SYNTHÈSE POUR L'HABILITATION À DIRIGER DES RECHERCHES

# Contributions à l'étude de quelques processus stochastiques à sauts

*par*


## Ronan LE GUÉVEL


Présentée et soutenue publiquement le 14 Février 2024 à l'Université Rennes 2 devant le jury ci-dessous

| | |
|---|---|
| Hermine BIERMÉ | Professeur (Université de Tours) |
| Alexandra CARPENTIER | Professeur (Universität Postdam) |
| Jean-François COEURJOLLY | Professeur (Université Grenoble Alpes) |
| Fabienne COMTE | Professeur (Université Paris Descartes) |
| Marc HOFFMANN | Professeur (Université Paris Dauphine) |
| Eva LÖCHERBACH | Professeur (Université Paris Panthéon Sorbonne) |

Au vu des rapports de :
Alexandra CARPENTIER, Jean-François COEURJOLLY et Marc HOFFMANN.

A Gaëlle, Antoine et Maxime

# Table des matières

# Introduction

Ce document propose une synthèse de mes travaux de recherche effectués depuis la fin de ma thèse. Ils portent globalement sur trois grandes classes de processus : les processus multistables, les processus de Markov jump-move et les processus de comptage (avec en particulier le processus de Poisson). Ces trois classes ont en commun de regrouper des processus à sauts en temps continu pouvant servir de modèles pour des données aléatoires présentant une inhomogénéité temporelle dans l'intensité des sauts. Ma contribution porte dans chaque cas sur l'étude de quelques propriétés probabilistes caractéristiques de ces modèles, ainsi que sur l'obtention de résultats de contrôle d'erreur pour des procédures statistiques adaptées à ces modèles.

Ce mémoire comporte essentiellement trois parties principales. La première partie concerne l'étude de régularité de processus multifractionnaires, avec en particulier celle des processus multistables. Ce thème de recherche fut au cœur de mon travail de thèse de doctorat, dirigé par Anne Philippe et Jacques Lévy-Véhel. J'ai poursuivi ces travaux après ma soutenance tout en démarrant deux grands projets de recherche alternatifs. Le premier projet fait l'objet de la deuxième partie de ce manuscrit et concerne l'étude des processus de type jump-move. Ces processus ont émergé à la suite de discussions avec Frédéric Lavancier dont le but était d'introduire du mouvement dans des modèles de données spatiales. L'étude de ces processus se poursuit encore maintenant avec notamment le co-encadrement de la thèse d'Emilien Manent. Le deuxième projet alternatif, objet de la troisième grande partie, aborde la question de la détection de rupture dans l'intensité d'un processus de Poisson sous l'angle des tests minimax. En collaboration avec Magalie Fromont, ce projet, qui a donné lieu au co-encadrement de la thèse de Fabrice Grela, s'inscrit plus largement dans ma démarche de caractérisation d'inhomogénéités dans l'observation d'un processus stochastique.

Ces activités de recherche sont décrites plus précisément sous la forme de quatre chapitres. Dans le premier chapitre je rappelle quelques définitions de modèles de processus stochastiques à sauts, en commençant par les processus de comptage et leur martingale carrée associée, processus dont le processus de Poisson est le représentant le plus célèbre. Ce chapitre sera l'occasion de présenter les inégalités exponentielles obtenues dans [A9] pour les queues de distribution des processus de comptage et leurs martingales associées, inégalités qui constituent un outil clé pour le contrôle d'erreurs $L^2$ des statistiques d'intérêt dans les problèmes relatifs au processus jump-move considéré dans le chapitre 3. Elles sont également au cœur des résultats minimax non-asymptotiques sur les tests de détection de rupture d'intensité d'un processus de Poisson étudiés dans le chapitre 4. Je rappelle également dans ce premier chapitre plusieurs définitions équivalentes des processus stables avec comme exemple le plus simple le processus de Lévy $\alpha$-stable, modèle de base des processus multifractionnaires. Ces définitions sont à l'origine des définitions (non équivalentes cette fois-ci) des processus multistables dont les études probabilistes et statistiques sont présentées dans le chapitre 2.

Ce deuxième chapitre porte sur l'étude de diverses notions de régularités de processus multifractionnaires avec en ligne de mire celle des processus multistables. Plusieurs exposants sont utilisés pour caractériser la régularité d'un processus. Le plus courant est l'exposant de Hölder ponctuel, dont l'étude précise est faite pour le processus de Lévy multistable dans [A3] et [A8]. D'autres exposants sont aussi largement étudiés, parmi lesquels l'exposant de Hölder local ou encore l'exposant 2-microlocal, qui sont également des exposants de régularité sur les trajectoires d'un processus. Outre l'exposant de Hölder ponctuel, je me suis aussi intéressé dans [A1], [A2] et [A3] à l'indice de localisabilité des processus multistables, indice caractéristique de la régularité en loi d'un processus. Les résultats obtenus montrent que cette régularité, qu'elle soit trajectorielle ou en loi, est intimement reliée à la fonction de stabilité $\alpha$ qui régit la hauteur typique des sauts du processus. D'un point de vue statistique, je me suis donc intéressé dans [A4] au problème d'estimation de ces deux fonctions qui gouvernent la régularité locale des processus multistables, la fonction de localisabilité $H$ et la fonction de stabilité $\alpha$. D'un point de vue pratique, il est naturel de se demander au préalable si le problème est de nature paramétrique ou non. Je propose ainsi dans [A7] un test statistique permettant de décider si la fonction de stabilité $\alpha$ est constante ou non lorsque l'on observe un processus de Lévy multistable. D'autres caractéristiques que ces exposants apparaissent avec les processus fractionnaires, comme des lignes de niveau fractales par exemple. Elles sont généralement mesurées grâce à la mesure de Hausdorff et la dimension fractale qu'elle induit. La régularité du temps local et de la mesure d'occupation d'un processus est également reliée à la régularité trajectorielle. Ces notions ont ainsi été à l'origine de plusieurs travaux. Je me suis intéressé dans [A6] au calcul de la dimension fractale de l'ensemble image d'un processus de Lévy multistable, calcul qui fait réapparaître la différence de comportement selon que le processus est à variations bornées ($\alpha < 1$) ou non. On retrouve cette différence dans la vitesse d'estimation de la mesure d'occupation d'un processus stable, problème que j'ai abordé avec Randolf Altmeyer dans [A11]. Dans ce cas la solution la plus simple est la meilleure, au sens où pour estimer l'intégrale du temps local, l'estimateur de Riemann est de taux de convergence $L^2$-optimal. Enfin je termine ce chapitre avec la caractérisation complète de la régularité ponctuelle d'un processus de Lévy multistable, donnée dans [A8] avec le calcul des spectres de singularités de Hausdorff et de Legendre.

Le troisième chapitre est consacré à la présentation du processus jump-move introduit avec Frédéric Lavancier dans [A10], processus qui constitue le centre d'intérêt de la thèse d'Emilien Manent. On introduit ce processus avec comme objectif la modélisation d'un système de particules en interaction, mais il peut néanmoins servir de modèle dans un cadre plus large. A cette fin le cas particulier du processus Birth-Death-Move généralise des processus classiques dans le domaine des dynamiques de population, introduisant un mouvement aléatoire markovien pour les individus de la population. Ce processus à sauts étant encore un processus de Markov, nous étudions dans [A13] les propriétés de Feller de son semi-groupe associé. Ses propriétés de convergence en temps long vers un régime stationnaire sont également considérées, afin d'assurer notamment l'accumulation d'observations de sauts similaires au cours du temps. Ces résultats de convergence sont obtenus aussi bien dans le cadre d'une population de taille bornée que dans celui d'une population non bornée. On estime alors l'intensité de sauts du processus pour une observation en temps continu grâce à un estimateur à noyau. On obtient des garanties théoriques de convergence avec des taux usuels en statistique non-paramétrique, en particulier lorsque les données sont projetées dans un espace de dimension finie. Des résultats semblables sont obtenus pour des observations effectuées le long d'une discrétisation de la trajectoire dans un cadre hautes fréquences. Afin de donner une méthode utilisable en pratique, je présente enfin une procédure de sélection de fenêtre par validation croisée basée sur une vraisemblance partielle.

Pour terminer, le quatrième chapitre porte sur le problème de la détection de rupture d'intensité d'un processus de Poisson inhomogène simple, problème considéré avec Magalie Fromont lors du co-encadrement de la thèse de Fabrice Grela. Plus précisément, en considérant des intensités de la forme $\lambda(t) = \lambda_0 + \delta \mathbf{1}_{(\tau,1]}(t)$ par rapport à la mesure $d\Lambda(t) = Ldt$, l'objectif consiste à détecter le cas $H_0 : \lambda \equiv \lambda_0$ avec l'observation d'une trajectoire du processus de Poisson sur l'intervalle $[0, 1]$. Ce problème est abordé dans un premier temps sous l'angle des tests simples, en considérant des tests répondant au critère d'optimalité défini au sens du minimax. Dans le cas d'une intensité de base $\lambda_0$ connue, nous considérons dans [A12] toutes les alternatives possibles, selon la connaissance ou non de la hauteur du saut $\delta$ ou de la position du saut $\tau$. Pour chaque problème de test, on donne la vitesse de séparation minimax (en $L$) accompagnée d'une statistique de test qui atteint cette vitesse minimax. Ce travail est également effectué dans le cas où l'intensité de base $\lambda_0$ n'est pas connue. Les outils de calcul, basés sur des inégalités de martingales dans le cas où $\lambda_0$ est connue, changent alors pour des méthodes basées sur le processus empirique, les vitesses de séparation minimax restant quant à elles inchangées. Dans un second temps, le problème de la localisation de l'instant de rupture $\tau$ est abordé dans [A14] sous l'angle des tests multiples. Plus précisément, il s'agit alors de considérer simultanément les hypothèses nulles de la forme $H_k : \lambda(t) = \lambda_0 + \delta \mathbf{1}_{(\tau,1]}(t)$ avec $\tau \in [k/M, 1]$. En utilisant le critère d'erreur de première espèce FWER, nous étudions la vitesse de séparation minimax de ce problème de tests multiples pour le critère de seconde espèce FWSR dans le cas d'une intensité de base $\lambda_0$ connue. Nous déterminons également dans ce contexte la vitesse de séparation minimax ainsi qu'une procédure de tests multiples qui atteint cette vitesse de séparation minimax, aussi bien pour des intensités dont la hauteur de saut $\delta$ est connue que pour des intensités de hauteur de saut inconnue. Je terminerai ce chapitre en établissant un lien entre ces procédures de tests multiples minimax et la détermination d'un intervalle de confiance de longueur minimale pour l'instant de rupture $\tau$.

# Chapitre 1

# Présentation de quelques modèles de processus à sauts

On rappelle ici quelques définitions et notations de processus à sauts apparaissant tout au long de ce manuscrit.

## 1 Les processus de comptage

Les processus à sauts les plus élémentaires sont certainement les processus de comptage.

**Définition 1.1** Un processus réel $N = (N_t)_{t \geq 0}$ est un processus de comptage si ses trajectoires sont croissantes par sauts d'amplitude 1, continues à droite et nulles à l'instant 0.

Pour une filtration $\mathcal{F}$ de l'espace $(\Omega, \mathbb{P})$, on note $\mathcal{P}(\mathcal{F})$ la tribu prévisible pour $\mathcal{F}$ sur $\Omega \times \mathbb{R}_+$, engendrée par les processus adaptés à $\mathcal{F}$ et continus à gauche. Un processus $Y$ sera dit $\mathcal{F}$-prévisible si $(\omega, t) \mapsto Y_t(\omega)$ est $\mathcal{P}(\mathcal{F})$-mesurable. On définit alors le compensateur d'un processus de comptage.

**Définition 1.2** Soit $N$ un processus de comptage. On appelle compensateur de $N$ un processus croissant, continu à droite et $\mathcal{F}$-prévisible $\Lambda$ tel que $N - \Lambda$ soit une martingale locale (par rapport à la filtration $\mathcal{F}$). Si $\mathbb{P}$-presque sûrement toute trajectoire de $\Lambda$ est Lebesgue absolument continue, la dérivée de Radon-Nikodym de $\Lambda$ s'appelle l'intensité de $N$ pour la filtration $\mathcal{F}$.

L'exemple de base de processus de comptage est le processus de Poisson, élément central du Chapitre 4, dont on rappelle la définition ici.

**Définition 1.3** Le processus de comptage $N$ est un processus de Poisson si et seulement si

i) $N$ est à accroissements indépendants,

ii) la fonction $\Lambda : t \mapsto \mathbb{E}[N_t]$ est continue.

La fonction $\Lambda$ est alors le compensateur pour la filtration naturelle de $N$.

Dans beaucoup de problèmes statistiques faisant intervenir des processus à sauts, et notamment pour des problèmes d'estimation d'intensité de sauts, on utilise des statistiques linéaires de la forme $\int_0^t H_s dN_s$ ou des statistiques quadratiques de type $U$-statistiques utilisant $(\int_0^t H_s dN_s)^2$ avec $H$ un processus prévisible borné. Bien que la définition de $N$ apparaisse assez générale, le comportement de ces statistiques est toujours intimement relié à celui de $\int_0^t H_s d\Lambda_s$. En posant

$$M_t = \int_0^t H_s d(N - \Lambda)_s$$

et

$$\tilde{M}_t = (\int_0^t H_s d(N - \Lambda)_s)^2 - \int_0^t H_s^2 dN_s,$$

on obtient dans [A9] des inégalités exponentielles pour $M$ et $\tilde{M}$ (Theorem 5 et 6 de [A9]) :

**Théorème 1.1** Pour tout $x > 0$ et $T > 0$,

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |M_t| \geq x\right) \leq 2\exp\left(-\frac{\|H\|_{2,T}^2}{\|H\|_{\infty,T}^2} I(\frac{\|H\|_{\infty,T}}{\|H\|_{2,T}^2} x)\right)$$

et

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |\tilde{M}_t| \geq x\right) \leq 6\exp\left(-\frac{\|H\|_{2,T}^2}{\|H\|_{\infty,T}^2} I(\frac{\|H\|_{\infty,T}}{\|H\|_{2,T}^2} \sqrt{\frac{x}{2}})\right)$$

avec $I(t) = (1+t)\log(1+t) - t$, $\|H\|_{\infty,T}$ un majorant presque sûr de $\sup_{0\leq t\leq T} |H_s|$ et $\|H\|_{2,T}^2$ un majorant presque sûr de $\int_0^T H_s^2 d\Lambda_s$.

Ces processus de comptage se généralisent en des processus à sauts de tailles autres que 1 avec le processus de Poisson composé par exemple, processus qui peut se décrire comme un processus de Lévy (de mesure de Lévy finie), dont on rappelle ici les propriétés élémentaires, et que l'on peut trouver par exemple dans Sato (1999).

**Définition 1.4** Un processus stochastique réel $(X_t)_{t\geq 0}$ est un processus de Lévy si

i) presque sûrement, $X$ est càdlàg

ii) $X$ est continu en probabilité

iii) $X$ est à accroissements indépendants

iv) $X$ est à accroissements stationnaires (la loi de $X_{s+t} - X_s$ est la loi de $X_t - X_0$).

La fonction caractéristique d'un processus de Lévy $(X_t)_{t\geq 0}$ est alors donnée par la formule de Lévy-Khintchine : pour $t \geq 0$ et $\theta \in \mathbb{R}$,

$$\mathbb{E}[e^{i\theta X_t}] = \exp\left(t\psi(\theta)\right)$$

avec

$$\psi(\theta) = i\gamma\theta - \frac{\sigma^2}{2}\theta^2 + \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{[-1,1]}(x))\nu(dx)$$

où $\nu$ est la mesure de Lévy vérifiant $\nu(\{0\}) = 0$ et $\int_{\mathbb{R}}(|x|^2 \wedge 1)\nu(dx) < \infty$. Les processus de Lévy à sauts font souvent office de référence pour des modèles de processus où l'homogénéité est présente. Dans le cas où $\nu(\mathbb{R}) = +\infty$, on obtient une activité de sauts infinie avec comme représentant le processus de Lévy symétrique $\alpha$-stable (pour lequel $\nu(dx) \propto |x|^{-1-\alpha}dx$), processus qui appartient aussi plus généralement à la famille des processus stables.

## 2 Les processus stables

On rappelle ici quelques définitions équivalentes des processus stables symétriques, présentés par exemple dans Samorodnitsky and Taqqu (1994), suivies ensuite de celles des processus multistables qui seront au cœur du Chapitre 2.

**Définition 1.5** Un processus stochastique réel $(X_t)_{t\in T}$ est symétrique stable si toutes ses lois finies dimensionnelles sont symétriques stables, i.e. s'il existe $\alpha \in (0, 2]$ tel que toute combinaison linéaire $\sum_{k=1}^{d} b_k X_{t_k}$ vérifie pour $\theta \in \mathbb{R}$

$$\mathbb{E}[e^{i\theta \sum_{k=1}^{d} b_k X_{t_k}}] = e^{-\sigma_d^\alpha |\theta|^\alpha}$$

pour un certain $\sigma_d > 0$ ne dépendant que de $(b_k)_k$ et $(t_k)_k$.

Pour construire des processus symétriques stables, on peut utiliser des mesures aléatoires stables ou de manière équivalente des intégrales aléatoires stables qui apparaissent par exemple dans Bretagnolle et al. (1966) ou dans Schilder (1970). On considère ainsi $(E, \mathcal{E}, m)$ un espace mesuré, $\mathcal{E}_0 = \{A \in \mathcal{E} : m(A) < +\infty\}$ formé des ensembles de mesure $m$ finie, et $L^0(\Omega)$ l'ensemble des variables aléatoires réelles sur $\Omega$.

**Définition 1.6** Le processus $M : \mathcal{E}_0 \to L^0(\Omega)$ est appelé mesure aléatoire symétrique $\alpha$-stable si pour tous ensembles disjoints $A_1, ..., A_k$ de $\mathcal{E}_0$, les variables $M(A_1), ..., M(A_k)$ sont indépendantes, si presque sûrement

$$M(\bigcup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} M(A_j)$$

avec $A_1, A_2, ...$ ensembles disjoints de $\mathcal{E}_0$ tels que $\bigcup_{j=1}^{+\infty} A_j \in \mathcal{E}_0$, et si pour tout $A \in \mathcal{E}_0$, $\theta \in \mathbb{R}$,

$$\mathbb{E}[e^{i\theta M(A)}] = e^{-m(A)|\theta|^\alpha}.$$

On écrira $M(A) = \int_E \mathbf{1}_A(x)M(dx)$ ce qui permet de définir $\int_E f(x)M(dx)$ (par linéarité et passage à la limite de fonctions étagées), variable aléatoire qui vérifie

$$\mathbb{E}[e^{i\theta \int_E f(x)M(dx)}] = \exp\left(-|\theta|^\alpha \int_E |f(x)|^\alpha m(dx)\right)$$

dès lors que $\int_E |f(x)|^\alpha m(dx) < +\infty$. Beaucoup de processus stables admettent une représentation intégrale de la forme $I(f_t) := \int_E f_t(x)M(dx)$ (en particulier ceux qui sont continus en probabilité, voir Bretagnolle et al. (1966)). En choisissant comme mesure de contrôle $m(dx) = dx$ la mesure de Lebesgue, on peut redéfinir par exemple le processus de Lévy symétrique $\alpha$-stable.

**Exemple 1.1** (Processus de Lévy symétrique $\alpha$-stable) C'est le processus défini pour $t \geq 0$ par

$$L_t^\alpha = \int_0^{+\infty} \mathbf{1}_{x \leq t} M(dx) = \int_0^t M(dx) = M([0,t]).$$

On a également comme exemple principal le mouvement linéaire fractionnaire stable.

**Exemple 1.2** (Mouvement Linéaire Fractionnaire Stable symétrique, LFSM) C'est le processus défini pour $H \in (0,1)$, $H \neq \frac{1}{\alpha}$, et $t \geq 0$ par

$$Y_t^{\alpha,H} = \int_{-\infty}^{+\infty} (|t-x|^{H-1/\alpha} - |x|^{H-1/\alpha})M(dx).$$

Ce processus, défini dans Taqqu and Wolpert (1983) et Maejima (1983) est $H$-autosimilaire à accroissements stationnaires. Pour $\alpha = 2$, c'est le mouvement Brownien fractionnaire, introduit par Kolmogorov (1940) ou Mandelbrot and Van Ness (1968).

**Représentation de Poisson des processus symétriques stables** Les processus stables admettent plusieurs représentations alternatives. On peut citer par exemple la représentation sous forme de somme de Poisson (Samorodnitsky and Taqqu, 1994), aussi connue comme la représentation de Lévy-Itô (Itô, 2004). Pour $(E, \mathcal{E}, m)$ un espace mesuré $\sigma$-fini, on travaille avec un processus de Poisson $\Pi$ sur $E \times \mathbb{R}$ de mesure d'intensité $m \times \mathcal{L}$ avec $\mathcal{L}$ mesure de Lebesgue. $\Pi$ est donc un sous-ensemble aléatoire dénombrable de $E \times \mathbb{R}$ tel que si $N(A)$ est le nombre de points de $\Pi \cap A$, $N(A)$ est de loi $\mathcal{P}((m \times \mathcal{L})(A))$ et $N(A_1)$, ...,$N(A_n)$ sont indépendantes pour $A_1, ..., A_n$ ensembles disjoints de $E \times \mathbb{R}$ (Kingman, 1993).

On considère ensuite $\alpha \in (0,2)$ et $g : E \times \mathbb{R} \to \mathbb{R}$ une fonction $(m \times \mathcal{L})$-mesurable vérifiant $g(x,-y) = -g(x,y)$ ainsi que $|g(x,y)| \leq h(x)|y|^{-1/\alpha}$ pour $h$ telle que $\int_E |h(x)|^\alpha m(dx) < +\infty$. En découpant l'espace $E$ sous la forme d'une union d'ensembles croissants $E = \bigcup_n E_n$ avec

$m(E_n) < +\infty$, on définit le rectangle $R_n = \{(x, y) : x \in E_n, |y| \leq n\} \subset E \times \mathbb{R}$. On peut alors considérer (Falconer and Lévy Véhel, 2009)

$$\Sigma = \sum_{(X,Y)\in\Pi} g(X,Y) = \lim_{n\to+\infty} \sum_{(X,Y)\in\Pi\cap R_n} g(X,Y)$$

qui vérifie pour $\theta \in \mathbb{R}$,

$$\mathbb{E}[e^{i\theta\Sigma}] = \exp\left(-2\int\int \sin^2(\frac{1}{2}\theta g(x,y))m(dx)dy\right).$$

En particulier le processus $X(t) = I(f_t)$ admet la représentation de Poisson

$$I(f_t) = \int_E f_t(x)M(dx) \overset{d}{=} c(\alpha) \sum_{(X,Y)\in\Pi} f_t(X)Y^{<-1/\alpha>} \qquad (0 < \alpha < 2)$$

avec $\alpha \in (0, 2)$, $c(\alpha) = (2\alpha^{-1}\Gamma(1-\alpha)\cos(\frac{\pi\alpha}{2}))^{-1/\alpha}$ et $a^{<b>} = \text{sign}(a)|a|^b$.

**Représentation de Ferguson-Klass-LePage des processus symétriques stables** On peut également citer la représentation sous forme de série de Ferguson-Klass et LePage (Ferguson and Klass, 1972; Le Page, 1980a,b). Pour $(E, \mathcal{E}, m)$ un espace mesuré $\sigma$-fini, on peut trouver une fonction $r : E \to \mathbb{R}_+$ telle que $\hat{m}(dx) = \frac{1}{r(x)}m(dx)$ soit une mesure de probabilité (on choisit $r \equiv m(E)$ si $m(E) < +\infty$). On considère alors trois suites indépendantes : $(\Gamma_i)_{i\geq1}$ les temps d'arrivée d'un processus de Poisson simple de taux 1 sur $\mathbb{R}_+$, $(V_i)_{i\geq1}$ une suite i.i.d. de loi $\hat{m}$ sur $E$ et $(\gamma_i)_{i\geq1}$ une suite i.i.d. de loi $\mathbb{P}(\gamma_i = 1) = \mathbb{P}(\gamma_i = -1) = \frac{1}{2}$. Pour $\alpha \in (0, 2)$, $M$ une mesure aléatoire symétrique $\alpha$-stable et une fonction $f_t : E \to \mathbb{R}$ telle que $\int_E |f_t(x)|^\alpha m(dx) < +\infty$, on a la représentation

$$I(f_t) = \int_E f_t(x)M(dx) \overset{d}{=} C_\alpha \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha} r(V_i)^{1/\alpha} f_t(V_i)$$

avec $C_\alpha = \left(\Gamma(1-\alpha)\cos(\frac{\pi\alpha}{2})\right)^{1/\alpha}$.

**Exemple 1.3** (Processus de Lévy symétrique $\alpha$-stable) Pour $\alpha \in (0, 2)$, le processus de Lévy symétrique $\alpha$-stable admet pour $t \in [0, 1]$ la représentation

$$(L_t^\alpha)_{t\in[0,1]} := \{\int_0^{+\infty} \mathbf{1}_{x\leq t} M(dx), 0 \leq t \leq 1\} \overset{d}{=} \{C_\alpha \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha} \mathbf{1}_{U_i\leq t}, 0 \leq t \leq 1\}$$

où $(U_i)_{i\geq1}$ est une suite i.i.d. de variables de loi uniforme sur $[0, 1]$.

# 3 Les processus multistables

Pour définir les processus multistables, on utilise dans [A2] la représentation de Ferguson-Klass-LePage des processus stables en remplaçant le paramètre $\alpha$ par une fonction $\alpha(t)$ afin de rendre le paramètre de stabilité évolutif dans le temps. Un processus multistable y est donc défini comme étant la somme d'une série de type Ferguson-Klass-LePage, à savoir

$$X_t = C_{\alpha(t)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(t)} r(V_i)^{1/\alpha(t)} f_t(V_i), \quad t \geq 0, \tag{1.1}$$

avec $C_\alpha = \left(\Gamma(1-\alpha)\cos(\frac{\pi\alpha}{2})\right)^{1/\alpha}$ et $f_t : E \to \mathbb{R}$ une fonction vérifiant $\int_E |f_t(x)|^{\alpha(t)} m(dx) < +\infty$, les trois suites $(\Gamma_i)_{i\geq 1}$, $(V_i)_{i\geq 1}$ et $(\gamma_i)_{i\geq 1}$ étant les trois suites indépendantes des séries de Ferguson-Klass-LePage. Il est alors immédiat avec cette définition de constater que pour tout $t \geq 0$, $X_t$ est une variable aléatoire $\alpha(t)$-stable vérifiant pour $\theta \in \mathbb{R}$,

$$\mathbb{E}[e^{i\theta X_t}] = \exp\left(-|\theta|^{\alpha(t)} \int_E |f_t(x)|^{\alpha(t)} m(dx)\right).$$

Avec cette représentation on peut définir par exemple une extension multistable du processus de Lévy symétrique $\alpha$-stable sur $[0, T]$.

**Exemple 1.4** (Processus de Lévy multistable) On considère une fonction $\alpha : (0, \infty) \to (0, 2)$. Le processus de Lévy symétrique multistable est défini pour $t \in [0, T]$ par

$$L_t = C_{\alpha(t)} T^{1/\alpha(t)} \sum_{i=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(t)} \mathbf{1}_{U_i \leq t}. \tag{1.2}$$

On peut définir également une version multistable multifractionnaire du Mouvement Linéaire Fractionnaire Stable.

**Exemple 1.5** (Mouvement Linéaire Multifractionnaire Multistable, LMMM) On considère une fonction $\alpha : [0, \infty) \to (0, 2)$ ainsi qu'une fonction $H : [0, \infty) \to (0, 1)$, une suite $(\Gamma_i)_{i\geq 1}$ de temps d'arrivée d'un processus de Poisson simple de taux 1 sur $\mathbb{R}_+$, $(\gamma_i)_{i\geq 1}$ une suite i.i.d. de loi $\mathbb{P}(\gamma_i = 1) = \mathbb{P}(\gamma_i = -1) = \frac{1}{2}$ et $(V_i)_{i\geq 1}$ une suite i.i.d. de loi $\hat{m}(dx) = \frac{3}{\pi^2} \sum_{j=1}^{\infty} j^{-2} \mathbf{1}_{[-j,-j+1[\cup[j-1,j[}(x)dx$ sur $\mathbb{R}$, $(\Gamma_i)_{i\geq 1}$, $(\gamma_i)_{i\geq 1}$ et $(V_i)_{i\geq 1}$ étant indépendantes. Le Mouvement Linéaire Multifractionnaire Multistable symétrique est alors défini pour $t \geq 0$ par

$$Y_t = C_{\alpha(t)} \sum_{i,j=1}^{\infty} \gamma_i \Gamma_i^{-1/\alpha(t)} \left(|t - V_i|^{H(t)-1/\alpha(t)} - |V_i|^{H(t)-1/\alpha(t)}\right) \left(\frac{\pi^2 j^2}{3}\right)^{1/\alpha(t)} \mathbf{1}_{[-j,-j+1[\cup[j-1,j[}(V_i).$$

**Représentation de Poisson des processus symétriques multistables** La représentation en somme de Poisson des processus stables donne également une représentation des processus multistables, qui correspond en réalité à la définition originelle des processus multistables par Falconer and Lévy Véhel (2009). Pour une fonction $\alpha : \mathbb{R}_+ \to (0,2)$, un processus multistable de fonction de stabilité $\alpha$ est un processus de la forme

$$X_t = \sum_{(X,Y)\in\Pi} f_t(X) Y^{<-1/\alpha(t)>}, \quad t \geq 0 \tag{1.3}$$

où $\Pi$ est un processus de Poisson sur $E \times \mathbb{R}$ d'intensité $m \times \mathcal{L}$. De part le calcul de ses lois finies-dimensionnelles, on sait que cette représentation conduit au même processus multistable que celui défini par la représentation en séries de Ferguson-Klass-LePage.

**Les mesures multistables** Inspirée de la construction des intégrales par rapport à une mesure aléatoire stable, une autre manière de définir des processus multistables est utilisée dans Falconer and Liu (2012). La mesure aléatoire multistable $M_{\alpha(.)}(A)$ d'un ensemble $A \in \mathcal{E}_0$ y est définie par la fonction caractéristique

$$\mathbb{E}[e^{i\theta M_{\alpha(.)}(A)}] = \exp\left(-\int_E |\theta \mathbf{1}_A(x)|^{\alpha(x)} m(dx)\right)$$

pour $\theta \in \mathbb{R}$. On écrira $M_{\alpha(.)}(A) = \int_E \mathbf{1}_A(x) M_{\alpha(.)}(dx)$, la loi de $\int_E f(x) M_{\alpha(.)}(dx)$ étant alors

$$\mathbb{E}[e^{i\theta \int_E f(x) M_{\alpha(.)}(dx)}] = \exp\left(-\int_E |\theta f(x)|^{\alpha(x)} m(dx)\right).$$

En particulier cette construction donne naissance à une autre extension du processus de Lévy symétrique $\alpha$-stable, notée $L^I$, différente de celle issue des représentations précédentes, notamment car $L^I$ est un processus à accroissements indépendants contrairement au processus $L$ défini par (1.2) . On montre dans [A5] que $L^I$ admet également une représentation sous la forme de série de Ferguson-Klass-LePage avec l'expression pour $t \in [0,1]$

$$L_t^I = \sum_{i=1}^{\infty} C_{\alpha(U_i)} \gamma_i \Gamma_i^{-1/\alpha(U_i)} \mathbf{1}_{U_i \leq t}. \tag{1.4}$$

Le lien (ou la différence) avec le processus de Lévy multistable $L$ défini par (1.2) fait l'objet du Theorem 8 de [A5], lien qui sera exploité pour le calcul du spectre d'analyse multifractal de ces processus (voir Chapitre 2).

**Théorème 1.2** ([A5]) Si $\alpha$ est de classe $\mathcal{C}^1$ alors presque sûrement, pour tout $t$,

$$L_t = L_t^I + A_t$$

avec

$$A_t = \int_0^t \sum_{i=1}^{+\infty} \gamma_i \frac{d(C_{\alpha(u)} \Gamma_i^{-1/\alpha(u)})}{du}(s) \mathbf{1}_{[0,s)}(U_i) ds.$$

# Chapitre 2

# Etudes de régularités

Cela fait longtemps que l'on s'intéresse à la régularité des trajectoires de processus stochastiques. On trouvera par exemple dans Lévy (1948) des résultats fondamentaux sur le mouvement Brownien, dont en particulier la loi du log-itéré dûe à A. Khintchine en 1924 qui indique que presque sûrement

$$\limsup_{t\to+\infty} \frac{B_t}{\sqrt{2t\log\log t}} = 1.$$

Cette loi du log-itéré inspirera ensuite bon nombre de résultats du même type. Chung (1948) détermine une loi du log-itéré pour le maximum du mouvement Brownien, loi qui sera ensuite obtenue par Monrad and Rootzén (1995) pour le mouvement Brownien fractionnaire, puis étendue au cas du mouvement Brownien bifractionnaire par Tudor and Xiao (2007) : on a ainsi presque sûrement

$$\liminf_{r\to 0^+} \frac{\max_{t\in[0,r]} |B_H(t)|}{r^H/(\log\log 1/r)^H} = c \in (0,+\infty).$$

Plus largement on trouvera des résultats sur la continuité, la régularité et le caractère borné des trajectoires de processus gaussiens dans Berman (1970, 1972), pour les processus infiniment divisibles avec composante brownienne dans Cambanis et al. (1990), sans composante brownienne dans Marcus and Rosinski (2005) ou encore pour les processus stables symétriques dans Nolan (1989a) pour ne citer qu'eux. L'étude de la régularité des trajectoires reste un domaine encore très actif, avec des développements récents comme dans Richard (2014) par exemple qui étudie la régularité de champs Browniens fractionnaires, ou encore dans Hannebicque (2021), où la notion de régularité est étudiée pour un processus $(X(t))_{t\in\mathcal{T}}$ indéxé par un ensemble général $\mathcal{T}$.

On sait également depuis longtemps que le long d'une trajectoire donnée d'un mouvement Brownien, certains points ne vérifient pas la loi du log-itérée. L'existence de points lents pour le mouvement Brownien, i.e. de points ne vérifiant pas la loi du log-itéré car $\limsup_{h\to 0} |B_{t+h} - B_t|/\sqrt{|h|} < +\infty$, est obtenue dans Kahane (1974), ces points pouvant aussi être éventuellement des zéros du processus (Kahane, 1976). L'existence de points rapides, i.e. tels que $\limsup_{h\to 0} |B_{t+h} - B_t|/\sqrt{2h \log\log 1/h} > 1$, est démontrée dans Orey and Taylor (1974). Ces résultats seront ensuite généralisés pour le mouvement Brownien fractionnaire par Khoshnevisan and Shi (2000) avec notamment l'égalité presque sûre $\limsup_{h\to 0^+} \sup_{0\le t\le 1} |B_H(t+h) - B_H(t)|/\sqrt{2h^H \log(1/h)} = 1$.

De manière générale le comportement typique d'un processus stochastique n'est donc pas exactement identique en chaque point, ce qui conduit naturellement à la notion de régularité locale d'un processus.

# 1 Plusieurs notions de régularité à travers divers exposants

### L'exposant de Hölder ponctuel

Une première caractérisation de la régularité locale se fait classiquement à l'aide de l'exposant de Hölder ponctuel. On considère $s > 0$, $t_0 \in \mathbb{R}$ et on dit qu'une fonction réelle $f \in \mathcal{C}^s(t_0)$ s'il existe un polynôme $P$ de degré inférieur à $s$ tel que

$$|f(t) - P(t - t_0)| \le C|t - t_0|^s.$$

L'exposant de Hölder ponctuel de $f$ en $t_0$ est alors défini par

$$h(t_0) = \sup\{s : f \in \mathcal{C}^s(t_0)\}.$$

Andersson (1997) donne une caractérisation des fonctions $h$ qui peuvent être des exposants de Hölder ponctuels (ce sont les lim inf de fonctions continues) ainsi que des espaces de Hölder $\mathcal{C}^s(t_0)$. L'exposant de Hölder ponctuel pour un processus $\alpha$-stable symétrique est obtenu dans Fristedt (1979) et Davis (1984a), et pour un processus de Lévy en général sans composante brownienne dans Pruitt (1981). On sait que l'on a $h(t) = 1/\beta$ presque partout avec $\beta$ correspondant à l'indice de Blumenthal and Getoor (1961). Takashima (1989) étend ces résultats au cas du mouvement linéaire fractionnaire stable (LFSM) : on sait que presque sûrement et pour tout $t$, $h(t) = H - 1/\alpha$ si $1 < \alpha < 2$ et $1/\alpha < H < 1$.

On obtient des résultats analogues pour les processus de Lévy symétriques multistables dans [A3]. La régularité typique du processus est reliée dans ce cas à la régularité de la fonction $\alpha$ elle-même. Ce phénomène se produit également dans le cas du mouvement Brownien multifractionnaire (Herbin, 2006). On note $h_\alpha(t)$ l'exposant de Hölder de $\alpha$ au point $t$.

**Théorème 2.1** ([A3]) Soit $t \in (0, 1)$.

1. Si $0 < \alpha(t) < 1$, presque sûrement

$$h(t) = \min\left(\frac{1}{\alpha(t)}, h_\alpha(t)\right)$$

au moins lorsque $\frac{1}{\alpha(t)} \neq h_\alpha(t)$.

2. Si $1 \leq \alpha(t) < 2$ et si $\alpha$ est de classe $\mathcal{C}^1$, presque sûrement
$$h(t) = \frac{1}{\alpha(t)}.$$

Ces résultats sont valables pour tout $t$, presque sûrement. On dispose en réalité d'une caractérisation plus fine de l'exposant de Hölder ponctuel pour un processus de Lévy multistable, avec la caractérisation presque sûre obtenue dans [A8].

**Théorème 2.2** ([A8]) Soit $L$ un processus de Lévy multistable symétrique avec $\alpha : [0,1] \mapsto (1,2)$ de classe $\mathcal{C}^1$. Alors, presque sûrement, pour tout $t \in [0,1]\backslash S$,
$$h(t) = \frac{\delta(t)}{\alpha(t)} = \frac{1}{\alpha(t)} \inf_{(V_{\phi(n)})_n \in \mathcal{R}_t} \liminf_{i\to\infty} -\frac{\log \phi(i)}{\log |V_{\phi(i)} - t|},$$
où $S$ est l'ensemble des points de saut du processus, $\mathcal{S} = S^{\mathbb{N}}$ et $\mathcal{R}_t = \{(r_n)_{n\in\mathbb{N}} \in \mathcal{S} : r_n \to t\}$.

### L'exposant de Hölder local

Seuret and Lévy Véhel (2002) définissent un autre exposant de régularité possédant la particularité d'être stable sous l'action d'opérateurs pseudo-différentiables contrairement à l'exposant de Hölder ponctuel . Pour un ouvert $U$, on dit que $f \in \mathcal{C}^s(U)$ avec $0 < s < 1$ s'il existe $C > 0$ tel que pour tout $x, y \in U$,
$$|f(x) - f(y)| \leq C|x-y|^s,$$
et pour $m < s < m+1$, $|\partial^m f(x) - \partial^m f(y)| \leq C|x-y|^{s-m}$. On définit ensuite
$$\alpha_l(U) = \sup\{s : f \in \mathcal{C}^s(U)\}.$$
Considérant à présent une suite décroissante d'ouverts $(U_n)_n$ telle que $\{t_0\} = \cap_n U_n$, l'exposant de Hölder local $h_l$ de la fonction $f$ au point $t_0$ est défini par
$$h_l(t_0) = \sup_n \alpha_l(U_n) = \lim_{n\to+\infty} \alpha_l(U_n).$$

Les exposants de Hölder $h$ et $h_l$ sont deux exposants distincts. Pour la fonction de type Cusp $x \mapsto |x|^\gamma$, on a $h(t) = h_l(t) = \gamma$ mais l'indice local permet de détecter un comportement de type Chirp. En effet pour la fonction $x \mapsto |x|^\gamma \sin(\frac{1}{|x|^\beta})$, on a également $h(t) = \gamma$ mais par contre $h_l(t) = \frac{\gamma}{1+\beta}$. Les liens entre les deux exposants sont explorés dans Seuret and Lévy Véhel (2002). De manière générale, on a toujours
$$h_l(t_0) \leq h(t_0)$$

et même $h_l(t_0) \leq h(t_0) \wedge \liminf_{t \to t_0} h(t)$ si $f$ est une fonction continue.

Pour un processus gaussien continue, $\alpha_l(U)$ et $h_l$ ne sont généralement pas aléatoires alors que la fonction $h$ peut l'être (voir Herbin and Lévy-Véhel (2009) ou Ayache (2011) où il est justement construit un processus gaussien possédant un exposant ponctuel aléatoire). Cependant pour les mouvements Brownien (multi)-fractionnaires, ce n'est pas le cas en général. On trouvera des propriétés trajectorielles du mouvement Brownien fractionnaire dans Adler (1981), Embrechts and Maejima (2002), Falconer (1990) ou encore dans Samorodnitsky and Taqqu (1994). Pour sa version multifractionnaire, si $\max_{t \in U} H(t) < \alpha_l(U)$, on sait que pour tout $t \in U$, $h(t) = H(t)$ presque sûrement tandis que si $\max_{t \in U} H(t) \geq \alpha_l(U)$ et $h_H(t) \neq H(t)$ (où $h_H$ est l'exposant de Hölder ponctuel de la fonction $H$), alors presque sûrement $h(t) = H(t) \wedge h_H(t)$.

Le cas des processus stables est considéré de manière générale dans Xiao (2009) ou Panigrahi et al. (2021) à travers l'étude du module de continuité uniforme. Dans le cas du mouvement linéraire fractionnaire stable (LFSM), la régularité Höldérienne des trajectoires est établie dans Takashima (1989) et Kôno and Maejima (1991) pour $\alpha \in (1, 2)$ et $H \in (\frac{1}{\alpha}, 1)$. On sait également que presque sûrement, pour tout $t$,

$$h_l(t) = H - \frac{1}{\alpha}.$$

Ayache and Hamonier (2013) montreront ensuite pour le mouvement linéaire multifractionnaire stable (LMSM) que si $h_{l,H}(t) > 1/\alpha$ (où $h_{l,H}$ est l'exposant de Hölder local de la fonction $H$), alors $h_l(t) = H(t) - \frac{1}{\alpha}$ (un encadrement était déjà fourni dans Stoev and Taqqu (2004a)). Ils établissent dans Ayache and Hamonier (2017) que le processus appartient à un espace de Hölder critique d'exposant $h_H \wedge (H_* - 1/\alpha)$ avec $H_* := \inf_{t \in U} H(t) \in (1/\alpha, 1)$.

Enfin pour la version multistable de ce processus, Biermé and Lacaux (2013) obtiennent une borne supérieure pour le module de continuité du mouvement linéaire multifractionnaire multistable (LMMM), améliorant également les résultats de Ayache and Hamonier (2013) dans le cas où $\alpha$ est constant. Le résultat étant que si la fonction $H$ est globalement Höldérienne d'exposant $H_* - 1/\alpha_*$ (avec $H_* > 1/\alpha_*$), alors

$$h_l(t) \geq H(t) - \frac{1}{\alpha(t)}. \tag{2.1}$$

### L'exposant 2-microlocal

Les exposants de Hölder sont contenus dans un environnement plus large, à savoir celui de l'analyse 2-microlocale introduite par Bony (1986) au travers d'une analyse en ondelettes de type Littlewood Paley, caractérisée dans le domaine temporel par Lévy Véhel and Seuret (2003) (voir aussi Kolwankar and Lévy Véhel (2002)) et développée dans un cadre aléatoire par Herbin and Lévy-Véhel (2009).

On peut ainsi définir les espaces 2-microlocaux pour une fonction $f$ non-différentiable de la manière suivante. Pour $D = \{(s, s') : 0 < s + s' < 1, s < 1, s' < 0\}$, on dit que $f \in \mathcal{C}^{s,s'}(t_0)$ avec $(s, s') \in D$ s'il existe $\delta \in \mathbb{R}$ et $C \in \mathbb{R}$ tels que pour tous $0 < |t_1 - t_0| < \delta$, $0 < |t_2 - t_0| < \delta$,

$$|f(t_1) - f(t_2)| \leq C|t_1 - t_2|^{s+s'}(|t_1 - t_0| + |t_1 - t_2|)^{-s'}.$$

On peut alors considérer le domaine 2-microlocal de $f$, $E(f, t_0) = \{(s, s') : f \in \mathcal{C}^{s,s'}(t_0)\}$ qui est un ensemble convexe. L'exposant 2-microlocal de $f$, appelé aussi frontière 2-microlocale, est en réalité la courbe décroissante concave

$$\sigma : s' \mapsto \sigma(s') = \sup\{s : f \in \mathcal{C}^{s,s'}(t_0)\}.$$

Dans le cas où $f \in \mathcal{C}^{\eta}(\mathbb{R})$ pour un $\eta > 0$, on peut déduire de la frontière 2-microlocale l'expression des deux exposants de Hölder grâce aux relations

$$h(t_0) = -\inf\{s' : \sigma(s') \geq 0\}$$

et

$$h_l(t_0) = \sigma(0).$$

L'analyse 2-microlocale permet entre autres choses de distinguer les deux fonctions de type Chirp $f_1(x) = |x|^{0.5}\sin(\frac{1}{|x|})$ et $f_2(x) = |x|^{0.8}\sin(\frac{1}{|x|^{2.2}}) + |x|^{0.5}\sin(\frac{1}{|x|^{0.25}})$, pour qui toutes les deux $h(0) = 0,5$ et $h_l(0) = 0,25$ mais qui ne présentent pas les mêmes frontières 2-microlocales. Dans le cadre aléatoire, cette analyse 2-microlocale est effectuée pour étudier la régularité de martingales et d'intégrales stochastiques continues dans Balança and Herbin (2012) par exemple, ou encore dans Balança (2014) pour les processus de Lévy et le mouvement linéraire multifractionnaire stable (LMSM).

### Régularité en loi : l'autosimilarité et l'indice de Hurst

Pour une notion de régularité plus faible, on peut considérer la loi des accroissements plutôt que leur comportement presque sûr. La notion qui apparaît alors naturellement est celle d'autosimilarité. On dit qu'un processus $X$ est autosimilaire d'indice de Hurst $H$ si pour tout $a > 0$,

$$X(at) \stackrel{d}{=} a^H X(t).$$

On sait que si $X$ est un processus non-trivial autosimilaire d'exposant $H$ alors $H > 0$ (Vervaat, 1985). Si le processus est de plus à accroissements stationnaires, alors

$$\frac{X(t+h) - X(t)}{h^H} \stackrel{d}{=} X(1),$$

et l'on a alors trivialement

$$\lim_{h \to 0} \frac{X(t+h) - X(t)}{h^H} \stackrel{d}{=} X(1).$$

La régularité globale en loi $H$ correspond bien souvent à la régularité trajectorielle presque partout de ces processus homogènes. On pourra se référer à Embrechts and Maejima (2002) pour une analyse approfondie des processus autosimilaires, leurs propriétés trajectorielles, leur simulation, ainsi que les méthodes classiques d'estimation de l'indice de Hurst $H$.

**Régularité en loi locale : l'indice de localisabilité**

La localisation de cette notion de régularité en loi conduira ensuite à celle de localisabilité, très utile en théorie et en pratique car elle permet de contrôler des caractéristiques locales d'un processus comme sa régularité höldérienne mais aussi l'intensité de ses sauts dans le cas des processus multistables. Cette notion a été introduite par Falconer (2002, 2003).

On dit qu'un processus réel $X$ est $h_d(t)$-localisable au point $t$ s'il existe un processsus limite non-trivial $X'_t = (X'_t(u))_{u\in\mathbb{R}}$ tel que

$$\lim_{r\to 0^+} \frac{X(t+ru)-X(t)}{r^{h_d(t)}} \stackrel{d}{=} X'_t(u),$$

où l'égalité est en loi finies-dimensionnelles. Si la convergence a lieu en loi par rapport à une métrique sur $C(\mathbb{R})$ ou $D(\mathbb{R})$, on parle de localisabilité forte. La limite $X'_t$ est appelée forme locale ou processus tangent de $X$ au point $t$. Cette notion est très proche de celle de processus localement asymptotiquement autosimilaire décrite dans Cohen (1999) ou encore dans Peltier and Lévy Véhel (1995); Benassi et al. (1997) pour le mouvement Brownien multifractionnaire.

Les processus localisables les plus simples sont les processus autosimilaires à accroissements stationnaires, pour qui la forme locale en $t$ est constante égale à $X$ : $X'_t(.) = X(.)$. La forme locale $X'_t$ au point $t$, si elle existe, doit elle-même être $h_d(t)$-autosimilaire, i.e. $X'_t(au) \stackrel{d}{=} a^{h_d(t)}X'_t(u)$ pour $a>0$. De plus, sous des conditions assez générales (Falconer, 2002, 2003), $X'_t$ est également à accroissements stationnaires. Les formes locales typiques sont donc des processus autosimilaires à accroissements stationnaires.

Un autre exemple est donné par la moyenne mobile symétrique $\alpha$-stable de la forme $X(t) = \int g(t-x)M(dx)$ où $M$ est une mesure $\alpha$-stable symétrique de mesure de contrôle la mesure de Lebesgue ([A1]). La condition sur le noyau

$$\lim_{r\to 0+} \int |\frac{g(r(u+x))-g(rx)}{r^\gamma} - h(u,x)|^\alpha dx = 0$$

pour $\gamma + 1/\alpha > 0$ assure que le processus $X$ a pour forme locale $X'_t(u) = \int h(u,x)M(dx)$ avec $h_d(t) = \gamma + 1/\alpha$. Ainsi le processus d'Ornstein-Uhlenbeck $X(t) = \int \exp(-\lambda(x-t))M(dx)$ admet pour forme locale $X'_t(u) = L^\alpha_u$ avec $h_d(t) = 1/\alpha$. On peut également trouver des conditions sur la transformée de Fourier du noyau $g$ pour obtenir le log-fractional stable motion comme forme locale. Par ailleurs, si $g$ est globalement höldérienne avec $g(r) \sim c_0^+ r^{-\gamma}$ et $g(-r) \sim c_0^- r^{-\gamma}$, la forme locale sera le mouvement linéaire fractionnaire stable (LFSM).

Cependant de manière générale la forme locale varie avec $t$. Dans le cas du mouvement Brownien multifractionnaire,

$$\lim_{r\to 0} \frac{X(t+ru)-X(t)}{r^{H(t)}} \stackrel{d}{=} B_{H(t)}(u)$$

où $B_H$ est le mouvement Brownien fractionnaire d'exposant $H$. Ainsi le mouvement Brownien multifractionnaire ressemble localement au point $t$ à un mouvement Brownien fractionnaire d'indice $H(t)$ qui varie avec $t$. On obtient un résultat identique en remplaçant la mesure gaussienne par une mesure $\alpha$-stable (Stoev and Taqqu, 2004b).

On retrouve de même pour des processus moyennes mobiles multistables ([A1]) que la condition

$$\lim_{r\to 0+} \int |\frac{g(r(u+x))-g(rx)}{r^{\gamma}} - h(u,x)|^{\alpha(t)} dx = 0$$

avec $0 < \gamma + 1/\alpha(t) < \eta \leq 1$ assure que le processus a pour forme locale $X'_t(u) = \int h(u,x) M_{\alpha(t)}(dx)$ avec $h_d(t) = \gamma + 1/\alpha(t)$ et $M_{\alpha(t)}$ une mesure aléatoire $\alpha(t)$-stable symétrique. Ainsi le processus d'Ornstein-Uhlenbeck multistable est tangent en chaque point $t$ à un processus de Lévy symétrique $\alpha(t)$-stable.

Dans Falconer and Lévy Véhel (2009), des processus de forme locale prescrite sont construits en collant des processus localisables dont la forme locale est connue. Ils considèrent ainsi la localisabilité d'un champ aléatoire $\{X(t,v) : (t,v) \in U \times U\}$ et de son processus diagonal $Y = \{X(t,t) : t \in U\}$, et donnent des conditions pour que $Y$ et $X(\cdot,u)$ aient la même forme locale. Cette approche permet de construire des processus localisables à partir de processus élémentaires qui sont connus pour être localisables (typiquement des processus autosimilaires à accroissements stationnaires). Cette méthode permet également de simuler les accroissements d'un processus localisable en utilisant la structure de la forme locale (voir [A1]). C'est cette idée qui est à la base de la construction des processus multistables dans Falconer and Lévy Véhel (2009), utilisant également la représentation des processus stables sous forme de somme de processus de Poisson. La localisabilité des processus moyennes mobiles multistables est abordée de manière directe dans [A1], mais cette construction est reprise dans [A2] pour proposer et étudier une version alternative des processus multistables sous forme de série de Ferguson-Klass-LePage.

Pour un processus multistable général écrit sous la forme (1.1), on obtient des conditions sur le noyau $f_t$ pour que le processus multistable $X$ soit localisable, de forme locale un processus $\alpha(t)$-stable dont l'indice de stabilité dépend ainsi du temps ([A2]). Si $\alpha$ est de classe $\mathcal{C}^1$, on sait alors que le processus de Lévy multistable a pour forme locale $X'_t = L^{\alpha(t)}$ avec $h_d(t) = 1/\alpha(t)$ et que le mouvement linéaire multifractionnaire multistable (LMMM) a pour forme locale $X'_t = Y^{\alpha(t),H(t)}$ avec $h_d(t) = H(t)$. On retrouve ainsi le mouvement linéaire fractionnaire stable (LFSM), le processus de Lévy $\alpha$-stable et le log-fractional stable motion comme formes locales des processus localisables stables et multistables. On trouvera d'autres exemples dans Falconer and Liu (2012) et [A2].

Généralement l'exposant de localisabilité $h_d$ est relié à la régularité ponctuelle de Hölder classique $h$ (voir Boufoussi et al. (2008) pour quelques conditions). Pour un processus multistable général écrit sous la forme (1.1), sous certaines conditions sur le noyau $f_t$, on a une majoration de $h$ pour tout $t$ presque sûrement par la fonction de localisabilité du processus $h_d$.

**Théorème 2.3** ([A3]) Pour $t > 0$, si $X$ est $h_d(t)$-localisable au point $t$, et sous certaines conditions sur le noyau $f_t$, on a presque sûrement

$$h(t) \leq h_d(t)$$

avec égalité dans le cas du mouvement de Lévy multistable et $\alpha$ de classe $\mathcal{C}^1$.

L'indice de régularité $h_d$ est également étroitement relié à la régularité ponctuelle du processus en espérance. En effet sous certaines conditions d'intégrabilité et de régularité du noyau $f_t$ (voir [A3]), on a pour $\eta < \alpha_*$

$$\lim_{\varepsilon \to 0+} \mathbb{E}[|\frac{X(t+\varepsilon)-X(t)}{\varepsilon^{h_d(t)}}|^\eta] = \mathbb{E}[|X'_t(1)|^\eta].$$

### Autres indices de régularité

D'autres indices ont été introduits afin de caractériser la régularité d'un processus stochastique, utilisant notamment la décomposition en ondelettes du signal. Par exemple en considérant le coefficient d'ondelette $d_{j,k} = \int f(t)\psi_{j,k}(t) = \int f(t)2^{j/2}\psi(2^j t - k)dt$ pour une ondelette mère $\psi$, on peut définir l'exposant discrétisé à base d'ondelettes

$$\omega_{k_n}^{(n)}(t) = -\frac{1}{n}\log_2 |2^{n/2} d_{n,k_n}|$$

avec $k_n(t) = \lfloor t2^n \rfloor$, et l'exposant de singularité ponctuel à base d'ondelettes

$$\omega(t) = \liminf_{n \to +\infty} \omega_{k_n}^{(n)}(t).$$

Celui-ci est alors relié à l'exposant de Hölder ponctuel $h$ dans certains cas (Jaffard, 1991). On peut également s'intéresser au coefficient d'oscillation défini dans Arneodo et al. (1998), ou encore aux exposants d'annulation et fractionnaires introduits et caractérisés dans Abry et al. (2017).

## 2 Quelques résultats statistiques

### Estimation de la régularité

La question de l'estimation de la régularité trajectorielle a fait l'objet de nombreuses études. La décomposition du signal en ondelettes fournit pour cela un outil performant en traitement du signal et a également donné lieu à beaucoup de méthodes pour l'estimation du coefficient local de Hölder d'un processus. On pourra citer en particulier la méthode des lignes maximales de la transformée continue en ondelettes, transformée introduite initialement par Grossmann and Morlet (1984) puis utilisée pour l'estimation par Mallat and Hwang (1992) (voir aussi Daubechies (1992)). Elle est également utilisée pour l'estimation du paramètre de Hurst $H$ du mouvement Brownien fractionnaire dans Flandrin (1992), Delbeke and Van Assche (1998), Abry and Veitch (1998), Goncalves et al. (1998), Veitch and Abry (1999), Bardet (2000), Bardet et al. (2000), Abry et al. (2001), Veitch and Abry (2001), Bardet (2002), Abry et al. (2003), Abry et al. (2011), Wu and Ding (2017) ou encore dans Abry et al. (1998) pour l'estimation du paramètre de longue mémoire d'un processus stationnaire.

Pour ce problème d'estimation du paramètre de longue mémoire, beaucoup d'estimateurs différents ont déjà été introduits. On pourra se référer à Bardet et al. (2003) pour une review sur le sujet, ou encore à Embrechts and Maejima (2002), mais nous pouvons citer en particulier l'estimateur de Whittle (Fox and Taqqu, 1986; Giraitis and Surgailis, 1990), l'utilisation de méthode spectrale (Flandrin, 1989; Moulines and Soulier, 2003; Beran, 1992; Robinson, 1995), de la variation quadratique (Istas and Lang, 1997; Coeurjolly, 2001) ou encore de l'estimateur du log-periodogramme (Geweke and Porter-Hudak, 1983; Hurvich et al., 1998; Lang and Azaïs, 1999; Moulines and Soulier, 1999).

Concernant les processus multifractionnaires et le mouvement Brownien multifractionnaire en particulier, de nombreuses références existent aussi, parmi lesquelles Ayache and Lévy Véhel (2004); Ayache et al. (2004), Begyn (2007a,b), Istas and Lang (1997), Benassi et al. (1998, 2000), Lacaux (2004). Bardet and Surgailis (2012) présentent en particulier deux estimateurs, inspirés respectivement de Coeurjolly (2005) et Bardet and Surgailis (2011), et montrent la consistance de ces estimateurs pour des processus gaussiens généraux. Ils obtiennent également la convergence presque sûre pour l'estimateur de $H_*$ avec une vitesse en probabilité qui est logarithmique. Une meilleure vitesse de convergence a été obtenue plus récemment par Lebovits and Podolskij (2017) pour ce problème d'estimation de la régularité globale d'un mouvement Brownien multifractionnaire.

C'est dans ce contexte que l'on a considéré dans [A4] le problème d'estimation de régularité pour les processus multistables, la régularité étant celle en loi. On considère $\alpha$ de classe $\mathcal{C}^1$ à valeurs dans $[\alpha_*, \alpha^*] \subset (0, 2)$ et l'on suppose observée une discrétisation du processus $(X(\frac{k}{N}))_{k\in\mathbb{Z},N\in\mathbb{N}}$. N'ayant pas de connaissance de la valeur de $\alpha_*$, une méthode des moments ne semble a priori pas adaptée pour ce problème d'estimation. On opte alors pour une utilisation du logarithme des accroissements qui possède des moments de tous ordres, ainsi qu'une utilisation de la structure locale des processus multistables. Partant du principe que pour $k/N$ proche de $t_0 \in \mathbb{R}$ fixé, $|X(\frac{k+1}{N}) - X(\frac{k}{N})| \overset{d}{\approx} N^{-h_d(t_0)}|X'_{t_0}(1)|$, on introduit un estimateur de $h_d(t_0)$ avec

$$\hat{h}_d(t_0) = -\frac{1}{n(N)\log N} \sum_{k=\lfloor Nt_0 \rfloor - \frac{n(N)}{2}}^{\lfloor Nt_0 \rfloor + \frac{n(N)}{2}+1} \log |X(\frac{k+1}{N}) - X(\frac{k}{N})|$$

où $(n(N))_{N\in\mathbb{N}}$ est une suite d'entiers. On suppose ensuite certaines conditions sur le noyau $f_t$ pour $t$ dans un ouvert $U$, conditions qui sont en particulier vérifiées pour les processus de Lévy multistables et pour le mouvement linéaire multifractionnaire multistable (LMMM). On a alors pour tout $r > 0$ et $t_0 \in U$, sous les conditions $\lim\limits_{N\to+\infty} n(N) = +\infty$ et $\lim\limits_{N\to+\infty} \frac{n(N)}{N} = 0$,

$$\lim_{N\to+\infty} \mathbb{E}[|\hat{h}_d(t_0) - h_d(t_0)|^r] = 0$$

et pour $[a, b] \subset U$,

$$\lim_{N\to+\infty} \mathbb{E}[\int_a^b |\hat{h}_d(t) - h_d(t)|^r dt] = 0.$$

Concernant la vitesse de convergence, on a un résultat supplémentaire pour le mouvement de Lévy multistable symétrique. En effet, si $\alpha \in (1, 2)$ est de classe $\mathcal{C}^1$, pour $n(N) = O(N^\delta)$ avec $\delta \in (0, \frac{2\alpha(t_0)-2}{3\alpha(t_0)+2})$, on obtient

$$\lim_{N\to+\infty} \sqrt{n(N)}(\log N(\hat{h}_d(t_0) - h_d(t_0)) + \mu_{\alpha(t_0)}) \stackrel{d}{=} \mathcal{N}(0, \sigma^2_{\alpha(t_0)}) \tag{2.2}$$

avec $\mu_\alpha = \mathbb{E}[\log |Z|]$ et $\sigma^2_\alpha = \mathbf{Var}(\log |Z|)$ pour $Z$ de loi $\alpha$-stable standard. Les expressions de $\mu_\alpha$ et $\sigma^2_\alpha$ étant explicites, on obtient de plus un intervalle de confiance de niveau asymptotique $1 - \beta$ avec

$$\left[\frac{\hat{h}_d(t_0)\log N + \Gamma'(1)}{\Gamma'(1) + \log N} \pm \frac{q_{1-\beta/2}\sqrt{\sigma^2_{\hat{h}_d(t_0)^{-1}}}}{(\Gamma'(1) + \log N)\sqrt{n(N)}}\right]$$

où $q_\beta$ est le $\beta$-quantile de la loi normale $\mathcal{N}(0, 1)$.

### Estimation de l'intensité de sauts

La régularité d'un processus à sauts est intimement reliée à son intensité de sauts. Dans le cas des processus multistables, cette intensité est complètement caractérisée par la fonction de stabilité $\alpha$, qui gouverne également la queue de distribution des marginales ou encore l'ordre de ses moments. L'estimation de cette fonction relève ainsi une importance particulière dans le cas des processus multistables.

De manière non-paramétrique, l'estimation de la queue de distribution d'une loi quelconque est abordée par Hill (1975), ainsi que par Gonçalves and Riedi (2005) qui proposent d'estimer $\lambda^+ = \sup\{r > 0 : \mathbb{E}[|X|^r] < +\infty\}$ dans le cas où $\lambda^+ \in (0, 2)$ à l'aide d'un estimateur de la fonction caractéristique et des ondelettes. Pour des processus de Lévy ou des semi-martingales à sauts, poursuivant les travaux de Aït-Sahalia and Jacod (2009a), Jing et al. (2012) et Todorov (2015), Mies (2020) introduit un estimateur de l'indice de Blumenthal et Getoor $\beta$ ainsi qu'un estimateur de la volatilité. Cet estimateur est de vitesse de convergence optimale, améliorant ainsi les résultats de Reiß (2013) pour les processus de Lévy et de Bull (2016) pour les semi-martingales.

Le problème devient paramétrique dans le cas des lois $\alpha$-stables et plusieurs estimateurs du paramètre $\alpha$ ont déjà été proposés. Ainsi un estimateur qui utilise la queue de distribution est proposé dans Koutrouvelis (1980), la fonction caractéristique est utilisée dans McCulloch (1986, 1997), l'estimateur du maximum de vraisemblance est considéré dans Nolan (2001) et Brouste and Masuda (2018), et on retrouve une estimation des paramètres d'une loi stable de manière combinée dans Lévy Véhel et al. (2021).

Pour les processus multistables, on s'est intéressé à la question de l'estimation de la fonction de stabilité $\alpha$ dans [A4]. On observe à nouveau une discrétisation du processus $(X(\frac{k}{N}))_{k\in\mathbb{Z},N\in\mathbb{N}}$. L'idée cette fois-ci est d'utiliser le fait que pour une loi stable $Z$ de paramètre $\alpha \in (0,2)$, $\mathbb{E}[|Z|^p] < +\infty$ si et seulement si $p < \alpha$, l'expression de $\mathbb{E}[|Z|^p]$ étant explicite en fonction de $p$ et $\alpha$ (voir Samorodnitsky and Taqqu (1994)) :

$$\mathbb{E}[|Z|^p] = \frac{2^{p-1}\Gamma(1-p/\alpha)}{p\int_0^\infty u^{-p-1}\sin^2(u)du}.$$

On introduit donc pour $t_0 \in \mathbb{R}$ les moments empiriques

$$S_N(p) = \left(\frac{1}{n(N)}\sum_{k=\lfloor Nt_0\rfloor-\frac{n(N)}{2}}^{\lfloor Nt_0\rfloor+\frac{n(N)}{2}+1} |X(\frac{k+1}{N}) - X(\frac{k}{N})|^p\right)^{1/p}$$

ainsi que les ratios de ces moments empiriques afin d'éviter tout problème d'échelle (inconnue) : pour $p_0 \in (0, \alpha_*)$, on pose

$$R_{\exp}(p) = \frac{S_N(p_0)}{S_N(p)} \quad \text{et} \quad R_\alpha(p) = \frac{\mathbb{E}[|Z|^{p_0}]^{1/p_0}}{\mathbb{E}[|Z|^p]^{1/p}}\mathbf{1}_{p<\alpha}.$$

On estime alors $\alpha(t_0)$ par

$$\hat{\alpha}(t_0) = \min\left(\operatorname*{arg\,min}_{\alpha\in[0,2]}(\int_{p_0}^2 |R_{\exp}(p) - R_\alpha(p)|^\gamma dp)^{1/\gamma}\right)$$

où $\gamma \in (0,1)$. Cet estimateur nécessite la connaissance au préalable de $\alpha_*$, ce qui n'est pas toujours réaliste. Cependant dans certains cas comme celui du mouvement linéaire multifractionnaire multistable (LMMM), on sait que si les trajectoires sont continues alors $\alpha \in (1,2)$. On obtient sous certaines conditions sur le noyau $f_t$,

$$\lim_{N\to+\infty} \mathbb{E}[|\hat{\alpha}(t_0) - \alpha(t_0)|^r] = 0$$

pour tout $r > 0$ à partir du moment où le processus $X$ devient autosimilaire à accroissements stationnaires si $\alpha$ est constant. Ces résultats s'appliquent ainsi aux cas du mouvement linéaire multifractionnaire multistable (LMMM) et du processus de Lévy multistable.

Avant d'estimer la fonction ou le paramètre de stabilité $\alpha$, il est préférable de savoir si le modèle présente une intensité de saut homogène ou non, afin de déterminer à l'avance si l'estimation sera paramétrique ou non-paramétrique et utiliser l'estimateur adéquat. C'est la question considérée dans [A7] pour les processus de Lévy multistables avec $\alpha \in (1,2)$ de classe $\mathcal{C}^1$. On y considère les deux versions du processus de Lévy multistable $X = L$ (voir (1.2)) ou $X = L^I$ (voir (1.4)), et l'on souhaite tester

$$(H_0) : \alpha \in \{\alpha : [a,b] \to [\alpha_*, \alpha^*] : \alpha(t_1) = \alpha(t_2), \forall t_1, t_2\}$$

vs

$$(H_1) : \alpha \in \{\alpha : [a,b] \to [\alpha_*, \alpha^*] : \exists t_1, t_2, \alpha(t_1) \neq \alpha(t_2) \text{ et } \|\alpha'\|_\infty \leq M\}$$

en observant une discrétisation $(X(\frac{k}{N}))_{k\in\mathbb{Z},N\in\mathbb{N}}$. Le test est basé sur l'estimation de la fonction de localisabilité $h_d$ avec $h_d(t) = \frac{1}{\alpha(t)}$ pour le processus de Lévy multistable. On pose ainsi

$$\hat{h}_d(t_0) = \frac{1}{n(N)} \sum_{k=\lfloor Nt_0\rfloor - \frac{n(N)}{2}}^{\lfloor Nt_0\rfloor + \frac{n(N)}{2}+1} \frac{\Gamma'(1) - \log|X(\frac{k+1}{N}) - X(\frac{k}{N})|}{\Gamma'(1) + \log N}$$

et on définit la statistique de test pour $(t_0, t_1) \in [a,b]^2$ tels que $(t_0 - \frac{n(N)}{2N}, t_0 + \frac{n(N)}{2N}) \cap (t_1 - \frac{n(N)}{2N}, t_1 + \frac{n(N)}{2N}) = \varnothing$ par

$$T_N = \frac{n(N)(\Gamma'(1) + \log N)^2}{(b-a)\hat{\sigma}^2(t_1)} \int_a^b |\hat{h}_d(t) - \hat{h}_d(t_0)|^2 dt$$

avec $\hat{\sigma}^2(t_1) = \Big(v(\hat{h}_d(t_1)^{-1}) \vee v(2)\Big) \wedge v(1)$ et $v(x) = \frac{\pi^2}{6x^2} + \frac{\pi^2}{12}$. $v(\alpha)$ étant la variance asymptotique $\sigma_\alpha^2$ de (2.2) si $\alpha \in (1,2)$, $\hat{\sigma}^2(t_1)$ est un estimateur de $\sigma^2_{\alpha(t_1)}$ qui devient $\sigma_\alpha^2$ sous $(H_0)$. La normalité asymptotique de $\hat{h}_d$ donne la convergence de $T_N$ vers la loi de $1 + \chi^2(1)$ où $\chi^2(1)$ suit une loi du $\chi^2$ à 1 degré de liberté. La zone de rejet du test est ainsi de la forme

$$R_c = [q_{1-\beta}, +\infty)$$

où $q_\beta$ est le $\beta$-quantile de la loi $1+\chi^2(1)$. Posons enfin $\xi_N = \sqrt{n(N)}(\Gamma'(1)+\log N)$ et la distance de Kolmogorov $\kappa(\mu_1, \mu_2) = \sup_{x\in\mathbb{R}} |F_{\mu_1}(x) - F_{\mu_2}(x)|$ où $F_\mu$ est la fonction de répartition de $\mu$. On obtient un contrôle uniforme sous $(H_0)$ de la distance de Kolmogorov entre $T_N$ et $1+\chi^2(1)$.

**Théorème 2.4** ([A7]) Soit $P_N^\alpha$ la loi de $T_N$ et $P_{1+\chi^2}$ la loi de $1+\chi^2(1)$. Il existe une constante $K^* > 0$ telle que pour tout $N \geq 1$,

$$\sup_{\alpha\in\Theta_0} \kappa(P_N^\alpha, P_{1+\chi^2}) \leq K^*(\frac{1}{\xi_N} + \frac{1}{\sqrt{n(N)}} + \left(\sqrt{\frac{n(N)}{N}} + \frac{1}{\xi_N} + \frac{1}{\sqrt{\lfloor Nb\rfloor - \lfloor Na\rfloor - n(N)}}\right)^{1/3}).$$

Par conséquent il existe une constante $K^* > 0$ telle que pour chaque niveau fixé $\beta \in (0,1)$ dans le problème de test et chaque $N \geq 1$,

$$\sup_{\alpha\in(H_0)} \mathbb{P}(T_N \in R_c) \leq \beta + K^*\left(\frac{1}{\xi_N} + \frac{1}{\sqrt{n(N)}} + \left(\sqrt{\frac{n(N)}{N}} + \frac{1}{\xi_N} + \frac{1}{\sqrt{\lfloor Nb\rfloor - \lfloor Na\rfloor - n(N)}}\right)^{1/3}\right).$$

De plus, si $\lim\limits_{N\to+\infty} n(N) = +\infty$ et $\lim\limits_{N\to+\infty} \frac{n(N)}{N} = 0$, alors pour tout $\alpha \in (H_0)$ on a la convergence en loi suivante :

$$\lim_{N\to+\infty} T_N \stackrel{d}{=} 1 + \chi^2(1).$$

Sous l'alternative $(H_1)$, on établit une convergence en loi dans le cas $X = L$ ainsi qu'un contrôle de la vitesse de convergence de $T_N$ vers $+\infty$ dans le cas $X = L^I$.

**Théorème 2.5** ([A7]) On suppose que l'on observe $X = L$. On définit

$$Z_N := \frac{1}{\xi_N}\left(T_N - \frac{\xi_N^2}{\hat{\sigma}^2(t_1)}\frac{1}{b-a}\int_a^b |h_d(t) - h_d(t_0)|^2 dt\right)$$

et on considère $Z$, une variable aléatoire normalement distribuée $\mathcal{N}(0,\sigma_Z^2)$ avec

$$\sigma_Z^2 = \frac{4\sigma^2_{\alpha(t_0)}}{(b-a)^2\sigma^4_{\alpha(t_1)}}\left(\int_a^b (h_d(t) - h_d(t_0))dt\right)^2.$$

Soit $P_{Z_N}$ la loi de $Z_N$, $P_Z$ la loi de $Z$ et $\eta \in (0, \frac{1}{2})$. Il existe une constante positive $K^*_\eta$ telle que pour tout $\alpha \in (H_1)$ et $N \geq 1$,

$$\kappa(P_{Z_N}, P_Z) \leq K^*_\eta\left(\frac{1}{\sqrt{n(N)}} + \frac{1}{\sqrt{\sigma_Z}}\sqrt{\frac{\xi_N}{N^{(1-\frac{1}{\alpha_*})\frac{\eta}{2}}} + \xi_N\frac{n(N)}{N} + (\frac{n(N)}{N})^{\frac{1}{2}} + \frac{1}{\xi_N}}\right).$$

Par conséquent si $n(N) = N^\gamma$ avec $\gamma \in (0, \frac{\alpha_*-1}{4\alpha_*})$, alors pour tout $\alpha \in (H_1)$,

$$\lim_{N\to+\infty} Z_N \stackrel{d}{=} Z$$

et pour un niveau fixé $\beta \in (0,1)$ du problème de test, il existe une constante $K^*_\gamma > 0$ telle que pour tout $\alpha \in (H_1)$ et $N \geq 1$,

$$\mathbb{P}(T_N \in R_c) \geq$$

$$\mathbb{P}\left(Z \leq \frac{\xi_N}{3(b-a)}\int_a^b |h_d(t) - h_d(t_0)|^2 dt - \frac{q_\beta}{\xi_N}\right) - K^*_\gamma \frac{\sqrt{\log N}}{N^{\gamma/4}}\left(\frac{1}{b-a}\left|\int_a^b (h_d(t) - h_d(t_0))dt\right|\right)^{-\frac{1}{2}}.$$

Si $h_d(t_0) = \frac{1}{b-a}\int_a^b h_d(t)dt$, $\lim\limits_{N\to+\infty} Z_N \overset{\mathbb{P}}{=} 0$ où la convergence est en probabilité. En fait on a

$$\lim_{N\to+\infty} T_N - \frac{Z_N^2}{\hat{\sigma}^2(t_1)}\frac{1}{b-a}\int\limits_a^b |h_d(t)-h_d(t_0)|^2 dt \overset{d}{=} \frac{1}{b-a}\int_a^b \frac{\sigma^2_{\alpha(t)}}{\sigma^2_{\alpha(t_1)}}dt + \frac{\sigma^2_{\alpha(t_0)}}{\sigma^2_{\alpha(t_1)}}\chi^2(1),$$

et on peut obtenir une borne similaire pour la distance de Kolmogorov entre les deux lois. En ce qui concerne le processus de Lévy multistable $L^I$, l'indépendance des accroissements permet d'utiliser des bornes de type Chernov, donnant des taux de convergence exponentiels.

**Théorème 2.6** ([A7]) On suppose que l'on observe $X = L^I$, que $(\frac{n(N)\log N}{N})_N$ est bornée et $n(N) \geq 256$. Alors il existe une constante $K^* > 0$ telle que pour tout $\alpha \in (H_1)$ et $x > 0$,

$$\mathbb{P}(T_N \leq x) \leq e^{K^*\sqrt{x}}(K^*)^{\sqrt{n(N)}} e^{-\frac{\xi_N}{b-a}\int\limits_a^b |h_d(t)-h_d(t_0)|dt}.$$

Par conséquent pour un niveau fixé $\beta \in (0,1)$ dans le problème de test, il existe une constante $K_\beta > 0$ telle que pour tout $\alpha \in (H_1)$,

$$\mathbb{P}(T_N \in R_c) \geq 1 - K_\beta^{\sqrt{n(N)}} e^{-\frac{\xi_N}{b-a}\int\limits_a^b |h_d(t)-h_d(t_0)|dt}.$$

L'indépendance des accroissements du processus de Lévy multistable $L^I$ permet ainsi d'obtenir des puissances de test plus grandes que dans le cas du processus $L$ qui possède seulement une indépendance asymptotique, au sens où par exemple $\hat{h}_d(t_0)$ et $\hat{h}_d(t_1)$ sont asymptotiquement indépendants pour $L$ et indépendants à partir d'un certain rang pour $L^I$.

# 3 Dimensions fractales

Les caractéristiques principales d'un processus (multi)fractionnaire en tant que fonction aléatoire, telles que ses lignes de niveau, son ensemble image ou encore son graphe font naturellement apparaître des ensembles fractals, dont on mesure généralement la taille grâce aux mesures de Hausdorff et la dimension qu'elles induisent. On rappelle ici la définition de cette notion introduite par Hausdorff en 1919 (voir Falconer (1990)).

On définit la mesure de Hausdorff $s$-dimensionnelle d'un ensemble $F$ par

$$\mathcal{H}^s(F) = \lim_{\delta\to 0}\mathcal{H}^s_\delta(F)$$

où

$$\mathcal{H}^s_\delta(F) = \inf\left\{\sum_{i=1}^\infty |U_i|^s : F \subset \bigcup_{i=1}^\infty U_i \text{ avec } 0 < |U_i| \leq \delta \text{ pour chaque } i\right\}.$$

La dimension de Hausdorff de $F$ (ou dimension de Hausdorff-Besicovitch) est alors

$$\dim_H(F) = \inf\{s : \mathcal{H}^s(F) = 0\} = \sup\{s : \mathcal{H}^s(F) = \infty\}.$$

Le calcul de la dimension de Hausdorff d'ensembles issus de processus fractionnaires commencent une nouvelle fois avec le mouvement Brownien, et en particulier avec Itô and McKean (1964) qui obtiennent la dimension de Hausdorff des lignes de niveau $\{t : B_t = 0\}$ pour le mouvement Brownien mais aussi plus largement pour des processus de diffusion. Les points lents et rapides du mouvement Browniens sont également mesurés dans Perkins (1983) (pour les points lents), et dans Orey and Taylor (1974), Khoshnevisan and Shi (2000) et Khoshnevisan et al. (2000) (pour les points rapides). On pourra se référer à Xiao (2004) et Khoshnevisan and Xiao (2004) pour un bon panorama sur les questions de calculs de dimension et de régularité dans le cas des processus de Markov, des processus de Lévy et des processus additifs.

### Dimension de l'image

De manière générale on sait qu'un processus de Lévy envoie un borélien $G$ sur un ensemble fractal aléatoire $X(G)$. En effet pour un processus de Lévy dans $\mathbb{R}^d$ d'exposant $\psi$, Pruitt (1969) donne la formule pour la dimension de l'image de l'intervalle $[0,1]$

$$\dim_H X([0,1]) = \sup\{\alpha \geq 0 : \limsup_{r\to 0} r^{-\alpha} \int_0^1 \mathbb{P}(|X(t)| \leq r)dt < \infty\}.$$

Pour un ensemble quelconque $E$, Blumenthal and Getoor (1961) obtient une majoration de $\dim_H X(E)$ dans le cas $\beta < 1$ et Millar (1971) obtient la même majoration pour $\beta$ quelconque. Le cas général sera établi plus tard : Khoshnevisan and Xiao (2005) caractérisent $\dim_H X(E)$ pour un Lévy en général avec la formule presque sûre

$$\dim_H X(E) = \sup\{\beta \in (0,1) : \inf_{\mu\in\mathcal{P}(E)} \int_{\mathbb{R}} \mathbb{E}[|e^{i\xi X(t)}\mu(dt)|^2]\xi^{\beta-1}d\xi < +\infty\}.$$

On dispose également de formules analogues pour un processus additif (Khoshnevisan et al., 2003) :

$$\gamma = \sup\{\alpha < d : \int_{\mathbb{R}^d} Re(\frac{1}{1+\psi(\xi)})\frac{d\xi}{|\xi|^{d-\alpha}} < \infty\},$$

et sous certaines conditions pour un processus de Markov (Xiao, 2004) :

$$\dim_H X([0,1]) = \sup\{\alpha \geq 0 : \limsup_{r\to 0} \frac{1}{r^\alpha} \int_0^1 P_t(0, B(0,r))dt < \infty\} \quad P_x\text{-p.s.}$$

Pour un processus $\alpha$-stable dans $\mathbb{R}^d$, la formule est obtenue par Blumenthal and Getoor (1960a,b) : pour tout borélien $E \subset \mathbb{R}_+$, presque sûrement

$$\dim_H X(E) = d \wedge \alpha \dim_H E.$$

Le cas inhomogène des processus stables de type-Lévy, cas particulier d'un processus de type-Lévy (voir Böttcher et al. (2013)), est considéré dans Yang (2015). On a dans ce cas la dimension (aléatoire) suivante pour un intervalle ouvert $U$ :

$$\dim_H X(U) = d \wedge \sup_{s\in U} \alpha(X_s).$$

On se pose la question dans [A6] de savoir ce que devient la dimension de l'image dans le cas du processus de Lévy multistable. On considère à nouveau les deux versions $L$ et $L^I$ du processus de Lévy multistable pour $\alpha$ est de classe $\mathcal{C}^1$ à valeurs dans $(0,2)$. On considère une partition $(A_i^n)_{i=1,\dots,n}$ de $[0,1]$ de taille $n$ : $[0,1] = \bigcup_{i=1}^n A_i^n$ avec $A_i^n \cap A_j^n = \varnothing$ si $i \neq j$. On pose également $|A^n| = \max_{i=1}^n |A_i^n|$ le pas de la partition, et pour $E$ un sous-ensemble de $[0,1]$ : $\alpha_*(E) = \inf_{t\in E} \alpha(t)$, $\alpha_*(E) = \sup_{t\in E} \alpha(t)$ et $d_*(E) = (1 \vee \alpha_*(E)) \dim_H E$. On obtient alors la dimension de l'image pour tout sous-ensemble de $[0,1]$.

**Théorème 2.7** ([A6]) Pour toute partition de $[0,1]$ telle que $\lim_{n\to+\infty} |A^n| = 0$ et tout sous-ensemble $E$ de $[0,1]$, on a presque sûrement

$$\dim_H L^I(E) = 1 \wedge \limsup_{n\to+\infty} \max_{i=1}^n [\alpha^*(E\cap A_i^n) \dim_H(E\cap A_i^n)].$$

Pour le processus $L$, on suppose qu'il s'agit réellement d'un processus multistable et non d'un processus de Lévy stable, i.e. on suppose

$$\sup_{(s,t)\in(E\cap \mathring{A}_i^n)^2} \frac{|t-s|}{|\alpha(t)-\alpha(s)|} < +\infty$$

pour tout $i$ à partir d'un certain rang $n$. On obtient alors également la dimension de l'image pour presque tout sous-ensemble de $[0,1]$.

**Théorème 2.8** ([A6]) Pour toute partition de $[0,1]$ telle que $\lim_{n\to+\infty} |A^n| = 0$ et tout sous-ensemble $E$ de $[0,1]$ tel que $\inf_{s\in E} s > 0$, presque sûrement

$$\dim_H L(E) = 1 \wedge \limsup_{n\to+\infty} \max_{i=1}^n [d^*(E\cap A_i^n)].$$

Si $\alpha$ est une constante, $L = L^I$ est un processus de Lévy $\alpha$-stable standard et l'on retrouve la formule déjà connue $\dim_H L(E) = \dim_H L^I(E) = 1 \wedge \alpha \dim_H E$. Si $\alpha$ n'est plus une constante, comme c'est le cas par exemple sous l'hypothèse $\max_i \sup_{(s,t)\in(E\cap \mathring{A}_i^n)^2} \frac{|t-s|}{|\alpha(t)-\alpha(s)|} < +\infty$, alors $L \neq L^I$. Si $\alpha^*(E) > 1$, presque sûrement $\dim_H L(E) = \dim_H L^I(E)$ mais pour $\alpha^*(E) < 1$, les deux processus présentent des comportements différents. En effet $\dim_H L(E) = \dim_H E$, ce qui n'est plus vrai dans le cas où $\alpha$ est constant. On tire de l'égalité $h(t) = \frac{1}{\alpha(t)} \wedge h_\alpha(t)$ que si $\alpha^*(E) < 1$ et $\alpha$ est de classe $\mathcal{C}^1$ alors $h(t) = 1$ pour $t \in E$, ce qui explique en partie pourquoi $L$ agit comme une fonction Lipschitz. En ce sens l'égalité $\dim_H L(E) = \dim_H E$ n'est pas si surprenante. $L^I$ par contre agit de manière similaire au cas $\alpha$ constant, à savoir comme un processus à sauts purs, et donc comme une somme de processus constants par morceaux.

# 4 Temps local

A l'instar de l'ensemble image, pour caractériser le comportement fractionnaire d'un processus, il est naturel de considérer pour un borélien $F$ l'ensemble réciproque $X^{-1}(F) = \{t : X(t) \in F\}$. Il est alors usuel de mesurer la taille de ces ensembles à l'aide de la dimension de Hausdorff. De tels calculs de dimension sont obtenus par exemple pour des processus $\alpha$-stables symétriques (Blumenthal and Getoor, 1962), pour des processus stables (Hawkes, 1971), des processus de Lévy (Hu and Taylor, 1997; Khoshnevisan and Xiao, 2005; Song et al., 2018; Park et al., 2020), des processus additifs (Khoshnevisan and Xiao, 2002; Khoshnevisan et al., 2008), pour le mouvement Brownien (bi)-fractionnaire (Tudor and Xiao, 2007) ou encore pour des champs gaussiens (Xiao, 2006; Biermé et al., 2009).

Dans le cas où $F$ est un singleton, la détermination des lignes de niveaux $\{t : X_t = x\}$ est fortement reliée à la notion de temps local lorsque celui-ci existe. En effet, en considérant la mesure de Lebesgue $\lambda$ en lieu et place de la mesure de Hausdorff, on définit la mesure d'occupation d'un processus $X \in \mathbb{R}^d$ :

$$\mu_I(A) = \lambda(\{t \in I : X_t \in A\}).$$

Si $\mu_I << \lambda$, alors le temps local $l(y, I)$ est défini par

$$l(y, I) = \frac{d\mu_I}{d\lambda}(y), \quad y \in \mathbb{R}^d.$$

Dans le cas où $I = [0, T]$, on note plus simplement le temps d'occupation de l'ensemble $A$ durant l'intervalle de temps $[0, T]$

$$\mathcal{O}_T(A) = \int_0^T \mathbf{1}_A(X_t)dt,$$

et son temps local

$$l_T(y) = l(y, [0, T]) = \frac{d\mathcal{O}_T}{dy}(y).$$

On a alors la formule d'occupation $\int_I f(X_t)dt = \int_{\mathbb{R}^d} f(y)l(y, I)dy$ pour toute fonction borélienne positive $f$ (Samorodnitsky, 2016) et donc pour $T > 0$

$$\int_0^T f(X_t)dt = \int_{\mathbb{R}^d} f(y)l_T(y)dy.$$

L'existence du temps local n'est pas garanti en toute généralité (on pourra se référer à Geman and Horowitz (1980) pour une review). Citons simplement que dans le cas d'un processus de Lévy réel d'exposant $\psi$, le temps local existe si et seulement si $\int_{\mathbb{R}} Re(\frac{1}{1+\psi(\xi)})d\xi < +\infty$ (Hawkes, 1974). Ce résultat est connu depuis 1968 pour des processus de Lévy $\alpha$-stables avec $\alpha > 1$ (Blumenthal and Getoor, 1968). Plus généralement pour des processus stables, dont par exemple des moyennes mobiles stables, des conditions d'existence d'un temps local continu sont établies dans Nolan (1989b).

Dans la lignée des travaux de Altmeyer and Chorowski (2017) et Altmeyer (2021), nous nous sommes posés la question dans [A11] de l'estimation de la mesure d'occupation et du temps local pour un processus de Lévy $\alpha$-stable à partir de l'observation d'une trajectoire discrétisée. On cherche ainsi à estimer $\mathcal{O}_T(A)$ et $l_T(y)$ en supposant que l'on observe un processus réel symétrique $\alpha$-stable $X = L^\alpha$ aux temps $t_k = k\Delta_n$, $k = 1, ..., n$, $\Delta_n = T/n$ pour un temps d'horizon $T > 0$ fixé et l'asymptotique $n \to +\infty$. L'erreur quadratique minimale pour ce problème d'estimation est atteinte par les espérances conditionnelles $\mathbb{E}[\mathcal{O}_T(A)|\mathcal{G}_n]$ et $\mathbb{E}[l_T(y)|\mathcal{G}_n]$ où $\mathcal{G}_n$ est la tribu engendrée par les $X_{t_k}$. Ces espérances n'étant pas calculables en pratique (sauf dans quelques rares exceptions), on introduit les estimateurs non-paramétriques basés sur les sommes de Riemann

$$\hat{\mathcal{O}}_{T,n}(A) = \Delta_n \sum_{k=1}^{n} \mathbf{1}_A(X_{t_{k-1}})$$

et

$$\hat{l}_{T,n}(y) = \frac{\Delta_n}{2h_n} \sum_{k=1}^{n} \mathbf{1}_{[y-h_n, y+h_n]}(X_{t_{k-1}}).$$

Ce problème d'estimation a été étudié dans un premier temps pour des processus de diffusion (voir par exemple Florens-Zmirou (1993)) et apparaît en particulier avec les schemas de discrétisation d'une équation différentielle stochastique (Gobet and Labart, 2008; Kohatsu-Higa et al., 2014; Neuenkirch and Szölgyenyi, 2020). La question de l'approximation du temps local (ou de ses dérivées comme dans Jaramillo et al. (2021)) est également considérée pour des diffusions dans Hofmann (1999), Kohatsu-Higa et al. (2014) ou encore Dion and Genon-Catalot (2016). Des estimateurs standards à noyaux de taux optimaux apparaissent en particulier dans Borodin (1986), Jacod (1998), Ngo and Ogawa (2011) et Kohatsu-Higa et al. (2014). Certains processus de Markov possédant des bornes de type noyau de la chaleur sont étudiés dans Ganychenko et al. (2015) et Ganychenko (2015), et dans Altmeyer and Chorowski (2017) pour les processus de Markov stationnaires.

On retrouve des résultats sur l'estimateur de Riemman pour l'estimation du temps d'occupation dans Gobet and Matulewicz (2017), Chorowski (2018), pour une diffusion dans Ngo and Ogawa (2011) et pour un processus de Lévy autosimilaire dans Ivanovs and Podolskij (2022) (voir aussi Ivanovs and Podolskij (2020)). Son optimalité dans le cas particulier du mouvement Brownien avec drift peut s'obtenir grâce à Ngo and Ogawa (2011), Altmeyer (2021) et Ivanovs and Podolskij (2022).

On prouve dans [A11] que les estimateurs basés sur les sommes de Riemann sont de taux de convergence optimaux au sens $L^2$. En effet les vitesses optimales sont obtenues par l'erreur quadratique des espérances conditionnelles $\mathbb{E}[\mathcal{O}_T(A)|\mathcal{G}_n]$ et $\mathbb{E}[l_T(y)|\mathcal{G}_n]$, dont on calcule la limite exacte dans le cas du processus de Lévy symétrique $\alpha$-stable.

**Théorème 2.9** ([A11]) Pour $y \in \mathbb{R}$ on note $\mathcal{O}_T(y) = \mathcal{O}_T([y, +\infty))$. Si $1 < \alpha \leq 2$ alors

$$\lim_{n\to\infty} \Delta_n^{-1-1/\alpha} \|\mathbb{E}[\mathcal{O}_T(y)|\mathcal{G}_n] - \mathcal{O}_T(y)\|^2_{L^2(\mathbb{P})} = 2\mathbb{E}[l_T(y)] \int_0^\infty \mathbb{E}\left[\mathbf{Var}\left(\mathcal{O}_1(x)|X_1\right)\right] dx.$$

Si $0 < \alpha < 1$ alors

$$\lim_{n\to\infty} \Delta_n^{-2} (\log n)^{-1} \|\mathbb{E}[\mathcal{O}_T(0)|\mathcal{G}_n] - \mathcal{O}_T(0)\|^2_{L^2(\mathbb{P})} = \frac{\Gamma(\alpha)\sin(\frac{\pi\alpha}{2})}{12\pi} \mathbb{E}[|X_1|^{-\alpha}].$$

Enfin pour $\alpha = 1$, on a

$$\lim_{n\to\infty} \Delta_n^{-2} (\log n)^{-2} \|\mathbb{E}[\mathcal{O}_T(0)|\mathcal{G}_n] - \mathcal{O}_T(0)\|^2_{L^2(\mathbb{P})} = \frac{1}{12\pi^2}.$$

Pour le temps local on a également l'expression exacte de la limite de l'erreur quadratique de l'espérance conditionnelle.

**Théorème 2.10** ([A11]) Soit $1 < \alpha \leq 2$, $y \in \mathbb{R}$ et $C(\alpha) = -(\alpha - 1)\Gamma(\alpha)\cos(\frac{\pi\alpha}{2})$. Alors

$$\lim_{n\to\infty} \Delta_n^{1/\alpha - 1} \|\mathbb{E}[l_T(y)|\mathcal{G}_n] - l_T(y)\|^2_{L^2(\mathbb{P})} = \frac{\mathbb{E}[l_T(y)]}{C(\alpha)^2} \int_{\mathbb{R}} \mathbb{E}[\mathbf{Var}(l_1(x)|X_1)] dx.$$

Ces vitesses sont atteintes par les sommes de Riemann dans le cas des processus de Lévy symétriques $\alpha$-stables. Là encore on obtient le comportement asymptotique exact des erreurs quadratiques de l'estimateur de Riemann pour le temps d'occupation.

**Théorème 2.11** ([A11]) Soit $\phi(x) := (x - \lfloor x \rfloor) - (x - \lfloor x \rfloor)^2$, $x \geq 0$, où $\lfloor x \rfloor$ est la partie entière de $x$. Si $1 < \alpha \leq 2$, alors

$$\lim_{n\to\infty} \Delta_n^{-1-1/\alpha} \|\hat{\mathcal{O}}_{T,n}(0) - \mathcal{O}_T(0)\|^2_{L^2(\mathbb{P})} = T^{1-1/\alpha} \frac{2\Gamma(1/\alpha)}{\pi\alpha^2} \mathbb{E}[|X_1|] \int_0^\infty \frac{\phi(x)}{x^{2-1/\alpha}} dx.$$

Si $0 < \alpha < 1$, alors

$$\lim_{n\to\infty} \Delta_n^{-2} (\log n)^{-1} \|\hat{\mathcal{O}}_{T,n}(0) - \mathcal{O}_T(0)\|^2_{L^2(\mathbb{P})} = \frac{\Gamma(\alpha)\sin(\frac{\pi\alpha}{2})}{12\pi} \mathbb{E}[|X_1|^{-\alpha}].$$

Enfin pour $\alpha = 1$, on a

$$\lim_{n\to\infty} \Delta_n^{-2} (\log n)^{-2} \|\hat{\mathcal{O}}_{T,n}(0) - \mathcal{O}_T(0)\|^2_{L^2(\mathbb{P})} = \frac{1}{12\pi^2}.$$

Pour le temps local on obtient une borne supérieure, prouvant que l'estimateur de Riemann est une nouvelle fois de vitesse de convergence optimale.

**Théorème 2.12** ([A11]) Soit $1 < \alpha \leq 2$ et $y \in \mathbb{R}$. Alors pour $\epsilon > 0$ suffisamment petit

$$\|\hat{l}_{T,n}(y) - l_T(y)\|^2_{L^2(\mathbb{P})} \leq C(1 \vee T^{-\epsilon})T^{1-1/\alpha}(h_n^{\alpha-1} + \Delta_n^{1+1/\alpha}h_n^{-2}),$$

où $C < \infty$ est indépendante de $a, b, T$ et $n$. Si $h_n = \Delta_n^{1/\alpha}$, alors

$$\Delta_n^{1/\alpha-1}\|\hat{l}_{T,n}(y) - l_T(y)\|^2_{L^2(\mathbb{P})} \leq C(1 \vee T^{-\epsilon})T^{1-1/\alpha}.$$

On constate ainsi que l'estimateur le plus simple et le plus naturel pour ce problème, à savoir la somme de Riemann, est de vitesse de convergence optimale pour tout $1 < \alpha \leq 2$ et tout $y \in \mathbb{R}$, tandis que pour $0 < \alpha \leq 1$, $\hat{\mathcal{O}}_{T,n}(0)$ est même asymptotiquement efficace et réalise donc le minimum d'erreur quadratique asymptotique possible. L'efficacité n'est en revanche plus vrai pour $1 < \alpha \leq 2$, en particulier pour le mouvement Brownien, comme le montre la proposition suivante.

**Proposition 2.1** ([A11]) Pour tout $1 < \alpha \leq 2$ on a

$$\frac{\lim_{n\to\infty} \Delta_n^{-1-1/\alpha}\|\hat{\mathcal{O}}_{T,n}(0) - \mathcal{O}_T(0)\|^2_{L^2(\mathbb{P})}}{\lim_{n\to\infty} \Delta_n^{-1-1/\alpha}\|\mathbb{E}[\mathcal{O}_T(0)|\mathcal{G}_n] - \mathcal{O}_T(0)\|^2_{L^2(\mathbb{P})}} > 1.$$

# 5 Formalisme multifractal

### Spectre des singularités

Pour un processus multifractionnaire, le calcul de $h(t_0)$ n'est pas toujours pertinent ni réalisable en pratique car $h$ peut varier très fortement d'un point à un autre. De très nombreux signaux sont très irréguliers, comme c'est le cas par exemple lors de l'analyse de la vitesse d'un flux turbulent en mécanique des fluides. Etabli dans ce cadre d'analyse des turbulences, le concept de multifractalité naît de l'introduction des cascades multiplicatives par Mandelbrot (1974), puis la notion de spectre fut introduite par Frisch and Parisi (1985). On retrouve ainsi une analyse multifractale des turbulences de Navier-Stokes dans plusieurs références, parmi lesquelles Hentschel and Procaccia (1983) qui considèrent des dimensions fractales de boîte généralisées, puis Frisch and Parisi (1985) qui définissent le spectre des singularités et introduisent la transformation de Legendre dans cette analyse des turbulences, Halsey et al. (1986) qui considèrent des transformées de Legendre de dimensions généralisées introduites par Hentschel and Procaccia (1983) et donnent des exemples de calculs, ou encore Chhabra et al. (1989).

L'analyse multifractale intervient dans beaucoup d'autres domaines, comme par exemple la modélisation des marchés financiers (Cont, 2001), l'étude des tremblements de terre (Harte, 2001) ou encore le traitement d'images médicales (Morales and Kolaczyk, 2002). L'analyse du traffic TCP fait également état d'observations semblables, à savoir des signaux d'une très grande irrégularité présentant un caractère (multi)fractionnaire. De nombreux articles montrent ainsi le caractère fractal du traffic au sens où il présente de l'autosimilarité ou de la longue mémoire (voir par exemple Leland et al. (1994), Paxson and Floyd (1995), Lévy Véhel and Riedi (1997), Jacquet (1998), Erramilli et al. (2000), Cappe et al. (2002), Dang et al. (2003), Doi et al. (2004)). Cette autosimilarité observée dans le traffic Ethernet (voir Leland et al. (1994)) et expliquée en partie de manière empirique dans Willinger et al. (1997), provient sans doute de l'aggrégation de sources individuelles qui engendrent le flux global des paquets. La superposition de plusieurs sources i.i.d. ON/OFF (appelées aussi "packet train") dont les périodes ON ou OFF présentent "l'effet Noah" (i.e. une variance infinie) peut produire un réseau aggrégé présentant "l'effet Joseph" (de l'autosimilarité ou de la longue mémoire). Lévy Véhel and Riedi (1997) et Lévy Véhel and Rams (2013) proposent en ce sens une modélisation et une analyse multifractale du traffic internet en accord avec les observations antérieures. Dans un modèle simplifié, une source $W(t)$ binaire stationnaire à accroissements indépendants, une période ON ($W(t) = 1$) de loi d'inter-arrivée $F_1$ succède à une période OFF de loi $F_2$. Pour $M$ sources avec un facteur d'échelle $T$, $(W_M(Tt))_t = (\int_0^{Tt}(\sum_{i=m}^{M} W^m(u))du)_t$ représente le nombre de paquets cumulés présents sur le réseau dans l'intervalle $[0, Tt]$. Correctement renormalisé, on montre que lorsque $M \to +\infty$ puis $T \to +\infty$, on obtient $\sigma_\infty B_H(t)$ avec $B_H$ un mouvement Brownien fractionnaire. En prenant $T \to +\infty$ en premier, si $F_1$ et $F_2$ sont à queue lourde comme dans Taqqu et al. (1998), on obtient un processus de Lévy $\alpha$-stable qui devient ainsi une première modélisation homogène d'un phénomène additif naturel, plus simple à étudier dans un premier temps. Le caractère homogène étant remis en cause en pratique, il semble opportun de se tourner vers des processus multifractionnaires inhomogènes comme les processus multistables.

Abandonnant l'idée de déterminer complètement la fonction $h(t)$ trop irrégulière, la question que l'on peut se poser alors est celle de savoir si la fonction $h(t)$ peut prendre une certaine valeur donnée $h$ et si oui quelle est la taille de l'ensemble des points où la fonction prend cette valeur fixée $h$. Cette taille n'est en général pas calculée avec la mesure de Lebesgue puisque bien souvent la régularité la plus probable apparaît Lebesgue presque partout et les points atypiques forment un ensemble de mesure de Lebesgue nulle. Il est donc courant de chercher à mesurer cet ensemble en terme de dimension de Hausdorff. Plusieurs notions distinctes de spectres sont apparues depuis l'introduction initiale de Frisch and Parisi (1985). Ces notions diffèrent selon que l'on utilise la notion de régularité höldérienne et son indice de régularité ponctuelle $h(t)$, que l'on utilise la régularité des coefficients d'ondelettes $\omega(t)$, mais également selon la manière d'approcher ces régularités. On définit tout d'abord le spectre des singularités.

**Définition 2.1** (Spectre des singularités) Soit $f$ une fonction réelle et $E_h$ l'ensemble $\{t : h(t) = h\}$, où $h(t)$ est l'exposant de Hölder ponctuel de $f$ au point $t$. Le spectre des singularités ou spectre de Hausdorff est la fonction

$$f_H : h \mapsto \dim_H(E_h).$$

Ce spectre a d'abord été introduit dans le but d'analyser des mesures possédant des propriétés fractales. On retrouve des calculs de spectres par exemple dans Falconer (1994), King (1995), Olsen (1996, 1998) ou Hu and Taylor (1997). On trouvera une synthèse des notions relatives aux spectres des singularités des mesures dans Falconer (1997), et une bibliographie dans Olsen (1994). Plus récemment on retrouve l'étude des cascades de Mandelbrot dans Barral (2000), des cascades de Poisson dans Barral and Mandelbrot (2004) et encore plus récemment, le spectre de la mesure d'occupation d'un processus de type-Lévy stable est obtenu dans Seuret and Yang (2017).

Sans le formaliser de la sorte, les résultats pour des processus stochastiques commencent avec Orey and Taylor (1974) qui calculent la dimension des points d'un processus $\alpha$-stable symétrique dont la régularité est au plus égale à $h$ fixé. Le spectre du mouvement Brownien fractionnaire sera ensuite obtenu dans Adler (1981), et celui d'un processus de Lévy en général dans Jaffard (1999). En notant $\beta$ l'indice de Blumenthal et Getoor du processus de Lévy, on obtient en particulier

$$f_H(h) = \begin{cases} \beta h & \text{si } h \in [0, 1/\beta] \\ -\infty & \text{sinon} \end{cases}$$

si $\beta > 0$ et si le processus n'a pas de composante brownienne. Si le processus possède une composante brownienne, on a alors

$$f_H(h) = \begin{cases} \beta h & \text{si } h \in [0, 1/2) \\ 1 & \text{si } h = 1/2 \\ -\infty & \text{sinon.} \end{cases}$$

Le cas du processus de Lévy en temps multifractionnaire est considéré dans Barral and Seuret (2007). Auparavant Barral and Lévy Véhel (2004) proposaient un modèle simplifié de protocole TCP possédant le même spectre qu'un processus de Lévy bien qu'ayant des accroissements corrélés et non stationnaires. Les résultats de Durand (2009) et Durand and Jaffard (2012) complètent ensuite ceux de Jaffard (1999) en considérant les $g$-mesures de Hausdorff d'ensembles $E_h \cap V$ pour $V$ un ouvert quelconque.

Barral et al. (2010) introduisent quant à eux la notion de spectre local pour étudier un processus de Markov purement à sauts positifs $(M_t)_{t\geq 0}$ sans composante brownienne qui ressemble localement à un Lévy subordinateur. A l'aide d'un théorème d'ubiquité de Barral and Seuret (2011) et d'une représentation de Poisson $M_t = \int_0^t \int_0^\infty \frac{1}{(1+z)^{1/\gamma(M_{s-})}} N(ds, dz)$, ils obtiennent pour tout intervalle $I$ un spectre aléatoire

$$f_H(I, h) = \sup\{h.\gamma(M_t) : t \in I, h.\gamma(M_t) < 1\}.$$

En comparaison la représentation de Poisson $X_t = \int\limits_0^\infty \int_{\mathbb{R}\setminus\{0\}} f_t(s) z^{<-1/\alpha(t)>} N(ds, dz)$ (voir (1.3)) d'un processus multistable général de noyau $f$, comme définie dans Falconer and Lévy Véhel (2009) pourrait permettre d'obtenir le spectre de processus multistables comme le mouvement linéaire multifractionnaire multistable (LMMM). Avant cela le spectre des singularités pour le mouvement linéaire fractionnaire stable (LFSM) a été obtenu dans Balança (2014) à l'aide du formalisme 2-microlocal. On a dans ce cas

$$f_H(h) = \begin{cases} \alpha(h-H)+1 & \text{si } h \in [H-\frac{1}{\alpha}, H] \\ -\infty & \text{sinon.} \end{cases}$$

Balança (2014) fournit également le spectre local pour le mouvement linéaire multifractionnaire stable (LMSM). D'autres exemples de processus sont également considérés dans la littérature, comme des superprocessus avec un mécanisme de branchement stable (Mytnik and Wachtel, 2012), les processus de Volterra-Lévy, i.e. des processus de la forme $X_t = \int_0^t f_t(u)dM(u)$ avec $M$ un processus de Lévy (Marquardt, 2006; Mytnik and Neuman, 2012; Neuman, 2014), les processus de Boltzmann qui décrivent l'évolution de la vitesse d'une particule dans le modèle de Boltzmann (Xu, 2016), des processus de diffusion avec sauts comme les processus de type-Lévy stables (Yang, 2018), ou encore des processus de type-Lévy avec des coefficients non bornés (Lörinczi and Yang, 2019).

### Spectre de Legendre

L'exemple historique d'application physique est l'étude de la vitesse des flots de turbulence. Ces flots ne sont pas spatialement homogènes et l'irrégularité du flot semble varier énormément d'un point à l'autre (ce phénomène est appelé "intermittency"). Il apparaît dès le départ avec la définition du spectre de Hausdorff faite à base d'infimums et de limites inférieures que son calcul s'avère impossible à mettre en oeuvre d'un point de vue pratique, ou du moins aboutit à des calculs numériquement très instables. On a alors proposé d'utiliser des quantités moyennes plus stables numériquement en formulant l'hypothèse que cela permettait de retrouver le spectre des singularités (ce qui est vrai dans certains cas, mais faux de manière générale). C'est ce constat que font Frisch and Parisi (1985) lors de leur introduction de la définition de $f_H$. Ils ont alors l'idée d'utiliser des quantités moyennes afin de stabiliser les calculs numériques, puis d'utiliser ensuite la transformée de Legendre pour retrouver $f_H$. Plus précisément on définit pour $q \in \mathbb{R}$ une fonction de structure $S_q$ par $S_q(l) = \int |X_{t+l} - X_t|^q dt$, ou dans sa version empirique $S_q(n) = \sum_{k=0}^{n-1} |X(\frac{k+1}{n}) - X(\frac{k}{n})|^q$, et on introduit la fonction de partition $\tau(q) = \liminf_{l \to 0} \frac{\log S_q(l)}{\log l}$, ou dans sa version empirique $\tau(q) = \liminf_{n\infty} \frac{\log S_q(n)}{-\log n}$. Le spectre de Legendre est alors défini par

$$f_l(h) = \tau^*(h) = \inf_{q\in\mathbb{R}} (hq - \tau(q)),$$

i.e. par la transformée de Legendre de $\tau$. C'est comme cela que la notion de spectre a été introduite historiquement par Hentschel and Procaccia (1983) et Frisch and Parisi (1985) pour des mesures. A priori chaque fonction de structure ou discrétisation $(X(\frac{k}{n}))_{k,n}$ conduit à une notion de spectre de Legendre différente. D'autres choix sont ainsi possibles, comme ceux faits par exemple dans Leonarduzzi et al. (2016).

### Spectre des grandes déviations

Une autre approche encore plus facile à implémenter que le spectre de Legendre consiste à utiliser un estimateur de la dimension de boîte. De manière générale, comme il est indiqué dans Lévy Véhel and Tricot (2004), on peut considérer $v_X(U)$ une mesure de la variation de $X$ sur l'intervalle $U$. Les choix classiques sont les accroissements $v_X(U) = |X(U_{\max}) - X(U_{\min})|$ si $U = [U_{\max}, U_{\min}]$, l'oscillation $v_X^{osc}(U) = \sup_{t\in U} X(t) - \inf_{t\in U} X(t)$, ou encore le coefficient d'ondelette : si $U_{k,j} = [\frac{k}{2^j} : \frac{k+1}{2^j}]$, on choisit $v_X^w(U_{k,j}) = d_{j,k} = \int X(t)2^{-j}\psi(2^{-j}t-k)dt$ (le choix dépend alors a priori de l'ondelette d'analyse $\psi$). Pour des intervalles du type $I_n^k = [\frac{k}{n}, \frac{k+1}{n}]$, le choix naturel devient

$$X_n^k = X(\frac{k+1}{n}) - X(\frac{k}{n})$$

qui donne les calculs les plus simples, ou encore $X_n^k = Osc(I_n^k) = \sup_{s\in I_n^k} X(s) - \inf_{s\in I_n^k} X(s)$, l'oscillation de $X$ sur $I_n^k$. On définit alors $A(U) = \frac{\log v_X(U)}{\log |U|}$, ce qui donne pour l'intervalle $I_n^k = [\frac{k}{n}, \frac{k+1}{n}]$

$$A(I_n^k) = -\frac{\log |X_n^k|}{\log n},$$

et l'on considère ensuite l'estimateur de la dimension de boîte

$$N^{(n)}(h,\varepsilon) = \mathrm{Card}\{k = 0, ..., n-1 : h - \varepsilon \leq A(I_n^k) < h + \varepsilon\}.$$

On définit enfin le spectre des grandes déviations $f_g$ par

$$f_g(h) = \lim_{\varepsilon\to 0} \limsup_{n\to+\infty} \frac{\log N^{(n)}(h,\varepsilon)}{\log n}.$$

L'avantage immédiat de cette définition repose dans sa simplicité calculatoire. A contrario la définition repose sur un choix de partitionnement du support du signal et sur un choix de mesure de variation du signal à l'intérieur de ce support : a priori différents choix conduisent à différents spectres. On peut définir aussi

$$\underline{f}_g(h) = \lim_{\varepsilon\to 0} \liminf_{n\to+\infty} \frac{\log N^{(n)}(h,\varepsilon)}{\log n}.$$

Canus et al. (1998) introduisent également deux autres définitions du spectre des grandes déviations plus adaptés à l'estimation. On trouvera des applications du calcul de $f_g$ à la modélisation du traffic internet dans Lévy Véhel and Riedi (1997) et Lévy Véhel and Rams (2013), ainsi que le calcul de $f_g$ et $f_l$ pour le mouvement Brownien fractionnaire.

### Formalisme multifractal

Le spectre de Legendre $f_l$ ayant été introduit à l'origine par Frisch and Parisi (1985) dans le but de fournir une méthode de calcul de $f_H$ plus facile à mettre en pratique, il est supposé dans leurs modélisations que les deux spectres sont égaux, justifiant ainsi leur méthode. On a un résultat (voir Ellis (1984), Riedi (1995)) qui relie le spectre des grandes déviations $f_g$ au spectre de Legendre $f_l$. Si la fonction de partition $\tau(q) = \lim_{n \to +\infty} \frac{\log S_q(n)}{\log n}$ existe et est différentiable, alors

$$f_g(h) = \tau^*(h) = f_l(h).$$

De manière générale, on a sous certaines conditions (voir Lévy Véhel and Vojak (1998))

$$f_H(h) \leq \underline{f}_g(h) \leq f_g(h) \leq f_l(h).$$

Lorsque les spectres coïncident on dit que le formalisme multifractal est vérifié. On parle aussi de formalisme multifractal fort, et de formalisme multifractal faible lorsque l'on a $f_g = f_l$ avec $f_H \neq f_g$. Dans le cas des processus de Lévy $\alpha$-stables par exemple, c'est le formalisme multifractal faible qui se produit (Riedi, 2002) :

$$\underline{f}_g(h) = f_g(h) = f_l(h) = \begin{cases} \alpha h & \text{si } h \in [0, \frac{1}{\alpha}] \\ 1 + 1/\alpha - h & \text{si } h \in [\frac{1}{\alpha}, 1 + \frac{1}{\alpha}] \\ -\infty & \text{sinon.} \end{cases}$$

Lors de leur analyse des données de turbulence, ces notions ont été introduites dans Frisch and Parisi (1985), mais également dans Chhabra and Jensen (1989) ou encore dans Chhabra et al. (1989), en supposant que le formalisme multifractal (fort) était vérifié. L'objectif de ce formalisme est précisément de déterminer le spectre des singularités par d'autres méthodes que le calcul de l'exposant de Hölder et la dimension de Hausdorff numériquement instables. Le spectre des grandes déviations apparaît ensuite en 1989 pour ces raisons computationelles.

Encore une fois le formalisme multifractal est étudié dans un premier temps pour des mesures. Rand (1989) montre le formalisme multifractal pour une mesure de Gibbs invariante par un cookie-cutters, Cawley and Mauldin (1992) pour des mesures de Moran. Peyrière (1992) et Brown et al. (1992) établissent des liens entre les spectres de mesure et prouvent le formalisme multifractal pour une large classe de mesures, dont les mesures multinomiales, résultats qui seront ensuite étendus par Mandelbrot and Riedi (1995).

Un autre formalisme pour des mesures autosimilaires est prouvé dans Riedi (1995) en utilisant le spectre de boîtes qui y est défini. Jaffard (1997) définit et considère également des fonctions autosimilaires qui vérifient le formalisme multifractal. Ben Nasr et al. (2002) étudient la validité du formalisme multifractal pour les mesures en général et donnent des conditions pour que le formalisme multifractal soit vérifié.

### D'autres formalismes

D'autres choix sont possibles afin de résoudre le problème de calcul de $f_H$ en pratique. Par exemple Lévy Véhel and Tricot (2004) définissent un autre spectre $f_\Delta^{limsup}$ vérifiant toujours

$$f_H \leq f_\Delta^{limsup} \leq f_g,$$

et qui sera donc plus précis que $f_g$ pour approcher $f_H$. Barral et al. (2012) introduisent un formalisme multifractal local pour des données dont le spectre dépend des observations i.e. un spectre aléatoire adapté par exemple aux processus définis dans Barral et al. (2010). Grahovac and Leonenko (2014) proposent de modifier la fonction de partition pour rendre plus robuste le formalisme qui sera vérifié en particulier par les processus Lévy $\alpha$-stables et le mouvement linéaire fractionnaire stable (LFSM).

Un autre formalisme utilise la décomposition en ondelettes du processus. C'est ce formalisme qui est souvent utilisé pour l'estimation du spectre des singularités quand le formalisme est vérifié comme dans Riedi et al. (1999) et Ribeiro et al. (2000). L'idée consiste à considérer l'indice de régularité $\omega(t)$ à la place de $h(t)$ et à utiliser le coefficient d'ondelettes $d_{j,k}$ comme mesure de régularité du signal au niveau de résolution $2^{-j}$. L'enjeu devient alors comme précédemment de savoir si les spectres de Legendre ou des grandes déviations ainsi définis vérifient le formalisme multifractal. La littérature à ce sujet est vaste. Là encore différents choix de fonctions de structure et de fonctions de partitions produisent différents formalismes. Citons parmi ces différents choix la Wavelet Transform Integrale Méthod et la Wavelet Transform Modulus Maxima Method introduites par Muzy et al. (1993), Bacry et al. (1993) et Muzy et al. (1994), méthodes qui présentent peu de garanties mathématiques et qui ont un coût calculatoire élevé, la méthode Multifractal Detrending Fluctuation Analysis (voir par exemple Serrano and Figliola (2009)), la méthode Wavelet Leaders (Lashermes et al., 2005) utilisée par exemple dans Wendt and Abry (2007), Wendt et al. (2007) et Wendt et al. (2012).

Les espaces 2-microlocaux peuvent également être utilisés pour étudier un formalisme basé sur l'exposant de régularité introduit par Meyer (1998), le weak scaling exposant (voir Lashermes et al. (2005)). Bastin et al. (2016) proposent un autre formalisme en définissant $f_g$ à partir des Wavelets Leaders, et Leonarduzzi et al. (2019) proposent une modification du spectre de Legendre basé sur les ondelettes pour détecter des spectres non-concaves et garder l'aspect pratique du spectre de Legendre au niveau des calculs.

De manière plus générale, Jaffard (2000) et Jaffard and Meyer (2000) montrent que le formalisme multifractal est vérifié pour quasiment toutes les fonctions, au sens où c'est le cas pour beaucoup d'espaces de Besov. Barral and Seuret (2020) cherchent alors la plus grande classe possible de fonctions telles que $f_H = f_l$, complétant ainsi les résultats de Jaffard (2000) qui trouve des espaces de Baire de fonctions dont les fonctions typiques ont un spectre de singularités prescrit et qui vérifient le formalisme multifractal. C'est ce problème qui est appelé Conjecture de Frisch et Parisi, dont une réponse partielle est apportée dans Jaffard (2000) grâce au formalisme développé à l'aide des Wavelets Leaders. Barral and Seuret (2020) introduisent des espaces fonctionnels de Baire appelés espaces de Besov en environnement aléatoire et concluent que la conjecture de Frisch et Parisi est vraie : pour chaque fonction $\sigma$ qui est admissible pour être une fonction spectre et qui vérifie le formalisme multifractal (en particulier qui est concave), il existe un espace fonctionnel de Baire tel que chaque élément vérifie $f_H = \sigma$ et tel que le formalisme multifractal est vérifié.

**Le cas des processus de Lévy multistables**

C'est dans ce contexte que l'on obtient dans [A8] le spectre des singularités des deux versions du processus de Lévy multistable, ainsi que le spectre des grandes déviations du processus $L^I$. Rappelons que l'on considère pour $t \in [0,1]$

$$L_t = \sum_{i=1}^{\infty} \gamma_i C_{\alpha(t)}^{1/\alpha(t)} \Gamma_i^{-\alpha(t)} \mathbf{1}_{[0,t]}(V_i)$$

et

$$L_t^I = \sum_{i=1}^{\infty} \gamma_i C_{\alpha(V_i)}^{1/\alpha(V_i)} \Gamma_i^{-\alpha(V_i)} \mathbf{1}_{[0,t]}(V_i)$$

avec $C_u = (\int_0^\infty x^{-u} \sin x dx)^{-1}$. Ici $\alpha : [0,1] \mapsto (1,2)$ est une fonction de classe $\mathcal{C}^1$ et l'on pose $\alpha_* = \inf_{u\in[0,1]} \alpha(u)$ et $\alpha^* = \sup_{u\in[0,1]} \alpha(u)$. En vertu de la relation unissant les deux processus $L$ et $L^I$ ([A5]), les deux spectres des singularités sont égaux.

**Théorème 2.13** ([A8]) Presque sûrement le spectre des singularités de $L$ et $L^I$ vérifie

$$f_H(h) = \begin{cases} -\infty & \text{pour } h < 0; \\ h\alpha^* & \text{pour } h \in [0, \frac{1}{\alpha^*}]; \\ 1 & \text{pour } h \in (\frac{1}{\alpha^*}, \frac{1}{\alpha_*}); \\ \dim_H(\{t \in [0,1] : \alpha(t) = \alpha_*\}) & \text{pour } h = \frac{1}{\alpha_*}; \\ -\infty & \text{pour } h > \frac{1}{\alpha_*}. \end{cases}$$

Si $\alpha_* \neq \alpha^*$, comme $f_H(h) = 1$ pour $h \in (\frac{1}{\alpha^*}, \frac{1}{\alpha_*})$, le spectre $f_H$ fournit en particulier une décomposition de $[0,1]$ en une union non dénombrable d'ensembles disjoints de dimension de Hausdorff égale à un. La preuve du Théorème 2.13 repose sur deux éléments essentiels. Le premier est la décomposition reliant $L$ et $L^I$ issue de [A5] : on a presque sûrement

$$L_t = L_t^I + \int_0^t \sum_{i=1}^{\infty} \gamma_i \frac{d(C_{\alpha(.)}^{1/\alpha(.)} \Gamma_i^{-1/\alpha(.)})}{dt}(s) \mathbf{1}_{[0,s)}(V_i) ds$$

ce qui assure en particulier que $L$ et $L^I$ ont tous deux même exposant de Hölder ponctuel et même spectre des singularités. Le deuxième élément est l'obtention de la formule exacte de l'exposant de Hölder ponctuel pour $L$ (et $L^I$), à savoir

$$h(t) = \frac{\delta(t)}{\alpha(t)} = \frac{1}{\alpha(t)} \inf_{V_{\phi(n)} \in R_t} \liminf_{i\to+\infty} -\frac{\log \phi(i)}{\log |V_{\phi(i)} - t|}$$

avec $\delta(t)$ qui ne dépend pas de $\alpha$ et qui correspond à la vitesse d'approximation d'un point $t$ par une suite de variables aléatoires i.i.d. de loi uniforme. Les lignes de niveaux de $h(t)$

$\{t : h(t) = h\}$ correspondent donc aux abscisses des points d'intersection de la courbe de $\delta$ avec la courbe de $h\alpha$. A la lecture du résultat du Théorème 2.13, on voit que $\delta$ est une fonction dont les lignes de niveaux vérifient $\dim_H\{t : \delta(t) = h\} = h$ pour $h \in [0, 1]$.

Concernant le spectre des grandes déviations, on le calcule pour $L^I$ en utilisant l'indépendance des accroisssements du processus.

**Théorème 2.14** ([A8]) Presque sûrement, le spectre des grandes déviations de $L^I$ vérifie

$$f_g(h) = f_l(h) = \begin{cases} -\infty & \text{pour } h < 0; \\ h\alpha^* & \text{pour } h \in [0, \frac{1}{\alpha^*}]; \\ 1 & \text{pour } h \in (\frac{1}{\alpha^*}, \frac{1}{\alpha_*}]; \\ 1 + \frac{1}{\alpha_*} - h & \text{pour } h \in (\frac{1}{\alpha_*}, 1 + \frac{1}{\alpha_*}]; \\ -\infty & \text{for } h > \frac{1}{\alpha_*}. \end{cases}$$

On a ainsi le formalisme multifractal faible pour $L^I$ mais pas le formalisme fort.

# 6 Perspectives de recherche

### Perspectives probabilistes

On a présenté quelques résultats de calculs d'exposants de Hölder pour les processus multistables, avec notamment une inégalité pour l'exposant local dans le cas du mouvement linéaire multifractionnaire multistable (2.1). Il serait intéressant de savoir si cette inégalité est optimale ou non. Pour ce faire, une analyse 2-microlocale des processus multistables en général et du mouvement linéaire multifractionnaire multistable (LMMM) en particulier permettrait d'établir des résultats de régularité plus fins que ceux obtenus jusqu'alors, et d'obtenir au passage les exposants de Hölder ponctuel et local si la frontière 2-microlocale est déterminée. Les travaux de Balança (2014) vont dans ce sens puisqu'ils répondent à cette question dans le cas où $\alpha$ est constant. Des liens y sont établis entre le mouvement linéaire fractionnaire stable et les processus de Lévy, liens qui mériteraient d'être creusés dans le cas multistable. Dans un second temps on pourrait alors envisager de calculer le spectre des singularités du mouvement linéraire multifractionnaire multistable.

On pourra également chercher à intégrer ces modèles dans l'étude des équations différentielles stochastiques dirigées par des processus à sauts. De tels modèles ont déjà été largement étudiés dans le cas où le bruit est un processus de Lévy par exemple. Les processus multistables fournissant un modèle où l'intensité des sauts est explicite car directement régie par la fonction de stabilité $\alpha$, il sera intéressant d'intégrer ces processus multistables dans une équation différentielle stochastique afin de mieux contrôler la hauteur typique des sauts d'une solution. Si l'intégration d'un processus de Lévy multistable ne pose pas de difficulté particulière car il s'agit d'une semi-martingale, il en est tout autrement dans le cas par exemple où le processus à sauts est le mouvement linéaire multifractionnaire multistable (LMMM), qui lui n'est a priori pas une semi-martingale. Il faudra alors se rapprocher des résultats d'intégration stochastique par rapport au mouvement Brownien (multi)fractionnaire (voir par exemple Lebovits (2012)).

### Perspectives statistiques

Dans la continuité du point précédent, on pourra considérer dans un premier temps des équations différentielles stochastiques dirigées par un processus de Lévy multistable et se poser la question de l'estimation de la fonction de stabilité dirigeant le bruit. On peut également se poser la question de savoir comment estimer le drift ou la volatilité dans ce cadre. Des résultats d'estimation paramétrique du drift et de la volatilité existent déjà par exemple dans le cas où le bruit est un processus de Lévy $\alpha$-stable (voir Clément and Gloter (2019)). L'estimation non-paramétrique par projection est également déjà considérée pour des bruits qui sont Lévy et mériterait d'être étudiée dans ce cadre d'EDS avec bruit multistable. La question de la pertinence du modèle pour les sauts est également naturelle et peut être abordée par exemple en terme de test d'inhomogénéité (ou de multistabilité), i.e. de test détectant si $\alpha$ est une fonction ou une constante. Une autre manière de procéder consiste à considérer des processus de type-Lévy, et en particulier les stable-like processus qui sont des modèles très proches du processus de Lévy multistable $L^I$. Les questions statistiques classiques autour de ces processus semblent relativement ouvertes et toutes les questions précédentes provenant de la caractérisation du modèle par une fonction $\alpha$ se posent également dans ce contexte. De manière similaire on peut aussi se demander si le processus observé appartient bien à la famille des processus (multi)stables. Là encore il semble raisonnable de considérer en premier lieu le cas du processus de Lévy, la question se rapprochant alors de celle de l'identification de loi, comme par exemple dans Fromont and Laurent (2006).

Il sera également intéressant d'essayer d'approfondir les résultats déjà obtenus dans le cas des processus multistables. On pourra en particulier élargir ces résultats à des classes plus larges que le processus de Lévy multistable. Ces processus étant définis en toute généralité par un noyau $f_t(x)$ (voir (1.1)), on pourra chercher à établir des tests de multistabilité pour des noyaux quelconques. Un autre résultat extrêmement utile en vue d'applications potentielles est d'estimer le noyau $f_t$ lui-même. Cette question semble toute fois assez difficile de prime abord. On pourra plus simplement dans un premier temps essayer de caractériser ce noyau à l'aide de test, avec par exemple un test d'identification de modèle du type $H_0 : f_t(x) = f_t^0(x)$ pour un noyau $f^0$ donné. Dans le cas particulier $f_t^0(x) = \mathbf{1}_{[0,t]}(x)$ du processus de Lévy, on pourra ainsi utiliser la propriété de semi-martingale et se rapprocher de tests existants comme par exemple celui de Aït-Sahalia and Jacod (2009b) ou Aït-Sahalia and Jacod (2018). De manière générale la question du choix de modélisation par des processsus à sauts est assez peu abordée d'un point vue statistique. Une réponse à cette question est apportée en partie respectivement par Duvernet et al. (2010) et Aït-Sahalia and Jacod (2018) qui observent une discrétisation d'une semi-martingale et testent asymptotiquement si la semi-martingale est de type Itô ou non. Concernant l'estimation de la régularité ou de l'intensité des sauts, le problème relève de l'estimation non-paramétrique. Il semble donc naturel de se demander si l'on ne peut pas définir des estimateurs minimax à l'aide d'autres méthodes, comme par exemple des estimateurs par projection, issus de la fonction caractéristique ou encore à l'aide d'une représentation par ondelettes. Pour l'estimation du temps local, il semble raisonnable d'essayer de savoir ce que l'on peut obtenir dans le cas du processus de Lévy multistable ou dans celui de processus stables plus généraux, avant éventuellement d'envisager la classe des processus multistables en général.

Citons enfin une autre utilisation potentielle de ces modèles multistables avec le cas où l'intensité de sauts du processus dépendrait d'une ou plusieurs covariables comme c'est souvent le cas en finance ou en cybersécurité par exemple. On pourrait ainsi considérer une fonction de stabilité de la forme $\alpha(t, Z_t)$ où $Z$ est un processus covariable. Si la définition d'un tel processus ne pose a priori pas de difficultés particulières, les questions statistiques autour de ce modèle (comme par exemple celle de l'estimation de la fonction $\alpha(t, x)$) ne sont en revanche pas abordées pour le moment, bien que pouvant modéliser un bon nombre de phénomènes réels.

# Chapitre 3

# Particules en interaction

Ce chapitre est en partie issu du co-encadrement de la thèse d'Emilien Manent avec Frédéric Lavancier.

## 1 Quelques modèles de systèmes de particules en interaction

On cherche depuis longtemps à modéliser des dynamiques de populations. Un des modèles les plus basiques pour cela est certainement celui du processus de naissances et morts simple (étudié au moins depuis Feller (1939) et Kendall (1949)), processus qui trouve néanmoins des applications encore de nos jours avec par exemple la description des apparitions/disparitions des restaurants dans le quartier de Shibuya à Tokyo (Sadahiro, 2019). Ce processus de comptage n'a cependant pour objectif que celui de décrire le cardinal de la population étudiée, occultant ainsi le comportement particulier des individus composant cette population. Dans bien des situations ce comportement individuel joue un rôle essentiel dans la dynamique globale et mérite d'être intégré dans le modèle servant à l'analyse.

Il en est ainsi par exemple en épidémiologie (Masuda and Holme, 2017) ou en analyse d'images (Wang and Zhu, 2002). Dans le domaine de la sylviculture, l'intégration des individus dans l'étude permet de faire ressortir des phénomènes de compétition/coopération entre les arbres, décrits à l'aide de processus du type Growth Interaction Model (voir par exemple Renshaw and Särkkä (2001); Särkkä and Renshaw (2006); Pommerening and Grabarnik (2019) ou encore Häbel et al. (2019)).

Les travaux présentés ici sont motivés à l'origine par l'analyse d'images en biologie et notamment l'étude du phénomène d'exocytose cellulaire. Les techniques de captation vidéo actuelles permettent de tracer précisément l'évolution de chaque protéine ciblée au sein d'une cellule (Costes et al., 2004; Bolte and Cordelieres, 2006; Lagache et al., 2015). On dispose alors d'images 3D de l'intérieur d'une cellule (Pécot et al., 2018), où des protéines ciblées apparaissent (ou naissent) régulièrement dans l'image, se déplacent aléatoirement puis disparaissent (ou meurent) de l'image une fois l'exocytose effectuée.

De nombreux modèles existent pour représenter un système de particules au comportement pré-établi. La généralisation du processus de naissance et morts simple au cas de particules sans mouvement naissant et disparaissant dans un espace quelconque a été étudiée par Preston (1975). Auparavant, le cas de particules évoluant indépendamment les unes des autres est abordé avec le processus de diffusion branchant de Skorokhod (1964), qui est l'un des premiers processus introduit pour décrire un système de particules dans son entier. Les processus de branchements généraux seront ensuite définis et étudiés dans une série de papiers de Ikeda et al. (1965, 1966a,b,c). La construction, qui consiste à accoler des trajectoires de processus de Markov forts, est séquentielle. Un autre point de vue est adopté par la construction des superprocessus qui fait intervenir des EDS pour des processus à valeurs mesures. On pourra se référer par exemple à Dynkin (1991); Li (2011); Athreya and Ney (2012) ou encore à Bansaye and Méléard (2015) pour cette construction. Citons enfin le modèle défini séquentiellement des processus de Markov déterministes par morceaux introduit par Davis (1984b) qui décrit un processus évoluant dans le temps de façon déterministe (le long d'un flot), sautant dans l'espace à un temps aléatoire avant de repartir de façon déterministe et continuer ainsi de suite.

**Le processus jump-move** Généralisant la construction de Davis (1984a), nous définissons dans [A10] un processus appelé Jump-Move qui permet en particulier d'introduire du mouvement pour les particules d'un processus spatial de naissances et morts (Preston, 1975).

On considère $E$ un espace polonais muni de la tribu borélienne $\mathcal{E}$ et d'une distance $d$. On se donne alors $(\Omega, \mathcal{F})$ un espace probabilisé, $(\mathbb{P}_x)_{x\in E}$ une famille de mesures de probabilités sur $(\Omega, \mathcal{F})$, ainsi que les trois ingrédients suivants :

1. une fonction $\alpha : E \to \mathbb{R}_+^*$ d'intensité des sauts,
2. un noyau de transition $K$ défini sur $E \times \mathcal{E}$,
3. un processus de Markov $((Y_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E})$ continu sur $E$ pour le mouvement entre les sauts, de noyau de transition

$$Q_t^Y(x, A) = \mathbb{P}_x(Y_t \in A), \quad x \in E, A \in \mathcal{E}.$$

On se donne ensuite $(Y_t^{(j)})_{t\geq 0}$ une suite de processus sur $E$ identiquement distribués selon $(Q_t^Y)_{t\geq 0}$. On pose $T_0 = 0$ et on prend $x_0 \in E$ comme état initial. On construit alors itérativement le processus jump-move $(X_t)_{t\geq 0}$ de la façon suivante : pour $j \geq 0$,

i) sachant $X_{T_j} = x_j$, simuler $(Y_t^{(j)})_{t\geq 0}$ conditionnellement à $Y_0^{(j)} = x_j$ selon le noyau $(Q_t^Y(x_j, .))_{t\geq 0}$.

ii) Sachant $X_{T_j} = x_j$ et $(Y_t^{(j)})_{t\geq 0}$, simuler $\tau_{j+1}$ selon la fonction de répartition

$$F_{j+1}(t) = 1 - \exp\left(-\int_0^t \alpha(Y_u^{(j)})du\right).$$

iii) Sachant $X_{T_j} = x_j$, $(Y_t^{(j)})_{t\geq 0}$ et $\tau_{j+1}$, simuler $x_{j+1}$ selon le noyau $K(Y^{(j)}_{\tau_{j+1}}, .)$.

iv) Poser $T_{j+1} = T_j + \tau_{j+1}$, $X_t = Y^{(j)}_{t-T_j}$ pour $t \in [T_j, T_{j+1}[$ et $X_{T_{j+1}} = x_{j+1}$.

On note $(\mathcal{F}_t)_{t\geq 0}$ la filtration naturelle complétée de $(X_t)_{t\geq 0}$ et $\mathcal{F}_{t^+} = \bigcap_{s>t} \mathcal{F}_s$. Le processus $((X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E}, (\mathcal{F}_t)_{t\geq 0})$ est alors un processus càdlàg de Markov homogène en temps ([A10]). On note également $N_t = \sum_{j\geq 1} \mathbf{1}_{T_j \leq t}$ le processus de comptage du nombre de sauts du processus $X$. Si $\alpha$ est bornée, $N$ est un processus non-explosif. On montre également dans [A10] (Proposition S7) que $M_t = N_t - \int_0^t \alpha(X_s)ds$ est une $\mathcal{F}_{t^+}$-martingale, i.e. qu'une version continue à gauche de l'intensité de $N$ par rapport à $(\mathcal{F}_{t^+})_{t\geq 0}$ est $\alpha(X_{t^-})$.

**Le cas particulier du processus Birth-Death-Move** En vue de modéliser un processus de type naissances et morts avec des individus en mouvement, on considère le cas du processus birth-death-move qui généralise celui de Preston (1975). On considère ainsi $E$ de la forme spécifique $E = \bigcup_{n=0}^{+\infty} E_n$ avec $((E_n)_{n\geq 0}, (\mathcal{E}_n)_{n\geq 0})$ une suite d'espaces polonais disjoints, muni de la tribu $\mathcal{E} = \sigma(\bigcup_{n=0}^{+\infty} \mathcal{E}_n)$. On suppose de plus que $E_0$ est réduit à un singleton noté $\varnothing$. On suppose alors que le mouvement $(Y_t)_{t\geq 0}$ vérifie

$$\mathbb{P}_x((Y_t)_{t\geq 0} \subset E_n) = \mathbf{1}_{E_n}(x), \quad x \in E, n \geq 0.$$

Pour l'intensité des sauts, on suppose que $\alpha$ s'écrit $\alpha = \beta + \delta$ avec $\beta : E \to \mathbb{R}_+$ une intensité de naissances et $\delta : E \to \mathbb{R}_+$ une intensité de morts (avec $\delta(\varnothing) = 0$). Afin de spécifier la dynamique des sauts, on suppose enfin que $K$ s'écrit de la forme

$$K(x, A) = \frac{\beta(x)}{\alpha(x)} K_\beta(x, A) + \frac{\delta(x)}{\alpha(x)} K_\delta(x, A)$$

avec $K_\beta : E \times \mathcal{E} \to [0, 1]$ un noyau de transition pour les naissances et $K_\delta : E \times \mathcal{E} \to [0, 1]$ un noyau de transition pour les morts, i.e. vérifiant

$$K_\beta(x, E_{n+1}) = \mathbf{1}_{E_n}(x) \quad \text{et} \quad K_\delta(x, E_n) = \mathbf{1}_{E_{n+1}}(x), \quad x \in E, n \geq 0.$$

Ce processus sert alors pour la modélisation d'un système de particules en interaction de type naissances/morts avec des particules en mouvement dans $\mathbb{R}^d$ si l'on considère le cas particulier où $E_n$ représente l'ensemble des configurations de points de cardinal $n$, ou plus précisément si $E_n$ s'écrit $E_n = \pi_n(W^n)$ pour $W \subset \mathbb{R}^d$ un fermé de $\mathbb{R}^d$ représentant l'espace où vivent les particules et $\pi_n : (x_1, ..., x_n) \in W^n \to \{x_1, ..., x_n\}$ est la projection naturelle associée (qui supprime simplement l'ordre des particules).

# 2 Propriétés du processus Jump-Move

Afin de caractériser le comportement du processus, il est naturel de considérer les propriétés de son semi-groupe de transition. Le noyau de transition $Q_t(x, A)$ pour $x \in E$ et $A \in \mathcal{E}$ est entièrement déterminé par l'équation de Kolmogorov backward (Theorem 1 de [A13]) :

$$Q_t(x, A) = \mathbb{E}_x^Y \left[ \mathbf{1}_{Y_t \in A} e^{-\int_0^t \alpha(Y_u)du} \right] + \int_0^t \int_E Q_{t-s}(y, A) \mathbb{E}_x^Y [K(Y_s, dy)\alpha(Y_s) e^{-\int_0^s \alpha(Y_u)du}] ds,$$

équation qui possède une unique solution stochastique si la fonction d'intensité $\alpha$ est bornée. Nous nous sommes également intéressés dans [A13] aux propriétés de continuité du semi-groupe de transition. Comme dans Dynkin (1965) et Øksendal (2013), on dit que le processus $(X_t)_{t\geq 0}$ de noyau de transition $Q_t$ sur $E$ est Feller continu si $Q_t C_b(E) \subset C_b(E)$ avec $C_b(E)$ l'ensemble des fonctions continues bornées sur $E$. On dit qu'il est Feller si $Q_t C_0(E) \subset C_0(E)$ avec $C_0(E)$ l'ensemble des fonctions continues qui s'annulent à l'infini et si $\lim_{t\to 0} \|Q_t f - f\|_\infty = 0$ pour tout $f \in C_0(E)$. On peut alors relier la continuité du semi-groupe $Q_t$ à celle du semi-groupe $Q_t^Y$ correspondant au processus $Y$ régissant le mouvement continu de $X$. On a en effet :

**Théorème 3.1** ([A13])
1. Si $Q_t^Y C_b(E) \subset C_b(E)$ et si $K C_b(E) \subset C_b(E)$ alors $(X_t)_{t\geq 0}$ est Feller continu.
2. Si $Q_t^Y C_0(E) \subset C_0(E)$ et si $K C_0(E) \subset C_0(E)$ alors $(X_t)_{t\geq 0}$ est Feller.

Pour un système de particules de $\mathbb{R}^d$ (dans ce cas $E = \bigcup_{n=0}^{+\infty} E_n$ avec $E_n = \pi(W^n)$, $W \subset \mathbb{R}^d$), le mouvement continu $(Y_t)_{t\geq 0}$ se décompose en mouvements continus $(Z_t^{|n})_{t\geq 0}$ sur chaque $E_n$, $Z_t^{|n}$ étant alors un déplacement continu de $n$ particules non-ordonnées de $\mathbb{R}^d$. Le caractère Feller de $Y$ est alors hérité des propriétés de continuité de chaque semi-groupe de $(Z_t^{|n})_{t\geq 0}$. Plus précisément, on a le résultat suivant ([A13], Proposition 9) :

**Théorème 3.2** Si $(Z_t^{|n})_{t\geq 0}$ est un processus de Feller continu (resp. Feller) sur $W^n$ pour chaque $n \geq 1$ alors $(Y_t)_{t\geq 0}$ est un processus de Feller continu (resp. Feller) sur $E$.

Enfin dans le cas où le mouvement $Y$ est Feller, le générateur infinitésimal du processus jump-move $X$ se décompose comme la somme du générateur infinitésimal de $Y$ (noté $\mathcal{A}^Y$) et du générateur infinitésimal d'un processus de Markov purement à sauts. En effet, si $\mathcal{A}$ est le générateur infinitésimal de $X$, on a la formule ([A13], Theorem 6)

$$\mathcal{A} f = \mathcal{A}^Y f + \alpha \times K f - \alpha \times f$$

pour toute fonction $f$ dans le domaine de $\mathcal{A}$.

**Propriétés de convergence** On s'est posé également la question de la convergence du processus vers un régime stationnaire, la convergence considérée ici étant celle en temps long. La référence pour cela dans le cas fondamental du processus de naissances et morts simple est très certainement l'article de Karlin and McGregor (1957) qui établit la classification de ses états en donnant des critères de transience, de récurrence nulle ou de récurrence positive, ainsi que des critères d'ergodicité sous forme de séries pour le processus. Dans le cas du processus de naissances et morts spatial, des taux de convergence vers la mesure stationnaire sont obtenus dans Møller (1989). D'autres résultats similaires ont été établis pour des systèmes de particules avec un mouvement dirigé par un flot (Çinlar and Kao, 1991), pour des processus de naissances et morts avec des contraintes sur les sauts (Manjrekar, 2018), ou encore plus récemment pour des processus de type naissances et morts dont les taux dépendent du temps (Bezborodov and Persio, 2022).

Dans le cas du processus birth-death-move, i.e. pour $E$ de la forme $E = \bigcup_{n=0}^{+\infty} E_n$ avec $(E_n)_{n\geq 0}$ une suite d'espaces polonais disjoints, on peut montrer la convergence en temps long du processus vers une mesure stationnaire sous l'hypothèse que l'espace d'états est constitué d'un nombre fini de $E_n$. Plus précisément on suppose que pour un certain $n^* > 0$, $\beta(x) = 0$ si $x \in E_n$ avec $n \geq n^*$, ou de manière équivalente que $E$ s'écrit $E = \bigcup_{n=0}^{n^*} E_n$ pour un certain $n^* > 0$. On montre alors sous cette hypothèse la convergence vers une mesure stationnaire :

**Théorème 3.3** ([A10], Proposition S4) $(X_t)_{t\geq 0}$ admet une unique mesure stationnaire $\mu_\infty$ et il existe $a > 0$ et $c > 0$ tels que pour toute fonction mesurable bornée $g$ et $t > 0$,

$$\sup_{y\in E} \left| \int_E g(y) Q_t(y, ds) - \int_E g(y)\mu_\infty(dy) \right| \leq a\|g\|_\infty e^{-ct}.$$

On obtient également dans ce cas un contrôle uniforme de la variance des intégrales de la forme $\int_0^t g(X_s)ds$, ce qui constitue un résultat clé pour l'étude de la convergence des estimateurs de l'intensité des sauts de $X$.

**Théorème 3.4** ([A10], Corollary S5) Il existe $a > 0$ et $c > 0$ tels que pour toute fonction positive mesurable bornée $g$ et $t > 0$,

$$\left| \mathbb{E}\left[ \frac{1}{t} \int_0^t g(X_s)ds \right] - \int_E g(z)\mu_\infty(dz) \right| \leq \frac{a\|g\|_\infty}{ct}$$

et

$$\mathbf{Var}\left( \int_0^t g(X_s)ds \right) \leq c_0 \|g\|_\infty \mathbb{E}\left[ \int_0^t g(X_s)ds \right]$$

avec $c_0$ une constante indépendante de $t$ et $g$.

Si on ne suppose plus l'existence d'un $n^* > 0$ tel que $E = \bigcup_{n=0}^{n^*} E_n$, la première étape pour montrer l'ergodicité du processus consiste à établir un couplage entre $(X_t)_{t\geq 0}$ et un processus de naissances et morts simple $(\eta_t)_{t\geq 0}$ sur $\mathbb{N}$, de taux de naissances $\beta_n = \sup_{x\in E_n} \beta(x)$, de taux de morts $\delta_n = \inf_{x\in E_n} \delta(x)$, dont le noyau de transition est noté $q_t$. Inspiré du couplage de Preston (1975), on construit un processus $\check{C}_t = (X'_t, \eta'_t)$ dont les marginales vérifient en particulier les égalités en loi $(X'_t)_{t\geq 0} = (X_t)_{t\geq 0}$ et $(\eta'_t)_{t\geq 0} = (\eta_t)_{t\geq 0}$, ainsi que le théorème suivant :

**Théorème 3.5** ([A13]) Supposons que $(Y_t)_{t\geq 0}$ soit un processus de Feller et que $KC_0(E) \subset C_0(E)$. Alors pour tout $t \geq 0$, $(x, n) \in E \times \mathbb{N}$, $A \in \mathcal{E}$ et $S \in \mathcal{P}(\mathbb{N})$, on a :

1. $\check{Q}_t((x, n); E \times S) = q_t(n, S)$
2. $\check{Q}((x, n); A \times \mathbb{N}) = Q_t(x, A)$
3. Si $x \in E$ avec $n(x) \leq n$ alors $\check{Q}_t((x, n); \Gamma) = 0$ pour

$$\Gamma = \{(y, n) \in E \times \mathbb{N}; n(y) > m\}.$$

On obtient en particulier pour $x \in E_m$ avec $m \leq n$,

$$\mathbb{P}_{(x,n)}\left((\check{C}_s)_{s\geq 0} \subset \Gamma^c\right) = 1,$$

ce qui traduit le fait que le processus $(\eta_t)_{t\geq 0}$ converge plus lentement vers 0 que le processus $(X_t)_{t\geq 0}$ vers l'état $E_0 = \{\varnothing\}$. On déduit alors des propriétés d'ergodicité du processus de naissances et morts simple sur $\mathbb{N}$ (étudiées par exemple dans Karlin and McGregor (1957)) des propriétés d'ergodicité identiques pour le processus birth-death-move $(X_t)_{t\geq 0}$.

**Théorème 3.6** ([A13], Theorem 11) Supposons que $(Y_t)_{t\geq 0}$ soit un processus de Feller avec $KC_0(E) \subset C_0(E)$, que $\beta_n\delta_n > 0$ pour tout $n \geq 1$ avec

$$\sum_{n=2}^{+\infty} \frac{\beta_1 \dots \beta_{n-1}}{\delta_1 \dots \delta_n} < \infty \quad \text{et} \quad \sum_{n=2}^{\infty} \frac{\delta_1 \dots \delta_n}{\beta_1 \dots \beta_n} = \infty.$$

Alors $\mu_\infty(A) = \lim_{t\to+\infty} Q_t(x, A)$ existe pour $x \in E$ et $A \in \mathcal{E}$, et est indépendante de $x$. De plus $\mu_\infty$ est l'unique mesure de probabilité invariante du processus, i.e. vérifiant

$$\mu_\infty(A) = \int_E Q_t(x, A)\mu_\infty(dx)$$

pour tout $A \in \mathcal{E}$ et $t \geq 0$.

Concernant le taux de convergence, on a des résultats de convergence identiques à ceux obtenus par Møller (1989) dans le cas du processus de naissances et morts spatial. Sans l'hypothèse de l'existence d'un entier $n^* > 0$ tel que $E = \bigcup_{n=0}^{n*} E_n$, la vitesse de convergence n'est plus uniforme en l'état initial. Elle s'exprime grâce au théorème suivant ([A13], Theorem 12, Corollary 13) :

**Théorème 3.7** Supposons que les hypothèses du Théorème 3.6 soient satisfaites. Soit $\gamma_1$ et $\gamma_2$ deux mesures de probabilité sur $(E, \mathcal{E})$ telles que pour $k = 1, 2$

$$\sum_{n=2}^{+\infty} \gamma_k(E_n) \sqrt{\frac{\delta_1 \ldots \delta_n}{\beta_1 \ldots \beta_{n-1}}} < \infty.$$

On peut trouver deux constantes $c > 0$ et $0 < r < 1$ (dépendantes de $\gamma_1$ et $\gamma_2$) telles que pour tout $t \geq 0$,

$$\sup_{A \in \mathcal{E}} \left| \int_E Q_t(x, A)\gamma_1(dx) - \int_E Q_t(x, A)\gamma_2(dx) \right| \leq cr^t.$$

Si de plus $\beta_n \leq \delta_{n+1}$ à partir d'un certain rang et

$$\sum_{n=2}^{+\infty} \sqrt{\frac{\beta_1 \ldots \beta_{n-1}}{\delta_1 \ldots \delta_n}} < \infty,$$

alors pour tout $y \in E$, on peut trouver deux constantes $c > 0$ et $0 < r < 1$ (dépendantes de $y$ et $\mu_\infty$) telles que

$$\sup_{A \in \mathcal{E}} |\mu_\infty(A) - Q_t(y, A)| \leq cr^t.$$

**Application à un système de particules en interaction** En guise d'application, nous pouvons modéliser par exemple un système de particules en interaction à valeurs dans $W \subset \mathbb{R}^d$ de la forme $W = I_1 \times \ldots \times I_d$ avec $I_i$ compact de $\mathbb{R}^d$. On considère alors

$$W_n = \{(x_1, \ldots, x_n) \in (\mathring{W})^n, i \neq j \Rightarrow x_i \neq x_j\}$$

l'ensemble des $n$ particules disjointes de $\mathring{W}$, et pour espace d'états de notre processus $X$ nous prenons $E_0 = \{\varnothing\}$, pour $n \geq 1$, $E_n = \pi_n(W_n)$ et enfin $E = \bigcup_{n=0}^{+\infty} E_n$.

Afin de définir les naissances de $X$, nous considérons un potentiel d'interaction par paires $V : E \to \mathbb{R} \cup \{\infty\}$ de la forme

$$V(x) = an(x) + \sum_{1 \leq i \neq j \leq n} \phi(x_i - x_j)$$

avec $a > 0$, $\phi(\xi) = \phi(-\xi)$ pour tout $\xi \in \mathbb{R}^d$ si $n \geq 2$, $V(\{\varnothing\}) = 0$ et $V(\{\xi\}) = a$ pour $\xi \in W$.

On peut alors définir le noyau de naissance $K_\beta$ pour $x \in E$ et $A \subset W$ par

$$K_\beta(x, A \cup x) = \frac{\int_A e^{-(V(x \cup \xi) - V(x))} d\xi}{\int_W e^{-(V(x \cup \xi) - V(x))} d\xi}$$

avec l'intensité de naissance associée

$$\beta(x) = \frac{1}{n(x) \vee 1} \int_W e^{-(V(x \cup \xi) - V(x))} d\xi.$$

Pour le noyau de mort, on choisit le noyau uniforme : pour $f$ mesurable bornée sur $E$ et $x \in E$,

$$K_\delta f(x) = \frac{1}{n(x)} \sum_{i=1}^{n(x)} f(x \backslash x_i)$$

avec l'intensité de mort associée uniforme également : $\delta(x) = \mathbf{1}_{n(x)\geq 1}$.

Pour le mouvement, nous considérons la diffusion de Langevin sur $W_n$

$$dZ_{t,i}^{|n} = -\sum_{j\neq i} \nabla\phi(Z_{t,i}^{|n} - Z_{t,j}^{|n})dt + \sqrt{2}dB_{t,i}, \quad 1 \leq i \leq n$$

avec des conditions de réflection au bord (Fattler and Grothaus (2007)), le mouvement $Y$ étant alors défini sur chaque $E_n$ par $Y_t^n = \pi_n(Z_t^{|n})$ (on ne conserve pas l'ordre des particules).

Sous des hypothèses raisonnables sur $\phi$, hypothèses vérifiées par les potentiels classiques de Lennard-Jones répulsif, soft-core, de Riesz ou encore de Strauss régularisé, le potentiel $V$ correspond au potentiel d'une mesure de Gibbs invariante pour le processus $X$.

**Théorème 3.8** ([A13]) Le processus birth-death-move $(X_t)_{t\geq 0}$ ainsi défini est un processus de Feller convergeant vers la mesure de probabilité invariante de Gibbs sur $W$ de potentiel $V$, i.e. de mesure à densité proportionnelle à $\exp(-V(x))$ par rapport au processus de Poisson simple sur $W$.

# 3 Estimation non-paramétrique de l'intensité de sauts

D'un point de vue statistique, nous avons cherché dans [A10] à estimer le taux de sauts $\alpha$ dans un modèle de type birth-death-move. Beaucoup d'articles se sont intéressés à cette question de l'estimation d'un taux de sauts pour des modèles variés plus ou moins complexes, à commencer par le processus de naissances et morts simple sur $\mathbb{N}$ de taux constants, pour lequel les travaux d'estimation sur les taux $\alpha, \beta$ et $\delta$ remontent aux années 1950 (Anscombe (1953), Kendall (1949, 1952), Moran (1951, 1953), Darwin (1956)). Pour des taux non constants, Wolff (1965) pose et étudie un modèle paramétrique sur les taux, tout comme Keiding (1975) qui obtient également des résultats sur l'estimateur du maximum de vraisemblance. Un modèle classique consiste à supposer que le taux $\alpha_n$ est proportionnel à la taille de la population (i.e. $\alpha_n = \alpha \times n$), hypothèse qui toute fois n'est pas reprise par Reynolds (1973) pour établir un modèle bayésien. Pour des modèles de systèmes de particules plus complexes de type branching processes, la vraisemblance est calculée dans Löcherbach (2002).

Les premières méthodes non-paramétriques à noyau apparaissent avec Ramlau-Hansen (1983) pour l'estimation de l'intensité d'un processus de comptage (avec une version lisse de l'estimateur de Nelson et Aalen), et Diggle (1985) pour l'estimation de l'intensité d'un processus de Cox stationnaire. L'extension des résultats de Diggle (1985) au cas d'un champ de $\mathbb{R}^2$ sera considéré par Guan (2008), le cas de l'estimateur de Nelson et Aalen pour un modèle avec covariables étant abordé dans Martinussen and Scheike (2007). L'estimateur à noyau est également utilisé par Van Es et al. (2007) pour estimer la densité de sauts d'un processus de Poisson composé, par Azaïs et al. (2013) pour l'estimation du taux de sauts d'un processus de renouvellement, ou encore par Azaïs et al. (2014) pour la loi des temps de sauts d'un processus de Markov déterministe par morceaux.

La méthode d'estimation non-paramétrique par projection est quant à elle utilisée par Reynaud-Bouret (2003) pour l'estimation de l'intensité d'un processus de Poisson inhomogène en dimension 1. Cela fournit une méthode adaptative qui sera reprise ensuie par Reynaud-Bouret and Schbath (2010) pour l'estimation de l'intensité d'un processus de Hawkes.

Enfin, en vue d'applications plus pratiques, l'observation d'une trajectoire le long d'une discrétisation du temps fait l'objet de travaux plus récents, comme ceux de Figueroa-López (2009) qui réutilise un estimateur par projection pénalisée pour estimer la densité de Lévy de la somme d'un processus de Poisson composé et d'un mouvement brownien, et pour des processus de Lévy plus généraux par Comte and Genon-Catalot (2011). L'estimateur à noyau est quant à lui privilégié par Bec and Lacour (2012) pour estimer la densité de Lévy d'un processus de Lévy purement à sauts à l'aide d'une discrétisation de la trajectoire.

Nous introduisons également dans [A10] un estimateur non-paramétrique à noyau pour estimer l'intensité de sauts $\alpha$ d'un processus jump-move $(X_t)_{t\geq 0}$. Dans le cas de l'observation d'une trajectoire complète en temps continu $(X_t)_{0\leq t\leq T}$ sur l'intervalle $[0,T]$ de temps de sauts $(T_j)_{j\geq 1}$, l'estimateur $\hat{\alpha}(x)$ de $\alpha(x)$ pour $x\in E$ est défini par

$$\hat{\alpha}(x) = \frac{\int_0^T k_T(x, X_{s^-})dN_s}{\int_0^T k_T(x, X_s)ds}$$

où $N_t = \text{Card}\{j\geq 1 : T\leq t\}$ et $(k_T)_{T\geq 0}$ est une famille de fonctions positives sur $E\times E$ vérifiant $\sup_{x,y,T}|k_T(x,y)| < +\infty$. Pour $x,y\in E$, le rôle de $k_T(x,y)$ est de mesurer la proximité entre $x$ et $y$. Un choix typique pour $k_T(x,y)$ est de la forme $k_T(x,y) = k(\frac{d(x,y)}{h_T})$ pour $k$ fonction bornée sur $\mathbb{R}$, une pseudo-distance $d$ de notre choix et $h_T$ une fenêtre. Lorsque le processus $(X_t)_{t\geq 0}$ converge (en temps long) vers une mesure invariante $\mu_\infty$, on montre que l'estimateur $\hat{\alpha}(x)$ est consistant lorsque $T\to+\infty$. Dans le cas d'un processus birth-death-move pour lequel l'espace d'états $E$ est de la forme $E = \bigcup_{n=0}^{n^*} E_n$ pour un certain $n^*$, on obtient (Proposition 1 [A10]) :

**Théorème 3.9** Pour tout $x\in E$ et $\varepsilon > 0$, on peut fixer $c > 0$ (ne dépendant que de $\varepsilon$) tel que

$$\mathbb{P}\left(|\hat{\alpha}(x) - \alpha(x)| > \varepsilon\right) \leq c\left(\frac{1}{Tv_T(x)} + w_T^2(x)\right)$$

avec $v_T(x) = \int_E k_T(x,y)\mu_\infty(dz)$ et $w_T(x) = \frac{1}{v_T(x)}\int_E(\alpha(z)-\alpha(x))k_T(x,y)\mu_\infty(dz)$. En particulier si $\lim_{T\to\infty} Tv_T(x) = +\infty$ et $\lim_{T\to+\infty} w_T(x) = 0$ alors

$$\lim_{T\to+\infty}\hat{\alpha}(x) \overset{\mathbb{P}}{=} \alpha(x).$$

Le terme $v_T(x)$ est relié au contrôle de $\mathbb{E}\left[\int_0^T k_T(x, X_s)ds\right]$ alors que $w_T(x)$ dépend de la régularité de la fonction $\alpha$ sous la mesure invariante $\mu_\infty$, régularité qui apparaît classiquement dans le contrôle de la vitesse de convergence en estimation non-paramétrique à noyau.

La consistance en norme $L^2(\Omega)$ est également discutée dans [A10], l'estimateur $\hat{\alpha}(x)$ ne possédant pas nécessairement de moment d'ordre 2. En effet, comme $\int_0^T k_T(x, X_s)ds$ peut être infiniment petit à horizon $T$ fini, on peut trouver des exemples pour lesquels $\mathbb{E}[\hat{\alpha}(x)] = +\infty$ (Lemme S1 [A10]). D'un point de vue pratique, il n'est pas raisonnable d'espérer estimer correctement $\alpha(x)$ avec une valeur trop faible de $\int_0^T k_T(x, X_s)ds$ (qui mesure le temps passé par $X$ proche de $x$). On peut donc considérer l'estimateur modifié

$$\hat{\alpha}_\xi(x) = \hat{\alpha}(x)\mathbf{1}_{\int_0^T k_T(x,X_s)ds>\xi}$$

pour lequel on contrôle le moment d'ordre 2 : pour tout $\xi > 0$ et $0 < \eta < 1/3$, lorsque $T$ tend vers $+\infty$,

$$\mathbb{E}\left[|\hat{\alpha}_\xi(x) - \alpha(x)|^2\right] = \mathcal{O}\left(\frac{1}{(Tv_T(x))^{1/3-\eta}} + w_T^2(x)\right).$$

**Différents choix de pseudo-distance $d$** L'estimateur $\hat{\alpha}$ recouvre un large éventail de situations diverses grâce aux différentes possibilités pour $k_T$. Un choix standard pour la famille $(k_T)_{T\geq 0}$ consiste à prendre $k_T$ de la forme

$$k_T(x,y) = k\left(\frac{d(x,y)}{h_T}\right)$$

où $k$ est une fonction bornée, $d$ est une pseudo-distance et $h_T > 0$ est un paramètre de fenêtre. Pour la consistance on demandera classiquement à ce que $\lim_{T\to+\infty} h_T = 0$ mais pas trop vite, de sorte que $\lim_{T\to+\infty} Tv_T(x) = +\infty$.

Lorsque $E = \bigcup_{n=0}^{n^*} E_n$ avec $E_n$ l'espace des configurations de points de $\mathbb{R}^d$ de cardinal $n$, plusieurs choix de pseudo-distance $d$ sont possibles. Un premier choix standard correspond à la distance de Hausdorff $d_H$ définie par

$$d_H(x,y) = \max\left(\max_{u\in x}\min_{v\in y}\|u-v\|, \max_{v\in y}\min_{u\in x}\|v-u\|\right).$$

Un autre choix classique dans ce contexte consiste à prendre la distance de transport optimal $d_\kappa$ introduite par Schuhmacher and Xia (2008), définie pour $\kappa > 0$, $x = \{x_1, ..., x_{n(x)}\}$ et $y = \{y_1, ..., y_{n(y)}\}$ avec $n(x) \leq n(y)$ par

$$d_\kappa(x,y) = \frac{1}{n(y)}\left(\min_{\pi\in S_n(y)}\sum_{i=1}^{n(x)}(\|x_i - y_{\pi(i)}\| \wedge \kappa) + \kappa(n(y) - n(x))\right)$$

où $S_n(y)$ désigne l'ensemble des permutations de $\{y_1, ..., y_{n(y)}\}$. En supposant que la fonction $\alpha(x)$ ne dépende que de la cardinalité de $x$, i.e. soit de la forme $\alpha(x) = \alpha_0(n(x))$

pour une fonction $\alpha_0$ définie sur $\mathbb{N}$, on peut exploiter cette hypothèse en choisissant par exemple $k_T(x,y) = \mathbf{1}_{n(x)=n(y)}$ (on retrouve alors l'estimateur utilisé pour le processus de naissances et morts simple) ou encore $d(x,y) = |n(x) - n(y)|$ qui fournit une version lissée du choix précédent. Plus généralement, si l'on suppose que la fonction $\alpha$ ne dépend que de caractéristiques encodées dans une fonction $f : E \to \mathbb{R}^p$ pour $p \geq 1$, i.e. $\alpha(x) = \alpha_0(f(x))$, on peut exploiter cette hypothèse en choisissant $d(x,y) = \|f(x) - f(y)\|$. On se ramène alors à un problème d'estimation non-paramétrique en dimension $p$ et l'on retrouve pour $\hat{\alpha}$ une vitesse de convergence standard d'estimation d'une fonction de $\mathbb{R}^p$ à savoir (Corollary 1 [A10]) :

$$\mathbb{P}\left(|\hat{\alpha}(x) - \alpha(x)| > \varepsilon\right) \leq c\left(\frac{1}{Th_T^p} + h_T^2\right).$$

**Estimation en temps discret** Si l'on suppose d'un point de vue plus réaliste que l'on observe le processus $(X_t)_{t\geq 0}$ en des temps discrets $t_0, ..., t_m$ avec $t_0 = 0$ et $t_m = T$, une version discrétisée de $\hat{\alpha}(x)$ consiste à remplacer les intégrales par des sommes de Riemann, ce qui donne

$$\frac{\sum_{j=0}^{m-1} \Delta N_{t_{j+1}} k_T(x, X_{t_j})}{\sum_{j=0}^{m-1} \Delta t_{j+1} k_T(x, X_{t_j})},$$

avec $\Delta N_{t_j} = N_{t_j} - N_{t_{j-1}}$ et $\Delta t_j = t_j - t_{j-1}$. $\Delta N_{t_j}$ n'étant pas nécessairement observé, il convient de l'estimer par une quantité observable $D_j$. Notre estimateur discrétisé devient

$$\hat{\alpha}_{(d)}(x) = \frac{\sum_{j=0}^{m-1} D_{j+1} k_T(x, X_{t_j})}{\sum_{j=0}^{m-1} \Delta t_{j+1} k_T(x, X_{t_j})}.$$

La consistance de $\hat{\alpha}_{(d)}(x)$ se déduit alors de celle de $\hat{\alpha}(x)$ dans un cadre d'observations hautes fréquences (Proposition 3 [A10]).

**Choix de fenêtre par validation croisée sur vraisemblance partielle** Toujours en vue d'applications pratiques, nous fournissons pour finir une procédure de choix de fenêtre basée sur une vraisemblance partielle du processus (toujours dans [A10]). Plus précisément, on se base sur la log-vraisemblance du processus $(N_t)_{0\leq t\leq T}$. L'intensité de $N_t$ étant $\alpha(X_{t^-})$, le théorème de Girsanov donne que la log-vraisemblance de $(N_t)_{0\leq t\leq T}$ par rapport au processus de Poisson simple sur $[0,T]$ est

$$T + \sum_{j=1}^{N_T} \log \alpha(X_{T_j^-}) - \int_0^T \alpha(X_s)ds.$$

Dans le cas d'observations en temps continu, on choisira alors $h_T$ comme

$$\hat{h} = \arg\max_h \left( \sum_{j=1}^{N_T} \log \hat{\alpha}_h^-(X_{T_j^-}) - \int_0^T \hat{\alpha}_h^-(X_s)ds \right),$$

$\hat{\alpha}_h^-(X_s)$ étant l'estimation de $\alpha(x)$ pour $x = X_s$ sans utiliser le processus entre $T_{N_s}$ et $T_{N_s+1}$ :

$$\hat{\alpha}_h^-(X_s) = \frac{\sum_{i=1, i\neq N_{s+1}}^{N_T} k_T\left(\frac{d(X_s, X_{T_i^-})}{h}\right)}{\int_{[0,T]\setminus[T_{N_s}, T_{N_{s+1}}]} k_T\left(\frac{d(X_s,X_u)}{h}\right) du}.$$

Enfin pour une procédure d'estimation complète, on observe le processus en temps discret et la discrétisation des intégrales par les sommes de Riemann conduit à choisir $h_T$ en utilisant la procédure de validation croisée sur la log-vraisemblance discrétisée

$$\hat{h}_{(d)} = \arg\max_h \left( \sum_{j=0}^{m-1} D_{j+1} \log \hat{\alpha}^{(-)}_{(d),h}(X_{t_j}) - \sum_{j=0}^{m-1} \Delta t_{j+1} \hat{\alpha}^{(-)}_{(d),h}(X_{t_j}) \right)$$

avec

$$\hat{\alpha}^{(-)}_{(d),h}(X_{t_j}) = \frac{\sum_{i=0, i\neq j}^{m-1} D_{i+1} k_T\left(\frac{d(X_{t_j}, X_{t_i})}{h}\right)}{\sum_{i=0, i\neq j}^{m-1} \Delta t_{i+1} k_T\left(\frac{d(X_{t_j}, X_{t_i})}{h}\right)}.$$

# 4 Perspectives de recherche

On a présenté ici des résultats d'estimation de la fonction d'intensité de sauts $\alpha$ à l'aide d'observation d'une trajectoire discrétisée. Le modèle des processus jump-move étant régi par la fonction $\alpha$, le noyau de transition des sauts $K$ et le mouvement $Y$, il est envisagé dans la lignée de ces travaux sur l'estimation de $\alpha$ de considérer le problème d'estimation du noyau $K$. Concernant le mouvement $Y$, on peut se demander comment retrouver statistiquement ses caractéristiques. Cette question d'estimation revenant peu ou prou à celle de l'estimation d'un processus de Markov continu (qui est vaste), néanmoins pertubé ici par la présence de sauts, celle de tester des caractéristiques de $Y$ peut se révéler aussi très pertinente. Dans le cas d'un système de particules, celles-ci sont principalement caractérisées par leur équation de mouvement. Mettre en place des tests utilisant le modèle des processus jump-move en vue par exemple d'une classification des particules comme dans Briane et al. (2019) peut permettre de mieux prédire le système. Le processus jump-move permet également de modéliser facilement des systèmes de particules marquées. Classifier et/ou estimer statistiquement ces marques est un enjeu également intéressant dans la compréhension d'un tel système.

Enfin, pour simplifier l'étude globale du processus, on peut également se demander si le modèle n'est pas mieux décrit par un modèle paramétrique. L'étude statistique de tels modèles commence à être menée depuis peu. Toujours dans l'espoir de simplifier le modèle, on pourra chercher à déterminer dans un premier temps un test pour savoir si la fonction $\alpha$ n'est pas en réalité constante. Ceci ouvrirait la voie dans un second temps à la question de la détection de rupture d'intensité des sauts, problème formalisé plus précisément dans le Chapitre 4 pour le processus de Poisson.

# Chapitre 4

# Rupture d'intensité dans un processus de Poisson

Ce chapitre est issu du co-encadrement de la thèse de Fabrice Grela avec Magalie Fromont.

## 1 Détection d'une rupture par test simple

Un nombre gigantesque de travaux de recherche, aussi bien appliqués que théoriques, considèrent un signal changeant brutalement de caractéristique suite à un événement quelconque. Cette rupture trouve des interprétations naturelles dans de nombreux domaines, comme l'analyse du réseau internet, l'analyse de notre ADN ou encore l'étude d'attaques cyberterroristes. On trouvera un panorama général de cette large thématique dans Basseville and Nikiforov (1993) par exemple. Dans leur analyse de séquences génétiques, Olshen et al. (2004) considèrent le problème de la détection de rupture dans un modèle gaussien appliqué au logarithme des intensités empiriques des marqueurs d'ADN considérés. Cependant c'est le processus de Poisson inhomogène avec une intensité constante par morceaux qui est introduit dans Shen and Zhang (2012) pour modéliser le nombre de copies de gènes et répondre au problème biologique de la détection d'anomalies. Le processus de Poisson est également fréquemment utilisé dans le processus de contrôle de la qualité d'un produit en sortie d'usine pour le comptage de produits non-conformes, la rupture correspondant alors à un changement dans le procédé de fabrication qu'il convient de détecter correctement (voir par exemple Perry and Pignatiello (2011) et ses références).

La motivation première des travaux présentés dans cette section se trouve dans la détection d'anomalies relatives à la cybersécurité. Là encore le processus de Poisson y joue un rôle important (voir par exemple Holm (2013) ou Daras (2014) et ses références), servant par exemple de modèle pour des communications secrètes (Wang, 2018; Soltani et al., 2015, 2020). Le changement d'intensité dans le flux de paquets internet pouvant signifier une activité suspecte (Soltani et al., 2008), de nombreuses méthodes sont mises en œuvre afin de détecter au mieux ce changement d'activité sur le réseau (Peng et al., 2004; Zhan et al., 2013; Liu et al., 2021).

Nous nous sommes intéréssés spécifiquement au cas du processus de Poisson qui fournit un premier modèle simplifié des échanges internet par protocole TCP, protocole dont l'étude est essentielle dans le domaine de la cybersécurité. Cependant de nombreuses études empiriques considèrent des signaux plus larges de différentes formes, allant de l'échantillon de loi quelconque pour la plus générale (Kendall and Kendall, 1980) à la loi exponentielle intervenant dans le processus de Poisson (Bain et al., 1985; Engelhardt et al., 1990; Cohen and Sackrowitz, 1993; Ho, 1993, 1995; Green, 1995), en passant par l'échantillon gaussien qui est à la base d'un nombre conséquent de résultats, dans des modèles de régression linéaire par exemple (Bai and Perron, 2003; Picard et al., 2011), ou encore pour l'analyse de signaux autosimilaires (Veitch and Abry, 2001).

### De l'estimation d'une rupture

Beaucoup de travaux abordent la question d'un changement de régime sous l'angle de l'estimation des paramètres modélisant la rupture du modèle, et prioritairement l'estimation de ses instants de rupture (on pourra se référer à Lee (2010) ou Aue and Horváth (2013) pour des éléments bibliographiques sur ce sujet). L'approche bayésienne notamment s'est beaucoup développée autour de cette question, proposant de nombreuses méthodes d'estimation pour différentes lois a priori et différents modèles de vraisemblance. Par exemple pour le modèle de Poisson, des lois a posteriori de l'instant de rupture sont étudiées dans Leonard (1978), Raftery and Akman (1986), Carlin et al. (1992), Raftery (1994), Assareh et al. (2016) ou encore plus récemment dans Farinetto et al. (2020) pour ne citer qu'eux, et par exemple dans Rigaill (2010) pour le modèle gaussien.

Plus classiquement, l'estimation du nombre et des instants de rupture dans une série temporelle est abordée dans Bai (1994) pour des processus ARMA et plus généralement pour des processus linéaires différence de martingales dans le cas d'une rupture, puis étendue au cas de plusieurs ruptures dans Bai and Perron (1998), Lavielle and Moulines (2000) ou encore Lavielle (2005) dans le cas i.i.d. Des intervalles de confiance sont obtenus dans Worsley (1986); Hinkley (1970); Hinkley and Hinkley (1970); Siegmund (1988) dans le cas plus spécifique d'une famille exponentielle de lois. Le cas du processus de Poisson est moins présent pour ce qui est du point de vue de l'estimation non bayésienne. On pourra citer néanmoins les travaux de Galeano (2007) et Perry and Pignatiello (2011) à ce sujet.

Comme souvent le cas de la loi normale fait l'objet d'un traitement particulier. Yao (1988) estime ainsi de manière asymptotiquement consistante le nombre de ruptures dans l'espérance d'un échantillon gaussien. Adoptant le point de vue de la sélection de modèle non-paramétrique de Birgé and Massart (2001), Lebarbier (2005) propose une estimation non-asymptotique pénalisée par moindre carrés du nombre et de la localisation de changements dans l'espérance d'un processus gaussien (voir aussi Harchaoui and Lévy-Leduc (2010) pour une amélioration quasi-optimale). Ce sont des méthodes de sélection forward qui sont utilisées le plus souvent pour l'estimation de plusieurs ruptures (Fryzlewicz, 2014, 2018), et non des méthodes backward (sauf dans Shin et al. (2020)). Le cas de la dimension supérieure est abordé quant à lui dans Wang and Samworth (2018), alors que les situations d'observations corrélées sont traitées plus rarement. On trouve néanmoins des résultats récents dans Dette et al. (2020) ou encore dans Benassi et al. (2000) dans le cas d'une observation en temps discret d'une trajectoire de mouvement Brownien fractionnaire.

On obtient également des résultats optimaux dans le cadre gaussien. C'est le cas par exemple dans Donoho and Johnstone (1998) qui propose une estimation minimax du drift dans le modèle de bruit blanc observé en des temps discrets $i/n$. Dans le cadre de la régression gaussienne, des premiers résultats sur des intervalles de confiance en lien avec des procédures de test minimax sont obtenues dans Nickl and van de Geer (2013), des taux d'estimation minimax sont donnés par Yao (1988) et Cai and Guo (2017), alors que le cas de l'estimation minimax du support d'un bump est traité dans Kou (2023). Wang et al. (2020) étendent ces résultats au cadre sous-gaussien avec des résultats minimax non-asymptotiques pour l'estimation des instants de rupture de l'espérance. Plus généralement, Garreau and Arlot (2018) fournissent une estimation minimax de changements multiples en grande dimension dans un cadre non-paramétrique, pour des caractéristiques d'une distribution qui peuvent être différentes de l'espérance ou de la variance.

#### A la détection de rupture

Il y a plusieurs manières de considérer la question de la détection d'un changement de comportement d'un processus stochastique. L'histoire commence principalement avec Wald (1945) qui aborde le problème de façon online, c'est-à-dire de manière séquentielle et en temps réel, et qui se donne comme objectif de construire un temps d'arrêt aussi proche que possible du temps inconnu de rupture dans la loi. On pourra se référer à Siegmund (1985); Csorgo and Horváth (1997); Lai (2001); Moustakides et al. (2008); Polunchenko and Tartakovsky (2012) pour une review sur les méthodes online.

Le point de vue bayésien a largement été adopté pour aborder ce problème, notamment par Shiryaev (1978) dans le cadre du modèle de bruit blanc gaussien, montrant qu'il existe une règle de décision minimisant le risque bayésien dans le cas où le changement s'effectue en un temps $\tau$ inconnu. Ces résultats seront ensuite étendus en particulier par Pollak (1985) dans un cadre de loi plus générale, puis également par Peskir and Shiryaev (2002), Herberts and Jensen (2004), Bayraktar et al. (2006) et Brown and Zacks (2006) dans le cadre du modèle de Poisson.

Sans considérer de modèle particulier, Wald (1945) expose une stratégie séquentielle de prise de décision qui consiste à décider $H_0$, $H_1$ ou une troisième décision qui augmente le nombre d'observations d'une unité, et réitérer jusqu'à décider finalement $H_0$ ou $H_1$, tout en respectant le principe de Neyman-Pearson : une erreur de première espèce contrôlée par $\alpha$ et une erreur de seconde espèce contrôlée par $\beta$, pour deux valeurs de $\alpha$ et $\beta$ prescrites. Cette stratégie est basée sur le processus Log Probability Ratio qui sera largement réutilisé ensuite, notamment par Page (1954), ou encore par Dvoretzky et al. (1953) qui étudient la méthode de Wald (1945) dans le cas d'un processus à temps continu comme le mouvement Brownien ou le processus de Poisson. Lorden (1971) introduit alors un critère d'optimalité de type minimax pour la détection online de rupture dans une série i.d.d. à temps discret, critère utilisé par Moustakides (1986) pour montrer que la procédure de test de Page (1954) est optimale en ce sens. Ce problème est également traité en temps continu, par Moustakides (2004) pour le modèle de bruit blanc et par El Karoui et al. (2017) dans le cas du processus de Poisson.

Un autre point de vue consiste à effectuer une analyse des données offline, c'est-à-dire a posteriori mais sans le formalisme bayésien, point de vue que nous avons adopté dans [A12] pour le modèle du processus de Poisson. De manière plus générale, Page (1955) et Page (1957) proposent des tests de changement de paramètre d'un échantillon de loi quelconque, tests dont les puissances seront ensuite comparées avec les modèles bayésiens binomiaux et gaussiens dans Chernoff and Zacks (1964), puis dans Kander and Zacks (1966) pour des familles exponentielles de lois. Le cas de la détection de rupture dans un échantillon de loi Bernoulli est abordé d'un point de vue non bayésien dans Hinkley and Hinkley (1970), le cas de la loi normale étant considéré par exemple dans Hinkley (1970); Hawkins (1977); James et al. (1987) et celui de la loi exponentielle dans Antoch and Jaruvskova (2007).

### La détection de rupture dans le cas du processus de Poisson

On trouve des premiers résultats de tests sur le modèle de Poisson, d'adéquation à la loi exponentielle ou directement au processus de Poisson avec les travaux de Epstein (1960), Lewis (1965), Cox and Lewis (1966) ou encore Davies (1977). La question à proprement parlée de la détection d'une rupture dans le modèle de Poisson est ensuite abordée par Deshayes (1984), qui considère le cas général où tous les paramètres sont inconnus. En conditionnant par le nombre de points observés, il étudie en réalité le test de Kolmogorov Smirnov pour la loi uniforme et le test du rapport de vraisemblance qui est asymptotiquement optimal lorsque l'intensité totale tend vers l'infini. Deux autres tests seront ensuite considérés dans Deshayes and Picard (1985), tests qui ne sont pas globalement optimaux. Leur optimalité est alors étudiée de manière asymptotique au sens de Bahadur (1966) et Brown (1971) (on recherche une décroissance exponentielle des erreurs de type I et II pour chaque couple d'hypothèses $H_0$ et $H_1$), ainsi qu'au sens de la théorie locale de Le Cam (1970). Plusieurs tests classiques pour tester une intensité constante contre une intensité croissante sont introduits dans Bain et al. (1985) pour une comparaison empirique de leur puissance. On y retrouve le test de Laplace, le test Z qui est le test UPP dans le cas paramétrique du processus de Weibull, le test du rapport de vraisemblance ou encore le test F. Une variante du test Z est également introduite et étudiée de manière non-conditionnelle en temps long dans Bhattacharjee et al. (2004) pour des alternatives croissantes. Une approximation asymptotique du niveau et de la puissance du test de rapport de vraisemblances est développée dans Loader (1990) pour ce cas où aucun paramètre n'est connu. Le test du rapport de vraisemblance enfin est utilisé par Dachian and Kutoyants (2006) dans le cas où l'intensité constante est connue sous $H_0$ pour construire des tests asymptotiquement uniformément plus puissants au sens de Le Cam pour des alternatives composées de processus ponctuels stationnaires auto-excitants.

### L'optimalité au sens du minimax

Afin de comparer les performances des différents tests de détection de rupture, on peut utiliser des critères d'optimalité au sens du minimax. On se fixe deux bornes $\alpha$ et $\beta$ pour les probabilités d'erreurs. D'un point de vue asymptotique, on considère une suite de tests paramétriques du type $H_0^n$ vs $H_1^n$. Pour une statistique de test $\psi_n$, on peut considérer ses erreurs de type $I$ et $II$

$$\alpha_n(\psi_n) = \sup_{\theta \in H_0^n} \mathbb{P}_\theta(\psi_n = 1) \quad \text{et} \quad \beta_n(\psi_n) = \sup_{\theta \in H_1^n} \mathbb{P}_\theta(\psi_n = 0).$$

On dit alors que la statistique $\psi_n$ est asymptotiquement minimax si $\alpha_n(\psi_n) \leq \alpha + o(1)$ et $\beta_n(\psi_n) \leq \beta + o(1)$. En supposant que $H_1^n$ dépende d'un paramètre de distinguabilité $\varepsilon_n = o(1)$ (i.e. tel que $H_1^n \cap H_0^n = \varnothing$ si $\varepsilon_n > 0$ et $H_1^n \cap H_0^n \neq \varnothing$ si $\varepsilon_n = 0$), on s'intéresse le plus souvent au problème de distinguabilité suivant : trouver $\underline{\varphi}_n = o(1)$ une borne inférieure de détection minimax (i.e. telle qu'il n'existe pas de test $\psi_n$ asymptotiquement minimax pour $H_1^n(\underline{\varphi}_n)$) et $\overline{\varphi}_n = o(1)$ une borne supérieure de détection minimax (i.e. telle qu'il existe un test $\psi_n$ asymptotiquement minimax pour $H_1^n(\overline{\varphi}_n)$). Si $\overline{\varphi}_n \backslash \underline{\varphi}_n = O(1)$, on parle alors de vitesse de séparation minimax pour $\varphi_n$.

D'un point de vue non-asymptotique, on suppose que les observations sont issues d'une loi $\mathbb{P}_\lambda$ indicée par une intensité $\lambda$, on se donne deux espaces $\mathcal{S}_0$ et $\mathcal{S}_1$ munis d'une distance $d$ et on considère un test du type $H_0 : \lambda \in \mathcal{S}_0$ vs $H_1 : \lambda \in \mathcal{S}_1$. Pour un test $\psi$, son erreur de type I est $\alpha(\psi) = \sup_{\lambda \in \mathcal{S}_0} \mathbb{P}_\lambda(\psi = 1)$. On définit alors, pour un test $\psi_\alpha$ de niveau $\alpha$ (i.e. vérifiant $\alpha(\psi_\alpha) \leq \alpha$), son taux de séparation $\beta$-uniforme pour l'alternative $\mathcal{S}_1$ par

$$SR_\beta(\psi_\alpha, \mathcal{S}_1) = \inf\{r > 0 : \sup_{\lambda \in \mathcal{S}_1, d(\lambda, \mathcal{S}_0) \geq r} \mathbb{P}_\lambda(\psi_\alpha = 0) \leq \beta\}.$$

La vitesse de séparation minimax pour l'alternative $\mathcal{S}_1$ est alors

$$mSR_{\alpha,\beta}(\mathcal{S}_1) = \inf\{SR_\beta(\psi_\alpha, \mathcal{S}_1) : \alpha(\psi_\alpha) \leq \alpha\}. \tag{4.1}$$

Les résultats d'optimalité au sens du minimax apparaissent principalement dans le cas du modèle gaussien, considéré notamment par Spokoiny (1996) et Lepski and Tsybakov (2000), où l'observation consiste en une trajectoire en temps continu d'une solution de l'équation $dX_t = f(t)dt + \varepsilon dB_t$ pour $B$ un mouvement Brownien. Se pose alors la question de l'adaptation en la régularité de $f$ pour tester $H_0 : f \equiv 0$ vs $H_1 : \|f\|_2 \geq q(\varepsilon)$ avec $f$ appartenant à un espace de Besov ou Sobolev. On cherche à déterminer ici $q(\varepsilon)$ tel que le problème soit soluble avec des erreurs de première et seconde espèce prescrites, la vitesse minimax étant pour l'asymptotique $\varepsilon \to 0$. Dümbgen and Spokoiny (2001) reprennent ce cadre en considérant des fonctions dans l'alternative $f$ négatives, décroissantes ou concaves. Des résultats minimax non-asymptotiques sont également établis dans Ermakov (1991), Lepski and Spokoiny (1999), Baraud (2002) ou encore Verzelen et al. (2020) selon l'espace d'alternatives considérées. Dans le cas de variables gaussiennes corrélées, le problème de la détection de rupture dans l'espérance est abordé par exemple dans Hall and Jin (2010) et Enikeeva et al. (2018).

Des bornes asymptotiques minimax pour le problème de la détection d'un bump dans l'espérance d'un signal gaussien non corrélé sont établies dans Arias-Castro et al. (2005), Chan and Walther (2013) et complétées par Gao et al. (2020) (voir aussi Frick et al. (2014)), le cas d'un processus gaussien stationnaire avec corrélations étant considéré dans Enikeeva et al. (2020). Ces résultats sont ensuite étendus au cas de plusieurs segments dans Jeng et al. (2010) et au cas d'un processus gaussien hétérogène dans Pein et al. (2017). Brunel (2014) établit la différence de vitesse de séparation minimax entre la détection d'une rupture et la détection d'un segment dans un cadre sous-gaussien. La vitesse $\log n/n$ apparaît ainsi dans le problème d'adaptation à la position de la rupture pour la détection d'un segment, au lieu de la vitesse $1/n$ pour une seule rupture.

**Une détection minimax non-asymptotique pour un processus de Poisson**

Pour le processus de Poisson, il existe semble-t-il peu de résultats minimax. Des tests minimax d'adéquation à un processus de Poisson homogène sont développés dans Ingster and Kutoyants (2007) pour le point de vue asymptotique et dans Fromont et al. (2011) pour le point de vue non-asymptotique. La détection d'un bump dans l'intensité du processus de Poisson observé est étudiée du point de vue minimax dans Rivera and Walther (2013), après conditionnement par le nombre de points observés on y retrouve une vitesse en $\log n/n$.

Dans [A12], nous nous sommes placés dans le cadre de l'étude minimax non-asymptotique du problème de détection de rupture d'intensité dans un processus de Poisson, pour une détection a posteriori (offline), aussi bien d'une seule rupture que d'un segment, en considérant les cas d'une intensité de base connue et inconnue. Nous déterminons en particulier la vitesse de séparation minimax pour les problèmes de détection d'une rupture et les problèmes de détection d'un segment, et ce dans tous les cas de figure : une intensité de base connue ou inconnue, avec une hauteur, une longueur et une position de saut connues et/ou inconnues.

On obtient ainsi une vision complète des différentes vitesses de séparation minimax dans ce problème de détection, faisant apparaître différentes transition de phases selon les différents espaces considérés. Pour la détection d'une seule rupture, la vitesse de séparation est $\sqrt{1/L}$ (en supposant que l'intensité soit proportionnelle à un paramètre d'échelle $L$) à l'exception du cas où ni la hauteur ni la position du saut ne sont connues, cas où la vitesse de séparation minimax devient $\sqrt{\log\log L/L}$. Un troisième régime apparaît dans le problème de détection d'un segment, régime pour lequel la vitesse est $\sqrt{\log L/L}$. Il correspond à la situation où ni la position ni la longueur du segment ne sont connues. Il existe également une situation intermédiaire avec une vitesse en $\sqrt{\log\log L/L}$ pour la détection d'un segment, lorsque ni la hauteur ni la longueur du segment ne sont connues mais avec une position initiale connue.

Enfin les autres cas présentent une vitesse paramétrique classique en $\sqrt{1/L}$.

On suppose donc que l'on observe un processus de Poisson $(N_t)_{t\in[0,1]}$ sur l'intervalle $[0,1]$, d'intensité $\lambda$ par rapport à la mesure de Lebesgue $Ldt$, où $L>0$. Lorsque $L$ est un entier, le processus $N$ peut s'interpréter comme la somme de $L$ processus de Poisson indépendants de même intensité $\lambda$ par rapport à $dt$. On peut donc penser à $L\to\infty$ si l'on souhaite se comparer aux résultats asymptotiques.

### La détection avec une intensité de base connue

Dans ce cas de figure on se donne $\lambda_0>0$ et on considère des tests pour lesquels $H_0$ correspond à un processus de Poisson homogène d'intensité $\lambda_0$ : $H_0 : \lambda\in\mathcal{S}_0[\lambda_0]=\{\lambda_0\}$. Pour les espaces d'alternatives $\mathcal{S}_1$, on introduit pour $\delta^*\in(-\lambda_0,+\infty)\backslash\{0\}$, $\tau^*\in(0,1)$ et $l^*\in(0,1-\tau^*]$ l'ensemble (qui ne contient qu'un seul élément) :

$$\mathcal{S}_{\delta^*,\tau^*,l^*}[\lambda_0]=\{\lambda:[0,1]\to(0,\infty),\forall t\in[0,1],\lambda(t)=\lambda_0+\delta^*\mathbf{1}_{(\tau^*,\tau^*+l^*]}(t)\}$$

correspondant à un segment de rupture de hauteur $|\delta^*|$ à l'instant $\tau^*$ pour une durée $l^*$. Ensuite en remplaçant $\delta^*,\tau^*$ et/ou $l^*$ par $\cdot$ on indique la connaissance ou non du paramètre. Ainsi par exemple

$$\mathcal{S}_{\delta^*,\cdot,\cdot}[\lambda_0]=\{\lambda:\exists\tau\in(0,1),\exists l\in(0,1-\tau)\text{ tq }\lambda\in\mathcal{S}_{\delta^*,\tau,l}[\lambda_0]\}$$

correspond au cas de l'alternative $\mathcal{S}_1=\mathcal{S}_{\delta^*,\cdot,\cdot}[\lambda_0]$ où la position et la longueur du segment sont inconnus, avec une hauteur (connue) égale à $|\delta^*|$. Le cas d'une seule rupture correspond au cas $l=1-\tau$ et s'écrit simplement $\mathcal{S}_{\delta^*,\cdot,1-\cdot}[\lambda_0]$. On introduit enfin

$$\mathcal{S}_{\cdot,\tau^*,\cdot}[\lambda_0,R]=\mathcal{S}_{\cdot,\tau^*,\cdot}[\lambda_0]\cap\{\lambda:[0,1]\to(0,R]\}$$

car $\mathcal{S}_{\cdot,\tau^*,\cdot}[\lambda_0]$, trop grand, vérifie $mSR_{\alpha,\beta}(\mathcal{S}_{\cdot,\tau^*,\cdot}[\lambda_0])=+\infty$.

Les vitesses de séparation minimax sont calculées sur ces alternatives pour la distance $L^2$ usuelle : $d_2^2(f,g)=\int_0^1|f(t)-g(t)|^2dt$. On sait alors que pour $\lambda=\lambda_0+\delta\mathbf{1}_{(\tau,\tau+l]}$,

$$d_2(\lambda,\mathcal{S}_0[\lambda_0])=|\delta|\sqrt{l}.$$

Pour obtenir des statistiques de tests minimax, on peut se servir dans chaque cas d'une statistique linéaire en $N$. On a ainsi utilisé selon les cas

$$N(\tau,\tau+l]=N_{\tau+l}-N_\tau,$$

$$\max_{\tau\in[0,1-l]}N(\tau,\tau+l]\text{ et }\min_{\tau\in[0,1-l]}N(\tau,\tau+l],$$

$$S_{\tau_1,\tau_2}=N(\tau_1,\tau_2]-\lambda_0(\tau_2-\tau_1)L$$

et

$$S_{\delta,\tau_1,\tau_2}=\operatorname{sgn}(\delta)S_{\tau_1,\tau_2}-|\delta|L(\tau_2-\tau_1)/2.$$

On peut également utiliser presque dans chaque cas une statistique quadratique en $N$, du type de

$$T_{\tau_1,\tau_2} = \frac{1}{L^2(\tau_2-\tau_1)}(N(\tau_1,\tau_2]^2 - N(\tau_1,\tau_2]) - \frac{2\lambda_0}{L}N(\tau_1,\tau_2] + \lambda_0^2(\tau_2-\tau_1)$$

qui correspond à une estimation sans biais de la norme $L^2$ au carré de la projection orthogonale de $\lambda-\lambda_0$ sur Vect$(\mathbf{1}_{(\tau_1,\tau_2]})$, et qui présente de meilleurs résultats empiriques dans certaines situations (voir [A12] pour plus de détails). On obtient alors les vitesses de séparation minimax suivantes :

| **Détection d'une rupture simple** | | |
|---|---|---|
| Alternative | $\mathrm{mSR}_{\alpha,\beta}$ | Statistique de test |
| $\mathcal{S}_{\delta^*,\tau^*,1-\tau^*}[\lambda_0]$ | - | $N(\tau^*,1]$ |
| $\mathcal{S}_{\cdot,\tau^*,1-\tau^*}[\lambda_0]$ | $L^{-1/2}$ | $N(\tau^*,1]$ |
| $\mathcal{S}_{\delta^*,\cdot,1-\cdot}[\lambda_0]$ | $L^{-1/2}$ | $\sup_{\tau\in(0,1)} S_{\delta^*,\tau,1}(N)$ |
| $\mathcal{S}_{\cdot,\cdot,1-\cdot}[\lambda_0,R]$ | $\sqrt{\frac{\log\log L}{L}}$ | $\max_{k\in\{1,\dots,\lfloor\log_2 L\rfloor\}}\left(\left\lvert S_{1-2^{-k},1}(N)\right\rvert - s_{\lambda_0,1-2^{-k},1}(1-u_\alpha)\right)$ |

| **Détection d'un segment** | | |
|---|---|---|
| Alternative | $\mathrm{mSR}_{\alpha,\beta}$ | Statistique de test |
| $\mathcal{S}_{\delta^*,\tau^*,\ell^*}[\lambda_0]$ | - | $N(\tau^*,\tau^*+\ell^*]$ |
| $\mathcal{S}_{\cdot,\tau^*,\ell^*}[\lambda_0]$ | $L^{-1/2}$ | $N(\tau^*,\tau^*+\ell^*]$ |
| $\mathcal{S}_{\delta^*,\cdot,\ell^*}[\lambda_0]$ | - | $\max_{\tau\in[0,1-\ell^*]} N(\tau,\tau+\ell^*]$, $\min_{\tau\in[0,1-\ell^*]} N(\tau,\tau+\ell^*]$ |
| $\mathcal{S}_{\cdot,\cdot,\ell^*}[\lambda_0]$ | $L^{-1/2}$ | $\max_{\tau\in[0,1-\ell^*]} N(\tau,\tau+\ell^*]$, $\min_{\tau\in[0,1-\ell^*]} N(\tau,\tau+\ell^*]$ |
| $\mathcal{S}_{\delta^*,\tau^*,\cdot}[\lambda_0]$ | $L^{-1/2}$ | $\sup_{\ell\in(0,1-\tau^*)} S_{\delta^*,\tau^*,\tau^*+\ell}(N)$ |
| $\mathcal{S}_{\cdot,\tau^*,\cdot}[\lambda_0,R]$ | $\sqrt{\frac{\log\log L}{L}}$ | $\max_{k\in\{1,\dots,\lfloor\log_2 L\rfloor\}}\left(\left\lvert S_{\tau^*,\tau^*+(1-\tau^*)2^{-k}}(N)\right\rvert - s_{\lambda_0,\tau^*,\tau^*+(1-\tau^*)2^{-k}}(1-u_\alpha)\right)$ |
| $\mathcal{S}_{\delta^*,\cdot,\cdot}[\lambda_0]$ | $\sqrt{\frac{\log L}{L}}$ | $\max_{k\in\{0,\dots,\lceil L\rceil-1\},k'\in\{1,\dots,\lceil L\rceil-k\}}\left(\left\lvert S_{\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}(N)\right\rvert - s_{\lambda_0,\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}(1-u_\alpha)\right)$ |
| $\mathcal{S}_{\cdot,\cdot,\cdot}[\lambda_0,R]$ | $\sqrt{\frac{\log L}{L}}$ | $\max_{k\in\{0,\dots,\lceil L\rceil-1\},k'\in\{1,\dots,\lceil L\rceil-k\}}\left(\left\lvert S_{\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}(N)\right\rvert - s_{\lambda_0,\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}(1-u_\alpha)\right)$ |

On retrouve les trois régimes avec la vitesse $\sqrt{1/L}$, la vitesse $\sqrt{\log\log L/L}$ dûe à l'adaptation à la hauteur et la longueur de la rupture, et la vitesse $\sqrt{\log L/L}$ dûe à l'adaptation à la position et la longueur du saut dans le cas d'un segment.

### La détection avec une intensité de base inconnue

Dans ce cas de figure, on considère des tests pour lesquels $H_0$ correspond à un processus de Poisson homogène avec une intensité $\lambda_0$ (inconnue) inférieure à un $R>0$ donné (pour ne pas avoir de vitesses de séparation infinies), $\lambda$ appartenant donc à cet ensemble noté $\mathcal{S}_0^u[R]$. Les espaces d'alternatives deviennent de la forme

$$\mathcal{S}^u_{\delta^*,\tau^*,l^*}[R] = \{\lambda : [0,1] \to (0,R] : \exists \lambda_0 \in (-\delta^* \vee 0, (R-\delta^*) \wedge R] \text{ tq } \lambda \in \mathcal{S}_{\delta^*,\tau^*,l^*}[\lambda_0, R]\}$$

accompagnés de ses équivalents avec des $\cdot$ dans le cas de paramètres inconnus. Pour $\lambda = \lambda_0 + \delta \mathbf{1}_{(\tau,\tau+l]}$, on a dans ce cas

$$d_2(\lambda, \mathcal{S}^u_0[R]) = |\delta|\sqrt{l(1-l)}.$$

Là encore on peut utiliser systématiquement des statistiques linéaires en $N$ avec un ajustement pour la statistique $S$ (du fait de la non-connaissance de $\lambda_0$) qui devient

$$S'_{\tau_1,\tau_2} = N(\tau_1,\tau_2] - (\tau_2-\tau_1)N_1$$

et

$$S'_{\delta,\tau_1,\tau_2} = \text{sgn}(\delta) S'_{\tau_1,\tau_2} - |\delta| L(\tau_2-\tau_1)(1-\tau_2+\tau_1)/2.$$

On peut également ajuster les statistiques quadratiques en $N$ et obtenir des statistiques minimax. On obtient dans ce cas les vitesses de séparation suivantes :

| **Détection d'une rupture simple** | | |
|---|---|---|
| Alternative | $\text{mSR}_{\alpha,\beta}$ | Statistique de test |
| $\mathcal{S}^u_{\delta^*,\tau^*,1-\tau^*}[R]$ | - | $N(\tau^*,1]$ |
| $\mathcal{S}^u_{\cdot,\tau^*,1-\tau^*}[R]$ | $L^{-1/2}$ | $N(\tau^*,1]$<br>$T'_{\tau^*,1}(N)$ |
| $\mathcal{S}^u_{\delta^*,\cdot\cdot,1-\cdot\cdot}[R]$ | $L^{-1/2}$ | $\sup_{\tau\in(0,1)} S'_{\delta^*,\tau,1}(N)$ |
| $\mathcal{S}^u_{\cdot,\cdot\cdot,1-\cdot\cdot}[R]$ | $\sqrt{\frac{\log\log L}{L}}$ | $\max_{\tau\in\mathcal{D}_L}\left(\left\lvert S'_{\tau,1}(N)\right\rvert - s'_{N_1,\tau,1}(1-u_\alpha)\right)$<br>$\mathcal{D}_L = \left\{2^{-k},\ k\in\{2,\dots,\lfloor\log_2(L)\rfloor\}\right\} \cup \left\{1-2^{-k},\ k\in\{1,\dots,\lfloor\log_2(L)\rfloor\}\right\}$ |

| **Détection d'un segment** | | |
|---|---|---|
| Alternative | $\text{mSR}_{\alpha,\beta}$ | Statistique de test |
| $\mathcal{S}^u_{\delta^*,\tau^*,\ell^*}[R]$ | - | $N(\tau^*,\tau^*+\ell^*]$ |
| $\mathcal{S}^u_{\cdot,\tau^*,\ell^*}[R]$ | $L^{-1/2}$ | $N(\tau^*,\tau^*+\ell^*]$ |
| $\mathcal{S}^u_{\delta^*,\cdot\cdot,\ell^*}[R]$ | - | $\max_{\tau\in[0,1-\ell^*\wedge(1/2)]} N(\tau,\tau+\ell^*\wedge(1/2)]$,<br>$\min_{\tau\in[0,1-\ell^*\wedge(1/2)]} N(\tau,\tau+\ell^*\wedge(1/2)]$<br>$M = \lceil 2/(\ell^*(1-\ell^*))\rceil$ |
| $\mathcal{S}^u_{\cdot,\cdot\cdot,\ell^*}[R]$ | $L^{-1/2}$ | $\max_{\tau\in[0,1-\ell^*\wedge(1/2)]} N(\tau,\tau+\ell^*\wedge(1/2)]$,<br>$\min_{\tau\in[0,1-\ell^*\wedge(1/2)]} N(\tau,\tau+\ell^*\wedge(1/2)]$ |
| $\mathcal{S}^u_{\delta^*,\tau^*,\cdot\cdot\cdot}[R]$ | $L^{-1/2}$ | $\sup_{\ell\in(0,1-\tau^*)} S'_{\delta^*,\tau^*,\tau^*+\ell}(N)$ |
| $\mathcal{S}^u_{\cdot,\tau^*,\cdot\cdot\cdot}[R]$ | $\sqrt{\frac{\log\log L}{L}}$ | $\max_{k\in\{1,\dots,\lfloor\log_2 L\rfloor\}}\left(\left\lvert S'_{\tau^*,\tau^*+\frac{1-\tau^*}{2^k}}(N)\right\rvert - s'_{N_1,\tau^*,\tau^*+\frac{1-\tau^*}{2^k}}(1-u_\alpha)\right)$ |
| $\mathcal{S}^u_{\delta^*,\cdot\cdot,\cdot\cdot\cdot}[R]$ | $\sqrt{\frac{\log L}{L}}$ | $\max_{k\in\{0,\dots,\lceil L\rceil-1\},k'\in\{1,\dots,\lceil L\rceil-k\}}\left(\left\lvert S_{\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}(N)\right\rvert - s_{\lambda_0,\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}(1-u_\alpha)\right)$ |
| $\mathcal{S}^u_{\cdot,\cdot\cdot,\cdot\cdot\cdot}[R]$ | $\sqrt{\frac{\log L}{L}}$ | $\max_{k\in\{0,\dots,\lceil L\rceil-1\},k'\in\{1,\dots,\lceil L\rceil-k\}}\left(\left\lvert S'_{\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}(N)\right\rvert - s'_{N_1,\frac{k}{\lceil L\rceil},\frac{k+k'}{\lceil L\rceil}}\left(1-u^{(1)}_\alpha\right)\right)$ |

On remarque que l'adaptation en $\lambda_0$ modifie les techniques de calcul mais n'affecte pas les vitesses de séparation minimax qui restent inchangées par rapport au cas où $\lambda_0$ est connu.

# 2 Localisation d'une rupture par tests multiples

Les procédures de test précédentes conduisent à la conclusion éventuelle qu'une rupture à eu lieu durant l'intervalle de temps $[0, 1]$. Nous nous sommes intéressés dans [A14] à la question de la localisation d'une éventuelle rupture formulée à l'aide d'un problème de test multiple. On suppose toujours que l'on observe une trajectoire $(N_t)_{t\in[0,1]}$ d'intensité $\lambda$ par rapport à la mesure $Ldt$ avec $\lambda$ appartenant à l'ensemble $\overline{\mathcal{S}}$ des fonctions de la forme $\lambda = \lambda_0 + \delta\mathbf{1}_{(\tau,1]}$ (et $\lambda = \lambda_0$ si $\tau \geq 1$). On considère pour un entier non nul $M$ une collection d'hypothèses $\mathcal{H} = \{H_k, k = 1, ..., M\}$ avec $H_k \subset \overline{\mathcal{S}}$. Pour $\lambda \in \overline{\mathcal{S}}$, on notera $\mathcal{T}(\lambda) = \{H_k \in \mathcal{H} : \lambda \in H_k\}$ l'ensemble des hypothèses vérifiées par $\lambda$ et $\mathcal{F}(\lambda) = \mathcal{H}\backslash\mathcal{T}(\lambda)$ l'ensemble des hypothèses fausses pour $\lambda$. Une procédure de test multiple $\mathcal{R}$ est alors un sous-ensemble aléatoire de $\mathcal{H}$ dont l'objectif est d'inférer $\mathcal{F}(\lambda)$, indiquant les hypothèses rejetées suite à l'observation de $N$.

D'un point de vue pratique, plusieurs études ont pour objectif d'estimer la proportion d'hypothèses vraies, i.e. d'estimer la quantité $\pi_0 = |\mathcal{T}(\lambda)|/M$ à l'aide de l'observation des p-values $P_i, ..., P_M$ des différents tests $H_i$. En supposant que $P_i$ possède la densité $h$ sous $H_{1i}$ (et la densité uniforme sous $H_{0i}$), les observations $P_i$ sont des variables i.i.d. de loi de mélange $f(p) = \pi_0 + (1 - \pi_0)h(p)$. Les premiers travaux à propos de l'estimation de $\pi_0$ à l'aide des observations $P_1, ..., P_M$ remontent à Schweder and Spjøtvoll (1982) et Storey (2002) pour un point de vue pratique, formalisé ensuite par Langaas et al. (2005) pour des fonctions $h$ décroissantes, ou encore par Celisse and Robin (2010) où l'optimalité asymptotique est considérée. On trouve également un intervalle de confiance pour $\pi_0$ dans Meinshausen and Rice (2006), l'estimation de la proportion d'espérances non-nulles dans un vecteur gaussien dans Jin (2008), ainsi que des lois de mélange plus générales dans Cai and Jin (2010) et Kumar Patra and Sen (2016).

Préférant l'idée de suivre le principe de Neyman Pearson, plusieurs critères d'erreur de première espèce ont été introduits afin de quantifier l'erreur commise par l'utilisation de $\mathcal{R}$ en lieu et place de $\mathcal{F}(\lambda)$. Le premier critère est le Family Wise Error Rate (FWER) défini par

$$FWER(\mathcal{R}) = \sup_{\lambda\in\overline{\mathcal{S}}} \mathbb{P}_\lambda(\mathcal{R} \cap \mathcal{T}(\lambda) \neq \varnothing).$$

On dit qu'une procédure de test multiple $\mathcal{R}$ contrôle le FWER si $FWER(\mathcal{R}) \leq \alpha$ pour $\alpha$ fixé au préalable. Faisant suite aux résultats de Bonferroni, Holm (1979) propose une procédure générale séquentielle forward de type Bonferroni qui contrôle le FWER, procédure qui sera ensuite améliorée et rendue moins conservatrice par Simes (1986), Shaffer (1986), Hommel (1988), Hochberg (1988) puis Rom (1990). Une généralisation au contrôle du $k$-FWER ($= \sup_{\lambda\in\overline{\mathcal{S}}} \mathbb{P}_\lambda(|\mathcal{R}\cap\mathcal{T}(\lambda)| \geq k)$) est donnée dans Sarkar (2007). Approfondissant une idée de Schweder and Spjøtvoll (1982), Finner and Gontscharuk (2009) proposent quant à eux une procédure qui contrôle le FWER basée sur l'estimateur de $\pi_0$ de Storey (2002), procédure qui sera ensuite développée dans Sarkar et al. (2012).

Le deuxième critère d'erreur de première espèce le plus couramment utilisé est le False Discovery Rate (FDR) introduit par Benjamini and Hochberg (1995). Il est défini par

$$FDR(\mathcal{R}) = \sup_{\lambda \in \overline{\mathcal{S}}} \mathbb{E}_\lambda \left[ \frac{|\mathcal{R} \cap \mathcal{T}(\lambda)|}{1 \vee |\mathcal{R}|} \right].$$

Il est facile de voir que l'on a toujours $FDR(\mathcal{R}) \leq FWER(\mathcal{R})$. Une procédure contrôlant le FDR est proposée dans Benjamini and Hochberg (1995) dans le cas de p-values indépendantes et dans Benjamini and Yekutieli (2001) dans un cadre de dépendance. Hao et al. (2013) ont l'idée de formuler le problème de détection de ruptures multiples dans l'espérance d'un vecteur gaussien en terme de problème de tests multiples et proposent une procédure qui contrôle le FDR. Ce modèle est repris par Li et al. (2016) qui fournissent une procédure permettant notamment la détection minimax d'un segment (car vérifiant la condition de Chan and Walther (2013)).

D'autres critères existent également comme le positive False Discovery Rate (pFDR) proposé par Storey (2002)

$$pFDR(\mathcal{R}) = \sup_{\lambda \in \overline{\mathcal{S}}} \mathbb{E}_\lambda \left[ \frac{|\mathcal{R} \cap \mathcal{T}(\lambda)|}{|\mathcal{R}|} \mid |\mathcal{R}| > 0 \right]$$

ou encore le $k$-FDR étudié dans Sarkar (2007)

$$k - FDR(\mathcal{R}) = \sup_{\lambda \in \overline{\mathcal{S}}} \mathbb{E}_\lambda \left[ \frac{|\mathcal{R} \cap \mathcal{T}(\lambda)|}{|\mathcal{R}|} \mathbf{1}_{|\mathcal{R} \cap \mathcal{T}(\lambda)| \geq k} \right].$$

Nous choisissons dans [A14] d'établir des procédures de tests multiples qui contrôlent le FWER ainsi qu'un critère de seconde espèce adapté au FWER afin de localiser une rupture dans un processus de Poisson. Plusieurs notions de puissances existent pour les tests multiples (voir Benjamini and Hochberg (1995); Chen et al. (2011); Westfall et al. (2011)) ainsi que quelques critères d'optimalité de type maximin (Lehmann et al., 2005; Romano et al., 2011). Nous utilisons un critère d'optimalité au sens du minimax introduit par Fromont et al. (2016), la vitesse de séparation minimax par famille (Family Wise Separation Rate FWSR). On considère la distance $d_2$ associée à la norme usuelle sur $L_2([0,1])$. Pour $r > 0$ et $\lambda \in \overline{\mathcal{S}}$, on définit

$$\mathcal{F}_r(\lambda) = \{H_k \in \mathcal{H} : d_2(\lambda, H_k) \geq r\}$$

l'ensemble des hypothèses à distance au moins $r$ de $\lambda$. On se fixe $\alpha$ et $\beta$ deux probabilités d'erreur. Pour $\mathcal{S} \subset \overline{\mathcal{S}}$ et $\mathcal{R}$ une procédure de tests multiples dont le FWER sur $\overline{\mathcal{S}}$ est contrôlé par $\alpha$, on définit la vitesse de séparation par famille de niveau $\beta$ de $\mathcal{R}$ sur $\mathcal{S}$ par

$$FWSR_\beta(\mathcal{R},\mathcal{S}) = \inf\{r>0 : \sup_{\lambda\in\mathcal{S}} \mathbb{P}_\lambda(\mathcal{F}_r(\lambda)\cap(\mathcal{H}\backslash\mathcal{R})\neq\varnothing)\leq\beta\}.$$

La vitesse de séparation par famille minimax de niveau $\alpha$ et $\beta$ sur $\mathcal{S}$ est alors

$$mFWSR_{\alpha,\beta}(\mathcal{S}) = \inf_{\mathcal{R}:FWER(\mathcal{R})\leq\alpha} FWSR_\beta(\mathcal{R},\mathcal{S}),$$

où l'infimum est pris sur toutes les procédures de tests multiples dont le FWER est contrôlé par $\alpha$ sur $\overline{\mathcal{S}}$. Lorsque $\mathcal{H}$ n'est composé que d'une seule hypothèse, on a alors $mFWSR_{\alpha,\beta}(\mathcal{S}) = mSR_{\alpha,\beta}(\mathcal{S})$.

**Tests multiples minimax de détection et localisation d'une rupture avec une hauteur de saut connue**

Pour $\lambda_0>0$ et $\delta^*\in(-\lambda_0,+\infty)\backslash\{0\}$ fixés, on considère dans [A14] l'ensemble des intensités

$$\mathcal{S}[\lambda_0,\delta^*] = \{\lambda:[0,1]\to(0,+\infty), \exists\tau\in(0,1), \forall t\in[0,1], \lambda(t)=\lambda_0+\delta^*\mathbf{1}_{(\tau,1]}(t)\}$$

présentant un saut de hauteur $\delta^*$ au temps $\tau$ (inconnu), et l'ensemble

$$\overline{\mathcal{S}}[\lambda_0,\delta^*] = \mathcal{S}[\lambda_0,\delta^*]\cup\{\lambda_0\}$$

des intensités possédant au plus un saut. Afin de localiser $\tau$ avec une précision de $1/M$, on considère pour $M\in\mathbb{N}^*$ fixé la collection d'hypothèses $\mathcal{H}=\{H_k, k\in\{1,...,M\}\}$ où

$$H_k = \{\lambda:[0,1]\to(0,+\infty), \exists\tau\in[k/M,1], \forall t\in[0,1], \lambda(t)=\lambda_0+\delta^*\mathbf{1}_{(\tau,1]}(t)\}$$

de sorte que

$$\{\lambda_0\} = H_M\subset H_{M-1}\subset ...\subset H_1\subset\overline{\mathcal{S}}[\lambda_0,\delta^*].$$

Les hypothèses étant emboîtées, on obtient (voir Fromont et al. (2016)) la borne inférieure suivante pour le mFWSR :

**Proposition 4.1** ([A14]) Soient $\alpha$ et $\beta$ deux niveaux de $(0,1)$ tels que $\alpha+\beta<1$, $\lambda_0>0$ et $\delta^*\in(-\lambda_0,+\infty)\backslash\{0\}$. Pour tout $M\in\mathbb{N}^*$ et pour tout $L\geq\lambda_0\log(C_{\alpha,\beta}/(\delta^*)^2)$ (avec $C_{\alpha,\beta}=1+4(1-\alpha-\beta)^2$)

$$mFWSR_{\alpha,\beta}(\mathcal{S}[\lambda_0,\delta^*]) \geq \sqrt{\frac{\lambda_0\log C_{\alpha,\beta}}{L}}.$$

Afin de construire une procédure de tests multiples qui atteint cette borne inférieure, on considère le test simple de $H_k$ vs $\mathcal{S}[\lambda_0,\delta^*]\backslash H_k$ de la forme

$$\phi_k(N) = \mathbf{1}_{S_{\delta^*,k}(N)>s_{\delta^*,k}(1-\alpha)}$$

avec

$$S_{\delta^*,k}(N) = \sup_{t\in(0,k/M)} \left( \mathrm{sign}(\delta^*)(N(t, \frac{k}{M}] - \lambda_0 L(\frac{k}{M} - t)) - \frac{|\delta^*|}{2} L(\frac{k}{M} - t) \right)$$

et $s_{\delta^*,k}(u)$ le u-quantile de $S_{\delta^*,k}(N)$ sous $H_k$. On introduit ensuite l'indice

$$\hat{k} = \sup\{k' \in \{1, ..., M\} : \phi_{k'} = 0\}$$

(avec $\sup \varnothing = -\infty$) ainsi que la procédure de tests multiples

$$\mathcal{R}_1 = \{H_k : k \geq \hat{k} + 1\}.$$

On montre alors dans [A14] la borne supérieure suivante :

**Théorème 4.1** ([A14]) Soit $L \geq 1$, $\alpha$ et $\beta$ deux niveaux fixés de $(0, 1)$, $\lambda_0 > 0$ et $\delta^* \in (-\lambda_0, +\infty)\backslash\{0\}$. Il existe une constante (explicite) $C(\alpha, \beta, \lambda_0, \delta^*) > 0$ telle que la procédure de test $\mathcal{R}_1$ vérifie pour tout $M \in \mathbb{N}^*$

$$FWER(\mathcal{R}_1) \leq \alpha \quad \text{et} \quad FWSR_\beta(\mathcal{R}_1, \mathcal{S}[\lambda_0, \delta^*]) \leq \min(|\delta^*|, \frac{C(\alpha, \beta, \lambda_0, \delta^*)}{\sqrt{L}}).$$

En particulier si $L \geq (C(\alpha, \beta, \lambda_0, \delta^*)/\delta^*)^2$,

$$mFWSR_{\alpha,\beta}(\mathcal{S}[\lambda_0, \delta^*]) \leq \frac{C(\alpha, \beta, \lambda_0, \delta^*)}{\sqrt{L}},$$

redonnant la borne inférieure à une constante près.

### Tests multiples minimax de détection et localisation d'une rupture avec une hauteur de saut inconnue

On considère ensuite le cas d'une hauteur de saut inconnue, i.e. des intensités appartenant à l'espace

$$\mathcal{S}[\lambda_0] = \{\lambda : [0, 1] \to (0, +\infty), \exists(\delta^* \times \tau) \in ((-\lambda_0, +\infty)\backslash\{0\}) \times (0, 1), \forall t \in [0, 1], \lambda(t) = \lambda_0 + \delta^* \mathbf{1}_{(\tau,1]}(t)\}.$$

Cependant cet espace étant trop grand (au sens où $mSR_{\alpha,\beta}(\mathcal{S}[\lambda_0]) = +\infty$), nous travaillons pour $R > \lambda_0$ fixé avec les espaces

$$\mathcal{S}[\lambda_0, R] = \{\lambda : [0, 1] \to (0, +\infty), \exists(\delta^* \times \tau) \in ((-\lambda_0, R-\lambda_0)\backslash\{0\}) \times (0, 1), \forall t \in [0, 1], \lambda(t) = \lambda_0 + \delta^* \mathbf{1}_{(\tau,1]}(t)\}$$

et

$$\overline{\mathcal{S}}[\lambda_0, R] = \mathcal{S}[\lambda_0, R] \cup \{\lambda_0\}.$$

On considère alors la collection d'hypothèses $\mathcal{H} = \{H_k, k \in \{1, ..., M\}\}$ avec

$$H_k = \{\lambda : [0,1] \to (0,+\infty), \exists(\delta^* \times \tau) \in ((-\lambda_0, R-\lambda_0)\backslash\{0\}) \times (k/M, 1), \forall t \in [0,1], \lambda(t) = \lambda_0 + \delta^* \mathbf{1}_{(\tau,1]}(t)\}$$

de sorte que comme précédemment

$$\{\lambda_0\} = H_M \subset H_{M-1} \subset ... \subset H_1 \subset \overline{\mathcal{S}}[\lambda_0, R].$$

On obtient ainsi directement la borne inférieure suivante :

**Proposition 4.2** ([A14]) Soient $\alpha$ et $\beta$ deux niveaux de $(0,1)$ tels que $\alpha + \beta < 1/2$, $\lambda_0 > 0$ et $R > \lambda_0$. Il existe $L_0(\alpha, \beta, \lambda, R) > 0$ (explicite) telle que pour tout $M \in \mathbb{N}^*$ et pour tout $L \geq L_0(\alpha, \beta, \lambda_0, R)$,

$$mFWSR_{\alpha,\beta}(\mathcal{S}[\lambda_0, R]) \geq \sqrt{\frac{\lambda_0 \log\log L}{L}}.$$

Pour établir une borne supérieure, on considère le test simple

$$\phi_k(N) = \mathbf{1}_{\left\{\max_{j\in\{1,\dots,\lfloor\log_2 L\rfloor\}}\left(N\left(\frac{k}{M}(1-2^{-j}), \frac{k}{M}\right] - p_{\frac{\lambda_0 kL}{M2^j}}\left(1-\frac{u_\alpha}{2}\right)\right)>0\right\}} \vee \mathbf{1}_{\left\{\max_{j\in\{1,\dots,\lfloor\log_2 L\rfloor\}}\left(p_{\frac{\lambda_0 kL}{M2^j}}\left(\frac{u_\alpha}{2}\right) - N\left(\frac{k}{M}(1-2^{-j}), \frac{k}{M}\right]\right)>0\right\}}$$

où $p_\xi(u)$ est le $u$-quantile de la loi de Poisson de paramètre $\xi$ et $u_\alpha = \alpha/\lfloor\log_2 L\rfloor$. Pour définir notre procédure de tests multiples, on utilise l'indice

$$\hat{k} = \sup\{k' \in \{1, ..., M\} : \phi_{k'} = 0\}$$

et on pose

$$\mathcal{R}_2 = \{H_k : k \geq \hat{k} + 1\}.$$

On obtient alors la borne supérieure suivante pour $\mathcal{R}_2$ :

**Théorème 4.2** ([A14]) Soit $L \geq 1$, $\alpha$ et $\beta$ deux niveaux de $(0,1)$, $\lambda_0 > 0$ et $R > \lambda_0$. Il existe une constante (explicite) $C(\alpha, \beta, \lambda_0, R) > 0$ telle que pour tout $M \in \mathbb{N}^*$

$$FWER(\mathcal{R}_2) \leq \alpha \quad \text{et} \quad FWSR_\beta(\mathcal{R}_2, \mathcal{S}[\lambda_0, R]) \leq \min(\max(\lambda_0, R-\lambda_0), C(\alpha, \beta, \lambda_0, R)\sqrt{\frac{\log\log L}{L}}).$$

En particulier il existe $L_0(\alpha, \beta, \lambda_0, R)$ tel que pour $L \geq L_0(\alpha, \beta, \lambda_0, R)$,

$$mFWSR_{\alpha,\beta}(\mathcal{S}[\lambda_0, R]) \leq C(\alpha, \beta, \lambda_0, R)\sqrt{\frac{\log\log L}{L}}.$$

On obtient ainsi une procédure de tests multiples minimax, la vitesse étant ici $\sqrt{\frac{\log\log L}{L}}$, i.e. identique à celle du problème de détection simple.

**Liens avec des intervalles de confiance de longueur minimale pour l'instant de rupture**

On s'intéresse enfin à la question de trouver un intervalle de confiance "optimal" pour l'instant de rupture d'intensité. On pose $\mathcal{MD}$ l'ensemble des fonctions mesurables de $\mathcal{D}[0,1]$ dans $[0,1]$ et on considère des intervalles de confiance de la forme $[\phi(N)-a;\phi(N)+b]$ pour $\phi \in \mathcal{MD}$ et $a,b \geq 0$. La longueur minimale d'un intervalle de confiance de niveau $1-\varepsilon$ d'un estimateur $\phi(N)$ de $\tau$ pour $\lambda \in \overline{\mathcal{S}}$ est définie par

$$\mathcal{L}_\varepsilon(\phi,\overline{\mathcal{S}}) = \inf\{a+b : a,b \geq 0, \inf_{\lambda\in\overline{\mathcal{S}}} \mathbb{P}_\lambda(\tau \in [\phi(N)-a;\phi(N)+b]) \geq 1-\varepsilon\},$$

et la longueur minimale d'un intervalle de confiance de niveau $1-\varepsilon$ pour $\tau$ sur la classe $\overline{\mathcal{S}}$ est

$$\mathcal{L}_\varepsilon(\overline{\mathcal{S}}) = \inf\{\mathcal{L}_\varepsilon(\phi,\overline{\mathcal{S}}) : \phi \in \mathcal{MD}\}.$$

**Intervalle de confiance pour l'instant de rupture dans le cas d'une hauteur de saut connue** La proposition suivante établit un lien entre la longueur minimale d'un intervalle de confiance pour $\tau$ et $mFWSR_{\alpha,\beta}(\mathcal{S}[\lambda_0,\delta^*])$ lorsque la hauteur du saut $\delta^*$ est connue.

**Proposition 4.3** ([A14]) Soit $L \geq 1$, $\alpha$ et $\beta$ deux niveaux de $(0,1)$, $\lambda_0 > 0$, $\delta^* \in (-\lambda_0,+\infty)\setminus\{0\}$ et $M \in \mathbb{N}^*$. On considère $a,b>0$ et $\phi \in \mathcal{MD}$ tels que

$$\inf_{\lambda\in\overline{\mathcal{S}}[\lambda_0,\delta^*]} P_\lambda(\tau \in (\phi(N)-a,\phi(N)+b]) \geq 1-\alpha \ .$$

La procédure de tests multiples $\mathcal{R}$ sur $\mathcal{H}$ définie par $\mathcal{R} = \{H_k \in \mathcal{H},\ k/M > \phi(N)+b\}$ vérifie

$$FWER(\mathcal{R}) \leq \alpha, \ \text{ et } \ FWSR_{\alpha,\beta}(\mathcal{R},\mathcal{S}[\lambda_0,\delta^*]) \leq |\delta^*|\sqrt{a+b}.$$

En particulier, on obtient pour tout $M \in \mathbb{N}^*$

$$\mathcal{L}_\alpha(\overline{\mathcal{S}}[\lambda_0,\delta^*]) \geq \frac{mFWSR_{\alpha,\alpha}(\mathcal{S}[\lambda_0,\delta^*])^2}{(\delta^*)^2}.$$

Réciproquement, soit $r>0$ tel que $r \geq mFWSR_{\alpha,\beta}(\mathcal{S}[\lambda_0,\delta^*])$ et une procédure de tests multiples $\mathcal{R}$ sur $\mathcal{H}$ vérifiant les deux inégalités $FWER(\mathcal{R}) \leq \alpha$ et $FWSR_{\alpha,\beta}(\mathcal{R},\mathcal{S}[\lambda_0,\delta^*]) \leq r$. Soit $\hat{k} = \sup\{k \in \{1,\ldots,M\},\ H_k \notin \mathcal{R}\}$ et $\hat{\tau} = \hat{k}/M$. L'intervalle $I_{r,\delta^*,M,\mathcal{R}}$ défini par $I_{r,\delta^*,M,\mathcal{R}} = (\hat{\tau} - r^2/(\delta^*)^2, \hat{\tau}+1/M]$ vérifie alors

$$\inf_{\lambda\in\overline{\mathcal{S}}[\lambda_0,\delta^*]} \mathbb{P}_\lambda(\tau \in I_{r,\delta^*,M,\mathcal{R}}) \geq 1-\alpha-\beta.$$

En particulier, on obtient pour tout $M \in \mathbb{N}^*$

$$\mathcal{L}_{\alpha+\beta}(\overline{\mathcal{S}}[\lambda_0,\delta^*]) \leq \frac{1}{M} + \frac{mFWSR_{\alpha,\beta}(\mathcal{S}[\lambda_0,\delta^*])^2}{(\delta^*)^2}.$$

Cette proposition permet de construire un intervalle de confiance pour $\tau$ de longueur minimale à partir de la procédure de tests multiples $\mathcal{R}_1$. En choisissant la collection d'hypothèses $\mathcal{H}$ telle que

$$M = \left\lfloor \frac{L(\delta^*)^2}{C\left(\alpha/2, \alpha/2, \lambda_0, \delta^*\right)^2} \right\rfloor$$

pour $L \geq \max(\lambda_0 \log C_{\alpha,\beta}), \frac{C(\alpha/2,\alpha/2,\lambda_0,\delta^*)^2}{(\delta^*)^2})$ et en posant

$$\hat{\tau} = \frac{\hat{k}}{M},$$

l'intervalle $(\hat{\tau} - \frac{1}{M}, \hat{\tau} + \frac{1}{M}]$ est un intervalle de confiance de niveau $1-\alpha$ atteignant la longueur minimale $\mathcal{L}_\alpha(\overline{\mathcal{S}}[\lambda_0, \delta^*])$ (à une constante près).

**Intervalle de confiance pour l'instant de rupture dans le cas d'une hauteur de saut inconnue** Dans le cas où la hauteur de saut $\delta^*$ est inconnue, on ne peut trouver de réponse uniformément optimale pour un intervalle de confiance de longueur minimale pour $\tau$. En effet,

**Proposition 4.4** ([A14]) Soit $\alpha \in (0, 1/2)$, $\lambda_0 > 0$ et $R > \lambda_0$. Pour tout $L \geq 1$, on a

$$\mathcal{L}_\alpha(\overline{\mathcal{S}}[\lambda_0, R]) = 1.$$

On peut alors restreindre l'espace des alternatives en considérant par exemple pour $\Delta \in (0, \lambda_0 \wedge (R - \lambda_0))$

$$\mathcal{S}[\lambda_0, \Delta, R] = \{\lambda : [0,1] \to (0, R],\ \exists \delta \in \{(-\lambda_0, -\Delta] \cup [\Delta, R - \lambda_0]\},\ \exists \tau \in (0,1),\ \forall t \in [0,1],$$
$$\lambda(t) = \lambda_0 + \delta \mathbf{1}_{(\tau,1]}(t)\}$$

l'ensemble des intensités dont la hauteur de saut est au moins égale à $\Delta$ fixé. On montre alors que ce dernier problème revient au cas où la hauteur du saut est connue, l'adaptation en $\delta$ ne modifiant ni la vitesse (en $1/\sqrt{L}$) ni la longueur d'intervalle de confiance en $1/L$ par rapport au cas où $\delta^* = \Delta$ est connu.

# 3 Perspectives de recherches

Les travaux présentés dans ce chapitre sont de nature non-asymptotiques et optimaux au sens du minimax dans ce qui constitue le modèle le plus simple de processus de comptage à savoir le processus de Poisson simple. Une première question naturelle consiste à se demander s'il est possible d'étendre les résultats de localisation obtenus au cas d'une intensité de base $\lambda_0$ inconnue, au prix vraisemblablement d'un conditionnement par le nombre de points observés. On pourra également chercher à savoir s'ils peuvent être étendus à des processus de comptage plus généraux, comme des processus de Poisson composés ou des processus de Cox par exemple, aussi bien pour la détection que pour la localisation par tests multiples d'un instant de rupture de continuité de l'intensité par exemple (lorsque celle-ci est continue).

Un peu plus par analogie avec la détection de rupture dans le modèle gaussien, il serait intéressant de considérer des problèmes de détection de rupture de drift dans des modèles d'Equations Différentielles Stochastiques de type

$$dX_t = b(t, X_t)dt + dZ_t,$$

avec des processus $Z$ plus généraux que le mouvement Brownien. On pourra par exemple considérer pour $Z$ un processus de Lévy ou de type-Lévy. Les exemples privilégiés sont les processus stables ou de type-stables (stable-like processes), ou encore les processus multistables, la hauteur et l'intensité des sauts étant alors régulées par un paramètre ou une fonction $\alpha$.

On pourra également considérer des problèmes de détection de rupture de l'intensité de saut elle-même (en abonnant probablement les critères non-asymptotiques dans un premier temps au profit d'une étude asymptotique). On pourra ainsi étudier la détection de ruptures d'une fonction de multistabilité $\alpha$ constante par morceaux dans un modèle de Lévy multistable (ou de processus multistables plus généraux). Il serait intéressant aussi d'étendre les travaux de Benassi et al. (2000) ou encore de Bardet and Kammoun (2008) par exemple qui considèrent une fonction $H$ constante par morceaux dans un modèle de mouvement Brownien multifractionnaire, avec la détection d'une rupture de la fonction $H$ et/ou de la fonction $\alpha$ dans un mouvement linéaire multifractionnaire multistable (LMMM).

# Liste des publications


**[A1]** Falconer K.J., Le Guével R. and Lévy Véhel J. (2009). Localisable moving average stable and multistable processes. *Stochastic Models* 25(4), 648-672.

**[A2]** Le Guével R. and Lévy Véhel J. (2012). A Ferguson-Klass-LePage series representation of multistable multifractional motions and related processes. *Bernoulli* 18(4), 1099-1127.

**[A3]** Le Guével R. and Lévy Véhel J. (2013). Incremental moments and Hölder exponents of multifractional multistable processes. *ESAIM : Probability and Statistics*, 17, 135-178.

**[A4]** Le Guével R. (2013). An estimation of the stability and the localisability functions of multistable processes. *Electronic Journal of Statistics*, 7, 1129-1166.

**[A5]** Le Guével R., Lévy Véhel J. and Liu L. (2013). On two multistable extensions of stable Lévy motion and their semi-martingale representations. *Journal of Theoretical Probability*, Vol. 28, 1125-1144.

**[A6]** Le Guével R. (2018) The Hausdorff dimension of the range of the Lévy multistable processes. *Journal of Theoretical Probability*, 32(2), 765-780.

**[A7]** Le Guével R. (2019). Goodness-of-fit test for multistable Lévy processes. *Communications in Statistics - Theory and Methods*, 50(8), 1807-1837.

**[A8]** Le Guével R. and Lévy Véhel J. (2020). Hausdorff, large deviation and Legendre multifractal spectra of Lévy multistable processes. *Stochastic Processes and their Applications*, vol. 130(4), 2032-2057.

**[A9]** Le Guével R. (2021). Exponential inequalities for the supremum of some counting processes and their square martingales. *Comptes Rendus. Mathématique*, Tome 359 (2021) no. 8, 969-982.

**[A10]** Lavancier F. and Le Guével R. (2021). Spatial birth-death-move processes : basic properties and estimation of their intensity functions. *Journal of the Royal Statistical Society (B)*, 83(4), 798-825.

**[A11]** Altmeyer R. and and Le Guével R. (2021). Optimal L2-approximation of occupation and local times for symmetric stable processes, *Electronic Journal of Statistics*, 16(1), 2859-2883.

[**A12**] Fromont M., Grela F. and Le Guével R. (2023) Minimax and adaptive tests for detecting abrupt and possibly transitory changes in a Poisson process, *Electronic Journal of Statistics*, 17(2), 2575-2744.

[**A13**] Lavancier F., Le Guével R. and Manent, M. (2022) Feller and ergodic properties of jump-move processes with applications to interacting particle systems, in revision, HAL : https ://hal.archives-ouvertes.fr/hal-03632962.

[**A14**] Fromont M., Grela F. and Le Guével R. (2022) Minimax multiple testing procedures for localising an abrupt change in a Poisson process with a known baseline intensity, Working paper, HAL : https ://hal-cnrs.archives-ouvertes.fr/hal-03702141.

# Bibliographie


Abry, P., Flandrin, P., and Taqqu, M. (1998). Wavelets for the analysis, estimation and synthesis of scaling Data.

Abry, P., Flandrin, P., Taqqu, M., and Veitch, D. (2003). Self-similarity and long-range dependence through the wavelet lens. In Doukhan, P., Oppenheim, G., and Taqqu, M., editors, *Theory and Applications of Long-Range Dependence*, pages 527–556. Birkhäuser.

Abry, P., Helgason, H., and Pipiras, V. (2011). Wavelet-based analysis of non-Gaussian long-range dependent processes and estimation of the Hurst parameter. *Lithuanian Mathematical Journal*, 51 :287–302.

Abry, P., Jaffard, S., Leonarduzzi, R., Melot, C., and Wendt, H. (2017). New exponents for pointwise singularity classification. In Seuret, S. and Barral, J., editors, *Recent Developments in Fractals and Related Fields : Conference on Fractals and Related Fields III*, volume FARF3 2015 : Recent Developments in Fractals and Related Fields of *Trends in Mathematics book series (TM)*, pages 1–37. Birkhäuser.

Abry, P. and Veitch, D. (1998). Wavelet analysis of long-range-dependent traffic. *IEEE Trans. Inf. Theory*, 44 :2–15.

Abry, P., Veitch, D., and Flandrin, P. (2001). Long-range dependence : revisiting aggregation with wavelets. *Journal of Time Series Analysis*, 19 :253 – 266.

Adler, R. (1981). *The geometry of random fields.* Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM).

Aït-Sahalia, Y. and Jacod, J. (2009a). Estimating the degree of activity of jumps in high frequency data. *The Annals of Statistics*, 37(5A) :2202 – 2244.

Aït-Sahalia, Y. and Jacod, J. (2009b). Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1) :184 – 222.

Aït-Sahalia, Y. and Jacod, J. (2018). Semimartingale : Itô or not ? *Stochastic Processes and their Applications*, 128(1) :233–254.

Altmeyer, R. (2021). Approximation of occupation time functionals. *Bernoulli*, 27(4) :2714–2739.

Altmeyer, R. and Chorowski, J. (2017). Estimation error for occupation time functionals of stationary Markov processes. *Stochastic processes and their applications*, 128(6) :1830–1848.

Andersson, P. (1997). Characterization of pointwise Hölder regularity. *Applied and Computational Harmonic Analysis*, 4 :429–443.

Anscombe, F. J. (1953). Sequential estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 15(1) :1–29.

Antoch, J. and Jaruvskova, D. (2007). Testing a homogeneity of stochastic processes. *Kybernetika*, 43(4) :415–430.

Arias-Castro, E., Donoho, D. L., and Huo, X. (2005). Near-optimal detection of geometric objects by fast multiscale methods. *IEEE Trans. Inform. Theory*, 51(7) :2402–2425.

Arneodo, A., Bacry, E., Jaffard, S., and Muzy, J. F. (1998). Singularity spectrum of multifractal functions involving oscillating singularities. *Journal of Fourier Analysis and Applications*, 4(2) :159–174.

Assareh, H., Noorossana, R., Mohammadi, M., and Mengersen, K. (2016). Bayesian multiple change-point estimation of Poisson rates in control charts. *Scientia Iranica. Transaction E, Industrial Engineering*, 23(1) :316.

Athreya, K. B. and Ney, P. E. (2012). *Branching processes*, volume 196. Springer Science & Business Media.

Aue, A. and Horváth, L. (2013). Structural breaks in time series. *Journal of Time Series Analysis*, 34(1) :1–16.

Ayache, A. (2011). Continuous Gaussian multifractional processes with random pointwise Hölder regularity. *Journal of Theoretical Probability*, 26.

Ayache, A., Benassi, A., Cohen, S., and Lévy Véhel, J. (2004). Regularity and identification of generalized multifractional Gaussian processes. *Lecture Notes in Mathematics*, 1857.

Ayache, A. and Hamonier, J. (2013). Linear multifractional stable motion : fine path properties. *Revista Matematica Iberoamericana*, 30 :1301–1354.

Ayache, A. and Hamonier, J. (2017). Behaviour of linear multifractional stable motion : membership of a critical Hölder space. *Stochastics*, 89(5) :709–725.

Ayache, A. and Lévy Véhel, J. (2004). On the identification of the pointwise Hölder exponent of the generalized multifractional Brownian motion. *Stochastic Processes and their Applications*, 111 :119–156.

Azaïs, R., Dufour, F., and Gégout-Petit, A. (2013). Nonparametric estimation of the jump rate for non-homogeneous marked renewal processes. *Annales de l'I.H.P. Probabilités et statistiques*, 49(4) :1204–1231.

Azaïs, R., Dufour, F., and Gégout-Petit, A. (2014). Non-parametric estimation of the conditional distribution of the interjumping times for piecewise-deterministic markov processes. *Scandinavian Journal of Statistics*, 41(4) :950–969.

Bacry, E., Muzy, J., and Arnéodo, A. (1993). Singularity spectrum of fractal signals from wavelet analysis : Exact results. *J Stat Phys*, 70 :635–674.

Bai, J. (1994). Least squares estimation of a shift in linear processes. *J. Time Ser. Anal.*, 15(5) :453–472.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1) :47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. In *Time series econometrics. Vol. 2*, pages 159–193. World Sci. Publ., Hackensack, NJ.

Bain, L. J., Engelhardt, M., and Wright, F. T. (1985). Tests for an increasing trend in the intensity of a Poisson process : a power study. *Journal of the American Statistical Association*, 80(390) :419–422.

Balança, P. (2014). Fine regularity of Lévy processes and linear (multi)fractional stable motion. *Electronic Journal of Probability*, 19 :1–37.

Balança, P. and Herbin, E. (2012). 2-microlocal analysis of martingales and stochastic integrals. *Stochastic Processes and their Applications*, 122 :2346–2382.

Bansaye, V. and Méléard, S. (2015). *Stochastic models for structured populations*, volume 16. Springer.

Baraud, Y. (2002). Non-asymptotic minimax rates of testing in signal detection. *Bernoulli*, 8(5) :577–606.

Bardet, J.-M. (2000). Testing for the presence of self-similarity of Gaussian time series having stationary increments. *Journal of Time Series Analysis*, 21(5) :497–515.

Bardet, J.-M. (2002). Statistical study of the wavelet analysis of fractional Brownian motion. *IEEE Transactions on Information Theory*, 48(4) :991–999.

Bardet, J.-M., Gabriel, L., Moulines, E., and Soulier, P. (2000). Wavelet estimator of long-range dependent processes. *Statistical Inference for Stochastic Processes*, 3 :85–99.

Bardet, J.-M. and Kammoun, I. (2008). Detecting abrupt changes of the long-range dependence or the self-similarity of a gaussian process. *Comptes Rendus. Mathematique*, 346(13) :789–794.

Bardet, J.-M., Philippe, A., Oppenheim, G., Taqqu, M. S., Stoev, S., and Lang, G. (2003). Semi-parametric estimation of the long-range dependence parameter : a survey. In *Theory and applications of long-range dependence.*, pages 557–577. Birkhäuser.

Bardet, J.-M. and Surgailis, D. (2011). Measuring the roughness of random paths by increment ratios. *Bernoulli*, 17(2) :749 – 780.

Bardet, J.-M. and Surgailis, D. (2012). Nonparametric estimation of the local Hurst function of multifractional Gaussian processes. *Stochastic Processes and their Applications*, 123 :1004–1045.

Barral, J. (2000). Continuity of the multifractal spectrum of a random statistically self-similar measure. *Journal of Theoretical Probability*, 13(4) :1027–1060.

Barral, J., Durand, A., Jaffard, S., and Seuret, S. (2012). Local multifractal analysis.

Barral, J., Fournier, N., Jaffard, S., and Seuret, S. (2010). A pure jump Markov process with a random singularity spectrum. *The Annals of Probability*, 38(5) :1924–1946.

Barral, J. and Lévy Véhel, J. (2004). Multifractal analysis of a class of additive processes with correlated non-stationary increments. *Electronic Journal of Probability*, 9(none) :508 – 543.

Barral, J. and Mandelbrot, B. (2004). Random multiplicative multifractal measures. *Fractal Geometry and Applications. Proc. Symp. Pure Math.*

Barral, J. and Seuret, S. (2007). The singularity spectrum of Lévy processes in multifractal time. *Advances in Mathematics*, 214(1) :437–468.

Barral, J. and Seuret, S. (2011). A localized Jarnik-Besicovitch theorem. *Advances in Mathematics*, 226(4) :3191–3215.

Barral, J. and Seuret, S. (2020). Besov spaces in multifractal environment and the Frisch-Parisi conjecture.

Basseville, M. and Nikiforov, I. V. (1993). *Detection of abrupt changes : theory and application.* Prentice Hall, Inc.

Bastin, F., Esser, C., and Jaffard, S. (2016). Large deviation spectra based on wavelet leaders. *Revista Matemática Iberoamericana*, 32 :859–890.

Bayraktar, E., Dayanik, S., and Karatzas, I. (2006). Adaptive Poisson disorder problem. *The Annals of Applied Probability*, 16(3) :1190–1261.

Bec, M. and Lacour, C. (2012). Adaptive pointwise estimation for pure jump lévy processes. *Statistical Inference for Stochastic Processes*, 18.

Begyn, A. (2007a). Asymptotic expansion and central limit theorem for quadratic variations of Gaussian processes. *Bernoulli*, 13(3) :712 – 753.

Begyn, A. (2007b). Functional limit theorems for generalized quadratic variations of Gaussian processes. *Stochastic Processes and their Applications*, 117 :1848–1869.

Ben Nasr, F., Bhouri, I., and Heurteaux, Y. (2002). The validity of the multifractal formalism : results and examples. *Advances in Mathematics*, 165(2) :264–284.

Benassi, A., Bertrand, P., Cohen, S., and Istas, J. (2000). Identification of the Hurst index of a step fractional Brownian motion. *Statistical Inference for Stochastic Processes*, 3 :101–111.

Benassi, A., Cohen, S., and Istas, J. (1998). Identifying the multifractional function of a Gaussian process. *Statistics & Probability Letters*, 39(4) :337–345.

Benassi, A., Jaffard, S., and Roux, D. (1997). Elliptic Gaussian random processes. *Revista matemática iberoamericana, ISSN 0213-2230, Vol. 13, Nº 1, 1997, pags. 19-90*, 13.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate : a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1) :289–300.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4) :1165–1188.

Beran, J. (1992). Statistical methods for data with long-range dependence. *Statistical Science*, 7(4) :404 – 416.

Berman, S. M. (1970). Gaussian processes with stationary increments : local times and sample function properties. *The Annals of Mathematical Statistics*, 41(4) :1260 – 1272.

Berman, S. M. (1972). Gaussian sample functions : uniform dimension and Hölder conditions nowhere. *Nagoya Mathematical Journal*, 46(none) :63 – 86.

Bezborodov, V. and Persio, L. D. (2022). Spatial birth-and-death processes with a finite number of particles. *Modern Stochastics : Theory and Applications*, pages 1–34.

Bhattacharjee, M., Deshpande, J. V., and Naik-Nimbalkar, U. V. (2004). Unconditional tests of goodness of fit for the intensity of time-truncated nonhomogeneous Poisson processes. *Technometrics*, 46(3) :330–338.

Biermé, H. and Lacaux, C. (2013). Linear multifractional multistable motion : LePage series representation and modulus of continuity. *Annals of the University of Bucharest (mathematical series)*, LXII(4) :345–360.

Biermé, H., Lacaux, C., and Xiao, Y. (2009). Hitting probabilities and the Hausdorff dimension of the inverse images of anisotropic Gaussian random fields. *Bulletin of the London Mathematical Society*, 41(2) :253–273.

Birgé, L. and Massart, P. (2001). Gaussian model selection. *J. Eur. Math. Soc. (JEMS)*, 3(3) :203–268.

Blumenthal, R. M. and Getoor, R. K. (1960a). A dimension theorem for sample functions of stable processes. *Illinois Journal of Mathematics*, 4(3) :370 – 375.

Blumenthal, R. M. and Getoor, R. K. (1960b). Some theorems on stable processes. *Transactions of the American Mathematical Society*, 95(2) :263–273.

Blumenthal, R. M. and Getoor, R. K. (1961). Sample functions of stochastic processes with stationary independent increments. *Journal of Mathematics and Mechanics*, 10(3) :493–516.

Blumenthal, R. M. and Getoor, R. K. (1962). The dimension of the set of zeros and the graph of a symmetric stable process. *Illinois Journal of Mathematics*, 6(2) :308 – 316.

Blumenthal, R. M. and Getoor, R. K. (1968). *Markov processes and potential theory.* New York, Academic Press.

Bolte, S. and Cordelieres, F. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J Microscopy*, 224 :213–232.

Bony, J.-M. (1986). Second microlocalization and propagation of singularities for semi-linear hyperbolic equations. In Mizohata, S., editor, *Hyperbolic Equations and Related Topics*, pages 11–49. Academic Press.

Borodin, A. N. (1986). On the character of convergence to Brownian local time. I. *Probability theory and related fields*, 72(2) :231–250.

Böttcher, B., Schilling, R., and Wang, J. (2013). *Lévy Matters III : Lévy-type processes : Construction, approximation and sample path properties.* Lecture Notes in Mathematics. Springer, Cham.

Boufoussi, B., Dozzi, M., and Guerbaz, R. (2008). Path properties of locally asymptotically self similar processes. *Electronic Journal of Probability*, 13 :898–921.

Bretagnolle, J., Dacunha Castelle, D., and Krivine, J.-L. (1966). Lois stables et espaces $L^p$. *Annales de l'I.H.P. Probabilités et statistiques*, 2(3) :231–259.

Briane, V., Vimond, M., Valades-Cruz, C. A., Salomon, A., Wunder, C., and Kervrann, C. (2019). A sequential algorithm to detect diffusion switching along intracellular particle trajectories. *Bioinformatics*, 36(1) :317–329.

Brouste, A. and Masuda, H. (2018). Efficient estimation of stable Lévy process with symmetric jumps. *Statistical Inference for Stochastic Processes*, 21.

Brown, G., Michon, G., and Peyrière, J. (1992). On the multifractal analysis of measures. *J Stat Phys*, 2(66) :775–790.

Brown, M. and Zacks, S. (2006). A note on optimal stopping for possible change in the intensity of an ordinary Poisson process. *Statistics and Probability Letters*, 76(13) :1417–1425.

Brunel, V.-E. (2014). Convex set detection. *arXiv :1404.6224*.

Bull, A. D. (2016). Near-optimal estimation of jump activity in semimartingales. *The Annals of Statistics*, 44(1) :58 – 86.

Cai, T. T. and Guo, Z. (2017). Confidence intervals for high-dimensional linear regression : minimax rates and adaptivity. *Ann. Statist.*, 45(2) :615–646.

Cai, T. T. and Jin, J. (2010). Optimal rates of convergence for estimating the null density and proportion of nonnull effects in large-scale multiple testing. *Ann. Statist.*, 38(1) :100–145.

Cambanis, S., Nolan, J., and Rosinski, J. (1990). On the oscillation of infinitely divisible and some other processes. *Stochastic Processes and their Applications*, 35(1) :87–97.

Canus, C., Lévy Véhel, J., and Tricot, C. (1998). Continuous large deviation multifractal spectrum : definition and estimation. In Novak, M. M., editor, *Fractals 98*, Valleta, Malta. World Scientific.

Cappe, O., Moulines, E., Pesquet, J.-C., Petropulu, A., and Yang, X. (2002). Long-range dependence and heavy-tail modeling for teletraffic data. *IEEE Signal Processing Magazine*, 19(3) :14–27.

Carlin, B. P., Gelfand, A. E., and Smith, A. F. (1992). Hierarchical Bayesian analysis of changepoint problems. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 41(2) :389–405.

Cawley, R. and Mauldin, R. (1992). Multifractal decompositions of Moran fractals. *Advances in Mathematics*, 92(2) :196–236.

Celisse, A. and Robin, S. (2010). A cross-validation based estimation of the proportion of true null hypotheses. *J. Statist. Plann. Inference*, 140(11) :3132–3147.

Chan, H. P. and Walther, G. (2013). Detection with the scan and the average likelihood ratio. *Statist. Sinica*, 23(1) :409–428.

Chen, J., Luo, J., Liu, K., and Mehrotra, D. (2011). On power and sample size computation for multiple testing procedures. *Computational Statistics & Data Analysis*, 55 :110–122.

Chernoff, H. and Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *The Annals of Mathematical Statistics*, 35(3) :999 – 1018.

Chhabra, A. and Jensen, R. V. (1989). Direct determination of the f($\alpha$) singularity spectrum. *Phys. Rev. Lett.*, 62 :1327–1330.

Chhabra, A. B., Meneveau, C., Jensen, R. V., and Sreenivasan, K. R. (1989). Direct determination of the f($\alpha$) singularity spectrum and its application to fully developed turbulence. *Phys. Rev. A*, 40 :5284–5294.

Chorowski, J. (2018). Nonparametric volatility estimation in scalar diffusions : optimality across observation frequencies. *Bernoulli*, 24(4A) :2934 – 2990.

Chung, K. L. (1948). On the maximum partial sums of sequences of independent random variables. *Transactions of the American Mathematical Society*, 64(2) :205–233.

Çinlar, E. and Kao, J. S. (1991). Particle systems on flows. *Applied Stochastic Models and Data Analysis*, 7(1) :3–15.

Clément, E. and Gloter, A. (2019). Estimating functions for SDE driven by stable Lévy processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(3) :1316 – 1348.

Coeurjolly, J.-F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Statistical Inference for Stochastic Processes*, 4 :199–227.

Coeurjolly, J.-F. (2005). Identification of multifractional Brownian motion. *Bernoulli*, 11(6) :987 – 1008.

Cohen, A. and Sackrowitz, H. B. (1993). Evaluating tests for increasing intensity of a Poisson process. *Technometrics*, 35(4) :446–448.

Cohen, S. (1999). From self-similarity to local self-similarity : the estimation problem. In Dekking, M., Véhel, J. L., Lutton, E., and Tricot, C., editors, *Fractals*, pages 3–16, London. Springer London.

Comte, F. and Genon-Catalot, V. (2011). Estimation for Lévy processes from high frequency data within a long time interval. *The Annals of Statistics*, 39(2) :803–837.

Cont, R. (2001). Empirical properties of asset returns : stylized facts and statistical issues. *Quantitative Finance*, 1 :223 – 236.

Costes, S., Daelemans, D., Cho, E., Dobbin, Z., Pavlakis, G., and Lockett, S. (2004). Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical J*, 86(6) :3993–4003.

Cox, D. R. and Lewis, P. A. (1966). *The statistical analysis of series of events*. Springer.

Csorgo, M. and Horváth, L. (1997). *Limit theorems in change-point analysis*. John Wiley & Sons Chichester.

Dachian, S. and Kutoyants, Y. A. (2006). Hypotheses Testing : Poisson versus self-exciting. *Scandinavian Journal of Statistics*, 33(2) :391–408.

Dang, T., Molnár, S., and Maricza, I. (2003). Queuing performance estimation for general multifractal traffic. *Int. J. Communication Systems*, 16 :117–136.

Daras, N. J. (2014). Stochastic analysis of cyber-attacks. In *Applications of Mathematics and Informatics in Science and Engineering*, pages 105–129. Springer.

Darwin, J. (1956). The behaviour of an estimator for a simple birth and death process. *Biometrika*, 43(1/2) :23–31.

Daubechies, I. (1992). *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics.

Davies, R. B. (1977). Testing the hypothesis that a point process is Poisson. *Advances in Applied Probability*, 9(4) :724–746.

Davis, B. (1984a). On the paths of symmetric stable processes. *Transactions of the American Mathematical Society*, 281(2) :785–785.

Davis, M. H. (1984b). Piecewise-deterministic markov processes : A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society : Series B (Methodological)*, 46(3) :353–376.

Delbeke, L. and Van Assche, W. (1998). A wavelet based estimator for the parameter of self-similarity of fractional Brownian motion.

Deshayes, J. (1984). Rupture de modèles pour des processus de Poisson. *Annales scientifiques de l'Université de Clermont-Ferrand 2. Série Probabilités et applications*, 78(2) :1–7.

Deshayes, J. and Picard, D. (1985). Off-line statistical analysis of change-point models using non parametric and likelihood methods. In *Detection of Abrupt Changes in Signals and Dynamical Systems*, pages 103–168. Springer.

Dette, H., Eckle, T., and Vetter, M. (2020). Multiscale change point detection for dependent data. *Scandinavian Journal of Statistics*, 47(4) :1243–1274.

Diggle, P. (1985). A kernel method for smoothing point process data. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 34(2) :138–147.

Dion, C. and Genon-Catalot, V. (2016). Bidimensional random effect estimation in mixed stochastic differential model. *Statistical Inference for Stochastic Processes*, 19(2) :131–158.

Doi, H., Matsuda, T., and Yamamoto, M. (2004). Performance evaluation of multi-fractal nature of TCP traffic with RED gateway. In *29th Annual IEEE International Conference on Local Computer Networks*, pages 400 – 401.

Donoho, D. L. and Johnstone, I. M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.*, 26(3) :879–921.

Dümbgen, L. and Spokoiny, V. G. (2001). Multiscale testing of qualitative hypotheses. *Ann. Statist.*, 29(1) :124–152.

Durand, A. (2009). Singularity sets of Lévy processes. *Probab. Theory Related Fields*, 114 :1432–2064.

Durand, A. and Jaffard, S. (2012). Multifractal analysis of Lévy fields. *Probability Theory and Related Fields*, 153 :45–96.

Duvernet, L., Robert, C., and Rosenbaum, M. (2010). Testing the type of a semi-martingale : Itō against multifractal. *Electronic Journal of Statistics*, 4(none) :1300 – 1323.

Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1953). Sequential decision problem for processes with continuous time parameter. Testing hypotheses. *The Annals of Mathematical Statistics*, 24(2) :254–264.

Dynkin, E. B. (1965). *Markov processes.* Springer.

Dynkin, E. B. (1991). Branching Particle Systems and Superprocesses. *The Annals of Probability*, 19(3) :1157 – 1194.

El Karoui, N., Loisel, S., and Salhi, Y. (2017). Minimax optimality in robust detection of a disorder time in doubly-stochastic Poisson processes. *The Annals of Applied Probability*, 27(4) :2515–2538.

Ellis, R. S. (1984). Large deviations for a general class of random vectors. *The Annals of Probability*, 12(1) :1 – 12.

Embrechts, P. and Maejima, M. (2002). *Selfsimilar processes.* Princeton series in applied mathematics. Princeton University Press.

Engelhardt, M., Guffey, J. M., and Wright, F. T. (1990). Tests for positive jumps in the intensity of a Poisson process : a power study. *IEEE transactions on reliability*, 39(3) :356–360.

Enikeeva, F., Munk, A., Pohlmann, M., and Werner, F. (2020). Bump detection in the presence of dependency : Does it ease or does it load ? *Bernoulli*, 26(4) :3280–3310.

Enikeeva, F., Munk, A., and Werner, F. (2018). Bump detection in heterogeneous Gaussian regression. *Bernoulli*, 24(2) :1266–1306.

Epstein, B. (1960). Tests for the validity of the assumption that the underlying distribution of life is exponential. Part I. *Technometrics*, 2(1) :83–101.

Ermakov, M. S. (1991). Minimax detection of a signal in a Gaussian white noise. *Theory of Probability & Its Applications*, 35(4) :667–679.

Erramilli, A., Narayan, O., Neidhardt, A., and Saniee, I. (2000). Performance impacts of multi-scaling in wide area TCP/IP traffic. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, volume 1, pages 352–359 vol.1.

Falconer, K. (1990). *Fractal geometry.* John Wiley & Sons Ltd., Chichester. Mathematical foundations and applications.

Falconer, K. (1994). The multifractal spectrum of statistically self-similar measures. *Journal of Theoretical Probability*, 7 :681–702.

Falconer, K. (1997). *Techniques in fractal geometry.* J. Wiley & Sons, Chichester, Royaume-Uni.

Falconer, K. (2002). Tangent fields and the local structure of random fields. *Journal of Theoretical Probability*, 15 :731–750.

Falconer, K. (2003). The local structure of random processes. *Journal of the London Mathematical Society*, 67(3) :657–672.

Falconer, K. and Lévy Véhel, J. (2009). Multifractional, multistable, and other processes with prescribed local form. *J. Theoret. Probab.*, 22 :375–401.

Falconer, K. and Liu, L. (2012). Multistable processes and localizability. *Stochastic Models*, 28 :503–526.

Farinetto, C., Kutoyants, Y. A., and Top, A. (2020). Poisson source localization on the plane : change-point case. *Annals of the Institute of Statistical Mathematics*, 72(3) :675–698.

Fattler, T. and Grothaus, M. (2007). Strong Feller properties for distorted Brownian motion with reflecting boundary condition and an application to continuous N-particle systems with singular interactions. *Journal of Functional Analysis*, 246(2) :217–241.

Feller, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheoretica*, 5(1) :11–40.

Ferguson, T. and Klass, M. (1972). A representation of independent increment processes without Gaussian components. *Ann. Math. Stat.*, 43 :1634–1643.

Figueroa-López, J. E. (2009). Nonparametric estimation for Lévy models based on discrete-sampling. *Lecture notes-monograph series*, pages 117–146.

Finner, H. and Gontscharuk, V. (2009). Controlling the familywise error rate with plug-in estimator for the proportion of true null hypotheses. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 71(5) :1031–1048.

Flandrin, P. (1989). On the spectrum of fractional Brownian motions. *IEEE Trans. Inf. Theory*, 35 :197–199.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional brownian motion. *IEEE Transactions on Information Theory*, 38(2) :910–917.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability*, 30(4) :790.

Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 14(2) :517 – 532.

Frick, K., Munk, A., and Sieling, H. (2014). Multiscale change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, (3) :495–580.

Frisch, U. and Parisi, G. (1985). Fully developed turbulence and intermittancy. In Ghil, M., Benzi, R., and Parisi, G., editors, *Turbulence and predictability of geophysical flows and climate dynamics*, pages 84–88. North Holland.

Fristedt, B. (1979). Uniform local behavior of stable subordinators. *The Annals of Probability*, 7(6) :1003–1013.

Fromont, M. and Laurent, B. (2006). Adaptive goodness-of-fit tests in a density model. *The Annals of Statistics*, 34(2) :680 – 720.

Fromont, M., Laurent, B., and Reynaud-Bouret, P. (2011). Adaptive tests of homogeneity for a Poisson process. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1) :176–213.

Fromont, M., Lerasle, M., and Reynaud-Bouret, P. (2016). Family-wise separation rates for multiple testing. *The Annals of Statistics*, 44(6) :2533–2563.

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Ann. Statist.*, 42(6) :2243–2281.

Fryzlewicz, P. (2018). Tail-greedy bottom-up data decompositions and fast multiple change-point detection. *Ann. Statist.*, 46(6) :3390–3421.

Galeano, P. (2007). The use of cumulative sums for detection of changepoints in the rate parameter of a Poisson process. *Computational Statistics & Data Analysis*, 51(12) :6151–6165.

Ganychenko, I. (2015). Fast L2-approximation of integral-type functionals of Markov processes. *Modern Stochastics : Theory and Applications*, 2 :165–171.

Ganychenko, I., Knopova, V., and Kulik, A. (2015). Accuracy of discrete approximation for integral functionals of Markov processes. *Modern Stochastics : Theory and Applications*, 2 :401–420.

Gao, C., Han, F., and Zhang, C.-H. (2020). On estimation of isotonic piecewise constant signals. *Ann. Statist.*, 48(2) :629–654.

Garreau, D. and Arlot, S. (2018). Consistent change-point detection with kernels. *Electron. J. Stat.*, 12(2) :4440–4486.

Geman, D. and Horowitz, J. (1980). Occupation densities. *The Annals of Probability*, 8(1) :1 – 67.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series model. *Journal of Time Series Analysis*, 4 :221 – 238.

Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. *Probability Theory and Related Fields*, 86 :87–104.

Gobet, E. and Labart, C. (2008). Sharp estimates for the convergence of the density of the Euler scheme in small time. *Electronic Communications in Probability*, 13(0) :352–363.

Gobet, E. and Matulewicz, G. (2017). Parameter estimation of Ornstein–Uhlenbeck process generating a stochastic graph. *Statistical Inference for Stochastic Processes*, 20 :211–235.

Goncalves, P., Riedi, R., and Baraniuk, R. (1998). A simple statistical analysis of wavelet-based multifractal spectrum estimation. In *Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284)*, volume 1, pages 287–291 vol.1.

Gonçalves, P. and Riedi, R. (2005). Diverging moments and parameter estimation. *Journal of the American Statistical Association*, 100(472) :1382–1393.

Grahovac, D. and Leonenko, N. (2014). Bounds on the support of the multifractal spectrum of stochastic processes. *Fractals*, 26(4).

Green, P. J. (1995). Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4) :711–732.

Grossmann, A. and Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. *Siam Journal on Mathematical Analysis*, 15 :723–736.

Guan, Y. (2008). On consistent nonparametric intensity estimation for inhomogeneous spatial point processes. *Journal of the American Statistical Association*, 103(483) :1238–1247.

Häbel, H., Myllymäki, M., and Pommerening, A. (2019). New insights on the behaviour of alternative types of individual-based tree models for natural forests. *Ecological modelling*, 406 :23–32.

Hall, P. and Jin, J. (2010). Innovated higher criticism for detecting sparse signals in correlated noise. *The Annals of Statistics*, 38(3) :1686–1732.

Halsey, T., Jensen, M., Kadanoff, L., Procaccia, I., and Shraiman, B. (1986). Fractal measures and their singularities : the characterization of strange sets. *Phys. Rev.*, A33 :1141–1151.

Hannebicque, B. (2021). *Regularity of generalized stochastic processes*. PhD thesis, Université Paris-Saclay.

Hao, N., Niu, Y. S., and Zhang, H. (2013). Multiple change-point detection via a screening and ranking algorithm. *Statist. Sinica*, 23(4) :1553–1572.

Harchaoui, Z. and Lévy-Leduc, C. (2010). Multiple change-point estimation with a total variation penalty. *J. Amer. Statist. Assoc.*, 105(492) :1480–1493.

Harte, D. (2001). Multifractals : Theory and applications.

Hawkes, J. (1971). On the Hausdorff dimension of the intersection of the range of a stable process with a Borel set. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 19 :90–102.

Hawkes, J. (1974). Local times and zero sets for processes with infinitely divisible distributions. *Journal of the London Mathematical Society*, 2-8(3) :517–525.

Hawkins, D. M. (1977). Testing a sequence of observations for a shift in location. *J. Amer. Statist. Assoc.*, 72(357) :180–186.

Hentschel, H. and Procaccia, I. (1983). The infinite number of generalized dimensions of fractals and strange attractors. *Physica D : Nonlinear Phenomena*, 8(3) :435–444.

Herberts, T. and Jensen, U. (2004). Optimal detection of a change point in a Poisson process for different observation schemes. *Scandinavian Journal of Statistics*, 31(3) :347–366.

Herbin, E. (2006). From $N$ parameter fractional Brownian motions to $N$ parameter multifractional Brownian motions. *Rocky Mountain Journal of Mathematics*, 36(4) :1249 – 1284.

Herbin, E. and Lévy-Véhel, J. (2009). Stochastic 2-microlocal analysis. *Stochastic Processes and their Applications*, 119 :2277–2311.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5) :1163 – 1174.

Hinkley, D. V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika*, 57(1) :1–17.

Hinkley, D. V. and Hinkley, E. A. (1970). Inference about the change-point in a sequence of binomial variables. *Biometrika*, 57 :477–488.

Ho, C.-H. (1993). Forward and backward tests for an abrupt change in the intensity of a Poisson process. *Journal of statistical computation and simulation*, 48(3-4) :245–252.

Ho, C.-H. (1995). A simulation study of a change-point Poisson process based on two well-known test statistics. In *Monte Carlo and quasi-Monte Carlo methods in scientific computing*, pages 228–238. Springer.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4) :800–802.

Hofmann, M. (1999). Lp estimation of the diffusion coefficient. *Bernoulli*, 5(3) :447 – 481.

Holm, H. (2013). A large-scale study of the time required to compromise a computer system. *IEEE Transactions on Dependable and Secure Computing*, 11(1) :2–15.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2) :65–70.

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, 75(2) :383–386.

Hu, X. and Taylor, S. (1997). The multifractal structure of stable occupation measure. *Stochastic Processes and their Applications*, 66(2) :283–299.

Hurvich, C. M., Deo, R. S., and Brodsky, J. (1998). The mean squared error of Geweke and Porte-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 19 :19–46.

Ikeda, N., Nagasawa, M., and Watanabe, S. (1965). On branching Markov processes. *Proceedings of the Japan Academy*, 41(9) :816 – 821.

Ikeda, N., Nagasawa, M., and Watanabe, S. (1966a). A construction of branching Markov processes. *Proceedings of the Japan Academy*, 42(4) :380 – 384.

Ikeda, N., Nagasawa, M., and Watanabe, S. (1966b). A construction of Markov processes by piecing out. *Proceedings of the Japan Academy*, 42(4) :370 – 375.

Ikeda, N., Nagasawa, M., and Watanabe, S. (1966c). Fundamental equations of branching Markov processes. *Proceedings of the Japan Academy*, 4(42) :252 – 257.

Ingster, Y. I. and Kutoyants, Y. A. (2007). Nonparametric hypothesis testing for intensity of the Poisson process. *Mathematical Methods of Statistics*, 16(3) :217–245.

Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Annales IHP - probabilités et statistiques*, 33 :407–436.

Itô, K. (2004). *Stochastic processes*. Springer, Berlin, Heidelberg.

Itô, K. and McKean, H. (1964). *Diffusion processes and their sample paths*. Springer, Berlin, Heidelberg.

Ivanovs, J. and Podolskij, M. (2020). Optimal estimation of some random quantities of a Lévy process. *arXiv preprint arXiv :2001.02517*.

Ivanovs, J. and Podolskij, M. (2022). Optimal estimation of the supremum and occupation times of a self-similar Lévy process. *Electronic Journal of Statistics*, 16(1) :892 – 934.

Jacod, J. (1998). Rates of convergence to the local time of a diffusion. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 34(4) :505–544.

Jacquet, P. (1998). Long term dependences and heavy tails in traffics and queues generated by memoriless on/off sources in series. Research Report RR-3516, INRIA.

Jaffard, S. (1991). Pointwise smoothness, 2-microlocalization and wavelet coefficients. *Publicacions matematiques, ISSN 0214-1493, Vol. 35, Nº 1, 1991, pags. 155-168*, 35.

Jaffard, S. (1997). Multifractal formalism for functions part II : self-similar functions. *Siam Journal on Mathematical Analysis*, 28 :971–998.

Jaffard, S. (1999). The multifractal nature of Lévy processes. *Probability Theory and Related Fields*, 114 :207–227.

Jaffard, S. (2000). On the Frisch-Parisi Conjecture. *Journal De Mathematiques Pures Et Appliquees*, 79 :525–552.

Jaffard, S. and Meyer, Y. (2000). On the pointwise regularity of functions in critical Besov spaces. *Journal of Functional Analysis*, 175 :415–434.

James, B., James, K. L., and Siegmund, D. (1987). Tests for a change-point. *Biometrika*, 74(1) :71–83.

Jaramillo, A., Nourdin, I., and Peccati, G. (2021). Approximation of fractional local times : zero energy and derivatives. *The Annals of Applied Probability*, 31(5) :2143–2191.

Jeng, X. J., Cai, T. T., and Li, H. (2010). Optimal sparse segment identification with application in copy number variation analysis. *J. Amer. Statist. Assoc.*, 105(491) :1156–1166.

Jin, J. (2008). Proportion of non-zero normal means : universal oracle equivalences and uniformly consistent estimators. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 70(3) :461–493.

Jing, B.-Y., Kong, X.-B., Liu, Z., and Mykland, P. (2012). On the jump activity index for semimartingales. *Journal of Econometrics*, 166(2) :213–223.

Kahane, J. (1974). Sur l'irrégulatité locale du mouvement Brownien. *C. R. Acad. Sciences Paris*, 278 :331–333.

Kahane, J. (1976). Sur les zéros et les instants de ralentissement du mouvement Brownien. *C. R. Acad. Sciences Paris*, 282 :431–433.

Kander, Z. and Zacks, S. (1966). Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points. *The Annals of Mathematical Statistics*, 37(5) :1196 – 1210.

Karlin, S. and McGregor, J. (1957). The classification of birth and death processes. *Transactions of the American Mathematical Society*, 86(2) :366–400.

Keiding, N. (1975). Maximum likelihood estimation in the birth-and-death process. *The Annals of Statistics*, 3(2) :363–372.

Kendall, D. (1952). Les processus stochastiques de croissance en biologie. *Annales de l'institut Henri Poincaré*, 13(1) :43–108.

Kendall, D. G. (1949). Stochastic processes and population growth. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(2) :230–282.

Kendall, D. G. and Kendall, W. S. (1980). Alignments in two-dimensional random sets of points. *Advances in Applied Probability*, 12(2) :380–424.

Khoshnevisan, D., Peres, Y., and Xiao, Y. (2000). Limsup random fractals. *Electronic Journal of Probability*, 5(none) :1 – 24.

Khoshnevisan, D. and Shi, Z. (2000). Fast sets and points for fractional Brownian motion. *Séminaire de probabilités de Strasbourg*, 34 :393–416.

Khoshnevisan, D., Shieh, N.-R., and Xiao, Y. (2008). Hausdorff dimension of the contours of symmetric additive Lévy processes. *Probability Theory and Related Fields*, 140 :129–167.

Khoshnevisan, D. and Xiao, Y. (2002). Level sets of additive Lévy processes. *The Annals of Probability*, 30(1) :62 – 100.

Khoshnevisan, D. and Xiao, Y. (2004). Additive Lévy processes : capacity and Hausdorff dimension. In Bandt, C., Mosco, U., and Zähle, M., editors, *Fractal Geometry and Stochastics III*, pages 151–170, Basel. Birkhäuser Basel.

Khoshnevisan, D. and Xiao, Y. (2005). Lévy processes : capacity and Hausdorff dimension. *The Annals of Probability*, 33(3) :841 – 878.

Khoshnevisan, D., Xiao, Y., and Zhong, Y. (2003). Measuring the range of an additive Lévy process. *The Annals of Probability*, 31(2) :1097 – 1141.

King, J. (1995). The singularity spectrum for general Sierpinski carpets. *Advances in Mathematics*, 116(1) :1–11.

Kingman, J. (1993). *Poisson processes.* Oxford University Press.

Kohatsu-Higa, A., Makhlouf, R., and Ngo, H.-L. (2014). Approximations of non-smooth integral type functionals of one dimensional diffusion processes. *Stochastic Processes and their Applications*, 124(5) :1881–1909.

Kolmogorov, A. (1940). Wienersche spiralen und einige andere interessante Kurven in Hilbertscen Raum, C. R. (doklady). 26 :115–118.

Kolwankar, K. and Lévy Véhel, J. (2002). A time domain characterization of the fine local regularity of functions. *Journal of Fourier Analysis and Applications*, 8 :319–334.

Kôno, N. and Maejima, M. (1991). Hölder continuity of sample paths of some self-similar stable processes. *Tokyo Journal of Mathematics*, 14(1) :93 – 100.

Kou, J. (2023). Identifying the support of rectangular signals in Gaussian noise. *Communications in Statistics - Theory and Methods*, 52(10) :3262–3289.

Koutrouvelis, I. A. (1980). Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 75 :918–928.

Kumar Patra, R. and Sen, B. (2016). Estimation of a two-component mixture model with applications to multiple testing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 78(4) :869–893.

Lacaux, C. (2004). Real harmonizable multifractional Lévy motions. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 40(3) :259–277.

Lagache, T., Sauvonnet, N., Danglot, L., and Olivo-Marin, J.-C. (2015). Statistical analysis of molecule colocalization in bioimaging. *Cytometry Part A*, 87(6) :568–579.

Lai, T. L. (2001). Sequential analysis : some classical problems and new challenges. *Statistica Sinica*, 11(2) :303–351.

Lang, G. and Azaïs, J.-M. (1999). Non-parametric estimation of the long-range dependence exponent for Gaussian processes. *Journal of Statistical Planning and Inference*, 80(1-2) :59–80.

Langaas, M., Lindqvist, B. H., and Ferkingstad, E. (2005). Estimating the proportion of true null hypotheses, with application to DNA microarray data. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 67(4) :555–572.

Lashermes, B., Jaffard, S., and Abry, P. (2005). Wavelet leader based multifractal analysis. In *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 4, pages iv/161–iv/164 Vol. 4.

Lavielle, M. (2005). Using penalized contrasts for the change-point problem. *Signal processing*, 85(8) :1501–1510.

Lavielle, M. and Moulines, E. (2000). Least-squares estimation of an unknown number of shifts in a time series. *J. Time Ser. Anal.*, 21(1) :33–59.

Le Page, R. (1980a). Multidimensional infinitely divisible variables and processes. I. Stable case. *Tech. Rep. 292, Dept. Stat*, 26 :115–118.

Le Page, R. (1980b). Multidimensional infinitely divisible variables and processes. II. *Probability in Banach Spaces III, Lecture notes in Math.*, 860 :279–284.

Lebarbier, É. (2005). Detecting multiple change-points in the mean of Gaussian process by model selection. *Signal processing*, 85(4) :717–736.

Lebovits, J. (2012). *Stochastic calculus with respect to multi-fractional Brownian motion and applications to finance.* Theses, Ecole Centrale Paris ; laboratoire probabilités et modèles aléatoires.

Lebovits, J. and Podolskij, M. (2017). Estimation of the global regularity of a multifractional Brownian motion. *Electronic Journal of Statistics*, 11(1) :78 – 98.

Lee, T.-S. (2010). Change-point problems : bibliography and review. *Journal of Statistical Theory and Practice*, 4(4) :643–662.

Lehmann, E. L., Romano, J. P., and Shaffer, J. P. (2005). On optimality of stepdown and stepup multiple test procedures. *The Annals of Statistics*, 33(3) :1084–1108.

Leland, W., Taqqu, M., Willinger, W., and Wilson, D. (1994). On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1) :1–15.

Leonard, T. (1978). Density estimation, stochastic processes and prior information. *J. Roy. Statist. Soc. Ser. B*, 40(2) :113–146.

Leonarduzzi, R., Abry, P., Wendt, H., Jaffard, S., and Touchette, H. (2019). A generalized multifractal formalism for the estimation of nonconcave multifractal spectra. *IEEE Transactions on Signal Processing*, 67(1) :110–119.

Leonarduzzi, R., Touchette, H., Wendt, H., Abry, P., and Jaffard, S. (2016). Generalized Legendre transform multifractal formalism for nonconcave spectrum estimation. In *IEEE Workshop on statistical signal processing (SSP 2016)*, pages pp. 1–5, Palma de Mallorca, Spain.

Lepski, O. V. and Spokoiny, V. G. (1999). Minimax nonparametric hypothesis testing : The case of an inhomogeneous alternative. *Bernoulli*, 5 :333–358.

Lepski, O. V. and Tsybakov, A. B. (2000). Asymptotically exact nonparametric hypothesis testing in sup-norm and at a fixed point. *Probab. Theory Related Fields*, 117(1) :17–48.

Lévy, P. (1948). *Processus stochastiques et mouvement brownien*. Gauthier-Vilars, Paris.

Lévy Véhel, J., Philippe, A., and Robet, C. (2021). Explicit and combined estimators for parameters of stable distributions. *Statistics*, 55(4) :736–764.

Lévy Véhel, J. and Rams, M. (2013). Large deviation multifractal analysis of a class of additive processes with correlated nonstationary increments. *IEEE/ACM Transactions on Networking*, 21(4) :1309–1321.

Lévy Véhel, J. and Riedi, R. (1997). Fractional Brownian motion and data traffic modeling : the other end of the spectrum. In Vehel, J. L., Lutton, E., and Tricot, C., editors, *Fractals in Engineering : from theory to industrial applications*. Springer.

Lévy Véhel, J. and Seuret, S. (2003). A time domain characterization of 2-microlocal spaces. *Journal of Fourier Analysis and Applications*, 9.

Lévy Véhel, J. and Tricot, C. (2004). On various multifractal spectra. In Bandt, C., Mosco, U., and Zähle, M., editors, *Fractal Geometry and Stochastics III*, volume 57 of *Progress in Probability*, pages 23–42. Springer.

Lévy Véhel, J. and Vojak, R. (1998). Multifractal analysis of Choquet capacities. *Advances in Applied Mathematics*, 20(1) :1–43.

Lewis, P. A. W. (1965). Some results on tests for Poisson processes. *Biometrika*, 52(1-2) :67–77.

Li, H., Munk, A., and Sieling, H. (2016). FDR-control in multiscale change-point segmentation. *Electron. J. Stat.*, 10(1) :918–959.

Li, Z. (2011). *Measure-Valued Branching Processes*, pages 29–56. Springer Berlin Heidelberg.

Liu, H., Gao, C., and Samworth, R. J. (2021). Minimax rates in sparse, high-dimensional change point detection. *The Annals of Statistics*, 49(2) :1081–1112.

Loader, C. R. (1990). *Change point problems for Poisson processes.* PhD thesis, Stanford University.

Löcherbach, E. (2002). Likelihood ratio processes for markovian particle systems with killing and jumps. *Statistical inference for stochastic processes*, 5(2) :153–177.

Lorden, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Statist.*, 42 :1897–1908.

Lörinczi, J. and Yang, X. (2019). Multifractal properties of sample paths of ground state-transformed jump processes. *Chaos, Solitons & Fractals*, 120(C) :83–94.

Maejima, M. (1983). A self-similar process with nowhere bounded sample paths. *Z. Wahrscheinlichkeitstheorie verw Gebiete*, 65 :115–119.

Mallat, S. and Hwang, W., L. (1992). Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2) :617–643.

Mandelbrot, B. and Riedi, R. (1995). Multifractal formalism for infinite multinomial measures. *Advances in Applied Mathematics*, 16(2) :132–150.

Mandelbrot, B. and Van Ness, J. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4) :422–437.

Mandelbrot, B. B. (1974). Intermittent turbulence in self-similar cascades : divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics*, 62(2) :331–358.

Manjrekar, M. (2018). A hard-core stochastic process with simultaneous births and deaths. *Stochastic Models*, 34(4) :375–396.

Marcus, M. and Rosinski, J. (2005). Continuity and boundedness of infinitely divisible processes : a Poisson point process approach. *Journal of Theoretical Probability*, 18 :109–160.

Marquardt, T. (2006). Fractional Lévy processes with an application to long memory moving average processes. *Bernoulli*, 12(6) :1099 – 1126.

Martinussen, T. and Scheike, T. H. (2007). *Dynamic regression models for survival data.* Springer Science & Business Media.

Masuda, N. and Holme, P. (2017). *Temporal network epidemiology.* Springer.

McCulloch, J. (1986). Simple consistent estimator of stable distribution parameters. *Communications in Statistics - Simulation and Computation - CSSC*, 15.

McCulloch, J. (1997). Measuring tail thickness to estimate the stable index $\alpha$ : a critique. *Journal of Business & Economic Statistics*, 15(1) :74–81.

Meinshausen, N. and Rice, J. (2006). Estimating the proportion of false null hypotheses among a large number of independently tested hypotheses. *Ann. Statist.*, 34(1) :373–393.

Meyer, Y. (1998). *Wavelets, vibrations and scalings.* Number 9. American Mathematical Soc.

Mies, F. (2020). Rate-optimal estimation of the Blumenthal–Getoor index of a Lévy process. *Electronic Journal of Statistics*, 14(2) :4165 – 4206.

Millar, P. W. (1971). Path behavior of processes with stationary independent increments. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 17 :53–73.

Møller, J. (1989). On the rate of convergence of spatial birth-and-death processes. *Annals of the Institute of Statistical Mathematics*, 41(3) :565–581.

Monrad, D. and Rootzén, H. (1995). Small values of Gaussian processes and functional laws of the iterated logarithm. *Probability Theory and Related Fields*, 101 :173–192.

Morales, C. and Kolaczyk, E. (2002). Wavelet-based multifractal analysis of human balance. *Annals of biomedical engineering*, 30 :588–97.

Moran, P. A. P. (1951). Estimation methods for evolutive processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(1) :141–146.

Moran, P. A. P. (1953). The estimation of the parameters of a birth and death process. *Journal of the Royal Statistical Society. Series B (Methodological)*, 15(2) :241–245.

Moulines, E. and Soulier, P. (1999). Broadband log-periodogram regression of time series with long-range dependence. *The Annals of Statistics*, 27(4) :1415 – 1439.

Moulines, E. and Soulier, P. (2003). Semiparametric spectral estimation for fractional processes. In Doukhan, P., Oppenheim, G., and Taqqu, M., editors, *Theory and Applications of Long-Range Dependence*. Birkhäuser.

Moustakides, G. V. (1986). Optimal stopping times for detecting changes in distributions. *Ann. Statist.*, 14(4) :1379–1387.

Moustakides, G. V. (2004). Optimality of the CUSUM procedure in continuous time. *Ann. Statist.*, 32(1) :302–315.

Moustakides, G. V. et al. (2008). Sequential change detection revisited. *The Annals of Statistics*, 36(2) :787–807.

Muzy, J., Bacry, E., and Arnéodo, A. (1993). Multifractal formalism for fractal signals : the structure-function approach versus the wavelet-transform modulus-maxima method. *Phys. Rev. E*, 47 :875–884.

Muzy, J., Bacry, E., and Arnéodo, A. (1994). The multifractal formalism revisited with wavelets. *International Journal of Bifurcation and Chaos*, 04(02) :245–302.

Mytnik, L. and Neuman, E. (2012). Sample path properties of Volterra processes. *Communications on Stochastic Analysis*, 6.

Mytnik, L. and Wachtel, V. (2012). Multifractal analysis of superprocesses with stable branching in dimension one. *The Annals of Probability*, 43(5) :2763–2809.

Neuenkirch, A. and Szölgyenyi, M. (2020). The Euler Maruyama scheme for SDEs with irregular drift : convergence rates via reduction to a quadrature problem. *IMA Journal of Numerical Analysis*, (draa007).

Neuman, E. (2014). The multifractal nature of Volterra Lévy processes. *Stochastic Processes and their Applications*, 124 :3121–3145.

Ngo, H.-L. and Ogawa, S. (2011). On the discrete approximation of occupation time of diffusion processes. *Electronic Journal of Statistics*, 5 :1374–1393.

Nickl, R. and van de Geer, S. (2013). Confidence sets in sparse regression. *Ann. Statist.*, 41(6) :2852–2876.

Nolan, J. (1989a). Continuity of symmetric stable processes. *Journal of Multivariate Analysis*, 29 :84–93.

Nolan, J. (1989b). *Local nondeterminism and local times for stable processes*, volume 82.

Nolan, J. (2001). Maximum likelihood estimation and diagnostics for stable distributions.

Øksendal, B. (2013). *Stochastic differential equations : an introduction with applications.* Springer Science & Business Media.

Olsen, L. (1994). *Random geometrically graph directed self-similar multifractals.* Chapman and Hall/CRC.

Olsen, L. (1996). Multifractal dimensions of product measures. *Mathematical Proceedings of the Cambridge Philosophical Society*, 120(4) :709–734.

Olsen, L. (1998). Self-affine multifractal Sierpinski sponges in $R^d$. *Pacific Journal of Mathematics*, 183 :143–199.

Olshen, A. B., Venkatraman, E. S., Lucito, R., and Wigler, M. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*, 5(4) :557–572.

Orey, S. and Taylor, S. J. (1974). How often on a Brownian path does the law of iterated logarithm fail ? *Proceedings of the London Mathematical Society*, s3-28(1) :174–192.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, 41(1/2) :100–115.

Page, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika*, 42(3/4) :523–527.

Page, E. S. (1957). On problems in which a change in a parameter occurs at an unknown point. *Biometrika*, 44(1/2) :248–252.

Panigrahi, S., Roy, P., and Xiao, Y. (2021). Maximal moments and uniform modulus of continuity for stable random fields. *Stochastic Processes and their Applications*, 136 :92–124.

Park, H., Xiao, Y., and Yang, X. (2020). Uniform dimension results for the inverse images of symmetric Lévy processes. *Journal of Theoretical Probability*, 33 :2213–2232.

Paxson, V. and Floyd, S. (1995). Wide area traffic : the failure of Poisson modeling. *IEEE/ACM Transactions on networking*, 3(3) :226–244.

Pécot, T., Zengzhen, L., Boulanger, J., Salamero, J., and Kervrann, C. (2018). A quantitative approach for analyzing the spatio-temporal distribution of 3d intracellular events in fluorescence microscopy. *eLife*, 7 :e32311.

Pein, F., Sieling, H., and Munk, A. (2017). Heterogeneous change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 79(4) :1207–1227.

Peltier, R.-F. and Lévy Véhel, J. (1995). Multifractional Brownian Motion : definition and preliminary results. Research Report RR-2645, INRIA. Projet FRACTALES.

Peng, T., Leckie, C., and Ramamohanarao, K. (2004). Proactively detecting distributed denial of service attacks using source IP address monitoring. In *International conference on research in networking*, pages 771–782. Springer.

Perkins, E. (1983). On the Hausdorff dimension of the Brownian slow points. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 64(3) :369–399.

Perry, M. B. and Pignatiello, J. J. (2011). Estimating the time of step change with Poisson CUSUM and EWMA control charts. *International Journal of Production Research*, 49(10) :2857–2871.

Peskir, G. and Shiryaev, A. N. (2002). Solving the Poisson disorder problem. In *Advances in finance and stochastics*, pages 295–312. Springer.

Peyrière, J. (1992). *Multifractal measures*, pages 175–186. Springer Netherlands, Dordrecht.

Picard, F., Lebarbier, É., Budinskà, E., and Robin, S. (2011). Joint segmentation of multivariate Gaussian processes using mixed linear models. *Comput. Statist. Data Anal.*, 55(2) :1160–1170.

Pollak, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.*, 13(1) :206–227.

Polunchenko, A. S. and Tartakovsky, A. G. (2012). State-of-the-art in sequential change-point detection. *Methodology and computing in applied probability*, 14(3) :649–684.

Pommerening, A. and Grabarnik, P. (2019). *Individual-Based Methods in Forest Ecology and Management*. Springer.

Preston, C. (1975). Spatial birth and death processes. *Advances in applied probability*, 7(3) :371–391.

Pruitt, W. E. (1969). The Hausdorff dimension of the range of a process with stationary independent increments. *Journal of Mathematics and Mechanics*, 19(4) :371–378.

Pruitt, W. E. (1981). The growth of random walks and Lévy Processes. *The Annals of Probability*, 9(6) :948 – 956.

Raftery, A. E. (1994). Change point and change curve modeling in stochastic processes and spatial statistics. *Journal of Applied Statistical Science*, 1(4) :403–423.

Raftery, A. E. and Akman, V. E. (1986). Bayesian analysis of a Poisson process with a change-point. *Biometrika*, 73(1) :85–89.

Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *The Annals of Statistics*, pages 453–466.

Rand, D. (1989). The singularity spectrum $f(\alpha)$ for cookie-cutters. *Ergodic Theory and Dynamical Systems*, 9(3) :527–541.

Reiß, M. (2013). Testing the characteristics of a Lévy process. *Stochastic Processes and their Applications*, 123(7) :2808–2828. A Special Issue on the Occasion of the 2013 International Year of Statistics.

Renshaw, E. and Särkkä, A. (2001). Gibbs point processes for studying the development of spatial-temporal stochastic processes. *Computational statistics & data analysis*, 36(1) :85–105.

Reynaud-Bouret, P. (2003). Adaptive estimation of the intensity of inhomogeneous poisson processes via concentration inequalities. *Probability Theory and Related Fields*, 126(1) :103–153.

Reynaud-Bouret, P. and Schbath, S. (2010). Adaptive estimation for hawkes processes ; application to genome analysis. *The Annals of Statistics*, 38(5) :2781–2822.

Reynolds, J. F. (1973). On estimating the parameters of a birth-death process. *Australian Journal of Statistics*, 15(1) :35–43.

Ribeiro, V., Riedi, R., Crouse, M., and Baraniuk, R. (2000). Multiscale queuing analysis of long-range-dependent network traffic. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, volume 2, pages 1026–1035 vol.2.

Richard, A. (2014). *Local regularity of some fractional Brownian fields.* These, Ecole Centrale Paris ; Bar-Ilan university (Ramat-Gan, Israël).

Riedi, R. (2002). Multifractal processes. In Doukhan, Oppenheim, and Taqqu, editors, *Long range dependence : theory and applications*, chapter 28, pages 625–715. Birkhäuser.

Riedi, R., Crouse, M., Ribeiro, V., and Baraniuk, R. (1999). A multifractal wavelet model with application to network traffic. *IEEE Transactions on Information Theory*, 45 :992–1018.

Riedi, R. H. (1995). An improved multifractal formalism and self similar measures. *Journal of Mathematical Analysis and Applications*, 189 :462–490.

Rigaill, G. (2010). Pruned dynamic programming for optimal multiple change-point detection. *arXiv preprint arXiv :1004.0887*, 17.

Rivera, C. and Walther, G. (2013). Optimal detection of a jump in the intensity of a Poisson process or in a density with likelihood ratio statistics. *Scand. J. Stat.*, 40(4) :752–769.

Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, 23(5) :1630 – 1661.

Rom, D. M. (1990). A sequentially rejective test procedure based on a modified Bonferroni inequality. *Biometrika*, 77(3) :663–665.

Romano, J. P., Shaikh, A., and Wolf, M. (2011). Consonance and the closure method in multiple testing. *The International Journal of Biostatistics*, 7(1).

Sadahiro, Y. (2019). Analysis of the appearance and disappearance of point objects over time. *International Journal of Geographical Information Science*, 33(2) :215–239.

Samorodnitsky, G. (2016). *Stochastic Processes and Long Range Dependence.* Springer.

Samorodnitsky, G. and Taqqu, M. (1994). *Stable non-Gaussian random processes.* Chapman and Hall.

Sarkar, S. K. (2007). Stepup procedures controlling generalized FWER and generalized FDR. *The Annals of Statistics*, 35(6) :2405 – 2420.

Sarkar, S. K., Guo, W., and Finner, H. (2012). On adaptive procedures controlling the familywise error rate. *Journal of Statistical Planning and Inference*, 142(1) :65–78.

Särkkä, A. and Renshaw, E. (2006). The analysis of marked point patterns evolving through space and time. *Computational Statistics & Data Analysis*, 51(3) :1698–1718.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press.

Schilder, M. (1970). Some structure theorems for the symmetric stable laws. *Ann. Math. Statist.*, 41(2) :412–421.

Schuhmacher, D. and Xia, A. (2008). A new metric between distributions of point processes. *Advances in applied probability*, 40(3) :651–672.

Schweder, T. and Spjøtvoll, E. (1982). Plots of p-values to evaluate many tests simultaneously. *Biometrika*, 69 :493–502.

Serrano, E. and Figliola, A. (2009). Wavelet leaders : a new method to estimate the multifractal singularity spectra. *Physica A : Statistical Mechanics and its Applications*, 388(14) :2793–2805.

Seuret, S. and Lévy Véhel, J. (2002). The local Hölder function of a continuous function. *Applied and Computational Harmonic Analysis*, 13 :263–276.

Seuret, S. and Yang, X. (2017). Multifractal analysis for the occupation measure of stable-like processes. *Electronic Journal of Probability*, 22(none) :1 – 36.

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, 81(395) :826–831.

Shen, J. J. and Zhang, N. R. (2012). Change-point model on nonhomogeneous Poisson processes with application in copy number profiling by next-generation DNA sequencing. *The Annals of Applied Statistics*, 6(2) :476–496.

Shin, S. J., Wu, Y., and Hao, N. (2020). A backward procedure for change-point detection with applications to copy number variation detection. *Canadian Journal of Statistics*, 48(3) :366–385.

Shiryaev, A. N. (1978). *Optimal stopping rules.* Springer-Verlag, New York-Heidelberg.

Siegmund, D. (1985). *Sequential analysis.* Springer Series in Statistics. Springer-Verlag, New York.

Siegmund, D. (1988). Confidence sets in change-point problems. *Internat. Statist. Rev.*, 56(1) :31–48.

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, 73(3) :751–754.

Skorokhod, A. V. (1964). Branching diffusion processes. *Theory of Probability & Its Applications*, 9(3) :445–449.

Soltani, R., Goeckel, D., Towsley, D., and Houmansadr, A. (2015). Covert communications on Poisson packet channels. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1046–1052.

Soltani, R., Goeckel, D., Towsley, D., and Houmansadr, A. (2020). Fundamental limits of covert packet insertion. *IEEE Transactions on Communications*, 68(6) :3401–3414.

Soltani, S., Khayam, S. A., and Radha, H. (2008). Detecting malware outbreaks using a statistical model of blackhole traffic. In *2008 IEEE International Conference on Communications*, pages 1593–1597.

Song, R., Xiao, Y., and Yang, X. (2018). Uniform Hausdorff dimension result for the inverse images of stable Lévy processes. *Electronic Communications in Probability*, 23 :1 – 10.

Spokoiny, V. G. (1996). Adaptive hypothesis testing using wavelets. *Ann. Statist.*, 24(6) :2477–2498.

Stoev, S. and Taqqu, M. (2004a). Path properties of the linear multifractional stable motion. *Fractals*, 13.

Stoev, S. A. and Taqqu, M. S. (2004b). Stochastic properties of the linear multifractional stable motion. *Advances in Applied Probability*, 36 :1085 – 1115.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society : Series B*, 64(3) :479–498.

Takashima, K. (1989). Sample path properties of ergodic self-similar processes. *Osaka Journal of Mathematics*, 26(1) :159 – 189.

Taqqu, M., Willinger, W., and Sherman, R. (1998). Proof of a fundamental result in self-similar traffic modeling. *Computer Communication Review*, 27.

Taqqu, M. and Wolpert, R. (1983). Infinite variance self-similar processes subordinate to a Poisson measure. *Z. Wahrscheinlichkeitstheorie verw Gebiete*, 62 :52–73.

Todorov, V. (2015). Jump activity estimation for pure-jump semimartingales via self-normalized statistics. *Annals of Statistics*, 43 :1831–1864.

Tudor, C. A. and Xiao, Y. (2007). Sample path properties of bifractional Brownian motion. *Bernoulli*, 13(4) :1023 – 1052.

Van Es, B., Gugushvili, S., and Spreij, P. (2007). A kernel type nonparametric density estimator for decompounding. *Bernoulli*, 13(3) :672–694.

Veitch, D. and Abry, P. (1999). A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Transactions on Information Theory*, 45(3) :878–897.

Veitch, D. and Abry, P. (2001). A statistical test for the time constancy of scaling exponents. *IEEE Transactions on Signal Processing*, 49 :2325–2334.

Vervaat, W. (1985). Sample path properties of self-similar processes with stationary increments. *The Annals of Probability*, 13(1) :1 – 27.

Verzelen, N., Fromont, M., Lerasle, M., and Reynaud-Bouret, P. (2020). Optimal change-point detection and localization. *arXiv preprint 10.48550/arxiv.2010.11470*.

Wald, A. (1945). Sequential tests of statistical hypotheses. *The annals of mathematical statistics*, 16(2) :117–186.

Wang, D., Yu, Y., and Rinaldo, A. (2020). Univariate mean change point detection : Penalization, CUSUM and optimality. *Electron. J. Stat.*, 14(1) :1917–1961.

Wang, L. (2018). The continuous-time Poisson channel has infinite covert communication capacity. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 756–760.

Wang, T. and Samworth, R. J. (2018). High dimensional change point estimation via sparse projection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 80(1) :57–83.

Wang, Y. and Zhu, S.-C. (2002). A generative method for textured motion : Analysis and synthesis. In *European Conference on Computer Vision*, pages 583–598. Springer.

Wendt, H. and Abry, P. (2007). Multifractality tests using bootstrapped wavelet leaders. *IEEE Transactions on Signal Processing*, 55(10) :4811–4820.

Wendt, H., Abry, P., and Jaffard, S. (2007). Bootstrap for empirical multifractal analysis. *IEEE Signal Processing Magazine*, 24(4) :38–48.

Wendt, H., Jaffard, S., and Abry, P. (2012). Multifractal analysis of self-similar processes. In *2012 IEEE Statistical Signal Processing Workshop (SSP)*, pages 69–72.

Westfall, P. H., Tobias, R. D., and Wolfinger, R. D. (2011). *Multiple comparisons and multiple tests using SAS*. SAS Institute.

Willinger, W., Taqqu, M., Sherman, R., and Wilson, D. (1997). Self-similarity through high-variability : statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Transactions on Networking*, 5(1) :71–86.

Wolff, R. W. (1965). Problems of statistical inference for birth and death queuing models. *Operations Research*, 13(3) :343–357.

Worsley, K. J. (1986). Confidence regions and test for a change-point in a sequence of exponential family random variables. *Biometrika*, 73(1) :91–104.

Wu, L. and Ding, Y. (2017). Wavelet-based estimator for the Hurst parameters of fractional Brownian sheet. *Acta Mathematica Scientia*, 37(1) :205–222.

Xiao, Y. (2004). Random fractals and Markov processes. *Fractal Geometry and Applications. Proc. Symp. Pure Math*, 72 :261–338.

Xiao, Y. (2006). Strong local nondeterminism and the sample path properties of Gaussian random fields.

Xiao, Y. (2009). Uniform modulus of continuity of random fields. *Monatshefte für Mathematik*, 159.

Xu, L. (2016). The multifractal nature of Boltzmann processes. *Stochastic Processes and their Applications*, 126(8) :2181–2210.

Yang, X. (2015). Hausdorff dimension of the range and the graph of stable-like processes. *Journal of Theoretical Probability*, 31 :2412–2431.

Yang, X. (2018). Multifractality of jump diffusion processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 54(4) :2042–2074.

Yao, Y.-C. (1988). Estimating the number of change-points via Schwarz' criterion. *Statistics & Probability Letters*, 6(3) :181–189.

Zhan, Z., Xu, M., and Xu, S. (2013). Characterizing honeypot-captured cyber attacks : Statistical framework and case study. *IEEE Transactions on Information Forensics and Security*, 8(11) :1775–1789.