

UNIVERSITE DE POITIERS

UFR Sciences Fondamentales et Appliquées

Document de synthèse en vue de

L'HABILITATION A DIRIGER DES RECHERCHES

Spécialité : Mathématiques appliquées

**Théorie spatiale des extrêmes et
propriétés des processus max-stables**
par
Clément Dombry

Soutenance le 8 Novembre 2012 devant le jury composé de :

M. Marc ARNAUDON	Université de Poitiers	Examineur
M. Jean-Christophe BRETON	Université Rennes 1	Examineur
M. Youri DAVYDOV	Université Lille 1	Examineur
M. Laurens DE HAAN	Erasmus University Rotterdam	Rapporteur
M. Christian MAZZA	Université de Fribourg	Examineur
M. Julien MICHEL	Université de Poitiers	Examineur
M. Gennady SAMORODNITSKY	Cornell University	Rapporteur
M. Philippe SOULIER	Université Paris Ouest	Rapporteur

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude aux rapporteurs de cette habilitation, Laurens de Haan, Gennady Samorodnitsky et Philippe Soulier qui m'ont fait l'honneur d'accepter cette tâche et d'expertiser mon travail. Je remercie également les membres du jury : Marc Arnaudon, Jean-Christophe Breton, Youri Davydov, Christian Mazza et Julien Michel.

Ma profonde reconnaissance va bien sûr ensuite à mes directeurs de thèse, Christian Mazza et Nadine Guillotin-Plantard, qui m'ont initié à la recherche mathématique et m'ont toujours soutenu, depuis mes débuts jusqu'à aujourd'hui. Je les en remercie chaleureusement.

Je tiens aussi à remercier vivement mes collègues du Laboratoire de Mathématiques et Applications de l'Université de Poitiers, qui m'ont accueilli dans une ambiance agréable et chaleureuse. Je pense en particulier à Pol Vanhaecke, directeur du LMA, et à Marc Arnaudon, ancien responsable de l'équipe de probabilités, pour leur disponibilité, leur soutien et leurs conseils lors de mon arrivée à Poitiers. Ils ont certainement contribué pour beaucoup au dynamisme et au plaisir de mes débuts en tant que jeune Maître de Conférences. Merci à Madalina Petcu, Jocelyne Attab et Julien Michel pour leur amitiés et les bons moments partagés au LMA - et en dehors ! Merci à Jocelyne Attab, Brigitte Brault, Nathalie Marlet, Benoît Métrot et Nathalie Mongin pour leur disponibilité et les nombreux services rendus au quotidien.

Un grand merci aux personnes qui m'ont accompagné sur les chemins de la recherche mathématiques, pour leurs collaborations passées, présentes et futures j'espère : Nadine Guillotin-Plantard, Christian Mazza, René Schott, Serge Cohen, Jean-Christophe Breton, Eli Upfal, Vincent Bansaye, Ingemar Kaj, Youri Davydov, Marc Arnaudon, Frédéric Eyi-Minko et Mathieu Ribatet. Des remerciements tout particuliers à Mathieu pour m'avoir fait découvrir le domaine passionnant de la théorie des valeurs extrêmes et pour les discussions enthousiastes dans la bonne humeur et l'amitié.

Enfin, merci à mes parents, et à Jean-Charles, pour leur soutien et leur affection de tous les jours.

Table des matières

Remerciements	iii
Table des matières	vi
Avant-propos	vii
Liste de publications	ix
1 Théorie spatiale des extrêmes et propriétés des processus max-stables	1
1.1 Introduction	2
1.2 Lois conditionnelles des processus max-infiniment divisibles [D19]	5
1.3 Simulations conditionnelles des processus de Brown-Resnick [D20]	13
1.4 Propriétés de mélange fort des processus max-infiniment divisibles [D21]	20
1.5 Extrêmes de champs aléatoires i.i.d. et processus ponctuels [D18]	25
1.6 Modèles de shot-noise extrémaux [D14]	30
1.7 Perspectives	35
Bibliographie	38
2 Théorèmes limites, dépendance à longue portée et variations régulières	43
2.1 Introduction	44
2.2 Théorèmes limites et marches aléatoires en scènes aléatoires [D5,D6,D7]	46
2.3 Modèles d'agrégation de sources et processus fractionnaires [D13,D17]	57
2.4 Théorèmes limites pour les modèles de boules aléatoires [D8,D12]	65
2.5 Convergence des séries de LePage dans l'espace de Skohorod [D15]	71
2.6 Perspectives	73
Bibliographie	75
3 Quelques études asymptotiques de modèles aléatoires issus de la géométrie, de la biologie ...	81
3.1 Introduction	82
3.2 Un algorithme stochastique de type gradient pour le calcul de p -moyenne sur une variété [D16]	83
3.3 Diversité phénotypique et croissance d'une population multitype en environnement aléatoire [D11]	88

3.4 Travaux de thèse [D1,D2,D3]	93
3.5 Autres travaux [D4,D9,D10]	94
Bibliographie	95

Avant-propos

Ce mémoire présente une vue d'ensemble de mes activités de recherche depuis mon doctorat en 2005. C'est pour moi l'occasion de retracer, dans un esprit de synthèse, mon parcours scientifique. Pour rendre compte de la grande variété des modèles abordés, le mémoire s'organise autour de trois thèmes :

- la théorie spatiale des extrêmes et les propriétés des processus max-stables ;
- les théorèmes limites en présence de dépendance à longue portée et de distributions à queues lourdes ;
- l'étude asymptotique de modèles aléatoires issus de la géométrie, de la biologie . . .

Ces trois thèmes sont traités de manière indépendante dans trois chapitres distincts. Bien sûr, l'organisation en chapitre indépendants est parfois artificielle et certaines notions transversales comme les théorèmes limites ou la théorie des variations régulières relient les différentes parties. L'ordre proposé est ici anti-chronologique : le premier chapitre rend compte de mes travaux les plus récents, le second chapitre regroupe des travaux commencés pendant ma dernière année à Lyon puis poursuivi pendant mes premières années à Poitiers, enfin le troisième chapitre est plus proche de mes travaux de thèse.

L'accent a été mis de manière volontaire sur la thématique « Théorie spatiale des extrêmes et propriétés des processus max-stables » qui donne son nom au mémoire. Plusieurs raisons à ce choix. Ce thème a été développé dans le cadre de la thèse de Frédéric Eyi-Minko, commencée sous ma direction en septembre 2010 et il me semble intéressant, dans le cadre de l'habilitation à diriger des recherches, de mettre en valeur ce travail d'encadrement doctoral. De plus, c'est principalement dans cette direction que s'orientent mes travaux actuels. La théorie des valeurs extrêmes marie probabilités, statistiques et applications dans une communauté dynamique. Mes projets et perspective de recherche s'inscrivent dans ce cadre.

Le premier chapitre autour de la théorie spatiale des extrêmes et des propriétés des processus max-stables est le plus détaillé. Après une introduction rappelant les enjeux et le contexte scientifique, je reprends les résultats des différents articles :

- les lois conditionnelles des processus max-infiniment divisibles ;
- la simulation conditionnelle des processus de Brown-Resnick ;
- les propriétés de mélange fort des processus max-infiniment divisibles ;
- les processus ponctuels et les extrêmes de champs aléatoires i.i.d. ;
- les modèles de shot-noise extrémaux.

En conclusion au chapitre, je présente une description des mes projets de recherche.

Le deuxième chapitre autour des théorèmes limites, de la dépendance à longue portée et des variations régulières est organisé autour de quatre axes, chacun correspondant à un ou plusieurs articles :

- les modèles de marche aléatoires en scène aléatoire ;
- les modèles d'agrégation de sources et les processus fractionnaires ;
- les modèles de boules aléatoires et leurs asymptotiques ;
- la convergence des séries de LePage dans l'espace de Skohorod.

Le chapitre se termine par quelques perspectives de recherche.

Le troisième chapitre autour de l'étude asymptotique de modèles aléatoires issus de la biologie ou la géométrie est plus hétéroclite. J'y détaille deux résultats récents :

- un algorithme stochastique de type gradient pour le calcul de p -moyenne sur une variété ;
- la croissance d'une population multitype en environnement aléatoire et la diversité phénotypique

Par souci d'exhaustivité, les travaux plus anciens, incluant les travaux de thèse, sont brièvement résumés.

Liste de publications

Voici la liste de mes publications. Toutes les références du type [Dxx] dans le mémoire font référence à cette liste.

- [D1] C.Dombry, A probabilistic model for DNA denaturation, *Journal of Statistical Physics*, Vol 120, No 3-4, 695-719, 2005.
- [D2] C.Dombry, N.Guillot-Plantard, B.Pinçon et R.Schott, Data structures with dynamical random transitions, *Random Structures and Algorithms*, Vol 28, No 4, 403-426, 2006.
- [D3] C.Dombry, A weighted random walk model - Application to a genetic algorithm, *Advances in Applied Probability*, Vol 39, No 2, 550-568, 2007.
- [D4] C.Dombry, N.Guillot-Plantard, The Curie-Weiss model with quasiperiodic external random field, *Markov Processes and Related Fields*, Vol 15, No 1, 1-30, 2009.
- [D5] C.Dombry, N.Guillot-Plantard, Discrete approximation of a stable selfsimilar stationary increments process, *Bernoulli*, Vol 15, No 1, 195-222, 2009.
- [D6] S.Cohen, C.Dombry, Convergence of dependent walks in a random scenery to fBm-local time fractionnal stable motions, *J. Math. Kyoto Univ.*, Vol 49, No 2, 267-286, 2009.
- [D7] C.Dombry, N.Guillot-Plantard, A fonctionnal approach for random walks in random sceneries, *Electronic Journal of Probability*, Vol 14, 1495-1512, 2009.
- [D8] J.-C. Breton, C.Dombry, Rescaled weighted random balls models and stable self-similar random fields, *Stochastic Process. Appl.*, Vol 119, No 10, 3633-3652, 2009.
- [D9] C.Dombry, C.Mazza, Some remarks on Betti numbers of random polygon spaces, *Random Structures Algorithms*, Vol. 37, No 1, 67-84, 2010.
- [D10] A.Anagnostopoulos, C.Dombry, N.Guillot-Plantard, I.Kontoyiannis, E.Upfal, Stochastic analysis of the k-server problem on the circle, In 21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10), *Discrete Math. Theor. Comput. Sci. Proc.*, AM, pages 21-34. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2010.
- [D11] C.Dombry, C.Mazza, V.Bansaye, Phenotypic diversity and population growth in a fluctuating environment, *Adv. in Appl. Probab.*, Vol. 43, No 2, 375-398, 2011.

- [D12] J.-C.Breton, C.Dombry, Functional macroscopic behavior of weighted random ball model, *ALEA Lat. Am. J. Probab. Math. Stat.*, Vol. 8, 177-196, 2011.
- [D13] C.Dombry, I.Kaj, The on-off network traffic model under intermediate scaling, *Queueing Systems*, Vol. 69, No. 1, 29-44, 2011.
- [D14] C.Dombry, Extremal shot noises, heavy tails and max-stable random fields, *Extremes*, Vol. 15, No. 2, 129–158, 2012.
- [D15] Y.Davydov, C.Dombry, On the convergence of LePage series in Skohorod space, *Statistics & Probability Letters*, Vol. 82, No. 1, 145–150, 2012.
- [D16] M.Arnaudon, C.Dombry, Y.Le, A.Phan, Stochastic algorithms for computing means of probability measures, *Stochastic Processes and their Applications*, Vol. 122, No. 4, 1437–1455, 2012.
- [D17] C.Dombry, I.Kaj, Moment measures of heavy tailed renewal point processes : asymptotics and applications, to appear in *ESAIM P&S*.
- [D18] C.Dombry, F.Eyi-Minko, Extremes of independent stochastic processes : a point process approach, Preprint hal-00627368.
- [D19] C.Dombry, F.Eyi-Minko, Regular conditional distributions of max infinitely divisible processes, Preprint hal-00627375.
- [D20] C.Dombry, F.Eyi-Minko, M.Ribatet, Conditional simulations of Brown-Resnick processes, to appear in *Biometrika*.
- [D21] C.Dombry, F.Eyi-Minko, Strong mixing properties of max-infinitely divisible random fields, to appear in *Stochastic Processes and their Applications*.

Chapitre 1

Théorie spatiale des extrêmes et propriétés des processus max-stables

Ce chapitre s'articule autour des articles [D14], [D18], [D19], [D20] et [D21]. Après une introduction à la thématique « Théorie spatiale des extrêmes et propriétés des processus max-stables », nous consacrons une section à chaque article avant de conclure le chapitre sur les perspectives de recherche relatives à la thématique.

1.1 Introduction

Motivations : Pourquoi la théorie des valeurs extrêmes ?

La théorie des valeurs extrêmes est un domaine de recherche très dynamique, à l'interface des probabilités, des statistiques et de leurs applications. Le problème de base est de modéliser et/ou prévoir l'occurrence d'événements de types extrêmes dans un système aléatoire donné. Des exemples d'applications classiques sont la modélisation :

- de la hauteur d'eau d'une rivière et des phénomènes de crues,
- des températures maximales lors de vagues de chaleur,
- de la force des vents lors d'une tempête,
- des pics de pollution . . .

Nous avons choisi à dessein des applications relevant des sciences de l'environnement. Elles ont joué un rôle historique important, notamment l'hydrologie au Pays-Bas, avec le problème du calibrage des digues permettant de protéger les terres situées au dessous du niveau de la mer d'inondations maritimes lors de tempêtes. Une question posée aux ingénieurs est la suivante : comment calibrer une digue afin que la probabilité d'une inondation pour une année donnée soit inférieure à 0,001 ? D'autres domaines d'applications classiques sont la finance et l'assurance, la fiabilité, les sciences de l'ingénieur . . .

D'un point de vue mathématique, il s'agit de modéliser le phénomène aléatoire, en s'intéressant principalement non pas au « corps » de la distribution, mais à la « queue » de la distribution. C'est à dire qu'on modélise le phénomène uniquement d'un point de vue asymptotique, dans le régime des valeurs extrêmes. Cette idée a donné lieu à de fructueux développements mathématiques, notamment :

- la notion de variation régulière qui joue un rôle important. En supposant que les distributions ont une certaine régularité dans le domaine des valeurs extrêmes, elle permet d'extrapoler le comportement des extrêmes d'un échantillon de taille modérée. Ceci est particulièrement important dans la pratique car les extrêmes sont par essence même peu fréquemment observés : comment prédire par exemple l'ampleur d'un phénomène arrivant tous les 100 ans à partir d'observation sur les 10 dernières années ? La notion de variation régulière permet une telle extrapolation.
- les théorèmes limites pour les maxima renormalisés, la caractérisation des lois GEV (Generalized Extreme Value distribution) et de leurs domaines d'attraction et les notions de max-stabilité et max-infinité-divisibilité ;
- la modélisation des excès au delà d'un seuil élevé, la méthode POT (Peak Over Threshold) et les lois GPD (Generalized Pareto Distribution) ;
- les liens étroits avec la théorie des processus ponctuels (loi fonctionnelle des petits nombres) et les processus ponctuels de Poisson.

Ces notions ont d'abord été développées pour la théorie univariée des extrêmes et les échantillons indépendants, puis étendues au cadre multivarié et aux échantillons dépendants. Aujourd'hui, une direction de recherche particulièrement active est celle de la théorie spatiale des extrêmes, où l'on quitte le contexte fini-dimensionnel pour se placer dans un cadre

fonctionnel et où les processus et champs aléatoires occupent une place prépondérante. Cette composante spatiale est présente naturellement dans les applications aux sciences de l'environnement citées plus haut, et un enjeu important est celui de la modélisation de la dépendance spatiale des extrêmes et des problèmes d'inférence relatifs. Mes recherches concernent la théorie spatiale des extrêmes et les propriétés des processus max-stables et max-infiniment divisibles.

Quelques repères historiques et bibliographiques

Ce bref aperçu forcément incomplet reflète mes propres intérêts et ne rend pas justice à de nombreux auteurs. Avant de commencer, mentionnons plusieurs monographies synthétisant les développements de la théorie dans différentes directions et dans lesquelles se trouvent des bibliographies plus complètes. Le livre de Resnick [55] est un classique et expose les liens entre variations régulières, processus ponctuels et extrêmes. Falk, Huesler and Reiss [32] développent les lois fonctionnelles des petits nombres en lien avec les extrêmes et les applications statistiques. Embrechts, Klüppelberg et Mikosch [31] proposent un exposé de la théorie des valeurs extrêmes en vue des applications en assurance et finance. Beirlant, Goegebeur, Teugels, et Segers [3] traitent plus spécifiquement des problèmes statistiques et des applications liées aux extrêmes. Enfin, le livre de de Haan et Ferreira [24] est un ouvrage de référence pour la théorie des lois max-stables, fini-dimensionnelle ou en dimension infinie.

La théorie des valeurs extrêmes a d'abord été développée dans un contexte univarié, *i.e.* pour des observations réelles unidimensionnelles, et pour les échantillons indépendants. La théorie probabiliste concernant la caractérisation des lois max-stables et de leurs domaines d'attraction débute en 1927 par les travaux de Fréchet [35], Fisher et Tippett [33], puis von Mises [67], Gnedenko [38], Gumbel [39] ... Les aspects statistiques de la théorie univariée ont été étudiés plus tardivement à partir de 1975 par Pickands [53] et Hill [42].

Le cas des extrêmes d'échantillons dépendants (*e.g.* suite stationnaires de variables aléatoires réelles) a été considéré dans les années 50 et a une histoire riche. Différentes conditions de dépendance ont été étudiées. Les livres de Galambos [36] et de Leadbetter, Lindgren et Rootzén [51] offrent une synthèse de ces développements.

La théorie multivariée des extrêmes débute dans les années 60 avec les travaux de Gefroy [39], Sibuya [62] et Tiago de Oliveira [65] dans un cadre bivarié. Mais elle ne prend son essor qu'à la fin des années 70 et est aujourd'hui encore un thème de recherche dynamique. La caractérisation des lois max-infiniment divisibles en dimension finie est due à Balkema et Resnick [9], le cas max-stable à de Haan [22] et Deheuvels [28, 29]. La caractérisation des extrêmes multivariés est également abordée par Heffernan et Tawn [41] et Heffernan et Resnick [40]. Les applications statistiques dans un cadre multivarié sont considérées entre autre par Coles et Tawn [11], Schlather et Tawn [60, 61] avec l'étude des propriétés des coefficients extrémaux, ou encore Fougères et Soulier [1, 34] avec des modèles de dépendance en présence de conditionnement extrêmes.

Le contexte fonctionnel et la théorie des processus max-stables connaissent un développement important du milieu des années 80 jusqu'à aujourd'hui. La structure des processus max-stables a été donnée par de Haan [23] et le cas max-infiniment divisible par Giné, Hahn et Vatan [37]. Citons également les travaux de de Haan et Pickands [26], Resnick et Roy [56]. Différents modèles de champs max-stables ont été introduits par Smith [63], Schlather [59], Kabluchko, Schlather et de Haan [49] avec les fameux processus de Brown-Resnick. Les contributions récentes sur les propriétés des processus max-stables incluent les travaux de Kabluchko et Schlather [48] sur les questions d'ergodicité et de mélange, de Kabluchko [47] en lien avec les extrêmes de processus gaussien ou encore de Stoev [64] et Kabluchko [45] sur les représentations des processus max-stables et le parallèle avec les processus α -stables ... Concernant les applications et les aspects statistiques dans un cadre spatial, voir Coles et Tawn [12], de Haan et Pereira [25], Buishand, de Haan et Zhou [10] ou encore Cooley, Nychka et Naveau [14] ... Davison, Padoan et Ribatet [19] présentent une revue récente des méthodes statistiques pour les extrêmes spatiaux et contient de nombreuses références bibliographiques sur cet aspect de la théorie.

Contributions au domaine

Nos travaux portent essentiellement sur les modèles de champs aléatoires max-stables ou max-infiniment divisibles (max-i.d.) et leurs propriétés. La représentation de ces champs en terme de processus ponctuels de Poisson y joue en rôle central. D'une manière générale, le point de vue fonctionnel et processus ponctuel est au coeur de notre approche. Rappelons qu'une partie importante du travail a été effectuée dans le cadre de la thèse de Frédéric Eyi-Minko.

Nos contributions sont les suivantes :

- la détermination des lois conditionnelles des champs aléatoires max-i.d. sachant des observations en différents points [D19] ;
- la simulation conditionnelle des processus de Brown-Resnick [D20] ;
- une estimation des coefficients de β -mélange des processus max-i.d. avec applications au théorème central limite et à la normalité asymptotique d'estimateurs du coefficient extrémal [D21] ;
- une approche « processus ponctuels sur espace fonctionnel » pour les extrêmes de champs aléatoires indépendants [D18] ;
- une étude des « shot-noise extrémaux » et de leurs propriétés [D14].

Une section est consacrée à l'exposé de chacun de ces travaux puis le chapitre se termine sur quelques problèmes ouverts et perspectives de recherche pour les années à venir.

1.2 Lois conditionnelles des processus max-infiniment divisibles [D19]

Problématique

Nous nous intéressons au problème de *prédiction* pour les processus max-stables. De manière très générale, dans un contexte de statistique spatiale, on est très souvent confronté à des observations partielles, c'est-à-dire que le champ aléatoire $\eta = (\eta(t))_{t \in T}$ n'est observé qu'en un nombre fini de stations s_1, \dots, s_k . Il s'agit alors d'utiliser les observations partielles

$$\eta(s_i) = y_i, \quad i = 1, \dots, k, \quad (1.1)$$

afin de prédire au mieux le comportement du champ η hors des stations observées. Autrement dit, comment prendre en compte de manière optimale l'information (1.1) pour inférer le comportement du champ η hors des sites d'observations? La réponse mathématique à cette question très naturelle réside dans la détermination des lois conditionnelles du champ η conditionnellement aux observations (1.1). Le conditionnement étant singulier (*i.e.* de probabilité nulle), on s'intéressera aux lois conditionnelles *régulières* définies comme des noyaux de transition. Nous n'insisterons pas trop sur cet aspect technique ici et utiliserons volontiers des notations semi-heuristiques pour introduire les lois conditionnelles. Le formalisme plus rigoureux est introduit dans [D19].

Dans le contexte très classique des champs gaussiens, il est bien connu que les lois conditionnelles restent gaussiennes et qu'elles sont calculables explicitement en déterminant les espérances et covariance conditionnelles par des calculs matriciels. Les calculs reposent sur des méthodes hilbertiennes et il s'agit de déterminer la projection orthogonale de la variable η sur l'espace des variables L^2 mesurables par rapport à la tribu engendrée par les variables conditionnantes $\eta(s_1), \dots, \eta(s_k)$. Ces propriétés simples sont à la base des techniques courantes de krigeage en géostatistique.

Le problème correspondant en théorie des valeurs extrêmes, c'est à dire lorsque η est un champ max-stable s'avère difficile et est resté ouvert pendant longtemps. Les premières contributions sont dues à Davis et Resnick [17, 18] qui imitent les techniques de projection gaussiennes et définissent les prédicteurs comme la projection sur l'espace de Fréchet engendrée par les variables conditionnantes $\eta(s_1), \dots, \eta(s_k)$ au sens d'une métrique L^1 basée sur la représentation en termes de fonctions spectrales. Mais, contrairement à ce qui se passe dans le contexte gaussien, ces prédicteurs obtenus comme projections ne permettent pas d'obtenir les vraies lois conditionnelles. Dans un travail récent, Wang et Stoev [68] apportent une contribution majeure au problème des lois conditionnelles des processus max-stables en déterminant des formules explicites pour les modèles max-linéaires, *i.e.* de la forme

$$\eta(t) = \bigvee_{j=1}^q Z_j f_j(t), \quad t \in T, \quad (1.2)$$

avec f_1, \dots, f_q des fonctions déterministes et Z_1, \dots, Z_q des variables aléatoires indépendantes et identiquement distribuées (v.a.i.i.d.) de loi Fréchet unitaire standard. Les auteurs

déterminent en fait la loi de $\mathbf{Z} = (Z_1, \dots, Z_q)$ sachant les observations (1.1). Le travail repose sur la notion importante de *hitting scenario* défini comme l'ensemble aléatoire des indices $j \in \llbracket 1, q \rrbracket$ tels que $\eta(s_i) = Z_j f_j(s_i)$ pour au moins un indice $i \in \llbracket 1, k \rrbracket$ - on note ici $\llbracket 1, n \rrbracket = \{1, \dots, n\}$. La loi conditionnelle de \mathbf{Z} sachant les observations (1.1) est explicitée comme un mélange sur tous les hitting scenarios admissibles et de rang minimal. Une des limites essentielles de ce travail est qu'il utilise de manière essentielle la structure (1.2) des processus max-linéaires et ne peut donc s'étendre aux processus max-stables génériques.

Notre résultat essentiel consiste en la résolution du problème des lois conditionnelles pour un conditionnement du type (1.1) pour la classe complète des processus max-stables et même celle des processus max-infiniment divisibles (max-i.d.)

Représentation spectrale des processus max-i.d.

L'élément clé de notre approche est la représentation due à Giné Hahn et Vatan [37] des processus max-i.d. en terme de processus ponctuel de Poisson. On suppose que l'espace des paramètres T est un espace métrique compact, que η est à trajectoires continues et tel que

$$\text{ess inf} \eta(t) \equiv 0, \quad t \in T.$$

Il est alors possible de définir une mesure de Borel μ sur $\mathcal{C}_0 = \mathcal{C}(T, [0, +\infty)) \setminus \{0\}$, appelée mesure spectrale, telle que

$$(\eta(t))_{t \in T} \stackrel{\mathcal{L}}{=} (\vee_{\phi \in \Phi} \phi(t))_{t \in T}$$

où Φ est un processus ponctuel de Poisson sur \mathcal{C}_0 d'intensité μ et $\stackrel{\mathcal{L}}{=}$ désigne l'égalité en loi des processus. La mesure spectrale μ vérifie alors

$$\forall \varepsilon > 0, \quad \mu(\{f \in \mathcal{C}_0; \sup_T f > \varepsilon\}) < \infty \quad (1.3)$$

et

$$\mu(\{f \in \mathcal{C}_0; \exists i \in \llbracket 1, p \rrbracket f(t_i) > z_i\}) = -\log \mathbb{P}[\eta(t_1) \leq z_1, \dots, \eta(t_p) \leq z_p]$$

avec $p \geq 1$, $t_1, \dots, t_p \in T$ et $z_1, \dots, z_p > 0$. Cette dernière relation permet d'exprimer les fonctions de répartition multivariée de η en terme de la mesure spectrale μ . Réciproquement, si μ vérifie la condition (1.3), alors il est possible de construire un processus max-i.d. continu sur T de mesure spectrale μ . On supposera dans la suite que μ vérifie (1.3) et que η est donné par

$$\eta(t) = \vee_{\phi \in \Phi} \phi(t), \quad t \in T. \quad (1.4)$$

L'idée essentielle sous-jacente à notre approche est d'étudier le processus ponctuel Φ conditionnellement aux observations (1.1). La loi conditionnelle de η se déduit ensuite de celle de Φ grâce à l'équation (1.4).

La figure 1.1 peut servir de support à l'intuition pour cette représentation en terme de processus ponctuel de Poisson. Il s'agit de simulation d'un processus max-stable de type

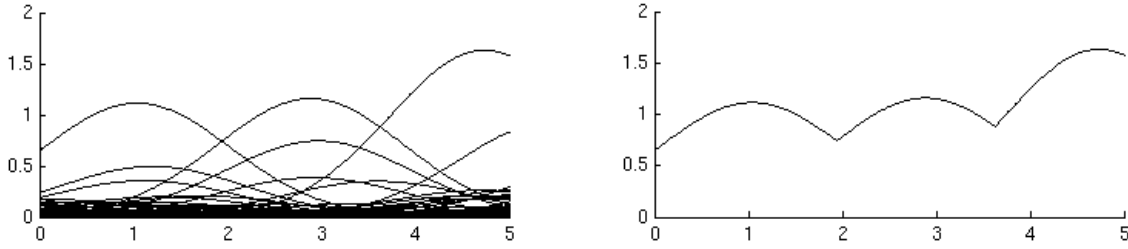


FIGURE 1.1 – Une réalisation du processus ponctuel de Poisson Φ (à gauche) et la réalisation du processus max-stable correspondante (à droite) pour le modèle de maximum mobile.

maximum mobile et du processus ponctuel sous-jacent. On note que la fonction nulle est l'unique point d'accumulation, ce qui traduit la condition (1.3).

Processus ponctuels extrémaux et sous-extrémaux

L'étude du processus ponctuel Φ conditionné aux observations (1.1) amène naturellement à définir la notion de fonctions extrémales et de processus ponctuel extrémal et sous-extrémal.

Définition 1.2.1 *Pour $S \subset T$ fermé, on écrit*

$$\Phi = \Phi_S^+ \cup \Phi_S^-,$$

avec Φ_S^+ processus ponctuel S -extrémal

$$\Phi_S^+ = \{\phi \in \Phi; \exists s \in S, \phi(s) = \eta(s)\},$$

et Φ_S^- processus ponctuel S -sous-extrémal

$$\Phi_S^- = \{\phi \in \Phi; \forall s \in S, \phi(s) < \eta(s)\}.$$

Autrement dit Φ_S^+ est l'ensemble des fonctions $\phi \in \Phi$ qui contribuent au maximum η sur S , alors que Φ_S^- est l'ensemble des fonctions $\phi \in \Phi$ n'ayant pas de contribution au maximum η sur S .

La figure 1.2 représente cette décomposition sur une réalisation du processus ponctuel associé au processus de type maximum mobile.

Il est important de noter que Φ_S^+ et Φ_S^- vérifient les propriétés suivantes

$$\forall \phi \in \Phi_S^-, \phi <_S \eta \quad \text{et} \quad \forall \phi \in \Phi_S^+, \phi \not<_S \eta,$$

et que la restriction de η à S ne dépend que de Φ_S^+ car

$$\eta =_S \bigvee_{\phi \in \Phi_S^+} \phi.$$

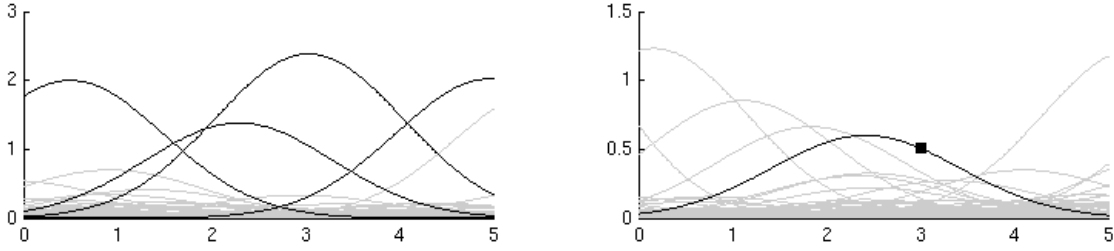


FIGURE 1.2 – Décomposition du processus ponctuel de Poisson Φ en sa partie S -extrémale Φ_S^+ (en noir) et sa partie S -sous-extrémale Φ_S^- (en gris). A gauche : $S = [0, 5]$. A droite : $S = \{3\}$ représenté par le carré noir.

On utilise ici les notations suivantes :

$$\begin{aligned} f <_S g & \text{ si et seulement si } \forall s \in S, f(s) < g(s), \\ f \not<_S g & \text{ si et seulement si } \exists s \in S, f(s) \geq g(s), \\ f =_S g & \text{ si et seulement si } \forall s \in S, f(s) = g(s). \end{aligned}$$

Un premier résultat important concerne la loi jointe de (Φ_S^+, Φ_S^-) . On note $M_p(\mathcal{C}_0)$ l'ensemble des mesures ponctuelles M localement finies sur \mathcal{C}_0 muni de la tribu engendrée par les applications

$$A \mapsto M(A), \quad A \in \mathcal{C}_0 \text{ Borel.}$$

Théorème 1.2.2 *La loi de (Φ_S^-, Φ_S^+) est donnée par*

$$\begin{aligned} & \mathbb{P}[\Phi_S^+ \in A, \Phi_S^- \in B, \Phi_S^+(\mathcal{C}_0) = k] \\ &= \frac{1}{k!} \int_{\mathcal{C}_0^k} \mathbf{1}_{\{\sum_{i=1}^k \delta_{f_i} \in A\}} \mathbf{1}_{\{\forall i \in [1, k], f_i \not<_S \vee_{j \neq i} f_j\}} \\ & \quad \mathbb{P}[\{\Phi \in B\} \cap \{\forall \phi \in \Phi, \phi <_S \vee_{i=1}^k f_i\}] \mu(df_1) \cdots \mu(df_k) \end{aligned}$$

pour tout $k \geq 1$ et $A, B \subset M_p(\mathcal{C}_0)$ mesurable.

La preuve de ce théorème repose essentiellement sur la formule de Slyvniak pour les processus ponctuel de Poisson. Il faut noter que le théorème caractérise la loi jointe de (Φ_S^+, Φ_S^-) lorsque Φ_S^+ est fini presque sûrement. On montre que cette dernière propriété est vérifiée si et seulement si $\inf_S \eta > 0$ ce qui est vrai par exemple dans le cas max-stable.

Fonctions extrémales et hitting scénario

On spécialise ici les résultats précédents lorsque $S = \{s_1, \dots, s_k\}$ est un ensemble fini et on commence par le cas d'un singleton. Lorsque $S = \{s\}$, on appelle fonction s -extrémale toute fonction $\phi \in \Phi$ vérifiant $\phi(s) = \eta(s)$, de sorte que $\Phi_{\{s\}}$ représente l'ensemble des

fonctions s -extrémales. En se basant sur le Théorème 1.2.2, on montre les équivalences suivantes :

- i) $\Phi_{\{s\}}^+(\mathcal{C}_0) = 1$ presque sûrement ;
- ii) la loi de $\eta(s)$ est sans atome ;
- iii) la fonction $y \mapsto \bar{\mu}_s(y) = \mu(\{f \in \mathcal{C}_0; f(s) > y\})$ est continue et telle que $\bar{\mu}_s(0) = +\infty$.

Sous ces conditions, il existe presque sûrement une unique fonction s -extrémale notée ϕ_s^+ de sorte que $\Phi_{\{s\}}^+ = \{\phi_s^+\}$. On montre grâce au Théorème 1.2.2 que la loi de ϕ_s^+ est donnée par

$$\mathbb{P}[\phi_s^+ \in A] = \int_A \exp[-\bar{\mu}_s(f(s))] \mu(df), \quad A \subset \mathcal{C}_0 \text{ mesurable.}$$

On considère ensuite le cas $S = \{s_1, \dots, s_k\}$ fini et on suppose que s_1, \dots, s_k satisfont les conditions i)-iii) ci-dessus. Ainsi les fonctions extrémales $\phi_{s_1}^+, \dots, \phi_{s_k}^+$ sont uniques. Clairement,

$$\Phi_S^+ \subseteq \{\phi_{s_1}^+, \dots, \phi_{s_k}^+\},$$

mais il y a des possibles répétitions dans l'écriture de droite. Pour prendre en compte ces répétitions, on introduit la notion de *hitting scenario* suivante.

Définition 1.2.3 *Il existe (presque sûrement) une unique partition aléatoire $\Theta = \{\theta_1, \dots, \theta_\ell\}$, $\ell \in \llbracket 1, k \rrbracket$, de $S = \{s_1, \dots, s_k\}$ et des fonctions aléatoires $\phi_1^+, \dots, \phi_\ell^+$ telles que :*

- si $s \in \theta_j$, alors $\phi_s^+ = \phi_j^+$.
- $\Phi_S^+ = \{\phi_1^+, \dots, \phi_\ell^+\}$;

Cette définition est illustrée par la figure 1.3 où le hitting scenario est explicité sur 2 exemples.

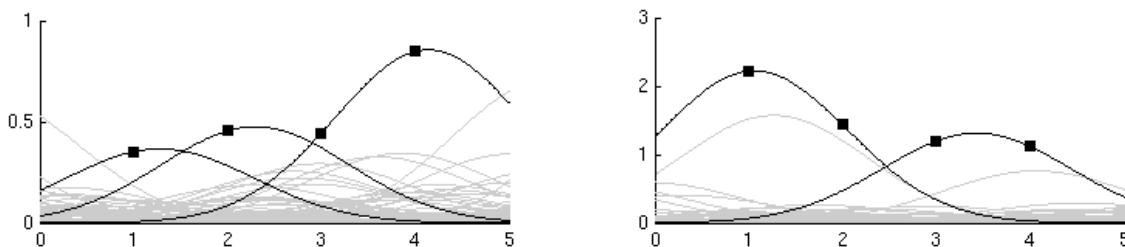


FIGURE 1.3 – Deux réalisations du processus ponctuel de Poisson Φ et du hitting scenario Θ et des fonctions extrémales $\phi_1^+, \dots, \phi_{\ell(\Theta)}^+$ avec $S = \{s_1, s_2, s_3, s_4\}$ représenté par les carrés noirs. A gauche : le hitting scenario est $\Theta = (\{s_1\}, \{s_2\}, \{s_3, s_4\})$, les fonctions extrémales sont $\phi_1^+ = \phi_{s_1}^+$, $\phi_2^+ = \phi_{s_2}^+$ et $\phi_3^+ = \phi_{s_3}^+ = \phi_{s_4}^+$. A droite : le hitting scenario est $\Theta = (\{s_1, s_2\}, \{s_3, s_4\})$, les fonctions extrémales sont $\phi_1^+ = \phi_{s_1}^+ = \phi_{s_2}^+$ et $\phi_2^+ = \phi_{s_3}^+ = \phi_{s_4}^+$.

Les fonctions extrémales et les hitting scenarii sont soumises aux contraintes suivantes :

$$\forall j \in \llbracket 1, \ell \rrbracket, \quad \phi_j^+ >_{\theta_j} \vee_{j' \neq j} \phi_{j'}^+.$$

Le Théorème suivant donne la loi jointe du hitting scénario, des fonctions extrémales, et du processus ponctuel sous-extrémal. On note \mathcal{P}_k l'ensemble des partitions de $\{s_1, \dots, s_k\}$.

Théorème 1.2.4 *Pour $\tau = (\tau_1, \dots, \tau_\ell) \in \mathcal{P}_k$, $A \subset \mathcal{C}_0^\ell$ et $B \subset M_p(\mathcal{C}_0)$ mesurables*

$$\begin{aligned} & \mathbb{P}[\Theta = \tau, (\phi_1^+, \dots, \phi_\ell^+) \in A, \Phi_S^- \in B] \\ = & \int_{\mathcal{C}_0^\ell} 1_{\{\forall j \in \llbracket 1, \ell \rrbracket, f_j >_{\tau_j} \vee_{j' \neq j} f_{j'}\}} 1_{\{(f_1, \dots, f_\ell) \in A\}} \\ & \mathbb{P}[\{\Phi \in B\} \cap \{\forall \phi \in \Phi, \phi <_S \vee_{j=1}^\ell f_j\}] \mu(df_1) \cdots \mu(df_\ell) \end{aligned}$$

Lois conditionnelles

Nous présentons les résultats de manière relativement heuristique. Pour plus de précision et de rigueur, se référer à [D19]. Afin d'expliciter les différentes lois conditionnelles, on introduit une désintégration de la mesure spectrale μ sur \mathcal{C}_0 :

– pour $\mathbf{t} \in T^k$, on définit

$$\mu_{\mathbf{t}}(A) = \mu(\{f \in \mathcal{C}_0; f(\mathbf{t}) \in A\}), \quad A \subset [0, +\infty)^k \text{ Borelien.}$$

– on note $\left\{ P_{\mathbf{t}}(\mathbf{z}, \cdot), \mathbf{z} \in [0, \infty)^k \setminus \{0\} \right\}$ le noyau de probabilité sur \mathcal{C}_0 tel que

$$\int_{\mathcal{C}_0} 1_{\{f(\mathbf{t}) \in A\}} 1_{\{f \in B\}} \mu(df) = \int_A P_{\mathbf{t}}(\mathbf{z}, B) \mu_{\mathbf{t}}(d\mathbf{z})$$

pour tout compact $A \subset [0, +\infty)^k \setminus \{0\}$ et Borélien $B \in \mathcal{C}_0$.

On interprète $\mu_{\mathbf{t}}$ comme une mesure marginale et $P_{\mathbf{t}}$ comme une mesure conditionnelle.

On introduit les notations vectorielles suivantes. La condition (1.1) se réécrit

$$\eta(\mathbf{s}) = \mathbf{y}, \quad \text{avec } \mathbf{s} = (s_1, \dots, s_k) \text{ et } \mathbf{y} = (y_1, \dots, y_k),$$

et l'on note

$$\bar{\mu}_{\mathbf{s}}(\mathbf{y}) = \mu(\{f \in \mathcal{C}_0; \exists i \in \llbracket 1, k \rrbracket, f(s_i) > y_i\}) = -\log \mathbb{P}[\eta(\mathbf{s}) \leq \mathbf{y}].$$

Pour $\tau = (\tau_1, \dots, \tau_\ell) \in \mathcal{P}_k$ et $j = 1, \dots, \ell$, on définit $I_j = \{i: s_i \in \tau_j\}$ et on note $\mathbf{s}_{\tau_j} = (s_i)_{i \in I_j}$, $\mathbf{y}_{\tau_j} = (y_i)_{i \in I_j}$, $\mathbf{s}_{\tau_j^c} = (s_i)_{i \notin I_j}$ et $\mathbf{y}_{\tau_j^c} = (y_i)_{i \notin I_j}$.

Avant de s'intéresser à la loi conditionnelle de η sachant les observations (1.1), nous nous

intéressons à la loi marginale de $\eta(\mathbf{s})$ notée $\nu_{\mathbf{s}}$. Pour toute partition $\tau \in \mathcal{P}_k$, on introduit la mesure

$$\nu_{\mathbf{s}}^{\tau}(C) = \mathbb{P}(\eta(\mathbf{s}) \in C; \Theta = \tau), \quad C \subset [0, +\infty)^k \text{ Borelien.}$$

Clairement la loi $\nu_{\mathbf{s}}$ de $\eta(\mathbf{s})$ vérifie

$$\nu_{\mathbf{s}} = \sum_{\tau \in \mathcal{P}_k} \nu_{\mathbf{s}}^{\tau}.$$

Le résultat suivant permet d'expliciter la loi de $\eta(\mathbf{s})$.

Théorème 1.2.5 *La mesure $\nu_{\mathbf{s}}^{\tau}$ est égale à*

$$\nu_{\mathbf{s}}^{\tau}(d\mathbf{y}) = \exp[-\bar{\mu}_{\mathbf{s}}(\mathbf{y})] \bigotimes_{j=1}^{\ell} \left\{ P_{\mathbf{s}_{\tau_j}}(\mathbf{y}_{\tau_j}, \{f(\mathbf{s}_{\tau_j^c}) < \mathbf{y}_{\tau_j^c}\}) \mu_{\mathbf{s}_{\tau_j}}(d\mathbf{y}_{\tau_j}) \right\}.$$

Après avoir donné la loi marginale de $\eta(\mathbf{s})$, nous explicitons la loi conditionnelle de η sachant $\eta(\mathbf{s}) = \mathbf{y}$. L'idée est de reconstruire la loi conditionnelle du processus ponctuel Φ en trois étapes :

1. construction du hitting scenario Θ sachant $\eta(\mathbf{s}) = \mathbf{y}$;
2. construction des fonctions extrémales $(\phi_1^+, \dots, \phi_{\ell}^+)$ sachant

$$\eta(\mathbf{s}) = \mathbf{y} \quad \text{et} \quad \Theta = \tau;$$

3. construction des fonctions sous-extrémales $\Phi_{\bar{S}}$ sachant

$$\eta(\mathbf{s}) = \mathbf{y}, \quad \Theta = \tau \quad \text{et} \quad (\phi_1^+, \dots, \phi_{\ell}^+).$$

On obtient alors une réalisation du processus ponctuel Φ conditionnel et du processus η conditionnel en posant

$$\Phi = \{\phi_1^+, \dots, \phi_{\ell}^+\} \cup \Phi_{\bar{S}},$$

et

$$\eta = \bigvee \left\{ \bigvee_{j=1}^{\ell} \phi_j^+, \bigvee_{\phi \in \Phi_{\bar{S}}} \phi \right\}.$$

Le résultat suivant explicite les différentes lois conditionnelles intervenant dans l'heuristique précédente.

Théorème 1.2.6 *Soit $\mathbf{s} \in T^k$ et $\mathbf{y} \in (0, +\infty)^k$.*

1. *La loi conditionnelle de Θ sachant $\eta(\mathbf{s}) = \mathbf{y}$ est donnée par*

$$\mathbb{P}(\Theta = \tau \mid \eta(\mathbf{s}) = \mathbf{y}) = \frac{d\nu_{\mathbf{s}}^{\tau}}{d\nu_{\mathbf{s}}}(\mathbf{y}).$$

2. La loi conditionnelle de $(\phi_1^+, \dots, \phi_\ell^+)$ sachant $\eta(\mathbf{s}) = \mathbf{y}$ et $\Theta = \tau$ est

$$\begin{aligned} & \mathbb{P}[(\phi_1^+, \dots, \phi_\ell^+) = (df_1, \dots, df_\ell) \mid \eta(\mathbf{s}) = \mathbf{y}, \Theta = \tau] \\ &= \bigotimes_{j=1}^{\ell} \left(\frac{\mathbf{1}_{\{f_j(\mathbf{s}_{\tau_j^c}) < \mathbf{y}_{\tau_j^c}\}}}{P_{\mathbf{s}_{\tau_j}}(\mathbf{y}_{\tau_j}, \{f(\mathbf{s}_{\tau_j^c}) < \mathbf{y}_{\tau_j^c}\})} P_{\mathbf{s}_{\tau_j}}(\mathbf{y}_{\tau_j}, df_j) \right). \end{aligned}$$

3. La loi conditionnelle de $\Phi_{\mathcal{S}}^-$ sachant $\eta(\mathbf{s}) = \mathbf{y}$, $\Theta = \tau$ et $(\phi_1^+, \dots, \phi_\ell^+) = (f_1, \dots, f_\ell)$ est la loi d'un processus ponctuel de poisson sur \mathcal{C}_0 d'intensité $\mathbf{1}_{\{f(\mathbf{s}) < \mathbf{y}\}} \mu(df)$.

En particulier, le Théorème précédent permet d'affirmer que :

- les différentes fonctions extrémales $(\phi_1^+, \dots, \phi_\ell^+)$ sont indépendantes conditionnellement à $\eta(\mathbf{s}) = \mathbf{y}$ et $\Theta = \tau$;
- la loi conditionnelle de ϕ_j^+ est égale à $P_{\mathbf{s}_{\tau_j}}(\mathbf{y}_{\tau_j}, df_j)$ sachant $f_j(\mathbf{s}_{\tau_j^c}) < \mathbf{y}_{\tau_j^c}$.
- conditionnellement à $\eta(\mathbf{s}) = \mathbf{y}$, les processus ponctuels $\Phi_{\mathcal{S}}^+$ et $\Phi_{\mathcal{S}}^-$ sont indépendants.

De manière intéressante, on peut également expliciter les fonctions de répartition conditionnelles en fonction des lois introduites ci-dessus :

$$\begin{aligned} & \mathbb{P}[\eta(\mathbf{t}) \leq \mathbf{z} \mid \eta(\mathbf{s}) = \mathbf{y}] \\ &= \frac{\mathbb{P}[\eta(\mathbf{t}) \leq \mathbf{z}, \eta(\mathbf{s}) \leq \mathbf{y}]}{\mathbb{P}[\eta(\mathbf{s}) \leq \mathbf{y}]} \sum_{\tau \in \mathcal{P}_k} \left[\left(\prod_{j=1}^{\ell} \mathbb{P}[\phi_j^+(\mathbf{t}) \leq \mathbf{z} \mid \eta(\mathbf{s}) = \mathbf{y}, \Theta = \tau] \right) \mathbb{P}[\Theta = \tau \mid \eta(\mathbf{s}) = \mathbf{y}] \right]. \end{aligned}$$

Prolongements

Le travail [D19] se poursuit en explicitant les formules du Théorème 1.2.6 pour différentes classes d'exemples :

- la classe des processus max-stables est importante dans la pratique et certains calculs peuvent être poussés en utilisant le fait que la mesure spectrale μ se factorise en une partie radiale et une partie polaire ;
- la classe des modèles réguliers permet de pousser les calculs en utilisant le fait que les lois fini-dimensionnelles de μ admettent des densités par rapport aux mesures de Lebesgue en dimension correspondante ;
- la classe des processus de type Brown-Resnick est très adaptée aux calculs de lois conditionnelles, et la désintégration de μ est alors simple grâce au fait que les processus gaussiens conditionnés demeurent gaussiens (*cf.* la section suivante) ;
- la classe des processus max-linéaires peut être traitée afin de comparer les expressions à celles obtenues par Wang et Stoev [68] ; la comparaison n'est cependant pas aisée et l'approche de Wang et Stoev simplifie grandement les calculs dans ce cas.

Par souci de concision, nous ne détaillons pas plus avant ces exemples dans le mémoire.

1.3 Simulations conditionnelles des processus de Brown-Resnick [D20]

Introduction

L'objectif de ce travail est de montrer que les formules théoriques relativement complexes pour les lois-conditionnelles des processus max-i.d. décrites dans la section précédente peuvent avoir une application pratique et être utilisées pour réaliser des simulations conditionnelles des processus max-stables. Notons que le problème de la simulation conditionnelle et son implémentation informatique a été également abordée par Wang et Stoev [68] pour les processus max-linéaires. Pour notre part, nous nous limitons au cadre des processus de type Brown-Resnick [9, 49]. La construction de ces processus max-stables est basée sur des processus gaussiens pour lesquels les lois-conditionnelles sont bien connues et simulables.

On commence par rappeler la définition des processus de Brown-Resnick. Soit $(W_i)_{i \geq 1}$ des copies i.i.d d'un processus Gaussien centré continu W sur $T = \mathbb{R}^d$ et de variance $\sigma^2(t)$ et indépendamment, soit $(\xi_i)_{i \geq 1}$ les points d'un processus de Poisson sur $(0, +\infty)$ d'intensité la mesure de Lebesgue. Alors

$$\eta(t) = \bigvee_{i=1}^{\infty} \xi_i^{-1} e^{W_i(t) - \sigma^2(t)/2}, \quad t \in \mathbb{R}^d, \quad (1.5)$$

est un processus max-stable continu sur \mathbb{R}^d . De plus, si W est à accroissements stationnaires, alors η est stationnaire sur \mathbb{R}^d et sa loi ne dépend que du semi-variogramme

$$\gamma(h) = \frac{1}{2} \mathbb{E}[(W(t+h) - W(t))^2], \quad h \in \mathbb{R}^d.$$

Le premier processus de ce type a été introduit par Brown et Resnick [9] dans le cas où W est un mouvement Brownien standard en dimension $d = 1$. Plus récemment, Kabluchko, Schlather et de Haan [49] ont généralisé cette construction en dimension supérieure et étudié les propriétés de ces processus baptisés de Brown-Resnick. Les liens avec les maxima de processus gaussiens indépendants sont étudiés par Kabluchko [47]. Dernièrement, Oesting, Kabluchko et Schlather [52] ont proposé des méthodes de simulation efficaces pour le processus « historique » construit à partir de mouvements Browniens standards ou fractionnaires.

Notre objectif ici est d'explicitier les lois conditionnelles données par le Théorème 1.2.6 pour les processus de Brown-Resnick et d'implémenter les simulations conditionnelles en R et C. Une difficulté majeure est l'explosion combinatoire de l'espace des partitions \mathcal{P}_k de l'ensemble des points conditionnants $\{s_1, \dots, s_k\}$ lorsque k devient grand et nous proposons une technique du type *Monte Carlo Markov Chain* pour contourner cet écueil.

Régularité du modèle et expression des densités

Soit W un processus gaussien W centré, à trajectoires continues sur $T \subset \mathbb{R}^d$ et tel que

$W(0) = 0$. On note γ le semi-variogramme. On définit η , $(W_i)_{i \geq 1}$ et $(\xi_i)_{i \geq 1}$ comme dans (1.5) et on note

$$\phi_i = \xi_i^{-1} e^{W_i(\cdot) - \gamma(\cdot)}.$$

Alors $\Phi = \{\phi_i, i \geq 1\}$ est un processus ponctuel de Poisson sur \mathcal{C}_0 d'intensité

$$\mu(A) = \int_0^\infty \mathbb{P}[r e^{W(\cdot) - \gamma(\cdot)} \in A] r^{-2} dr, \quad A \in \mathcal{C}_0 \text{ mesurable}$$

et clairement l'équation (1.4) est satisfaite.

Une propriété important de ce modèle est sa régularité, au sens où la mesure spectrale μ a des marginales fini-dimensionnelles absolument continues par rapport à la mesure de Lebesgue. En effet, si $\mathbf{t} = (t_1, \dots, t_k) \in T^k$ est tel que la matrice de covariance $\Sigma_{\mathbf{t}} = \mathbb{E}[W(\mathbf{t})W(\mathbf{t})^T]$ est définie positive, alors le vecteur aléatoire gaussien $W(\mathbf{t})$ admet une densité si bien que la mesure d'intensité

$$\mu_{\mathbf{t}} = \int_0^\infty \mathbb{P}[r e^{W(\mathbf{t}) - \gamma(\mathbf{t})} \in A] r^{-2} dr, \quad A \in (0, +\infty)^k \text{ mesurable}$$

associée au processus ponctuel de Poisson $\{\phi_i(\mathbf{t}), i \geq 1\}$ est absolument continue par rapport à la mesure de Lebesgue. Un calcul basé sur les lois gaussiennes montre que

$$\mu_{\mathbf{t}}(d\mathbf{z}) = \lambda_{\mathbf{t}}(\mathbf{z}) d\mathbf{z}$$

et la densité $\lambda_{\mathbf{t}}$ est donnée par

$$\lambda_{\mathbf{t}}(\mathbf{z}) = C_{\mathbf{t}} \exp\left(-\frac{1}{2} \log \mathbf{z}^T Q_{\mathbf{t}} \log \mathbf{z} + L_{\mathbf{t}} \log \mathbf{z}\right) \prod_{i=1}^k z_i^{-1}, \quad \mathbf{z} \in (0, \infty)^k,$$

avec

$$\begin{aligned} Q_{\mathbf{t}} &= \Sigma_{\mathbf{t}}^{-1} - \frac{\Sigma_{\mathbf{t}}^{-1} \mathbf{1}_k \mathbf{1}_k^T \Sigma_{\mathbf{t}}^{-1}}{\mathbf{1}_k^T \Sigma_{\mathbf{t}}^{-1} \mathbf{1}_k}, \\ L_{\mathbf{t}} &= \frac{1}{2} \left(\frac{\mathbf{1}_k^T \Sigma_{\mathbf{t}}^{-1} \sigma_{\mathbf{t}}^2 - 2}{\mathbf{1}_k^T \Sigma_{\mathbf{t}}^{-1} \mathbf{1}_k} \mathbf{1}_k^T - \sigma_{\mathbf{t}}^{2T} \right) \Sigma_{\mathbf{t}}^{-1}, \\ C_{\mathbf{t}} &= (2\pi)^{(1-k)/2} |\Sigma_{\mathbf{t}}|^{-1/2} (\mathbf{1}_k^T \Sigma_{\mathbf{t}}^{-1} \mathbf{1}_k)^{-1/2} \exp \left\{ \frac{1}{2} \frac{(\mathbf{1}_k^T \Sigma_{\mathbf{t}}^{-1} \sigma_{\mathbf{t}}^2 - 1)^2}{\mathbf{1}_k^T \Sigma_{\mathbf{t}}^{-1} \mathbf{1}_k} - \frac{1}{2} \sigma_{\mathbf{t}}^{2T} \Sigma_{\mathbf{t}}^{-1} \sigma_{\mathbf{t}}^2 \right\}, \end{aligned}$$

et $\mathbf{1}_k = (\mathbf{1})_{i=1, \dots, k}$, $\sigma_{\mathbf{t}}^2 = \{\sigma^2(t_i)\}_{i=1, \dots, k}$. Notons que $\lambda_{\mathbf{t}}$ ressemble à une densité log-normale, mais que sa masse totale est infinie car la matrice $Q_{\mathbf{t}}$ est positive mais non définie car le vecteur $\mathbf{1}_k$ annule la forme quadratique associée, *i.e.* $\mathbf{1}_k^T Q_{\mathbf{t}} \mathbf{1}_k = 0$.

Désintégration de la mesure spectrale μ

Dans le Théorème 1.2.6 explicitant les lois conditionnelles, la désintégration de la mesure

spectrale μ intervient de manière essentielle. Nous avons déjà explicité la densité des mesures marginales $\mu_{\mathbf{t}}$, il reste à considérer le noyau $P_{\mathbf{t}}(\mathbf{y}, df)$ sur \mathcal{C}_0 . Le point essentiel est que dans le modèle de Brown-Resnick et sous des hypothèses de régularité, le noyau $P_{\mathbf{t}}(\mathbf{y}, df)$ correspond à la loi d'un processus log-normal dont on peut déterminer explicitement les paramètres par des calculs matriciels. Une manière de le faire est de considérer les lois fini-dimensionnelles du noyau $P_{\mathbf{t}}(\mathbf{y}, df)$, c'est à dire la loi

$$P_{\mathbf{t}}(\mathbf{y}, f(\mathbf{s}) \in d\mathbf{u}), \quad \mathbf{s} \in T^m, \quad \mathbf{u} \in (0, +\infty)^m.$$

Sous l'hypothèse que la matrice de covariance $\Sigma_{(\mathbf{s}, \mathbf{t})}$ soit définie positive, on montre que

$$P_{\mathbf{t}}(\mathbf{y}, f(\mathbf{s}) \in d\mathbf{u}) = \lambda_{\mathbf{s}|\mathbf{t}, \mathbf{z}}(\mathbf{u}) d\mathbf{u} \quad \text{avec} \quad \lambda_{\mathbf{s}|\mathbf{t}, \mathbf{z}}(\mathbf{u}) = \frac{\lambda_{(\mathbf{s}, \mathbf{t})}(\mathbf{u}, \mathbf{z})}{\lambda_{\mathbf{t}}(\mathbf{z})}.$$

De plus, la fonction $\lambda_{\mathbf{s}|\mathbf{t}, \mathbf{z}}(\mathbf{u})$ est une densité log-normale multivariée, *i.e.* de la forme

$$\lambda_{\mathbf{s}|\mathbf{t}, \mathbf{z}}(\mathbf{u}) = (2\pi)^{-m/2} |\Sigma_{\mathbf{s}|\mathbf{t}}|^{-1/2} \exp \left\{ -\frac{1}{2} (\log \mathbf{u} - m_{\mathbf{s}|\mathbf{t}, \mathbf{z}})^T \Sigma_{\mathbf{s}|\mathbf{t}}^{-1} (\log \mathbf{u} - m_{\mathbf{s}|\mathbf{t}, \mathbf{z}}) \right\} \prod_{i=1}^m u_i^{-1},$$

avec $m_{\mathbf{s}|\mathbf{t}, \mathbf{z}} \in \mathbb{R}^m$ et $\Sigma_{\mathbf{s}|\mathbf{t}}$ les moyennes et covariance de la loi normale sous-jacente données par

$$\begin{aligned} \Sigma_{\mathbf{s}|\mathbf{t}}^{-1} &= J_{m,k}^T Q_{(\mathbf{s}, \mathbf{t})} J_{m,k}, \\ m_{\mathbf{s}|\mathbf{t}, \mathbf{z}} &= \left(L_{(\mathbf{s}, \mathbf{t})} J_{m,k} - \log \tilde{\mathbf{z}}^T \tilde{J}_{m,k}^T Q_{(\mathbf{s}, \mathbf{t})} J_{m,k} \right) \Sigma_{\mathbf{s}|\mathbf{t}}, \end{aligned}$$

avec

$$J_{m,k} = \begin{bmatrix} \text{Id}_m \\ \mathbf{0}_{k,m} \end{bmatrix}, \quad \tilde{\mathbf{z}} = \begin{bmatrix} \mathbf{z} \\ \mathbf{1}_m \end{bmatrix}, \quad \text{and} \quad \tilde{J}_{m,k} = \begin{bmatrix} \mathbf{0}_{m,k} \\ \text{Id}_k \end{bmatrix},$$

où Id_k est la matrice identité de taille $k \times k$ et $\mathbf{0}_{m,k}$ la matrice nulle de taille $m \times k$.

Principe de la simulation conditionnelle

La simulation conditionnelle s'effectue alors sur la base du Théorème 1.2.6 et des calculs précédents. On obtient la procédure en trois étapes suivante pour simuler selon la loi conditionnelle de $\eta(\mathbf{t})$ sachant $\eta(\mathbf{s}) = \mathbf{y}$. Les notations sont celles des Théorèmes 1.2.5 et 1.2.6.

Soit $(\mathbf{s}, \mathbf{t}) \in T^{k+m}$ tels que la matrice de covariance $\Sigma_{(\mathbf{s}, \mathbf{t})}$ soit définie positive. On considère la procédure en trois étapes :

1. Tirer une partition $\tau \in \mathcal{P}_k$ de taille $|\tau|$ de loi

$$\mathbb{P}[\theta = \tau \mid \eta(\mathbf{s}) = \mathbf{y}] = \frac{1}{C(\mathbf{s}, \mathbf{y})} \prod_{j=1}^{|\tau|} \lambda_{\mathbf{s}_{\tau_j}}(\mathbf{y}_{\tau_j}) \int_{\{\mathbf{u}_j < \mathbf{y}_{\tau_j^c}\}} \lambda_{\mathbf{s}_{\tau_j^c} | \mathbf{s}_{\tau_j}, \mathbf{y}_{\tau_j}}(\mathbf{u}_j) d\mathbf{u}_j,$$

avec la constante de normalisation

$$C(\mathbf{s}, \mathbf{y}) = \sum_{\tilde{\tau} \in \mathcal{P}_k} \prod_{j=1}^{|\tilde{\tau}|} \lambda_{\mathbf{s}_{\tilde{\tau}_j}}(\mathbf{y}_{\tilde{\tau}_j}) \int_{\{\mathbf{u}_j < \mathbf{y}_{\tilde{\tau}_j^c}\}} \lambda_{\mathbf{s}_{\tilde{\tau}_j^c} | \mathbf{s}_{\tilde{\tau}_j}, \mathbf{y}_{\tilde{\tau}_j}}(\mathbf{u}_j) d\mathbf{u}_j.$$

2. Sachant $\tau = (\tau_1, \dots, \tau_\ell)$, tirer ℓ vecteurs aléatoires indépendants $\phi_1^+(\mathbf{t}), \dots, \phi_\ell^+(\mathbf{t})$ de loi

$$\mathbb{P}[\phi_j^+(\mathbf{t}) \in d\mathbf{v} \mid \eta(\mathbf{s}) = \mathbf{y}, \theta = \tau] = \frac{1}{C_j} \left(\int 1_{\{\mathbf{u} < \mathbf{y}_{\tau_j^c}\}} \lambda_{(\mathbf{t}, \mathbf{s}_{\tau_j^c}) | \mathbf{s}_{\tau_j}, \mathbf{y}_{\tau_j}}(\mathbf{v}, \mathbf{u}) d\mathbf{u} \right) d\mathbf{v}$$

avec

$$C_j = \int 1_{\{\mathbf{u} < \mathbf{y}_{\tau_j^c}\}} \lambda_{(\mathbf{t}, \mathbf{s}_{\tau_j^c}) | \mathbf{s}_{\tau_j}, \mathbf{y}_{\tau_j}}(\mathbf{v}, \mathbf{u}) d\mathbf{u} d\mathbf{v}.$$

Définir $\eta^+(\mathbf{t}) = \max_{j=1, \dots, \ell} \phi_j^+(\mathbf{t})$.

3. Indépendamment des étapes 1 et 2, tirer $\{\xi_i\}_{i \geq 1}$ les points d'un processus ponctuel de Poisson sur $(0, \infty)$ d'intensité Lebesgue et $\{W_i(\cdot)\}_{i \geq 1}$ des copies i.i.d. de $W(\cdot)$. Définir

$$\eta^-(\mathbf{t}) = \max_{i \geq 1} \xi_i^{-1} e^{W_i(\mathbf{t}) - \gamma(\mathbf{t})} 1_{\{\xi_i^{-1} e^{W_i(\mathbf{s}) - \gamma(\mathbf{s})} < \mathbf{y}\}}.$$

Alors, le vecteur aléatoire $\tilde{\eta}(\mathbf{t}) = \max\{\eta^+(\mathbf{t}), \eta^-(\mathbf{t})\}$ suit la loi conditionnelle de $\eta(\mathbf{t})$ sachant $\eta(\mathbf{s}) = \mathbf{y}$.

Notons que dans cette procédure, les différentes fonctions sont explicites. Pour l'étape 1, les intégrales du type $\int_{\{\mathbf{u}_j < \mathbf{y}_{\tau_j^c}\}} \lambda_{\mathbf{s}_{\tau_j^c} | \mathbf{s}_{\tau_j}, \mathbf{y}_{\tau_j}}(\mathbf{u}_j) d\mathbf{u}_j$ ne sont rien d'autres que des fonctions de répartition de vecteurs log-normaux si bien que, pour des petites valeurs de k , le hitting scénario conditionnel se simule aisément comme une loi discrète sur un espace de taille raisonnable. Dans l'étape 2, les lois conditionnelles $\mathbb{P}[\phi_j^+(\mathbf{t}) \in d\mathbf{v} \mid \eta(\mathbf{s}) = \mathbf{y}, \theta = \tau]$ ne sont rien d'autres que des lois log-normales conditionnées et se simulent aisément grâce à une simple technique de rejet. La procédure précédente permet donc une simulation des lois conditionnelles lorsque le nombre k de points conditionnants n'est pas trop élevé.

Techniques MCMC pour la simulation conditionnelle du hitting scénario

Une difficulté subsiste dès que k n'est pas petit, car l'espace \mathcal{P}_k des partitions de $\{s_1, \dots, s_k\}$ devient énorme, si bien que la constante de normalisation $C(\mathbf{s}, \mathbf{y})$ ne peut pas être évaluée par une énumération naïve de tous les cas possibles. Déjà lorsque $k = 10$, \mathcal{P}_k a un cardinal de l'ordre de 10^5 et les capacités de stockage d'un PC standard sont dépassées lorsque $k > 12$. Il paraît alors naturel de se tourner vers des techniques de type *Monte Carlo Markov Chain* et de simuler la loi conditionnelle $\mathbb{P}[\theta \in \cdot \mid \eta(\mathbf{s}) = \mathbf{y}]$ à l'aide d'une chaîne de Markov sur \mathcal{P}_k admettant exactement cette distribution cible comme mesure stationnaire. De nombreux choix sont possibles pour déterminer les transitions d'une telle chaîne de Markov et nous présentons ici une adaptation de l'échantillonneur de Gibbs.

Pour $\tau \in \mathcal{P}_k$, et $j \in \{1, \dots, k\}$, on note τ_{-j} la restriction de τ à l'ensemble $\{s_1, \dots, s_k\} \setminus \{s_j\}$. Dans l'échantillonneur de Gibbs, les transitions de la chaîne de Markov sont basées sur les lois conditionnelles

$$\mathbb{P}[\theta \in \cdot \mid \theta_{-j} = \tau_{-j}], \quad (1.6)$$

où $\theta \in \mathcal{P}_k$ est une partition aléatoire distribuée selon la loi cible et τ est typiquement l'état courant de la chaîne de Markov. Comme le nombre de mises à jours possibles est toujours inférieur à k , l'écueil de l'explosion combinatoire est évité. En effet pour $\tau \in \mathcal{P}_k$ de taille ℓ , le nombre de partitions $\tau^* \in \mathcal{P}_k$ telles que $\tau_{-j}^* = \tau_{-j}$ est donné par

$$b^+ = \begin{cases} \ell & \text{si } \{s_j\} \text{ est une composante de la partition } \tau, \\ \ell + 1 & \text{sinon,} \end{cases}$$

puisque le point s_j peut être réaffecté à n'importe quelle composante de τ_{-j} ou bien à une nouvelle composante. A titre d'exemple, on considère $k = 3$ et $\tau = (\{s_1, s_2\}, \{s_3\})$. Alors les choix possibles pour τ^* tels que $\tau_{-2}^* = \tau_{-2}$ sont

$$(\{s_1, s_2\}, \{s_3\}), \quad (\{s_1\}, \{s_2\}, \{s_3\}), \quad (\{s_1\}, \{s_2, s_3\}),$$

alors qu'il n'y a que deux partitions telles que $\tau_{-3}^* = \tau_{-3}$,

$$(\{s_1, s_2\}, \{s_3\}), \quad (\{s_1, s_2, s_3\}).$$

La distribution (1.6) a de bonnes propriétés : si $\tau^* \in \mathcal{P}_k$ est tel que $\tau_{-j}^* = \tau_{-j}$ alors

$$\mathbb{P}[\theta = \tau^* \mid \theta_{-j} = \tau_{-j}] \propto \frac{\prod_{j=1}^{|\tau^*|} w_{\tau^*,j}}{\prod_{j=1}^{|\tau|} w_{\tau,j}}, \quad (1.7)$$

avec

$$w_{\tau,j} = \lambda_{\mathbf{s}_{\tau_j}}(\mathbf{y}_{\tau_j}) \int_{\{\mathbf{u} < \mathbf{y}_{\tau_j^c}\}} \lambda_{\mathbf{s}_{\tau_j^c} | \mathbf{s}_{\tau_j}, \mathbf{y}_{\tau_j}}(\mathbf{u}) d\mathbf{u}.$$

Dans le membre de droite de (1.7), la constante de proportionnalité est déterminée grâce à une somme sur tous les choix τ^* possibles. Il y a b^+ termes possibles dans cette somme, et comme $b^+ \leq k$, l'explosion combinatoire est évitée. De plus, dans le membre de droite de (1.7), la plupart des facteurs se simplifient si bien qu'il reste au plus quatre facteurs. D'un point de vue computationnel, cela rend l'échantillonneur de Gibbs particulièrement satisfaisant.

Mise en oeuvre

Les différentes procédures présentées ci-dessus ont été implémentées sous R et C et leurs performances évaluées. Nous avons utilisé deux types de modèles pour le travail numérique :

- une classe de modèle très simple en dimension 1, où le processus Gaussien sous-jacent est un mouvement Brownien fractionnaire d'indice de Hurst $H \in (0, 1)$;
- un modèle en dimension 2 basé sur des données réelles de températures en Suisse et étudié par Davison, Padoan et Ribatet [19].

Pour les détails des études numériques, se référer à [D20] où un certain nombre de tests sont effectués permettant de valider les simulations.

Nous présentons ici brièvement quelques simulations conditionnelles pour les processus de Brown-Resnick dirigés par des mouvements Browniens fractionnaires en dimension 1. Pour une meilleure lisibilité des figures, les processus sont standardisés avec des marginales Gumbel. La figure 1.4 présente des simulations (non-conditionnelles) de ces processus pour différentes valeurs de l'indice de Hurst : $H = 1/4$, $H = 1/2$ et $H = 3/4$. Notons que la valeur de H a une forte influence sur la régularité des trajectoires. La figure 1.5 présente des simulations conditionnelles de ces processus, pour les différentes valeurs de H (de gauche à droite) et différents nombres de points conditionnants (de haut en bas). Sont également évalués en chaque point les quantiles conditionnels d'ordre 0,025, 0,5 et 0,975. Pour ce faire, on utilise les quantiles empiriques d'une population de 1000 simulations conditionnelles indépendantes. La médiane conditionnelle permet d'obtenir un prédicteur en chaque point, tandis que les quantiles conditionnels fournissent l'intervalle de confiance de niveau 0,95.

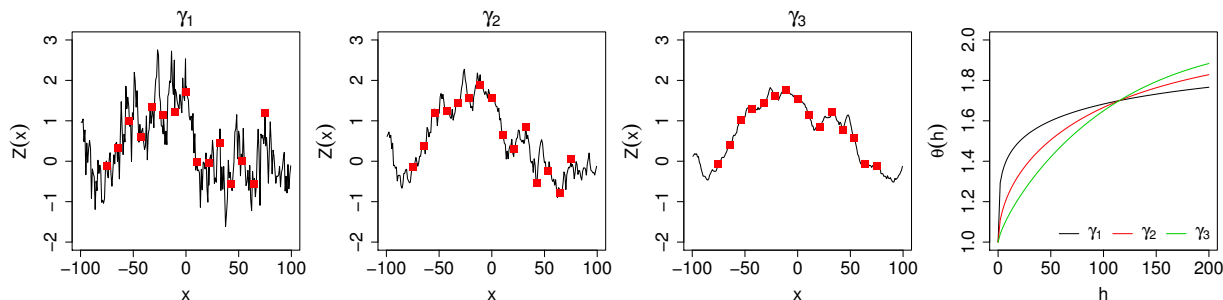


FIGURE 1.4 – Réalisations de processus de Brown Resnick dirigés par des mouvements Browniens fractionnaires d'indice de Hurst $H = 1/4$, $H = 1/2$ et $H = 3/4$ (de gauche à droite). Les semi-variogrammes correspondants γ_1 , γ_2 et γ_3 sont donnés dans le graphe de droite. Les points rouges sont les points conditionnants utilisés dans la figure 1.5.

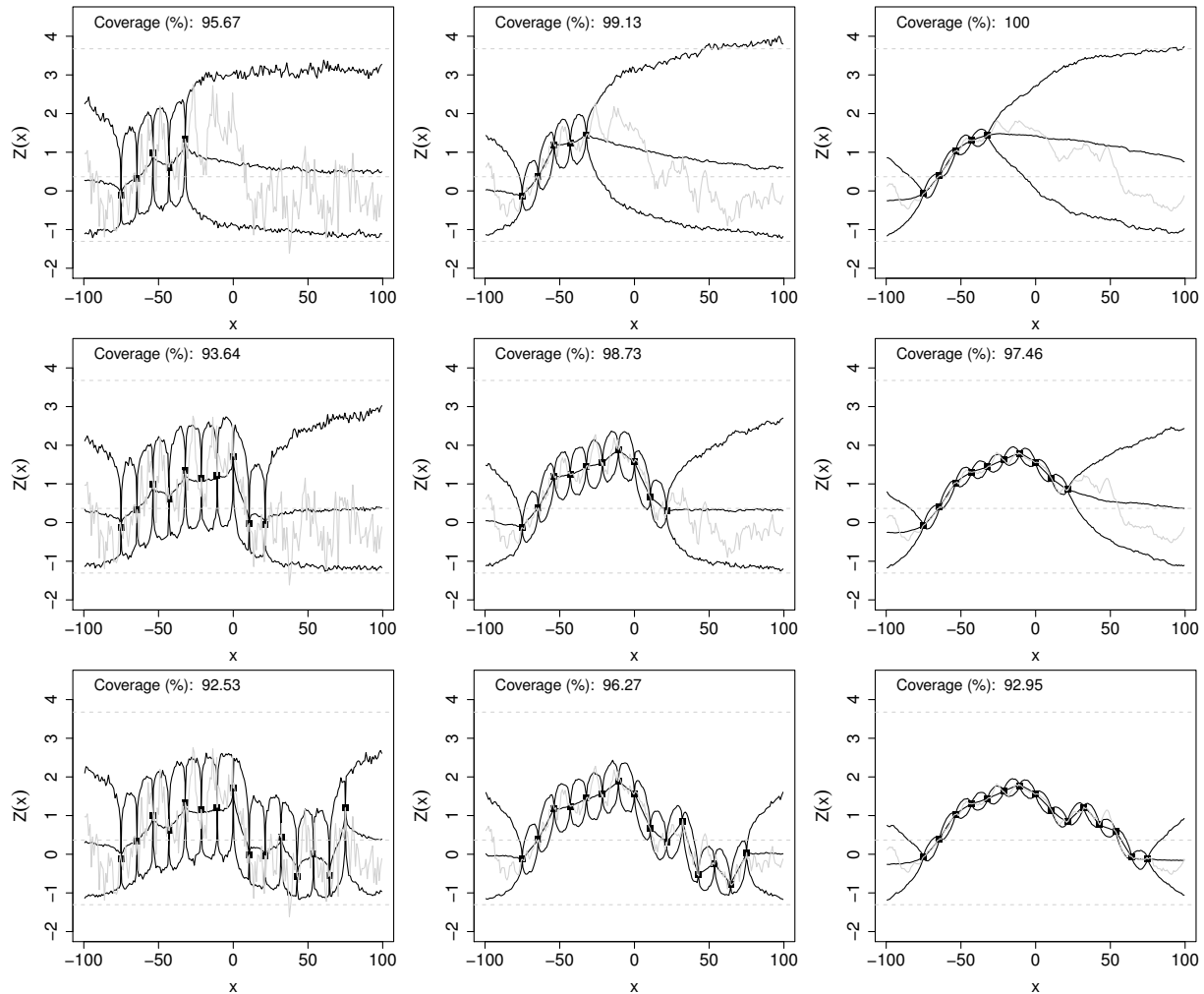


FIGURE 1.5 – Quantiles conditionnels ponctuels estimés sur la base de 1000 simulations conditionnelles de processus de Brown-Resnick à marginales Gumbel et semi-variogrammes γ_1 , γ_2 et γ_3 (de gauche à droite) et avec $k = 5, 10, 15$ points conditionnants (de haut en bas). Les lignes noires continues représentent les quantiles conditionnels d'ordre 0,025, 0,5 et 0,975 et les lignes pointillées grises les quantiles correspondants de la loi de Gumbel. Les carrés indiquent les valeurs conditionnantes.

1.4 Propriétés de mélange fort des processus max-infiniment divisibles [D21]

Contexte

Les propriétés ergodiques des processus max-stables jouent un rôle central dans les questions de loi des grands nombres et de consistance d'estimateur dès que l'on quitte le cadre des échantillons indépendants. Les premiers résultats sur l'ergodicité des processus max-stables sont dus à Weintraub [69] et ont été complétés par Stoev [64] qui a obtenu des conditions nécessaires et suffisantes de mélange pour les processus max-stables en se basant sur leur représentation spectrale. Plus récemment, Kabluchko et Schlather [48] ont élargi ces résultats à la classe des processus max-i.d. et également considéré la question de l'ergodicité. Ils considèrent un processus $\eta = (\eta(t))_{t \in \mathbb{Z}}$ max-i.d. stationnaire sur \mathbb{Z} et définissent la fonction de dépendance

$$\tau_a(h) = \log \frac{\mathbb{P}[\eta(0) \leq a, \eta(h) \leq a]}{\mathbb{P}[\eta(0) \leq a] \mathbb{P}[\eta(h) \leq a]}, \quad a > \text{ess inf } \eta(0), \quad h \in \mathbb{Z}.$$

Le critère est le suivant, avec $\ell = \text{ess inf } \eta(0)$:

- η est mélangeant si et seulement si pour tout $a > \ell$, $\tau_a(n) \rightarrow 0$ quand $n \rightarrow +\infty$;
- η est ergodique si et seulement si pour tout $a > \ell$, $n^{-1} \sum_{h=1}^n \tau_a(h) \rightarrow 0$ quand $n \rightarrow +\infty$.

Leur résultat se généralise au cas \mathbb{Z}^d ou \mathbb{R}^d .

Les propriétés d'ergodicité sont importantes pour obtenir des lois des grands nombres. Si l'on veut obtenir un théorème central limite, des estimées plus fines sont nécessaires et un outil classique est la notion de mélange fort. Il s'agit de mesurer de manière quantitative à quelle vitesse décroît la dépendance entre des régions éloignées de l'espace. De nombreux coefficients de mélange ont été introduits (*cf.* les travaux de Rosenblatt [58], Bradley [6, 7, 8], Doukhan *et al.* [30, 27], Rio [57]). Nous nous sommes concentrés sur le coefficient de β -mélange (ou coefficient d'absolue régularité) de Volkonskii et Rozanov [66] défini de la manière suivante. Si X_1 et X_2 sont deux variables aléatoires à valeurs dans les espaces mesurables (E_1, \mathcal{E}_1) et (E_2, \mathcal{E}_2) respectivement, on note

$$\beta(X_1, X_2) = \sup \left\{ |\mathcal{P}_{(X_1, X_2)}(C) - \mathcal{P}_{X_1} \otimes \mathcal{P}_{X_2}(C)|; C \in \mathcal{E}_1 \otimes \mathcal{E}_2 \right\}$$

où $\mathcal{P}_{(X_1, X_2)}$, \mathcal{P}_{X_1} et \mathcal{P}_{X_2} désignent respectivement la loi de (X_1, X_2) , X_1 et X_2 .

Inégalités pour les coefficients de β -mélange

Notre principal résultat consiste en une majoration des coefficients de β -mélange associés aux processus max-i.d. Dans la suite, η désigne un processus max-i.d. à trajectoires continues sur un espace métrique localement compact T tel que

$$\text{ess inf } \eta(t) \equiv 0, \quad t \in T.$$

On désigne par μ la mesure spectrale de η sur $\mathcal{C}_0 = \mathcal{C}(T, [0, +\infty)) \setminus \{0\}$.

Si $S_1, S_2 \subset T$ sont deux parties disjointes de T , on note η_{S_i} la restriction de η à S_i et on considère le coefficient de β -mélange

$$\beta(S_1, S_2) = \beta(\eta_{S_1}, \eta_{S_2}).$$

Théorème 1.4.1 *On a la majoration*

$$\beta(S_1, S_2) \leq 2 \int_{\mathcal{C}_0} \mathbb{P}[f \not\prec_{S_1} \eta, f \not\prec_{S_2} \eta] \mu(df).$$

Le corollaire immédiat suivant est utile dans le cas discret.

Corollaire 1.4.2 *Lorsque S_1 et S_2 sont des parties finies ou dénombrables (par exemple lorsque $T = \mathbb{Z}^d$), on en déduit*

$$\beta(S_1, S_2) \leq 2 \sum_{s_1 \in S_1} \sum_{s_2 \in S_2} \int \mathbb{P}[\eta(s_1) \leq f(s_1), \eta(s_2) \leq f(s_2)] \mu(df).$$

De manière intéressante, dans le cas max-stable simple, on obtient des majorations en terme des coefficients extrémaux. On suppose η max-stable à marginales Fréchet unitaire. Pour $S \subset T$ compact, le coefficient extrémal $\theta(S)$ est défini par la relation

$$\mathbb{P}[\sup_{s \in S} \eta(s) \leq y] = \exp[-\theta(S)y^{-1}], \quad y > 0.$$

Le Théorème 1.4.1 permet alors de prouver que si S_1, S_2 sont deux compacts disjoints,

$$\beta(S_1, S_2) \leq 2 \left[C(S_1) + C(S_2) \right] \left[\theta(S_1) + \theta(S_2) - \theta(S_1 \cup S_2) \right]$$

avec

$$C(S_i) = \mathbb{E}[\sup\{\eta(s)^{-1}; s \in S_i\}] < \infty.$$

Lorsque S_1 et S_2 sont finis ou dénombrables, on trouve l'inégalité simple

$$\beta(S_1, S_2) \leq 4 \sum_{s_1 \in S_1} \sum_{s_2 \in S_2} [2 - \theta(\{s_1, s_2\})]$$

reliant les coefficients de β -mélange au coefficient extrémaux par paires.

La preuve du Théorème 1.4.1 repose sur la représentation du processus max-i.d.

$$\eta(t) = \bigvee_{\phi \in \Phi} \phi(t), \quad t \in T,$$

et sur la notion de processus S extrémal

$$\Phi_S^+ = \{\phi \in \Phi; \exists s \in S, \phi(s) = \eta(s)\}.$$

Comme η_S ne dépend que de Φ_S^+ , on a

$$\beta(S_1, S_2) = \beta(\eta_{S_1}, \eta_{S_2}) \leq \beta(\Phi_{S_1}^+, \Phi_{S_2}^+).$$

La majoration du coefficient de β mélange $\beta(\Phi_{S_1}^+, \Phi_{S_2}^+)$ est obtenue grâce à un couplage non trivial entre les lois $\mathcal{P}_{(\Phi_{S_1}^+, \Phi_{S_2}^+)}$ et $\mathcal{P}_{\Phi_{S_1}^+} \otimes \mathcal{P}_{\Phi_{S_2}^+}$ permettant de prouver

$$\beta(\Phi_{S_1}^+, \Phi_{S_2}^+) \leq 2\mathbb{P}[\Phi_{S_1}^+ \cap \Phi_{S_2}^+ = \emptyset].$$

L'inégalité précédente à une interprétation probabiliste claire puisque le terme de droite est la probabilité qu'il n'existe pas de fonction $\phi \in \Phi$ qui soit à la fois extrémale sur S_1 et S_2 . Le lien avec le Théorème 1.4.1 se fait grâce à l'inégalité suivante

$$\mathbb{P}[\Phi_{S_1}^+ \cap \Phi_{S_2}^+ = \emptyset] \leq \int_{\mathcal{C}_0} \mathbb{P}[f \not\prec_{S_1} \eta, f \not\prec_{S_2} \eta] \mu(df),$$

obtenue grâce à la formule de Slyvniak. Lorsque le processus ponctuel Φ est simple (par exemple dans le cas max-stable), on prouve également la minoration

$$\beta(\Phi_{S_1}^+, \Phi_{S_2}^+) \geq \mathbb{P}[\Phi_{S_1}^+ \cap \Phi_{S_2}^+ = \emptyset]$$

ce qui fournit un encadrement précis de $\beta(\Phi_{S_1}^+, \Phi_{S_2}^+)$. Il serait intéressant d'obtenir une borne inférieure similaire pour $\beta(S_1, S_2)$.

Applications : un théorème central limite pour les champs max-i.d. stationnaires

Les estimées précédentes sont utilisées pour obtenir un théorème central limite pour les champs max-i.d. stationnaires. Le résultat est obtenu comme une application du théorème central limite de Bolthausen [4] pour les champs stationnaires faiblement dépendants. Nous supposons donc $T = \mathbb{Z}^d$ et le champ $\eta = (\eta_t)_{t \in \mathbb{Z}^d}$ max-i.d. et stationnaire. On définit

$$\gamma(h) = \int \mathbb{P}[\eta(0) \leq f(0), \eta(h) \leq f(h)] \mu(df), \quad h \in \mathbb{Z}^d.$$

Supposons que $g : \mathbb{R}^p \rightarrow \mathbb{R}$ est une fonction mesurable et $t_1, \dots, t_p \in \mathbb{Z}^d$ sont tels que

$$\mathbb{E}[g(\eta(t_1), \dots, \eta(t_p))^{2+\delta}] < \infty \quad \text{avec } \delta > 0.$$

Supposons également que

$$\sum_{|h| \geq m} \gamma(h) = o(m^{d-1}) \quad \text{et} \quad \sum_{m=1}^{\infty} m^{d-1} \sup_{|h| \geq m} \gamma(h)^{\delta/(2+\delta)} < \infty. \quad (1.8)$$

Alors le champ aléatoire stationnaire X défini par

$$X(t) = g(\eta(t_1 + t), \dots, \eta(t_p + t)), \quad t \in \mathbb{Z}^d,$$

satisfait le théorème central limite dans le sens où

- i) la série $\sigma^2 = \sum_{t \in \mathbb{Z}^d} \text{Cov}[X(0), X(t)]$ converge absolument (éventuellement vers 0) ;
 ii) pour toute suite Λ_n de parties finies de \mathbb{Z}^d croissant vers \mathbb{Z}^d et telle que

$$\lim_{n \rightarrow \infty} |\partial\Lambda_n|/|\Lambda_n| = 0,$$

on a la convergence en loi

$$|\Lambda_n|^{-1/2} \sum_{t \in \Lambda_n} (X(t) - \mathbb{E}[X(t)]) \Longrightarrow \mathcal{N}(0, \sigma^2).$$

Ici $|\Lambda|$ désigne le cardinal de la partie $\Lambda \subset \mathbb{Z}^d$ et $\partial\Lambda$ son bord pour la structure de graphe standard sur \mathbb{Z}^d . La condition (1.8) requiert que γ tende vers 0 suffisamment vite à l'infini. Elle est satisfaite par exemple si

$$\gamma(h) \leq C \cdot |h|^{-b} \quad \text{pour } b > d \max(2, (2 + \delta)/\delta) \text{ et } C > 0.$$

Dans le cas max-stable, notons que $\gamma(h) \leq 2(2 - \theta(\{0, h\}))$. Une condition du type $\theta(h) = 2 + o(|h|^{-b})$ assure donc le théorème central limite.

Applications statistiques

Afin d'illustrer l'utilisation du théorème central limite précédent, nous proposons l'étude de la normalité asymptotique de plusieurs estimateurs simples du coefficient extrémal d'un champ aléatoire max-stable (à marginal Fréchet unitaire) stationnaire sur \mathbb{Z}^d . La relation

$$\theta(h) = -y \log p(h, y) \quad \text{avec } p(h, y) = \mathbb{P}(\eta(0) \leq y, \eta(h) \leq y), \quad y > 0,$$

suggère l'estimateur naïf

$$\hat{\theta}_n^{(1)}(h) = -y \log \hat{p}_n(h, y) \quad \text{avec } \hat{p}_n(h, y) = |\Lambda_n|^{-1} \sum_{t \in \Lambda_n} \mathbf{1}_{\{\eta(t) \leq y, \eta(t+h) \leq y\}}$$

et Λ_n une suite croissante de parties finies convergeant vers \mathbb{Z}^d et telle que $|\partial\Lambda_n|/|\Lambda_n| \rightarrow 0$ lorsque $n \rightarrow \infty$. Cet estimateur naïf $\hat{\theta}_n^{(1)}(h)$ dépend du seuil $y > 0$ choisi ce qui n'est pas très satisfaisant. De manière alternative, Smith [63] remarque que $\min(\eta(0)^{-1}, \eta(h)^{-1})$ suit une loi exponentielle de moyenne $\theta(h)^{-1}$ et propose l'estimateur

$$\hat{\theta}_n^{(2)}(h) = \frac{|\Lambda_n|}{\sum_{t \in \Lambda_n} \min(\eta(t)^{-1}, \eta(t+h)^{-1})}.$$

Cooley *et al.* [13] introduisent le F -madogramme défini par

$$\nu_F(h) = \mathbb{E}[|F(\eta(0)) - F(\eta(h))|] \quad \text{avec } F(y) = \exp(-1/y) \mathbf{1}_{\{y > 0\}}$$

et prouvent les relations

$$\nu_F(h) = \frac{1}{2} \frac{\theta(h) - 1}{\theta(h) + 1} \quad \text{et} \quad \theta(h) = \frac{1 + 2\nu_F(h)}{1 - 2\nu_F(h)}.$$

Ceci suggère l'estimateur

$$\hat{\theta}_n^{(3)}(h) = \frac{|\Lambda_n| + 2 \sum_{t \in \Lambda_n} |F(\eta(t)) - F(\eta(t+h))|}{|\Lambda_n| - 2 \sum_{t \in \Lambda_n} |F(\eta(t)) - F(\eta(t+h))|}.$$

Le théorème central limite mentionné précédemment permet de montrer la normalité asymptotique de ces estimateurs sous l'hypothèse

$$\forall h \in \mathbb{Z}^d \setminus \{0\}, \quad 2 - \theta(h) \leq C \cdot |h|^{-b} \quad \text{avec } b > 2d \text{ et } C > 0.$$

Alors les trois estimateurs de $\theta(h)$ introduits ci-dessus sont asymptotiquement normaux : quand $n \rightarrow \infty$,

$$|\Lambda_n|^{1/2} (\hat{\theta}_n^{(i)}(h) - \theta(h)) \implies \mathcal{N}(0, \sigma_i^2)$$

avec les variances limites données par

$$\begin{aligned} \sigma_1^2 &= y^2 \sum_{t \in \mathbb{Z}^d} (\exp[(2\theta(h) - \theta(\{0, h, t, t+h\}))y^{-1}] - 1), \\ \sigma_2^2 &= \theta(h)^4 \sum_{t \in \mathbb{Z}^d} \text{Cov}[\min(\eta(0)^{-1}, \eta(h)^{-1}), \min(\eta(t)^{-1}, \eta(t+h)^{-1})], \\ \sigma_3^2 &= (\theta(h) + 1)^4 \sum_{t \in \mathbb{Z}^d} \text{Cov}[|F(\eta(0)) - F(\eta(h))|, |F(\eta(t)) - F(\eta(t+h))|]. \end{aligned}$$

De manière intéressante, on peut voir que la fonction $y \mapsto \sigma_1^2$ est strictement convexe, avec limite $+\infty$ lorsque $y \rightarrow 0^+$ ou $+\infty$. On a donc un unique minimiseur y^* correspondant à un seuil optimal en terme de variance de l'estimateur $\hat{\theta}_n^{(1)}$. Malheureusement, les variances limites σ_2^2 et σ_3^2 n'ont pas une forme très explicite rendant la comparaison des trois estimateurs difficile.

1.5 Extrêmes de champs aléatoires i.i.d. et processus ponctuels [D18]

Motivations

Un problème classique en théorie des valeurs extrêmes est de déterminer le comportement asymptotique du maximum de variables aléatoires i.i.d. $(Z_i)_{i \geq 1}$: quelles hypothèses assurent la convergence du maximum renormalisé

$$\max_{1 \leq i \leq n} \frac{Z_i - b_n}{a_n}, \quad a_n > 0, \quad b_n \in \mathbb{R},$$

et quelles sont les lois limites possibles ? Ces questions sont à la base du développement de la théorie des valeurs extrêmes et trouvent leur réponse dans le théorème de Fisher et Tippett [33] caractérisant toutes les lois max-stables et dans la description de leurs domaines d'attraction par Gnedenko [38] (voir aussi de Haan [21]). Lamperti [50] a étudié les processus des maxima partiels

$$u \mapsto \max_{1 \leq i \leq [nu]} \frac{Z_i - b_n}{a_n}, \quad u \geq 0.$$

et montré leur convergence vers une limite connue sous le nom de processus extrémal. C'est un processus max-stable autosimilaire et un processus de Markov à saut pur (voir Resnick [54]).

Nous considérons les questions analogues dans un contexte fonctionnel et étudions l'asymptotique des maxima partiels construit à partir d'un tableau infini de processus indépendants. Soit T un espace métrique compact et pour $n \geq 1$, soient $\{X_{i,n}, i \geq 1\}$ des copies indépendantes d'un processus à trajectoires continues $(X_n(t))_{t \in T}$. Sans perdre de généralité, on supposera toujours que X_n est à valeurs positives, et donc dans l'espace $\mathcal{C} = \mathcal{C}(T, [0, +\infty))$. On considère le maximum

$$M_n(t) = \max\{X_{i,n}(t), 1 \leq i \leq n\}, \quad t \in T, \quad (1.9)$$

et les maxima partiels

$$\widetilde{M}_n(u, t) = \max\{X_{i,n}(t), 1 \leq i \leq [nu]\}, \quad u \geq 0, \quad t \in T. \quad (1.10)$$

Notre approche se base sur les mesures empiriques

$$\beta_n = \sum_{i=1}^n \delta_{X_{i,n}} \quad \text{et} \quad \widetilde{\beta}_n = \sum_{i \geq 1} \delta_{(X_{i,n}, i/n)} \quad (1.11)$$

sur \mathcal{C} et $\mathcal{C} \times [0, +\infty)$ respectivement. Clairement les processus M_n et \widetilde{M}_n s'écrivent comme des fonctionnelles des mesures empiriques β_n and $\widetilde{\beta}_n$ respectivement. Il est donc possible

d'étudier le comportement asymptotique des maxima à travers celui des mesures empiriques.

Les relations entre la théorie des valeurs extrêmes et la théorie des processus ponctuels sont bien connues, au moins dans le cadre des vecteurs aléatoires de dimension finie. Dans le cas d'un espace d'état localement compact, on a (*cf.* Proposition 3.21 dans Resnick [55]) : si $n\mathbb{P}[X_n \in \cdot]$ converge vaguement vers μ , alors les mesures empiriques β_n et $\tilde{\beta}_n$ convergent vers des mesures aléatoires de Poisson d'intensité μ et $\mu \otimes \ell$ respectivement, avec ℓ la mesure de Lebesgue sur $[0, +\infty)$. Dans notre contexte, l'espace fonctionnel \mathcal{C} n'est pas localement compact et le résultat précédent doit être adapté. A cette fin, on suit l'approche de Davis et Mikosch [16] basée sur la notion de mesures finies sur les bornés et la convergence \sharp -faible correspondante (*cf.* Daley et Vere-Jones [15]).

Convergence des mesures empiriques

L'approche par processus ponctuels sur espace fonctionnel s'avère relativement puissante pour traiter notre problème mais elle nécessite un travail préliminaire assez important et technique sur des espaces de mesures.

Les notions de mesure finie sur les bornées et de convergence \sharp -faible jouent un rôle important (*cf.* Daley et Vere-Jones [15] ainsi que Davis et Mikosch [16]). On rappelle ici très brièvement les grandes lignes. Si (E, d) est un espace métrique séparable muni de sa tribu borélienne \mathcal{E} , on note $\mathcal{M}_b^\sharp(E)$ l'ensemble des mesure μ finies sur les bornées, *i.e.* telles que $\mu(B) < \infty$ pour tout borné $B \in \mathcal{E}$. On dit qu'une suite $\{\mu_n, n \geq 1\}$ converge \sharp -faiblement vers μ dans $\mathcal{M}_b^\sharp(E)$ et on note $\mu_n \xrightarrow{w^\sharp} \mu$ si et seulement si $\int f d\mu_n \rightarrow \int f d\mu$ pour toute fonction f continue bornée à support borné dans E . Cette notion de convergence est métrisable grâce à une métrique sur $\mathcal{M}_b^\sharp(E)$ rendant l'espace complet et séparable. L'espace des mesures ponctuelles finies sur les bornés noté $\mathcal{M}_{(b,p)}^\sharp(E)$ est un fermé de $\mathcal{M}_b^\sharp(E)$ et, muni de la métrique induite, c'est un espace métrique complet séparable. Un processus ponctuel sur E fini sur les bornés est une variable aléatoire à valeurs dans $\mathcal{M}_{(b,p)}^\sharp(E)$, un exemple typique étant un processus ponctuel de Poisson Π_ν d'intensité $\nu \in \mathcal{M}_b^\sharp(E)$. On notera enfin \Rightarrow la convergence en loi des processus ponctuels.

On considère les processus ponctuels empiriques β_n et $\tilde{\beta}_n$ définis par (1.11), où $\{X_{i,n}, i \geq 1\}$ sont des copies i.i.d d'une variable aléatoire X_n à valeurs dans E . Grâce à une adaptation de la Proposition 3.21 dans Resnick [55] (voir aussi Mikosch [16] et Davydov, Molchanov et Zuyev [20]), on obtient l'équivalence des trois assertions suivantes, où les limites sont prises lorsque $n \rightarrow \infty$:

- i) $n\mathbb{P}[X_n \in \cdot] \xrightarrow{w^\sharp} \nu$;
- ii) $\beta_n \Rightarrow \Pi_\nu$ avec Π_ν processus ponctuel de Poisson sur E d'intensité ν ;
- iii) $\tilde{\beta}_n \Rightarrow \tilde{\Pi}_\nu$ avec $\tilde{\Pi}_\nu$ processus ponctuel de Poisson sur $E \times [0, +\infty)$ d'intensité $\nu \otimes du$.

La condition i) est une condition de type variations régulières (*cf.* Hult et Lindskog [44]). Un cas particulier important est le cas où $X_n = a_n^{-1}X_1$, avec $a_n > 0$. Cela suppose que E a une structure de cône, *i.e.* est muni d'une multiplication par les scalaires positifs. Dans ce cas, la condition i) devient

$$n\mathbb{P}[a_n^{-1}X_1 \in \cdot] \xrightarrow{w^\#} \nu$$

et si $\nu \neq 0$, on dit que X_1 est à variations régulières. Sous quelques hypothèses techniques sur la structure du cône, on prouve que la mesure limite ν est homogène :

$$\exists \alpha > 0 \quad \text{tel que} \quad \nu(\lambda A) = \lambda^{-\alpha} \nu(A), \quad \text{pour tout } \lambda > 0, A \in \mathcal{E}.$$

Extrêmes de champs aléatoires i.i.d.

On se replace dans notre cadre initial où les $X_{i,n}$ sont des processus continus sur T et à valeurs positives. L'espace $\mathcal{C} = \mathcal{C}(T, [0, +\infty))$ est naturellement muni de la norme uniforme $\|x\| = \sup_{t \in T} |x(t)|$. Mais cette métrique uniforme n'est pas adaptée à notre problème, en grande partie car les ensembles de la forme $\{x \in \mathcal{C}; \|x\| \geq \varepsilon\}$ ne sont pas bornés. Pour cette raison on introduit $\bar{\mathcal{C}}_0 = (0, +\infty) \times \mathcal{S}_{\mathcal{C}}$ avec $\mathcal{S}_{\mathcal{C}} = \{x \in \mathcal{C} : \|x\|_{\infty} = 1\}$ la sphère unité. On définit la métrique

$$d((r_1, s_1), (r_2, s_2)) = |1/r_1 - 1/r_2| + \|s_1 - s_2\|, \quad (r_1, s_1), (r_2, s_2) \in \bar{\mathcal{C}}_0$$

assurant que $(\bar{\mathcal{C}}_0, d)$ est complet et séparable. La « décomposition polaire »

$$T : \begin{cases} \mathcal{C}_0 & \rightarrow (0, \infty) \times \mathcal{S}_{\mathcal{C}} \\ x & \mapsto (\|x\|, x/\|x\|) \end{cases} .$$

permet d'identifier \mathcal{C}_0 et $(0, +\infty) \times \mathcal{S}_{\mathcal{C}}$. Pour cette métrique, un ensemble $B \subset \mathcal{C}_0$ est borné si et seulement si il est de la forme $\{x \in \mathcal{C}; \|x\| \geq \varepsilon\}$ avec $\varepsilon > 0$.

La méthode est alors la suivante. On écrit les processus de maxima M_n et \widetilde{M}_n comme des fonctionnelles des processus empiriques β_n et $\tilde{\beta}_n$, puis on vérifie de bonnes propriétés de continuité de ces fonctionnelles pour les métriques introduites ci-dessus. Le théorème de l'application continue permet alors de déduire les convergences en loi des processus M_n et \widetilde{M}_n de convergences en loi au niveau des mesures empiriques β_n et $\tilde{\beta}_n$. On obtient les résultats suivants.

Théorème 1.5.1 *Supposons que $n\mathbb{P}[X_n \in \cdot] \xrightarrow{w^\#} \nu$ dans $\mathcal{M}_b^\#(\bar{\mathcal{C}}_0^+)$ et que $\nu(\bar{\mathcal{C}}_0^+ \setminus \mathcal{C}_0^+) = 0$. Alors :*

- *le processus M_n converge en loi dans \mathcal{C} vers le processus M défini par*

$$M(t) = \sup_{i \geq 1} Y_i(t), \quad t \in T,$$

avec $\sum_{i \geq 1} \delta_{(Y_i)}$ un processus ponctuel de Poisson sur \mathcal{C}_0 d'intensité ν .

- le processus \widetilde{M}_n converge en loi dans $\mathbb{D}([0, +\infty), \mathcal{C})$ vers le processus super-extrémal \widetilde{M} défini par

$$\widetilde{M}(u, t) = \sup\{Y_i(t)\mathbf{1}_{[U_i, +\infty)}(u); i \geq 1\}, \quad u > 0, t \in T, \quad (1.12)$$

avec $\sum_{i \geq 1} \delta_{(Y_i, U_i)}$ un processus ponctuel de Poisson sur $\mathcal{C}_0 \times [0, +\infty)$ d'intensité $\nu \otimes du$.

L'article [D18] détaille aussi les propriétés des processus limites M et \widetilde{M} : max-stabilité, autosimilarité, propriété de Markov ... La puissance de l'approche par processus ponctuel nous permet aussi d'obtenir des résultats similaires pour les statistiques d'ordre plus générales (*i.e.* les r plus grandes valeurs, le maximum correspondant à $r = 1$.)

Application aux champs aléatoires Gaussiens

En géostatistique, les champs Gaussiens sont la base de nombreux modèles spatiaux. Dans un cadre fini-dimensionnel, les extrêmes de vecteurs aléatoires Gaussiens ont été étudiés par Huesler and Reiss [43]. Plus récemment, Kabluchko, Schlather and de Haan [49] et Kabluchko [47] considèrent les extrêmes de champs i.i.d. Gaussien et montrent leur convergence vers les champs max-stable du type Brown-Resnick sous des conditions portant sur l'asymptotique de la structure de covariance. Nous revisitons ces résultats à la lumière de notre approche par processus ponctuel.

On suppose que (T, d) est un espace métrique compact satisfaisant la condition d'entropie

$$\int_0^1 (\log N(\varepsilon))^{1/2} d\varepsilon < \infty,$$

avec $N(\varepsilon)$ le nombre minimal de boules de rayon ε nécessaires pour recouvrir T . Soit $Z_n = \{Z_n(t); t \in T\}$, $n \geq 1$, une suite de champs aléatoires Gaussiens centrés réduits de covariance

$$r_n(t_1, t_2) = \mathbb{E}[Z_n(t_1)Z_n(t_2)], \quad t_1, t_2 \in T.$$

On considère la suite de normalisation

$$b_n = (2 \log n)^{1/2} - (2 \log n)^{-1/2}((1/2) \log \log n + \log(2\sqrt{\pi})), \quad n \geq 1,$$

ainsi que les processus renormalisés

$$Y_n(t) = b_n(Z_n(t) - b_n) \quad \text{et} \quad X_n(t) = \exp Y_n(t), \quad t \in T.$$

On suppose les deux conditions suivantes réalisées :

- uniformément en $t_1, t_2 \in T$,

$$\Gamma(t_1, t_2) = \lim_{n \rightarrow \infty} 4 \log n (1 - r_n(t_1, t_2)) \in [0, \infty);$$

ii) pour tout $t_1, t_2 \in T$, il existe $C > 0$ tel que

$$\sup_{n \geq 1} \log n(1 - r_n(t_1, t_2)) \leq Cd(t_1, t_2).$$

Alors

$$\lim_{n \rightarrow \infty} n\mathbb{P}[X_n \in \cdot] \xrightarrow{w^\sharp} \nu(\cdot) \quad \text{dans } \mathcal{M}_b^\sharp(\bar{\mathcal{C}}_0),$$

avec

$$\nu(A) = \int_0^\infty \mathbb{P}[re^{W(\cdot) - \frac{1}{2}\Gamma(t_0, \cdot)} \in A] r^{-2} dr, \quad A \subseteq \mathcal{C}_0 \text{ borélien,}$$

où $t_0 \in T$ et $\{W(t), t \in T\}$ est un processus gaussien centré de variogramme Γ et tel que $W(t_0) = 0$. On retrouve la mesure spectrale des processus de type Brown-Resnick.

La preuve de ce résultat repose sur un critère de \sharp -convergence dans $\mathcal{M}_b^\sharp(\bar{\mathcal{C}}_0)$ dû à Hult et Lindskog [44]. On doit vérifier les trois points suivants :

- $n\mathbb{P}[X_n \in \cdot] \rightarrow \nu$ au sens de la convergence des marginales fini-dimensionnelles ;
- pour tout $r > 0$,

$$\sup_{n \geq 1} n\mathbb{P}[\sup_{t \in K} X_n(t) > r] < \infty;$$

- pour tout $\varepsilon > 0$,

$$\limsup_{\delta \rightarrow 0} \sup_{n \geq 1} n\mathbb{P}[\omega_{X_n}(\delta) \geq \varepsilon] = 0,$$

avec

$$\omega_f(\delta) = \sup\{|f(t_1) - f(t_2)|; t_1, t_2 \in T, d(t_1, t_2) \leq \delta\}$$

le module de continuité de $f \in \mathcal{C}$;

1.6 Modèles de shot-noise extrémaux [D14]

Motivations

Les shot-noises extrémaux ou *moving maximum processes* forment une classe de modèles assez simple et flexible pour modéliser les extrêmes dans un contexte spatial. Ces champs aléatoires sont définis de la manière suivante. Soit Φ un processus ponctuel de Poisson sur $\mathbb{R}^d \times (0, +\infty)$ d'intensité $dxG(dm)$, avec G mesure sur $(0, +\infty)$ de queue $\bar{G}(u) = G((u, +\infty))$ finie pour tout $u > 0$. Un point générique du processus ponctuel Φ est noté $\phi = (x_\phi, m_\phi)$. Soit $h : \mathbb{R}^d \rightarrow [0, +\infty]$ une fonction mesurable et pour $\phi \in \Phi$, on note h_ϕ la fonction $h_\phi(y) = m_\phi h(y - x_\phi)$. On considère

$$M(y) = \sup\{h_\phi(y) ; \phi \in \Phi\}, \quad y \in \mathbb{R}^d. \quad (1.13)$$

On note $\text{ESN}(h, G)$ sa loi, où ESN signifie *extremal shot noise*. Ce modèle se généralise aisément au cas de fonctions de forme aléatoires, on obtient alors les processus du type *mixed moving maximum process*, mais on se limite ici aux fonctions de formes déterministes.

Mentionnons rapidement deux exemples où ces champs apparaissent naturellement. Le premier est l'analyse des maxima annuels de précipitations journalières. La modélisation des précipitation est une tâche complexe et de nombreux auteurs ont étudié ce sujet (*cf.* Bandorff-Nielsen [2] et les références mentionnées). Les précipitations de type convectif apportent généralement de grandes quantités d'eau sur une surface réduite, avec une superposition de nombreux événements. Comme modèle de ce type de précipitation, Smith [63] propose le « processus de tempête » défini par (1.13), et dans ce contexte chaque point (x_ϕ, m_ϕ) du processus de Poisson représente un événement de tempête localisé en x_ϕ et d'intensité m_ϕ . Le choix $G(dm) = m^{-2}dm$ est standard et permet d'obtenir un champ aléatoire max-stable à marginales Fréchet unitaire. La fonction h donne la répartition spatiale des pluies lors d'un orage. Smith [63] propose la densité gaussienne multivariée de matrice de covariance Σ comme choix typique de la fonction h . Une simulation du processus de Smith en dimension 2 est donnée en Figure 1.6. Schlather [59] considère le cas des fonctions de forme aléatoire ainsi que d'autres classes de processus max-stables plus adaptées au cas des précipitations cycloniques.

Un second domaine d'application est celui des télécommunications. Dans ce contexte, le processus ponctuel de Poisson Φ représente un ensemble d'émetteurs dans l'espace euclidien, le point (x, m) représentant un émetteur en position x de puissance m . La fonction h représente alors l'atténuation du signal entre l'émetteur et le récepteur. Un choix standard dans ce cadre est une fonction de type puissance

$$h(u) = (A \max(r_0, |u|))^{-\beta} \quad \text{ou} \quad h(u) = (1 + A|u|)^{-\beta}$$

avec $A > 0$, $r_0 > 0$ et $\beta > d$. Le shot-noise extrémal (1.13) représente alors la puissance maximale recue par un récepteur positionné en y .

Ces shot-noise extrémaux sont très proches des *mixed moving maximum processes* étudiés par Schlather [59], de Haan et Pereira [25] ou encore Zhang et Smith [70]. Cependant dans ces travaux, seul le cas max-stable est traité. Nous autorisons ici une mesure G générale, si bien que le shot-noise extrémal M est seulement max-infiniment divisible et non max-stable. Les liens entre les deux se font lorsque G est une mesure de probabilité à variations régulières.

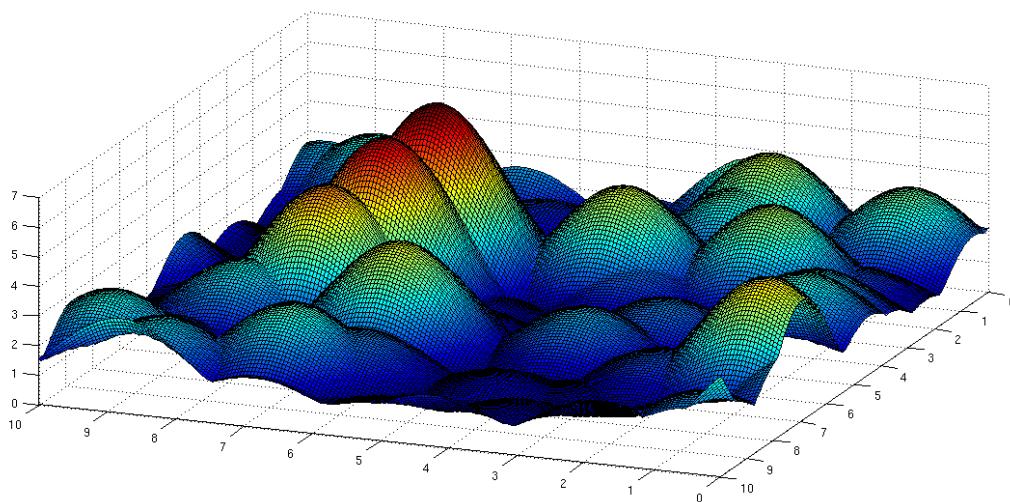


FIGURE 1.6 – Une réalisation du processus de Smith en dimension 2.

Propriétés générales des shot-noise extrémaux

L'article [D14] commence par rappeler quelques propriétés de base des shot-noise extrémaux : mesurabilité, stationnarité, max-infini-divisibilité, fonction de répartition multivariées. Une quantité importante est

$$\alpha(h, G) = \inf \left\{ u > 0; \int_{\mathbb{R}^d} \bar{G}(u/h(x)) dx < +\infty \right\},$$

qui correspond à l'infimum essentiel des variables aléatoires $M(y)$, lorsque M suit la loi $\text{ESN}(h, G)$. Lorsque $\bar{G}(u) = u^{-\alpha}$ avec $\alpha > 0$ et $\int_{\mathbb{R}^d} h(x)^\alpha dx < +\infty$, alors $\alpha(h, G) = 0$ et on retrouve le modèle de champ max-stable de Smith [63] et Schlather [59] : les fonctions de répartition multivariées s'écrivent alors

$$\mathbb{P}(M(y_1) \leq u_1, \dots, M(y_k) \leq u_k) = \exp \left(- \int_{\mathbb{R}^d} (\wedge_{1 \leq i \leq k} \{h(y_i - x)^\alpha u_i^{-\alpha}\}) dx \right).$$

On s'intéresse ensuite aux propriétés trajectorielles du champ aléatoire M et commence par déterminer sous quelles conditions le champ est borné (sur les compacts). Soit M de loi $\text{ESN}(h, G)$ et A partie mesurable non vide de \mathbb{R}^d . Soit

$$M_A(y) = \sup_{z \in A} M(y + z) \in [0, +\infty].$$

On vérifie facilement que M_A suit le loi $\text{ESN}(h_A, G)$ avec

$$h_A(x) = \sup_{z \in A} h(x + z), \quad x \in \mathbb{R}^d.$$

On en déduit que si h est localement bornée et $\alpha(h_A, G) < +\infty$, alors $\sup_{z \in A} M(z) < +\infty$ presque sûrement. En particulier, il est important de donner un critère pour que le champ M soit borné sur les compacts. Pour cela on définit

$$h^+(x) = \sup\{h(x + z), |z| \leq 1\}, \quad x \in \mathbb{R}^d, \quad (1.14)$$

et $\alpha^+(h, G) = \alpha(h^+, G)$. On montre que si h est localement bornée, alors M est presque sûrement borné sur les compacts si et seulement si $\alpha^+(h, G) < +\infty$. Des exemples « pathologiques » sont étudiés où $\alpha(h, G) = 0$ et $\alpha^+(h, G) = +\infty$. Le shot-noise extrémal correspondant M est alors fini p.s. en chaque point mais non borné sur tout ensemble ouvert.

On s'intéresse ensuite aux propriétés de régularité du shot-noise extrémal M de loi $\text{ESN}(h, G)$. Pour $\varepsilon > 0$, on définit

$$h_\varepsilon^-(x) = \inf\{h(x + z); |z| \leq \varepsilon\}, \quad x \in \mathbb{R}^d,$$

et $\alpha^-(h, G) = \lim_{\varepsilon \rightarrow 0} \alpha(h_\varepsilon^-, G)$. On a toujours :

$$0 \leq \alpha^-(h, G) \leq \alpha(h, G) \leq \alpha^+(h, G).$$

On montre que :

1. Si h est semi-continue inférieurement, alors M est semi-continue inférieurement.
2. Si $\alpha^-(h, G) = \alpha^+(h, G)$ et h est semi-continue supérieurement, alors M est semi-continue supérieurement.
3. Si $\alpha^-(h, G) = \alpha^+(h, G)$ et h est continue, alors M est continue.

La condition $\alpha^-(h, G) = \alpha^+(h, G)$ jouant un rôle important, nous donnons des conditions suffisantes pour qu'elle soit vérifiée. En supposant G finie ou h bornée, on montre que $\alpha^-(h, G) = \alpha^+(h, G)$ dès qu'une des conditions suivantes est réalisée :

- il existe $\gamma > d$ et $C > 0$ tel que

$$\lim_{x \rightarrow \infty} h(x) \bar{G}^{\leftarrow}(C \|x\|^{-\gamma}) = 0 \quad \text{avec} \quad \bar{G}^{\leftarrow}(t) = \inf\{u > 0; \bar{G}(u) \leq t\};$$

– pour tout $\delta \in (0, 1)$, il existe $R > 0$ tel que pour $\|x\| > R$ et $\|z\| \leq 1$

$$(1 - \delta)h(x) \leq h(x + z) \leq (1 + \delta)h(x).$$

On considère ensuite la propriété de séparabilité des shot-noise extrémaux. Soit D un ensemble dénombrable dense de \mathbb{R}^d . Une fonction $f : \mathbb{R}^d \rightarrow [0, +\infty]$ est dite D -séparable si pour tout $t \in \mathbb{R}^d$, il existe une suite $(t_n)_{n \geq 1}$ dans D telle que $t_n \rightarrow t$ et $f(t_n) \rightarrow f(t)$. On dit que f est universellement séparable (US) si f est D -séparable pour toute partie D dénombrable dense de \mathbb{R}^d . Par exemple, toute fonction continue est US puisque $f(x) = \lim f(x')$ lorsque $x' \rightarrow x, x' \in D$. Dans le cadre des shot-noise extrémaux, l'observation suivante joue un rôle important : la classe USUSC des fonctions universellement séparables et semi-continues supérieurement est fermée pour la topologie de la convergence uniforme sur les compacts. Elle est également invariante par translation et stable pour l'opération de maximum ponctuel. On en déduit que si M est de loi ESN(h, G) avec h USUSC et $\alpha^-(h, G) = \alpha^+(h, G)$, alors M est presque sûrement USUSC.

Pour conclure, on considère les propriétés de mélange des shot-noises extrémaux. On montre que si M est de loi ESN(h, G), alors M est mélangeant dans $[0, +\infty]^{\mathbb{R}^d}$. De plus, si h est continue et $\alpha^-(h, G) = \alpha^+(h, G)$, alors M est mélangeant dans $\mathcal{C}(\mathbb{R}^d, [0, +\infty])$.

Asymptotique des shot-noises extrémaux à queues lourdes

On suppose ici que G est une loi de probabilité sur $(0, +\infty)$ de queue \bar{G} à variations régulières d'ordre $-\alpha < 0$ en $+\infty$ et l'on note $\bar{G} \in RV_{-\alpha}$. En particulier, G est dans le domaine d'attraction de la loi de Fréchet de fonction de répartition

$$F_\alpha(x) = \exp(-x^{-\alpha})1_{\{x>0\}}.$$

Pour $\lambda > 0$, on pose

$$a_\lambda = G^{\leftarrow}(1 - \lambda^{-1}), \quad \text{avec} \quad G^{\leftarrow}(u) = \sup\{x > 0; G((0, x]) \leq u\}, \quad 0 < u < 1.$$

On a pour $x > 0$,

$$\lim_{\lambda \rightarrow +\infty} G(a_\lambda x)^\lambda = F_\alpha(x) \quad \text{et} \quad \lim_{\lambda \rightarrow +\infty} \lambda \bar{G}(a_\lambda x) = \bar{G}_\alpha(x)$$

où $\bar{G}_\alpha(x) = x^{-\alpha}$ est fonction de queue de la mesure $G_\alpha(dx) = \alpha x^{-\alpha-1}1_{\{x>0\}}dx$.

On considère le shot-noise extrémal M_λ de loi ESN($h, \lambda G$) avec $h : \mathbb{R}^d \rightarrow [0, +\infty]$ continue, $\lambda > 0$ et G mesure de probabilité à queue lourde. On considère les conditions suivantes :

$$\text{il existe } \delta \in (0, \alpha) \text{ tel que } \int_{\mathbb{R}^d} h^+(x)^{\alpha-\delta} dx < +\infty, \quad (\mathbf{C}_\alpha)$$

il existe $\gamma > d/\alpha$ et $C > 0$ tels que $|h(x)| \leq C(|x|^{-\gamma} \wedge 1)$, $x \in \mathbb{R}^d$. (\mathbf{C}'_α)

Notons que la condition (\mathbf{C}'_α) implique (\mathbf{C}_α) . On montre que si $\bar{G} \in RV_{-\alpha}$ et h satisfait la condition (\mathbf{C}_α) , alors $\alpha^+(h, \lambda G) = 0$ et aussi $\alpha^+(h, G_\alpha) = 0$. On en déduit que le champ M_λ de loi $\text{ESN}(h, \lambda G)$ est stationnaire, mélangeant et à trajectoires p.s. continues.

Notre résultat principal donne le comportement asymptotique de M_λ lorsque l'intensité λ tend vers l'infini.

Théorème 1.6.1 *Supposons $\bar{G} \in RV_{-\alpha}$ et que est h continue et satisfait (\mathbf{C}_α) . Pour $\lambda > 0$, on considère le shot-noise extrémal M_λ de loi $\text{ESN}(h, \lambda G)$. On a la convergence en loi dans $\mathcal{C}(\mathbb{R}^d, [0, +\infty))$:*

$$a_\lambda^{-1} M_\lambda \Longrightarrow \text{ESN}(h, G_\alpha) \quad \text{quand } \lambda \rightarrow +\infty.$$

Deux preuves sont proposées : une preuve directe classique, par la convergence des marginales de dimensions finies et la tension ; une preuve plus conceptuelle utilisant la convergence du processus ponctuel empirique $N_\lambda = \sum_{\phi \in \Phi_\lambda} \delta_{a_\lambda^{-1} h_\phi}$ dans un espace fonctionnel approprié. Cette preuve plus puissante permet d'obtenir des résultats non seulement pour le maximum mais pour toutes les statistiques d'ordre.

On obtient également un résultat pour le supremum des shot-noises extrémaux sur de grandes boules.

Théorème 1.6.2 *Soit M_λ de loi $\text{ESN}(h, \lambda G)$ avec $\bar{G} \in RV_{-\alpha}$ et h satisfaisant la condition (\mathbf{C}'_α) . Alors on a la convergence en loi suivante, lorsque $R \rightarrow +\infty$:*

$$\frac{1}{C(R)} \sup_{|y| \leq R} M_\lambda(y) \Longrightarrow F_\alpha,$$

avec $C(R) = \|h\|_\infty c_d^{1/\alpha} \lambda^{1/\alpha} G^{\leftarrow}(1 - R^{-d})$ et c_d le volume de la boule euclidienne unité dans \mathbb{R}^d .

1.7 Perspectives

Nous proposons dans cette section des perspectives de recherche à plus ou moins longue échéance. La plupart se place dans le prolongement des travaux [D19], [D20] et [D21]. Il nous semble que les notions de processus ponctuel extrémal et sous-extrémal constituent un outil théorique offrant de nouvelles perspectives dans la théorie des processus max-stable.

Simulations conditionnelles pour d'autres classes de processus max-stables

Dans l'article [D19], nous avons donné des formules théoriques pour les lois conditionnelles des processus max-infinitement divisibles. Dans l'article [D20], nous montrons que pour le modèle de Brown-Resnick, ces formules peuvent être implémentées sous R de façon concrète afin d'obtenir des simulations conditionnelles. Il nous semble intéressant de déterminer d'autres classes de processus max-stables effectives du point de vue des simulations.

Nous pensons entre autres au processus de type *mixed moving maximum process* de la forme

$$\eta(t) = \bigvee_{i \geq 1} \xi_i^{-1} W_i(t - t_i), \quad t \in \mathbb{R}^d,$$

où $\Phi = \{(\xi_i, t_i, W_i), i \geq 1\}$ est un processus ponctuel de Poisson sur $(0, +\infty) \times \mathbb{R}^d \times \mathcal{C}_0$ d'intensité $d\xi dt \mathbb{Q}(dW)$. Pour quel choix de loi \mathbb{Q} les lois conditionnelles de η peuvent-elles être simulées ?

Une autre classe de processus à l'étude est celle des processus de la forme

$$\eta(t) = \bigvee_{i \geq 1} \xi_i^{-1} \max(0, W_i(t)), \quad t \in T,$$

où $\{\xi_i, i \geq 1\}$ est un processus ponctuel de Poisson sur $(0, +\infty)$ d'intensité $d\xi$ et W_i une suite de processus Gaussien i.i.d. (voir Schlather [59]).

Maximum de vraisemblance pour les lois max-stables

Le Théorème 1.2.5 nous montre que la loi du vecteur max-stable k -dimensionnel $\eta(\mathbf{s})$ est égale à

$$\nu_{\mathbf{s}}(d\mathbf{y}) = \exp[-\bar{\mu}_{\mathbf{s}}(\mathbf{y})] \sum_{\tau \in \mathcal{P}_k} \bigotimes_{j=1}^{\ell} \left\{ P_{\mathbf{s}_{\tau_j}}(\mathbf{y}_{\tau_j}, \{f(\mathbf{s}_{\tau_j^c}) < \mathbf{y}_{\tau_j^c}\}) \mu_{\mathbf{s}_{\tau_j}}(d\mathbf{y}_{\tau_j}) \right\}. \quad (1.15)$$

On en déduit une condition pour que la loi de $\eta(\mathbf{s})$ soit absolument continue par rapport à la mesure de Lebesgue en terme d'absolue continuité de la mesure spectrale $\mu_{\mathbf{s}}(d\mathbf{y})$. La densité peut alors s'exprimer de manière explicite.

La formule (1.15) est nouvelle et il nous semble qu'elle ouvre la porte à l'étude systématique des méthodes de maximum de vraisemblance pour les lois max-stables. Deux questions semblent particulièrement importantes :

- Peut-on trouver l'estimateur du maximum de vraisemblance dans un modèle statistique paramétrique de lois max-stables ? Quels modèles et quels algorithmes permettent une implémentation efficace ?
- Quelles sont les propriétés théoriques de cet estimateur : consistance, efficacité, normalité asymptotique . . . ? Cela pose naturellement la question de la normalité asymptotique locale du modèle.

Différents coefficients de mélange fort

Nous avons obtenu dans [D21] des estimations des coefficients de β -mélange pour les processus max-infinitement divisibles. Dans la théorie des champs faiblement dépendants, de nombreux coefficients de mélange sont utilisés. Les techniques introduites pour estimer les coefficients de β -mélange semblent pouvoir s'adapter pour comprendre d'autres notions de mélange. On pense en particulier aux notions de m -dépendance ou encore de ϕ -mélange. Rappelons (*cf.* Bradley [5]) que le coefficient de ϕ -mélange entre deux tribus \mathcal{A} et \mathcal{B} est donné par

$$\phi(\mathcal{A}, \mathcal{B}) = \sup \left\{ |\mathbb{P}(B|A) - \mathbb{P}(B)|; A \in \mathcal{A}, B \in \mathcal{B}, \mathbb{P}(A) > 0 \right\}.$$

Absolue régularité des champs aléatoires max-stables stationnaires

Les estimées portant sur le coefficient de β -mélange nous amènent naturellement à considérer la question de l'absolue régularité des champs aléatoires max-stables stationnaires. Il semblerait dans cette optique important de fournir également une minoration du coefficient de β -mélange $\beta(S_1, S_2)$. Par ailleurs, il nous semble pertinent d'étudier les liens éventuels entre l'absolue régularité du champ et la décomposition du processus max-stable en composantes dissipative, conservative récurrente, et conservative nulle (voir Kabluchko [46]). Des calculs préliminaires montrent que les processus purement dissipatifs sont absolument réguliers.

Mosaïques aléatoires associés aux champs aléatoires max-stables

On peut associer naturellement à un processus max stable

$$\eta(t) = \bigvee_{i \geq 1} \xi_i^{-1} W_i(t), \quad t \in \mathbb{R}^d,$$

une décomposition de l'espace \mathbb{R}^d en cellules de manière à obtenir une mosaïque aléatoire.

La cellule du point 0 est ainsi définie comme le fermé aléatoire

$$C(0) = \{t \in \mathbb{R}^d; \phi_t^+ = \phi_0^+\},$$

avec ϕ_t^+ la fonction extrémale associée au point t .

Des calculs préliminaires montrent que l'on peut calculer différentes quantités relatives à cette cellule : volume moyen, moments d'ordre supérieur du volume, diverses probabilité de recouvrement . . . Les mosaïques aléatoires sont un sujet d'étude classique en géométrie aléatoire et il nous semble que l'étude systématique des propriétés des mosaïques associées au processus max-stable serait pertinente. On a déjà remarqué que la mosaïque associée au processus de Smith s'identifie à la mosaïque de Laguerre. De plus, il semble que le fait que les cellules soient bornées ou non ait un lien avec les propriétés d'ergodicité du champ max-stable sous-jacent.

Classification des processus de Brown-Resnick

Comme nous l'avons vu dans le travail [D20], les processus de Brown-Resnick sont des processus max-stables stationnaires, dont la loi est caractérisée de manière simple par un variogramme. Le modèle a de bonnes propriétés d'un point de vue computationnel (formules, simulations) grâce aux propriétés des processus Gaussiens sous-jacents.

L'idée serait ici de fournir une classification de tous les processus de Brown-Resnick reposant sur le théorème de Bochner permettant une caractérisation de tous les variogrammes possibles via leur mesure spectrale. Plus précisément, on se sert du fait que si W est un champ Gaussien centré, à accroissements stationnaires, continu en probabilité, et tel que $W(0) \equiv 0$, alors il existe

- une unique mesure de Borel σ sur $\mathbb{R}^d \setminus \{0\}$ vérifiant

$$\int_{\mathbb{R}^d} (\|x\|^2 \wedge 1) \sigma(dx) < \infty$$

- une unique matrice Σ positive de taille $d \times d$
tels que pour tout $x_1, x_2 \in \mathbb{R}^d$,

$$\mathbb{E}[W(x_1) \cdot W(x_2)] = \int_{\mathbb{R}^d} (e^{-ix_1 \cdot x} - 1) \overline{(e^{-ix_2 \cdot x} - 1)} \sigma(dx) + x_1 \cdot \Sigma x_2.$$

L'idée naturelle est alors d'étudier les propriétés du processus de Brown-Resnick en fonction de la paire (σ, Σ) .

Bibliographie

- [1] B. Abdous, A.-L. Fougères, K. Ghoudi, and P. Soulier. Estimation of bivariate excess probabilities for elliptical models. *Bernoulli*, 14(4) :1065–1088, 2008.
- [2] O.E. Bandorff-Nielsen, V.K. Gupta, V. Pérez-Abreu, and E. Waymare. *Stochastic Methods in Hydrology*. World Scientific, Singapore, New Jersey, London, Hong Kong, 1998.
- [3] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers. *Statistics of extremes*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2004. Theory and applications, With contributions from Daniel De Waal and Chris Ferro.
- [4] E. Bolthausen. On the central limit theorem for stationary mixing random fields. *Ann. Probab.*, 10(4) :1047–1050, 1982.
- [5] R.C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2 :107–144 (electronic), 2005. Update of, and a supplement to, the 1986 original.
- [6] R.C. Bradley. *Introduction to strong mixing conditions. Vol. 1*. Kendrick Press, Heber City, UT, 2007.
- [7] R.C. Bradley. *Introduction to strong mixing conditions. Vol. 2*. Kendrick Press, Heber City, UT, 2007.
- [8] R.C. Bradley. *Introduction to strong mixing conditions. Vol. 3*. Kendrick Press, Heber City, UT, 2007.
- [9] B.M. Brown and S.I. Resnick. Extreme values of independent stochastic processes. *J. Appl. Probability*, 14(4) :732–739, 1977.
- [10] T. A. Buishand, L. de Haan, and C. Zhou. On spatial extremes : with application to a rainfall problem. *Ann. Appl. Stat.*, 2(2) :624–642, 2008.
- [11] S.G. Coles and J.A. Tawn. Modelling extreme multivariate events. *J. Roy. Statist. Soc. Ser. B*, 53(2) :377–392, 1991.
- [12] S.G. Coles and J.A. Tawn. Modelling extremes of the areal rainfall process. *J. Roy. Statist. Soc. Ser. B*, 58(2) :329–347, 1996.
- [13] D. Cooley, P. Naveau, and P. Poncet. Variograms for spatial max-stable random fields. In *Dependence in probability and statistics*, volume 187 of *Lecture Notes in Statist.*, pages 373–390. Springer, New York, 2006.
- [14] D. Cooley, D. Nychka, and P. Naveau. Bayesian spatial modeling of extreme precipitation return levels. *J. Amer. Statist. Assoc.*, 102(479) :824–840, 2007.
- [15] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II*. Probability and its Applications (New York). Springer, New York, second edition, 2008. General theory and structure.
- [16] R.A. Davis and T. Mikosch. Extreme value theory for space-time processes with heavy-tailed distributions. *Stochastic Process. Appl.*, 118(4) :560–584, 2008.

-
- [17] R.A. Davis and S.I. Resnick. Basic properties and prediction of max-ARMA processes. *Adv. in Appl. Probab.*, 21(4) :781–803, 1989.
- [18] Richard A. Davis and Sidney I. Resnick. Prediction of stationary max-stable processes. *Ann. Appl. Probab.*, 3(2) :497–525, 1993.
- [19] A. Davison, S. Padoan, and M. Ribatet. Statistical modelling of spatial extremes. *Statistical Science*. To appear.
- [20] Y. Davydov, I. Molchanov, and S. Zuyev. Strictly stable distributions on convex cones. *Electron. J. Probab.*, 13 :no. 11, 259–321, 2008.
- [21] L. de Haan. *On regular variation and its application to the weak convergence of sample extremes*, volume 32 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1970.
- [22] L. de Haan. A characterization of multidimensional extreme-value distributions. *Sanhyā Ser. A*, 40(1) :85–88, 1978.
- [23] L. de Haan. A spectral representation for max-stable processes. *Ann. Probab.*, 12(4) :1194–1204, 1984.
- [24] L. de Haan and A. Ferreira. *Extreme value theory*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006. An introduction.
- [25] L. de Haan and T.T. Pereira. Spatial extremes : models for the stationary case. *Ann. Statist.*, 34(1) :146–168, 2006.
- [26] L. de Haan and J. Pickands, III. Stationary min-stable stochastic processes. *Probab. Theory Relat. Fields*, 72(4) :477–492, 1986.
- [27] J. Dedecker, P. Doukhan, G. Lang, J.R. León R., S. Louhichi, and C. Priour. *Weak dependence : with examples and applications*, volume 190 of *Lecture Notes in Statistics*. Springer, New York, 2007.
- [28] P. Deheuvels. Caractérisation complète des lois extrêmes multivariées et de la convergence aux types extrêmes. *Publ. Inst. Statist. Univ. Paris*, 23, 1978.
- [29] P. Deheuvels. Probabilistic aspects of multivariate extremes. In *Statistical extremes and applications (Vimeiro, 1983)*, volume 131 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 117–130. Reidel, Dordrecht, 1984.
- [30] P. Doukhan. *Mixing*, volume 85 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1994. Properties and examples.
- [31] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. For insurance and finance.
- [32] M. Falk, J. Hüsler, and R.-D. Reiss. *Laws of small numbers : extremes and rare events*, volume 23 of *DMV Seminar*. Birkhäuser Verlag, Basel, 1994. With an appendix by S. Haßmann, Reiss and M. Thomas, With 1 IBM-PC floppy disk (3.5 inch ; HD).
- [33] R.A. Fisher and L.H.C. Tippett. Limiting forms of the frequency of the largest or smallest member of a sample. *Proc. Cambridge Philos. Soc.*, 24 :180–190, 1928.

- [34] A.-L. Fougères and P. Soulier. Limit conditional distributions for bivariate vectors with polar representation. *Stoch. Models*, 26(1) :54–77, 2010.
- [35] M. Fréchet. Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Math. Polon.*, 6 :93–116, 1927.
- [36] J. Galambos. *The asymptotic theory of extreme order statistics*. John Wiley & Sons, New York-Chichester-Brisbane, 1978. Wiley Series in Probability and Mathematical Statistics.
- [37] E. Giné, M.G. Hahn, and P. Vatan. Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields*, 87(2) :139–165, 1990.
- [38] B. Gnedenko. Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math. (2)*, 44 :423–453, 1943.
- [39] E. J. Gumbel. *Statistics of extremes*. Columbia University Press, New York, 1958.
- [40] J.E. Heffernan and S.I. Resnick. Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.*, 17(2) :537–571, 2007.
- [41] J.E. Heffernan and J.A. Tawn. A conditional approach for multivariate extreme values. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 66(3) :497–546, 2004. With discussions and reply by the authors.
- [42] B.M. Hill. A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3(5) :1163–1174, 1975.
- [43] J. Huesler and R.-D. Reiss. Maxima of normal random vectors : between independence and complete dependence. *Statist. Probab. Lett.*, 7(4) :283–286, 1989.
- [44] H. Hult and F. Lindskog. Extremal behavior of regularly varying stochastic processes. *Stochastic Process. Appl.*, 115(2) :249–274, 2005.
- [45] Z. Kabluchko. Spectral representations of sum- and max-stable processes. *Extremes*, 12(4) :401–424, 2009.
- [46] Z. Kabluchko. Spectral representations of sum- and max-stable processes. *Extremes*, 12(4) :401–424, 2009.
- [47] Z. Kabluchko. Extremes of independent gaussian processes. *Extremes*, To appear, DOI : 10.1007/s10687-010-0110-x.
- [48] Z. Kabluchko and M. Schlather. Ergodic properties of max-infinitely divisible processes. *Stochastic Process. Appl.*, 120(3) :281–295, 2010.
- [49] Z. Kabluchko, M. Schlather, and L. de Haan. Stationary max-stable fields associated to negative definite functions. *Ann. Probab.*, 37(5) :2042–2065, 2009.
- [50] J. Lamperti. On extreme order statistics. *Ann. Math. Statist*, 35 :1726–1737, 1964.
- [51] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York, 1983.
- [52] M. Oesting, Z. Kabluchko, and M. Schlather. Simulation of Brown–Resnick processes. *Extremes*, 2012.

-
- [53] III J. Pickands. Statistical inference using extreme order statistics. *Ann. Statist.*, 3 :119–131, 1975.
- [54] S.I. Resnick. Weak convergence to extremal processes. *Ann. Probability*, 3(6) :951–960, 1975.
- [55] S.I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.
- [56] S.I. Resnick and R. Roy. Random usc functions, max-stable processes and continuous choice. *Ann. Appl. Probab.*, 1(2) :267–292, 1991.
- [57] E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2000.
- [58] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U. S. A.*, 42 :43–47, 1956.
- [59] M. Schlather. Models for stationary max-stable random fields. *Extremes*, 5(1) :33–44, 2002.
- [60] M. Schlather and J. Tawn. Inequalities for the extremal coefficients of multivariate extreme value distributions. *Extremes*, 5(1) :87–102, 2002.
- [61] M. Schlather and J. Tawn. A dependence measure for multivariate and spatial extreme values : properties and inference. *Biometrika*, 90(1) :139–156, 2003.
- [62] M. Sibuya. Bivariate extreme statistics. I. *Ann. Inst. Statist. Math. Tokyo*, 11 :195–210, 1960.
- [63] R. Smith. Max-stable processes and spatial extremes. unpublished manuscript, 1990.
- [64] S.A. Stoev. Max-stable processes : representations, ergodic properties and statistical applications. In *Dependence in probability and statistics*, volume 200 of *Lecture Notes in Statist.*, pages 21–42. Springer, Berlin, 2010.
- [65] J. Tiago de Oliveira. La représentation des distributions extrémales bivariées. *Bull. Inst. Internat. Statist.*, 39(livraison 2) :477–480, 1962.
- [66] V.A. Volkonskiĭ and Y.A. Rozanov. Some limit theorems for random functions. I. *Theor. Probability Appl.*, 4 :178–197, 1959.
- [67] R. von Mises. La distribution de la plus grande de n valeurs. *Selected Papers II, American Math. Soc.*, pages 271–294, 1936.
- [68] Y. Wang and S.A. Stoev. Conditional sampling for spectrally discrete max-stable random fields. *Adv. in Appl. Probab.*, 43(2) :461–483, 2011.
- [69] K.S. Weintraub. Sample and ergodic properties of some min-stable processes. *Ann. Probab.*, 19(2) :706–723, 1991.
- [70] Z. Zhang and R.L. Smith. The behavior of multivariate maxima of moving maxima processes. *J. Appl. Probab.*, 41(4) :1113–1123, 2004.

Chapitre 2

Théorèmes limites, dépendance à longue portée et variations régulières

Nous reprenons dans ce chapitre les résultats des articles [D5], [D6], [D7], [D8], [D12], [D13], [D15] et [D17]. Après une brève introduction à la thématique « Théorèmes limites, dépendance à longue portée et variations régulières », nous développons quatre thèmes différents : les théorèmes limites associés aux marches aléatoires en scènes aléatoires, les modèles d'agrégation de sources et les processus fractionnaires, les modèles de boules aléatoires et leurs asymptotiques et enfin les séries de LePage. Le chapitre se termine sur quelques perspectives de recherche.

2.1 Introduction

La dépendance à longue portée, encore appelée longue mémoire, est une notion importante en modélisation comme en témoignent de nombreuses publications récentes dans des domaines d'applications variés : finance [45], économétrie [59], hydrologie [56], climatologie [70], séquençage de l'ADN [39] . . . Cette notion a été introduite dans les années 60 par Mandelbrot dans une série de travaux [49, 51, 52] cherchant à expliquer la loi climatologique de Hurst [33]. Hurst, hydrologue anglais, s'est intéressé au niveau du Nil pour lequel des données historiques existent sur plusieurs siècles, et a remarqué que ces séries chronologiques ont un comportement surprenant : la statistique R/S mesurant l'étendue standardisée des fluctuations des sommes partielles de la série chronologique est anormalement grande ; pour une série de n d'observations, on attendrait typiquement des fluctuations d'ordre n^H avec $H = 1/2$ mais Hurst met en évidence un ordre de grandeur plus grand avec $H \approx 0,74$. Mandelbrot a montré qu'une série chronologique basée sur les accroissements d'un mouvement Brownien fractionnaire permet de rendre compte de ce phénomène. Il a mis en évidence le phénomène de persistance, popularisé sous le terme de longue mémoire ou encore l'« effet Joseph », en référence à la Génèse où Joseph annonce à Pharaon sept années de vaches grasses suivies sept années de vaches maigres.

Le mouvement Brownien fractionnaire a été introduit par Kolmogorov [41] en 1940 comme moyen d'engendrer des « spirales gaussiennes » dans les espaces de Hilbert. Pour $H \in (0, 1)$, on définit le processus Gaussien centré de covariance

$$\text{Cov}[B_H(s), B_H(t)] = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}), \quad s, t \geq 0.$$

C'est un processus à accroissements stationnaires qui a des propriétés de dépendance particulières : indépendance des accroissements pour $H = 1/2$ (on retrouve alors un mouvement Brownien standard), dépendance à longue portée pour $H > 1/2$, et intermittence pour $H < 1/2$. Une autre caractéristique importante est son invariance en loi par changement d'échelle, appelée autosimilarité : pour tout $c > 0$, les processus $(B_H(ct))_{t \geq 0}$ et $(c^H B_H(t))_{t \geq 0}$ ont la même loi. Cette notion d'autosimilarité a connu un succès important avec le développement de la géométrie fractale [50].

Il serait impossible dans cette courte introduction de retracer les nombreux développements autour de ces notions très riches. On pourra se référer aux articles de revue de Cox [15], Rosenblatt [60], Taqqu [69] Samorodnitsky [63] et aux monographies de Beran [7] et Doukhan [23]. Mentionnons cependant quelques directions importantes :

- la théorie des processus de variance infinie α -stable, la longue mémoire et l'autosimilarité dans ce cadre avec les travaux de Samorodnitsky [62], Samorodnitsky et Taqqu [64, 65, 66], Maejima [46, 47] ;
- les théorèmes limites en présence de longue mémoire, souvent atypiques et très riches, avec les travaux de Davydov [18], Dobrushin et Major [20], Surgailis [67] ;
- les méthodes statistiques originales en présence de longue mémoire, avec les travaux

de Soulier [5, 34, 35, 55] ou Doukhan [21, 22] ;

- les différentes généralisations des processus fractionnaires (draps fractionnaires, processus multifractionnaires ...) avec les travaux de Cohen [2, 6] ou Ayache [3, 4] ;

Dans ce chapitre, nous présentons nos travaux autour des théorèmes limites, de la dépendance à longue portée et des variations régulières. Nous considérons différentes constructions de processus stochastiques mettant en jeu une forme de dépendance à longue portée et considérons leur convergence vers des processus limites ayant des propriétés remarquables telles que l'autosimilarité, la stabilité . . . En particulier, dans de nombreux cas, nous retrouvons le mouvement Brownien fractionnaires à la limite, mais aussi de nouveaux processus stables autosimilaires et à accroissements stationnaires. Le chapitre s'organise autour de quatre axes :

- les modèles de marches aléatoires en scène aléatoires ;
- les modèles d'agrégation de sources de type ON/OFF ou renouvellement ;
- les modèles de boules aléatoires et leurs asymptotiques ;
- les séries de LePage.

Nous proposons ensuite quelques perspectives de recherche.

2.2 Théorèmes limites et marches aléatoires en scènes aléatoires [D5,D6,D7]

Motivations

Nous présentons les résultats des articles [D5], [D6] et [D7] qui portent sur les théorèmes limites relatifs aux modèles de « marches aléatoires en scènes aléatoires » en lien avec une classe de processus stables autosimilaires et à accroissements stationnaires appelés *local time fractional stable motion*.

Le point de départ de ce travail est un papier de Cohen et Samorodnitsky [14] introduisant une nouvelle classe de processus stables, autosimilaires et à accroissements stationnaires, appelé LTFSM (pour *local time fractional stable motion*) et construit comme l'intégrale stable du temps local d'un mouvement brownien fractionnaire. Les auteurs étudient les propriétés de ce processus (autosimilarité, régularité, représentation en série, système dynamique associé ...) et présentent également un modèle discret qui, sous les renormalisations adéquates, devrait converger vers le LTFSM. Le schéma discret baptisé « schéma de récompenses aléatoires » s'interprète comme la récompense cumulée de marcheurs se déplaçant sur un réseau. La preuve de la convergence de ce schéma discret vers le LTFSM n'est donnée dans [14] que dans le cas où l'indice de Hurst du Brownien fractionnaire est égal à $1/2$. Dans ce cas, le Brownien fractionnaire est un Brownien standard et les auteurs utilisent un principe d'invariance fort du mouvement brownien par la marche aléatoire. Le cas $H \neq 1/2$ est alors laissé à l'état de conjecture.

Le schéma de récompenses aléatoires introduit par Cohen et Samorodnitsky [14] est fortement lié aux marches aléatoires en scène aléatoire introduite en 1979 par Kesten et Spitzer [40]. En effet, le schéma de récompenses aléatoires n'est rien d'autre qu'une somme de copies indépendantes de marches aléatoires en scène aléatoire. Kesten et Spitzer [40] traitent le cas d'une marche dans le domaine d'attraction d'un Lévy α -stable, $\alpha \in (1, 2]$ et d'une scène dans le domaine d'attraction d'un Lévy β -stable, $\beta \in (0, 2]$. Cela suggère d'étendre le principe d'invariance pour le schéma de récompenses aléatoires à ces hypothèses et d'obtenir à la limite de nouveaux processus stables autosimilaires et à accroissements stationnaires, construit comme l'intégrale α -stable du temps local d'un Lévy β -stable. Cela fait l'objet de l'article [D5] en collaboration avec N.Guillot-Plantard.

L'analogie du théorème de Kesten et Spitzer pour les marches fortement corrélées et dans le domaine d'attraction d'un Brownien fractionnaire a été établi par Wang [71]. En collaboration avec S.Cohen, nous utilisons dans [D6] les résultats de Wang pour étudier le schéma de récompenses aléatoires avec marches fortement corrélées et ainsi traiter le cas du mouvement Brownien fractionnaire d'indice de Hurst $H \neq 1/2$.

Les modèles de marches aléatoires en scènes aléatoires ont donné lieu à de nombreuses variantes, selon les types de marches ou de scènes considérées. On peut alors adopter un

point de vue très général et se poser la question suivante : quelles hypothèses sur la marche et sur la scène permettent d'obtenir un théorème limite pour la marche aléatoire en scène aléatoire ? Le travail [D7] en collaboration avec N. Guillin-Plantard répond à cette question par une approche fonctionnelle du problème.

Travaux antérieurs

Ces résultats s'inscrivent dans une littérature assez riche autour des marches aléatoires en scène aléatoire. Les travaux fondateurs sont ceux de Kesten et Spitzer [40] et Borodin [11, 12] motivés par la construction d'une nouvelle classe de processus autosimilaires et à accroissements stationnaires. Ils traitent le cas d'un marche à accroissements i.i.d. en dimension 1 et d'une scène i.i.d. dans le domaine d'attraction d'une loi stable. Bolthausen [10] considère le cas d'une marche récurrente dans \mathbb{Z}^2 . Maejima [48] traite le cas d'une marche en dimension 1 et d'une scène vectorielle en lien avec la propriété « d'autosimilarité par opérateur ». Tous ces travaux sont basés sur des propriétés fortes d'indépendance. Wang [71] s'affranchit de telles hypothèse et considère le cas d'une marche fortement corrélée dans le domaine d'attraction du Brownien fractionnaire. Lang et Xahn [42] considère le cas d'une marche i.i.d. dans une scène fortement corrélée vérifiant un théorème limite non central de Dobrushin et Major [20]. Le cas des scènes faiblement dépendantes vérifiant une condition de θ -mélange est traité par Guillin-Plantard et Prieur [30, 31]. Les marches aléatoires interviennent dans l'étude des marches aléatoires sur le réseau orienté \mathbb{Z}^2 , où les orientations sont aléatoires (*cf.* Guillin-Plantard et Le Ny [28, 29]) et dans le contexte des polymères dirigés (*cf.* Chen et Khoshnevisan [13]).

Modèle et notations

Avant de décrire nos résultats, nous précisons le cadre général et les notations. La *scène* $\xi = \{\xi_x; x \in \mathbb{Z}\}$ est une collection de variables aléatoires indexées par \mathbb{Z} . La *marche* $S = \{S_n, n \geq 0\}$ est un processus à temps discret à valeur dans \mathbb{Z} . On supposera toujours l'indépendance de la marche S et de la scène ξ . La *marche aléatoire en scène aléatoire* est le processus cumulatif des valeurs de la scène (récompenses) rencontré par le marcheur

$$Z_n = \sum_{k=0}^n \xi_{S_k}, \quad n \geq 0.$$

On étend la définition au temps $t \geq 0$ par interpolation linéaire, le processus $(Z_t)_{t \geq 0}$ est donc continu.

Pour définir le *schéma de récompenses aléatoires*, on considère des copies i.i.d. de la marche aléatoire en scène aléatoire, marquées par un indice supérieur $i \geq 1$: la i -ème scène est $\xi^i = \{\xi_x^i; x \in \mathbb{Z}\}$, la i -ème marche $S^i = \{S_n^i, n \geq 0\}$ et la i -ème marche aléatoire en scène aléatoire $Z^i = \{Z_n^i, n \geq 0\}$. Le schéma de récompenses aléatoires d'ordre c correspond

alors à la somme indépendante

$$\sum_{i=1}^c Z_n^i = \sum_{i=1}^c \sum_{k=0}^n \xi_{S_k^i}^i, \quad n \geq 0.$$

Le cas des marches à accroissements indépendants [D5]

Nous décrivons ici les résultats de [D5]. On se place dans les hypothèses du théorème de Kesten et Spitzer [40] et l'on suppose que :

- la scène ξ est i.i.d. et dans le domaine normal d'attraction d'un Lévy β -stable, $\beta \in (0, 2]$;
- la marche S est issue de 0, à accroissements i.i.d et dans le domaine d'attraction normal d'un Lévy α -stable, $\alpha \in (1, 2]$.

Sous ces conditions, on a la convergence en loi dans l'espace des fonctions continues :

$$(n^{-\delta} Z_{nt})_{t \geq 0} \Longrightarrow \left(\Delta(t) = \int_{-\infty}^{+\infty} L_t(x) dW(x) \right)_{t \geq 0} \quad \text{quand } n \rightarrow \infty,$$

où $\delta = 1 - \alpha^{-1} + (\alpha\beta)^{-1}$, L_t est le temps local d'un processus de Lévy α -stable associé à la marche et W un processus de Lévy β -stable bilatère associé à la scène. Le processus Δ est continu, δ -autosimilaire et à accroissements stationnaires. Il peut être vu comme un mélange de processus β -stables mais n'est pas un processus stable.

Afin d'obtenir un processus stable, on considère des sommes indépendantes de tels processus. Soient $\Delta^{(i)}$, $i \geq 1$ des copies indépendantes du processus Δ .

Théorème 2.2.1 *On a la convergence en loi dans $\mathcal{C}([0, \infty), \mathbb{R})$:*

$$\left(n^{-\frac{1}{\beta}} \sum_{i=1}^n \Delta^{(i)}(t) \right)_{t \geq 0} \Longrightarrow (\Gamma(t))_{t \geq 0} \quad \text{quand } n \rightarrow \infty,$$

où le processus limite Γ est continu, β -stable, δ -autosimilaire et à accroissements stationnaires.

Remarquons que nécessairement, dans le cas $\beta = 2$ où la scène a une variance finie, le processus Γ est gaussien δ -autosimilaire et à accroissements stationnaires, c'est donc un mouvement brownien fractionnaire d'indice de Hurst $\delta = 1 - \frac{1}{2\alpha} \in (1/2, 3/4]$.

En remplaçant les processus Δ^i par les marches aléatoires en scènes aléatoires correspondantes, on retrouve une variante du schéma de récompenses aléatoires proposé par Cohen et Samorodnitsky [14]. Le principe d'invariance correspondant est le suivant :

Théorème 2.2.2 Soit c_n une suite d'entiers telle que $\lim c_n = +\infty$. On a la convergence en loi dans $\mathcal{C}([0, \infty), \mathbb{R})$:

$$\left(c_n^{-\frac{1}{\beta}} n^{-\delta} \sum_{i=1}^{c_n} Z_{nt}^i \right)_{t \geq 0} \Longrightarrow (\Gamma(t))_{t \geq 0} \quad \text{quand } n \rightarrow \infty,$$

où le processus limite Γ est le même que dans le Théorème 2.2.1.

Les preuves des Théorèmes 2.2.1 et 2.2.2 reposent sur la convergence des lois fini-dimensionnelles (méthode de fonction caractéristique) et la tension (critère standard). Des résultats fins sur les temps locaux sont utilisés (cf. [D5] Lemme 3.1.).

L'étude se poursuit en donnant quelques propriétés de processus limite Γ . On donne une représentation intégrale de Γ analogue à celle du *local time fractional stable motion* proposée par Cohen et Samorodnitsky [14] en remplaçant le temps local du Brownien fractionnaire le temps local d'un Lévy β -stable. Plus précisément, on montre que Γ admet la représentation intégrale :

$$\Gamma(t) = \int_{\Omega' \times \mathbb{R}} \sigma L_t(x)(\omega') M(d\omega', dx), \quad t \geq 0$$

avec

- $(\Omega', \mathcal{F}', \mathbb{P}')$ un espace probabilisé sur lequel est défini un processus de Lévy α -stable de temps local bicontinu $L_t(x)$,
- $M(d\omega', dx)$ une mesure aléatoire β -stable sur $\Omega' \times \mathbb{R}$ de mesure de contrôle $\mathbb{P}' \otimes dx$ et biais égal à ν (le paramètre de biais est déterminé par la scène ξ).

Pour les définitions et propriétés des mesures aléatoires stables, on pourra se référer à la monographie de Samorodnitsky et Taqqu [66]. Cette représentation intégrale s'avère un outil puissant pour étudier certaines propriétés trajectoires du processus Γ . Les résultats trajectoires sont donnés dans le cas $\beta \in (0, 2)$. Dans le cas $\beta = 2$, Γ est un mouvement Brownien fractionnaire bien connu. On considère d'abord l'asymptotique pour la queue du supremum de Γ sur $[0, T]$: pour $\beta \in (0, 2)$ et $T > 0$,

$$\lim_{u \rightarrow \infty} u^\beta \mathbb{P} \left(\sup_{t \in [0, T]} \Gamma(t) \geq u \right) = C_\beta \frac{1 + \nu}{2} \mathbb{E} \left(\int_{\mathbb{R}} L_T(x)^\beta dx \right)$$

et

$$\lim_{u \rightarrow \infty} u^\beta \mathbb{P} \left(\sup_{t \in [0, T]} |\Gamma(t)| \geq u \right) = C_\beta \mathbb{E} \left(\int_{\mathbb{R}} L_T(x)^\beta dx \right),$$

avec

$$C_\beta = \left(\int_0^\infty x^{-\beta} \sin x dx \right)^{-1}.$$

On s'intéresse enfin à la régularité Hölder des trajectoires : si $\beta \in (0, 2)$ et $\nu = 0$ lorsque $\beta \in [1, 2)$, alors on a presque sûrement

$$\sup_{0 \leq s < t \leq 1/2} \frac{|\Gamma(t) - \Gamma(s)|}{(t - s)^{1 - \frac{1}{\alpha}} |\log(t - s)|^{\frac{1}{\alpha} + \varepsilon}} < \infty,$$

avec

$$\varepsilon = \begin{cases} 0 & \text{si } 1 \leq \beta < 2 \\ \frac{1}{2} & \text{si } 0 < \beta < 1 \end{cases} .$$

Le cas des marches à accroissements fortement corrélés [D6]

Nous décrivons ici les résultats de [D6]. L'objectif est de démontrer la convergence du schéma de récompenses aléatoires, vers le *local time fractional stable motion* défini par Cohen et Samorodnitsky [14] de la façon suivante :

$$\Gamma_H(t) \stackrel{\mathcal{L}}{=} \int_{\Omega' \times \mathbb{R}} L_t(x)(\omega') M(d\omega', dx) , \quad t \geq 0$$

avec

- $(\Omega', \mathcal{F}', \mathbb{P}')$ est un espace probablisé sur lequel est défini un mouvement Brownien fractionnaire B_H d'indice de Hurst $H \in (0, 1)$ et de temps local bicontinu $L_t^H(x)$,
- $M(d\omega', dx)$ est une mesure aléatoire β -stable sur $\Omega' \times \mathbb{R}$ de mesure de contrôle $\mathbb{P}' \otimes dx$ et biais égal à ν .

Les propriétés d'autosimilarité, de stationnarité des accroissements et certaines propriétés trajectoires du processus Γ_H ont été étudiées dans [14]. La convergence du schéma est prouvée dans [14] lorsque $H = 1/2$. Ce cas est également un cas particulier des résultats de la section précédente : en effet, $B_{1/2}$ est un mouvement Brownien standard et un schéma avec marche à accroissements i.i.d. de variance finie convient.

Dans le cas $H \neq 1/2$, l'idée est d'approcher le mouvement Brownien fractionnaire par une marche à accroissement fortement corrélés. On se place sous les hypothèses du travail de Wang [71] et l'on suppose que :

- la scène ξ est i.i.d. et dans le domaine normal d'attraction d'un Lévy β -stable, $\beta \in (0, 2]$;
- la marche S est construite à partir d'une suite gaussienne stationnaire et dans le domaine d'attraction d'un mouvement Brownien fractionnaire B_H d'indice de Hurst $H \in (0, 1)$. Plus précisément, on pose

$$\begin{cases} S_0 = 0, \\ S_n = [\sum_{k=1}^n X_k] , \quad n \geq 1, \end{cases}$$

avec $X_i, i \geq 1$ suite gaussienne stationnaire centrée de covariance $r(i-j) = \mathbb{E}[X_i X_j]$ satisfaisant

$$\sum_{i=1}^n \sum_{j=1}^n r(i-j) \sim n^{2H} \quad \text{quand } n \rightarrow \infty.$$

Ici $[\cdot]$ désigne la partie entière.

Sous ces conditions, on a la convergence en loi dans l'espace de Skohorod $\mathcal{D}([0, +\infty), {}^b bR)$

$$\frac{1}{n^H} (S_{[nt]})_{t \geq 0} \Longrightarrow (B_H(t))_{t \geq 0} \quad \text{quand } n \rightarrow \infty,$$

avec B_H mouvement brownien fractionnaire d'indice de Hurst H . L'analogie du théorème de Kesten et Spitzer, démontré par Wang [W03], s'écrit alors

$$(n^{-\delta_H} Z_{nt})_{t \geq 0} \implies \left(\Delta_H(t) = \int_{-\infty}^{+\infty} L_t^H(x) dW(x) \right)_{t \geq 0}$$

avec $\delta_H = 1 - H + H\beta^{-1}$, L_t^H temps local du mouvement Brownien fractionnaire B_H associé à la marche et W processus de Lévy β -stable bilatère associé à la scène. Le processus Δ_H est continu, δ_H -autosimilaire et à accroissements stationnaires. Il peut être vu comme un mélange de processus β -stables mais n'est pas un processus stable.

De manière analogue au Théorème 2.2.1, un premier principe d'invariance pour Γ_H s'obtient à partir de copies indépendantes du processus Δ_H , notées Δ_H^i , $i \geq 1$.

Théorème 2.2.3 *On a la convergence en loi dans $\mathcal{C}([0, \infty), \mathbb{R})$*

$$\left(n^{-\frac{1}{\beta}} \sum_{i=1}^n \Delta^{(i)}(t) \right)_{t \geq 0} \implies (\Gamma_H(t))_{t \geq 0} \quad \text{quand } n \rightarrow \infty.$$

On en vient ensuite au schéma de récompenses aléatoires proprement dit, qui fournit une approximation discrète de Γ_H par des marches fortement corrélées en scènes aléatoires. On prouve ainsi la conjecture proposée par Cohen et Samorodnitsky [14].

Théorème 2.2.4 *Soit c_n une suite d'entiers telle que $\lim c_n = +\infty$. On a la convergence en loi dans $\mathcal{C}([0, \infty), \mathbb{R})$:*

$$\left(c_n^{-\frac{1}{\beta}} n^{-\delta} \sum_{i=1}^{c_n} Z_{nt}^i \right)_{t \geq 0} \implies (\Gamma_H(t))_{t \geq 0} \quad \text{quand } n \rightarrow \infty.$$

Si les Théorèmes 2.2.3 et 2.2.4 sont formellement très proches des Théorèmes 2.2.1 et 2.2.2, les preuves sont sensiblement plus difficiles, car la dépendance forte des accroissements de la marche entraîne des difficultés techniques supplémentaires. Pour donner une idée des techniques mises en jeu, mais sans trop rentrer dans les détails, notons que le temps local, le temps local d'auto-intersection et la taille du support de la marche fortement corrélée S jouent un rôle important dans la preuve.

On définit le temps local au point $x \in \mathbb{Z}$ et au temps $n \geq 1$ de la marche S comme le nombre de visites de la marche entre les temps 0 et n :

$$N(n, x) = \sum_{k=0}^n 1_{\{S_k=x\}}.$$

Le maximum en espace du temps local au temps n est noté L_n :

$$L_n = \sup_{x \in \mathbb{Z}} N(n, x).$$

Le nombre d'auto-intersections de la marche entre les temps 0 et n est défini par

$$V_n = \sum_{0 \leq i, j \leq n} 1_{\{S_i = S_j\}} = \sum_{x \in \mathbb{Z}} N(n, x)^2.$$

Enfin, la taille du support de la marche entre les temps 0 et n est

$$R_n = \sum_{x \in \mathbb{Z}} 1_{\{N(n, x) \neq 0\}}.$$

Les estimées suivantes jouent un rôle important dans la preuve du Théorème 2.2.4 :

- Pour tout $\delta' > 1 - H$ (et en particulier pour $\delta' = \delta_H$), on a la convergence en probabilité

$$n^{-\delta'} L_n \xrightarrow[n \rightarrow \infty]{P} 0.$$

- Pour tout $p \in [1, +\infty)$, il existe $C > 0$ tel que pour tout $n \geq 1$,

$$\mathbb{E}(V_n^p) \leq Cn^{p(2-H)}.$$

- Pour tout $p \in [1, +\infty)$, il existe $C > 0$ tel que pour tout $n \geq 1$,

$$\mathbb{E}(R_n^p) \leq Cn^{pH}.$$

Le cas général : une approche fonctionnelle [D7]

Dans le travail [D7], nous essayons de développer un cadre général assurant un principe d'invariance pour une marche aléatoire en scène aléatoire. La méthode permet de découpler les hypothèses sur la marche et sur la scène, ce qui est assez naturel étant donné que marche et scène sont supposées indépendantes. On ramène alors le problème à deux problèmes plus simples :

- la convergence dans L^p du temps local de la marche d'une part ;
- la convergence de mesures aléatoires liées à la scène d'autre part.

Critère de convergence dans L^p du temps local de la marche

Rappelons que le temps local de la marche est défini par

$$N(n, x) = \sum_{k=0}^n 1_{\{S_k = x\}}.$$

Nous obtenons le résultat suivant :

Proposition 2.2.5 *Supposons la condition (RW) suivante satisfaite :*

- (RW1) *Il existe une suite a_n telle que $a_n \rightarrow \infty$, $n^{-1}a_n \rightarrow 0$ et que la marche renormalisée $(a_n^{-1}S_{[nt]})_{t \geq 0}$ converge dans $\mathcal{D}([0, \infty), \mathbb{R})$ vers un processus $(Y_t)_{t \geq 0}$ possédant un temps local $(L(t, x))_{t \geq 0, x \in \mathbb{R}}$.*

(RW2) Il existe $p \geq 1$ tel que pour tout $M > 0$:

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \int_{[-M, M]} \mathbb{E} |L_n(t, x) - L_n(t, [x]_\delta)|^p dx = 0$$

avec $[x]_\delta = \delta[\delta^{-1}x]$ et $L_n(t, x)$ le temps local discret renormalisé

$$L_n(t, x) = n^{-1} a_n N([nt], [a_n x]).$$

Alors, pour tout $m \geq 1$, $\theta_1, \dots, \theta_m \in \mathbb{R}$ et $0 < t_1 < \dots < t_m$, on a la convergence en loi dans L^p :

$$\sum_{i=1}^m \theta_i L_n(t_i, \cdot) \Longrightarrow \sum_{i=1}^m \theta_i L(t_i, \cdot) \quad \text{quand } n \rightarrow \infty.$$

On peut montrer que ce résultat s'applique aux cas particuliers considérés précédemment : marche i.i.d. dans le domaine normal d'attraction d'un Lévy β -stable ou encore marche aléatoire gaussienne fortement corrélée et dans le domaine d'attraction d'un mouvement Brownien fractionnaire (cf. Proposition 2.2 de [D7]).

Critère de convergence de mesures aléatoires associées à la scène

A la scène $\xi = \{\xi_x, x \in \mathbb{Z}\}$, on associe le processus cumulatif $(w_x)_{x \in \mathbb{Z}}$

$$w_x = \begin{cases} \sum_{i=0}^{x-1} \xi_i & \text{si } x > 0 \\ \sum_{i=x}^{-1} \xi_k & \text{si } x < 0 \\ 0 & \text{si } x = 0 \end{cases}, \quad x \in \mathbb{Z},$$

étendu à \mathbb{R} par interpolation linéaire, et le processus cumulatif renormalisé

$$W_h(x) = \gamma_h w_{h^{-1}x}, \quad x \in \mathbb{R},$$

avec $\gamma_h > 0$ constante de normalisation à préciser. On définit également la mesure signée μ_h sur \mathbb{R} par sa densité

$$\frac{d\mu_h}{dx}(x) = \gamma_h h^{-1} \sum_{k \in \mathbb{Z}} \xi_k 1_{[hk, h(k+1))}(x).$$

On a alors de manière formelle, $dW_h(x) = \mu_h(dx)$, autrement dit pour tout $x_1 \leq x_2$,

$$\mu_h([x_1, x_2]) = W_h(x_2) - W_h(x_1).$$

Ainsi, si f est telle que l'intégrale $\int_{\mathbb{R}} f(x) \mu_h(dx)$ ait un sens, on note

$$\int_{\mathbb{R}} f(x) \mu_h(dx) = \int_{\mathbb{R}} f(x) dW_h(x).$$

Lorsque le processus cumulatif renormalisé W_h converge vers une limite W , on s'attend à ce que

$$\int_{\mathbb{R}} f(x) dW_h(x) \Longrightarrow \int_{\mathbb{R}} f(x) dW(x)$$

pour f dans une classe raisonnable de fonctions. Nous montrons que cela est bien vérifié sous les hypothèses adéquates.

Le premier cas considéré est le cas i.i.d. et donne lieu à la condition suivante :

(RS1) la scène ξ est i.i.d. dans le domaine d'attraction normal d'un processus de Lévy β -stable W , avec $\beta \in (1, 2]$.

Proposition 2.2.6 *Sous l'hypothèse (RS1) et avec la normalisation $\gamma_h = h^{\frac{1}{\beta}}$:*

- l'intégrale $\int_{\mathbb{R}} f(x) dW_h(x)$ converge presque sûrement pour tout $f \in L^\beta(\mathbb{R})$;
- si $f_n \rightarrow f$ dans L^β et $h_n \rightarrow 0$ lorsque $n \rightarrow \infty$,

$$\int_{\mathbb{R}} f(x) dW_{h_n}(x) \implies \int_{\mathbb{R}} f(x) dW(x).$$

Ce premier résultat permet de considérer le cas classique des scènes à la Kesten et Spitzer dans le cas $\beta \in (1, 2]$. Le cas $\beta \in (0, 1)$ pourrait être envisagé mais des difficultés techniques apparaissent du fait que L^β n'est alors pas un espace de Banach.

Le second cas considéré est le cas des scènes de carré intégrable et repose sur des méthodes L^2 . Soit \mathcal{E} la classe des fonctions en escalier sur \mathbb{R} de la forme $f = \sum_{i=1}^K \theta_i 1_{]x_i, x_{i+1}]}$ avec $\theta_1, \dots, \theta_K$ et $x_1 < \dots < x_{K+1}$. On peut définir pour $f \in \mathcal{E}$, les variables aléatoires de carré intégrable

$$W_h[f] = \int_{\mathbb{R}} f(x) dW_h(x) = \sum_{i=1}^K \theta_i (W_h(x_{i+1}) - W_h(x_i))$$

et

$$W[f] = \int_{\mathbb{R}} f(x) dW(x) = \sum_{i=1}^K \theta_i (W(x_{i+1}) - W(x_i)).$$

Lorsque le processus limite W est de carré intégrable et non dégénéré (dans le sens où les variables aléatoires $\{W(x); x \in \mathbb{R}\}$ sont linéairement indépendantes dans $L^2(\Omega)$), on définit un produit scalaire sur \mathcal{E} par

$$\langle f_1, f_2 \rangle_W = \text{Cov}(W[f_1], W[f_2]), \quad f_1, f_2 \in \mathcal{E}.$$

On note L_W^2 l'espace de Hilbert obtenue par complétion de \mathcal{E} pour $\langle \cdot, \cdot \rangle_W$. L'isométrie naturelle

$$\begin{cases} \mathcal{E} & \rightarrow L^2(\Omega) \\ f & \mapsto W[f] \end{cases}$$

s'étend en une injection isométrique $W : L_W^2 \rightarrow L^2(\Omega)$. La fonctionnelle obtenue s'interprète comme l'intégrale de f contre dW et on note

$$W[f] = \int_{\mathbb{R}} f(x) dW(x).$$

Notre deuxième jeu d'hypothèses (RS2) portant sur la scène est :

- (RS2.a) la scène ξ est centrée et de carré intégrable ;
 (RS2.b) les distributions fini-dimensionnelles du processus cumulatif renormalisé W_h convergent vers un processus W de carré intégrable non dégénéré ;
 (RS2.c) il existe $C_1 > 0$ tel que pour tout $h > 0$ et $f \in \mathcal{E}$

$$\text{Var}(W_h[f]) \leq C_1 \text{Var}(W[f]);$$

- (RS2.d) on a l'inclusion $L^1 \cap L^2 \subset L^2_W$ et l'injection correspondante est bornée, *i.e.* il existe $C_2 > 0$ tel que pour tout $f \in L^1 \cap L^2$,

$$\|f\|_{L^2_W} \leq C_2 \max(\|f\|_{L^1}, \|f\|_{L^2}).$$

La dernière hypothèse (RS2.d) est un peu technique mais vérifiée dans les principaux exemples considérés. Notre résultat est le suivant.

Proposition 2.2.7 *Sous l'hypothèse (RS2) :*

- l'intégrale $\int_{\mathbb{R}} f(x) dW_h(x)$ converge dans $L^2(\Omega)$ pour tout $f \in L^1 \cap L^2$;
- si $f_n \rightarrow f$ dans $L^1 \cap L^2$ et $h_n \rightarrow 0$ lorsque $n \rightarrow \infty$,

$$\int_{\mathbb{R}} f(x) dW_{h_n}(x) \implies \int_{\mathbb{R}} f(x) dW(x).$$

Ce résultat permet de considérer des scènes non i.i.d. et s'applique aux scènes faiblement dépendantes satisfaisant des conditions de θ -mélange (*cf.* Guillin-Plantard et Prieur [GPP]), ou encore à des scènes fortement dépendantes dans le domaine d'attraction d'un mouvement Brownien fractionnaire et construite par filtrage linéaire à partir d'une séquence i.i.d.

Application au principe d'invariance pour la marche aléatoire en scène aléatoire

Nous expliquons succinctement comment les résultats précédents permettent d'obtenir des principes d'invariance pour la marche aléatoire en scène aléatoire. L'idée principale est la réécriture de $Z_{[nt]}$ en fonction du temps local discret renormalisé $L_n(\cdot, \cdot)$ de la marche S et des mesures μ_h associées à la scène ξ . On a en effet

$$Z_{[nt]} = \sum_{k=0}^{[nt]} \xi_{S_k} = \sum_{x \in \mathbb{Z}} N([nt], x) \xi_x = \frac{n}{a_n \gamma_{a_n^{-1}}} \int_{\mathbb{R}} L_n(t, x) \mu_{a_n^{-1}}(dx)$$

et donc

$$n^{-1} a_n \gamma_{a_n^{-1}} Z_{[nt]} = \int_{\mathbb{R}} L_n(t, x) dW_{a_n^{-1}}(x)$$

avec $a_n^{-1} \rightarrow 0$.

Les Propositions 2.2.5, 2.2.6 et 2.2.7 précédentes permettent alors d'obtenir le Théorème suivant.

Théorème 2.2.8 *Supposons que la marche S et la scène ξ sont indépendantes et satisfont l'une des hypothèses suivantes :*

- ξ vérifie (RS1) pour $\beta \in (1, 2]$ et S vérifie (RW) avec $p = \beta$;
- ξ vérifie (RS2) et S vérifie (RW) avec $p = 2$.

Alors, on a la convergence au sens des distributions fini-dimensionnelles :

$$(n^{-1}a_n\gamma_{a_n^{-1}}Z_{[nt]})_{t \geq 0} \Longrightarrow \left(\int_{\mathbb{R}} L(t, x) dW(x) \right)_{t \geq 0}$$

avec $(W(x))_{x \in \mathbb{R}}$ et $(L(t, x))_{t \geq 0; x \in \mathbb{R}}$ indépendants et définis respectivement dans (RS) et (RW).

Ce résultat permet de retrouver les principaux principes d'invariance pour les marches aléatoires en scènes aléatoires rencontrés dans la littérature. La méthode est robuste et nous permet de traiter dans [D7] le cas de plusieurs marcheurs évoluant dans une même scène et l'asymptotique lorsque le nombre de marcheurs tend vers l'infini.

2.3 Modèles d'agrégation de sources et processus fractionnaires [D13,D17]

Motivations

Les travaux [D13] et [D17] en collaboration avec I.Kaj sont consacrés à l'étude asymptotique de modèles issus des télécommunications et leurs liens avec les processus fractionnaires. Des mesures ont fait apparaître que dans de nombreux réseaux de télécommunication le trafic a des caractéristiques proches de la dépendance à longue portée et de l'autosimilarité (*cf.* Willinger *et al.* [72]). Afin d'explicitier les mécanismes sous-jacents à ce comportement, divers modèles stochastiques de trafic ont été développés où la présence de distributions à queue lourde joue un rôle central - la monographie de Resnick intitulée *Heavy-tail phenomena* [58] fournit un bon aperçu de la variété des modèles et résultats considérés dans la littérature. La plupart de ces modèles représente le trafic comme une agrégation d'un grand nombre de sources indépendantes. Nous nous concentrons ici sur le modèle simple des sources ON/OFF où chaque source alterne entre un état ON actif où elle transmet des données à vitesse unité, et un état OFF inactif où elle demeure silencieuse. Un processus de renouvellement alterné permet de modéliser la succession des états ON et OFF. Les différentes sources sont supposées indépendantes. Il s'agit alors d'évaluer le trafic total, vu comme l'agrégation des trafics créés par chaque source. On s'intéresse au comportement asymptotique du système correctement renormalisé lorsque le nombre de sources tend vers l'infini et lorsque le temps est changé d'échelle. Une compétition s'instaure entre les deux paramètres dirigeant l'asymptotique, à savoir le nombre de sources et l'échelle de temps et différents régimes apparaissent.

Un premier résultat dû à Taqqu, Willinger et Sherman [68] introduit une méthode de double limite pour le modèle ON/OFF. L'ordre des limites successives joue un rôle essentiel. Si l'on considère d'abord la limite infinie pour le nombre de sources puis pour l'échelle de temps, le processus de trafic total renormalisé converge vers un mouvement Brownien fractionnaire, processus Gaussien autosimilaire et à longue mémoire. Mais si les limites sont considérées dans l'ordre inverse, *i.e.* l'échelle de temps puis le nombre de sources tendent vers l'infini, le processus limite est un Lévy stable, processus autosimilaire, à accroissements indépendants et de variance infinie. Dans le travail de Mikosch, Resnick, Rootzén et Stegeman [53], la double limite est remplacée par une limite simple où l'échelle de temps et le nombre de sources tendent simultanément vers l'infini. Deux régimes limites différents apparaissent selon la croissance relative de ces deux paramètres et on retrouve les deux asymptotiques précédentes vers le mouvement Brownien fractionnaire (régime de trafic lourd) ou le Lévy stable (régime de trafic léger).

L'objectif de notre travail est de montrer l'existence d'un régime intermédiaire, où la croissance des deux paramètres s'équilibre et où on obtient un nouveau processus limite baptisé mouvement de Poisson fractionnaire. Ce résultat complète donc la description des asymptotiques pour le modèle ON/OFF.

Il faut mentionner deux autres modèles de sources où ces différents régimes ont été identifiés et où les asymptotiques sont analogues. Le régime intermédiaire pour le modèle de trafic basé sur des sources de type « processus de renouvellement » a été étudié par Gaigalas et Kaj [27]. Le processus de Poisson fractionnaire a alors été identifié et ses propriétés étudiées : rappelons que le processus de Poisson fractionnaire a la même structure de covariance que le mouvement Brownien fractionnaire et qu'il réalise un pont entre le mouvement Brownien fractionnaire et le processus stable de Lévy (*cf.* Gaigalas [26]). Mentionnons également que des résultats similaires ont été établis pour le modèle du type « infinite source Poisson » par Kaj et Taqqu [38]. L'étude de ce dernier modèle basé sur des processus ponctuel de Poisson est grandement facilitée par le calcul pour les intégrales de Poisson.

Mentionnons encore quelques travaux récents sur ces modèles d'agrégation de sources. Mikosch et Samorodnitsky [54] considèrent les limites d'échelles pour des sources très générales, autorisant plus de dépendance temporelle ainsi que des phénomènes de cluster. Ils montrent que l'asymptotique en régime de trafic lourd est plutôt robuste alors que le régime de trafic léger peut donner lieu à de nombreux autres processus limites. Voir aussi Fasen et Samorodnitsky [24] pour un modèle où le mouvement fractionnaire apparaît dans le régime de trafic léger. Pour un modèle de file d'attente fluide alimenté par une source ON/OFF à queue lourde, voir D'Aurio et Samorodnitsky [16]. La problématique de l'estimation statistique du paramètre de longue mémoire pour les processus du type « infinite source Poisson » est abordée par Fay, Roueff et Soulier [25].

Description du modèle ON/OFF et résultats connus

Nous commençons par introduire le modèle ON-OFF. Les notations sont celles de Mikosch *et al.* [53] dont nous rappelons les résultats.

Soit X_1, X_2, \dots des variables aléatoires i.i.d. positives de loi F_{on} et moyenne μ_{on} et Y_1, Y_2, \dots v.a.i.i.d. positives de loi F_{off} et moyenne μ_{off} . On suppose les deux suites indépendantes et les distributions sont à variations régulières : les fonctions de queues vérifient

$$\bar{F}_{\text{on}}(x) = x^{-\alpha_{\text{on}}} L_{\text{on}}(x) \quad \text{et} \quad \bar{F}_{\text{off}}(x) = x^{-\alpha_{\text{off}}} L_{\text{off}}(x), \quad x \rightarrow \infty, \quad (2.1)$$

avec

$$1 < \alpha_{\text{on}} < \alpha_{\text{off}} < 2 \quad (2.2)$$

et $L_{\text{on}}, L_{\text{off}}$ à variations lentes à l'infini. Les distributions ont une moyenne finie mais une variance infinie. La condition initiale (X_0, Y_0) est choisie de sorte à ce que le système soit en régime stationnaire (nous ne donnons pas la forme explicite de la loi initiale ici).

Le modèle ON/OFF est le modèle de renouvellement alterné généré par l'alternance des périodes ON et OFF de longueurs données par les X_i (resp. Y_i). Au temps $t = 0$, le système démarre dans l'état ON pendant une durée X_0 , puis passe à l'état OFF pendant une durée

Y_0 , puis passe à l'état ON pour une durée $X_1 \dots$

On pose $Z_i = X_i + Y_i$ la longueur du i -ème cycle ON/OFF. Les variables Z_i sont indépendantes, et pour $i \geq 1$, de loi $F = F_{\text{on}} * F_{\text{off}}$ et moyenne $\mu = \mu_{\text{on}} + \mu_{\text{off}}$. On introduit la suite de renouvellement associée $(T_n)_{n \geq 1}$ par

$$T_n = \sum_{i=0}^n Z_i.$$

Le modèle ON-OFF est représenté par le processus « input »

$$I(t) = 1_{[0, X_0)}(t) + \sum_{n \geq 0} 1_{[T_n, T_n + X_{n+1})}(t), \quad t \geq 0.$$

La source est active lorsque $I(t) = 1$ et inactive sinon. Le choix de la loi initiale (X_0, Y_0) assure que I est un processus stationnaire et on montre que $\mathbb{P}(I(t) = 1) = \mu_{\text{on}}/\mu$.

Le trafic cumulé associé ou « workload »

$$W(t) = \int_0^t I(s) ds, \quad t \geq 0$$

est un processus à accroissements stationnaires de moyenne $\mathbb{E}[W(t)] = t\mu_{\text{on}}/\mu$.

Le nombre de cycles ON/OFF jusqu'au temps t est donné par le processus de renouvellement

$$N(t) = \sum_{n \geq 0} 1_{[0, t]}(T_n), \quad t \geq 0.$$

Dans la figure 2.1, nous présentons une réalisation de ces différents processus afin d'illustrer les différentes notations .

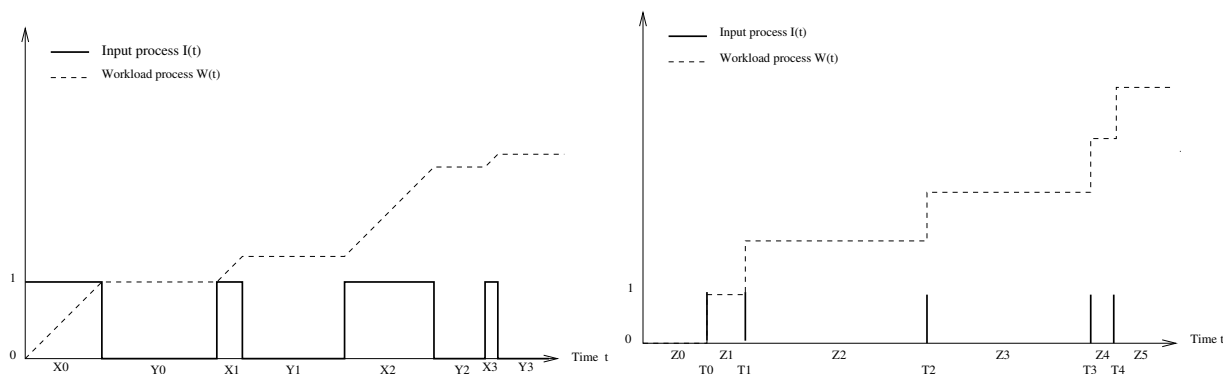


FIGURE 2.1 – Réalisation de l' *input process* $I(t)$ et du *workload process* $W(t)$ correspondant pour le modèle ON/OFF (à gauche) et de renouvellement (à droite).

Pour définir le modèle à plusieurs sources, on considère des copies i.i.d. de la construction précédente : soit $(I^j, W^j, N^j)_{j \geq 1}$ des copies i.i.d. de (I, W, N) , l'indice j représentant

la j -ème source. Le trafic cumulé associé à un réseau à m sources est

$$W_m(t) = \sum_{j=1}^m W^j(t), \quad t \geq 0, \quad m \geq 1,$$

et le processus de comptage associé

$$N_m(t) = \sum_{j=1}^m N^j(t), \quad t \geq 0, \quad m \geq 1.$$

Nous nous intéressons aux propriétés asymptotiques du modèle lorsque le nombre de sources m est grand et l'échelle de temps $a > 0$ est grande. On considère ainsi le processus centré et renormalisé

$$\frac{W_m(at) - mat\mu_{\text{on}}/\mu}{b(a, m)} = \frac{1}{b(a, m)} \sum_{j=1}^m \int_0^{at} (I^j(s) - \frac{\mu_{\text{on}}}{\mu}) ds, \quad t \geq 0,$$

où la normalisation $b(a, m)$ sera précisée dans la suite.

Résultats asymptotiques connus

L'asymptotique est considéré lorsque m et a tendent simultanément vers l'infini. La croissance relative de m et a a un grand impact sur la limite et nous considérons les régimes suivants :

- trafic lourd ou « fast connection rate »

$$m \gg \frac{a^{\alpha_{\text{on}}-1}}{L_{\text{on}}(a)}; \quad (\text{FCR})$$

- trafic léger ou « slow connection rate »

$$m \ll \frac{a^{\alpha_{\text{on}}-1}}{L_{\text{on}}(a)}; \quad (\text{SCR})$$

- trafic intermédiaire ou « intermediate connection rate »

$$m \sim \mu c^{\alpha_{\text{on}}-1} \frac{a^{\alpha_{\text{on}}-1}}{L_{\text{on}}(a)}, \quad 0 < c < \infty. \quad (\text{ICR})$$

L'asymptotique du modèle ON-OFF pour les régimes (FCR) et (SCR) est précisée dans Mikosch *et al.* [53]. Sous les conditions (2.1) et (2.2), on a les asymptotiques suivantes :

- dans le régime (FCR) avec la normalisation $b(a, m) = (a^{3-\alpha_{\text{on}}} L_{\text{on}}(a) m)^{1/2}$, on a la convergence en loi dans l'espace des fonctions continues sur \mathbb{R}^+

$$\frac{W_m(at) - mat\mu_{\text{on}}/\mu}{b(a, m)} \implies \sigma_{\alpha_{\text{on}}} \frac{\mu_{\text{on}}}{\mu^{3/2}} B_H(t), \quad t \geq 0$$

avec

$$\sigma_{\alpha_{\text{on}}}^2 = \frac{2}{(\alpha_{\text{on}} - 1)(2 - \alpha_{\text{on}})(3 - \alpha_{\text{on}})}$$

et $B_H(t)$ mouvement Brownien fractionnaire standard d'indice $H = (3 - \alpha_{\text{on}})/2$.

- dans le régime (SCR) avec la normalisation

$$b(a, m) = \inf\{x \geq 0 : \bar{F}_{\text{on}}(x) \leq 1/am\},$$

on a la convergence des lois fini-dimensionnelles

$$\frac{W_m(at) - mat\mu_{\text{on}}/\mu}{b(a, m)} \xrightarrow{fdd} \sigma_0 \frac{\mu_{\text{off}}}{\mu^{1+1/\alpha_{\text{on}}}} X_{\alpha_{\text{on}},1,1}(t), \quad t \geq 0,$$

avec $X_{\alpha_{\text{on}},1,1}(t)$ processus de Lévy α_{on} -stable totalement asymétrique, *i.e.* tel que

$$X_{\alpha_{\text{on}},1,1}(1) \sim S_{\alpha_{\text{on}}}(1, 1, 0),$$

et

$$\sigma_0^{\alpha_{\text{on}}} = \frac{\Gamma(2 - \alpha_{\text{on}}) \cos(\pi\alpha_{\text{on}}/2)}{1 - \alpha_{\text{on}}}.$$

Le régime intermédiaire pour le modèle ON/OFF [D8]

Notre principal résultat donne l'asymptotique dans le régime intermédiaire (ICR).

Théorème 2.3.1 *Dans le régime (ICR) avec normalisation $b(a, m) = a$, on a la convergence en loi dans l'espace des fonctions continues sur \mathbb{R}^+*

$$\frac{W_m(at) - mat\mu_{\text{on}}/\mu}{a} \implies \sigma_{\alpha_{\text{on}}} \frac{\mu_{\text{off}}}{\mu} c P_H(t/c)$$

avec $P_H(t)$ processus de Poisson fractionnaire d'indice de Hurst $H = (3 - \alpha_{\text{on}})/2$ défini par

$$P_H(t) = \frac{1}{\sigma_{\alpha_{\text{on}}}} \int_{\mathbb{R} \times \mathbb{R}^+} \int_0^t 1_{[x, x+u]}(y) dy (N(dx, du) - dx \alpha_{\text{on}} u^{-\alpha_{\text{on}}-1} du), \quad (2.3)$$

où $N(dx, du)$ est une mesure aléatoire de Poisson sur $\mathbb{R} \times \mathbb{R}^+$ d'intensité $dx \alpha_{\text{on}} u^{-\alpha_{\text{on}}-1} du$.

La preuve de ce théorème repose sur un couplage entre le modèle ON-OFF et le modèle de renouvellement pur. L'heuristique est la suivante : le trafic total pour une source se décompose en

$$W(t) = X_0 \wedge t + \sum_{i=1}^{N(t)} X_i - (T_{N(t)-1} + X_{N(t)} - t)_+$$

ou encore

$$t - W(t) = Y_0 \wedge t + \sum_{i=1}^{N(t)} Y_i - (T_{N(t)} - t) \wedge Y_{N(t)}.$$

Ainsi

$$\begin{aligned} W(t) - \frac{\mu_{\text{on}}}{\mu}t &= -(t - W(t)) + \frac{\mu_{\text{off}}}{\mu}t \\ &= -\mu_{\text{off}}(N(t) - t/\mu) - \sum_{i=1}^{N(t)} (Y_i - \mu_{\text{off}}) + R(t) \end{aligned}$$

avec

$$R(t) = (T_{N(t)} - t) \wedge Y_{N(t)} - Y_0 \wedge t.$$

On a donc, pour m sources,

$$W_m(t) - \frac{\mu_{\text{on}}}{\mu}mt = -\mu_{\text{off}}(N_m(t) - mt/\mu) - \sum_{j=1}^m \sum_{i=1}^{N^j(t)} (Y_i^j - \mu_{\text{off}}) + \sum_{j=1}^m R^j(t) \quad (2.4)$$

avec des notations évidentes. La preuve consiste à montrer que dans le régime (ICR), les termes

$$\frac{1}{a} \sum_{j=1}^m \sum_{i=1}^{N^j(at)} (Y_i^j - \mu_{\text{off}}), \quad \frac{1}{a} \sum_{j=1}^m R^j(at)$$

disparaissent à la limite, si bien que asymptotiquement

$$\frac{W_m(at) - \mu_{\text{on}}mat/\mu}{a} \sim -\mu_{\text{off}} \frac{N_m(at) - mat/\mu}{a}. \quad (2.5)$$

On fait ensuite appel à un résultat de Gaigalas et Kaj [27] donnant le régime intermédiaire pour le modèle de renouvellement pur : sous la condition (ICR) avec normalisation $b(a, m) = a$, on a la convergence faible dans l'espace de Skorokhod

$$\frac{N_m(at) - mat/\mu}{a} \Longrightarrow -\frac{1}{\mu} \sigma_{\alpha_{\text{on}}} c P_H(t/c). \quad (2.6)$$

La comparaison (2.5) permet alors de déduire le Théorème 2.3.1 des résultats de Gaigalas et Kaj [27]. Il faut bien sur rendre la comparaison (2.5) rigoureuse et également traiter les questions de tension.

Moments d'un processus de renouvellement à queue lourde [D17]

Dans le travail [D17], nous proposons une nouvelle approche pour l'étude du modèle de renouvellement, qui comme nous venons de le voir, est à la base de nos résultats pour le modèle ON-OFF. Nous obtenons en effet une nouvelle preuve du résultat (2.6) de Gaigalas et Kaj [27]. Rappelons les notations du modèle de renouvellement.

Soit $(T_n)_{n \geq 0}$ une suite de renouvellement stationnaire basée sur la loi F pour le temps d'attente entre deux arrivées successives. On suppose que F est à queue lourde :

$$\bar{F}(x) = x^{-1-\beta} \ell(x), \quad 1 < \alpha < 2.$$

On note N le processus de comptage associé

$$N(t) = \sum_{n \geq 0} 1_{\{T_n \leq t\}}, \quad t \geq 0$$

ainsi que ξ la mesure aléatoire ponctuelle

$$\xi = \sum_{n \geq 0} \delta_{T_n}.$$

Notons que $N(t) = \xi([0, t])$. On se donne $\{N^i, i \geq 1\}$ des copies i.i.d. de N . On se place sous la condition de régime intermédiaire

$$m \sim \mu c^\beta \frac{a^\beta}{\ell(a)}, \quad 0 < c < \infty. \quad (\text{ICR})$$

Il s'agit de montrer sous cette condition la convergence fini-dimensionnelle

$$\frac{\sum_{i=1}^m N^i(at) - mat/\mu}{a} \xrightarrow{f.d.d.} -\frac{1}{\mu} \sigma c P_H(t/c), \quad t \geq 0, \quad (2.7)$$

avec $\sigma^2 = \frac{2}{\beta(1-\beta)(2-\beta)}$.

On utilise la méthode des cumulants et on note $c_k(X)$ le k -ème cumulant d'une variable aléatoire X . On montre que pour $k \geq 2$,

$$c_k(P_H(t)) = \frac{1}{\beta \sigma^k} \int_{[0, t]^k} \frac{dx_1 \cdots dx_k}{|\max(x_i) - \min(x_i)|^\beta}$$

et plus généralement, pour t_1, \dots, t_p et $\theta_1, \dots, \theta_p$

$$c_k\left(\sum_{j=1}^p \theta_j P_H(t_j)\right) = \frac{1}{\beta \sigma^k} \int f(x_1) \cdots f(x_k) \frac{dx_1 \cdots dx_k}{|\max(x_i) - \min(x_i)|^\beta} \quad (2.8)$$

avec $f = \sum_{j=1}^p \theta_j 1_{[0, t_j]}$. D'un autre côté, les moments associés au processus de comptage N vérifient, pour $k \geq 2$,

$$\mathbb{E}\left[\left(N(t) - t/\mu\right)^k\right] = \widetilde{M}_k([0, t]^k)$$

avec \widetilde{M}_k le k -ème moment central de la mesure ponctuelle ξ défini par

$$\widetilde{M}_k(A_1 \times \cdots \times A_k) = \mathbb{E}\left[\prod_{i=1}^k (\xi(A_i) - \lambda(A_i)/\mu)\right], \quad A_1, \dots, A_k \subset [0, +\infty),$$

avec λ mesure de Lebesgue sur $[0, +\infty)$. Plus généralement, pour t_1, \dots, t_p et $\theta_1, \dots, \theta_p$,

$$\mathbb{E}\left[\left(\sum_{j=1}^p \theta_j (N(t_j) - t_j/\mu)\right)^k\right] = \int f(x_1) \cdots f(x_k) \widetilde{M}_k(dx_1 \cdots dx_k) \quad (2.9)$$

avec $f = \sum_{j=1}^p \theta_j 1_{[0, t_j]}$.

Notre approche est basée sur l'asymptotique de ces mesures de moments centraux. Pour $k \geq 1$, on définit

$$\mathcal{F}_k = \left\{ f : [0, +\infty)^k \rightarrow \mathbb{R}; \text{ support compact, } \frac{\partial^k f}{\partial x_1 \cdots \partial x_k} \text{ exists and continuous} \right\}.$$

On dit qu'une suite de mesures m_n sur $[0, +\infty)^k$ \mathcal{F}_k -converge vers m (notation $m_n \xrightarrow{\mathcal{F}_k} m$) lorsque $\int f dm_n \rightarrow \int f dm$ pour tout $f \in \mathcal{F}_k$. Notre résultat est le suivant :

Théorème 2.3.2 *Pour $k \geq 2$ et lorsque $a \rightarrow \infty$,*

$$\frac{\widetilde{M}_k(a \cdot)}{a^{k-\beta} \ell(a)} \xrightarrow{\mathcal{F}_k} \frac{(-1)^k}{\beta \mu^{k+1}} \widetilde{P}_k(\cdot)$$

où les mesures limites \widetilde{P}_k sont définies par

$$\widetilde{P}_k(dx_1 \cdots dx_k) = \frac{dx_1 \cdots dx_k}{|\max(x_i) - \min(x_i)|^\beta}.$$

Au vu de ce résultat, la comparaison des équations (2.8) et (2.9) rend plausible la convergence (2.7). Malheureusement, la fonction f combinaison linéaire d'indicatrice n'a pas de bonnes conditions de régularité ce qui nous empêche de déduire (2.7) du Théorème 2.3.2. On montre cependant une convergence similaire pour une classe de fonction plus régulières grâce à la méthode des cumulants et au Théorème 2.3.2. La méthode s'avère suffisamment robuste pour s'adapter au cas non stationnaire et également à un modèle continu analogue au renouvellement discret et construit grâce à la mesure d'occupation d'un subordonateur de Lévy (*cf.* [D17] pour plus de détails).

2.4 Théorèmes limites pour les modèles de boules aléatoires [D8,D12]

Motivations

Nous présentons ici les principaux résultats des articles [D8] et [D12] en collaboration avec J.-C. Breton. Nous considérons le modèle de boules aléatoires pondérées (ou *weighted random ball model*) et étudions ses limites d'échelles. Nous montrons que l'asymptotique est particulièrement riche et que trois régimes différents apparaissent :

- un champ aléatoire invariant par translation et isotrope, auto-similaire sur \mathbb{R}^d ;
- un champ aléatoire Poissonien fractionnaire ;
- un champ aléatoire stable avec indépendance.

Le modèle de boules aléatoires pondérées est construit de la manière suivante : les centres sont distribués aléatoirement selon un processus ponctuel de Poisson, et chaque centre x est étiqueté avec un rayon r et un poids m aléatoire. On considère ensuite pour chaque point y de l'espace la somme des poids des boules qui recouvrent y . Les intersections possibles de boules donnent lieu à une structure de dépendance non triviale. Cette dépendance peut persister lors des limites d'échelles dans le cas où le rayon des boules suit une loi à queue lourde.

Cette construction géométrique simple a trouvé de nombreuses applications et s'avère pertinente dans la modélisation de divers phénomènes. Kaj [36] considère des modèles analogues pour des modèles simplifiés de réseaux sans fil où une collection de stations est disposée aléatoirement dans l'espace et partage un canal de transmission. Dans ce contexte, le point x représente la position de la station, r sa portée et m sa puissance. Le champ aléatoire permet alors de mesurer la puissance reçue en chaque point de l'espace. Biermé et Estrade [8] ont considéré des modèles similaires en dimension $d = 2$ dans le contexte de l'imagerie médicale (dans ce cas, le poids représente une intensité de couleur en niveau de gris) et en dimension $d = 3$ pour la modélisation de milieux poreux hétérogène (ici le poids représente la densité du milieu). Les propriétés microscopiques du milieu sont étudiées grâce à un changement d'échelle permettant de zoomer sur de petites régions de l'espace (*zoom-in*). D'un point de vue plus mathématique, Kaj *et al.* [37] étudient des modèles similaires de type germe/grain et leur limite d'échelles et ont montré que le champ vu de loin grâce à une procédure de *zoom-out* exhibe des propriétés de dépendance à longue portée lorsque la loi du volume des grains est à queue lourde. Un travail commun de Biermé, Estrade et Kaj [9] a permis d'unifier les deux approches en traitant de manière similaire les régimes *zoom-in* [8] et *zoom-out* [37].

Dans ces différents travaux, il n'y a pas de poids aléatoire, ce qui revient à prendre $m \equiv 1$. Les modèles à poids ont reçu moins d'attention, à l'exception de la dimension $d = 1$ où l'on retrouve un modèle classique dans le cadre des télécom connu sous le nom *infinite source Poisson model*. Dans ce cadre, chaque « boule » correspond à une session de transmission, le rayon correspond à l'intensité de la transmission et le poids à son débit. Kaj et Taqqu

[38] ont étudié les limites d'échelles de ce modèle unidimensionnel et identifié les trois régimes avec des limites de type Gaussiennes, Poissonniennes ou stables.

Notre contribution consiste à étudier de manière systématique les modèles de boules aléatoires avec poids en dimension $d \geq 1$. Nous identifions trois régimes et les limites correspondantes. En particulier, lorsque les poids sont à queues lourdes, nous obtenons dans de nouveaux champs généralisés stables, invariants par translation et rotation et auto-similaires. La dimension supérieure pose également de nouveaux problèmes liés à la convergence fonctionnelle et la tension.

Description du modèle

Nous considérons une mesure aléatoire de Poisson d'intensité $N_\lambda(dx, dr, dm)$ sur $\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}$ d'intensité

$$n(dx, dr, dm) = \lambda dx F(dr) G(dm)$$

avec $\lambda > 0$, F mesure sur \mathbb{R}^+ et G loi de probabilité sur \mathbb{R} . A chaque triplet (x, r, m) est associée la boule $B(x, r)$ dans \mathbb{R}^d de centre x et rayon r ainsi que le poids m . On note $\|\cdot\|$ la distance euclidienne sur \mathbb{R}^d et $B(x, r) = \{y \in \mathbb{R}^d : \|y - x\| < r\}$.

On suppose que la mesure F est absolument continue, *i.e.* $F(dr) = f(r)dr$ avec

$$\int_{\mathbb{R}^+} r^d F(dr) < +\infty$$

et telle que pour $\epsilon = +1$ ou $\epsilon = -1$,

$$f(r) \sim_{r \rightarrow 0^\epsilon} C_\beta r^{-1-\beta}$$

avec la convention $0^{+1} = 0$ and $0^{-1} = +\infty$. Comme nous le verrons par la suite, cette convention non standard nous permettra d'unifier deux cas, le cas *zoom-in* $\epsilon = +1$ et le cas *zoom-out* $\epsilon = -1$. La conjonction de ces deux hypothèses sur F impose que $\beta < d$ lorsque $\epsilon = +1$, et $\beta > d$ lorsque $\epsilon = -1$.

Par ailleurs, on suppose que G est dans le domaine normal (normal) d'attraction d'une loi α -stable $S_\alpha(\sigma, b, 0)$, avec $\alpha \in (1, 2]$, $\sigma > 0$ et $b \in [-1, 1]$. Dans ce cas, le comportement en 0 de la fonction caractéristique est connu et sera utilisé dans les asymptotiques.

Soit \mathcal{M} l'ensemble des mesures signées sur \mathbb{R}^d de variation totale $|\mu|$ finie. On considère le champ aléatoire M

$$M(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}} m \mu(B(x, r)) N_\lambda(dx, dr, dm)$$

indexé par les mesures signées $\mu \in \mathcal{M}$. Remarquons que lorsque μ est la masse de dirac en $y \in \mathbb{R}^d$, $\mu = \delta_y$, alors $M(\delta_y)$ est le poids total en y du système de boules pondérées, c'est à

dire la somme algébrique des poids des boules contenant le point y . L'intégrale de Poisson $M(\mu)$ est bien définie et d'espérance

$$\mathbb{E}[M(\mu)] = \lambda |B(0, 1)| \int_{\mathbb{R}} mG(dm) \int_{\mathbb{R}^+} r^d F(dr) \mu(\mathbb{R}^d).$$

Différentes asymptotiques pour le modèle de boules aléatoires [D8]

Nous nous intéressons aux fluctuations de $M(\mu)$ au niveau macroscopique ou microscopique. Pour cela, on considère le changement d'échelle en espace $x \mapsto \rho x$, et on considère l'asymptotique lorsque $\rho \rightarrow \infty$ (point de vue microscopique, ou *zoom-in*) ou lorsque $\rho \rightarrow 0$ (point de vue macroscopique ou *zoom-out*). Lors de ce changement d'échelle, la mesure régissant le rayon des boules devient $F_\rho(dr) = f(r/\rho)dr/\rho$, mesure image de $F(dr)$ par le changement d'échelle $r \mapsto \rho r$. Pour obtenir des asymptotiques non triviales, nous considérons également une intensité $\lambda(\rho)$ dépendant de ρ . Après changement d'échelle, le modèle devient

$$M_\rho(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}} m\mu(B(x, r)) N_{\lambda(\rho), \rho}(dx, dr, dm)$$

avec $N_{\lambda(\rho), \rho}(dx, dr, dm)$ mesure aléatoire de Poisson d'intensité $\lambda(\rho)dx F_\rho(dr)G(dm)$. L'asymptotique $\rho \rightarrow 0$ correspond au cas *zoom-out* et requiert une hypothèse sur f au voisinage de $+\infty$ (cas $\epsilon = -1$). De manière symétrique, l'asymptotique $\rho \rightarrow +\infty$ correspond au cas *zoom-in* et requiert une hypothèse sur f au voisinage de $+\infty$ (cas $\epsilon = +1$).

Les résultats que nous obtenons sont des résultats de convergence fini-dimensionnelle du processus $M_\rho(\mu)$ pour μ dans des sous-espaces stricts de \mathcal{M} . L'espace $\mathcal{M}_{\alpha, \beta} \subset \mathcal{M}$ suivant joue un rôle important : pour $1 < \alpha \leq 2$ et $\beta > 0$, on note $\mathcal{M}_{\alpha, \beta}$ l'ensemble des mesures $\mu \in \mathcal{M}$ vérifiant

$$\forall r > 0, \quad \int_{\mathbb{R}^d} |\mu(B(x, r))|^\alpha dx \leq C \min(r^p, r^q)$$

pour des constantes $C > 0$ et $0 < p < \beta < q$. Les Propositions 2.2 et 2.3 de [D8] donnent quelques propriétés élémentaires de ces espaces que nous résumons ici brièvement :

- pour tout $\mu \in \mathcal{M}_{\alpha, \beta}$,

$$\int_{\mathbb{R}^d \times \mathbb{R}^+} |\mu(B(x, r))|^\alpha r^{-\beta-1} dx dr < \infty$$

- $\mathcal{M}_{\alpha, \beta}$ est un sous-espace vectoriel de \mathcal{M} , stable par translation, rotation et dilatation ;
- si $\alpha \leq \alpha'$, $\mathcal{M}_{\alpha, \beta} \subseteq \mathcal{M}_{\alpha', \beta}$;
- si $\beta \geq d$, alors toute mesure $\mu \in \mathcal{M}_{\alpha, \beta}$ est diffuse, et de plus si $d < \beta < \alpha d$, $\mathcal{M}_{\alpha, \beta}$ contient $L^1(\mathbb{R}^d) \cap L^\alpha(\mathbb{R}^d)$;
- si $\beta \leq d$, alors toute mesure $\mu \in \mathcal{M}_{\alpha, \beta}$ est centrée, et de plus si $d - 1 < \beta < d$, $\mathcal{M}_{\alpha, \beta}$ contient les mesures centrées $\mu \in L^1(\mathbb{R}^d)$ et les mesures centrées μ à support fini.

Nous présentons maintenant les principaux résultats obtenus : il s'agit de préciser le comportement asymptotique du champ aléatoire centré et renormalisé

$$\frac{M_\rho(\mu) - \mathbb{E}[M_\rho(\mu)]}{n(\rho)}$$

lorsque $\rho \rightarrow 0$ ou $\rho \rightarrow +\infty$ et μ est restreint à un sous-espace adéquat. Comme nous allons le voir, différents régimes se présentent selon le comportement de $\lambda(\rho)$, c'est-à-dire de l'intensité pour le nombre de boules. Ces régimes donnent lieu à des renormalisations $n(\rho)$ et à des champs limites variés. On utilise la convention $0^{-\epsilon} = 0$ si $\epsilon = -1$ et $0^{-\epsilon} = +\infty$ si $\epsilon = +1$.

Théorème 2.4.1 • Supposons $\rho^\beta \lambda(\rho) \rightarrow +\infty$ lorsque $\rho \rightarrow 0^{-\epsilon}$. Alors avec la normalisation $n(\rho) = \lambda(\rho)^{1/\alpha} \rho^{\beta/\alpha}$,

$$\frac{M_\rho(\cdot) - \mathbb{E}[M_\rho(\cdot)]}{n(\rho)} \xrightarrow{f.d.d.} Z_\alpha(\cdot) \quad \text{sur } \mathcal{M}_{\alpha,\beta} \text{ quand } \rho \rightarrow 0^{-\epsilon}$$

où $Z_\alpha(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^+} \mu(B(x, r)) M_\alpha(dr, dx)$ est une intégrale stable par rapport à la mesure α -stable M_α de mesure de contrôle $\sigma^\alpha C_\beta r^{-1-\beta} dr dx$ et biais constant b .

- Supposons $\lambda(\rho) \rho^\beta \rightarrow a > 0$ lorsque $\rho \rightarrow 0^{-\epsilon}$. Alors avec la normalisation $n(\rho) = 1$,

$$M_\rho(\cdot) - \mathbb{E}[M_\rho(\cdot)] \xrightarrow{f.d.d.} J_a(\cdot) \quad \text{sur } \mathcal{M}_{\alpha,\beta} \text{ quand } \rho \rightarrow 0^{-\epsilon}$$

où J_a est l'intégrale de Poisson compensée

$$J_a(\mu) = \int_{\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^+} m \mu(B(x, r)) \tilde{N}_{\beta,a}(dx, dr, dm)$$

par rapport à la mesure de Poisson compensée $\tilde{N}_{\beta,a}$ d'intensité $a C_\beta r^{-\beta-1} dx dr G(dm)$.

- Supposons $\epsilon = -1$ (cas zoom-out), $d < \beta < \alpha d$. Supposons de plus $\lambda(\rho) \rightarrow +\infty$ et $\lambda(\rho) \rho^\beta \rightarrow 0$ lorsque $\rho \rightarrow 0$. Alors avec la normalisation $n(\rho) = \lambda(\rho)^{d/\beta} \rho^d$,

$$\frac{M_\rho(\cdot) - \mathbb{E}[M_\rho(\cdot)]}{n_2(\rho)} \xrightarrow{f.d.d.} \tilde{Z}_\gamma(\cdot) \quad \text{sur } L^1(\mathbb{R}^d) \cap L^\alpha(\mathbb{R}^d) \text{ quand } \rho \rightarrow 0$$

où $\gamma = \beta/d$ et $\tilde{Z}_\gamma(\phi(x) dx) = \int_{\mathbb{R}^d} \phi(x) \tilde{M}_\gamma(dx)$ est une intégrale stable par rapport à la mesure γ -stable M_γ de mesure de contrôle $\sigma_\gamma^\gamma dx$ et biais b_γ donnés par

$$\sigma_\gamma^\gamma = \frac{c_d^\gamma C_\beta}{d} \int_{\mathbb{R}_+} \frac{1 - \cos(r)}{r^{1+\gamma}} dr \int_{\mathbb{R}} |m|^\gamma G(dm)$$

et

$$b_\gamma = - \frac{\int_{\mathbb{R}} \text{sign}(m) |m|^\gamma G(dm)}{\int_{\mathbb{R}} |m|^\gamma G(dm)}.$$

Convergence fonctionnelle et questions de tension [D12]

Au vu de ces résultats de convergence fini-dimensionnelle, la question de la tension se pose naturellement. Nous apportons des réponses dans l'article [D12]. Motivés par des modèles en télécommunication, nous considérons un modèle légèrement différent en introduisant une fonction d'atténuation h . Pour μ mesure signée de variation totale finie sur \mathbb{R}^d , on pose

$$M(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}} m\mu(\tau_{x,r}h) N_\lambda(dx, dr, dm)$$

avec $\tau_{x,r}h(\cdot) = h(\frac{\cdot-x}{r})$ et $\mu[f] = \int_{\mathbb{R}^d} f(y)\mu(dy)$ pour $f \in L^1(\mathbb{R}^d, \mu)$. D'un point de vue physique, il est naturel de supposer que la fonction d'atténuation est positive, continue, décroissante radialement, et telle que $h(0) = 1$ et $\lim_{x \rightarrow \infty} h(x) = 0$. D'un point de vue mathématique, il s'avère suffisant pour notre propos que h soit continue presque partout et que

$$h^*(\mathbf{x}) := \sup\{|h(r\mathbf{x})| : r \geq 1\} \in L^1(\mathbb{R}^d) \cap L^\alpha(\mathbb{R}^d).$$

Nous retrouvons le modèle des boules aléatoires lorsque h est l'indicatrice de la boule $B(0,1)$, ce qui dans le modèle telecom, s'interprète comme une transmission parfaite à courte distance puis transmission nulle à distance plus longue. Avec les mêmes notations que précédemment, le modèle changé d'échelle s'écrit

$$M_\rho(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}} m\mu(\tau_{x,r}h) N_{\lambda(\rho),\rho}(dx, dr, dm).$$

Nous nous intéressons à son asymptotique dans le cas *zoom-out* $\rho \rightarrow 0$.

Pour traiter les questions de tension, nous nous plaçons dans un cadre fini-dimensionnel et considérons une famille paramétrique de mesures $(\mu_t)_{t \in \mathbb{R}^p}$ telles que

$$\mu_t(dx) = \phi_t(x)dx \quad \text{avec } \phi_t \in L^1(\mathbb{R}^d) \cap L^\alpha(\mathbb{R}^d).$$

Pour $s, t \in \mathbb{R}^p$ tels que $s_i \leq t_i, 1 \leq i \leq p$, on définit le pavé $[s, t] = \prod_{i=1}^p [s_i, t_i]$. On définit l'accroissement de ϕ_t sur le bloc $[s, t]$ par

$$\phi_{[s,t]} := \sum_{\epsilon} (-1)^{p - \sum_{i=1}^p \epsilon_i} \phi_{s_1 + \epsilon_1(t_1 - s_1), \dots, s_p + \epsilon_p(t_p - s_p)}$$

où la somme parcourt l'ensemble des $\epsilon \in \{0, 1\}^p$ tels que $\epsilon_i = 0$ si $s_i = t_i$.

La condition principale est la suivante : pour $\gamma \geq 1$, on dit que la famille $(\phi_t)_{t \in \mathbb{R}^p}$ satisfait la propriété (P_γ) si pour tout $T > 0$, il existe une constante $C_T > 0$ tel que pour tout $[s, t] \subset [-T, T]^p$,

$$\|\phi_{[s,t]}\|_\gamma^\gamma \leq C_T \prod_{i: s_i < t_i} |t_i - s_i|. \quad (P_\gamma)$$

Ici $\|\cdot\|_\gamma$ est la norme $L^\gamma(\mathbb{R}^p)$. On donne dans [D12] des exemples montrant que cette condition (P_γ) est assez naturelle et vérifiée par de larges classes de fonctions.

On peut alors renforcer le Théorème 2.4.1 et obtenir des convergences fonctionnelles plutôt que de simples convergences fini-dimensionnelles.

Théorème 2.4.2 *Soit $\mu_t(dy) = \phi_t(y)dy$, $t \in \mathbb{R}^p$, une famille paramétrique de mesures dans $L^1(\mathbb{R}^d) \cap L^\alpha(\mathbb{R}^d)$ satisfaisant les conditions (P_1) et (P_α) .*

1. *Si $\lambda(\rho)\rho^\beta \rightarrow +\infty$, alors avec $n(\rho) = (\lambda(\rho)\rho^\beta)^{1/\alpha}$, on a la convergence en loi dans l'espace des fonctions continues sur \mathbb{R}^p :*

$$\widetilde{M}_\rho(\mu_t) \Longrightarrow \int_{\mathbb{R}^d \times \mathbb{R}^+} \mu_t[\tau_{x,r}h] M_\alpha(dx, dr), \quad t \in \mathbb{R}^p,$$

avec M_α comme dans le Théorème 2.4.1.

2. *Si $\lambda(\rho)\rho^\beta \rightarrow a$, alors avec $n(\rho) = 1$, on a la convergence en loi dans l'espace des fonctions continues sur \mathbb{R}^p :*

$$\widetilde{M}_\rho(\mu_t) \Longrightarrow \int_{\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}} m\mu_t[\tau_{x,r}h] \widetilde{N}_{\beta,a}(dx, dr, dm), \quad t \in \mathbb{R}^p,$$

avec $\widetilde{N}_{\beta,a}$ comme dans le Théorème 2.4.1.

La technique utilisée repose sur le critère de tension de Kolmogorov et des majorations assez fine des moments fractionnaires de $\widetilde{M}_\rho(\mu)$ (cf. [D12] Proposition 3.1). Ces estimées de moment permettent aussi d'obtenir des résultats de régularité Hölder des trajectoires (cf. [D12] Proposition 2.5).

Nous étudions enfin des résultats de tension dans l'espace des distributions. On notera $\mathcal{D}(\mathbb{R}^d)$ l'espace des fonctions lisses à support compact et $\mathcal{D}'(\mathbb{R}^d)$ son dual, *i.e.* l'espace des distributions. On montre que pour tout $\rho > 0$, M_ρ induit une distribution aléatoire, *i.e.* la forme linéaire

$$M_\rho : \begin{cases} \mathcal{D}(\mathbb{R}^d) & \rightarrow \mathbb{R} \\ \phi & \mapsto M_\rho(\phi(y)dy) \end{cases}$$

est presque sûrement continue. On a le résultat suivant de tension dans $\mathcal{D}'(\mathbb{R}^d)$:

Théorème 2.4.3 *1. Si $\lambda(\rho)\rho^\beta \rightarrow +\infty$ et $n(\rho) = (\lambda(\rho)\rho^\beta)^{1/\alpha}$, on a la convergence en loi dans $\mathcal{D}'(\mathbb{R}^d)$:*

$$\widetilde{M}_\rho(\mu) \Longrightarrow \int_{\mathbb{R}^d \times \mathbb{R}^+} \mu[\tau_{x,r}h] M_\alpha(dx, dr), \quad \mu \in \mathcal{D}(\mathbb{R}^d), \quad \text{lorsque } \rho \rightarrow 0.$$

2. *Si $\lambda(\rho)\rho^\beta \rightarrow a$ et $n(\rho) = (\lambda(\rho)\rho^\beta)^{1/\alpha}$, on a la convergence en loi dans $\mathcal{D}'(\mathbb{R}^d)$:*

$$\widetilde{M}_\rho(\mu) \Longrightarrow \int_{\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}} m\mu[\tau_{x,r}h] \widetilde{N}_{\beta,a}(dx, dr, dm), \quad \mu \in \mathcal{D}(\mathbb{R}^d), \quad \text{lorsque } \rho \rightarrow 0.$$

2.5 Convergence des séries de LePage dans l'espace de Skohorod [D15]

Motivations

En collaboration avec Y.Davydov, on s'intéresse dans [D15] à la convergence dans l'espace de Skohorod $\mathcal{D}^d = \mathcal{D}([0, 1], \mathbb{R}^d)$ muni de la topologie J_1 des séries aléatoires de la forme

$$X(t) = \sum_{i=1}^{\infty} \Gamma_i^{-1/\alpha} \varepsilon_i Y_i(t), \quad t \in [0, 1], \quad (2.10)$$

avec $\alpha \in (0, 2)$,

- $(\Gamma_i)_{i \geq 1}$ les points d'un processus ponctuel de Poisson homogène sur $[0, +\infty)$,
- $(\varepsilon_i)_{i \geq 1}$ une suite i.i.d. de variables aléatoires réelles,
- $(Y_i)_{i \geq 1}$ une suite i.i.d. à valeurs dans \mathcal{D}^d ,

et où les suites (Γ_i) , (ε_i) et (Y_i) sont indépendantes. Les séries de la forme (2.10) sont connues sous le nom de séries de LePage. Pour $t \in [0, 1]$ fixé, la convergence dans \mathbb{R}^d de la série (2.10) est assurée dès que l'une des conditions suivantes est satisfaite :

- $0 < \alpha < 1$, $\mathbb{E}|\varepsilon_1|^\alpha < \infty$ et $\mathbb{E}|Y_1(t)|^\alpha < \infty$,
- $1 \leq \alpha < 2$, $\mathbb{E}\varepsilon_1 = 0$, $\mathbb{E}|\varepsilon_1|^\alpha < \infty$ et $\mathbb{E}|Y_1(t)|^\alpha < \infty$,

où $|\cdot|$ désigne la norme euclidienne usuelle sur \mathbb{R} ou \mathbb{R}^d . Alors $X(t)$ suit une loi α -stable sur \mathbb{R}^d . Réciproquement, il est bien connu que toute loi α -stable sur \mathbb{R}^d admet un développement en série de LePage (*cf.* Samorodnitsky et Taqqu [66] section 3.9).

La littérature sur les lois symétriques α -stable dans un espace de Banach séparable est assez riche (*cf.* Ledoux et Talagrand [43] ou Araujo et Giné [1]). En particulier, toute loi symétrique α -stable sur un Banach séparable admet un développement en série de LePage. Mais l'existence d'une loi symétrique α -stable avec une mesure spectrale donnée n'est pas automatique et est liée à la notion du type de l'espace de Banach (*cf.* Théorème 9.27 de Ledoux et Talagrand [43]). Plus récemment, Davydov, Molchanov et Zuyev [19] ont considéré les lois α -stables à valeurs dans un cône convexe abstrait.

L'espace de Skohorod \mathcal{D}^d muni de la norme uniforme $\|x\| = \sup\{|x(t)|; t \in [0, 1]\}$ est un espace de Banach mais n'est pas séparable. La topologie uniforme associée à cette norme est strictement moins fine que la topologie J_1 . Par ailleurs, l'espace \mathcal{D}^d muni de la topologie J_1 est Polonais (*i.e.* la topologie est métrisée par une métrique rendant l'espace complet et séparable), mais cette métrique n'est pas compatible avec la structure d'espace vectoriel, car l'addition n'est pas continue pour la topologie J_1 . Ces propriétés expliquent que la théorie générale des lois stables sur un Banach séparable ne s'appliquent pas à l'espace de Skohorod.

Cependant, lorsque la série (2.10) converge, la loi de X est une loi α -stable sur \mathcal{D}^d de

mesure spectrale σ sur la sphère unité $\mathcal{S}^d = \{x \in \mathcal{D}^d; \|x\| = 1\}$ donnée par

$$\sigma(A) = \frac{\mathbb{E}\left(|\varepsilon_1|^\alpha \|Y_1\|^\alpha \mathbf{1}_{\{\text{sign}(\varepsilon_1)Y_1/\|Y_1\| \in A\}}\right)}{\mathbb{E}(|\varepsilon_1|^\alpha \|Y_1\|^\alpha)}, \quad A \subset \mathcal{S}^d \text{ Borélien.}$$

De plus, l'élément X est à variations régulières sur \mathcal{D}^d (cf. Hult et Lindskog [32] ou Davis et Mikosch [17]) : pour tout $r > 0$ et tout Borélien $A \subset \mathcal{S}^d$ tel que $\sigma(\partial A) = 0$, on a

$$\lim_{n \rightarrow \infty} n\mathbb{P}\left(\frac{X}{\|X\|} \in A \mid \|X\| > rb_n\right) = r^{-\alpha} \sigma(A),$$

avec

$$b_n = \inf\{r > 0; \mathbb{P}(\|X\| < r) \leq n^{-1}\}, \quad n \geq 1.$$

On dit alors que X est à variations régulières sur \mathcal{D}^d d'indice α et mesure spectrale σ .

Ces propriétés intéressantes conduisent naturellement à se demander quelles conditions assurent la convergence de la série de LePage (2.10). Seulement quelques cas particuliers étaient connus :

- lorsque $0 < \alpha < 1$, $\mathbb{E}|\varepsilon_1|^\alpha < \infty$ et $\mathbb{E}\|Y_1\|^\alpha < \infty$, la série (2.10) converge presque sûrement uniformément sur $[0, 1]$ (cf. Davis et Mikosch [17]) ;
- lorsque $1 \leq \alpha < 2$, que les ε_i sont symétriques et ont un moment d'ordre α et que $Y_i(t) = \mathbf{1}_{[0,t]}(U)$ avec $(U_i)_{i \geq 1}$ i.i.d. de loi uniforme sur $[0, 1]$, la série (2.10) converge presque sûrement uniformément sur $[0, 1]$ et le processus limite X est un Lévy symétrique α -stable (see Rosinski [61]).

Contribution

On propose une condition suffisante pour la convergence presque sûre de la série de LePage (2.10) dans l'espace de Skorohod \mathcal{D}^d . Le résultat est le suivant :

Théorème 2.5.1 *Supposons $1 \leq \alpha < 2$, $\mathbb{E}\varepsilon_1 = 0$, $\mathbb{E}|\varepsilon_1|^\alpha < \infty$ et $\mathbb{E}\|Y_1\|^\alpha < \infty$. Si il existe de plus $\beta_1, \beta_2 > \frac{1}{2}$ et F_1, F_2 continues et croissantes sur $[0, 1]$ tels que*

$$\mathbb{E}|Y_1(t_2) - Y_1(t_1)|^2 \leq |F_1(t_2) - F_1(t_1)|^{\beta_1}, \quad 0 \leq t_1 \leq t_2 \leq 1 \quad (2.11)$$

$$\mathbb{E}|Y_1(t_2) - Y_1(t)|^2 |Y_1(t) - Y_1(t_1)|^2 \leq |F_2(t_2) - F_2(t_1)|^{2\beta_2}, \quad 0 \leq t_1 \leq t \leq t_2 \leq 1, \quad (2.12)$$

alors la série de LePage (2.10) converge presque sûrement dans \mathcal{D}^d .

Nous donnons dans [D15] plusieurs exemples d'applications de ce résultat. Notons seulement ici que par exemple si $Y_1(t)$ est un processus de Poisson d'intensité $\lambda > 0$ sur $[0, 1]$, alors

$$\mathbb{E}(Y_1(t_2) - Y_1(t_1))^2 = \lambda|t_2 - t_1| + \lambda^2|t_2 - t_1|^2, \quad 0 \leq t_1 \leq t_2 \leq 1,$$

et pour $0 \leq t_1 \leq t \leq t_2 \leq 1$,

$$\mathbb{E}(Y_1(t_2) - Y_1(t))^2 (Y_1(t) - Y_1(t_1))^2 = (\lambda|t_2 - t| + \lambda^2|t_2 - t|^2)(\lambda|t - t_1| + \lambda^2|t - t_1|^2).$$

et on vérifie bien que les conditions (2.11) et (2.12) sont satisfaites.

2.6 Perspectives

Les perspectives de recherche sur cette thématique se placent dans le prolongement des travaux [D13] et [D17] sur les modèles de trafic de type « agrégation de sources » et leurs asymptotiques, notamment le régime intermédiaire.

Moments des mesures aléatoires régénératives

Nous réfléchissons actuellement à un cadre général permettant de généraliser les calculs de [D17] sur les moments du processus ponctuel de renouvellement $\xi = \sum_{n \geq 0} \delta_{T_n}$, avec $(T_n)_{n \geq 0}$ suite de renouvellement non stationnaire vérifiant $T_0 = 0$. La masse de Dirac $\delta_{T_n} = \delta_0(\cdot - T_n)$ représente ici un paquet de taille 1 qui arrive au début de la session débutant au temps T_n .

Plus généralement, lorsque la n -ème session donne lieu à une activité représentée par une mesure aléatoire $\eta_n(\cdot - T_n)$, on est amené à considérer la mesure aléatoire

$$\xi(\cdot) = \sum_{n \geq 0} \eta_n(\cdot - T_n),$$

avec la suite $(\eta_n, T_{n+1} - T_n)_{n \geq 0}$ indépendantes et identiquement distribuées. Ces mesures aléatoires sont dites régénératives, car elles vérifient la propriété de renouvellement

$$\xi(\cdot) = \eta_0(\cdot) + \tilde{\xi}(\cdot - T_1) \quad \text{avec} \quad \tilde{\xi}(\cdot) = \sum_{n \geq 1} \eta_n(\cdot - (T_n - T_1))$$

où $\tilde{\xi}$ a la même loi que ξ et (η_0, T_1) est indépendant de $\tilde{\xi}$. De nombreux modèles de trafic peuvent s'inscrire dans ce contexte des mesures régénératives : *infinite source Poisson*, renouvellement, ON/OFF, renouvellement/récompense ...

Des calculs préliminaires montre que la propriété de renouvellement des mesures régénératives permet d'obtenir des relations de récurrence entre les différents moments associés. Les moments de la mesure aléatoire ξ peuvent alors être calculés de manière récursive en exprimant le moment M_k (au sens mesure) d'ordre k en fonction des moments d'ordres inférieurs M_1, \dots, M_{k-1} . Une étude plus systématique de ces mesures aléatoires régénératives nous semble pertinente.

Régime intermédiaire pour le modèle de renouvellement/récompense

Le modèle de renouvellement/récompense est le modèle de trafic où l'activité d'une source est représenté par le processus

$$I(t) = \sum_{n \geq 0} R_n \mathbf{1}_{[T_n, T_{n+1}]}(t)$$

avec $(T_n)_{n \geq 0}$ une suite de renouvellement et $(R_n)_{n \geq 0}$ une suite de variables i.i.d. (les récompenses) indépendante de $(T_n)_{n \geq 0}$. L'interprétation de ce modèle est que le débit de transmission lors de la n -ème session est aléatoire et égal à R_n .

Ce modèle est étudié par Levy et Taqqu [44] qui étudient l'asymptotique lors de l'agrégation de sources dans les régimes de trafic lourd ou léger. Voir aussi Pipiras et Taqqu [57] pour les propriétés des processus limites. Nous envisageons d'appliquer notre approche par les mesures régénératives afin d'obtenir le régime intermédiaire pour ce modèle. La mesure régénérative correspondant au modèle est donnée par

$$\xi(dt) = \sum_{n \geq 0} R_n 1_{[T_n, T_{n+1}]}(t) dt.$$

Nous pensons que les moments peuvent être calculés récursivement et qu'on peut ensuite déterminer leur asymptotique pour obtenir un analogue du Théorème 2.3.2 pour ce modèle. Nous anticipons qu'une variante du processus de Poisson fractionnaire devrait apparaître à la limite. En particulier, la loi des récompenses persisterait à la limite, ce qui n'est pas le cas dans les régimes de trafic lourd et léger.

Etude de la propriété d'autosimilarité par agrégation

Cette question a été suggérée par Serge Cohen. La notion d'auto-similarité par agrégation a été introduite par Biermé, Estrade et Kaj [9] lors de l'étude des champs aléatoires de type « intégrale de Poisson » apparaissant à la limite dans le régime intermédiaire des modèles de boules aléatoires. Cette propriété exprime le fait qu'une somme de copies indépendantes de champs aléatoires suit la même loi que le champ de départ changé d'échelle. Pour le processus de Poisson fractionnaire introduit dans le Théorème 2.3.1, on obtient ainsi que pour $m \geq 1$ et avec $c_m = m^{1/(2H-2)}$, on a

$$\sum_{i=1}^m P_H^i(t) \stackrel{fdd}{=} c_m P_H(t/c_m),$$

avec P_H^1, P_H^2, \dots des copies i.i.d. de P_H .

Il serait intéressant d'étudier de manière plus systématique la notion d'auto-similarité par agrégation (ce problème a été soulevé par Serge Cohen). Clairement, un processus auto-similaire par agrégation est infiniment divisible. On a ainsi pour le processus de Poisson fractionnaire

$$P_H(t) \stackrel{fdd}{=} \sum_{i=1}^m c_m^{-1} P_H^i(c_m t).$$

On pourrait alors s'appuyer sur les travaux de Rosinski et les représentations des processus infiniment divisibles pour étudier plus systématiquement les propriétés des processus auto-similaires par agrégation.

Bibliographie

- [1] A. Araujo and E. Giné. *The central limit theorem for real and Banach valued random variables*. John Wiley & Sons, New York-Chichester-Brisbane, 1980. Wiley Series in Probability and Mathematical Statistics.
- [2] A. Ayache, A. Benassi, S. Cohen, and J. Lévy Véhel. Regularity and identification of generalized multifractional Gaussian processes. In *Séminaire de Probabilités XXXVIII*, volume 1857 of *Lecture Notes in Math.*, pages 290–312. Springer, Berlin, 2005.
- [3] A. Ayache, F. Roueff, and Y. Xiao. Linear fractional stable sheets : wavelet expansion and sample path properties. *Stochastic Process. Appl.*, 119(4) :1168–1197, 2009.
- [4] A. Ayache, D. Wu, and Y. Xiao. Joint continuity of the local times of fractional Brownian sheets. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(4) :727–748, 2008.
- [5] J.M. Bardet, G. Lang, E. Moulines, and P. Soulier. Wavelet estimator of long-range dependent processes. *Stat. Inference Stoch. Process.*, 3(1-2) :85–99, 2000. 19th “Rencontres Franco-Belges de Statisticiens” (Marseille, 1998).
- [6] A. Benassi, S. Cohen, and J. Istas. Identifying the multifractional function of a Gaussian process. *Statist. Probab. Lett.*, 39(4) :337–345, 1998.
- [7] J. Beran. *Statistics for long-memory processes*, volume 61 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, New York, 1994.
- [8] H. Biermé and A. Estrade. Poisson random balls : self-similarity and x-ray images. *Adv. in Appl. Probab.*, 38(4) :853–872, 2006.
- [9] H. Biermé, A. Estrade, and I. Kaj. Self-similar random fields and rescaled random balls models. *J. Theoret. Probab.*, 23(4) :1110–1141, 2010.
- [10] E. Bolthausen. A central limit theorem for two-dimensional random walks in random sceneries. *Ann. Probab.*, 17(1) :108–115, 1989.
- [11] A.N. Borodin. Limit theorems for sums of independent random variables defined on a transient random walk. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 85 :17–29, 237, 244, 1979. Investigations in the theory of probability distributions, IV.
- [12] A.N. Borodin. Limit theorems for sums of independent random variables defined on a recurrent random walk. *Teor. Veroyatnost. i Primenen.*, 28(1) :98–114, 1983.
- [13] X. Chen and D. Khoshnevisan. From charged polymers to random walk in random scenery. In *Optimality*, volume 57 of *IMS Lecture Notes Monogr. Ser.*, pages 237–251. Inst. Math. Statist., Beachwood, OH, 2009.
- [14] S. Cohen and G. Samorodnitsky. Random rewards, fractional Brownian local times and stable self-similar processes. *Ann. Appl. Probab.*, 16(3) :1432–1461, 2006.
- [15] D. Cox. Long range dependence : A review. In *Statistics : An Appraisal*, (H. David and H. David, eds), pages 55–74. Iowa State University Press, 1984.
- [16] B. D’Auria and G. Samorodnitsky. Limit behavior of fluid queues and networks. *Oper. Res.*, 53(6) :933–945, 2005.

- [17] R.A. Davis and T. Mikosch. Extreme value theory for space-time processes with heavy-tailed distributions. *Stochastic Process. Appl.*, 118(4) :560–584, 2008.
- [18] Y. Davydov. The invariance principle for stationary processes. *Teor. Veroyatnost. i Primenen.*, 15 :498–509, 1970.
- [19] Y. Davydov, I. Molchanov, and S. Zuyev. Strictly stable distributions on convex cones. *Electron. J. Probab.*, 13 :no. 11, 259–321, 2008.
- [20] R.L. Dobrushin and P. Major. Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete*, 50(1) :27–52, 1979.
- [21] P. Doukhan, A. Khezour, and G. Lang. Nonparametric estimation for long-range dependent sequences. In *Theory and applications of long-range dependence*, pages 303–311. Birkhäuser Boston, Boston, MA, 2003.
- [22] P. Doukhan, G. Lang, and D. Surgailis. Asymptotics of weighted empirical processes of linear fields with long-range dependence. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(6) :879–896, 2002. En l’honneur de J. Bretagnolle, D. Dacunha-Castelle, I. Ibragimov.
- [23] P. Doukhan, G. Oppenheim, and M.S. Taqqu, editors. *Theory and applications of long-range dependence*. Birkhäuser Boston Inc., Boston, MA, 2003.
- [24] V. Fasen and G. Samorodnitsky. A fluid cluster Poisson input process can look like a fractional Brownian motion even in the slow growth aggregation regime. *Adv. in Appl. Probab.*, 41(2) :393–427, 2009.
- [25] G. Fay, F. Roueff, and P. Soulier. Estimation of the memory parameter of the infinite-source Poisson process. *Bernoulli*, 13(2) :473–491, 2007.
- [26] R. Gaigalas. A Poisson bridge between fractional Brownian motion and stable Lévy motion. *Stochastic Process. Appl.*, 116(3) :447–462, 2006.
- [27] R. Gaigalas and I. Kaj. Convergence of scaled renewal processes and a packet arrival model. *Bernoulli*, 9(4) :671–703, 2003.
- [28] N. Guillin-Plantard and A. Le Ny. Transient random walks on 2D-oriented lattices. *Teor. Veroyatn. Primen.*, 52(4) :815–826, 2007.
- [29] N. Guillin-Plantard and A. Le Ny. A functional limit theorem for a 2D-random walk with dependent marginals. *Electron. Commun. Probab.*, 13 :337–351, 2008.
- [30] N. Guillin-Plantard and C. Priour. Central limit theorem for sampled sums of dependent random variables. *ESAIM Probab. Stat.*, 14 :299–314, 2010.
- [31] N. Guillin-Plantard and C. Priour. Limit theorem for random walk in weakly dependent random scenery. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(4) :1178–1194, 2010.
- [32] H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)*, 80(94) :121–140, 2006.
- [33] H. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116 :770–808, 1951.

-
- [34] C. Hurvich, G. Lang, and P. Soulier. Estimation of long memory in the presence of a smooth nonparametric trend. *J. Amer. Statist. Assoc.*, 100(471) :853–871, 2005.
- [35] C.M. Hurvich and P. Soulier. Testing for long memory in volatility. *Econometric Theory*, 18(6) :1291–1308, 2002.
- [36] I. Kaj. Aspects of wireless network modeling based on poisson point processes. Fields Institute Workshop on Applied Probability, Carleton University, 2006.
- [37] I. Kaj, L. Leskelä, I. Norros, and V. Schmidt. Scaling limits for random fields with long-range dependence. *Ann. Probab.*, 35(2) :528–550, 2007.
- [38] I. Kaj and M.S. Taqqu. Convergence to fractional Brownian motion and to the Telecom process : the integral representation approach. In *In and out of equilibrium. 2*, volume 60 of *Progr. Probab.*, pages 383–427. Birkhäuser, Basel, 2008.
- [39] D. Karmeshu and A. Krishnamachari. Sequence variability and long range dependence in dna : An information theoretic perspective. *Neural Information Processing*, pages 1354–1361, 2004.
- [40] H. Kesten and F. Spitzer. A limit theorem related to a new class of self-similar processes. *Z. Wahrsch. Verw. Gebiete*, 50(1) :5–25, 1979.
- [41] A. Kolmogorov. Wiener'sche spiralen und einige andere interessante kurven in hilbert'schen raum. *Dokl. Akad. Nauk SSSR*, 26 :115–118, 1940.
- [42] R. Lang and X.-X. Nguyen. Strongly correlated random fields as observed by a random walker. *Z. Wahrsch. Verw. Gebiete*, 64(3) :327–340, 1983.
- [43] M. Ledoux and M. Talagrand. *Probability in Banach spaces*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991. Isoperimetry and processes.
- [44] J.B. Levy and M.S. Taqqu. Renewal reward processes with heavy-tailed inter-renewal times and heavy-tailed rewards. *Bernoulli*, 6(1) :23–44, 2000.
- [45] A. Lo. Fat tails, long memory, and the stock market since the 1960s. *Economic notes*, 2001. Banca Monte dei Paschi di Siena.
- [46] M. Maejima. On a class of self-similar processes. *Z. Wahrsch. Verw. Gebiete*, 62(2) :235–245, 1983.
- [47] M. Maejima. A self-similar process with nowhere bounded sample paths. *Z. Wahrsch. Verw. Gebiete*, 65(1) :115–119, 1983.
- [48] M. Maejima. Limit theorems related to a class of operator-self-similar processes. *Nagoya Math. J.*, 142 :161–181, 1996.
- [49] B. Mandelbrot. Une classe de processus stochastiques homothétiques à soi ; application à la loi climatologique de h.e.hurst. *Comptes Rendus de l'Académie des Sciences de Paris*, 240 :3274–3277, 1965.
- [50] B. Mandelbrot. *The fractal geometry of nature*. W. H. Freeman and Co., San Francisco, Calif., 1982. Schriftenreihe für den Referenten. [Series for the Referee].

- [51] B. Mandelbrot and J. Wallis. Noah, Joseph and operational hydrology. *Water Resources Research*, 4 :909–918, 1968.
- [52] B.B. Mandelbrot and J.W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10 :422–437, 1968.
- [53] T. Mikosch, S. Resnick, H. Rootzén, and A. Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.*, 12(1) :23–68, 2002.
- [54] T. Mikosch and G. Samorodnitsky. Scaling limits for cumulative input processes. *Math. Oper. Res.*, 32(4) :890–918, 2007.
- [55] E. Moulines and P. Soulier. Semiparametric spectral estimation for fractional processes. In *Theory and applications of long-range dependence*, pages 251–301. Birkhäuser Boston, Boston, MA, 2003.
- [56] S. Painter. Long-range dependence in the subsurface : Empirical evidence and simulation methods. 1998. Invited paper at the American Geophysical Union 1998 Fall Meeting.
- [57] V. Pipiras and M.S. Taqqu. The limit of a renewal reward process with heavy-tailed rewards is not a linear fractional stable motion. *Bernoulli*, 6(4) :607–614, 2000.
- [58] S.I. Resnick. *Heavy-tail phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.
- [59] Peter M. Robinson, editor. *Time series with long memory*. Advanced Texts in Econometrics. Oxford University Press, Oxford, 2003.
- [60] M. Rosenblatt. Stochastic processes with short-range and long-range dependence. In *Statistics : An Appraisal*, (H. David and H. David, eds), pages 509–520. Iowa State University Press, 1984.
- [61] J. Rosiński. Series representations of Lévy processes from the perspective of point processes. In *Lévy processes*, pages 401–415. Birkhäuser Boston, Boston, MA, 2001.
- [62] G. Samorodnitsky. Long memory and self-similar processes. *Ann. Fac. Sci. Toulouse Math. (6)*, 15(1) :107–123, 2006.
- [63] G. Samorodnitsky. Long range dependence. *Found. Trends Stoch. Syst.*, 1(3) :163–257, 2006.
- [64] G. Samorodnitsky and M.S. Taqqu. The various linear fractional Lévy motions. In *Probability, statistics, and mathematics*, pages 261–270. Academic Press, Boston, MA, 1989.
- [65] G. Samorodnitsky and M.S. Taqqu. $(1/\alpha)$ -self similar α -stable processes with stationary increments. *J. Multivariate Anal.*, 35(2) :308–313, 1990.
- [66] G. Samorodnitsky and M.S. Taqqu. *Stable non-Gaussian random processes*. Stochastic Modeling. Chapman & Hall, New York, 1994. Stochastic models with infinite variance.
- [67] D. Surgailis. Convergence of sums of nonlinear functions of moving averages to self-similar processes. *Dokl. Akad. Nauk SSSR*, 257(1) :51–54, 1981.

- [68] M.S. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self similar traffic modeling. *Comput. Comm. Rev.*, 27 :5–23, 1997.
- [69] Murad S. Taqqu. A bibliographical guide to self-similar processes and long-range dependence. In *Dependence in probability and statistics (Oberwolfach, 1985)*, volume 11 of *Progr. Probab. Statist.*, pages 137–162. Birkhäuser Boston, Boston, MA, 1986.
- [70] C. Varotsos and D. Kirk-Davidoff. Long-memory processes in global ozone and temperature variations at the region 60 ° s- 60 ° n. *Atmospheric Chemistry and Physics*, 6 :4093–4100, 2006.
- [71] W. Wang. Weak convergence to fractional Brownian motion in Brownian scenery. *Probab. Theory Related Fields*, 126(2) :203–220, 2003.
- [72] W. Willinger, M.S. Taqqu, R. Sherman, and D. Wilson. Self-similarity through high variability : statistical analysis of ethernet lan traffic at the source level. *Comput. Comm. Rev.*, 25 :100–113, 1995.

Chapitre 3

Quelques études asymptotiques de modèles aléatoires issus de la géométrie, de la biologie . . .

Dans ce chapitre, nous regroupons des résultats développés depuis la thèse et correspondant aux articles [D1],[D2],[D3],[D4],[D9],[D10],[D11],[D16]. Ces résultats plus épars et abordant des modèles variés sont regroupés dans la thématique « Quelques études asymptotiques de modèles aléatoires issus de la géométrie, de la biologie ». Comme il est difficile d'effectuer une synthèse détaillant chaque travail tant les sujets abordés sont variés, nous avons choisi de présenter deux travaux récents de manière plus détaillée, tandis que les travaux plus anciens seront résumés de manière très succincte.

3.1 Introduction

Ce chapitre intitulé « étude asymptotique de modèles stochastiques issus de la géométrie, de la biologie ... » regroupe des travaux plus éparés. Nous considérons des modèles stochastiques issus de domaines variés, allant de la biologie à la géométrie, en passant par l'informatique et la dynamique des populations.

Derrière la diversité des modèles considérés et des applications se cache une même démarche : l'étude du comportement asymptotique du système. Dans un modèle complexe, le calcul exact des quantités d'intérêt est souvent hors d'atteinte, ou au prix de formules théoriques inexploitable. Il peut alors être pertinent de considérer le comportement asymptotique des observables lorsqu'un paramètre (*e.g.* la taille du système ou l'échelle de temps) tend vers l'infini. Dans de nombreuses situations, le comportement asymptotique du système peut être décrit de manière précise grâce aux théorèmes classiques des probabilités :

- une loi des grands nombres assure la convergence des observables correctement renormalisées vers des constantes ; lorsque la taille du système grandit, des phénomènes de moyennisation s'opèrent et on observe un comportement quasi-déterministe ;
- un théorème central limite permet d'étudier les fluctuations du système autour de sa limite ; les fluctuations ont généralement un ordre de grandeur donné par l'inverse de la racine carrée de la taille du système et suivent approximativement une loi normale ;
- un principe de grandes déviations permet de quantifier dans une échelle logarithmique la probabilité que le système dévie de son comportement typique.

Ma thèse intitulée « Quelques applications de la théorie des grandes déviations » portait sur l'analyse asymptotique de trois modèles aléatoires issus de la biologie, de l'informatique et de l'algorithmique. L'unité du travail repose sur les méthodes utilisées : un travail de modélisation permettant de traduire le problème appliqué dans le formalisme des probabilités suivi d'une étude théorique du modèle et de ses propriétés asymptotiques. L'outil principal utilisé pendant la thèse est la théorie des grandes déviations où l'étude des minimas de la fonctionnelle de taux permet d'obtenir des lois des grands nombres pour les quantités d'intérêt et de déterminer le comportement asymptotique du système lorsque sa taille tend vers l'infini. Dans ce mémoire, les résultats de thèses et autres résultats plus anciens sont rappelés très succinctement à la fin du chapitre. Nous insistons ici seulement sur deux travaux plus récents :

- un travail en collaboration avec l'équipe de Poitiers (M. Arnaudon, A. Phan et L. Yang) portant sur un algorithme stochastique de type gradient pour le calcul de p -moyenne sur une variété ; nous déterminons la convergence presque sûre de l'algorithme ainsi qu'un principe d'invariance fonctionnelle pour ses fluctuations ;
- un travail en collaboration avec V. Bansaye et C. Mazza portant sur le lien entre croissance d'une population en environnement fluctuant et diversité phénotypique ; l'étude repose sur un modèle de population multitype évoluant selon un processus de branchement en environnement aléatoire.

3.2 Un algorithme stochastique de type gradient pour le calcul de p -moyenne sur une variété [D16]

Motivations et contexte

Etant donné un ensemble de points $\{x_1, \dots, x_n\}$ d'un espace euclidien E de métrique d , le barycentre géométrique noté e_2 de ces points est l'unique minimiseur de la distance quadratique moyenne :

$$e_2 = \arg \min_{x \in E} \frac{1}{n} \sum_{i=1}^n d^2(x, x_i).$$

Il est égal à la moyenne empirique $e_2 = \frac{1}{n} \sum_{i=1}^n x_i$ et est l'estimateur le plus commun en statistique. Mais il est sensible aux valeurs extrêmes, et il est naturel de remplacer la puissance 2 par $p \in [1, 2)$. On définit ainsi la notion de p -moyenne : pour $p \geq 1$, un minimiseur de la fonctionnelle

$$H_p : \begin{array}{l} E \rightarrow \mathbb{R}_+ \\ x \mapsto \frac{1}{n} \sum_{i=1}^n d^p(x, x_i) \end{array}$$

est appelé p -moyenne de l'ensemble de points $\{x_1, \dots, x_n\}$. Pour $p = 1$, e_1 est la médiane de l'ensemble de points et est utilisé couramment en statistique pour sa robustesse. Il est bien connu qu'en dimension 1, la médiane d'un ensemble de réels n'est pas toujours définie de manière unique. Ce cas est cependant exceptionnel : la p -moyenne d'un ensemble de points est uniquement déterminée dès que $p > 1$ où que $p = 1$ et les points ne sont pas alignés. Dans ces cas, la fonctionnelle H_p est en effet strictement convexe ce qui assure l'unicité du minimiseur.

La notion de p -moyenne est naturellement étendue aux probabilités sur une variété riemannienne. Soit μ une mesure de probabilité sur une variété Riemannienne M de distance ρ . Pour $p \geq 1$, une p -moyenne de μ est un minimiseur de la fonctionnelle

$$H_p : \begin{array}{l} M \rightarrow \mathbb{R}_+ \\ x \mapsto \int_M \rho^p(x, y) \mu(dy) \end{array} .$$

Contrairement au cas Euclidien, la fonctionnelle H_p n'est pas nécessairement convexe lorsque $p \geq 2$ et la p -moyenne n'est pas toujours uniquement déterminée. Le cas $p = 2$ correspond au barycentre Riemannien ou *Karcher mean* et a été étudiée par Karcher [13], Kendall [15], [16], Emery et Mokobodzki [8], Picard, [18] Arnaudon et Li, [4] ... où les questions d'existence, d'unicité, de stabilité, les liens avec les martingales et les flots stochastiques sont considérés. Dans le cas général $p \geq 1$, l'existence et l'unicité des p -moyennes est prouvée par Afsari [1] dans le cas où le support de la loi μ est « petit ». Plus précisément, si $\text{inj}(M)$ désigne le rayon d'injectivité de la variété M et $\alpha^2 > 0$ une borne supérieure pour les courbures sectionnelles sur M , l'existence et l'unicité de la p -moyenne est prouvée pour les mesures μ à support inclus dans un compact convexe K_μ d'une boule géodésique

$B(a, r)$ de rayon

$$r < r_{\alpha, p} \quad \text{with} \quad r_{\alpha, p} = \begin{cases} \frac{1}{2} \min \left\{ \text{inj}(M), \frac{\pi}{2\alpha} \right\} & \text{if } p \in [1, 2) \\ \frac{1}{2} \min \left\{ \text{inj}(M), \frac{\pi}{\alpha} \right\} & \text{if } p \in [2, \infty) \end{cases} . \quad (3.1)$$

Etant donné l'existence et l'unicité de la p -moyenne, la question de sa détermination pratique se pose naturellement. Dans le cas euclidien avec $p = 1$, le problème de la détermination de la médiane e_1 d'un ensemble de points est connu sous le nom de problème de Fermat-Weber et de nombreux algorithmes ont été proposés. Le premier par Weiszfeld [22] et ensuite adapté par Fletcher *et al.* [10] au cas des variétés à courbure négative. Yang [23] a élargi ces résultats au cadre des de la courbure positive ou négative. Dans le cas $p = 2$, le barycentre Riemanien est calculé par Le [17] à l'aide d'un algorithme de descente du gradient.

Contributions

Nous considérons un algorithme stochastique du type gradient pour déterminer la p -moyenne sur une variété Riemanienne dans le cas général $p \geq 1$ et sous les conditions de Asfari [1] assurant l'existence et l'unicité de la p -moyenne. Cet algorithme est plus facile à implémenter qu'un algorithme de descente du gradient classique, car on n'a pas à calculer à chaque pas le gradient de la fonctionnelle H_p à minimiser. Plus précisément, on construit une chaîne de Markov inhomogène $(X_k)_{k \geq 0}$ de la manière suivante : à chaque pas de temps $k \geq 0$, on tire un point P_{k+1} de loi μ et l'on déplace le point courant X_k en X_{k+1} selon la géodésique de X_k à P_{k+1} d'une distance dépendant de p et d'un paramètre déterministe t_{k+1} . Nous prouvons la convergence de la chaîne de Markov $(X_k)_{k \geq 0}$ vers e_p et étudions aussi les fluctuations autour de la limite.

L'algorithme du gradient stochastique et sa convergence

Soit M une variété riemanienne à courbure pincée : soit $\alpha, \beta > 0$ tels que α^2 soit un majorant de la courbure sectionnelle sur M et $-\beta^2$ un minorant. La distance riemanienne sur M est notée ρ et le rayon d'injectivité de M est noté $\text{inj}(M)$. Soit $B(a, r)$ une boule géodésique dans M avec $a \in M$, $r > 0$ et μ une mesure de probabilité à support inclus dans un compact convexe K_μ de $B(a, r)$. Soit $p \in [1, \infty)$. On suppose la condition suivante sur (r, p, μ) :

Hypothèse A_1 :

- le support de μ n'est pas réduit à un point, et si $p = 1$, on suppose de plus que le support de μ n'est pas inclus dans une droite ;
- le rayon r vérifie (3.1).

Notons que la boule $B(a, r)$ est convexe lorsque $r < \frac{1}{2} \min \left\{ \text{inj}(M), \frac{\pi}{\alpha} \right\}$.

Il est prouvé dans [1] (Theorem 2.1) que, sous la condition A_1 , la fonction

$$H_p : M \rightarrow \mathbb{R}_+$$

$$x \mapsto \int_M \rho^p(x, y) \mu(dy)$$

a un unique minimiseur e_p sur M appelé la p -moyenne de μ , et de plus $e_p \in B(a, r)$. Si $p = 1$, e_1 est la médiane de μ . On vérifie facilement que si $p \in [1, 2)$, alors H_p est strictement convexe sur $B(a, r)$ et que si $p \geq 2$, alors H_p est de classe C^2 sur $B(a, r)$ (mais pas nécessairement convexe).

Dans la suite on fixe

$$K = \bar{B}(a, r - \epsilon) \quad \text{avec} \quad \epsilon = \frac{\rho(K_\mu, B(a, r)^c)}{2}.$$

Nous définissons un algorithme stochastique $(X_k)_{k \geq 0}$ de type gradient stochastique pour approcher la p -moyenne e_p de la manière suivante. On se donne $(P_k)_{k \geq 1}$ une suite i.i.d de variables aléatoires à valeurs dans $B(a, r)$ et de loi μ et $(t_k)_{k \geq 1}$ une suite de réels positifs satisfaisant

$$\forall k \geq 1, \quad t_k \leq \delta_1 = \min \left(\frac{1}{C_{p, \mu, K}}, \frac{\rho(K_\mu, B(a, r)^c)}{2p(2r)^{p-1}} \right), \quad (3.2)$$

$$\sum_{k=1}^{\infty} t_k = +\infty \quad \text{et} \quad \sum_{k=1}^{\infty} t_k^2 < \infty. \quad (3.3)$$

Pour $x_0 \in K$ fixé, on définit la suite $(X_k)_{k \geq 0}$ de manière inductive par

$$X_0 = x_0 \quad \text{et pour } k \geq 0 \quad X_{k+1} = \exp_{X_k} \left(-t_{k+1} \text{grad}_{X_k} F_p(\cdot, P_{k+1}) \right) \quad (3.4)$$

avec $F_p(x, y) = \rho^p(x, y)$ et la convention $\text{grad}_x F_p(\cdot, x) = 0$.

On montre la convergence de l'algorithme du gradient stochastique.

Théorème 3.2.1 *La suite $(X_k)_{k \geq 1}$ converge presque sûrement et dans L^2 vers e_p .*

L'algorithme se simplifie de manière radicale lorsque $M = \mathbb{R}^d$ et $p = 2$. Lorsque $M = \mathbb{R}^d$ et μ loi de probabilité sur \mathbb{R}^d à support compact, l'algorithme (3.4) devient

$$X_0 = x_0 \quad \text{et pour } k \geq 0 \quad X_{k+1} = X_k - t_{k+1} \text{grad}_{X_k} F_p(\cdot, P_{k+1}).$$

Lorsque $p = 2$, $e_2 = \mathbb{E}[P_1]$ et $\text{grad}_x F_p(\cdot, y) = 2(x - y)$, si bien que

$$X_{k+1} = (1 - 2t_{k+1})X_k + 2t_{k+1}P_{k+1}, \quad k \geq 0$$

et on obtient

$$X_k = x_0 \prod_{j=0}^{k-1} (1 - 2t_{k-j}) + 2 \sum_{j=0}^{k-1} P_{k-j} t_{k-j} \prod_{\ell=0}^{j-1} (1 - 2t_{k-\ell}), \quad k \geq 1.$$

Avec $t_k = \frac{1}{2k}$, on a

$$\prod_{j=0}^{k-1} (1 - 2t_{k-j}) = 0 \quad \text{et} \quad \prod_{\ell=0}^{j-1} (1 - 2t_{k-\ell}) = \frac{k-j}{k}$$

et

$$X_k = \sum_{j=0}^{k-1} P_{k-j} \frac{1}{k} = \frac{1}{k} \sum_{j=1}^k P_j.$$

Dans ce cas, l'algorithme du gradient stochastique pour la moyenne e_2 de μ coïncide avec la moyenne empirique d'un échantillon i.i.d. de loi μ . Dans ce cas simple, le résultat du Théorème 3.2.1 n'est rien d'autre que la loi des grands nombres. De plus, les fluctuations autour de la moyenne sont données par le Théorème de Donsker.

Fluctuations de l'algorithme du gradient

Nous nous penchons sur les fluctuations de la suite $(X_k)_{k \geq 0}$ autour de la p -moyenne e_p . On considère la chaîne de Markov inhomogène et à valeurs dans M définie par (3.4) dans le cas particulier où

$$t_k = \min \left(\frac{\delta}{k}, \delta_1 \right), \quad k \geq 1 \quad (3.5)$$

avec $\delta > 0$. La suite $(t_k)_{k \geq 1}$ vérifie alors (3.2) et (3.3), si bien que le Théorème 3.2.1 s'applique et que $(X_k)_{k \geq 0}$ converge p.s. et dans L^2 vers e_p . Afin d'étudier ses fluctuations, on définit pour $n \geq 1$ la chaîne de Markov $(Y_k^n)_{k \geq 0}$ à valeurs dans $T_{e_p} M$ par

$$Y_k^n = \frac{k}{\sqrt{n}} \exp_{e_p}^{-1} X_k.$$

Nous allons montrer la convergence de $(Y_{[nt]}^n)_{0 \leq t \leq 1}$ vers un processus de diffusion non-homogène. Le processus limite est défini dans la proposition suivante :

Proposition 3.2.2 *Supposons que H_p soit C^2 dans un voisinage de e_p et que $\delta > C_{p,\mu,K}^{-1}$. Soit*

$$\Gamma = \mathbb{E} \left[\text{grad}_{e_p} F_p(\cdot, P_1) \otimes \text{grad}_{e_p} F_p(\cdot, P_1) \right]$$

et $G_\delta(t)$ le générateur

$$G_\delta(t)f(y) := \langle d_y f, t^{-1}(y - \delta \nabla d H_p(y, \cdot)^\#) \rangle + \frac{1}{2} \text{Hess}_y f(\Gamma)$$

où $\nabla dH_p(y, \cdot)^\sharp$ est le vecteur dual de la forme linéaire $\nabla dH_p(y, \cdot)$. Alors il existe une unique diffusion inhomogène $(y_\delta(t))_{t>0}$ sur $T_{e_p}M$ de générateur $G_\delta(t)$ et convergeant en probabilité vers 0 lorsque $t \rightarrow 0^+$. Le processus y_δ est continu, converge p.s. vers 0 lorsque $t \rightarrow 0^+$ et admet la représentation intégrale suivante :

$$y_\delta(t) = \sum_{i=1}^d t^{1-\delta\lambda_i} \int_0^t s^{\delta\lambda_i-1} \langle \delta\sigma dB_s, e_i \rangle e_i, \quad t \geq 0, \quad (3.6)$$

avec B_t mouvement Brownien standard sur $T_{e_p}M$, $\sigma \in \text{End}(T_{e_p}M)$ satisfaisant $\sigma\sigma^* = \Gamma$, $(e_i)_{1 \leq i \leq d}$ base orthonormale diagonalisant la forme bilinéaire symétrique $\nabla dH_p(e_p)$ et $(\lambda_i)_{1 \leq i \leq d}$ les valeurs propres associées.

La représentation intégrale (3.6) implique que y_δ est un processus Gaussien centré de covariance

$$\mathbb{E} [y_\delta^i(t_1) y_\delta^j(t_2)] = \frac{\delta^2 \Gamma(e_i^* \otimes e_j^*)}{\delta(\lambda_i + \lambda_j) - 1} t_1^{1-\delta\lambda_i} t_2^{1-\delta\lambda_j} (t_1 \wedge t_2)^{\delta(\lambda_i + \lambda_j) - 1},$$

avec $y_\delta^i(t) = \langle y_\delta(t), e_i \rangle$, $1 \leq i, j \leq d$ et $t_1, t_2 \geq 0$.

Notre principal résultat concernant les fluctuations de l'algorithme du gradient stochastique est le suivant :

Théorème 3.2.3 *Supposons que H_p est C^2 sur un voisinage de e_p et que $\delta > C_{p,\mu,K}^{-1}$. Alors la suite de processus $(Y_{[nt]}^n)_{t \geq 0}$ converge faiblement dans $\mathcal{D}((0, \infty), T_{e_p}M)$ vers y_δ .*

Notons que dans ce résultat, la limite dépend fortement du paramètre $\delta > 0$ choisi pour régler le pas de l'algorithme (cf. équation (3.5)). La preuve de ce résultat repose sur la convergence des générateurs associés aux chaînes de Markov et des calculs différentiel en géométrie riemannienne sur les champs de Jacobi.

Perspectives

Il semblerait intéressant d'obtenir des principes de grandes déviations ou de déviations modérées permettant de quantifier dans une échelle logarithmique la vitesse de convergence de X_n vers la p -moyenne e_p . Une façon d'attaquer le problème serait d'utiliser l'approche par générateur infinitésimal de la théorie des grandes déviations pour les chaînes de Markov développée par Feng et Kurtz [9]. Il faudrait donc commencer par vérifier si les générateurs associés à la chaîne de Markov $(Y_{[nt]}^n)_{t \geq 0}$ « convergent » dans l'échelle des grandes déviations ou des déviations modérées. La théorie générale développée dans [9] est cependant assez lourde et l'application à ce problème précis pas encore claire à ce jour. On s'attend néanmoins à ce que les déviations modérées soient plus faciles à appréhender grâce à la simplicité des fonctionnelles de taux quadratiques en général.

3.3 Diversité phénotypique et croissance d'une population multitype en environnement aléatoire [D11]

L'article [D11] en collaboration avec V.Bansaye et C.Mazza traite de dynamique de populations. Nous étudions le lien entre diversité phénotypique et variations de l'environnement et pour ce faire, nous considérons un modèle de branchement multitype en environnement aléatoire et considérons la dépendance entre le taux de croissance de la population et les taux de transition inter-types.

Motivations

La diversité phénotypique est un facteur important d'adaptabilité des populations aux fluctuations de l'environnement dans lequel elles évoluent. Ainsi, des observations montrent que la plupart des colonies de cellules sont hétérogènes, c'est-à-dire composées de phénotypes variés. Récemment, plusieurs auteurs ont proposés des modèles expliquant le rôle de cette hétérogénéité. Citons notamment les travaux de Haccou et Iwasa [12], Kussel et Leibler [14], Thattai et van Oudenaarden [21] ou encore Gander, Mazza et Rummeler [11]. Dans ce papier, des processus stochastiques à temps continu représentent l'évolution de la population et reposent sur des modèles de naissance couplés par un mécanisme de migration, où les taux de naissance et de migration dépendent d'un environnement fluctuant. Les auteurs utilisent le fait que l'espérance mathématique du vecteur représentant l'évolution de la population est solution d'une équation différentielle linéaire à coefficient variables et dépendant de l'environnement et ils étudient le comportement asymptotique en temps long de ces solutions. Sans rentrer dans les détails, l'idée principale de ces travaux est d'étudier l'impact des taux de migrations entre différents types sur la croissance globale de la population et de déterminer les stratégies permettant la meilleure adaptation globale de la population. On montre ainsi le rôle fondamental des individus mal adaptés à une date donnée mais ayant un bon potentiel dans un environnement susceptible d'arriver dans le futur. Différents types de stratégie sont considérés : les stratégies de type *stochastic switching* où les individus n'ont pas connaissance de l'environnement dans lequel ils évoluent et ne peuvent donc pas adapter leur stratégie de migration à l'environnement courant, et les stratégies de type *responsive switching* où les individus disposent de capteurs leur permettant de connaître l'environnement dans lequel ils évoluent et d'adapter leur stratégie migratoire en conséquence.

Dans l'article [D11], nous considérons des questions similaires en nous appuyant sur des processus de branchements multitypes en environnement aléatoire plutôt que sur des processus de naissance avec migration. On s'intéresse à la détermination des stratégies optimales permettant de maximiser l'exposant de Lyapounov représentant le taux de croissance géométrique de la population sur le long terme. On s'appuie en particulier sur un modèle résoluble de manière analytique et basé sur des noyaux gaussiens (*cf.* Haccou and Iwasa [12]).

Modèle

Commençons par préciser le modèle de processus de branchement multitype en environnement aléatoire. On se donne :

- un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$,
- un espace métrique \mathcal{T} (espace des types),
- un espace polonais \mathcal{E} (espace de l'environnement).

L'environnement est donné par une suite stationnaire ergodique $\omega = (\omega_n)_{n \geq 0}$ à valeurs dans \mathcal{E} et de distribution ν sur $\mathcal{E}^{\mathbb{N}}$. Ici ω_n désigne la valeur de l'environnement au temps $n \geq 0$. La population au temps n est représentée par un vecteur $Z_n = (Z_n^t)_{t \in \mathcal{T}}$, où Z_n^t est le nombre d'individus ayant le trait t dans la population au temps n . La taille de la population totale est notée $|Z_n|$. On suppose qu'à chaque pas de temps, les individus se reproduisent et laisse place à une nouvelle génération. La loi du nombre de descendants d'un individu de type t en environnement e est notée $\Upsilon_{t,e}$, les descendants choisissent leur trait avec la loi $\pi_{t,e}$ sur \mathcal{T} . Soit $\mathbb{N}^{\mathcal{T}}$ l'ensemble des fonctions de \mathcal{T} dans \mathbb{N} à support fini et $(1_t)_{t \in \mathcal{T}}$ la base canonique. On définit le processus $Z_n = (Z_n^t)_{t \in \mathcal{T}}$ à valeurs dans $\mathbb{N}^{\mathcal{T}}$ de la manière suivante :

$$\begin{cases} Z_0 = N_0, \\ Z_{n+1} = \sum_{t \in \mathcal{T}} \sum_{i=1}^{Z_n^t} \sum_{j=1}^{\xi_{i,n}^{t,\omega_n}} 1_{\tau_{i,j,n}^{t,\omega_n}}, \quad n \geq 0, \end{cases} \quad (3.7)$$

avec

- N_0 variable aléatoire à valeurs dans $\mathbb{N}^{\mathcal{T}}$ donnant la population au temps 0 ;
- $\{\xi_{i,n}^{t,e}, i \geq 1, n \geq 0, t \in \mathcal{T}, e \in \mathcal{E}\}$ et $\{\tau_{i,j,n}^{t,e}; i \geq 1, j \geq 1, n \geq 0, t \in \mathcal{T}, e \in \mathcal{E}\}$ variables aléatoires indépendantes de loi

$$\xi_{i,n}^{t,e} \sim \Upsilon_{t,e} \quad \text{et} \quad \tau_{i,j,n}^{t,e} \sim \pi_{t,e}.$$

- $N_0, \{\xi_{i,n}^{t,e}\}, \{\tau_{i,j,n}^{t,e}\}$ et $\omega = \{\omega_n\}$ indépendantes.

Notre étude se concentre sur l'influence de la distribution de la distribution des traits $\pi_{t,e}$ sur la croissance de la population. On considérera la loi de l'environnement ν et la loi du nombre de descendant $\Upsilon_{t,e}$ fixées, alors que π sera considéré comme un paramètre. L'idée intuitive est que :

- i) le fitness d'un individu est déterminé par son phénotype et son environnement,
- ii) la variabilité phénotypique intergénérationnelle influe sur le fitness global de la population et doit être optimisée afin de maximiser le fitness global.

La distribution des traits π peut ainsi être vue comme la stratégie utilisée par la population pour maximiser sa croissance. La performance de la stratégie est mesurée par le taux de croissance à long terme

$$\gamma(\pi) = \lim_{n \rightarrow \infty} n^{-1} \log \mathbb{E}_{\omega, N_0} |Z_n|.$$

Sous des hypothèses assez générale, cette limite existe et ne dépend pas de l'environnement ni de la population initiale. Notons qu'il est important de considérer le modèle gelé

(*quenched*, i.e. conditionné à l'environnement) et non pas le modèle moyenné sur tous les environnements (*annealed*). Ainsi le critère $\gamma(\pi)$ donne le taux de croissance de la population dans un environnement typique. Un critère moyenné sur l'environnement serait biaisé par des environnements peu probable où la population croît très vite.

Quelques résultats

La plupart de nos résultats sont donnés dans le cas non héréditaire où $\pi_{t,e}$ ne dépend pas de t . La suite $(|Z_n|)_{n \geq 0}$ a alors une structure simple de processus de branchement (à un seul type) en environnement aléatoire. L'environnement par paires $\omega^{(2)} = ((\omega_{n-1}, \omega_n))_{n \geq 1}$ forme une suite stationnaire et ergodique sur \mathcal{E}^2 de loi stationnaire ν_2 , c'est-à-dire que ν_2 est la loi de (ω_1, ω_2) . On note également $m_{t,e}$ le moment d'ordre 1 de $\Upsilon_{t,e}$, i.e. le nombre moyen d'enfants d'un individu de type t en environnement e .

Proposition 3.3.1 *Dans le cas non-héréditaire $\pi_{t,e} \equiv \pi_e$, le processus de taille $(|Z_n|)_{n \geq 1}$ est un processus de branchement simple en environnement aléatoire $\omega^{(2)}$ avec loi de reproduction*

$$\Upsilon_{\pi, (e_1, e_2)} = \int_{\mathcal{T}} \Upsilon_{t, e_2} \pi_{e_1}(dt).$$

Conditionnellement à ω , la population au temps n a pour taille moyenne

$$\mathbb{E}_\omega |Z_n| = \mathbb{E}_{\omega_0} |Z_1| \prod_{k=1}^{n-1} m_{p_{\omega_{k-1}, \omega_k}}$$

avec $m_{\pi_{e_1, e_2}} = \int_{\mathcal{T}} m_{t, e_2} p_{e_1}(dt)$ le moment d'ordre 1 de $\Upsilon_{\bar{p}, (e_1, e_2)}$. Si l'intégrale

$$\gamma(\pi) = \int_{\mathcal{E}^2} \log(m_{p_{e_1, e_2}}) \nu_2(de_1, de_2),$$

existe, alors

$$\lim_{n \rightarrow \infty} n^{-1} \log \mathbb{E}_\omega |Z_n| = \gamma(\bar{p}) \quad \text{presque sûrement.}$$

Le critère générale d'extinction des processus de branchements en environnement aléatoire (cf. Athreya et Karlin [2, 3] et Tanny [20]) affirme alors que :

- dans les cas sous-critique $\gamma(\pi) < 0$ et critique $\gamma(\pi) = 0$, la population s'éteint avec probabilité 1 dans presque tous les environnements ;
- dans le cas sur-critique (et sous une hypothèse technique supplémentaire), la population peut survivre avec une probabilité positive dans presque tous les environnements.

Un exemple très simple permet de mettre en évidence l'intérêt de la diversité phénotypique. Dans le cas 2×2 avec $\mathcal{E} = \{e_1, e_2\}$ et $\mathcal{T} = \{t_1, t_2\}$, considérons un environnement i.i.d. tel que $\omega_0 \sim 1/2(\delta_{e_1} + \delta_{e_2})$. Supposons

$$m_{t_1, e_1} = 1.5, \quad m_{t_2, e_1} = 0.6, \quad m_{t_1, e_2} = 0.6, \quad m_{t_2, e_2} = 1.5. \quad (3.8)$$

Ici, les deux environnements sont équiprobables, avec le type t_1 mieux adapté à l'environnement e_1 et une symétrie dans le modèle. Pour les stratégies pures $\pi_{t,e} \equiv \delta_{t_1}$ ou $\pi_{t,e} \equiv \delta_{t_2}$, c'est à dire une population sans diversité phénotypique, on a

$$\gamma(\delta_{t_1}) = \gamma(\delta_{t_2}) = 0.5(\log(1.5) + \log(0.6)) \approx -0.053.$$

En particulier, on est dans le régime sous-critique et la population s'éteint p.s. On peut montrer que la stratégie $\pi_{t,e} \equiv 0.5\delta_{t_1} + 0.5\delta_{t_2}$ est optimale et que

$$\gamma(0.5\delta_{t_1} + 0.5\delta_{t_2}) = \log(0.5 * (1.5 + 0.6)) \approx 0.049.$$

Le régime est surcritique et la population survit avec probabilité non nulle p.s. Cela met en lumière le rôle de la diversité phénotypique dans la population.

On donne ensuite une caractérisation des stratégies non-héréditaires optimales ainsi que des conditions suffisantes d'existence et d'unicité des stratégies optimales. On note $\nu_{e_1}(de_2)$ la loi conditionnelle de ω_2 sachant $\omega_1 = e_1$.

Proposition 3.3.2 *Dans le cas non-héréditaire $\pi_{t,e} \equiv \pi_e$:*

- *Une stratégie est optimale si et seulement si*

$$\sup_{t \in \mathcal{T}} \int_{\mathcal{E}} \frac{m_{t,e_2}}{m_{\pi_{\omega_1}, e_2}} \nu_{\omega_1}(de_2) \leq 1 \quad \text{presque sûrement.}$$

- *Les conditions suivantes assurent l'existence et l'unicité de la stratégie optimale :*
 - *il existe $M > 0$ tel que $m_{t,e} \leq M$ pour tout $t \in \mathcal{T}, e \in \mathcal{E}$;*
 - *pour tout $e \in \mathcal{E}$, l'application $t \mapsto m_{t,e}$ est continue et tend vers 0 en l'infini ;*
 - *la famille $\{t \mapsto m_{t,e}; e \in \mathcal{E}\}$ est linéairement indépendante.*

Les résultats précédents permettent d'analyser les cas où l'espace des traits \mathcal{T} est fini et également un modèle basé sur des distributions Gaussiennes dû à Haccou et Iwasa [12]. Nous détaillons seulement ce deuxième exemple ici. On se donne $\mathcal{E} = \mathcal{T} = \mathbb{R}$ et on suppose l'environnement donné par une suite Gaussienne stationnaire et ergodique de loi stationnaire $\nu_1 = \mathcal{N}(\mu, \sigma_2^2)$ et de corrélation $\rho = \text{corr}(\omega_0, \omega_1)$. On suppose que le nombre moyen d'enfants d'un individu de trait t en environnement e est donné par le noyau Gaussien

$$m_{t,e} = \frac{C}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(t-e)^2}{2\sigma_1^2}\right).$$

En environnement e , les individus de trait $t = e$ sont les mieux adaptés. La proposition précédente permet d'affirmer l'existence et l'unicité de la stratégie optimale et de la calculer. Elle est donnée par

$$\pi_e = \begin{cases} \delta_{\mu + \rho(e - \mu)} & \text{si } (1 - \rho^2)\sigma_2^2 \leq \sigma_1^2 \\ \mathcal{N}(\mu + \rho(e - \mu), (1 - \rho^2)\sigma_2^2 - \sigma_1^2) & \text{si } (1 - \rho^2)\sigma_2^2 \geq \sigma_1^2 \end{cases}$$

et l'exposant de Lyapounov optimal est alors

$$\gamma(\pi_e) = \begin{cases} \log C - \frac{1}{2} \log(2\pi\sigma_1^2) - \frac{(1-\rho^2)\sigma_2^2}{2\sigma_1^2} & \text{si } (1-\rho^2)\sigma_2^2 \leq \sigma_1^2 \\ \log C - \frac{1}{2} \log(2\pi(1-\rho^2)\sigma_2^2) - \frac{1}{2} & \text{si } (1-\rho^2)\sigma_2^2 \geq \sigma_1^2 \end{cases} .$$

Des calculs analogues permettent de déterminer la stratégie optimale lorsque l'individu n'a pas connaissance de l'environnement dans lequel il évolue. On peut alors discuter selon les paramètres du modèle $(C, \sigma_1, \sigma_2, \rho)$ la pertinence des différentes stratégies : pure, *stochastic switching* ou *responsive switching* et déterminer les stratégies amenant à l'extinction de la population ou au contraire assurant sa survie.

Dans le papier [D11], nous développons également le cas où l'espace des traits \mathcal{T} est fini. L'étude se termine par quelques résultats dans le cas héréditaire qui se révèle nettement plus complexe : on donne la loi en champ moyen de la généalogie typique d'un individu ainsi qu'une formule pour l'espérance de la taille de la population au temps n . Le modèle de Haccou et Iwasa permet là encore des formules analytiques.

Perspectives

Notre travail s'est concentré sur le cas non-héréditaire, la taille de la population évolue alors comme un processus de branchement simple en environnement aléatoire ce qui simplifie grandement les calculs. Le cas héréditaire, c'est à dire lorsque la loi du type des enfants dépend du type des parents est beaucoup plus difficile. Nous réfléchissons actuellement à une approche par la théorie des grandes déviations du problème, qui nous permettrait d'exprimer l'exposant de Lyapounov comme solution d'un problème variationnel. Notre approche reposerait sur les travaux de Seppalainen [19] étudiant les grandes déviations des chaînes de Markov en environnement aléatoire et sur la méthode de Laplace permettant de traiter la structure de branchement. Cette approche par la théorie des grandes déviations des problèmes de branchement (sans environnement aléatoire) est développée par Bansaye et Lambert [5]

3.4 Travaux de thèse [D1,D2,D3]

Un modèle stochastique de dénaturation de l'ADN [D1]

On considère le modèle de Benham pour la séparation des deux brins (ou dénaturation) de l'ADN qui prend en compte des données physiques (température) ou géométrique (superhélicité du polymère d'ADN). C'est un modèle de spins en champ moyen avec champ extérieur hétérogène et conditionné aux configurations présentant un faible nombre de « bulles de dénaturation ». Sous des conditions portant sur le champ extérieur, nous prouvons un principe de grandes déviations pour la dénaturation totale, où la fonction de taux dépend fortement de la structure du champ extérieur. Cela nous permet d'obtenir une loi des grands nombres pour la dénaturation et d'étudier le taux de dénaturation en fonction de la température et de la superhélicité. On met en évidence des phénomènes de seuil et de transitions de phase.

Structure de données dynamiques [D2]

On présente une analyse probabiliste des structures de données dynamiques dont la taille évolue selon une marche aléatoire dynamique. Les opérations basiques d'insertion, suppression, requêtes ... sont considérées et ont une loi dépendant du temps. Le modèle constitue un premier pas vers l'étude de ces modèles classiques avec une hétérogénéité temporelle introduite par le système dynamique. Nous illustrons nos résultats sur les structures de listes classiques (listes linéaires, queues de priorité, dictionnaires ...).

Un algorithme génétique de type « sélection/mutation » [D3]

Nous considérons un modèle de marches aléatoires pondérée défini comme suit : la trajectoire (S_0, \dots, S_n) de la marche simple symétrique sur \mathbb{Z} est affectée d'une probabilité proportionnelle au produit $\prod_{k=1}^n f(S_k)$, où f désigne la fonction de *fitness*. Grâce à des méthodes de grandes déviations, on prouve la convergence à vitesse exponentielle de la trajectoire renormalisée vers une fonction déterministe. Cette fonction limite est solution d'un problème variationnel que l'on peut résoudre explicitement dans le cas des fonctions *fitness* de type puissance. Ces résultats sont ensuite appliqués à l'étude d'une dynamique de mutation-sélection sur les entiers, obtenue comme la limite en population infinie d'un algorithme génétique de mutation-sélection. On montre que cette dynamique a un comportement ballistique et on explicite sa vitesse.

3.5 Autres travaux [D4,D9,D10]

Le modèle de Curie-Weiss en champ aléatoire quasi-périodique [D4]

On étudie le modèle de Curie Weiss soumis à un champ extérieur engendré par un système dynamique. On prouve que la magnétisation vérifie les théorèmes classiques de probabilité (loi des grands nombres, théorème central limite, principe de grandes déviations). Ces résultats étendent ceux obtenus par Ellis et Newman [7, 6] dans le cas de champ extérieur homogène.

Nombres de Betti des surfaces aléatoires associés aux espaces de polygones [D9]

Les espaces de polygones du type

$$M_\ell = \left\{ (u_1, \dots, u_n) \in S^1 \times \dots \times S^1 ; \sum_{i=1}^n l_i u_i = 0 \right\} / SO(2)$$

ou leurs analogues trois-dimensionnels jouent un rôle important en géométrie et en topologie et interviennent également dans des modèles de robotique où l_i représente les longueurs des différents segments d'un bras articulé. On considère le cas où les réels l_i sont des variables aléatoires et donc M_ℓ est une surface aléatoire. La complexité de la surface peut être approchée par des grandeurs telles que les nombres de Betti, la caractéristique d'Euler ou le polynôme de Poincaré. On étudie ces quantités d'un point de vue asymptotique lorsque n tend vers l'infini et on obtient des résultats du type loi des grands nombres et théorème central limite.

Analyse du problème stochastique des serveurs sur le cercle [D10]

On analyse une version stochastique du problème des serveurs dans lequel des serveurs se déplacent sur un cercle pour satisfaire des requêtes arrivant de manière aléatoire. On suppose les requêtes indépendantes et identiquement distribuées. Le coût de service est mesuré par la distance effectuée par un serveur pour honorer une requête. Le but est de minimiser le coût moyen de service par requête à l'équilibre. On étudie la performance de la stratégie glouton, où une requête est systématiquement servie par le serveur le plus proche.

Bibliographie

- [1] B. Afsari. Riemannian L^p center of mass : existence, uniqueness, and convexity. Proceedings of the American Mathematical Society, S 0002-9939(2010)10541-5, Article electronically published on August 27, 2010.
- [2] K.B. Athreya and S.Karlin. On branching processes with random environments, I : extinction probabilities. *Ann. Math. Stat.* 42, No. 5, 1499–1520, 1971.
- [3] K.B. Athreya and S. Karlin. On branching processes with random environments, II : limit theorems. *Ann. Math. Stat.* 42, No.6, 1843–1858, 1971.
- [4] M. Arnaudon and X.M. Li. Barycenters of measures transported by stochastic flows. *Ann. Probab.* 33, no. 4, 1509–1543, 2005.
- [5] V. Bansaye et A. Lambert. Past, growth and persistence of source-sinks metapopulations. Preprint, 2011.
- [6] R.S. Ellis. *Entropy, Large Deviations and Statistical Mechanics*. New-York, Springer-Verlag, 1985.
- [7] R.S. Ellis and C.M. Newman. Limit theorems for sums of dependent random variables occuring in statistical mechanics. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 44, 117–139, 1978.
- [8] M. Emery and G. Mokobodzki. Sur le barycentre d’une probabilité dans une variété. *Séminaire de Probabilités XXV, Lecture Notes in Mathematics 1485*, Springer, Berlin, 220–233, 1991.
- [9] J. Feng and T. Kurtz. *Large deviations for stochastic processes. Mathematical Surveys and Monographs*, 131. American Mathematical Society, Providence, RI, 2006.
- [10] P.T. Fletcher, S. Venkatasubramanian and S. Joshi. The geometric median on Riemannian manifolds with application to robust atlas estimation. *NeuroImage* 45, S143–S152, 2009.
- [11] M. Gander, C. Mazza and H.Rummler. Stochastic gene expression in switching environments. *J. Math. Biol.* 55, 249-269, 2007.
- [12] P. Haccou and Y. Iwasa Y. Optimal mixed strategies in stochastic environments. *Theoretical population biology* 47, 212-243, 1995.
- [13] H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, vol XXX, 509–541, 1977.
- [14] E. Kussel and S. Leibler. Phenotypic diversity, population growth and information in fluctuating environments. *Science* 309, 2075-2078, 2005.
- [15] W.S. Kendall. Probability, convexity and harmonic maps with small image I : uniqueness and fine existence. *Proc. London Math. Soc.* (3) 61 no. 2, 371–406, 1990.
- [16] W.S. Kendall. Convexity and the hemisphere. *J. London Math. Soc.*, (2) 43 no.3, 567–576, 1991.

- [17] H. Le. Estimation of Riemannian barycentres. *LMS J. Comput. Math.* 7, 193–200, 2004.
- [18] J. Picard. Barycentres et martingales dans les variétés. *Ann. Inst. H. Poincaré Probab. Statist.* 30, 647–702, 1994.
- [19] T. Seppalainen. Large deviations for Markov chains with random transitions. *Ann. Probab.* 22, no. 2, 713–748, 1994.
- [20] D. Tanny. Limit theorems for branching process in a random environment. *Ann. Probab.* 5, No. 1, 100–116, 1994.
- [21] M. Thattai and A. van Oudenaarden. Stochastic gene expression in fluctuating environments. *Genetics* 167, 523–530, 2004.
- [22] E. Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Math. J.* 43, 355–386, 1937.
- [23] L. Yang, Riemannian median and its estimation, to appear in *LMS Journal of Computation and Mathematics*.