functions, then used the Diagram theorem to express the products of expansions as a sum of Wiener-Itô integrals (see Dobrushin, 1979 [53] or Major, 1981 [105], p.42; another version will also be given and recalled below, in terms of Hermite polynomials, by Arcones, 1994 [5]), and finally used Fourier transforms for the computations.

• *Kratz and León (1997) : Hermite polynomial expansion.*
At the end of the 1990's, Kratz and León (cf. [88]) proposed a new and direct method to get, under some assumptions on the spectral moments of the process, the Hermite polynomial expansion of crossings of any level by a stationary Gaussian process :

Proposition 1.1.2 *(Kratz and León, 1997)*
Let X be a mean zero stationary Gaussian process, with variance one, satisfying

$$-r''(0) = 1 \ and \ r^{(iv)}(0) < \infty. \tag{1.30}$$

Then the following expansion holds in $L^2(\Omega)$

$$N_t(x) = \sum_{q=0}^{\infty} \sum_{l=0}^{[\frac{q}{2}]} b_{q-2l}(x) a_{2l} \int_0^t H_{q-2l}(X_s) H_{2l}(\dot{X}_s) ds, \tag{1.31}$$

where $b_k(x) := \frac{1}{k!\sqrt{2\pi}} e^{-x^2/2} H_k(x).$

This approach is based on an analytical formula involving the "Dirac function", which defines formally the number of crossings $N_t(x)$ as

$$N_t(x) = \int_0^t \delta_x(X_s) |\dot{X}_s| ds \tag{1.32}$$

43

and which can be made precise when approximating the Dirac function; it makes then explicit formulas for MWI expansions much easier to obtain than was true in the papers of Slud, mainly because expanding $|\dot{X}_s|$ in Hermite polynomials in $\dot{X}_s$ rather than in X_s quite simplify the calculations (note that at s fixed, X_s and $\dot{X}_s$ are independent).

Note that the condition (1.30), stronger than the Geman condition (1.9), can from now on be replaced by the only Geman condition because of Kratz and León's recent result (Theorem 1.1.9), which makes even more attractive the chosen method.

Moreover this approach is natural in the sense that formally the Dirac function δ_x has the generalized Hermite expansion $\delta_x(u) = \sum_{k=0}^{\infty} b_k(x) H_k(u)$ with $b_k(x) = \int_{-\infty}^{\infty} \delta_x(y) \frac{1}{k!} H_k(y) \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = \frac{1}{k!} H_k(x) \frac{e^{-x^2/2}}{\sqrt{2\pi}}$; then (1.32) has the corresponding development given by Lemma 1.1.6 below, made precise by approximating δ_x by $\Phi'_{\sigma,x} := \varphi_{\sigma,x}$.

Lemma 1.1.6 *(Kratz and León, 1997)*
Let X satisfying the conditions of Proposition 1.1.2.
Let $f \in L^2(\phi(x)dx)$ and let $(c_k , k \geq 0)$ be its Hermite coefficients. One has the following expansion

$$\int_0^t f(X_s)|\dot{X}_s|ds = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} c_k a_{2l} \int_0^t H_k(X_s) H_{2l}(\dot{X}_s) ds$$

$$= \sum_{q=0}^{\infty} \sum_{l=0}^{[q/2]} c_{q-2l} a_{2l} \int_0^t H_{q-2l}(X_s) H_{2l}(\dot{X}_s) ds,$$

where $(a_k , k \geq 0)$ are the Hermite coefficients of the function $|.|$, defined by $a_0 = \left(\frac{2}{\pi}\right)^{1/2}$ and $a_{2l} = \left(\frac{2}{\pi}\right)^{1/2} \frac{(-1)^{l+1}}{2^l l!(2l-1)}$ if $l \geq 1$.

By defining $\zeta^{K,L} = \sum_{k=0}^{K} \sum_{l=0}^{L} c_k a_{2l} \int_0^t H_k(X_s) H_{2l}(\dot{X}_s) ds$, we note that $(\zeta^{K,L})_{K,L}$ is a Cauchy sequence in $L^2(\Omega)$ and we deduce from the Hermite expansions of $|x|$ and $f(x)$ that $\zeta^{K,L}$ converges to $\int_0^t f(X_s)|\dot{X}_s|ds$ in $L^2(\Omega)$. Then to conclude the proof of the lemma, just notice that the second expansion is a consequence of the orthogonality. $\square$

Let us be more explicit about the proof of Proposition 1.1.2, by giving the main steps. We will apply the previous result (Lemma 1.1.6) to the function

$$f := \varphi_{\sigma,x}(y) = \sum_{k=0}^{\infty} b_k^{\sigma}(x) H_k(y), \quad \text{with} \quad b_k^{\sigma}(x) = \frac{1}{2\pi\sigma k!} \int_{-\infty}^{\infty} e^{-z^2/2} e^{-\frac{(z-x)^2}{2\sigma^2}} H_k(z) dz,$$

44

to get the Hermite expansion of $N_t^\sigma(x) := \int_0^t \varphi_{\sigma,x}(X_s)|\dot{X}_s|ds$, namely, in $L^2(\Omega)$:

$$N_t^\sigma(x) = \sum_{q=0}^\infty \sum_{l=0}^{[q/2]} b_{q-2l}^\sigma(x)a_{2l} \int_0^t H_{q-2l}(X_s)H_{2l}(\dot{X}_s)ds.$$

Let $\hat{N}_t(x) = \sum_{q=0}^\infty \sum_{l=0}^{[q/2]} b_{q-2l}(x)a_{2l} \int_0^t H_{q-2l}(X_s)H_{2l}(\dot{X}_s)ds$,

with $b_k(x) := \lim_{\sigma \to 0} b_k^\sigma(x) = \frac{1}{k!}\varphi(x)H_k(x)$.
Now we can write

$$I\!E[(N_t(x) - \hat{N}_t(x))^2] \leq 2(I\!E[(N_t(x) - N_t^\sigma(x))^2] + I\!E[(N_t^\sigma(x) - \hat{N}_t(x))^2])$$

and prove by Fatou's lemma and by Jensen inequality that $E[(N_t(x) - N_t^\sigma(x))^2] \to 0$ and via the chaos decomposition that $E[(N_t^\sigma(x) - \hat{N}_t(x))^2]$, as $\sigma \to 0$, to conclude to Proposition 1.1.2.
□

By using now the method of regularization of Wschebor, the result of Proposition 1.1.2 can be extended to a larger class of processes :

Proposition 1.1.3 *(Kratz and León, 1997).*
Let X be a mean zero stationary Gaussian with variance one and satisfying the Geman condition (1.9). In addition we will assume that

$$|\theta''(t+h) - \theta''(t)| \leq |h|L_1(h), \quad where \quad \theta''(t) := tL(t), \tag{1.33}$$

$L_1(h)$ being an even function belonging to $L^1([0,\delta], dx)$.
Then the following expansion holds

$$\frac{N_t(x)}{\sqrt{-r''(0)}} = \sum_{q=0}^\infty \sum_{l=0}^{[\frac{q}{2}]} b_{q-2l}(x)a_{2l} \int_0^t H_{q-2l}(X_s)H_{2l}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right)ds. \tag{1.34}$$

Indeed, Wschebor's regularization method allows us to drop the strong condition involving the fourth derivative of the covariance function of the process to replace it by a new smoother condition, which could be named "uniform Geman condition", constituted by the conditions (1.9) and (1.33) together, in the following way.
By introducing the regularised process X^ε defined by

$$X_t^\varepsilon = \frac{1}{\varepsilon}\int_{-\infty}^\infty \varphi\left(\frac{t-s}{\varepsilon}\right)X_s ds, \quad \text{where } \varphi \text{ is a } \mathcal{C}^2 \text{ function with support in } [-1,1],$$

we can check that Proposition 1.1.2 applies to the number $N_t^\varepsilon(x)$ of x-crossings associated to X^ε, then we can prove that, under the uniform Geman condition, $N_t^\varepsilon(x) \to N_t(x)$ in L^2 as $\varepsilon \to 0$, to conclude on Proposition 1.1.3. Note that for this last convergence, an important step is the use of the diagram formula that we already mentioned (Major, 1981) to show that the

45

partial finite developments of $\dfrac{N_t^\varepsilon(x)}{\sqrt{-r_\varepsilon''(0)}}$ converge to the same developments in the right hand side of (1.34), from which we deduce that for each fixed q

$$\sum_{l=0}^{[\frac{q}{2}]} b_{q-2l}(x) a_{2l} \int_0^t H_{q-2l}\left(\frac{X_s^\varepsilon}{\sigma_\varepsilon}\right) H_{2l}\left(\frac{\dot{X}_s^\varepsilon}{\sigma_\varepsilon\sqrt{-r''(0)}}\right) ds \rightarrow$$

$$\sum_{l=0}^{[\frac{q}{2}]} b_{q-2l}(x) a_{2l} \int_0^t H_{q-2l}(X_s) H_{2l}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right) ds, \text{ in probability as } \varepsilon \to 0. \quad \square$$

To be more explicit, let us give a version of Major Diagram formula in terms of Hermite polynomials, which provides the expectations of product of Hermite polynomials over a Gaussian vector, version which can be found in Breuer and Major (1983, [31]) or in Arcones (1994, [5]). We need to introduce some notations.

Let $G := \{(j,l) : 1 \le j \le p, 1 \le l \le l_j\}$ be the diagram of order $(l_1, \cdots, l_p)$, $V(G)$ the set of vertices (j,l) of the diagram G, $\Gamma\{l_1, \cdots, l_p\}$ the set of diagrams of order $(l_1, \cdots, l_p)$, $L_j = \{(j,l) : 1 \le l \le l_j\}$ the jth level of G, $\{((i,l),(j,m)) : 1 \le i < j \le p, 1 \le l \le l_i, 1 \le m \le l_j\}$ the set of edges $((i,l),(j,m))$ s.t. every vertex is on only one edge. The edges connect vertices of different levels. Given an edge $w = ((i,l),(j,m))$, let $d_1(w) = i$, $d_2(w) = j$.

Theorem 1.1.19 : *Diagram formula*
Let $(X_1, \cdots, X_p)$ *be a Gaussian vector mean zero, variance 1 and with* $\mathbb{E}[X_i X_j] = r(i,j)$, $1 \le i,j \le p$. *Then*

$$\mathbb{E}\left[\prod_{i=1}^p H_{l_i}(X_i)\right] = \sum_{G \in \Gamma\{l_1, \cdots, l_p\}} \prod_{w \in V(G)} r(d_1(w), d_2(w)).$$

• Let us come back on the heuristic formula (1.32) for the number of crossings.
Suppose that $X = (X_s, s \ge 0)$ is a mean zero stationary Gaussian process with variance one, that the function of covariance r has two derivatives and satisfies (1.30).
Let $0 = \alpha_0 < \alpha_1 < \ldots < \alpha_{k-1} < \alpha_k = t$ be the points where the change of sign of the derivative of X occurs. They are in a finite number because the process has a finite fourth spectral moment (condition (1.30)).
There is a x-crossing of X between α_i and α_{i+1} if

$$\left|\mathbb{1}_{(X_{\alpha_{i+1}}>x)} - \mathbb{1}_{(X_{\alpha_i}>x)}\right| = 1,$$

which implies that

$$N_t(x) = \sum_{i=0}^{k-1} \left|\mathbb{1}_{(X_{\alpha_{i+1}}>x)} - \mathbb{1}_{(X_{\alpha_i}>x)}\right|,$$

$$= \sum_{i=0}^{k-1} \left|Y_x(X_{\alpha_{i+1}}) - Y_x(X_{\alpha_i})\right|,$$

46

where $Y_x(u) = \mathbb{1}_{[x,\infty)}(u)$ is the Heaviside function whose the generalized derivative is the "Dirac function" $\delta_x(u) = \infty$ if $u = x$ and 0 if not.

So we can write formally

$$N_t(x) = \sum_{i=0}^{k-1} \int_{\alpha_i}^{\alpha_{i+1}} \delta_x(X_s)|\dot{X}_s|ds,$$

which gives the formula (1.32).

Note also that (1.32) allows to get back formally some classical results, as :

- the Rice formula :

$$\mathbb{E}[N_t(x)] = \int_0^t \mathbb{E}[\delta_x(X_s)]\mathbb{E}\mid \dot{X}_s \mid ds = te^{-x^2/2}\sqrt{-r''(0)}/\pi ;$$

- more generally if g is a positive function on $\mathbb{R}$, if G is a primitive of g, then (cf. Cabaña (1985), [34])

$$\int_{\mathbb{R}} N_t(x)g(x)dx = \int_0^t \left[\int_{\mathbb{R}} \delta_{X_s}(x)g(x)dx\right]\mid \dot{X}_s \mid ds = \int_0^t g(X_s)\mid \dot{X}_s \mid ds$$

$$= \sum_{i=0}^{k-1}\mid G(X_{\alpha_{i+1}}) - G(X_{\alpha_i})\mid .$$

- the Kac formula :
  since $\hat{\delta}_0(t) = 1$, by applying formally the Fourier inversion formula, we have $\delta_0(u) = \frac{1}{2\pi}\int_{\mathbb{R}} cos(tu)dt$,

$$N_t(0) = \int_0^t \delta_0(X_s)\mid \dot{X}_s \mid ds = \frac{1}{2\pi}\int_{\mathbb{R}}\int_0^t cos(\xi X_s)\mid \dot{X}_s \mid d\xi ds.$$

• *Kratz (2000)*

If we choose for a Gaussian process the Brownian motion B, then its number of x-crossings is such that a.s. either $N_t(x) = \infty$ or $N_t(x) = 0$, but we can define its local time $L_t(x)$ formally by

$$L_t(x) = \int_0^t \delta(x)(B_s)ds,$$

that we can make precise by a Gaussian approximation (see the array below) or by a uniform one, as $L_t(x) = \lim_{\varepsilon\to 0}\int_0^t \frac{1}{2\varepsilon}\mathbb{1}_{[-\varepsilon,\varepsilon]}(B_s)ds$ p.s. (cf. for instance [60]).

More generally, let us define the local time $L_t(x)$ of a Gaussian process X as the density of the occupation measure of X, as for instance in Berman ([23]). Its construction by limiting processes is based on the sample path properties of X, as first introduced by Levy in the case of the Brownian motion (see for instance [60]).

We may then notice that the notions of number of crossings and local time present some "analogies" in their formulae (heuristic and non heuristic), as it appears in the following array :

47

Local time L_t^x for $\tilde{X} = (\tilde{X}_s, s \geq 0)$ (for which L_t^x does exist)	Crossings number N_t^x for $X = (X_s, s \geq 0)$ (for which N_t^x does exist)
Sojourn time of $\tilde{X}$ on a level x in $[0, t]$: $$S_x(t) = \int_0^t \mathbb{1}_{(\tilde{X}_s \geq x)} ds := \int_0^t Y_x(\tilde{X}_s) ds$$	
Local time = density of $S_x(t)$	Number of crossings

Formally

$$L_t^x = \int_0^t \delta_x(\tilde{X}_s) ds \qquad \qquad N_t^x = \int_0^t \delta_x(X_s)|\dot{X}_s| ds$$

Mathematically
the Heaviside function Y_x and the "Dirac function" δ_x are approximated
respectively by the distribution function and the density of a Gaussian
r.v. mean x and variance σ^2, with $\sigma \to 0$.

Moreover, $\forall f \in \mathcal{C}^\infty$,

$$\int_0^t f(\tilde{X}_s) ds = \int_{-\infty}^\infty f(x) L_t^x dx \qquad \qquad \int_0^t f(X_s)|\dot{X}_s| ds = \int_{-\infty}^\infty f(x) N_t(x) dx$$

Banach-Kac (1925-43) formula
(cf. [16] and [82])

Because of this type of correspondance with the local time, a new goal is to 'validate mathematically' the heuristic formula (1.32) for crossings-counts by choosing the right space, as it has been done for Brownian local time (see Nualart and Vives ([125],[126]), Imkeller et al. ([78])). An appropriate framework is the Sobolev spaces $I\!D^{2,\alpha}$ over canonical Wiener space obtained via completion of the set of polynomials F with respect to the norms $||F||_{2,\alpha} = ||(I - L)^{\alpha/2} F||_2$, $\alpha \in I\!R$.

Whereas the distribution $\delta_x(X_s)$ and the random variable $|\dot{X}_s|$ belong to $I\!D^{2,\alpha}$ for any $\alpha < -1/2$ and for any $\alpha \in I\!R$ respectively, the integral in the definition of $N_t(x)$ would appear to have a smoothing effect, showing $N_t(x)$ as a random variable which would belong to $I\!D^{2,\alpha}$ for any $\alpha < 1/4$ (cf. [87]). The proof uses the Malliavin calculus, quite powerful tool when dealing with such non-smooth functionals.

Thus we should be able to complete the previous array with the following line :

$L_t^x \in \left\{ \begin{array}{ll} I\!\!D^{2,\alpha}, & 0 < \alpha < 1/2 \\ I\!\!D^{p,\alpha}, & \forall p \geq 2,\ 0 < \alpha < 1/2 \end{array} \right.$	$N_t(x) \in I\!\!D^{2,\alpha},\ \forall 0 < \alpha < 1/4$
Nualart & Vives (1992), Imkeller et al. (1995)	Kratz (2000)

More explicitely, let $X = (X_t; t \geq 0)$ be a real stationary Gaussian process with mean zero, variance one. Assume that its covariance function r is twice differentiable and that $-r''(0) = 1$. r has the well-known representation

$$r(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda) \ , \ -\infty < t < \infty,$$

where F is the spectral distribution.
Suppose that F is absolutely continuous, then $r(t)$ has also the representation

$$r(t) = \int_{-\infty}^{\infty} b(t+s)b(s)ds$$

for some $b \in L^2$, which can be chosen as the Fourier transform in L^2 of $(F'(\lambda))^{1/2}$, F' being the Radon-Nikodym derivative of F with respect to $d\lambda$ (see Berman [23], p.149), and X has the stochastic integral representation

$$X_t = \int_{-\infty}^{\infty} b(t+s)W(ds)$$

where W is the standard Brownian Motion (see Berman [23], p.157).
We have

$$X_t = W(h_t) \ , \ \text{with } h_t(\cdot) = b(t+\cdot) \in L^2(I\!\!R).$$

Note that $\forall t, \ ||h_t||_2 = 1$ (since we suppose $r(0) = 1$).
The following results can be established. Let p_ε be the centered Gaussian density with variance ε. On one hand,

Proposition 1.1.4 *(Kratz, 2000)*
$\delta_x(X_s) = \lim_{\varepsilon \to 0} p_\varepsilon(X_s - x)$ *exists in* $I\!\!D^{2,\alpha}$ *for any* $\alpha < -1/2$, *and is given by the series*

$$
\begin{aligned}
\delta_x(X_s) &= \sum_{n=0}^{\infty} c_n(x)H_n(X_s) = \sum_{n=0}^{\infty} c_n(x)H_n(W(h_s)) \\
&= \sum_{n=0}^{\infty} c_n(x)I_n(h_s^{\otimes n}),
\end{aligned}
$$

$$\text{with } c_n(x) = \frac{1}{n!}H_n(x)p_1(x), \forall n \geq 0. \tag{1.35}$$

On the other hand,

49

Proposition 1.1.5 *(Kratz, 2000)*
$| \dot{X}_s |$ exists in $\mathbb{D}^{2,\alpha}$ for any α and is given by the series

$$| \dot{X}_s | = \sum_{n=0}^{\infty} a_{2n} H_{2n}(\dot{X}_s) = \sum_{n=0}^{\infty} a_{2n} H_{2n}(W(\dot{h}_s))$$

$$= \sum_{n=0}^{\infty} a_{2n} I_{2n}(\dot{h}_s^{\otimes 2n}),$$

$$\text{with } a_{2n} = \sqrt{\frac{2}{\pi}} \frac{(-1)^{n+1}}{2^n n !(2n-1)}, \forall n \geq 0. \tag{1.36}$$

And finally,
Conjecture *(Kratz, 2000)*
The number of crossings $N_t(x)$ at any level x by the process X on the interval $[0,t]$ is such that

$$N_t(x) = \int_0^t \delta_x(X_s) \, | \dot{X}_s | \, ds \quad \text{exists in } \mathbb{D}^{2,\alpha} \text{ for any } \alpha < 1/4$$

and is given by the series

$$N_t(x) = \sum_{n=0}^{\infty} \sum_{m=0}^{[n/2]} a_{2m} c_{n-2m}(x) \int_0^t H_{n-2m}(X_s) H_{2m}(\dot{X}_s) ds$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{[n/2]} a_{2m} c_{n-2m}(x) \int_0^t I_{n-2m}(h_s^{\otimes n-2m}) I_{2m}(\dot{h}_s^{\otimes 2m}) ds,$$

where $h_s(.) = b(s + .)$, a_{2m} and $c_{n-2m}(x)$ are given respectively in (1.36) and (1.35).

Let us give the main ideas of the proof.

We consider $N_t^\varepsilon(x) := \int_0^t p_{\varepsilon,x}(X_s) \, | \dot{X}_s | \, ds$, with $p_{\varepsilon,x}(y) := p_\varepsilon(y - x)$, and we want to show that $N_t^\varepsilon(x)$ converges to $N_t(x)$ in $\mathbb{D}^{2,\alpha}$, $\forall \alpha < 1/4$.
Fix $0 < \alpha < 1$.
Let us prove that $(I - L)^{\frac{\alpha}{2}} N_t^\varepsilon(x)$ is bounded in $L^2(\Omega)$, uniformly in $\varepsilon > 0$, or equivalently that

$$(I - L)^{\frac{\alpha-1}{2}} D_\theta N_t^\varepsilon(x)$$

is bounded in $L^2(\Omega, L^2(\mathbb{R}, dx))$, uniformly in $\varepsilon > 0$, where D_θ is the derivative operator, $\theta \in \mathbb{R}$.
Let us compute

$$||(I - L)^{\frac{\alpha-1}{2}} D_\theta N_t^\varepsilon(x)||^2_{L^2(\Omega, L^2(\mathbb{R}, dx))} = \mathbb{E}\left[||(I - L)^{\frac{\alpha-1}{2}} D_\theta N_t^\varepsilon(x)||^2_2 \right]$$

$$= E\left[\int_0^t \left| (I - L)^{\frac{\alpha-1}{2}} D_\theta N_t^\varepsilon(x) \right|^2 d\theta \right].$$

We have

$$D_\theta N_t^\varepsilon(x) = \int_0^t D_\theta \left(| \dot{X}_s | \, p_{\varepsilon,x}(X_s) \right) ds$$

$$= \int_0^t D_\theta(| \dot{X}_s |) p_{\varepsilon,x}(X_s) ds + \int_0^t | \dot{X}_s | \, D_\theta \left(p_{\varepsilon,x}(X_s) \right) ds,$$

therefore

$$E\left|(I-L)^{\frac{\alpha-1}{2}}D_\theta N_t^\varepsilon(x)\right|^2 \le 2(E_1(\theta)+E_2(\theta)),$$

with $E_1(\theta)=E_{1,\alpha,\varepsilon}(\theta):=E\left|\int_0^t (I-L)^{\frac{\alpha-1}{2}}\left(D_\theta(|\dot{X}_s|)p_{\varepsilon,x}(X_s)\right)ds\right|^2$

and $E_2(\theta)=E_{2,\alpha,\varepsilon}(\theta):=E\left|\int_0^t (I-L)^{\frac{\alpha-1}{2}}\left(|\dot{X}_s|D_\theta(p_{\varepsilon,x}(X_s))\right)ds\right|^2$,

hence

$$||(I-L)^{\frac{\alpha-1}{2}}D_\theta N_t^\varepsilon(x)||_{L^2(\Omega,L^2(R,dx))}^2 \le 2\int_{I\!R}(E_1(\theta)+E_2(\theta))d\theta.$$

First let us consider $E_1(\theta)$.

We have $E_1(\theta) \le \int_0^t E\left|(I-L)^{\frac{\alpha-1}{2}}\left(D_\theta(|\dot{X}_s|)p_{\varepsilon,x}(X_s)\right)\right|^2 ds.$

Then Proposition 1.2.3 in Nualart ([124], 1995), Lemma 1.1 in Nualart & Vives ([126], 1992) and some new technical lemmas ([87]) allow to prove that $\forall \alpha < 1/2$,

$$\int_{I\!R} E_{1,\alpha,\varepsilon}(\theta)d\theta < \infty \qquad \text{uniformly in } \varepsilon > 0.$$

Now let us consider $E_2(\theta)$ and let $g_\varepsilon(s):=|\dot{X}_s|\dfrac{x-X_s}{\varepsilon}p_{\varepsilon,x}(X_s).$
After some computations, we have that

$$\int_{I\!R} E_2(\theta)d\theta = 2(I_\delta+I_t), \text{ for some } \delta > 0,$$

where

$$
\begin{aligned}
I_t &= \int_\delta^t (t-s)r(s)I\!E\left[g_\varepsilon(s)(I-L)^{\alpha-1}(g_\varepsilon(0))\right]ds = I\!E\left[((I-L)^{\alpha-1}g_\varepsilon(0))\int_\delta^t (t-s)r(s)g_\varepsilon(s)ds\right] \\
&\le E^{1/2}\left[((I-L)^{\alpha-1}g_\varepsilon(0))^2\right]E^{1/2}\left[\left(\int_\delta^t (t-s)r(s)g_\varepsilon(s)ds\right)^2\right],
\end{aligned}
$$

and

$$I_\delta = \int_0^\delta (t-s)r(s)I\!E\left[g_\varepsilon(s)(I-L)^{\alpha-1}(g_\varepsilon(0))\right]ds. \tag{1.37}$$

On what concerns I_t, it can be proved that

$$
\begin{aligned}
I\!E\left[((I-L)^{\alpha-1}g_\varepsilon(0))^2\right] &= E[g_\varepsilon(s)(I-L)^{2(\alpha-1)}(g_\varepsilon(s))] \\
&= \frac{1}{\Gamma(2(1-\alpha))}\int_0^\infty e^{-u}u^{1-2\alpha}E[g_\varepsilon(s)T_u(g_\varepsilon(s))]du
\end{aligned}
$$

and the explicit computation of the right hand side helps to get that

$$\sup_{\varepsilon\to 0} I\!E\left[((I-L)^{\alpha-1}g_\varepsilon(0))^2\right] < \infty \quad \text{whenever} \quad \alpha < 1/4.$$

51

Now, since

$$\mathbb{E}\left[\left(\int_\delta^t (t-s)r(s)g_\varepsilon(s)ds\right)^2\right] = 2t\,\mathbb{E}\left[\int_\delta^t (t-s)^2 r(s)g_\varepsilon(0)g_\varepsilon(s)\right]ds$$

$$= 2t\int_\delta^t (t-s)^2 r(s)\int_{\mathbb{R}^4} |\dot{x}_1||f_\varepsilon(x_1)||\dot{x}_2|f_\varepsilon(x_2)\phi_s(x_1,\dot{x}_1,x_2,\dot{x}_2)dx_1 d\dot{x}_1 dx_2 d\dot{x}_2 ds,$$

some change of variables allows to prove that $\mathbb{E}^{1/2}\left[\left(\int_\delta^t (t-s)r(s)g_\varepsilon(s)ds\right)^2\right] < \infty$ uniformly in $\varepsilon > 0$.

We can conclude that $\forall \alpha < 1/4$, $I_t < \infty$ uniformly in $\varepsilon > 0$.

The conjecture relies on the behavior of I_δ.

If it can be proved that $\mathbb{E}^{1/2}\left[\left(\int_0^\delta (t-s)r(s)g_\varepsilon(s)ds\right)^2\right] < \infty$ uniformly in $\varepsilon > 0$, then the same argument as for I_t applies and we have the enunciated result.

If this assertion isn't true, to consider directly (1.37) without applying Cauchy Schwarz inequality would hopefully lead to the same result. $\square$

Approximation of the local time by the number of crossings

Several authors got interested in the problem of approximating the local time of an irregular process X by the number of crossings of a regularization of this process. One classical regularization is the one obtained by convolution, already presented, defined by

$$X_t^\varepsilon = \psi_\varepsilon * X_t, \quad \text{where} \quad \psi_\varepsilon(.) = \frac{1}{\varepsilon}\psi\left(\frac{\cdot}{\varepsilon}\right) \quad \text{and } \psi \text{ some smooth function.} \tag{1.38}$$

• *Wschebor (1984, 1992).*

Wschebor considered this problem for a specific Gaussian process, the Brownian motion (in the multiparametric case, but we will only present the one parameter case).

Let $W = (W_t, t \geq 0)$ be a Brownian motion and let W_ε be the convolution approximation of W defined by $W_\varepsilon := W * \psi_\varepsilon$, where ψ_ε is defined in (1.38) with ψ, a non-negative $\mathcal{C}^\infty$ function with compact support.

Let $N_\varepsilon^x([a,b])$ be the number of crossings of level x by the regularized process W_ε on the interval $[a,b]$.

Wschebor in [179] showed that, for $x \in \mathbb{R}$,

Theorem 1.1.20 *(Wschebor, 1984)*

$$\sqrt{\frac{\pi}{2}}\frac{\varepsilon^{1/2}}{||\psi||_2}N_\varepsilon^x([a,b]) \xrightarrow{L^k} L(x,[a,b]), \quad as \quad \varepsilon \to 0, \quad for \quad k = 1,2,\cdots \tag{1.39}$$

where $L(x,[a,b])$ is the Brownian motion's local time at level x on $[a,b]$.

Later (cf. [181]), he proved the following :

Theorem 1.1.21 *(Wschebor, 1992)*
For any continuous and bounded function f, for a.e. fixed w,

$$\frac{1}{||\psi||_2}\sqrt{\frac{\varepsilon\pi}{2}}\int_{-\infty}^{\infty}f(x)N_\varepsilon(x)dx \;-\; \int_{-\infty}^{\infty}f(x)L(x)dx \;\rightarrow\; 0, \quad as \quad \varepsilon \rightarrow 0,$$

where $N_\varepsilon(x) := N_\varepsilon^x([0,1])$ and $L(x) := L(x,[0,1])$.

• *Azaïs, Florens-Zmirou (1987, 1990)*
Azaïs and Florens-Zmirou, in their 87 paper ([8]), extended Wschebor's result (1.39) to a large class of stationary Gaussian processes, when considering the L^2 convergence and the zero crossings. Under some technical conditions on the Gaussian stationary process X and its regularization $X_\varepsilon := X * \psi_\varepsilon$ (requiring a bit more than the non-derivability of X and giving some bounds on the second and fourth spectral moments of X_ε), on its correlation function r (namely r twice differentiable outside a neighborhood of zero, with bounded variation at zero, r' and r'' bounded at infinity) and on the convolution kernel ψ (i.e. $\psi(u)$ and $\psi'(u)$ bounded by a constant times $|u|^{-2}$), they proved that

Theorem 1.1.22 *(Azaís and Florens-Zmirou, 1987)*
If X admits a local time $L(x,[0,T])$ continuous en x at zero, then

$$\sqrt{\frac{\pi}{2\lambda_{2,\varepsilon}}}N_\varepsilon^0([0,T]) \;\underset{\varepsilon\to 0}{\rightarrow}\; L(0,[0,T]) \quad in \;\; L^2,$$

where $\lambda_{2,\varepsilon}$ denotes the second spectral moment of X_ε.

Note that Azaís in 1990 ([7]) considered more general stochastic processes and provided sufficient conditions for L^2-convergence of the number of crossings of some smooth approximating process X_ε of X (which converges in some sense to X) to the local time of X, after normalization.

• *Berzin, León, Ortega (1992, 1998), Azaís and Wschebor (1996).*
Whereas Wschebor proved the a.s. convergence of the variable given in the theorem 1.1.21, León and Ortega studied the L^2 convergence of this variable slightly modified, with the rates of convergence, first in 1992 (see [97]), then in a collaboration with Berzin, in 1998 (see [25]), under slightly different hypotheses. Indeed, for $X = (X(t), t \in [0,1])$ a centered stationary Gaussian process which covariance function r satisfies $r(t) \sim_{t\to 0} 1 - C|t|^{2\alpha}$, $0 < \alpha < 1$, they considered the regularized process $X_\varepsilon = \dfrac{\psi_\varepsilon * X}{var(\psi_\varepsilon * X)}$ (where the kernel ψ_ε approaches the Dirac function as $\varepsilon \to 0$), and proved that *the convergence of*

$$\frac{1}{\varepsilon^{a(\alpha)}}\int_{-\infty}^{\infty}f(x)\left(\frac{N_\varepsilon(x)}{c(\varepsilon)}-L(x)\right)dx, \qquad with \;\; c(\varepsilon)=\left(\frac{2}{\pi}\frac{var(\dot{X}_\varepsilon(t))}{var(X_\varepsilon(t))}\right)^{1/2},$$

($f \in L^4(\phi(x)dx)$ satisfying certain regularity conditions), is in L^2 or is the weak convergence, depending on the values of α ($0 < \alpha < 1/4$, $1/4 < \alpha < 3/4$ and $3/4 < \alpha < 1$).

Meantime, Azaïs and Wschebor considered in 1996 (see [11]) the a.s. convergence of $\int_{-\infty}^{\infty} f(x) N_\varepsilon(x) dx$, for any continuous function f and for X belonging to a larger class of Gaussian processes than the previous one of Berzin, León and Ortega.

Note that we didn't mention all the papers on this subject or closely related one; indeed much work on the approximation of local times and occupation measures has been done also on specific Gaussian processes such as the fractional brownian motion (we can quote for instance the work by Révész et al. in the 80's on invariance principles for random walks and the approximation of local times and occupation measures, as well as the work by Azaïs and Wschebor (1997, [12]) on the first order approximation for continuous local martingales). Nevertheless, we chose not to develop that here.

1.1.3   Asymptotic results

Law of small numbers and rate of convergence

• Due to the extension of the Poisson limit theorem from independent to dependent Bernoulli random variables (see Chen, 1975 [38]), there has been the same extension in extreme value theory, in particular for the study of the number of (up-)crossings.
As the level $x = x(t)$ increases with the length t of the time interval, the upcrossings tend to become widely separated in time provided that there is a finite expected number in each interval. Under some suitable mixing property satisfied by the process X to assure that the occurrences of upcrossings in widely separated intervals can be considered as asymptotically independent events, we get a limiting Poisson distribution for the number of x-upcrossings N_x^+ by using the Poisson theorem for dependent random variables (see Volkonskii & Rozanov (1961, [175]), Cramér (1966, [40], chap. 12), Qualls (1968, [140]), Pickands (1969, [130], [131]), Berman (1971, [23]), and, for a survey and some further minor improvement, Leadbetter et al. (1978, [94], chap. 8 and 9). We shall give here a version of such a result, proposed in Leadbetter et al. (Theorem 9.1.2 p. 174).
Under the condition that

$$\mu(\mathbf{x}) := I\!\!E[\mathbf{N}_{\mathbf{x}}^+((\mathbf{0}, \mathbf{1}))] < \infty,$$

then $N_x^+(I) < \infty$ a.s. for bounded I, and the upcrossings form a stationary point process N_x^+ with intensity parameter $\mu = \mu(x)$. The point process of upcrossings has properties analogous to those of exceedances in discrete parameter cases, namely :

Theorem 1.1.23 *(Leadbetter et al.)*
Let assume that the correlation function r of the centered stationary normal process X satisfies

$$r(s) = 1 - \frac{\lambda_2}{2} s^2 + o(s^2) \quad as \quad s \to 0 \tag{1.40}$$

and the Berman condition given by

$$r(s) \log s \underset{s \to \infty}{\to} 0 . \tag{1.41}$$

Suppose that x and t tend to ∞ in a coordinated way such that

$$I\!E N_x^+(t) = t\mu(x) \underset{t\to\infty}{\to} \tau \quad \text{for some constant} \quad \tau \geq 0. \qquad (1.42)$$

Define a time-normalized point process N_*^+ of upcrossings having points at s/t when X has an upcrossing of x at s.
Then N_*^+ converges to a Poisson process with intensity τ as $t \to \infty$.

Note that the asymptotic Poisson character of upcrossings applies also to nondifferentiable normal processes with covariance functions satisfying

$$r(\tau) = 1 - C|\tau|^\alpha + o(|\tau|^\alpha), \quad \text{as} \quad \tau \to 0 \quad (\text{with } 0 < \alpha < 2, \text{ and } C \text{ positive constant}), \qquad (1.43)$$

if we consider ε-upcrossings, introduced by Pickands (1969, [130]) and defined below, instead of ordinary upcrossings (for which under (1.43) their mean number of any level per unit time is infinite).
For a given $\varepsilon > 0$, a process ξ is said to have an ε-upcrossing of the level x at t if $\xi(s) \leq x$, $\forall s \in (t - \varepsilon, t)$ and, $\forall \eta > 0$, $\exists s \in (t, t + \eta)$, $\xi(s) > x$;
so an ε-upcrossing is always an upcrossing while the reverse is not true.
The expectation of the number of ε-upcrossings of x by ξ satisfying (1.43) can be evaluated, and it can be proved that asymptotically this mean number is independent of the choice of ε for a suitably increasing level x, which leads, under the Berman condition and (1.43), to the Poisson result of Lindgren et al. (1975) for the time-normalized point process of ε-upcrossings (see [130] and [94], chap.12 for references and more details).

● Remark : Another notion related to the one of upcrossings of the level x by the process X is the time spent over x, called also sojourn of X above x, defined in the array of the section "Representations of the number of crossings" by $S_x(t) = \int_0^t I\!I_{(\tilde{X}_s \geq x)} ds$. An upcrossing of the level x marks the beginning of a sojourn above x. If there is a finite number of expected upcrossings in each interval, then the number of sojourns above x is the same as the number of upcrossings. Under some mixing conditions (recalled in the result of Bermangiven below) and the assumption of a finite expected number of upcrossings in each interval, Volkonskii & Rozanov (1961, [175]), then Cramér & Leadbetter (1967, [40], under weaker conditions) showed by using the reasoning used in the proof of the Poisson limit for the distribution of upcrossings, that the sojourn above X has a compound Poisson limit distribution (as the sum of a Poisson distributed random number of nearly independent r.v. which are the durations of the sojourns). From the 70's, Berman (see [23] for the review of these topics) proposed an alternative to discussing upcrossings; he introduced a method based on Hermite polynomial expansions to study the asymptotic form of the sojourns of X above a level x, on one hand for $x \to \infty$ with fixed t, on the other hand for $x, t \to \infty$ in a coordinated way. He considered a larger class of processes (with sample functions not necessarily differentiable), allowing a possible infinite expected number of upcrossings in each finite interval, to prove the compound Poisson limit theorem.
More specifically, by using arguments on moments, Berman (see [23], chap.7) proved that :
if the level function $x = x(t)$ satisfies the asymptotic condition $x(t) \sim \sqrt{2 \log t}$ as $t \to \infty$ and for a covariance function r such that $1 - r$ is a regularly varying function of index α ($0 < \alpha \leq 2$) when the time tends to 0, then

the limiting distribution of $v(t)S_x(t)$ is a compound Poisson distribution,
where v is an increasing positive function determined by the asymptotic form of $1 - r(t)$ for
$t \to 0$ and the compounding distribution is uniquely defined in terms of the index of regular
variation of $1 - r(t)$ for $t \to 0$.

• For practical use of the asymptotic theory, it is rather important to know how faithful and accurate these Poisson approximations are, and in view of applications, we must then study the rate of convergence carefully. Many years passed before getting some information about those rates.

Finally, Pickersgill and Piterbarg (cf. [134]) established rates of order $t^{-\nu}$ for the point probabilities in Theorem 1.1.23, but without giving any information on the size of ν.

Kratz and Rootzén (cf. [92]) proposed, under the condition that $r(.)$ and $r'(.)$ decay at a specified polynomial rate (quite weak assumption, even if more restrictive than $r(s) \log s \underset{s \to \infty}{\to} 0$), bounds for moments of the number of upcrossings and also for the rate of convergence in Theorem 1.1.23 roughly of the order $t^{-\delta}$, with

$$\delta = \frac{1}{2} \wedge \inf_{s \geq 0} \rho(s), \tag{1.44}$$

where $\rho(.)$ has been defined by Piterbarg (cf. [135]) by

$$\rho(s) = \frac{(1 - r(s))^2}{1 - r(s)^2 + r'(s)|r'(s)|}. \tag{1.45}$$

Their approach proceeds on one hand by discretization and blocking to go back to discrete parameter cases, and on the other hand by combining the normal comparison method and the Stein-Chen method (cf. [166] and [38]) as developed by Barbour, Holst and Janson (cf. [68] and [17]) (see also Falk, Hüsler and Reiss [56]), this last method imposing to measure the rate of convergence with the total variation distance, defined between integer valued r.v.'s. X and Y by

$$d(X, Y) = \frac{1}{2} \sum_k |I\!P(X = k) - I\!P(Y = k)| = \sup_A |I\!P(X \in A) - I\!P(Y \in A)|.$$

More precisely, the authors got on what concerns the rate of convergence the following result, throughout assuming for convenience that $x \geq 1$ and $t \geq 1$:

Theorem 1.1.24 *(Kratz and Rootzén, 1997)*
Suppose $\{X(s); s \geq 0\}$ is a continuous stationary normal process which satisfies (1.11), (1.12) with $\gamma = 2$, (1.13) and (1.14) and that $r(s) \geq 0$ for $0 \leq s \leq t$. Then there are constants K and K' which depend on $r(s)$ but not on x or t such that

$$d(N_x^+(t), \mathcal{P}(t\mu(x))) \leq K \frac{x^{2+2/\alpha}}{t^\delta} \tag{1.46}$$

and, for $t \geq t_0 > 1$,

$$d(N_x^+(t), \mathcal{P}(t\mu(x))) \leq K' \frac{(\log t)^{1+1/\alpha}}{t^\delta}. \tag{1.47}$$

56

The constants K in (1.46) and K' in (1.47) are specified in the paper of the authors.

Let us present briefly the method used to prove such result.

The discretisation consists in replacing the continuous process $\{X(s); s \geq 0\}$ by a sampled version $\{X(jq); j = 0, 1, ...\}$, with q roughly chosen equal to $s^{-1/2}$, which makes extremes of the continuous time process sufficiently well approximated by extremes of its sampled version. Nevertheless it brings a new problem : we may count too many exceedances. So a new parameter θ is introduced to divide the interval $(0, t]$ into blocks $((k-1)\theta, k\theta]$, $k = 1, ..., t/\theta$ with θ roughly equal to $t^{1/2}$ (that is what is designed by blocking method). This choice of θ makes the blocks long enough to ensure approximate independence of extremes over disjointed blocks.

It brings to consider W the number of blocks with at least one exceedance : $W := \sum_{k=1}^{t/\theta} I_k$, where $I_k := \mathbb{1}_{\left(\max_{jq \in ((k-1)\theta, k\theta]} X(jq) > x\right)}$. Let $\lambda = \mathbb{E}[W]$.

Then by the triangle inequality for the total variation distance d, we have

$$d(N_x(t), \mathcal{P}(t\mu)) \leq d(N_x(t), W) + d(W, \mathcal{P}(\lambda)) + d(\mathcal{P}(\lambda), \mathcal{P}(t\mu)), \tag{1.48}$$

and we estimate the three terms in the righthand side (RHS) of (1.48) separately.

To estimate $d(N_x(t), W)$, we introduce $N_x^{(q)}(\theta)$ the number of exceedances of x by the sampled process $\{X(jq); jq \in (0, t]\}$ upon an interval of length θ, and the essential part concerns the evaluation of the difference between $N_x^{(q)}(\theta)$ and the probability that we have at least one exceedance by the sampled process on the same interval, i.e. $\mathbb{E}[N_x^{(q)}(\theta)] - \mathbb{E}[I_1]$. To evaluate it, we use Lemma 1.1.3.

The last term of the RHS is bounded in the same way as the first one, using that $d(\mathcal{P}(\lambda), \mathcal{P}(t\mu)) \leq |\lambda - t\mu|$ (see e.g. Holst and Janson (1990), [68] p.12), where $\lambda - t\mu = \dfrac{t}{\theta}(\mathbb{E}[I_1] - \theta\mu)$.

To study the middle term of the RHS, we are using the Stein-Chen method described below. It means that we have to evaluate double sums of covariances of indicators which are grouped into blocks, which implies, because of stationarity, the study of $cov(I_1, I_k), k \geq 2$. A version of the normal comparison lemma given in Theorem 1.1.16 (by taking $x = \min(x_i, 1 \leq i \leq n)$) helps to treat the case $k \geq 3$. When $k = 2$, a new parameter θ^* is introduced to assure a quasi-independence between blocks. Then classical normal technics are used.

Concerning references about speed of approximation in related problems, we could mention the works by Borodin and Ibragimov in the 80's and 90' ([30]), by Jacod (1998,[81]), by Perera and Wschebor (1998,[129]), by Berzin and León (1997,[24]), etc ...

• *The Stein-Chen method.*

In 1970, Stein proposed a new method to obtain Central Limit Theorems for dependent r.v. In 1975, Chen adapted this method for the Poisson approximation.

Let $(\Omega, \mathcal{B}, P)$ be a probability space, $\mathbb{E}$ the expectation under $\mathbb{P}$ and $Z : \Omega \to \mathbb{R}$ a r.v. such that $\mathbb{E}[Z] < \infty$.

The purpose of such method is to approximate $\mathbb{E}[Z]$ in order to approximate the c.d.f. of a real r.v. W defined on $(\Omega, \mathcal{B}, P)$ by $\mathbb{1}_{(W \leq w_0)} := Z$.

The basic approach of Stein method consists in :

1. determining $Ker\,I\!E := \{Y : I\!E[Y] = 0\}$;

2. searching in $Ker\,I\!E$ for $Z - c$ for some constant c ;

3. concluding $I\!E[Z] \simeq c$.

Given an integer valued r.v. W, the problem is then how to determine $Ker\,I\!E$, i.e. to characterize the set of functions $h : I\!N \to I\!R$ such that $I\!E[h(W)] = 0$. To solve this problem, a method of "exchangeability" (cf. [166]) is applied. Let us recall the definition of a pair of exchangeable variables.

Let $(\Omega, \mathcal{B}, I\!P)$ and $(\Omega_1, \mathcal{B}_1, I\!P_1)$ be two probability spaces. (X, X') is an "exchangeable" pair of mappings of $(\Omega, \mathcal{B}, I\!P)$ into $(\Omega_1, \mathcal{B}_1, I\!P_1)$ if

$$\begin{cases} I\!P[X \in A] = I\!P_1(A), \ \forall A \in \mathcal{B}_1 \\ I\!P[X \in A \cap X' \in A'] = I\!P[X \in A' \cap X' \in A], \ \forall A, A' \in \mathcal{B}_1. \end{cases}$$

A pair (X, X') of exchangeable variables satisfies the following property :

$$\begin{aligned} & If \quad \mathcal{F} = \{F : \Omega^2 \to I\!R, F \ antisymmetric \ (i.e. \ F(x, x') = -F(x', x)), \ s.t. \ I\!E|F(X, X')| < \infty\}, \\ & \quad \mathcal{X} = \{h : \Omega \to I\!R \ measurable, \ \ s.t. \ I\!E|h(X)| < \infty\}, \\ & and \quad T : \mathcal{F} \to \mathcal{X} \ s.t. \ (TF)(X) = I\!E[F(X, X')|X], \end{aligned}$$

then (X, X') satisfies $I\!E \circ T = 0$, with the composition defined by $(I\!E \circ T)(F) := I\!E[TF]$.
To apply this method to the Poisson approximation, let us recall the following characterization for a Poisson law :
W is an integer valued r.v., Poisson $\mathcal{P}(\lambda)$ distributed, if and only if
for any bounded f, $I\!E[\lambda f(W + 1) - W f(W)] = 0$.
Hence if $I\!E[\lambda f(W + 1) - W f(W)] \simeq 0$ for all bounded function f defined on $I\!N$, then W is nearly distributed as $\mathcal{P}(\lambda)$.
This approximative equality is often easier to check than the direct approximation $I\!E[f(W)] \simeq I\!E[f(W)]$ where Z is $\mathcal{P}(\lambda)$ distributed.
Therefore, let us illustrate *the Stein-Chen method when W is the number of occurrences of a large number of independent events*, the advantage of this method being that the dependent case involves only minor transformations.
Let $(X_i)_{1 \le i \le n}$ be n independent r.v. with values in $\{0, 1\}$, with (for fixed n) $p_i := I\!P[X_i = 1]$, $1 \le i \le n$, and $\lambda := \sum_{i=1}^{n} p_i$.

Let $W := \sum_{i=1}^{n} X_i$. We proceed in four steps :

1. Construction of an exchangeable pair (W, W').
   Let $(X_i^*)_{1 \le i \le n}$ be n independent r.v., independent of the X_i and s.t. $X_i =^d X_i^*, \forall 1 \le i \le n$.
   Let U be a uniformly distributed in $\{1, \cdots, n\}$ r.v., independent of the X_i and X_i^*.
   Let $W' := W - X_U + X_U^*$. Then (W, W') is exchangeable.

2. Definition of an antisymmetric function F.
   Let $F(X, X') := f(X') 1\!\!1_{(X'=X+1)} - f(X) 1\!\!1_{(X=X'+1)}$, with $f : I\!N \to I\!R$.

58

Then we can apply the property of exchangeable mappings to (W, W'), namely $I\!E\left[I\!E^X[F(W, W')]\right] = 0$, i.e.

$$I\!E[\lambda f(W+1) - Wf(W)] = I\!E\left[(f(W+1) - f(W))\sum_{j=1}^n p_j 1\!\!1_{(X_j=1)}\right]. \qquad (1.49)$$

3. Artificial solution $U_\lambda h$ (particular case of the general method of Stein).
Let $h : I\!N \to I\!R$ be bounded and let the function $U_\lambda h$ be defined by

$$U_\lambda h(w) := -\sum_{j=w}^\infty \frac{(w-j)!}{j!}\lambda^{j-w}\left(h(j) - I\!E[h(Z)]\right),$$

with Z Poisson $\mathcal{P}(\lambda)$ distributed r.v., solution of the equation

$$\lambda f(w+1) - wf(w) = h(w) - I\!E[h(Z)].$$

Taking $f := U_\lambda h$ in (1.49) provides that for all bounded function h defined on $I\!N$,

$$I\!E[h(W)] = I\!E[h(Z)] + \sum_{j=1}^n p_j^2 I\!E[V_\lambda h(W_j)],$$

where $W_j := \sum_{j' \neq j} X_{j'}$ and $V_\lambda h(w) := U_\lambda h(w+2) - U_\lambda h(w+1)$.

4. Estimation of $d_{tv}(\mathcal{L}(W), \mathcal{P}(\lambda))$.
Let choose h s.t. $h := h_A$, $A \subset I\!N$, with $h_A(w) := 1\!\!1_{(w \in A)}$; a bound can then be deduced for $V_\lambda h$ and we obtain

$$|I\!P[W \in A] - I\!P[Z \in A]| \leq \left(1 \wedge \frac{1}{\lambda}\right)\sum_{j=1}^n p_j^2$$

which gives the same upper bound for the distance in variation $d_{tv}(\mathcal{L}(W), \mathcal{P}(\lambda))$ defined by $d_{tv}(\mathcal{L}(X), \mathcal{L}(Y)) = d_{tv}(X, Y) := \sup_{A \subset I\!N} |I\!P[X \in A] - I\!P[Y \in A]|$.

Note that Barbour (1992) (cf. [17]) extended the Stein-Chen method by combining it with coupling techniques (Serfling, 1975,1978 (cf. [154] and [155])) to solve the general problem of Poisson approximation for the distribution of a sum of r.v., not necessarily independent, $\{0, 1\}$-valued. In the case of independent indicators, Deheuvels et al. (1988) (cf. [50] and references therein) combined semigroup theory (Le Cam, 1960 ([154])) with coupling techniques to obtain results for the Poisson approximation of sums of independent indicators; it can be shown that it gives sharper results than with Barbour's method (cf. [86]).

Central Limit Theorems

• *Malevich (1969), Cuzick (1976).*
In the 70's, some work has been done to prove Central Limit Theorems (CLT) for the number of zero crossings $N_t(0)$ as $t \to \infty$, for instance by Malevich (cf. [108]) and by Cuzick (cf. [43]). Cuzick gave conditions on the covariance function of X which ensure on one hand a mixing condition at infinity for X and on the other hand a local condition for the sample paths of X; those conditions are weaker than the ones given by Malevich to prove the same result, although the same type of proofs is used.

Theorem 1.1.25 *(Cuzick, 1976)*
Assume that r'' exists and take $r(0) = 1$.
If the following conditions are satisfied :
i) r, $r'' \in L^2$,
ii) the Geman condition (1.9) is verified,
iii) $\liminf\limits_{t \to \infty} \dfrac{var(N_t(0))}{t} = \sigma^2 > 0,$
then
$$\frac{(N_t(0) - I\!E[N_t(0)])}{\sqrt{t}} \xrightarrow{d} \mathcal{N}(0, \sigma^2) \quad , \text{ as } t \to \infty,$$

$$\text{where } \sigma^2 = \frac{1}{\pi} \left\{ \sqrt{\lambda_2} + \int_0^\infty \left(\frac{I\!E[|X_0' X_s'| \mid X_0 = X_s = 0]}{\sqrt{1 - r^2(s)}} - I\!E^2|X'(0)| \right) ds \right\}.$$

Remark : A sufficient condition to have the assumption (iii) and which is directly related to the covariance structure of X can be expressed by
$$\int_0^\infty \frac{r'(t)^2}{1 - r^2(t)} dt \; < \; \frac{\pi}{2} \sqrt{\lambda_2}.$$

The method used to prove the theorem followed an idea of Malevich, which consists to approximate the underlying Gaussian process X (and its derivatives when they exist) by an m-dependent process X_m (i.e. such that $I\!E[X_m(s)X_m(t)] = 0$ if $|s - t| > m$).

• *The m-dependent method.*
Let f be the spectral density of X and let consider the spectral representation of X given by $X_t = \int_{-\infty}^{\infty} \cos(\lambda t) \sqrt{f(\lambda)} \, dB(\lambda)$, where B is a Gaussian white noise.

Define, for each positive m, the process $X_m(t) := \int_{-\infty}^{\infty} \cos(\lambda t) \sqrt{f * P_m} \, dB(\lambda)$, where $*$ denotes convolution and $P_m(\lambda) = mK \dfrac{\sin(m\lambda)}{m\lambda}$, with K such that $\int_{-\infty}^{\infty} P_m(\lambda) d\lambda \equiv 1$.

Then $(X(t), X_m(t), \dot{X}(t), \dot{X}_m(t))$ are jointly Gaussian and stationary, and the number of zero crossings of the process X_m, denoted by $N_t^{X_m}(0)$, is an m-dependent process.

Define $Z_m(t) := \dfrac{N_t^{X_m}(0) - I\!E[N_t^{X_m}(0)]}{\sqrt{t}}$ and $Z(t) := \dfrac{N_t^X(0) - I\!E[N_t^X(0)]}{\sqrt{t}}$.

In order to establish that Z is asymptotically normal $\mathcal{N}(0, \sigma^2)$, it suffices to show that

$\cdot\ Z_m(t) \overset{L^2}{\to} Z(t)$ uniformly in $t > t_0$ as $m \to \infty$,

$\cdot\ Z_m(t) \overset{d}{\to} \mathcal{N}(0, \sigma_m^2)$ for each m as $t \to \infty$,

$\cdot\ \lim\limits_{t\to\infty} I\!E[(Z_m(t))^2] := \sigma_m^2 \ \to\ \lim\limits_{t\to\infty} I\!E[(Z(t))^2] := \sigma^2$ as $m \to \infty$.

Then the fact that $\lim\limits_{t\to\infty} \dfrac{var(N_t^{X_m}(0))}{t} = \sigma_m^2 > 0$, deduced from (iii) and from the first of the three conditions above, will allow to apply the CLT for m-dependent process given in Hoeffding and Robbins (cf. [67]).

● *Piterbarg (1978-1996).*
Applying both the method of comparison and the method of discretization, Piterbarg provided a central limit theorem for the number $N_x^+(t)$ of upcrossings at level x by a Gaussian stationary process X (see [132] or [136]), namely

Theorem 1.1.26 *(CLT for the number $N_x^+(t)$ of upcrossings, Piterbarg, 1978)*
Let $X = (X_s, s \in I\!R)$ be a stationary Gaussian process, mean zero, unit variance, with covariance function r satisfying the Geman condition (1.9) and

$$\int_0^\infty s\left(|r(s)| + |r'(s)| + |r''(s)|\right) ds \ < \infty. \tag{1.50}$$

Then

$$var N_x^+(t) = \sigma^2 t(1 + o(1)) \quad as \quad t \to \infty,$$

where

$$\sigma^2 = \int_0^\infty \int_0^\infty \int_0^\infty yz \left(\phi_s(x, y, x, z) - \frac{e^{-x^2}\sqrt{-r''(0)}}{2\pi} exp\left\{\frac{y^2 + z^2}{2r''(0)}\right\}\right) dydzds + \frac{\sqrt{-r''(0)}e^{-x^2}}{2\pi} > 0.$$

Also, the central limit theorem holds for $N_x^+(t)$.

For the computation of the variance of $N_x^+(t)$, it suffices to write it as
$var(N_x^+(t)) = I\!E[N_x^+(t)(N_x^+(t)-1)] + I\!E[N_x^+(t)] - I\!E[(N_x^+(t))^2]$, and to use results on moments, which combined with the condition (1.50) give the convergence of σ^2.
To obtain the CLT, Piterbarg proceeded both by discretization and by approximation, introducing a smooth enough process $X_\triangle^\delta$ discretized in time, in order to apply to this process a known result, a CLT (under some conditions) for the number of upcrossings for stationary Gaussian process in a discrete time (see for instance [136]). The approximating process $X_\triangle^\delta$ has been chosen as

$$X_\triangle^\delta(s) := X^\delta(k\triangle) + (s - k\triangle)\frac{X^\delta((k+1)\triangle) - X^\delta(k\triangle)}{\triangle},$$

for $s \in [k\triangle, (k+1)\triangle)$, $k = 0, 1, \cdots$, $\triangle$ such that $1/\triangle$ is an integer and $\triangle \to 0$, and with the smoothed process X^δ defined by

$$X^\delta(s) := \frac{1}{\delta}\int_0^\delta X(s + v)dv.$$

Note that approximating the process X by a discretized in time process $X(\triangle k)$ wouldn't have allow $X(\triangle k)$ to satisfy the conditions of the CLT in discrete time, in particular the condition

(1.50) does not imply that the approximated process would satisfy one of the conditions on the covariance function r, namely $\sum_{j=0}^{\infty} j|r(j)| < \infty$; that's why smoothing the process was also needed.

• *Berman (1971-1992).*
Let us come back briefly to the notion of sojourn time. When choosing the level function $x = x(t)$ of the asymptotic order of $\sqrt{2\log t}$ (as $t \to \infty$), we recalled in the section *Law of small numbers*, that the sojourn time above the level x tends to a compound Poisson limit, since the sojourns are relatively infrequent and their contributions are few but individually relatively substantial. In this case, the local behavior of the correlation function r was determining in the form of the limiting distribution. Now when choosing the level function rising at a slower rate (such that $x(t)/\sqrt{2\log t}$ is bounded away from 1), the sojourns become more frequent and their contributions more uniform in magnitude, implying with the customary (in the application of CLT) normalization, namely $\dfrac{S_t(x) - I\!\!E[S_t(x)]}{\sqrt{var(S_t(x))}}$ (with $I\!\!E[S_t(x)]$ not depending on r), a normal limit distribution. Here the local behavior of r does play in the normalization function, but not in the form of the limiting distribution.
Berman proved this CLT for two different types of mixing conditions, namely when the covariance function r decays sufficiently rapidly to 0 as $t \to \infty$, or on the contrary at a sufficiently small rate.
In the case of a rapid rate of decay of r, the mixing condition is based not on the function r itself but on the function b appearing in its spectral representation given by

$$r(t) = \int_{-\infty}^{\infty} b(t+s)\bar{b}(s)ds,$$

when supposing that the spectral distribution is absolutely continuous with derivative $f(\lambda)$, and where b is the Fourier transform in the L^2-sense of $\sqrt{f(\lambda)}$.
This mixing condition is given by $b \in L^1 \cap L^2$, which means in particular that the tail of b is sufficiently small so that r tends to 0 sufficiently rapidly. The proof of the CLT is based on the m-dependent method. Indeed, Berman introduced a family of m-dependent stationary Gaussian processes $\{X_m(t); -\infty < t < \infty\}$ to approach uniformly in t the original process X in the mean square sense for large m. Then he deduced a CLT to the normalized sojourn of X from the CLT of the m-dependent process (established by adapting a blocking method used in the proofs of CLT for dependent r.v.).
In the case of slowly decreasing covariances, the proof of the CLT relies on a method specific to Gaussian processes, based on the expansion of $S_t(x)$ in a series of integrated Hermite polynomials and the method of moments. In fact it is a special case of what is known as a non-central limit theorem for Gaussian processes with long-range dependence (see Dobrushin and Major (1979, [54]), Taqqu (1979, [170]) and the next section *Noncentral limit theorems*).

• *Slud (1991, 1994)*
Introducing the theoretical tool of Multiple Wiener Itô integrals (MWI's) allowed some authors as Taqqu (1975, [168]), Dobrushin and Major (1981, [54]), Giraitis and Surgailis (1985, [61]), Maruyama (1976, [114]), Chambers and Slud (1989, [36]), ... to prove general functional central

limit theorems (FCLT) (and non central limit theorems) for MWI expansions. In the 1990's, Slud (cf. [161]) applied a general central limit theorem of Chambers and Slud (1989, [36]) to provide Cuzick's CLT for the zero crossings $N_t(0)$ by the process X, without needing Cuzick's additional assumptions to get a strictly positive limiting variance. Then he generalized the result to constant levels (cf. Theorem 3.1 in [163]), as follows.

Theorem 1.1.27 *(Slud, 1994)*
Let $x \in I\!R$ be arbitrary, and let $X = (X_t, t \geq 0)$ be a mean 0, variance 1, stationary Gaussian process with continuous spectral measure σ and twice-differentiable correlation function r. Suppose $var(N_t(x)) < \infty$. Assume that $\int_{-\infty}^{\infty} r^2(s)ds < \infty$ if $x = 0$ and that $\int_{-\infty}^{\infty} r(s)ds < \infty$ if $x \neq 0$. Then, as $t \to \infty$,

$$\frac{1}{\sqrt{t}} \left(N_x(t) - e^{-x^2/2} t \sqrt{-r''(0)}/\pi \right) \xrightarrow{d} \mathcal{N}(0, \alpha^2),$$

where $\alpha^2 > 0$ is given by the expansion

$$\alpha^2 = \sum_{n=1}^{\infty} \frac{1}{n!} I\!\!E_{m^n}[|f_n(\Lambda)|^2 \mid \Lambda_1 + \cdots + \Lambda_n = 0] \int_{-\infty}^{\infty} r^n(s)ds,$$

with $f_n \in L_s^2(I\!R^n, m^n)$ (introduced in (1.24)) defined by

$$f_n(\underline{\lambda}_n) := \frac{e^{-x^2/2}(e^{i\sum_{j=1}^n \lambda_j} - 1)}{\pi \sum_{j=1}^n \lambda_j} \sum_{k=0}^{[n/2]} \frac{(-r''(0))^{1-2k} H_{n-2k}(x) H_{2k}(0) i^{1+2k}}{(2k)!(2k-1)} \sum_{1 \leq n_1 < \cdots < n_{2k} \leq n} \lambda_{n_1} \cdots \lambda_{n_{2k}}.$$

Note that Slud extended this result to curve crossings in cases where the curve is periodic and the underlying process has rapidly decaying correlations (cf. Theorem 6.4 in [163]).

• *Kratz and León (2001) : a general method.*
In 2001, Kratz and León (cf. [90]) introduced a general method, which can be applied to many different cases, in particular when the dimension of the index set T is bigger than one, to provide a CLT as general as possible for level functionals of Gaussian processes $X = (X_t, t \in T)$. This method is a combination of two approaches, the one developed by the authors in 1997 (cf. [88]) and one derived by the work of Malevich (cf. [108]), Cuzick (cf. [43]) and mainly Berman (cf. [23]), that consists in approaching the process X by a m-dependent process, in order to be able to use well-known results on m-dependent processes. Applying this method, a CLT is given for functionals of $(X_t, \dot{X}_t, \ddot{X}_t, t \geq 0)$, which allows in particular to get immediately the CLT for the number of crossings $N_t(x)$ of X, given in Slud (cf. [163], Theorem 3.1).
We suppose that the correlation function r of our (stationary Gaussian) process mean zero, variance one, satisfies

$$r \in L^1 \text{ and } r^{(iv)} \in L^2. \tag{1.51}$$

Note that (1.51) implies that r', r'' and r''' belong to L^2 as well.
Let Z_s be a r.v. independent of X_s and $\dot{X}_s$ (for each s fixed) such that

$$\frac{\ddot{X}_s}{\sqrt{r^{(iv)}(0)}} = \rho_1 X_s + \rho_2 Z_s, \quad \text{with} \quad \rho_1 = \frac{r''(0)}{\sqrt{r^{(iv)}(0)}}, \rho_2 = \sqrt{1 - \rho_1^2}. \tag{1.52}$$

63

Note that X_s and $\dot{X}_s$ are independent, as well as $\dot{X}_s$ and $\ddot{X}_s$, which ensures the existence of Z_s with the stated properties.

Let F_t^X be the Hermite expansion given in $H(X)$ by

$$F_t^X = \sum_{q=0}^{\infty} \sum_{0 \leq n+m \leq q} d_{qnm} \int_0^t H_n(X_s) H_{q-(n+m)} \left(\frac{\dot{X}_s}{\sqrt{-r''(0)}} \right) H_m(Z_s) ds, \qquad (1.53)$$

with d_{qnm} such that $\forall q \geq 0$,

$$\sum_{0 \leq n+m \leq q} d_{qnm}^2 n! m! (q - (n+m))! < C(q), \qquad (1.54)$$

with $(C(q))_q$ some bounded sequence.

Besides the property of stationarity of the process and the orthogonality of the chaos, which will be used to simplify the computations whenever it is possible, let us give the basic points constituting this general method.

▷ Let $\mathcal{F}_{Q,t}(X)$ be the finite sum deduced from $\mathcal{F}_t(X)$ for $q = 1$ to Q, i.e.

$$\mathcal{F}_{Q,t}(X) = \sum_{q=1}^{Q} \sum_{0 \leq n+m \leq q} \frac{d_{qnm}}{\sqrt{t}} \int_0^t I_{qnm}(s) ds, \quad \text{with } I_{qnm}(s) = H_n(X_s) H_{q-(n+m)} \left(\frac{\dot{X}_s}{\sqrt{-r''(0)}} \right) H_m(Z_s).$$

▷ First we show that $\mathcal{F}_t(X)$ can be approximated in L^2 by the r.v. $\mathcal{F}_{Q,t}(X)$:

$$\lim_{Q \to \infty} \lim_{t \to \infty} I\!E[\mathcal{F}_t(X) - \mathcal{F}_{Q,t}(X)]^2 = 0. \qquad (1.55)$$

To prove this convergence, we follow the method developed in Kratz & León (cf. [89], proof of Theorem 1) where two results (cited below), one of Arcones (1994, [5], Lemma 1) and the other of Taqqu (1977, [169], Lemma 3.2), helped respectively for the computation of expectations of the form $I\!E \left[\sum_{0 \leq n_1+m_1 \leq q} \sum_{0 \leq n_2+m_2 \leq q} d_{qn_1m_1} d_{qn_2m_2} I_{qn_1m_1}(s) I_{qn_2m_2}(s') \right]$ and $I\!E \left[\sum_{0 \leq n+m \leq q} d_{qnm} I_{qnm}(0) \right]^2$.

Lemma 1.1.7 *(Arcones inequality, 1994)*
Let $X = (X_i, 1 \leq i \leq d)$ and $Y = (Y_i, 1 \leq i \leq d)$ be two Gaussian vectors on $I\!R^d$, mean zero, such that $I\!E[X_i X_j] = I\!E[Y_i Y_j] = \delta_{ij}, 1 \leq i, j \leq d$.
Let $r^{(i,j)} := I\!E[X_i Y_j]$ and f a function on $I\!R^d$ with finite second moment and Hermite rank τ, $1 \leq \tau < \infty$, w.r.t. X. Recall that the Hermite rank of a function f is defined by

$$rank(f) := \inf \left\{ \tau : \exists l_j \text{ with } \sum_{j=1}^{d} l_j = \tau \text{ and } I\!E \left[(f(X) - I\!E[f(X)]) \prod_{j=1}^{d} H_{l_j}(X_j) \right] \neq 0 \right\}.$$

Suppose $\Psi := \left(\sup_{1 \leq i \leq d} \sum_{j=1}^{d} |r^{(i,j)}| \right) \vee \left(\sup_{1 \leq j \leq d} \sum_{i=1}^{d} |r^{(i,j)}| \right) \leq 1$. Then

$$I\!E \left[(f(X) - I\!E[f(X)]) (f(Y) - I\!E[f(Y)]) \right] \leq \psi^\tau I\!E \left[(f(X) - I\!E[f(X)])^2 \right].$$

In our case, ψ denotes the supremun of the sum of the absolute values of the off-diagonal terms in the column vectors belonging to the covariance matrix for the Gaussian vector

$$\left(X_0, \frac{\dot{X}_0}{\sqrt{-r''(0)}}, Z_0, X_u, \frac{\dot{X}_u}{\sqrt{-r''(0)}}, Z_u \right).$$

Under the conditions on r, we have $\int_0^\infty \psi^2(u)du < \infty$.

Lemma 1.1.8 *(Taqqu, 1977)*
Let $p \geq 2$ and $(X_1, \cdots, X_p)$ be standard Gaussian. Then

$$\mathbb{E}[H_{k_1}(X_1) \cdots H_{k_p}(X_p)] = \begin{cases} \dfrac{k_1! \cdots k_p!}{2^q(q!)} \displaystyle\sum_\Gamma r_{i_1 j_1} \cdots r_{i_q j_q} & if \displaystyle\sum_{l=1}^p k_l = 2q \ and \ 0 \leq k_1, \cdots, k_p \leq q \\ 0 & otherwise \end{cases}$$

where $\displaystyle\sum_\Gamma$ is a sum over all indices $i_1, j_1, \cdots, i_q, j_q \in \{1, 2, \cdots, p\}$ such that $i_l \neq j_l, \forall l = 1, \cdots, p$ and there are k_1 indices 1, k_2 indices 2, $\cdots$, k_p indices p.

▷ Now by classical tools (the dominated convergence theorem, Fatou Lemma, Cauchy-Schwarz inequality, ...) and Arcones inequality, we can get the limit variance σ^2 of $\mathcal{F}_{Q,t}(X)$ and prove that it is finite :

$$\lim_{Q\to\infty} \lim_{t\to\infty} \mathbb{E}[\mathcal{F}_{Q,t}(X)]^2 := \sigma^2 = \sum_{q=1}^\infty \sigma^2(q) < \infty, \quad \text{where} \tag{1.56}$$

$$\sigma^2(q) = \sum_{0 \leq n_1+m_1 \leq q} \sum_{0 \leq n_2+m_2 \leq q} d_{qn_1m_1} d_{qn_2m_2} \int_0^\infty \mathbb{E}[I_{qn_1m_1}(0)I_{qn_2m_2}(s)]ds.$$

▷ A version of the m-dependent method : Berman's method.
This method consists in approaching the process X by a m-dependent process, in order to be able to use well-known results on m-dependent processes.
Define X^ε as an $(1/\varepsilon)$-dependent process to approach the process X, as follows.
Let ψ be defined by $\psi(x) = \varphi * \varphi(x)$, where $*$ denotes the convolution and where φ is a twice differentiable even function with support contained in $[-1/2, 1/2]$ (so ψ has support in $[-1, 1]$). We can suppose w.l.o.g. that $||\varphi||^2 = 1$.
By using the Fourier inversion formula we can write

$$\psi(x) = \frac{1}{2\pi} \int_{-\infty}^\infty e^{-i\lambda x} \hat{\psi}(\lambda)d\lambda = \frac{1}{2\pi} \int_{-\infty}^\infty e^{-i\lambda x} |\hat{\varphi}(\lambda)|^2 d\lambda.$$

Let $\hat{\varphi}_\varepsilon(\lambda) = \frac{1}{2\pi\varepsilon} |\hat{\varphi}(\frac{\lambda}{\varepsilon})|^2$, and

$$X_t^\varepsilon = \int_{-\infty}^\infty e^{i\lambda t} (f * \hat{\varphi}_\varepsilon(\lambda))^{\frac{1}{2}} dW(\lambda),$$

(note that X_t^ε for fixed t is a standard Gaussian r.v.) and its derivatives denoted by $X_t^{(j)\varepsilon}$, $0 \le j \le 2$, such that

$$X_t^{(j)\varepsilon} = \int_{-\infty}^{\infty} e^{i\lambda t}(i\lambda)^j (f * \hat{\varphi}_\varepsilon(\lambda))^{\frac{1}{2}} dW(\lambda).$$

So we can prove that

$$\mathbb{E}[\mathcal{F}_{Q,t}(X) - \mathcal{F}_{Q,t}(X^\varepsilon)]^2 \underset{t\to\infty}{\to} 0 \ , \ \text{with} \ \varepsilon(t) \underset{t\to\infty}{\to} 0, \tag{1.57}$$

by using Arcones inequality and some results on the correlation between the process (respectively its derivatives) and the $(1/\varepsilon)$-dependent process associated (respectively its derivatives), obtained when working with the spectral representation of the correlation functions, namely

Proposition 1.1.6 *(Kratz and León, 2001)*
(i) *For all $0 \le j, k \le 2$, $\mathbb{E}[X_{t+.}^{(j)\varepsilon} X_t^{(k)\varepsilon}] = (-1)^k r_\varepsilon^{(j+k)}(.)$ converge uniformly over compacts and in L^2 as $\varepsilon \to 0$ towards $r_{j,k}(.) := (-1)^k r^{(j+k)}(.).$*
For $j = k = 0$, the convergence takes place in L^1 as well.
(ii) *For all $0 \le j, k \le 2$, $r_{j,k}^\varepsilon(.) := \mathbb{E}[X_{t+.}^{(j)\varepsilon} X_t^{(k)}]$ converge uniformly over compacts and in L^2 as $\varepsilon \to 0$ towards $r_{j,k}(.).$*

▷ Then it is enough to consider the weak convergence of the sequence $\mathcal{F}_{Q,t}(X^\varepsilon)$ towards a Gaussian r.v. as $t \to \infty$ to get the CLT for F_t^X.
We can write

$$
\begin{aligned}
\mathcal{F}_{Q,t}(X^\varepsilon) &= \sum_{q=1}^{Q} \sum_{0 \le n+m \le q} d_{qnm}^\varepsilon \frac{1}{\sqrt{t}} \int_0^t I_{qnm}^\varepsilon(s) ds \\
&= \frac{1}{\sqrt{t}} \int_0^t \sum_{q=1}^{Q} \sum_{0 \le n+m \le q} d_{qnm}^\varepsilon I_{qnm}^\varepsilon(s) ds := \frac{1}{\sqrt{t}} \int_0^t f_Q\left(X_s^\varepsilon, \frac{\dot{X}_s^\varepsilon}{\sqrt{-r_\varepsilon''(0)}}, Z_s^\varepsilon\right) ds \\
&= \frac{1}{\sqrt{t}} \int_0^t \theta_s \left[f_Q\left(X_0^\varepsilon, \frac{\dot{X}_0^\varepsilon}{\sqrt{-r_\varepsilon''(0)}}, Z_0^\varepsilon\right) \right] ds,
\end{aligned}
$$

where θ is the shift operator associated to the process
and $f_Q(x_1, x_2, x_3) = \displaystyle\sum_{q=1}^{Q} \sum_{0 \le n+m \le q} d_{qnm}^\varepsilon H_n(x_1) H_{q-(n+m)}(x_2) H_m(x_3).$

Hence the weak convergence of $\mathcal{F}_Q(X^\varepsilon)$ towards a Gaussian r.v. is a direct consequence of the CLT for sums of m-dependent r.v. (cf. [67] and [22]), which combined with (1.55), (1.56) and (1.57) provides the CLT for $\mathcal{F}_t(X)$, namely

Theorem 1.1.28 *(Kratz and León, 2001)*
Under the above conditions, we have

$$\mathcal{F}_t(X) := \frac{F_t^X - \mathbb{E}[F_t^X]}{\sqrt{t}} \longrightarrow \mathcal{N}(0, \sigma^2) \ \text{as} \ t \to \infty,$$

where $I\!E[F_t^X] = td_{000}$ *and* $\sigma^2 = \sum_{q=1}^{\infty} \sigma^2(q) < \infty$, *with*

$$\sigma^2(q) = \sum_{0 \leq n_1 + m_1 \leq q} \sum_{0 \leq n_2 + m_2 \leq q} d_{qn_1m_1} d_{qn_2m_2} \int_0^{\infty} I\!E[I_{qn_1m_1}(0)I_{qn_2m_2}(s)]ds,$$

$$and \quad I_{qnm}(s) = H_n(X_s)H_{q-(n+m)}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right)H_m(Z_s).$$

Remark. Condition (1.51) can of course be weakened to :

$$r \in L^1 \quad \text{and} \quad r'' \in L^2 \tag{1.58}$$

when considering the process X and its first derivative only, as for the number of crossings, and to

$$r \in L^1 \tag{1.59}$$

when considering the process X only.

$\bullet$ As an application of Theorem 1.1.28, we get back, under the condition (1.58), Slud's CLT for the number of crossings of X enunciated in Theorem 1.1.27.
Indeed, let us consider the Hermite expansion of $N_t(x)$ given in (1.34), which corresponds to F_t^X with $m := 0$, $d_{qn} := d_{qn0} = b_{q-2l}(x)a_{2l}1_{n=2l}$, $\forall q, n, l \in I\!N$ and with $I\!E[F_t^X] = td_{000} = tb_0(x)a_0 = \frac{t}{\pi}e^{-x^2/2}$. We have $\forall x \in I\!R$, $\sum_{l=0}^{[q/2]} b_{q-2l}^2(x)a_{2l}^2(2l)!(q-2l)! < C$, with C some constant independent of q, which is a consequence of some result in Imkeller et al. (cf. [78], Proposition 3) given by

Proposition 1.1.7 *(Imkeller, Perez-Abreu and Vives, 1995)*
Let $1/4 \leq \alpha \leq 1/2$. *Then there exists a constant* c *such that for any* $n \in I\!N$,

$$\sup_{x \in I\!R} |H_n(x)e^{-\alpha x^2}| \leq cn^{-(8\alpha-1)/12}.$$

Therefore we can apply Theorem 1.1.28 to obtain

$$\sqrt{t}\left(\frac{N_t(x)}{t\sqrt{-r''(0)}} - \frac{e^{-x^2/2}}{\pi}\right) \xrightarrow[t\to\infty]{} \mathcal{N}(0, \sigma^2),$$

which can be written as

$$\sqrt{t}\left(\frac{N_t(x)}{t} - \frac{\sqrt{-r''(0)}}{\pi}e^{-x^2/2}\right) \xrightarrow[t\to\infty]{} \mathcal{N}(0, -r''(0)\sigma^2),$$

with σ^2 $(0 < \sigma^2 < \infty)$ given by

$$\sigma^2 = \sum_{q=1}^{\infty} \sigma^2(q) = \sum_{q=1}^{\infty} \sum_{n_1=0}^{[q/2]} \sum_{n_2=0}^{[q/2]} b_{q-2n_1}(x) a_{2n_1} b_{q-2n_2}(x) a_{2n_2} \int_0^{\infty} \mathbb{E}\left[H_{2n_1}(X_0) H_{q-2n_1}\left(\frac{\dot{X}_0}{\sqrt{-r''(0)}} \right) \right.$$
$$\left. H_{2n_2}(X_s) H_{q-2n_2}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}} \right) \right] ds.$$

We can easily check that $\sigma^2 \neq 0$, by proving that $\sigma^2(2) \neq 0$ (showing that the determinant of the matrix (expressed in the domain of frequencies) associated to $\sigma^2(2)$ is striclty positive). We conclude to Slud's CLT.

Non central limit theorems

From the CLT given in the previous section, or even more generally from the literature of the CLT for non-linear functionals of Gaussian processes, we can ask what happens when some of the conditions of those theorems are violated, in particular when we are under a condition of regular long-range dependence.
This problem interested many authors, among whom we can cite in a chronological order Taqqu ([168],[170]), Rosenblatt ([147], [148], [149]), Dobrushin and Major ([54]), Major ([106]), Giraitis and Surgailis ([61]), Ho and Sun ([70]), Slud ([162] or [163]). MWI's expansions proved to be quite useful in defining the limiting behavior in non central limit theorems for functionals of Gaussian processes with regular long-range dependence. It is what allows Slud to prove, by using techniques proper to MWI's, in particular Major's noncentral limit theorem for stationary Gaussian fields with regular long-range dependence (see theorem 8.2 in Major, [105]), the following non CLT for level-crossing counts.

Theorem 1.1.29 *(Slud, 1994)*
Let X be a stationary Gaussian process, mean 0, variance 1, with continuous spectrum and twice differentiable correlation function with regular long range dependence, i.e. $r(s) = (1+|s|)^{-\alpha} L(s)$, where L is slowly varying at ∞ and $0 < \alpha < 1/2$. Assume also that for some $\delta \in (-\infty, \alpha)$ and constant $C < \infty$, for $k \geq 1, 2$ and all $x \geq 0$,

$$\frac{1}{|r(x)|} \left| \frac{d^k}{dx^k} r(x) \right| \leq C(1+|x|)^{\delta}.$$

i) Let $c \neq 0$ be fixed arbitrarily. Then

$$\frac{N_t(c) - e^{-c^2/2} t \sqrt{-r''(0)}/\pi}{\sqrt{t^{2-\alpha} L(t)}} \xrightarrow{d} \mathcal{N}\left(0, \frac{2e^{-c^2} c^2 (-r''(0))}{(1-\alpha)(2-\alpha)\pi^2} \right), \quad as \ t \to \infty.$$

ii) When $c = 0$,

$$\frac{t^{\alpha-1}}{L(t)} \left(N_t(0) - t\sqrt{-r''(0)}/\pi \right) \xrightarrow{d} \frac{-r''(0)}{\pi} \tilde{I}_2 \left(\frac{e^{i(\lambda_1+\lambda_2)} - 1}{i(\lambda_1 + \lambda_2)} \right), \quad as \ t \to \infty,$$

where $\tilde{I}_2$ denotes the second order MWI integral operator for a stationary Gaussian process $\tilde{X}$ with correlation function $r_0(s) := \int e^{isx} \sigma_0(dx)$ uniquely determined by

$$\int e^{isx} \frac{(1 - \cos x)^2}{x^2} \sigma_0(dx) = \int_0^1 (1-x)|x+s|^{-\alpha} dx, \quad s > 0.$$

1.2 Extensions

The aim of this section is to present some examples of possible applications of some of the methods reviewed in the previous section; so it won't provide an exhaustive review on the possible extensions.

1.2.1 CLT for other non linear functionals of stationary Gaussian processes

As an application of the heuristic of the previous section, in particular of the theorem 1.1.28, we may look at various functionals related to crossing functionals of a stationary Gaussian process X, as for instance the sojourn time of X in some interval, the local time of X when it exists, or the number of maxima of the process in an interval.

● *Time occupation functionals.*
The simplest application of Theorem 1.1.28 is when the integrand appearing in F_t^X, defined in (1.53), depends only on one variable, without needing other condition than the smooth one $r \in L^1$ (see the remark following the theorem).

▷ A first example of this type is when F_t^X represents the sojourn of the process X above a level x in an interval $[0,t]$, i.e. when $F_t^X := S_x(t)$.
It is easy to obtain the Hermite expansion of $S_x(t)$ as

$$S_x(t) = \sum_{q=0}^{\infty} d_q \int_0^t H_q(X_s)ds,$$

with $d_0 = 1 - \Phi(x)$ and $d_q = \dfrac{1}{q!} \displaystyle\int_x^{+\infty} H_q(u)\phi(u)du = -\dfrac{1}{q!}H_{q-1}(x)\phi(x)$, $\forall q \geq 1$.

Then an application of Theorem 1.1.28 yields to the CLT for the sojourn time under the condition $r \in L^1$, namely

$$\frac{S_x(t) - td_0}{\sqrt{t}} \xrightarrow[t \to \infty]{} \mathcal{N}(0, \sigma^2) ,$$

with $\sigma^2 = \displaystyle\sum_{q=1}^{\infty} d_q^2 \int_0^\infty \mathbb{E}[H_q(X_0)H_q(X_s)]ds = \sum_{q=1}^{\infty} q! d_q^2 \int_0^\infty r^q(s)ds.$

▷ Another example already discussed is the local time L_t^x for X in the level x (when it exists! see [23]).
Its Hermite expansion is given by

$$L_t^x = \phi(x) \sum_{k=0}^{\infty} \frac{H_k(x)}{k!} \int_0^t H_k(X_s)ds = \sum_{k=0}^{\infty} l_k \int_0^t H_k(X_s)ds.$$

and again Theorem 1.1.28 allows to get back its asymptotical normal behavior under the condition $r \in L^1$, namely

$$\frac{L_t^x - t\phi(x)}{\sqrt{t}} \xrightarrow[t \to \infty]{} \mathcal{N}(0, \tilde{\sigma}^2) ,$$

with $\tilde{\sigma}^2 = \sum_{k=1}^{\infty} l_k^2 \int_0^{\infty} I\!E[H_k(X_0)H_k(X_s)]ds = \phi^2(x) \sum_{k=1}^{\infty} \frac{1}{k!} \int_0^{\infty} r^k(s)ds.$

• *Number of maxima in an interval.*

One of the main concerns of extreme value theory is the study of the maximum $\left(\max_{t \in [0,T]} X_t\right)$ of a real-valued stochastic process $X = (X_t, t \in [0, T])$ having continuous paths, in particular the study of its distribution F.

There is an extensive literature on this subject, going mainly in three directions, according to Azaïs and Wschebor ([13]) : one looking for general inequalities for the distribution F, the other for describing the behavior of F under various asymptotics, and the last one for studying the regularity of the distribution F. For more references, see [13].

In the discussion of maxima of continuous processes, the upcrossings of a level play an important role, as it was the case in the discrete case between maxima of sequences and exceedances of a level u_n through the equivalence of the events $(M_n^{(k)} \le u_n) = (N_n^+ < k)$, $M_n^{(k)}$ being the kth largest value of the r.v. $X_1, \cdots, X_n$ and N_n^+ the point process of exceedances on $(0, 1]$ (i.e. $N_n^+ = \#\{i/n \in (0, 1] : X_i > u_n, 1 \le i \le n\}$).

In the continuous case, we have already seen that crossings and maxima are closely related when describing the Rice method, in particular with the Lemma 1.1.4 providing bounds of $I\!P[\max_{0 \le s \le t} X_s \ge x]$ in terms of factorial moments of the level x-crossings of the process X.

Note that it is Cramér (1965) who noted the connection between u- upcrossings by X and its maximum, e.g. by $\{N_u(T) = 0\} = \{M(T) \le u\} \cup \{N_u(T) = 0, X(0) > u\}$, which led to the determination of the asymptotic distribution of the maximum $M(T)$.

Recently, when looking for a reasonable way based upon natural parameters of the considered process X to compute the distribution of the maximum of X, Azaïs and Wschebor ([14], 2001) established a nice method, based upon expressing the distribution of the maximum $I\!P[\max_{0 \le s \le 1} X_s > x]$ of the process satisfying some regularity conditions, by means of Rice series, whose main kth term is given by $(-1)^{k+1} \nu_k/k!$, ν_k denoting the kth factorial moment of the number of upcrossings. This method, named 'Rice method revisited' because inspired by previous works such that the one of Miroshin ([119], 1974), can be applied to a large class of processes, and allows a numerical computation of the distribution in Gaussian cases more powerful in many respects than the widely used Monte-Carlo method, based on the simulation of the paths of the continuous parameter process.

Note also another useful connection with the maximum $\left(\max_{0 \le s \le t} X_s\right)$ of a process X, which is the sojourn time. Indeed, we can write

$$\left(\max_{0 \le s \le t} X_s \le x\right) \quad \Leftrightarrow \quad (S_x(t) = 0).$$

Berman uses this equivalence between the events $\left(\max_{0 \le s \le t} X_s > x\right)$ and $(S_x(t) > 0)$ to study the maximum of the process X (see [23], chapter 10), going in the second direction mentioned above .

Here we are interested in the number of local maxima of a stationary Gaussian process lying in some interval, and more specifically in its asymptotical behavior, one of the motivations being the applications in hydroscience.

More precisely let $M^X_{[\beta_1,\beta_2]}$ be the number of local maxima of $X_s, 0 \le s \le t$, lying in the real interval $[\beta_1, \beta_2]$ and let $r(0) = 1$.

Kratz & León (1997) provided under the condition $r^{(vi)}(0) < \infty$ the Hermite expansion of $M^X_{[\beta_1,\beta_2]}$, by adapting the proof of Proposition 1.1.2 on the number of crossings, and with the change of variables (1.52).

Formally, $M^X_{[\beta_1,\beta_2]} = \displaystyle\int_0^t 1_{[\beta_1,\beta_2]}(X_s)\delta_0(\dot{X}_s) \mid \ddot{X}_s \mid 1_{[0,\infty)}(\ddot{X}_s)ds$; more precisely :

Theorem 1.2.1 *(Kratz and León, 1997)*
Under the condition $-r^{(vi)}(0) < \infty$*, we have*

$$M^X_{[\beta_1,\beta_2]} = -\sqrt{\frac{r^{(iv)}(0)}{-r''(0)}} \sum_{q=0}^{\infty} \sum_{0 \le n+m \le q} \delta_{nm} \frac{H_{q-(m+n)}(0)}{(q-(m+n))!\sqrt{2\pi}}$$

$$\int_0^t H_n(X_s)H_{q-(n+m)}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right) H_m(Z_s)ds$$

where δ_{nm} *is defined by*

$$\delta_{nm} = \frac{1}{n!m!} \int_{\beta_1}^{\beta_2} \int_{I\!R} (\rho_1 x + \rho_2 z)1_{(-\infty,0)}(\rho_1 x + \rho_2 z)H_n(x)H_m(z)\phi(x)\phi(z)dxdz, \qquad (2.60)$$

ρ_1 *and* ρ_2 *satisfying (1.52).*

As for the number of crossings, the condition appearing in the theorem could be weakened by taking a similar condition to (1.33) for the fourth derivative of $r(.)$; it would then provide that $E(M^X_{[\beta_1,\beta_2]})^2 < \infty$.

Previously two cases of application of Theorem 1.1.28 were considered, on one hand when the integrand appearing in F_t^X, defined in (1.53), depends only on one variable (under the simple condition $r \in L^1$), on the other hand when the integrand depends on two variables (under the conditions $r \in L^1$ and $r'' \in L^2$), case of our main study, the number of crossings, but which also concerns any convex combination of it as, for instance, Cabaña estimator ([34]) of the second spectral moment (slightly modified) given by $\gamma = \dfrac{\pi}{t} \displaystyle\int_{-\infty}^{\infty} N_t(x)d\alpha(x)$ (with $\alpha(.)$ a distribution function on $I\!R$), and which has been studied by Kratz & León (2000, [89]).

To apply Theorem 1.1.28 to obtain a CLT for the number of local maxima $M^X_{[\beta_1,\beta_2]}$ made us consider the last possible case, i.e. when the integrand appearing in F_t^X depends on three variables.

Note that the condition $r^{(vi)}(0) < \infty$ and $r \in L^1$ imply the condition (1.51) of the theorem 1.1.28 and that we can write $M^X_{[\beta_1,\beta_2]} = F_t^X$ when taking $d_{qnm} = -\sqrt{\dfrac{r^{(iv)}(0)}{-r''(0)}} \dfrac{\delta_{nm}H_{q-(m+n)}(0)}{(q-(m+n))\sqrt{2\pi}}$.

Moreover we can easily check that $\sum_{q=0}^{\infty} d_{qnm}^2 n!\,m!\,(q-(n+m))! < C$, with C some constant, by using again Proposition 1.1.7 (Imkeller et al.).

Therefore Theorem 1.1.28 yields to a CLT for $M_{[\beta_1,\beta_2]}^X$, result obtained by Kratz & León in 2000 ([89]) with another method, which consisted mainly on adapting and verifying conditions (a1)-(a3) in the 1994 Slud's paper (see [163], p.1362). But, as already noticed by Berman (cf. [22], pp.62-63), it wouldn't have been possible to benefit of this method when working in a higher dimension than one, which reinforced us to find a more general method.

Theorem 1.2.2 *(Kratz and León, 2000)*
Under the conditions $-r^{(vi)}(0) < \infty$ *and* $r \in L_1$,

$$\frac{M_{[\beta_1,\beta_2]}^X - I\!\!E[M_{[\beta_1,\beta_2]}^X]}{\sqrt{t}} \;\xrightarrow{d}\; N(0,\sigma^2) \ as \ t \longrightarrow \infty,$$

where

$$I\!\!E\left(M_{[\beta_1,\beta_2]}^X\right) = \frac{t}{\sqrt{2\pi}} \frac{\sqrt{r^{(iv)}(0)}}{\sqrt{-r''(0)}} \left[\rho_1 \left(\phi(\beta_2)\Phi\left(-\frac{\rho_1}{\rho_2}\beta_2\right) - \phi(\beta_1)\Phi\left(-\frac{\rho_1}{\rho_2}\beta_1\right) \right) \right. \tag{2.61}$$

$$\left. + \frac{1}{\sqrt{2\pi}}\sqrt{\rho_2^2+\rho_1^2}\left\{ \Phi\left(\beta_2\sqrt{1+\frac{\rho_1^2}{\rho_2^2}}\right) - \Phi\left(\beta_1\sqrt{1+\frac{\rho_1^2}{\rho_2^2}}\right) \right\} \right],$$

ρ_i $(i=1,2)$ *being defined in (1.52), and* σ^2 *is given by*

$$\sigma^2 = \frac{r^{(iv)}(0)}{-r''(0)2\pi} \sum_{q=1}^{\infty} \sum_{0\le n_1+m_1\le q} \sum_{0\le n_2+m_2\le q} \delta_{n_1 m_1}\delta_{n_2 m_2} \frac{H_{q-(m_1+n_1)}(0)H_{q-(m_2+n_2)}(0)}{(q-(m_1+n_1))!(q-(m_2+n_2))!} \times \tag{2.62}$$

$$\int_0^{+\infty} I\!\!E\left[H_{n_1}(X_0)H_{q-(n_1+m_1)}\left(\frac{\dot{X}_0}{\sqrt{-r''(0)}}\right) H_{m_1}(Z_0)H_{n_2}(X_s)H_{q-(n_2+m_2)}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right) H_{m_2}(Z_s) \right] ds.$$

This last result has applications to hydroscience, in particular in the study of random seas. Ocean waves have been studied for many years by many authors, in particular through a stochastic approach by M. Ochi (cf. [128], for references also). In this approach, the randomly changing waves are considered as a stochastic process (so that it is possible to evaluate the statistical properties of waves through the frequency and probability domains) and in deep water as a Gaussian random process. This Gaussian property was first found by Rudnick (1951) through analysis of measured data obtained in the Pacific Ocean. Moreover we will assume that this process is stationary. It is a common hypothesis for the study of ocean waves since the sea does behave stationary when observed in short periods of time (cf. Longuet-Higgins, Ochi).

We can deduce from the general CLT for the number of maxima of a stationary Gaussian process given in Theorem 1.2.2, the asymptotical behavior of functionals related to the ocean waves, as the number of local positive maxima of waves (studied when having a non-narrow-band-spectrum random process (cf. Ochi (1998), section 3.3)), or the waves amplitude, or even the amplitude associated to the acceleration (cf. Ochi (1998), section 4.3)). As an example, let us give the asymptotic behavior of the number of local positive maxima of waves. Indeed, as

noticed by Ochi, consideration of the negative maxima is unnecessary in the maximum-point-process of waves for engineering problems, because the negative maxima do not contribute to the largest peak value (the extreme wave amplitude) expected to occur in a certain number of observations.

So let us give the asymptotic behavior of the ratio $\dfrac{M^X_{[\beta,+\infty)}}{M^X_{[0,+\infty)}}$, with $\beta \geq 0$.

Corollary 1.2.1 *(Number of local positive maxima of waves, Kratz and León, 2000)*

$$\sqrt{T}\left(\frac{M^X_{[\beta,+\infty)}}{M^X_{[0,+\infty)}} - \frac{I\!E[M^X_{[\beta,+\infty)}]}{I\!E[M^X_{[0,+\infty)}]}\right) \xrightarrow{d} N(0,\sigma^2) \ as\ T \longrightarrow \infty,$$

where σ^2 is given by

$$
\begin{aligned}
\sigma^2 \ = \ & \frac{r^{(iv)}(0)}{-r''(0)2\pi C_0^4}\sum_{q=1}^{\infty}\sum_{0\leq n_1+m_1\leq q}\sum_{0\leq n_2+m_2\leq q}(C_0\delta_{n_1m_1}(\beta)-C_\beta\delta_{n_1m_1}(0))(C_0\delta_{n_2m_2}(\beta)-C_\beta\delta_{n_2m_2}(0))\\
& \times\frac{H_{q-(m_1+n_1)}(0)H_{q-(m_2+n_2)}(0)}{(q-(m_1+n_1))!(q-(m_2+n_2))!}\int_0^{+\infty}I\!E\left[H_{n_1}(X_0)H_{q-(n_1+m_1)}\left(\frac{\dot{X}_0}{\sqrt{-r''(0)}}\right)H_{m_1}(Z_0)\right.\\
& \left.\times H_{n_2}(X_s)H_{q-(n_2+m_2)}\left(\frac{\dot{X}_s}{\sqrt{-r''(0)}}\right)H_{m_2}(Z_s)\right]ds, \qquad (2.63)
\end{aligned}
$$

with $C_y := I\!E[M^X(y)] = \dfrac{\sqrt{r^{(iv)}(0)}}{2\pi\sqrt{-r''(0)}}\left(1-\Phi\left(y\sqrt{1+\dfrac{\rho_1^2}{\rho_2^2}}\right)\right)+\sqrt{\dfrac{-r''(0)}{2\pi}}\phi(y)\Phi\left(-\dfrac{\rho_1}{\rho_2}y\right),$

and in particular $C_0 = \dfrac{\sqrt{r^{(iv)}(0)-r''(0)}}{4\pi\sqrt{-r''(0)}}.$

Remark : for statistical motivation, a consistent estimator of the asymptotic variance can also be proposed (cf. [89]).

1.2.2 Multidimensional case : dimension n for the index set

In most cases, for applications in other fields, results in dimension larger than one are required. Generalizing results obtained in dimension one seems often possible in theory, but, depending upon the method, can reveal itself more difficult than expected.

There are different ways to consider the multidimensional case, one when choosing the stochastic process indexed by a set of dimension larger than one, the other when considering the process such that each of its components is a random vector, and finally when combining those two cases.

We will be interested in the first case, on which Wschebor ([180], 1985) has been working, generalizing results of Cramér et al. and Marcus for a d-parameter stochastic process $X = (X_t, t \in I\!R^d)$ having C^1 paths. Wschebor studied problems related to the level sets of the paths of X defined by $C^X_u := \{t : X_t = u\}$, proving at the same time some type of Rice formula.

For more details, you can refer to his 1985's book.

It is in this context that takes place the work of Kratz and León ([90], 2001) on Gaussian fields, in which they considered the problem of the asymptotic behavior of the length of a level curve of a Gaussian field, by adapting the method settled in the one dimensional case, for random processes indexed by a set of dimension larger than one.

Note that once again this study was motivated because of its various applications, in particular for the random modelisation of the sea (see for instance Azaïs, León and Ortega, [9]).

Consider a mean zero stationary Gaussian randon field $(X_{s,t}; (s,t) \in I\!\!R^2)$ with variance one and correlation function r having partial derivatives $\partial_{ij}r$, for $1 \le i + j \le 2$.

Assume that $r \in L^1$ satisfies $r^2(0,0) - r^2(s,t) \ne 0$, for $(s,t) \ne (0,0)$, and that $\partial_{02}r$, $\partial_{20}r$ are both in L^2.

Let $H(X)$ be the space of real square integrable functionals of the field $(X_{s,t}; (s,t) \in I\!\!R^2)$.

Let $\mathcal{L}^X_{Q(T)}(u)$ be the length of $\{(s,t) \in Q(T) : X_{s,t} = u\}$ the level curve at level u for the random field X, $Q(T)$ being the square $[-T,T] \times [-T,T]$ and $|Q(T)|$ its Lebesgue measure.

By theorem 3.2.5 of Federer (see [57], p. 244), we have for $g \in \mathbf{C}(I\!\!R)$

$$\int_{-\infty}^{\infty} g(u)\mathcal{L}^X_{Q(T)}(u)du = \int\int_{Q(T)} g(X_{s,t})||\nabla X_{s,t}||dsdt.$$

We will assume w.l.o.g. that $\lambda := I\!\!E[\partial_{10}^2 X_{s,t}] = I\!\!E[\partial_{01}^2 X_{s,t}] = 1$ and that X is isotropic, i.e. that the matrix of covariance of $(\partial_{10}X_{s,t}, \partial_{01}X_{s,t})$ is of the form $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ (see Adler, [1], §6.2).

(Note that in the non isotropic case, i.e. when the density function of $(\partial_{10}X_{s,t}, \partial_{01}X_{s,t})$ is $\mathcal{N}(0, \Sigma)$ with $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}$, $\Sigma_{i,j} \ne 0$ for $i \ne j$, we can consider the isotropic process $Y_{s,t}$ defined by

$Y_{s,t} = \dfrac{1}{\sqrt{r(0,0)}}X_{u,v}$, where $\begin{pmatrix} u \\ v \end{pmatrix} = \dfrac{1}{\sqrt{r(0,0)}}\Sigma^{-1/2}\begin{pmatrix} s \\ t \end{pmatrix}$, and deduce the results for $X_{s,t}$ from the ones of $Y_{s,t}$.)

Under these conditions, we can get the chaos expansion in $H(X)$ for $\mathcal{L}^X_{Q(T)}(u)$, as well as a CLT for it, namely :

Theorem 1.2.3 *(Kratz and León, 2001)*
In $H(X)$, we have

$$\mathcal{L}^X_{Q(T)}(u) = \sum_{q=0}^{\infty} \sum_{0 \le m+l \le [q/2]} d_{q-2(m+l)}(u)c_{2m,2l} \int\int_{Q(T)} H_{q-2(m+l)}(X_{s,t})H_{2m}(\partial_{10}X_{s,t})H_{2l}(\partial_{01}X_{s,t})dsdt$$

with $d_k(u) = \frac{1}{k!}H_k(u)\phi(u)$ (ϕ being the standard normal density) and

$$c_{2m,2l} = \frac{(-1)^{m+l}\sqrt{2\pi}}{m!l!2^{m+l}}\sum_{p_1=0}^{l}\sum_{p_2=0}^{m}\binom{l}{p_1}\binom{m}{p_2}\frac{(-1)^{p_1+p_2}}{B(p_1+p_2+1,1/2)} \text{ (B being the Beta function).}$$

The asymptotical behavior of $\mathcal{L}^X_{Q(T)}(u)$ is described by

$$\frac{\mathcal{L}^X_{Q(T)}(u) - I\!\!E[\mathcal{L}^X_{Q(T)}(u)]}{|Q(T)|^{1/2}} \xrightarrow[T\to\infty]{} \mathcal{N}(0,\sigma^2),$$

74

with $I\!E[\mathcal{L}^X_{Q(T)}(u)] = |Q(T)|\dfrac{\sqrt{2\pi}\phi(u)}{B(1,1/2)}$ *and*

$$\sigma^2 = \sum_{q=0}^{\infty} \sum_{0 \le m_1+l_1 \le [q/2]} \sum_{0 \le m_2+l_2 \le [q/2]} \delta_{q,2m_1,2l_1}(u)\delta_{q,2m_2,2l_2}(u) \int\!\!\int_{I\!R^2} I\!E[I_{q,2m_1,2l_1}(0,0)I_{q,2m_2,2l_2}(s,t)]dsdt,$$

where $\delta_{q,2m,2l}(u) := d_{q-2(m+l)}(u)c_{2m,2l}$, *and* $I_{q,2m,2l}(s,t) := H_{q-2(m+l)}(X_{s,t})H_{2m}(\partial_{10}X_{s,t})H_{2l}(\partial_{01}X_{s,t})$.

The proof is an adaptation to dimension 2 of the methods used to get Proposition 1.1.2 and Theorem 1.1.28 respectively.

Indeed we introduce $\mathcal{L}^X_{Q(T)}(u,\sigma) := \dfrac{1}{\sigma}\displaystyle\int_{I\!R} \mathcal{L}^X_{Q(T)}(v)\phi\left(\dfrac{u-v}{\sigma}\right) dv$ and prove that it converges to $\mathcal{L}^X_{Q(T)}(u)$ in L^2. Then the generalization to dimension 2 of Lemma 1.1.6 will give us the Hermite expansion of $\mathcal{L}^X_{Q(T)}(u,\sigma)$ with coefficients $d^\sigma_k(u)c_{2m,2l}$, with $d^\sigma_k(u) \underset{\sigma\to 0}{\to} d_k(u)$.

Lemma 1.2.1 *(Kratz and León, 2001)*
Let $f \in L^2(\phi(u,v)dudv)$ *and let* $(d_k, k \ge 0)$ *be its Hermite coefficients.*
One has the following expansion

$$\int\!\!\int_{Q(T)} f(X_{s,t})||\nabla X_{s,t}||dsdt =$$

$$\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{l=0}^{\infty} d_k c_{2m,2l} \int\!\!\int_{Q(T)} H_k(X_{s,t})H_{2m}(\partial_{10}X_{s,t})H_{2l}(\partial_{01}X_{s,t})dsdt =$$

$$\sum_{k=0}^{\infty}\sum_{0 \le m+l \le [q/2]} d_{k-2(m+l)}c_{2m,2l} \int\!\!\int_{Q(T)} H_{k-2(m+l)}(X_{s,t})H_{2m}(\partial_{10}X_{s,t})H_{2l}(\partial_{01}X_{s,t})dsdt,$$

where $c_{2m,2l}$ *is given in Theorem* 1.2.3.

The first part of theorem 1.2.3, i.e. the chaos expansion follows then, exactly as in dimension one.
On what concerns the CLT, we define

$$L_{Q(T)}(u) = \frac{\mathcal{L}^X_{Q(T)}(u) - I\!E[\mathcal{L}^X_{Q(T)}(u)]}{|Q(T)|^{1/2}}$$

and $L^K_{Q(T)}$ as the finite sum deduced from $L_{Q(T)}(u)$ for $q = 1$ to K, i.e.

$$L^K_{Q(T)}(u) = \frac{1}{|Q(T)|^{1/2}} \sum_{q=1}^{K} \sum_{0 \le m+l \le [q/2]} \delta_{q,2m,2l}(u) \int\!\!\int_{Q(T)} I_{q,2m,2l}(s,t)dsdt.$$

We define in the same way $L^K_{Q(T)}(u,\varepsilon)$ in which we consider the $1/\varepsilon$-dependent random field X^ε defined as

$$X^\varepsilon_{s,t} = \int_{I\!R^2} e^{i(s\lambda_1+t\lambda_2)}((f * \hat{\varphi}_\varepsilon)(\lambda_1,\lambda_2))^{\frac{1}{2}}dW(\lambda_1,\lambda_2).$$

We must however normalize the partial derivatives, dividing by the constant $(I\!E[\partial_{10}X^\varepsilon_{0,0}]^2)^{1/2} = (I\!E[\partial_{01}X^\varepsilon_{0,0}]^2)^{1/2} = (-\partial_{20}r(0,0))^{1/2}$ to get random variables with variance one.

The proof follows then the one of Theorem 1.1.28, using our combined method.

With respect to Rice formulae and the distribution of the maximum of Gaussian fields, let us also mention the recent works by Taylor et al. (2002 [171], 2004 [172]), and by Azaïs and Wschebor (2005, [15]).

1.3   Conclusion and references

In this work, attention was paid for this specific subject on stationary Gaussian case. The non-stationary case has also received many contributions, for instance by Hüsler and Bräker when dealing with the notion of local stationarity, by Azaïs and Mercadier when dealing with a constant variance, by Florens-Zmirou for diffusions, by Rudzkis, by Sjö, etc ...
and more specifically for classes of processes such that Brownian motions, fractional Brownian motions by Berzin, León and Ortega, etc ...

The non-Gaussian case has been considered by many authors, as Azaïs, Besson, Marcus and Wschebor, who provided very general proofs of Rice formulae. In particular, the stable processes have been studied for instance by Adler, Samorodnitsky and Gadrich (cf. [3], 1993, and [2], 1997), Azaïs (cf. [7], 1990), Davis and Resnick ([7], 1994), Michna and I. Rychlik (cf. [116], 1995), etc ...

For completeness, I tried to make an inventory of the maximum of papers and books dealing with the subject of level crossings, even though all the references are not explicitely mentioned in the synopsis above.
Note that papers published by authors of books before the publishing date of the books are not given since you can refer to the book's bibliography.

Bibliographie

[1] R. Adler, The Geometry of Random Fields, Wiley (1981).

[2] R. Adler and G. Samorodnitsky, Level crossings of absolutely continuous stationary symmetric α-stable processes, Ann. Appl. Probab. 7 (1997) 460-493.

[3] R. Adler, G. Samorodnitsky and T. Gadrich, The expected number of level crossings for stationary, harmonisable, symmetric, stable processes, Ann. Appl. Probab. 3 (1993) 553-575.

[4] J. Albin, Extremes and crossings for differentiable stationary processes with application to Gaussian processes in $I\!\!R^m$ and Hilbert space, Stochastic Processes Appl. 42 (1992) 119-147.

[5] M. Arcones, Limit theorems for non linear functionals of a stationary Gaussian sequence of vectors, Ann. Probab. 22 (1994) 2242-2274.

[6] J.-M. Azais, Approximation des trajectoires et temps local des diffusions, Ann. I.H.P. Vol 25 (2) (1989) 175-194.

[7] J-M. Azaïs, Conditions for convergence of number of crossings to the local time, Probab. Math. Stat. 11 (1990) 19-36.

[8] J-M. Azaïs and D. Florens-Zmirou, Approximations du temps local des processus gaussiens stationnaires par régularisation des trajectoires, Probab. Th. Rel. Fields 76 (1987) 121-132.

[9] J-M. Azaïs, J.R. león and J. Ortega, Geometrical characteristics of Gaussian sea waves, J. Appl. Probab. 42 (2005) 1-19.

[10] J-M. Azaïs and C. Mercadier, Asymptotic Poisson character of extremes in nonstationary Gaussian model, preprint (2003).

[11] J-M. Azaïs and M. Wschebor, Almost sure oscillation of certain random processes, Bernoulli 2 (1996) 257-270.

[12] J-M. Azaïs and M. Wschebor, Oscillation presque sûre de martingales continues, Séminaire de Probabilités XXXI, L.N.M.1655 (1997) 69-76.

[13] J-M. Azaïs and M. Wschebor, On the regularity of the distribution of the maximum of one-parameter Gaussian processes, Probab. Th. Rel. Fields 119 (2001a) 70-98.

[14] J-M. Azaïs and M. Wschebor, The distribution of the maximum of a Gaussian process : Rice method revisited, Prog. Probab. 51 (2001b) 321-348.

[15] J-M. Azaïs and M. Wschebor, On the distribution of the maximum of a Gaussian field with d parameters, Ann. Appl. Probab. 15 (2005) 254-278.

[16] S. Banach, Sur les lignes rectifiables et les surfaces dont l'aire est finie, Fund. Math. 7 (1925) 225-237.

[17] A. Barbour, L. Holst and S. Janson, Poisson approximation, Oxford Science Pub. (1992).

[18] I. Bar-David and A. Nemirovsky, Level crossings of nondifferentiable shot processes, IEEE Trans. Inform. Theory 18 (1972) 27-34.

[19] A. Barbé, A measure of the meanlevel-crossing activity of stationary normal processes, IEEE Trans. Inform. Theory 22 (1976) 96-102.

[20] Y.K. Belyaev, On the number of crossings of a level by a Gaussian random process, Th. Probab. Appl. 12 (1967) 392-404.

[21] Y.K. Belyaev, On the number of exits across the boundary of a region by a vector stochastic process, Th. Probab. Appl. 13 (1968) 320-324.

[22] S, Berman, A Central Limit Theorem for the renormalized self-intersection local time of a stationary vector Gaussian process, Ann. Probab. 20 (1992) 61-81.

[23] S. Berman, Sojourn and Extremes of Stochastic Processes, Wadsworth & Brooks-Cole (1992).

[24] C. Berzin and J.R. León, Weak convergence of the integrated number of level crossings to the local time for Wiener processes, Theory of Probab. and its Applic. 42 (1997) 568-579.

[25] C. Berzin, J.R. León and J. Ortega, Level crossings and local time for regularized Gaussian processes, Probab. Math. Stat. 18 (1998) 39-81.

[26] J.Besson, Sur le processus ponctuel des passages par un niveau d'une fonction aléatoire à trajectoires absolument continues, Thèse d'état de l'Université Lyon I (1984).

[27] J.-L. Besson and M. Wschebor, Sur la finitude des moments du nombre des passages par un niveau d'une fonction aléatoire, C.R.A.S. I, 297 (1983) 361-363.

[28] I. Blake and W. Lindsey, Level-crossing problems for random processes, IEEE Trans. Inform. Theory 19 (1973) 295-315.

[29] P. Boatwright and S.Riter, Level selection for a level crossing receiver, IEEE Trans.Comm. 21 (1973) 953-954.

[30] A.N. Borodin and I.A.Ibragimov, Limit theorems for functionals of random walk, A.M.S. Providence (1995).

[31] J. Breuer and P. Major, Central limit theorems for non-linear functionals of Gaussian fields, J. Mult. Anal. 13 (1983) 425-441.

[32] K. Breitung, Crossing rates for functions of Gaussian vector processes, Information theory, stat. dec. funct., r.proc., Trans. 11th Prague Conf. vol.A (1990) 303-314.

[33] E.V. Bulinskaya, On the mean number of crossings of a level by a stationary Gaussian process, Theory Probab. Appl. 6 (1961) 435-438.

[34] E. Cabaña, Estimation of the spectral moment by means of the extrema, Trabajos de Estadística y de investigación operativa 36 (1985) 71-80.

[35] E. Castillo, Extreme Value Theory in Engineering, Academic Press (1988)

[36] D. Chambers and E. Slud, Central limit theorems for nonlinear functionals of stationary Gaussian processes, Probab. Th. Rel. Fields 80 (1989) 323-346.

[37] D. Chambers and E. Slud, Necessary conditions for nonlinear functionals of stationary Gaussian processes to satisfy central limit theorems, Stoch. Proc. Applic. 32 (1989) 93-107.

[38] H. Chen, Poisson approximation for dependent trials, Ann. Probab. 3 (1975) 534-545.

[39] V. Chuberkis, Crossings of fixed levels by two stationary random processes, Automat. Remote Control 32 (1971) 1581-1587.

[40] H. Cramér and M.R. Leadbetter, Stationary and Related Stochastic Processes, New York : Wiley (1967).

[41] J. Cuzick, On the moments of the number of curve crossings by a stationary Gaussian process, Ph.D. Dissertation, Claremont Graduate School (1974).

[42] J. Cuzick, Conditions for finite moments of the number of zero crossings for Gaussian processes, Ann. Probab. 3 (1975) 849-858.

[43] J. Cuzick, A central limit theorem for the number of zeros of a stationary Gaussian process, Ann. Probab. 4 (1976) 547-556.

[44] J. Cuzick, A lower bound for the prediction error of stationary Gaussian process, Indiana Univ. Math. J. 26 (1977) 577-584.

[45] J. Cuzick, Local nondeterminism and the zeros of Gaussian processes, Ann. Probab. 6 (1978) 72-84.

[46] J. Cuzick, Correction :"Local nondeterminism and the zeros of Gaussian processes"', Ann. Probab. 15 (1987) 1229.

[47] J. Cuzick and G. Lindgren, Frequency estimation from crossings of un unknown level, Biometrika 67 (1980) 65-72.

[48] R. Davis and S. Resnick, Crossings of max-stable processes, J. Appl. Probab. 31 (1994) 130-138.

[49] P. Diananda, The central limit theorem for m-dependent variables, Proc. Cambridge Philos. Soc. 51 (1955) 92-95.

[50] P. Deheuvels, A. Karr, D. Pfeifer, R. Serfling, Poisson approximations in selected metrics by coupling and semigroup methods with applications, J. Stat. Planning Inf. 20 (1988) 1-22.

[51] P. Deheuvels , D. Pfeifer, Poisson approximations of mutinomial distributions and point processes, J. Mult. Anal. 25 (1988) 65-89.

[52] P. Deheuvels, D. Pfeifer and M. Puri, A new semigroup technique in Poisson approximation, Semigroup forum 38 (1989) 189-201.

[53] R. Dobrushin, Gaussian and their subordinated self-similar random generalized fields, Ann. Probab. 7 (1979) 1-28.

[54] R. Dobrushin and P. Major, Non-central limit theorems for non-linear functionals of Gaussian fields, Z. Wahrsch. verw. Geb. 50 (1979) 27-52.

[55] M. Ershov, A limit theorem for the number of level-crossings by a stationary Gaussian process, Dokl. Akad. Nauk SSSR. 177 (1967) 1259-1262.

[56] M. Falk, J. Hüsler and R.-D. Reiss, Laws of Small Numbers : Extremes and Rare Events, Birkhäuser (1994).

[57] H. Federer, Geometric Measure Theory, Springer Verlag (1969).

[58] D. Florens-Zmirou, Statistics on crossings of discretized diffusions and local time, Stoch. Proc. Applic. 39 (1991) 139-151.

[59] D. Geman, On the variance of the number of zeros of a stationary Gaussian process, Ann.Math.Stat. 43 (1972) 977-982.

[60] D. Geman and J. Horowitz, Occupation densities, Ann. Probab.8 (1980) 1-67.

[61] L. Giraitis and D. Surgailis, CLT and other Limit Theorems for Functionals of Gaussian Processes, Z. Wahrsch. verw. Geb. 70 (1985) 191-212.

[62] M. Graça Temido, Maxima and minima of stationary random sequences under a local dependence restriction, Port. Math. 56 (1999) 59-72.

[63] P. Hall, On the rate of convergence of normal extremes, J.Appl.Probab.16 (1979) 433-439.

[64] H. Harter, A bibliography of extreme-value theory, Intern. Statist. Rev. 46 (1978) 279-306.

[65] A. Hasofer, The upcrossing rate of a class of stochastic processes, In Studies in Probability and Statistics (1976) 153-159.

[66] C. Helstrom, The distribution of the number of crossings of a Gaussian stochastic process, IEEE Trans. Inform. Theory 3 (1957) 232-237.

[67] W. Hoeffding and H. Robbins, The Central Limit Theorem for dependent random variables, Duke Math. J. 15 (1948) 773-780.

[68] L. Holst and S. Janson, Poisson approximation using the Stein-Chen method and coupling : number of exceedances of Gaussian random variables, Ann. Probab. 18 (1990) 713-723.

[69] H-C Ho and T-C Sun, A central limit theorem for non-instanataneous filters of a stationary Gaussian process, J. Multivariate Anal. 22 (1987) 144-155.

[70] H-C Ho and T-C Sun, On central and non-central limit theorems for non-linear functions of a stationary Gaussian process, In Dependence in Probab. and Statist., Oberwolfach (1985) 3-19.

[71] T. Hsing, J. Hüsler and M.R. Leadbetter, On the exceedance point process for a stationary sequence, Probab. Th. Rel. Fields 78 (1988) 97-112.

[72] J. Hüsler, On limit distributions of first crossing points of Gaussian sequences, Stochastic Process. Appl. 6 (1977) 65-75.

[73] J. Hüsler, Asymptotic approximations of crossing probabilities of random sequences, Z. Wahrsch. verw. Geb. 63 (1983) 257-270.

[74] J. Hüsler, Extreme values of non-stationary random sequences, J. Appl. Probab. 23 (1986) 937-950.

[75] J. Hüsler, Extreme values and high boundary crossings of locally stationary Gaussian processes, Ann. Probab. 18 (1990) 1141-1158.

[76] J. Hüsler and M. Kratz, On the convergence of the number of exceedances of nonstationary normal sequences, Journal of Research of the NIST, vol. 99 (1994) 539-542.

[77] J. Hüsler and M. Kratz, Rate of Poisson approximation of the number of exceedances of nonstationary normal sequences, Stoch. Proc. Applic. 55 (1995) 301-313.

[78] P. Imkeller, V. Perez-Abreu and J. Vives, Chaos expansions of double intersection local time of Brownian motion in R^d and renormalization, Stoch. Proc. Applic. 56 (1995) 1-34.

[79] V.A. Ivanov, On the average number of crossings of a level by sample functions of a stochastic process, Theory Probab. Appl. 5 (1960) 319-323.

[80] K. Itô, The expected number of zeros of continuous stationary Gaussian processes, J. Math. Kyoto Univ. 3 (1964) 207-216.

[81] J.Jacod, Rates of convergence to the local time of a diffusion, Ann. Inst. H. Poincaré, Probab. Stat. 34 (1998) 505-544.

[82] M. Kac, On the average number of real roots of a random algebraic equation, Bull.Amer.Math.Soc. 49 (1943) 314-320.

[83] G. Kalliampur, Stochastic filtering theory, (1980) Springer.

[84] B. Kedem, Sufficiency and the number of level crossings by a stationary process, Biometrika 65 (1978), 207-210.

[85] A. Kostyukov, Curve-crossing problems for Gaussian stochastic processes and its application to neural modeling, Biol. Cybernetics 29 (1978) 187-191.

[86] M.F. Kratz, Statistique des queues de distribution et approximation Poissonnienne, Thèse de doctorat de l'Université Pierre et Marie Curie, Paris VI (1993).

[87] M. Kratz, Chaos expansions and level crossings, preprint Samos 127 (non submitted, 2000).

[88] M. Kratz and J. León, Hermite polynomial expansion for non-smooth functionals of stationary Gaussian processes : crossings and extremes, Stoch. Proc. Applic. vol 66 (1997) 237-252.

[89] M. Kratz and J. León, Central limit theorems for the number of maxima and some estimator of the second spectral moment of a stationary Gaussian process. Applications in hydroscience, Extremes 3 :1 (2000) 57-86.

[90] M. Kratz and J. León, Central limit theorems for level functionals of stationary Gaussian processes and fields, J.Theoretical Probab.14 (2001) 639-672.

[91] M. Kratz and J. León, On the second moment of the number of crossings by a stationary Gaussian process, preprint Samos 203 ou Map5-21 (2004).

[92] M. Kratz and H. Rootzén, On the rate of convergence for extremes of mean square differentiable stationary normal processes, J. Applied Probab.34 (1997) 908-923.

[93] L. le Cam, An approximation theorem for the Poisson binomial distribution, Pacific J. Math. 10 (1960) 1181-1197.

[94] M.R. Leadbetter, G. Lindgren and H. Rootzén, Extremes and related properties of random sequences and processes, Springer Series in Statistics, Springer, New York, 1983.

[95] M.R. Leadbetter and H. Rootzén, Extremal theory for stochastic processes, Ann. Probab. 16 (1988) 431-478.

[96] M. Ledoux and M. Talagrand, Probability on Banach spaces, Wiley, New-York (1991).

[97] J.R. León and J. Ortega, Level crossings and local times for regularized Gaussian processes : L^2 convergence, C.R.A.S. 314, I (1992) 227-331.

[98] W.V. Li and Q.-M. Shao, A normal comparison inequality and its applications, Probab. Theor. Rel. Fields 122 (2002) 494-508.

[99] M.A. Lifshits, Gaussian random functions, Math. and its Applic. 322, Kluwer (1995).

[100] G. Lindgren, Use and structure of Slepian model processes for prediction and detection in crossing and extreme value theory, In Statistical Extremes and Applications (J. Tiago de Oliveira, ed.), Reidel, Dordrecht (1984) 261-284.

[101] G. Lindgren and I. Rychlik, Slepian models and regression approximations in crossing and extreme value theory, Int. stat. Rev. 59 (1991) 195-225.

[102] M.S. Longuet-Higgins, Reflection and refraction at a random surface. I,II,III, Journal of the Optical Society of America 50 (1960) 838-856.

[103] I. Macak, On the number of crossings of a curve by a random process, Theory Probab. math. Statist. 17 (1979) 101-110.

[104] I. Macak, On the tangencies and intersections of a random process with a continuous curve, Theory Probab. math. Statist. 18 (1979) 103-107.

[105] P. Major, Multiple Wiener-Itô integrals, Lecture Notes Math. 849 (1981) Springer.

[106] P. Major, Limit theorems for non-linear functionals of Gaussian sequences, Probab. Theor. Rel. Fields 57 (1981b) 129-158.

[107] P. Major, A note on the approximation of the uniform empirical process, Ann. Probab. 18 (1990) 129-139.

[108] T. Malevich, Asymptotic normality of the number of crossings of level 0 by a Gaussian process, Theory Probab. Applic. 14 (1969) 287-295.

[109] T. Malevich, Some bounds for the probabilities of events generated by Gaussian processes and fields, and applications to problems of level crossings, Theory Probab. Appl. 19 (1974) 142-153.

[110] T. Malevich, On conditions for the moments of the number of zeros of Gaussian stationary processes to be finite, Theory Probab. Appl. 24 (1979) 741-753.

[111] M.B. Marcus, Level crossings of a stochastic process with absolutely continuous sample paths, Ann. Probab. 5 (1977) 52-71.

[112] M.B. Marcus, Some bounds for the expected number of level crossings of harmonizable p-stable processes, Stoch. Proc. Appl. 33 (1989) 217-231.

[113] G. Maruyama, Nonlinear functionals of Gaussian stationary processes and their applications, In Maruyama and Prokhorov (eds.) Lecture Notes Math. 550 (1976) 375-378.

[114] G. Maruyama, Wiener functionals and probability limit theorems I : the centtral limit theorems, Osaka J. Math. 22 (1985) 697-732.

[115] Z. Michna, Ruin probabilities and first passage times for self-similar processes, Lund University (1998).

[116] Z. Michna and I. Rychlik, The expected number of level crossing for certain symmetric α-stable processes, Commun. Stat. Stochastic Models 11 (1995) 1-19.

[117] R. Miroshin, On the finiteness of the number of zeros of a differentiable Gaussian stationary process, Dokl. Akad. Nauk SSSR. 200 (1971) 1321-1324.

[118] R. Miroshin, A sufficient condition for finiteness of moments of the number of zeros of a differentiable stationary Gaussian process, Theory Probab. Applic. 18 (1973) 454-463.

[119] R. Miroshin, A necessary condition for finiteness of moments of the number of zeros of a differentiable stationary Gaussian process, Theory Probab. Applic. 19 (1974) 570-577.

[120] Y. Mittal, A new mixing condition for stationary Gaussian process, Ann. Probab. 7, (1979) 724-730.

[121] Y. Mittal, Comparison technique for highly dependent stationary Gaussian processes, in Statistical Extremes and Applications (J. Tiago de Oliveira, ed.), Reidel, Dordrecht (1984) 181-195.

[122] Y. Mittal and D. Ylvisaker, Limit distributions for the maxima of stationary Gaussian processes, Stoch. Proc. Appl. 3 (1975) 1-18.

[123] S. Nayak and K. Wali, On the number of boundary crossings related to LIL and SSLN for record values and partial maxima of i.i.d. sequences and extremes of uniform spacings, Stochastic Processes Appl. 43 (1992) 317-329.

[124] D. Nualart, The Malliavin Calculus and Related Topics, Springer-Verlag (1995).

[125] D. Nualart and J. Vives, Chaos expansions and local times, Publicacions Matemàtiques 36 (1992a) 827-836.

[126] D. Nualart and J. Vives, Smoothness of Brownian local times and related functionals, Potential Analysis 1 (1992b) 257-263.

[127] D. Nualart and M. Wschebor, Intégration par parties dans l'espace de Wiener et approximation du temps local, Probab. Th. Rel. Fields 90 (1991) 83-109.

[128] Ochi, M. (1998) : *Ocean waves. The stochastic approach.*, Cambridge ocean technology series 6.

[129] G.Perera and M.Wschebor, Crossings and occupation measures for a class of semimartingales, Ann. Probab. 26 (1998) 253-266.

[130] J. Pickands, Upcrossings probabilities for stationary Gaussian processes, Trans. Amer. Math. Soc. 145 (1969) 51-73.

[131] J. Pickands, Asymptotic properties of the maximum in a stationary Gaussian process, Trans. Amer. Math. Soc. 145 (1969) 75-86.

[132] V.I. Piterbarg, The central limit theorem for the number of level crossings of a stationary Gaussian process, Theory Probab. Appl. 23 (1978) 178-182.

[133] V.I. Piterbarg, Asymptotic expansions for the probability of large excursions of Gaussian processes, Soviet Math. Dokl. 19 (1978) 1279-1283.

[134] Pickersgill, and V.I. Piterbarg, Estimation of the rate of convergence in a Poisson limit theorem for the excursions of stationary Gaussian processes, Theory of Probab. and Math. Stat. t.36 (1987) (in Russian).

[135] V.I. Piterbarg, Large deviations of random processes close to Gaussian ones, Theory Prob. Appl. 27 (1982) 504-524.

[136] V.I. Piterbarg, Asymptotic methods in the theory of Gaussian processes and fields, Trans. Math. Monographs, 148 (1996).

[137] V.I. Piterbarg, Rice method for Gaussian random fields, Fundam. Prikl.Mat. 2 (1996) 187-204.

[138] V.I. Piterbarg, Central limit theorem for wave-functionals of Gaussian processes, Adv. Appl. Probab. 31 (1999) 158-177.

[139] R. Plackett, A reduction formula for normal multivariate integrals, Biometrika 41 (1954) 351-360.

[140] C. Qualls, On a limit distribution of high level crossings of a stationary Gaussian process, Ann. Math. Statist. 39 (1968) 2108-2113.

[141] C. Qualls, On the joint distribution of crossings of high nultiple levels by a stationary Gaussian process, Ark. Mat. 8 (1969) 129-137.

[142] S.I. Resnick, Extreme values, Regular Variation and Point processes, Springer (1987).

[143] S.O. Rice, Mathematical analysis of random noise, Bell System Tech. J. 23-24 (1944,1945).

[144] P. Roes, On the expected number of crossings of a level in certain stochastic processes, J. Appl. Probab. 7 (1970) 766-770.

[145] H. Rootzén, The rate of convergence of extremes of stationary random sequences, Adv. Appl. Probab. 15 (1983) 54-80.

[146] H. Rootzén, Extreme value theory for stochastic processes, Scan. Actuar. J. 1 (1995) 54-65.

[147] M. Rosenblatt, Independence and dependence, Proc. 4th Berkeley Sympos. Math. Statist. Probab., Berkeley Univ. Calif. Press (1961) 411-443.

[148] M. Rosenblatt, Some limit theorems for partial sums of quadratic forms in stationary Gaussian variables, Z. Wahrsch. verw. 49 (1979) 125-132.

[149] M. Rosenblatt, Limit theorems for Fourier transforms of functionals of Gaussian sequences, Z. Wahrsch. verw. Geb. 55 (1981) 123-132.

[150] R. Rudzkis, Probability of large rejection of a nonstationary Gaussian process. I, Liet. Mat. Rinkinys 25. 1 (1985) 143-154.

[151] R. Rudzkis, Probability of large rejection of a nonstationary Gaussian process. I, Liet. Mat. Rinkinys 25. 2 (1985) 148-163.

[152] I. Rychlick, On some reliability applications of Rice's formula for the intensity of level crossings, Extremes 3 :4 (2000), 331-348.

[153] H. Sato and M. Tanabe, Level dependence of the variance of the number of level-crossing points of a stationary Gaussian process, IEEE Trans. Inform. Theory 25 (1979) 127-129.

[154] R. Serfling, A general Poisson approximation theorem, Ann. Probab. 3 (1975) 726-731.

[155] R. Serfling, Some elementary Results on Poisson approximation in a sequence of Bernoulli trials, Siam review 20 (1978) 567-579.

[156] R. Serfozo, High level exceedances of regenerative and semi-stationary processes, J. Appl. Probab. 17 (1980) 423-431.

[157] K. Sharpe, Level crossings of a constructed process, J. Appl. Probab. 16 (1979) 206-212.

[158] E. Sjö, Crossings and maxima in Gaussian fields and seas, Phd thesis, Lund University (2000).

[159] M. Sköld, Kernel intensity estimation for marks and crossing of differentiable stochastic processes, Theory Stoch. Process. 2(18) (1996) 273-284.

[160] D. Slepian, The one-side barrier problem for Gaussian noise, Bell Syst. Tech. J. 41 (1962) 463-501.

[161] E. Slud, Multiple Wiener-Itô integral expansions for level-crossing-count functionals, Prob. Th. Rel. Fields 87 (1991) 349-364.

[162] E. Slud, MWI expansions for functionals related to level-crossing counts, in : Chaos Expansions, Multiple Wiener-Ito Integrals and their Applications by Christian Houdré and Victor Pérez-Abreu, Probability and Stochastics Series (1994a) pp. 125-143.

[163] E. Slud, MWI representation of the number of curve-crossings by a differentiable Gaussian process, with applications, Ann.Probab.22 (1994b) 1355-1380.

[164] R. Smith, Extreme value theory for dependent sequences via the Stein-Chen method of Poisson approximation, Stoch. Process. Appl. 30 (1988) 317-327.

[165] J. Tiago de Oliveira, Bibliography of Extreme Value Theory, (1987).

[166] C. Stein, Approximate computation of expectations, IMS-Lecture Notes 7 (1986).

[167] M. Talagrand, The generic chaining, Springer (2005).

[168] M. Taqqu, Weak convergence to fractional Brownian motion and to the Rosenblatt process, Z. Wahrsch. verw. Geb. 31 (1975) 287-302.

[169] M. Taqqu, Law of Iterated Logarithm for Sums of Non-Linear Functions of Gaussian Variables that Exhibit a Long Range Dependence, Z. Wahrsch. verw. Geb. 40 (1977) 203-238.

[170] M. Taqqu, Convergence of integrated processes of arbitrary Hermite rank, Z. Wahrsch. verw. Geb. 50 (1979) 53-83.

[171] J. Taylor and R. Adler, Euler characteristics for Gaussian fields and manifolds, Ann. Probab. 30 (2002) 533-563.

[172] J. Taylor, A. Takemura and R. Adler, Validity of the expected Euler characteristic heuristic, Ann. Probab., to appear.

[173] Y. Tong, The multivariate normal distribution, Springer, New-York (1990).

[174] V.S. Tsyrelson, The density of the maximum of a Gaussian process, Theory Probab. Appl. 20 (1975) 817-856.

[175] V. Volkonskii and Y. Rozanov, Some limit theorems for random functions II, Theory Probab. Appl. 6 (1961) 186-198.

[176] S. Watanabe, Lectures on Stochastic Differential Equations and Malliavin Calculus, Springer Berlin (1984).

[177] H. Witte, A unification of some approaches to Poisson approximation, J. Appl. Probab. 27 (1990) 611-621.

[178] R. Woods and R. Gonzales, Sampling considerations for multilevel-crossing analysis, IEEE Trans. Pattern anal. Machine Intelligence 4 (1982), 117-123.

[179] M. Wschebor, Régularisation des trajectoires et approximation du temps local, CRAS I-29 (1984) 209-212.

[180] M. Wschebor, Surfaces Aléatoires, L.N.M. 1147 (1985) Springer Verlag.

[181] M. Wschebor, Sur les accroissements du processus de Wiener, CRAS 315 I (1992) 1293-1296.

[182] N. Ylvisaker, The expected number of zeros of a stationary Gaussian process, Ann. Math. Stat. 36 (1965) 1043-1046.

[183] N. Ylvisaker, On a theorem of Cramér and Leadbetter, Ann. Math. Statist. 37 (1966) 682-685.

Chapitre 2

Some contributions in statistics of extremes and in statistical mechanics

This part will be divided into two independent sections, one devoted to the heavy tail modeling, the other to the random energy model bringing us back in a Gaussian background.

2.1   Heavy tail modeling

Those past two decades, the constant evolution in computer science has generated a new approach in modelisation since we have started to deal with large amounts of data, for instance in ethernet traffic. Thus the classical queuing assumption of thin tails and independence had to be dropped for this type of data, replaced by the notions of long range dependence (characterized by a slow rate of decay of the covariance function), self similarity (invariance in distribution under some scaling of time and space) and heavy tails. Because of numerous applications and of the fact that heavy tails are one of the causes of long range dependence, the interest in studying heavy tailed phenomena has growed sensibly those past years, in theory as in practice. For a brief review of the subject and some references, see the special invited paper by S. Resnick (*Ann. Stat. 1997, [22]*). We will focuse here in heavy tailed modeling.

We say that X has a heavy tailed distribution F if

$$1 - F(x) = x^{-\alpha} L(x), \quad \text{as } x \to \infty, \tag{1.1}$$

where L is a slowly varying (at infinity) function, i.e.

$$\forall x \geq 1, \quad \lim_{t \to \infty} \frac{L(tx)}{L(t)} = 1. \tag{1.2}$$

Recall also the Karamata's representation for L, namely

$$L(t) = a(t) \exp\left\{\int_{t_0}^t \frac{b(s)}{s} ds\right\}, \ t \geq t_0, \ \text{with } \lim_{t \to \infty} a(t) = a_0 \in]0, \infty[\ \text{ and } \ \lim_{t \to \infty} b(t) = 0.$$

Another way to define a heavy tailed distribution is to say that $1 - F$ is regularly varying with index α (denoted by $RV_{-\alpha}$), i.e. for all $x > 0$,

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}. \tag{1.3}$$

This notion of regular variation constitutes a basic analytical tool in extreme value theory as well as for stable processes; moreover, in this context, to establish asymptoptical results and to calculate rates of convergence, we need to introduce the notion of second order regular variation:

$1 - F$ is a second order regular varying function with first order parameter $-\alpha$ and second order parameter ρ (denoted by $1 - F \in 2RV(-\alpha, \rho)$) if there exists a function A with $\lim_{t \to \infty} A(t) = 0$ such that

$$\lim_{t \to \infty} \frac{\frac{1-F(tx)}{1-F(t)} - x^{-\alpha}}{A(t)} = cx^{-\alpha} \int_1^x u^{\rho-1} du, \qquad \text{for } x > 0 \text{ and } c \neq 0. \tag{1.4}$$

Note that necessarily $A(\cdot)$ is regularly varying of index ρ and $1 - F$ is regularly varying of index $-\alpha$. The form of the limit is discussed in de Haan and Stadmüller (1996, [17]); see also Geluk and de Haan (1987, [15]).

Having the definitions in hand, we will present two contributions to the study on the detection and modeling of heavy tailed phenomena. The first one gives a graphical technique which can be helpful for the detection of heavy tailed phenomena, and provides the estimation of the first order parameter α of a regular varying function defined in (1.3) when trying to fit a heavy tailed model to a data set; the second one deals with linear programming estimators in heavy tailed time series models.

2.1.1   The QQ–Estimator and Heavy Tails

Introduction

Many empirical phenomena can be characterized by the mathematical form of the tails of distribution, at least approximately, whereas the global form of the distribution is unknown.

When being interested in establishing the form of the upper tail distribution of a random variable, arises the problem of the estimation of the parameters determining the upper tail (but not the entire distribution). Appropriate methods used to solve this problem are quite different than those of conventional parametric and nonparametric statistical inference. They rely on the conditionning upon the upper r order statistics of the data, for some r that can be determined in a variety of ways, such that data-analytic or decision-theoretic ways. A brief survey of those methods, as well as of the problems of modeling, testing and forecasting in this area, has been proposed by Hill (cf. [19] and [20]).

We are restricting our focus on estimating the tail index α ($\alpha > 0$) of a distribution F with a regularly varying upper tail satisfying (1.1), where L is an unknown nuisance function slowly varying at infinity. Given a n-sample $Z_1, \cdots, Z_n$ with distribution F satisfying (1.1) and with order statistics $Z_{1,n} \leq \cdots \leq Z_{n,n}$, Hill proposed in 1975 ([18]) an estimator of α^{-1} (commonly named after him)

$$\widehat{\alpha_H^{-1}} = \frac{1}{k_n} \sum_{i=1}^{k_n} \log \left(\frac{Z_{n+1-i,n}}{Z_{n-k_n,n}} \right), \tag{1.5}$$

where the k_n are positive integers which, for theoretical asymptotic considerations, satisfy the

conditions

$$1 \leq k_n < n, \quad k_n \to \infty, \quad \frac{k_n}{n} \to 0, \quad \text{as} \quad n \to \infty. \tag{1.6}$$

The asymptotic properties of Hill's estimator have been investigated by many authors as Mason (1982) for the weak consistency, Hall (1982) for the asymptotic normality, Deheuvels, Haeusler and Mason (1988) for the almost sure behavior (with the additional condition $k_n/\log\log n \to \infty$ as $n \to \infty$), only to mention the pioneers ...
Since Hill's paper (1975, [18]), many more estimators related to Hill's estimator have been proposed, also in the case of dependent data, and studied, using the theory of order statistics. It has been motivated by a large amount of applications in economics, business, insurance, social sciences, industry, hydrology, reliability and teletraffic engineering, etc...
For a review on this study of Hill's estimator and related estimators, with a long list of references (for applications also), we refer the reader to Csörgő and Viharos (1998, [3]). Note also that there are alternative methods to avoid order statistics to construct estimators, but it won't be our purpose here.

The qq-estimator

Much effort at detecting heavy tails has centered around the Hill estimator (1.5), which is based on using $k = k(n)$ upper order statistics when the sample size is n. Choosing k in a sensible manner is difficult, the estimator can exhibit outrageous bias and graphical aids are often very difficult to interpret accurately. So it is wise to consider alternative methods to supplement information given by the Hill estimator and associated plots. One way is to turn to the qq-plot.

The qq-method is based on the following simple observation :
if $U_{1,n} \leq U_{2,n} \leq \ldots U_{n,n}$ are the order statistics from n iid observations which are uniformly distributed on $[0,1]$, then the plot of $\{(i/(n+1), U_{i,n}), 1 \leq i \leq n\}$ should be roughly linear, since $U_{i,n}$ should be close to its mean $i/(n+1)$.
Now suppose that $X_{1,n} \leq X_{2,n} \leq \ldots X_{n,n}$ are the order statistics from an iid sample of size n which we suspect comes from a particular continuous distribution G;
the plot of $\{(i/(n+1), G(X_{i,n})), 1 \leq i \leq n\}$ should be approximately linear and hence also the plot of $\{(G^{\leftarrow}(i/(n+1)), X_{i,n}), 1 \leq i \leq n\}$ should be approximately linear. Note $G^{\leftarrow}(i/(n+1))$ is a theoretical quantile and $X_{i,n}$ is the correponding quantile of the empirical distribution function and hence the name *qq-plot*.

Suppose we suspect the data comes from a location–scale family

$$G_{\mu,\sigma}(x) = G_{0,1}\left(\frac{x-\mu}{\sigma}\right), \quad \text{where } \mu, \sigma \text{ are unknown.}$$

The plot of $\{(G_{\mu,\sigma}^{\leftarrow}(i/(n+1)), X_{i,n}), 1 \leq i \leq n\}$ should be approximately a line through 0 of slope 1 and since

$$G_{\mu,\sigma}^{\leftarrow}(y) = \sigma G_{0,1}^{\leftarrow}(y) + \mu$$

the plot of

$$\{(G_{0,1}^{\leftarrow}(i/(n+1), X_{i,n}), 1 \leq i \leq n\}$$

should be approximately a line of slope σ and intercept μ. Thus visually we can assess the goodness of fit of the location–scale family and provide estimates of μ, σ.

The relevance of this technique to heavy tails is the following.

Suppose we suspect that $Z_{1,n} \leq \cdots \leq Z_{n,n}$ are the order statistics from a random sample from a Pareto family indexed by its shape parameter $\alpha > 0$:

$$F_\alpha(x) = 1 - x^{-\alpha}, \quad x \geq 1.$$

Then of course for $y > 0$

$$G_{0,\alpha}(y) := P[\log Z_1 > y] = e^{-\alpha y}$$

and the plot of

$$\{(G_{0,1}^{\leftarrow}(i/(n+1))), \log Z_{i,n}), 1 \leq i \leq n\} = \{(-\log(1 - i/(n+1)), \log Z_{i,n}), 1 \leq i \leq n\}$$

should be approximately a line with intercept 0 and slope α^{-1}.

Thus for the Pareto example, if we set

$$x_i = -\log\left(1 - \frac{i}{n+1}\right) \quad \text{and} \quad y_i = \log Z_{n,i},$$

then an estimator of α^{-1} corresponds to the slope of the least squares line through the points $\{(x_i, y_i), 1 \leq i \leq n\}$, i.e.

$$\widehat{\alpha^{-1}} = \frac{\sum_{i=1}^n -\log(\frac{i}{n+1})\{n \log Z_{n-i+1,n} - \sum_{j=1}^n \log Z_{n-j+1,n}\}}{n\sum_{i=1}^n(-\log(\frac{i}{n+1}))^2 - (\sum_{i=1}^n -\log(\frac{i}{n+1}))^2} \tag{1.7}$$

which we call the **qq–estimator**.

Suppose now that the underlying distribution is not exactly Pareto and to make the more general and less ad-hoc assumption, we consider regularly varying tails.

Let $Z_1, \ldots, Z_n$ be a random sample from a heavy tailed distribution F satisfying (1.1).

We now modify (1.7) to make it suitable for the regularly varying case. Observe that since $1 - F$ is regularly varying, we have for large t,

$$\frac{1 - F(tx)}{1 - F(t)} = P\left[\frac{Z_1}{t} > x \;\middle|\; Z_1 > t\right] \approx x^{-\alpha}. \tag{1.8}$$

Now choose $k = k(n) \to \infty$ such that $k/n \to 0$. Then the $(k+1)$st largest order statistic $Z_{n-k,n}$ satisfies $Z_{n-k,n} \overset{P}{\to} \infty$ as $n \to \infty$. This follows, for example, from Smirnov (1952). Conditional on $Z_{n-k,n}$ we have that $Z_{n,n}, Z_{n-1,n}, \ldots, Z_{n-k+1,n}$ have the distribution of the order statistics from a random sample of size k from a distribution concentrating on $(Z_{n-k,n}, \infty)$ of the form $F(\cdot)/(1 - F(Z_{n-k,n}))$. Thus conditional on $Z_{n-k,n}$,

$$\left(\frac{Z_{n-k+i,n}}{Z_{n-k,n}}, i = 1, \ldots, k\right)$$

behave like the order statistics from a sample of size k from the distribution concentrating on $(1, \infty)$ with tail

$$\frac{1 - F(Z_{n-k,n}x)}{1 - F(Z_{n-k,n})} \approx x^{-\alpha}$$

where the approximation follows from (1.8). The qq–estimator can be defined from the upper k order statistics, as the slope of the points $\{(-\log(1 - i/(k+1)), \log(Z_{n-k+i,n}/Z_{n-k,n})),\ 1 \leq i \leq k\}$, or also of the points

$$\{(-\log(1 - i/(k+1)), \log Z_{n-k+i,n}),\ 1 \leq i \leq k\}.$$

In practice we would make a qq–plot of all the data and choose k based on visual observation of the portion of the graph which looked linear. Then we would compute the slope of the line through the chosen upper k order statistics and the corresponding exponential quantiles. Choosing k is very difficult and the estimate of α is usually rather sensitive to the choice of k. Alternatively, one can plot $\{(k, \widehat{\alpha^{-1}}(k)), 1 \leq k \leq n\}$ and look for a stable region of the graph as representing the true value of α^{-1}. This is analogous to what is done with the Hill estimator of α^{-1} defined in (1.5).

Under the regular variation assumption (1.1), we can prove on one hand the weak consistency of the qq-estimator $\widehat{\alpha^{-1}}$, on the other hand the asymptotic normality when considering a second order strengthening of (1.1).

Theorem 2.1.1 *(Kratz and Resnick, 1996)*
Suppose $k = k(n) \to \infty$ in such a way that as $n \to \infty$ we have $k/n \to 0$. Suppose $Z_1, \ldots, Z_n$ are a random sample from F, a distribution with regularly varying tail satisfying (1.1). Then the qq–estimator $\widehat{\alpha^{-1}}$ is weakly consistent for $1/\alpha$:

$$\widehat{\alpha^{-1}} \overset{P}{\to} \alpha^{-1} \qquad as \quad n \to \infty.$$

To prove this theorem, we write the qq-estimator $\widehat{\alpha^{-1}}$ as a function of the Hill estimator $\widehat{\alpha_H^{-1}} = \widehat{\alpha_H^{-1}}(k)$ (defined in (1.5)) as

$$\widehat{\alpha^{-1}} = \frac{\frac{1}{k}\sum_{i=1}^{k}(-\log(1 - \frac{i}{k+1}))\log(\frac{Z_{n-k+i,n}}{Z_{n-k,n})}) - \frac{1}{k}\sum_{i=1}^{k}(-\log(1 - \frac{i}{k+1}))\widehat{\alpha_H^{-1}}(k)}{\frac{1}{k}\sum_{i=1}^{k}(-\log(1 - \frac{i}{k+1}))^2 - (\frac{1}{k}\sum_{i=1}^{k}(-\log(1 - \frac{i}{k+1}))^2},$$

so that, by using the weak consistency of the Hill estimator (Mason, 1982), it suffices to show that

$$\frac{1}{k}\sum_{i=1}^{k}(-\log(1 - i/(k+1)))\log(Z_{n-k+i,n}/Z_{n-k,n}) \overset{P}{\to} \frac{2}{\alpha}.$$

To this end, we used some classical important results as the Potter's inequalities, the Renyi's representation of order statistics, and one simple lemma, which might be of some interest on its own ; let us recall these results.
Potter's inequalities take the following form : since $1/(1 - F)$ is regularly varying with index α, the inverse $U = (1/(1 - F))^{\leftarrow}$ is regularly varying with index $1/\alpha$ and for $\epsilon > 0$, there exists $t_0 = t_0(\epsilon)$ such that if $y \geq 1$ and $t \geq t_0$

$$(1 - \epsilon)y^{\alpha^{-1}-\epsilon} \leq \frac{U(ty)}{U(t)} \leq (1 + \epsilon)y^{\alpha^{-1}+\epsilon}.$$

The Renyi representation states that if $E_{1,n} \le E_{2,n} \le \cdots \le E_{n,n}$ are the order statistics associated to $E_1, E_2, \ldots E_n$, then

$$(E_{1,n}, E_{2,n} - E_{1,n}, \ldots, E_{n,n} - E_{n-1,n}) \overset{d}{=} \left(\frac{E_n}{n}, \frac{E_{n-1}}{n-1}, \ldots, \frac{E_n}{1} \right).$$

Lemma 2.1.1 *(Kratz and Resnick, 1996)*
If $\{E_j, j \ge 1\}$ are iid unit exponentially distributed random variables, then

$$\frac{1}{k} \sum_{j=1}^{k} E_j \log j = o_p(1) + \left(-1 + \log k + \frac{\log \sqrt{2\pi} + (\log k)/2}{k} \right).$$

To investigate now the limit distribution of the qq-estimator $\widehat{\alpha^{-1}}$ that can also be written as

$$\widehat{\alpha^{-1}} = \frac{\sum_{i=1}^{k} -\log\left(\frac{i}{k+1}\right) \left(k \log\left(Z_{n-i+1,\,n}\right) - \sum_{j=1}^{k} \log\left(Z_{n-j+1,\,n}\right) \right)}{k \sum_{i=1}^{k} \left(-\log\left(\frac{i}{k+1}\right)\right)^2 - \left(\sum_{i=1}^{k} -\log\left(\frac{i}{k+1}\right)\right)^2},$$

we need to impose a second order regular variation condition and a condition which restricts the growth of $k = k(n)$, namely :

Let $U(t) = \left(\dfrac{1}{1-F} \right)^{\overleftarrow{}} (t)$, $t > 0$ and set $\gamma = \alpha^{-1}$.

Suppose U is $2RV(\gamma, \rho)$ (see (1.4)) with $\rho \le 0$, i.e. there exists a positive function $\lim_{t\to\infty} A(t) = 0$
such that for all $x > 1$

$$\frac{\frac{U(tx)}{U(t)} - x^{\gamma}}{A(t)} \quad \to \quad c x^{\gamma} \left(\frac{x^{\rho} - 1}{\rho} \right) \quad \text{as} \quad t \to \infty, \tag{1.9}$$

for some $c \in \mathbb{R}^*$. If $\rho = 0$, interpret $(x^{\rho} - 1)/\rho$ as $\log x$.
The second additional condition, commonly used in the literature, restricts the growth of $k = k(n)$ as

$$k \to \infty, \qquad k/n \to 0, \qquad \sqrt{k}\, A(n/k) \to 0. \tag{1.10}$$

Note this condition depends on the underlying (unknown) distribution F since the function A depends on F.

Theorem 2.1.2 *(Kratz and Resnick, 1996)*
If (1.9) and (1.10) hold, then

$$\sqrt{k} \left(\widehat{\alpha^{-1}} - \alpha^{-1} \right) \overset{d}{\to} \mathcal{N} \left(0, 2\alpha^{-2} \right).$$

For the proof, we use Potter's inequalities, de Haan's method ([16]) and some results and method of Csörgő (M. and S.), Horváth and Mason ([5]) and of Csörgő, Deheuvels and Mason ([6]).

Remarks :
The asymptotic variance of $\sqrt{k}\left(\widehat{\alpha^{-1}} - \alpha^{-1}\right)$ is thus $2\alpha^{-2}$. In contrast, the Hill estimator $\widehat{\alpha^{-1}}_H = \widehat{\alpha^{-1}}_H(k,n)$ satisfies under (1.9) and (1.10)

$$\sqrt{k}\left(\widehat{\alpha^{-1}}_H - \alpha^{-1}\right) \xrightarrow{d} \mathcal{N}\left(0,\, \alpha^{-2}\right)$$

and hence has an asymptotic variance of α^{-2}. However, the Hill estimator exhibits considerable bias in certain circumstances and thus asymptotic variance is not a good criterion for superiority. Some examples are given in Kratz and Resnick (see [21], section 4) with a comparison between the plot of $\{(k,(\widehat{\alpha^{-1}})^{-1}(k)), 1 \le k \le n\}$ and the corresponding Hill plot $\{(k, \widehat{\alpha^{-1}}_H^{-1}(k,n)), 1 \le k \le n\}$: the qq–plot might seem to be a bit less volatile than the Hill plot. Moreover one of the advantages of qq–plotting over the Hill estimator is that the residuals contain information which potentially can be utilized to combat the bias in the estimates when the tail is not Pareto.

Conclusion

Independently of Kratz and Resnick, Schultze and Steinebach ([26], 1996) and Csörgő and Viharos ([4], 1997) proposed and studied, from ordinary least-squares considerations, some estimators of the tail index α^{-1}, out of which what we called the *qq*-estimator. Schultze and Steinebach proved its weak consistency under the conditions of our theorem 2.1.1 and the extra growth condition that $k_n/\log^2 n$, wheras its asymptotic normality was considered by Csörgő and Viharos ([4]), without forcing the centering at α^{-1}, which allows to deduce the theorem 2.1.2 ; considering one of the form of the *qq*-estimator given by

$$\widetilde{\alpha_n^{-1}}(k_n) := \frac{\frac{1}{k_n}\sum_{j=1}^{k_n}\log\left(\frac{k_n}{j}\right)\log Z_{n+1-j,n} - \frac{1}{k_n^2}\left(\sum_{j=1}^{k_n}\log Z_{n+1-j,n}\right)\left(\sum_{j=1}^{k_n}\log\left(\frac{k_n}{j}\right)\right)}{\frac{1}{k_n}\sum_{j=1}^{k_n}\log^2\left(\frac{k_n}{j}\right) - \left(\frac{1}{k_n}\sum_{j=1}^{k_n}\log\left(\frac{k_n}{j}\right)\right)^2},$$

they got the following result :

Theorem 2.1.3 *(Csörgő and Viharos, 1997)*
Provided $\dfrac{k_n}{\log^4 n} \underset{n\to\infty}{\to} \infty$, *we have*

$$\sqrt{k_n}\left(\widetilde{\alpha_n^{-1}}(k_n) - \alpha^{-1}\right) \xrightarrow{d} \mathcal{N}\left(0,\, 2\alpha^{-2}\right) \quad \text{if and only if}$$

$$\sqrt{k_n}\left(\int_0^1 (-1 - \log t)\log F^{\leftarrow}\left(1 - \frac{k_n t}{n}\right)dt \;-\; \alpha^{-1}\right) \underset{n\to\infty}{\to} 0.$$

Generalizing the least-squares procedure, Viharos ([27], 1999) extended this study on the estimation of the tail index. Indeed, whereas S. Csörgő, Deheuvels and Mason ([6]) proposed in 1985 a whole class of kernel estimators which contains the Hill's estimator defined in (1.5), Viharos ([27]) proposed in 1999 a whole class of weighted least-squares estimators containing an asymptotically equivalent form of the *qq*-estimator. So arised the problem not only of the choice of k_n but also of the kernel or the weight functions. This problem has been discussed a

year later by the author and Csörgő (see [3]). Notice that they chose to call the qq-estimator, the Zipf estimator, since Zipf used the least-squares estimator from the late 1920's. In fact, it should be called the Pareto-Levy estimator (see [2]).

As already said, many more methods have been proposed to construct estimators , some avoiding order statistics as for instance the method introduced by Gawronski and Stadtmüller (see [14]), based on the empirical Laplace transform, to introduce a naive estimator ; another example is the very simple method of linear fitting proposed by Barbut (see [2]) to come up with another naive estimator. And we could go on with other examples (see [3])...

2.1.2   Heavy tailed time series

For the type of data we consider (network traffic data), for which an infinite variance phenomenon or more generally heavy tailed distributions can be observed, there are mainly two approaches for the modeling, one using structural models, the other time series.

When choosing the traditional time series approach in the heavy tailed context, it was natural first to experiment linear models.

Recall that classical time series analysis concerns stationary processes, in particular linear processes defined by

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \quad t \in \mathbb{Z}, \quad \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty,$$

where $(Z_t, t \in \mathbb{Z})$ is a white noise with finite variance.

Dealing with heavy-tailed linear processes means dropping the condition of finite variance of the process $(Z_t, t \in \mathbb{Z})$ and supposing that this process has a symmetric α-stable distribution with characteristic function $\mathbb{E}[e^{isZ}] = e^{-c|s|^\alpha}$, $s \in \mathbb{R}$, $\alpha \in (0,2)$, and $c > 0$.

In fact, many statistical methods of classical time series analysis can be adapted to the heavy tailed case. It has been the subject of a certain number of papers in the literature of the past decade, theoretically, with the study of estimators for coefficients in heavy tailed ARMA models, as well as in practice, with software tools for those models.

The work presented in the following section belongs to this axis of theoretical investigation on heavy tailed linear models, with the estimation of the coefficients of heavy tailed ARMA processes. Different classes of estimators have been proposed, such that Yule-Walker estimators, spectral density estimators, least gamma deviation estimators, for the autoregressive case, Hsing estimator based on extreme value considerations for the MA case, linear programming estimators and Whittle estimators for the general case (see [22], §4 and [8], §7.5 for references).

While Feigin and Resnick ([10],[11]) studied linear programming estimators (lp-estimators) for autoregressive coefficients and proved that they have a good rate of convergence which is frequently superior to those achieved by Yule Walker or maximum likelihood estimators, we will be concerned by the moving average case, necessary step to being able to estimate parameters in more general ARMA processes which combine both autoregressive and moving average components.

Parameter Estimation for Moving Averages with Positive Innovations

The process under consideration is the finite order moving average of order q, denoted MA(q) and specified as follows.

Let $\{Z_t\}$ be an iid sequence of non-negative random variables. For a positive integer $q \geq 1$, suppose we have parameters $\theta_1, \ldots, \theta_q$ such that $\theta_i \geq 0$ for $1 \leq i \leq q$. The MA(q) process $\{X_t\}$ is

$$X_t = Z_t + \sum_{i=1}^{q} \theta_i Z_{t-i} := \Theta(B) Z_t, \quad t \in \mathbb{Z}, \tag{1.11}$$

where Θ is the moving average polynomial defined by $\Theta(z) = \sum_{i=0}^{q} \theta_i z^i$, $\theta_0 = 1$, and B is the backward shift operator, and we are interested in estimating $\theta_1, \ldots, \theta_q$.

For a pure autoregressive process of order p, denoted by AR(p), with positive innovations $\{Z_t\}$, and with autoregressive coefficients $\phi_1, \ldots, \phi_p$, $(\phi_p \neq 0, \sum_{i=1}^{p} \phi_i < 1)$, of the form

$$X_t = \sum_{k=1}^{p} \phi_k X_{t-k} + Z_t \; ; \; t \in \mathbb{Z} \tag{1.12}$$

Feigin and Resnick (1994) defined the linear programming estimators $\hat{\phi}$ based on observing $X_1, \ldots, X_n$ as

$$\hat{\phi} = \arg \max_{\delta \in D_n} \delta' \mathbf{1} \tag{1.13}$$

where $\mathbf{1}' = (1, \ldots, 1)$ and where the feasible region D_n is defined as

$$D_n = \{\delta \in \mathbb{R}^p : X_t - \sum_{i=1}^{p} \delta_i X_{t-i} \geq 0, t = p+1, \ldots, n\}. \tag{1.14}$$

Assuming regular variation conditions on either the left or right tails of the innovations was sufficient to show that a limit distribution existed for $\hat{\phi}$ and that rates of convergence were often superior to the Yule–Walker estimators. So a natural approach to the estimation problem for moving averages is to see what results from the autoregressive case can be brought to bear and thus we assume the moving average in (1.11) is invertible, i.e. that $\Theta(z) \neq 0$, $|z| \leq 1$. This allows us to write

$$\Pi(z) := \frac{1}{\Theta(z)} = \sum_{i=0}^{\infty} \pi_k z^k, \quad |z| \leq 1,$$

and to convert (1.11) into an infinite order autoregression

$$\Pi(B) X_t = Z_t, \quad t = 0, \pm 1, \pm 2, \ldots.$$

If we now try to apply the lp estimators we find we have a nice objective function but the constraints involve an infinite number of variables. If we truncate the constraints suitably, we should obtain an estimator with worthwhile properties. The precise definition of our estimator of the moving average coefficients in the MA(q) process is

$$\hat{\theta} := \arg \max_{D_n} \sum_{i=1}^{q} \theta_i \tag{1.15}$$

where

$$D_n := \{\boldsymbol{\theta} : \sum_{i=0}^{2l} (I - \Theta(B))^i X_t \geq 0, t = 2lq + 1, \ldots, n\} \tag{1.16}$$

and l is the first integer such that $2l \geq q$.

We need conditions which specify the model. In order to obtain a limit distribution for our estimators, we impose regular variation and moment conditions on the distribution of the innovation sequence.

Condition M (model specification) : The process $\{X_t : t \in \mathbb{Z}\}$ satisfies the equations (1.11) where $\{Z_t\}$ is an independent and identically distributed sequence of random variables with essential infimum (left endpoint) equal to 0 and common distribution function F. The coefficients $\theta_1, \ldots, \theta_q$ satisfy the invertibility condition that the moving average polynomial $\Theta(z) = \sum_{i=0}^{q} \theta_i z^i$ has no root in the unit disk $\{z : |z| \leq 1\}$.

Condition L (left tail) : The distribution F of the innovations Z_t satisfies, for some $\alpha > 0$:

$$\lim_{s \downarrow 0} \frac{F(sx)}{F(s)} = x^\alpha \text{ for all } x > 0; \tag{1.17}$$

$$E(Z_t^\beta) = \int_0^\infty u^\beta F(du) < \infty \text{ for some } \beta > \alpha. \tag{1.18}$$

Condition R (right tail) : The distribution F of the innovations Z_t satisfies, for some $\alpha > 0$:

$$\lim_{s \to \infty} \frac{1 - F(sx)}{1 - F(s)} = x^{-\alpha} \text{ for all } x > 0; \tag{1.19}$$

$$E(Z_t^{-\beta}) = \int_0^\infty u^{-\beta} F(du) < \infty \text{ for some } \beta > \alpha. \tag{1.20}$$

Remarks : Condition L is rather mild. It is satisfied if a density f of F exists which is continuous at 0 and with $f(0) > 0$. In this case $\alpha = 1$. Other common cases where Condition L holds are the Weibull distributions of the form $F(x) = 1 - \exp\{-x^\alpha\}$ where $F(x) \sim x^\alpha$, as $x \downarrow 0$ and the gamma densities $f(x) = ce^{-x}x^{r-1}$, $r > 0, x > 0$ so that $f(x) \sim cx^{r-1}$ as $x \downarrow 0$ and therefore the associated Gamma distribution function satisfies $F(x) \sim cr^{-1}x^r$, as $x \downarrow 0$. Examples of distributions satisfying condition R include positive stable densities and the Pareto density.

Suppose the true value of the moving average coefficients is $\boldsymbol{\theta}^{(0)}$. In inverted form, the model can be written as the AR(∞) process

$$\Pi(B)X_t = Z_t, \quad t \in \mathbb{Z} \quad \text{where} \quad \frac{1}{\Theta(z)} = \Pi(z), \quad |z| \leq 1. \tag{1.21}$$

For a finite order autoregression (1.12), the linear programming estimator of autoregressive coefficients is given by (1.13) and (1.14). If in (1.12) we write as usual the autoregressive polynomial as

$$\Phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i,$$

then the objective function in (1.13) can be written as $1 - \Phi(1)$ and the constraints in (1.14) can be expressed as

$$\Phi(B)X_t \geq 0, \tau = p+1, \ldots, n. \tag{1.22}$$

If we try to write down an analogous expression for the parameter estimators for the AR(∞) process in (1.21), we obtain as objective function

$$1 - \frac{1}{\Theta(1)} = \frac{\sum_{i=1}^q \theta_i}{1 + \sum_{i=1}^q \theta_i}$$

which is monotone in $\sum_{i=1}^q \theta_i$. So we try to maximize $\sum_{i=1}^q \theta_i$. For the constraints, (1.22) suggests the set of conditions

$$\Pi(B)X_t \geq 0, \quad t = 1, \ldots, n.$$

A problem arises in that this constraint set requires knowledge of $X_t, X_{t-1}, \ldots$ with the index extending back to $-\infty$ and since we only have knowledge of $X_1, \ldots, X_n$ we must somehow truncate this constraint set.

A suggestion for how to construct a truncated set of constraints comes from symbolically expanding $1/\Theta$:

$$\frac{1}{\Theta(B)} = \frac{1}{I - (I - \Theta(B))} = \sum_{k=0}^{\infty}(I - \Theta(B))^k \approx \sum_{k=0}^{2l}(I - \Theta(B))^k,$$

where $l \geq 1$ is an integer to be specified. Note that

$$\sum_{k=0}^{2l}(I - \Theta(B))^k\Theta(B) = I - (I - \Theta(B))^{2l+1} = I + Q(B)^{2l+1},$$

where $Q(B) = \Theta(B) - I = \sum_{i=1}^q \theta_i B^i$. Let $\Theta^{(0)}(B) = \sum_{i=0}^q \theta_i^{(0)} B^i$ and $Q_0(B) = \sum_{i=1}^q \theta_i^{(0)} B^i$ and thus

$$\sum_{k=0}^{2l}(I - \Theta^{(0)}(B))^k X_t = \sum_{k=0}^{2l}(I - \Theta^{(0)}(B))^k \Theta^{(0)}(B)Z_t = (I + Q_0^{2l+1}(B))Z_t \geq 0,$$

since all $\theta_i^{(0)}$'s are assumed non-negative. So by truncating the series expansion for $1/\Theta$ in a judicious manner, the truncated expansion is always positive at the correct value of the parameter vector. Thus we have :

Proposition 2.1.1 *(Feigin, Kratz and Resnick, 1996)*
Assume that the invertible model $MA(q)$ X is specified by Condition M and that the true value of the moving average coefficients is $\boldsymbol{\theta}^{(0)}$. Then our estimator is

$$\hat{\boldsymbol{\theta}} = arg \ \max_{D_n} \sum_{i=1}^q \eta_i \tag{1.23}$$

where the constraint set is

$$D_n = \{\boldsymbol{\eta} \in \mathbb{R}_+^q : \sum_{k=0}^{2l}(I - \sum_{i=0}^q \eta_i B^i)^k X_t \geq 0, t = 2lq+1, \ldots, n; \ \eta_0 = 1, \sum_{i=0}^q \eta_i z^i \neq 0, \ |z| \leq 1\}. \tag{1.24}$$

The choice of l suggested by the limit theory is to choose l to be the first integer such that $2l \geq q$.

We seek a limit distribution for $q_n(\hat{\boldsymbol\theta} - \boldsymbol\theta^{(0)})$ where q_n is an appropriate scaling satisfying $q_n \to \infty$. It turns out that

Proposition 2.1.2 *(Feigin, Kratz and Resnick, 1996)*

- *Under Condition R, the right choice of q_n is $q_n = b_n = F^{\leftarrow}\left(1 - \dfrac{1}{n}\right) = \left(\dfrac{1}{1 - F}\right)^{\leftarrow}(n)$;*

- *under Condition L, the appropriate choice of q_n is $q_n = a_n = F^{\leftarrow}\left(\dfrac{1}{n}\right)$.*

Thus

$$q_n(\hat{\theta} - \theta^{(0)}) = arg\ \max_{\Lambda_n} \delta' \mathbf{1}$$

where

$$\Lambda_n = \{\delta \in \mathbb{R}_+^q : 1 + \sum_{i=1}^q \left(\frac{\delta_i}{q_n} + \theta_i^{(0)}\right) z^i \neq 0,\ |z| \leq 1,$$

$$\sum_{k=0}^{2l} (-1)^k \left(Q_0(B) + \frac{\delta(B)}{q_n}\right)^k X_t \geq 0, t = 2lq + 1, \ldots, n\}.$$

Indeed, we observe that $q_n(\hat{\boldsymbol\theta} - \boldsymbol\theta^{(0)})$ satisfies

$$q_n(\hat{\boldsymbol\theta} - \boldsymbol\theta^{(0)})'\mathbf{1} \geq q_n(\boldsymbol\eta - \boldsymbol\theta^{(0)})'\mathbf{1}$$

for all $\boldsymbol\eta \in D_n$. Let $\boldsymbol\delta = q_n(\eta - \theta^{(0)})$ so that $q_n^{-1}\delta + \theta^{(0)} = \eta$. Then $q_n(\hat{\boldsymbol\theta} - \boldsymbol\theta^{(0)})$ satisfies $q_n(\hat{\boldsymbol\theta} - \boldsymbol\theta^{(0)})'\mathbf{1} \geq \delta'\mathbf{1}$ for all δ such that

$$1 + \sum_{i=1}^q \left(\frac{\delta_i}{q_n} + \theta_i^{(0)}\right) z^i \neq 0, \quad |z| \leq 1$$

and

$$\sum_{k=0}^{2l} (-1)^k \left(Q_0(B) + \frac{\delta(B)}{q_n}\right)^k X_t \geq 0$$

for $t = 2lq + 1, \ldots, n$, where $\delta(B) = \sum_{i=1}^q \delta_i B^i$, hence the result.

In the case $\mathbf{q = 1}$, i.e. when $l = 1$ and $\Theta(B) = I + \theta B$, the estimator is

$$\hat{\theta} = \bigwedge_{t=3}^n \sup\{\eta \in [0, 1) : X_t - \eta X_{t-1} + \eta^2 X_{t-2} \geq 0\},$$

which gives

$$q_n(\hat{\boldsymbol\theta} - \boldsymbol\theta^{(0)}) = arg\ \max_{\Lambda_n} \delta = \sup\{\delta \geq 0 : \delta \in \Lambda_n\}$$

where $\quad \Lambda_n = \left\{\delta \geq 0 : 1 + \left(\dfrac{\delta}{q_n} + \theta^{(0)}\right) z \neq 0,\ |z| \leq 1,\right.$

$$\left. X_t - \left(\theta^{(0)} + \frac{\delta}{q_n}\right) X_{t-1} + \left(\theta_0 + \frac{\delta}{q_n}\right)^2 X_{t-2} \geq 0,\ t = 3, \ldots, n\right\}.$$

We can also write

$$q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)}) = q_n \bigwedge_{t=3}^{n} \sup\{0 \leq \eta < 1 - \theta^{(0)} : A_t\eta^2 + B_t\eta + C_t \geq 0\}, \qquad (1.25)$$

where

$$A_t = Z_{t-2} + \theta^{(0)}Z_{t-3}, \ B_t = -Z_{t-1} + \theta^{(0)}Z_{t-2} + 2(\theta^{(0)})^2 Z_{t-3}, \ C_t = Z_t + (\theta^{(0)})^3 Z_{t-3}.$$

So the limit distribution depends on the behavior of random parabolas and from extreme value theory we expect the limit distribution to be in the Weibull family.
The analyze of the limit distribution in (1.25) comes back to the study of the random parabola $p_t(\eta) = A_t\eta^2 + B_t\eta + C_t$.

Theorem 2.1.4 *(Feigin, Kratz and Resnick, 1996)*
Suppose $\{X_t\}$ is the MA(1) process given in (1.11) and that Conditions M, R hold. Suppose the true parameter is $\theta^{(0)} \in (0,1)$ and that F is continuous. Let $q_n = b_n$ be the quantile function

$$b_n = \left(\frac{1}{1-F}\right)^{\leftarrow}(n) = F^{\leftarrow}(1 - \frac{1}{n})$$

where F is the distribution of Z_1. The estimator $\hat{\theta}$ given in (1.23) has a Weibull limit distribution : In $[0, \infty)$

$$b_n(\hat{\theta} - \theta^{(0)}) \Rightarrow \bigwedge_{k=1}^{\infty} \Gamma_k^{1/\alpha}(Y_k + (\theta^{(0)})^3 Y_k'), \qquad (1.26)$$

where $\{Y_k, Y_k', k \geq 1\}$ are iid with common distribution F and

$$\Gamma_k = E_1 + \cdots + E_k, \quad k \geq 1,$$

is a sum of iid unit exponentially distributed random variables independent of $\{(Y_k, Y_k')\}$. The limit distribution of $\hat{\theta}$ is Weibull :

$$\lim_{n\to\infty} P[b_n(\hat{\theta} - \theta^{(0)}) \leq x] = 1 - \exp\{-cx^\alpha\}, \quad x > 0, \qquad (1.27)$$

where

$$c = E|Y_k + (\theta^{(0)})^3 Y_k'|^{-\alpha}$$

which is finite by the second statement of Condition R.

Theorem 2.1.5 *(Feigin, Kratz and Resnick, 1996)*
Suppose $\{X_t\}$ is the MA(1) process given in (1.11) and that Conditions M, L hold. Suppose the true parameter is $\theta^{(0)} \in (0,1)$ and that F, the distribution of Z_1, is continuous. Let $q_n = a(n)^{-1}$ where $a(n)$ is the quantile function

$$a(n) = F^{\leftarrow}(1/n).$$

Note $a(n) \to 0$. The estimator $\hat{\theta}$ given in (1.23) has a Weibull limit distribution : In $[0, \infty)$

$$a(\sqrt{n})^{-1}(\hat{\theta} - \theta^{(0)}) \Rightarrow \frac{(\theta^{(0)})^{3/2\alpha}}{c(\alpha)^{1/2\alpha}} \bigwedge_{\substack{1 \leq k < \infty \\ Y_{k,1} > Y_{k,2}}} \frac{\Gamma_k^{1/2\alpha}}{|Y_{k,1} - (\theta^{(0)})^3 Y_{k,}|}, \tag{1.28}$$

where $\{Y_{k,1}, Y_{k,2}, k \geq 1\}$ are iid with common distribution F and

$$\Gamma_k = E_1 + \cdots + E_k, \quad k \geq 1,$$

is a sum of iid unit exponentially distributed random variables. The constant $c(\alpha)$ is defined by the Beta integral

$$c(\alpha) = \int_0^1 (1 - s)^\alpha \alpha s^{\alpha - 1} ds.$$

The limit distribution of $\hat{\theta}$ is Weibull :

$$\lim_{n \to \infty} P[a(\sqrt{n})^{-1}(\hat{\theta} - \theta^{(0)}) \leq x] = 1 - \exp\{-kx^{2\alpha}\}, \quad x > 0, \tag{1.29}$$

where

$$k = (\theta^{(0)})^{-3\alpha} c(\alpha) E\left(|Y_{k,1} - (\theta^{(0)})^3 Y_{k,2}|^{2\alpha} 1_{[Y_{k,1} > Y_{k,2}]}\right)$$

which is finite by Condition L. The convergence rate is $1/a(\sqrt{n})$.

The proofs are based on point process techniques and weak convergence theory. $\square$

It is noteworthy that in contrast to the autoregressive case, the moving average estimators in the left tail case suffer a performance degradation depending on the order q of the model; no such degradation is present under condition R. From the results, we see that the convergence rate for the estimator of the MA(1) parameter is $1/a(\sqrt{n})$ which is a regularly varying function of index $\alpha/2$. Contrast this to the convergence rate of the lp estimators in the autoregressive case which is regularly varying of index α. We anticipate that the convergence rate in the left tail case for MA(q) parameters will have index $\alpha/(q+1)$. Thus under Condition L, a sharp penalty is paid for using models which have moving average components and the penalty increases as the order of the model increases. This is in contrast to results under Condition R and to the results found for lp estimators for autoregressive parameters.

The next challenge would have been to extend these results from the MA(1) case to more general moving average processes and then on to the general ARMA model, but linear models are not enough pertinent for dependent heavy tailed data (see for instance the discussion of Resnick "Why non-linearities can ruin the heavy tailed modeler's day" in [1]).

Hence, the focus has been turned to non-linear models as the ARCH models and related models (there exists more than 50 different such models!), very popular in econometrics, and which provides an abundant literature (see [8], §7.6 and §8.4 for references).

2.1.3   References

Bibliographie

[1] R. Adler, J. Feldman and M. Taqqu, *A user's guide to heavy tails : statistical techniques for analyzing heavy tailed distributions and processes*, Birkhäuser, Boston (1997).

[2] M. Barbut, Une famille de distributions : des paretiennes aux "contra-paretiennes", Math. Inf. Sci. Hum., **141**, 43-72 (1998).

[3] S. Csörgő and L. Viharos, Estimating the tail index, In *Asymptotic Methods in Probability and statistics*, Elsevier, Amsterdam (1998).

[4] S. Csörgő and L. Viharos, Asymptotic normality of least squares estimators of tail indices, *Bernoulli* **3**, 351-370 (1997).

[5] M. Csörgő, S. Csörgő, L. Horváth and D. Mason, Weighted Empirical and quantile processes, *Ann. Probability* **14**, 1-85 (1986).

[6] S. Csörgő, P. Deheuvels and D. Mason, , Kernel Estimates of the tail index of a distribution, *Ann. Statist.* **13**, 1050-1077 (1985).

[7] R. Davis, and S. Resnick, Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution, *Stoch. Proc. Applic.* **30**, 41-68 (1988).

[8] P. Embrechts, C. Kluppelberg and T. Mikosh, *Modelling extremal events*, Springer Verlag (1997).

[9] P. Feigin, M. Kratz and S. Resnick, Parameter Estimation for Moving Averages with Positive Innovations, *Ann. Appl. Probab.* **6**, 1157-1190 (1996).

[10] P. Feigin, and S. Resnick, Estimation for autoregressive processes with positive innovations, *Stoch. Mod.* **8**, 479-498 (1992).

[11] P. Feigin, and S. Resnick, Limit distributions for linear programming time series estimators, *Stoch. Proc. Applic.* **51**, 135-166 (1994).

[12] P. Feigin, and S. Resnick, Linear programming estimators and bootstrapping for heavy-tailed phenomena, *Adv. in Appl. Probab.* **29**, 759-807 (1997).

[13] P. Feigin, S. Resnick and C. Stărică, Testing for Independence in Heavy Tailed and Positive Innovation Time Series, *Comm. Statist. Stoch. Mod.* **11**, 587-612 (1995).

[14] W. Gawronski and U. Stadtmüller, Parameter estimation for distributions with regularly varying tails, *Statist. Decisions* **3**, 297-316 (1985).

[15] J. Geluk and L. de Haan, *Regular Variation, Extensions and Tauberian Theorems*, CWI Tract 40 (1987).

[16] L. de Haan, *Extreme Value Statistics*, Lecture Notes, Econometric Institute, Erasmus University, Rotterdam (1991).

[17] L. de Haan and U. Stadtmüller, Generalized regular variation of second order, *J. Austral. Math. Soc. Ser. A* **61**, 381-395 (1996).

[18] B. Hill, A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **3**, 1163-1174 (1975).

[19] B. Hill, Tail probabilities, In *Encyclopedia of Stat. Sciences* **9**, Wiley N.Y., 170-173 (1988).

[20] B. Hill, Bayesian forecasting of extreme values in an exchangeable sequence, *J. Res. Inst. Stand. Technol.* **99**, 521-538 (1994).

[21] M. Kratz and S. Resnick, The QQ-estimator and heavy tails, *Stoch. Models*, **12**, 699-724 (1996).

[22] S. Resnick, Heavy tail modeling and teletraffic data (special invited paper), *Ann. Statist.* **25**, 1805-1869 (1997).

[23] S. Resnick, Why non-linearities can ruin the heavy tailed modeler's day, In *A practical guide to heavy tails : statistical* , Springer Verlag (1997).

[24] S. Resnick and C. Stărică, Consistency of Hill's estimator for dependent data, *J. Appl. Probab.* **32**, 139-167 (1995).

[25] S. Resnick and C. Stărică, Smoothing the Hill estimator, *Adv. Appl. Probab.* **29**, 271-293 (1997).

[26] J. Schultze and J. Steinebach, On least squares estimates of an exponential tail coefficient, *Statist. Decisions* **14**, 353-372 (1996).

[27] L. Viharos, Weighted least squares estimators of tail indices, *Probab. Math. Statist.* **19**, 249-265 (1999).

2.2   On a representation of Gibbs measure for R.E.M.

2.2.1   Introduction

Let us present the general framework of the spin systems.

Historical framework

• In 1907, following P. Curie's discovery of the critical temperature for ferromagnetic order a few years before, P. Weiss came up with the idea of a system of 'spins' or elementary magnetic moments. It is what constitutes the origin of the Curie-Weiss or mean-field model for ferromagnetism.
The ferromagnetism phenomena is due to the alignment of the spins of the (iron) atoms;
⋆ this phenomena depends on the heat temperature;
⋆ the ferromagnetic interaction becomes attractive between sufficiently closed spins.
• The first simple model of the ferromagnetic system has been invented and setted up with the formalism of statistical mechanics by Lenz, then analized in dimension one by his student Ising in 1923, hence the name of Ising's model.
This model consists in placing the atoms on the sites of a regular lattice of Z^d, represented by the spin variables taking only the values ± 1. Spins would interact only when being at neighbouring sites and in the sense of an identity of signs. It can also exist an external magnetic field favouring the orientation of the spins towards the plus or the minus sign.
All those interactions are represented by an energy function or 'Hamiltonian' function : to a given spin-configuration σ corresponds an Hamiltonian $H(\sigma)$.

Notations and definitions

• The basic axiom of statistical mechanics is that the equilibrium properties of a system are described by specifying a probability measure on the space of configurations, $\{-1, 1\}^{\mathbb{Z}^d}$ in the case of Ising, or $\{-1, 1\}^{\Lambda}$ in the case of a finite set $\Lambda \subset \mathbb{Z}^d$.
The probability measure chosen in our context is the **Gibbs measure**, defined formally by

$$\mu_\beta(d\sigma) = \frac{1}{Z_\beta} e^{-\beta H(\sigma)} \rho(d\sigma),$$

where $\beta = 1/T$, Z_β is a constant of normalization and ρ is the uniform measure on the configuration space.
• Let us now define the general framework for the study of spin systems with regular interactions.
The general lattice will be $\mathbb{Z}^d$ and Λ will denote a finite subset of $\mathbb{Z}^d$.

  We define $\sigma :$ $\mathbb{Z} \to \{-1, 1\} := S_0$
                $i \to \sigma_i =$ spin at site i.
On S_0, we take the discrete topology, and so $(S_0, \mathcal{F}_0)$ denotes a measure space.
On $\mathbb{Z}^d$, we consider the state space $S = S_0^{\mathbb{Z}^d}$, complete separable space when equipped with the product topology; $\mathcal{F} = \mathcal{F}_0^{\mathbb{Z}^d}$ denotes the product σ-algebra.

We shall use the same notation for any subset $\Lambda \subset \mathbb{Z}^d : (S^\Lambda = S_0^\Lambda, \mathcal{F}_\Lambda = \mathcal{F}_0^\Lambda)$.

• Let us come back now to simple models as the so called mean-field models, which prototype is the Curie-Weiss model, trivialization of the Ising model, as seen previously. We will classify those models into two classes :
⋆ the Gaussian processes
⋆ the models of Hopfield (1982), which take in account the fact that the network passed through a certain number M of states (for the interaction variables) (notion of associative memory).

Let us define the notion of disordered **mean-field** model.
This model is characterized by the fact that the spatial structure of the lattice $\mathbb{Z}^d$ is given up and replaced by an hypercube of finite dimension N, where the sites are indexed by the natural numbers and where all spins interact with each other independently of the distance between them.

To study more easily those models, a point of view consists in considering the **Hamiltonian as a Gaussian process**. In such a class of models figure :
a) *Derrida systems* (1980's), namely :
⋆ the Random Energy Model (*R.E.M.*), which Hamiltonian is given by

$$H_{rem}(\sigma) = -\frac{N^{1/2}}{2^{N/2}} \sum_{\alpha \subset \{1,\dots,N\}} J_\alpha \sigma_\alpha,$$

where the sum is over the 2^N subsets α of$\{1,\dots,N\}$, $\{J_\alpha, \alpha \subset \{1,\dots,N\}\}$ is a family of standard Gaussian r.v. iid (defined on a same probability space $(\Omega, \mathcal{F}, I\!P)$) and $\sigma_\alpha = \prod_{i \in \alpha} \sigma_i$
(with $\sigma_\emptyset = 1$) ;
it is the most disordered model because of independent interactions ;
⋆ the Generalized REM (*G.R.E.M.*) : the hypothesis of independence of the J in the REM is replaced by the assumption of positive correlations between the r.v.
b) *Sherrington-Kirkpatrick system* (1976), which Hamiltonian is given by

$$H_{SK}(\sigma) = -N^{-1/2} \sum_{i,j \in \{1,\dots,N\}} J_{ij} \sigma_i \sigma_j,$$

where the sum is over all the pairs of distinct sites and $\{J_{ij}, i, j\}$ is a family of standard Gaussian r.v. iid.
Note that we used the Hamiltonian defined for a configuration $\sigma \in \{-1, +1\}^N$ of spins ± 1 in a volume $\{1, \dots, N\}$, to characterize those different models.

In what follows, we will consider the simplest system, i.e. the R.E.M., caricature of the S.K. model.

2.2.2 The Random Energy Model

In the case of the REM, we consider an hypercube of dimension N;
$\Lambda = \Lambda_N = \{1, 2, \cdots, N\} \subset I\!N$ denotes a subset of $I\!N$, and $S_N = S_\Lambda = \{-1, 1\}^N$ the state space.
To the function $\quad \sigma : \quad \Lambda \to \{-1, 1\}$
$$\qquad\qquad\qquad i \;\to\; \sigma_i$$
we associate the Hamiltonian $H_N(\sigma) = -\dfrac{\sqrt{N}}{2^{N/2}} \sum_{\alpha \subset \Lambda} J_\alpha \sigma_\alpha$.

The random coefficients J are defined on a probability space $(\Omega, \mathcal{F}, I\!P)$.
Note that N being fixed, $(H_{rem}(\sigma))_\sigma$ is a family of 2^N r.v. i.i.d. $\mathcal{N}(0, N)$ defined on $(\Omega, \mathcal{F}, I\!P)$,
i.e. $(H_N(\sigma))$ is of the form $(-\sqrt{N} X_\sigma)$, where (X_σ) are iid $\mathcal{N}(0, 1)$ (see [9]).

Let us give some characteristics of this model.
Given $\beta \geq 0$, the inverse temperature, let us denote by

$$Z_N \equiv Z_N(\beta) = \sum_{\sigma \in \{-1, +1\}^N} e^{-\beta H_{rem}(\sigma)} \qquad\qquad (2.30)$$

the finite volume partition function and by

$$F_N(\beta) = -\frac{1}{\beta N} \log Z_N(\beta) \qquad\qquad (2.31)$$

the finite volume free energy.
The asymptotic behavior of F_N has been quite studied, in particular we know that $\forall \beta \geq 0$,
$\lim_{N \to \infty} F_N(\beta) = F(\beta)$ exists $I\!P$-a.s. and in $L^p(\Omega, I\!P)$ for $1 \leq p < \infty$, where $F(\beta)$ is a non random
function twice differentiable in β, whose second derivative has a jump at $\beta_c := \sqrt{2 \log 2}$. More
precisely, we have

$$F(\beta) = \begin{cases} -\beta/2 - \beta_c^2/(2\beta) & \text{si } \beta < \beta_c \\ -\beta_c & \text{si } \beta \geq \beta_c. \end{cases}$$

What is also important to know is that when $\beta > \beta_c$, the main contribution to $F_N(\beta)$ comes
from the terms in $Z_N(\beta)$ that have the lowest possible energies, that are the σ whose $H_{rem}(\sigma)$
are of order $-N\sqrt{2 \log 2} \equiv -N\beta_c$. In the physic literature, it is said that such a system is
"frozen" in the sense that in the whole range of $\beta \geq \beta_c$, the lowest possible energies give the
main contribution to the free energy, while in a "non frozen" system, this occurs only at zero
temperature. This can be seen on $F(\beta)$ since when $\beta > \beta_c$, the free energy does not depend on
β.
Note that another way of studying the fluctuations of the free energy $F_N(\beta)$, is to consider the
number of random variables of the sample that are below some well chosen non random energy,
and so to come down to the classical convergence to a Poisson Point Process (see Ruelle (87,
[19]), Galves, Martinez et Picco (89, [10]), Bovier, Kurkova et Loewe (02, [2])).

The aim of this work was to characterize (physically) the finite volume Gibbs measure $\mu_{N,\beta}$ on
$(S_N, \mathcal{F}_N)$ when $\beta > \beta_c = \sqrt{2 \log 2}$ and N is large enough; so we studied its support.

Since $\mu_{N,\beta}$ is defined for each N as the random probability measure on $\{-1, +1\}^N$ which gives to the configuration σ the weight

$$\mu_{N,\beta}(\sigma) \equiv \frac{e^{-\beta H_{rem}(\sigma)}}{Z_N}, \tag{2.32}$$

the main question was, when $\beta > \beta_c$, for a given sample of $(H_{rem}(\sigma), \sigma \in \{-1, +1\}^N)$, what are the sample dependent configurations σ where the Gibbs measure is concentrated, when N is very large ?

Note first that the sample dependent configuration $\sigma^{(1)}$ which corresponds to the minimal value of the sample $H_{rem}(\sigma)$ is clearly among these configurations. So instead of considering the variables that are below a non random energy as it was done to get the Poisson Point Process mentioned above, it is better to consider the variables that are above the minimal one that is a sample dependent energy. It is clear that the order statistics of the random variables $H_{rem}(\sigma)$, namely

$$H_{rem}(\sigma^{(2^N)}) \geq H_{rem}(\sigma^{(2^N-1)})) \cdots \geq H_{rem}(\sigma^{(2)}) \geq H_{rem}(\sigma^{(1)}) \tag{2.33}$$

come into play. To have information on the support of the Gibbs measure, we can subtract from all the Hamiltonians the minimal energy, therefore the spacings $H_{rem}(\sigma^{(k+1)}) - H_{rem}(\sigma^{(k)})$ and the sums of spacings $H_{rem}(\sigma^{(k)}) - H_{rem}(\sigma^{(1)})$ are the basic objects under study.

We were then faced to different problems :

1) How many successive terms $k = k_N$ in the sum of spacings do we need to take to have, as $N \to \infty$, a "good" (in the sense, with probability 1) approximation of the Gibbs measure, when considering the probability measure on $\{-1, +1\}^N$ that have only these k terms ?

2) Does the total variation distance d_{tv} between the Gibbs measure and its approximation (based only on $k = k_N$ terms) tend to 0 a.s. as $N \to \infty$? (Recall that the total variation distance between two measures is given by $d_{TV}(\mu, \nu) = \sup_{\{\Psi : ||\Psi||_\infty = 1\}} |\mu(\Psi) - \nu(\Psi)|$).

3) When considering this approximation, what does represent this new measure ? Is it the uniform measure on k_N points chosen without replacement within 2^N points ? a point mass at the minimum ? or... ?

4) Does the Gibbs measure depend on β, even if the system is frozen ? For instance, can we estimate the distance d_{tv} between the Gibbs measure for finite $\beta > \beta_c$ and the limit measure as $\beta \to \infty$?

5) Finally, is there an easy way and no costly, to construct this approximated Gibbs measure ?

Let $\Psi : S_N = \{-1, +1\}^N \to I\!\!R$, $||\Psi||_\infty := \sup_{\sigma \in S_N} |\Psi(\sigma)| < \infty$. We have

$$\mu_{N,\beta}(\Psi) = \sum_\sigma \Psi(\sigma) \frac{e^{-\beta H_N(\sigma)}}{Z_{\beta,N}} = \frac{\sum_{i=1}^{2^N} \Psi((\sigma)_i) e^{-\beta \sqrt{N} X_i}}{\sum_{i=1}^{2^N} e^{-\beta \sqrt{N} X_i}},$$

where $(X_i)_i$ are iid standard Gaussian r.v. (associated to $(\sigma)_i$).
Let $(\tilde{\sigma})_j = (\sigma)_i$ to which $X_i = X_{j,2^N}$ is associated. Then we have

$$\mu_{N,\beta}(\Psi) = \frac{\sum_{j=1}^{2^N} \Psi((\tilde{\sigma})_j) e^{-\frac{\beta}{\beta_c} \sqrt{2 \log 2^N} X_{j,2^N}}}{\sum_{j=1}^{2^N} e^{-\frac{\beta}{\beta_c} \sqrt{2 \log 2^N} X_{j,2^N}}}.$$

First we defined how many terms of the order statistics $X_{i,2^N}$ are needed to ensure a good approximation of the Gibbs measure $\mu_{N,\beta}$.

Theorem 2.2.1 *(Kratz and Picco, 2004)*
Let k_N satisfying

$$k_N \uparrow \infty, \quad \frac{k_N}{N} \uparrow \infty \quad and \quad \frac{\log k_N}{N} \downarrow 0, \quad as \quad N \to \infty. \tag{2.34}$$

There exists $\Omega_N \subset \Omega$ and N_0 such that for all $N > N_0$,

$$I\!P[\Omega_N] \geq 1 - \frac{4}{(N \log 2)^{1+\delta}} \tag{2.35}$$

and on Ω_N we have

$$\mu_{\mathbf{N},\beta}(\boldsymbol{\Psi}) = \frac{\sum_{\mathbf{i=1}}^{\mathbf{k_N}} \boldsymbol{\Psi}((\tilde{\sigma})_{\mathbf{i}}) \mathbf{e}^{-\frac{\beta}{\beta_{\mathbf{c}}}\sqrt{\mathbf{2 \log 2^N}}(\mathbf{X_{i,2^N}-X_{1,2^N}})} + \mathbf{B_N}(\boldsymbol{\Psi})}{\sum_{\mathbf{i=1}}^{\mathbf{k_N}} \mathbf{e}^{-\frac{\beta}{\beta_{\mathbf{c}}}\sqrt{\mathbf{2 \log 2^N}}(\mathbf{X_{i,2^N}-X_{1,2^N}})} + \mathbf{B_N}(\mathbf{1})}, \tag{2.36}$$

with $|B_N(\Psi)| \leq \dfrac{2||\Psi||_\infty}{\frac{\beta}{\beta_c} - 1} \times \dfrac{1}{(k_N - 1)^{\frac{1}{2}(\frac{\beta}{\beta_c}-1)}}.$

An example of such k_N is $k_N = N \log_p(N)$, where $\log_p = \log \log_{p-1}$.

Let define the random probability measure on $\{-1, +1\}^N$ by

$$\mu^{(1)}_{k_N,\beta}(\Psi) := \frac{\sum_{k=1}^{k_N} \Psi((\tilde{\sigma})_k) e^{-\frac{\beta}{\beta_c}\sqrt{2 \log 2^N}(X_{k,2^N}-X_{1,2^N})}}{\sum_{k=1}^{k_N} e^{-\frac{\beta}{\beta_c}\sqrt{2 \log 2^N}(X_{k,2^N}-X_{1,2^N})}}. \tag{2.37}$$

We proved, by using the total variation distance between two measures and the first Borell-Cantelli lemma, that this measure is a good approximation of the Gibbs measure :

Corollary 2.2.1 *(Kratz and Picco, 2004)*
For all β such that $\frac{\beta}{\beta_c} > 1$, for all k_N satisfying (2.34),
we have with probability 1, for all but a finite number of indices N,

$$d_{TV}(\mu_{N,\beta}, \mu^{(1)}_{k_N,\beta}) \leq \frac{4}{\frac{\beta}{\beta_c} - 1} \times \frac{1}{(k_N - 1)^{\frac{1}{2}(\frac{\beta}{\beta_c}-1)}}. \tag{2.38}$$

It simply means that the measure $\mu^{(1)}_{k_N,\beta}$ is a.s. very closed to the Gibbs measure and provides an estimate of the distance.

Another possible approximation we proposed for the simulation, is the following random measure on $\{-1, +1\}^N$:

$$\mu^{(2)}_{k_N, \beta}(\Psi) := \frac{\sum_{k=1}^{k_N} \Psi((\tilde{\sigma})_k) e^{-\frac{\beta}{\beta_c} \sum_{\ell=1}^{k} \frac{W_\ell}{\ell}}}{\sum_{k=1}^{k_N} e^{-\frac{\beta}{\beta_c} \sum_{\ell=1}^{k} \frac{W_\ell}{\ell}}}, \tag{2.39}$$

where (W_i) are iid exponential $\mathcal{E}(1)$ r.v. ;

in order to have a good approximation (considering $d_{TV}(\mu_{N,\beta}, \mu^{(2)}_{k_N, \beta})$), a supplementary condition on k_N was added, namely $\frac{(\log k_N)(\log N)}{N} \downarrow 0, \ N \to \infty$,

i.e. k_N must satisfy, as $N \to \infty$,

$$k_N \uparrow \infty, \quad \frac{k_N}{N} \uparrow \infty \quad \text{and} \quad \frac{(\log k_N)(\log N)}{N} \downarrow 0. \tag{2.40}$$

This construction comes, in particular, from the fact that we can prove that for all k_{2^N} satisfying (2.40), on Ω_n^* (s.t. $P[\Omega_n^*] \to 1$ as $n := 2^N \to \infty$),

$$\forall j, \ 1 \leq j \leq k_n, \quad \sqrt{2 \log n}(X_{j,n} - X_{1,n}) = \sum_{i=1}^{j} \frac{W_i}{i} + O\left(\frac{\log(k_n) \log \log n}{\log n}\right).$$

Those various results allowed to deduce some properties of the Gibbs measure, namely :

(i) $\forall \varepsilon > 0, \exists k_N, \ \limsup_{N \uparrow \infty} \mu_{N,\beta}\left(\{(\sigma)_1, \ldots, (\sigma)_{k_N}\}^c\right) \leq \epsilon \quad a.s.$

(ii) When $\beta_c < \beta < \infty$, the Gibbs measure cannot be concentrated on the minimum, i.e. $\mu_{K_N, \beta} \neq \delta_{(\tilde{\sigma})_1}$.

(iii) At zero temperature, the Gibbs measure is a.s. the point mass at the minimum.

(iv) With probability 1, the Gibbs measure is not the uniform measure on the k_N first minima $\{(\tilde{\sigma})_1, \ldots, (\tilde{\sigma})_{k_N}\}$.

(v) With probability 1, the Gibbs measure is not the measure $\mu_{k_0, \beta}$ for a finite k_0.

(vi) With probability 1, for $\beta_c < \beta < \infty$, the Gibbs measure has not all his mass concentrated on $\{(\tilde{\sigma})_1, \ldots, (\tilde{\sigma})_{k_0}\}$, for a finite k_0.

Now let us consider the simulation of the Gibbs measure, namely :

- start with a 2^N-sample of independent uniformly distributed random variables on $(0, 1)$, $(U_1, \ldots, U_{2^N})$, and order it : $(U_{1,2^N} \leq \ldots \leq U_{2^N, 2^N})$;
- construct $\exp\left\{-\frac{\beta}{\beta_c} \sqrt{2 \log 2^N}(X_{k,2^N} - X_{1,2^N})\right\}$,

  by using $X_{i,n} - X_{1,n} \overset{d}{=} G(U_1, n) - G(U_i, n)$, where $1 - \Phi(G(u)) = u, \ 0 < u < 1$;
- consider only the last k_N terms $U_{2^N, 2^N}, \ldots, U_{2^N - k_N + 1, 2^N}$, with k_N satisfying (2.34) ;
- independently, choose one after the other and without replacement, k_N spins configurations in $\{-1, +1\}^N$; thus an ordered sequence of configurations is obtained and denoted by $((\sigma)_1, \ldots, (\sigma)_{k_N})$;
- then

$$\tilde{\mu}^{(1)}_{k_N, \beta}(\Psi) := \frac{\Psi((\sigma)_1) + \Sigma_{k=2}^{k_N} \Psi((\sigma)_k) e^{-\frac{\beta}{\beta_c} \sqrt{2 \log 2^N}(G(U_{1,2^N}) - G(U_{k,2^N}))}}{1 + \Sigma_{k=2}^{k_N} e^{-\frac{\beta}{\beta_c} \sqrt{2 \log 2^N}(G(U_{1,2^N}) - G(U_{k,2^N}))}} \overset{d}{=} \mu^{(1)}_{k_N, \beta}.$$

An alternative and easy way to proceed is to consider $\mu^{(2)}_{k_N,\beta}$ defined in (2.39), instead of $\mu^{(1)}_{k_N,\beta}$; in this case, we just have to consider a k_N-sample of independent uniformly distributed random variables $U_1,\ldots,U_{k_N}$, with k_n satisfying (2.40), then to choose as before the k_N spin configurations, $(\sigma)_1,\ldots,(\sigma)_{k_N}$, to construct the resulting measure

$$\tilde{\mu}^{(2)}_{k_N,\beta}(\Psi) := \frac{\Psi((\sigma)_1) + \Sigma_{k=2}^{k_N}\Psi((\sigma)_k)e^{+\frac{\beta}{\beta_c}\sum_{l=1}^{k}\frac{\log U_\ell}{\ell}}}{1 + \Sigma_{k=2}^{k_N}e^{+\frac{\beta}{\beta_c}\sum_{l=1}^{k}\frac{\log U_\ell}{\ell}}}, \tag{2.41}$$

which has the same law as $\mu^{(2)}_{k_N,\beta}$. Note that this second procedure needs only two independent samples : one of k_N spin configurations chosen without replacement within 2^N, and one of k_N independent uniform random variables on $(0,1)$, so we don't have to work with the k_n largest order statistics.

For instance, k_n could be chosen as $k_N = N \log \log \log N$.

2.2.3 Some results involving order statistics and sums of spacings

As told previously, all the results of the previous section were applications of some new results concerning order statistics and sums of spacings, that we will enumerate below because of their independent interest, after recalling some notations and some well known results very useful : $(X_i, i \in \{1,\ldots,n\})$ are i.i.d. standard Gaussian random variables with distribution function Φ, $(U_i, i \in \{1,\ldots,n\})$ are i.i.d. $\mathcal{U}(0,1)$, uniformly distributed on $(0,1)$, random variables, and $(W_i, i \in \{1,\ldots,n\})$ are i.i.d. exponentially distributed r.v. $\mathcal{E}(1)$. Let $S_0 := 0$ and $S_m := \sum_{k=1}^{m} W_k$, $m \in \mathbb{N}^*$.

We keep the same notation for the order statistics, i.e. $Y_{1,n} \leq \ldots \leq Y_{n,n}$ denote the order statistics associated to some random variables $(Y_i, i \in \{1,\ldots,n\})$.

Set $U_{0,n} := 0$ and $U_{n+1,n} := 1$. Let G be defined by $1 - \Phi(G(u)) = u$, $0 < u < 1$. Then

Lemma 2.2.1 *(Deheuvels, 1985)*
G satisfies, as $u \downarrow 0$,

$$G(u) = \sqrt{2\log(1/u)} - \frac{\log\log(1/u) + \log(4\pi)}{2\sqrt{2\log(1/u)}} + O\left(\frac{(\log\log(1/u))^2}{(\log(1/u))^{3/2}}\right).$$

As it is standard, we can construct the Gaussian random variables by using $X_i = G(U_i)$ and since G is decreasing, we have $X_{i,n} = G(U_{n-i+1,n})$. Then, by symmetry, we have the identity in distribution (denoted by $\overset{d}{=}$)

$$X_{i,n} - X_{1,n} \overset{d}{=} G(U_{1,n}) - G(U_{i,n}).$$

A classical representation of the Gaussian random variables in term of uniform random variables is used mainly because a lot of explicit distributions for the spacings of uniform random variables are available.

Let us mention two last well known results, namely

Lemma 2.2.2 *(Pyke, 1965)*

$$\{U_{i,n}, 0 \le i \le n+1\} \overset{d}{=} \{S_i/S_{n+1}, 0 \le i \le n+1\}$$

and also

Lemma 2.2.3 *(Malmquist, 1950)*

$$\left\{ \xi_i^{(n)} := \left(\frac{U_{i,n}}{U_{i+1,n}} \right)^i ; \quad 1 \le i \le n \right\} \text{ are i.i.d. } \mathcal{U}(0,1) \text{ random variables.}$$

Then we proved the following inequalities :

Lemma 2.2.4 *(Kratz and Picco, 2004)*
For all $\delta > 0$, there exists $n_0 = n_0(\delta)$ such that for all $n > n_0$,

$$I\!P\left[|\log(1/U_{1,n}) - \log n| \le 2\log(\log n)^{1+\delta} \right] \ge 1 - \frac{4}{(\log n)^{1+\delta}}$$

Lemma 2.2.5 *(Kratz and Picco, 2004)*
For all positive x, for all positive integers j_0, j_1 (with $j_1 \ge j_0$), and for all positive t such that $\frac{t^2}{j_0^2} \le 3/4$, we have

$$I\!P\left[\sum_{l=j_0}^{j_1} \frac{W_l - E[W_l]}{l} \ge x \right] \le e^{-tx + \frac{t^2}{j_0}\left(1 + 2\frac{t^2}{j_0^2}\right)\left(1 + \frac{1}{j_0}\right)}.$$

In particular, for $t^2 \le 3/4$,

$$I\!P\left[\sum_{l=1}^{j} \frac{W_l - E[W_l]}{l} \ge x \right] \le e^{-tx + 2t^2(1+2t^2)}.$$

It allowed us to obtain :

Proposition 2.2.1 *(Kratz and Picco, 2004)*
We have

$$\log(U_{1,n}/U_{j,n}) \overset{d}{=} -\sum_{i=1}^{j-1} W_i/i$$

$$\overset{d}{=} -Z + \sum_{i=j}^{\infty} \left(\frac{W_i}{i} - \frac{E[W_i]}{i} \right) - \sum_{i=1}^{j-1} \frac{E[W_i]}{i},$$

where Z is a random variable such that, $\forall x > 0$,

$$P[Z \ge x] \le e^{-x\frac{\sqrt{3}}{2} + \frac{15}{4}} \quad \text{and} \quad P[Z \le -x] \le e^{-x\frac{\sqrt{3}}{2} + \frac{15}{4}}.$$

Moreover, for all $j > e$, we have

$$\mathbb{P}\left[\log(U_{1,n}/U_{j,n}) \leq -Z + \sqrt{\frac{3\log j}{j}} - \log j\right] \geq 1 - \frac{e^{3\frac{(\log j)^2}{j}}}{j^{3/4}};$$

More generally, $\forall c > 0$, $\forall 0 < \varepsilon < 1$, $\exists j_0 = j_0(c, \varepsilon)$, $\forall j \geq j_0$,

$$\mathbb{P}\left[\log(U_{1,n}/U_{j,n}) \leq -Z + \sqrt{\frac{c\log j}{j}} - \log j\right] \geq 1 - \frac{1}{j^{(1-\varepsilon)c/4}}.$$

The main result needed for the sum of spacings is then given by :

Proposition 2.2.2 *(Kratz and Picco, 2004)*
For all $0 < \delta$, for all $0 < \varepsilon < 1$, and for all k_n satisfying (2.34),
by defining, for $0 < \lambda < 1$, $0 < \alpha < 1$,

$$\lambda_n := \lambda\frac{\log n}{n^\alpha}, \tag{2.42}$$

and

$$\tilde{\lambda}_n := \lambda_n\left(1 + 2\sqrt{\frac{2}{\lambda}}\frac{(1 - \lambda_n)}{n^{\frac{1-\alpha}{2}}}\right), \tag{2.43}$$

there exists $n_0 = n_0(\varepsilon, \delta, \lambda, \alpha)$ such that $\forall n \geq n_0$, there exists $\Omega_n \subset \Omega$, with

$$\mathbb{P}[\Omega_n] \geq 1 - 4\left(\frac{1}{(\log n)^{1+\delta}} + \frac{e^{-k_n/16}}{1 - e^{-1/16}}\right),$$

such that on Ω_n, for all j such that $k_n \leq j \leq n\tilde{\lambda}_n$, we have

$$\sqrt{2\log n}(X_{j,n} - X_{1,n}) \geq 2\log n\left(1 - \sqrt{1 - \frac{\log j}{\log n}}\right)(1 - \varepsilon) \tag{2.44}$$

while for all j such that $n\tilde{\lambda}_n \leq j \leq n$,

$$\sqrt{2\log n}(X_{j,n} - X_{1,n}) > 2\log n\sqrt{1 - \varepsilon} - G(\lambda_n)\sqrt{2\log n} \tag{2.45}$$

In particular, on Ω_n,

$$for\ \frac{\log j}{\log n} \uparrow 1, \quad \sqrt{2\log n}(X_{j,n} - X_{1,n}) > 2\log n\sqrt{1 - \varepsilon} - G(\lambda_n)\sqrt{2\log n}. \tag{2.46}$$

Moreover, for $0 < \eta < 1$, there exists $\Omega_{k_n}^*$ with $\mathbb{P}(\Omega_{k_n}^*) \geq 1 - 2\exp\left\{-\frac{\sqrt{3}}{2}(\log k_n)^\eta\right\}$, such that

on $\Omega_{k_n}^*$, for all j such that $k_n \leq j \leq n\tilde{\lambda}_n$ and $\frac{\log j}{\log n} \downarrow 0$, (2.44) can be refined as

$$\sqrt{2\log n}(X_{j,n} - X_{1,n}) \geq \log j\left(1 - \frac{1}{(\log j)^{1-\eta}}\right). \tag{2.47}$$

Remarks :

1) If λ_n is chosen as a constant λ independent of n, $0 < \lambda < 1$, (2.45) gives that : $\forall j > n\lambda$,
$\sqrt{2 \log n}(X_{j,n} - X_{1,n}) \geq 2 \log n \sqrt{1 - \varepsilon} - G(\lambda)(\sqrt{2 \log n}) > 2 \log n(1 - \epsilon)$ if $n > n_0(\lambda, \epsilon))$ for
some $n_0(\lambda, \epsilon)$. When one enters in various regimes as $n\tilde{\lambda}_n^{(1)} < j < n\tilde{\lambda}_n^{(2)}$ with $\lambda_n^{(i)} \downarrow 0$, for $i = 1, 2$,
a cancellation could occur between the two terms in the right hand side of (2.45). The choice of
λ_n in (2.42) allows then to see such cancellation, since for $\alpha < 1/4$ it provides that $\forall j \geq n\tilde{\lambda}_n$,
on Ω_n

$$\sqrt{2 \log n}(X_{j,n} - X_{1,n}) \geq 2(1 - \sqrt{\alpha})(1 - \varepsilon) \log n \geq (1 - \varepsilon) \log n. \qquad (2.48)$$

2) Note that the lower bounds in (2.44) and (2.45), even if obtained by two completely different
methods, are of the same order $2(1 - \sqrt{\alpha}) \log n$ when $j = n\tilde{\lambda}_n$.
3) Note that for $n = 2^N$, under (2.34), we have $\sum_{N=1}^{\infty} I\!P[\Omega_{2^N}^c] < \infty$. It is what allowed us, in
the previous section, to get some results true $I\!P$–almost surely for all but a finite number of
indices N (by the first Borel-Cantelli Lemma).

The last result given on the sum of spacings showed that the first k_N sums of spacings can be
represented as successive partial weighted sums of exponential random variables.

Proposition 2.2.3 *(Kratz and Picco, 2004)*
*For all $\delta > 0$, for all k_n satisfying (2.34), for $\Omega_n^* \subset \Omega$, given by $\Omega_n^* := \Omega_{n,k} \cap \Omega_{n,\delta} \cap \tilde{\Omega}_{k_n}$ defined
respectively by*

$$\Omega_{n,k} := \left\{ \sup_{1 \leq j \leq k_n} U_{j,n} = U_{k_n,n} \leq (1 + \gamma)\frac{k_n}{n} \right\}, \quad P[\Omega_{n,\delta}] \geq 1 - \frac{4}{(\log n)^{1+\delta}},$$

and

$$\Omega_{k_n} := \left(\cap_{j=k_n}^{n\tilde{\lambda}_n} \tilde{\Omega}_j \right) \cap \Omega_{n,\delta}, \quad where \quad \tilde{\Omega}_j := \left\{ \left| \frac{S_j}{j} - 1 \right| < 1/2 \right\},$$

then $I\!P[\Omega_n^*] \geq 1 - \dfrac{6}{(\log n)^{1+\delta}}$, and on Ω_n^*, we have, $\forall j$ s.t. $1 \leq j \leq k_n$,

$$\sqrt{2 \log n}(X_{j,n} - X_{1,n}) = \sum_{i=1}^{j} \frac{W_i}{i} + O\left(\frac{\log(k_n) \log \log n}{\log n} \right)$$

2.2.4 Conclusion and references

In this sudy, we got interested in a problem related to the equilibrium Statistical Mechanics of
Spin Glasses, namely the study of the Gibbs measure of the Random Energy Model. To give a
precise description of the support of the Gibbs measure below the critical temperature required
quite a lot of work, in particular to establish some new results on sums of spacings for i.i.d.
Gaussian random variables, of independent interest too.
The next and natural step will be to deal with the Generalized Random Energy Model, when
the assumption of independence of the variables is dropped in favor of the hypothesis of positive
correlation between the r.v.

Bibliographie

[1] Bovier A., Statistical Mechanics of Disordered Systems, Lecture Notes by A. Bovier, Ma-PhySto program, May 2001.

[2] Bovier A., Kurkova I., Loewe M., Fluctuations of the free energy in the p-spin S.K. Model and in the R.E.M., *Annals of Prob.* **30**, 605-651 (2002).

[3] David, H. *Order statistics.* Wiley, N.Y. (2nd ed.)(1981).

[4] Deheuvels, P., The limiting behavior of the maximal spacing generated by an i.i.d. sequence of Gaussian random variables, *J. Appl. Prob.* **22**, 816-208 (1985).

[5] Deheuvels, P., On the influence of the extremes of an i.i.d. sequence on the maximal spacings. *Ann. Probab.* **14**, 194–208 (1986).

[6] Derrida B., Random Energy Model : Limit of a Family of Disordered Models, *Phys. Rev. Lett.* **45**, 79-82 (1980).

[7] Derrida B., Random Energy Model : An Exactly Solvable Model of Disordered Systems, *Phys. Rev. B* **24**, 2613-2626 (1981).

[8] Eisele Th., On a Third Order Phase Transition, *Comm. Math. Phys.* **90**, 125-159 (1983).

[9] Fontes L.R.G., Isopi M., Kohayakawa Y., Picco P., The spectral gap of the REM under Metropolis dynamics, *Annals of App. Prob.* **8**, 917-943 (1998).

[10] Galves J.A., Martinez S. and Picco P., Fluctuations in Derrida's Random Energy and Generalized Random Energy Models, *J. Stat. Phys.* **54**, 515-529 (1989).

[11] Kratz, M. and Picco, P., On a representation of Gibbs measure for R.E.M., *Ann. Applied Probab.* **14**, 651-677 (2004).

[12] Leadbetter, M.R., Lindgren, G. and Rootzen, H., *Extremes and Related Properties of Random Sequences and Processes,* Springer-Verlag (1983).

[13] Malmquist, S., On the properties of order statistics from a rectangular distribution, *Skand. Aktuarietidskr.* **33**, 214-222 (1950).

[14] Mathieu P., Picco P., Metastability and convergence to equilibrium for the random field Curie-Weiss model, *J. Stat. Phys.* **91**, 679-732 (1998).

[15] Mezard M., Parisi G. and Virasoro M., *Spin Glass theory and beyond,* World Scientific (1980).

[16] Olivieri E., Picco P., On the Existence of Thermodynamics for the Random Energy Model, *Comm. Math. Phys.* **96**, 125-144 (1984).

[17] Petrov V.V., *Sums of Independent Random Variables*, Springer (New York) (1978).

[18] Pyke (1965), Spacings, *J. Roy. Statist. Soc.* **27**, 395-436 (1965).

[19] Ruelle D., A mathematical reformulation of Derrida's REM and GREM, *Com. Math. Phys.* **108**, 225-239 (1987).