



**HAL**  
open science

**Comportement oscillatoire presque sûr des processus empiriques et browniens : lois fonctionnelles, déviations et approximations, utilisation en statistique asymptotique.**

Philippe Berthet

► **To cite this version:**

Philippe Berthet. Comportement oscillatoire presque sûr des processus empiriques et browniens : lois fonctionnelles, déviations et approximations, utilisation en statistique asymptotique.. Mathématiques [math]. IRMAR - Université Rennes 1, 2005. tel-01761027

**HAL Id: tel-01761027**

**<https://hal.science/tel-01761027>**

Submitted on 7 Apr 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse d'Habilitation à Diriger des Recherches  
Université de Rennes 1  
Institut de Recherche MATHématique de Rennes<sup>1</sup>  
*Spécialité : Mathématiques*

Comportement oscillatoire presque sûr des processus  
empiriques et browniens : lois fonctionnelles,  
déviations et approximations,  
utilisation en statistique asymptotique.

**Philippe BERTHET**<sup>2</sup>

**Soutenue le 8 décembre 2005 devant le jury composé de Messieurs les Professeurs:**

**Bernard DELYON** (Université Rennes 1)  
**Uwe EINMAHL** (Vrije Universiteit Brussel, Belgique)  
**Yves GUIVARC'H** (Université Rennes 1)  
**Pascal MASSART** (Université Paris 11 Orsay)  
**Jean MÉMIN** (Université Rennes 1)  
**Michel TALAGRAND** (Académie des Sciences, Universités Paris 6 et Ohio)  
**Alexandre TSYBAKOV** (Université Paris 6)

**Au vu des rapports de Messieurs les Professeurs:**

**Evarist GINÉ** (University of Connecticut, USA)  
**Pascal MASSART** (Université Paris 11 Orsay)  
**Yakov NIKITIN** (University of St-Petersbourg, Russie)

---

<sup>1</sup>IRMAR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes (FRANCE)

<sup>2</sup>philippe.berthet@univ-rennes1.fr



## Avant-propos.

Notre champ d'investigation privilégié est situé quelque part dans la jeune et déjà vaste contrée des mathématiques de l'aléatoire, le long du fleuve stochastique, sinueux et brumeux, parfois tumultueux, séparant vaguement probabilité et statistique théoriques.

Nous y menons l'étude systématique de quelques animaux épurés, surprenantes entités étherées, que nous appelons processus fondamentaux. Bien étrange est leur comportement, à la fois imprévisible et déterminé, régulé et aléatoire.

Pour disséquer cet aléa et en découvrir quelques mécanismes nous zoomons toujours plus près, ajustant d'une main l'horloge logarithmique et cueillant avec l'autre de mystérieuses constantes.

Cependant, notre humble rôle consiste aussi à leur faire traverser le fleuve, dans les deux sens, afin qu'ils s'alimentent sur chaque berge et enrichissent chaque territoire. Ainsi bâtissons-nous des ponts, conceptuels, à l'aide d'outils, mathématiques, et de techniques, bien spécifiques.

Qu'ils soient empiriques, browniens ou poissoniens, ces ponts enjambent la brume. Puissent-ils relier les temples et inviter au voyage leurs gardiens les moins sectaires !

Car les lois du hasard sont les mêmes pour tou(te)s, là-bas, aux sources du fleuve.

A quoi bon, dès lors, jurer par Theta ou par Omega ?

*À mon Arc-En-Lune.*



## Remerciements.

Je remercie vivement les membres de mon jury de me faire le grand plaisir d'y participer. Tous sont garants du meilleur niveau d'expertise scientifique dans leur spécialité respective.

En premier lieu son président, Michel TALAGRAND. Ses nombreuses contributions à la théorie générale des processus, empiriques ou gaussiens, comme la régularité et la concentration, ont été novatrices, élégantes et décisives. Je cite souvent ses articles, en conférence comme en cours. Il est donc le troisième membre de l'Académie des Sciences à me faire l'honneur d'examiner mon travail.

En second lieu, les rapporteurs de mes travaux. Evarist GINÉ, fameux représentant de l'école nord américaine de probabilité dans les espaces de Banach. Il a beaucoup contribué à la théorie des processus empiriques abstraits depuis plus de vingt ans et à ses applications en statistique non paramétrique, notamment en estimation fonctionnelle. Pascal MASSART, représentant incontournable de l'école française de statistique mathématique. Fervent défenseur de l'approche non asymptotique, il étudie la concentration, l'approximation gaussienne et la pénalisation de la mesure empirique et en développe différents champs d'application. Yakov NIKITIN, digne représentant de l'école russe sur les méthodes asymptotiques en probabilité et en statistique mathématique dans la lignée d'Ildar IBRAGIMOV. Il étudie les statistiques non paramétriques en développant des outils analytiques de petites et grandes déviations des processus gaussiens, notamment pour le contrôle des tests.

Ma satisfaction est vraiment très grande d'avoir la reconnaissance scientifique motivée de ces trois éminents chercheurs de classe internationale, dans des domaines distincts.

En troisième lieu, les examinateurs. Bernard DELYON, dont la curiosité scientifique iconoclaste est si rafraîchissante. Son champ de compétence intersecte l'analyse, les probabilités et les statistiques. Uwe EINMAHL, probabiliste reconnu à la croisée des écoles américaines et européennes. Il est notoirement spécialiste en grandes déviations, construction hongroise, lois générales du logarithme itéré, sommes partielles et processus empiriques. Yves GUIVARC'H et Jean MÉMIN, probabilistes rennais émérites bien connus. Leurs travaux respectifs sur les marches aléatoires et les processus stochastiques ont très fortement contribué à l'excellente réputation de l'aléatoire à Rennes et leur vitalité mathématique s'avère inaltérable, donc ergodique ou stationnaire. Alexandre TSYBAKOV, spécialiste renommé en statistique non et semi paramétrique, problèmes inverses, adaptativité, minimax et classification. Sa présence atteste d'une nouvelle orientation dans mes investigations liées aux processus empiriques.

Je remercie très chaleureusement mes co-auteurs et amis, qui m'ont beaucoup appris et motivé, chacun à sa manière : Charles EL-NOUTY, Andrei FROLOV, Mikhaïl LIFSHITS, David MASON, Zhan SHI.

Je dois également évoquer l'école hongroise sur les théorèmes limites, non représentée ici. Je pense en particulier à Endre CSÁKI, Miklós et Sándor CSÖRGŐ, Pál ERDÖS, Lajos HORVÁTH, János KOMLÓS, Péter MAJOR, Alfred RÉNYI, Pál RÉVÉSZ et Gábor TUSNÁDY. Ils m'ont transmis la passion du logarithme itéré, de l'approximation forte, des propriétés trajectorielles et des méthodes probabilistes en statistique non paramétrique. Je n'oublie pas que ces directions de recherche initiales ont pris source dans le cours mémorable de Paul DEHEUVELS et ses articles avec David MASON.

Je salue aussi Jean BERTOIN, Paul DEHEUVELS, Gérard KERKYACHARIAN, Michel LEDOUX, Jean-François LEGALL, Dominique PICARD, Bernard ROYNETTE, Zhan SHI, Wendelin WERNER qui ont accepté de participer au jury mais que mon choix tardif de date a rendu indisponibles.

Je salue amicalement mes jeunes collègues des trois équipes de l'aléatoire à l'IRMAR, dont j'apprécie le dynamisme, la disponibilité et la capacité à dépasser les anciens clivages. Car une bonne ambiance de travail n'a pas de prix. Plus généralement, un grand merci aux enthousiastes, aux curieux, aux vecteurs d'intuition et aux élabousseurs d'idées qui colorient nos savoirs et affûtent notre perception. Qu'ils soient chercheurs de renom ou étudiants, j'aime les écouter, les lire, les contre-dire.

Je salue enfin les chercheurs libres et indépendants, plus particulièrement ceux qui fuient les chapelles et persévèrent sur des sentiers peu fréquentés ou mal éclairés, sans capitaine d'industrie pour défricher ou sous-traiter. Leur résistance est un précieux soutien moral.

# Sommaire.

1 – INTRODUCTION . . . . .	P 4
<b>1.1 – Motivation</b>	
1.1.1 – Aléa et déterminisme	
1.1.2 – Première droite de Kolmogorov	
1.1.3 – Seconde droite de Kolmogorov	
<b>1.2 – Plan</b>	
<b>1.3 – Ligne directrice</b>	
1.3.1 – Approche générale	
1.3.2 – Approche spécifique	
2 – CONVERGENCE DES INCRÉMENTS : LOIS FONCTIONNELLES DU LOGARITHME . . . . .	P 10
<b>2.1 – Recouvrement de Strassen en topologie uniforme</b>	
2.1.1 – Vitesses de concentration	
2.1.2 – Vitesses d’approche ponctuelles	
2.1.3 – Vitesses pour les oscillations browniennes	
2.1.4 – Autres processus	
2.1.5 – Vitesses poissoniennes et principe d’invariance	
2.1.6 – Recouvrement complet et temps d’atteinte	
<b>2.2 – Lois et vitesses en topologie de type Höldérienne</b>	
2.2.1 – Ensembles de trajectoires browniennes	
2.2.2 – Incréments browniens	
2.2.3 – Compacité relative et espaces pré-Höldériens	
2.2.4 – Vitesses intérieures le long de $\mathcal{B}_{n,\alpha}$	
2.2.5 – Vitesses extérieures le long de $\mathcal{B}_{n,\alpha}$	
<b>2.3 – Lois fonctionnelles intermédiaires</b>	
2.3.1 – Interpolation entre deux résultats de Lévy	
2.3.2 – Quelques vitesses browniennes intermédiaires	
2.3.3 – Processus empirique réel	
2.3.4 – Processus empirique multivarié	
<b>2.4 – Autres axes de recherche</b>	
2.4.1 – Incréments retardés	
2.4.2 – Ensembles fractals aléatoires	
2.4.3 – Convergences ensemblistes dans la loi d’Erdős-Rényi	
2.4.4 – Marche aléatoire dans le domaine d’attraction stable	
2.4.5 – Perspectives	
<b>2.5 – Références citées en partie 2</b>	
3 – OUTILS : PETITES ET GRANDES DÉVIATIONS, APPROXIMATION FORTE . . . . .	P 44
<b>3.1 – Petites déviations browniennes</b>	
3.1.1 – Frontière de Strassen	
3.1.2 – Vitesse à la frontière irrégulière	
3.1.3 – Un lemme de petite déviation	
3.1.4 – Exemples de constantes explicites	

<b>3.2 – Petites boules browniennes en norme sup pondérée</b>	
3.2.1 – Probabilité de petite boule pondérée	
3.2.2 – Applications et constantes explicites	
<b>3.3 – Grandes déviations de l’argmin</b>	
3.3.1 – Brownien avec dérive polynomiale	
3.3.2 – Brownien avec dérive régulière	
<b>3.4 – Approximation forte du processus empirique indexé par des fonctions</b>	
3.4.1 – Problème de l’approximation gaussienne	
3.4.2 – Vitesse sous condition d’entropie uniforme	
3.4.3 – Vitesse sous condition d’entropie à crochet	
3.4.4 – Deux résultats de Dudley et Philipp revisités	
3.4.5 – Quelques conséquences	
3.4.6 – Chaînage fort	
<b>3.5 – Autres axes de recherche</b>	
3.5.1 – Petites déviations poissoniennes et empiriques	
3.5.2 – Processus composés	
3.5.3 – Perspectives	
<b>3.6 – Références citées en partie 3</b>	
4 – UTILISATION EN STATISTIQUE NON PARAMÉTRIQUE . . . . .	P 66
<b>4.1 – Estimation de la densité par delta-suites</b>	
4.1.1 – Delta-suites de fonctions	
4.1.2 – Forme des delta-fonctions	
4.1.3 – Déviations stochastiques exactes	
4.1.4 – Quelques vitesses explicites	
<b>4.2 – Estimation du shorth</b>	
4.2.1 – Notations et hypothèses	
4.2.2 – Lois du logarithme itéré	
4.2.3 – Une nouvelle méthode	
<b>4.3 – Estimation du mode</b>	
4.3.1 – Position du problème	
4.3.2 – Représentation quantile	
4.3.3 – Oscillations exactes	
4.3.4 – Convergence d’estimateurs	
<b>4.4 – Autres axes de recherche</b>	
4.4.1 – Détection de singularités	
4.4.2 – Sélection de modèles	
4.4.3 – Régression et mode conditionnel	
4.4.4 – Perspectives	
<b>4.5 – Références citées en partie 4</b>	
5 – CONCLUSION . . . . .	P 87
6 – LISTE DE TRAVAUX . . . . .	P 90



# 1 – INTRODUCTION

Notre domaine de recherche est celui des théorèmes limites presque sûrs en probabilité et statistique. Nos objets d'étude favoris sont la mesure empirique, le mouvement brownien et les processus associés. Ce document esquisse une synthèse homogène de nos travaux terminés [1–17] et place en perspective ceux en cours de finalisation [18–25]. A ce jour [12–17] ne sont pas encore publiés, car dans notre plan de recherche, global mais cohérent, nous avons mené de front un grand nombre de projets.

Nous mettons en lumière trois aspects de notre contribution scientifique, distincts et complémentaires. L'étude fine des fluctuations trajectorielles, les outils probabilistes attenants, et quelques conséquences statistiques. Afin de préciser ce découpage nous évoquons par l'exemple nos motivations en 1.1, situons ces trois angles d'éclairage en 1.2 et formalisons notre approche systématique en 1.3.

## 1.1 – Motivation

### 1.1.1 – Aléa et déterminisme

Notre motivation profonde est tant mathématique que philosophique. Elle concerne la recherche de l'ordre dans le chaos, du déterminisme dans l'aléatoire. Aussi les objets déterministes, rigides, ne nous intéressent-ils que lorsqu'ils apparaissent à la limite, par la force même d'un aléa qui, pour ainsi dire, leur donne vie. Aussi l'aléa ne nous intéresse-t-il que lorsqu'il s'ordonne inéluctablement, tout en oscillant inlassablement, sans se soumettre complètement. Les phénomènes ainsi révélés sont autant de caractères du hasard pur. Bien qu'ils contrarient parfois notre perception trop cartésienne, ils donnent le plus souvent un sens rigoureux à l'intuition fréquentiste qui tend à postuler de tels déterminismes.

Le contrôle déterministe des fluctuations du hasard est une branche de la théorie des probabilités qui s'est beaucoup développée ces dernières décennies, notamment depuis [S]. Il joue naturellement un rôle primordial en théorie de la décision statistique. C'est à l'interface de ces deux théories que nous situons notre activité de recherche et lui donnons sens. Par ailleurs cette problématique n'est pas sans affinité avec le point de vue dual, celui de la théorie ergodique, qui recherche le désordre imprévisible dans le déterminisme, des comportements typiquement probabilistes dans des systèmes dynamiques. Approfondir cette dualité est une perspective qui nous tient à coeur, repoussée faute de temps.

Deux types d'aléas déterminés nous motivent particulièrement, trajectoriel (1.1.2) et spatial (1.1.3).

En premier lieu, les phénomènes probabilistes presque sûrs liés aux oscillations de certains processus stochastiques élémentaires. Les techniques développées s'appliquent à d'autres trajectoires plus à la mode pourvu que les outils appropriés soient forgés. Car dans notre optique, l'extravagance doit rester improbable et il faut évaluer des probabilités d'événements rares de nature variée, souvent non disponibles. Au degré de précision recherché, ces contrôles doivent se faire de façon non asymptotique, ce qui est très exigeant. C'est une des justifications de notre démarche, qui n'est pas seulement dédiée à l'asymptotique puisque la plupart de nos théorèmes limites presque sûrs admettent en puissance une version à distance finie, sous la forme d'une borne explicite de la probabilité d'un événement d'intérêt.

En second lieu, les phénomènes statistiques presque sûrs liés à la répartition d'un grand nombre  $n$  de données. La problématique concerne tout ce qui arrivera au statisticien immortel,  $n \rightarrow \infty$ . Les objets étudiés sont ici moins épurés. Notre approche consiste alors à se ramener aux processus fondamentaux qui gouvernent la partie purement stochastique de certaines statistiques, les termes de déviation et non de biais, tout en contenant les quantités et les transformations secondaires. De nouveau, pour traiter l'asymptotique, nous produisons des bornes à distance finie assez puissantes pour dévoiler, en corollaire, des probabilités d'erreur, des régions de confiance ou des convergences en loi.

### 1.1.2 – Première droite de Kolmogorov

Précisons la notion de comportement trajectoriel déterministe via l'exemple universel de la marche aléatoire sur  $\mathbb{R}$  qui du reste, ne sera pas repris ensuite. Comment en effet ne pas évoquer la loi forte

des grands nombres de Kolmogorov ? Additionner indépendamment du hasard pur s'avère, vu de trop près, bien chaotique. Pourtant, vu de l'infini, nous obtenons à chaque fois une unique et même droite. Chacun reconnaissant l'équivalence presque sûre  $\sum_{i \leq n} X_i \sim n\mathbb{E}X$  avec  $X_i$  indépendantes et de même loi que  $X \in \mathbb{L}_1$ , d'espérance  $\mathbb{E}X \neq 0$ . Par le simple fait des axiomes de probabilité,  $\mathbb{E}X$  est bien la limite déterministe de la moyenne empirique  $\sum_{i \leq n} X_i/n$ , en accord avec l'intuition objectiviste.

**Question 1.1.** Vu de plus près, mais pas trop, dans quelle mesure la marche aléatoire est-elle déterminée à quitter la droite  $n\mathbb{E}X$  ?

(i) Évoquons quelques réponses précises à la grande échelle de temps  $[0, n]$ . Si  $X \in \mathbb{L}_2$  les mêmes axiomes règlent le comportement de  $S_n = \sum_{i \leq n} X_i - n\mathbb{E}X$  et la loi du logarithme itéré atteste que

$$\limsup_{n \rightarrow \infty} \frac{\max_{m \leq n} S_m}{\sqrt{2n \log \log n}} = -\liminf_{n \rightarrow \infty} \frac{\min_{m \leq n} S_m}{\sqrt{2n \log \log n}} = 1 \quad p.s.$$

Considérons alors le processus d'écart à la droite de Kolmogorov, qui n'est presque jamais nul,

$$S_n(t) = S_{[nt]} + (nt - [nt])(X_{[nt]+1} - \mathbb{E}X).$$

La loi fonctionnelle et le principe d'invariance de Strassen [S] déterminent l'ensemble  $\mathcal{S}_1$  des fonctions presque sûrement adhérentes de la suite de trajectoires  $\{S_n(t)/\sqrt{2n \log \log n} : t \in [0, 1]\}$ . Depuis lors, un grand nombre de théorèmes limites ont décrit  $S_n(t)$  plus précisément, y compris si  $X$  est à valeurs dans un espace de Banach séparable. Par exemple, appliquons à  $S_n(t)$  quelques résultats fins sur le mouvement brownien via [KMT]. D'après [S],  $S_n(t)/\sqrt{2n \log \log n}$  ressemble infiniment souvent à  $\gamma t^{1/2+\varepsilon} \in \mathcal{S}_1$  si  $\varepsilon \in ]0, 1/2]$  et  $|\gamma| \leq \gamma_\varepsilon = \sqrt{2\varepsilon}/(1/2+\varepsilon)$ . Lorsque  $X \in \mathbb{L}_\delta$ ,  $\delta > 2$ , les vitesses d'approche en métrique uniforme ne dépendent que de  $\varepsilon$  et  $\gamma$  car nous tirons facilement de [A2,Cs,Cs2],

$$\liminf_{n \rightarrow \infty} (\log \log n) \sup_{t \in [0,1]} \left| \frac{S_n(t)}{\sqrt{2n \log \log n}} - \gamma t^{1/2+\varepsilon} \right| = \frac{\pi}{4\sqrt{1 - \gamma^2/\gamma_\varepsilon^2}} \quad p.s. \quad (|\gamma| < \gamma_\varepsilon)$$

$$\liminf_{n \rightarrow \infty} (\log \log n)^{2/3} \sup_{t \in [0,1]} \left| \frac{S_n(t)}{\sqrt{2n \log \log n}} - t \right| = \left( \frac{\pi^2}{32} \right)^{1/3} \quad p.s. \quad (\varepsilon = 1/2, \gamma = \gamma_{1/2})$$

et de [6], pour un  $C_\varepsilon$  explicite (voir 3.1.4. avec  $\alpha = 1/2 + \varepsilon$ ) et une puissance entre 2/3 et 1,

$$\liminf_{n \rightarrow \infty} (\log \log n)^{(1+2\varepsilon)/(1+4\varepsilon)} \sup_{t \in [0,1]} \left| \frac{S_n(t)}{\sqrt{2n \log \log n}} - \gamma_\varepsilon t^{1/2+\varepsilon} \right| = C_\varepsilon \quad p.s.$$

D'après [Gr,T]  $S_n(t)/\sqrt{2n \log \log n}$  se maintient, en topologie uniforme, à distance  $O_{p.s.}(1/(\log \log n)^{2/3})$  de  $\mathcal{S}_1$ . Ainsi, dans le domaine d'attraction gaussien, les fluctuations autour de la droite de Kolmogorov à l'échelle  $[0, n]$  sont asymptotiquement très bien décrites. Une littérature moins dense explore les cas frontières, comme [EM], ou les domaines d'attraction  $\alpha$ -stables en topologie adaptée, comme [W].

(ii) Pour découvrir d'autres organisations prévisibles nous zoomons localement sur la trajectoire  $S_n$  aux échelles  $[t, t + A_n]$  intermédiaires entre le court terme oscillant et le long terme régulé comme ci-dessus. Bien que seul le Corollaire 2.10 soit formellement écrit, nos théorèmes de la partie 2 admettent une version pour  $S_n$ , sans changement si  $\mathbb{E}e^{sX} < \infty$ ,  $s \in ]-s_0, s_0[$ . Ces résultats décrivent très précisément les oscillations locales de longueur  $A_n$  telle que  $A_n \gg \log n$ ,  $A_n \ll n$ , définies par

$$\Delta S_n(t, A_n) = \{S_n(t + A_n \cdot) - S_n(t) : \cdot \in [0, 1]\}$$

que produit nécessairement le processus  $S_n$ . L'ensemble  $\{\Delta S_n(t, A_n)/\sqrt{2A_n L_n} : t \in [0, n - A_n]\}$  de ces fluctuations trajectorielles normalisées par  $L_n = \log(n/A_n) + \log \log n$  est relativement compact dans diverses topologies, d'ensembles limites déterministes, avec probabilité un. Les déviations locales à la droite de Kolmogorov sont donc très structurées. Nous produisons en outre, au second ordre, les vitesses de convergence vers ces structures adhérentes ou limites, les vitesses de concentration et

de recouvrement des ensembles limites et fournissons divers compléments, liés au renforcement de la métrique, aux temps d'attente d'événements, aux sous-ensembles etc. Par exemple, d'après 2.1.1, si  $A_n$  n'est pas trop lente alors, pour  $\sigma$  assez grand, les oscillations locales de  $S_n$  se concentrent à vitesse

$$\mathbb{P} \left( \exists n_0 : \sup_{t \in [0, n-A_n]} \inf_{f \in \mathcal{S}_1} \sup_{\cdot \in [0,1]} \left| \frac{\Delta S_n(t, A_n)(\cdot)}{\sqrt{2A_n L_n}} - f(\cdot) \right| \leq \frac{\sigma}{L_n^{2/3}}, n > n_0 \right) = 1.$$

Nos conditions suffisantes sur  $A_n$  s'adaptent si  $\mathbb{E}\varphi(X) < \infty$  pour  $\varphi$  sous-exponentielle et régulière. Établir des outils de déviation dans le cas des  $A_n$  trop rapides est l'une de nos directions récentes de recherche, de même que la description de quelques cas  $X \notin \mathbb{L}_2$  comme dans [23], ainsi que les vitesses dans la loi d'Erdős-Rényi pour  $X \in \mathbb{R}^d$  avec moment exponentiel [22] et  $X$  dans un espace de Banach.

### 1.1.3 – Seconde droite de Kolmogorov

Précisons la notion de répartition spatiale déterministe via un autre exemple universel, celui de la mesure empirique sur  $\mathbb{R}$ , qui sera intensément repris en partie 2. Si  $F(t) = \mathbb{P}(X \leq t)$  est continue en  $t$  alors avec probabilité un, la fonction de répartition empirique uniforme

$$F_n(t) = \frac{1}{n} \sum_{1 \leq i \leq n} \mathbb{I}_{\{F(X_i) \leq t\}}, \quad t \in [0, 1]$$

converge uniformément vers l'identité de  $[0, 1]$ . Divers tests d'adéquation à  $F$  exploitent ce fait.

**Question 1.2.** Vu de plus près, dans quelle mesure  $F_n$  est-elle déterminée à quitter la droite  $t$  ?

Qualifions  $t$  de seconde droite de Kolmogorov et définissons le processus d'écart  $\alpha_n(t) = \sqrt{n}(F_n(t) - t)$ . Kolmogorov a en effet obtenu la convergence faible de  $\sup_{t \in [0,1]} |\alpha_n(t)|$ , Smirnov la loi du logarithme itéré et Donsker la convergence faible de  $\alpha_n$  vers le pont brownien  $B$ , à une célèbre correction topologique près. Depuis lors, le processus fondamental  $\alpha_n$ , son inverse  $\beta_n$  ou leurs généralisations ont été intensément étudiés et utilisés, voir les références de [CH,CS,Du,GS,GZ,SW,VW].

Dans [1,2,3,8,9,14,17,19,20,24] nous étudions le comportement de  $\alpha_n$ ,  $\beta_n$  et  $B$  sur des intervalles  $[t, t + a_n]$ ,  $t \in T_n$ , à différentes échelles  $a_n \rightarrow 0$ . Nos résultats décrivent ainsi divers phénomènes autour de la seconde droite de Kolmogorov et quantifient à quel point la distribution des variables uniformes  $F(X_i)$  est nécessairement non uniforme et bien organisée à la fois, donc  $\alpha_n$  ou  $\beta_n$  différents et proches de toute suite approximante de ponts browniens  $B_n$ . Bien loin de la précision extrême du cas réel, les résultats de [11,25] s'appliquent à la mesure empirique sur un espace quelconque. C'est ce cadre général, évoqué en 1.3.1, 2.3.4, 2.4.5, 3.4 et 4.4.2, qui nous a occupé récemment.

La différence essentielle entre  $\alpha_n$  et  $S_n$  est que  $\alpha_n$  effectue des sauts déterministes à des instants aléatoires, à l'inverse littéral de  $S_n$ . Ceci pose différents problèmes, comme la mesurabilité, mais leurs oscillations sont similaires aux échelles gaussiennes. Ces deux processus renvoient alors au mouvement et au pont browniens, d'où la nécessité de développer des principes d'invariance et des outils précis de contrôle des déviations et des trajectoires browniennes. C'est l'objet en particulier de [4,5,6,7,13,16,18].

Au plan statistique notre compréhension fine de l'écart à la seconde droite de Kolmogorov est exploitée dans [1,10,12,15,21], des corollaires de nos théorèmes pour  $\alpha_n$  ou  $\beta_n$  et quelques résultats latents.

## 1.2 – Plan

(i) Nos travaux portent sur les réponses précises aux questions de type 1.1 ou 1.2 et leurs implications statistiques. D'où l'organisation de ce document en trois catégories de résultats intimement liées.

La première catégorie rassemble des théorèmes décrivant la convergence d'ensembles de processus d'accroissement locaux, sous la forme de lois fonctionnelles du logarithme, itéré ou non, avec vitesses de recouvrement et de concentration exactes, changement de topologie et quelques autres raffinements. Nous nous limitons au cas de processus réels, les résultats étant les plus précis. Sont concernés les

processus empiriques, de quantile, browniens, poissoniens et de sommes partielles, que nous qualifions, pour faire court, de fondamentaux. Cette partie 2 étant la plus longue, nous n'évoquons pas, ou si peu, les variations autour de notre thème principal, telles que la dépendance faible, le cas multivarié, les domaines d'attraction non gaussiens, les extensions trop spécifiques.

La seconde catégorie concerne les outils probabilistes indispensables, qui ont un intérêt en eux-mêmes et relèvent de trois domaines clairement identifiés. Il s'agit des grandes déviations, et de la concentration de la mesure ; des petites déviations, et des probabilités de petites boules ; des principes d'invariance faibles et forts. Dans cette partie 3 nous affichons quelques contributions à chacun de ces domaines.

La troisième catégorie rassemble en partie 4 des théorèmes limites en statistique, principalement en estimation non paramétrique de la densité, de son maximum ou de ses singularités. Bien que ces résultats soient des conséquences des lois fonctionnelles ou des outils précédents, leur déduction n'est pas directe et demande de bien comprendre les phénomènes sous-jacents pour inventer les preuves. Certains ont d'ailleurs demandé de développer des résultats spécifiques figurant en partie 2 et 3.

(ii) Ce plan ne reflète aucune relation d'ordre, car notre démarche scientifique n'est résolument pas à sens unique, chacun des trois aspects se nourrissant des deux autres. En effet, il est satisfaisant pour le probabiliste de déployer les conséquences de ses résultats les plus fins jusqu'à obtenir des théorèmes limites inédits en statistique mathématique, et de motiver ainsi l'étude fine de ses processus favoris. Par exemple, une partie de nos résultats et outils concernent le mouvement brownien, processus fondamental par excellence, et omniprésent à la limite. Certes, mais il n'est pas moins satisfaisant pour le statisticien de découvrir *comment* un problème classique est dirigé par certains processus fondamentaux, ou un objet s'y rapportant, qui peuvent être traités en eux-mêmes et donner lieu à des résultats originaux, tant en probabilité que dans le champ statistique initial. Par exemple, nous avons dû poser à la trajectoire brownienne des questions nouvelles, et raffiner les lois fonctionnelles sur le processus empirique pour pouvoir les utiliser réellement.

Or ce *comment* est d'une complexité désarçonnante qui fait la richesse et la saveur des statistiques mathématiques. Traçons néanmoins notre ligne directrice, en partant du problème fondateur, et pour nous central, des fluctuations d'échantillonnage. L'essentiel de nos axes de recherche s'y rattachent.

## 1.3 – Ligne directrice

### 1.3.1 – Approche générale

Soit  $X, X_1, \dots, X_n$  des variables aléatoires définies sur un espace de probabilité  $(\Omega, \mathcal{T}, \mathbb{P})$ , à valeur dans un espace mesurable  $(\mathcal{X}, \mathcal{A})$ , indépendantes et identiquement distribuées, de loi  $P = \mathbb{P}^X$ . Considérons une statistique d'intérêt sous la forme d'une variable aléatoire  $Y_n = \varphi_n(X_1, \dots, X_n)$  à valeur dans un espace de Banach séparable  $(\mathcal{Y}, \|\cdot\|)$  muni de ses boréliens  $\mathcal{B}$ .

(i) Quelle que soit la problématique statistique initiale menant à  $Y_n$  les deux questions incontournables concernent la loi de  $Y_n$  et le contrôle de  $Y_n - \mathbb{E}Y_n$ . Pour y répondre il faut évaluer, en éludant ici tout problème de mesurabilité, des probabilités de grande déviation du type

$$\mathbb{P} \left( \inf_{y \in \mathcal{B}} \left\| \frac{Y_n}{c_n} - y \right\| > \varepsilon_n \right)$$

et des probabilités de petite déviation du type

$$\mathbb{P} (\|Y_n - c_n y\| < \varepsilon_n)$$

avec  $y \in B \in \mathcal{B}$ ,  $c_n \rightarrow \infty$  et  $\varepsilon_n \rightarrow 0$ . Afin d'éviter de traiter les  $Y_n$  un par un, une approche plus intrinsèque consiste à réduire le problème à l'action d'un *processus élémentaire* et de voir en quoi l'aléa d'une famille de tels  $Y_n$  est guidé par les oscillations de cet unique processus sous-jacent.

Or  $Y_n$  est d'autant plus éloigné de ses valeurs typiques que les  $X_1, \dots, X_n$  sont mal répartis dans  $\mathcal{X}$  par rapport à  $P$ . Appelons ce phénomène inévitable les *fluctuations d'échantillonnage*. Il constitue notre

préoccupation majeure, bien plus, en général, que les valeurs typiques  $\mathbb{E}Y_n$ , qui relèvent de chaque situation particulière et souvent d'arguments déterministes. Ces écarts incompressibles, et parfois extrêmes, à la répartition idéale inconnue  $P$  sont liés au comportement du processus empirique

$$\alpha_n = \sqrt{n}(\mathbb{P}_n - P), \quad \mathbb{P}_n = \frac{1}{n} \sum_{i \leq n} \delta_{X_i}$$

où  $\delta_X(A) = \mathbb{I}_{X \in A}$  pour  $A \in \mathcal{A}$ . Regardons  $\alpha_n$  comme indexé par une classe d'ensembles  $\mathcal{C} \subset \mathcal{A}$  ou de fonctions  $\mathcal{F} \subset \mathbb{L}_2(P)$  en posant  $f = \mathbb{I}_C$  si  $C \in \mathcal{C}$  puis, si  $f \in \mathcal{F}$ ,

$$\mathbb{P}_n f = \frac{1}{n} \sum_{i \leq n} f(X_i), \quad Pf = \mathbb{E}f(X).$$

Le fait que  $\alpha_n(f)$  s'éloigne de 0 signifie que les  $X_i$  sont si mal répartis que les moyennes empiriques  $\mathbb{P}_n g$  estiment au plus mal les espérances  $Pg$  d'une infinité de  $g$  voisins de  $f$ .

**Question 1.3.** Notre problème général consiste à mesurer à quel point, et de traduire en quoi, la mesure aléatoire  $\mathbb{P}_n$  peut et doit être une mauvaise approximation de  $P$  en terme de  $Y_n - \mathbb{E}Y_n$ .

(ii) Pour contrôler les variations de  $Y_n$  par celles de  $\alpha_n$ , associons à  $Y_n$  une classe  $\mathcal{F}$  et  $\psi_n$  opérant sur  $\mathbb{L}_\infty(\mathcal{F})$  tels qu'en notant  $Z_n = \psi_n(\alpha_n)$ , la quantité  $\|Y_n - Z_n\|$  soit stochastiquement négligeable. Dans cette optique, quantifier les oscillations du  $P$ -processus empirique  $\alpha_n$  autour de 0 donne une clef pour étudier simultanément un grand nombre de  $Z_n$  donc de  $Y_n$ . Soit  $h \in \mathcal{H} \subset \mathbb{L}_\infty(\mathcal{F})$  et  $\|h\|_{\mathcal{F}} = \sup_{\mathcal{F}} |h|$ . Les probabilités de déviations du type

$$\mathbb{P} \left( \inf_{h \in \mathcal{H}} \left\| \frac{\alpha_n}{c_n} - h \right\|_{\mathcal{F}} > \varepsilon_n \right), \quad \mathbb{P} (\|\alpha_n - c_n h\|_{\mathcal{F}} < \varepsilon_n)$$

sont alors à évaluer une fois pour toutes, de même que l'amplitude du module de continuité,

$$\mathbb{P}(\omega_n(\varepsilon_n) > c_n \sqrt{\varepsilon_n}), \quad \omega_n(\varepsilon) = \sup_{\rho(f,g) < \varepsilon} |\alpha_n(f) - \alpha_n(g)|$$

pour une distance  $\rho$  sur  $\mathcal{F}$ . En toute généralité ceci n'est pas possible avec précision car tout dépend de  $P$  et  $\mathcal{F}$ . Néanmoins, bien que des hypothèses s'imposent sur  $P$ ,  $\mathcal{F}$  et  $\psi_n$  cette démarche est générique, requiert des outils utiles par ailleurs et fait appel à l'étude d'autres processus dignes d'intérêt.

Les théorèmes limites pour  $\alpha_n$  obtenus hors contexte donnent accès, après plus ou moins d'efforts selon la nature de  $\psi_n$ , à la loi forte des grands nombres et au théorème limite central pour  $Z_n$ , donc  $Y_n$ , avec parfois les vitesses associées, presque sûres ou en loi. Nous ne présentons que des résultats presque sûrs avec vitesses exactes, les convergences en loi s'en déduisant en rectifiant les preuves.

(iii) A ce stade, une technique avantageuse consiste à substituer un  $P$ -pont brownien  $\mathbb{G}_n$  à  $\alpha_n$  pour approcher  $Z_n$  par  $G_n = \psi_n(\mathbb{G}_n)$ . C'est utiliser un outil puissant, le principe d'invariance fort qui, lorsqu'il est disponible, stipule l'existence sur un même  $(\Omega, \mathcal{T}, \mathbb{P})$  de versions de  $(\mathbb{G}_n, X_1, \dots, X_n)$  vérifiant

$$\mathbb{P} (\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > \eta_n) \leq \lambda_n$$

avec  $\eta_n \rightarrow 0$  et  $\lambda_n \rightarrow 0$  explicites, voir à ce propos 3.4 et 2.1.5. Lorsque  $\mathcal{X} = \mathbb{R}^d$  et  $\mathcal{Y} = \mathbb{R}$  cette étape d'approximation forte est déjà intégrée dans la preuve des résultats sur  $\alpha_n$  présentés en partie 2. Cependant, il convient parfois de se ramener d'emblée au processus gaussien  $\mathbb{G}_n$  par invariance forte, comme en 4.2 et 4.3 où l'événement clef est plus facile à évaluer dans le cas brownien grâce à 3.3.

### 1.3.2 – Approche spécifique

(i) Supposons que  $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ , de sorte que  $\mathcal{F} = \{\mathbb{I}_{(-\infty, t]} : t \in \mathbb{R}\}$ . Notre réponse à la question 1.3 concerne les  $Y_n$  sensibles aux fluctuations locales d'échantillonnage, décrites par les incréments de  $\alpha_n$

$$\Delta \alpha_n(a_n, t) = \{\alpha_n(t + a_n \cdot) - \alpha_n(t) : \cdot \in [0, 1]\}$$

avec  $a_n \rightarrow 0$  déterminé. Plus  $\Delta\alpha_n(a_n, t)$  reste proche de 0, plus les  $X_i \in [t, t + a_n]$  sont bien répartis par rapport à  $P$ , au sens de l'évaluation simultanée des espérances de  $\{\mathbb{I}_{[t, t+a_n]} : \cdot \in [0, 1]\}$ . Posons

$$\Theta_{j,n} = \left\{ \frac{\Delta\alpha_n(a_n, t)}{b_n} : t \in T_{j,n} \right\}$$

avec  $b_n > 0$ . Supposons que  $|Y_n - Z_n|$  est stochastiquement négligeable avec

$$Z_n = \Gamma(\Theta_{1,n}) + \Phi(\Theta_{2,n})$$

où  $T_{1,n}$  désigne les localisations sensibles et  $T_{2,n}$  tout ou partie des localisations en jeu. En partie 2 nous présentons des théorèmes de relative compacité presque sûre avec vitesse pour  $\Theta_{j,n}$  dont l'un des intérêts est de permettre de contrôler très précisément de tels  $Z_n$ . En partie 4 nous avons typiquement besoin de  $T_{1,n}$  petit ou aléatoire dans le terme oscillant principal  $\Gamma(\Theta_{1,n})$  et de  $T_{2,n}$  grand dans le terme d'erreur  $\Phi(\Theta_{2,n})$ . Par nos arguments nous pouvons aussi traiter d'autres types d'ensembles aléatoires de trajectoires que les accroissements  $\Theta_{1,n}$  et  $\Theta_{2,n}$  d'un même processus, voir 2.2.1.

**(ii)** Même dans la situation de  $Z_n$  ci-dessus, très simplifiée en apparence, de nombreux problèmes se posent. D'une part,  $\alpha_n$  n'est pas à accroissements indépendants, et nous devons approcher  $\alpha_n$  ou  $\Delta\alpha_n(a_n, t)$  par des processus ponctuels de Poisson puis si possible, par des processus gaussiens, comme un pont ou un accroissement brownien. Certaines circonstances peuvent demander d'introduire une marche aléatoire. D'autre part,  $\Gamma$  et  $\Phi$  peuvent s'avérer si délicats que les ensembles aléatoires  $\Theta_{j,n}$  doivent s'étudier à travers des variantes spécifiques, en topologie forte ou avec des  $T_{j,n}$  exotiques, comme en 4.3. Enfin, le terme principal  $\Gamma(\Theta_{1,n})$  peut exiger des théorèmes de déviations particuliers, comme en 4.2. Ces besoins relèvent des parties 2 et 3.

Une intuition typique de notre approche est la suivante. Imaginons que  $Z_n = \psi_n(\alpha_n) = \Gamma(\Theta_{1,n}) \in \mathbb{R}^d$  s'éloigne d'autant plus de  $\mathbb{E}Z_n$  que  $\alpha_n$  présente une certaine structure locale  $\Delta\alpha_n(a_n, t)/b_n \sim g$  critique au voisinage d'au moins une localisation  $t \in T_{1,n}$ , critique au sens de  $\Gamma$ . Alors un théorème limite très utile consiste à déterminer à quelle vitesse presque sûre  $\Theta_{1,n}$  approche la cible  $g$ . C'est l'objet en particulier de 2.1.2, 2.1.5, 2.1.6, 2.2.4 et 2.3.3. Si au contraire  $\Gamma$  dévie lorsqu'un élément au moins de  $\Theta_{1,n}$  s'éloigne sensiblement de la structure déterministe limite  $\mathcal{S}$  alors la vitesse de concentration de  $\Theta_{1,n}$  dans  $\mathcal{S}$  est cruciale. C'est l'objet notamment de 2.1.1, 2.1.6, 2.2.5 et 2.3.3.

Ainsi la partie 2 résume-t-elle notre connaissance du comportement déterministe des fluctuations d'échantillonnage sur  $\mathbb{R}$ . Elle va bien au delà puisque nos méthodes et résultats se transposent à d'autres processus, mais nous focalisons sur le domaine d'attraction brownien.

**(iii)** Évidemment, il y a autant de problèmes en statistique que d'espaces  $\mathcal{X}$ ,  $\mathcal{Y}$  et les variables d'intérêt  $Y_n$  sont formidablement variées. D'où la richesse des questionnements évoquée plus haut. Cependant il n'y a qu'un seul processus empirique  $\alpha_n$  sous-jacent et son  $P$ -pont brownien associé, et leur utilisation systématique a porté des fruits uniques en leur genre depuis une quarantaine d'années. En outre, même si le passage de  $\alpha_n$  à  $Y_n$  reste un problème en soi, la compréhension fine des processus associés à  $\alpha_n$ , de leurs oscillations et des outils appropriés est bénéfique. La partie 4 s'en veut l'illustration.

Pour conclure cette entrée en matière, revenons sur notre motivation 1.1 sous un angle statistique. Un déterminisme omniprésent en statistique asymptotique est le comportement limite gaussien tant recherché. A tel point que nombre d'utilisateurs, à la fois bien et mal instruits, ne soupçonnent pas qu'il existe des phénomènes non gaussiens, et ce d'autant moins lorsqu'ils sont engendrés par des processus presque gaussiens. Dans une partie de nos travaux nous déployons des efforts pour déterminer les échelles auxquelles en effet l'aspect gaussien prévaut pour tout  $P$ . A l'inverse, un aléa persistant et non gaussien se manifeste dès que l'asymptotique n'est pas de mise ou que les échelles critiques entrent en jeu. Une autre partie de nos travaux explore cette frontière de non invariance et les différences structurelles avec les phénomènes gaussiens. Ainsi les écarts aux droites de Kolmogorov se comportent-ils de façon gaussienne ou révélatrice de la loi sous-jacente, selon l'échelle d'observation, mais dans les deux cas nous nous attachons avant tout à décrire l'inéluctable.

## 2 – CONVERGENCE DES INCRÉMENTS : LOIS FONCTIONNELLES DU LOGARITHME

La statistique mathématique s'interroge sur la répartition de variables aléatoires (v.a.)  $X_n$  dans leur espace  $(\mathcal{X}, \mathcal{A}, P)$  : qu'en déduire sur  $P$ ? Les objets élémentaires sont donc les mesures aléatoires, telles que  $\mathbb{P}_n - P$ , qui sont indexées par  $\mathcal{A}$  ou  $\mathcal{A}^- \subset \mathcal{A}$ , et les processus privilégiés sont fonction de l'espace et non d'un temps. Néanmoins, sur  $\mathcal{X} = \mathbb{R}$ , en identifiant les intervalles  $[0, t]$  à un temps  $t$ , les deux aléas évoqués en 1.1.1 coïncident de façon trompeuse mais fructueuse, et nous avons intensivement exploré l'aspect trajectorien de cette situation singulière. D'où l'usage d'arguments de type martingale, temps d'arrêt, franchissement de frontière, déviation trajectorielle, temps d'occupation, tension etc.

Dans le cas de v.a. réelles  $X_1, \dots, X_n$  indépendantes et identiquement distribuées (i.i.d.) nous nous ramenons, par la transformation des quantiles, à l'étude de v.a.  $U_1, \dots, U_n$  i.i.d. de loi uniforme sur  $]0, 1[$ . Soit  $I$  l'identité de  $[0, 1]$  et  $\mathcal{B}$  l'espace de Banach des fonctions réelles bornées sur  $[0, 1]$ , nulles en 0, muni de la norme  $\|f\| = \sup_{[0,1]} |f|$ . Notons  $B_0 = \{f : \|f\| < 1\}$  sa boule unité. Considérons le processus empirique uniforme, dont l'étude approfondie est le leitmotif de la littérature classique [CS,GS,SW,V] et le guide de ses multiples généralisations [CH,Du,DP,G,GZ,VW],

$$\alpha_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbb{I}_{\{U_i \leq I\}} - I) \in \mathcal{B}.$$

Son incrément local en  $t \in [0, 1]$ , de longueur  $a > 0$  est défini par  $\Delta\alpha_n(a, t) = \alpha_n(t + aI) - \alpha_n(t) \in \mathcal{B}$ . Dans cette partie nous présentons nos résultats décrivant le comportement limite presque sûr, et notamment les vitesses de convergence, des suites d'ensembles aléatoires du type

$$\Theta_n = \left\{ \frac{\Delta\alpha_n(a_n, t)}{\sqrt{2a_n L_n}} : t \in T_n \right\}$$

avec  $a_n > 0$ ,  $L_n > 0$ ,  $T_n \subset [0, 1]$  en général déterministes. Lorsque les ensembles limites de  $\Theta_n$  ne sont pas dégénérés,  $\Theta_n$  décrit précisément la répartition des  $U_i$  dans  $[0, 1]$ , qui est bien loin d'être uniforme. Plus  $\|\Delta\alpha_n(a_n, t)\|$  est détaché de 0 moins les  $U_i$  s'équi-répartissent dans  $[t, t + a_n]$ . Comprendre cette non uniformité structurelle, purement stochastique mais asymptotiquement bien organisée, c'est se donner les moyens de contrôler les fluctuations d'échantillonnage de  $X_1, \dots, X_n$  et les déviations de  $Y_n = \varphi_n(X_1, \dots, X_n)$  pour certains  $\varphi_n$ , notamment ceux de la forme 1.3.2. et de la partie 4.

En 2.1 ([2,7,8,9,18,20]) nous présentons nos théorèmes limites fins en topologie uniforme pour  $\Theta_n$  avec  $T_n$  grand, ainsi que pour les incréments de processus associés, notamment le mouvement brownien, et en tirons quelques enseignements, comme la vitesse optimale d'approximation gaussienne.

En 2.2 ([3,5,13,17]) nous décrivons nos résultats en topologie de type höldérienne pour  $\Theta_n$  ou d'autres ensembles de trajectoires, en mettant l'accent sur notre remède unique en son genre à l'impossibilité de contrôler le processus de saut  $\alpha_n$  en métrique renforcée : nous évaluons convergence et vitesses de  $\Theta_n$  le long d'une suite d'espaces métriques réglés pour que  $\alpha_n$  imite le mouvement brownien.

En 2.3 ([14,16]) nous exposons quelques extensions des résultats précédents lorsque  $T_n$  est petit, en détaillant en particulier l'extrapolation qui s'ensuit entre les versions fonctionnelles de deux résultats de Lévy, sur les modules global et local d'oscillation brownienne.

En 2.4 ([19,22,23,24]) nous évoquons brièvement nos autres recherches en cours sur les oscillations des trajectoires stochastiques, notamment des marches aléatoires, et les points exceptionnels.

### 2.1 – Recouvrement de Strassen en topologie uniforme

**Tailles d'incrément.** Posons  $c_n = |\log a_n| / \log \log n$ ,  $d_n = na_n / \log n$  et imposons que

$$a_n \leq 1, \quad a_n \downarrow 0, \quad na_n \uparrow \infty, \quad c_n \rightarrow c, \quad d_n \rightarrow d \tag{1}$$

avec  $c \in [0, \infty]$  et  $d \in [0, \infty]$ . Nous parlons d'incrémentes *très grands* si  $|\log a_n| \ll \log \log \log n$ , *grands* si  $c < \infty$ , *moyens* si  $c = \infty$  et  $na_n \gg (\log n)^3$ , *petits* si  $na_n = O((\log n)^3)$  et  $d = \infty$ , *très petits* si  $d < \infty$ . Dans ce document nous laissons de côté les très grands et les très petits incréments.

**Fonctions limites.** Si  $f$  est absolument continue par rapport à la mesure de Lebesgues  $\lambda$  sur  $[0, 1]$  notons  $f'$  une version de sa dérivée et  $J(f) = \int_{[0,1]} f'^2 d\lambda$ . Sinon, posons  $J(f) = \infty$ . Lorsque  $d = \infty$  les ensembles limites naturels de  $\Theta_n$  sont les boules de Strassen [S]

$$\mathcal{S}_\theta = \{f \in \mathcal{B} : J(f) \leq \theta\}. \quad (2)$$

Les intérieurs  $\mathcal{S}_\theta^-$  et les frontières  $\mathcal{S}_\theta^{vb} \cup \mathcal{S}_\theta^{vli}$ , non topologiques, jouent un rôle important et sont définis en 3.1.1. Soit  $h(x) = x \log x - x + 1$ . Dans le cas des petits incréments s'opère une transition entre  $\mathcal{S}_1$  et la boule d'Orlicz  $\mathcal{O}_d = \{(g - dI)/\sqrt{2d} : g \uparrow, d \int_{[0,1]} h(g'/d) \leq 1\}$  liée à une loi de type [ERé], voir [DM,D]. L'une de nos perspectives tenaces est de décrire complètement cette transition, comme amorcé dans [1,8] et évoqué en 2.1.3 et 3.5.3(ii-iii).

**Tailles d'ensemble.** Nous parlons de lois *locales* ou *globales* selon que  $\Theta_n$  se comporte comme si  $T_n = \{0\}$  ou  $T_n = [0, 1 - a_n]$ . Nos lois *intermédiaires* 2.3 interpolent entre ces deux situations. La loi locale de [M2] atteste que si  $T_n = \{0\}$  la suite  $\Delta\alpha_n(a_n, 0)/\sqrt{2a_n \log \log n}$  adhère presque sûrement à  $\mathcal{S}_1$ . D'après les lois globales de [DM,D2], pour  $T_n = [0, 1 - a_n]$  et  $L_n = |\log a_n| + \log \log n$  la suite d'ensembles  $\Theta_n$  est presque sûrement relativement compacte dans  $\mathcal{B}$  avec, si  $d = \infty$ , les points adhérents  $\mathcal{S}_1$  et les points limites  $\mathcal{S}_{c/(1+c)}$  ou, si  $d \in ]0, \infty[$ , les points limites  $\mathcal{O}_d$ .

Dans cette section 2.1 nous précisons les vitesses de recouvrement et les similitudes ou différences avec les ensembles de trajectoires gaussiennes. Nos vitesses exactes sont les premières jamais obtenues dans des théorèmes de convergence d'ensembles de processus. Nous avons ainsi développé nos propres techniques et astuces de blocage, la précision recherchée demandant non seulement des ingrédients puissants du type 3.1 mais aussi une certaine dextérité dans la manipulation des événements.

### 2.1.1 – Vitesses de concentration

Supposons que  $T_n = [0, 1 - a_n]$  et  $L_n = |\log a_n| + \log \log n$ . Rappelons le phénomène de concentration et de recouvrement de [DM,D2], inspiré de [R], qui synthétise élégamment diverses assertions.

**Théorème 2A.** *Sous (1),  $\mathcal{S}_1$  et  $\mathcal{S}_{c/(1+c)}$  sont le plus petit et le plus grand compacts de  $\mathcal{B}$  tels que pour tout  $\varepsilon > 0$ ,  $\mathbb{P}(\exists n_0 : \mathcal{S}_{c/(1+c)} \subset \Theta_n + \varepsilon B_0 \subset \mathcal{S}_1 + 2\varepsilon B_0, n > n_0) = 1$  et  $\mathbb{P}(\mathcal{S}_{c/(1+c)} \subset \Theta_n + \varepsilon B_0) \rightarrow 1$ .*

(i) Dans [1,2] nous avons obtenu les vitesses de concentration autour des ensembles limites. Elles sont lentes et non sans rapport avec les vitesses d'approche des fonctions les plus inaccessibles, voir 2.1.2 et 2.1.6(i). Outre l'aspect très technique des preuves, l'argument final repose sur un outil plus précis que le principe de grandes déviations, l'inégalité isopérimétrique de [T] et ses variantes de type [Le].

**Théorème 2.1.** *Il existe une constante universelle  $\sigma_1 \in ]0, \infty[$  telle que si (1),  $na_n/(\log n)^{7/3} \rightarrow \infty$  et  $L_n = |\log a_n| + \log \log n$  nous avons, pour tout  $\sigma > \sigma_1$ ,*

$$\mathbb{P}\left(\exists n_0 : \Theta_n \subset \mathcal{S}_1 + \frac{\sigma}{L_n^{2/3}} B_0, n > n_0\right) = 1.$$

Cette vitesse est optimale car en mêlant l'argumentation de la réciproque de [T] à notre technique de preuve nous avons exhibé depuis [1,2] une constante universelle  $\sigma_0 < \sigma_1$  telle que si  $\sigma < \sigma_0$

$$\mathbb{P}\left(\exists n_k \uparrow \infty : \Theta_{n_k} \not\subset \mathcal{S}_1 + \frac{\sigma}{L_{n_k}^{2/3}} B_0, k \geq 1\right) = 1.$$

La preuve de l'inégalité isopérimétrique gaussienne repose sur l'identification de  $\mathcal{S}_1$ , pour la mesure de Wiener, à la boule unité de l'espace auto-reproduisant associé au produit infini de gaussiennes sur  $\ell_2$ , et sur le contrôle de  $\sigma B_0$  par un cube inclus dans son image par l'isomorphisme de [Ci2]. Le passage



d'une dilatation par boule sur la base de Haar-Schauder à une dilatation par cube infini dans  $\ell_2$  génère une petite perte géométrique qui nous empêche de combler le trou entre  $\sigma_0$  et  $\sigma_1$ .

**Application.** Considérons deux statistiques dépendant de la pire fluctuation d'échantillonnage. Soit  $\varphi$  suffisamment régulière de  $\mathcal{B}$  dans  $\mathbb{R}$  et  $Y_n = \sup_{g \in \Theta_n} \varphi(g) \geq Z_n = \inf_{g \in \Theta_n} \varphi(g)$ . D'après le Théorème 2.1, avec probabilité un il se produit, pour un  $\zeta_\varphi > 0$  déterministe et lorsque  $n \rightarrow \infty$ ,

$$\inf_{f \in \mathcal{S}_1} \varphi(f) - \frac{\zeta_\varphi}{L_n^{2/3}} < Z_n \leq Y_n < \sup_{f \in \mathcal{S}_1} \varphi(f) + \frac{\zeta_\varphi}{L_n^{2/3}}.$$

De même, en utilisant les modules de continuité  $\pi_n = \sup_{\|f-g\| < I} |\varphi_n(f) - \varphi_n(g)|$  d'une suite de fonctionnelles  $\varphi_n$  uniformément continues sur  $\mathcal{B}$  nous avons presque sûrement, si  $\sigma > \sigma_1$  et  $n \rightarrow \infty$ ,

$$\inf_{f \in \mathcal{S}_1} \varphi_n(f) - \pi_n \left( \frac{\sigma}{L_n^{2/3}} \right) < \inf_{t \in T_n} \varphi_n \left( \frac{\Delta \alpha_n(a_n, t)}{\sqrt{2a_n L_n}} \right) \leq \sup_{t \in T_n} \varphi_n \left( \frac{\Delta \alpha_n(a_n, t)}{\sqrt{2a_n L_n}} \right) < \sup_{f \in \mathcal{S}_1} \varphi_n(f) + \pi_n \left( \frac{\sigma}{L_n^{2/3}} \right).$$

(ii) Pour des petits incréments, tels que  $d = \infty$  et  $na_n/(\log n)^{7/3} \rightarrow 0$ , la contrainte d'approximation forte interdit de conclure à une vitesse optimale, et nous n'avons pas encore trouvé d'alternative à l'isopérimétrie gaussienne. La concentration prouvée peut alors s'avérer arbitrairement lente.

**Théorème 2.2.** *Il existe une constante universelle  $\sigma_2 \in ]0, \infty[$  telle que si (1),  $na_n < d_0(\log n)^{7/3}$ ,  $d = \infty$  et  $L_n = \log n$  nous avons, pour tout  $\sigma > \sigma_2$ ,*

$$\mathbb{P} \left( \exists n_0 : \Theta_n \subset \mathcal{S}_1 + \sigma \sqrt{d_n} B_0, n > n_0 \right) = 1.$$

Cependant les résultats de 2.1.5 nous font conjecturer que cette vitesse d'inclusion est la bonne et quantifie la transition vers l'ensemble limite d'Erdős-Rényi  $\mathcal{O}_d$  pour la loi de Poisson centrée, approprié au cas  $d_n \rightarrow d$ . Cette question est à l'étude, sous la forme d'un principe de concentration mêlant des termes gaussiens et poissonniens en compétition selon l'intensité de la partie poissonnienne.

(iii) Du point de vue de la convergence faible nous obtenons, sous les conditions du Théorème 2.1,

$$\mathbb{P} \left( \mathcal{S}_{c/(1+c)} \subset \Theta_n + \frac{\sigma}{L_n^{2/3}} B_0 \right) \rightarrow \begin{cases} 1 & \text{si } \sigma > \sigma_1, \\ 0 & \text{si } \sigma < \sigma_0. \end{cases}$$

De même, pour la concentration récurrente nous produisons des vitesses qui s'appliquent à chaque suite  $\Theta_{n_j}$  dont nous extrayons une sous-suite  $\Theta_{n_k}$  à vitesse d'inclusion optimale. Sous (1) avec  $c < \infty$  il vient, pour tout  $b \in ]c/(1+c), 1]$ ,

$$\mathbb{P} \left( \exists n_k \uparrow \infty : \Theta_{n_k} \subset \mathcal{S}_b + \frac{\sigma}{L_{n_k}^{2/3}} B_0, \Theta_{n_k} \not\subset \mathcal{S}_b, k \geq 1 \right) = 1$$

et nous pouvons encore croiser avec des événements de type petite déviation du 2.1.2 ci-après.

(iv) Considérons enfin les grands incréments tels que (1),  $c = 0$  et  $a_n \ll 1/(\log \log n)^\gamma$  pour tout  $\gamma > 1$ . Si  $L_n = \log \log n$  et  $|\log a_n| \gg (\log \log \log n)^3$  il vient, pour tout  $\sigma > \sigma_1$

$$\mathbb{P} \left( \exists n_k \uparrow \infty : \Theta_{n_k} \subset \sqrt{c_{n_k}} \mathcal{S}_1 + \frac{\sigma}{\sqrt{\log \log n_k} |\log a_{n_k}|^{1/6}} B_0, k \geq 1 \right) = 1.$$

Si  $L_n = \log \log n$  et  $\log \log \log n \ll |\log a_n| = O((\log \log \log n)^3)$  et  $\sigma$  assez grand,

$$\mathbb{P} \left( \exists n_k \uparrow \infty : \Theta_{n_k} \subset \sqrt{c_{n_k}} \mathcal{S}_1 + \frac{\sigma \log \log \log n_k}{\sqrt{\log \log n_k} |\log a_{n_k}|} B_0, k \geq 1 \right) = 1.$$

La concentration des très grands incréments est plus délicate à décrire car les  $\Delta \alpha_n(a_n, t)/\sqrt{2a_n L_n}$  ne s'organisent pas autour de 0 selon  $\mathcal{S}_1$  et un désordre persiste, voir [1]. Une version fonctionnelle de

[CsR,Sh] pour le processus de Wiener est à l'étude dans [18] et s'applique intégralement ici, avec une renormalisation en  $L_n = |\log a_n| - \log \log \log n$ .

**(v) Principe de démonstration.** Au Théorème 2.1 nous mettons au point une technique de blocage encadrant le phénomène le long d'une suite  $a_{n_k}$  en distinguant différentes situations, par exemple  $c < \infty$  et  $c = \infty$ . Un des arguments permettant de maîtriser simultanément les  $\Theta_n$  pour  $n \in ]n_k, n_{k+1}]$  consiste à regrouper les incréments voisins et contrôler leurs écarts uniforme par une variante de [St,DEMR] affirmant que si  $a_n^* \downarrow 0$  vérifie  $\lim_{\theta \downarrow 1} \lim_{k \rightarrow \infty} a_{\theta^{k+1}}^*/a_{\theta^k}^* = 1$  et  $na_n^* \gg \log n$  si  $n \rightarrow \infty$  alors

$$\limsup_{n \rightarrow \infty} \frac{\sup_{0 < t \leq 1 - a_n^*} \|\Delta \alpha_n(a_n, t)\|}{\sqrt{2a_n^* (|\log a_n^*| + \log \log n)}} \leq 1 \quad p.s.$$

mais si  $d_n^* = na_n^*/\log n \rightarrow d < \infty$  et  $|\log a_n^*| \sim \log n$  il vient au contraire, pour  $\sigma_d$  explicite,

$$\limsup_{n \rightarrow \infty} \frac{\sqrt{n} |\log d_n^*|}{\log n} \sup_{0 < t \leq 1 - a_n^*} \|\Delta \alpha_n(a_n, t)\| \leq \sigma_d < \infty \quad p.s.$$

Parmi les autres outils nous utilisons la poissonisation de [DM], le principe d'invariance fort [KMT] et des bornes exponentielles telles que [MSW] qui garantit, pour  $a \in (0, 1/2]$ ,  $c_\eta > 0$ ,  $\eta > 0$  et  $\delta > 0$ ,

$$\mathbb{P} \left( \sup_{0 < t \leq 1 - a} \|\Delta \alpha_n(a, t)\| \geq \delta \sqrt{a} \right) \leq \frac{c_\eta}{a} \exp \left( -\frac{na}{1 + \eta} h \left( 1 + \frac{\delta}{\sqrt{na}} \right) \right).$$

Mais surtout, lorsque  $\gamma \rightarrow \infty$ ,  $\varepsilon \rightarrow 0$  et  $\gamma\varepsilon \rightarrow \infty$ , en repondérant [T,Le] via  $1/p + 1/q = 1$  la concentration autour de  $\bigcup_{\theta > 0} \mathcal{S}_\theta$  de la mesure de Wiener  $w$  sur  $\mathcal{B}$  est bien contrôlée par

$$1 - w(\gamma \mathcal{S}_1 + \varepsilon B_0) \leq 2 \exp \left( \frac{q^2 \sigma_1}{4\varepsilon^2} - \frac{\gamma\varepsilon}{q} - \frac{\gamma^2}{2} \right). \quad \square$$

### 2.1.2 – Vitesses d'approche ponctuelles

Notons  $d(f, S) = \inf_{g \in S} \|f - g\|$ . La loi de Kuelbs-Chung-Mogulsk'iï [Ch,K,Mo]

$$\liminf_{n \rightarrow \infty} \sqrt{\log \log n} \|\alpha_n\| = \frac{\pi}{\sqrt{8}} \quad p.s.$$

signifie que les  $U_i$  ne peuvent pas s'équi-répartir exactement dans  $[0, 1]$ , puisque  $\|\alpha_n\|$  n'atteint jamais 0. Cette loi quantifie donc une forme de désordre incompressible et se réécrit, en posant  $a_n = 1$ ,  $T_n = \{0\}$  et  $L_n = \log \log n$  dans  $\Theta_n$ ,

$$\liminf_{n \rightarrow \infty} (\log \log n) d(0, \Theta_n) = \frac{\pi}{4} \quad p.s.$$

Signalons au passage que la structure des  $U_i$  reste mal comprise quand  $\|\alpha_n\| < (1 + \varepsilon)\pi/\sqrt{8 \log \log n}$ . Lorsque  $a_n \rightarrow 0$  donnons le résultat local valable pour  $T_n$  fini, tiré de [M2].

**Théorème 2B.** *Si  $T_n = \{t\}$ ,  $a_n \downarrow 0$ ,  $na_n \uparrow \infty$ ,  $na_n \gg \log \log n$ ,  $L_n = \log \log n$  et  $f \in \mathcal{B}$  alors*

$$d(f, \mathcal{S}_1) = \liminf_{n \rightarrow \infty} d(f, \Theta_n) \leq \limsup_{n \rightarrow \infty} d(f, \Theta_n) = \sup_{g \in \mathcal{S}_1} d(f, g) \quad p.s.$$

Plus précisément, les points adhérents de  $d(f, \Theta_n)$  sont  $[d(f, \mathcal{S}_1), \sup_{g \in \mathcal{S}_1} d(f, g)]$ . A l'inverse, lorsque  $T_n$  est grand, certains  $f$  sont éternellement visités par les incréments, d'après [DM,D2].

**Théorème 2C.** *Sous (1),  $L_n = |\log a_n| + \log \log n$ ,  $d = \infty$  et  $|\log \lambda(T_n)| \ll L_n$  avec  $T_1 \subset T_2 \subset \dots$  ou  $T_1 \supset T_2 \supset \dots$ , tout  $f \in \mathcal{B}$  vérifie*

$$d(f, \mathcal{S}_1) = \liminf_{n \rightarrow \infty} d(f, \Theta_n) \leq \limsup_{n \rightarrow \infty} d(f, \Theta_n) = d(f, \mathcal{S}_{c/(1+c)}) \quad p.s.$$

Par conséquent, si  $T_n = [0, 1 - a_n]$  chaque  $f \in \mathcal{S}_1$  est presque sûrement valeur d'adhérence d'une suite  $f_n \in \Theta_n$  et chaque  $f \in \mathcal{S}_{c/(1+c)}$  point limite d'une suite  $f_n \in \Theta_n$ . En  $f = 0$  il s'ensuit

$$\inf_{t \in T_n} \|\Delta\alpha_n(a_n, t)\| \ll \sqrt{2a_n L_n} \quad p.s.$$

Se pose alors la question d'une version locale et uniforme de la loi de Chung-Mogulsk'ii donnant les vitesses exactes dans 2B, 2C et pour  $\inf_{t \in T_n} \|\Delta\alpha_n(a_n, t)\|$ . Ceci afin de contrôler de façon déterministe les schémas de répartition presque sûrs des  $U_i$  dans  $T_n + [0, a_n]$  et de préciser le désordre.

(i) Tout d'abord, un résultat local obtenu dans [9] si  $f \in \mathcal{S}_1^{vb} \cup \mathcal{S}_1^{vli}$  et [D3] pour  $f \in \mathcal{S}_1^-$ . Pour les  $na_n \gg \log \log n$  plus petits voir [Sm]. Les fonctions  $\nabla_f$  et  $\chi_f$  sont décrites en 3.1.3.

**Théorème 2.3.** *Si  $f \in \mathcal{S}_1^*$ ,  $T_n = \{t\}$ ,  $L_n = \log \log n$ , (1) et  $na_n \gg (\log \log n) \nabla_f^2(\log \log n)$  alors*

$$\liminf_{n \rightarrow \infty} \nabla_f(\log \log n) d(f, \Theta_n) = \chi_f(1) \quad p.s.$$

**Application.** Si  $Y_n = \varphi(\Delta\alpha_n(a_n, t)/\sqrt{2a_n \log \log n})$  avec  $\varphi$  assez contrastée en  $f_0 = \arg \min_{\mathcal{S}_1} \varphi \in \mathcal{S}_1^*$  et  $f_1 = \arg \max_{\mathcal{S}_1} \varphi \in \mathcal{S}_1^*$  il vient presque sûrement, pour  $\zeta_\varphi > 0$  et  $n$  assez grand,

$$\varphi(f_0) - \frac{\zeta_\varphi}{\nabla_{f_0}(\log \log n)} < Y_n < \varphi(f_1) + \frac{\zeta_\varphi}{\nabla_{f_1}(\log \log n)}$$

mais aussi, pour deux suites aléatoires  $n_{0,k} \rightarrow \infty$  et  $n_{1,k} \rightarrow \infty$ ,

$$Y_{n_{0,k}} < \varphi(f_0) + \frac{\zeta_\varphi}{\nabla_{f_0}(\log \log n_{0,k})}, \quad \varphi(f_1) - \frac{\zeta_\varphi}{\nabla_{f_1}(\log \log n_{1,k})} < Y_{n_{1,k}}.$$

(ii) Les versions globales du Théorème 2.3 sont prouvées dans [7,9] et demandent de distinguer les longueurs d'incrément  $a_n$ . Les démonstrations sont longues et techniques, voir (v) ci-après.

Il est instructif de caractériser au passage les  $T_n$  se comportant comme  $[0, 1 - a_n]$  à l'aide de conditions suffisantes minimales, au sens où nous savons construire des  $T_n$  les vérifiant "presque" mais pour lesquels nos théorèmes sont faux. Soit  $\kappa_n$  le plus petit entier tel que  $T_n \subset \bigcup_{j \leq \kappa_n} [t_{j,n}, t_{j,n} + a_n]$  avec  $t_{j,n} \in T_n$  et  $t_{j,n} - t_{j-1,n} \geq a_n$  pour  $j = 2, \dots, \kappa_n$ . Si  $a_n$  est moyen ou petit ( $c = \infty$ ) nous imposons une condition identique à celle sur  $\lambda(T_n)$  en 2C lorsque  $f \in \mathcal{S}_1^-$ , et un peu plus forte sinon, à savoir

$$\sqrt{1 - \frac{\log \kappa_n}{|\log a_n|}} \ll \frac{\nabla_f(|\log a_n|)}{|\log a_n|}.$$

**Théorème 2.4.** *Soit  $f \in \mathcal{S}_1^*$ . Supposons vérifiés la condition ci-dessus, (1),  $c = \infty$ ,  $d_n \gg \nabla_f^2(\log n)$ , et ou bien  $na_n > (\log n)^{3+\varepsilon}$  ou bien  $na_n < (\log n)^k$ ,  $k > 0$ . De plus, si  $f \in \mathcal{S}_1^{vli} \cup \mathcal{S}_1^{vb}$ , renforçons  $c = \infty$  en  $\nabla_f^2(|\log a_n|) \ll |\log a_n| \log_2 n$ . Alors, pour  $L_n = |\log a_n| + \log \log n$  ou  $L_n = |\log a_n|$  il vient*

$$\lim_{n \rightarrow \infty} \nabla_f(|\log a_n|) d(f, \Theta_n) = \chi_f(1) \quad p.s.$$

Nos arguments montrant aussi que si  $f \in \mathcal{S}_1 \setminus \mathcal{S}_1^*$  la vitesse est plus rapide que  $|\log a_n|^{2/3}$ , les vitesses d'approche s'échelonnent entre  $|\log a_n|$  et  $|\log a_n|^{2/3}$  avec des points lents  $\mathcal{S}_1^{vb}$  et des points rapides  $\mathcal{S}_1^-$ .

**Remarque.** Bien que les deux normalisations fonctionnent, les preuves diffèrent. En effet,  $f$  ne peut pas être approchée simultanément par les deux types d'incrément au même  $t$  car

$$\left\| \frac{\Delta\alpha_n(a_n, t)}{\sqrt{2a_n |\log a_n|}} - \frac{\Delta\alpha_n(a_n, t)}{\sqrt{2a_n |\log a_n| (1 + 1/c_n)}} \right\|$$

est de l'ordre de  $1/c_n$  qui ne résiste pas à  $\nabla_f(|\log a_n|)$ . Les changements de normalisations équivalentes sont donc très sensibles, et nous pouvons fournir pour nos incréments des critères de type [Bu,BL].

Du Théorème 2.4 appliqué à  $f = 0$  nous tirons le module exact de non-différentiabilité de  $\alpha_n$ , par analogie avec celui du mouvement brownien [CR2],

$$\lim_{n \rightarrow \infty} \sqrt{a_n |\log a_n|} \inf_{0 \leq t \leq 1-a_n} \sup_{|t-s| \leq a_n} \frac{|\alpha_n(t) - \alpha_n(s)|}{a_n} = \frac{\pi}{\sqrt{8}} \quad p.s.$$

**Interprétation.** Ceci mesure le désordre persistant. L'oscillation la plus plate, c'est-à-dire là où les  $U_i$  se répartissent le mieux, est toujours d'amplitude supérieure à  $(\log n)/\sqrt{n}$  alors que, par définition, le minimum possible est  $1/2\sqrt{n}$ . L'uniformité ne s'atteint toujours pas, même localement.

Le rapport entre les oscillations maximale et minimale est donc, en comparant [St] au Théorème 2.4,

$$\frac{\sup_{0 \leq t \leq 1-a_n} \sup_{|t-s| \leq a_n} |\alpha_n(t) - \alpha_n(s)|}{\inf_{0 \leq t \leq 1-a_n} \sup_{|t-s| \leq a_n} |\alpha_n(t) - \alpha_n(s)|} \sim \frac{4}{\pi} |\log a_n| \quad p.s.$$

Au delà des  $T_n$  déterministes nous avons obtenu des extensions où des  $t_j^n$  aléatoires dépendent peu de  $\{U_i : i \leq n\} \cap [t_j^n, t_j^n + a_n]$ . En particulier, le Théorème 2.4 reste vrai pour certains sous-ensembles des statistiques d'ordre, tels que par exemple  $\{U_i : i \leq n\} \cap T_n$  avec  $T_n$  ci-dessus. Ceci sert en 2.1.4.

**Application.** Si  $a_n \in ]1/n^b, 1/n^a[$  avec  $0 < a < b < 1$ ,  $Y_n = \sup_{g \in \Theta_n} \varphi(g)$  et  $Z_n = \inf_{g \in \Theta_n} \varphi(g)$  avec  $\varphi$  suffisamment contrastée en  $f_0 = \arg \min_{\mathcal{S}_1} \varphi \in \mathcal{S}_1^*$  et  $f_1 = \arg \max_{\mathcal{S}_1} \varphi \in \mathcal{S}_1^*$  le Théorème 2.4 implique, pour  $0 < \zeta_0 < \zeta_1 < \infty$ ,  $(\zeta_{0,n}, \zeta_{1,n}) \in [\zeta_0, \zeta_1]^2$  et presque sûrement en  $n \rightarrow \infty$ ,

$$|Z_n - \varphi(f_0)| \sim \frac{\zeta_{0,n}}{\nabla_{f_0}(\log n)}, \quad |Y_n - \varphi(f_1)| \sim \frac{\zeta_{1,n}}{\nabla_{f_1}(\log n)}.$$

(iii) Pour les grands incréments nous utilisons  $L_n = |\log a_n| + \log \log n$ . La condition sur  $T_n$  s'écrit

$$\sqrt{1 - \frac{\log \kappa_n}{|\log a_n|}} \ll \frac{\nabla_f(\log \log n)}{\sqrt{|\log a_n| \log \log n}}$$

où  $|\log a_n|$  est de l'ordre de  $\log \log n$  si  $c \in (0, \infty)$  et négligeable si  $c = 0$ , auquel cas la condition est vide pour  $f \in \mathcal{S}_1^-$ . Pour des  $a_n$  trop grands la structure de pont de  $\alpha_n$  réapparaît. Considérons la lim inf pour  $a_n = O(1/(\log \log n)^\alpha)$  avec  $\alpha > \alpha_f \in [4/3, 2]$  et la lim sup pour  $c > 0$ .

**Théorème 2.5.** Sous (1),  $c < \infty$ ,  $a_n \ll 1/\nabla_f^2(\log \log n)$  et l'hypothèse ci-dessus il vient, si  $f \in \mathcal{S}_1^*$

$$\liminf_{n \rightarrow \infty} \nabla_f((1+c) \log \log n) d(f, \Theta_n) = \chi_f(1) \quad p.s.$$

La condition d'entropie sur  $T_n$  contrôlant la minoration de la lim sup devient sensible à  $c_n$

$$\sqrt{\left( \frac{\log \kappa_n}{|\log a_n|} - \frac{1 + 1/c_n}{1 + 1/c} \right)^+} \ll \frac{\nabla_f(\log \log n)}{\log \log n}$$

où  $(x)^+ = \max(0, x)$ , ce qui est vide si  $c_n \leq c$  et limite la taille de  $T_n$  si  $c_n$  décroît lentement. Pour  $n_k = \lceil \exp(k^\theta) \rceil$ ,  $\theta \in ]0, 1/2[$  supposons  $R_k = \bigcap_{n_k \leq n < n_{k+1}} T_n$  de  $a_{n_{k+1}}$ -entropie  $\tau_k$  telle que

$$\sqrt{\left( \frac{1 + 1/c_{n_k}}{1 + 1/c} - \frac{\log \tau_k}{|\log a_{n_k}|} \right)^+} \ll \frac{\nabla_f(\log \log n_k)}{\log \log n_k}.$$

**Théorème 2.6.** Sous (1),  $c \in ]0, \infty[$ ,  $f \in \mathcal{S}_{c/(1+c)}^*$  et les deux conditions précédentes,

$$\limsup_{n \rightarrow \infty} \nabla_f((1+c) \log \log n) d(f, \Theta_n) = \chi_f\left(\frac{c}{1+c}\right) \quad p.s.$$

Au besoin, nos conditions sur  $T_n$  peuvent s'améliorer à la marge mais elles sont quasiment nécessaires, au sens où les vitesses ou normalisations changent pour certains  $T_n$  réguliers ne les vérifiant pas, soit

trop petits soit trop grands, selon  $c_n$ . Les Théorèmes 2.5 et 2.6 appliqués à  $f = 0$  montrent qu'avec probabilité un, pour tout  $\varepsilon > 0$  il existe  $n_\varepsilon$  aléatoire tel que pour tout  $n > n_\varepsilon$ ,

$$\frac{(\pi - \varepsilon)}{\sqrt{8(1+c)\log\log n}} < \inf_{0 \leq t \leq 1-a_n} \sup_{|t-s| \leq a_n} \frac{|\alpha_n(t) - \alpha_n(s)|}{\sqrt{a_n}} < \frac{(\pi + \varepsilon)}{\sqrt{8c\log\log n}}.$$

Le cas  $c = 0$  s'avère délicat car plusieurs phénomènes apparaissent selon l'échelle de normalisation. Nous obtenons dans [18] des points limites  $\mathcal{S}_1$  lorsque  $L_n = \log\log n$ , avec des vitesses lentes mais explicites si  $|\log a_n| \gg \log\log\log n$ . La description exacte pose problème si  $a_n$  est plus grande, en rapport avec le phénomène de [CsR] évoqué en fin de 2.1.1.

**Application.** Si  $a_n = (\log n)^c$  les statistiques  $Y_n$  et  $Z_n$  ci-avant vérifient presque sûrement, pourvu que  $f_0 \in \mathcal{S}_{c/(1+c)}^*$  et  $f_1 \in \mathcal{S}_1^*$ , pour tout  $n$  assez grand,

$$\frac{\zeta_0}{\nabla_{f_0}(\log\log n)} < |Z_n - \varphi(f_0)| < \frac{\zeta_1}{\nabla_{f_0}(\log\log n)}, \quad \frac{\zeta_0}{\nabla_{f_1}(\log\log n)} < |Y_n - \varphi(f_1)|$$

et  $|Y_{n_k} - \varphi(f_1)| < \zeta_1/\nabla_{f_1}(\log\log n_k)$  pour une suite  $n_k$  que nous pouvons en fait choisir déterministe d'après la preuve du Théorème 2.5.

**(iv) Principe de démonstration.** Pour la borne supérieure du Théorème 2.4, par passage à un processus de Poisson  $\{\Pi(t) : t \geq 0\}$  et un mouvement brownien  $\{W(t) : t \geq 0\}$  couplés nous obtenons

$$\begin{aligned} & \frac{1}{2} \mathbb{P} \left( \bigcap_{j \leq \kappa_n} \left\{ \left\| \frac{\Delta\alpha_n(a_n, t_{j,n})}{\sqrt{2a_n L_n}} - f \right\| > \varepsilon_n \right\} \right)^{1/\kappa_n} \\ & \leq \mathbb{P} \left( \left\| \frac{\Pi(na_n I) - na_n I - W(na_n I)}{\sqrt{2na_n L_n}} \right\| > \frac{\varepsilon_n}{p} \right) + \mathbb{P} \left( \left\| \frac{W(na_n I)}{\sqrt{2na_n L_n}} - f \right\| > \frac{\varepsilon_n}{q} \right) \end{aligned}$$

avec  $1/p + 1/q = 1$  bien choisis et  $\varepsilon_n$  la vitesse dilatée de  $1 + \varepsilon$ . Conclure par Borel-Cantelli via [KMT] pour  $\Pi_\rho = (\Pi - \rho I)/\sqrt{\rho}$  et le Lemme 3.3 pour  $W_\rho = W(\rho I)/\sqrt{\rho}$  à l'aide des hypothèses sur  $a_n$  et  $T_n$ . Pour la borne inférieure du Théorème 2.4, un bon choix de  $\theta$  dans  $n_k = \lfloor \exp(k^\theta) \rfloor$  assure, par un raffinement de [St] pour les petits accroissements,

$$\max_{n_{k-1} < n \leq n_k} \sup_{t \in T_n} \left\| \frac{\Delta\alpha_n(a_n, t)}{\sqrt{2a_n L_n}} - \sqrt{\frac{n}{n_k}} \frac{\Delta\alpha_n(a_{n_k}, t)}{\sqrt{2a_{n_k} L_{n_k}}} \right\| \ll \varepsilon_{n_k} \quad p.s.$$

puis, quand  $k \rightarrow \infty$  et  $\mu_k \rightarrow 0$ , nous établissons séparément chacune des majorations

$$\begin{aligned} & \frac{1}{2} \mathbb{P} \left( \bigcup_{n=1+n_{k-1}}^{n_k} \bigcup_{t \in T_n} \left\{ \left\| \sqrt{\frac{n}{n_k}} \frac{\Delta\alpha_n(a_{n_k}, t)}{\sqrt{2a_{n_k} L_{n_k}}} - f \right\| < \varepsilon_{n_{k-1}} \right\} \right) \\ & \leq \mathbb{P} \left( \bigcup_{t \in T_{n_{k-1}+1} \cup \dots \cup T_{n_k}} \left\{ \left\| \frac{\Delta\alpha_{n_k}(a_{n_k}, t)}{\sqrt{2a_{n_k} L_{n_k}}} - f \right\| < p \varepsilon_{n_{k-1}} \right\} \right) \\ & \leq \frac{2}{\mu_k} \mathbb{P} \left( \left\| \frac{\Delta\alpha_{n_k}(a_{n_k}, 0)}{\sqrt{2a_{n_k} L_{n_k}}} - f \right\| < p_1 p_2 \varepsilon_{n_{k-1}} \right) + \mathbb{P} \left( \sup_{|t-s| \leq \mu_k} \frac{|\alpha_{n_k}(t) - \alpha_{n_k}(s)|}{\sqrt{2a_{n_k} L_{n_k}}} \geq \frac{p_1 p_2}{2q_2} \varepsilon_{n_{k-1}} \right) \end{aligned}$$

avant de poissoniser le premier terme, lui appliquer [KMT] et le Lemme 3.3, déterminer un  $\mu_k$  qui laisse tout juste sommable le second terme de grande déviation réglé par [MSW]. Ici  $\varepsilon_n$  est la vitesse optimale contractée par  $1 - \varepsilon$  et chaque étape induit des contraintes contradictoires sur  $n_k$  donc  $\theta$ .

Pour les Théorèmes 2.5 et 2.6 il y a quatre types de preuve, avec diverses techniques de blocage, divers  $n_k$  et des choix délicats de paramètres intermédiaires. Voir les nombreux lemmes de [9].  $\square$

### 2.1.3 – Vitesses pour les oscillations browniennes

(i) Dans [9] nous approfondissons aussi les résultats de [CR,BS,CR2,R] et la vitesse générale de [B,GK]. Soit  $\{W(t) : t \geq 0\}$  un mouvement brownien réel standard issu de 0 et  $A_T > 0$  une fonction telle que

$$A_T \uparrow \infty, \quad \frac{A_T}{T} \downarrow 0, \quad C_T = \frac{\log(T/A_T)}{\log \log T} \rightarrow c \quad (3)$$

si  $T \rightarrow \infty$ . Considérons une collection d'ensembles  $\mathcal{T}_T \subset [0, T - A_T]$  d'entropie  $\kappa_T$  par rapport à  $A_T$  et notons  $\tau_k$  l'entropie de  $\bigcap_{n_k \leq T < n_{k+1}} \mathcal{T}_T$  pour les réels  $n_k = \exp(k^\theta)$ ,  $\theta \in ]0, 1/2[$ . Les ensembles d'incrément browniens  $\Delta W(A, t) = W(t + AI) - W(t)$  normalisés sont notés

$$\Lambda_T = \left\{ \frac{\Delta W(A_T, t)}{\sqrt{2A_T L_T}} : t \in \mathcal{T}_T \right\}.$$

**Théorème 2.8.** *Soit  $f \in \mathcal{S}_1^*$  et  $L_T = \log(T/A_T) + \log \log T$ . Supposons (3) et, lorsque  $c = \infty$  et  $f \in \mathcal{S}_1^{vli} \cup \mathcal{S}_1^{vb}$ ,  $\nabla_f^2(\log(T/A_T)) \ll \log(T/A_T) \log \log T$ . Si*

$$\sqrt{1 - \frac{\log \kappa_T}{\log(T/A_T)}} \ll \frac{\nabla_f(L_T)}{L_T}$$

nous avons

$$\liminf_{T \rightarrow \infty} \nabla_f(L_T) d(f, \Lambda_T) = \chi_f(1) \quad p.s.$$

et cette  $\liminf$  est une limite si, et seulement si,  $c = \infty$ . Si  $c \in ]0, \infty[$ ,  $f \in \mathcal{S}_{c/(1+c)}^*$  et

$$\frac{L_{n_k}}{\nabla_f(L_{n_k})} \sqrt{\left( \frac{1 + 1/C_{n_k}}{1 + 1/c} - \frac{\log \tau_k}{\log(n_k/A_{n_k})} \right)^+} \rightarrow 0, \quad \frac{B_T}{\nabla_f(B_T)} \sqrt{\left( \frac{\log_1 \kappa_T}{\log_1(T/A_T)} - \frac{1 + 1/C_T}{1 + 1/c} \right)^+} \rightarrow 0$$

nous avons

$$\limsup_{T \rightarrow \infty} \nabla_f(L_T) d(f, \Lambda_T) = \chi_f\left(\frac{c}{1+c}\right) \quad p.s.$$

Par rapport à 2.1.2 des conditions ont disparu car l'unique trajectoire directrice  $W$  n'est pas un pont, ce qui élimine quelques contraintes dans les preuves.

(ii) Dans [8] nous montrons un résultat similaire se comparant mieux à ceux de 2.1.2 et 2.1.5. Pour une suite quelconque de ponts browniens  $B_n$  considérons  $\Delta B_n(a_n, t) = B_n(t + a_n I) - B_n(t)$  et

$$\Gamma_n = \left\{ \frac{\Delta B_n(a_n, t)}{\sqrt{2a_n |\log a_n|}} : t \in [0, 1 - a_n] \right\}.$$

Un cas particulier important est donné par  $B_n^W = (W(nI) - W(n)I)/\sqrt{n}$ . Omettons les  $f \in \mathcal{S}_1^{vli}$ .

**Proposition 2.9.** *Supposons que  $a_n$  vérifie (1) avec  $c = d = \infty$  et  $f \in \mathcal{S}_1$  avec  $V(f') < \infty$ . Pour toute suite  $B_n$  de ponts browniens, nous avons, si  $J(f) < 1$*

$$\limsup_{n \rightarrow \infty} |\log a_n| d(f, \Gamma_n) \leq \frac{\pi}{4\sqrt{1 - J(f)}} \quad p.s.$$

et, si  $J(f) = 1$

$$\limsup_{n \rightarrow \infty} |\log a_n|^{2/3} d(f, \Gamma_n) \leq \chi_f \quad p.s.$$

Pour  $B_n = B_n^W$  il y a égalité et de vraies limites dans les deux cas.

### 2.1.4 – Autres processus

Les résultats de la partie 2 s'étendent à tout processus réel pouvant s'approximer par  $\alpha_n$ ,  $B_n$ ,  $\Pi$  ou  $W$ . Prenons comme dans [7,9] l'exemple des quantiles et de la marche aléatoire.

(i) Tout d'abord, le processus empirique uniforme des quantiles  $\beta_n = \sqrt{n}(F_n^{-1} - I)$  associé à  $U_1, \dots, U_n$  via  $F_n^{-1} = \sup \{s \in [0, 1] : F_n(s) < I\}$  sur  $]0, 1[$ . Le processus  $\beta_n$  est, par nature, plus oscillant que  $\alpha_n$ . Pour décrire ses incréments nous devons nous ramener à ceux de  $\alpha_n$  par la représentation de Bahadur-Kiefer [Ki], et cela en des positions aléatoires. Cet aléa supplémentaire ne perturbe pas si

$$a_n \gg \nabla_f^2(\log n) \frac{\sqrt{\log \log n}}{\sqrt{n} \log n}$$

mais pour les  $a_n$  plus petits nous avons dû, dans [9], restreindre  $T_n$  à être un intervalle. C'est pourquoi une statistique dirigée par  $\beta_n$  est plus délicate à maîtriser, voir 4.2 et 4.3 pour une autre illustration. Les résultats [M] sur le module de continuité de  $\beta_n$  se généralisent comme en 2.2 au module d'oscillation fonctionnel  $d(f, \Xi_n)$  où

$$\Xi_n = \left\{ \frac{\Delta\beta_n(a_n, t)}{\sqrt{2a_n L_n}} : t \in T_n \right\}, \quad \Delta\beta_n(a_n, t) = \beta_n(t + a_n I) - \beta_n(t).$$

**Corollaire 2.10.** *Soit  $f \in \mathcal{S}_1^*$ ,  $L_n = |\log a_n| + \log \log n$  et  $T_n$  vérifiant la contrainte ci-dessus et celles des Théorèmes 2.4, 2.5 et 2.6 selon que  $c = \infty$  ou  $c \in ]0, \infty[$ . Sous (1),  $d_n \ll \nabla_f^4(\log n)$ ,  $a_n \ll 1/\nabla_f^2(\log \log n)$  et ou bien  $na_n > (\log n)^{3+\varepsilon}$  ou bien  $na_n < (\log n)^k$ , avec  $\varepsilon > 0$  et  $k > 0$ ,*

$$\liminf_{n \rightarrow \infty} \nabla_f(L_n) d(f, \Xi_n) = \chi_f(1) \quad p.s.$$

Si de plus  $f \in \mathcal{S}_{c/(1+c)}^*$  alors

$$\limsup_{n \rightarrow \infty} \nabla_f(L_n) d(f, \Xi_n) = \chi_f\left(\frac{c}{1+c}\right) \quad p.s.$$

Si  $f \in \mathcal{S}_1^-$ ,  $d = \infty$  dans (1),  $na_n = O((\log n)^5)$ ,  $T_n$  est un intervalle et  $\log \lambda(T_n) \sim |\log a_n|$ ,

$$\limsup_{n \rightarrow \infty} d_n^{1/4} d(f, \Xi_n) \leq 2^{1/4} \quad p.s.$$

Nous ne savons pas si le  $d_n^{1/4}$  est optimal, mais le Théorème 2.13 plaide en faveur d'une vitesse de ce type pour certains  $f$  au moins. Trouver lesquelles donnerait une borne inférieure à [KMT] pour  $\beta_n$ .

(ii) Reformulons les résultats de 2.1.2 pour la marche aléatoire  $S_n$  de 1.1.2 et les fonctions de Strassen non frontières. Soit  $\Delta S_n(t, a_n) = \sum_{[nt] < i \leq [n(t+a_n I)]} X_i + (nt + na_n I - [nt + na_n I])X_{[n(t+a_n I)]+1}$  avec  $X_i$  des v.a. réelles i.i.d. de même loi que  $X$ ,  $\mathbb{E}X = 0$  et  $\mathbb{E}e^{tX} < \infty$  pour  $t \in ]-s, s[$ . Diminuer la condition de moments en  $\mathbb{E}|X|^{2+\varepsilon} < \infty$  restreint le choix des plus petits  $a_n$  à  $a_n > 1/n^a$ ,  $a < \varepsilon/(2+\varepsilon)$ .

**Corollaire 2.11.** *Supposons que  $a_n$  vérifie (1),  $c > 0$ ,  $na_n > (\log n)^{3+\varepsilon}$  et  $T_n = [0, 1 - a_n]$ . Si donc  $J(f) < 1$  et  $L_n = |\log a_n| + \log \log n$  nous avons*

$$\liminf_{n \rightarrow \infty} L_n \inf_{t \in T_n} \left\| \frac{\Delta S_n(t, a_n)}{\sqrt{2na_n L_n}} - f \right\| = \frac{\pi}{4\sqrt{1 - J(f)}} \quad p.s.$$

et, si de plus  $J(f) < c/(1+c)$ ,

$$\limsup_{n \rightarrow \infty} L_n \inf_{t \in T_n} \left\| \frac{\Delta S_n(t, a_n)}{\sqrt{2na_n L_n}} - f \right\| = \frac{\pi}{4\sqrt{c/(1+c) - J(f)}} \quad p.s.$$

En particulier, si  $f = 0$ ,  $T_n = [0, 1 - a_n]$ ,  $a_n = b/n^a$ ,  $b > 0$  et  $a \in ]0, 1[$  nous obtenons que la plus petite oscillation locale autour de la première droite de Kolmogorov, sur une portion de longueur  $bn^{1-a}$ , est presque sûrement équivalente à  $\pi\sqrt{bn^{1-a}}/\sqrt{8a \log n}$ .

C'est une nécessité incompressible du hasard que de désobéir ainsi, et partout, à l'espérance.

### 2.1.5 – Vitesses poissoniennes et principe d'invariance

D'après [KMT] il existe, sur un espace  $(\Omega, \mathcal{T}, \mathbb{P})$  commun, une version de la suite  $\{U_n\}$  et une suite de ponts browniens  $B_n$  tels que  $\|\alpha_n - B_n\| = O_{p.s.}((\log n)/\sqrt{n})$  mais pour l'approximation par un unique processus de Kiefer  $K(n, I)$  la vitesse devient  $(\log n)^2/\sqrt{n}$ . Par l'argument de poissonisation de nos preuves, la borne exponentielle utilisée agit au final comme si nous avions  $O_{p.s.}((\log n)/\sqrt{n})$ .

(i) Cependant, en zoomant trop près sur les trajectoires empiriques et browniennes, leurs différences structurelles apparaissent. Les méthodes utilisant la construction hongroise nous conduisent dans [1,8] au cas limite suivant, avec des vitesses moins bonnes qu'à la Proposition 2.9 pour  $B_n^W$ .

**Proposition 2.12.** *Soit  $T_n$  tel que  $\log \kappa_n \sim \log n$ ,  $L_n = |\log a_n|$  et  $a_n$  vérifiant (1) avec  $d = \infty$ . Si  $f \in \mathcal{S}_1^* \setminus \{0\}$  et  $d_n = na_n/\log n = O(\nabla_f^2(\log n))$  ou  $f \in \mathcal{S}_1 \setminus \mathcal{S}_1^*$  et  $d_n = O((\log n)^{4/3})$  il vient*

$$\limsup_{n \rightarrow \infty} \sqrt{d_n} d(f, \Theta_n) < \infty \quad p.s.$$

Si  $f = 0$  nous avons

$$\limsup_{n \rightarrow \infty} \frac{n^{1/6}}{a_n^{1/3}} \inf_{t \in T_n} \|\Delta \alpha_n(t, a_n)\| < \infty \quad p.s.$$

La vitesse  $\sqrt{d_n}$  est arbitrairement lente si  $a_n$  est proche de  $(\log n)/n$ . Dans la preuve, ces bornes sont contraintes à la fois par [KMT] et le contrôle du module de continuité. D'où l'interrogation suivante, puisque [KMT] est optimal. Est-ce bien une différence trajectorielle entre  $\alpha_n$  et  $B_n$  qui se manifeste à ce niveau de zoom critique ? Pour répondre il a fallu contourner l'approximation gaussienne.

Dans [1], l'idée donnée page 126 et l'outil esquissé page 362 étaient basés sur la perturbation des déviations gaussiennes et un développement opposant termes gaussiens et poissoniens était conjecturé. Puis [DL2, Sm] ont justifié ce développement et obtenu une borne exponentielle presque exacte pour les petites boules centrées en  $f$  telle que  $f' \in \mathcal{VB}$ , du processus de Poisson en haute intensité. Ceci permet de s'affranchir des gaussiens sur  $\mathcal{S}_1^{vb}$  et, en retravaillant le changement de mesure de Poisson comme signalé en 3.5.1, sur une partie de  $\mathcal{S}_1^{li}$ .

(ii) Dans [8] nous obtenons les vitesses exactes dans la Proposition 2.12. Soit

$$\theta_f = \int_{[0,1]} f'^3 d\lambda, \quad \phi_f = \int_{[0,1]} f'^4 d\lambda, \quad c_f = |f'(1)| + \text{Var}(f').$$

La constante  $c_f$  est typique des petites déviations gaussiennes,  $\theta_f$  et  $\phi_f$  sont typiques du comportement poissonien. Pour distinguer les  $a_n$  trop petits posons  $\mu^-$  et  $\mu^+$  la  $\liminf$  et la  $\limsup$  de  $na_n/(\log n)^{7/3}$ . La condition supplémentaire sur  $\log d_n$  est sans importance.

**Théorème 2.13.** *Soit  $f \in \mathcal{S}_1$  telle que  $\text{Var}(f') < \infty$  et  $a_n$  tel que (1),  $d = \infty$  et  $na_n < d_0(\log n)^3$ . Si  $J(f) < 1$  alors  $d(f, \Theta_n) \sim \pi/(4\sqrt{1-J(f)} \log n)$  p.s. Si  $J(f) = 1$  et  $\mu^- = \infty$  alors  $d(f, \Theta_n) \sim \chi_f/(\log n)^{2/3}$  p.s. Si  $J(f) = 1$  et  $\mu^+ = 0$  alors nous avons, lorsque  $\theta_f < 0$*

$$\lim_{n \rightarrow \infty} \sqrt{d_n} d(f, \Theta_n) = \frac{|\theta_f|}{3\sqrt{2} c_f} \quad p.s.$$

et, lorsque  $\theta_f > 0$  et, ou bien  $d_n > (\log n)^\varepsilon$  ou bien  $\log d_n \sim \log d_{\exp(\log n)^{1+|\circ(1)|}}$ ,

$$\lim_{n \rightarrow \infty} \frac{\log n}{d_n^{1/4}} d(f, \Theta_n) = \frac{\pi}{29/4} \sqrt{\frac{3}{\theta_f}} \quad p.s.$$

De façon assez inattendue, la vitesse de 2.12 ([9]) était parfois optimale. De plus, la première vitesse du Théorème 2.13 est toujours plus lente que le  $1/(\log n)^{2/3}$  de la Proposition 2.9 appliquée à  $B_n^W$  et la seconde toujours plus rapide. Nous avons ainsi identifié un déterminisme trajectoriel qui distingue



les ponts brownien et empirique dans le domaine d'invariance. Il est en outre surprenant de constater que certaines fonctions de Strassen sont approchées plus rapidement par un processus non gaussien.

**Interprétation.** Il est plus facile pour le processus de saut  $\Delta\alpha_n$  de suivre une fonction extrémale majoritairement croissante,  $\theta_f > 0$  que pour un accroissement brownien  $\Delta B_n^W$ . A l'inverse, la structure de décroissance linéaire de  $\Delta\alpha_n$  entre ces sauts trop espacés l'empêche de suivre aussi facilement que  $\Delta B_n^W$  une fonction d'énergie maximale et majoritairement décroissante,  $\theta_f < 0$ .

Nous apprenons donc qu'à petite échelle, les accroissements de  $\alpha_n$  ne sont pas symétriques. Nous interprétons cela comme la transition, au second ordre, vers la non-invariance au sens de [Sb].

**Exemple.** Soit  $\Gamma_n$  et  $B_n^W$  de 2.1.3(ii). Pour l'accroissement maximum, à savoir l'identité, il vient

$$\begin{aligned} \frac{\pi\sqrt{3}}{2^{9/4}} \frac{d_n^{1/4}}{\log n} &\sim d(I, \Theta_n) \ll d(I, \Gamma_n) \sim \frac{\pi^{2/3}}{4(\log n)^{2/3}} \quad p.s. \\ \frac{1}{6\sqrt{2}} \frac{1}{\sqrt{d_n}} &\sim d(-I, \Theta_n) \gg d(-I, \Gamma_n) \sim \frac{\pi^{2/3}}{4(\log n)^{2/3}} \quad p.s. \end{aligned}$$

**Exemple.** Pour la fonction triangulaire  $h'_a = \mathbb{I}_{[0,a]} - \mathbb{I}_{(a,1]}$  nous obtenons, avec probabilité un

$$\begin{aligned} \frac{\pi}{2^{9/4}} \sqrt{\frac{3}{2a-1}} \frac{d_n^{1/4}}{\log n} &\sim d(h_a, \Theta_n) \ll d(h_a, \Gamma_n) \sim \frac{\chi_{h_a}}{(\log n)^{2/3}} \quad \text{si } a > \frac{1}{2}, \\ \frac{1-2a}{12\sqrt{2}} \frac{1}{\sqrt{d_n}} &\sim d(h_a, \Theta_n) \gg d(h_a, \Gamma_n) \sim \frac{\chi_{h_a}}{(\log n)^{2/3}} \quad \text{si } a < \frac{1}{2}. \end{aligned}$$

Enfin, le Théorème 2.13 nous informe également sur le comportement des espacements locaux, qu'ils soient maximums, minimums ou moyens, et non globaux comme [Dv]. Ce serait long à développer.

(iii) De ce type de considérations particulières issues de la comparaison des vitesses 2.9 et 2.13 nous déduisons dans [8] une borne inférieure au principe d'invariance fort, de plusieurs façons différentes.

**Corollaire 2.14.** Soit  $U_n$  des v.a. uniformes i.i.d. et  $B_n$  des ponts browniens sur  $(\Omega, \mathcal{T}, \mathbb{P})$ . Il vient

$$\liminf_{n \rightarrow \infty} \frac{\sqrt{n}}{\log n} \|\alpha_n - B_n\| \geq \frac{1}{6} \quad p.s.$$

Nous fournissons ainsi l'explication trajectorielle de [KMT] qui nous hantait depuis [1] et les vitesses de type  $\sqrt{d_n}$ . Comme démontré dans [BM] la construction hongroise permet de coupler  $(\alpha_n, B_n)$  de sorte que  $\|\alpha_n - B_n\| \leq 12(\log n)/\sqrt{n}$  mais nous ne savons pas construire de suites atteignant la liminf  $1/6$ . Nous obtenons cette constante, entre autres écritures, comme

$$\frac{1}{6} = \sup \left\{ \frac{|\theta_f|}{6 c_f} : J(f) = 1, c_f = |f'(1)| + \text{Var}(f') < \infty, \theta_f = \int_{[0,1]} f'^3 d\lambda \neq 0 \right\}.$$

(iv) Lorsque  $J(f) = 1$ ,  $\text{Var}(f') < \infty$  et  $\theta_f \neq 0$  nous montrons dans [8] qu'à l'échelle transitoire où  $\mu^\pm \in ]0, \infty[$  les valeurs d'adhérence de  $(\log n)^{2/3} d(f, \Theta_n)$  sont presque sûrement incluses dans  $[C, D]$  avec  $0 < C < D < \infty$  solutions de certaines équations fonction de  $\mu^\pm$ ,  $c_f$  et  $\theta_f$ .

Lorsque  $\theta_f = 0$  notons  $\phi_f = \int_{[0,1]} f'^4 d\lambda \geq 1$  et  $\nu^- \leq \nu^+$  la lim inf et la lim sup de  $na_n/(\log n)^{5/3}$ .

**Proposition 2.15.** Supposons  $J(f) = 1$ ,  $\text{Var}(f') < \infty$ ,  $\theta_f = 0$ , (1),  $d = \infty$  et  $na_n < d_0(\log n)^3$ . Si  $0 < \nu^- < \nu^+ < \infty$  alors  $(\log n)^{2/3} d(f, \Theta_n)$  adhère presque sûrement dans  $[C, D]$  avec  $0 < C < D < \infty$  solutions de certaines équations fonction de  $\nu^\pm$ ,  $c_f$  et  $\phi_f$ . Si  $\nu^- = \infty$  alors  $(\log n)^{2/3} d(f, \Theta_n)$  adhère presque sûrement dans  $(\pi^2/32c_f)^{1/3} [1, 3]$ . Si  $\nu^+ = 0$  alors

$$\lim_{n \rightarrow \infty} d_n d(f, \Theta_n) = \frac{\phi_f}{6 c_f} \quad p.s.$$

Ainsi pour les fonctions telles que  $\int_{[0,1]} f'^2 d\lambda = 1$  et  $\int_{[0,1]} f'^3 d\lambda = 0$  les vitesses rapides se dégradent et les lentes s'améliorent mais toutes sont moins bonnes que dans le cas gaussien.

**Exemple.** Pour  $h'_{1/2} = \mathbb{I}_{[0,1/2]} - \mathbb{I}_{(1/2,1]}$  et  $a_n = a_0(\log n)^\chi/n$  avec  $\chi \in ]1, 5/3[$  nous avons

$$\frac{1}{24a_0(\log n)^{\chi-1}} \sim d(h_{1/2}, \Theta_n) \gg d(h_{1/2}, \Gamma_n) \sim \frac{\pi^{2/3}}{2^{7/3}(\log n)^{2/3}} \quad p.s.$$

**(v) Principe de démonstration.** Il s'agit de pousser nos techniques de blocages à leur extrême limite en utilisant des  $n_k$  implicites, de les adapter aux différents cas et d'introduire des astuces pour contrôler le module de continuité à des échelles sur-critiques  $a_n^*$  telles que  $na_n^* \ll \log n$ . Pour le choix des paramètres et des sous-suites la contrainte de déviation gaussienne est remplacée par la version poissonnienne suivante, due à [Sm] et basée sur [DL2,Mo]. Soit  $\rho > 0$  et  $\Pi_\rho = (\Pi(\rho I) - \rho I)/\sqrt{\rho}$  la version standardisée d'un processus de Poisson  $\{\Pi(t) : t \in [0, \rho]\}$  d'intensité 1. Si  $f \in \mathcal{S}_1^{vb}$ ,  $\delta \in [0, 1[$  et  $\rho \rightarrow \infty$ ,  $\lambda \rightarrow \infty$ ,  $\varepsilon \rightarrow 0$ ,  $\rho\lambda^2\varepsilon^2 \rightarrow \infty$ ,  $\rho/\lambda^2 \rightarrow \infty$  alors

$$\begin{aligned} & \exp\left(-\frac{\lambda^2}{2}J(f) - \frac{\pi^2 + o(1)}{8\lambda^2\varepsilon^2(1-\delta)^2} + \lambda^2\varepsilon(2\delta-1)(c_f + o(1)) + \frac{\lambda^3}{6\sqrt{\rho}}(\theta_f + o(1))\right) \\ & \leq \mathbb{P}\left(\frac{\Pi_\rho}{\lambda} \in f + \varepsilon\mathcal{B}\right) \\ & \leq \exp\left(-\frac{\lambda^2}{2}J(f) - \frac{\pi^2 + o(1)}{8\lambda^2\varepsilon^2} + \lambda^2\varepsilon(c_f + o(1)) + \frac{\lambda^3}{6\sqrt{\rho}}(\theta_f + o(1))\right) \end{aligned}$$

avec des  $o(1)$  en  $\rho$  fois un polynôme en  $\lambda/\sqrt{\rho}$  de terme dominant  $-\lambda^4\phi_f/12\rho = o(\lambda^3/\sqrt{\rho})$ .  $\square$

Depuis [8] nous retravaillons [Sm] par la méthode [6] pour démontrer les versions du Théorème 2.13 précisant vitesse et constante dans la Proposition 2.12 pour certains  $f \in \mathcal{S}_1^{vli}$ , voir 3.5.1.

### 2.1.6 – Recouvrement complet et temps d'atteinte

(i) A la suite de [14,16] nous avons calculé les vitesses de recouvrement complet de  $\mathcal{S}_{c/(1+c)}$  par des suites d'ensembles d'incrémentes. Il ne s'agit pas d'une conséquence des vitesses ponctuelles 2.1.2 car les fonctions du  $\varepsilon_n$ -recouvrement optimal de  $\mathcal{S}_1$  changent avec  $n$ .

**Théorème 2.16.** *Sous les hypothèses du Théorème 2.1 il vient, pour  $\sigma > \chi_I$*

$$\mathbb{P}\left(\exists n_0 : \mathcal{S}_1 \subset \Theta_n + \frac{\sigma}{L_n^{2/3}} B_0, n > n_0\right) = 1.$$

Soit  $d_H(A, B) = \max(\sup_{f \in A} d(f, B), \sup_{g \in B} d(g, A))$  la distance de Hausdorff. Quand  $c = \infty$  les Théorèmes 2.1 et 2.16 entraînent le comportement déterministe précis

$$0 < \liminf_{n \rightarrow \infty} |\log a_n|^{2/3} d_H(\Theta_n, \mathcal{S}_1) \leq \limsup_{n \rightarrow \infty} |\log a_n|^{2/3} d_H(\Theta_n, \mathcal{S}_1) < \infty \quad p.s.$$

qui constitue la première vitesse optimale de recouvrement complet obtenue dans des lois fonctionnelles du logarithme. De même, si  $0 < c < \infty$  alors il existe  $0 < \sigma_3 < \sigma_4 < \infty$  tels que

$$\begin{aligned} & \mathbb{P}\left(\exists n_k \uparrow \infty : d_H(\Theta_{n_k}, \mathcal{S}_{c/(1+c)}) < \frac{\sigma_4}{(\log \log n_k)^{2/3}}, k \geq 1\right) \\ & = \mathbb{P}\left(\exists n_0 : d_H(\Theta_n, \mathcal{S}_{c/(1+c)}) > \frac{\sigma_3}{(\log \log n)^{2/3}}, n \geq n_0\right) = 1. \end{aligned}$$

**Principe de démonstration.** Il convient de se placer sur un  $(\varepsilon_n/M)$ -recouvrement optimal de  $\mathcal{S}_1$  par des boules centrées en des  $f \in \mathcal{S}_1$ , qui est de taille  $2 \exp(M/\varepsilon_n)$  puis, en chacun des centres  $f$ , d'appliquer les bornes exactes de petite déviation menant au Lemme 3.3. L'argument clef est alors

une minoration universelle de l'économie d'énergie qui permet de nous ramener aux vitesses les plus lentes, et de conclure par Borel-Cantelli en faisant  $n \rightarrow \infty$  puis  $M \rightarrow \infty$ . Pour  $a_n$  trop lent, combiner à une technique de blocage classique qui complique la preuve.  $\square$

(ii) Dans [E,ER] les temps d'attente entre deux violations de la loi du logarithme itéré pour  $W$  sont évalués, puis [R3] évalue de même les inter-temps de concentration de  $\Lambda_T$  autour de  $\mathcal{S}_{c/(1+c)}$ . Dans [20] nous explorons les versions fonctionnelles de [E,ER] en tenant compte des vitesses d'approche, les étendons à  $\alpha_n$  et ses accroissements, et approfondissons [R3]. Il s'agit de déterminer les temps d'attente de divers événements extrêmes ou rares, liés au recouvrement des boules de [S] ou [F]. Dans le cas d'événements du type  $d(f, \Theta_n) \leq (1 + \varepsilon)\varepsilon_n$  nous parlons de *temps d'atteinte presque sûrs*.

Notre résultat de base évalue le temps nécessaire à la formation d'une configuration d'échantillonnage donnée, sans vitesse  $\varepsilon_n$ . D'après [F],  $\alpha_n$  est presque sûrement relativement compacte dans  $\mathcal{B}$  de points adhérents  $\mathcal{F}_1 = \mathcal{S}_1 \cap \{f : f(1) = 0\}$ . Combien de temps met  $\alpha_n$  à rencontrer  $f \in \mathcal{F}_1$  ? Soit  $\varepsilon > 0$ ,  $J(f, \varepsilon) = \inf_{\|h-f\| < \varepsilon} J(h)$  et

$$N_1(f, \varepsilon, n) = \inf \left\{ m > n : d \left( f, \frac{\alpha_m}{\sqrt{2 \log \log m}} \right) < \varepsilon \right\}.$$

De même, si  $f \in \mathcal{S}_1 \setminus \{0\}$ , combien de temps met  $\Delta\alpha_n(a_n, 0)$  à rencontrer  $f$  ? Posons

$$N_2(f, \varepsilon, n) = \inf \left\{ m > n : d \left( f, \frac{\Delta\alpha_m(a_m, 0)}{\sqrt{2a_m \log \log m}} \right) < \varepsilon \right\}.$$

**Proposition 2.17.** *Si  $f \in \mathcal{F}_1 \setminus \{0\}$  et  $\varepsilon < \|f\|$  alors*

$$\limsup_{n \rightarrow \infty} \frac{\log \log (N_1(f, \varepsilon, n)/n)}{\log \log n} = J(f, \varepsilon) \quad p.s.$$

*Si  $f \in \mathcal{S}_1 \setminus \{0\}$  et  $\varepsilon < \|f\|$  alors*

$$\limsup_{n \rightarrow \infty} \frac{\log \log (N_2(f, \varepsilon, n)/n)}{\log \log n} = J(f, \varepsilon) \quad p.s.$$

En d'autres termes, si  $\mu > 0$  nous avons avec probabilité un, pour tout  $n$  assez grand,

$$n \leq N_2(f, \varepsilon, n) < n \exp \left( (\log n)^{J(f, \varepsilon) + \mu} \right)$$

et pour une suite  $n_{k, \mu} \uparrow \infty$ ,

$$N_2(f, \varepsilon, n_{k, \mu}) > n_{k, \mu} \exp \left( (\log n_{k, \mu})^{J(f, \varepsilon) - \mu} \right).$$

**Exemple.** Comme  $J(I, \varepsilon) = 1 - 2\varepsilon + \varepsilon^2$  nous apprenons, par comparaison avec l'adaptation de [ER] pour  $\Delta\alpha_n(a_n, 0)$ , qu'il est toujours plus long pour  $\Delta\alpha_n(a_n, 0)$  d'approcher  $I$  que de franchir l'enveloppe  $\sqrt{I}$ . Mais le pire temps mis pour entrer dans  $I + (\varepsilon/2)B_0$  est comparable au pire temps mis pour franchir  $(1 - \varepsilon)\sqrt{I}$ . Comme le premier événement implique le second ceci montre que si  $\Delta\alpha_n(a_n, 0)$  franchit  $(1 - \varepsilon)\sqrt{I}$  le plus tard possible après  $n$  c'est nécessairement en étant presque linéaire, donc avec les  $na_n + (1 \pm \varepsilon)\sqrt{2na_n \log \log n}$  variables  $U_i \in [0, a_n]$  presque équidistribuées.

Nous considérons de même des temps d'atteinte de type  $N_3(f, \varepsilon, n) = \inf \{m > n : d(f, \Theta_m) < \varepsilon\}$ , qui se comporte comme  $N_2$  pourvu que  $T_n$  soit suffisamment stable et petit. Puis nous tenons compte des vitesses,  $N_4(A, \varepsilon, n) = \inf \{m > n : d(A, \Theta_m) < (1 + \varepsilon)\varepsilon_m\}$  pour  $A = \{f\}$  ou  $A = S_1$ .

Nous recherchons aussi, dans [20], de tels temps d'attente conditionnels à un événement de départ fonction de  $\alpha_n$  lui-même. Des changements de topologie de type 2.2 et le cadre intermédiaire 2.3.2 et 2.3.3 sont à la portée de nos techniques de preuve, avec un asymptotique différent. Des applications en estimation non paramétrique sont envisagées, comme expliqué en 4.4.4.

## 2.2 – Lois et vitesses en topologie de type Höldérienne

Nous avons généralisé les résultats du 2.1 dans un cadre topologique renforcé, et nous présentons le cas des métriques de type Höldériennes. Le but est de mieux comprendre la structure des incréments locaux de  $\alpha_n$  et  $W$ , leurs propres sous-structures oscillantes.

Soit  $\alpha \in [0, 1/2[$ ,  $\beta = 1/2 - \alpha$  et  $\mathcal{B}_\alpha$  l'espace de Banach  $\mathcal{B}$  muni de la norme  $\alpha$ -Höldérienne

$$\|f\|_\alpha = \sup_{0 < |t-s| \leq 1} \frac{|f(t) - f(s)|}{|t-s|^\alpha}, \quad f \in \mathcal{B}.$$

Notons les distances  $d_\alpha(f, S) = \inf_{g \in S} \|f - g\|_\alpha$  et les voisinages ouverts  $V_\alpha(S, \varepsilon) = \{f : d_\alpha(f, S) < \varepsilon\}$ . Introduisons les modules de régularité  $\alpha$ -Höldériens

$$\omega_{\alpha, f}(\delta) = \sup_{0 < |t-s| \leq \delta} \frac{|f(t) - f(s)|}{|t-s|^\alpha}, \quad \omega_{\alpha, f}^+(\delta) = \sup_{|t-s| > \delta} \frac{|f(t) - f(s)|}{|t-s|^\alpha}, \quad f \in \mathcal{B}, \quad 0 < \delta \leq 1.$$

L'espace  $\mathcal{H}_\alpha = \{f \in \mathcal{B} : \|f\|_\alpha < \infty\}$  n'est pas  $d_\alpha$ -séparable mais  $\mathcal{C}_\alpha = \{f \in \mathcal{H}_\alpha : \lim_{\delta \rightarrow 0} \omega_{\alpha, f}(\delta) = 0\}$  l'est, voir [Ci]. Les modules se comparent via  $\omega_f = \omega_{0, f} \leq \omega_{\alpha, f} \leq \|f\|_\alpha \leq \max(\|f\|_0 / I^\alpha, \omega_{\alpha, f})$  où  $\|f\|_0 = \omega_f(1)$  est l'amplitude de  $f$ .

L'ensemble  $\bigcup_{\theta > 0} \mathcal{S}_\theta \subset \mathcal{H}_\alpha$  est de fermeture  $\mathcal{C}_\alpha$  dans  $\mathcal{B}_\alpha$  et l'injection canonique de  $\mathcal{H}_\alpha$  dans  $\mathcal{C}_\alpha$  est compacte. La fonctionnelle  $J$  est semi-continue inférieurement de  $\mathcal{B}_\alpha$  dans  $\mathbb{R}$  et  $\mathcal{S}_\theta$  est compact d'intérieur vide dans  $\mathcal{B}_\alpha$ . Il est de ce fait naturel de tenter d'étendre les résultats de 2.1 de  $\mathcal{B}$  à  $\mathcal{B}_\alpha$ .

**Remarque.** Notre observation initiale est que pour certaines classes de fonctions, donc de processus,  $d_\alpha$  est contrôlée par  $\omega_{\alpha, f}^+(\delta)$  quand  $\delta \rightarrow 0$ , donc par la métrique uniforme  $d$ . Si  $|f(t) - f(s)| \leq \varphi(|t-s|)$  alors  $\omega_{\alpha, f}(\delta) \leq \sup_{(0, \delta]} (\varphi / I^\alpha)$  et  $\|f\|_\alpha \leq \|\varphi / I^\alpha\|$ . En particulier  $\varphi = \|f\|_\rho I^\rho$  pour  $f \in \mathcal{H}_\rho$  et  $\rho > \alpha$ . Puisque  $\mathcal{S}_1 \subset \mathcal{H}_{1/2}$  les fonctions de Strassen ont aussi un petit module de  $\alpha$ -régularité,

$$\omega_f \leq \omega_{\alpha, f} \leq \min(\|f\|_\alpha, \sqrt{\theta} I^\beta), \quad f \in \mathcal{S}_\theta.$$

Clairement, dans ces deux cas, pour contrôler  $\|f\|_\alpha = \max(\omega_{\alpha, f}, \omega_{\alpha, f}^+)$  il suffit de considérer  $\omega_{\alpha, f}^+(\delta)$  en  $\delta$  assez petit. Opérer un contrôle similaire pour nos processus d'intérêt, mais en probabilité et en maîtrisant  $\delta = \delta_n \rightarrow 0$ , est l'un des arguments clef de nos preuves.

Ce travail est initié dans [1,3] puis les développements ont abouti à [5,13,17]. En 2.2.1 nous décrivons le comportement de suites d'ensembles de mouvement browniens en topologie forte. En 2.2.2 les vitesses de la section 2.1 sont généralisées de  $d$  à  $d_\alpha$  pour les incréments browniens  $\Lambda_T$ . Comme  $\alpha_n \notin \mathcal{H}_\alpha$  nous avons dû approcher  $\mathcal{B}_\alpha$  par une suite d'espaces métriques  $\mathcal{B}_{n, \alpha}$  afin d'obtenir des lois fonctionnelles similaires pour  $\alpha_n$ . Les notions adéquates sont introduites en 2.2.3. et les résultats pour  $\Theta_n$  sont énoncés en 2.2.4 le long de  $\mathcal{B}_{n, \alpha}$  avec le même degré de précision que pour  $\Lambda_T$ .

### 2.2.1 – Ensembles de trajectoires browniennes

Les travaux de [LS] et, plus généralement, [DR] concernent les ensembles de trajectoires browniennes, en nombre croissant, pouvant par exemple représenter des particules dans le sillage d'une fusée. En quoi l'allure de ce sillage est-elle déterminée ? Il s'agit, à chaque instant  $t$ , de

$$DR_t = \left\{ \frac{W_i(\tau_t I)}{\sqrt{2\tau_t (\log \eta_t + \log \log \tau_t)}} : i = 1, 2, \dots, [\eta_t] \right\}$$

où les  $\{W_i(t) : t \geq 0\}$  sont des mouvements browniens standards,  $\tau_t > 0$  et  $\eta_t \geq 1$  sont croissantes. Nous cherchons à raffiner le résultat principal de [DR], qui s'énonce comme suit.

**Théorème 2D.** *Supposons que  $\varepsilon > 0$  et, si  $t \rightarrow \infty$ ,  $\log \eta_t / \log \log \tau_t \rightarrow c \in [0, \infty]$ ,  $\max(\eta_t, \tau_t) \rightarrow \infty$ . Alors  $\mathbb{P}(DR_t \subset \mathcal{S}_{c/(1+\varepsilon)} + \varepsilon B_0) \rightarrow 1$  et, si de plus  $\tau_t / \eta_t \downarrow$  ou  $(\log \eta_t) / \tau_t \downarrow$  lorsque  $c = \infty$ , il vient*

$\mathbb{P}(\exists t_0 : DR_t \subset \mathcal{S}_1 + \varepsilon B_0, t > t_0) = 1$ . Si les  $W_i$  sont indépendants alors  $\mathbb{P}(\mathcal{S}_{c/(1+c)} \subset DR_t + \varepsilon B_0) \rightarrow 1$  et, sous l'hypothèse ci-dessus lorsque  $c = \infty$ ,  $\mathcal{S}_1$  et  $\mathcal{S}_{c/(1+c)}$  sont le plus grand et le plus petit compacts de  $\mathcal{C}_0$  tels que  $\mathbb{P}(\exists t_0 : \mathcal{S}_{c/(1+c)} \subset DR_t + \varepsilon B_0 \subset \mathcal{S}_1 + 2\varepsilon B_0, t > t_0) = 1$ .

Cet énoncé recouvre, entre autres, [S] quand  $\tau_t = t$  et  $\eta_t = 1$ , [La] quand  $\tau_t = 1$  et  $\eta_t = t$ , [AK] et [LS] quand  $\tau_t = t$  et  $\eta_t$  est lent. Supposons alors que  $\tau_t \rightarrow \infty$  est continue et strictement croissante. En prenant  $\tau_t$  comme référence de temps, substituons  $t$  à  $\tau_t$  et  $\kappa_t = \eta_{\tau^{-1}(t)}$  à  $\eta_t$  pour étudier

$$\Omega_t = DR_{\tau^{-1}(t)} = \left\{ \frac{W_i(tI)}{v_t} : i = 1, \dots, [\kappa_t] \right\}, \quad v_t = \sqrt{2t(\log \kappa_t + \log \log t)}$$

avec  $\kappa_t \geq 1$  croissante. Précisons enfin notre triple problématique.

A quelles conditions sur  $\kappa_t$  la suite  $\Omega_t$  est-elle presque sûrement relativement compacte dans  $\mathcal{B}_\alpha$  ? Avec quelles vitesses ? Quels résultats supplémentaires entraîne l'indépendance des  $W_i$  ?

(i) La taille  $\kappa_t$  n'étant pas limitée comme en 2.1 par le nombre  $1/a_n$  d'incrémentés disjoints nous sommes ici confrontés à des ensembles plus grands et plus diversifiés que  $\Theta_n$ . Dans [5] nous aboutissons aux conditions ci-après. Pour les traduire en terme de  $DR_t$  il suffit de substituer  $(\tau_t, \eta_t)$  à  $(t, \kappa_t)$ .

$$H_1 : \kappa_t \uparrow \text{ et ou bien (i) } \frac{\kappa_t}{t^a} \uparrow \quad \text{ou bien (ii) } \frac{\log \kappa_t}{t^a} \downarrow \text{ pour un } a > 0.$$

$$H_2 : \frac{\log \kappa_t}{\log \log t} \rightarrow c \in [0, \infty].$$

$$H_3 : \limsup_{t \rightarrow \infty} \frac{\log \log \kappa_t}{\log \log t} = s < \infty.$$

$$H_4 : \text{ou bien (i) } \liminf_{t \rightarrow \infty} \frac{\log \log \kappa_t}{\log \log \log t} > 3 - 2\alpha \quad \text{ou bien (ii) } \limsup_{t \rightarrow \infty} \frac{\log \log \kappa_t}{\log \log \log t} < \infty.$$

Les quatre conditions  $H_{1,4}(i, ii)$  ne sont pas exclusives l'une de l'autre, sauf  $H_1(i)$  et  $H_4(ii)$ . Si  $\kappa_t$  est à variation régulière alors  $H_3$  contient  $H_1(ii)$ . Nous disons que  $\Omega_t$  est *très grand* si  $H_1(i)$  est vraie, ce qui implique  $H_4(i)$  et  $s \in [1, \infty]$ . Nous disons que  $\Omega_t$  est *grand* sous  $H_1(ii)$ ,  $H_3$  et  $H_4(i)$ . Dans ces deux situations,  $c = \infty$ . Ainsi  $\Omega_t$  est *petit* lorsqu'il vérifie  $H_1(ii)$ ,  $H_4(ii)$  et donc  $H_3$  avec  $s = 0$ . En particulier il suffit d'avoir  $H_1(ii)$  et  $H_2$  avec  $c < \infty$ .

D'après le Théorème 2D,  $H_{1,2}$  avec  $a = 1$  implique que  $\Omega_t$  admet  $\mathcal{S}_{c/(1+c)}$  comme limite faible et  $(\mathcal{S}_{c/(1+c)}, \mathcal{S}_1)$  comme limite forte, à vitesse inconnue, dans  $\mathcal{C}_0$ . D'après [5] ceci reste vrai dans  $\mathcal{C}_\alpha$ .

**Proposition 2.16.** *Sous  $H_{1,2}$  la famille  $\Omega_t$  est presque sûrement relativement compacte dans  $\mathcal{C}_\alpha$  avec un ensemble de fonctions adhérentes  $A \subset \mathcal{S}_1$  et un ensemble de fonctions limites  $L \subset \mathcal{S}_{c/(1+c)}$ . Si les processus  $W_i$  sont indépendants nous avons  $L = \mathcal{S}_{c/(1+c)} \subset \mathcal{S}_1 = A$ .*

L'énoncé complet est identique à 2D. Par conséquent, quelle que soit sa vitesse de croissance, qui doit être régulière, l'ensemble des trajectoires correctement normalisées se concentre dans un ensemble non aléatoire et en recouvre complètement un sous-ensemble. C'est un déterminisme, au premier ordre.

En outre, dans [5] nous montrons que des vitesses d'inclusion et de recouvrement sont garanties par la régularité  $H_{1,2,3,4}$ . Même pour  $\alpha = 0$  elles apportent une information nouvelle.

(ii) Soit  $\gamma(\alpha) = (1 - \alpha)/(3/2 - \alpha)$ , qui varie dans  $]1/2, 2/3]$ , et  $\sigma_\alpha$  une certaine constante universelle.

**Théorème 2.17.** Sous  $H_1(i)$  ou bien  $H_{1,3,4}$  nous avons, pour tout  $\sigma > \sigma_\alpha$ ,

$$\mathbb{P} \left( \exists t_0 : \Omega_t \subset V_\alpha \left( \mathcal{S}_1, \frac{\sigma}{(\log \kappa_t + \log \log t)^{\gamma(\alpha)}} \right), t > t_0 \right) = 1.$$

Ainsi, plus l'ensemble  $\Omega_t$  est grand, plus il se concentre vite, en partie à cause de  $v_t$ . Quand  $H_1(i)$  n'est pas de mise, la taille de  $\Omega_t$  est contrainte et les vitesses sont au mieux de l'ordre de  $1/(\log t)^p$

avec  $\rho > 0$  et au pire de l'ordre de  $1/(\log \log t)^{\gamma(\alpha)}$ . Si  $\kappa_t = 1$  et  $\alpha = 0$  le Théorème 2.17 correspond à [T], qui est optimal, et caractérise les  $\kappa_t \rightarrow \infty$  pour lesquels la vitesse  $1/(\log \log t)^{\gamma(\alpha)}$  reste inchangée.

Si les  $W_i$  sont indépendants nous savons démontrer que notre vitesse d'inclusion est optimale, comme au Théorème 2.1. Il est par ailleurs évident qu'une forte dépendance produit une concentration dans un sous-ensemble de  $\mathcal{S}_1$ . Par exemple, après  $\kappa_t$  trajectoires indépendantes, injectons  $\kappa_t^+$  répliques de  $W_1$  dans  $\Omega_t$  et supposons que le Théorème 2.17 s'applique à  $\kappa_t$ . Dans ce cas, lorsque  $t \rightarrow \infty$  il vient

$$\Omega_t \subset \sqrt{\frac{\log \kappa_t}{\log(\kappa_t + \kappa_t^+)}} V_\alpha \left( \mathcal{S}_1, \frac{\sigma}{(\log \kappa_t)^{\gamma(\alpha)}} \right)$$

et  $\Omega_t$  peut alors s'enfermer dans une boule arbitrairement petite si  $\log \kappa_t^+ \gg \log \kappa_t$ .

(iii) Pour  $\kappa_t \uparrow \infty$  sans autre contrainte asymptotique il se produit une concentration récurrente de

$$\Omega'_t = \left\{ \frac{W_i(tI)}{\sqrt{2t \log \kappa_t}} : i = 1, 2, \dots, [\kappa_t] \right\}.$$

**Proposition 2.18.** *Si  $\kappa_t \rightarrow \infty$  et  $\sigma > \sigma_\alpha$ ,  $\Omega'_t$  se concentre faiblement dans  $V_\alpha(\mathcal{S}_1, \sigma/(\log \kappa_t)^{\gamma(\alpha)})$  et*

$$\mathbb{P} \left( \exists t_k \uparrow \infty : \Omega'_{t_k} \subset V_\alpha \left( \mathcal{S}_1, \frac{\sigma}{(\log \kappa_{t_k})^{\gamma(\alpha)}} \right), k \geq 1 \right) = 1.$$

En réalité, avec probabilité un, de toute suite  $t_n \uparrow \infty$  nous pouvons extraire une sous-suite  $t_{n_k}$  vérifiant la concentration asymptotique à cette vitesse. Sous  $H_2$  avec  $c \in (0, \infty)$  il s'ensuit

$$\mathbb{P} \left( \exists t_k \uparrow \infty : \Omega_{t_k} \subset V_\alpha \left( (1 + \xi_{t_k}) \mathcal{S}_{c/(1+c)}, \frac{\sigma(c/(1+c))^{1/2-\gamma(\alpha)}}{(\log \kappa_{t_k} + \log \log t_k)^{\gamma(\alpha)}} \right), k \geq 1 \right) = 1$$

avec  $\xi_t \sim (\log \kappa_t / c \log \log t - 1) / 2(1+c)$  pouvant perturber la vitesse de concentration récurrente autour de  $\mathcal{S}_{c/(1+c)}$ . Si  $c = 0$  il ressort que  $\Omega_t$  se concentre dans une petite boule topologique ou de Strassen selon la vitesse dans  $H_2$ . Si  $c = 0$  et  $\xi_t = (\log \log \log t) / 2 \log \kappa_t \rightarrow 0$  nous montrons alors que

$$\Omega''_t = \left\{ \frac{W_i(tI)}{\sqrt{2t(\log \kappa_t - \log \log \log t)}} : i = 1, 2, \dots, [\kappa_t] \right\}$$

se concentre infiniment souvent dans  $\mathcal{S}_1 + V_\alpha(\xi_t \mathcal{S}_1, \sigma/(\log \kappa_t)^{\gamma(\alpha)})$ . Ceci est une version fonctionnelle du phénomène exhibé dans [CsR] pour le plus grand increment de longueur  $t/\kappa_t$  du processus de Wiener dans  $\mathcal{B}$ . Si  $\log \kappa_t \sim \kappa \log \log \log t$  avec  $\kappa > 1$  nous obtenons des inclusions  $\Omega''_t \subset (1 + \varepsilon) \mathcal{S}_{\kappa/(\kappa-1)}$  au lieu de  $(1 + \varepsilon) B_0$  comme espéré. Ces résultats incomplets pour les trop petits ensembles, tout comme ceux de taille  $\log \log \log n$  ( $c = 0$ ) de 2.1, seront affinés à la lumière de [18].

(iv) Dans [5] nous apprenons aussi que la dépendance des  $W_i$  ne peut qu'être néfaste aux vitesses de convergence ponctuelle, à l'inverse de la concentration (ii) ci-dessus. C'est un phénomène général, la corrélation réduit la variété. Soit  $C_\alpha > 0$  la constante universelle de 2.2.2 et  $\beta = 1/2 - \alpha$ .

**Théorème 2.19.** *Supposons  $H_1(i)$  ou  $H_{1,2,3}$ . Si  $J(f) < 1$  nous avons*

$$\liminf_{t \rightarrow \infty} (\log \kappa_t + \log \log t)^{1-\alpha} d_\alpha(f, \Omega_t) \geq \frac{1}{\sqrt{2}} \left( \frac{C_\alpha}{1 - J(f)} \right)^\beta \quad p.s.$$

et, si de plus  $c > 0$  et  $J(f) < c/(1+c)$ ,

$$\limsup_{t \rightarrow \infty} (\log \kappa_t + \log \log t)^{1-\alpha} d_\alpha(f, \Omega_t) \geq \frac{1}{\sqrt{2}} \left( \frac{C_\alpha}{c/(1+c) - J(f)} \right)^\beta \quad p.s.$$

Si les processus  $W_i$  sont indépendants, les deux inégalités sont des égalités.

Lorsque  $c = \infty$  nous prouvons de même, après renormalisation et en cas d'indépendance des  $W_i$ ,

$$\inf_{i \leq \kappa_t} \left\| \frac{W_i(tI)}{\sqrt{2t \log \kappa_t}} - f \right\|_\alpha \sim \frac{1}{\sqrt{2}} \left( \frac{C_\alpha}{1 - J(f)} \right)^\beta \frac{1}{(\log \kappa_t)^{1-\alpha}} \quad p.s.$$

**(v) Principe de démonstration.** Au Théorème 2.17 nous exploitons la régularité Höldérienne de  $W$  et  $\mathcal{S}_1$  dans le contrôle des événements par blocs de temps  $T \in ]t_n, t_{n+1}]$  en vu d'appliquer l'inégalité de concentration suivante, issue de la méthode [Le] en calculant les seconds ordres,

$$\mathbb{P}(W_1 \notin V_\alpha(\gamma \mathcal{S}_1, \varepsilon)) \leq D \varepsilon^{(1+\beta)/2\beta} \exp \left( \frac{q^2 \sigma_\alpha}{4\varepsilon^{1/\beta}} - \frac{\gamma \varepsilon}{p} - \frac{\gamma^2}{2} - \frac{\varepsilon^2}{2p^2} \right)$$

où  $1/p + 1/q = 1$ ,  $D > 0$  et  $\sigma_\alpha > 0$ . Les conditions  $H_{1,3,4}$  peuvent s'affaiblir pour tenir compte des irrégularités de croissance de  $\Omega_t$ . Sous l'hypothèse  $H_1(ii)$  d'ensembles pas trop grands il suffit qu'il existe une suite  $\theta_n \uparrow \infty$  telle que, pour  $C > 0$  et  $L^{-1}$  la fonction inverse de  $L(t) = \log \kappa_t + \log \log t$ ,

$$\theta_n \left( \frac{L^{-1}(\theta_n)}{L^{-1}(\theta_{n-1})} - 1 \right)^{\beta/\gamma(\alpha)} \rightarrow 0, \quad \sum_{n=1}^{\infty} \frac{1}{\log L^{-1}(\theta_n)} \exp \left( -C \theta_n^{1-\gamma(\alpha)} \right) < \infty.$$

Par exemple, sous  $H_4(ii)$  la suite  $\theta_n = \log \kappa_{\exp(n/(\log n)^\delta)} + \log n - \delta \log \log n$ , avec  $\delta$  grand, convient. Au Théorème 2.19 nous tirons profit des petites déviations de [BR] et pouvons aussi affaiblir  $H_{1,2,3}$ . Sous  $H_1(ii)$ ,  $H_2$  avec  $c = \infty$  et  $\kappa^{-1}(x) = \inf \{t : \kappa_t = x\}$  il suffit de trouver  $\theta_n \uparrow \infty$  telle que

$$\theta_n \left( \frac{\kappa^{-1}(\exp(\theta_n))}{\kappa^{-1}(\exp(\theta_{n-1}))} - 1 \right)^{\beta(\alpha)/(1-\alpha)} \rightarrow 0, \quad \sum_{n=1}^{\infty} \exp(-C \theta_n) < \infty.$$

Par exemple  $H_3$  autorise  $\theta_n = \log \kappa_{\exp(n^\lambda)}$  avec  $\lambda$  petit et  $H_1(i)$  autorise  $\theta_n = n^\lambda$ .  $\square$

### 2.2.2 – Incréments browniens

Les résultats précédents s'appliquent aux ensembles d'incrément browniens  $\Lambda_T$  de 2.1.3 comme si seuls les  $\kappa_T \sim T/A_T$  incréments disjoints comptaient. Le problème dans les preuves consiste dès lors à contrôler les incréments voisins dans  $\mathcal{C}_\alpha$ . Une borne exponentielle s'appliquant directement aux déviations du module de régularité Höldérien  $\omega_{\alpha,W}(\delta)$  est à l'étude mais nous parvenons dans [13] à nous en passer par diverses astuces exploitant le fait que  $\mathbb{P}(W \in \mathcal{C}_\rho) = 1$ ,  $\alpha < \rho < 1/2$ .

**(i)** Nos énoncés sont valables pour  $L_T = \log(T/A_T) + \log \log T$  et  $\mathcal{T}_T = [0, T - A_T]$  ou sous certaines conditions sur l'entropie  $\kappa_T$  de  $\mathcal{T}_T$  par rapport à  $A_T$ , la moins exigeante étant  $\log \kappa_T \sim \log(T/A_T)$ .

**Proposition 2.20.** *Supposons (3) et  $\varepsilon > 0$ . Alors  $\mathbb{P}(\Lambda_T \subset V_\alpha(\mathcal{S}_{c/(1+c)}, \varepsilon) \subset V_\alpha(\Lambda_T, 2\varepsilon)) \rightarrow 1$  et  $\mathbb{P}(\exists T_0 : \Lambda_T \subset V_\alpha(\mathcal{S}_1, \varepsilon), T > T_0) = 1$ . De plus  $\mathcal{S}_1$  et  $\mathcal{S}_{c/(1+c)}$  sont le plus grand et le plus petit compacts de  $\mathcal{C}_\alpha$  tels que pour tout  $\varepsilon > 0$ ,  $\mathbb{P}(\exists T_0 : \mathcal{S}_{c/(1+c)} \subset V_\alpha(\Lambda_T, \varepsilon) \subset V_\alpha(\mathcal{S}_1, 2\varepsilon), T > T_0) = 1$ .*

**(ii)** Les vitesses de recouvrement correspondent à celles des petits ensembles de 2.2.1 et sont de nouveau parfaitement déterminées par quelques principes fondamentaux du hasard.

**Théorème 2.21.** *Supposons (3),  $L_T = \log(T/A_T) + \log \log T$  et  $\mathcal{T}_T = [0, T - A_T]$ . Si  $f \in \mathcal{S}_1^-$  alors*

$$\liminf_{T \rightarrow \infty} L_T^{1-\alpha} d_\alpha(f, \Lambda_T) = \frac{1}{\sqrt{2}} \left( \frac{C_\alpha}{1 - J(f)} \right)^\beta \quad p.s.$$

et, si  $f \in \mathcal{S}_{c/(1+c)}^-$ ,

$$\limsup_{T \rightarrow \infty} L_T^{1-\alpha} d_\alpha(f, \Lambda_T) = \frac{1}{\sqrt{2}} \left( \frac{C_\alpha}{c/(1+c) - J(f)} \right)^\beta \quad p.s.$$

Si  $f \in \mathcal{S}_{c/(1+c)}$  alors, pour  $\gamma(\alpha) = (1 - \alpha)/(3/2 - \alpha)$ ,

$$\limsup_{T \rightarrow \infty} L_T^{\gamma(\alpha)} d_\alpha(f, \Lambda_T) < \infty \quad p.s.$$

Si  $\sigma > \sigma_\alpha$  et  $c > 0$ ,

$$\begin{aligned} \mathbb{P} \left( \exists T_0 : \Lambda_T \subset V_\alpha \left( \mathcal{S}_1, \frac{\sigma}{L_T^{\gamma(\alpha)}} \right), T > T_0 \right) &= 1, \\ \mathbb{P} \left( \exists T_k \uparrow \infty : \Lambda_{T_k} \subset V_\alpha \left( \mathcal{S}_{c/(1+c)}, \frac{\sigma}{L_{T_k}^{\gamma(\alpha)}} \right), k \geq 1 \right) &= 1. \end{aligned}$$

**(iii) Principe de démonstration.** Par exemple, pour évaluer  $d_\alpha(f, \Lambda_T)$  dans 2.21 nous utilisons la régularité Höldérienne de  $W_1 = \{W(t) : t \in [0, 1]\}$  et  $\mathcal{S}_1$  pour contrôler les événements par blocs  $T \in ]T_k, T_{k+1}]$  et appliquer de façon optimale le principe de petite déviation [BR] affirmant que

$$\lim_{\xi \rightarrow \infty} \frac{1}{\xi^2} \log \mathbb{P} \left( \left\| \frac{W_1}{\xi} - f \right\|_\alpha \leq \frac{r}{\xi^{2(1-\alpha)}} \right) = -\frac{C_\alpha}{r^{1/\beta}} - \frac{J(f)}{2}$$

pour tout  $f \in \mathcal{S}_\theta$  et  $r > 0$ , où  $C_\alpha$  est la constante inconnue de la probabilité de petite boule de [KL],

$$\lim_{\xi \rightarrow \infty} \frac{1}{\xi^{1/\beta}} \log \mathbb{P} \left( \|W_1\|_\alpha \leq \frac{1}{\xi} \right) = -C_\alpha. \quad \square$$

### 2.2.3 – Compacité relative et espaces pré-Höldériens

L'incompatibilité entre mesurabilité et topologie posée par la convergence faible de  $\alpha_n$  vers le pont brownien  $B$  a trouvé différents paliatifs : affaiblissement de la topologie, de type Skorohod, ajustement des définitions de convergence, recours aux mesures extérieures. Dans ce type de problème où les processus courants et limite ne vivent pas dans le même espace fonctionnel, l'approche suivante est originale. Nous réglons la convergence le long d'une suite d'espaces. La difficulté initiale reste la même, les points de saut de  $\alpha_n$  sont aléatoires et  $\alpha_n \notin \mathcal{H}_\alpha$  incontrôlable dans  $\mathcal{C}_\alpha$  alors que  $\mathbb{P}(B \in \mathcal{C}_\alpha) = 1$ .

(i) A toute suite déterministe  $\delta_n \downarrow 0$  associons la suite  $\mathcal{B}_{\alpha,n}$  des espaces de Banach  $\mathcal{B}$  équipés de

$$\|f\|_{\alpha,n} = \omega_{\alpha,f}^+(\delta_n) = \sup_{\delta_n < |t-s| \leq 1} \frac{|f(t) - f(s)|}{|t-s|^\alpha} < \infty$$

que nous qualifions de norme pré-höldérienne. Si  $f \in \mathcal{C}_\alpha$  et  $n \rightarrow \infty$  alors  $\|f\|_{\alpha,n} \uparrow \|f\|_\alpha$ .

Notons  $V_{\alpha,n}$  les voisinages et  $d_{\alpha,n}(S, A) = \inf \{\varepsilon > 0 : S \subset V_{\alpha,n}(A, \varepsilon)\}$  la distance par inclusion.

**Remarque.** Appliquer 3.4.5 à la métrique tronquée  $d_{\alpha,n}$  montre que  $\alpha_n$  vérifie un principe d'invariance fort le long de  $\mathcal{B}_{\alpha,n}$ , non optimal, mais nous exploitons [KMT] pour ne pas contraindre  $a_n$  de (1).

Notre problème se formule ainsi. Comment utiliser les résultats gaussiens sur  $\mathcal{B}_\alpha$  et empiriques sur  $\mathcal{B}_0$  pour régler  $\delta_n$  de sorte que  $\alpha_n$  oscille dans  $\mathcal{B}_{\alpha,n}$  comme un mouvement brownien dans  $\mathcal{B}_\alpha$  ?

**Définitions.** Nous disons que  $\Theta_n$  est presque sûrement relativement compacte le long de  $\mathcal{B}_{\alpha,n}$  si, avec probabilité un, il existe un ensemble compact, éventuellement aléatoire,  $\Theta$  de  $\mathcal{B}_\alpha$  tel que  $d_{\alpha,n}(\Theta_n, \Theta) \rightarrow 0$ . Dans ce cas, écrivons  $\Theta_n \rightsquigarrow_{\mathcal{B}_{\alpha,n}} (\underline{\Theta}, \bar{\Theta})$  p.s. et disons que  $\Theta_n$  a un ensemble limite extérieur  $\bar{\Theta}$  et intérieur  $\underline{\Theta}$  si  $\bar{\Theta} \neq \emptyset$  et  $\underline{\Theta}$  sont deux compacts déterminés dans  $\mathcal{B}_\alpha$  tels qu'avec probabilité un,

$$\underline{\Theta} = \bigcap_{\varepsilon > 0} \bigcup_{m > 0} \bigcap_{n > m} V_{\alpha,n}(\Theta_n, \varepsilon) \subset \bar{\Theta} = \bigcap_{\varepsilon > 0} \bigcap_{m > 0} \bigcup_{n > m} V_{\alpha,n}(\Theta_n, \varepsilon).$$



**Lemme 2.22.** Soit  $\underline{\Theta}, \bar{\Theta} \neq \emptyset$  deux compacts de  $\mathcal{B}_\alpha$ . Il y a équivalence entre  $\Theta_n \rightsquigarrow_{\mathcal{B}_{\alpha,n}} (\underline{\Theta}, \bar{\Theta})$  p.s. et

$$(Con_\alpha) \quad \mathbb{P}(\exists n_0 \geq 1 : \Theta_n \subset V_{\alpha,n}(\bar{\Theta}, \varepsilon), n \geq n_0) = 1, \quad \varepsilon > 0.$$

$$(Out_\alpha) \quad \mathbb{P}(\exists n_k \uparrow \infty : f \in V_{\alpha,n_k}(\Theta_{n_k}, \varepsilon), k \geq 1) = 1, \quad f \in \bar{\Theta}, \quad \varepsilon > 0.$$

$$(Cov_\alpha) \quad \mathbb{P}(\exists n_0 \geq 1 : \underline{\Theta} \subset V_{\alpha,n}(\Theta_n, \varepsilon), n \geq n_0) = 1, \quad \varepsilon > 0.$$

$$(In_\alpha) \quad \mathbb{P}(\exists n_k \uparrow \infty : f \notin V_{\alpha,n_k}(\Theta_{n_k}, \varepsilon), k \geq 1) = 1, \quad f \notin \underline{\Theta}, \quad \text{un } \varepsilon > 0.$$

La nécessité de  $(Con_\alpha)$  n'est pas immédiate et requiert un procédé diagonal. Nous disons que  $\Theta_n$  recouvre faiblement  $\tilde{\Theta}$  le long de  $\mathcal{B}_{\alpha,n}$  si  $\tilde{\Theta}$  est le plus petit compact tel que

$$(Wk_\alpha) \quad \mathbb{P}(\exists n_k \uparrow \infty : \Theta_{n_k} \subset V_{\alpha,n_k}(\tilde{\Theta}, \varepsilon), k \geq 1) = 1, \quad \varepsilon > 0.$$

Clairement  $\underline{\Theta} \subset \tilde{\Theta} \subset \bar{\Theta}$  et en cas d'égalités nous écrivons  $\Theta_n \rightsquigarrow_{\mathcal{B}_{\alpha,n}} \bar{\Theta}$  p.s. et ceci équivaut à exiger  $(Con_\alpha)$  et  $(Cov_\alpha)$  de  $\bar{\Theta}$ .

Nous avons les inégalités  $\|f\|_{\alpha,n} \leq \|f\|_{\alpha,n+1} \leq \|f\|_\alpha = \max(\omega_{\alpha,f}(\delta_n), \|f\|_{\alpha,n})$  et, si  $\delta_n < 1/(2+2^\alpha)^{1/\alpha}$ ,  $\max(\|f\|, \|f\|_{0,n}) \leq \|f\|_{\alpha,n} \leq \|f\|_{0,n}/\delta_n^\alpha \leq \|f\|_0/\delta_n^\alpha \leq 2\|f\|/\delta_n^\alpha$ . À l'aide de contre exemples il s'ensuit que  $V_\alpha(f, \varepsilon) \subsetneq V_{\alpha,n}(f, \varepsilon) \subsetneq V(f, \varepsilon) \subsetneq V_{\alpha,n}(f, 2\varepsilon/\delta_n^\alpha)$ . En prenant soin de travailler sur  $\mathcal{B}_{\alpha+\varepsilon}$  pour contrôler fonctions et processus, ceci nous permet de raccorder la topologie uniforme de  $\mathcal{B}_{\alpha,n}$  à la topologie forte de  $\mathcal{B}_\alpha$ . Les raisonnements déjà longs des preuves de type 2.1 se corsent encore.

L'énergie de  $A \subset \mathcal{B}$  est  $J(A) = \inf_A J$ . Il est clair que  $J(V(f, \varepsilon)) \leq J(V_{\alpha,n}(f, \varepsilon)) \leq J(V_\alpha(f, \varepsilon)) < J(f)$  si  $\varepsilon > 0, n \geq 1$  et  $0 < J(f) \leq \theta$ . Si  $0 < J(A) < \infty$  il vient, en notant  $V_\alpha^+$  la fermeture de  $V_\alpha$  dans  $\mathcal{B}_\alpha$ ,  $J(V_\alpha(A, 2\varepsilon)) < J(A) - \rho_A(\varepsilon) < J(V_\alpha^+(A, \varepsilon)) \leq J(V_\alpha(A, \varepsilon)) < J(A)$  avec  $\rho_A(\varepsilon) \rightarrow 0$  quand  $\varepsilon \rightarrow 0$ .

(ii) L'analogie de [DM] obtenu dans [3,17] s'énonce ainsi, avec  $L_n = |\log a_n|$  et  $\kappa_n$  comme en 2.1.2.

**Proposition 2.23.** Sous (1) avec  $c = d = \infty$  et  $\log \kappa_n \sim |\log a_n|$ , si  $\delta_n \gg ((\log n)^2 / na_n |\log a_n|)^{1/2\alpha}$  alors  $\mathcal{S}_1$  vérifie  $(Cov_\alpha)$ . Si de plus  $\delta_n (na_n)^{1/2\alpha} \uparrow \infty$ ,  $\limsup_{n \rightarrow \infty} \log(na_n) / \log \log n = 1 + d_0 > 1$  et  $\delta_n \gg 1/(\log n)^\delta$ ,  $\delta < d_0/4\alpha$  alors  $\mathcal{S}_1$  vérifie  $(Con_\alpha)$ . Ces conditions impliquent donc  $\Theta_n \rightsquigarrow_{\mathcal{B}_{\alpha,n}} \mathcal{S}_1$  p.s.

Il ressort que la concentration est plus exigeante que le recouvrement, le long de  $\mathcal{B}_{n,\alpha}$ . Dans le premier cas  $\delta_n$  doit être une puissance de  $\log n$ , ce qui s'améliore en  $\exp(-(\log n)^\delta)$ ,  $\delta < 1$  quand  $d_0 = \infty$ . Dans le second,  $\delta_n = 1/n^\gamma$  est autorisé si  $a_n$  le permet et cela suffit à approcher tout  $\mathcal{S}_1$  dans  $\mathcal{B}_{\alpha,n}$ .

(iii) Dans [3,17] notre analogue de [D2] s'énonce de façon similaire, avec  $L_n = |\log a_n| + \log \log n$ . Soit  $\zeta_k$  la  $a_{n_k}$ -capacité de  $\bigcap_{n_k < n \leq n_{k+1}} T_n$  pour  $\theta$  assez petit et  $n_k = \lfloor \exp k^\theta \rfloor$ .

**Proposition 2.24.** Sous (1) avec  $0 < c < \infty$ , si  $\log \zeta_k \sim |\log a_{n_k}|$  et  $\delta_n > 1/(\log n)^\delta$ ,  $\delta < \infty$  alors  $\mathcal{S}_{c/(1+c)}$  vérifie  $(Cov_\alpha)$  et  $(Wk_\alpha)$ . Si  $\delta < 1/2\alpha$  et  $\delta_n (na_n)^{1/2\alpha} \uparrow \infty$  alors  $\mathcal{S}_{c/(1+c)}$  et  $\mathcal{S}_1$  vérifient  $(In_\alpha)$  et  $(Out_\alpha)$  respectivement. Sous  $\delta_n (na_n)^{1/2\alpha} \uparrow \infty$  et  $\delta_n > 1/(\log \log n)^\delta$ ,  $\delta < 1/2\alpha$  alors  $\mathcal{S}_1$  vérifie  $(Con_\alpha)$ . Les conditions précédentes impliquent donc  $\Theta_n \rightsquigarrow_{\mathcal{B}_{\alpha,n}} (\mathcal{S}_{c/(1+c)}, \mathcal{S}_1)$  p.s.

La concentration demande cette fois que  $\delta_n$  soit une puissance de  $\log \log n$  et le recouvrement autorise une puissance de  $\log n$ . La petite condition de monotonie semble pouvoir se contourner.

Dans les Propositions 2.23 et 2.24 les modules pré-höldériens de non-différenciabilité et de continuité

$$\underline{\omega}_{\alpha,n} = \inf_{t \in T_n} \|\Delta \alpha_n(a_n, t)\|_{\alpha,n} \leq \bar{\omega}_{\alpha,n} = \sup_{t \in T_n} \|\Delta \alpha_n(a_n, t)\|_{\alpha,n}$$

vérifient donc, pour  $L_n = |\log a_n| + \log \log n$  et avec probabilité un,

$$\underline{\omega}_{\alpha,n} \ll \sqrt{a_n L_n} \quad \text{et} \quad \liminf_{n \rightarrow \infty} \frac{\bar{\omega}_{\alpha,n}}{\sqrt{2a_n L_n}} = \sqrt{\frac{c}{1+c}} \leq \limsup_{n \rightarrow \infty} \frac{\bar{\omega}_{\alpha,n}}{\sqrt{2a_n L_n}} = 1.$$

**(iv) Principe de démonstration.** Nous mettons au point une technique de preuve pour chaque condition du Lemme 2.22 et l'adaptions selon  $c$ . Quand un blocage est nécessaire, circuler entre les topologies nous permet de contrôler le module de régularité de  $\alpha_n$  via [MSW] sur  $\mathcal{B}_0$ , si  $na_n\delta_n^{2\alpha} \rightarrow \infty$ ,

$$\mathbb{P}\left(\omega_{\alpha_n}(\delta_n a_n) \geq \varepsilon \delta_n^\beta \sqrt{a_n}\right) \leq \frac{C_\eta}{\delta_n a_n} \exp\left(-\frac{\varepsilon^2}{2 + \eta} \frac{1}{\delta_n^{2\alpha}}\right)$$

pour finalement obtenir, suite au découpage usuel en trois termes,

$$\max_{n_k < n \leq n_{k+1}} \sup_{0 \leq t \leq 1 - a_n} \left\| \frac{\Delta \alpha_{n_{k+1}}(a_{n_{k+1}}, t)}{\sqrt{2a_{n_{k+1}} L_{n_{k+1}}}} - \frac{\Delta \alpha_n(a_n, t)}{\sqrt{2a_n L_n}} \right\|_{\alpha, n} \rightarrow 0 \quad p.s.$$

Après les arguments de poissonisation, d'approximation forte et de contrôle par les propriétés de  $\mathcal{S}_1$  et  $\mathcal{B}_\rho$  pour  $\rho > \alpha$ , nous utilisons la continuité des énergies et le principe de grande déviations de [BBK] qui atteste que si  $A_\alpha \subset \mathcal{C}_\alpha$  est d'intérieur  $A_\alpha^-$  et de fermeture  $A_\alpha^+$  dans  $\mathcal{B}_\alpha$  alors

$$-J(A_\alpha^-) = \liminf_{\xi \rightarrow \infty} \frac{1}{\xi^2} \log \mathbb{P}\left(\frac{W_1}{\sqrt{2\xi}} \in A_\alpha\right) \leq \limsup_{\xi \rightarrow \infty} \frac{1}{\xi^2} \log \mathbb{P}\left(\frac{W_1}{\sqrt{2\xi}} \in A_\alpha\right) = -J(A_\alpha^+). \quad \square$$

## 2.2.4 – Vitesses extérieures le long de $\mathcal{B}_{n, \alpha}$

(i) Dans [1] nous remarquons qu'en appliquant les propriétés de  $\mathcal{B}_{\alpha, n}$  et la concentration [T],

$$\mathbb{P}(\exists n_0 : \Theta_n \subset V_{\alpha, n}(\mathcal{S}_1, \varepsilon_n), \quad n > n_0) = 1$$

pourvu que (1) soit satisfaite avec  $c = d = \infty$ ,  $L_n = |\log a_n|$  et  $L_n^{2/3} \varepsilon_n \delta_n^\alpha \rightarrow \infty$ . Par suite, lorsque  $\delta_n \gg 1/|\log a_n|^{\gamma_1}$  avec  $\gamma_1 < 2/3\alpha$  la vitesse induite est au mieux en  $1/|\log a_n|^{2/3 - \alpha\gamma_1}$ . Autrement dit, pour les très lents  $\delta_n$  nous sommes assurés d'une vitesse en puissance  $2/3$  de type 2.1, dans  $\mathcal{B}_0$ .

Ce constat est incomplet. Que se passe-t-il si la métrique est plus proche de  $\mathcal{B}_\alpha$  ?

(ii) Dans [17] nous améliorons notablement l'argument (i) pour  $\gamma_1$  grand. Une constante  $\nu_\alpha > 0$  apparaît, fonction de  $\alpha$ , qui décide des  $\delta_n$  pour lesquels la vitesse est de type uniforme ou höldérienne.

**Théorème 2.25.** *Supposons (1) avec  $c = d = \infty$ ,  $L_n = |\log a_n|$ ,  $\log(na_n) \ll \log \log n$ ,  $\delta_n \downarrow 0$  et  $\delta_n < 1/|\log a_n|^{\gamma_1}$  avec  $\gamma_1 \in ]\nu_\alpha, \infty[$  alors il vient, pour  $\gamma(\alpha) = (1 - \alpha)/(3/2 - \alpha)$  et  $\sigma$  assez grand,*

$$\mathbb{P}\left(\exists n_0 : \Theta_n \subset V_{\alpha, n}\left(\mathcal{S}_1, \frac{\sigma}{L_n^{\gamma(\alpha)}}\right), \quad n > n_0\right) = 1.$$

Par comparaison avec le Théorème 2.21 nous voyons que la vitesse est aussi bonne le long de  $\mathcal{B}_{\alpha, n}$  pour  $\alpha_n$  que dans  $\mathcal{B}_\alpha$  pour  $W$  pourvu que  $\mathcal{B}_{\alpha, n}$  s'éloigne assez rapidement de  $\mathcal{B}_0$ . Nous conjecturons l'optimalité de cette vitesse, ainsi qu'une vitesse de recouvrement  $d_{\alpha, n}(\mathcal{S}_1, \Theta_n) = O_{p.s.}(1/|\log a_n|^{\gamma(\alpha)})$ . Concernant  $\delta_n$  la contrainte de concentration est renforcée par rapport à la Proposition 2.23 en  $d_0 = \infty$  puisque seules les grandes puissances de  $\log n$  sont autorisées pour garantir la vitesse de type  $\mathcal{B}_\alpha$ .

Ceci s'étend aux grands incréments,  $c < \infty$ , et dans ce cas  $\delta_n < 1/(\log \log n)^{\gamma_1}$ . Le réglage de  $\delta_n$  s'apparente au choix d'une suite de paramètres dans une collection d'espaces emboîtés de type [MS].

**Application.** Si  $Y_n = \sup_{g \in \Theta_n} \varphi(g)$  et  $Z_n = \inf_{g \in \Theta_n} \varphi(g)$  où  $\varphi$  est continue de  $\mathcal{B}_{\alpha, n}$  dans  $\mathbb{R}$  de module  $\sup_{\|f-g\|_{\alpha, n} \leq \varepsilon} |\varphi(f) - \varphi(g)| \leq \pi_n(\varepsilon)$ . Sous les hypothèses du Théorème 2.25 nous avons

$$\mathbb{P}\left(\exists n_0 : \inf_{f \in \mathcal{S}_1} \varphi(f) - \pi_n\left(\frac{\sigma}{L_n^{\gamma(\alpha)}}\right) < Z_n \leq Y_n < \sup_{f \in \mathcal{S}_1} \varphi(f) + \pi_n\left(\frac{\sigma}{L_n^{\gamma(\alpha)}}\right), \quad n > n_0\right) = 1.$$

**(iii) Principe de démonstration.** Nous nous ramenons à la concentration brownienne évoquée en 2.2.1(v) par invariance en circulant entre les trois métriques, en contrôlant  $W$  et  $\mathcal{S}_1$  dans  $\mathcal{B}_{\alpha+\varepsilon}$ , en jouant sur les comparaisons de module et en adaptant la technique de preuve du Théorème 2.1.  $\square$

### 2.2.5 – Vitesses intérieures le long de $\mathcal{B}_{n,\alpha}$

(i) Considérons le module d'oscillation fonctionnel pré-höldérien de  $\alpha_n$  centré en  $f \in \mathcal{S}_1$ , c'est-à-dire  $d_{\alpha,n}(f, \Theta_n)$ . Soit  $C_\alpha$  la constante des probabilités de petites boules dans  $\mathcal{B}_\alpha$  de [BR,KL].

**Théorème 2.26.** *Sous (1) avec  $c = \infty$ ,  $\log \kappa_n \sim L_n = |\log a_n|$ ,  $\delta_n \downarrow 0$ ,*

$$\liminf_{n \rightarrow \infty} \frac{\log(na_n)}{\log \log n} = \gamma > \frac{3 - 4\alpha}{1 - 2\alpha}$$

et

$$1 + \frac{1}{1 - 2\alpha} < \liminf_{n \rightarrow \infty} \frac{|\log \delta_n|}{\log |\log a_n|} = \gamma_1 \leq \limsup_{n \rightarrow \infty} \frac{|\log \delta_n|}{\log |\log a_n|} = \gamma_2 < 1 + \frac{\gamma - 3}{2\alpha}$$

il vient, pour  $f \in \mathcal{S}_1^-$ ,

$$\lim_{n \rightarrow \infty} |\log a_n|^{1-\alpha} d_{\alpha,n}(f, \Theta_n) = \frac{1}{\sqrt{2}} \left( \frac{C_\alpha}{1 - J(f)} \right)^\beta \quad p.s.$$

et, si  $J(f) = 1$  et  $\gamma(\alpha) = (1 - \alpha) / (3/2 - \alpha)$ ,

$$\limsup_{n \rightarrow \infty} |\log a_n|^{\gamma(\alpha)} d_{\alpha,n}(f, \Theta_n) < \infty \quad p.s.$$

En  $f = 0$  nous obtenons la non-différentiabilité  $\underline{\omega}_{\alpha,n} \sim C_\alpha^\beta \sqrt{a_n} / |\log a_n|^\beta$  et un rapport des modules

$$\frac{\bar{\omega}_{\alpha,n}}{\underline{\omega}_{\alpha,n}} \sim \frac{\sqrt{2}}{C_\alpha^\beta} |\log a_n|^{1-\alpha} \quad p.s.$$

moins important que dans  $\mathcal{B}_0$ . La vitesse optimale  $|\log a_n|^{1-\alpha}$  correspond à celle du cas brownien dans  $\mathcal{B}_\alpha$ , ce qui était notre but. De même, la jonction avec le cas  $\alpha = 0$  étudié en 2.1.2 est satisfaisante, notamment la contrainte  $\gamma > 3$ . Les  $\delta_n$  autorisés sont des puissances de  $\log n$  particulières.

L'interdiction des trop grands  $\delta_n$  et la condition  $\gamma_2 \geq 1$  s'expliquent par le fait que les limites

$$0 = \lim_{n \rightarrow \infty} |\log a_n|^{1-\alpha} d(f, \Theta_n) \leq \lim_{n \rightarrow \infty} |\log a_n|^{1-\alpha} d_{\alpha,n}(f, \Theta_n) \leq \lim_{n \rightarrow \infty} \frac{2 |\log a_n|}{\delta_n^\alpha |\log a_n|^\alpha} d(f, \Theta_n)$$

sont toutes nulles si  $\delta_n |\log a_n| \rightarrow \infty$ . Ainsi nos hypothèses sur  $\delta_n$  assurent que  $\mathcal{B}_{\alpha,n}$  est assez proche de  $\mathcal{B}_\alpha$  pour le contrôle des petites déviations et assez proche de  $\mathcal{B}_0$  pour le contrôle des oscillations.

Le Théorème 2.26 décrit la convergence des incréments moyens. Celle des grands incréments s'obtient de façon similaire, avec le cortège usuel de difficultés et d'hypothèses supplémentaires.

**(ii) Application.** Si  $a_n \in ]1/n^b, 1/n^a[$  avec  $0 < a < b < 1$ ,  $\delta_n = 1/(\log n)^{\gamma_1}$ ,  $Y_n = \sup_{g \in \Theta_n} \varphi_n(g)$  et  $Z_n = \inf_{g \in \Theta_n} \varphi_n(g)$  avec  $\varphi_n$  continue de  $\mathcal{B}_{\alpha,n}$  dans  $\mathbb{R}$  de module  $\sup_{\|f-g\|_{\alpha,n} \leq \varepsilon} |\varphi_n(f) - \varphi_n(g)| \leq \phi_n(\varepsilon)$  alors le Théorème 2.26 et la conjecture 2.2.4 impliquent, pour un  $\zeta > 0$  et presque sûrement en  $n \rightarrow \infty$ ,

$$\left| Z_n - \sup_{f \in \mathcal{S}_1} \varphi_n(f) \right| \leq \phi_n \left( \frac{\zeta}{(\log n)^{\gamma(\alpha)}} \right), \quad \left| Y_n - \sup_{f \in \mathcal{S}_1} \varphi_n(f) \right| \leq \phi_n \left( \frac{\zeta}{(\log n)^{\gamma(\alpha)}} \right).$$

Il s'agit d'une interpolation entre le cas exigeant  $\varphi_n = \varphi$  uniformément continue et le cas peu exigeant  $\varphi_n = \varphi$   $\alpha$ -Hölder continue, ce dernier rendant incontrôlables les statistiques  $Y_n$  et  $Z_n$ . Ceci reste vrai sans la conjecture de recouvrement si  $\bigcap_{n \geq 0} \arg \inf_{f \in \mathcal{S}_1} \varphi_n(f) \neq \emptyset$  et  $\bigcap_{n \geq 0} \arg \sup_{f \in \mathcal{S}_1} \varphi_n(f) \neq \emptyset$ .

**(iii) Principe de démonstration.** Nous pouvons tirer profit des techniques mises au point pour les Théorèmes 2.4, 2.5 et 2.6 en naviguant entre  $\mathcal{B}_0$ ,  $\mathcal{B}_{\alpha,n}$ ,  $\mathcal{B}_\alpha$  et  $\mathcal{B}_{\alpha+\varepsilon}$  selon les besoins, pour exploiter  $\|f\|_\alpha = \max(\omega_{\alpha,f}(\delta_n), \|f\|_{\alpha,n})$  et les outils disponibles sur  $\mathcal{B}_0$  et  $\mathcal{B}_\alpha$  tels que [MSW,St,BR].  $\square$

## 2.3 – Lois fonctionnelles intermédiaires

Les résultats de 2.1 portent sur les lois fonctionnelles locales,  $T_n = \{t\}$  et surtout globales, avec la recherche des conditions pour que tout se passe comme si  $T_n = [0, 1 - a_n]$ . Se pose alors la question des  $T_n$  quelconques. Comment normaliser les incréments pour en déterminer la forme asymptotique lorsqu'ils sont moins nombreux ? Comment stabiliser des ensembles de trajectoires pour en contrôler les excès, grands et petits ? C'est l'objet de [14,16]. L'importance de ces questions est apparue, au plan technique, en étudiant des estimateurs non paramétriques dans [10,15,21].

Nous évoquons tout d'abord le cas d'une trajectoire unique et fixée, ce qui inverse notre problématique usuelle sur les incréments mais révèle un principe important dans le mécanisme de normalisation. Ensuite nous transposons ce principe pour répondre au problème initialement posé.

### 2.3.1 – Interpolation entre deux résultats de P. Lévy

(i) Soit  $\{W(t) : t \in [0, 1]\}$  un mouvement brownien standard. Pour  $0 < a < 1$  et  $T \subset [0, 1 - a]$  notons

$$\omega_T(a) = \sup_{t \in T} \sup_{t \leq s \leq t+a} |W(t) - W(s)|$$

son module de continuité restreint à  $T$ . Les lois du logarithme, *locale* et *globale*, de [Lé]

$$\limsup_{a \rightarrow 0} \frac{\omega_{\{t\}}(a)}{\sqrt{2a \log |\log a|}} = \lim_{a \rightarrow 0} \frac{\omega_{[0,1-a]}(a)}{\sqrt{2a |\log a|}} = 1 \quad p.s.$$

ont été généralisées sous forme fonctionnelle dans [S] et [A] respectivement. Tout ensemble fini  $T$  peut remplacer  $\{t\}$  et tout intervalle  $T \neq \emptyset$  peut remplacer  $[0, 1 - a]$ . Dans [16] nous laissons varier  $T = T_a$  avec  $a \rightarrow 0$  et décrivons la trajectoire brownienne observée en des petits  $T_a \neq [0, 1 - a]$ .

Que voyons nous en promenant ainsi un zoom en peu de points le long du brownien figé ?

Ceci n'a été exploré que par [Mu] et seulement du point de vue des valeurs d'adhérence. Noter qu'en 2.3 une vitesse de zoom  $A_T/T$  est imposée et la trajectoire de référence  $W_T$  varie, ce qui renouvelle progressivement les incréments. Ici nous avons  $a \rightarrow 0$  et la variété provient uniquement des points  $T_a$  et de l'autosimilarité, sans renouvellement trajectorien, d'où des phénomènes d'une autre nature.

**Motivation.** D'une part, nous souhaitons croiser les deux points de vue,  $a_n$  imposé et  $T_n$  petit dans  $\Theta_n$ , voir 2.3.2. D'autre part, de nombreux modèles probabilistes dépendent d'une seule trajectoire  $W$ . Par exemple, dans une équation différentielle stochastique le comportement localisé de  $W$  peut jouer un rôle dans le contrôle fin du processus solution, via les schémas de discrétisation et d'approximation, faibles ou forts, qui promènent des points le long de  $W$ . D'où l'intérêt de savoir ce qui peut ou doit s'y passer. Une perspective d'application concerne les propriétés trajectoires des diffusions.

(ii) Soit  $\sharp(T, a)$  le plus petit nombre d'intervalles de longueur  $a$  nécessaires au recouvrement de  $T$ . Il est évident que  $T_a$  et  $\sharp(T, a)$  ne doivent pas varier trop brutalement, et l'une de nos principales difficultés consiste à exhiber des conditions suffisantes minimales. Pour un intervalle  $A \subset [0, 1]$  posons

$$T_A^- = \bigcap_{a \in A} T_a \subset T_A^+ = \bigcup_{a \in A} T_a, \quad \sharp_a = \sharp(T_a, a), \quad \sharp_A^- = \sharp(T_A^-, \inf A), \quad \sharp_A^+ = \sharp(T_A^+, \inf A).$$

La stabilité élémentaire exige l'existence d'une suite  $a_n$  et  $A_n = ]a_{n+1}, a_n]$  telle que, lorsque  $n \rightarrow \infty$

$$(H^-) \quad a_n \in ]0, 1[, \quad a_n \downarrow 0, \quad a_n \sim a_{n+1}, \quad \sup_{a \in A_n} \log \sharp_a \sim \log \sharp_{A_n}^- \rightarrow \infty$$

$$(H^+) \quad a_n \in ]0, 1[, \quad a_n \downarrow 0, \quad a_n \sim a_{n+1}, \quad \inf_{a \in A_n} \log \sharp_a \sim \log \sharp_{A_n}^+ \rightarrow \infty.$$

Notons  $(H_\theta^\pm)$  l'existence de  $a_n$  vérifiant  $(H^\pm)$  et (i)  $a_n = O(\exp(-n^\theta))$  ou bien (ii)  $n^{\theta/(1-\theta)} = O(\sharp_{a_n})$ . Notre condition fondamentale de régularité  $(C_1)$  correspond à  $(C_1^-)$  et  $(C_1^+)$  où

$$(C_1^\pm) \quad \text{pour un } \theta_0 < 1 \text{ et tout } \theta \in ]\theta_0, 1[, (H_\theta^\pm) \text{ est vraie.}$$

Concernant la taille de  $T_a$  la condition  $(C_2)$  que  $\log \#_a / \log |\log a| \rightarrow c \in [0, \infty]$  joue un rôle similaire à celui de  $c$  en 2.1. Si  $T_a$  est trop petit  $\mathcal{S}_{c/(1+c)}$  apparaît, ce qui n'était pas attendu et absent de [Mu].

**Remarque.** En effet, nous avons une unique trajectoire  $W$  alors qu'en 2.1.3 le renouvellement de la suite  $W_T$  impose, par la loi de Chung, que  $W_T$  est plat infiniment souvent, donc tous ces grands incréments aussi. Nous pouvions alors penser que les Théorèmes 2.2, 2.5, 2.6 étaient dus à la taille des incréments ou au renouvellement de  $W_T$ . Mais l'apparition de  $\mathcal{S}_{c/(1+c)}$  pour  $W$  fixé prouve que l'unique raison expliquant le comportement des grands incréments est leur trop petit nombre.

Les trop petits ensembles requièrent, dans le contexte du phénomène de [CsR,Sh], la condition  $(C_3)$  que ou bien (i)  $\log \#_a \gg \log \log |\log a|$  ou bien (ii)  $\#_a \gg \log |\log a|$  et  $T_a = \bigcup_{j \leq \#_a} [t_{a,j}, t_{a,j} + a]$ .

(iii) Dans [16] nous étendons [Lé] avec une condition minimale pour chaque borne, non précisée ici.

**Proposition 2.27.** *Si  $\#_a \rightarrow \infty$  alors  $\liminf_{a \rightarrow 0} \omega_{T_a}(a) / \sqrt{2a \log \#_a} \leq 1$  p.s. Sous  $(C_1)$  et  $(C_2)$  il vient*

$$\liminf_{a \rightarrow 0} \frac{\omega_{T_a}(a)}{\sqrt{2a(\log \#_a + \log |\log a|)}} = \sqrt{\frac{c}{1+c}} \leq \limsup_{a \rightarrow 0} \frac{\omega_{T_a}(a)}{\sqrt{2a(\log \#_a + \log |\log a|)}} = 1 \quad p.s.$$

Sous  $(C_2)$  avec  $c = 0$ ,  $(C_3)$  et un  $(H_\theta^\pm)$  avec  $\#_{a_n} \sim \#_{A_n}^\pm$  il vient

$$\liminf_{a \rightarrow 0} \frac{\omega_{T_a}(a)}{\sqrt{2a(\log \#_a - \log \log |\log a|)}} = 1 \quad p.s.$$

Si  $c > 0$  la première liminf vaut 1. Sous  $c = 0$  et  $(C_3)(i)$  elle peut prendre toutes les valeurs de  $[0, 1]$ .

(iv) La version fonctionnelle de la Proposition 2.27 montre qu'avec probabilité un, la collection

$$\mathcal{I}_a = \left\{ \frac{\Delta W(a, t)}{\sqrt{2aL_a}} : t \in T_a \right\}, \quad \Delta W(a, t) = W(t + aI) - W(t), \quad L_a = \log \#_a + \log |\log a|$$

est relativement compacte dans  $\mathcal{B}$ , d'ensembles limites  $\mathcal{S}_{c/(1+c)}$  et  $\mathcal{S}_1$ .

**Théorème 2.28.** *Sous  $(C_1)$  et  $(C_2)$  nous avons, pour tout  $\varepsilon > 0$*

$$\mathbb{P} \left( \bigcup_{0 < b < 1} \bigcap_{0 < a < b} \{\mathcal{I}_a \subset \mathcal{S}_1 + \varepsilon B_0\} \right) = 1$$

et  $\mathbb{P}(\mathcal{I}_a \subset \mathcal{S}_{c/(1+c)} + \varepsilon B_0) \rightarrow 1$  si  $a \rightarrow 0$ , alors que tout  $f \in \mathcal{S}_1$  vérifie

$$\mathbb{P} \left( \bigcap_{0 < b < 1} \bigcup_{0 < a < b} \{f \in \mathcal{I}_a + \varepsilon B_0\} \right) = 1.$$

Si de plus  $\#_a \rightarrow \infty$ ,  $\mathcal{S}_{c/(1+c)}$  est le plus grand sous-ensemble de  $\mathcal{B}$  tel que

$$\mathbb{P} \left( \bigcup_{0 < b < 1} \bigcap_{0 < a < b} \{\mathcal{S}_{c/(1+c)} \subset \mathcal{I}_a + \varepsilon B_0\} \right) = 1.$$

Par conséquent,  $d(f, \mathcal{S}_1) = \liminf_{a \rightarrow 0} d(f, \mathcal{I}_a) \leq \limsup_{a \rightarrow 0} d(f, \mathcal{I}_a) = d(f, \mathcal{S}_{c/(1+c)})$  p.s. si  $f \in \mathcal{B}$ . Lorsque  $c = \infty$ , le recouvrement est complet,  $d_H(\mathcal{I}_a, \mathcal{S}_1) \rightarrow 0$  p.s. et si  $c < \infty$  il y a une concentration récurrente,  $\liminf_{a \rightarrow 0} d_H(\mathcal{I}_a, \mathcal{S}_{c/(1+c)}) = 0$  p.s. avec  $d_H$  la distance de Hausdorff.

(v) **Critères.** Si  $T_a$  est monotone pour l'inclusion quand  $a \rightarrow 0$  alors  $(H^\pm)$  est réduite à  $a_n \sim a_{n+1}$  et  $\log \#_{a_n} \sim \log \#_{a_{n+1}}$ . Si  $T_a$  est monotone et  $\log \#_a$  à variation lente quand  $a \rightarrow 0$  alors  $(C_1)$  est vraie. La condition  $(C_1^\pm)$  est impliquée par l'existence de  $a_n$  vérifiant  $(H^\pm)$  et  $\log \#_{a_n} \gg \log n$ . Sous  $(C_2)$  avec  $c = \infty$ , l'existence de  $a_n < \exp(-n^\theta)$  vérifiant  $(H^\pm)$  pour un  $\theta > 0$  implique  $(C_1^\pm)$ . Sous  $(C_2)$  et  $c = 0$ , l'existence de  $a_n$  vérifiant  $(H^\pm)$  et  $\#_{a_n} > n^{1/\theta}$  pour un  $\theta > 0$  implique  $(C_1^\pm)$ . Sous  $(C_2)$  et  $(H_\theta^\pm)$  avec  $1 < c < \infty$  et  $1/c < \theta < 1$ ,  $(C_1^\pm)$  est vraie.

Les exemples typiques de tels  $T_a$  sont des intervalles mobiles  $[t_a, t_a + a\sharp_a]$  de longueur  $a\sharp_a \rightarrow 0$ .

**(vi) Principe de démonstration.** Concernant le module de continuité restreint il est facile d'établir que si  $0 < a < 1$ ,  $T \subset [0, 1]$  et  $\delta \in (0, 1)$  nous avons, pour tout  $x > 0$ ,

$$\mathbb{P} \left( \frac{\omega_T(a)}{\sqrt{a}} \geq x \right) \leq \frac{64}{\delta^2} \exp \left( -\frac{1}{2} (1 - \delta)^2 x^2 + \log \sharp(T, a) - \log x \right).$$

Combinée aux propriétés de déviation de  $W$  via des techniques de blocage d'événements indicés par  $a \in A_n$  les exigences sur  $a_n$  apparaissent naturellement, notamment (i) et (ii) dans  $(H_\theta^\pm)$ . La difficulté nouvelle est que les parties de  $W$  visitées par  $T_{a_n}$  et  $\bigcup_{k>n} T_{a_k}$  empêchent d'appliquer Borel-Cantelli à des événements indépendants, même après le blocage. Nous devons alors contrôler les  $\Delta W(a, t)$  présents amputés des  $\Delta W(b, t)$  futurs en veillant à ce que cela induise le résultat escompté. Cette élimination du futur alourdit drastiquement le cas des petits ensembles  $\mathcal{I}_a$ ,  $c < \infty$ .  $\square$

### 2.3.2 – Quelques vitesses browniennes intermédiaires

**(i)** Dans [16] le contrôle des vitesses demande plus de régularité à  $\sharp_{A_n}^\pm$ . Nous remplaçons  $(C_1)$  par l'exigence  $(C_4)$ , pour une fonction  $\nabla$  à préciser, qu'il existe  $a_n = \exp(-n/\psi(n))$  telle que  $a_n \downarrow 0$  et

$$\limsup_{n \rightarrow \infty} \frac{L_{a_n} \log \psi(n)}{\nabla^2(L_{a_n})} < \gamma, \quad \left( 1 - \frac{\log \sharp_{A_n}^-}{\log \sharp_{A_n}^+} \right) \nabla(L_{a_n}) \rightarrow 0, \quad \left( 1 - \frac{a_{n-1}}{a_n} \right) \nabla^2(L_{a_n}) \rightarrow 0.$$

Nous utilisons les fonctions  $\nabla = \nabla_f$  de 3.1.3. Comme  $\eta_1 I^{2/3} < \nabla < \eta_2 I$  la suite régularisante  $a_n$  doit être sous-géométrique mais pas trop lente. Les situations typique utilisent  $\psi(n) = (\log n)^\delta$ ,  $\delta > 0$  ou  $\psi(n) = n^{1-\theta}$ ,  $\theta \in ]0, 1[$ . Étant donné  $T_a$  les contraintes explicites de  $(C_4)$  sont donc faciles à traduire.

**Exemple.** Si  $T_a = [0, a \exp(|\log a|^\beta)]$  avec  $0 < \beta \leq 1$  il vient  $L_a \sim |\log a|^\beta$  et  $(H_\theta^\pm)$ ,  $(C_1)$ ,  $(C_2)$ ,  $c = \infty$  sont vérifiées, ainsi que  $(C_4)$  avec  $\psi(n) = (\log n)^\delta$  et  $\delta$  assez grand.

**(ii)** En zoomant sur  $W$  en suffisamment de positions  $T_a$  et en normalisant par  $\sqrt{2a \log \sharp_a}$  toute forme limite  $f \in \mathcal{S}_1^*$  est visitée à la vitesse  $\nabla_f(\log \sharp_a)$ . De nouveau le hasard s'ordonne de façon prévisible.

**Théorème 2.29.** Soit  $f \in \mathcal{S}_1^*$  et  $L_a = \log \sharp_a$ . Sous  $(C_2)$ ,  $c = \infty$  et  $(C_4)$  avec  $\nabla = \nabla_f$  nous avons

$$\lim_{a \rightarrow 0} \nabla_f(\log \sharp_a) d(f, \mathcal{I}_a) = \chi_f(1) \quad p.s.$$

Si de plus  $\nabla_f(\log \sharp_a) \gg \sqrt{(\log \sharp_a) \log |\log a|}$  alors ceci reste vrai avec  $L_a = \log \sharp_a + \log |\log a|$  dans  $\mathcal{I}_a$ .

En  $f = 0$  nous obtenons le module de non-différentiabilité restreint exact de  $W$ ,

$$\lim_{a \rightarrow 0} \sqrt{\frac{\log \sharp_a}{a}} \inf_{t \in T_a} \sup_{|t-s| \leq a} |W(t) - W(s)| = \frac{\pi}{\sqrt{8}} \quad p.s.$$

qui étend le cas  $T_a = [0, 1 - a]$  de [CR2]. En comparant à la Proposition 2.27 il vient

$$\inf_{t \in T_a} \sup_{|t-s| \leq a} |W(t) - W(s)| \sim \frac{\pi}{4 \log \sharp_a} \sup_{t \in T_a} \sup_{|t-s| \leq a} |W(t) - W(s)| \quad p.s.$$

**(iii)** La vitesse de concentration des observations trajectorielles de [16] reprend les constantes de 2.1.1.

**Théorème 2.30.** Sous  $(C_2)$ ,  $c = \infty$  et  $(C_4)$  avec  $\nabla = I^{2/3}$  alors, si  $L_a = \log \sharp_a$  et  $\sigma > \sigma_1$ ,

$$\mathbb{P} \left( \bigcup_{0 < b < 1} \bigcap_{0 < a < b} \left\{ \mathcal{I}_a \subset \mathcal{S}_1 + \frac{\sigma}{(\log \sharp_a)^{2/3}} B_0 \right\} \right) = 1$$

et, si  $\sigma < \sigma_0$ ,

$$\mathbb{P} \left( \bigcap_{0 < b < 1} \bigcup_{0 < a < b} \left\{ \mathcal{I}_a \not\subset \mathcal{S}_1 + \frac{\sigma}{(\log \sharp_a)^{2/3}} B_0 \right\} \right) = 1.$$

Enfin, sous les hypothèses  $(C_2)$ ,  $c = \infty$  et  $(C_4)$  avec  $\nabla = I^{2/3}$  et  $\nabla = I$  nous obtenons la vitesse de recouvrement complet, pour  $\sigma > \chi_I = \pi^{2/3}/2^{5/3}$ ,

$$\mathbb{P} \left( \bigcup_{0 < b < 1} \bigcap_{0 < a < b} \left\{ \mathcal{S}_1 \subset \mathcal{I}_a + \frac{\sigma}{(\log \#_a)^{2/3}} B_0 \right\} \right) = 1.$$

Nous en déduisons l'ordre de grandeur asymptotique de la distance de Hausdorff,

$$\sigma_0 \leq \liminf_{a \rightarrow 0} L_a^{2/3} d_H(\mathcal{I}_a, \mathcal{S}_1) \leq \limsup_{a \rightarrow 0} L_a^{2/3} d_H(\mathcal{I}_a, \mathcal{S}_1) \leq \sigma_1 \vee \chi_I \quad p.s.$$

(iv) Pour les petits ensembles nous exigeons que  $c_a = \log \#_a / \log |\log a| \rightarrow c$  assez vite. Les situations se diversifient mais le principal enseignement est que la rareté stabilisée réduit les limites à  $\mathcal{S}_{c/(1+c)}$ .

**Théorème 2.31.** Soit  $f \in \mathcal{S}_1^*$  et  $L_a = \log \#_a + \log |\log a|$ . Sous  $(C_2)$ ,  $c < \infty$  et  $(C_4)$  avec  $\nabla = \nabla_f$

$$\liminf_{a \rightarrow 0} \nabla_f(L_a) d(f, \mathcal{I}_a) = \chi_f(1) \quad p.s.$$

et, si de plus  $f \in \mathcal{S}_{c/(1+c)}^*$  et  $\sqrt{(c_a - c)^\pm} \ll \nabla_f(\log |\log a|) / \log |\log a|$ ,

$$\limsup_{a \rightarrow 0} \nabla_f(L_a) d(f, \mathcal{I}_a) = \chi_f \left( \frac{c}{1+c} \right) \quad p.s.$$

Les hypothèses précédentes entraînent aussi la concentration récurrente

$$\mathbb{P} \left( \bigcap_{0 < b < 1} \bigcup_{0 < a < b} \left\{ \mathcal{I}_a \subset \mathcal{S}_{c/(1+c)} + \frac{\sigma}{(\log \#_a + \log |\log a|)^{2/3}} B_0 \right\} \right) = 1.$$

(v) **Principe de démonstration.** Les ingrédients brièvement évoqués en section 2.1 restent les mêmes et des techniques de blocage sont rendues possibles par les hypothèses sur  $T_a$  mais avec une difficulté supplémentaire lourde, induite par le non renouvellement de  $W$ . En effet nous travaillons les chaînes d'événements par des coupes dans  $W$  pour éliminer l'influence des blocs d'incrément futurs, localisés en  $T_b$ ,  $b > a$ , et appliquons nos outils à des trajectoires browniennes reconstruites.  $\square$

### 2.3.3 – Processus empirique réel

Dans [14] les résultats de 2.1 pour  $\Theta_n$  sont étendus aux situations où  $T_n \subset [0, 1 - a_n]$  est petit. Les principes ci-avant d'entropie et de normalisation, qui résultent de [16], règlent le conflit entre événements rares et nombre de tentatives garanti. Nous les adaptons à des contraintes différentes.

(i) Pour une suite d'entiers  $n_k$  posons  $N_k = \{n_k + 1, \dots, n_{k+1}\}$  et  $T_k^- = \bigcap_{n \in N_k} T_n \subset T_k^+ = \bigcup_{n \in N_k} T_n$ . Les nombres d'entropie de 2.1.2 s'écrivent

$$\kappa_n = \#(T_n, a_n), \quad \tau_k^- = \#(T_k^-, a_{n_k}), \quad \tau_k^+ = \#(T_k^+, a_{n_{k+1}})$$

avec  $\#(T, a) = \inf\{m : T \subset \bigcup_{1 \leq j \leq m} [t_j, t_j + a[$  pour  $t_j \in T$ ,  $t_{j+1} \geq t_j + a\}$ . Désignons par  $(H^\pm)$  les hypothèses que  $\log \kappa_{n_k} \sim \log \kappa_{n_{k+1}} \sim \log \tau_k^\pm$  quand  $k \rightarrow \infty$  et  $(H)$  les deux. Lorsque  $(H^\pm)$  est vérifiée pour une suite telle que  $n_{k+1} \sim n_k$  et  $\log \tau_k^\pm \gg \log k$  nous écrivons  $(H_G^\pm)$  pour *grand ensemble*. Lorsque  $(H^\pm)$  est vérifiée pour chaque  $m$  et  $n_k = \varphi(k, m)$  tel que

$$\lim_{m \rightarrow \infty} \lim_{k \rightarrow \infty} \frac{\varphi(k+1, m)}{\varphi(k, m)} = 1, \quad \lim_{m \rightarrow \infty} \liminf_{k \rightarrow \infty} \frac{\log \log \varphi(k, m)}{\log k} = 1$$

nous écrivons  $(H_P^\pm)$  pour *petit ensemble*. Nous avons des critères et des formulations alternatives.

(ii) La première étape de [14] consiste à étendre [St] au module restreint  $\omega_n(a, T) = \sup_{t \in T} \|\Delta \alpha_n(a, t)\|$ . Le phénomène de 2.3.1 réapparaît de nouveau pour les petits ensembles,  $c < \infty$ , et chacune des bornes, inférieure ou supérieure, demande en fait une seule hypothèse minimale sur  $T_n$ .

**Proposition 2.32.** Soit  $L_n = \log \kappa_n + \log \log n$ ,  $\kappa_n \rightarrow \infty$ ,  $a_n \downarrow 0$ ,  $na_n \uparrow \infty$ ,  $na_n \log \kappa_n \ll (\log n)^2$  et  $(\log \kappa_n)/\log \log n \rightarrow c$ . Alors

$$\liminf_{n \rightarrow \infty} \frac{\omega_n(a_n, T_n)}{\sqrt{2a_n \log \kappa_n}} \leq 1 \leq \limsup_{n \rightarrow \infty} \frac{\omega_n(a_n, T_n)}{\sqrt{2a_n L_n}} \quad p.s.$$

et, sous  $(H_G^+)$  ou  $(H_P^+)$ ,

$$\limsup_{n \rightarrow \infty} \frac{\omega_n(a_n, T_n)}{\sqrt{2a_n L_n}} \leq 1 \quad p.s.$$

alors que sous  $(H_G^-)$  ou  $(H_P^-)$ ,

$$\liminf_{n \rightarrow \infty} \frac{\omega_n(a_n, T_n)}{\sqrt{2a_n L_n}} = \sqrt{\frac{c}{1+c}} \quad p.s.$$

**Exemple.** Le cas  $T_n = [0, t_n]$ ,  $na_n \uparrow \infty$ ,  $t_n/a_n \uparrow \infty$ ,  $na_n/\log \log n \rightarrow \infty$  et  $|\log(t_n/a_n)|/\log \log n \rightarrow c$  restitue [Ho]. Il est clair que la taille  $a_n$  ne gouverne plus seule le comportement du module restreint. Ainsi la frontière du domaine d'invariance bouge et tous les  $a_n \gg (\log n)/n$  ne sont pas autorisés. Lorsque  $c = 0$  et  $\kappa_n/\log |\log a_n| \rightarrow \infty$ , et si  $\kappa_n$  est aussi la mesure de boîte de  $T_n$ , le résultat devient

$$\liminf_{n \rightarrow \infty} \frac{\omega_n(a_n, T_n)}{\sqrt{2a_n(\log \kappa_n - \log \log |\log a_n|)}} = 1 \quad p.s.$$

**Principe de démonstration.** En adaptant la borne de déviation du module de  $\alpha_n$  [MSW,SW] il vient, en posant  $\Psi(x) = 2((1+x)\log(1+x) - x)/x^2$ , pour tout  $n$ ,  $a < \gamma \leq 1/2$  et tout  $x > 0$ ,

$$\mathbb{P} \left( \sup_{t \in T} \sup_{t \leq s \leq t+a} \frac{|\alpha_n(s) - \alpha_n(t)|}{\sqrt{a}} \geq x \right) \leq \frac{20}{\gamma^3} \exp \left( -\frac{x^2}{2} (1-\gamma)^4 \Psi \left( \frac{x}{\sqrt{na}} \right) + \log \#(T, a) \right).$$

Nous appliquons le lemme de Borel-Cantelli à des suites d'événements par blocs en surveillant les fluctuations de  $T_n$  et en privant les incréments futurs des points de l'échantillon du bloc présent, par renouvellement trajectorien. Les contraintes induites varient selon  $a_n$  et imposent la stabilité de  $T_n$ .  $\square$

(iii) Rappelons que  $d_H(A, B) = \max(\sup_{f \in A} d(f, B), \sup_{g \in B} d(g, A))$  désigne la distance de Hausdorff. Bien que la vitesse [KMT] d'approximation gaussienne suggère la condition minimale  $na_n/\log \kappa_n \rightarrow \infty$  nous utilisons dans [14] les hypothèses

$$a_n \downarrow 0, \quad na_n \uparrow \infty, \quad \frac{\log \kappa_n}{\log \log n} \rightarrow c, \quad \frac{na_n \log \kappa_n}{(\log n)^2} \rightarrow \infty \quad (4)$$

Nous généralisons alors le recouvrement complet de [DM,D2] rappelé au Théorème 2A,  $c = \infty$ . Cette loi intermédiaire concerne les ensembles moyens d'incrémentes moyens ou petits.

**Théorème 2.33.** Posons  $L_n = \log \kappa_n$  dans  $\Theta_n$  et choisissons  $a_n$  et  $T_n$  vérifiant (4) avec  $c = \infty$  et  $(H_P)$  ou  $(H_G)$ . Alors, avec probabilité un,  $d_H(\Theta_n, \mathcal{S}_1) \rightarrow 0$ .

(iv) Une seconde version fonctionnelle de la Proposition 2.32 est la loi intermédiaire pour les petits ensembles d'incrémentes, eux-mêmes grands, moyens ou petits. Nous omettons quelques assertions.

**Théorème 2.34.** Soit  $L_n = \log \kappa_n + \log \log n$  dans  $\Theta_n$  et  $(a_n, T_n)$  vérifiant (4) avec  $c < \infty$  et  $(H_P)$  ou  $(H_G)$ . Alors  $\mathcal{S}_1$  et  $\mathcal{S}_{c/(1+c)}$  sont le plus petit et le plus grand compacts de  $\mathcal{B}$  tels que pour tout  $\varepsilon > 0$ ,  $\mathbb{P}(\exists n_0 : \mathcal{S}_{c/(1+c)} \subset \Theta_n + \varepsilon B_0 \subset \mathcal{S}_1 + 2\varepsilon B_0, n > n_0) = 1$  et  $\mathbb{P}(\mathcal{S}_{c/(1+c)} \subset \Theta_n + \varepsilon B_0) \rightarrow 1$ .

Cette extension du Théorème 2A admet des versions pour les processus de 2.1.3 et 2.1.4.

**Application.** Soit  $T_n = [t - t_n + \varepsilon_n^-, t + t_n + \varepsilon_n^+]$ ,  $t \in ]0, 1[$ ,  $na_n \uparrow \infty$ ,  $t_n/a_n \uparrow \infty$ ,  $na_n/\log \log n \rightarrow \infty$  et  $|\log(t_n/a_n)|/\log \log n \rightarrow c$  avec  $\varepsilon_n^\pm$  aléatoires mais tels que  $\limsup_{n \rightarrow \infty} |\varepsilon_n^\pm|/t_n < 1$  p.s. Considérons une statistique basée sur les  $U_i \in T_n$ , par exemple  $Y_n = \sup_{g \in \Theta_n} \varphi(g)$  avec  $\varphi$  continue de  $\mathcal{B}$  dans  $\mathbb{R}$ . Alors  $\liminf_{n \rightarrow \infty} Y_n = \sup_{\mathcal{S}_{c/(1+c)}} \varphi \leq \limsup_{n \rightarrow \infty} Y_n = \sup_{\mathcal{S}_1} \varphi$  p.s., en vertu de 2.33 et 2.34.



La version du Théorème 2.34 pour les processus de Kiefer  $K(n, t)$  et de quantile  $\beta_n$  sert en 4.2 et 4.3.

(v) Nous produisons également dans [14] des conditions suffisantes minimales sur  $a_n$  et  $T_n$ , outre  $(H_P)$ ,  $(H_G)$  et (4), garantissant les vitesses du type suivant. Sous (4) avec  $c = \infty$ , en normalisant par  $L_n = \log \kappa_n$  ou  $L_n = \log \kappa_n + \log \log n$  il vient, si  $\sigma$  est assez grand,

$$\mathbb{P} \left( \exists n_0 : \Theta_n \subset \mathcal{S}_1 + \frac{\sigma}{(\log \kappa_n)^{2/3}} B_0, n > n_0 \right) = 1.$$

et, pour  $f \in \mathcal{S}_1^*$ ,

$$\lim_{n \rightarrow \infty} \nabla_f(\log \kappa_n) d(f, \Theta_n) = \chi_f(1) \quad p.s.$$

Par conséquent, le module de non-différentiabilité restreint  $\inf_{t \in T_n} \sup_{|t-s| \leq a_n} |\alpha_n(t) - \alpha_n(s)| / a_n$  de  $\alpha_n$  équivaut presque sûrement à  $\pi / \sqrt{8a_n \log \kappa_n}$  et, en comparant à la Proposition 2.32,

$$\frac{\sup_{t \in T_n} \sup_{|t-s| \leq a_n} |\alpha_n(t) - \alpha_n(s)|}{\inf_{t \in T_n} \sup_{|t-s| \leq a_n} |\alpha_n(t) - \alpha_n(s)|} \sim \frac{4}{\pi} \log \kappa_n \quad p.s.$$

Au contraire, si  $c < \infty$  l'entropie influence la normalisation  $L_n = \log \kappa_n + \log \log n$  et les vitesses, mais uniquement au niveau des constantes puisque

$$\liminf_{n \rightarrow \infty} \nabla_f(\log \log n) d(f, \Theta_n) = \chi_f(1) \quad p.s.$$

$$\limsup_{n \rightarrow \infty} \nabla_f(\log \log n) d(f, \Theta_n) = \chi_f \left( \frac{c}{1+c} \right) \quad p.s.$$

Pour  $T_n$  très petit nous retrouvons les résultats locaux avec vitesses, qui renforcent [B,EM,M,M2].

**Exemple.** Soit  $T_n = [t_n, t_n + a_n \exp((\log n)^{2\beta})]$  avec  $\beta < 1/2$  et  $t_n$  ne bougeant pas trop vite. Si  $a_n = n^{-\gamma}$ ,  $\gamma < 1$ ,  $\alpha \in ]1/2, 1[$  et  $c_\alpha$  est la constante de l'exemple suivant le Corollaire 3.4, alors pour la fonction cible  $f_\alpha = (\sqrt{2\alpha - 1}/\alpha) I^\alpha \in \mathcal{S}_1^{vli}$  nous obtenons la normalisation et la vitesse atypiques

$$\lim_{n \rightarrow \infty} (\log n)^{4\alpha\beta/(4\alpha-1)} \inf_{t \in T_n} \left\| \frac{\Delta \alpha_n(a_n, t)}{\sqrt{2a_n} (\log n)^\beta} \pm f_\alpha \right\| = c_\alpha \quad p.s.$$

La statistique  $Y_n = \sup_{g \in \Theta_n} \varphi_\alpha(g)$  est fonction des fluctuations d'échantillonnage dans ce  $T_n$  et, si  $\varphi_\alpha$  est assez régulière et  $f_\alpha = \arg \min_{\mathcal{S}_1} \varphi_\alpha$  unique, il s'ensuit  $|Y_n - \varphi_\alpha(f_\alpha)| = O_{p.s.}(1/(\log n)^{4\alpha\beta/(4\alpha-1)})$ .

### 2.3.4 – Processus empirique multivarié

(i) Soit  $X_n$  une suite i.i.d. de vecteurs aléatoires dans  $\mathbb{R}^p$  de densité  $f_X$  par rapport à la mesure de Lebesgue  $\lambda_p$ . Dans [M3] le Théorème 2A est généralisé au cas multivarié en indexant les incréments locaux par des fonctions. Soit  $\mathcal{F}$  une famille de fonctions boréliennes de  $\mathbb{R}^p$  dans  $\mathbb{R}$ , uniformément bornées par  $M > 0$ , ayant leur support inclus dans le cube unité centré  $\Lambda^p = [-1/2, 1/2]^p$ . L'incrément local  $\Delta \alpha_n(a_n, x)$  de volume  $a_n$ , observé en  $x$  et indexé par  $\mathcal{F}$  est la fonction aléatoire

$$\Delta \alpha_n(a_n, x)(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left( f \left( \frac{X_i - x}{a_n^{1/p}} \right) - \mathbb{E} f \left( \frac{X_i - x}{a_n^{1/p}} \right) \right)$$

qui mesure sur la classe  $\mathcal{F}$  les déviations de la densité empirique locale. Pour  $p = 1$  et  $X_i = U_i$  de loi uniforme sur  $[0, 1]$  nous avons  $\Delta \alpha_n(a_n, t - a_n/2)(\mathbb{I}_{[0,1]}) = \Delta \alpha_n(a_n, t)$  en notation 2.1. Supposons que  $\inf_{V+\varepsilon B_p} f_X > 0$  avec  $V \subset \mathbb{R}^p$  et  $B_p$  la boule unité fermée de  $\mathbb{R}^p$ . Pour  $V_n \subset V$  posons

$$\Theta_n(\mathcal{F}) = \left\{ \frac{\Delta \alpha_n(a_n, x)(\cdot)}{\sqrt{2f_X(x)a_n L_n}} : x \in V_n \right\}.$$

Soit  $\mathcal{H}$  le sous-espace de Hilbert de  $\mathbb{L}_2(\Lambda^p)$  engendré par  $\mathcal{F}$  et

$$\mathcal{S}_{p,\theta} = \left\{ \varphi_h : h \in \mathcal{H}, \int_{\Lambda^p} h^2 d\lambda_p \leq \theta \right\}$$

les boules de l'espace auto-reproduisant  $\mathcal{S}$  de  $\mathcal{H}$ . Chaque  $h \in \mathcal{H}$  est donc associé à  $\varphi_h \in \mathbb{L}_\infty(\mathcal{F})$  définie par  $\varphi_h(f) = \int_{\Lambda^p} fhd\lambda_p$  et  $\langle \varphi_h, \varphi_g \rangle = \int_{\Lambda^p} gh d\lambda_p$ . Soit  $B_\infty = \{\varphi : \sup_{f \in \mathcal{F}} |\varphi(f)| < 1\}$ .

(ii) Si  $V_n = \{x\}$  la loi fonctionnelle locale généralisant [M] est obtenue dans [EM,EM2] avec un théorème limite central et un résultat d'approximation forte locale.

**Théorème 2E.** *Si  $V_n = \{x\}$ ,  $L_n = \log \log n$  et  $a_n$  vérifie les hypothèses du Théorème 2B, la suite  $\Theta_n(\mathcal{F})$  admet presque sûrement  $\mathcal{S}_{p,1}$  comme ensemble de points adhérents dans  $\mathbb{L}_\infty(\mathcal{F})$ .*

La loi globale de [M3] généralise le Théorème 2A pris en  $c = \infty$ , et s'énonce ainsi.

**Théorème 2F.** *Si  $V_n = V$  est compact,  $\inf_{V+\varepsilon B_p} f_X > 0$ ,  $L_n = |\log a_n| + \log \log n$  et  $a_n$  vérifie (1) avec  $c = d = \infty$ , les ensembles  $\Theta_n(\mathcal{F})$  recouvrent complètement  $\mathcal{S}_{p,1}$  dans  $\mathbb{L}_\infty(\mathcal{F})$  avec probabilité un.*

(iii) Comme en 2.3.3 nous interpolons entre 2E et 2F à l'aide des ingrédients [Ar,GG,GMZ,T2,Z] et de leur alchimie inventée dans [M3], couplée à la stratégie de contrôle des localisations futures de [14]. Pour cela nous considérons des petits ensembles  $V_n \subset V$  et leurs nombres d'entropie cubiques. Pour toute suite de blocs  $N_k = \{n_k + 1, \dots, n_{k+1}\}$  introduisons  $V_k^- = \bigcap_{n \in N_k} V_n \subset V_k^+ = \bigcup_{n \in N_k} V_n$  et

$$\kappa_n = \sharp_p(V_n, a_n), \quad \tau_k^- = \sharp_p(V_k^-, a_{n_k}), \quad \tau_k^+ = \sharp_p(V_k^+, a_{n_{k+1}})$$

avec  $\sharp_p(V, a) = \inf\{m : V \subset \bigcup_{1 \leq j \leq m} (t_j + a^{1/p} \Lambda^p)\}$ . La preuve change de nature, et s'affranchit de tout recours aux processus gaussiens, mais au final les conditions suffisantes sont préservées.

**Théorème 2.35.** *Soit  $L_n = \log \kappa_n + \log \log n$  dans  $\Theta_n(\mathcal{F})$  et  $(a_n, V_n)$  vérifiant (4) et  $(H_P)$  ou  $(H_G)$ . Alors  $\mathcal{S}_{p,1}$  et  $\mathcal{S}_{p,c/(1+c)}$  sont le plus petit et le plus grand compacts de  $\mathbb{L}_\infty(\mathcal{F})$  tels que pour tout  $\varepsilon > 0$ ,  $\mathbb{P}(\exists n_0 : \mathcal{S}_{p,c/(1+c)} \subset \Theta_n + \varepsilon B_\infty \subset \mathcal{S}_{p,1} + 2\varepsilon B_\infty, n > n_0) = 1$  et  $\mathbb{P}(\mathcal{S}_{p,c/(1+c)} \subset \Theta_n + \varepsilon B_\infty) \rightarrow 1$ .*

Cette loi intermédiaire fournit aussi la version de la loi globale [M3] pour les grands incréments, en  $\kappa_n = 1/a_n$  et  $a_n$  de l'ordre de  $1/(\log n)^c$ . L'article associé à 2.3.4 est en préparation. La recherche de nouveaux outils donnant accès à des vitesses fait l'objet d'une thèse sous notre direction (3.5.1(ii)).

(iii) L'auteur de [M3] a aussi transposé le Théorème 2F dans le cadre des processus gaussiens en interchangeants les outils de théorie des processus empiriques et ceux d'analyse gaussienne. Il en va donc de même pour le Théorème 2.35 et dans ce cadre, les principes d'invariance 3.4 sont très utiles.

## 2.4 – Autres axes de recherche

### 2.4.1 – Incréments retardés

Dans [CKL,HR,LL] des incréments browniens d'un type différent de [CR] et 2.1.3 sont considérés, ainsi que les incréments dits retardés, de la marche aléatoire. Nos techniques de 2.1 nous permettent dans [19] de produire des versions fonctionnelles et des vitesses de recouvrement dans ce cadre, puisqu'il s'agit encore d'ensembles de trajectoires browniennes, et nous l'étendons à  $\alpha_n$  dans deux directions.

(i) D'une part, nous décrivons la convergence presque sûre d'ensembles du type, pour  $W$ ,

$$\Lambda_{1,T} = \left\{ \frac{\Delta W(A, t-A)}{\sqrt{2AL_{T,A}}} : A \leq T, t \in [A, T] \right\}, \quad \Lambda_{2,T} = \left\{ \frac{\Delta W(A, T-A)}{\sqrt{2AL_{T,A}}} : A \leq T \right\}$$

avec  $L_{T,A} = \log(T/A) + \log |\log A|$  et du type, pour le processus empirique uniforme  $\alpha_n$ ,

$$\Theta_{1,n} = \left\{ \frac{\Delta \alpha_n(a, t)}{\sqrt{2a(|\log a| + \log |\log na|)}} : a_n \leq a \leq 1, t \leq 1-a \right\}.$$

(ii) D'autre part, nous explorons dans [19] une situation de retard utile pour les applications en statistique adaptative. Il s'agit d'étudier le comportement joint de  $\alpha_n$  et  $\alpha_{m_n}$  avec  $m_n < n$  un décalage de temps, en décrivant la convergence des ensembles de couples de processus du type

$$\Theta_{2,n} = \left\{ \left( \frac{\Delta\alpha_n(a_n^+, t)}{\sqrt{2a_n^+ L_n^+}}, \frac{\Delta\alpha_{m_n}(a_n^-, t)}{\sqrt{2a_n^- L_n^-}} \right) : t \in [0, 1 - a_n^+] \right\}$$

où  $a_n^- \ll a_n^+$ ,  $L_n^+ = |\log a_n^+| + \log \log n$  et  $L_n^- = |\log a_n^-| + \log \log m_n$ .

### 2.4.2 – Ensembles fractals aléatoires

Les dimensions fractales presque sûres d'ensembles de points exceptionnels, aléatoires, engendrés par les trajectoires de  $W_T$  et  $\alpha_n$  sont calculées dans [OT,DM2] respectivement. Dans [24] nous avons entrepris de faire de même en considérant les points  $t \in [0, 1]$  tels qu'infiniment souvent (i.s.)

$$\left\| \frac{\Delta\alpha_n(a_n, t)}{\sqrt{2a_n L_n}} - f \right\| \leq \xi_n$$

avec, typiquement,  $\xi_n = (1 + \varepsilon)\varepsilon_n$ . Il s'agit donc d'évaluer la dimension fractale de tels ensembles de points, disons-les chanceux, et de montrer qu'ils sont partout denses dans  $[0, 1]$ . Leur dimension de Hausdorff  $\dim_H$ , ou celle de Hausdorff-Besicovich lorsque  $\dim_H = 0$ , est fonction de la vitesse  $\xi_n \rightarrow 0$ , de  $c$  dans (1), de  $J(f)$  et de  $\nabla_f$  donc du type de la fonction de Strassen  $f$  considérée. Nos preuves deviennent d'autant plus techniques qu'elles entremêlent les longues démonstrations de [9,DM2].

Par exemple, le long d'une grille  $T_n = \{t_{n,k} = k\delta_n : k < 1/\delta_n\}$  avec  $\delta_n > a_n$  posons  $t_n(t) = t_{n,k} = k\delta_n$  le voisin de  $t \in [t_{n,k}, t_{n,k+1})$  et considérons pour  $f \in \mathcal{S}_1^-$  et  $\varepsilon > 0$  fixés les ensembles aléatoires

$$\begin{aligned} L(f) &= \left\{ t \in [0, 1] : \liminf_{n \rightarrow \infty} \left\| \frac{\Delta\alpha_n(a_n, t_n(t))}{\sqrt{2a_n |\log \delta_n|}} - f \right\| = 0 \right\} \\ L_\varepsilon(f) &= \left\{ t \in [0, 1] : \liminf_{n \rightarrow \infty} \left\| \frac{\Delta\alpha_n(a_n, t_n(t))}{\sqrt{2a_n |\log \delta_n|}} - f \right\| |\log \delta_n| \leq \frac{\pi(1 + \varepsilon)}{4\sqrt{1 - J(f)}} \right\} \subset L(f). \end{aligned}$$

**Théorème 2.36.** *Si  $\delta_n \vee a_n \downarrow 0$ ,  $na_n \uparrow \infty$ ,  $|\log \delta_n| / \log \log n \rightarrow \infty$  et  $na_n |\log \delta_n| / (\log n)^4 \rightarrow \infty$  alors, avec probabilité un,  $\dim_H L(f) = 1 - J(f)$  et  $\dim_H(L_\varepsilon(f)) = (1 - J(f))(1 - (1 + \varepsilon)^{-2})$ .*

Les points de  $[0, 1] \setminus L(f)$  voient leur voisin visité par  $f$  un nombre fini de fois. Nos résultats pour  $f \in \mathcal{S}_1^{vb} \cup \mathcal{S}_1^{vi}$  explicitent la rupture d'Hausdorff-Besicovich. Les  $T_n$  de 2.3.3 sont aussi considérés.

Ces phénomènes ont une incidence en statistique. En particulier, les points de forte oscillation des estimateurs à noyau de la densité correspondent à une oscillation de  $\alpha_n$  proche d'un  $f$  fonction du noyau. Nous montrons alors que le noyau  $\mathbb{I}_{[-1/2, 1/2]}$  de Rosenblatt génère le moins de points où la déviation de l'estimateur approche infiniment souvent sa borne supérieure d'un facteur  $(1 - \varepsilon)$ .

Dans le cadre 2.3.4 des incréments du processus empirique multivarié, indexés par  $\mathcal{F}$ , choisissons une cible  $\varphi_h \in \mathcal{S}_d$  et considérons dans le compact  $V \subset \mathbb{R}^p$  les points

$$L(h) = \left\{ x \in V : \liminf_{n \rightarrow \infty} \left\| \frac{\Delta\alpha_n(a_n, x)}{\sqrt{2f_X(x)a_n |\log a_n|}} - \varphi_h \right\|_{\mathcal{F}} = 0 \right\}$$

visités infiniment souvent par un incrément mimant sur  $\mathcal{F}$  la configuration  $\varphi_h(f) = \int_{\Lambda^p} f h d\lambda_d$ .

**Théorème 2.37.** *Sous les conditions du Théorème 2F,  $\dim_H(L(h)) = d - \int_{\Lambda^p} h^2 d\lambda_d$  presque sûrement et  $\dim_H \left( \bigcup_{\varphi_h \in \mathcal{H}_0} L(h) \right) = d - \inf_{\varphi_h \in \mathcal{H}_0 \cap \mathcal{S}_d} \int_{\Lambda^p} h^2 d\lambda_d$ .*

### 2.4.3 – Convergences ensemblistes dans la loi d’Erdős-Rényi-Shepp

(i) Soit  $X_n$  une suite i.i.d. de vecteurs aléatoires centrés à valeurs dans un espace de Banach séparable  $\mathcal{X}$  de dual  $\mathcal{X}^*$ , tels que  $\mathbb{E}(\exp(t\|X\|)) < \infty$  pour  $t \in \mathbb{R}$  si  $\mathcal{X}$  est de dimension infinie et  $t$  au voisinage de 0 sinon. Pour  $\gamma \in ]0, \infty[$  et  $A_n = \lceil \gamma \log n \rceil$  considérons les moyennes mobiles de  $S_n = \sum_{i \leq n} X_i$ ,

$$F_n = \left\{ \frac{S_{i+A_n} - S_i}{A_n} : i \leq n \right\} \subset E_n = \left\{ \frac{S_{i+j} - S_i}{A_n} : i \leq n, j \leq A_n \right\}.$$

Soit  $B_{\mathcal{X}}$  la boule unité de  $\mathcal{X}$ . En notant  $\mu(\varphi) = \mathbb{E}(\exp \varphi(X))$ ,  $\varphi \in \mathcal{X}^*$ , la transformée de Laplace de  $\mathbb{P}^X$ , et  $h(x) = \sup_{\varphi \in \mathcal{X}^*} (\varphi(x) - \log \mu(\varphi))$ ,  $x \in \mathcal{X}$ , sa transformée de Crámer-Chernoff, les ensembles  $K_\gamma = \{x \in \mathcal{X} : h(x) \leq 1/\gamma\}$  sont convexes et compacts.

Les lois du logarithme itéré sont bien étudiées sur  $\mathcal{X}$ , voir [LT,EK], les lois du logarithme beaucoup moins. Reformulons la généralisation de la loi d’[ERé,Sp] obtenue dans [AK] de la façon suivante.

**Théorème 2F.** Les suites  $E_n$  et  $F_n$  sont presque sûrement relativement compactes dans  $\mathcal{X}$  d’ensemble limite  $K_\gamma$ . De plus,  $\mathbb{P}(\exists n_0 : E_n \subset K_\gamma + \varepsilon B_{\mathcal{X}} \subset F_n + 2\varepsilon B_{\mathcal{X}}, n > n_0) = 1$  pour tout  $\varepsilon > 0$ .

La motivation initiale de [22] est de rechercher les vitesses de convergence dans 2F à l’aide d’outils de déviations modérées tels que [EK2], dans la perspective également du calcul des vitesses de [AK2].

(ii) Or le cas  $\mathcal{X} = \mathcal{X}^* = \mathbb{R}^d$  n’est pas résolu, pas même pour des vecteurs gaussiens  $X_n$ . Notre intention dans [22] est donc de généraliser à  $\mathbb{R}^d$  la vitesse obtenue sur  $\mathbb{R}$  pour [ERé,Sp] dans [DD,DDL],

$$\limsup_{n \rightarrow \infty} \frac{A_n}{\log \log n} (\max E_n - m_\gamma^+) = \limsup_{n \rightarrow \infty} \frac{A_n}{\log \log n} (\max F_n - m_\gamma^+) = \frac{1}{2\varphi_{m_\gamma^+}} \quad p.s.$$

avec  $m_\gamma^+ = \max \{x : h(x) = 1/\gamma\}$ . Les  $\liminf$  centrées en  $m_\gamma^- = \min \{x : h(x) = 1/\gamma\}$  sont égales à  $-1/2\varphi_{m_\gamma^-}$ . Le vecteur  $\varphi_x \in \mathbb{R}^d$  est l’unique solution de

$$\text{grad } \mu(\varphi) = \mu(\varphi) x.$$

Dans [N] la formule de représentation de  $\mathbb{P}(S_n \in nK)$  pour tout convexe  $K \not\ni 0_d$  de  $\mathbb{R}^d$  fait apparaître le point dominant  $x = \arg \min_K h \in \partial K$  ainsi que la direction  $\varphi_x$  de l’énergie.

Pour mesurer la vitesse de recouvrement de  $K_\gamma$  par  $E_n$  nous cherchons les suites déterministes  $C_n \subset \mathbb{R}^d$ , proches de  $K_\gamma$ , de la classe supérieure  $\mathbb{P}(\exists n_0 : E_n \subset C_n, n > n_0) = 1$ , de la classe inf-supérieure  $\mathbb{P}(\exists n_k \uparrow \infty : E_{n_k} \subset C_{n_k}, k \geq 1) = 1$ , de la classe sup-inférieure  $\mathbb{P}(\exists n_k \uparrow \infty : E_{n_k} \not\subset C_{n_k}, k \geq 1) = 1$  et de la classe inférieure  $\mathbb{P}(\exists n_0 : E_n \not\subset C_n, n > n_0) = 1$ . Ces notions ensemblistes généralisent les classes de Lévy pour les processus réels. Posons  $\Delta_\gamma^- = \inf_{x \in \partial K_\gamma} \|\varphi_x\|_{\mathbb{R}^d}$ ,  $\Delta_\gamma^+ = \sup_{x \in \partial K_\gamma} \|\varphi_x\|_{\mathbb{R}^d}$  et

$$\varepsilon_n^+ = \frac{d}{2\gamma\Delta_\gamma^-} \frac{\log \log n}{\log n} \geq \varepsilon_n^- = \frac{d}{2\gamma\Delta_\gamma^+} \frac{\log \log n}{\log n}.$$

Dans les cas réguliers nous obtenons, pour tout  $\varepsilon > 0$ , que  $K_\gamma + (1+\varepsilon)\varepsilon_n^+ B_{\mathcal{X}}$  est dans la classe supérieure de  $F_n$  alors que  $K_\gamma + (1-\varepsilon)\varepsilon_n^- B_{\mathcal{X}}$  est dans la classe sup-inférieure. Ceci généralise [DD,DDL] ci-dessus puisqu’alors, par convexité de  $h$ ,  $\Delta_\gamma^\pm = \varphi_{m_\gamma^\pm}$  et  $K_\gamma = [m_\gamma^-, m_\gamma^+]$ . D’autres résultats sont en cours.

### 2.4.4 – Loi universelle pour la marche dans un domaine d’attraction stable

Dans [23] nous étudions l’extension de [W] dans le cadre des  $a_n$ -incrément du processus de somme partielle de v.a.r.  $X_1, \dots, X_n$  i.i.d. de loi  $\mathbb{P}^X$  telle que  $\mathbb{E}(e^{tX} \mathbb{1}_{X>0}) < \mathbb{E}(X^2 \mathbb{1}_{X<0}) = \infty$  pour un  $t > 0$  et  $X \mathbb{1}_{X<0}$  est dans le domaine d’attraction d’une loi  $\alpha$ -stable. En travaillant dans l’espace de [W], c’est-à-dire en topologie  $M_1$  de Skorokhod à une transformation près pour éliminer les sauts infinis, nous recherchons une normalisation universelle des incréments afin d’obtenir en un seul résultat la loi des grands nombres, la loi du logarithme itéré de [W] et les lois du logarithme de type [CS] et [ERé], dans l’optique de [Fr]. En notant  $h$  la transformée de Crámer-Chernoff de  $X \mathbb{1}_{X>0}$  la normalisation est du type  $b_n = a_n h^{-1}(L_n/a_n)$  avec  $L_n = |\log a_n| + \log \log n$ . Ces lois fonctionnelles sont ardues et demandent de développer un nouveau principe de grandes déviations, dans l’esprit de [LyS,Ar].

### 2.4.5 – Perspectives

(i) Nous avons obtenu, ou sommes en mesure de le faire, une version fonctionnelle, des vitesses et des applications à la plupart des lois du logarithme des livres [LL,R2] et des chapitres concernés dans [SW]. Nous rédigeons une telle compilation, en deux volumes, dont le présent texte constitue le squelette.

(ii) Les résultats de 2.1.5 décrivent en partie la transition de l'adhérence de  $\Theta_n$  à  $\mathcal{S}_1$  lorsque  $(\log n)/n \ll a_n \ll (\log n)^3/n$  vers l'adhérence à l'ensemble de type Erdős-Rényi  $\mathcal{O}_d$  lorsque  $a_n = d(\log n)/n$ . La description complète de ce passage vers la structure poissonnienne d' $\mathcal{O}_d$  manque encore à l'édifice 2.1. Nous recherchons pour cela des inégalités de concentration et de grandes déviations fonctionnelles, au seconde ordre près, directement pour le processus de Poisson standardisé  $(\pi_\rho - \rho)/\sqrt{\rho}$  en grande intensité  $\rho \rightarrow \infty$  à l'image de [Sm], en exploitant [Ar] et le principe de perturbation suggéré dans [1].

(iii) Dans l'approche de [14,16] relatée en 2.3 nous avons commencé à explorer des hypothèses du type intégrale comme dans [Mu], bien adaptées à la recherche de points adhérents. Il s'agit de caractériser les points limites sous cet angle. De plus nous avons rencontré une compétition entre  $a_n$  et  $\kappa_n$  pouvant interférer dans les lois de type Erdős-Rényi pour  $a_n$  petit, il faudrait approfondir ce phénomène.

(iv) Des lois fonctionnelles précises restent à établir pour diverses variantes de  $\alpha_n$  utiles en statistique non paramétrique : processus empirique pondéré avec loi des poids non bornée, processus empirique issu du bootstrap ou de ses itérations, étude des incréments avec longueur  $a_n$  aléatoire ou fonction de la position, étude conjointe des ensembles d'incréments utilisant plusieurs tailles  $a_{1,n}, \dots, a_{k(n),n}$ . Nous avons quelques pièces de ce puzzle mais d'autres restent largement ouvertes aux bonnes idées.

(v) Pour la version fonctionnelle de [Ws], comment mesurer la répartition des oscillations dans  $\Theta_n$  ?

(vi) Nous pouvons tirer de [17] des conditions générales sur une métrique tronquée ou une suite de métriques assurant des résultats du type 2.2.3. En particulier nous avons une partie des résultats pour la norme uniforme pondérée tronquée, en lien direct avec 3.2. Plus généralement, dans l'esprit de [DL] il est possible de produire des conditions suffisantes sur une semi-métrique pour exprimer nos vitesses de convergence en fonction des vitesses de petites et grandes déviations dans la topologie associée.

(vii) Le cas des variables  $X_n$  dépendantes, sous des conditions de mélange, peut s'aborder avec des outils comparables, quand ils sont disponibles. Les techniques de blocage sont tout aussi efficaces pour renouveler la trajectoire empirique à longues échelles de temps mais la dépendance gêne le contrôle du module de continuité, qui doit alors s'appréhender par des inégalités de moment et de couplage.

### 2.5 – Références citées en partie 2

- [A] De Acosta A. (1985). On the functional form of Lévy modulus of continuity for Brownian motion. *Z. Wahrsch. Verw. Gebiete* **69**, 567-579.
- [AK] De Acosta A. and Kuelbs J. (1983). Limit theorems for moving averages of independent random vectors. *Z. Wahrsch. und Verw. Gebiete* **64**, 67-123.
- [AK2] De Acosta A. and Kuelbs J. (1983). Some results on the cluster set  $C(S_n/a_n)$  and the law of the iterated logarithm. *Annals of Probability* **11**, 102-122.
- [Ar] Arcones M. (2002). The large deviation principle of stochastic processes. I. *Teorija Veroyatnost. i Primenen* **47**, 727-746.
- [BBK] Baldi P., Ben Arous G. and Kerkycharian G. (1992). Large deviations and the Strassen theorem in Hölder norm. *Stochastic Processes and their Applications* **42**, 171-180.
- [BR] Baldi P. and Roynette B. (1992). Some exact equivalents for the Brownian motion in Hölder norm. *Probability Theory and Related Fields* **93**, 457-484.
- [B] Bolthausen E. (1978). On the speed of convergence in Strassen's law of the iterated logarithm. *Annals of Probability* **6**, 668-672.
- [BS] Book S. A. and Shore T. R. (1978). On large intervals in the Csörgő-Révész theorem on increments of a Wiener process. *Z. Wahrsch. Verw. Gebiete* **46**, 1-11.

- [Bu] Bulinskii A. V. (1980). A new variant of the functional law of the iterated logarithm. *SIAM Theory of Probability and its Applications* **25**, 493-503.
- [BL] Bulinskii A. V. and Lifshits M. A. (1995). Rate of convergence in the functional law of the iterated logarithm with non-standard normalizing factors. *Uspekhi Mat. Nauk* **50**, 83-102 (in Russian). *Russian Math. Surveys* **50**, 925-944 (in English).
- [CKL] Chen G. J., Kong F. C. and Lin Z.Y. (1986). Answers to some questions about increments of a Wiener process. *Annals of Probability* **14**, 1252-1261.
- [Ch] Chung K. L. (1948). On the maximum partial sums of sequences of independent random variables. *Transactions of the American Mathematical Society* **64**, 205-233.
- [Ci] Ciesielski Z. (1960). On the isomorphism of the space  $H_\alpha$  and  $m$ . *Bulletin de l'Académie Polonaise des Sciences* **7**, 217-222.
- [Ci2] Ciesielski Z. (1960). Some properties of Schauder basis of the spaces  $C(0,1)$ . *Bulletin de l'Académie Polonaise des Sciences*, **8**, 141-144.
- [CsR] Csáki E. and Révész P. (1979). How big must be the increments of a Wiener process ? *Acta Mathematica Academiae Scientiarum Hungaricae* **33**, 37-49.
- [CR] Csörgő M. and Révész P. (1979). How big are the increments of a Wiener process ? *Annals of Probability* **7**, 731-737.
- [CR2] Csörgő M. and Révész P. (1979). How small are the increments of a Wiener process ? *Stochastic Processes and their Applications* **8**, 119-129.
- [D] Deheuvels P. (1991). Functional Erdős-Rényi laws. *Studia Scient. Math. Hungaricae* **26**, 261-295.
- [D2] Deheuvels P. (1992). Functional laws of the iterated logarithm for large increments of empirical and quantile processes. *Stochastic Processes and their Applications* **43**, 133-163.
- [D3] Deheuvels P. (2000). Chung-type functional laws of the iterated logarithm for tail empirical processes. *Annales de l'Institut Henri Poincaré, Prob. et Stat.* **36**, 583-616.
- [DD] Deheuvels P. and Devroye L. (1987). Limit laws of the Erdős-Rényi-Shepp type. *Annals of Probability* **15**, 1363-1386.
- [DDL] Deheuvels P., Devroye L. et Lynch J. (1986). Exact convergence rates in the limit theorems of Erdős-Rényi and Shepp. *Annals of Probability* **14**, 209-223.
- [DEMR] Deheuvels P., Einmahl J. H. J., Mason D.M., Ruymgaart F.H. (1987). The almost sure behavior of maximal and minimal  $k_n$  spacings. *Journal of Multivariate Analysis* **24**, 155-176.
- [DL] Deheuvels P. and Lifshits M. A. (1994). Necessary and sufficient conditions for the Strassen law of the iterated logarithm in nonuniform topologies. *Annals of Probability* **22**, 1838-1856.
- [DL2] Deheuvels P. and Lifshits M. A. (2001). Probabilities of hitting of shifted small balls by a centered Poisson process. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 278, Veroyatn. i Stat. 4, 63-85, in Russian (translated in 2003, *J. Math. Sci.* 118(6), 5541-5554).
- [DM] Deheuvels P. and Mason D. M. (1992). Functional laws of the iterated logarithm for the increments of empirical and quantile processes. *Annals of Probability* **20**, 1248-1287.
- [DM2] Deheuvels P. and Mason D. M. (1998). Random fractals functional laws of the iterated logarithm. *Studia Scientiarum Mathematica Hungaricae* **34**, 89-106.
- [DR] Deheuvels P. and Révész P. (1993). On the coverage of Strassen-type sets by sequences of Wiener processes. *Journal of Theoretical Probability* **6**, 427-449.
- [Dv] Devroye L. (1982). Upper and lower class sequences for minimal uniform spacings. *Z. Wahrsch. Verw. Gebiete* **61**, 237-254.
- [Du] Dudley R. M. (1984). A course on empirical processes. *Ecole d'été de Probabilités de Saint-Flour XII 1982, Lecture Notes in Mathematics* **1097**, 1-142, Springer-Verlag.
- [E] Erdős P. (1942). On the law of iterated logarithm. *Annals of Mathematics* **43**, 419-436, MR 4-16.

- [ERé] Erdős P. and Rényi A. (1970). A new law of large numbers. *Journal of Analysis and Mathematics* **23**, 103-111.
- [ER] Erdős P. and Révész P. (1990). A new law of the iterated logarithm. *Acta Mathematica Hungarica* **55**, 125-131.
- [EK] Einmahl U. and Kuelds J. (2001). Cluster sets for a generalized law of the iterated logarithm in Banach spaces. *Annals of Probability* **29**, 1451-1475.
- [EK2] Einmahl U. and Kuelds J. (2004). Moderate deviation probabilities for open convex sets : non logarithmic behavior. *Annals of Probability* **32**, 1316-1355.
- [EM] Einmahl U. and Mason D. M. (1993). Rates of clustering in Strassen's LIL for partial sum processes. *Probability Theory and Related Fields* **97**, 479-487.
- [EM2] Einmahl U. and Mason D. M. (1994). A universal Chung-type law of the iterated logarithm. *Annals of Probability* **22**, 1803-1825.
- [F] Finkelstein H. (1971). The law of the iterated logarithm for empirical distributions. *Annals of Mathematical Statistics* **23**, 103-111.
- [Fv] Frolov A. N. (2004). Universal limit theorems for increments of processes with independent increments. (Russian) *Teor. Veroyatn. Primen.* **49**, 601-609.
- [GMZ] Giné E., Mason D.M. and Zaitsev A. (2003). The  $L_1$ -norm density estimator process. *Annals of Probability* **31**, 719-768.
- [Gr] Grill K. (1992). Exact rate of convergence in Strassen's law of the iterated logarithm. *Journal of Theoretical Probability* **5**, 197-204.
- [GK] Goodman V. and Kuelbs J. (1991). Rates of clustering in Strassen's LIL for Brownian motion. *Journal of Theoretical Probability* **4**, 285-309.
- [GS] Gaenssler P. and Stute W. (1979). Empirical processes : A survey of results for independent and identically distributed random variables. *Annals of Probability* **7**, 193-243.
- [HR] Hanson D. L. and Russo R. P. (1983). Some results on increments of the Wiener process with applications to lag sums of i.i.d. random variables. *Annals of Probability* **11**, 609-623.
- [Ho] Hong S. Y. (1992). Local oscillation modulus of the uniform empirical process. *Syst. Sciences and Math. Sciences* **5**, 164-179.
- [Ki] Kiefer J. (1967). On Bahadur's representation of sample quantiles. *Annals of Mathematical Statistics* **38**, 1323-1342.
- [KMT] Komlós J., Major P. and Tusnády G. (1975). An approximation of partial sums of independent r.v.'s and the sample d.f. I. *Z. Wahrsch. und Verw. Gebiete* **32**, 111-131.
- [K] Kuelbs J. (1979). Rates of growth for Banach space valued increment processes. *Lectures Notes in Mathematics* **709**, 151-169.
- [KL] Kuelbs J. and Li W. (1993). Small ball estimates for Brownian motion and the Brownian sheet. *Journal of Theoretical Probability* **6**, 547-577.
- [La] Lai T. L. (1974). Reproducing kernel Hilbert spaces and the law of the iterated logarithm for Gaussian processes. *Z. Wahrsch. und Verw. Gebiete* **29**, 7-19.
- [Le] Ledoux M. (1994). Isoperimetry and Gaussian Analysis. Lectures on probability theory and statistics (Saint-Flour). *Lecture Notes in Mathematics* **1648**, 165-294.
- [LT] Ledoux M. and Talagrand M. (1991). Probability in Banach Spaces. *Springer-Verlag*.
- [Lé] Lévy P. (1948). Processus stochastiques et mouvement Brownien. *Gauthier-Villars, Paris*.
- [LS] Le Page R. and Schreiber B. M. (1985). An iterated logarithm law for families of Brownian paths. *Z. Wahrsch. und Verw. Gebiete* **70**, 341-344.
- [LL] Lin Z. Y. and Lu C. R. (1992). Strong Limit Theorems. *Mathematics and Its Applications, Kluwer*.

- [LyS] Lynch J. and Sethuraman J. (1987). Large deviations for processes with independent increments. *Annals of Probability* **15**, 610-627.
- [M] Mason D. M. (1984). A strong limit theorem for the oscillation modulus of the uniform empirical quantile process. *Stochastic Processes and their Applications* **17**, 127-136.
- [M2] Mason, D. M. (1988). A strong invariance theorem for the tail empirical process. *Annales de l'Institut Henri Poincaré, Prob. et Stat.* **24**, 491-506.
- [M3] Mason, D. M. (2004). A uniform functional law of the iterated logarithm for the local empirical process. *Annals of Probability* **32**, 1391-1418.
- [MSW] Mason D. M., Shorack G. R. and Wellner J. A. (1983). Strong limit theorems for oscillation moduli of the uniform empirical process. *Z. Wahrsch. Verw. Gebiete* **65**, 83-97.
- [Mo] Mogulskii, A. A. (1980). On the law of the iterated logarithm in Chung's form for functional spaces. *SIAM Theory Probability and its Applications* **24**, 405-413.
- [Mu] Mueller C. (1981). An unification of Strassen's law and Lévy's modulus of continuity. *Z. Wahrsch. Verw. Gebiete* **50**, 257-264.
- [MS] Musielak J. and Semadeni Z. (1961). Some classes of Banach spaces depending on a parameter. *Studia Mathematica* **20**, 271-284.
- [N] Ney P. (1983). Dominating points and the asymptotics of large deviations for random walk on  $\mathbb{R}^d$ . *Annals of Probability* **11**, 158-167.
- [OT] Orey S. and Taylor S. I. (1974). How often on a Brownian path does the law of the iterated logarithm fail? *Proceedings of the London Mathematical Society* **28**, 174-192.
- [R] Révész P. (1979). A generalization of Strassen's functional law of the iterated logarithm. *Z. Wahrsch. und Verw. Gebiete* **50**, 257-264.
- [R2] Révész P. (1990). Random walk in random and non-random environments. *World Scientific Publishing, Singapore*.
- [R3] Révész P. (1991). Waiting for the coverage of the Strassen's set. *Studia Scientiarum Mathematicarum Hungaricae* **26**, 379-391.
- [Sh] Shao Q. M. (1989). On a problem of Csáki and Révész. *Annals of Probability* **17**, 809-812.
- [Sp] Shepp L. A. (1964). A limit theorem concerning moving averages. *Annals of Mathematical Statistics* **35**, 424-428.
- [Sm] Shmileva E. (2003). Small ball probability for centered Poisson process of high intensity. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI)* **298**, 280-303, in Russian; translation in *J. Math. Sci.* **128** (2005), 2656-2668.
- [Sb] Steinebach J. (1982). Between invariance principles and Erdős-Rényi laws. In : "Limit theorems in Probability and Statistics", Vezsprém. Coll. Math. Soc. János Bolyai **36**, 981-1005.
- [S] Strassen V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. und Verw. Gebiete* **3**, 211-226.
- [St] Stute W. (1982). The oscillation behavior of empirical processes. *Annals of Probability* **10**, 86-107.
- [T] Talagrand M. (1992). On the rate of clustering in Strassen's law of the iterated logarithm for Brownian motion. *Probability in Banach Spaces, Birkhäuser.* **8**, 339-347.
- [V] V. van der Vaart (1998). Asymptotic Statistics. *Cambridge University Press*.
- [W] Wichura M. (1974). Functional laws of the iterated logarithm for the partial sums of i.i.d. random variables in the domain of attraction of a completely asymmetric stable laws. *Annals of Probability* **2**, 1108-1138.
- [Ws] Wschebor M. (1995). Almost sure weak convergence of the increments of Lévy processes. *Stochastic Processes and their Applications* **55**, 253-270.



### 3 – OUTILS : PETITES ET GRANDES DÉVIATIONS, APPROXIMATION FORTE

Le développement des lois fonctionnelles de la partie 2 ainsi que leur application dans la partie 4 nécessitent de nouveaux outils qui ont aussi, et avant tout, un grand intérêt en eux-mêmes.

En 3.1 ([6]) nous évaluons précisément la propension du mouvement brownien à visiter les fonctions localement à variation infinie de la frontière de l'ensemble de Strassen. Ceci résoud une partie d'un problème de petites déviations bien connu, sinon le plus connu, qui résiste depuis plus de vingt ans.

En 3.2 ([4]) nous abordons un autre problème de petites déviations, celui des petites boules browniennes en métrique uniforme pondérée. Le résultat principal a diverses applications intéressantes, notamment l'évaluation de la mesure de Wiener de faisceaux de trajectoires non symétriques. Ceci participe à la preuve de 3.1 et nous permet, en puissance, d'étendre certains résultats de la partie 2.1 à cette topologie forte, de la même façon que nous l'avons fait pour la topologie höldérienne en 2.2.

En 3.3 ([10,15]) nous résolvons un problème de grandes déviations qui constitue l'ingrédient nécessaire et décisif de la nouvelle méthode 4.2-4.3 pour contrôler l'argmax empirique d'une densité. Il s'agit d'un principe de grande déviation pour l'argmin d'un mouvement brownien avec forte dérive positive. La vitesse qui apparaît est atypique et fonction de la dérive. Notre méthode s'étend à d'autres processus.

En 3.4 ([11,25]) nous présentons une série de résultats dans le domaine difficile de l'approximation gaussienne des processus empiriques. Ils fournissent des vitesses dans les principes d'invariance faible et fort pour les mesures empiriques indexées par des fonctions, sous des conditions d'entropie générales ou des conditions géométriques plus fortes. Les méthodes mises en place sont inédites, tirent profit d'inégalités récentes et montrent la puissance d'un outil jusqu'ici négligé. C'est une porte ouverte à la modernisation et la simplification des résultats fort techniques des années quatre-vingt, y compris sur l'approximation des sommes de v.a. à valeurs dans des espaces de Banach.

Dans chaque cas il a fallu inventer de nouvelles méthodes, sans rapport les unes avec les autres, et s'immerger dans une littérature très dense, car les trois domaines en question sont trois branches distinctes de la théorie des probabilités, avec des listes d'auteurs bien différentes.

#### 3.1 – Petites déviations browniennes

Soit  $I$  l'identité de  $[0, 1]$ . Considérons un processus de Wiener  $\{W(t) : t \geq 0\}$  et la trajectoire

$$W_T = \frac{W(TI)}{\sqrt{2T \log \log T}}$$

qui, d'après le théorème de Strassen [S], est presque sûrement relativement compacte dans  $\mathcal{B}$  d'ensemble limite  $\mathcal{S}_1$  de (2). Comme  $\mathcal{S}_1$  est fermé dans  $\mathcal{B}$  ceci prouve que

$$\liminf_{T \rightarrow \infty} \|W_T - f\| \begin{cases} = 0 & \text{si } f \in \mathcal{S}_1 \\ > 0 & \text{si } f \notin \mathcal{S}_1 \end{cases} \quad p.s.$$

Dans cette section nous évoquons le problème de la vitesse pour  $f \in \mathcal{S}_1$ . Il se résout en évaluant très exactement les probabilités de petite déviation. Les bornes ainsi calculées servent en section 2.1. Les théorèmes de 3.1.2 restent vrais pour le processus  $S_n/\sqrt{n}$  de l'introduction 1.12, avec  $X \in \mathbb{L}_2$ .

##### 3.1.1 – Frontière de Strassen

(i) Soit  $\mathcal{VB}$  l'ensemble des fonctions  $g$  à variation  $Var(g) < \infty$ . Décomposons la frontière, énergétique mais non topologique, de la boule de Strassen  $\mathcal{S}_\theta$  en posant  $\mathcal{S}_\theta^* = \mathcal{S}_\theta^- \cup \mathcal{S}_\theta^{vb} \cup \mathcal{S}_\theta^{vli} \subset \mathcal{S}_\theta \subset \sqrt{\theta}B_0$  avec

$$B_0 = \{f : \|f\| < 1\}, \quad \mathcal{S}_\theta^- = \{f : J(f) < \theta\}, \quad \mathcal{S}_\theta^{vb} = \left\{f : J(f) = \theta, f' \in \mathcal{VB}\right\}$$

et  $\mathcal{S}_\theta^{vi}$  l'ensemble des  $f \in \mathcal{S}_\theta$  telles que  $J(f) = \theta$  et  $f'$  est à *variation localement infinie*, au sens où pour certains  $x_1, \dots, x_k$  dans  $[0, 1]$  il vient, pour tout  $\eta > 0$  et  $A_\eta = \bigcup_{j \leq k} [x_j - \eta, x_j + \eta]$ ,

$$Var(f' \mathbb{I}_{[0,1] \setminus A_\eta}) < Var(f' \mathbb{I}_{[0,1] \cap A_\eta}) = \infty.$$

Notons que  $\{f : J(f) = \theta\} \setminus (\mathcal{S}_\theta^{vb} \cup \mathcal{S}_\theta^{vi}) \neq \emptyset$ . Pour être exact, à  $f \in \mathcal{S}_\theta^{vb}$  il convient d'associer à  $f'$  une mesure signée  $\mu_f$  à variation bornée telle que  $f = - \int_0^1 \mu_f([x, 1]) dx$ . Ceci permet d'utiliser l'intégration par parties  $\int_0^1 f' dW = W f' - \int_0^1 W d\mu_f$  dans la formule de Cameron-Martin-Girsanov pour  $W$ . Dans les cas où  $f \in \mathcal{S}_\theta^{vi}$  nous nous y ramenons par l'astuce principale de [6], car sans ce calcul précis, les petites déviations ne sont plus optimales. Pour  $f \in \mathcal{S}_\theta^-$  ce passage n'est pas critique.

(ii) Fixons  $\theta = 1$ . L'énergie  $J(f)$  quantifie au premier ordre l'effort que fournit  $W_T$  pour ressembler à  $f$ , comme l'atteste le principe de grande déviation usuel [DS] sur  $(\mathcal{C}, \|\cdot\|)$  avec bonne fonction de taux  $J/2$  et le principe de petite déviation de [A2].

La loi de Chung [Ch], qui correspond à  $f = 0$ , a été généralisée dans [Cs,A2] comme suit.

**Proposition 3A.** *Si  $f \in \mathcal{S}_1^-$  alors*

$$\liminf_{T \rightarrow \infty} (\log \log T) \|W_T - f\| = \frac{\pi}{4\sqrt{1 - J(f)}} \quad p.s.$$

Inversement, si  $J(f) = 1$  il est montré dans [GK] que

$$\liminf_{T \rightarrow \infty} (\log \log T)^{2/3} \|W_T - f\| = \chi_f < \infty \quad p.s.$$

La vitesse dans cette  $\liminf$  est liée à la fonction cruciale  $\varphi_f(\varepsilon) = 1 - \inf_{\|g-f\| \leq \varepsilon} J(g)$  d'économie d'énergie globale de  $f$ , l'infimum étant pris parmi les  $g$  absolument continues. Noter que  $\varphi_f$  est aussi en rapport avec la longueur et la régularité de  $f$ . La constante  $\chi_f$  ci-dessus vérifie

$$d_f = \lim_{\varepsilon \rightarrow 0} \frac{\varphi_f(\varepsilon)}{\varepsilon} \begin{cases} \in [2, \infty[ & \text{si } \chi_f > 0 \\ = \infty & \text{si } \chi_f = 0, \end{cases}$$

où  $d_f$  existe toujours, d'après [KLT], et détermine  $\chi_f$ . D'après [Gr2] il y a équivalence entre  $\chi_f > 0$  et  $f \in \mathcal{S}_1^{vb}$ . Dans [Cs2] les valeurs de  $\chi_f$  sont évaluées pour  $f$  linéaire par morceaux ou quadratique puis la méthode a été étendue par [GL] et  $\chi_f$  caractérisée pour tout  $f \in \mathcal{S}_1^{vb}$  comme unique solution d'une équation. Enfin, [Gr2] montre que si  $f \in \{f : J(f) = 1, Var(f') = \infty\}$  alors

$$\liminf_{T \rightarrow \infty} \frac{\|W_T - f\|}{\varepsilon(T)} \in [1, 2] \quad p.s. \quad (5)$$

où  $\varepsilon(T)$  est l'unique solution de  $\varphi_f(\varepsilon) = (\pi/4\varepsilon \log \log T)^2$ . Ainsi, lorsque  $T \rightarrow \infty$  nous avons  $\varepsilon(T) \downarrow 0$ ,  $\varepsilon(T) \gg 1/\log \log T$  et  $\varepsilon(T) \ll 1/(\log \log T)^{2/3}$ , c'est-à-dire des vitesses intermédiaires. L'équation en  $\varphi_h$  reste valable sur  $\mathcal{S}_1^-$  avec une constante limite égale à 1. Elle reste aussi valable pour  $f \in \mathcal{S}_1^{bv}$  mais la constante limite n'est plus 1, car  $(16d_f/\pi^2)^{1/3} \neq \chi_f$  pour certains  $f \in \mathcal{S}_1^{vb}$ .

### 3.1.2 – Vitesse à la frontière irrégulière de Strassen

(i) Au bilan, si  $J(f) = 1$  et  $Var(f') = \infty$  la constante de (5) résiste depuis des années, y compris pour la fonction puissance  $f_1(x) = bx^\alpha$ . Elle dépend nécessairement des rectifications de  $f$

$$\mathcal{R}_f(\varepsilon) = \{f_\varepsilon : f_\varepsilon \in \mathcal{S}_1, \|f_\varepsilon - f\| \leq \varepsilon, J(f_\varepsilon) = 1 - \varphi_f(\varepsilon)\}$$

qui sont proches de  $f$  mais de moindre longueur. L'étude directe de  $\mathcal{R}_f$  n'est pas envisageable car il n'y a pas de méthode analytique pour comprendre  $\varphi_f$ . Intuitivement, il est d'autant plus difficile

pour la trajectoire irrégulière  $W_T$  de produire l'effort de ressembler à  $f$  d'énergie  $J(f) = 1$  maximale si  $f$  est en outre régulière. Supposons que  $f \in \mathcal{S}_1^{vli}$ , de sorte que l'irrégularité à imiter est locale.

Dans [6] nous parvenons à calculer la probabilité de petite boule  $\mathbb{P}(\|W_T - f\| < \varepsilon)$  en la minorant par des probabilités du type  $\mathbb{P}(W_T \in D(f, \varepsilon))$  où  $D(f, \varepsilon)$  n'est qu'une partie de la boule uniforme centrée en  $f$ . Il ne s'agit pas d'une boule mais d'un faisceau de trajectoires guidées dans un petit tube très resserré aux points critiques de  $f$  et relâché ailleurs. Notre astuce est donc de nature géométrique.

(ii) Pour un borélien  $A \subset [0, 1]$  et  $f$  absolument continue posons  $\|f\|_A = \sup_A |f|$ ,  $J(f, A) = \int_A f'^2 d\lambda$  et  $Var(f', A)$  la variation totale de  $f'$  sur  $A$ . Introduisons la fonction d'économie d'énergie locale

$$\varphi_f(\varepsilon, A) = J(f, A) - \inf_{\|f-g\| \leq \varepsilon} J(g, A)$$

qui joue un rôle essentiel dans notre problème. Ainsi  $\varphi_f(\varepsilon, [0, 1]) = \varphi_f(\varepsilon)$ . Considérons le cas où l'énergie de  $f \in \mathcal{S}_1^{vli}$  doit être économisée en 0. La vitesse  $\varepsilon(T)$  correspond à  $\nabla_f(\log \log T)$  de 3.1.3.

**Théorème 3.1.** *Soit  $f$  vérifiant  $J(f) = 1$  et  $Var(f', [x, 1]) < Var(f', [0, x]) = \infty$  pour tout  $x \in ]0, 1[$ . Nous avons alors  $\varphi_f(\varepsilon, [0, x]) \sim \varphi_f(\varepsilon)$  lorsque  $\varepsilon \rightarrow 0$  et*

$$\liminf_{T \rightarrow \infty} \frac{\|W_T - f\|}{\varepsilon(T)} = 1 \quad p.s.$$

où  $\varepsilon(T)$  est l'unique solution en  $\varepsilon$  de  $\varphi_f(\varepsilon) = (\pi/4\varepsilon \log \log T)^2$ .

Par conséquent, il suffit de trouver un équivalent à  $\varphi_f(\varepsilon, [0, x])$  plutôt que  $\varphi_f(\varepsilon)$  et la rectification optimale n'est pas requise. Des exemples simples figurent en 3.1.4.

Illustrons le lien entre  $\mathcal{R}_f(\varepsilon)$  et le fait que le point à rectifier soit 0.

**Exemple.** Supposons que  $f$  du Théorème 3.1 soit ou bien (i) concave sur  $[0, x_0]$  avec  $f'(0) = \infty$  ou bien (ii) convexe sur  $[0, x_0]$  avec  $f'(0) = -\infty$ . Soit  $\delta_\varepsilon$  la plus petite solution de (i)  $f(\delta) \geq \varepsilon + \delta f'(\delta)$  ou (ii)  $f(\delta) \leq -\varepsilon + \delta f'(\delta)$  de sorte que  $\delta_\varepsilon \downarrow 0$  si  $\varepsilon \rightarrow 0$ . Alors pour tout  $\varepsilon > 0$  assez petit il existe  $f_\varepsilon \in \mathcal{R}_f(\varepsilon)$  et  $x \in ]0, x_0[$  tel que  $\mathbb{I}_{[\delta_\varepsilon, x]} f'_\varepsilon = \mathbb{I}_{[\delta_\varepsilon, x]} f'$  presque partout. Dans ce cas nous avons aussi  $\varphi_f(\varepsilon) = J(f, [0, \delta_\varepsilon]) - (|f(\delta_\varepsilon)| - \varepsilon)^2 / \delta_\varepsilon + O(\varepsilon)$ .

(iii) La présence de plusieurs points critiques occasionne des difficultés techniques dans la preuve de [6] mais le résultat reste le même, les économies d'énergie s'additionnent dans  $\varphi_f(\varepsilon)$ .

**Théorème 3.2.** *Si  $f \in \mathcal{S}_1^{vli}$  alors, pour tout  $\eta > 0$  fixé nous avons  $\varphi_f(\varepsilon, A_\eta) \sim \varphi_f(\varepsilon)$  quand  $\varepsilon \rightarrow 0$ , avec  $A_\eta = \bigcup_{j \leq k} [x_j - \eta, x_j + \eta]$ , et la conclusion du Théorème 3.1 reste vraie.*

Clairement, les points  $x_1, \dots, x_k$  ne jouent pas tous le même rôle. Si  $x_i$  est sous-critique, au sens où  $\varphi_f(\varepsilon, [x_i - \eta, x_i + \eta]) \ll \varphi_f(\varepsilon)$  quand  $\varepsilon \rightarrow 0$  pour un  $\eta > 0$  assez petit, alors il n'influence pas la vitesse  $\varepsilon(T)$ , du moins asymptotiquement, et peut donc être ignoré.

Nous apprenons aussi que la position du point le plus critique  $x_i$  peut influencer la constante dans  $\varepsilon(T)$  mais pas l'ordre de sa vitesse, au sens de la permutation de morceaux de  $f$  et notamment d'un voisinage de  $x_i$ . En particulier, placer ce point critique en  $x_1 = 0$  génère un  $\varepsilon(T)$  plus grand puisque  $f(0) = 0$  ne peut pas être rectifié en  $f(0) = \pm\varepsilon$  donc fait passer de  $\varphi_f(\varepsilon, [0, \eta])$  à  $\varphi_f(2\varepsilon, [x_i, x_i + \eta])$ .

(iv) **Principe de démonstration.** Nous établissons en premier lieu que

$$\varphi_f(\varepsilon, A_\eta) \gg \varepsilon, \quad \varphi_f(\varepsilon, [0, 1] \setminus A_\eta) = O(\varepsilon), \quad \varphi_f(\varepsilon) \geq \varphi_f(\varepsilon, A_\eta) \geq (1 - \rho_f(\varepsilon))\varphi_f(\varepsilon)$$

avec  $\rho_f(\varepsilon) > 0$  et  $\rho_f(\varepsilon) \rightarrow 0$  si  $\varepsilon \rightarrow 0$ . Le passage clef consiste alors à se recentrer en une combinaison de bonnes rectifications locales  $f_{c, \varepsilon, i}$  de  $f$  pour forcer le mouvement brownien  $W$  à passer par certains points d'économie d'énergie aux bords de la petite boule initiale. Si  $\varepsilon$  est assez petit, pour  $i \geq 1$  et  $|c| \leq \varepsilon$  il existe  $f_{c, \varepsilon, i} \in \mathcal{S}_1$  telle que

$$f_{c, \varepsilon, i}(x_i - \eta) = f(x_i - \eta) + c, \quad \|f_{c, \varepsilon, i} - f\|_{A_{\eta, i}} \leq \varepsilon, \quad J(f_{c, \varepsilon, i}, A_{\eta, i}) \leq J(f, A_{\eta, i}) - (1 - \beta_i)\varphi_f(\varepsilon, A_{\eta, i}).$$

La formule de translation des boules browniennes fait ensuite apparaître, par changement de mesure, un produit d'espérances conditionnelles et nous concluons en majorant les exponentielles grâce aux rectifications  $f_{c, \varepsilon, i}$  et aux relations entre les intégrales représentant leurs énergies.  $\square$

### 3.1.3 – Un lemme de petites déviations

(i) Fixons  $\theta > 0$ . Nous avons synthétisé les différents sous-résultats de la section 2.1.5 grâce aux notations suivantes. Les vitesses  $\nabla_f(L)$  et les constantes  $\chi_f(\theta)$  sont définies par

$$\begin{aligned} \text{si } f \in S_\theta^-, \quad \nabla_f(L) = L \text{ et } \chi_f(\theta) &= \frac{\pi}{4\sqrt{\theta - J(f)}} \\ \text{si } f \in S_\theta^{vli}, \quad \nabla_f(L) \text{ est l'unique solution de (6) et } \chi_f(\theta) &= 1 \\ \text{si } f \in S_\theta^{vb}, \quad \nabla_f(L) = L^{2/3} \text{ et } \chi_f(\theta) &= \chi_f. \end{aligned}$$

Le réel  $\chi_f > 0$  est la solution de l'équation (3.1) de [GL]. Lorsque  $f \in S_\theta^{vli}$ ,  $\nabla_f(L)$  est la solution en  $\nabla$  de l'équation

$$\theta - \inf_{\|g-f\| \leq 1/\nabla} J(g) = \frac{\pi^2 \nabla^2}{16L^2} \quad (6)$$

Dans ce cas nous avons  $L^{2/3} \ll \nabla_f(L) \ll L$  lorsque  $L \rightarrow \infty$ . Des exemples  $\nabla_f(L) = c_\nu L^\nu$  avec  $\nu \in (2/3, 1)$  sont donnés en 3.1.4. Nous montrons, en analysant (6), que  $\nabla_f((1+o(1))L) \sim \nabla_f(L)$  si  $L \rightarrow \infty$  et  $\nabla_f(L) \leq \nabla_f(\delta L) \leq \delta \nabla_f(L)$  pour tout  $\delta \geq 1$  et tout  $L$  assez grand.

(ii) Dans [9] nous condensons l'ensemble de ces connaissances dans un lemme de petites déviations, en retravaillant les bornes manquantes et en compilant les résultats de [Cs,A2,GL,BL].

**Lemme 3.3.** *Soit  $\{W(t) : t \in [0, 1]\}$  un mouvement brownien standard,  $\theta > 0$  et  $f \in \mathcal{S}_\theta^*$ . Pour tout  $\delta > 0$  suffisamment petit, il existe  $\gamma^\pm = \gamma^\pm(\theta, f, \delta) \in ]0, \infty[$  tel que, si  $\xi$  est assez grand*

$$\begin{aligned} \mathbb{P} \left( \left\| \frac{W}{\xi} - f \right\| \leq \frac{(1+\delta)\chi_f(\theta)}{\nabla_f(\xi^2/2)} \right) &\geq \exp \left( -\frac{\theta\xi^2}{2} + \gamma^+ \frac{\nabla_f^2(\xi^2/2)}{\xi^2} \right) \\ \mathbb{P} \left( \left\| \frac{W}{\xi} - f \right\| \leq \frac{(1-\delta)\chi_f(\theta)}{\nabla_f(\xi^2/2)} \right) &\leq \exp \left( -\frac{\theta\xi^2}{2} - \gamma^- \frac{\nabla_f^2(\xi^2/2)}{\xi^2} \right). \end{aligned}$$

Ces deux bornes sont extrêmement précises car si  $\delta \rightarrow 0$  nous avons  $\gamma^\pm(\theta, f, \delta) \rightarrow 0$  dans chacune des six preuves incluses dans l'énoncé, selon que  $f \in S_\theta^-$ ,  $f \in S_\theta^{vb}$  ou  $f \in S_\theta^{vli}$ .

Sur la frontière très irrégulière  $f \in \mathcal{S}_\theta \setminus \mathcal{S}_\theta^*$  nous avons la même borne supérieure mais, via [Gr2], la borne inférieure hérite d'un facteur 1/2 au second terme dans l'exponentielle.

### 3.1.4 – Exemples de constantes explicites

(i) Pour illustrer 3.1.2 et 2.1.5 par des vitesses et constantes explicites, considérons quelques fonctions de  $\mathcal{S}_1^{vli}$  localement Hölderiennes d'indice  $\alpha \in ]1/2, 1[$ . Elles vérifient toutes  $\varphi_f(\varepsilon) = O(\varepsilon^{(2\alpha-1)/\alpha})$  avec des constantes exactes calculées par des arguments d'optimisation et les propriétés de  $J(f, \varepsilon)$  bien connues. Posons  $\Delta_\alpha = \alpha^2(1-\alpha)^{(1-\alpha)/\alpha}/(2\alpha-1)$  et fixons  $\varepsilon$  aussi petit que nécessaire.

**Exemple.** La fonction  $f_1(x) = bx^\alpha = \pm\sqrt{2\alpha-1}x^\alpha/\alpha$  vérifie  $\varphi_{f_1}(\varepsilon) = 2|b|^{1/\alpha} \Delta_\alpha \varepsilon^{(2\alpha-1)/\alpha}$  avec le facteur constant dans  $]1, 2[$ . La fonction  $f_2(x) = cx^\alpha + ax$ ,  $0 < c^2 < c_\alpha^2 = (2\alpha-1)/(1-\alpha)^2$ ,  $a = c(-1 \pm \sqrt{1/c^2 - 1/c_\alpha^2})$  vérifie  $\varphi_{f_2}(\varepsilon) = 2|c|^{1/\alpha} \Delta_\alpha \varepsilon^{(2\alpha-1)/\alpha} + 2a\varepsilon$ .

Plaçons à présent le point critique en  $y \in ]0, 1[$  avec un petit intervalle à gauche de  $y$  (resp. à droite) de longueur  $l_- \in ]0, y[$  (resp.  $l_+ \in ]0, 1-y[$ ). Noter que  $y$  et  $l_\pm$  ne jouent aucun rôle.

**Exemple.** La fonction  $f_3 = \mathbb{I}_{[0,y]}f_{3-} + \mathbb{I}_{[y,y+l_+]}f_{3\alpha} + \mathbb{I}_{[y+l_+,1]}f_{3+}$  avec  $f_{3\alpha}(x) = f_{3-}(y) + b(x-y)^\alpha$ ,  $b \neq 0$  et  $f_{3\pm}'$  à variation bornée, vérifie  $\varphi_{f_3}(\varepsilon) = 2^{(3\alpha-1)/\alpha} |b|^{1/\alpha} \Delta_\alpha \varepsilon^{(2\alpha-1)/\alpha} + O(\varepsilon)$ .

**Exemple.** La fonction  $f_4 = \mathbb{I}_{[0,y-l_-]}f_{4-} + \mathbb{I}_{[y-l_-,y]}f_{4\alpha} + \mathbb{I}_{[y,y+l_+]}f_{4\alpha} + \mathbb{I}_{[y+l_+,1]}f_{4+}$  avec  $a^2 < y$ ,  $b_\pm \neq 0$ ,  $f_{4\pm}(x) = a + b_\pm|y-x|^\alpha$  et  $f_{4\pm}'$  à variation bornée est doublement Hölder en  $y$  et vérifie d'une part, si  $b_-b_+ > 0$ ,  $\varphi_{f_4}(\varepsilon) = 2^{(3\alpha-1)/\alpha} (|b_-|^{1/\alpha} + |b_+|^{1/\alpha}) \Delta_\alpha \varepsilon^{(2\alpha-1)/\alpha} + O(\varepsilon)$  et d'autre part, si

$b_- b_+ < 0$ ,  $\varphi_{f_4}(\varepsilon) = 2^{(3\alpha-1)/\alpha} (|b_-|^{1/(1-\alpha)} + |b_+|^{1/(1-\alpha)})^{(1-\alpha)/\alpha} \Delta_\alpha \varepsilon^{(2\alpha-1)/\alpha} + O(\varepsilon)$  auquel cas moins d'énergie s'économise.

**Exemple.** La fonction  $f_5 = \sum_{i=1}^n \mathbb{I}_{[x_i, x_i+l_i]} f_{5,i} + \mathbb{I}_{I_0} f_{5,0}$  avec, pour  $i = 0, \dots, k$ ,  $\alpha_i \in ]1/2, 1[$ ,  $b_i \neq 0$ ,  $x_i \in ]0, 1[$ ,  $x_i < x_{i+1}$ ,  $l_i \in ]0, x_{i+1} - x_i]$ ,  $f_{5,i}(x) = a_i + b_i(x - x_i)^{\alpha_i}$ ,  $I_0 = [0, 1] \setminus \bigcup_{i=1}^k [x_i, x_i + l_i]$  et  $f'_{5,0}$  à variation bornée vérifie, pour  $\alpha = \min_i \alpha_i$ ,  $\varphi_{f_5}(\varepsilon) = 2^{(3\alpha-1)/\alpha} \sum_{i:\alpha_i=\alpha} |b_i|^{1/\alpha} \Delta_\alpha \varepsilon^{(2\alpha-1)/\alpha} + O(\varepsilon)$ .

(ii) En particulierisant le Théorème 3.2 comme suit, il s'applique aux fonctions  $f_j$  avec  $\rho = 2 - 1/\alpha$ .

**Corollaire 3.4.** Si  $f \in \mathcal{S}_1^{vli}$  et  $\varphi_f(\varepsilon) \sim \gamma \varepsilon^\rho$  avec  $\rho < 1$  et  $\gamma \in ]0, \infty[$  alors

$$\liminf_{T \rightarrow \infty} (\log \log T)^{2/(\rho+2)} \|W_T - f\| = \left( \frac{\pi^2}{16\gamma} \right)^{1/(\rho+2)} \quad p.s.$$

**Exemple.** Pour  $f_1$  nous obtenons

$$\liminf_{T \rightarrow \infty} (\log \log T)^{2\alpha/(4\alpha-1)} \left\| W_T \pm \frac{\sqrt{2\alpha-1}}{\alpha} I^\alpha \right\| = \frac{(\pi^2/32)^{\alpha/(4\alpha-1)} (2\alpha-1)^{(2\alpha-1)/(8\alpha-2)}}{\alpha^{(2\alpha-1)/(4\alpha-1)} (1-\alpha)^{(1-\alpha)/(4\alpha-1)}} \quad p.s.$$

La puissance  $2\alpha/(4\alpha-1)$  interpole entre  $2/3$  et  $1$ . Si  $\alpha \rightarrow 1/2$  la constante est de limite  $\pi/4$ , en accord avec la loi de Chung pour  $f = 0$ . Et dans ce cas  $f_1$  et  $W$  dépendent leur énergie en  $0$  puis stationnent dans une bande  $[-\varepsilon, \varepsilon]$ . Même commentaire pour  $f_3$  si  $f'_{3\pm} = 0$  mais avec une constante limite plus petite que  $\pi/4$ . A l'inverse, si  $\alpha \rightarrow 1$  la constante limite converge vers  $(\pi^2/32)^{1/3}$  ce qui correspond bien à la vitesse exacte pour  $f = I$  obtenue dans [Cs2].

## 3.2 – Petites déviations browniennes en norme sup pondérée

### 3.2.1 – Probabilité de petite boule pondérée

(i) Le problème des grandes déviations d'un processus  $\mathbb{X}$  concerne l'évaluation de  $\log \mathbb{P}(\|\mathbb{X}\| > \gamma)$  lorsque  $\gamma \rightarrow \infty$  et celui de ses petites déviations l'évaluation de  $\log \mathbb{P}(\|\mathbb{X}\| < \varepsilon)$  lorsque  $\varepsilon \rightarrow 0$ . Dans le premier cas, la métrique  $\|\cdot\|$  n'a que peu d'insidence sur la nature du problème car il s'agit principalement de maîtriser la queue de distribution de  $\mathbb{X}$  loin de sa trajectoire moyenne et une théorie unifiée a pu se développer [DS]. A l'inverse, dans le cas de la probabilité d'une petite boule centrée  $\{\|x\| < \varepsilon\}$  il est clair que beaucoup dépend de la topologie et aucune théorie générale n'a pu être développée. Même dans le cas des processus gaussiens c'est sans espoir, comme le montre [KL2,LLi].

Pour une suite de processus  $\mathbb{X}_t$  indicée par un temps  $t$  et  $\mathbb{X}_t$  avec une propriété de changement d'espace-temps vers  $\mathbb{X}$ , les deux problèmes deviennent respectivement l'évaluation de

$$\log \mathbb{P}(\mathbb{X}_t \in \gamma_t A)$$

avec  $\mathbb{E}(\mathbb{X}_t) \notin \gamma_t A$ ,  $\gamma_t \rightarrow \infty$  et celle de

$$\log \mathbb{P}(\|\mathbb{X}_t - \gamma_t x\| < \varepsilon_t)$$

avec  $\gamma_t x \notin \mathbb{E}(\mathbb{X}_t)$ ,  $\varepsilon_t \rightarrow 0$ . Le second type de probabilité est dite de petites boules translattées et se ramène souvent à l'évaluation des petites boules centrées. La différence de contexte est patente car dans un cas nous regardons typiquement la probabilité d'atteindre un demi espace qui se translate vers l'infini (grande déviation) et dans l'autre, celle d'atteindre une petite boule qui se translate vers l'infini mais dont le rayon tend vers  $0$  (petite déviation). Lorsque le demi-espace se recourbe plus progressivement, c'est le champ des déviations modérées.

(ii) Soit  $\{W(t) : t \in [0, 1]\}$  un mouvement brownien réel standard. Nous avons

$$\log \mathbb{P}(\|W\| < \varepsilon) \sim -\frac{c}{\varepsilon^b}, \quad \varepsilon \rightarrow 0$$

avec  $b = 2$  et  $c = \pi^2/8$  en norme sup [Ch],  $b = 2$  et  $c = 1/8$  en norme  $\mathbb{L}_2$  [CM],  $b = 2/(1 - 2\alpha)$  et  $c = C_\alpha$  de 2.2 en norme Höldérienne d'indice  $\alpha < 1/2$  [BR],  $b = 2$  et  $c = c_p$  en norme  $\mathbb{L}_p$  d'indice  $p \geq 1$  [DV],  $b$  contrôlée et  $c$  inconnue en norme de type Bésou ou Sobolev [Sz].

Dans [4] nous recherchons  $b$  et  $c$  en métrique uniforme pondérée. Ceci a un lien avec l'évaluation du temps local de certains processus, voir [DV,Shi], nous permet d'envisager des lois fonctionnelles précises dans cette topologie forte, à l'image de 2.2 dans le cadre Höldérien, et s'avère utile dans l'étude de certains théorèmes limites en probabilité et statistique [CH]. Pour  $\varphi \geq 0$  notons

$$\|W\|_\varphi = \sup_{[0,1]} \frac{|W|}{\varphi}.$$

Les deux résultats [Mo,Shi] antérieurs à [4] s'énoncent respectivement comme suit.

**Proposition 3B.** *Si  $\varphi$  est continue, linéaire par morceaux et  $\inf_{[0,1]} \varphi > 0$  alors*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P}(\|W\|_\varphi < \varepsilon) = -\frac{\pi^2}{8} \int_{[0,1]} \frac{1}{\varphi^2}.$$

**Proposition 3C.** *Si  $\varphi = I^\alpha$  avec  $\alpha < 1/2$  alors*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P}(\|W\|_\varphi < \varepsilon) = -\frac{\pi^2}{8(1 - 2\alpha)}.$$

Dans ces deux situations,  $b = 2$  et  $c = \pi^2 \int_{[0,1]} 1/8\varphi^2$ . Est-ce le cas général, quand  $\inf_{[0,1]} \varphi = 0$  ?

(iii) Pour que  $\|W\|_\varphi$  soit définie il faut d'une part que pour tout  $\varepsilon > 0$ ,

$$\inf_{[\varepsilon,1]} \varphi > 0$$

et d'autre part, que

$$\limsup_{t \rightarrow 0} \frac{|W(t)|}{\varphi(t)} < \infty \quad p.s.$$

Pour les fonctions  $\varphi$  croissantes au voisinage de 0 cette dernière condition, par un test intégral bien connu, voir [CH], est équivalente au fait que pour un  $\rho > 0$

$$\int_{[0,1]} \frac{1}{t} \exp\left(-\rho \frac{\varphi^2(t)}{t}\right) dt < \infty.$$

Le cas  $\varphi(t) = \gamma \sqrt{t \log |\log t|}$  est la frontière de validité naturelle du fait même de la loi du logarithme itéré [Lé]. Dans la suite nous ne considérons que des fonctions de poids vérifiant ces deux conditions.

(iv) Le résultat principal de [4] généralise les Propositions 3B et 3C en supprimant la contrainte de continuité par morceaux et surtout en caractérisant les fonctions de poids pour lesquelles  $b = 2$ .

**Théorème 3.5.** *Soit  $\varphi \geq 0$ . Si  $\inf_{[0,1]} \varphi > 0$  ou si  $\varphi$  est croissante au voisinage de 0 alors*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P}(\|W\|_\varphi < \varepsilon) = -\frac{\pi^2}{8} \int_{[0,1]} \frac{1}{\varphi^2}.$$

Nous ne savons pas dans quelle mesure la monotonie de  $\varphi$  en 0 est une nécessité mais nous l'imposons pour vérifier le test intégral, dont la validité en dehors du cas croissant reste un problème ouvert.

En conséquence, dans le problème de petite boule pondérée, nous avons  $b = 2$  si, et seulement si,

$$\int_{[0,1]} \frac{1}{\varphi^2} < \infty.$$

C'est le cas de  $\varphi(t) = \sqrt{t}|\log t|^\beta$  pour  $\beta > 1/2$ , ce qui prolonge la Proposition 3C à l'ordre  $\alpha = 1/2$ . Cependant, des cas critiques subsistent, tels que  $\varphi(t) = \gamma\sqrt{t\log|\log t|}$ , pour lesquels  $b \neq 2$  ou bien la log-probabilité n'est pas polynomiale. Notre preuve donne en fait l'évaluation incomplète suivante.

**Proposition 3.6.** *Soit  $\varphi \geq 0$  croissante au voisinage de 0. Si  $\int_{[0,1]} 1/\varphi^2 = \infty$  alors*

$$\limsup_{\varepsilon \rightarrow 0} \frac{\varepsilon^2}{\mu_\varphi(\varepsilon)} \log \mathbb{P}(\|W\|_\varphi < \varepsilon) \leq -\frac{\pi^2}{8}$$

où

$$\mu_\varphi(\varepsilon) = \int_{[\varepsilon,1]} \frac{1}{\varphi^2}.$$

Si de plus  $\varphi(t) > \rho\sqrt{t}$  nos arguments montrent que  $\mu_\varphi(\varepsilon)$  peut être remplacé par  $\int_{[\varepsilon^2,1]} 1/\varphi^2$ . En particulier la Proposition 3.6 donne

$$\limsup_{\varepsilon \rightarrow 0} \frac{\varepsilon^2 \log |\log \varepsilon|}{|\log \varepsilon|} \log \mathbb{P} \left( \sup_{[0,1]} \frac{|W|}{\sqrt{I \log |\log I|}} < \varepsilon \right) < 0.$$

Certains lim sup de 3.6 laissent un écart avec les lim inf. Par exemple, en  $\varphi(t) = \sqrt{t|\log t|}$  nous obtenons

$$\limsup_{\varepsilon \rightarrow 0} \frac{\varepsilon^2}{\log |\log \varepsilon|} \log \mathbb{P} \left( \sup_{[0,1]} \frac{|W|}{\sqrt{I \log I|}} < \varepsilon \right) \leq -\frac{\pi^2}{8}$$

mais, en retournant nos arguments,

$$\liminf_{\varepsilon \rightarrow 0} \frac{\varepsilon^2}{|\log \varepsilon|} \log \mathbb{P} \left( \sup_{[0,1]} \frac{|W|}{\sqrt{I \log I|}} < \varepsilon \right) > -\infty.$$

(v) Dans [L] notre technique est optimisée pour résoudre en partie les cas critiques.

**Théorème 3D.** *Pour  $\varphi \geq 0$  croissante sur  $[0, a]$  soit  $N(\varepsilon) = \max(n : t_n > 0)$  où  $t_{n+1} = t_n - \varepsilon^2 \varphi^2(t_n)$ . S'il existe  $c_i > 0$  tels que  $c_1 N(\varepsilon) \leq N(\varepsilon/2) \leq c_2 N(\varepsilon)$  alors il existe des constantes  $C_i > 0$  telles que  $-C_2 N(\varepsilon) \leq \log \mathbb{P}(\sup_{[0,a]} |W|/\varphi < \varepsilon) \leq -C_1 N(\varepsilon)$ .*

Ainsi, dans le cas  $\varphi(t) = \sqrt{t|\log t|}$  ci-dessus, la vitesse est bien de l'ordre de  $|\log \varepsilon|/\varepsilon^2$ .

(vi) **Principe de démonstration.** Pour la partie facile, la borne supérieure, nous découpons le temps en blocs  $]t_{k,n}, t_{k+1,n}]$  et bornons la probabilité de l'intersection des événements par un produit de probabilités conditionnelles, chaque événement conditionnel successif faisant apparaître une borne en  $\varphi(t_{k,n})$  puis l'évaluation de petite boule uniforme s'appliquant, nous majorons ce produit par une borne exponentielle avec la série des  $(t_{k+1,n} - t_{k,n})/\varphi^2(t_{k,n})$  qui reconstruit l'intégrale.

La borne inférieure est bien plus délicate et révèle la contrainte de gaussianité. Dès l'étape initiale, qui concerne le cas  $\inf_{[0,1]} \varphi > 0$  sans hypothèse de continuité, l'argument de départ crucial est la minoration donnée par [Kh] qui affirme que pour des vecteurs gaussiens  $Z_1, \dots, Z_n$  et des réels  $a_k > 0$ ,

$$\mathbb{P} \left( \bigcap_{k=1}^n \{|Z_k| < a_k\} \right) \geq \prod_{k=1}^n \mathbb{P}(|Z_k| < a_k). \quad \square$$

### 3.2.2 – Applications et constantes explicites

(i) Soit  $\{M(t) : t \in [0, 1]\}$  un processus de Markov gaussien réel centré. Si  $\mathbb{E}(M(s)M(t)) \neq 0$  pour  $0 < s \leq t < 1$  alors il existe deux fonctions  $L > 0$  et  $H > 0$  telles que

$$\mathbb{E}(M(s)M(t)) = L(s)H(t), \quad 0 < s \leq t < 1$$

et  $L/H$  est croissante sur  $]0, 1[$ . Ces fonctions sont uniques à une constante multiplicative près.

**Corollaire 3.7.** *Supposons que  $L$  et  $H$  sont absolument continues et  $L/H$  strictement croissante. Si  $\sup_{[0,1]} H > 0$  ou si  $H$  est croissante au voisinage de 0 alors*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P}(\|M\| < \varepsilon) = -\frac{\pi^2}{8} \int_{[0,1]} (L'H - LH').$$

Pour obtenir ce résultat il suffit d'écrire, comme dans [L], la preuve de [4] en remplaçant  $W(t)$  par  $M(t) = H(t)W(L(t)/H(t))$  et de remarquer que la norme  $\sup \|M\|$  de  $M$  est la norme  $\sup$  pondérée  $\|W \circ (L/H)\|_{1/H}$  de  $W$  changé de temps par la transformation croissante  $L/H$ .

**Exemple.** Un processus gaussien à accroissements indépendants a la constante  $c = \pi^2(L(1) - L(0))/8$ .

**Exemple.** Pour un brownien changé d'échelle,  $M(t) = W(t)/t^\alpha$  ou  $M(t) = B(t)/t^\alpha$  nous avons les covariances  $s^{1-\alpha}t^{-\alpha}$  et  $s^{1-\alpha}t^{-\alpha}(1-t)$  pour  $s < t$  donc la même constante  $c = \pi^2/8(1 - 2\alpha)$ .

**Exemple.** Pour un processus de Markov gaussien stationnaire ou le processus d'Ornstein-Uhlenbeck la covariance est  $\sigma^2 \exp(-\gamma t + \gamma s)$  et la constante devient  $c = \pi^2 \sigma^2 \gamma / 4$ .

(ii) Du Théorème 3.5 nous tirons classiquement, par la formule de translation [CM] et les arguments de [DM2], la loi fonctionnelle de type Chung pour  $W$  en temps long.

**Corollaire 3.8.** *Si  $\inf_{[0,1]} \varphi > 0$  ou si  $\varphi$  est croissante au voisinage de 0 et  $\int_{[0,1]} 1/\varphi^2 < \infty$  alors pour tout  $f \in \mathcal{S}_1$*

$$\liminf_{T \rightarrow \infty} (\log \log T) \left\| \frac{W(TI)}{\sqrt{2T \log \log T}} - f \right\|_\varphi = \frac{\pi}{4} \frac{\sqrt{\int_{[0,1]} 1/\varphi^2}}{\sqrt{1 - J(f)}} \quad p.s.$$

Le Théorème 3.5 est l'une des clefs de l'extension à  $\|\cdot\|_\varphi$  du Théorème 2.8 sur les incréments browniens  $\Lambda_T$  et des résultats de 2.2 sur les incréments empiriques  $\Theta_n$ , en rapport avec [Ja, CCHM, EjM, CH]. La seconde clef, dont nous avons commencé l'étude, est une nouvelle borne exponentielle de type grandes déviations contrôlant sous  $\|\cdot\|_\varphi$  le module de continuité de ces trajectoires.

(iii) Soit  $\{W_d(t) : t \in [0, 1]\}$  un mouvement brownien standard sur  $\mathbb{R}^d$  muni de sa norme euclidienne  $\|\cdot\|_{\mathbb{R}^d}$ ,  $d \geq 1$ . L'analogie du Théorème 3.5 se prouve de la même façon. Noter que  $j_{-1/2} = \pi/2$ .

**Théorème 3.9.** *Si  $\inf_{[0,1]} \varphi > 0$  ou si  $\varphi$  est croissante au voisinage de 0 alors*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P} \left( \sup_{[0,1]} \frac{\|W_d\|_{\mathbb{R}^d}}{\varphi} < \varepsilon \right) = -\frac{j_{(d-2)/2}^2}{2} \int_{[0,1]} \frac{1}{\varphi^2}$$

avec  $j_{(d-2)/2}^2$  la plus petite racine positive de la fonction de Bessel  $J_{(d-2)/2}$ .

(iv) Considérons un pont Brownien  $\{B_d(t) : t \in [0, 1]\}$  sur  $\mathbb{R}^d$ , qui est l'approximation naturelle du processus empirique uniforme multivarié  $\alpha_n$  construit à partir des  $d$  composantes d'un échantillon de vecteurs aléatoires de  $\mathbb{R}^d$ . Une méthode standard permet de transformer le Théorème 3.9 comme suit.

**Corollaire 3.10.** *Si  $\inf_{[s,t]} \varphi > 0$  pour tout  $0 < s \leq t < 1$  et si  $\varphi$  et  $\varphi \circ (1 - I)$  sont croissantes au voisinage de 0 alors*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P} \left( \sup_{[0,1]} \frac{\|B_d\|_{\mathbb{R}^d}}{\varphi} < \varepsilon \right) = -\frac{j_{(d-2)/2}^2}{2} \int_{[0,1]} \frac{1}{\varphi^2}$$

avec  $j_{(d-2)/2}^2$  la plus petite racine positive de la fonction de Bessel  $J_{(d-2)/2}$ .

Ce résultat s'applique donc aux processus empiriques pondérés via [CH] et 3.4.5(iv).

(v) Une autre conséquence importante du Théorème 3.5 concerne les probabilités de franchissement de frontière, dans la lignée de [Mo]. Il permet d'évaluer la mesure de Wiener de petits ensembles de



trajectoires qui ne sont pas des boules, comme des faisceaux issus de 0. Soit  $\varphi_1 \leq \varphi_2$  deux fonctions réelles mesurables définies sur  $[0, 1]$ . Considérons l'événement

$$\Xi_\varepsilon = \{\varepsilon\varphi_1(t) \leq W(t) \leq \varepsilon\varphi_2(t) : t \in [0, 1]\}$$

quand  $\varepsilon \rightarrow 0$  et autorisons le cas délicat  $\varphi_1(0) = \varphi_2(0) = 0$ , contrairement à [Mo]. Puisque

$$\Xi_\varepsilon = \left\{ \|W - \varepsilon f\|_\varphi < \varepsilon \right\}$$

avec  $f = (\varphi_1 + \varphi_2)/2$  et  $\varphi = (\varphi_2 - \varphi_1)/2$  nous pouvons combiner le Théorème 3.5 à la formule de [CM] pour obtenir la proposition suivante.

**Proposition 3.11.** *Supposons que  $\inf_{[0,1]}(\varphi_2 - \varphi_1) > 0$  ou bien que  $\varphi_2 - \varphi_1 \geq 0$  est croissante au voisinage de 0. Si de plus  $\varphi_1 + \varphi_2$  est absolument continue et de dérivée de Lebesgue dans  $\mathbb{L}_2$  alors*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P}(\Xi_\varepsilon) = -\frac{\pi^2}{2} \int_{[0,1]} \frac{1}{(\varphi_2 - \varphi_1)^2}.$$

Le cas  $\varphi_1 = -\varphi_2$  fonctionne toujours, en convenant que  $\log(0) = -\infty$ . Remarquons l'impossibilité d'avoir  $\varphi_1 \geq 0$  ou  $\varphi_2 \leq 0$  pour une limite finie. En effet,  $\varphi_2 - \varphi_1$  doit être rapide en 0 et  $\varphi_2 + \varphi_1$  lent, et cette double contrainte entraîne que  $\min(-\varphi_1, \varphi_2) \gg \sqrt{T \log |\log T|}$  en 0. Nous utilisons la Proposition 3.11 dans un passage technique de la preuve du Théorème 3.1 ([6]).

### 3.3 – Grandes déviations de l'argmin

#### 3.3.1 – Brownien avec dérive polynomiale

(i) Soit  $\{W(t) : t \in \mathbb{R}\}$  un mouvement brownien réel issu de  $W(0) = 0$  et  $d > 0$ . Dans [Gn] la densité de  $Y = \arg \max_{t \in \mathbb{R}} \{W(t) - dt^2\}$  est calculée et équivaut en  $|y| \rightarrow \infty$  à

$$f(y) \sim \frac{1}{2a} (4d)^{4/3} |y| \exp\left(-\frac{2}{3} d^2 |y|^3 (1 + o(1))\right)$$

où  $a \sim 0.7022$  est relié aux "Airy functions". D'où la queue de distribution

$$C_1 \exp\left(-\frac{2}{3} d^2 T^3 (1 + \varepsilon_T)\right) \leq \mathbb{P}\left(\left|\arg \max_{t \in \mathbb{R}} \{W(t) - dt^2\}\right| > T\right) = C_1 \exp\left(-\frac{2}{3} d^2 T^3 (1 - \varepsilon_T)\right)$$

avec  $\varepsilon_T \rightarrow 0$  si  $T \rightarrow \infty$ . Cette estimation très précise implique le résultat de grande déviation suivant.

**Proposition 3E.** *Pour tout  $d > 0$ ,*

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{1}{T^3} \log \mathbb{P}\left(\left|\arg \max_{t \in \mathbb{R}} \{W(t) - dt^2\}\right| > T\right) \\ &= \lim_{T \rightarrow \infty} \frac{1}{T^3} \log \mathbb{P}\left(\left|\arg \min_{t \in \mathbb{R}} \{W(t) + dt^2\}\right| > T\right) = -\frac{2}{3} d^2. \end{aligned}$$

(ii) L'étude de [Gn] est très complète mais propre à la dérive quadratique. Or, dans [10,15], l'argument final repose sur un contrôle de l'arg min d'un processus empirique avec dérive, obtenu comme différence de paires d'incrémentes locaux du processus de quantile et approximé par

$$\arg \min_{|t| < T_n} \{W_n(t) + d_n(\omega, t) |t|^\alpha + \varepsilon_n(\omega, t)\}$$

avec  $W_n$  un processus distribué comme un mouvement brownien et deux suites de v.a.  $d_n(\omega, t)$  et  $\varepsilon_n(\omega)$  telles que  $\sup_{|t| < T_n} |d_n(\omega, t) - d| = o_{p.s.}(1)$  et  $\sup_{|t| < T_n} |\varepsilon_n(\omega, t)| = O_{p.s.}(\varepsilon_n)$  où  $T_n \rightarrow \infty$  et  $\varepsilon_n \rightarrow 0$  sont deux suites déterministes. D'où le besoin de généraliser la Proposition 3E au cas  $\alpha \neq 2$ .

(iii) Dans [10] nous mettons au point une méthode d'évaluation des grandes déviations de

$$\arg \min_{t \in \mathbb{R}} \{W(t) + d_n |t|^\alpha + \varepsilon_n\}$$

avec  $d_n \rightarrow d > 0$  et  $\varepsilon_n \rightarrow 0$ . Nous ne cherchons pas à calculer la loi jointe du min et de l'arg min comme dans [Gn], qui tire profit des spécificités du cas parabolique. Nous visons plutôt la classique évaluation logarithmique, par une méthode probabiliste originale et généralisable.

**Théorème 3.12.** *Soit  $\alpha > 1$  et  $d_i$  des fonctions réelles déterministes telles que  $d_i(T) \rightarrow d > 0$  si  $T \rightarrow \infty$ . Alors, si  $T \rightarrow \infty$  il vient*

$$\mathbb{P} \left( \min_{|t| \leq T} (W(t) + d_1(T) |t|^\alpha) \geq \min_{|t| > T} (W(t) + d_2(T) |t|^\alpha) + o(T^\alpha) \right) = \exp \left( -\frac{d^2 \alpha^2 (1 + o(1))}{2(2\alpha - 1)} T^{2\alpha - 1} \right).$$

Nous en déduisons le résultat de grandes déviations suivant, qui contient bien la Proposition 3E.

**Corollaire 3.13.** *Si  $\alpha > 1$  et  $d > 0$  alors*

$$\lim_{T \rightarrow \infty} \frac{1}{T^{2\alpha - 1}} \log \mathbb{P} \left( \arg \min_{t \in \mathbb{R}} (W(t) + d |t|^\alpha) \notin [-T, T] \right) = -\frac{\alpha^2 d^2}{2(2\alpha - 1)}.$$

Ce principe pour  $W$  se généralise aux processus réels à accroissements indépendants pour lesquels les ingrédients suivants sont disponibles. D'où un résultat analogue pour les marches aléatoires.

(iv) **Principe de démonstration.** La méthode de preuve inventée dans [10] est technique mais fort instructive quant au traitement des arg min de processus. Elle repose sur une manipulation astucieuse des principes de grande et petite déviations browniennes bien connus.

Pour la minoration nous conditionnons par un passage de  $W$  le long d'une petite boule longeant  $d |t|^\alpha$  et appliquons [Ac2]. C'est la situation de déviation de l'arg min la plus probable au vu de ce qui suit.

Pour la majoration nous conditionnons par les faisceaux de trajectoires possibles en fixant une grille de points de passage  $W(T)$  et de points de dépassement pour comparer les min sur  $\{W(t) : t \in ]0, T[ \}$ ,  $\{W(t) : t \in ]T, \eta(T)[ \}$  et  $\{W(t) : t > \eta(T)\}$  avec  $\eta(T)$  convenable. Le choix des paramètres est fait de sorte que chaque sous-événement, après changement d'échelle et de temps dans le faisceau, soit contrôlé par des probabilités de type  $\mathbb{P}(W_1 \notin v(T)A)$  avec  $v(T) \rightarrow \infty$  et  $A \subset \mathcal{B}$  un ensemble de trajectoires dont nous savons calculer le point dominant donc  $J(A) = \inf_A J$ . En appliquant les bornes usuelles de grandes déviations [DS] à un nombre fini de tels événements, les moins probables sont éliminés.

Enfin, le produit de déconditionnement des événements rares les plus probables révèle la plus grande borne exponentielle en faisant apparaître une équation non linéaire qu'il faut miniser pour majorer le tout. Au final, les constantes issues des petites et grandes déviations, la minoration simple et la majoration complexe, coïncident exactement.  $\square$

### 3.3.2 – Brownien avec dérive régulière

(i) Avec une dérive en  $|t|^\alpha$ ,  $\alpha \in ]1/2, 1]$ , la trajectoire type de  $W$  assurant la grande déviation est linéaire et non plus puissance. Nous ne savons pas traiter le cas  $\alpha \leq 1/2$  car le problème revient à décrire ce qui se passe dans une petite boule brownienne et change totalement de nature. En adaptant la preuve du Théorème 3.12 nous obtenons un résultat utile pour le travail en cours 4.4.1.

**Proposition 3.14.** *Si  $\alpha \in ]1/2, 1]$  et  $d > 0$  alors*

$$\lim_{T \rightarrow \infty} \frac{1}{T^{2\alpha - 1}} \log \mathbb{P} \left( \arg \min_{t \in \mathbb{R}} (W(t) + d |t|^\alpha) \notin ]-T, T[ \right) = -\frac{d^2}{2}.$$

(ii) Perturbons un peu le drift. Soit  $\varphi(t)$  une fonction à variation régulière d'ordre  $\alpha$  en  $+\infty$  et  $-\infty$ , au sens où  $\lim_{|t| \rightarrow \infty} \varphi(\theta t)/\varphi(t) = \theta^\alpha$  pour tout  $\theta > 0$ . Nos arguments montrent que si  $\alpha$  diffère en  $+\infty$  et  $-\infty$  seule la plus petite valeur compte. D'après [Sé],  $\varphi(t) = t^\alpha \exp(\eta(t) + \int_{(m, \infty)} \chi(t)/t)$  où  $\eta(t)$  est

mesurable bornée,  $\eta(t) \rightarrow \eta$  si  $t \rightarrow \infty$  et  $\chi(t)$  est  $\mathcal{C}_k$  pourvu que  $m > m_k$  soit assez grand. En outre la convergence vers  $\theta^\alpha$  est uniforme en  $\theta \in [a, b]$ , ce qui s'avère essentiel dans nos changements de temps. La généralisation suivante vise à étendre [15] et 4.3 dans les directions décrites en 4.4.1, et sera publiée à cette occasion avec la proposition 3.14 en corollaire.

**Théorème 3.15.** *Si  $\varphi$  est telle que  $\lim_{|t| \rightarrow \infty} \varphi(\theta t)/\varphi(t) = \theta^\alpha$  pour tout  $\theta > 0$ , avec  $\alpha > 1/2$ , alors*

$$\lim_{T \rightarrow \infty} \frac{T}{\varphi^2(T)} \log \mathbb{P} \left( \arg \min_{t \in \mathbb{R}} (W(t) + \varphi(t)) \notin [-T, T] \right) = -\frac{1}{2} \left( \mathbb{I}_{\alpha \leq 1} + \frac{\alpha^2}{2\alpha - 1} \mathbb{I}_{\alpha > 1} \right).$$

A la lumière de [LS] l'étude des temps lents de sortie d'un mouvement brownien plan d'un domaine  $\{(x, y) : y > g(x)\}$ , notamment parabolique, s'appuie sur la compétition entre  $W(t)$  et une fonction  $\varphi(t)$ . Dans ce contexte la méthode de preuve du Théorème 3.15 peut aider.

**(iii) Principe de démonstration.** Les propriétés de  $\varphi$  rappelées ci-dessus permettent d'ajuster les découpages et les changements d'échelle opérés dans la preuve par conditionnements exhaustifs de la borne supérieure du Théorème 3.12. Finalement, l'équation à minimiser pour identifier le plus probable faisceau réalisant la grande déviation absorbe la perturbation. Il y a cependant une difficulté supplémentaire à régler au départ des trajectoire lorsque  $\varphi$  n'est ni convexe ni concave.  $\square$

### 3.4 – Approximation forte du processus empirique indexé par des fonctions

#### 3.4.1 – Problème de l'approximation gaussienne

Soit  $n \geq 1$  et  $X, X_1, \dots, X_n$  des v.a. i.i.d. définies sur l'espace de probabilité  $(\Omega, \mathcal{T}, \mathbb{P})$  à valeurs dans un espace mesurable  $(\mathcal{X}, \mathcal{A})$ . Notons  $\mathbb{E}$  l'espérance sous  $\mathbb{P}$  des v.a. réelles construites sur  $(\Omega, \mathcal{T})$ ,  $P = \mathbb{P}^X$  et  $\mathcal{M}$  l'ensemble des fonctions mesurables de  $(\mathcal{X}, \mathcal{A})$  dans  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , où  $\mathcal{B}(\mathbb{R})$  est la tribu borélienne. Soit  $\mathcal{F} \subset \mathcal{M}$  une classe de fonctions. Le  $P$ -processus empirique indexé par  $\mathcal{F}$  est

$$\alpha_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{f(X_i) - \mathbb{E}(f(X))\}, \quad f \in \mathcal{F}.$$

Le  $P$ -pont brownien  $\mathbb{G}$  indexé par  $\mathcal{F}$  est le processus gaussien centré de même covariance que  $\alpha_n$ ,

$$\langle f, h \rangle = \text{cov}(\mathbb{G}(f), \mathbb{G}(h)) = \mathbb{E}(f(X)h(X)) - \mathbb{E}(f(X))\mathbb{E}(h(X)), \quad (f, g) \in \mathcal{F}^2.$$

Sous nos hypothèses  $\mathbb{G}$  admet une version séparable, presque sûrement continue pour la semi-métrique

$$d_P(f, h) = \sqrt{\mathbb{E}(f(X) - h(X))^2}, \quad (f, g) \in \mathcal{F}^2$$

donc nous incluons la  $d_P$ -continuité dans la définition de  $\mathbb{G}$ .

Le principe d'invariance faible stipule, pour certains  $\mathcal{F}$  et  $P$ , que  $\alpha_n$  converge en loi vers  $\mathbb{G}$  dans  $\mathbb{L}_\infty(\mathcal{F})$ , au sens des mesures extérieures. Une telle classe  $\mathcal{F}$  est dite  $P$ -Donsker [Po,Du,Du3,V,VW].

**(i)** Le problème de l'approximation gaussienne est de construire une version de  $X_1, \dots, X_n$  et  $\mathbb{G}$  sur  $(\Omega, \mathcal{T}, \mathbb{P})$  supposé assez riche, de telle sorte qu'avec grande probabilité

$$\|\alpha_n - \mathbb{G}\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |\alpha_n(f) - \mathbb{G}(f)|$$

soit le plus petit possible. Un tel principe d'invariance, dit fort, induit le principe d'invariance faible et en précise la vitesse. Réaliser cela en toute généralité est très utile en vue des applications.

**Application.** Pour une statistique de la forme  $Y_n = \Gamma(\alpha_n)$  avec  $\Gamma$  de  $\mathbb{L}_\infty(\mathcal{F})$  dans  $\mathbb{R}$  supposée régulière ou contrôlable, l'approximation forte permet d'encadrer  $\mathbb{P}(Y_n \in A)$  par  $\mathbb{P}(\mathbb{G} \in \Gamma^{-1}(A)) \pm v_n$ ,  $v_n \rightarrow 0$ , ce qui est valable  $n$  par  $n$  donc peut s'avérer bien plus satisfaisant que la convergence en loi.

Le meilleur couplage connu de ce type est celui de la construction hongroise [KMT,CR3,BM,CL,MZ], une technique très fine valable sur  $\mathcal{X} = \mathbb{R}$  et  $\mathcal{F} = \{\mathbb{I}_{[0,t]} : t \in [0, 1]\}$ . Le raffinement et l'exploitation

fructueuse de cette approche ont fait l'objet de nombreux travaux depuis trente ans. Les extensions concernent le passage à  $\mathbb{R}^d$ , qui reste toujours sans vitesse optimale, le passage à la dépendance faible, et la recherche des structures de  $\mathcal{F}$  permettant d'utiliser [KMT] via des bases de fonctions.

(ii) Il y a deux manières d'attaquer le problème de l'approximation gaussienne. D'une part, pour obtenir des vitesses optimales, ou du moins assez pertinentes, il convient de bien spécifier la géométrie de  $\mathcal{F}$  par rapport à  $P$ , c'est l'usage courant. D'autre part, pour obtenir des vitesses en toute généralité il convient de spécifier uniquement la taille de  $\mathcal{F}$ , qui seule régule le principe d'invariance faible.

La seconde direction, peu explorée depuis [DP], fait l'objet de [11] et du travail [25] en cours d'écriture. Nos vitesses et nos conditions générales sur  $\mathcal{F}$  et  $P$  sont données en 3.4.2 et 3.4.3, qui sont des extensions de [11], puis commentées en 3.4.5. Le résultat principal de [11] figure en 3.4.4 et répond à l'exigence plus forte, comme dans [DP], d'approcher  $\alpha_1, \dots, \alpha_n$  par les sommes partielles d'une suite de processus  $\mathbb{G}$  indépendants. En 3.4.6 nous contribuons à la première optique par des nouvelles vitesses.

Dans [11,25] nous exploitons, entre autres outils, un couplage dérivé de [Z,Z2], méconnu mais puissant.

(iii) Par commodité nous évitons l'usage des mesures extérieures en rendant les normes sup mesurables par l'hypothèse qu'il existe une sous-classe dénombrable  $\mathcal{F}_\infty \subset \mathcal{F}$  telle qu'à tout  $f \in \mathcal{F}$  correspond une suite  $\{f_m\} \in \mathcal{F}_\infty$  telle que  $\lim_{m \rightarrow \infty} f_m(x) = f(x)$  pour tout  $x \in \mathcal{X}$ .

Pour des raisons bien comprises, notamment l'existence de certaines intégrales clef, il est nécessaire d'imposer que  $\|f\|_{\mathcal{X}} = \sup_{x \in \mathcal{X}} |f(x)| < M$  pour tout  $f \in \mathcal{F}$ .

Pour traiter des classes non uniformément bornées, notre point de vue devient asymptotique et nous imposons que  $\mathcal{F}$  admette une enveloppe  $F \geq \sup_{f \in \mathcal{F}} |f|$ ,  $F \in \mathbb{L}_2(P)$ . Nos résultats concernent alors  $\mathcal{F}_n = \{f \in \mathcal{F} : \|f\|_{\mathcal{X}} < M_n\}$  avec  $M_n \rightarrow \infty$ . Ceci donne accès à l'approximation en métrique forte tronquée, voir 3.4.5, et prend en compte des classes de type  $\mathcal{F} = \bigcup_{n \geq 1} \mathcal{G}_n$ . Dans le cas borné,  $F = M$ .

### 3.4.2 – Vitesse sous condition d'entropie uniforme

(i) Considérons des classes  $\mathcal{F}$  dites polynomiales, en référence à [VC,Po]. A toute probabilité  $Q$  sur  $(\mathcal{X}, \mathcal{A})$  associons  $d_Q^2(f, h) = \int_{\mathcal{X}} (f - h)^2 dQ$  et  $Q(F^2) = \int_{\mathcal{X}} F^2 dQ$ , avec  $F$  l'enveloppe de  $\mathcal{F}$ . Soit  $N(\varepsilon, \mathcal{F}, d_Q)$  le nombre minimum de boules de  $d_Q$ -rayon  $\varepsilon > 0$  nécessaires au recouvrement de  $\mathcal{F}$ . Le nombre d'entropie uniforme de  $\mathcal{F}$  est

$$N_F(\varepsilon, \mathcal{F}) = \sup_Q N\left(\varepsilon \sqrt{Q(F^2)}, \mathcal{F}, d_Q\right)$$

avec un supremum pris parmi les probabilités  $Q$  sur  $(\mathcal{X}, \mathcal{A})$  telles que  $0 < Q(F^2) < \infty$ .

Faisons l'hypothèse que pour  $c_0 > 0$  et  $\nu_0 > 0$ ,

$$(VC) \quad N_F(\varepsilon, \mathcal{F}) \leq \frac{c_0}{\varepsilon^{\nu_0}}, \quad \varepsilon \in ]0, 1[.$$

Étant donné  $\mathcal{F}$ , nous montrons dans [11] que ceci garantit un principe d'invariance fort pour tout  $P$ . Cette uniformité apporte une information inédite, comme le montrent les exemples 3.4.5.

**Proposition 3.16.** *Sous (VC) il existe  $(\Omega, \mathcal{T}, \mathbb{P})$  sur lequel nous pouvons construire, pour tout  $\lambda > 1$  et un  $\rho_\lambda > 0$  associé, une version de  $\mathbb{G}$  et  $X_1, \dots, X_n$  telle que, pour  $\tau_0 = (4 + 5\nu_0)/(4 + 10\nu_0)$ ,*

$$\mathbb{P} \left\{ \|\alpha_n - \mathbb{G}\|_{\mathcal{F}} > \rho_\lambda \left(\frac{1}{n}\right)^{1/(2+5\nu_0)} (\log n)^{\tau_0} \right\} \leq \frac{1}{n^\lambda}.$$

Par Borel-Cantelli il s'ensuit l'existence d'une suite de  $P$ -ponts browniens  $\mathbb{G}_n$  indexés par  $\mathcal{F}$  telle que

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{p.s.} \left( \frac{(\log n)^{\tau_0}}{n^{1/(2+5\nu_0)}} \right).$$

Nous en déduisons la même vitesse pour la distance de Levy-Prohorov entre les lois de  $\alpha_n$  et  $\mathbb{G}_n$ . Par suite,  $\mathcal{F}$  est universellement  $P$ -Donsker et  $P$ -Glivenko-Cantelli, avec cette vitesse de convergence.

(ii) Évoquons une extension aux classes non bornées. Ici  $(\Omega, \mathcal{T}, \mathbb{P})$  est le produit des espaces de 3.16.

**Proposition 3.17.** *Sous (VC) et  $\lim_{n \rightarrow \infty} M_n \sqrt{(\log n)/n} = 0$  nous pouvons construire sur  $(\Omega, \mathcal{T}, \mathbb{P})$  une suite de v.a.  $X_n$  i.i.d. de loi  $P = \mathbb{P}^X$  et une suite  $\mathbb{G}_n$  de  $P$ -ponts browniens tels que*

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_n} = O_{\mathbb{P}} \left( \left( \frac{M_n^2}{n} \right)^{1/(2+5\nu_0)} (\log n)^{\tau_0} \right).$$

De plus  $O_{p.s.}$  peut remplacer  $O_{\mathbb{P}}$  si  $\limsup_{n \rightarrow \infty} (\log M_n)/\log n < 1/2$ .

(iii) Dans [25] nous étendons ces résultats à des classes de Donsker plus grandes ou plus petites, en utilisant ce que nous appelons l'intégrale de Dudley, voir [Du2]. Pour  $\phi > 1$  sur  $]0, 1[$  définissons

$$J_\phi(\varepsilon) = \int_{[0, \varepsilon]} \sqrt{\log \phi(s)} ds, \quad \Psi_\phi(\varepsilon) = \frac{\varepsilon}{\phi^{5/2}(\varepsilon)}, \quad \varepsilon \in ]0, 1[.$$

Supposons qu'il existe une fonction croissante  $\phi$  telle que  $J_\phi(1) < \infty$ ,  $\lim_{\varepsilon \rightarrow 0} \phi(\varepsilon) = \infty$  et

$$(VC') \quad N_F(\varepsilon, \mathcal{F}) \leq \phi(\varepsilon) < \infty, \quad \varepsilon \in ]0, 1[.$$

Puisque  $P(F^2) < \infty$ ,  $(VC')$  assure que  $\mathcal{F}$  est  $P$ -Donsker, voir [VW]. La fonction  $J_\phi \circ \Psi_\phi^{-1}$  décroît rapidement en 0 si  $\mathcal{F}$  a un petit nombre d'entropie  $N_F(\varepsilon, \mathcal{F})$ . Du point de vue asymptotique il suffit d'étudier  $J_\phi(\varepsilon) \sim \varepsilon \sqrt{\log \phi(\varepsilon)}$  mais pour  $n$  fini  $J_\phi$  apparaît elle-même dans le contrôle exact des probabilités de déviation de  $\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_n}$  dans notre preuve.

**Théorème 3.18.** *Sous  $(VC')$  et  $\limsup_{n \rightarrow \infty} (\log M_n)/\log n < 1/2$  il existe sur un espace  $(\Omega, \mathcal{T}, \mathbb{P})$  une suite de v.a.  $X_n$  i.i.d. de loi  $P = \mathbb{P}^X$  et une suite  $\mathbb{G}_n$  de  $P$ -ponts browniens tels que*

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_n} = O_{p.s.} \left( J_\phi \circ \Psi_\phi^{-1} \left( M_n \sqrt{\frac{\log n}{n}} \right) \right).$$

Nous apprenons ainsi que la vitesse de décroissance de l'intégrale de Dudley en 0 est intrinsèquement liée à la pire vitesse, donc à la pire loi  $P$ , dans le théorème de Donsker pour  $\mathcal{F}$ . Pour les plus grands  $\mathcal{F}$  autorisés la vitesse est arbitrairement lente et pour les plus petits  $\mathcal{F}$  elle s'approche de  $\sqrt{(\log n)/n}$ . Notons que plus  $\mathcal{F}$  est petite, plus la troncature  $M_n$  a un impact important sur la vitesse, voir 3.4.3.

Heuristiquement,  $J_\phi$  régule le module de continuité de  $\mathbb{G}$  et  $\alpha_n$  donc même une loi  $P$  défavorable ne suffit pas à les faire osciller exagérément.

(iv) **Principe de démonstration.** Nous combinons des outils d'analyse gaussienne [Bo, Du2, Du3, Lf], de théorie des processus empiriques [EM3, T2, VW] et le couplage de [Z] sur  $\mathbb{R}^d$  pour évaluer chaque terme de la décomposition

$$\begin{aligned} \mathbb{P}(\|\alpha_n - \mathbb{G}\|_{\mathcal{F}} > A\mu_n(\varepsilon) + \mu(\varepsilon) + \delta + (A+1)t) &\leq \mathbb{P} \left( \sup_{(f, f') \in \mathcal{F}(\varepsilon)} |\mathbb{G}(f) - \mathbb{G}(f')| > t + \mu(\varepsilon) \right) \\ &+ \mathbb{P} \left( \sup_{(f, f') \in \mathcal{F}(\varepsilon)} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n (f - f')(X_i) \right| > A\mu_n(\varepsilon) + At \right) + \mathbb{P} \left( \max_{h \in \mathcal{H}(\varepsilon)} |\alpha_n(h) - \mathbb{G}(h)| > \delta \right) \end{aligned}$$

avec  $\mathcal{F}(\varepsilon) = \{(f, f') \in \mathcal{F}^2 : d_P(f, f') < \varepsilon\}$  les incréments,  $\mathcal{H}(\varepsilon) = \{h_i : 1 \leq i \leq N(\varepsilon, \mathcal{F}, d_P)\}$  la grille de recouvrement et, pour des poids i.i.d.  $\varepsilon_i$  de Rademacher  $\mathbb{P}(\varepsilon_i = \pm 1) = 1/2$  indépendantes des  $X_i$ ,

$$\mu_n(\varepsilon) = \mathbb{E} \left( \frac{1}{\sqrt{n}} \sup_{(f, f') \in \mathcal{F}(\varepsilon)} \left| \sum_{i=1}^n \varepsilon_i (f - f')(X_i) \right| \right), \quad \mu(\varepsilon) = \mathbb{E} \left( \sup_{(f, f') \in \mathcal{F}(\varepsilon)} |\mathbb{G}(f) - \mathbb{G}(f')| \right).$$

Par exemple, le second terme est borné via [T2] donc nous devons contrôler le processus  $\alpha_n$  symétrisé en généralisant à (VC') la borne de moment prouvée dans [EM3] sous (VC). D'où le Lemme 3.19, prouvé en partie à l'aide de [JM,GZ], qui s'applique à  $\mathcal{G} = \mathcal{F}(\varepsilon)$ . Puis [BP] permet d'étendre  $\mathbb{G}$  à  $\mathcal{F}$  en aggrandissant le  $\Omega$  de [Z] de sorte que [Bo,Du2] soient vérifiés par le processus couplé  $\mathbb{G}$  ainsi reconstruit. La vitesse émerge en optimisant les bornes supérieures, par choix des paramètres dans les sommes d'exponentielles.  $\square$

**Lemme 3.19.** *Soit  $\mathcal{G}$  d'enveloppe  $G \in \mathbb{L}_2(P)$  et  $\gamma_3, \gamma_8, \sigma_0$  et  $\beta_0$  des constantes universelles. Si  $0 < \sigma \leq \sigma_0 < \beta_0 \leq \beta < \infty$  et  $\mathbb{E}(G^2(X)) \leq \beta^2$ ,  $N_G(\varepsilon, \mathcal{G}) \leq \phi(\varepsilon)$ ,  $\sup_{g \in \mathcal{G}} \mathbb{E}(g^2(X)) \leq \sigma^2$  et  $\sup_{g \in \mathcal{G}} \|g\|_{\mathcal{X}} \leq \gamma_8 \sigma \sqrt{n / \log \phi(\sigma / \gamma_3 \beta)}$  alors pour un  $A_2$  dépendant seulement de  $\phi$  il vient*

$$\mathbb{E} \left( \left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n \varepsilon_i g(X_i) \right\|_{\mathcal{G}} \right) \leq A_2 \beta J_{\phi} \left( \frac{\sigma}{\beta} \right).$$

### 3.4.3 – Vitesse sous condition d'entropie à crochet

Une autre mesure usuelle de la taille de  $\mathcal{F}$  compte les  $\mathbb{L}_2(P)$ -crochets plutôt que les boules, c'est-à-dire les  $\{f \in \mathcal{F} : l \leq f \leq u\}$  avec  $l \leq u$  mesurables telles que  $d_P(l, u) < \varepsilon$ . Soit  $N_{[\cdot]}(\varepsilon, \mathcal{F}, d_P)$  le nombre minimum de  $\varepsilon$ -crochets recouvrant  $\mathcal{F}$ , de sorte que  $N(\varepsilon, \mathcal{F}, d_P) \leq N_{[\cdot]}(\varepsilon/2, \mathcal{F}, d_P)$ .

(i) Considérons des classes de taille exponentielle, satisfaisant pour  $b_0 > 0$  et  $r_0 \in ]0, 1[$ ,

$$(BR) \quad \log N_{[\cdot]}(\varepsilon, \mathcal{F}, d_P) \leq \frac{b_0^2}{\varepsilon^{2r_0}}, \quad \varepsilon \in ]0, 1[.$$

**Proposition 3.20.** *Sous (BR) il existe un espace  $(\Omega, \mathcal{T}, \mathbb{P})$  sur lequel nous pouvons construire, pour tout  $\lambda > 1$  et un  $\rho_{\lambda} > 0$  associé, une version de  $\mathbb{G}$  et  $X_1, \dots, X_n$  telle que*

$$\mathbb{P} \left\{ \|\alpha_n - \mathbb{G}\|_{\mathcal{F}} > \frac{\rho_{\lambda}}{(\log n)^{(1-r_0)/2r_0}} \right\} \leq \frac{1}{n^{\lambda}}.$$

Comme  $\rho_{\lambda}$  est indépendant de  $n$ , il existe sur un espace  $(\Omega, \mathcal{T}, \mathbb{P})$  assez gros une suite i.i.d.  $X_n$  et  $\mathbb{G}_n$  de  $P$ -ponts browniens telles que  $\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{p.s.}(1/(\log n)^{(1-r_0)/2r_0})$ .

(ii) La vitesse universelle dans le principe d'invariance semble contrôlée par l'intégrale de Dudley

$$J_{[\varphi]}(\varepsilon) = \int_{[0, \varepsilon]} \sqrt{\varphi(s)} ds, \quad \Psi_{[\varphi]}(\varepsilon) = \varepsilon \exp \left( -\frac{5}{2} \varphi(\varepsilon) \right).$$

Dans [25] nous étendons la Proposition 3.20 à la condition

$$(BR') \quad \log N_{[\cdot]}(\varepsilon, \mathcal{F}, d_P) \leq \varphi(\varepsilon) < \infty, \quad \varepsilon \in ]0, 1[$$

avec  $\varphi > 0$  décroissante,  $J_{[\varphi]}(1) < \infty$  et  $\lim_{\varepsilon \rightarrow 0} \varphi(\varepsilon) = \infty$ . D'après [Du3,VW],  $\mathcal{F}$  est  $P$ -Donsker.

**Théorème 3.21.** *Sous (BR'),  $M_n > M > 0$  et  $\lim_{n \rightarrow \infty} M_n \sqrt{(\log n)/n} = 0$  il existe sur  $(\Omega, \mathcal{T}, \mathbb{P})$  assez riche une suite de v.a.  $X_n$  i.i.d. de loi  $P = \mathbb{P}^X$  et une suite  $\mathbb{G}_n$  de  $P$ -ponts browniens tels que*

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_n} = O_{p.s.} \left( J_{[\varphi]} \circ \Psi_{[\varphi]}^{-1} \left( M_n \sqrt{\frac{\log n}{n}} \right) \right).$$

Si  $J_{[\varphi]}$  est lente,  $J_{[\varphi]} \circ \Psi_{[\varphi]}^{-1}$  aussi donc l'approximation est d'autant meilleure que  $\varphi$  est petite.

**Exemple.** Pour  $M_n = M$  et  $\varphi$  très lente en  $\varepsilon \rightarrow 0$  la vitesse approche  $\sqrt{(\log n)/n}$  qui est meilleure que [KMT] mais s'adresse à des  $\mathcal{F}$  plus petites.

**Exemple.** Si  $\varphi(\varepsilon) = 1/\varepsilon^2 \zeta^2(\varepsilon)$  avec  $\zeta(\varepsilon) \rightarrow 0$  à variation lente quand  $\varepsilon \rightarrow 0$  alors  $J_{[\varphi]}(\varepsilon) \sim 1/\zeta(\varepsilon)$  et  $\Psi_{[\varphi]}^{-1}(x) \sim \sqrt{5/2} |\log x| / \zeta(\sqrt{|\log x|})$  en  $x \rightarrow 0$  donc la vitesse peut être arbitrairement lente. Prendre  $M_n = M$  et  $\zeta(\varepsilon) = |\log \varepsilon|^{\gamma}$  donne une vitesse  $O_{p.s.}(1/(\log \log n)^{\gamma})$ .

**Exemple.** Si  $\varphi(\varepsilon) = 1/\varepsilon^{2r_0}$  nous obtenons  $O_{p.s.}(1/(\log(\sqrt{n}/M_n\sqrt{\log n}))^{(1-r_0)/2r_0})$  qui est la version de 3.20 avec la troncature  $M_n$  prise en compte. Pour une entropie aussi grande, la vitesse de  $M_n \rightarrow \infty$  importe peu et n'affecte que les constantes, sauf si  $M_n$  est de l'ordre de  $\sqrt{n}$ .

**Exemple.** Si  $\varphi(\varepsilon) = \nu_0 \log(c_0/\varepsilon)$  et  $\limsup_{n \rightarrow \infty} \log M_n / \log n < 1/2$  nous obtenons comme en 3.17,  $O_{p.s.}((M_n\sqrt{\log n}/\sqrt{n})^{1/(1+5\nu_0/2)}\sqrt{\log n})$ . Dans ce cas l'approximation est très sensible à  $M_n \rightarrow \infty$ .

**Exemple.** Si  $\varphi(\varepsilon) = (\log(1/\varepsilon))^{2\beta}$  la vitesse du Théorème 3.21 vaut, si  $\beta > 1/2$ ,

$$\exp\left(-\left(\left(\frac{5}{2} + o(1)\right)\log n\right)^{1/2\beta}\right)$$

alors que si  $\beta < 1/2$  elle s'avère plus rapide et très dépendante de  $M_n$ ,

$$\frac{M_n}{\sqrt{n}}(\log n)^{\beta+1/2} \exp\left(\frac{5}{2}(\log n)^{2\beta}\right).$$

### 3.4.4 – Deux résultats de Dudley et Philipp revisités

Le but de [11] est de révéler l'importance et la portée de [Z] comme outil d'approximation en théorie générale des processus empiriques. Nous montrons comment la technique menant aux Propositions 3.16 et 3.20 grâce à [BP,Bo,Du2,EM3,T2,VW,Z] complétés par [DP,Ms,P] permet au final d'actualiser et de simplifier les résultats historiquement très importants de [DP]. Il s'agit de coupler  $\alpha_n$  à une somme de processus  $\mathbb{G}_n$  indépendants, ce qui est plus exigeant que le principe d'invariance.

(i) Ainsi le Théorème 7.4 de [DP], qui reposait sur un principe général de déconvolution [Y,De], s'étend aux classes de fonctions grâce à la Proposition 3.16 qui entraîne l'approximation forte suivante.

**Théorème 3.22.** *Soit  $\tau = 1/(2 + 5\nu_0)$ . Sous (VC) il existe  $(\Omega, \mathcal{T}, \mathbb{P})$  sur lequel nous pouvons construire, pour tout  $\lambda > 1$ ,  $\alpha \in ]1/2\tau, 1/\tau[$  et  $\rho_{\lambda,\alpha} > 0$  associé, une suite i.i.d.  $X_1, X_2, \dots$  de loi  $P$  et une suite de copies indépendantes  $\mathbb{G}_1, \mathbb{G}_2, \dots$  de  $\mathbb{G}$  vérifiant*

$$\mathbb{P}\left\{\max_{1 \leq m \leq n} \left\| \sqrt{m} \alpha_m - \sum_{i=1}^m \mathbb{G}_i \right\|_{\mathcal{F}} > \frac{\rho_{\lambda,\alpha} (\log n)^{\tau_0}}{n^{\tau(\alpha)-1/2}}\right\} \leq \frac{1}{n^\lambda}$$

avec  $\tau(\alpha) = (\alpha\tau - 1/2)/(1 + \alpha)$  et

$$\max_{1 \leq m \leq n} \left\| \sqrt{m} \alpha_m - \sum_{i=1}^m \mathbb{G}_i \right\|_{\mathcal{F}} = O_{p.s.} \left( \frac{(\log n)^{\tau_0}}{n^{\tau(\alpha)-1/2}} \right).$$

Par croissance de  $\tau(\alpha)$  le  $(\log n)^{\tau_0}$  peut disparaître de la vitesse presque sûre. Ici  $\Omega$  est un produit infini d'espaces de type 3.16, construits comme dans [BP,DP]. Puisque  $\tau(\alpha) < \tau/(2+2\tau)$  et  $\sum_{i=1}^m \mathbb{G}_i/\sqrt{m}$  est distribué comme  $\mathbb{G}$ , la vitesse est moins bonne qu'à la Proposition 3.16, en conséquence de l'uniformité en  $m \in [1, n]$ . Du Théorème 3.18 nous tirons aussi la version du Théorème 3.22 sous (VC').

(ii) De la même façon, nous généralisons le Théorème 7.1 de [DP].

**Théorème 3.23.** *Sous (BR) avec  $r_0 \in ]1/2, 1[$  donc  $\kappa_0 = (1 - r_0)/2r_0 < 1/2$  il existe  $(\Omega, \mathcal{T}, \mathbb{P})$  sur lequel nous pouvons construire, pour tout  $h > 0$  et  $\kappa < \kappa_0(1/2 - \kappa_0)$  avec  $\delta_{h,\kappa} > 0$  associé, une suite i.i.d.  $X_1, X_2, \dots$  de loi  $P$  et une suite de copies indépendantes  $\mathbb{G}_1, \mathbb{G}_2, \dots$  de  $\mathbb{G}$  vérifiant*

$$\mathbb{P}\left\{\max_{1 \leq m \leq n} \left\| \sqrt{m} \alpha_m - \sum_{i=1}^m \mathbb{G}_i \right\|_{\mathcal{F}} > \frac{\delta_{h,\kappa} \sqrt{n}}{(\log n)^\kappa}\right\} \leq \frac{1}{(\log n)^h}$$

et

$$\max_{1 \leq m \leq n} \left\| \sqrt{m} \alpha_m - \sum_{i=1}^m \mathbb{G}_i \right\|_{\mathcal{F}} = O_{p.s.} \left( \frac{\sqrt{n}}{(\log n)^\kappa} \right).$$

**(iii) Principe de démonstration.** Il s'agit de découper les  $\max_{1 \leq m \leq n}$  en blocs  $m_k < m \leq m_{k+1}$  de taille convenable  $n_k$  en contrôlant le maximum global par les  $\alpha_{m_k}$ , les déviations intra-blocs et les déviations inter-blocs via l'inégalité maximale [Ms]. Comme les  $\sqrt{m_{k+1}}\alpha_{m_{k+1}} - \sqrt{m_k}\alpha_{m_k}$  sont indépendants de loi  $\sqrt{n_k}\alpha_{n_k}$  nous les approximons via les Propositions 3.16 et 3.20 par des ponts browniens  $\mathbb{G}_{(k)}$  qui ensuite peuvent se représenter, d'après [BP,P], comme  $\sum_{i=m_k+1}^{m_{k+1}} \mathbb{G}_i/\sqrt{n_k}$  quitte à redéfinir l'espace  $\Omega$ . La vitesse finale apparaît sous la forme d'une série convergente de vitesses, qui doit être optimisée, d'où nous tirons  $\tau(\alpha)$  et  $\kappa$ .  $\square$

### 3.4.5 – Quelques conséquences

**(i)** Sur le cube  $\mathcal{X}_d = [0, 1]^d$  les meilleures vitesses connues concernent la loi uniforme  $P = \mathbb{U}_d$  et les classes  $\mathcal{F}_d^- = \{\mathbb{I}_{[0,t]} : t \in \mathcal{X}_d\} \subset \mathcal{F}_d = \{\mathbb{I}_{[s,t]} : s \leq t, (s,t) \in \mathcal{X}_d^2\}$ . Si  $d = 1, 2$  la vitesse optimale est  $(\log n)^d/\sqrt{n}$ , d'après [KMT,Tu,BM,CL]. Pour  $d > 2$  il n'est toujours pas conjecturé de vitesse mais la borne inférieure  $(\log n)^{(d-1)/2}/\sqrt{n}$  est obtenue par [Be] et [Ri] obtient un ordre sans dimension  $1/n^{1/12}$  que nous améliorons en  $1/n^{1/10}$  en 3.4.6. Lorsque  $P \neq \mathbb{U}_d$  la Proposition 3.16 implique

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_d} \leq 2^d \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_d^-} = O_{p.s.} \left( \frac{(\log n)^{(2+5d)/(2+10d)}}{n^{1/(2+10d)}} \right)$$

et cette borne universelle ne dépend que de  $d$ . En effet  $\mathcal{F}_d^-$  est une VC-classe d'indice  $d+1$  au sens de [VC] donc vérifie (VC) avec  $\nu_0 = 2d$ , voir [VW].

**(ii)** Soit  $S_d = \{\mathbb{I}_{B(s,\varepsilon)} : s \in \mathcal{X}_d, \varepsilon < 1\}$  la classe des sphères euclidiennes  $B(s, \varepsilon)$  de centre  $s$  et rayon  $\varepsilon$ . D'après [Be], si  $P = \mathbb{P}^X$  est absolument continue la vitesse est au mieux  $O_{p.s.}(n^{-1/2d})$  et pour  $P = \mathbb{U}_d$  [Ma] obtient  $O_{p.s.}((\log n)^{3/2}/n^{1/2d})$ . Comme  $S_d$  a un VC-indice  $d+2$  la Proposition 3.16 appliquée avec  $\nu_0 = 2d+2$  donne, pour tout  $P$ , une vitesse universelle en  $(\log n)^m/n^{1/(12+10d)}$ .

**(iii)** Soit  $P = \mathbb{U}_d$  et  $\mathcal{C}_d$  une famille de boréliens de  $\mathcal{X}_d$  telle que  $\mathcal{F}(\mathcal{C}_d) = \{\mathbb{I}_C : C \in \mathcal{C}_d\}$  vérifie (BR) avec  $r_0 < 1/2$ . Si les  $C \in \mathcal{C}_d$  ont une frontière régulière au sens de Minkowski, [Ma] produit une vitesse  $O_{p.s.}(1/(\log n)^{(1-2r_0)/2d})$  que la Proposition 3.20 améliore en  $O_{p.s.}((\log n)^{-(1-r_0)/2r_0})$ . Si les  $C \in \mathcal{C}_d$  ont une frontière différentiable à l'ordre  $\alpha > d-1$ , (BR) est vraie avec  $r_0 = (d-1)/2\alpha$  et notre vitesse

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}(\mathcal{C}_d)} = O_{p.s.} \left( \frac{1}{(\log n)^{\alpha/(d-1)-1/2}} \right)$$

devient d'autant meilleure que  $\alpha \rightarrow \infty$ . Le cas particulier  $\mathcal{F}(\mathcal{C}_d) = S_d$  montre que cette vitesse peut passer en puissance de  $n$  si  $\alpha = \infty$ .

**(iv)** Considérons une fonction  $\omega > 0$  sur  $\mathcal{X}_d = [0, 1]^d$  telle que

$$\inf_{\mathcal{X}_d} \omega = 0, \quad \int_{\mathcal{X}_d} \frac{1}{\omega^2} dP < \infty.$$

Les hypothèses 3.4.1 sont vérifiées par la classe  $\mathcal{F}_{\omega,d} = \{\mathbb{I}_{[0,t]}/\omega(t) : t \in \mathcal{X}_d\}$  avec  $F = 1/\omega$ . Les graphes de  $\mathcal{F}_{\omega,d}$  étant inclus dans les quadrants de  $\mathcal{X}_d \times \mathbb{R}_+$  il s'agit d'une VC-classe d'indice au plus  $d+2$  et 3.17 s'applique avec  $\nu_0 = 2d+2$ . Soit  $T_n$  une suite croissante de sous-ensembles de  $\mathcal{X}_d$  telle que  $M_n = \|1/\omega\|_{T_n} \rightarrow \infty$  et  $\inf_{T_n} \omega^2 > (\log n)/n$ . Posons  $\mathcal{F}_{\omega,n,d} = \{\mathbb{I}_{[0,t]}/\omega(t) : t \in T_n\}$ . Les processus empirique et brownien multivariés,  $\alpha'_n(t) = \alpha_n(\mathbb{I}_{[0,t]})$  et  $\mathbb{G}'_n(t) = \mathbb{G}_n(\mathbb{I}_{[0,t]})$  vérifient donc, par 3.17

$$\left\| \frac{\alpha'_n - \mathbb{G}'_n}{\omega} \right\|_{T_n} = \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_{\omega,n,d}} = O_{\mathbb{P}} \left( \left( \frac{M_n^2}{n} \right)^{1/(12+10d)} (\log n)^{(7+5d)/(12+10d)} \right)$$

pour tout  $P = \mathbb{P}^X$ . Ceci consitue un principe d'invariance en norme uniforme pondérée. La meilleure disponible, celle de [CH], concerne  $P = \mathbb{U}_1$  sur  $\mathcal{X} = \mathbb{R}$  et  $\omega_\alpha = I^\alpha(1-I)^\alpha$  avec  $\alpha < 1/2$ . Elle implique

$$\left\| \frac{\alpha'_n - \mathbb{G}'_n}{\omega_\alpha} \right\|_{T_{c,n}} = O_{\mathbb{P}} \left( \frac{1}{n^{1/2-\alpha}} \right)$$



pour tout  $P$  de fonction de répartition  $F$ ,  $\omega_\alpha = F^\alpha(1-F)^\alpha$  et  $T_{c,n} = (F^{-1}(c/n), F^{-1}(1-c/n))$ ,  $c > 0$ . C'est plus rapide que notre  $O_{p.s.}((\log n)^{6/11}/n^{(1/2-\alpha)/11})$  obtenu pour  $M_n = (n/c)^\alpha$  mais cette situation est singulière puisqu'à la fois le domaine  $T_{c,n}$  et la fonction de poids  $\omega_\alpha$  dépendent de  $P$ .

### 3.4.6 – Chaînage fort

Dans [25] nous calculons aussi une vitesse d'approximation forte via un contrôle uniforme des  $f \in \mathcal{F}$  par chaînage, en fonction de la longueur des chaînes, du nombre de couches de maillons et de la qualité de représentation sur  $\mathbb{L}_2(P)$ . Cela nécessite de préciser la géométrie de  $\mathcal{F}$  dans  $\mathbb{L}_1(P)$ ,  $\mathbb{L}_2(P)$  et  $\mathbb{L}_\infty$ .

(i) Faisons les deux hypothèses suivantes. D'une part, il existe  $\mathcal{F}_n^*$  avec  $n^{a_0} \leq \text{card}(\mathcal{F}_n^*) = c_n < \infty$  et  $\lambda > 1$  tels que  $\mathcal{F}_\infty = \bigcup_{n \geq 1} \mathcal{F}_n^*$  et chaque  $f \in \mathcal{F}$  est approchée par  $\theta_n^*(f) \in \mathcal{F}_n^*$  de telle sorte que

$$\mathbb{P}(\|\alpha_n - \alpha_n \circ \theta_n^*\|_{\mathcal{F}} \geq \varrho_n) \leq \frac{1}{n^\lambda}, \quad \mathbb{P}(\|\mathbb{G} - \mathbb{G} \circ \theta_n^*\|_{\mathcal{F}} \geq \varrho_n) \leq \frac{1}{n^\lambda}.$$

D'autre part, il existe  $\{f_{j,k} : k \leq k(j)\}$ ,  $\|f_{j,k}\|_{\mathcal{X}} = 1$  et  $\mathcal{F}_j^- \subset \{f = \sum_{k=1}^{k(j)} b_{j,k}(f) f_{j,k} : b_{j,k}(f) \in \mathbb{R}\}$  avec  $\sup_{f \in \mathcal{F}_j^-} \sum_{k=1}^{k(j)} b_{j,k}^2(f) < \infty$ , tels que tout  $f \in \mathcal{F}_\infty$  s'approche par

$$\theta_m^{**}(f) \in \mathcal{F}(m) = \left\{ f_1 + \dots + f_m : f_j \in \mathcal{F}_j^- \right\}$$

de telle sorte que

$$d_P^2(f, \theta_m^{**}(f)) \leq h(m) \downarrow 0.$$

Pour  $c_n$ ,  $k(j)$ ,  $\varrho_n$  et  $h(m)$  assez grands ces conditions sont remplies. Le problème est de diminuer ces quantités. Le chaînage consiste à passer de  $f \in \mathcal{F}$  à  $\theta_n^*(f) \in \mathcal{F}_n^*$  puis à  $\theta_m^{**} \circ \theta_n^*(f) \in \mathcal{F}(m)$ .

La dimension intrinsèque du problème d'approximation gaussienne se cache dans  $\Psi(m)$  défini par

$$b_m = \sup_{f \in \mathcal{F}} \sum_{j=1}^m \sum_{k=1}^{k(j)} b_{j,k}^2(f), \quad d_m = \sum_{j=1}^m k(j), \quad S_m = \sum_{j=1}^m k(j) \max_{k \leq k(j)} \mathbb{E}^2(\mathbb{I}_{f_{j,k} \neq 0}(X)),$$

$$r(j) = \sup_{x \in \mathcal{X}} \sum_{k=1}^{k(j)} \mathbb{I}_{f_{j,k} \neq 0}(x), \quad R_m = \sum_{j=1}^m r(j), \quad \Psi(m) = \frac{h(m)}{b_m(R_m + S_m) d_m^4 (\log d_m)^2} \downarrow 0.$$

**Théorème 3.24.** *Nous pouvons construire sur un même espace  $(\Omega, \mathcal{T}, \mathbb{P})$  une suite de v.a.  $X_n$  i.i.d. de loi  $P = \mathbb{P}^X$  et une suite de  $P$ -ponts browniens  $\mathbb{G}_n$  telles que*

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{p.s.}(v_n), \quad v_n = \max \left( \varrho_n, \frac{\log c_n}{\sqrt{n}}, \sqrt{h \circ \Psi^{-1} \left( \frac{(\log n)^2}{n \log c_n} \right) \log c_n} \right).$$

Comme  $h \circ \Psi^{-1} \gg I$  en 0, il vient  $v_n \gg (\log n)/\sqrt{n}$ . Nous prouvons que  $\mathbb{P}\{\|\alpha_n - \mathbb{G}\|_{\mathcal{F}} > \rho_\lambda v_n\} \leq 1/n^\lambda$  pour  $\lambda > 0$ . Typiquement la chaîne  $\mathcal{F}_j^-$  est générée par une base de fonctions contrôlée par  $P$ .

**Exemple.** Soit  $\mathcal{H} = \bigcup_{j \geq 0} \mathcal{H}_j$  une base de Haar de  $\mathbb{L}_2(P)$  avec  $\mathcal{H}_j = \{2^{j/2} f_{j,k} : k = 0, \dots, 2^j - 1\}$ ,  $f_{j,k} = \mathbb{I}_{C_{j+1,2k}} - \mathbb{I}_{C_{j+1,2k+1}}$ ,  $\mathbb{P}(C_{j,k}) = 1/2^j$  et  $C_{j+1,2k} \cup C_{j+1,2k+1} = C_{j,k}$  les partitions emboîtées. Notons  $a_{j,k}(f)$  les coordonnées de  $f \in \mathbb{L}_2(P)$  sur  $\mathcal{H}$ ,  $k(j) = 2^j$  et  $\mathcal{F}_j^- = \{\sum_{k=1}^{k(j)} a_{j,k} 2^{j/2} f_{j,k} : |a_{j,k}| \leq a_j\}$  avec  $a_j \rightarrow 0$  si  $j \rightarrow \infty$ . Alors  $2^m \leq d_m \leq 2^{m+1}$ ,  $R_m = m$ ,  $S_m \leq 1$ ,  $b_m = \sum_{j=0}^m a_j^2$ . En supposant  $a_j = 2^{-\alpha j}$ ,  $\alpha > 1/2$ , nous avons  $\Psi(m) = \gamma_1/m^3 2^{(4+\beta)m}$  et la classe

$$\mathcal{F} = \left\{ f : |a_{j,k}(f)| \leq \frac{1}{2^{\alpha j}}, \quad k \leq 2^j, \quad j \geq 0 \right\}$$

de [Ko] vérifie le principe d'invariance fort avec un ordre principal  $1/n^{(\alpha-1/2)/(7+2\alpha)}$ .

(ii) Si  $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{C}\}$  avec  $\mathcal{C} \subset \mathcal{A}$  une classe d'ensembles, nos conditions s'écrivent ainsi. Il existe  $n^{a_0} \leq \text{card}(\mathcal{C}_n^*) = c_n < \infty$  et  $\theta_n^* : \mathcal{C} \rightarrow \mathcal{C}_n^*$  telle que

$$\mathbb{P}(\|\alpha_n - \alpha_n \circ \theta_n^*\|_{\mathcal{F}} \geq \varrho_n) \leq \frac{1}{n^\lambda}, \quad \mathbb{P}(\|\mathbb{G} - \mathbb{G} \circ \theta_n^*\|_{\mathcal{F}} \geq \varrho_n) \leq \frac{1}{n^\lambda}.$$

Il existe des partitions de  $\mathcal{X}$  finies  $\mathcal{P}_j = \{C_{j,k} : k \leq k(j)\} \subset \mathcal{C}_\infty$ ,  $j \leq m$ , telles qu'à tout ensemble  $C \in \mathcal{C}_n^*$  correspond  $\theta_n^{**}(C) \in \mathcal{C}(m) = \{C_1 \cup \dots \cup C_m : C_j \in \mathcal{P}_j, C_j \cap C_{j'} = \emptyset \text{ si } j \neq j'\}$  vérifiant

$$\mathbb{P}(C \setminus \theta_n^{**}(C)) \leq h(m) \downarrow 0.$$

Dans ce cas  $f_{j,k} = \mathbb{I}_{C_{j,k}}$ ,  $R_m = b_m = m$  et souvent  $S_m \leq a_\infty m$ . D'où une expression simple de la fonction clef donnant la vitesse dans le Théorème 3.24 en fonction de l'approximation  $\mathbb{L}_1(P)$  de  $\mathcal{F}$ ,

$$\Psi(m) = \frac{h(m)}{m^2 d_m^4 (\log d_m)^2} \downarrow 0.$$

**Exemple.** Soit  $\mathcal{F}_d$  la classe des cubes de  $\mathcal{X}_d = [0, 1]^d$  et  $P = \mathbb{P}^X$  admettant une densité par rapport à  $\mathbb{U}_d$  bornée et détachée de 0. La partition dyadique de [Ri] remplit les conditions avec  $\varrho(n) = \log n$ ,  $c_n \leq (2n)^d$ ,  $\sup_m S_m < \infty$ ,  $d_m \leq 2^m (m+d)^{d-1}$ ,  $h(m) = c_0 m^{d-1}/2^m$  donc, en notant  $\log_2$  le logarithme en base 2,  $\Psi(m) = 1/2^{5m+5(d-1)\log_2 m}$  et  $h(\Psi^{-1}(y)+1) = O(y^{1/5} |\log_2 y|^{d-1})$  si  $y \rightarrow 0$ . Le Théorème 3.24 donne alors  $\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{p.s.}((\log n)^{d-1/2}/n^{1/10})$  qui améliore l'ordre  $1/n^{1/12}$  de [Ri].

(iii) Dans (i,ii) la classe  $\mathcal{F}$  est uniformément bornée. Pour intégrer le fait que l'approximation par couche se dégrade ou s'améliore lorsque  $\mathcal{F} = \mathcal{F}_n$  varie avec  $n$  et explorer l'interaction entre la qualité de représentation de  $\mathcal{F}_n$  sur  $\mathbb{L}_2(P)$  et  $\mathbb{L}_\infty$ , supposons que

$$d_P^2(f, \theta_m^{**}(f)) \leq h_n(m), \quad \|f - \theta_m^{**}(f)\|_{\mathcal{X}} \leq \pi_n(m)$$

où, pour  $n$  fixé,  $h_n \leq \pi_n^2$ ,  $h_n \downarrow 0$ ,  $\pi_n \downarrow$  et  $\pi_n^2/h_n \uparrow$  lorsque  $m \rightarrow \infty$ . Les fonctions de  $m$

$$\Psi_{h,n} = \frac{\sqrt{h_n}}{\Psi}; \quad \Psi_{\pi,n} = \frac{\pi_n}{\Psi}; \quad \Psi_{\pi,h,n} = \frac{\pi_n}{\sqrt{h_n}} = \frac{\Psi_{\pi,n}}{\Psi_{h,n}}; \quad \Psi(m) = \sqrt{b_m(R_m + S_m)} d_m^2 \log d_m;$$

servent à équilibrer la qualité d'approximation de  $\mathcal{F}_n$  et la dimension de la chaîne. Si  $b_m < \infty$  nous pouvons coupler sur  $(\Omega, \mathcal{T}, \mathbb{P})$  la suite de ponts empiriques  $\alpha_n$  à une suite de ponts browniens  $\mathbb{G}_n$  d'une telle façon que, pour  $C > 0$ ,

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_n} = O_{a.s.}(\max(\varrho(n), v_n))$$

avec  $\varrho(n)$  le terme de module de continuité de (i) et  $v_n$  définie comme suit. Si  $c_n > e^{a_0 n}$ ,  $a_0 > 0$ ,

$$v_n = \pi_n \circ \Psi_{\pi,n}^{-1} \left( \frac{C \log n}{\sqrt{n \log(nc_n)}} \right) \sqrt{\log(nc_n)}$$

ce qui n'est pertinent que pour des classes  $\mathcal{F}_n$  extrêmement bien approximées dans  $\mathbb{L}_\infty$  par la chaîne  $\mathcal{F}_j^-$ ,  $j \geq 1$ , pour lesquelles  $\pi_n$  décroît très rapidement avec  $m$ . Si  $c_n < e^{a_0 n}$ ,  $a_0 > 0$ ,

$$v_n = \sqrt{h_n} \circ \Psi_{h,n}^{-1} \left( \frac{C \log n}{\sqrt{n \log(nc_n)}} \right) \sqrt{\log(nc_n)} \\ + \mathbb{I} \left\{ \Psi_{h,n}^{-1} \left( \frac{C \log n}{\sqrt{n \log(nc_n)}} \right) > \Psi_{\pi,h,n}^{-1} \left( \sqrt{\frac{C n}{\log(nc_n)}} \right) \right\} \pi_n \circ \Psi_{\pi,n}^{-1} \left( \frac{C \log n}{\log(nc_n)} \right) \frac{\log(nc_n)}{\sqrt{n}}$$

où le premier terme dû à la qualité d'approximation dans  $\mathbb{L}_2(P)$  est pénalisé lorsque l'approximation dans  $\mathbb{L}_\infty$  n'est pas de qualité suffisante. Ceci autorise la classe  $\mathcal{F}_n$  de 3.4.1, uniformément bornée par des constantes  $\pi_n = M_n \rightarrow \infty$  mais avec un second terme  $M_n \log(nc_n)/\sqrt{n}$  tendant vers 0 uniquement si  $c_n < e^{o(\sqrt{n}/M_n)}$  et  $M_n(\log n)/\sqrt{n} \rightarrow 0$ . Comme  $\pi_n \circ \Psi_{\pi,n}^{-1} \gg I$  et  $h_n \circ \Psi_{h,n}^{-1} \gg I$  en 0 nous avons  $v_n \gg (\log n)/\sqrt{n}$ . Ici nous exploitons au mieux les inégalités de Hoeffding, Bennett, Bernstein.

(iv) Cette approche par chaînage analytique ou géométrique, permet de construire explicitement de nouvelles classes de Donsker à partir de couches de fonctions, ou d'ensembles, élémentaires.

### 3.5 – Autres axes de recherche

#### 3.5.1 – Petites déviations poissoniennes et empiriques

En 2.1.5 nous avons vu qu'à l'échelle même du principe d'invariance les oscillations du processus empirique réel  $\alpha_n$  sont étroitement liées aux petites déviations du processus de Poisson centré.

(i) Depuis [8] nous avons entrepris l'extension de la borne de [DL2,Sm] aux  $f \in \mathcal{S}_1$  avec  $f'$  à variation non bornée, en couplant la technique de [6] exposée en 3.1 au changement de mesure de Poisson. La difficulté nouvelle consiste à maîtriser l'économie d'énergie poissonienne et non plus gaussienne. En corollaire, les résultats de [8] sont prolongés pour certaines fonctions de  $\mathcal{S}_1^{pli}$ .

Par exemple, supposons que  $Var(f', [x, 1]) < Var(f', [0, x]) = \infty$  pour tout  $x > 0$  et  $f \in \mathcal{S}$  concave croissante ( $s_f = 1$ ) ou convexe décroissante ( $s_f = -1$ ) sur  $]0, x_0[$  avec  $x_0 > 0$ . Soit

$$p_\varepsilon = \frac{f - s_f \varepsilon}{I}, \quad x_\varepsilon = \inf \left\{ x : \frac{p_\varepsilon(x)}{f'(x)} \geq 1 \right\}, \quad k_\varepsilon = p_\varepsilon(x_\varepsilon),$$

et les paramètres d'économie d'énergie gaussienne et de vitesse de fuite du terme poissonien,

$$J_{f,\varepsilon} = \int_{[0,x_\varepsilon]} f'^2 - k_\varepsilon^2 x_\varepsilon, \quad \theta_{f,\varepsilon} = \int_{[x_\varepsilon,1]} f'^3.$$

**Lemme 3.25.** *Soit  $\Pi$  un processus de Poisson sur  $[0, \rho]$  et  $\Pi_\rho = (\Pi(\rho I) - \rho I)/\sqrt{\rho}$ . Si*

$$\rho \rightarrow \infty, \quad \lambda \rightarrow \infty, \quad \varepsilon \rightarrow 0, \quad \rho \lambda^2 \varepsilon^2 \rightarrow \infty, \quad \frac{\lambda^2 k_\varepsilon^2}{\rho} \rightarrow 0$$

*il vient, pour tout  $\delta \in [0, 1[$ ,*

$$\begin{aligned} & \exp \left( -\frac{\lambda^2}{2} (1 - J_{f,\varepsilon}) - \frac{\pi^2 + o(1)}{8\lambda^2 \delta^2 \varepsilon^2} + s_f \frac{\lambda^3 (k_\varepsilon^3 x_\varepsilon + \theta_{f,\varepsilon})}{(6 + o(1)) \sqrt{\rho}} - (2 + o(1)) \lambda^2 \delta \varepsilon k_\varepsilon \right) \\ & \leq \mathbb{P} \left( \frac{\Pi_\rho}{\lambda} \in f + (1 + \delta) \varepsilon \mathcal{B} \right) \\ & \leq \exp \left( -\frac{\lambda^2}{2} (1 - J_{f,\varepsilon}) - \frac{\pi^2 + o(1)}{8\lambda^2 \delta^2 \varepsilon^2} + s_f \frac{\lambda^3 (k_\varepsilon^3 x_\varepsilon + \theta_{f,\varepsilon})}{(6 + o(1)) \sqrt{\rho}} + (2 + o(1)) \lambda^2 \delta \varepsilon k_\varepsilon \right). \end{aligned}$$

Si  $f \sim c_0 I^\alpha$  sur  $V(0)$  nous avons  $J_{f,\varepsilon} \sim c_0 \varepsilon^{2-1/\alpha}$ ,  $k_\varepsilon^3 x_\varepsilon + \theta_{f,\varepsilon} \sim c_2 \varepsilon^{3-2/\alpha}$  et  $\varepsilon k_\varepsilon \sim c_3 \varepsilon^{2-1/\alpha}$ . Dans ce cas la méthode de preuve du 2.1.5 montre que la vitesse presque sûre au Théorème 2.13 pour  $d(f, \Theta_n)$  est gaussienne lorsque  $na_n \gg (\log n)^{3/(4\alpha-1)}$ , de l'ordre de  $(1/\log n)^{2\alpha/(4\alpha-1)}$ , et cela sans utiliser [KMT]. Sinon, lorsque  $c_0 > 0$  le Lemme 3.25 admet une version pour  $\lambda^2 k_\varepsilon^2 / \rho \rightarrow \infty$  avec un premier terme poissonien dominant qui montre que  $d(f, \Theta_n)$  est de l'ordre de  $((\log n)/na_n)^{\alpha/2(1-\alpha)}$ .

(ii) Une direction de recherche vierge concerne les petites déviations de  $\alpha_n$  indexé par des fonctions et, plus généralement, des processus indexés par un espace métrique. Nous avons entrepris d'approfondir les relations mises en évidence dans [KL] entre le problème des petites déviations et l'action de  $\mathbb{P}^X$  sur  $\mathcal{F}$  pour les répercuter sur  $\alpha_n$  par approximation poissonienne ou gaussienne. Le premier objectif est d'exhiber des vitesses dans 2.3.4 quand  $X \in \mathbb{R}^d$  et les incréments de  $\alpha_n$  sont indexés par des ensembles. A terme il s'agit de traiter les incréments de type [G].

Pour évaluer des probabilités de petite boule pour  $\mathbb{G}$  vu comme une mesure sur  $\mathbb{L}_\infty(\mathcal{F})$  nous devons, d'après [KL], relier l'entropie et la géométrie de  $\mathcal{F}$  à l'entropie de la boule unité  $H_{\mathcal{F}}$  de l'espace de Hilbert auto-reproduisant associé à  $\mathbb{G}$ . Dans 2.3.4 le transfert à  $\alpha_n$  s'envisage par poissonisation, changement de mesure de Poisson pour translater les boules  $\{c\varphi + \varepsilon H_{\mathcal{F}}\}$ ,  $c \rightarrow \infty$ ,  $\varepsilon \rightarrow 0$  et application de 3.4 pour récupérer l'évaluation gaussienne en généralisant l'outil de [Mo] conduisant à [Sm]. Dans le cas général une stratégie consiste à combiner les outils sous-jacents à 3.4 avec un principe de grandes déviations de type [Ar]. Ce travail fait l'objet d'une thèse sous notre direction.

(iii) Pour traiter directement les petites déviations  $\mathbb{P}(\|\alpha_n - c_n h\|_{\mathcal{F}} \leq \varepsilon_n)$  évoquées en 1.3.1(ii) une autre approche consiste à affiner les techniques de chaînage dans  $\mathcal{F}$  en introduisant de nouveaux nombres de recouvrement, avec une contrainte sur les boules ou partitions liée à la fonctionnelle  $h$  via  $\mathbb{P}^X$ . Il est aussi envisageable d'étudier les processus de Poisson indexés par un espace quelconque et quitter l'approche par trajectoires pour l'approche par mesure. Le but est ici de détecter des différences avec les petites déviations gaussiennes pour produire une borne inférieure dans les principes d'invariance 3.4, comme en 2.1.5, au moins dans des cas particuliers.

### 3.5.2 – Petites déviations pour les processus composés

(i) En vue des applications il serait utile d'étendre les résultats de la partie 2 au cas où la mesure empirique  $\mathbb{P}_n$  est construite en affectant un poids aléatoire  $Y_i > 0$  à chaque  $X_i$ . Quand les poids  $Y_1, \dots, Y_n$  sont i.i.d.  $\mathbb{P}^Y$  et indépendants des  $X_1, \dots, X_n$  nous parlons de processus empirique composé. Or chaque ingrédient des preuves de 2.1 requiert un nouvel outil pour  $\alpha_n$  composé. Ceci ne pose pas de problème si  $\mathbb{P}(|Y_i| < M) = 1$ . Sinon, une technique de découpage doit faire intervenir les grandes déviations de  $\mathbb{P}^Y$  rapportées à l'intensité courante des sauts, et demande que  $\mathbb{E}(Y^2) < \infty$ .

(ii) En particulier, le problème des petites boules pour le processus de Poisson composé en haute intensité, ou plus généralement des processus de Lévy stables, se traite à l'aide d'une méthode de troncature des sauts et en supposant la mesure d'intensité dans  $\mathbb{L}_2$ . La borne de type [Sm] voit alors apparaître un terme perturbateur lié à la queue de distribution de la mesure de Lévy. Des lois de type Chung (2.1.2, 2.1.5) se déduisent de sa confrontation avec les termes gaussiens et poissoniens, via le niveau  $\zeta$  de troncature, fonction du temps, et à l'autosimilarité. La description des lois fonctionnelles de type Chung et Strassen pour les processus de Lévy  $\alpha$ -stables  $X_{<\zeta(t)}^\alpha(t)$  privés des sauts au dessus de  $\zeta(t)$  est en cours, avec vitesses et normalisations fonction de  $\zeta(t)$ . Ceci fait le lien entre le domaine gaussien et purement stable, où les théorèmes limites usuels échouent.

### 3.5.3 – Perspectives

(i) Au vu des Théorèmes 3.18, 3.21 et de nos commentaires, les vitesses des Propositions 3.16 et 3.18 nous semblent porteuses de sens. C'est pourquoi nous conjecturons qu'il est possible de construire une loi  $\mathbb{P}^X$  défavorable à  $\mathcal{F}$  pour laquelle la vitesse d'approximation s'approche de la borne donnée par la méthode de [11,25]. Une piste consiste à exploiter les lois vectorielles les moins "gaussiennes" dans le couplage [Z]. Si l'argumentation fait intervenir la structure de  $\mathcal{F}$ , il faudra dans un premier temps construire  $\mathcal{F}$  simultanément à la pire loi  $\mathbb{P}^X$ . Ce problème est entièrement ouvert.

(ii) Dans [EM2] les moyennes mobiles  $\Delta S_n(t, a_n)/na_n$  de la marche aléatoire  $S_n$  sur  $\mathbb{R}$  de 1.1.2 sont observées aux échelles qui révèlent un comportement non gaussien, de type  $na_n = (\log n)^\gamma$  avec  $\gamma$  fonction des hypothèses de moment sur  $\mathbb{P}^X$ . Ce phénomène de petites déviations, non fonctionnelles dans [EM2], correspond à la frontière entre les domaines d'invariance et de non-invariance [Sb]. Il serait intéressant d'en donner une version fonctionnelle, comme dans [8], afin d'expliquer la vitesse optimale d'approximation forte gaussienne [KMT,CR3] par des différences trajectoires précises. D'où la perspective motivante d'un analogue de 2.1.5 pour  $S_n$  lorsque  $na_n \sim (\log n)^p$  ou légèrement plus large, et  $\mathbb{E}(e^{tY}) < \infty$  avec  $s \in ]-s_0, s_1[$  et  $Y = |X|^{1/p} \text{signe}(X)$ . Les déviations exactes de  $S_n$  sont également en rapport avec la normalisation universelle de 2.4.4. Plus généralement, les petites déviations non gaussiennes restent à explorer dans le cas  $\mathbb{R}^d$  de 2.4.3 et le cas banachique de [AK].

(iii) Dans (ii) le cas particulier où  $\mathbb{E}e^{sX} < \infty$  doit être traité comme en 2.1.5 car l'écart au cas gaussien est ténu. L'échelle critique est  $\gamma = 1$  du fait de la loi d'Erdős-Rényi-Shepp [D,ERé,Sp] pour  $na_n = [d \log n]$  et l'ensemble limite  $\mathcal{S}_1$  devient la boule de Orlicz associée à la transformée de Crámer de  $\mathbb{P}^X$ , du type  $\mathcal{O}_d$ . D'après [8] ce comportement non gaussien est encore palpable aux échelles  $\gamma > 1$ . Comment décrire cette différence, et la transition entre les ensembles limites ?

Il convient pour cela de raffiner les propriétés de petites et grandes déviations de  $S_n$  à ces petites échelles, ce qui suppose de découvrir les deux outils suivants. Tout d'abord, l'analogue de [DL2,Sm]

et du Lemme 3.3 mêlant la loi gaussienne au premier ordre à la loi  $\mathbb{P}^X$  de  $X$  au second ordre selon l'intensité déterministe de l'échelle de temps. Ensuite, la borne de grandes déviations de  $S_n$  au delà du premier ordre logarithmique gaussien, en prouvant la conjecture de concentration de [1].

(iv) Les techniques de 3.4 demandent à être adaptées au cas des sommes de v.a. à valeurs dans un espaces de Banach, dont les bornes d'approximation sont sans progrès depuis plus de vingt ans.

(v) L'étude des systèmes dynamiques conduit, en théorie ergodique, à des résultats de type théorème limite central et principe d'invariance faible. Il est donc envisageable de découvrir des principes d'invariance forts. Le lien avec nos travaux se fait en considérant une suite  $X_n$  faiblement ou fortement mélangeante. La technologie de 3.4 basée sur [Z] peut servir aux côtés des outils de découplage [DG], de ceux de théorie ergodique et des inégalités de moments sous conditions de mélange.

### 3.6 – Références citées en partie 3

- [A2] De Acosta A. (1983). Small deviations in the functional central limit theorem with applications to functional law of the iterated logarithm. *Annals of Probability* **11**, 78-101.
- [Be] Beck J. (1985). Lower bounds on the approximation of the multivariate empirical process. *Z. Wahrsch. verw. Gebiete*. **70**, 289-306.
- [BP] Berkes I. and Philipp W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Annals of Probability* **7**, 29-54.
- [Bo] Borell C. (1975). The Brunn-Minkowski inequality in Gauss space. *Inventiones Mathematicae* **30**, 207-216.
- [BM] Bretagnolle J. and Massart P. (1989). Hungarian construction from the non asymptotic view point. *Annals of Probability* **17**, 239-256.
- [CM] Cameron R.H. and Martin W.T. (1944). The Wiener measure of Hilbert neighborhoods in the space of real continuous functions. *J. Math. Phys. Mass. Inst. Tech.* **23**, 195-209.
- [CL] Castelle N. and Laurent-Bonvalot F. (1998). Strong approximation of bivariate uniform empirical processes. *Ann. Inst. H. Poincaré* **34**, 425-480.
- [Cs] Csáki E. (1980). A relation between Chung's and Strassen's law of the iterated logarithm. *Z. Wahrsch. und Verw. Gebiete* **54**, 287-301.
- [Cs2] Csáki E. (1989). A liminf result in Strassen's law of the iterated logarithm. *Colloq. Math. Soc. János Bolyai* **57**, 83-93.
- [CCHM] Csörgő M., Csörgő S., Horváth L. and Mason D. M. (1986). Weighted empirical and quantile processes. *Annals of Probability* **14**, 31-85.
- [CH] Csörgő M. and Horváth L. (1993). Weighted strong approximations in Probability and Statistics. *Wiley*.
- [CR3] Csörgő M. and Révész P. (1981). Strong approximations in probability and statistics. *Academic Press, New York*.
- [DG] De la Peña V. H. and Giné E. (1999). Decoupling. From dependence to independence. Randomly stopped processes. U-Statistics and processes. Martingals and beyond. *Probability and its Applications. Springer-Verlag*.
- [De] Dehling H. (1983). Limit theorems for sums of weakly dependent Banach space valued random variables. *Z. Wahrsch. verw. Gebiete*. **63**, 393-432.
- [DS] Deuschel J. D. and Strook D. W. (1989). Large deviations. *Academic Press, New York*.
- [DV] Donsker M. D. and Varadhan S. R. S. (1977). On laws of the iterated logarithm for local times. *Comm. Pure Appl. Math.* **30**, 707-753.
- [Du2] Dudley R. M. (1967). The sizes of compact subsets of an Hilbert space and continuity of Gaussian processes. *Journal of Functional Analysis* **1**, 290-330.
- [Du3] Dudley R. M. (1999). Uniform central limit theorems. *Cambridge Univ. Press*.

- [DP] Dudley R. M. and Philip W. (1983). Invariance principle for sums of Banach space valued random elements and empirical processes. *Z. Wahrsch. verw. Gebiete* **62**, 509-552.
- [EjM] Einmahl J. H. J. et Mason D.M. (1988). Strong limit theorems for weighted quantile processes. *Annals of Probability* **16**, 1623-1643.
- [EM3] Einmahl U. and Mason D. M. (1997). Gaussian approximation of local empirical processes indexed by functions. *Probability Theory and Related Fields* **107**, 283-311.
- [G] van de Geer S. (2000). Empirical Processes in M-Estimation. *Cambridge University Press*.
- [GZ] Giné E. and Zinn J. (1984). Some limit theorems for empirical processes. *Annals of Probability* **12**, 929-989.
- [GK] Goodman V. and Kuelbs J. (1991). Rates of clustering in Strassen's LIL for Brownian motion. *Journal of Theoretical Probability* **4**, 285-309.
- [GL] Gorn N. and Lifshits M. A. (1999). Chung's law and the Csáki function. *Journal of Theoretical Probability* **12**, 399-420.
- [Gr2] Grill K. (1991). A liminf result in Strassen's law of the iterated logarithm. *Probability Theory and Related Fields* **89**, 149-157.
- [Gn] Groeneboom P. (1989). Brownian motion with a parabolic drift and airy functions. *Probability Theory and Related Fields* **81**, 79-109.
- [JM] Jain, M. C. and Marcus, M. B. (1978). Continuity of sub-Gaussian processes. Dekker, New York, *Advances in Probability* **4**, 81-196.
- [Ja] James B. R. (1975). A functional law of the iterated logarithm for weighted empirical processes. *Annals of Probability* **3**, 767-772.
- [Kh] Khatri C. G. (1967). On certain inequalities for normal distributions and their applications to simultaneous confidence bounds. *Annales of Mathematical Statistics* **38**, 1853-1867.
- [Ko] Koltchinskii V. I. (1994). Komlós-Major-Tusnády approximation for the general empirical process and Haar expansions of classes of functions. *Journal of Theoretical Probability* **7**, 73-118.
- [KL2] Kuelbs J. and Li W. (1993). Metric entropy and the small ball problem for Gaussian measures. *Journal of Functional Analysis* **116**, 133-157.
- [KLT] Kuelbs J., Li W. V. and Talagrand M. (1994). Liminf results for Gaussian samples and Chung's functional LIL. *Annals of Probability* **22**, 1879-1903.
- [L] Li W. V. (1999). Small deviations for gaussian Markov processes under the sup-norm. *Journal of Theoretical Probability* **12**, 971-984.
- [LLi] Li W. V. and Linde W. (1999). Approximation, metric entropy and small ball estimates for gaussian measures. *Annals of Probability* **27**, 1556-1578.
- [Lf] Lifshits M. A. (1995). Gaussian random functions. *Kluwer, Dordrecht*.
- [LS] Lifshits M. A. and Shi Z. (2002). The first exit time of Brownian motion from a parabolic domain. *Bernoulli* **8**, 745-765.
- [MZ] Mason D. M. and Van Zwet W. R. (1987). A refinement of the KMT inequality for the uniform empirical process. *Annals of Probability* **15**, 871-884.
- [Ma] Massart P. (1989). Strong approximations for multidimensional empirical and related processes, via KMT constructions. *Annals of Probability* **17**, 266-291.
- [Ms] Montgomery-Smith S. (1993). Comparison of sums of independent identically distributed random variables. *Probability and Mathematical Statistics* **14**, 281-285.
- [Mo] Mogulskii A. A. (1974). Small deviations in a space of trajectories. *SIAM Theory of Probability and Application* **19**, 726-736.
- [P] Philipp W. (1986). Invariance principles for dependent and weakly dependent random variables. In: Dependence in Probability and Statistics. *Progress in Probability and Statistics* **11**, 225-268.

- [Po] Pollard D. (1984). Convergence of stochastic processes. *Springer-Verlag*.
- [R] Révész, P. (1976). On strong approximation of the multidimensional empirical process. *Annals of Probability* **4**, 729-743.
- [Ri] Rio, E. (1996). Vitesses de convergence dans le principe d'invariance faible pour la fonction de répartition empirique multivariée. *C. R. Acad. Sci. Paris* **322**, Série I, 169-172.
- [Sé] Seneta E. (1976). Regularly varying functions. *Lecture Notes in Mathematics* **508**, Springer-Verlag.
- [Shi] Shi Z. (1996). Small ball probabilities for a Wiener process under weighted sup-norms, with an application to the supremum of Bessel local times. *Journal of Theoretical Probability* **9**, 915-929.
- [Sz] Stoltz W. (1996). Some small ball probabilities for Gaussian processes under non-uniform norms. *Journal of Theoretical Probability* **9**, 613-630.
- [T2] Talagrand M. (1994). Sharper bounds for Gaussian and empirical processes. *Annals of Probability* **22**, 28-76.
- [Tu] Tusnády G. (1977). A remark on the approximation of the sample DF in the multidimensional case. *Period. Math. Hungar.* **8**, 53-55.
- [VW] van der Vaart A. W. and Wellner J. A. (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag.
- [VC] Vapnik V. N. and Červonekis A. Ya. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *SIAM Theory of Probability and its Application* **16**, 264-280.
- [Y] Yurinskii V. V. (1977). On the error of the gaussian approximation for convolutions. *SIAM Theory of Probability and its Application* **22**, 236-247.
- [Z] Zaitsev A. Yu. (1987). Estimates of the Lévy-Prokhorov distance in the multivariate central limit theorem for random variables with finite exponential moments. *SIAM Theory of Probability and its Application* **31**, 203-220.
- [Z2] Zaitsev A. Yu. (1987). On the Gaussian approximation of convolutions under multidimensional analogues of S. N. Bernstein's inequality conditions. *Probability Theory and Related Fields* **74**, 534-566.

## 4 – UTILISATION EN STATISTIQUE NON PARAMÉTRIQUE

Dans cette partie nous entrons dans le domaine prolifique de l'estimation non paramétrique d'une densité  $f$  sur  $\mathbb{R}$  et de ses points particuliers, tels que les extréma locaux. Nos résultats s'adaptent au cadre conditionnel de l'estimation de la régression. Nous adoptons le point de vue de la convergence presque sûre et du contrôle déterministe des pires écarts possibles. En fait, vu sous cet angle le domaine paraît bien plus petit car la plupart des auteurs proposent des méthodes en s'appuyant au mieux sur un théorème de convergence en loi, au pire sur une loi des grands nombres sans vitesse, et le plus souvent sur une borne du risque. En effet, les critères usuels de comparaison d'estimateurs portent sur l'erreur moyenne plutôt que sur l'erreur elle-même et son comportement asymptotique.

Notre contribution consiste à préciser à quelle vitesse l'erreur va dévier de son espérance, donnant ainsi la clef probabiliste dans l'équilibre entre le biais et l'aléa. Pour comprendre les oscillations nous devons étudier la statistique ou la méthode en question avec l'arsenal des parties 2 et 3, comme expliqué en 1.3.2. Mais l'horizon à explorer est vaste et étoilé d'hypothèses had-hoc, en général faites pour contrôler les valeurs moyennes, le biais, bref la partie analytique. Ceci tend à dissimuler ou parasiter ce qui gouverne la partie purement aléatoire du problème, qui principalement nous intéresse. Un exemple flagrant est donné en 4.2 et 4.3 où depuis longtemps les hypothèses en estimation de

$$\arg \min_{(s,t)} \left\{ t - s : \int_s^t f = q \right\}$$

et

$$\arg \max_{\mathbb{R}} f$$

portent sur  $f$  elle-même et non sur la fonction de quantile  $F^{-1}$  alors que le processus fondamental pilotant le phénomène oscillatoire sous-jacent est ici le processus de quantile  $\beta_n = \sqrt{n}(F_n^{-1} - F^{-1})$ . Et dans ce cas, trop de régularité sur  $f$  dénature et appauvrit le problème alors qu'il est si riche. Pour développer ce point nous détaillons quelques exemples en 4.3

Dans nos travaux en statistique asymptotique nous faisons systématiquement appel aux processus empiriques et browniens pour produire des théorèmes limites. C'est aussi le point de vue de [SW] qui aborde de cette façon un grand nombre de problèmes non paramétriques. Deux décennies après, les outils disponibles sont plus affûtés, à l'image des parties 2 et 3. C'est pourquoi, en 4.1, 4.2 et 4.3 nous résolvons entre autre des problèmes ouverts de [SW] et évoquons en 4.4 des projets de ce type. L'approche par processus empiriques nécessite de développer de nouvelles méthodes et de nouveaux outils mais nous sommes convaincus qu'elle donnera bien d'autres résultats.

### 4.1 – Estimation de la densité par delta-suites

Dans [1,12,21] nous utilisons les lois fonctionnelles présentées en 2.1 pour contrôler le terme aléatoire associé aux estimateurs non paramétriques de la densité basés sur des delta-suites. Ceci produit des lois du logarithme pour une classe d'estimateurs assez large et généralise des résultats classiques basés sur [St]. En outre nous améliorons les conditions suffisantes, précisons les seconds ordres et relaxons la contrainte d'estimation sur un compact. Ces résultats valent en particulier pour les suites de noyaux et les delta-fonctions convergeant vers un ou plusieurs noyaux, comme les séries trigonométriques.

#### 4.1.1 – Delta-suites de fonctions

Soit à estimer une densité  $f > 0$ , continue sur  $\text{supp } f = [m, M]$  avec  $-\infty \leq m < M \leq \infty$ , au vu d'une suite  $X_1, X_2, \dots$  de v.a.r. réelles i.i.d. de loi absolument continue par rapport à  $\lambda$ , de densité  $f$ .

(i) Nous considérons l'estimateur de  $f$  sur  $\mathbb{R}$  défini par

$$f_n(x) = \frac{1}{n} \sum_{i=1}^n \delta_n(X_i - x)$$



avec  $\delta_n$  une delta-suite de fonctions au sens des conditions  $A$  suivantes, introduites dans [LW].

$$A(i) \quad \limsup_{n \rightarrow \infty} \|\delta_n\|_{\mathbb{L}_1} < \infty.$$

$$A(ii) \quad \int_{\mathbb{R}} \delta_n(x) dx = 1.$$

$$A(iii) \quad \sup_{|x| > \lambda} |\delta_n(x)| \rightarrow 0, \quad \lambda > 0.$$

$$A(iv) \quad \int_{|x| > \lambda} |\delta_n(x)| dx \rightarrow 0, \quad \lambda > 0.$$

Ainsi  $\delta_n$  est une suite d'approximations de  $\delta_0$  la masse de Dirac en 0 et  $f_n$  est une suite de densités aléatoires, dirigées à la fois par  $\delta_n$  et par  $X_1, \dots, X_n$  donc  $\mathbb{P}_n$ .

**Exemple.** Les estimateurs à noyau [P,Rb,Pr,Dv2,BLc] correspondent à  $\delta_n(x) = K(x/h_n)/h_n$  où  $h_n \rightarrow 0$  et  $K$  vérifie  $A(i, iv)$  si  $K \in \mathbb{L}_1$ ,  $A(ii)$  si  $\int_{\mathbb{R}} K = 0$  et  $A(iii)$  si  $xK(x) \rightarrow 0$  quand  $|x| \rightarrow \infty$ .

**Exemple.** Certains estimateurs par projection orthogonale sont  $A$ -admissibles.

(ii) La définition plus générale de [BW] concerne les delta-familles  $\delta_n^*$  de fonctions mesurables bornées sur un ouvert  $O$  de  $\mathbb{R}^2$  telles que  $\int_{y:(x,y) \in O} \delta_n^*(x_0, y) g(y) dy \rightarrow g(x_0)$  pour tout  $(x_0, y_0) \in O$  et  $g$  continue en  $x_0$ . C'est bien le cas de  $\delta_n^*(x, y) = \delta_n(y - x)$  sous les conditions  $A$  ou des estimateurs par projection. La convergence des critères de performance usuels [BR] s'étudie, dans le cadre de cette généralisation, par des techniques reposant sur la régularité des courbes  $\delta_n^*(x_0, \cdot)$  et  $\delta_n^*(\cdot, y_0)$ , qui ont pris un essor particulier dans le cas des ondelettes [HKPT]. En imposant de la régularité à  $f$ , souvent de type Bésou ou Sobolev [BW,CL,Wa], il s'agit d'évaluer les risques sous différentes normes, le risque minimax associé, et de localiser les maxisets de densités  $f$  atteignant des vitesses données.

Le point de vue non paramétrique strict refuse au contraire ces hypothèses sur  $f$ . Pour comprendre ce à quoi s'expose le statisticien face à  $f$  continue, nous adoptons un point de vue stochastique plutôt qu'analytique, celui des mécanismes intrinsèques d'oscillation de  $f_n$  autour de  $\mathbb{E}f_n$ .

(iii) Les résultats de [1,12,21] et leurs prolongements concernent  $f_n - \mathbb{E}f_n$  et généralisent [BL,DM,St2]. Nous n'étudions pas  $\mathbb{E}f_n - f$ , qui nécessite des conditions sur  $f$  dépendant de chaque  $\delta_n$ .

Notre problème est de caractériser le comportement asymptotique pour  $J_n \subset ]m, M[$  de

$$D_n = \sup_{J_n} \frac{|f_n - \mathbb{E}f_n|}{\sqrt{f}}, \quad D_n^\pm = \sup_{J_n} \pm \frac{f_n - \mathbb{E}f_n}{\sqrt{f}}$$

pour  $f > 0$  continue. Nous cherchons d'une part à caractériser les formes admissibles des  $\delta_n$ . Nous explorons d'autre part les cas atypiques de l'estimation sur  $J_n \rightarrow \mathbb{R}$  ou au contraire de  $J_n \downarrow$  très petit.

Les déviations de  $f_n^*(x) = \sum_{i=1}^n \delta_n^*(x, X_i)/n$  se traitent sous des conditions similaires, via les lois fonctionnelles pour  $\alpha_n$  réel, et d'autant mieux si  $\delta_n^*(x, y) = \delta_{1,n}(x) \delta_{2,n}(y - x)$ . Ceci fait partie de nos travaux non rédigés. L'étude de

$$\sup_{J_n} |f_n - \mathbb{E}f_n|$$

se conduit à travers les mêmes preuves mais apporte une information non relative donc un peu moins pertinente quand  $f < 1$ . Si  $J_n = [a, b]$  alors il apparaît mécaniquement un facteur  $\sup_{[a,b]} \sqrt{f}$  dans les constantes limites de 4.1.3. Si  $J_n$  est au contraire de petite entropie, tout dépend de  $J_n$  et les résultats adéquats sont à l'étude en prolongement de [21].

(iv) Les propriétés élémentaires de  $f_n$  établies dans [LW] montrent l'importance de

$$\mu_n^2 = \int_{\mathbb{R}} \delta_n^2.$$

Sous  $A(ii, iii, iv)$  le fait que  $\delta_n \in \mathbb{L}_2$ ,  $n \geq 1$  équivaut à  $A(i)$ . Si  $\delta_n \in \mathbb{L}_2$ ,  $n \geq 1$  alors  $\mu_n \rightarrow \infty$ , la suite  $\delta_n^2/\mu_n^2$  est  $A$ -admissible et  $\mathbb{V}(f_n(x)) \sim f(x)\mu_n/n$ .

**Théorème 4A.** Si  $\delta_n \in L_{2+\rho}$  est  $A$ -admissible,  $\mu_n \ll n^\rho \int_{\mathbb{R}} |\delta_n|^{2+\rho}$ ,  $\rho > 0$  et  $x \in ]m, M[ = \mathbb{R}$  alors la suite  $\sqrt{n/f(x)\mu_n}(f_n(x) - \mathbb{E}f_n(x))$  converge en loi vers une loi normale standard.

Ceci reste vrai si et sup  $|\delta_n| < C\mu_n \ll n$ . En 4.1.2 nous adjoignons à  $\mu_n$  un autre paramètre  $\lambda_n$ , dans le but de généraliser le résultat suivant, dû à [St2,DM]. Dans la suite  $D_n^*$  désigne  $D_n$ ,  $D_n^+$  et  $D_n^-$ .

**Théorème 4B.** Si  $K$  est une densité bornée, à support compact et à variation bornée et si  $h_n$  vérifie (1) avec  $c = d = \infty$  alors pour  $f_n(x) = \sum_{i=1}^n K((X_i - x)/h_n)/nh_n$  et  $J_n = [a, b] \subset ]m, M[$  il vient

$$\lim_{n \rightarrow \infty} \sqrt{\frac{nh_n}{2|\log h_n|}} D_n^* = \|K\|_{\mathbb{L}_2} \quad p.s.$$

Si  $h_n$  vérifie (1) avec  $c < d = \infty$  alors, avec probabilité un, les trois suites

$$\sqrt{\frac{nh_n}{2 \log \log n}} D_n^*$$

admettent  $[\sqrt{c}\|K\|_{\mathbb{L}_2}, \sqrt{1+c}\|K\|_{\mathbb{L}_2}]$  comme ensemble de valeurs d'adhérence.

#### 4.1.2 – Forme des delta-fonctions

(i) Afin de mesurer l'impact des fluctuations d'échantillonnage sur  $D_n^\pm$  nos hypothèses portent sur la manière et la vitesse avec lesquelles  $\delta_n \rightarrow \delta_0$ . Quantifions l'aspect Dirac de  $\delta_n$  par un rayon d'action effectif  $\lambda_n \downarrow 0$  et une surface balayée à l'origine

$$\sigma_n^2 = \int_{[-\lambda_n, \lambda_n]} \delta_n^2 \leq \mu_n^2 = \int_{]m-M, M-m[} \delta_n^2.$$

Il s'agit de garder un équilibre entre  $\lambda_n \downarrow 0$  et  $\sigma_n^2 \uparrow \infty$  tout en donnant une définition implicite à  $\lambda_n$ . Sous les conditions  $B$  nos vitesses sont de l'ordre  $\sqrt{2\sigma_n^2 L_n/n}$  avec  $L_n = |\log \lambda_n| + \log \log n$ .

$$B(i) \quad \frac{\sqrt{\log \log n}}{\sigma_n \sqrt{L_n}} \sup_{|x| > \lambda_n} |\delta_n(x)| \rightarrow 0.$$

$$B(ii) \quad \frac{\sqrt{\log \log n}}{\sigma_n \sqrt{L_n}} \int_{|x| > \lambda_n} |d\delta_n| \rightarrow 0.$$

$$B(iii) \quad \limsup_{n \rightarrow \infty} \frac{\sqrt{\lambda_n}}{\sigma_n} \int_{[-\lambda_n, \lambda_n]} |d\delta_n| < \infty.$$

Les conditions  $B(i, ii)$  renforcent  $A(iii, iv)$  et signifient que la queue relative de  $\delta_n$  est négligeable et peu oscillante. La condition  $B(iii)$  garantit la formation du pic sur  $[-\lambda_n, \lambda_n]$  au sens où la valeur moyenne  $\sigma_n^2/\lambda_n$  de  $\delta_n^2$  est comparable au carré de la variation de  $\delta_n$ . Sous les conditions  $A$  il existe  $\lambda_n$  suffisamment lente pour que  $B(i, ii)$  soient satisfaites mais l'exigence  $B(iii)$  est antagoniste.

**Exemple.** Si  $K$  est un noyau tel que  $\text{supp } K \subset [-k, k]$ ,  $\kappa^2 = \int_{\mathbb{R}} K^2 < \infty$  et  $\int_{\mathbb{R}} |dK| < \infty$  alors  $\lambda_n = kh_n$  est  $B$ -admissible pour  $\delta_n(x) = K(x/h_n)/h_n$  puisque  $\sigma_n^2 = \kappa^2/h_n$  et  $\int_{[-\lambda_n, \lambda_n]} |d\delta_n| = \int_{\mathbb{R}} |dK|/h_n$  donc les suites de  $B(i, ii)$  sont nulles et celle de  $B(iii)$  constante.

**Exemple.** Si  $K_n$  est une suite de noyaux tels que  $\text{supp } K_n \subset [-k_n, k_n]$ ,  $\kappa_n^2 = \int_{\mathbb{R}} K_n^2 < \infty$  et  $\theta_n = \int_{\mathbb{R}} |dK_n| < \infty$  alors  $\lambda_n = k_n h_n$  est  $B$ -admissible pour  $\delta_n(x) = K_n(x/h_n)/h_n$  si  $k_n h_n \rightarrow 0$  et  $k_n \theta_n^2 < C\kappa_n^2$ . Ceci limite les oscillations dans le pic relativement à sa surface. Si  $K_n$  est le triangle de base  $[-k_n, k_n]$  il vient  $C = 1/6$  mais si  $K_n^2$  est le triangle de base  $[-k_n, k_n]$  alors la suite des sommets doit être bornée.

**Contre-exemple.** Aucun noyau  $K$  à support non compact n'est  $B$ -admissible, d'où (iii) ci-après. Les meilleurs résultats sont atteints s'il existe  $\delta \in \mathbb{L}_2$  tel que

$$B(iv) \quad \sup_{x \in \mathbb{R}} |\lambda_n \delta_n(\lambda_n x) - \delta(x)| \rightarrow 0.$$

(ii) Reformulons les conditions  $B$  à l'aide de

$$K_{\delta,n}(x) = \lambda_n \delta_n(\lambda_n x) \mathbb{I}_{]m, M[} \left( \frac{x}{\lambda_n} \right), \quad \kappa_{\delta,n}^2 = \int_{[-1,1]} K_{\delta,n}^2 < \infty$$

donc  $\sigma_n^2 \sim \kappa_{\delta,n}^2 / \lambda_n$ . Sous cet angle,  $B(i)$  devient  $\sup_{|x| \geq 1} |K_{\delta,n}(x)| \ll \kappa_{\delta,n} \sqrt{\lambda_n L_n} / \sqrt{\log \log n}$ ,  $B(ii)$  devient  $\int_{|x| > 1} |dK_{\delta,n}(x)| \ll \kappa_{\delta,n} \sqrt{\lambda_n L_n} / \sqrt{\log \log n}$  et  $B(iii)$  s'écrit  $\int_{|x| \leq 1} |dK_{\delta,n}(x)| < C \kappa_{\delta,n}$ , ce qui illustre bien comment  $B(iii)$  exprime la contrainte de formation du pic relatif.

De même  $B(iv)$  devient  $\sup_{\mathbb{R}} |K_{\delta,n} - K_\delta| \rightarrow 0$  et d'autres conditions recouvrent des cas dignes d'intérêt, par exemple les suivantes.

$$B(v) \quad \sup_{t \in [0, 1/2]} \int_{[-1, 2t+1]} |K_{\delta,n} - K_\delta| \rightarrow 0.$$

$$B(vi) \quad K_{\delta,n} \text{ est adhérent à } \{K_1, \dots, K_m\} \text{ en norme sup.}$$

(iii) Les cas intéressants non  $B$ -admissibles incitent à fournir des conditions suffisantes d'approximation. Pour être  $C$ -admissible il faut qu'il existe pour tout  $\varepsilon > 0$  une suite  $(\delta_{\varepsilon,n}, \lambda_{\varepsilon,n})$  satisfaisant à la fois aux conditions  $B$  et  $C$  suivantes. Soit  $\sigma_{\varepsilon,n}^2 = \int_{[-\lambda_{\varepsilon,n}, \lambda_{\varepsilon,n}]} \delta_{\varepsilon,n}^2$  et  $L_{\varepsilon,n} = |\log \lambda_{\varepsilon,n}| + \log \log n$ .

$$C(i) \quad \|\delta_n - \delta_{\varepsilon,n}\|_{\mathbb{L}_2} < \sqrt{\varepsilon} \sigma_{\varepsilon,n}.$$

$$C(ii) \quad \frac{\sigma_{\varepsilon,n} \sqrt{L_{\varepsilon,n}} \|\delta_n - \delta_{\varepsilon,n}\|}{\sqrt{n} \|\delta_n - \delta_{\varepsilon,n}\|_{\mathbb{L}_2}} \rightarrow 0.$$

$$C(iii) \quad \frac{\sigma_{\varepsilon,n} \sqrt{L_{\varepsilon,n}} \int_{\mathbb{R}} |d(\delta_n - \delta_{\varepsilon,n})|}{\sqrt{n} \|\delta_n - \delta_{\varepsilon,n}\|_{\mathbb{L}_2}} \rightarrow 0.$$

**Exemple (noyaux, cas général).** Soit  $K_n$  une suite de noyaux tels que  $\kappa_n^2 = \int_{\mathbb{R}} K_n^2 < \infty$  et  $\int_{\mathbb{R}} |dK_n| < \infty$ . Les supports sont quelconques. Si  $h_n \rightarrow 0$ ,  $nh_n / \kappa_n^2 \log n \rightarrow \infty$  et

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{\kappa_n^2} \int_{|x| > k} K_n^2(x) dx = \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \int_{|x| > k} |dK_n(x)| = 0$$

alors  $\delta_n(x) = K_n(x/h_n)/h_n$  est  $C$ -admissible. Car il est possible de choisir des noyaux approximatifs

$$\delta_{\varepsilon,n}(x) = K_{\varepsilon,n} \left( \frac{x}{h_n} \right), \quad \lambda_{\varepsilon,n} = k_\varepsilon h_n$$

avec  $K_{\varepsilon,n}$  de support  $[-k_\varepsilon, k_\varepsilon]$ ,  $\int_{\mathbb{R}} (K_n - K_{\varepsilon,n})^2 < \varepsilon$  et  $k_\varepsilon \rightarrow \infty$  si  $\varepsilon \rightarrow 0$ .

**Exemple (Fourier).** Les conditions  $C$  sont vérifiées par les noyaux de Dirichlet et de Fejér,

$$\delta_n^D(x) = \frac{\sin((2j_n + 1)x/2)}{2\pi \sin(x/2)}, \quad \delta_n^F(x) = \frac{\sin^2((j_n + 1)x/2)}{4\pi(j_n + 1) \sin^2(x/2)}$$

obtenus en conservant  $j_n \rightarrow \infty$  coefficients dans les séries trigonométriques associées [Wt, Wi].

**Exemple.** Les estimateurs associés aux fonctions de Hermite [W1] sont  $C$ -admissibles, tout comme certains estimateurs associés à des polynômes orthogonaux [BLc].

### 4.1.3 – Déviations stochastiques exactes

(i) Le résultat principal de [12] s'énonce ainsi.

**Théorème 4.1.** *Si  $J_n = [a, b] \subset ]m, M[$  et  $(\delta_n, \lambda_n)$  est  $B$ -admissible et  $\lambda_n$  vérifie (1) avec  $d = \infty$  alors pour  $\sigma_n^2 = \int_{[-\lambda_n, \lambda_n]} \delta_n^2$  et  $L_n = |\log \lambda_n| + \log \log n$  nous avons*

$$\limsup_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_n^2 L_n}} D_n^\pm \leq \limsup_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_n^2 L_n}} D_n \leq 1 \quad p.s.$$

et

$$\liminf_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_n^2 L_n}} D_n^\pm = \liminf_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_n^2 L_n}} D_n = \sqrt{\frac{c}{1+c}} \quad p.s.$$

Le Théorème 4.1 fournit une bande de confiance exacte pour  $\mathbb{E}f_n$ . Si donc  $c = \infty$  alors

$$D_n^* \sim \sqrt{\frac{2\sigma_n^2 |\log \lambda_n|}{n}} \quad p.s.$$

donc, en supposant que le biais vérifie simplement  $\sup_{[a,b]} |f - \mathbb{E}f_n| \rightarrow 0$ , nous avons avec probabilité un, pour tout  $\varepsilon > 0$  et tout  $n$  assez grand, uniformément sur  $[a, b]$

$$f_n - \sqrt{\frac{(2+\varepsilon)f_n \sigma_n^2 |\log \lambda_n|}{n}} < \mathbb{E}f_n < f_n + \sqrt{\frac{(2+\varepsilon)f_n \sigma_n^2 |\log \lambda_n|}{n}}.$$

**Utilisation.** Pour estimer  $\sup_{[a,b]} |f_n - f|$  il reste à considérer  $\sup_{[a,b]} |f - \mathbb{E}f_n|$  sans plus d'hypothèse sur  $f$ . En pratique, nous disposons pour cela de méthodes suffisamment performantes. Par exemple, en procédant par simulation de  $N$  échantillons  $Y_{j,1}, \dots, Y_{j,n}$  de  $n$  v.a. i.i.d. de densité  $f_n^X$  alors la v.a.  $\sup_{[a,b]} |f - \mathbb{E}f_n|$  s'estime par  $\sup_{[a,b]} |g_n^X - f_n^X|$  avec  $g_n^X = (\sum_{j=1}^N f_{y,n}^Y)/N \sim \mathbb{E}f_n$  où  $f_{y,n}^Y$  et  $f_n^X = \mathbb{E}(f_{y,n}^Y | X_1, \dots, X_n) \sim f$  sont l'estimateur  $f_n$  appliqué respectivement aux  $Y_i$  simulés et aux  $X_i$  initiaux. En effet le biais est bien plus facile à estimer par bootstrap et Monte-Carlo que la déviation stochastique ne le serait si elle n'était pas donnée précisément au Théorème 4.1. Le théorème limite associé, précisant ce bootstrap de la densité, est l'une de nos perspectives dans ce domaine (4.4.4(iii)).

De simples adaptations de nos preuves donnent accès à diverses variantes et extensions. Par exemple, si  $c = \infty$  alors toute fonction de poids  $\varphi > 0$  telle que  $0 < \inf_{[a,b]} \varphi < \sup_{[a,b]} \varphi < \infty$  vérifie

$$\lim_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_n^2 L_n}} \sup_{[a,b]} \pm \frac{f_n - \mathbb{E}f_n}{\varphi} = \sup_{[a,b]} \frac{\sqrt{f}}{\varphi} \quad p.s.$$

De même, les vitesses au second ordre sont en  $O_{p.s.}(1/L_n^{2/3})$ , dans la lim sup et la lim inf.

Si la forme de  $\delta_n$  se stabilise, la lim sup du Théorème 4.1 s'identifie à 1.

**Proposition 4.2.** *Sous les conditions du Théorème 4.1 en  $c < \infty$  et sous  $B(iv)$  ou  $B(vi)$ , l'intervalle  $[\sqrt{c/(1+c)}, 1]$  est presque sûrement l'ensemble des valeurs d'adhérence des suites  $\sqrt{n/2\sigma_n^2 L_n} D_n^*$ .*

Concernant les delta-suites lentes il vient, lorsque  $c = 0$  et  $|\log \lambda_n| > \nu \log \log \log n$  pour  $\nu > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{\sqrt{n} D_n^*}{\sqrt{2\sigma_n^2 \log \log n}} \leq 1 \quad p.s.$$

et

$$\liminf_{n \rightarrow \infty} \frac{\sqrt{n} D_n^*}{\sqrt{2\sigma_n^2 (|\log \lambda_n| - \log \log \log n)}} = 1 \quad p.s.$$

En  $\delta_n(x) = K(x/h_n)/h_n$  les résultats 4.1 et 4.2 restituent bien le Théorème 4B.

**Principe de démonstration.** Après découpage de  $[a, b]$  nous avons, pour  $x \in [x_k, x_k + \varepsilon]$ ,

$$\sqrt{n}(f_n(x) - \mathbb{E}f_n(x)) = \int_{[-\lambda_n, \lambda_n]} \delta_n(y) d\alpha_n^X(x+y) + \int_{\mathbb{R} \setminus [-\lambda_n, \lambda_n]} \delta_n(y) d\alpha_n^X(x+y)$$

avec  $\alpha_n^X = \sqrt{n}(F_n - F)$  et  $F_n$  la fonction de répartition empirique de  $X_1, \dots, X_n$ . Après plusieurs intégrations par parties, une application des lois du logarithme sous les conditions  $B$  montre que le second terme est presque sûrement négligeable. Comme le premier terme est dirigé par les incréments  $\Delta\alpha_n^X(2\lambda_n, x - \lambda_n)$  nous sommes réduits à considérer les incréments uniformes  $\Delta\alpha_n(2f(x_k)\lambda_n, t)$  de  $\Theta_{k,n} \subset \Theta_n$  considérés aux points  $t \in T_{k,n} = [F(x_k) - f(x_k)\lambda_n, F(x_k) + f(x_k)(\varepsilon - \lambda_n)]$ . A ce stade nous disposons de tous les résultats de la partie 2 et nous aboutissons à une situation de type 1.3.2.

Il s'agit alors de bien évaluer les termes d'erreurs stochastiques dans les intégrales et d'identifier la pire oscillation. Celle-ci se produit dès lors que  $d(f_n^*, \Theta_{k,n}) < \varepsilon$  où  $f_n^* \in \mathcal{S}_1$  maximise l'intégrale dominante  $\int_{[-1,1]} K_{\delta,n} df_n^*((1+x)/2)$ . Cette fonction  $f_n^*$  est explicite et dépend uniquement d'une primitive de  $\delta_n$  donc nous pouvons appliquer les théorèmes de 2.1.1 et 2.1.2. Typiquement,  $f_n^* \in \mathcal{S}_1^{vb}$ . Si  $c = \infty$  alors  $d(f_n^*, \Theta_{k,n}) \rightarrow 0$  avec probabilité un et sinon, comme  $J(f_n^*) = 1$ , il est nécessaire de mieux localiser  $f_n^*$  pour contrôler la  $\liminf$ , ce qui est possible sous  $B(iv)$  ou  $B(vi)$ .  $\square$

(ii) Le cadre plus général des delta-suites approximables établit la vitesse de consistance forte pour une famille bien plus large d'estimateurs.

**Proposition 4.3.** *Si  $J_n = [a, b] \subset ]m, M[$  et  $\delta_n$  est  $C$ -admissible avec  $(\delta_{\varepsilon,n}, \lambda_{\varepsilon,n})$  telles que chaque  $\lambda_{\varepsilon,n}$  vérifie (1) avec  $d = \infty$  et un  $c$  unique alors pour tout  $\varepsilon > 0$ ,*

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_{\varepsilon,n}^2 L_{\varepsilon,n}}} D_n^* &\leq 1 + \varepsilon \quad p.s. \\ \liminf_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_{\varepsilon,n}^2 L_{\varepsilon,n}}} D_n^* &\in \left[ \sqrt{\frac{c}{1+c}} - \varepsilon, \sqrt{\frac{c}{1+c}} + \varepsilon \right] \quad p.s. \end{aligned}$$

Par conséquent une delta-suite  $C$ -admissible  $\delta_n$  vérifie le Théorème 4.1 si  $\sigma_{\varepsilon,n} \sim \sigma_n$  et  $L_{\varepsilon,n} \sim L_n$  pour tout  $\varepsilon > 0$  et une suite de référence  $(\sigma_n, L_n)$ , ce qui est le cas dans nos exemples de 4.1.4.

**Principe de démonstration.** Il s'agit de mêler la méthode ci-dessus de réduction aux incréments de  $\alpha_n$  à une extension de la technique d'approximation de [H3].  $\square$

(iii) Le cas des  $J_n$  grands se traite de différentes façons en cours d'étude, la plus simple reposant sur un contrôle de  $f$ . Si  $\inf_{J_n} f \gg 1/L_n^{4/3}$  alors les résultats précédents restent valides.

Le cas des  $J_n$  petits fait appel dans [21] aux lois intermédiaires de 2.3. Prenons l'exemple simple d'une grille de points  $a + k\varepsilon_n$  où  $f$  est estimée,  $k = 0, 1, \dots, [(b-a)/\varepsilon_n]$ .

**Proposition 4.4.** *Si  $(\delta_n, \lambda_n)$  est  $B$ -admissible avec  $\lambda_n \rightarrow 0$ ,  $n\lambda_n \rightarrow \infty$ ,  $\varepsilon_n \rightarrow 0$ ,  $\varepsilon_n/\lambda_n \rightarrow \infty$  et  $|\log \varepsilon_n|/\log \log n \rightarrow c$  alors les conclusions du Théorème 4.1 valent pour  $L_n = |\log \varepsilon_n| + \log \log n$  et  $J_n = \{a + k\varepsilon_n : k \leq (b-a)/\varepsilon_n\}$ .*

Le cas  $c = 0$  se développe plus avant. Si  $|\log \varepsilon_n| \sim |\log \lambda_n|$  les déviations de  $f_n$  ne diminuent pas. Par contre, pour un estimateur à noyau de fenêtre  $h_n = 1/n^\gamma$  et une grille en  $\varepsilon_n = 1/n^\beta$  avec  $\beta < \gamma$  la déviation presque sûre est contractée d'un facteur  $\sqrt{\beta/\gamma}$ .

Si  $J_n$  est un intervalle, déterministe ou aléatoire, respectant certaines conditions alors nous pouvons aussi contrôler  $D_n^*$ . Prenons le cas le plus simple.

**Proposition 4.5.** *Si  $(\delta_n, \lambda_n)$  est  $B$ -admissible avec  $\lambda_n \rightarrow 0$ ,  $n\lambda_n \rightarrow \infty$ ,  $\varepsilon_n \rightarrow 0$ ,  $\varepsilon_n/\lambda_n \rightarrow \infty$  et  $|\log(\varepsilon_n/\lambda_n)|/\log \log n \rightarrow c$  alors les conclusions du Théorème 4.1 valent pour  $L_n = |\log(\varepsilon_n/\lambda_n)| + \log \log n$  et  $J_n = [a, a + \varepsilon_n]$ .*

Plus généralement, sous des conditions sur  $J_n$ ,  $b_n = b_n(J_n, \lambda_n)$  et  $\varphi$  il existe  $C = C(f, \{J_n\}, \varphi)$  tel que

$$\limsup_{n \rightarrow \infty} \sqrt{\frac{n}{2\sigma_n^2(\log b_n + \log \log n)}} \sup_{J_n} \pm \frac{f_n - \mathbb{E}f_n}{\varphi} = C \quad p.s.$$

(iv) Nous avons des résultats similaires pour l'estimation de la dérivée  $p$ -ième de  $f \in \mathcal{C}_p$  par

$$f_n^{(p)}(x) = \frac{(-1)^p}{n} \sum_{i=1}^n \delta_n^{(p)}(X_i - x)$$

avec  $\delta_n \in \mathcal{C}_p$ , ce qui généralise [Si,Pr]. Soit  $\kappa_{p,\delta,n}^2 = \int_{[-1,1]} (K_{\delta,n}^{(p)})^2$  pour  $K_{\delta,n}$  de 4.1.2(ii). Si  $\lambda_n$  est  $B$ -admissible pour  $f_n^{(p)}$  et vérifie (1) avec  $c = \infty$  alors le Théorème 4.1 devient

$$\lim_{n \rightarrow \infty} \frac{\sqrt{n} \lambda_n^{p+1/2}}{\sqrt{2\kappa_{p,\delta,n}^2 L_n}} \sup_{[a,b]} \pm \frac{f_n^{(p)} - \mathbb{E}f_n^{(p)}}{\sqrt{f}} = 1 \quad p.s.$$

(v) Dans les résultats précédents, si  $c < \infty$  nous savons estimer le temps d'attente de la prochaine déviation maximale. Sous les hypothèses de la Proposition 4.2 nous avons, pour tout  $\varepsilon > 0$  et un  $\rho_\varepsilon > 0$  ne dépendant que de  $\delta$  dans  $B(iv)$  ou des  $\{K_1, \dots, K_2\}$  dans  $B(vi)$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{\log \log n} \log \log \left( \frac{1}{n} \inf \left\{ m > n : D_m^* > \sqrt{2(1-\varepsilon)\sigma_m^2 L_m/m} \right\} \right) = \rho_\varepsilon \quad p.s.$$

La constante  $\rho_\varepsilon$  est en rapport avec le problème des petites déviations browniennes 3.1 pour la primitive de  $K_\delta$  ou des  $K_i$ .

#### 4.1.4 – Quelques vitesses explicites

(i) Nous généralisons donc [H3] à une suite de noyaux, par des énoncés du type suivant.

**Corollaire 4.6. (Noyaux, cas général.)** *Soit  $K_n$  une suite de noyaux vérifiant les conditions de l'exemple 4.1.2(iii),  $B(vi)$  et  $h_n$  vérifiant (1). Alors, avec probabilité un, la suite*

$$\sqrt{\frac{nh_n}{2\kappa_n^2(|\log h_n| + \log \log n)}} \sup_{[a,b]} \frac{|f_n - \mathbb{E}f_n|}{\sqrt{f}}$$

est relativement compacte dans  $\mathbb{R}$  d'ensemble de valeurs d'adhérence  $[\sqrt{c/(1+c)}, 1]$ . Si de plus  $\varepsilon_n \rightarrow 0$ ,  $\varepsilon_n/h_n \rightarrow \infty$  et  $|\log \varepsilon_n|/\log \log n \rightarrow c_0$  alors la suite

$$\sqrt{\frac{nh_n}{2\kappa_n^2(|\log \varepsilon_n| + \log \log n)}} \sup_{[a,b] \cap \{a+k\varepsilon_n\}} \frac{|f_n - \mathbb{E}f_n|}{\sqrt{f}}$$

est relativement compacte d'ensemble de valeurs d'adhérence  $[\sqrt{c_0/(1+c_0)}, 1]$  et la suite

$$\sqrt{\frac{nh_n}{2\kappa_n^2 f(a)(|\log h_n| - |\log \varepsilon_n| + \log \log n)}} \sup_{[a, a+\varepsilon_n]} |f_n - \mathbb{E}f_n|$$

est de limite 1 si  $c_0 = \infty$  et d'ensemble limite  $[\sqrt{(c-c_0)/(1+c-c_0)}, 1]$  sinon.

(ii) Nous obtenons la loi du logarithme pour les estimateurs basés sur les séries de Fourier [Wt,Wi]. Supposons que  $[m, M] = [-\pi, \pi]$ . Les coefficients de Fourier empiriques sont définis par

$$\widehat{c}_{j,n} = \frac{1}{n} \sum_{i=1}^n \cos(jX_i), \quad \widehat{s}_{j,n} = \frac{1}{n} \sum_{i=1}^n \sin(jX_i).$$

Soit  $-\pi < a < b < \pi$ . Considérons les  $j_n$  premiers coefficients, avec

$$j_n \rightarrow \infty, \quad \frac{j_n}{n} \rightarrow 0, \quad \frac{\log j_n}{\log \log n} \rightarrow c \in [0, \infty]. \quad (7)$$

**Corollaire 4.7. (Dirichlet.)** *Soit  $c_D^2 = \int_{[-\pi,\pi]} \sin^2(x)/x^2 dx$ . Sous (7),*

$$f_n^D(x) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{j=1}^{j_n} (\widehat{c}_{j,n} \cos(jx) + \widehat{s}_{j,n} \sin(jx))$$

est tel qu'avec probabilité un, la suite

$$\sup_{x \in [a,b]} \pm \frac{\sqrt{n} (f_n^D(x) - \mathbb{E}f_n^D(x))}{\sqrt{j_n f(x) (\log j_n + \log \log n)}}$$

est relativement compacte dans  $\mathbb{R}$  et adhère à  $[c_D \sqrt{2c}/\pi \sqrt{1+c}, c_D \sqrt{2}/\pi]$ .

**Corollaire 4.8. (Fejér.)** Soit  $c_F^2 = \int_{[-\pi, \pi]} \sin^4(x)/x^4 dx$ . Sous (7),

$$f_n^F(x) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{j=1}^{j_n} \left(1 - \frac{j}{j_n + 1}\right) (\hat{c}_{j,n} \cos(jx) + \hat{s}_{j,n} \sin(jx))$$

est tel qu'avec probabilité un, la suite

$$\sup_{x \in [a,b]} \pm \frac{\sqrt{n} (f_n^F(x) - \mathbb{E}f_n^F(x))}{\sqrt{j_n f(x) (\log j_n + \log \log n)}}$$

est relativement compacte dans  $\mathbb{R}$  et adhère à  $[c_F \sqrt{2c}/\pi \sqrt{1+c}, c_F \sqrt{2}/\pi]$ .

(iii) Nos résultats peuvent servir à calibrer des procédures d'estimation adaptatives. Considérons des delta-suites  $\delta_{g,n}$  paramétrées par  $g \in \mathcal{G}$  comme évoqué en 4.4.4(ii). En comparant les vitesses du Théorème 4.1 à une estimation de  $\sup_{[a,b]} |f - \mathbb{E}f_{g,n}|$  par simulation basée sur chaque  $f_{g,n}$  ou une sous-famille finie approximante, il suffit d'ajuster les paramètres  $g$  pour réaliser l'équilibre entre biais estimé et déviation presque sûre exacte, après quelques itérations dans l'adaptation de  $\delta_{g,n}$ . De cette façon nous sélectionnons le  $\delta_{g,n}$  le mieux adapté parmi les autorisés et contrôlons  $\sup_{[a,b]} |f_{g,n} - f|$ . Le prix à payer réside dans la perte de précision due à l'estimation du terme de biais.

Ainsi, étant donné les  $\delta_{g,n}$  nos résultats fournissent le comportement exact de  $\sup_{[a,b]} |f_{g,n} - \mathbb{E}f_{g,n}|$  qui est difficilement estimable et de nombreuses stratégies adaptatives, par bootstrap ou autre, basées sur une estimation du biais, peuvent choisir  $g$  et révéler la meilleure vitesse presque sûre exacte de  $\sup_{[a,b]} |f_{g,n} - f|$ . Ceci s'applique en particulier à la sélection automatique de fenêtre  $h_n$  et de noyau dans l'estimation à noyau, mais sans hypothèse spécifique sur  $f$  pour déterminer d'avance une borne de biais  $\sup_{\mathcal{G}} \sup_{[a,b]} |f - \mathbb{E}f_{g,n}|$ .

Ce point de vue, qui consiste à déterminer une fois pour toutes la déviation stochastique inévitable pour autoriser différents types d'induction pour sélectionner automatiquement la meilleure vitesse de convergence, est aussi celui de 4.2 et 4.3.

## 4.2 – Estimation du shorth

### 4.2.1 – Hypothèses et notations

Soit  $\{X_n\}$  une suite i.i.d. de v.a. réelles de fonction de répartition  $F$  et de densité  $f$  par rapport à  $\lambda$ . Dans [10] nous considérons le problème de la localisation presque sûre du plus petit intervalle de masse  $q \in ]0, 1[$  fixée, appelé  $q$ -shorth. Pour en appréhender le mécanisme d'oscillation nous avons mis au point une technique tirant profit des résultats fins sur les processus empiriques et browniens.

(i) Nous faisons sur  $f$  les hypothèses peu restrictives suivantes. Il existe un unique couple

$$\{(x^-, x^+)\} = \arg \min_{(x,y)} \{y - x : F(y) = F(x) + q\}.$$

Il existe des intervalles ouverts  $A^-$  et  $A^+$  contenant  $x^-$  et  $x^+$  respectivement tels que sur  $A = A^- \cup A^+$ ,  $F$  est différentiable de dérivée  $f$  et  $\inf_{x \in A} f(x) > 0$ . La fonction  $f$  est continue en  $x^-$  et  $x^+$ . Enfin, la fonction de quantile  $F^{-1} = \inf \{x : F(x) \geq I\}$  est continue sur  $]0, 1[$ . Notons ses incréments

$$d(t) = F^{-1}(t + q) - F^{-1}(t), \quad t \in [0, 1 - q].$$

Posons  $m = f(x^\pm)$  et  $p = F(x^-)$ , de sorte que  $p + q = F(x^+)$ . Appelons  $p$  la hauteur du shorth et  $d = x^+ - x^-$  sa longueur. Ainsi  $d(t)$  atteint son unique minimum  $d(p) = d$  en  $p$  et  $y$  est dérivable, de dérivée  $d'(p) = 0$ . De plus, sur un voisinage  $B^- \subset F(A^-)$  de  $p$ ,  $d(t) = o(|t - p|)$ . Au contraire, pour un  $\rho > 0$  et  $t \notin F(A^-)$  il vient  $d(t) \geq d + \rho$ .

(ii) Soit  $(X_{(1)}, \dots, X_{(n)})$  la statistique d'ordre de  $X_1, \dots, X_n$  et  $d_{n,j} = X_{(j+[nq])} - X_{(j)}$  les espacements. Les estimateurs non paramétriques naturels de  $d$  et  $p$  sont respectivement

$$\widehat{d}_n = \min_{1 \leq j \leq n-[nq]} d_{n,j}$$

et

$$\widehat{p}_n = \frac{1}{n} \min \left\{ j : d_{n,j} = \widehat{d}_n \right\}.$$

Le comportement local de  $F^{-1}$  en  $x^\pm$  détermine le comportement asymptotique de  $\widehat{d}_n$  et  $\widehat{p}_n$ , comme observé initialement dans [JN], dont nous reprenons les hypothèses.

Pour contrôler  $\widehat{d}_n$  nous demandons l'existence de  $\beta > 0$  tel que

$$(H_d) \quad \liminf_{t \rightarrow p} \frac{d(t) - d}{|t - p|^{1+\beta}} > 0.$$

Pour contrôler  $\widehat{p}_n$  nous demandons, plus précisément, que pour tout  $|\varepsilon|$  assez petit,

$$(H_p^-) \quad F^{-1}(p + \varepsilon) - x^- = \frac{\varepsilon}{m} - (r^- + g_-(\varepsilon)) |\varepsilon|^{1+\beta_-}$$

$$(H_p^+) \quad F^{-1}(p + q + \varepsilon) - x^+ = \frac{\varepsilon}{m} + (r^+ + g_+(\varepsilon)) |\varepsilon|^{1+\beta_+}$$

avec  $\beta_\pm > 1/2$ ,  $r^\pm \geq 0$ ,  $r^- + r^+ > 0$  et  $g_\pm$  continues telles que  $\lim_{\varepsilon \rightarrow 0} g_\pm(\varepsilon) = 0$ . Dans  $(H_p^\pm)$  nous pouvons aussi laisser  $\beta_\pm$  et  $r^\pm$  dépendre du signe de  $\varepsilon$  mais ceci alourdit les notations. Posons  $\beta = \min(\beta^-, \beta^+)$ ,  $r = r^- + r^+$  si  $\beta^- = \beta^+$ ,  $r = r^-$  si  $\beta^- < \beta^+$  et  $r = r^+$  si  $\beta^+ < \beta^-$ .

#### 4.2.2 – Lois du logarithme itéré

(i) La convergence faible de  $(\widehat{d}_n, \widehat{p}_n)$  est obtenue dans [JN] en approximant les quantiles empiriques par des sommes de v.a. exponentielles, contrôlant les grandes déviations de ces sommes et reconstruisant le processus limite par Donsker. Soit  $\{W(t) : t \in \mathbb{R}\}$  un mouvement brownien standard.

**Théorème 4C.** *Sous  $(H_d)$  la suite  $\sqrt{n}(\widehat{d}_n - d)$  a pour limite en loi la v.a. gaussienne centrée de variance  $q(1 - q)/m^2$ . Sous  $(H_p^-)$  et  $(H_p^+)$  la suite  $n^{1/(2\beta+1)}(\widehat{p}_n - p)$  converge en loi vers*

$$\arg \min_{t \in \mathbb{R}} \left( W(t) + \frac{rm}{\sqrt{2}} |t|^{1+\beta} \right).$$

*Ces deux suites aléatoires sont asymptotiquement indépendantes.*

Sous des conditions plus restrictives sur  $f$ , [Gb] explore la convergence forte de  $\widehat{d}_n$  et [Lc] de  $\widehat{p}_n$ .

(ii) La méthode [10] clarifie et approfondit [JN,Lc] en se ramenant aux processus fondamentaux, comme expliqué en 4.2.3. Ceci fournit des bornes de déviation plus précises, la loi du logarithme itéré avec constante exacte et, accessoirement, une preuve alternative du Théorème 4C.

**Théorème 4.9.** *Sous  $(H_d)$  nous avons*

$$\limsup_{n \rightarrow \infty} \pm \frac{\sqrt{n}(\widehat{d}_n - d)}{\sqrt{\log \log n}} = \frac{\sqrt{2q(1 - q)}}{m} \quad p.s.$$

*Sous  $(H_p^-)$  et  $(H_p^+)$  il vient*

$$\limsup_{n \rightarrow \infty} \pm \left( \frac{n}{\log \log n} \right)^{1/(2\beta+1)} (\widehat{p}_n - p) = c_{F,q} = \left( \frac{4(2\beta + 1)}{r^2(\beta + 1)^2 m^2} \right)^{1/(2\beta+1)} \quad p.s.$$



Lorsque  $f \in \mathcal{C}_1$  est symétrique, donc  $\beta = 1$ , nous retrouvons [Lc]. Pour des cas où  $f$  est irrégulière, les exemples de 4.3 fournis pour le mode s'adaptent ici. Donnons deux applications du Théorème 4.9.

(iii) Supposons observés  $X_1, \dots, X_n$  i.i.d. de fonction de répartition  $F_\theta = F(\cdot - \theta)$  avec  $\theta \in \mathbb{R}$  un paramètre de localisation à estimer et  $F$  vérifiant les hypothèses précédentes, dont  $(H_p^-)$  et  $(H_p^+)$ . Si seul le  $q$ -shorth  $x_-$  de  $F$  est connu alors celui de  $F_\theta$  vaut  $x_- + \theta$  et il est naturel d'invoquer

$$\widehat{\theta}_n = X_{([n\widehat{p}_n])} - x_-.$$

Par la transformation des quantiles, la loi du logarithme itéré pour  $F^{-1}$  et le Théorème 4.9 il vient

$$X_{([n\widehat{p}_n])} = F^{-1}\left(p + (\widehat{p}_n - p) + (U_{([n\widehat{p}_n])} - \widehat{p}_n)\right) = x_- + \frac{\widehat{p}_n - p}{m} + o_{p.s.}\left(v_n^{1/(2\beta+1)}\right) + O_{p.s.}(\sqrt{v_n})$$

où  $v_n = (\log \log n)/n$  et  $U_{(1)} \leq \dots \leq U_{(n)}$  est un échantillon i.i.d. ordonné de loi uniforme sur  $[0, 1]$ .

**Corollaire 4.10.** *Avec probabilité un,*

$$\limsup_{n \rightarrow \infty} \pm \left(\frac{n}{\log \log n}\right)^{1/(2\beta+1)} (\widehat{\theta}_n - \theta) = \frac{c_{F,q}}{m}.$$

(iv) Supposons que  $f$  vérifie toutes les hypothèses précédentes et possède un mode  $\theta$  dont seul l'indice de position  $\delta$  dans le  $q$ -shorth est connu, au sens où  $\theta = x_- + \delta d = (1 - \delta)x_- + \delta x_+$  avec  $(x_-, x_+, d)$  inconnus. Le Théorème 4.9 permet de contrôler comme ci-dessus l'estimateur non paramétrique

$$\widehat{\theta}_{0,n} = X_{([n\widehat{p}_n])} + \delta \widehat{d}_n.$$

**Corollaire 4.11.** *Avec probabilité un,*

$$\limsup_{n \rightarrow \infty} \pm \left(\frac{n}{\log \log n}\right)^{1/(2\beta+1)} (\widehat{\theta}_{0,n} - \theta) = \frac{c_{F,q}}{m}.$$

Le cas où  $f \in \mathcal{C}_2$  est symétrique par rapport à  $\theta$  est étudié dans [A+,SW] sans vitesse de convergence presque sûre, et ce problème laissé ouvert. Or nos hypothèses sont vérifiées, avec  $\beta = 1$ , et  $\widehat{\theta}_{0,n} = (X_{([n\widehat{p}_n]+[nq])} + X_{([n\widehat{p}_n])})/2$  est plus simple que l'estimateur  $\widetilde{\theta}_{0,n}$  de [A+]. Le Corollaire 4.11 fournit pour  $\widehat{\theta}_{0,n}$  la vitesse  $((\log \log n)/n)^{1/3}$  et la constante  $c_{F,q}/m = (3m/s^2)^{1/3}$  avec  $s = |f^{(1)}(x^\pm)|$ . La même vitesse et une autre constante s'en déduisent pour  $\theta_{0,n}$  via la transformation utilisée dans [SW]. De plus, vitesse et constante pour  $\widehat{\theta}_{0,n}$  restent inchangées si  $f$  n'est ni symétrique ni  $\mathcal{C}_2$  pourvu que  $f \in \mathcal{C}_1$  et  $f(x^\pm + \varepsilon) = m \pm s\varepsilon + g^\pm(\varepsilon)$  avec  $g^\pm(\varepsilon)$  régulière puisqu'alors  $(H_p^-)$  et  $(H_p^+)$  sont identiques.

### 4.2.3 – Une nouvelle méthode

La preuve des Théorèmes 4.9 et 4.12 est fastidieuse mais innovante et très prometteuse, voir 4.4.1. Elle nous a longtemps résisté donc beaucoup appris sur le mécanisme oscillatoire au coeur de ce problème.

(i) L'argument réglant  $\sqrt{n}(\widehat{d}_n - d)$  et  $\sqrt{n}(\widehat{\Delta}_n - \Delta)$  provient directement des lois fonctionnelles de type 2.1 pour  $\beta_n$  couplées à une adaptation des bornes de grandes déviations de [JN]. En travaillant plus nous avons accès au second ordre de convergence, en tenant compte de nos vitesses 2.1, et au temps d'attente presque sûr de la prochaine déviation maximum, en tenant compte de 2.1.6(ii).

(ii) La difficulté provient de  $\widehat{p}_n$  et  $\widehat{x}_n^-$ . Contrairement à l'usage nous ne cherchons pas à appliquer un théorème d'argmax, argument confus quand tout est aléatoire et imprécis quant aux vitesses, mais à contrôler directement une v.a. de type argmax plus simple que  $\widehat{p}_n$ .

Présentons brièvement cette manipulation, en posant  $p_n = p$ ,  $q_n = q$  et  $r_n = r$  pour le  $q$ -shorth. Soit  $v_n = (r_n^2 n)^{1/(2\beta+1)}$  et  $t \in I_n = v_n] - p_n, 1 - p_n - q[$ . Par hypothèse,  $I_n \rightarrow \mathbb{R}$ . Nous cherchons à localiser l'entier  $\widehat{j} = [n(p_n + \widehat{t}/v_n)] = n\widehat{p}_n$  qui atteint le minimum en  $j$ , négatif, de

$$\widehat{\Delta}_{n,j} - \widehat{\Delta}_{n,[np_n]} = X_{(j+[nq_n])} - X_{(j)} - (X_{([np_n]+[nq_n])} - X_{([np_n])})$$

où, en notant  $\beta_n = \sqrt{n}(F_n^{-1} - F^{-1})$  le processus de quantile uniforme et  $X_j = F^{-1}(U_j)$ ,

$$X_{(j)} = F^{-1}(U_{(j)}) = F^{-1}\left(\frac{\beta_n(p_n + t/v_n)}{\sqrt{n}} + p_n + \frac{t}{v_n}\right).$$

La première étape consiste à utiliser la représentation locale de  $F^{-1}$ , l'approximation forte de  $\beta_n$  via [KMT] et un changement d'échelle de temps pour montrer que  $\hat{t}$  minimise

$$\frac{\Delta K(t, n)}{nf(x_n^\pm)} + |t|^{1+\beta} + \xi(t, n)$$

où  $\Delta K(t, n)$  est la différence de deux incréments d'un processus de Kiefer, l'un voisin de  $p_n$ , l'autre près de  $p_n + q_n$ , et  $\xi(t, n)$  est le processus de somme des erreurs successives d'approximation (analytiques, stochastique, espacements, loi du logarithme itéré). Un point crucial est que le contenu exact de  $\xi(t, n)$  change selon les  $t$ . L'événement considéré est alors la grande déviation de l'arg min d'un mouvement brownien avec dérive aléatoire dépendante. Nous découpons ce processus selon des zones déterministes  $I_{\kappa, n}$  bien choisies et comparons deux à deux les variables

$$\min_{t \in I_{\kappa, n}} \left( \frac{\Delta K(t, n)}{nf(x_n^\pm)} + |t|^{1+\beta} + \xi_\kappa(t, n) \right)$$

où les termes dominants dans la dérive  $\xi_\kappa$  dépendent de la zone  $\kappa = 1, 2, 3$ , etc. Par exemple, décrivons informellement la zone critique. Avec grande probabilité, si  $|t| \ll 1/(\log \log n)^{1/(2\beta+1)}$  alors

$$|\xi(t, n)| \ll |t|^{1+\beta} \ll \left| \frac{\Delta K(t, n)}{nf(x_n^\pm)} \right|$$

et, si  $|t| \gg 1/(\log \log n)^{1/(2\beta+1)}$  et  $|t| < t_n$  alors

$$|\xi(t, n)| \ll \left| \frac{\Delta K(t, n)}{nf(x_n^\pm)} \right| \ll |t|^{1+\beta}$$

donc, en rechangeant l'échelle de temps via  $v_n$ , l'événement rare le plus probable est distribué comme la déviation de  $\arg \min_{|t| < t_n} (W + |t|^{1+\beta})$ , où  $W$  est un mouvement brownien. D'où la nécessité du principe de grande déviation 3.3.1. Les changements d'espace-temps imposent l'ordre de grandeur de  $v_n$ , du fait de la probabilité de comparaison des deux  $\min_{t \in I_{\kappa, n}}$  les plus proches,  $\kappa = 1, 2$ .

Pour les  $t$  plus éloignés de 0,  $|t| \gg t_n$ , les probabilités de déviations se contrôlent par les bornes de type [JN]. Le second séparateur de zones étant  $t_n = 1/(\log n)^{1/(2\beta+1)}$  la convergence de la série des probabilités secondaires nécessite des arguments de blocage en temps  $n \in [n_k, n_{k+1}[$ . Or ceci requiert un contrôle en espace-temps de tous les processus en jeu, donc l'évaluation des déviations de

$$\sup_{n \in [n_k, n_{k+1}[} \sup_{t \in I_{\kappa, n}} |\Delta K(t, n) - \Delta K(t, n_k)|, \quad \sup_{n \in [n_k, n_{k+1}[} \sup_{t \in I_{\kappa, n}} |\xi(t, n)|$$

avec  $\Delta K(t, n)$  de longueur  $h_n$ . C'est l'usage dans toutes nos techniques, mais ici survient une difficulté inattendue. Il faut disposer, pour le blocage en espace, d'un théorème de convergence uniforme des incréments du processus de Kiefer ou, de façon alternative, de  $\beta_n$ , localisés dans une suite d'unions de très petits intervalles  $T_n \rightarrow \{p, p+q\}$  ou  $T_n \rightarrow \{\theta\}$  et normalisés par  $\sqrt{2a_n L_n}$  avec  $L_n = \log \log n$  et non pas  $L_n = \log n$  comme au Théorème 2B. D'où le besoin impérieux, pour résoudre cette difficulté technique profonde, d'une loi du logarithme intermédiaire sur les oscillations, le Théorème 2.34.

Une organisation judicieuse des événements permet enfin de conclure par Borel-Cantelli cette preuve de la borne supérieure. La borne inférieure est moins délicate, il suffit de considérer sans blocage les événements le long d'une suite  $n_k$  assez rapide pour se ramener par renouvellement à des événements indépendants, en maîtrisant de nouveau les écarts induits via les lois fonctionnelles sur les incréments des processus en jeu, et en concluant par le Théorème 3.12.  $\square$

### 4.3 – Estimation du mode

#### 4.3.1 – Position du problème

Soit  $\{X_n\}$  une suite i.i.d. de v.a. réelles de loi absolument continue  $P = \mathbb{P}^X$ , de fonction de répartition  $F$ . Supposons que la densité  $f$  est continue sur un voisinage de son mode, supposé unique,

$$\theta = \arg \max f.$$

Le mode est un paramètre de localisation usuel [A+] qui s'estime à des vitesses atypiques [KP].

(i) Les estimateurs de type  $\hat{\theta} = \arg \max \hat{f}$  où  $\hat{f}$  est un estimateur de la densité  $f$  sont les plus populaires. Dans le cas d'estimateurs à noyau [Ed,Ro],  $f$  est supposée très régulière pour contrôler uniformément  $\mathbb{E}\hat{f} - f$  mais seules des bornes supérieures de type  $|\hat{\theta} - \theta| = O_{p.s.}(1/n^\gamma)$  sont accessibles, avec  $\gamma$  fonction de la régularité de  $f$ . Cette approche est donc restrictive pour  $P$  et n'est pas adaptée à l'estimation d'un mode prononcé. En outre, calculer  $\hat{f}$  puis rechercher  $\arg \max \hat{f}$  au degré de précision souhaité est un algorithme de calcul gourmand. Les estimateurs que nous étudions sont plus simples, font aussi bien dans les cas réguliers, ont des vitesses exactes et recouvrent des situations irrégulières.

(ii) Dans [15] nous étudions des estimateurs de  $\theta$  basés sur le plus court intervalle de mesure empirique  $q_n \rightarrow 0$ , appelé  $q_n$ -shorth. A partir des accroissements de la fonction de quantile inconnue  $F^{-1}$ ,

$$\Delta_{n,s} = F^{-1}(s + q_n) - F^{-1}(s), \quad s \in ]0, 1 - q_n[$$

nous définissons la hauteur du plus court intervalle de  $P$ -masse  $q_n$ ,

$$p_n = \arg \min_{s \in ]0, 1 - q_n[} \Delta_{n,s}$$

ses bornes

$$F^{-1}(p_n) = x_n^- < x_n^+ = F^{-1}(p_n + q_n)$$

et sa longueur

$$\Delta_n = \min_{s \in ]0, 1 - q_n[} \Delta_{n,s} = x_n^+ - x_n^-.$$

Ainsi  $p_n \rightarrow F(\theta)$ ,  $x_n^\pm \rightarrow \theta$  et  $\Delta_n \rightarrow 0$  lorsque  $q_n \rightarrow 0$ . Introduisons les estimateurs naturels de ces quantités. Soit  $(X_{(1)}, \dots, X_{(n)})$  la statistique d'ordre de  $(X_1, \dots, X_n)$ . Les espacements

$$\hat{\Delta}_{n,j} = X_{(j+[nq_n])} - X_{(j)}, \quad j = 1, \dots, n - [nq_n]$$

déterminent l'intervalle le plus court de  $\mathbb{P}_n$ -masse  $q_n$  via sa longueur

$$\hat{\Delta}_n = \min_{1 \leq j \leq n - [nq_n]} \hat{\Delta}_{n,j}$$

puis sa hauteur

$$\hat{p}_n = \frac{1}{n} \min \left\{ j : \hat{\Delta}_{n,j} = \hat{\Delta}_n \right\}$$

et sa localisation

$$X_{(n\hat{p}_n)} = \hat{x}_n^- < \hat{x}_n^+ = \hat{x}_n^- + \hat{\Delta}_n = X_{(n\hat{p}_n + [nq_n])}.$$

Il suffit de considérer  $(\hat{\Delta}_n, \hat{x}_n^-)$  puisque  $\hat{x}_n^- = F_n^{-1}(\hat{p}_n) \sim \hat{p}_n/f(\theta)$ . Les temps de calcul sont dérisoires par rapport à ceux d'un estimateur de  $f$  sur un large intervalle. Cependant l'estimateur  $\hat{f}$  dit des  $nq_n$ -voisins est lié au  $q_n$ -shorth par  $(\hat{x}_n^- + \hat{x}_n^+)/2 = \hat{\theta} = \arg \max \hat{f}$  et oscille comme  $\hat{x}_n^-$ . Mais il n'est pertinent que si  $f$  est localement symétrique par rapport à  $\theta$ , voir 4.3.4.

(iii) Les travaux sur le  $q_n$ -shorth sont peu nombreux, sa définition par minimisation rendant difficile son étude. Dans [15] et des extensions en cours nous abordons les trois problèmes suivants.

Une première question concerne les hypothèses sur  $P$  et notamment les conditions de régularité sur  $f$ . Contrairement à la plupart des travaux en estimation de la densité et du mode nous ne demandons à  $f$  ni différentiabilité ni développement limité sur  $\mathcal{V}(\theta)$ . Nos conditions prennent en compte les singularités en  $\theta$ , par exemple  $f^{(1)}(\theta - \varepsilon) \sim -f^{(1)}(\theta + \varepsilon) \sim \infty$ . Conformément à l'intuition, les vitesses de convergence vers  $\theta$  sont d'autant plus rapides que  $f$  est pointue en  $\theta$ . L'exigence que  $q_n$  doit être lent quand  $f$  est plate en  $\theta$  et rapide en cas de pic en  $\theta$  est quantifiée en 4.3.2.

Notre second problème, et principale difficulté, consiste à déterminer le comportement presque sûr exact de  $\widehat{x}_n^\pm$  et  $\widehat{\Delta}_n$  étant donnée la suite déterministe  $q_n \rightarrow 0$ . Pour cela nous avons dû raffiner nos résultats sur le comportement oscillatoire de  $\beta_n = \sqrt{n}(F_n^{-1} - F^{-1})$  et du mouvement brownien dans des petites zones aléatoires. C'est le prix de la définition implicite du  $q_n$ -shorth. Notre travail se scinde en deux, car dans les cas plats  $\widehat{x}_n^-$  oscille avec plus grande amplitude que  $\widehat{\Delta}_n$  et dans les cas pointus  $\widehat{\Delta}_n$  est plus incertain que  $\widehat{x}_n^-$ . Cette cassure dépend de  $q_n$  et  $F$ , voir 4.2.3, 4.3.3 et 4.4.1.

La troisième interrogation, et principale motivation, concerne le choix de  $q_n$  et le comportement limite d'estimateurs  $\widehat{\theta}_n = \varphi_n(\widehat{\Delta}_n, \widehat{x}_n^-)$  de  $\theta$ , que ce soient ceux proposés dans la littérature ou ceux suggérés par notre compréhension des fluctuations. Il s'agit d'identifier dans chaque cas ce qui, dans les hypothèses, permet de localiser le mode dans l'intervalle  $[\widehat{x}_n^-, \widehat{x}_n^+]$ . Ceci afin de déterminer le biais  $\widehat{\theta}_n - \theta$  et de le comparer à la déviation stochastique due à  $\widehat{\Delta}_n$  et  $\widehat{x}_n^-$ . L'optimisation des deux termes identifie  $q_n$ , voir 4.3.4. Ainsi notre contrôle stochastique exact de  $[\widehat{x}_n^-, \widehat{x}_n^+]$  ouvre un champ aussi large qu'il y a de jeux d'hypothèses sur  $\mathbb{P}^X$  et d'estimateurs de type  $\widehat{\theta}_n = \varphi(\widehat{x}_1^-, \widehat{x}_1^+, \dots, \widehat{x}_n^-, \widehat{x}_n^+)$ , voir 4.3.4.

### 4.3.2 – Représentation quantile

Notre approche fait naturellement porter la régularité à  $F^{-1}$  plutôt qu'à  $f$ , ce qui autorise des  $f$  irrégulières. Les hypothèses usuelles sur  $f$  impliquent les nôtres sur  $F^{-1}$ .

(i) Supposons  $f$  continue et monotone sur  $] \theta^-, \theta^+ [ \ni \theta$  et posons

$$\begin{aligned} f(x) &= f(\theta) - \varphi^-(\theta - x), & x \in ] \theta^-, \theta [ \\ f(x) &= f(\theta) - \varphi^+(x - \theta), & x \in ] \theta, \theta^+ [. \end{aligned}$$

Par définition, les  $x_n^\pm$  s'expriment en fonction de  $q_n$  via les deux équations du  $q_n$ -shorth,

$$\varphi^-(\theta - x_n^-) = \varphi^+(x_n^+ - \theta)$$

et

$$q_n = f(\theta)\Delta_n - \int_{[0, x_n^-]} \varphi^- - \int_{[0, x_n^+]} \varphi^+.$$

(ii) Notre représentation de base demande l'existence de  $\beta^\pm > 1/2$  et  $0 < r_n^\pm = O(\sqrt{n/\log \log n})$  tels que pour  $\varepsilon_n \ll \min(q_n, (r_n^\pm)^{-1/\beta^\pm})$  et  $\sup_{|\varepsilon| < \varepsilon_n} g_n^\pm(\varepsilon) \rightarrow 0$  il vient

$$\begin{aligned} F^{-1}(p_n + \varepsilon) - x_n^- &= \frac{\varepsilon}{f(x_n^\pm)} - r_n^- |\varepsilon|^{1+\beta^-} (1 + g_n^-(\varepsilon)), & |\varepsilon| < \varepsilon_n \\ F^{-1}(p_n + q_n + \varepsilon) - x_n^+ &= \frac{\varepsilon}{f(x_n^\pm)} + r_n^+ |\varepsilon|^{1+\beta^+} (1 + g_n^+(\varepsilon)), & |\varepsilon| < \varepsilon_n. \end{aligned}$$

En injectant (i) dans (ii) il est possible d'exhiber une fonction explicite  $\phi$  telle que  $r_n^\pm \sim \phi(q_n)$ . Si le mode est prononcé alors  $r_n^\pm \rightarrow \infty$  et  $\phi(q_n) = O(\sqrt{n/\log \log n})$  impose à  $q_n$  de dominer une masse critique  $q_n^* \rightarrow 0$ . Si le mode est plat alors  $r_n^\pm \rightarrow 0$  et cette contrainte disparaît,  $q_n^* = 0$ .

Le rôle de  $\beta^\pm$  est essentiel dans la régularisation du problème du fait de 3.3.1, mais une généralisation à d'autres représentations est envisagée. De même, la contrainte sur  $r_n^\pm$  s'affaiblit selon  $\beta^\pm$ .

La plupart des auteurs supposent  $f$  symétrique par rapport à  $\theta$ , ou  $f \in \mathcal{C}_k$  donc quasiment symétrique.

**Exemple.** Dans le cas, symétrique ou non, où  $f \notin \mathcal{C}_1(\cdot] \theta^-, \theta^+ \cdot]$  mais  $f \in \mathcal{C}_1(\cdot] \theta^-, \theta \cdot] \cup \mathcal{C}_1(\cdot] \theta, \theta^+ \cdot]$  et  $|f^{(1)}(x_n^\pm)| = O(\sqrt{n/\log \log n})$  nos hypothèses sont vérifiées avec

$$\beta^- = \beta^+ = 1, \quad r_n^\pm = \frac{f^{(1)}(x_n^\pm)}{2f^3(\theta)}.$$

**Exemple.** Si  $f \in \mathcal{C}_k(\cdot] \theta^-, \theta^+ \cdot]$ ,  $f^{(k)}(\theta) \neq 0 = f^{(m)}(\theta)$  pour  $m < k$  et  $f^{(k)}(x_n^\pm) = O(\sqrt{n/\log \log n})$ ,

$$\beta^- = \beta^+ = k, \quad r_n^\pm = \frac{f^{(k)}(x_n^\pm)}{(k+1)!f^{2+k}(\theta)}.$$

### 4.3.3 – Oscillations exactes

(i) Posons  $\beta = \min(\beta^-, \beta^+)$ ,  $r_n = r_n^- + r_n^+$  si  $\beta^- = \beta^+$ ,  $r_n = r_n^-$  si  $\beta^- < \beta^+$  et  $r_n = r_n^+$  si  $\beta^+ < \beta^-$ .

**Théorème 4.12.** *Sous les hypothèses ci-dessus,*

$$\limsup_{n \rightarrow \infty} \pm \frac{\sqrt{n}(\widehat{\Delta}_n - \Delta_n)}{\sqrt{2q_n \log \log n}} = \frac{1}{f(\theta)} \quad p.s.$$

et, si  $q_n \gg \max(q_n^*, \sqrt{\log \log n}/\sqrt{n}, ((\log \log n)/nr_n^2)^{1/(2\beta+1)})$ ,

$$\limsup_{n \rightarrow \infty} \pm \left( \frac{nr_n^2}{\log \log n} \right)^{1/(2\beta+1)} (\widehat{x}_n^- - x_n^-) = m_\beta = \left( \frac{4(2\beta+1)}{(\beta+1)^2 f^{2\beta+3}(\theta)} \right)^{1/(2\beta+1)} \quad p.s.$$

Le même résultat vaut pour  $\widehat{x}_n^+ - x_n^+$ .

**Principe de démonstration.** Ce théorème et ses variantes utilisent la technique résumée en 4.2.3. Il s'agit de contrôler soigneusement les représentations en utilisant notamment les lois intermédiaires de 1.3 pour débusquer le processus qui oscille le plus et lui appliquer 3.3, ce qui détermine vitesses et constantes. En concluant différemment la preuve nous obtenons aussi la convergence en loi de  $\sqrt{n}(\widehat{\Delta}_n - \Delta_n)$  vers une limite gaussienne et celle de  $(nr_n^2)^{1/(2\beta+1)}(\widehat{x}_n^- - x_n^-)$  vers une limite de type  $\arg \min_{t \in \mathbb{R}} (W(t) + c|t|^{1+\beta})$ . Si besoin, nos arguments permettent de borner précisément les probabilités de dévier par des sommes de termes exponentiels.  $\square$

Ces vitesses exactes quantifient le comportement oscillatoire incompressible de la suite  $[\widehat{x}_n^-, \widehat{x}_n^+]$ . Si  $r_n$  est borné, donc pour un mode plat, les statistiques oscillant selon  $\widehat{x}_n^-$  ont plus d'amplitude que  $1/\sqrt{n}$ . Dans le cas contraire, le Théorème 4.12 autorise jusqu'à des vitesses  $1/n^{1-\varepsilon}$ .

**Interprétation.** C'est un phénomène typiquement non paramétrique. En effet, si  $f$  est entièrement déterminée par  $(\theta, \theta^\pm, \beta^\pm, f(\theta))$  ou un  $\theta^* \in \mathbb{R}^d$  l'ordre de grandeur paramétrique  $1/\sqrt{n}$  est atteint car chaque  $X_i$  apporte la même quantité d'information sur  $f$  donc son mode. A l'inverse, au Théorème 4.12  $f$  n'est spécifiée que localement et déterminée ni par  $(\theta, \theta^\pm, \beta^\pm, f(\theta), r_n)$  ni par  $\theta^* \in \mathbb{R}^d$ , d'où la dimension intrinsèquement infinie du problème posé. Ici, de nombreux  $X_i$  ne renseignent que peu sur  $\theta$ ,  $\Delta_n$  ou  $x_n^\pm$  puisqu'ils ne participent à la définition des estimateurs qu'en permettant de localiser les  $X_i$  proches de  $\theta$ , plus riches en information que dans le cas paramétrique mais en proportion  $q_n \rightarrow 0$ . Cette proportion est plus ou moins grande et permet de localiser le  $q_n$ -shorth plus ou moins vite selon que le mode est prononcé ou non, d'où la variété des vitesses, meilleures ou pires que  $1/\sqrt{n}$ .

**Exemple.** Si  $f \in \mathcal{C}_1(\cdot] \theta^-, \theta \cdot] \cup \mathcal{C}_1(\cdot] \theta, \theta^+ \cdot]$  n'est pas dérivable en  $\theta$  et  $\sqrt{(\log \log n)/nq_n^3} \ll |f^{(1)}(x_n^\pm)| = O(\sqrt{n/\log \log n})$  alors

$$\limsup_{n \rightarrow \infty} \pm \left( \frac{n(f^{(1)}(x_n^\pm))^2}{\log \log n} \right)^{1/3} (\widehat{x}_n^- - x_n^-) = (3f(\theta))^{1/3} \quad p.s.$$

**Exemple.** Si  $f \in \mathcal{C}_{k+1}(\theta^-, \theta^+]$ ,  $f^{(m)}(\theta) = 0 \neq f^{(k)}(\theta)$ ,  $m < k$ ,  $|f^{(k)}(x_n^\pm)| = O(\sqrt{n/\log \log n})$  et  $q_n \gg ((\log \log n)/n)^{1/(2k+1)}$  alors

$$\limsup_{n \rightarrow \infty} \pm \left( \frac{n}{\log \log n} \right)^{1/(2k+1)} (\widehat{x}_n^- - x_n^-) = \left( \frac{(2k+1)((k+1)!)^2 f(\theta)}{(k+1)^2 (f^{(k)}(\theta))^2} \right)^{1/(2k+1)} \quad p.s.$$

(ii) Illustrons l'interprétation précédente à l'aide d'un cas facile à visualiser. Pour  $\alpha > 0$  et  $a > 0$ ,

$$f(x) = f(\theta) - a|x - \theta|^\alpha (1 + o(1)), \quad x \in ]\theta^-, \theta^+ [.$$

Les équations du  $q_n$ -shorth impliquent  $|x_n^\pm - \theta| \sim q_n/2f(\theta)$  et justifient la représentation 4.3.2 avec

$$\beta = 1, \quad r_n = \frac{a\alpha q_n^{\alpha-1}}{2^{\alpha-1} f^{2+\alpha}(\theta)}.$$

Distinguons alors deux situations.

**Corollaire 4.13 (mode non prononcé).** Si  $\alpha > 1/2$  et  $q_n \gg ((\log \log n)/n)^{1/(2\alpha+1)}$ ,

$$\limsup_{n \rightarrow \infty} \pm (q_n)^{2(\alpha-1)/3} \left( \frac{n}{\log \log n} \right)^{1/3} (\widehat{x}_n^- - x_n^-) = \left( \frac{3f^{2\alpha-1}(\theta)}{4^{1-\alpha} a^2 \alpha^2} \right)^{1/3} \quad p.s.$$

Lorsque  $\alpha \in ]1/2, 1[$  les petits  $q_n$  atteignent des vitesses  $1/n^{1/(2\alpha+1)}$  avec  $1/(2\alpha+1) \in ]1/3, 1/2[$ . Si  $\alpha = 1$  alors tous les  $q_n$  autorisés approchent la vitesse  $((\log \log n)/n)^{1/3}$ . Si  $\alpha > 1$ ,  $q_n$  doit être au contraire aussi grand que possible et la vitesse ralentie d'autant. Le cas lent limite, c'est-à-dire  $q_n = q$ , est consistant avec le Théorème 4.9.

**Corollaire 4.14 (mode prononcé).** Si  $\alpha \leq 1/2$  et  $q_n \gg \sqrt{(\log \log n)/n}$ ,

$$\limsup_{n \rightarrow \infty} \pm \left( \frac{1}{q_n} \right)^{2(1-\alpha)/3} \left( \frac{n}{\log \log n} \right)^{1/3} (\widehat{x}_n^- - x_n^-) = \left( \frac{3f^{2\alpha-1}(\theta)}{4^{1-\alpha} a^2 \alpha^2} \right)^{1/3} \quad p.s.$$

Pour  $q_n$  proche de sa borne inférieure la vitesse est en  $1/n^{(2-\alpha)/3}$  avec  $(2-\alpha)/3 \in [1/2, 2/3[$ .

Notre représentation n'est pas adaptée à des modes trop prononcés. Dans le cas d'un pic exponentiel, qui sort de nos conditions, [Sa] a montré que des très petits  $q_n$  atteignent des vitesses  $1/n^{1-\varepsilon}$ . D'où le projet de recherche 4.4.1(ii).

#### 4.3.4 – Convergence d'estimateurs du mode

Le Théorème 4.12 concentre en un unique énoncé de quoi contrôler efficacement divers estimateurs du mode  $\theta$  basés sur  $(\widehat{x}_n^-, \widehat{\Delta}_n)$ . Par la méthode suivante, des vitesses connues sont prouvées optimales ou améliorées, et des estimateurs inédits fournis avec vitesse exacte. Nous pouvons en particulier trancher parmi des  $\widehat{\theta}_n \in [\widehat{x}_n^-, \widehat{x}_n^+]$  concurrents ou montrer très simplement qu'ils sont équivalents.

**Méthode.** Les équations de 4.3.2 et les hypothèses courantes sur  $f$  permettent d'évaluer ou identifier les fonctions  $\theta = \varphi_n(x_n^-, \Delta_n)$ . Il convient alors d'étudier le biais  $\varepsilon_n = \mathbb{E}\varphi_n(\widehat{x}_n^-, \widehat{\Delta}_n) - \theta$  en fonction de  $q_n$ , ses corrections éventuelles et les contraintes induites sur  $q_n$ . La comparaison de  $\varepsilon_n$  à la vitesse du Théorème 4.12 sous les différentes contraintes permet d'optimiser  $q_n$  dans  $\widehat{\theta}_n = \varphi_n(\widehat{x}_n^-, \widehat{\Delta}_n)$ . Si les paramètres du  $q_n$  optimal sont inconnus il faut soit en contrôler l'estimation et tenir compte des nouvelles contraintes sur  $q_n$  soit opter pour un  $q_n$  garantissant une vitesse non optimale mais explicite. Le  $q_n$  optimal ainsi obtenu minimise la suite déterministe  $v_n$  bornant presque sûrement  $|\widehat{\theta}_n - \theta|$  donc en fait, il minimise la probabilité de dévier pour  $n$  grand. Appliquons cette méthode.

(i) Supposons qu'il existe  $\delta \in ]0, 1[$  tel que  $\theta = x_n^- + \delta\Delta_n + \varepsilon_n$ . Sous les hypothèses du Théorème 4.12 pour  $q_n$  les estimateurs  $X_{(n\widehat{p}_n + \delta q_n)}$  et  $\widehat{x}_n^- + \delta\widehat{\Delta}_n = (1-\delta)\widehat{x}_n^- + \delta\widehat{x}_n^+$  oscillent comme  $\widehat{x}_n^- = X_{(n\widehat{p}_n)}$ . Si nous disposons d'estimateurs  $\widehat{\delta}_n$  convergeant assez vite vers  $\delta$  et si  $\varepsilon_n$  est suffisamment lent,

$$\widehat{x}_n^- + \widehat{\delta}_n \widehat{\Delta}_n - \theta = O_{p.s.} \left( \left( \frac{\log \log n}{n r_n^2} \right)^{1/(2\beta+1)} \right)$$

Mais en général,  $\widehat{\delta}_n$  reste à construire et le biais  $\varepsilon_n$  est fonction de  $q_n$  selon que  $f$  est plus ou moins néfaste à la localisation dans le  $q_n$ -shorth. En précisant le modèle, des solutions tombent.

(ii) Identifions le  $q_n$  optimal dans l'exemple historique de [Ve] où, comme suggéré dans [Ch],

$$f(x) = f(\theta) - a(x - \theta)^2 + b(x - \theta)^k (1 + o(1)), \quad x \in ]\theta^-, \theta^+[$$

avec  $k > 2$  un entier impair,  $a > 0$  et  $b \neq 0$ . La symétrie autour de  $\theta$  suggère l'estimateur naïf  $\widehat{\theta}_{1,n} = \widehat{x}_n^- + \widehat{\Delta}_n/2$  ou bien la médiane empirique de  $[\widehat{x}_n^-, \widehat{x}_n^+]$ , qui oscillent ici tous deux comme  $\widehat{x}_n^-$  avec un biais  $\varepsilon_n$  comparable. Les équations du shorth donnent

$$\Delta_n \sim \frac{q_n}{f(\theta)}, \quad \delta = \frac{1}{2}, \quad \varepsilon_n = \theta - x_n^- - \frac{\Delta_n}{2} = \frac{b\Delta_n q_n^{k-2}}{4a(2f(\theta))^{k-2}}$$

donc  $\widehat{\theta}_{1,n} - \theta = \widehat{x}_n^- - x_n^- + (\widehat{\Delta}_n - \Delta_n)/2 - q_n^{k-1}(b + o(1))/4a(2f(\theta))^{k-1}$  alors que le Théorème 4.12 implique  $\limsup_{n \rightarrow \infty} \pm(nq_n^2/\log \log n)^{1/3}(\widehat{x}_n^- - x_n^-) = (3/a^2)^{1/3}f(\theta)$  p.s. L'équilibre entre ces deux déviations contradictoires, déterministe et aléatoire, dévoile la vitesse optimale.

**Corollaire 4.15.** *Si  $q_n = \xi((\log \log n)/n)^{1/(3k-1)}$  nous avons, avec probabilité un,*

$$\limsup_{n \rightarrow \infty} \left( \frac{n}{\log \log n} \right)^{\frac{k-1}{3k-1}} \left| \widehat{\theta}_{1,n} - \theta \right| = \left( \frac{3}{\xi^2 a^2} \right)^{1/3} f(\theta) + \frac{b\xi^{k-1}}{4a(2f(\theta))^{k-1}}.$$

Pour le statisticien, tout dépend de  $k$ , qui détermine l'ordre de grandeur optimal de  $q_n$ . Si  $(k, f(\theta), a, b)$  sont connus, l'estimateur débiaisé  $\widehat{\theta}_{2,n} = \widehat{\theta}_{1,n} + bq_n^{k-1}/4a(2f(\theta))^{k-1}$  permet d'approcher  $\theta$  à la vitesse  $((\log \log n)/n)^{1/3}$  pour  $q_n$  lent. Dans  $\widehat{\theta}_{2,n}$  un estimateur usuel de  $f(\theta)$  peut s'y substituer sans perte. Curieusement, dès que  $a$  ou  $b$  est inconnu la meilleure vitesse accessible est  $((\log \log n)/n)^{(k-1)/(3k-1)}$ . En l'ignorance de  $k \geq 3$  le fait que  $\varepsilon_n = O(q_n^2)$  montre que le choix minimax  $q_n = ((\log \log n)/n)^{1/8}$  assure une vitesse  $O_{p.s.}((\log \log n)/n)^{1/4}$ . Tout ceci recouvre et précise [Ch, Ve, Lc].

Nous résolvons avec la même facilité le cas plus général où  $f(\theta) - a|x - \theta|^\alpha + b|x - \theta|^\gamma$ ,  $\alpha < \gamma$  sont non entiers et  $(a, b, \alpha, \gamma)$  diffèrent selon que  $x < \theta$  ou  $x > \theta$ .

(iii) Recherchons le  $q_n$  optimal dans un cas non symétrique. Pour  $a > 0$ ,  $b > 0$  et  $1/2 < \gamma < \alpha$

$$\begin{aligned} f(x) &= f(\theta) - a(\theta - x)^\alpha (1 + o(1)), & x \in ]\theta^-, \theta[ \\ f(x) &= f(\theta) - b(x - \theta)^\gamma (1 + o(1)), & x \in ]\theta, \theta^+[ \end{aligned}$$

de sorte que  $(\theta, f(\theta))$  est un point bien visible de la densité. Estimons  $\theta$  par  $\widehat{\theta}_{3,n} = \widehat{x}_n^+$ . Puisque

$$\delta = 1, \quad \varepsilon_n = - \left( \frac{a}{b} \right)^{1/\gamma} \left( \frac{q_n}{f(\theta)} \right)^{\alpha/\gamma}$$

le Théorème 4.12 donne  $\limsup_{n \rightarrow \infty} \pm q_n^{2(\gamma-1)/3} (n/\log \log n)^{1/3} (\widehat{x}_n^+ - x_n^+) = (3/b^2)^{1/3} f(\theta)$  p.s.

**Corollaire 4.16.** *Si  $\rho = 3\alpha/\gamma + 2(\gamma - 1)$  et  $q_n = \xi((\log \log n)/n)^{1/\rho}$  il vient, avec probabilité un,*

$$\limsup_{n \rightarrow \infty} \left( \frac{n}{\log \log n} \right)^{\frac{\alpha}{3\alpha + 2\gamma(\gamma-1)}} \left| \widehat{\theta}_{3,n} - \theta \right| \in ]0, \infty[.$$

Si  $(a, b, \alpha, \gamma)$  sont connus, la vitesse  $\sqrt{(\log \log n)/n}$  est approchable pour  $q_n$  du même ordre et  $\widehat{\theta}_{3,n} + \varepsilon_n$ . Si un paramètre au moins est inconnu l'ordre  $\alpha/(3\alpha + 2\gamma(\gamma - 1))$  est indépassable et non garanti. Cet ordre est quasi paramétrique si  $\alpha \rightarrow 1/2$  et au mieux  $1/3$  si  $\alpha \rightarrow \infty$ . Lorsque seul  $\alpha$  est connu, comme  $\varepsilon_n = O(q_n)$  la meilleure vitesse garantie est  $O_{p.s.}((1/n)^{1/(2\alpha+1)})$  et s'obtient pour  $q_n = (1/n)^{1/\rho_+}$  où  $\rho_+ = 2(\alpha - 1) + 3$ . Si  $\alpha$  est inconnu, l'estimer en s'inspirant de 4.3.4(v).

(iv) Puisque  $\widehat{\theta}_{1,n}$  ou  $\widehat{\theta}_{3,n}$  peuvent servir à étudier des estimateurs de  $\theta$  basés sur un estimateur à noyau de  $f$ , le Théorème 4.12 donne accès à la loi du logarithme itéré dans [We, Ed, Ro]. Par exemple

l'estimateur de [Ch] est le milieu  $\tilde{\theta}_{1,n}$  de l'intervalle de longueur  $2h_n \rightarrow 0$  contenant le plus de  $X_i$ . Si  $f \in \mathcal{C}_2(\] \theta^-, \theta^+ [)$  alors  $\tilde{\theta}_{1,n}$  avec  $2h_n$  est contrôlé par  $\hat{\theta}_{1,n}$  avec  $q_n = 2h_n f(\theta)$  et nous obtenons les vitesses évoquées ci-dessus en  $\alpha = 1$ ,  $\gamma = 2$ . Le Corollaire 4.14 précise la convergence forte de [We], sa version faible restitue la convergence en loi en  $n^{1/3} h_n^{2/3}$  de [Ch], et cela même si  $f \notin \mathcal{C}_2(\] \theta^-, \theta^+ [)$ .

(v) Notre approche conduit naturellement à de nouveaux estimateurs, sous les hypothèses sur  $F$  appropriées. En particulier nous pouvons évaluer la perte de vitesse induite par l'estimation de paramètres inconnus dans la représentation locale de  $F$ . Prenons l'exemple du crucial  $\delta = 1/(1 + \kappa)$  ci-dessus et supposons que pour des réels  $\alpha > 1$  et  $a \neq b$ ,

$$\begin{aligned} f(x) &= f(\theta) - a(\theta - x)^\alpha (1 + o(1)), & x \in \] \theta^-, \theta [ \\ f(x) &= f(\theta) - b(x - \theta)^\alpha (1 + o(1)), & x \in \] \theta, \theta^+ [. \end{aligned}$$

Soit  $m = m(n)$  une suite telle que  $m \uparrow \infty$  et  $m/n \rightarrow 0$ . Les équations du  $q_n$ -shorth impliquent

$$\kappa = \left(\frac{a}{b}\right)^{1/\alpha} = \frac{x_n^+ - \theta}{\theta - x_n^-} = \frac{x_m^+ - x_n^+}{x_n^- - x_m^-}, \quad \theta = x_n^- + \frac{\Delta_n}{1 + \kappa}$$

et le Théorème 4.12 donne accès au contrôle asymptotique exact des estimateurs

$$\hat{\kappa}_n = \frac{\hat{x}_m^+ - \hat{x}_n^+}{\hat{x}_n^- - \hat{x}_m^-}, \quad \hat{\theta}_n = \hat{x}_n^- + \frac{\hat{\Delta}_n}{1 + \hat{\kappa}_n}.$$

En posant  $m(n) = n/k_n$  la vitesse de  $\hat{\theta}_n - \theta$  est en  $((k_n \log \log n)/(n q_n^{2\alpha-2} (q_{n/k_n}/q_n)^{2\alpha+1}))^{1/3}$ . Au lieu de comparer séquentiellement le passé et le présent nous pouvons utiliser  $q_{1,n} \ll q_{2,n}$ , considérer

$$\hat{\kappa}_n = \frac{x_n^+(q_{2,n}) - x_n^+(q_{1,n})}{x_n^-(q_{1,n}) - x_n^-(q_{2,n})}$$

et choisir  $q_{2,n}$  pour le contrôle optimal de  $\hat{\theta}_n = \hat{x}_{1,n}^- + \hat{\Delta}_{1,n}/(1 + \hat{\kappa}_n)$  via la comparaison au biais. Enfin, pour estimer  $\alpha$  il est envisageable de contrôler, via une estimation de la densité et le Théorème 4.12,

$$\alpha_n = \frac{\log(\hat{f}(\hat{\theta}_n) - \hat{f}(\hat{x}_n^-))}{\log(\hat{\theta}_n - \hat{x}_n^-)}, \quad \hat{\theta}_n = \hat{x}_n^- + \frac{\hat{\Delta}_n}{1 + \hat{\kappa}_n}.$$

Des procédures adaptatives découlent de ces considérations.

## 4.4 – Autres axes de recherche

### 4.4.1 – Détection de singularités

(i) Au Théorème 4.9 les cas  $\beta = 1/2$  et  $\beta < 1/2$  sont à la portée de la méthode 4.2.3. De nouvelles vitesses émergent car des termes aléatoires liés aux espacements et jusqu'ici secondaires dirigent à présent les déviations, induisent une renormalisation et changent la nature des processus gaussiens limites. Une fois maîtrisé, le cas  $\beta \leq 1/2$  devra être adapté au Théorème 4.12 et combiné à (ii).

(ii) A la suite de [15] et au vu du Corollaire 4.14 nous explorons plus spécifiquement les modes très prononcés, qui requièrent une nouvelle représentation de  $F^{-1}$  fusionnant les rôles de  $\beta$  et  $r_n$ . Ce faisant, dans la preuve du Théorème 4.12, à l'échelle  $q_n < q_n^*$  un nouveau comportement prend le pas et les incréments locaux des quantiles empiriques jouent un rôle différent.

(iii) Les deux cas extrêmes ci-dessus correspondent à la détection de pics de  $f$ . Nous envisageons d'adapter nos arguments au contrôle d'estimateurs non paramétriques d'autres types de singularités de  $f$ . Par exemple une discontinuité, une inflexion, un tronçon de forme donnée ou un minimum local. Parfois en se basant sur le  $q_n$ -shorth, parfois en combinant deux  $q_n$ -shorth dans deux domaines séparés et en recherchant le meilleur séparateur, parfois en utilisant un triplet  $(x_n, y_n, z_n)$  maximisant le ratio  $(y_n - x_n)/(z_n - y_n)$  sous la contrainte que  $[x_n, y_n]$  et  $[y_n, z_n]$  sont de masse  $q_n$ .

(iv) Enfin, la procédure consistant à itérer le  $q_n$ -shorth plusieurs fois dans lui-même peut s'adapter dans chacun des cas ci-avant. C'est un autre défi à la compréhension des fluctuations d'échantillonnage.



#### 4.4.2 – Sélection de modèles

(i) Les théories de l'apprentissage statistique, voir [BBL], de la reconnaissance de formes, voir [DGL], de la classification et de la sélection de modèles, voir [MNé], sont très liées à celle du processus empirique indexé par  $\mathbb{R}^d$ , des ensembles ou des fonctions. Prenons l'exemple de la classification. La problématique centrale concerne la recherche de bornes de déviation précises sur l'écart entre le meilleur choix possible et le modèle sélectionné  $f \in \mathcal{F}$  par minimisation de  $\mathbb{P}_n f$  modulo des fonctions de coût et de pénalité. Or les outils usuels sont basés sur des comportement moyens, comme le contrôle de  $\mathbb{E} \|\alpha_n\|_{\mathcal{F}}$ , plutôt que presque sûrs. Ainsi  $\alpha_n$  joue un rôle crucial mais il est rarement utilisé en tant que processus, et jamais approché par un processus gaussien, ce qui serait utile pour faire des tests.

C'est pourquoi nous commençons une recherche visant à exporter nos savoirs du type 3.4 et 2.3.4 dans ce cadre. L'objectif est d'affiner les bornes, d'obtenir des limites gaussiennes via les Théorèmes 3.18 à 3.23 et d'étudier le comportement presque sûr de certaines stratégies de sélection, comme le bootstrap, la validation croisée ou le risque empirique. Par exemple, [MNé] s'étend aux classes de 3.18 et 3.21.

(ii) Plus précisément, considérons une v.a.  $(X, Y)$  à valeurs dans  $(\mathcal{X}, \mathcal{A}) \times [0, 1]$  et la mesure empirique  $\mathbb{P}_n$  associée à un échantillon  $(X_i, Y_i)$  i.i.d. de loi  $P = \mathbb{P}^{(X, Y)}$ . Soit  $\mathcal{F}$  une classe de modèles  $y = f(x)$  où  $f \in \mathcal{F}$  est mesurable de  $(\mathcal{X}, \mathcal{A})$  dans  $[0, 1]$ . Le meilleur modèle  $f_*$  pour une perte  $g$  mesurable de  $\mathcal{F} \times \mathcal{X} \times [0, 1]$  dans  $\mathbb{R}_+$  minimise le risque  $r(f) = \mathbb{E}(g(f, (X, Y)))$  qui est fonction de  $\mathbb{P}^{(X, Y)}$  donc inconnu. Le cas  $g_0((f, (X, Y))) = (y - f(x))^2$  correspond au cadre usuel de la régression non-paramétrique, où  $f_*(x) = \mathbb{E}(Y | X = x)$ . Le cas  $Y \in \{0, 1\}$  et  $g_1((f, (X, Y))) = \mathbb{I}_{Y \neq f(x)}$  correspond au cadre usuel de la classification binaire, où  $f_*(x) = \mathbb{P}(Y = 1 | X = x)$ .

Pour prévoir la réponse  $Y$  en fonction de la variable explicative  $X$  il est naturel de sélectionner  $f_n \in \mathcal{F}$  minimisant le risque empirique  $r_n(f) = \mathbb{P}_n(g(f, (X, Y)))$  éventuellement pénalisé par une quantité  $p(f) \geq 0$  liée à la complexité de  $f$ , par exemple sa dimension. Le prédicteur de  $Y_{n+1}$  est alors  $f_n(X_{n+1})$ . Or  $(r_n - r)(f) = \alpha_n(f)/\sqrt{n}$  est le  $P$ -processus empirique indexé par  $\mathcal{G} = \{g(f, \cdot) : f \in \mathcal{F}\}$ . Pour  $g_0$ ,  $\mathcal{G}$  est une classe de fonctions et pour  $g_1$ ,  $\mathcal{G}$  est une classe d'ensembles. Si  $\mathcal{G}$  vérifie la condition (VC) du 3.4.2 il s'ensuit par la Proposition 3.16 que  $f_n$  minimise en  $f$  le processus

$$r_n(f) = \frac{\mathbb{G}_n(g(f, \cdot))}{\sqrt{n}} + r(f) + \varepsilon_n(f), \quad f \in \mathcal{F}, \quad \|\varepsilon_n\|_{\mathcal{F}} = O_{p.s.} \left( \frac{(\log n)^{\tau_0}}{n^{1/2+1/(2+5\nu_0)}} \right)$$

donc  $g_n = g(f_n, \cdot)$  converge en loi vers l'arg min d'un  $P$ -pont brownien indexé par  $\mathcal{G}$  avec dérive le risque  $r$ . Dans cette version généralisée de 4.2 et 4.3 nous étudions les oscillations de  $g_n - g_*$  avec  $g_* = g(f_*, \cdot)$  en attaquant la question des incréments locaux de  $\mathbb{G}_n$ . Pour un risque  $r$  régulier, ceci réalise le contrôle presque sûr de  $r(f_n) - r(f_*)$  et donne accès à des bornes de déviations gaussiennes.

Il est intéressant de constater que selon la taille de  $\mathcal{G}$  l'erreur  $\|\varepsilon_n\|_{\mathcal{F}}$  issue de l'approximation forte varie, d'après 3.4, entre  $(\log n)/n$  et  $1/\sqrt{n}$  et interfère ou non avec le contrôle classique de  $\mathbb{E}(r(f_n) - r(f_*))$ .

(iii) En estimation non ou semi paramétrique de la densité, une méthode émergente consiste à choisir la densité dans un grand ensemble  $\mathcal{F}$  de candidates par minimisation d'un critère basé sur  $\mathbb{P}_n$ , comme par exemple  $f_n = \arg \inf_{\mathcal{F}} \sup_{C \in \mathcal{C}} |\mathbb{P}_n(C) - \int_C f|$ , et d'appliquer les bornes du type (i), voir [DvL]. Nous commençons à étudier ce type d'approche sous l'angle plus intrinsèque du processus empirique, la bonne connaissance des fluctuations stochastiques pouvant s'intégrer aux critères et suggérer des méthodes bien contrôlées. Ceci pourra révéler des besoins exotiques en bornes de déviation, et les recherches de type 3.5.1 seront probablement utiles dans ce contexte.

#### 4.4.3 – Régression et mode conditionnel

(i) En statistique inférentielle, les modèles avec variables explicatives se placent naturellement dans un cadre conditionnel. Supposons observés des vecteurs aléatoires  $(X_1, Y_1), \dots, (X_n, Y_n)$  i.i.d. de même loi que  $(X, Y) \in \mathbb{R}^p \times \mathbb{R}$ . Un grand nombre de travaux, théoriques ou méthodologiques, s'intéressent à l'estimation de la fonction de régression  $\varphi(x) = \mathbb{E}(Y | X = x)$ , la densité conditionnelle  $f_x = f_{Y|X=x}$  et de son mode  $\theta_x = \arg \max f_x$  par des méthodes non paramétriques. Proches de notre optique, des

articles tels que [St3,EM4] étudient la convergence presque sûre d'estimateurs de  $f_x$ . Nos résultats de 4.1 sur les delta-suites peuvent se prolonger dans ce sens.

(ii) Concernant  $\theta_x$  nous envisageons d'étendre le Théorème 4.12 au  $q_n$ -shorth conditionnel naïvement défini comme suit. Soit  $\mathcal{P}_n = \{A_k : k \leq k_n\}$  une suite de partitions du support de  $\mathbb{P}^X$ . Pour chaque  $k \leq k_n$  notons  $Y_{(1)}^{(k)} \leq \dots \leq Y_{(n_k)}^{(k)}$  la statistique d'ordre des  $n_k = \sum_{i=1}^n \mathbb{I}_{X_i \in A_k}$  valeurs  $Y_j^{(k)}$  des  $Y_i$  tels que  $X_i \in A_k$ . En considérant pour  $k \leq k_n$  les espacements conditionnels à  $A_k$

$$\widehat{\Delta}_{n,k}(j) = Y_{(j+[n_k q_{n_k}])}^{(k)} - Y_{(j)}^{(k)}, \quad j = 1, \dots, N_{n,k} - [N_{n,k} q_{N_{n,k}}]$$

posons

$$\widehat{j}_{n,k} = \arg \min_{1 \leq j \leq N_{n,k} - [N_{n,k} q_{N_{n,k}}]} \left\{ \widehat{\Delta}_{n,k}(j) \right\}.$$

La longueur et la position des  $k_n$  shorth conditionnels sont ainsi données par

$$\widehat{\Delta}_{n,k} = \widehat{\Delta}_{n,k}(\widehat{j}_{n,k}), \quad \widehat{y}_{n,k} = Y_{(\widehat{j}_{n,k})}^{(k)}, \quad k \leq k_n.$$

A l'aide de la méthode 4.2-4.3 nous cherchons à dégager un théorème limite contrôlant uniformément en  $k$  les déviations de ces  $q_n$ -shorths. Notre but est de construire des estimateurs du mode  $\theta_x$  de type  $\widehat{\theta}_n(x) = \varphi_n(\widehat{\Delta}_{n,k}, \widehat{y}_{n,k})$  pour  $x \in A_k$  tels que  $\sup_x |\widehat{\theta}_n(x) - \theta_x| = O_{p.s.}(v_n)$  avec une vitesse déterministe  $v_n$  très précise, dans l'esprit des résultats de 4.3.3. Une difficulté supplémentaire réside dans le fait que les  $N_{n,k}$  sont aléatoires, la partition  $\mathcal{P}_n$  bouge et dans l'exigence de contrôle uniforme des  $A_k$ , dont le choix s'avère crucial et requiert de nouvelles hypothèses sur la variation en  $x$  des fonctions de quantile conditionnelles  $F_x^{-1}$  où  $F_x(y) = \mathbb{P}(Y \leq y | X = x)$ .

#### 4.4.4 – Perspectives

(i) Dans la partie 4 nous avons eu accès à des résultats via l'approche 1.3.2. Ils peuvent encore être renforcés à l'aide de nos raffinements, notamment ceux de [20]. Par exemple, dans le Théorème 4.9 nous pouvons montrer, à l'aide du Théorème 2.17, que pour tout  $\varepsilon > 0$  et un  $\rho_\varepsilon > 0$  explicite,

$$\limsup_{n \rightarrow \infty} \frac{1}{\log \log n} \log \log \left( \frac{1}{n} \inf \left\{ k > n : \widehat{d}_k > d + \frac{\sqrt{2(1-\varepsilon)q(1-q)}}{m} \sqrt{\frac{\log \log k}{k}} \right\} \right) = \rho_\varepsilon \quad p.s.$$

Ceci contrôle le temps d'attente de la plus mauvaise estimation de  $d$ . Pour  $\widehat{\Delta}_m$  de 4.3 cela semble également possible. Mieux, nous pensons pouvoir décrire la gamme des temps d'attentes du type

$$N_n^+ = \inf \left\{ k > n : \widehat{d}_k > d + \pi_k \right\} \quad \text{ou} \quad N_n^- = \inf \left\{ k > n : \left| \widehat{d}_k - d \right| > \pi_k \right\}$$

avec  $\pi_k$  plus ou moins rapide donc  $N_n^\pm$  plus ou moins petit. Nous visons des résultats analogues dans le cadre de 4.1 avec des grandes fenêtres  $h_n$  ou trop peu de points d'estimation, comme en [21].

L'évaluation presque sûre des temps d'attente des pires déviations est donc une direction nouvelle qui mérite d'être explorée plus avant. C'est un critère de comparaison des estimateurs non paramétriques typiquement dans notre optique de contrôle déterministe du stochastique.

(ii) En 4.1 nous envisageons de considérer des classes de delta-suites de fonctions  $\delta_{g,n}$  indexées par des fonctions  $g \in \mathcal{G}$  qui modèlent leur forme et d'utiliser  $\alpha_n$  indexé par  $\mathcal{G}$ , dans l'esprit de [EM4] pour les noyaux et la régression et de [M3] pour les incréments. Il conviendrait d'inclure simultanément la reformulation récente uniforme en  $a_n \in [a_n^-, a_n^+]$  des lois fonctionnelles de la partie 2. Ainsi nous obtiendrions, par exemple pour deux suites  $\lambda_n^\pm$  de référence vérifiant (1) avec  $c = d = \infty$ ,  $\lambda_n^- < \lambda_n^+$ , et une collection  $\{\delta_{g,n} : g \in \mathcal{G}, n \geq 1\}$  de delta-suites de fonctions telles que pour tout  $g \in \mathcal{G}$  la suite  $\delta_{g,n}$  est  $B$  ou  $C$ -admissible avec  $\lambda_n \in [\lambda_n^-, \lambda_n^+]$ , un contrôle uniforme de ces estimateurs,

$$\lim_{n \rightarrow \infty} \sup_{g \in \mathcal{G}} \frac{\sqrt{n}}{\sqrt{2\sigma_{g,n}^2 L_{g,n}}} \sup_{[a,b]} \pm \frac{f_{g,n} - \mathbb{E}f_{g,n}}{\sqrt{f}} = c_g \quad p.s.$$

(iii) La littérature en estimation non paramétrique de la densité et de la régression propose une multitude de procédures : adaptativité, bootstrap, multifenêtrage, plug-in itéré, validation croisée, apprentissage etc. Cependant, relativement peu de théorèmes limites sont disponibles. En principe ces procédures peuvent s'analyser précisément en se ramenant aux processus sous-jacents mais tout ce travail reste à faire. En particulier, il semble possible de prolonger 4.1 pour donner des vitesses à [MNo]. De même, la compréhension de l'impact des fluctuations de l'échantillon initial sur les méthodes de ré-échantillonnage devrait permettre de développer de nouveaux théorèmes limites, curieusement assez rares dans le domaine du bootstrap, pourtant riche en applications.

#### 4.5 – Références citées en partie 4

- [A+] Andrews D. F., Bickel P. J., Hampel F. R., Hampel P. J., Rogers W. H. and Tukey J. W. (1972). Robust estimates of location. *Princeton Univ. Press, New Jersey*.
- [BR] Bickel P. J. and Rosenblatt M. (1973). On some global measures of the deviations of density function estimates. *Annals of Statistics* **1**, 1071-1095.
- [BW] Blum J. and Walter G. G. (1979). Probability density estimation using Delta Sequences. *Annals of Statistics* **7**, 328-340.
- [BLc] Bosq D. and Lecoutre J. P. (1987). Théorie de l'estimation fonctionnelle. *Economie et Statistiques Avancées, Economica*.
- [BBL] Bousquet O., Boucheron S. and Lugosi G. (2005). Théory of classification : a survey of recent advances. *ESAIM: Probability and Statistics*, to appear.
- [CL] Castellana J. V. et Leadbetter M. R. (1986). On smoothed probability density estimation for stationary processes. *Stochastic Processes and their Applications* **21**, 179-193.
- [Ch] Chernoff H. (1964). Estimation of the mode. *Annals of the Institute of Statistical Mathematics* **16**, 31-41.
- [Dv2] Devroye L. (1987). A course in density estimation. *Birkhäuser*.
- [DGL] Devroye L., Györfi L. and Lugosi G. (1996). A Probabilistic Theory of Pattern Recognition. *Springer-Verlag*.
- [DvL] Devroye L. and Lugosi G. (2001). Combinatorial Methods in Density Estimation. *Springer*.
- [Ed] Eddy W. F. (1982). The asymptotic distributions of kernel estimators of the mode. *Z. Wahrsch. verw. Gebiete*. **59**, 279-290.
- [EM4] Einmahl U. and Mason D. M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *Journal of Theoretical Probability* **13**, 1-37.
- [Fr] Fryer M. J. (1977). A review of some non-parametric methods of density estimation. *Journal of the Institute of Mathematics and Applications*, **20**, 335-354.
- [GG] Giné E. and Guillou A. (2001). On consistency of kernel density estimators for randomly censored data : rates holding uniformly over adaptative intervals. *Annales de l'Institut Henri Poincaré, Prob. Stat.* **37**, 503-522.
- [GKW] Giné E., Koltchinskii V. and Wellner J. (2005). Ratio limit theorems for empirical processes. *Preprint*.
- [GKZ] Giné E., Koltchinskii V. and Zinn J. (2004). Weighted uniform consistency of kernel density estimators. *Annals of Probability* **38**, 2570-2605.
- [Gb] Grübel R. (1988). The length of the shorth. *Annals of Statistics* **16**, 616-628.
- [H] Hall P. (1981). Laws of the iterated logarithm for nonparametric density estimation. *Z. Wahrsch. verw. Gebiete*. **56**, 247-261.
- [H2] Hall P. (1983). Measuring the efficiency of trigonometric series estimates of a density. *Journal of Multivariate Analysis* **13**, 234-256.

- [H3] Hall P. (1990). On the law of the iterated logarithm for density estimators. *Statistic and Probability Letters* **9**, 237-240.
- [HKPT] Härdle W., Kerkycharian G., Picard D. and Tsybakov A. (1998). Wavelets approximation, and stactical applications. *Lecture notes in Statistics* **129**, Springer-Verlag.
- [JN] Janaszewska E. and Nagaev A. V. (1998). On the joint distribution of the shorth height and length. *Mathematical Methods of Statistics* **7**, 210-229.
- [KP] Kim J. and Pollard D. (1990). Cube root asymptotics. *Annals of Statistics* **18**, 191-219.
- [LW] Leadbetter M. R. and Watson G. S. (1964). Hazard Analysis II. *Sankhya : the Indian Journal of Statistics, ser. A* **26**, 101-116.
- [Lc] Leclerc J. (2000). Strong limiting behavior of two estimates of the mode: the shorth and the naive estimator. *Statistics and Decisions* **19**, 413-427.
- [MNo] Marron J. S. and Nolan D. (1989). Uniform consistency of automatic and location-adaptative delta-sequence estimators. *Probability Theory and Related Fields* **80**, 619-632.
- [MNé] Massart P. and Nédélec E. (2005). Risk bounds for statistical learning. *Annals of Statistics*, to appear.
- [P] Parzen E. (1962). On estimating probability density function and mode. *Annals of Mathematical Statistics* **33**, 1065-1076.
- [Pr] Prakasa Rao B. L. S. (1983). Nonparametric Functional Estimation. *Academic Press*.
- [Ri2] Rio E. (1994). Local invariance principles and their application to density estimation. *Probability Theory and Related Fields* **98**, 21-45.
- [Ro] Romano J. (1988). On weak convergence and optimality of kernel density estimation of the mode. *Annals of Statistics* **16**, 629-647.
- [Rb] Rosenblatt M. (1971). Curve estimates. *Annals of Math. Statistics* **42**, 1815-1842.
- [Sa] Sager T. (1975). Consistency in nonparametric estimation of the mode. *Annals of Statistics* **3**, 698-706.
- [Si] Silverman B. (1978). Weak and strong uniform consistency of the Kernel estimate of a density and its derivatives. *Annals of Statistics* **6**, 177-184.
- [SW] Shorack G. R. and Wellner J. A. (1986). Empirical Processes with Applications to Statistics. *Wiley*.
- [St2] Stute W. (1982). A law of the logarithm for kernel density estimators. *Annals of Probability* **10**, 414-422.
- [St3] Stute W. (1986). On the almost sure convergence of conditional empirical distribution functions. *Annals of Probability* **14**, 891-901.
- [Ve] Venter J. H. (1967). On estimation of the mode. *Annals of Mathematical Statistics* **38**, 15-29.
- [Wa] Wahba G. (1975). Optimal convergence properties of variable knot, kernel, and orthogonal series methods for density estimation. *Annals of Statistics* **3**, 15-29.
- [Wl] Walter G. G. (1977). Properties of Hermite series estimation of a probability density. *Annals of Statistics* **5**, 1258-1264.
- [Wt] Watson G. S. (1969). Density estimation by orthogonal series. *Annals of Mathematical Statistics* **40**, 1496-1498.
- [We] Wegman E. J. (1971). A note on the estimation of the mode. *Annals of Mathematical Statistics* **42**, 1909-1915.
- [Wi] Winter B. B. (1975). Rate of strong consistency of two nonparametric density estimators. *Annals of Statistics* **3**, 759-766.

## 5 – CONCLUSION

Avant de souligner partialement notre apport dans nos différents domaines de recherche, évoquons notre travail en tant qu'activité humaine. Comme tout chercheur, nous sommes avant tout redevable à nos prédécesseurs et à l'ensemble de la communauté au sein de laquelle nous cultivons notre propre originalité et sans laquelle nous, et nos découvertes, sommes insignifiants. Les idées mathématiques se transmettent, se transforment, murissent, avant de s'incarner en des résultats pertinents et revêtir ainsi une apparence durable et riche de sens. Cette alchimie de la connaissance naît de la rencontre des bons outils, de la bonne facette du problème, de la recette efficace et de l'habileté mathématique. Elle se réalise parfois par chance ou par génie, le plus souvent par un savant dosage entre persévérance, perspicacité, précieuses méprises et large culture pour multiplier les pistes malgré les doutes. Mais un tel travail ne serait que bien peu créatif sans les touches essentielles que sont la curiosité intellectuelle, l'imagination et l'intuition. Que de qualités universelles, en rien propres aux mathématiciens !

Pour clore ce document de synthèse, résumons nos principales contributions, toutes liées à la maîtrise déterministe des fluctuations du hasard. Les processus empiriques ou browniens étant étudiés depuis longtemps, contribuer à leur compréhension nécessite à chaque fois d'inventer des outils plus fins et des méthodes plus poussées. C'est pourquoi nous posons le plus souvent des questions difficiles à des processus simples, nos travaux sont donc très techniques.

### **Concernant les lois fonctionnelles du logarithme (partie 2).**

Notre première contribution est d'avoir approfondi les ingrédients et les technologies de preuves au point d'obtenir, en topologie uniforme, les vitesses exactes dans la convergence de suites d'ensembles de processus, notamment d'incrémentaux locaux, empiriques ou browniens (voir 2.1). Si nos changements de topologie pour les trajectoires browniennes résultent naturellement de cette compréhension fine des mécanismes en jeu, la compétition entre la métrique pré-hölderienne et les vitesses browniennes orchestrée pour le processus empirique constitue un second apport inédit, qui a demandé des inventions spécifiques (voir 2.2). Notre troisième contribution marquante réside en l'utilisation aux petites échelles des petites déviations poissonniennes pour distinguer les oscillations browniennes et empiriques afin d'en déduire une borne inférieure trajectorielle dans le principe d'invariance fort (voir 2.1.5), ce qui ouvre la voie à une généralisation de ce raisonnement. Nos lois intermédiaires forment un quatrième apport bien identifié (voir 2.3), en réponse à une problématique de notre cru, et garantir ainsi des points limites aux petits ensembles de trajectoires est exploitable en statistique non paramétrique.

Nos directions récentes dans les raffinements techniques, de type temps d'atteinte presque sûr ou comportement joint des incréments passés et futur pour le processus empirique, s'organisent autour de besoins ressentis en statistique asymptotique. À l'inverse, c'est notre connaissance des phénomènes oscillatoires stochastiques qui oriente notre recherche des normalisations universelles et des topologies adaptées aux fluctuations de sommes de variables dans un domaine  $\alpha$ -stable, ou multidimensionnelles, et guide notre étude trajectorielle des processus empiriques et gaussiens indexés par des fonctions.

### **Concernant les outils probabilistes (partie 3).**

En probabilités de petites déviations, la compréhension des rectifications de la trajectoire brownienne contrainte à passer dans un tube avec des singularités a réglé une question pointue sur la vitesse d'approche de la frontière de Strassen (voir 3.1). Nous réutilisons actuellement l'argument géométrique clef dans le cas du processus de Poisson, simple ou composé. L'obtention de la probabilité de petites boules browniennes en norme uniforme pondérée (voir 3.2) nous permet à présent de contrôler finement les oscillations empiriques directement dans cette topologie forte. Notre principe de grandes déviations pour les argmin de processus est un apport méthodologique qui dépasse le seul cas du mouvement brownien (voir 3.3), même si nous l'utilisons et l'affinons dans ce cas. Enfin, notre étude des vitesses de convergence universelles dans le principe d'invariance pour le processus empirique indexé par une classe de Donsker (voir 3.4) est prometteuse, avec de nombreux prolongements et un fort potentiel d'applications. Notre apport ici est d'avoir réuni les inégalités récentes pour élaborer un principe de

démonstration simple, général et efficace, dans un domaine délicat et propice aux complications. Nous recherchons à présent les pires situations, atteignant nos vitesses, et l'extension de notre technique au cas voisin de l'approximation gaussienne des sommes, indépendantes ou non, de variables aléatoires à valeurs dans un espace de Banach.

#### **Concernant les applications en statistique non paramétrique (partie 4).**

Nous avons décrit les fluctuations presque sûres d'estimateurs de la densité grâce aux lois fonctionnelles sur les incréments empiriques qui nous ont notamment permis d'ajuster les hypothèses générales réglant la déviation stochastique des estimateurs par delta-suites (voir 4.1). Notre apport ici est d'avoir relié la forme des delta-fonctions au phénomène oscillatoire occasionné par la mesure empirique, et dégagé des vitesses pour des estimateurs à noyau perturbés ou par projection. Réciproquement, ce domaine de l'estimation fonctionnelle est une source d'inspiration pour les raffinements des lois sur les incréments locaux (voir 2.6, 2.3.3). En estimation du shorth nous avons introduit une méthode complète, originale et intrinsèque (voir 4.2) pouvant aussi s'adapter pour résoudre des problèmes connexes (voir 4.4). Notre contribution est d'avoir disséqué le problème de la localisation d'un argmin par une approche générique et des outils appropriés, construits pour l'occasion. Cette méthode se prolonge dans le cadre de l'estimation du mode, ou d'autres singularités locales. Nous avons ainsi caractérisé les déviations du mode empirique par un unique résultat (4.3), assez général pour englober les situations usuelles et bien au delà, assez précis pour induire les vitesses optimales par équilibre entre biais et variance. La compréhension des oscillations de l'argmin permet de construire de nouvelles procédures d'estimation. Dans le domaine de la statistique asymptotique notre approche résolument axée sur un contrôle précis des processus sous-jacent a donc porté ses fruits, et devrait en fournir d'autres (voir 4.4).

#### **Bilan.**

Au final, comprendre toujours mieux les processus fondamentaux ouvre des perspectives d'application mais le chemin inverse qui consiste à extraire d'un besoin de compréhension en statistique une nouvelle question sur un processus stochastique est particulièrement motivant. C'est celui que nous privilégions dorénavant, et nos aspirations récentes nous poussent à poser des questions plus simples à des processus plus complexes, en tirant profit de notre spécialisation dans les fluctuations d'échantillonnage.

Concluons en manifestant notre satisfaction d'avoir répondu à des questions qui nous préoccupent depuis longtemps. D'une part, celle du rapport entre les propriétés trajectorielles sur  $\mathbb{R}$  et le principe d'invariance. D'autre part, trois problèmes soulevés en congrès international par des spécialistes de renom il y a quelques années : celui de la rectification brownienne, celui des vitesses déconcertantes en estimation du mode, celui de l'approximation gaussienne dans le cas indexé par des fonctions.



## Publications

- [1] *Vitesses de recouvrement dans les lois fonctionnelles du logarithme itéré pour les accroissements du processus empirique uniforme avec applications statistiques* (1996). Thèse de doctorat de l'Université Paris 6, 396 pp, félicitations du jury : J. Beirland, D. Bosq, J. Bretnolle, P. Deheuvels, P. Massart présidé par M. Yor (rapporteurs : L. Devroye, D.M. Mason).
- [2] On the rate of clustering to the Strassen set for increments of the uniform empirical process (1997). *Journal of Theoretical Probability* **10**, 557-579.
- [3] Comportement oscillatoire du processus empirique et topologie Höldérienne (1998). *C. R. Acad. Sci. Paris* **324**, 105-110.
- [4] Small ball estimates for Brownian motion under a weighted sup-norm (2000, avec **Z. Shi**, Paris 6). *Studia Scientiarum Mathematicarum Hungaricae* **36**, 275-289.
- [5] Functional laws and clustering rates for sets of Wiener processes in Hölder topology (2002, dedicated to Professor Pál Révész), *Limit Theorems in Probability and Statistics IV, Vol. 1*, 159-179, Budapest 1999, I. Berkes, E. Csáki, M. Csörgö eds.
- [6] Some exact rates in the functional law of the iterated logarithm (2002, avec **M. Lifshits**, St-Petersbourg). *Annales de l'Institut Henri Poincaré, Prob. et Stat.* **38**, 811-824.
- [7] Module d'oscillation fonctionnel de quelques processus réels (2003). *C. R. Acad. Sci. Paris* **337**, 415-418.
- [8] Functional Chung laws for small increments of the empirical process and a remark on the strong invariance principle (2004). *Annales de l'Institut Henri Poincaré, Prob. et Stat.*, 32 pp, à paraître.
- [9] Inner rates of coverage of Strassen type sets by increments of the uniform empirical and quantile processes (2005). *Stochastic Processes and their Applications* **115**, 493-537.
- [10] Almost sure asymptotic behaviour of the shorth estimators (2005, avec **C. El Nouty**, Paris 6). *Mathematical Methods of Statistics*, 30 pp, à paraître.
- [11] Revisiting two strong approximation results of Dudley and Philipp (2005, avec **D.M. Mason**, Delaware). *High Dimensional Probability IV, IMS Lecture Notes-Monograph Series*, 18 pp, à paraître.

## Prépublications (date première version)

- [12] Exact rate of strong uniform consistency for density estimators based on Delta sequences (1999).
- [13] Almost sure oscillations of a Wiener process in Hölder topology (2001).
- [14] Between local and global functional laws for the empirical process (2002).
- [15] Almost sure deviations of some estimators of the mode (2003).
- [16] Convergence of small sets of Brownian increments (2004).
- [17] On the oscillation behavior of the uniform empirical process under Hölder-type metric (2005).

## Articles en préparation (date début travaux)

- [18] A note on very large increments of a Wiener process (2000).
- [19] Functional laws for lag increments of the empirical process (2001).
- [20] Waiting for the next visit of empirical and Brownian increment processes (2003).
- [21] The largest deviation of density estimators evaluated at a few points (2003).
- [22] Inner and outer rates of coverage in the Erdős-Rényi law for the random walk in  $\mathbb{R}^d$  (2003, avec **U. Einmahl**, Bruxelles).
- [23] Universal functional laws in weak topology for sums of random variables attracted by a stable law (2004, avec **A. Frolov**, St-Petersbourg).
- [24] Fractal dimension of some random points generated by empirical increments (2005).
- [25] Strong invariance principles for empirical processes (2005, avec **D.M. Mason**, Delaware).