



HAL
open science

Théorèmes limites fonctionnels pour des U-statistiques échantillonnées par une marche aléatoire. Étude de modèles stochastiques de repliement des protéines

Véronique Ladret

► **To cite this version:**

Véronique Ladret. Théorèmes limites fonctionnels pour des U-statistiques échantillonnées par une marche aléatoire. Étude de modèles stochastiques de repliement des protéines. Mathématiques [math]. Université Claude Bernard - Lyon I, 2004. Français. NNT: . tel-00008740

HAL Id: tel-00008740

<https://theses.hal.science/tel-00008740>

Submitted on 10 Mar 2005

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.

UNIVERSITÉ CLAUDE BERNARD LYON I

**Théorèmes limites fonctionnels pour des U -statistiques
échantillonnées par une marche aléatoire.**

**Étude de modèles stochastiques de repliement des
protéines.**

Thèse présentée et soutenue le 2 juillet 2004

par

Véronique LADRET

en vue de l'obtention du

Diplôme de Doctorat

(arrêté du 30 mars 1992)

Spécialité : Mathématiques

Composition du jury :

Jean-Dominique DEUSCHEL	Professeur (TU Berlin), Rapporteur
André GOLDMAN	Professeur (Université Lyon 1)
Nadine GUILLOTIN	MCF (Université Lyon 1), Directrice de thèse
Alice GUIONNET	CNRS (ENS Lyon)
Christian MAZZA	Professeur (Université de Genève), Directeur de thèse
Dimitri PÉTRITIS	Professeur (Université de Rennes), Rapporteur

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude à mes directeurs de thèse, Christian Mazza et Nadine Guillotin. Je remercie Christian Mazza pour m'avoir communiqué son enthousiasme pour des problèmes situés à l'interface entre les probabilités et la biologie pour lesquels la théorie n'est encore pas écrite. Son expérience, ses conseils judicieux et sa bonne humeur m'ont beaucoup apporté. Je suis particulièrement reconnaissante à Nadine Guillotin de m'avoir initiée aux marches aléatoires en scène aléatoire en me faisant profiter de ses connaissances étendues. Sa disponibilité, son attention et son soutien sont sans doute des éléments majeurs qui m'ont permis de mener à bien ce travail.

J'adresse également mes plus sincères remerciements à Jean-Dominique Deuschel et Dimitri Pétritis pour l'honneur qu'ils me font de s'être intéressés à mon travail, en acceptant d'être les rapporteurs de cette thèse.

Mes remerciements vont naturellement à André Goldman pour m'avoir donné le goût des probabilités, lors de ses enseignements mémorables, et pour m'avoir accueillie dans son laboratoire. Je suis très honorée de voir son nom figurer parmi les membres du jury.

Alice Guionnet a toute ma gratitude pour avoir accepté de faire partie de ce jury.

Je tiens également à remercier chaleureusement Didier Piau et Jean Bérard pour leur aide et leurs conseils précieux. Ils ont eu un rôle déterminant au début de ma thèse et je peux les assurer de mon admiration et de ma sympathie.

Ce travail a été élaboré dans une ambiance détendue grâce aux occupants du bureau 174 du LaPCS : Stéphane Bessy et Pierre Charbit. Je remercie par ailleurs tous les membres du laboratoire et j'en profite pour saluer plus particulièrement Anne, Frédérique, Gabriela, Jean-Baptiste, Stephan, Mariam, Pierre, Clément, Fabien, Christine, Madame Lefranc et Pierre Calka.

RÉSUMÉ

Cette thèse se décompose en deux parties indépendantes. Notre objectif dans la première partie est d'étudier le comportement asymptotique des U -statistiques, basées sur des noyaux d'ordre 2, échantillonnées par une marche aléatoire. Plus précisément, on se donne $(S_n)_{n \in \mathbb{N}}$ une marche aléatoire sur \mathbb{Z}^d , $d \geq 1$ et $(\xi_x)_{x \in \mathbb{Z}^d}$ une collection de variables aléatoires indépendantes, identiquement distribuées, indépendante de $(S_n)_{n \in \mathbb{N}}$. On note μ la loi de ξ_0 et l'on désigne par $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, une fonction mesurable, symétrique, telle que $h \in L^2(\mu \otimes \mu)$. On s'intéresse au comportement asymptotique de la suite de processus,

$$\mathcal{U}_n(t) = \sum_{i,j=0}^{[nt]} h(\xi_{S_i}, \xi_{S_j}), \quad t \in [0, 1], \quad n = 0, 1, \dots,$$

à valeurs dans $\mathcal{D}([0, 1])$, l'espace des fonctions c.à.l.à.g. définies sur $[0, 1]$, muni de la topologie de Skorohod. Cabus et Guillin ont obtenu la distribution asymptotique de ces objets, dans le cas où la marche aléatoire, $(S_n)_{n \in \mathbb{N}}$, est récurrente sur \mathbb{Z}^2 , ainsi que dans le cas où elle est transiente sur \mathbb{Z}^d , pour $d \geq 3$. Elles ont également conjecturé la forme de la distribution limite, dans le cas de la marche aléatoire simple, symétrique, sur \mathbb{Z} . Dans le cas où $(S_n)_{n \in \mathbb{N}}$ appartient au domaine d'attraction d'une loi stable d'indice $1 < \alpha \leq 2$, nous prouvons deux théorèmes limites fonctionnels, décrivant le comportement asymptotique de $\{\mathcal{U}_n, n = 1, 2, \dots\}$. Nous démontrons ainsi, la conjecture de Cabus et Guillin. Par ailleurs, nous donnons une nouvelle preuve de leurs résultats.

Dans une seconde partie, nous étudions le comportement asymptotique du temps d'atteinte de deux versions d'un algorithme d'évolution simplifié, modélisant le repliement d'une protéine : le (1+1)-EA sur le problème LeadingOnes. Pour chaque algorithme nous donnons une loi des grands nombres, un théorème central limite et nous comparons la performance des deux modèles.

MOTS-CLÉS

Marches aléatoires, scènes aléatoires, processus stochastiques, U -statistiques, théorèmes limites fonctionnels, chaînes de Markov, algorithmes d'évolution, stratégies d'évolution, repliement des protéines.

American Mathematical Society 2000 subject classifications

60J15, 60F05, 60J10, 60F05, 92D20, 92C05.

Table des matières

Introduction générale	9
1. Présentation de la première partie	10
2. Présentation de la seconde partie	15
Partie 1. Théorèmes limites fonctionnels pour des U-statistiques échantillonnées par une marche aléatoire	21
Chapitre 1. Marches aléatoires en scène aléatoire	23
1. Description du modèle	24
2. Cas d'une marche aléatoire récurrente sur \mathbb{Z}	25
3. Cas d'une scène aléatoire vectorielle	26
4. Cas d'une marche aléatoire récurrente dans \mathbb{Z}^2	36
5. Cas d'une marche aléatoire transiente	40
Chapitre 2. U -Statistiques	43
1. Introduction	44
2. La H -décomposition	46
3. Quelques résultats concernant le comportement asymptotique des U -statistiques de type théorème central limite	48
Chapitre 3. U -statistiques indexées par une marche aléatoire à valeurs dans \mathbb{Z}^d , $d \geq 2$	55
1. Description du modèle	56
2. Cas de la marche aléatoire récurrente dans \mathbb{Z}^2	58
3. Cas de la marche aléatoire transiente dans $\mathbb{Z}^d, d \geq 3$	66
4. Commentaires et Conjecture	68
Chapitre 4. U -statistiques indexées par une marche aléatoire récurrente sur \mathbb{Z}	71
1. Description du Modèle	72
2. Résultats	73
3. Propriétés des temps d'occupation de la marche aléatoire	75
4. Preuve du Théorème 2.1	77
5. Propriétés du processus limite dans le cas dégénéré.	88
6. Preuve du Théorème 2.2	90
7. Commentaires	91
8. Problème ouvert	95

1. Une nouvelle preuve des théorèmes 2.1 et 2.2 du chapitre 3	100
2. Une nouvelle preuve des théorèmes 3.1 et 3.2 du chapitre 3	106
Partie 2. Étude de modèles stochastiques de repliement des protéines	107
Chapitre 6. Algorithmes $(1 + 1)$ et modèles de repliement des protéines	109
1. Introduction	110
2. Preuve du Théorème 1.1	116
3. Preuve du Théorème 1.2	118
4. Preuve du Théorème 1.3	123
5. Conclusion	125
6. Problèmes ouverts	126
Bibliographie	129

Introduction générale

1. Présentation de la première partie

THÉORÈMES LIMITES FONCTIONNELS POUR DES U -STATISTIQUES ÉCHANTILLONNÉES PAR UNE MARCHÉ ALÉATOIRE

Cette première partie se situe à l'interface entre les marches aléatoires en scène aléatoire et les U -statistiques. Elle concerne l'étude du comportement asymptotique de U -statistiques échantillonnées par une marche aléatoire. Nous nous intéressons plus particulièrement à leur distribution asymptotique.

Marches aléatoires en scène aléatoire

Les marches aléatoires en scène aléatoire sont des modèles de diffusion simples en milieu désordonné, à dépendances longue portée. C'est dans la perspective de répondre à certaines questions de physique théorique, liés à la modélisation de systèmes présentant une invariance d'échelles à l'aide de processus auto-similaires, à accroissements stationnaires, qu'elles ont été introduites par Kesten et Spitzer [44], en 1979. A travers les limites faibles de processus stochastiques, correctement renormalisés, définis à partir d'une interpolation linéaire de ces objets, ils ont obtenu une nouvelle classe de processus auto-similaires. Les indices d'auto-similarité de ces processus variant, en fonction des propriétés d'intégrabilité de la chaîne et de la marche.

Le modèle mathématique

Soit $(S_n)_{n \geq 0}$ une chaîne de Markov définie sur un espace d'état E et soit $(\xi_x)_{x \in E}$ un champ aléatoire indexé par E , défini sur le même espace de probabilité que $(S_n)_{n \geq 0}$. On suppose que $(S_n)_{n \geq 0}$ est indépendante de $(\xi_x)_{x \in E}$. La marche aléatoire en scène aléatoire $(Z_n)_{n \geq 0}$ est définie par

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

Résultats existants

Dans le cas où $(S_n)_{n \geq 0}$ est la marche aléatoire simple symétrique sur \mathbb{Z} et où la scène aléatoire $(\xi_x)_{x \in \mathbb{Z}}$ est constituée de variables aléatoires i.i.d., centrées, de variance finie, Kesten et Spitzer [44] ont exhibé la limite faible non gaussienne, de la suite de processus $\{(n^{-3/4} Z_{[nt]})_{t \geq 0}, n = 1, 2, \dots\}$, quand $n \rightarrow \infty$. Le processus qui apparaît à la limite, appelé *mouvement brownien en scène aléatoire brownienne*, est auto-similaire d'indice $3/4$, à accroissements stationnaires. Il ne s'agit en fait là, que d'une partie des résultats

qu'ils ont obtenus, puisqu'ils ont également considéré le cas où la marche (resp. la scène) est asymptotiquement α -stable, $1 < \alpha \leq 2$ (resp. asymptotiquement β -stable, $0 < \beta \leq 2$). Nous y reviendrons en détail, au chapitre 1. Ces résultats de convergence faible ont été améliorés récemment par des principes d'invariance fort, par Khoshnevisan et Lewis [45] pour une scène gaussienne puis par Csáki et al. [23] pour une scène plus générale. Voir aussi Maejima [48], pour l'analogie de [44], dans le cas d'une scène vectorielle.

Dans leur article, Kesten et Spitzer [44] ont aussi conjecturé que, dans le cas où $(S_n)_{n \in \mathbb{N}}$ est une marche aléatoire récurrente sur \mathbb{Z}^2 , $\{(n \log n)^{-1/2} Z_{[nt]}\}_{t \geq 0, n = 1, 2, \dots}$ converge faiblement, quand n tend vers l'infini, vers un mouvement brownien défini sur $[0, \infty[$. Cette conjecture a été prouvée par Bolthausen [13] (voir également Borodin [14]). Plus récemment, Csáki et al. [24] ont complété ce résultat en prouvant un principe d'invariance fort.

Pour $d \geq 3$, Bolthausen [13] a remarqué que, quand $n \rightarrow \infty$, la suite de processus $\{(n^{-1/2} Z_{[nt]})_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement, à une constante multiplicative près, vers un mouvement brownien (voir encore Borodin [15]). Révész et Shi [58] ont renforcé ce résultat en prouvant un principe d'invariance fort.

Citons également den Hollander [27] et den Hollander et al. [28] qui se sont intéressés aux temps d'inter-arrivée et aux propriétés mélangeantes de certains exemples de ces marches.

Enfin, Piau [53] a introduit de nouveaux exemples de marches aléatoires en scène aléatoire. Pour chacun des modèles, il a étudié la taille typique, n^α , de Z_n , où, α , l'*exposant de renormalisation* de Z_n , est tel que : $E(Z_n^2) = n^{2\alpha+o(1)}$ quand $n \rightarrow \infty$. Il a donné les exposants de renormalisation pour une marche récurrente nulle sur l'arbre binaire, puis pour une marche évoluant sur les sommets de l'arbre de Galton-Watson surcritique conditionné par la non-extinction. Le cas des processus de Bessel discrets de dimension $d \geq 0$ ainsi que la marche simple symétrique sur des graphes spécifiques sont également envisagés. Dans un article à paraître [54], il a complété ces résultats en construisant, pour tout $1/2 \leq \alpha \leq 1$, des exemples de marches aléatoires en scène aléatoire d'exposant de renormalisation α .

***U*-statistiques**

Soit ξ_1, ξ_2, \dots une suite de variables aléatoires indépendantes, identiquement distribuées. Les *U*-statistiques, introduites par Hoeffding [39] dans les années 50, désignent la classe

des estimateurs de la forme

$$\Theta_n = \binom{n}{k}^{-1} \sum_{(n,k)} h(\xi_{i_1}, \dots, \xi_{i_k}),$$

où la somme $\sum_{(n,k)}$ est prise sur tous les sous-ensembles $1 \leq i_1 < \dots < i_k \leq n$ de l'ensemble $\{1, 2, \dots, n\}$, et où $h : \mathbb{R}^k \rightarrow \mathbb{R}$, appelé noyau d'ordre k , désigne une fonction réelle symétrique (invariante par permutation de ses arguments).

Désormais, nous nous restreindrons, à un noyau d'ordre deux :

$$\Theta_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} h(\xi_i, \xi_j).$$

Nous renvoyons à l'ouvrage de Lee [46] pour une revue détaillée de la littérature sur le sujet. En particulier, la distribution asymptotique de ces estimateurs, quand $n \rightarrow \infty$, a été largement étudiée. Quand la fonction $f(x) = E(h(x, \xi_1))$ possède une variance strictement positive, Hoeffding [39] a montré que, quand $n \rightarrow \infty$, la distribution asymptotique de $n^{\frac{1}{2}}\Theta_n$ est normale. Dans le cas d'un noyau dégénéré (quand cette même variance est nulle), Gregory [38] et Serfling [66] ont montré que la suite $(n\Theta_n)_{n \geq 0}$ converge en loi vers une variable aléatoire qui s'écrit, en terme des valeurs propres d'un certain opérateur associé à h , comme une somme pondérée de variables du Chi-deux indépendantes. Notons $T_n = \binom{n}{2}\Theta_n$. Neuhaus [52] a, entre autre, exhibé la limite faible dans l'espace de Skorohod, de la suite de processus $\{(n^{-1}T_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$, qui s'écrit en terme des mêmes valeurs propres, comme somme pondérée de carrés de mouvements browniens indépendants. Voir aussi Rubin et Vitale [60] et Dynkin et Mandelbaum [32] pour des résultats sur la distribution asymptotique des U -statistiques basées, sur des noyaux d'ordre $k \geq 2$.

U -statistiques échantillonnées par une marche aléatoire

Le modèle mathématique

Soit $(X_i)_{i \geq 1}$ une suite de vecteurs aléatoires centrés, indépendants, identiquement distribués, à valeurs dans \mathbb{Z}^d , $d \geq 2$ et $S_n = X_1 + \dots + X_n$, $n = 1, 2, \dots$ la marche aléatoire associée. Soit $(\xi_x)_{x \in \mathbb{Z}^d}$ une suite de variables aléatoires indépendantes identiquement distribuées, indépendante de $(X_i)_{i \in \mathbb{N}}$. Enfin $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ désigne un noyau symétrique de degré 2. On s'intéresse au comportement asymptotique, quand $n \rightarrow \infty$, de la suite de

processus

$$U_{[nt]} = \sum_{i,j=0}^{[nt]} h(\xi_{S_i}, \xi_{S_j}), \quad t \in [0, 1], \quad n = 1, 2, \dots,$$

Résultats existants

Cabus et Guillin [18] ont étudié la convergence faible, dans l'espace de Skorohod, de la suite $\{(U_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$, quand $n \rightarrow \infty$. Dans le cas d'un noyau dégénéré, elles ont montré qu'à une constante multiplicative près, dans le cas où $(S_n)_{n \in \mathbb{N}}$ est récurrente à valeurs dans \mathbb{Z}^2 (resp. dans le cas où $(S_n)_{n \in \mathbb{N}}$ est transiente, à valeurs dans \mathbb{Z}^d , $d \geq 3$) la suite de processus $\{(n \log n)^{-1} U_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$ (resp. $\{(n^{-1} U_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$) converge faiblement vers le même processus que celui obtenu par Neuhaus [52], comme limite faible de $\{(n^{-1} T_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$. Leur preuve qui est basée sur la méthode des moments, ne s'applique plus, dans le cas de la marche aléatoire simple symétrique sur \mathbb{Z} . Cependant, Cabus et Guillin [18] ont conjecturé la forme de la limite faible, quand $n \rightarrow \infty$, de $\{(n^{-3/2} U_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$, en lien avec des sommes pondérées (par les valeurs propres d'un certain opérateur lié à h) de carrés de *mouvements browniens en scène aléatoire brownienne* (cf. Kesten et Spitzer [44]).

Objectifs

Notre objectif, dans cette première partie, est de résoudre la conjecture de Cabus et Guillin [18] et de décrire la distribution asymptotique de la suite de processus $\{(U_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$, dans le cadre où la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, à valeur dans \mathbb{Z} , vérifie les mêmes hypothèses que dans l'article de Kesten et Spitzer [44], à savoir : les $(X_i)_{i \in \mathbb{N}}$ appartiennent au domaine d'attraction d'une loi stable d'indice α , avec $1 < \alpha \leq 2$.

Plan de la partie 1

La première partie se décompose en cinq chapitres.

Dans le premier chapitre nous détaillons certains résultats classiques sur les marches aléatoires en scène aléatoire. Nous nous intéressons, plus spécifiquement, aux scènes vectorielles, puisque c'est un cadre dans lequel nous nous placerons par la suite. Nous décrivons l'analogue vectoriel du Théorème de Kesten et Spitzer [44], dû à Maejima [48]. Nous détaillons une partie de la preuve, correspondant à un cas particulier de scène vectorielle, qui est suggérée mais omise dans [48], par Maejima. Il s'agit en fait d'une simple adaptation de la preuve du résultat général. Nous donnons ensuite la version en scène-vectorielle du

Théorème de Bolthausen [13]. Nous donnons enfin l’analogue de ces résultats, pour une marche aléatoire transiente dans \mathbb{Z}^d , $d \geq 3$.

Au deuxième chapitre nous présentons brièvement quelques éléments de la théorie des U -statistiques. En particulier, nous décrivons leur structure simple, très utile en pratique dès qu’il s’agit de décrire leur comportement asymptotique. Nous rappelons notamment parmi leurs propriétés classiques, celles qui concernent leur distribution asymptotique.

Nous abordons, au chapitre trois, le thème des U -statistiques échantillonnées par une marche aléatoire. Nous détaillons les résultats existants dans ce domaine, dus à Cabus et Guillin [18]. Ce sont des résultats de type théorèmes limites fonctionnels qui concernent le cas où la marche aléatoire, qui échantillonne la U -statistique, est soit récurrente dans \mathbb{Z}^2 , soit transiente dans \mathbb{Z}^d , $d \geq 3$. Pour chacun des théorèmes de [18], nous décrivons succinctement l’idée de la démonstration. Nous expliquons pour finir, la raison pour laquelle, la technique de la preuve de Cabus et Guillin [18], ne s’applique plus dans le cas la marche aléatoire simple symétrique sur \mathbb{Z} . Nous décrivons également la conjecture de Cabus et Guillin, pour cette dernière configuration.

Le quatrième chapitre aborde enfin le coeur de nos préoccupations dans cette partie : les U -statistiques échantillonnées par une marche aléatoire asymptotiquement α -stable, d’indice $1 < \alpha \leq 2$. Nous démontrons, sous des hypothèses appropriées pour le noyau, deux théorèmes limites fonctionnels décrivant le comportement asymptotique de processus qui leur sont associés et nous prouvons ainsi la conjecture de Cabus et Guillin [18].

Pour finir, au chapitre cinq, nous adaptons la technique de preuve adoptée au chapitre quatre, au cas où la marche aléatoire est soit récurrente sur \mathbb{Z}^2 , soit transiente sur \mathbb{Z}^d , $d \geq 3$. Nous donnons alors une nouvelle démonstration des résultats de Cabus et Guillin [18].

2. Présentation de la seconde partie

ÉTUDE D'UN ALGORITHME D'ÉVOLUTION SPÉCIFIQUE MODÉLISANT LE REPLIEMENT D'UNE PROTÉINE

La seconde partie de ce travail concerne l'étude de deux versions d'un algorithme d'évolution simple modélisant le repliement d'une protéine.

Algorithmes d'évolution et optimisation

Les algorithmes d'évolution (EA) sont des algorithmes d'optimisation stochastique, basés sur une analogie avec les mécanismes biologiques de l'évolution (voir Holland [40] et Golberg [37] pour les algorithmes génétiques, Rechenberg [57] et Schwefel [65] pour les stratégies d'évolution et Fogel et al. [33] pour la programmation évolutionnaire). Ils sont largement utilisés dans une grande variété de problèmes allant de la génétique des populations, à l'apprentissage en passant par l'optimisation. La tâche de l'EA est d'explorer un espace d'états, E , afin de maximiser une certaine fonction *fitness*, $f : E \rightarrow \mathbb{R}$, encore appelée *fonction d'adaptation*. Pratiquement, l'EA procède en faisant évoluer une population d'individus, considérés comme des solutions potentielles au problème d'optimisation, sous l'action répétée d'opérations de mutations, croisements et sélections. Chaque individu est évalué numériquement selon son fitness et les individus de fitness maximum sont recherchés. La dynamique de ces algorithmes stochastiques, modélisée mathématiquement par des chaînes de Markov, simule, à l'image de ce qui se passe dans les systèmes naturels, la survie des individus les mieux adaptés.

- 1. Mutation :** La mutation modifie aléatoirement chacun des individus de la population.
- 2. Re-combinaison :** La re-combinaison, ou *crossover*, crée une nouvelle population d'"enfants", formés à partir de paires de "parents" présents à la génération précédente.
- 2. Sélection :** L'étape de sélection, ré-échantillonne la population, en sélectionnant de préférence les individus ayant de grandes valeurs relatives de fitness. Les individus présentant de faibles valeurs relatives de fitness ont tendance à être éliminés.

L'opération de sélection joue un rôle majeur dans l'algorithme puisque c'est elle qui dirige l'exploration vers les individus ayant les "meilleures" valeurs de fitness, parmi les individus déjà présents dans la population. Sans sélection, l'algorithme d'évolution ne ferait guère mieux qu'une simple recherche aléatoire. D'un autre côté, les procédures de mutations et de croisement, assurent la "diversité" de la population tout en permettant la découverte de nouveaux individus, potentiellement meilleurs du point de vue du problème d'optimisation. Dans la suite de ce travail, nous mettrons de côté l'opérateur de re-combinaison pour nous focaliser essentiellement sur des algorithmes de type mutation-sélection.

Résultats existants

En dépit du grand nombre de résultats heuristiques existants dans ce domaine (voir par exemple J. Holland [40], H.G. Beyer [8], A. Prugel-Benett et J.L. Shapiro [55]), les résultats mathématiques rigoureux décrivant le comportement des EA restent relativement clairsemés. En effet, la dynamique généralement complexe de ces algorithmes rend les résultats de complexité relativement difficiles à obtenir, et une approche commune consiste à considérer des cas simplifiés. Parmi les exceptions se trouvent un certain nombre de travaux rigoureux, que nous décrivons ici, brièvement.

Recuit simulé généralisé

R. Cerf [20, 19] a étudié des algorithmes génétiques modélisés par des chaînes de Markov à transitions rares, qui apparaissent comme des généralisations de processus de recuit simulé. Il a, en particulier, décrit la mesure invariante du régime à basse température, c.à.d. le comportement asymptotique de la population. Ses résultats mettent notamment en lumière, l'existence d'une taille de population critique (liée au problème d'optimisation) au delà de laquelle la convergence vers les maxima globaux de la fonction d'adaptation devient possible.

Population infinie

P. Del Moral et A. Guionnet [26] ont considéré la limite en population infinie d'algorithmes de mutation-sélection en liaison avec le filtrage non-linéaire. Ils ont montré que la distribution empirique de la population converge vers la solution d'un *système dynamique à valeurs mesures* sur l'espace d'état. Voir aussi M. Vose [68] pour une étude de systèmes dynamiques similaires. Toujours en population infinie, Y. Rabinovich et A. Wigderson [56] ont obtenu des résultats concernant la vitesse de convergence de plusieurs algorithmes génétiques définis sur des chaînes binaires, et le taux de croissance du fitness moyen,

d’algorithmes définis sur les entiers, pour différents schémas de re-combinaison. Enfin, C. Mazza et D. Piau [49] ont étudié l’effet de la sélection par rapport à la mutation et la re-combinaison, sur deux algorithmes génétiques en population infinie, ayant pour but de maximiser une fonction d’adaptation linéaire sur la demi droite réelle ou discrète. Ils ont calculé leurs taux de croissance et ont montré, dans les deux cas, que la sélection était capable de surmonter une certaine tendance à converger vers zéro.

Population finie

J. Bérard et A. Bienvenue [6, 7] ont décrit le comportement asymptotique d’algorithmes de mutation-sélection simples faisant évoluer une population de taille finie. Toujours dans le cadre d’une population finie, J. Bérard [5] s’est intéressé à deux exemples d’algorithmes génétiques en environnement aléatoire. Dans le premier exemple, l’espace d’état est l’arbre de Galton-Watson surcritique conditionné par la non-extinction et la fonction d’adaptation correspond à la distance à la racine. Dans le deuxième algorithme, l’espace d’état est un arbre régulier, sur lequel la fonction d’adaptation est définie par une marche aléatoire indexée par l’arbre. Dans les deux cas, il a montré qu’après n étapes, l’algorithme trouve la valeur maximum possible de la fonction d’adaptation, à une constante aléatoire, finie, près.

Stratégies d’évolution (ES)

Citons enfin les résultats de complexité de H. Muhlenbein [51], T. Bäck [2], G. Rudolph [63], J. Garnier et al. [34], et S. Droste et al. [30, 31] concernant, certains algorithmes (1+1)-ES, ou (1+1)-EA, qui font partie de la classe des *stratégies d’évolution* (ES), définie ci-après. Les ES désignent une classe d’EA simplifiés pour lesquels le schéma de sélection est déterministe. Ils ont été introduits en Allemagne par Rechenberg [57] et Schwefel [65]. Citons également Beyer et Schwefel [9] pour une revue détaillée de la littérature sur le sujet. On en distingue deux types selon la procédure de sélection utilisée :

- Les (λ, μ) -ES, ou (λ, μ) -EA : μ parents donnent naissance à λ enfants et les μ meilleurs enfants (en terme de fitness) sont choisis pour former la nouvelle population. Le meilleur fitness de la population peut donc, éventuellement, décroître.
- Les $(\lambda + \mu)$ -ES, encore notés $(\lambda + \mu)$ -EA : μ parents donnent naissance à λ enfants, les μ meilleurs parmi parents et enfants sont sélectionnés pour former la population suivante. Cet algorithme est dit *élitiste* : contrairement à ce qui se passe dans le cadre des algorithmes génétiques classiques où le remplacement est “générationnel”, c.à.d. que tous les parents sont remplacés par leurs descendants,

les $(\lambda + \mu)$ -ES utilisent une procédure de remplacement “modifiée” qui met en compétition “parents” et “enfants”.

Des EA simples modélisant le repliement des protéines

Dans cette partie, nous nous intéressons à une modélisation probabiliste simple du phénomène de repliement des protéines (ou chaînes polypeptidiques) : le modèle *fermeture éclair* ou *zipper model*, des biophysiciens A. Bakk et al. [3]. Notre cadre d’étude est celui des algorithmes d’évolution dont la population est réduite à un unique individu (qui, ici, code la structure de la protéine). Dans ce modèle simplifié, la dynamique du repliement fait évoluer la protéine vers sa configuration “native”, ou biologiquement active, en empruntant un chemin aléatoire guidé, de repliements successifs. Ceux-ci s’effectuent le long d’une suite de points de contacts, situés à différents endroits de la protéine. Plus précisément, la chaîne polypeptidique, que l’on suppose linéaire, est divisée en n sous-structures, ou sous-unités, indépendantes, et l’on considère qu’il y a deux états possibles pour chaque sous-structure : la position pliée ou la position dépliée. On peut alors coder sa structure spatiale, à l’aide d’une chaîne binaire de longueur n ,

$$\phi = (\phi_1, \dots, \phi_n),$$

où $\phi_i = 1$ (resp. 0) si la i -ème sous-structure se trouve en position pliée (resp. dépliée). La dynamique de A. Bakk et al. [3], n’est autre que la dynamique de Métropolis associée à l’Énergie

$$\mathcal{H} = -L(\phi),$$

où

$$L(\phi) = (\phi_1 + \phi_1\phi_2 + \dots + \phi_1 \dots \phi_n)$$

compte la longueur du plus grand préfixe de uns dans la chaîne de longueur n .

Nous étudierons, au chapitre 6, le modèle de A. Bakk et al. [3], dans le cadre de la température nulle, sous deux schémas de mutations. Cette dynamique est connue, dans un cadre d’optimisation, sous le nom d’algorithme $(1 + 1)$ -ES sur le problème LeadingOnes.

Objectif

Notre objectif dans cette partie est d’étudier le comportement asymptotique du temps d’atteinte de l’état natif, pour deux versions de cet algorithme $(1 + 1)$, sous les deux scénarios de mutations suivants: le scénario “one flip” qui change un unique bit à la fois

dans la chaîne de longueur n et le scénario “Bernoulli flip” qui change chaque bit de la chaîne indépendamment les uns des autres avec probabilité c/n .

Plan de la partie 2

La deuxième partie se compose d’un unique chapitre : le chapitre 6. Dans ce chapitre, nous décrivons le modèle physique de A. Bakk et al. [3], ainsi que différents exemples, à température nulle, d’algorithmes $(1 + 1)$ associés. Nous, nous intéressons plus particulièrement, pour chacun, au comportement asymptotique de la vitesse du processus de repliement, quand la taille de la protéine tend vers l’infini. Pour chaque algorithme, nous donnons, pour cette vitesse, une loi des grands nombres, un théorème central limite et nous comparons la performance des modèles.

Partie 1

Théorèmes limites fonctionnels pour des U -statistiques échantillonnées par une marche aléatoire

CHAPITRE 1

Marches aléatoires en scène aléatoire

Nous rappelons ici, les résultats classiques concernant les marches aléatoires en scène aléatoire. Nous rédigeons en particulier une preuve, suggérée par Maejima [48], dans le cas où la marche aléatoire est récurrente dans \mathbb{Z} et où la scène est d -dimensionnelle, centrée de matrice de covariance la matrice identité et dont nous aurons besoin par la suite. Nous donnerons également un équivalent du résultat de Bolthausen [13] dans le cas vectoriel.

1. Description du modèle

Les marches aléatoires en scène aléatoire, introduites par Kesten et Spitzer [44], sont des sommes partielles de la forme

$$(1) \quad Z_n = \sum_{k=0}^n \xi_{S_k}, \quad n \geq 0,$$

où $S_n = X_1 + \dots + X_n, n = 1, 2, \dots$ est une marche aléatoire commençant en $S_0 = 0$, associée à la suite $(X_i)_{i \geq 1}$ de variables aléatoires centrées, indépendantes, identiquement distribuées, à valeurs dans \mathbb{Z}^d et où $(\xi_x)_{x \in \mathbb{Z}^d}$ désigne une suite de variables aléatoires indépendantes, identiquement distribuées de mesure de probabilité μ , indépendante de la marche aléatoire $(S_n)_{n \geq 0}$.

A partir de la suite des sommes partielles $(Z_n)_{n \in \mathbb{N}}$, on peut de façon naturelle définir la suite des processus $(Z_{[nt]})_{t \in [0, \infty[}, n = 1, 2, \dots$, à valeurs dans $D([0, \infty[, \mathbb{R})$, l'ensemble des fonctions à valeurs réelles, continues à droite avec limite à gauche (c.à.d.l.à.g.), en posant

$$Z_{[nt]} = \sum_{k=0}^{[nt]} \xi_{S_k},$$

où $[t]$ désigne la partie entière de t .

Nous définirons également la suite de processus stochastiques $(Z_{nt})_{t \geq 0}, n = 1, 2, \dots$ à valeurs dans $C([0, \infty[, \mathbb{R})$, au moyen de l'interpolation linéaire

$$(2) \quad Z_{nt} = Z_{[nt]} + (nt - [nt])(Z_{[nt]+1} - Z_{[nt]}).$$

2. Cas d'une marche aléatoire récurrente sur \mathbb{Z}

Dans le cas où la marche aléatoire $(S_n)_{n \in \mathbb{N}}$ évolue dans \mathbb{Z} , Kesten et Spitzer [44] ont montré que si les $(X_i)_{i \geq 1}$ et les $(\xi_x)_{x \in \mathbb{Z}}$ appartiennent au domaine d'attraction de Z_α et Z_β respectivement, deux variables aléatoires de lois stables d'indices respectifs $1 < \alpha \leq 2$ et $0 < \beta \leq 2$, il existe $\delta = 1 - \alpha^{-1} + (\alpha\beta)^{-1} > \frac{1}{2}$ tel que la suite de processus $\{(n^{-\delta} Z_{nt})_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement, quand $n \rightarrow \infty$, vers $(\Delta_t)_{t \geq 0}$, un processus auto-similaire d'indice δ , à accroissements stationnaires.

Plus précisément, les hypothèses de Kesten et Spitzer sont les suivantes :

- pour la marche aléatoire, les $(X_i)_{i \geq 1}$ satisfont :
 - $EX_i = 0$,
 - $\frac{1}{n^{1/\alpha}} S_n = \frac{1}{n^{1/\alpha}} \sum_{i=1}^n X_i$ converge en distribution vers Z_α , quand $n \rightarrow \infty$,

Observons que si $\alpha = 2$, Z_α est une variable gaussienne et si $\alpha < 2$, sa fonction caractéristique s'écrit sous la forme

$$\phi(\theta) = \exp[-|\theta|^\alpha (C_1 + iC_2 \operatorname{sgn} \theta)],$$

pour $0 < C_1 < \infty$ et $|C_1^{-1}C_2| \leq \tan(\frac{\pi}{2}\alpha)$.

- en ce qui concerne la scène aléatoire $(\xi_x)_{x \in \mathbb{Z}}$, on supposera que :
 - $n^{-\frac{1}{\beta}} \sum_{k=1}^n \xi_k$ converge en loi vers Z_β , quand $n \rightarrow \infty$
 - dans le cas où $\beta = 1$, on ajoute la condition de symétrie suivante : il existe un $K > 0$ tel que

$$|E(\xi_x \mathbf{1}_{\{|\xi_x| \leq \rho\}})| \leq K < \infty, \text{ pour tout } \rho > 0.$$

Afin de décrire le processus limite $(\Delta_t)_{t \geq 0}$, on introduit $(Z_+(t))_{t \geq 0}$ et $(Z_-(t))_{t \geq 0}$ qui désignent deux processus de Lévy β -stables à trajectoires continues à droite, tels que $Z_+(1)$ et $Z_-(1)$ aient la même distribution que Z_β et l'on introduit aussi $(Y(t))_{t \geq 0}$, un processus de Lévy α -stable à trajectoires continues à droite, tel que $Y(1)$ ait la même distribution que Z_α . On suppose que ces processus sont définis sur le même espace de probabilité et qu'ils sont indépendants. Notons $L_t(x)$ le *temps local* de $Y(\cdot)$ au point $x \in \mathbb{R}$. Comme $x \rightarrow L_t(x)$ est continue avec probabilité 1 (cf. [16, 35]) et que $Z_\pm(x)$ est

une semi-martingale ([50] sect. IV 15), on peut définir l'intégrale stochastique

$$\Delta_t = \int_0^\infty L_t(x) dZ_+(x) + \int_0^\infty L_t(-x) dZ_-(x).$$

Pour des raisons de simplicité on écrira

$$(3) \quad \Delta_t = \int_{-\infty}^\infty L_t(x) dZ(x).$$

Notons $\delta = 1 - \alpha^{-1} + (\alpha\beta)^{-1}$. Kesten et Spitzer [44] ont montré que le résultat suivant, concernant limite faible de $\{(Z_{nt})_{t \geq 0}, n = 1, 2, \dots\}$:

Théorème 2.1. [44] *La suite de processus $\{(n^{-\delta} Z_{nt})_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement dans $C([0, \infty[, \mathbb{R})$, quand $n \rightarrow \infty$, vers le processus $(\Delta_t)_{t \geq 0}$ défini par (3).*

Remarques : Ainsi, le processus limite, Δ_t , possède une *version continue* qui est

- (i) à *accroissements stationnaires*.
- (ii) *auto-similaire d'indice δ* , c.à.d. que pour tout $c > 0$, les deux processus $(\Delta_{ct})_{t \geq 0}$ et $(c^\delta \Delta_t)_{t \geq 0}$ sont équivalents en distribution :

$$(4) \quad (\Delta_{ct})_{t \geq 0} \stackrel{d}{=} c^\delta (\Delta_t)_{t \geq 0}.$$

3. Cas d'une scène aléatoire vectorielle

Dans le cas où la scène aléatoire, $(\xi_x)_{x \in \mathbb{Z}}$, évolue dans \mathbb{R}^d , Maejima [48] a prouvé l'analogie du Théorème 2.1.

Afin de décrire le modèle étudié par Maejima [48], nous aurons besoin de quelques préliminaires.

Soit Z_B une variable aléatoire à valeurs dans \mathbb{R}^d , *opérateur stable d'exposant B* , c'est-à-dire qu'il existe un opérateur linéaire inversible B sur \mathbb{R}^d tel que la loi μ de Z_B , soit infiniment divisible, de fonction caractéristique ϕ vérifiant pour tout $t > 0$

$$\phi(\theta)^t = \phi(t^{B^*} \theta), \forall \theta \in \mathbb{R}^d$$

où B^* désigne l'adjoint de B .

Notons $\lambda_B = \min\{\operatorname{Re} \sigma : \sigma \in \sigma(B)\}$ et $\Lambda_B = \max\{\operatorname{Re} \sigma : \sigma \in \sigma(B)\}$ où $\sigma(B)$ désigne l'ensemble des valeurs propres de B . On rappelle que μ désigne la loi de ξ_0 . Plaçons nous dans l'un ou l'autre des cas suivants :

- (i) μ est gaussienne, et dans ce cas on peut prendre $B = \frac{1}{2}I_d$ comme exposant.
- (ii) μ est purement non-gaussienne. Dans ce cas, $\lambda_B > \frac{1}{2}$ et quand μ est une mesure α -stable d -dimensionnelle, on peut prendre $B = \frac{1}{\alpha}I_d$.

Les hypothèses sur la scène aléatoire $(\xi_x)_{x \in \mathbb{Z}}$ et sur la marche aléatoire $(S_n)_{n \in \mathbb{N}}$ sont :

- les $(\xi_x)_{x \in \mathbb{Z}}$ sont des vecteurs aléatoires i.i.d., à valeurs dans \mathbb{R}^d . Ils appartiennent au domaine d'attraction de Z_B , c'est-à-dire

$$(5) \quad n^{-B} \sum_{k=1}^n \xi_k \text{ converge en distribution vers } Z_B, \text{ quand } n \rightarrow \infty.$$

où

$$n^{-B} = \exp\{\ln n(-B)\} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} (\ln n)^k B^k$$

- pour la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, les hypothèses sont les mêmes que dans la section précédente : la suite $(X_i)_{i \geq 1}$ est à valeurs dans \mathbb{Z} , elle est indépendante des $(\xi_x)_{x \in \mathbb{Z}}$ et les $(X_i)_{i \geq 1}$ appartiennent au domaine d'attraction de Z_α , de loi α -stable d'indice $1 < \alpha \leq 2$.

Comme dans la section précédente, $(Y(t))_{t \geq 0}$ désigne un processus de Lévy α -stable à trajectoires continues à droite tel que $Y(1)$ soit égal en loi à Z_α . $L_t(x)$ est le *temps local* de $(Y(t))_t$ au point x , dont on choisit une version continue en (t, x) . On désigne par $(Z_{+,B}(t))_{t \geq 0}$ et $(Z_{-,B}(t))_{t \geq 0}$ deux processus de Lévy indépendants, à valeurs dans \mathbb{R}^d , indépendants de $(Y(t))_{t \geq 0}$ tels que $Z_{+,B}(1)$ et $Z_{-,B}(1)$ aient la même la distribution que Z_B . Pour simplifier on écrira $(Z_B(t))_{t \in \mathbb{R}}$, à la place des composantes $(Z_{+,B}(t))_{t \geq 0}$ et

$(Z_{-,B}(t))_{t \geq 0}$) et on dira que $(Z_B(t))_{t \in \mathbb{R}}$ est un processus de Lévy bilatère (défini sur \mathbb{R}). Ce processus $(Z_B(t))_{t \in \mathbb{R}}$ est appelé processus de Lévy *opérateur stable*, d'exposant B . Chaque composante $Z_B^{(i)}(t), i = 1, 2, \dots, d$ est alors un processus de Lévy (non nécessairement stable) à valeurs réelles. On peut ainsi définir pour tout i , comme dans Kesten et Spitzer [44], l'intégrale stochastique

$$\Delta_t^{(i)} = \int_{-\infty}^{\infty} L_t(x) dZ_B^{(i)}(x).$$

Le processus, Δ_t , à valeurs dans \mathbb{R}^d dont la i -ème composante est $\Delta_t^{(i)}$ sera noté

$$(6) \quad \Delta_t = \int_{-\infty}^{\infty} L_t(x) dZ_B(x).$$

Nous pouvons maintenant énoncer le Théorème de Maejima [48].

Théorème 3.1. [48] *Soit $D = (1 - \alpha^{-1})I_d + \alpha^{-1}B$. Alors la suite de processus $\{(n^{-D}Z_{nt})_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement dans l'espace $C([0, \infty[, \mathbb{R}^d)$, quand $n \rightarrow \infty$, vers le processus $(\Delta_t)_{t \geq 0}$, pourvu que, dans le cas où $\lambda_B \leq 1 \leq \Lambda_B$, $\xi(0)$ soit symétrique au sens où $\xi(0) \stackrel{d}{=} -\xi(0)$. D'autre part, $(\Delta_t)_{t \geq 0}$ est auto-similaire d'exposant D , à accroissements stationnaires.*

Remarques :

- Si B est diagonalisable sur \mathbb{R} alors $Z_B^{(i)}$ est un processus uni-dimensionnel stable [41]. Ainsi, dans ce cas la i -ème composante $\Delta_t^{(i)}$ du processus Δ_t n'est rien d'autre que le processus auto-similaire qui apparaît dans [44]. Par contre si B n'est pas semi-simple, $Z_B^{(i)}(t)$ n'est pas stable. Et $\Delta_t^{(i)}$ n'est pas couvert par [44]. Dans ce cas $\Delta_t^{(i)}$ n'est pas auto-similaire et le processus $\Delta_t^{(i)}$ est différent du processus introduit par Kesten et Spitzer [44].
- Dans son article, Maejima [48] détaille la preuve du théorème précédent uniquement dans le cas (ii), où la loi, μ , de Z_B est purement non gaussienne. La preuve dans le cas gaussien étant laissée au lecteur. Cependant, le résultat qui nous servira directement, au chapitre 4, concerne le cas gaussien. Nous allons donc détailler l'équivalent de la preuve de Maejima appliqué au cas qui nous intéresse. Plus précisément, nous allons donner la démonstration ici, du théorème de Maejima dans le cadre plus restrictif où les vecteurs aléatoires $(\xi_x)_{x \in \mathbb{Z}}$ sont centrés, de matrice de covariance la matrice identité, I_d .

Désignons par $(B^{(1)}(t))_{t \in \mathbb{R}}, \dots, (B^{(d)}(t))_{t \in \mathbb{R}}$, d mouvements browniens bilatères indépendants, indépendants du processus de Lévy α -stable $(Y(t))_{t \geq 0}$ défini ci-dessus. Définissons pour tout $t \geq 0$, pour tout $n \geq 1$,

$$(7) \quad D_t^n = n^{-\delta_0} Z_{nt}, \quad \text{où} \quad \delta_0 = 1 - \frac{1}{2\alpha}.$$

Théorème 3.2. *Supposons que la scène, $(\xi_x)_{x \in \mathbb{Z}}$, soit composée de vecteurs aléatoires i.i.d., à valeurs dans \mathbb{R}^d , centrés de matrice de covariance la matrice identité, I_d . Quand $n \rightarrow \infty$, la suite $\{(D_t^n)_{t \geq 0} := n^{-\delta_0} (Z_{nt})_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement dans $C([0, \infty[, \mathbb{R}^d)$ vers le processus $(\Delta_t := (\Delta_t^{(1)}, \dots, \Delta_t^{(d)}))_{t \geq 0}$ dont la i -ème composante de Δ_t est définie par*

$$\Delta_t^{(i)} = \int_{-\infty}^{+\infty} L_t(x) dB^{(i)}(x).$$

Remarque : La démonstration est largement inspirée par les preuves données par Kesten et Spitzer [44] et Maejima [48]. Il existe cependant une légère différence que nous allons détailler, qui se situe entre autres, au niveau du calcul de la loi conjointe des Δ_{t_i} et qui concerne les Lemmes 3, 4 et 5 de [48] qui sont les équivalents, dans le cas vectoriel, des Lemmes 5 et 6 de [44].

Preuve du Théorème 3.2 : Dans la suite de cette section, $\|\cdot\|$ désignera la norme euclidienne et C désignera une constante absolue pouvant varier d'une inégalité à l'autre.

Remarquons, dans un premier temps, que les hypothèses sur la marche aléatoire entraînent (cf [36], Théorème 2 p. 480) que le processus

$$n^{-\frac{1}{\alpha}} (S_{[nt]})_{t \geq 0}$$

converge faiblement dans $D([0, \infty[, \mathbb{R})$, quand n tend vers l'infini, vers $(Y(t))_{t \geq 0}$.

Désignons par $N_n(u)$ le nombre de visites au point u de la marche aléatoire S_n , dans l'intervalle de temps $[0, n]$. On peut alors représenter Z_n sous la forme suivante

$$Z_n = \sum_{k=0}^n \xi_{S_k} = \sum_{u \in \mathbb{Z}} N_n(u) \xi_u.$$

On considère comme pour Z_t , l'interpolation linéaire de $N_n(u)$:

$$N_t(u) = N_{[t]}(u) + (t - [t])(N_{[t]+1}(u) - N_{[t]}(u)).$$

Pour $-\infty < a < b < \infty$, on pose

$$T_t^n(a, b) = \frac{1}{n} \sum_{n^{\frac{1}{\alpha}} a \leq u < n^{\frac{1}{\alpha}} b} N_{nt}(u).$$

Il s'agit de la fraction de temps dans l'intervalle $[0, nt]$, pendant laquelle le processus $n^{-\frac{1}{\alpha}} S_{[nt]}$ séjourne dans l'intervalle $[a, b]$. L'analogie de cette quantité pour le processus $Y(\cdot)$ correspond au temps d'occupation de $[a, b]$ durant $[0, t]$, c'est-à-dire,

$$(8) \quad \Lambda_t(a, b) = \int_0^t \mathbf{1}_{[a, b]}(Y(\sigma)) d\sigma.$$

Comme le processus $Y(\cdot)$ possède un temps local $L_t(x)$ continu en (t, x) , Λ_t défini par (8) est presque sûrement égal à

$$(9) \quad \Lambda_t(a, b) = \int_a^b L_t(u) du.$$

Les deux lemmes suivants sont dus à Kesten et Spitzer [44] et restent inchangés dans notre démonstration. Le premier découle directement de la convergence de $n^{-\frac{1}{\alpha}} S_{[nt]}$ vers $Y(t)$ et concerne la distribution asymptotique des $T_{t_i}^n(a_i, b_i)$, $i = 1, \dots, k$, pour des temps $t_i \geq 0$:

Lemme 3.1 ([44]). *Pour tous $t_1, \dots, t_k \geq 0$, pour tous $a_1, \dots, a_k, b_1, \dots, b_k \in \mathbb{R}$,*

$$\{T_{t_j}^n(a_j, b_j), 1 \leq j \leq k\} \xrightarrow{d} \{\Lambda_{t_j}(a_j, b_j), 1 \leq j \leq k\}.$$

Preuve : Voir [44] p.12.

De façon générale, nous définirons pour toute marche aléatoire $(S_n)_{n \in \mathbb{N}}$ et pour tout n , le temps local d'auto-intersection au temps n de $(S_n)_{n \in \mathbb{N}}$, défini par

$$(10) \quad V_n = \sum_{i, j=0}^n \mathbf{1}_{\{S_i=S_j\}}.$$

Dans le cas présent, cette quantité s'écrit encore

$$V_n = \sum_{u \in \mathbb{Z}} N_n^2(u).$$

Le second lemme concerne les temps locaux de la marche aléatoire $(S_n)_{n \in \mathbb{N}}$.

Lemme 3.2 ([44]). *Pour tout $p \geq 1$,*

$$(11) \quad \sup_{u \in \mathbb{Z}} E[N_n(u)^p] = \mathcal{O}(n^{p(1-\frac{1}{\alpha})}).$$

De plus,

$$(12) \quad P(N_n(u) > 0 \text{ pour un } u \text{ vérifiant } |u| > An^{\frac{1}{\alpha}}) \leq \varepsilon(A) \text{ pour } n \geq 1,$$

où $\varepsilon(A) \rightarrow 0$ quand $A \rightarrow +\infty$ et $\varepsilon(A)$ est indépendant de n .

Enfin,

$$(13) \quad E(V_s) = \sum_{u \in \mathbb{Z}} EN_s^2(u) = \mathcal{O}(s^{2\delta_0}).$$

Preuve : Il s'agit du Lemme 1 de [44].

C'est dans les Lemmes 3.3 et 3.4 suivants que réside la différence avec les preuves données par Maejima [48] et Kesten & Spitzer [44]. Nous allons donner ici les équivalents des Lemmes 3 et 5, [48] (analogues des Lemmes 5 et 6, [44]), qui concernaient exclusivement le cas où la scène est purement non gaussienne ($\beta < 2$). Les preuves données ici sont inspirées par les méthodes de preuves de [48] et [44]; la différence résidant principalement dans l'écriture de la fonction caractéristique du processus $(Z_B(t))_t$.

Lemme 3.3 (La loi conjointe de Δ_t). *Pour tous $t_1, t_2, \dots, t_k \geq 0$ et $\theta_1, \dots, \theta_k \in \mathbb{R}^d$,*

$$E \left[\exp \left\{ i \sum_{j=1}^k < \theta_j, \Delta_{t_j} > \right\} \right] = E \left[\exp \left\{ -\frac{1}{2} \int_{-\infty}^{+\infty} \left\| \sum_{j=1}^k L_{t_j}(u) \theta_j \right\|^2 du \right\} \right].$$

Preuve : Il s'agit ici de l'équivalent de [44], Lemme 5.

Notons d'abord qu'il existe une suite $0 = x_0^n < x_1^n < \dots$ vérifiant

$$\lim_{l \rightarrow +\infty} x_l^n = +\infty, \quad \lim_{n \rightarrow \infty} \max_l (x_{l+1}^n - x_l^n) = 0$$

telle que

$$\int_0^\infty L_{t_j}(x) dB_+(x) = \lim_{n \rightarrow \infty} \sum_{l=0}^\infty L_{t_j}(x_l^n) [B_+(x_{l+1}^n) - B_+(x_l^n)], j = 1, \dots, k \quad \text{p.s.}$$

Ceci découle du fait que $x \rightarrow L_t(x)$ est presque sûrement continue à support compact.

Notons ensuite que les $B_+(x_{l+1}^n) - B_+(x_l^n)$, $l = 1, 2, \dots$ sont indépendants de fonction caractéristique

$$\exp\left(-\frac{1}{2}(x_{l+1}^n - x_l^n) \|\theta\|^2\right).$$

Ainsi,

$$\begin{aligned}
E \left[\exp \left\{ i \sum_{j=1}^k < \theta_j, \int_0^\infty L_{t_j}(x) dB_+(x) > \right\} \right] \\
&= \lim_{n \rightarrow \infty} E \left[\exp \left\{ -\frac{1}{2} \sum_l (x_{l+1}^n - x_l^n) \left\| \sum_{j=1}^k \theta_j L_{t_j}(x_l^n) \right\|^2 \right\} \right] \\
&= E \left[\exp \left\{ -\frac{1}{2} \int_0^{+\infty} \left\| \sum_{j=1}^k L_{t_j}(u) \theta_j \right\|^2 du \right\} \right].
\end{aligned}$$

On traite enfin B_- de la même manière que B_+ . \square

Lemme 3.4. *Pour tous $t_1, t_2, \dots, t_k \geq 0$ et $\theta_1, \dots, \theta_k \in \mathbb{R}^d$,*

$$\sum_{u \in \mathbb{Z}} \left\| n^{-\delta_0} \sum_{j=1}^k N_{nt_j}(u) \theta_j \right\|^2 \xrightarrow{d} \int_{-\infty}^{+\infty} \left\| \sum_{j=1}^k L_{t_j}(u) \theta_j \right\|^2 du, \quad \text{quand } n \rightarrow \infty.$$

Preuve : On définit pour $\tau > 0$ petit et M grand

$$A_{n,l} = \{u \in \mathbb{Z} : l\tau n^{\frac{1}{\alpha}} \leq u < (l+1)\tau n^{\frac{1}{\alpha}}\}, \quad l \in \mathbb{Z},$$

$$U(\tau, M, n) = \sum_{|u| > M\tau n^{\frac{1}{\alpha}}} n^{-2\delta_0} \left\| \sum_{j=1}^k N_{nt_j}(u) \theta_j \right\|^2$$

et

$$V(\tau, M, n) = \sum_{|l| \leq M} |A_{n,l}| n^{-2\delta_0} \left\| \frac{1}{\tau n^{\frac{1}{\alpha}}} \sum_{y \in A_{n,l}} \sum_{j=1}^k N_{nt_j}(y) \theta_j \right\|^2$$

où $|A_{n,l}|$ désigne le nombre d'entiers dans $A_{n,l}$. On définit ensuite

$$\begin{aligned}
I &= \sum_{u \in \mathbb{Z}} n^{-2\delta_0} \left\| \sum_{j=1}^k N_{nt_j}(u) \theta_j \right\|^2 - U(\tau, M, n) - V(\tau, M, n) \\
&= \sum_{|l| \leq M} \sum_{u \in A_{n,l}} n^{-2\delta_0} \left\{ \left\| \sum_{j=1}^k N_{nt_j}(u) \theta_j \right\|^2 - \left\| \frac{1}{\tau n^{\frac{1}{\alpha}}} \sum_{y \in A_{n,l}} \sum_{j=1}^k N_{nt_j}(y) \theta_j \right\|^2 \right\}.
\end{aligned}$$

Pour simplifier momentanément les notations, on pose

$$g_j = N_{nt_j}(u) \quad \text{et} \quad h_j = \frac{1}{\tau n^{\frac{1}{\alpha}}} \sum_{y \in A_{n,l}} N_{nt_j}(y).$$

En utilisant l'inégalité suivante, vraie pour tous θ_1 et θ_2 dans \mathbb{R}^d

$$(14) \quad \left| \|\theta_1\|^2 - \|\theta_2\|^2 \right| \leq \|\theta_1 - \theta_2\| \left(\|\theta_1\| + \|\theta_2\| \right),$$

ainsi que l'inégalité de Cauchy-Schwartz, il suit que

$$\begin{aligned}
E(|I|) &\leq (2M+1)|A_{n,l}|n^{-2\delta_0} \sup_{u \in A_{n,l}} \left\{ E \left[\left\| \sum_{j=1}^k (g_j - h_j) \theta_j \right\| \left(\left\| \sum_{j=1}^k g_j \theta_j \right\| + \left\| \sum_{j=1}^k h_j \theta_j \right\| \right) \right] \right\} \\
&\leq C\tau M n^{-2(1-\frac{1}{\alpha})} \sup_{u \in A_{n,l}} \left\{ \left(E \left[\left\| \sum_{j=1}^k (g_j - h_j) \theta_j \right\|^2 \right] \right)^{\frac{1}{2}} \right. \\
&\quad \left. \left(\left(E \left[\left\| \sum_{j=1}^k g_j \theta_j \right\|^2 \right] \right)^{\frac{1}{2}} + \left(E \left[\left\| \sum_{j=1}^k h_j \theta_j \right\|^2 \right] \right)^{\frac{1}{2}} \right) \right\} \\
&\leq C\tau M n^{-2(1-\frac{1}{\alpha})} \sup_{u \in A_{n,l}} \left\{ \left(E \left[\sum_{j=1}^k (g_j - h_j)^2 \right] \sum_{j=1}^k \|\theta_j\|^2 \right)^{\frac{1}{2}} \right. \\
&\quad \left. \left(\left(E \left[\sum_{j=1}^k g_j^2 \right] \sum_{j=1}^k \|\theta_j\|^2 \right)^{\frac{1}{2}} + \left(E \left[\sum_{j=1}^k h_j^2 \right] \sum_{j=1}^k \|\theta_j\|^2 \right)^{\frac{1}{2}} \right) \right\}.
\end{aligned}$$

En utilisant l'inégalité suivante prouvée dans ([44])

$$\sup_{u \in A_{n,l}} E \left[|g_j - h_j|^2 \right] \leq C\tau^{\alpha-1} n^{2-\frac{2}{\alpha}}$$

ainsi que (12), on a

$$E(|I|) \leq CM\tau^{\frac{\alpha}{2}+\frac{1}{2}} = CM\tau(\tau^{\frac{1}{2}(\alpha-1)}).$$

D'autre part, comme dans [44] et [48], en utilisant (12) on voit que pour n suffisamment grand et pour tout $\eta > 0$, on peut choisir $M\tau$ assez grand pour que

$$P(U(\tau, M, n) \neq 0) \leq \eta.$$

On rappelle que $\alpha > 1$. On peut donc prendre τ de telle sorte que

$$CM\tau(\tau^{\frac{1}{2}(\alpha-1)}) \leq \eta^2.$$

Comme

$$\sum_{u \in \mathbb{Z}} \left\| n^{-\delta_0} \sum_{j=1}^k N_{nt_j}(u) \theta_j \right\|^2 - V(\tau, M, n) = I + U(\tau, M, n),$$

on peut conclure que pour de tels τ, M et pour n grand,

$$\begin{aligned}
P \left(\left| \sum_{u \in \mathbb{Z}} \left\| n^{-\delta_0} \sum_{j=1}^k N_{nt_j}(u) \theta_j \right\|^2 - V(\tau, M, n) \right| > \eta \right) &\leq P(|U(\tau, M, n)| > \frac{\eta}{2}) + P(|I| > \frac{\eta}{2}) \\
(15) &\leq 3\eta.
\end{aligned}$$

Nous allons maintenant nous intéresser à la convergence en loi de $V(\tau, M, n)$. En utilisant les notations et les résultats du Lemme 3.1, on voit que

$$V(\tau, M, n) = \sum_{|l| \leq M} \frac{|A_{n,l}|}{n^{\frac{1}{\alpha}}} \left\| \frac{1}{\tau} \sum_{j=1}^k T_{t_j}^n(l\tau, (l+1)\tau) \theta_j \right\|^2$$

et que $V(\tau, M, n)$ converge en loi, quand n tend vers l'infini, vers

$$(16) \quad \tau \sum_{|l| \leq M} \left\| \sum_{j=1}^k \frac{\theta_j}{\tau} \int_{l\tau}^{(l+1)\tau} L_{t_j}(y) dy \right\|^2$$

Pour finir, la continuité de $\sum_{j=1}^k L_{t_j}(u) \theta_j$ en tant que fonction de u et le fait que $L_{t_j}(\cdot)$ soit p.s. à support compact impliquent que, quand $\tau \rightarrow 0$ et $M \rightarrow \infty$, (16) converge vers

$$\int_{-\infty}^{\infty} \left\| \sum_{j=1}^k L_{t_j}(u) \theta_j \right\|^2 du.$$

On conclut alors la preuve du lemme en combinant ce résultat à (15). \square

Pour terminer la preuve du théorème on a besoin des lemmes suivants.

Lemme 3.5 ([44]).

$$\lim_{n \rightarrow \infty} n^{-\delta_0} \sup_{u \in \mathbb{Z}} N_n(u) = 0 \text{ en probabilité.}$$

Preuve : Ceci découle directement de (11). \square

Notons λ la fonction caractéristique de ξ_u ,

$$\lambda(\theta) = E[e^{i\langle \theta, \xi_u \rangle}], \theta \in \mathbb{R}^d.$$

Lemme 3.6. *Au voisinage de 0, $\log \lambda(\theta) \sim -\frac{1}{2} \|\theta\|^2$.*

Preuve : Cela découle du fait que ξ_u est centrée de matrice de covariance identité.

\square

On peut maintenant revenir à la preuve du théorème.

$$\begin{aligned}
I_n &:= E \left[\exp \left\{ i \sum_{j=1}^k \langle \theta_j, n^{-\delta_0} Z_{nt_j} \rangle \right\} \right] \\
&= E \left[\exp \left\{ i \sum_{j=1}^k \langle \theta_j, n^{-\delta_0} \sum_{u \in \mathbb{Z}} N_{nt_j}(u) \xi_u \rangle \right\} \right] \\
(17) \quad &= E \left[\prod_{u \in \mathbb{Z}} \lambda \left(n^{-\delta_0} \sum_{j=1}^k N_{nt_j}(u) \theta_j \right) \right]
\end{aligned}$$

Par les Lemmes 3.5 et 3.6,

$$\begin{aligned}
\lim_{n \rightarrow \infty} I_n &= \lim_{n \rightarrow \infty} E \left[\prod_{u \in \mathbb{Z}} \lambda \left(n^{-\delta_0} \sum_{j=1}^k N_{nt_j}(u) \theta_j \right) \right] \\
&= \lim_{n \rightarrow \infty} E \left[\exp \left\{ -\frac{1}{2} \sum_{u \in \mathbb{Z}} \left\| n^{-\delta_0} \sum_{j=1}^k N_{nt_j}(u) \theta_j \right\|^2 \right\} \right] \\
&= E \left[\exp \left\{ -\frac{1}{2} \int_{-\infty}^{\infty} \left\| \sum_{j=1}^k L_{t_j}(u) \theta_j \right\|^2 du \right\} \right] \text{ par le Lemme 3.4} \\
&= E \left[\exp \left\{ i \sum_{j=1}^k \langle \theta_j, \Delta_{t_j} \rangle \right\} \right] \text{ par le Lemme 3.3.}
\end{aligned}$$

Ce qui termine la preuve de la convergence des lois de dimension finie.

Afin de prouver la convergence de $\{D_t^n = n^{-\delta_0} Z_{nt}, t \geq 0\}$ vers $\{\Delta_t, t \geq 0\}$ dans $C([0, \infty[, \mathbb{R}^d)$, il reste encore à prouver la tension de la suite $\{(D_t^n)_{t \geq 0}, n = 1, 2, \dots\}$. Ici, la preuve est considérablement simplifiée par rapport à [48] en raison des hypothèses plus fortes sur les $(\xi_x)_x$ (centrés et de matrice de covariance inversible).

Lemme 3.7. *Pour tout $T < \infty$, il existe une constante $C > 0$ telle que pour tous $0 \leq t_1 < t_2 \leq T$,*

$$E \left[\left\| D_{t_2}^n - D_{t_1}^n \right\|^2 \right] \leq C(t_2 - t_1)^{2\delta_0}$$

Preuve : D'après la définition de D_t^n , (7), et l'indépendance des temps locaux d'occupation $N_{nt}(x)$ par rapport à la scène $(\xi_u)_{u \in \mathbb{Z}}$, il suit

$$\begin{aligned}
E \left[\left\| D_{t_1}^n - D_{t_2}^n \right\|^2 \right] &= E \left[\left\| n^{-\delta_0} \sum_{u \in \mathbb{Z}} (N_{nt_2}(u) - N_{nt_1}(u)) \xi_u \right\|^2 \right] \\
(18) \quad &= n^{-2\delta_0} E \left[\left\| \xi_0 \right\|^2 \right] \sum_{u \in \mathbb{Z}} E \left[(N_{nt_2}(u) - N_{nt_1}(u))^2 \right]
\end{aligned}$$

Par hypothèse, $E[|\xi_0|^2] < +\infty$ et par Kesten et Spitzer (cf. [44], p.24)

$$\sum_{u \in \mathbb{Z}} E \left[\left(N_{nt_2}(u) - N_{nt_1}(u) \right)^2 \right] \leq C[(t_2 - t_1)n]^{2 - \frac{1}{\alpha}}.$$

En se rappelant la définition (7) de δ_0 ,

$$(19) \quad \sum_{u \in \mathbb{Z}} E \left[\left(N_{nt_2}(u) - N_{nt_1}(u) \right)^2 \right] \leq C[(t_2 - t_1)n]^{2\delta_0}.$$

Ainsi on conclut le lemme, en combinant (18) et (19).

□

Comme $2\delta_0 > 1$, la tension de la suite $(D_t^n)_{t \geq}$ découle du théorème 12.5 de Billingsley [11], ce qui achève la preuve du Théorème 3.2.

□

4. Cas d'une marche aléatoire récurrente dans \mathbb{Z}^2

Le cas où les $(\xi_x)_{x \in \mathbb{Z}}$ admettent un moment d'ordre 2 et où $\alpha = 1$ (i.e., la suite $(X_i)_{i \in \mathbb{N}}$ appartient au domaine d'attraction d'une loi de Cauchy) n'est pas couvert par la section précédente. Il se traite de la même façon que le cas où la scène aléatoire, $(\xi_x)_{x \in \mathbb{Z}^2}$, admet un moment d'ordre 2 et où $(S_n)_{n \in \mathbb{N}}$ est une marche aléatoire *récurrente* sur \mathbb{Z}^2 . Dans ce cas, Bolthausen [13] a prouvé, ce qui avait été conjecturé par Kesten et Spitzer [44], à savoir, que la suite de processus $\{(Z_{[nt]})_{t \in [0,1]}, n = 1, 2, \dots\}$ converge faiblement, quand $n \rightarrow \infty$ dans $\mathcal{D}([0, \infty[, \mathbb{R})$ vers un mouvement brownien, avec une normalisation "non standard" en $(n \log n)^{\frac{1}{2}}$ (voir aussi l'article de Borodin [14]).

Plaçons nous dans les conditions définies ci-dessous :

- la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, à valeurs dans \mathbb{Z}^2 , est *apériodique* : i.e., il n'existe pas de sous-groupe propre, L , de \mathbb{Z}^2 tel qu'on puisse trouver un $x \in \mathbb{Z}^2$ avec $P(X_i = x) > 0$, vérifiant $P(X_i - x \in L) = 1$. Elle est aussi *récurrente* puisqu'on suppose ici, les $(X_i)_{i \in \mathbb{N}}$ centrés, de matrice de covariance non singulière, Σ .
- la scène $(\xi_x)_{x \in \mathbb{Z}^2}$ est constituée d'une suite de v.a.i.i.d. à valeurs réelles, indépendante des $(X_i)_{i \in \mathbb{N}}$, centrées de variance σ^2 finie,

Bolthausen [13] a prouvé le théorème central limite fonctionnel qui suit. On posera,

$$c = \frac{1}{\pi(\det \Sigma)^{1/2}}$$

Théorème 4.1. [13]. *La suite de processus $\{(\sigma^{-1}(cn \log n)^{-1/2} Z_{[nt]})_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement, quand $n \rightarrow \infty$, dans $D([0, \infty[, \mathbb{R})$ vers le mouvement brownien standard, $(B_t)_{t \geq 0}$,*

$$\left(\frac{1}{\sigma(cn \log n)^{1/2}} Z_{[nt]} \right)_{t \geq 0} \xrightarrow{\mathcal{D}} (B_t)_{t \geq 0},$$

Preuve : Voir [13].

Nous aurons besoin, au chapitre 5, de l'équivalent de ce résultat dans le cas où les $(\xi_x)_{x \in \mathbb{Z}^2}$ sont des vecteurs aléatoires de \mathbb{R}^d , $d \geq 1$, centrés, de matrice de covariance, la matrice identité I_d . Le résultat sur lequel nous nous appuyerons, et dont la preuve est une adaptation de celle de [13] est le suivant :

Théorème 4.2. *Si les $(\xi_x)_{x \in \mathbb{Z}^2}$ sont des vecteurs aléatoires i.i.d., de dimension d , centrés, de matrice de covariance I_d , alors la suite de processus $\{(Y_n(t) := (cn \log n)^{-1/2} Z_{[nt]})_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement, quand $n \rightarrow \infty$, dans $D([0, \infty[, \mathbb{R}^d)$ vers $(B_t)_{t \geq 0}$, le mouvement brownien standard d -dimensionnel,*

$$\left(Y_n(t) := \frac{1}{(cn \log n)^{1/2}} Z_{[nt]} \right)_{t \geq 0} \xrightarrow{\mathcal{D}} (B_t := (B_t^{(1)}, \dots, B_t^{(d)}))_{t \geq 0}.$$

Preuve : Convergence des lois de dimension finie. La démonstration suit le plan de [13] et s'appuie sur les résultats qui suivent, dus à Bolthausen [13] et Cabus et Guillotin [18]. On notera comme dans les sections précédentes, $N_n(x)$ le temps local d'occupation et V_n le temps local d'auto-intersection de la marche $(S_n)_{n \in \mathbb{N}}$.

Lemme 4.1. [18] *Quand $n \rightarrow \infty$, $V_n/(cn \log n)$ converge presque sûrement vers 1,*

$$\frac{V_n}{cn \log n} \xrightarrow{p.s.} 1, \quad n \rightarrow \infty.$$

Lemme 4.2. [13]

$$\sup_{x \in \mathbb{Z}^2} N_n(x) = o(n^\varepsilon), \quad p.s. \text{ pour tout } \varepsilon > 0.$$

Lemme 4.3. [13] Si $0 < a < b$, alors,

$$\sum_{j=1}^{[an]} \sum_{i=[an]+1}^{[bn]} P(S_i = S_j) = o(\log n).$$

Nous pouvons maintenant prouver la convergence des lois de dimension finie. Comme dans [13], on définit pour tout $a_1, \dots, a_m \in \mathbb{R}$, $0 = t_0 < t_1 < \dots < t_m$,

$$T_n := \sum_{j=1}^m a_j (Y_n(t_j) - Y_n(t_{j-1})) = \sum_{j=1}^m \sum_{x \in \mathbb{Z}^2} a_j (N_{[nt_j]}(x) - N_{[nt_{j-1}]}(x)) \xi_x / d_n,$$

en posant

$$d_n = (cn \log n)^{1/2}.$$

On désignera par \mathcal{A} la σ -algèbre engendrée par X_1, X_2, \dots , on notera $\|\cdot\|$ la norme euclidienne sur \mathbb{R}^d et pour tout $n \geq 2$ et pour tout $x \in \mathbb{Z}^2$, on appellera $\xi_{n,x}$ le vecteur aléatoire

$$\xi_{n,x} = \frac{\beta_{n,x}}{d_n} \xi_x,$$

où

$$\beta_{n,x} = \sum_{j=1}^m a_j (N_{[nt_j]}(x) - N_{[nt_{j-1}]}(x)).$$

On s'intéresse au comportement asymptotique, quand $n \rightarrow \infty$, de

$$T_n = \sum_{x \in \mathbb{Z}^2} \xi_{n,x}.$$

Conditionné par \mathcal{A} , T_n est la somme d'une suite triangulaire, de vecteurs aléatoires indépendants, non identiquement distribués (les $(\xi_{n,x})$). Nous allons voir que le Théorème 2.8.42 de [25] qui constitue une version multi-dimensionnelle, pour des suites triangulaires de vecteurs aléatoires, du Théorème central limite de Lindeberg, s'applique ici. Dans ce théorème la condition de Lindeberg est remplacée par la notion de *suite triangulaire* de vecteurs *asymptotiquement négligeables*.

Lemme 4.4. *Conditionnellement à \mathcal{A} , la suite triangulaire, de vecteurs aléatoires, $(\xi_{n,x})_{n,x}$, est asymptotiquement négligeable, dans le sens où pour tout $\epsilon > 0$,*

$$\sum_{x \in \mathbb{Z}^2} P(\|\xi_{n,x}\| > \epsilon | \mathcal{A}) \xrightarrow{p.s.} 0, \quad \text{quand } n \rightarrow \infty.$$

Preuve : Notons $\xi_x^{(1)}, \dots, \xi_x^{(d)}$, les d composantes du vecteur aléatoire ξ_x . Comme

$$\{\|\xi_{n,x}\| > \epsilon\} = \{\|\xi_x\| > \epsilon d_n / |\beta_{n,x}|\} \subset \bigcup_{i=1}^d \{|\xi_x^{(i)}| > (d^{-1/2} \epsilon) d_n / |\beta_{n,x}|\},$$

il suffit de prouver que pour tout $i = 1, 2, \dots, d$, la suite $(\xi_x^{(i)} \beta_{n,x} / d_n)_{n,x}$ est asymptotiquement négligeable. On écrit

$$P(|\xi_x^{(i)}| > \epsilon d_n / |\beta_{n,x}| | \mathcal{A}) \leq \frac{\beta_{n,x}^2}{\epsilon^2 d_n^2} E(\xi_x^2 \mathbf{1}_{\{|\xi_x| > \epsilon d_n / |\beta_{n,x}|\}} | \mathcal{A}),$$

$$(20) \quad \sum_{x \in \mathbb{Z}^2} P(|\xi_x^{(i)}| > \epsilon d_n / |\beta_{n,x}| | \mathcal{A}) \leq \left(\sum_{x \in \mathbb{Z}^2} \frac{\beta_{n,x}^2}{\epsilon^2 d_n^2} \right) E(\xi_{(0,0)}^2 \mathbf{1}_{\{|\xi_{(0,0)}| > \epsilon z_n\}} | \mathcal{A}),$$

où

$$z_n = d_n / \sup_{x \in \mathbb{Z}^2} |\beta_{n,x}|.$$

D'après les lemmes 4.1, 4.3 et la propriété de Markov,

$$(21) \quad \frac{1}{d_n^2} \sum_{x \in \mathbb{Z}^2} \beta_{n,x}^2 \xrightarrow{p.s.} \sum_{j=1}^m a_j^2 (t_j - t_{j-1}).$$

Pour finir, on conclut en utilisant ce résultat ainsi que les les Lemmes 4.1 et 4.2 dans l'équation (20). \square

Revenons au théorème. La suite de vecteurs aléatoires $(\xi_{n,x})_{n,x}$ vérifie les trois hypothèses suivantes,

- conditionnellement à \mathcal{A} , la suite $(\xi_{n,x})_{n,x}$ est, presque sûrement, asymptotiquement négligeable (Lemme 4.4),
- pour tout $n \in \mathbb{N}^*$,

$$\sum_{x \in \mathbb{Z}^2} E(\xi_{n,x} | \mathcal{A}) = 0,$$

- quand $n \rightarrow \infty$, (d'après (21)), en désignant par $V(\cdot)$ la matrice de covariance,

$$\Gamma_n = \sum_{x \in \mathbb{Z}^2} V(\xi_{n,x} | \mathcal{A}) = \frac{1}{d_n^2} \sum_{x \in \mathbb{Z}^2} \beta_{n,x}^2 \cdot I_d \xrightarrow{p.s.} \sum_{j=1}^m a_j^2 (t_j - t_{j-1}) \cdot I_d.$$

Le Théorème 2.8.42 de [25] s'applique alors : quand $n \rightarrow \infty$, la suite de vecteurs aléatoires $\{T_n := \sum_{j=1}^m a_j(Y_n(t_j) - Y_n(t_{j-1})), n = 1, 2, \dots\}$ est asymptotiquement distribuée selon un vecteur gaussien d -dimensionnel, de moyenne nulle et de matrice de covariance $\sum_{j=1}^m a_j^2(t_j - t_{j-1})I_d$. Par le théorème de Cramér-Wold (cf [2], Théorème 7.7), on conclut à la convergence des lois de dimension finie de $(Y_n(t))_{t \geq 0}$ vers celles du brownien standard d -dimensionnel.

Tension de la suite $(Y_n(t))_{t \in [0,1]}$. La preuve de la tension de la suite $\{(Y_n(t))_{t \geq 0}, n = 1, 2, \dots\}$ dans $\mathcal{D}([0, \infty[, \mathbb{R}^d)$ est la même que celle de [13]. Il suffit de prouver, par le théorème 8.4 de [11] que pour tout $i = 1, 2, \dots, d$, $Z_n^{(i)} := \sum_{k=0}^n \xi_{S_k}^{(i)}$, la $i^{\text{ème}}$ composante du vecteur Z_n , vérifie : pour tout $\varepsilon > 0$, il existe un $\lambda > 0$, tel que pour tout $n \in \mathbb{N}$, suffisamment grand,

$$P\left(\sup_{k \leq n} |Z_k^{(i)}| \geq \lambda(n \log n)^{1/2}\right) \leq \frac{\varepsilon}{\lambda^2}.$$

On renvoie à [13] pour une preuve détaillée de ce résultat. □

5. Cas d'une marche aléatoire transiente

Dans le cas d'une marche aléatoire $(S_n)_{n \in \mathbb{N}}$ transiente sur \mathbb{Z}^m , $m \geq 1$ et d'une scène vectorielle d -dimensionnelle, $(\xi_x)_{x \in \mathbb{Z}^m}$, constituée de vecteurs aléatoires i.i.d., centrés de matrice de covariance I_d , on obtient le même type de résultat que dans la section précédente avec une normalisation en $n^{1/2}$.

- Le cas d'une scène aléatoire unidimensionnelle ($d = 1$), où les $(X_i)_{i \in \mathbb{N}}$ évoluent dans \mathbb{Z} et appartiennent au domaine d'attraction d'une loi stable d'indice $0 < \alpha < 1$ est traité dans l'article de Kesten et Spitzer, [44], p.10.
- Bolthausen [13] a résolu, de la même façon, le cas d'une marche aléatoire transiente sur \mathbb{Z}^m , $m \geq 3$, dans le cadre d'une scène à valeurs réelles ($d=1$).

Notons,

$$G(0, 0) = \sum_{k=0}^{\infty} P(S_k = 0),$$

et posons

$$\gamma = 2G(0, 0) - 1.$$

D'après Bolthausen [13], quand, $m \geq 3$, la suite de processus $\{(Z_{[nt]}/(\sqrt{\gamma n}))_{t \geq 0}, n = 1, 2, \dots\}$ converge faiblement, dans $D([0, \infty[, \mathbb{R})$, quand $n \rightarrow \infty$, vers le mouvement

brownien standard, $(B_t)_{t \geq 0}$:

$$\left(\frac{1}{\sqrt{\gamma n}} Z_{[nt]} \right)_{t \geq 0} \xrightarrow{\mathcal{D}} (B_t)_{t \geq 0}.$$

• Si l'on considère, maintenant, une scène aléatoire $(\xi_x)_{x \in \mathbb{Z}^m}$ vectorielle de dimension $d \geq 1$, constituée de vecteurs i.i.d. centrés de matrice de covariance I_d , on peut procéder comme dans la section précédente. En s'appuyant sur la convergence presque sûre des temps locaux d'auto-intersections (voir [13] et [18]) :

$$\frac{V_n}{\gamma n} \xrightarrow{p.s.} 1, \quad \text{quand } n \rightarrow \infty,$$

on montre que, quand $n \rightarrow \infty$, la suite de processus $\left\{ (Z_{[nt]}/\sqrt{\gamma n})_{t \geq 0}, n = 1, 2, \dots \right\}$ converge faiblement, dans $D([0, \infty[, \mathbb{R}^d)$, vers le mouvement brownien standard d -dimensionnel, $(B_t^{(1)}, \dots, B_t^{(d)})_{t \geq 0}$:

$$\left(\frac{1}{\sqrt{\gamma n}} Z_{[nt]} \right)_{t \geq 0} \xrightarrow{\mathcal{D}} (B_t^{(1)}, \dots, B_t^{(d)})_{t \geq 0}, \quad n \rightarrow \infty.$$

CHAPITRE 2

U-Statistiques

Nous introduisons les objets sur lesquels nous travaillerons par la suite: les U -statistiques. Les résultats non exhaustifs exposés ici, concernent certaines propriétés asymptotiques de ces estimateurs, que nous généraliserons, aux chapitres 4 et 5, au cas des U -statistiques échantillonnées par des marches aléatoires. Il s'agit de résultats classiques que l'on peut trouver dans divers ouvrages de référence tels [66] et [46].

1. Introduction

Ce chapitre présente, de manière succincte, certaines caractéristiques des U -statistiques, notamment certaines propriétés structurelles spécifiques qui sont particulièrement utiles pour décrire leur comportement asymptotique. Nous renvoyons aux ouvrages de Serfling [66] et de Lee [46] pour une description détaillée de la théorie des U -statistiques.

Soit \mathcal{F} un ensemble de lois de probabilité définies sur \mathbb{R} . Notons $\theta(\mu)$ une fonctionnelle sur \mathcal{F} , c.à.d. un paramètre réel lié à \mathcal{F} :

$$(22) \quad \theta = \theta(\mu), \quad \mu \in \mathcal{F}.$$

Tout au long de ce chapitre, (ξ_1, \dots, ξ_n) désignera un échantillon de n variables aléatoires indépendantes, identiquement distribuées selon la loi μ .

Dans une large classe de problèmes statistiques, la fonctionnelle à estimer, θ , s'écrit sous la forme

$$(23) \quad \theta(\mu) = E\{h(\xi_1, \dots, \xi_k)\},$$

où $h : \mathbb{R}^k \rightarrow \mathbb{R}$ est une fonction borélienne *symétrique* (c.à.d. invariante par permutation de ses arguments), qui satisfait $E(|h(\xi_1, \dots, \xi_k)|) < \infty$, pour tout $\mu \in \mathcal{F}$.

Dans ce cas, il est facile de voir que l'estimateur

$$(24) \quad \Theta_n = \binom{n}{k}^{-1} \sum_{(n,k)} h(\xi_{i_1}, \dots, \xi_{i_k}),$$

où la somme $\sum_{(n,k)}$ est prise sur tous les sous-ensembles $1 \leq i_1 < \dots < i_k \leq n$ de $\{1, 2, \dots, n\}$, est un estimateur *sans biais* de θ (c.à.d. $E(\Theta_n) = \theta(\mu)$, pour tout $\mu \in \mathcal{F}$).

Définition 1.1. *La statistique, Θ_n , définie par (24) est appelée U -statistique de noyau h et d'ordre k .*

Remarque : Ces estimateurs doivent leur nom au fait qu'ils sont *sans biais* (U pour *unbiasedness* en anglais) et ont été ainsi nommés par Hoeffding [39]. Ils sont souvent l'unique estimateur sans biais symétrique de la fonctionnelle θ et sont très souvent de variance minimum parmi la classe des estimateurs sans biais de θ . On peut en citer quelques exemples.

Exemple 1 : Si $k = 1$, Θ_n est une simple somme de variables aléatoires i.i.d.. On peut prendre comme exemples la *moyenne empirique*, $\bar{\xi}_n$, ainsi que tous les *moments empiriques* d'ordre m , $m \geq 2$

$$\bar{\xi}_n = \frac{1}{n} \sum_{i=1}^n \xi_i \quad \text{et} \quad \frac{1}{n} \sum_{i=1}^n \xi_i^m, \quad m = 2, 3, \dots$$

Exemple 2 : Considérons l'estimation de ν^k où $\nu = E(\xi_1)$, k étant un entier ≥ 1 . En utilisant le noyau $h(x_1, \dots, x_k) = x_1 \cdots x_k$, on obtient la U -statistique, sans biais pour $\theta = \nu^k$, suivante :

$$\Theta_n = \binom{n}{k}^{-1} \sum_{(n,k)} \xi_{i_1} \cdots \xi_{i_k}.$$

Exemple 3 : Remarquons que la variance $\sigma^2 = \text{Var}(\xi_1)$ s'écrit

$$\sigma^2 = \frac{1}{2} E((\xi_1 - \xi_2)^2).$$

Elle est estimée par la U -statistique, sans biais pour σ^2 , basée sur le noyau $h(x_1, x_2) = (x_1 - x_2)^2/2$:

$$\Theta_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \frac{1}{2} (\xi_i - \xi_j)^2 = \frac{1}{n-1} \sum_{i=1}^n (\xi_i^2 - \bar{\xi}_n^2).$$

Exemple 4 : Dans certains cas on a besoin d'estimer $\theta = E|\xi_1 - \xi_2|$. En introduisant le noyau $h(x_1, x_2) = |x_1 - x_2|$, on obtient la U -statistique, sans biais, pour θ :

$$\Theta_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} |\xi_i - \xi_j|,$$

connue sous le nom de différence empirique de Gini.

Exemple 5 : Soit $\theta = P(\xi_1 + \xi_2 \leq 0)$. En posant $h(x_1, x_2) = \mathbf{I}_{]-\infty, 0]}(x_1 + x_2)$, la U -statistique, définie ci-dessous, est sans biais pour θ :

$$\Theta_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \mathbf{I}_{]-\infty, 0]}(\xi_i + \xi_j).$$

Il s'agit de la statistique de Wilcoxon.

2. La H -décomposition

La H -décomposition, ou décomposition de Hoeffding, correspond à une représentation des U -statistiques de degré k en tant que somme de k U -statistiques décorréélées, de degrés respectifs $1, 2, \dots, k$. Elle est fondamentale en ce qui concerne la description de leur comportement asymptotique. Commençons par introduire les espérances conditionnelles suivantes :

Définition 2.1. On définit pour tout $c = 1, 2, \dots, k$ les espérances conditionnelles

$$h_c(x_1, \dots, x_c) = E\{h(x_1, \dots, x_c, \xi_{c+1}, \dots, \xi_k)\}$$

et leurs variances

$$\sigma_c^2 = \text{Var}\{h_c(\xi_1, \dots, \xi_c)\}.$$

Pour $c = 0$, on pose

$$\sigma_0^2 = 0.$$

On dira que le noyau h , ou encore que sa U -statistique associée, présente une dégénérescence d'ordre c si $0 = \sigma_0 = \sigma_1 = \dots = \sigma_c < \sigma_{c+1}^2$.

On définit ensuite, récursivement, les noyaux $h^{(1)}, h^{(2)}, \dots, h^{(k)}$ de degrés $1, 2, \dots, k$, par les équations

$$h^{(1)}(x_1) = h_1(x_1) - \theta$$

et pour $c = 2, 3, \dots, k$,

$$(25) \quad h^{(c)}(x_1, \dots, x_c) = h_c(x_1, \dots, x_c) - \sum_{j=1}^{c-1} \sum_{(c,j)} h^{(j)}(x_{i_1}, \dots, x_{i_j}) - \theta.$$

Théorème 2.1. *Pour tout $j = 1, \dots, k$, notons $H_n^{(j)}$ la U -statistique basée sur le noyau $h^{(j)}$. Alors,*

$$(26) \quad \Theta_n = \theta + \sum_{j=1}^k \binom{k}{j} H_n^{(j)}.$$

Il s'agit de la H -décomposition de Θ_n .

L'utilité de cette écriture réside dans le fait que les $H_n^{(j)}$ sont décorréelées.

Théorème 2.2. *Les $(H_n^{(j)})_{j=1, \dots, k}$ sont décorréelées: pour tout $j < j'$,*

$$\text{Cov}(H_n^{(j)}, H_n^{(j')}) = 0.$$

D'autre part,

$$\text{Var}(H_n^{(j)}) = \binom{n}{j}^{-1} \delta_j^2,$$

$$\text{où } \delta_j^2 = \sum_{c=1}^j (-1)^{j-c} \binom{j}{c} \sigma_c^2.$$

Par conséquent, si Θ_n est une U -statistique dégénérée d'ordre d , les d premiers termes de sa H -décomposition s'annulent et sa variance est d'ordre $n^{-(d+1)}$.

Nous verrons dans la section suivante que la distribution asymptotique d'une U -statistique dépend directement de son ordre de dégénérescence. Mais tout d'abord, nous allons donner une interprétation géométrique à la H -décomposition en terme de projections orthogonales sur certains sous-espaces.

Interprétation géométrique de la H -décomposition.

Notons L_2 l'espace de Hilbert constitué des variables aléatoires de carré intégrable, muni du produit scalaire usuel. Considérons le sous-espace $L_2^{(1)}$ constitué des v.a. de la forme $\sum_{i=1}^n \psi(\xi_i)$ où ξ_1, \dots, ξ_n est une suite fixée de variables aléatoires i.i.d. telle que $E\psi^2(\xi_1)$ soit fini. Il est clair que $H_n^{(1)}$ dans la H -décomposition de Θ_n est de cette forme. En utilisant (26), on peut voir que $\Theta_n - \theta$ s'écrit comme la somme d'un élément de $L_2^{(1)}$ et d'un élément orthogonal à $L_2^{(1)}$ puisque par le Théorème 2.2, on peut écrire

$$(27) \quad \Theta_n - \theta = kH_n^{(1)} + R_n^{(1)}$$

avec $R_n^{(1)}$ orthogonal à $H_n^{(1)}$. Il s'en suit que $kH_n^{(1)}$ correspond à la projection orthogonale de Θ_n sur $L_2^{(1)}$.

Si maintenant, on définit $L_2^{(d)}$ comme étant l'espace de Hilbert composé des U -statistiques de degré d , basées sur l'échantillon ξ_1, \dots, ξ_n , de noyau de carré intégrable, alors la H -décomposition apparaît comme la représentation de $\Theta_n - \theta$ (qui appartient à $L_2^{(k)}$) en tant que somme de ses projections orthogonales sur les sous-espaces $\mathcal{M}_j = L_2^{(j)} \cap (L_2^{(j-1)})^\perp$, j allant de 1 à k .

3. Quelques résultats concernant le comportement asymptotique des U -statistiques de type théorème central limite

La plupart des résultats classiques de la théorie des probabilités qui concernent les sommes de variables aléatoires indépendantes, identiquement distribuées, ont des généralisations aux U -statistiques. Ainsi il existe des versions “ U -statistiques” du théorème central limite, de la loi des grands nombres, de la loi du logarithme itéré, de la convergence d'une loi binômiale vers une loi de Poisson, du principe d'invariance fonctionnel, etc...

Les lois limites pour les suites de U -statistiques ne sont en fait que partiellement analogues à celles des suites de sommes de v.a.i.i.d.. Cependant beaucoup de U -statistiques peuvent s'écrire, grâce à la H -décomposition, comme la somme i.i.d. de variables aléatoires, plus une petite perturbation. Dans ce cas, la théorie limite est plus ou moins la même que celle des sommes de suites de variables aléatoires i.i.d.. En revanche, dans les autres cas

la théorie limite est plus complexe.

Dans tous les cas, l'outil principal, pour le calcul des lois limites, est la H -décomposition, dont les propriétés déterminent la nature des distributions asymptotiques. Nous nous focaliserons essentiellement ici sur la version U -statistiques du théorème central limite, puisque c'est le résultat que nous généraliserons par la suite dans le cadre des U -statistiques échantillonnées par une marche aléatoire. Commençons d'abord par la classe ayant le comportement limite le plus simple : les U -statistiques asymptotiquement normales.

3.1. Cas non dégénéré.

Théorème 3.1. *Si $\sigma_1^2 > 0$. Alors $n^{\frac{1}{2}}(\Theta_n - \theta)$ est asymptotiquement normale d'espérance nulle et de variance $k^2\sigma_1^2$.*

Idée de la preuve [46] : En utilisant la H -décomposition, $n^{1/2}(\Theta_n - \theta)$ peut s'écrire sous la forme

$$\begin{aligned} n^{1/2}(\Theta_n - \theta) &= n^{1/2}(kH_n^{(1)} + R_n) \\ &= \frac{k}{n^{1/2}} \sum_{i=1}^n h^{(1)}(\xi_i) + \sqrt{n}R_n \end{aligned}$$

avec $\text{Var}(n^{1/2}R_n) = \mathcal{O}(n^{-1})$. Par le théorème de Slutsky, le comportement asymptotique de $n^{1/2}(\Theta_n - \theta)$ est le même que celui de la somme $kn^{1/2} \sum_{j=1}^n h^{(1)}(\xi_j)$. Comme $E(h^{(1)}(\xi_1)) = 0$ et $\text{Var}(h^{(1)}(\xi_i)) = \sigma_1^2 > 0$, le résultat découle alors du Théorème Central Limite.

Exemple : la variance empirique.

Notons $\mu = E(\xi_1)$ et $\sigma^2 = \sigma_1^2 = \text{Var}(\xi_1)$. La variance empirique est asymptotiquement normale de moyenne σ^2 et de variance asymptotique $(\mu_4 - \sigma^4)/n$ pourvu que $\mu_4 := E(\xi_1 - \mu)^4 > \sigma^4$.

3.2. Dégénérescence de 1ère espèce.

Quand $\sigma_1^2 = 0$, mais $\sigma_2^2 > 0$, on dit que la U -statistique a une dégénérescence de première espèce. Dans ce cas de figure, le premier terme de la H -décomposition s'annule presque

sûrement. Le même argument que celui utilisé dans la preuve du théorème précédent, implique que Θ_n se comporte asymptotiquement comme une U -statistique, basée sur noyaux de degré 2, dégénéré. On se ramène donc à l'étude des estimateurs, Θ_n , de la forme

$$(28) \quad \begin{aligned} \Theta_n &= \binom{n}{2}^{-1} \sum_{(n,k)} h(\xi_i, \xi_j) \\ &= \frac{1}{n(n-1)} \sum_{i \neq j} h(\xi_i, \xi_j), \end{aligned}$$

où $(\xi_i)_{i \in \mathbb{N}}$ désigne une suite de v.a.i.i.d., de mesure de probabilité μ et où $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ est un noyau *symétrique* ($h(x, y) = h(y, x)$), *dégénéré*, c.à.d.

$$E(h(\xi_1, \xi_2 | \xi_2)) = 0,$$

tel que

$$E(h^2(\xi_1, \xi_2)) < \infty, \quad Eh(\xi_1, \xi_2) = 0.$$

• **Décomposition des noyaux dégénérés d'ordre deux.**

Afin de décrire le comportement asymptotique de la suite $(\Theta_n)_{n \geq 2}$, nous aurons besoin de faire un rappel de théorie spectrale concernant la décomposition des noyaux h .

Si h est tel que défini ci-dessus (symétrique et dans $L^2(\mu \otimes \mu)$), on peut le plonger dans l'ensemble \mathcal{H} des fonctions $h : \mathbb{R}^2 \rightarrow \mathbb{C}$, *hermitiennes*, i.e.

$$h(x, y) = \overline{h(y, x)} \quad \forall x, y \in \mathbb{R}.$$

Notons H l'espace de Hilbert $L^2(\mu)$ et $L_{H.S.}(H)$ l'ensemble des opérateurs de Hilbert-Schmidt définis sur H . Désignons par $(e_i)_{i \in \mathbb{N}}$, une base hilbertienne quelconque de H . On rappelle que, muni du produit scalaire

$$(S|T)_{H.S.} = tr(T^*S) = \sum_{i \in \mathbb{N}} (Se_i | Te_i),$$

$L_{H.S.}(H)$ est un espace de Hilbert. Le théorème suivant (cf. par exemple [62], [17]) permet d'identifier h à un opérateur de Hilbert-Schmidt sur H .

Théorème 3.2. [62] Soit $h \in L^2(\mu \otimes \mu)$ un noyau hermitien. Alors l'opérateur intégral T_h , associé à h , défini par

$$(29) \quad \begin{aligned} T_h : L^2(\mu) &\rightarrow L^2(\mu) \\ f &\mapsto T_h(f)(x) = \int_{-\infty}^{+\infty} h(y, x) f(y) d\mu(y), \end{aligned}$$

est un opérateur de Hilbert-Schmidt sur l'espace de Hilbert $H = L^2(\mu)$.

De plus, l'application

$$(30) \quad \begin{aligned} L^2(\mu \otimes \mu) &\rightarrow L_{H.S.}(H) \\ h &\mapsto T_h \end{aligned}$$

est une isométrie surjective permettant l'identification isométrique: $L_{H.S.}(H) = L^2(\mu \otimes \mu)$.

La théorie spectrale des opérateurs compacts (cf. [62]) ainsi que l'identification isométrique (30) permet de conclure, dans le cas général, à l'existence, pour tout noyau hermitien h de $L^2(\mu \otimes \mu)$, d'une suite orthonormale $(\psi_\nu)_{\nu \geq 1}$ dans $L^2(\mu)$ de fonctions propres de T_h associées à la suite des valeurs propres non nulles de T_h , $(\lambda_\nu)_{\nu \geq 1} \in \ell^2(\mathbb{N}^*)$, (qui sont d'ailleurs réelles), telle que

$$\|h\|^2 = \sum_{\nu=1}^{+\infty} \lambda_\nu^2 \quad \text{et} \quad h(x, y) = \sum_{\nu=1}^{+\infty} \lambda_\nu \psi_\nu(x) \overline{\psi_\nu(y)} \quad \text{dans } L^2(\mu \otimes \mu).$$

Dans le cadre des U -statistiques, où le noyau h est *symétrique* réel, on remplace la suite orthonormale fonctions propres complexes $(\psi_\nu)_{\nu \geq 1}$ par une suite orthonormale de fonctions propres réelles $(\phi_\nu)_{\nu \geq 1}$ associées aux $(\lambda_\nu)_{\nu \geq 1}$.

Théorème 3.3. Soit $h \in L^2(\mu \otimes \mu)$ un noyau symétrique réel. Alors il existe $(\phi_\nu)_{\nu \geq 1}$ une suite orthonormale dans $L^2(\mu)$, de fonctions propres réelles de l'opérateur T_h , associées à la suite, $(\lambda_\nu)_{\nu \geq 1}$, des valeurs propres non nulles de T_h , telle que

$$(31) \quad \|h\|^2 = \sum_{\nu=1}^{+\infty} \lambda_\nu^2$$

et

$$(32) \quad h(x, y) = \sum_{\nu=1}^{+\infty} \lambda_\nu \phi_\nu(x) \phi_\nu(y), \quad \text{dans } L^2(\mu \otimes \mu).$$

On appellera valeurs propres de h , les valeurs propres non nulles $(\lambda_\nu)_{\nu \geq 1}$ associées à l'opérateur T_h . Notons au passage, que le caractère orthonormal de la suite $(\phi_\nu)_{\nu \geq 1}$ implique que

$$(33) \quad E(\phi_\nu(\xi_0)\phi_\mu(\xi_0)) = \delta_{\nu,\mu}, \quad \text{où} \quad \delta_{\nu,\mu} = \begin{cases} 1 & \text{si } \nu = \mu, \\ 0 & \text{sinon.} \end{cases}$$

Par ailleurs, la dégénérescence du noyau entraîne que les v.a. $(\phi_\nu(\xi_x))_{\nu \geq 1}$ sont centrées :

$$(34) \quad E(\phi_\nu(\xi_0)) = 0, \quad \text{pour tout } \nu \geq 1.$$

• Nous pouvons maintenant décrire la distribution asymptotique des U -statistiques basées sur un noyau d'ordre 2, dégénéré (cf [46]) :

Théorème 3.4. *Soit Θ_n une U -statistique d'espérance nulle basée sur un noyau, h , dégénéré ($Eh(x_1, \xi_2) = 0$) tel que $Eh^2(\xi_1, \xi_2) < \infty$, et $Eh(\xi_1, \xi_2) = 0$. Alors la statistique normalisée $n\Theta_n$ converge en loi vers la variable aléatoire*

$$(35) \quad \sum_{\nu=1}^{\infty} \lambda_\nu (Z_\nu^2 - 1)$$

où Z_1, Z_2, \dots désignent des variables aléatoires gaussiennes, centrées, réduites, indépendantes.

Idée de la preuve [46] : L'idée est de se ramener à des noyaux tronqués de la forme $h_K(x, y) = \sum_{\nu=1}^K \lambda_\nu \phi_\nu(x)\phi_\nu(y)$ et à leurs U -statistiques associées, notées $U_{n,K}$. La preuve se décompose en deux temps: on montre d'abord que quand $K \rightarrow \infty$, $U_{n,K}$ converge vers Θ_n en moyenne quadratique, uniformément en n , puis on calcule la distribution limite des $U_{n,K}$.

Neuhaus [52] a donné une version "théorème limite fonctionnel" de ce résultat. Elle concerne les processus définis comme suit :

$$(36) \quad \Theta_n(t) = \frac{1}{n} \sum_{1 \leq i \neq j \leq [nt]} h(\xi_i, \xi_j), \quad t \in [0, 1].$$

Ici, Θ_n est une variable aléatoire à valeurs dans l'espace des fonctions c.à.d.l.à.g., $(D[0, 1], \mathcal{S})$, muni de la topologie de Skorohod. Sous les hypothèses précédentes, concernant h et les $(\xi_i)_{i \in \mathbb{N}}$, le comportement asymptotique de Θ_n est décrit par le résultat suivant :

Théorème 3.5. [52]

Soit $B_t^{(1)}, B_t^{(2)}, \dots$ une suite de mouvements browniens indépendants définis sur $[0, 1]$. Alors, quand $n \rightarrow \infty$, la suite de processus $\{(\Theta_n(t))_{t \in [0, 1]}, n = 1, 2, \dots\}$ converge faiblement dans $(D[0, 1], \mathcal{S})$ vers

le processus

$$(37) \quad U(t) = \sum_{\nu=1}^{\infty} \lambda_{\nu} \left((B_t^{(\nu)})^2 - t \right),$$

à valeurs dans $(C[0, 1], \|\cdot\|_{\infty})$, l'espace des fonctions continues sur $[0, 1]$, muni de la norme de la convergence uniforme, $\|\cdot\|_{\infty}$.

Preuve : cf. Neuhaus [52].

3.3. Cas général.

Le cas général concerne les U -statistiques basées sur un noyau de degré k et dont l'ordre de dégénérescence vaut $d - 1$, i.e. $\sigma_c^2 = 0$, pour $c = 1, \dots, d - 1$ et $\sigma_d^2 > 0$. Grâce à un argument similaire à celui de la sous-section précédente utilisant la H -décomposition, on peut voir que pour obtenir le comportement asymptotique des U -statistiques d'ordre de dégénérescence supérieur, il suffit de se restreindre au cas où l'ordre de la dégénérescence est inférieur de un, au degré du noyau.

Voici l'énoncé des résultats dans le cadre général, dus à Rubin et Vitale [60] et Dynkin et Mandelbaum [32] :

Théorème 3.6. Soit Θ_n une U -statistique basée sur le noyau h et sur l'échantillon aléatoire ξ_1, \dots, ξ_n de mesure de probabilité μ . Supposons que $0 = \sigma_1^2 = \dots = \sigma_{d-1}^2 < \sigma_d^2$. Alors la distribution asymptotique de $n^{d/2}(\Theta_n - \theta)$ est celle de

$$\binom{k}{d} \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} (h^{(d)}, e_{i_1} \dots e_{i_d}) \prod_{l=1}^{\infty} H_{r_l(i)}(Z_l)$$

où e_1, e_2, \dots est une base orthonormale de $L_2(\mu)$, $h^{(d)}$ est le noyau de la U -statistique $H_n^{(d)}$ dans la H -décomposition de Θ_n , Z_1, Z_2, \dots est une suite de variables aléatoires normales centrées réduites, indépendantes, $r_l(i)$ le nombre d'indices parmi les $i = (i_1, \dots, i_d)$ égaux à l , H_r est le r -ième polynôme de Hermite et où $(h^{(d)}, e_{i_1} \dots e_{i_d})$ désigne le produit scalaire suivant

$$(h^{(d)}, e_{i_1} \dots e_{i_d}) = \int \dots \int h^{(d)}(x_1, \dots, x_d) e_{i_1}(x_1) \dots e_{i_d}(x_d) d\mu(x_1) \dots d\mu(x_d).$$

Voir Lee [46] pour la preuve de ce théorème. On remarquera que le théorème général contient les Théorèmes 3.1 et 3.4 précédents.

Nous allons, dans les chapitres suivants aborder le coeur de nos préoccupations pour la partie 1 de ce travail, les U -statistiques indexées par des marches aléatoires à valeurs dans $\mathbb{Z}^d, d \geq 1$.

CHAPITRE 3

U-statistiques indexées par une marche aléatoire à valeurs dans
 $\mathbb{Z}^d, d \geq 2$

Dans ce chapitre, nous présentons les résultats existants, dus à Cabus et Guillin [18], concernant les U -statistiques, basées sur des noyaux de degré deux, échantillonnées par des marches aléatoires récurrentes dans \mathbb{Z}^2 puis dans un second modèle, par des marches transientes à valeurs dans \mathbb{Z}^d , $d \geq 3$. Pour chaque résultat, nous donnons une idée de la preuve, avec, pour certains théorèmes, notamment ceux qui traitent du cas des noyaux non dégénérés, quelques variantes par rapport à la version de [18].

1. Description du modèle

Nous appellerons U -statistiques échantillonnées, les sommes partielles de la forme,

$$(38) \quad \sum_{1 \leq i_1, \dots, i_k \leq n} \psi(\xi_{S_{i_1}}, \dots, \xi_{S_{i_k}}),$$

où le noyau ψ désigne une fonction $\psi : \mathbb{R}^k \rightarrow \mathbb{R}$, mesurable, symétrique et où la scène $(\xi_x)_{x \in \mathbb{Z}^d}$ est indexée par une marche aléatoire $(S_n)_{n \in \mathbb{N}}$ évoluant dans \mathbb{Z}^d , avec $d \geq 1$. Ces objets apparaissent comme une extension naturelle à la théorie classique des U -statistiques, au même titre que les marches aléatoires en scène aléatoire généralisent la notion standard de sommes de v.a.i.i.d.. Nous allons détailler ici, les résultats existants dans ce domaine, dus à Cabus et Guillin [18] et qui concernent des modèles basés sur des noyaux d'ordre 2.

Soit $(\xi_x)_{x \in \mathbb{Z}^d}$ une suite de variables aléatoires i.i.d. à valeurs réelles, de mesure de probabilité μ et soit $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ un noyau (une fonction mesurable, symétrique) appartenant à $L^2(\mu \otimes \mu)$.

Soit enfin $(X_i)_{i \geq 1}$ une suite de vecteurs aléatoires i.i.d. à valeurs dans \mathbb{Z}^d , indépendants de la scène $(\xi_x)_{x \in \mathbb{Z}^d}$, centrés et dont on notera S_n , la marche aléatoire associée, i.e.

$$S_0 = 0 \quad \text{et} \quad S_n = \sum_{i=1}^n X_i, \quad \text{pour tout } n \geq 1.$$

Tout-au-long de cette partie, nous supposons que la marche aléatoire $(S_n)_{n \in \mathbb{N}}$ est fortement apériodique au sens donné par Spitzer [67], c'est-à-dire que pour tout $x \in \mathbb{Z}^d$, le sous-groupe engendré par l'ensemble

$$\{y \in \mathbb{Z}^d \mid y = x + z, \text{ où } P(X_1 = z) > 0\}$$

est égal à \mathbb{Z}^d lui-même.

On désignera comme au chapitre 2, par $N_n(x)$, le nombre de visites dans l'intervalle de temps $[0, n]$ de la marche $(S_n)_{n \in \mathbb{N}}$ au point $x \in \mathbb{Z}^d$ (que l'on appelle temps local) et par V_n , le temps local d'auto-intersection de $(S_n)_{n \in \mathbb{N}}$ défini par

$$V_n = \sum_{x \in \mathbb{Z}^d} N_n^2(x).$$

Dans leur article [18], Cabus et Guillotin se placent successivement dans les deux cas de figures suivants :

- $d = 2$ et les $(X_i)_{i \geq 1}$ (centrés) possèdent une matrice de covariance Σ non singulière; dans ce cas $(S_n)_{n \in \mathbb{N}}$ est une marche récurrente sur \mathbb{Z}^2 .
- $d \geq 1$ et $(S_n)_{n \in \mathbb{N}}$ est une marche transiente sur \mathbb{Z}^d .

Leurs résultats concernent la distribution asymptotique des suites de sommes partielles $(U_n)_{n \in \mathbb{N}}$ et $(\widehat{U}_n)_{n \in \mathbb{N}}$ définies pour tout $n \in \mathbb{N}$ par

$$U_n = \sum_{i,j=0}^n h(\xi_{S_i}, \xi_{S_j}) \quad \text{et} \quad \widehat{U}_n = \sum_{i,j=0}^n h(\xi_{S_i}, \xi_{S_j}) \mathbf{1}_{\{S_i \neq S_j\}}$$

et traitent plus spécifiquement de la limite faible de la suite de processus

$$U_{[nt]} \quad \text{et} \quad \widehat{U}_{[nt]}, \quad t \in [0, 1], \quad n = 1, 2, \dots$$

à valeurs dans $(D[0, 1], \mathcal{S})$ l'espace des fonctions c.à.d.l.à.g. définies sur l'intervalle $[0, 1]$, muni de la topologie de Skorohod.

Rappelons tout d'abord que le fait que $h \in L^2(\mu \otimes \mu)$, permet (d'après (32), chapitre 3), l'identification

$$h(x, y) \stackrel{L^2(\mu \otimes \mu)}{=} \sum_{\nu=1}^{+\infty} \lambda_\nu \phi_\nu(x) \phi_\nu(y),$$

où $(\phi_\nu)_{\nu \geq 1}$ désigne une suite orthonormale de fonctions propres de l'opérateur intégral T_h défini par (29), associées aux valeurs propres $(\lambda_\nu)_{\nu \geq 1}$.

Par ailleurs $(\lambda_\nu)_{\nu \geq 1} \in \ell^2(\mathbb{N})$ puisque d'après (31),

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h^2(x, y) d\mu(x) d\mu(y) = \sum_{\nu=1}^{\infty} \lambda_\nu^2.$$

2. Cas de la marche aléatoire récurrente dans \mathbb{Z}^2

Dans cette section nous nous plaçons dans le cas où $(S_n)_{n \in \mathbb{N}}$ est une marche aléatoire récurrente dans \mathbb{Z}^2 .

2.1. Cas dégénéré.

Ici, le noyau h est dégénéré, c'est-à-dire

$$E(h(\xi_{(0,0)}, x)) = 0, \quad \text{pour } (\mu)\text{-p.t. } x \in \mathbb{R}.$$

Nous allons d'abord énoncer le théorème limite qui concerne le processus $\widehat{U}_n(t)$ défini par

$$(39) \quad \widehat{U}_n(t) = \frac{\pi(\det \Sigma)^{1/2}}{n \log(n)} \widehat{U}_{[nt]}.$$

Cabus et Guillin [18] ont prouvé qu'avec une normalisation non-standard en $n \log n$, $\widehat{U}_n(t)$ converge faiblement vers le processus-limite $U(t)$ obtenu par Neuhaus [52] dans le cadre des U -statistiques classiques et qui est défini en (37). On notera que cette normalisation correspond au carré du facteur normalisant qui apparaît dans le résultat de Bolthausen [13] sur les marches aléatoires en scène aléatoire (cf. chap. 2, Théorème 4.1).

Au lieu de l'hypothèse de L^2 -intégrabilité par rapport à la mesure produit $\mu \otimes \mu$, retenue dans le cadre des U -statistiques classiques, nous aurons besoin ici, pour assurer la tension de la suite $\widehat{\mathcal{U}}_n$, d'une condition plus forte de L^4 -intégrabilité :

$$(40) \quad h \in L^4(\mu \otimes \mu),$$

Sous les hypothèses précédentes, la distribution asymptotique de $\widehat{\mathcal{U}}_n$ est décrite dans le Lemme 3.2, [18] cité ci-dessous :

Théorème 2.1. [18] *La suite de processus $\{(\widehat{\mathcal{U}}_n(t))_{t \in [0,1]}, n = 1, 2, \dots\}$ converge faiblement, quand $n \rightarrow \infty$, dans $D[0, 1]$ vers $(U(t))_{t \in [0,1]}$, le processus défini ci-dessous,*

$$\left(\widehat{\mathcal{U}}_n(t) := \frac{\pi(\det \Sigma)^{1/2}}{n \log(n)} \widehat{U}_{[nt]}\right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(U(t) := \sum_{\nu=1}^{\infty} \lambda_{\nu} \left((B_t^{(\nu)})^2 - t \right)\right)_{t \in [0,1]},$$

où $(B_t^{(\nu)})_{\nu \in \mathbb{N}}$ désigne une suite de mouvements browniens indépendants.

Idée de la preuve [18]

Fixons $\theta_1, \dots, \theta_m \in \mathbb{R}$ et $0 = t_0 < t_1 < \dots, < t_m \leq 1$ et posons,

$$(41) \quad \mathcal{W}_n = \sum_{j=1}^m \theta_j \widehat{\mathcal{U}}_n(t_j).$$

Les deux arguments clés de la preuve sont contenus d'une part, dans les quatre lemmes suivants qui viennent compléter les Lemmes 2.3, 2.4 et 2.6 de [13] et qui concernent la convergence presque sûre des temps locaux d'auto-intersection renormalisés de la marche et d'autre part, dans le Lemme 2.5 qui permet véritablement d'identifier, grâce à la méthode des moments, la distribution asymptotique de \mathcal{W}_n à celle de $\sum_{j=1}^m \theta_j U(t_j)$.

On rappelle la notation,

$$c = \frac{1}{\pi(\det \Sigma)^{1/2}},$$

Lemme 2.1. [18] *Quand $n \rightarrow \infty$, $V_n/(cn \log n)$ converge presque sûrement vers 1,*

$$\frac{V_n}{cn \log n} \xrightarrow{p.s.} 1, \quad n \rightarrow \infty.$$

Lemme 2.2. [18] *Pour tout $p \geq 1$ et tout $\delta > 0$, $\frac{1}{(n \log n)^{1+\delta}} \sum_{x \in \mathbb{Z}^2} N_n(x)^p$ converge presque sûrement vers 0, quand n tend vers l'infini,*

$$\frac{1}{(n \log n)^{1+\delta}} \sum_{x \in \mathbb{Z}^2} N_n(x)^p \xrightarrow{p.s.} 0, \quad n \rightarrow \infty.$$

Lemme 2.3. [18] *Pour tout $0 \leq a < b$,*

$$\sum_{i=0}^{[an]} \sum_{j=[an]+1}^{[bn]} 1_{\{S_i=S_j\}} \stackrel{p.s.}{=} o(n \log n)$$

Lemme 2.4. [18] *Pour tout $p \geq 1$,*

$$E(V_n^p) = \mathcal{O}((n \log n)^p).$$

Le plan de la preuve est le suivant (cf. [18]) :

- **Convergence des lois de dimension finie :**

Cette partie de la preuve de Cabus et Guillotin [18] se décompose en deux étapes :

– On considère dans un premier temps un noyau h borné par une constante $M > 0$.

La technique de la preuve est basée sur la méthode des moments appliquée à la suite

$$\mathcal{W}_n = \sum_{j=1}^m \theta_j \widehat{\mathcal{U}}_n(t_j) = \frac{1}{cn \log n} \sum_{x \neq y} \left(\sum_{j=1}^m \theta_j N_{[nt_j]}(x) N_{[nt_j]}(y) \right) h(\xi_x, \xi_y),$$

conditionnée par la σ -algèbre \mathcal{A} engendrée par X_1, X_2, \dots et dont les moments $\mu_k^{(n)} = E((\mathcal{W}_n)^k | \mathcal{A})$ s'écrivent pour tout $k \geq 1$,

$$(42) \quad \mu_k^{(n)} = \frac{1}{(cn \log n)^k} \sum_{(x_i, y_i); x_i \neq y_i} \left[\prod_{i=1}^k \sum_{j=1}^m \theta_j N_{[nt_j]}(x_i) N_{[nt_j]}(y_i) \right] E\left(\prod_{i=1}^k h(\xi_{x_i}, \xi_{y_i}) \right).$$

Le cas $k = 1$ se traite directement : $E(\mathcal{W}_n | \mathcal{A}) = 0$ puisque $E(h(\xi_x, \xi_y)) = 0$ pour tout $x \neq y$.

Regardons maintenant le cas $k \geq 2$: par la dégénérescence de h et par convergence presque sûre des temps locaux d'auto-intersections (Lemmes 2.1 et 2.2), on peut voir (cf. [18] pour une plus de détails) que pour déterminer la limite des $\mu_k^{(n)}$ quand $n \rightarrow \infty$, il suffit de considérer dans la somme de droite, les $\{(x_i, y_i); x_i \neq y_i; 1 \leq i \leq k\}$ tels que chaque indice x_i (indifféremment y_i) apparaît exactement deux fois parmi $\{x_1, y_1, \dots, x_k, y_k\}$. On va donc se restreindre dans le terme de droite de (42), à la quantité notée $\hat{\mu}_k^{(n)}$, correspondant à la somme sur les $\{(x_i, y_i); x_i \neq y_i; 1 \leq i \leq k\}$ qui s'écrivent, pour des paires

$(n_i, e_i), 1 \leq i \leq r$ d'entiers vérifiant $n_1 e_1 + \dots + n_r e_r = k$, comme réunion disjointe de e_1 cycles d'ordre n_1, \dots, e_r cycles d'ordre n_r , où un cycle d'ordre m désigne une suite de points de \mathbb{Z}^2 de longueur $2m$ de la forme $(x_1, x_2)(x_2, x_3) \dots (x_{m-1}, x_m)(x_m, x_1)$, constituée de $x_i, 1 \leq i \leq m$, deux à deux distincts.

Afin de calculer la limite quand $n \rightarrow \infty$, des $\mu_k^{(n)}$ (c'est-à-dire des $\hat{\mu}_k^{(n)}$), on a besoin d'évaluer pour tout $1 \leq p \leq k$ les quantités

$$I_p = E(h(\xi_{x_1}, \xi_{x_2}) \dots h(\xi_{x_p}, \xi_{x_1}))$$

calculées sur les cycles $(x_1, x_2)(x_2, x_3) \dots (x_p, x_1)$ d'ordre p .

Lemme 2.5. *Pour tout noyau $h \in L^2(\mu \otimes \mu)$ dégénéré, non nécessairement borné,*

$$(43) \quad I_k = \sum_{\nu=1}^{\infty} \lambda_{\nu}^k, \quad \text{quel que soit } k \geq 2.$$

Preuve : Nous allons d'abord montrer que pour tout noyau $h \in L^2(\mu \otimes \mu)$ dégénéré, non nécessairement borné et pour tout cycle d'ordre $k \geq 2, (x_1, x_2)(x_2, x_3) \dots (x_k, x_1)$,

$$(44) \quad E(|h(\xi_{x_1}, \xi_{x_2})h(\xi_{x_2}, \xi_{x_3}) \dots h(\xi_{x_k}, \xi_{x_1})|) \leq \left(E(h(\xi_{x_1}, \xi_{x_2})^2) \right)^{k/2} < \infty.$$

Il suffit pour cela d'appliquer une première fois l'inégalité de Cauchy-Schwartz,

$$E(|h(\xi_{x_1}, \xi_{x_2}) \dots h(\xi_{x_k}, \xi_{x_1})|) \leq \left(E(h(\xi_{x_1}, \xi_{x_2})^2) \right)^{1/2} \left(E(h(\xi_{x_2}, \xi_{x_3})^2 \dots h(\xi_{x_k}, \xi_{x_1})^2) \right)^{1/2},$$

puis de conditionner le terme de gauche par rapport à ξ_{x_2} et d'utiliser l'indépendance des $(\xi_x)_{x \in \mathbb{Z}^2}$ et ainsi de suite, ...

Dans le même ordre d'idée, on montre que

$$\lim_{n \rightarrow \infty} E \left(\left| h(\xi_{x_1}, \xi_{x_2}) \dots h(\xi_{x_k}, \xi_{x_1}) - \sum_{i_1=1}^n \lambda_{i_1} \phi_{i_1}(\xi_{x_1}) \phi_{i_1}(\xi_{x_2}) \dots \sum_{i_k=1}^n \lambda_{i_k} \phi_{i_k}(\xi_{x_k}) \phi_{i_k}(\xi_{x_1}) \right| \right) = 0$$

et ainsi, en utilisant les propriétés (33) et (34) des fonctions propres $(\phi_{\nu})_{\nu \geq 1}$ associées au noyau dégénéré h , il suit que

$$I_k = \sum_{\nu=1}^{\infty} \lambda_{\nu}^k.$$

□

Maintenant que l'on sait calculer l'espérance de h sur les cycles d'ordre k , pour tout $k \geq 2$ en fonction des valeurs propres $(\lambda_{\nu})_{\nu \geq 1}$ et que l'on connaît le comportement asymptotique des temps locaux d'auto-intersection, on montre que (cf. [18]),

- (i) par les Lemmes 2.1, 2.3 et 2.2 et 2.5, il existe une suite de réels $(\mu_k)_{k \geq 1}$, dont on calcule explicitement pour tout $k \geq 1$ le k -ième terme, μ_k , en fonction de $(\lambda_\nu)_{\nu \geq 1}$, de $(\theta_i)_{1 \leq i \leq m}$ et de $(t_i)_{1 \leq i \leq m}$, telle que pour tout $k \geq 1$,

$$\mu_k^{(n)} = E((\mathcal{W}_n)^k | \mathcal{A}) \rightarrow \mu_k \quad \text{presque sûrement, quand } n \rightarrow \infty,$$

- (ii) la fonction définie par

$$\phi(u) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} u^k,$$

possède un rayon de convergence strictement positif,

- (iii) en réappliquant la même méthode au processus $\sum_{j=1}^m \theta_j \Theta_n(t_j)$ (cf.(36)) défini par Neuhaus à partir d'une U -statistique classique, ϕ , grâce au calcul explicite des $(\mu_k)_{k \geq 1}$, n'est autre que la fonction génératrice de la variable aléatoire

$$\mathcal{W} = \sum_{j=1}^m \theta_j \sum_{i=1}^{\infty} \lambda_i ((B_{t_j}^{(i)})^2 - t_j) = \sum_{j=1}^m \theta_j U(t_j),$$

qui apparaît comme limite faible de $\sum_{j=1}^m \theta_j \Theta_n(t_j)$ (cf [52]).

– On traite enfin le cas où le noyau h n'est pas borné.

On construit une suite $(h_M)_{M \geq 1}$ de noyaux dégénérés, bornés, tendant en norme L_2 vers h , de la façon suivante :

$$h_M(x, y) = f_M(x, y) - Ef_M(x, \xi_1) - Ef_M(\xi_0, y) + Ef_M(\xi_0, \xi_1),$$

avec,

$$f_M(x, y) = h(x, y)_{\{|h(x, y)| \leq M\}}.$$

On note alors $\mathcal{W}_{n, M}$ l'équivalent, basé sur le noyau h_M , de la suite \mathcal{W}_n . Le point clé de la démonstration est le suivant : on montre que, quand M tend vers l'infini, $\mathcal{W}_{n, M} - \mathcal{W}_n$ converge en probabilité vers 0, uniformément par rapport à n . On conclut alors en utilisant le résultat sur les noyaux bornés (voir [18]).

• **Tension de $(\widehat{\mathcal{U}}_n)_{n \in \mathbb{N}}$:**

En ce qui concerne la tension de la suite $\widehat{\mathcal{U}}_n$, le schéma de la preuve de [18] consiste à montrer que pour tout $0 \leq t_1 < t < t_2 \leq 1$,

$$(45) \quad E((\widehat{\mathcal{U}}_{[nt_2]} - \widehat{\mathcal{U}}_{[nt]})^2 (\widehat{\mathcal{U}}_{[nt]} - \widehat{\mathcal{U}}_{[nt_1]})^2) \leq C(t_2 - t_1)^2.$$

L'idée étant, en fait, de calculer l'espérance conditionnelle

$$(46) \quad E((\widehat{\mathcal{U}}_{[nt_2]} - \widehat{\mathcal{U}}_{[nt]})^2 (\widehat{\mathcal{U}}_{[nt]} - \widehat{\mathcal{U}}_{[nt_1]})^2 | \mathcal{A}).$$

et de vérifier (45) en utilisant la dégénérescence de h ainsi que les résultats sur les temps locaux de la marche $(S_n)_{n \in \mathbb{N}}$ (Lemmes 2.1, 2.3 et 2.4). Pour finir, la tension de $(\widehat{\mathcal{U}}_n)_{n \in \mathbb{N}}$ découle d'une simple application de [11], Théorème 15.6.

□

Le résultat concernant le comportement limite fonctionnel de \mathcal{U}_n , nécessite une condition supplémentaire. Notons $\tilde{h} : \mathbb{R} \rightarrow \mathbb{R}$ la fonction définie par $\tilde{h}(x) = h(x, x)$ et supposons que

$$(47) \quad \tilde{h} \in L^4(\mu).$$

La distribution asymptotique de \mathcal{U}_n s'obtient alors à partir du Théorème 2.1.

Théorème 2.2. *La suite de processus $\{(\mathcal{U}_n(t))_{t \in [0,1]}, n = 1, 2, \dots\}$ converge faiblement, quand $n \rightarrow \infty$, dans $D[0, 1]$, vers le processus $(U(t) + E(\tilde{h}(\xi_x))t)_{t \in [0,1]}$, i.e.*

$$\left(\mathcal{U}_n(t) := \frac{\pi(\det \Sigma)^{1/2}}{n \log(n)} U_{[nt]} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(\sum_{\nu=1}^{\infty} \lambda_{\nu} \left((B_t^{(\nu)})^2 - t \right) + I_1 t \right)_{t \in [0,1]},$$

en posant $I_1 = E(\tilde{h}(\xi_x))$.

Si de plus, on suppose que $I_1 = \sum_{\nu=1}^{\infty} \lambda_{\nu}$, ce qui sous-entend que la série converge, alors

$$(\mathcal{U}_n(t))_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(\sum_{\nu=1}^{\infty} \lambda_{\nu} (B_t^{(\nu)})^2 \right)_{t \in [0,1]}, \quad n \rightarrow \infty.$$

Idée de la preuve [18] :

On définit $T_{[nt]}$, la composante de $U_{[nt]}$ associée aux auto-intersections de la marche, c'est-à-dire

$$T_{[nt]} = \sum_{i,j=0}^{[nt]} \mathbf{1}_{\{S_i=S_j\}} h(\xi_{S_i}, \xi_{S_j}) = \sum_{x \in \mathbb{Z}^2} N_{[nt]}^2(x) \tilde{h}(\xi_x).$$

La suite $U_{[nt]}$ s'écrit alors comme la somme de $\widehat{U}_{[nt]}$ et de $T_{[nt]}$,

$$U_{[nt]} = \widehat{U}_{[nt]} + T_{[nt]}.$$

Il est alors facile de voir que, conditionnellement à \mathcal{A} , $\sum_{j=1}^m \theta_j \frac{T_{[nt_j]}}{cn \log n}$ converge en probabilité vers $I_1 \sum_{j=1}^m \theta_j t_j$. Ce qui, combiné avec le Théorème 2.1, entraîne la convergence des lois de dimensions finies. La méthode de la preuve, pour la tension de \mathcal{U}_n est identique à celle utilisée précédemment. On obtient pour \mathcal{U}_n la même inégalité que (45).

Remarques :

- La condition de L_4 intégrabilité de h et de \tilde{h} , ne sert que pour la tension des suites $\widehat{\mathcal{U}}_n$ et \mathcal{U}_n . Seule une condition L_2 sur h et L_1 sur \tilde{h} est nécessaire pour prouver les convergences des lois de dimension finie.
- Sans rajouter de conditions supplémentaires sur h , il n'y a pas de raison, contrairement à ce qui se passe sur un cycle de longueur $k \geq 2$ (cf (2.5)), pour que $\tilde{h}(\xi_x) = h(\xi_x, \xi_x)$ soit intégrable et encore moins pour que $E(\tilde{h}(\xi_x))$ soit égal à la somme des valeurs propres $\sum_{\nu=1}^{\infty} \lambda_{\nu}$, au cas où cette quantité serait définie. En effet, le fait que h se décompose selon (32) dans l'espace $L^2(\mu \otimes \mu)$ ne nous donne à priori aucune indication sur l'intégrabilité de $\tilde{h}(\xi_x)$.

2.2. Cas non dégénéré.

Nous présentons une version du résultat de Cabus et Guillin [18], sur les noyaux non dégénérés et dont la preuve diffère légèrement par rapport à celle proposée dans [18]. On supposera toujours ici, $h \in L^4(\mu \otimes \mu)$

Notons pour tout $x \neq y$,

$$(48) \quad m = Eh(\xi_x, \xi_y), \quad g(\xi_x) = E(h(\xi_x, \xi_y)|\xi_x), \quad \text{et } \sigma^2 = \text{Var}(g(\xi_x)).$$

Alors \mathcal{U}_n , à une renormalisation en $1/n$ près, se comporte comme une marche aléatoire en scène aléatoire grâce à la H -décomposition de Hoeffding [39]. Le résultat de Bolthausen [13], rappelé au chapitre 2, Théorème 4.1, s'applique et l'on obtient les deux théorèmes suivants (cf [18]).

Théorème 2.3. *La suite de processus $\left\{ \left((2\sigma[nt](cn \log n)^{1/2})^{-1} (\widehat{U}_{[nt]} - m[nt]^2) \right)_{t \in [0,1]}, n = 1, 2, \dots \right\}$ converge faiblement, quand $n \rightarrow \infty$, dans $(D([0,1]), \mathcal{S})$ vers le mouvement brownien standard $(B_t)_{t \in [0,1]}$,*

$$\left(\frac{\widehat{U}_{[nt]} - m[nt]^2}{2\sigma[nt](cn \log n)^{1/2}} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (B_t)_{t \in [0,1]}, \quad n \rightarrow \infty,$$

où l'on rappelle que c est défini par $c = \frac{1}{\pi(\det \Sigma)^{1/2}}$.

Théorème 2.4. *Supposons que $\tilde{h} \in L^4(\mu)$. Alors, quand $n \rightarrow \infty$, la suite de processus $\left\{ \left((2\sigma[nt](cn \log n)^{1/2})^{-1} (U_{[nt]} - m[nt]^2) \right)_{t \in [0,1]}, n = 1, 2, \dots \right\}$ converge faiblement dans $(D([0,1]), \mathcal{S})$ vers le mouvement brownien standard $(B_t)_{t \in [0,1]}$,*

$$\left(\frac{U_{[nt]} - m[nt]^2}{2\sigma[nt](cn \log n)^{1/2}} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (B_t)_{t \in [0,1]}, \quad n \rightarrow \infty.$$

Preuve : Nous donnons la preuve du Théorème 2.4 qui contient celle du Théorème 2.3. En utilisant la décomposition de Hoeffding, pour tout $x \neq y$, posons

$$g(\xi_x) = E(h(\xi_x, \xi_y) | \xi_x)$$

et notons ϕ le noyau dégénéré de degré 2 tel que

$$h(\xi_x, \xi_y) = m + (g(\xi_x) - m) + (g(\xi_y) - m) + \phi(\xi_x, \xi_y).$$

En utilisant cette écriture, il vient

$$\begin{aligned} U_{[nt]} &= m([nt] + 1)^2 - mV_{[nt]} + \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m)) \\ &\quad + 2([nt] + 1) \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x) (g(\xi_x) - m) + \sum_{x \neq y} N_{[nt]}(x) N_{[nt]}(y) \phi(\xi_x, \xi_y) \\ &= m([nt] + 1)^2 - mV_{[nt]} + \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m)) \\ &\quad + 2([nt] + 1) \sum_{k=0}^{[nt]} (g(\xi_{S_k}) - m) + \sum_{x \neq y} N_{[nt]}(x) N_{[nt]}(y) \phi(\xi_x, \xi_y). \end{aligned}$$

On voit comme dans la preuve du Théorème 2.1 par dégénérescence de ϕ que

$$\begin{aligned} E \left(\left| \sum_{x \neq y} N_{[nt]}(x) N_{[nt]}(y) \phi(\xi_x, \xi_y) \right|^2 \right) &= \sum_{x \neq y} E \left(N_{[nt]}^2(x) N_{[nt]}^2(y) \right) E(\phi^2(\xi_x, \xi_y)) \\ &\leq \|\phi\|_2^2 E(V_n^2) \end{aligned}$$

Ainsi, par le Lemme 2.4

$$E\left(\left|\frac{1}{n(n \log n)^{1/2}} \sum_{x \neq y} N_{[nt]}(x) N_{[nt]}(y) \phi(\xi_x, \xi_y)\right|^2\right) = \mathcal{O}\left(\frac{\log n}{n}\right)$$

et

$$E\left(\frac{1}{n(n \log n)^{1/2}} \left| \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m)) \right|\right) = \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right).$$

En vertu des deux équations précédentes, quand n tend vers l'infini, le terme

$$(49) \frac{1}{n(n \log n)^{1/2}} \left[\sum_{x \neq y} N_{[nt]}(x) N_{[nt]}(y) \phi(\xi_x, \xi_y) + \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m)) \right]$$

converge en probabilité vers 0.

On peut enfin conclure grâce au théorème de Bolthausen [13] appliqué à la suite de sommes partielles,

$$\sum_{x \in \mathbb{Z}^2} N_{[nt]}(x) (g(\xi_x) - m) = \sum_{k=1}^{[nt]} (g(\xi_{S_k}) - m).$$

Remarque :

- Dans la preuve de la convergence en probabilité vers 0 de la quantité (49) ci-dessus, nous ne faisons pas intervenir ici le Théorème 2.1. Ce qui nous évite d'avoir éventuellement à faire une hypothèse supplémentaire sur la valeur de $E(\tilde{\phi})$.

3. Cas de la marche aléatoire transiente dans $\mathbb{Z}^d, d \geq 3$

Dans cette section, nous donnons une version des résultats de Cabus et Guillotin [18] dans le cas où $(S_n)_{n \in \mathbb{N}}$ est une marche aléatoire transiente évoluant dans $\mathbb{Z}^d, d \geq 3$. Les hypothèses concernant la scène $(\xi_x)_{x \in \mathbb{Z}^d}$ et le noyau h restent inchangées par rapport à la section précédente.

3.1. Cas dégénéré.

Nous nous plaçons ici, dans le cas où le noyau h est dégénéré et comme dans la section précédente, cas dégénéré, nous supposons que $h \in L^4(\mu \otimes \mu)$. En vertu de la transience de $(S_n)_{n \in \mathbb{N}}$, la constante $G(0, 0)$ donnée par

$$G(0, 0) = \sum_{k=0}^{\infty} P(S_k = 0),$$

est finie. Notons $\gamma = 2G(0, 0) - 1$. Alors, dans le cas où $(S_n)_{n \in \mathbb{N}}$ est une marche aléatoire transiente sur \mathbb{Z}^d , Cabus et Guillotin [18] ont prouvé :

Théorème 3.1. *Quand n tend vers l'infini, la suite de processus $\left\{ \left(\frac{1}{\gamma n} \widehat{U}_{[nt]} \right)_{t \in [0,1]}, n = 1, 2, \dots \right\}$ converge faiblement dans $(D[0, 1], \mathcal{S})$ vers $(U(t))_{t \in [0,1]}$, i.e.*

$$\left(\frac{1}{\gamma n} \widehat{U}_{[nt]} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(U(t) := \sum_{\nu=1}^{\infty} \lambda_{\nu} \left((B_t^{(\nu)})^2 - t \right) \right)_{t \in [0,1]}.$$

Théorème 3.2. [18] *Si $\tilde{h} \in L^4(\mu)$, la suite de processus $\left\{ \left(\frac{1}{\gamma n} U_{[nt]} \right)_{t \in [0,1]}, n = 1, 2, \dots \right\}$ converge faiblement, quand $n \rightarrow \infty$, dans $(D[0, 1], \mathcal{S})$ vers $(U(t) + I_1 t)_{t \in [0,1]}$, c.à.d.*

$$\left(\frac{1}{\gamma n} U_{[nt]} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (U(t) + I_1 t)_{t \in [0,1]}.$$

D'autre part, si, $I_1 = \sum_{\nu=1}^{\infty} \lambda_{\nu}$,

$$\left(\frac{1}{\gamma n} U_{[nt]} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(\sum_{\nu=1}^{\infty} \lambda_{\nu} (B_t^{(\nu)})^2 \right)_{t \in [0,1]}, \quad n \rightarrow \infty.$$

Idée des preuves [18] :

Dans ces deux cas de figure, les méthodes sont très similaires à celles déjà utilisées pour les théorèmes 2.1 et 2.2. En effet, le comportement asymptotique du temps local d'auto-intersection V_n de la marche est, ici encore, régi par une loi des grands nombres. Par une autre méthode que celle suggérée par Kesten et Spitzer [44] p.10 [44], Cabus et Guillotin montrent que

$$\frac{V_n}{n} \xrightarrow{\text{p.s.}} \gamma \quad \text{quand } n \rightarrow \infty.$$

Le reste des preuves s'appuie alors sur les équivalents, pour les marches transientes, des lemmes 2.2, 2.3 et 2.4 et suit exactement le plan de celles des Théorèmes 2.1 et 2.2.

3.2. Cas non dégénéré.

Ici, comme dans les sous-sections précédentes, on suppose que $h \in L^4(\mu \otimes \mu)$ et, en ce qui concerne le résultat portant sur le comportement asymptotique de $U_{[nt]}$, que $\tilde{h} \in L^4(\mu)$. De la même façon, on constate que grâce à la H -décomposition, quand h est non dégénéré, à une renormalisation près, \mathcal{U}_n se comporte comme une marche aléatoire en scène aléatoire dans le cas d'une marche transiente ([13]).

Théorème 3.3. *Quand n tend vers l'infini, la suite de processus $\left\{ \left(\frac{\widehat{U}_{[nt]} - m[nt]^2}{2\sigma[nt]\sqrt{\gamma n}} \right)_{t \in [0,1]}, n = 1, 2, \dots \right\}$ converge faiblement dans $D([0, 1], \mathbb{R})$, vers le mouvement brownien standard $(B_t)_{t \in [0,1]}$, i.e.*

$$\left(\frac{\widehat{U}_{[nt]} - m[nt]^2}{2\sigma[nt]\sqrt{\gamma n}} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (B_t)_{t \in [0,1]}, \quad \text{quand } n \rightarrow \infty,$$

où m et σ sont définis comme en (48).

Théorème 3.4. *Si de plus $\tilde{h} \in L^4(\mu)$, la suite de processus $\left\{ \left(\frac{U_{[nt]} - m[nt]^2}{2\sigma[nt]\sqrt{\gamma n}} \right)_{t \in [0,1]}, n = 1, 2, \dots \right\}$ converge faiblement dans $D([0, 1], \mathbb{R})$, vers le mouvement brownien standard $(B_t)_{t \in [0,1]}$, i.e.*

$$\left(\frac{U_{[nt]} - m[nt]^2}{2\sigma[nt]\sqrt{\gamma n}} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (B_t)_{t \in [0,1]}, \quad \text{quand } n \rightarrow \infty.$$

4. Commentaires et Conjecture

Commentaires :

- En ce qui concerne la suite de processus $U_{[nt]}$, il est à noter l'importance de la condition sur la fonction \tilde{h} : $\tilde{h} \in L^1(\mu)$. En effet le résultat de convergence faible de $U_{[nt]}$ dépend directement, à travers la valeur de $E(\tilde{h}(\xi_x))$, de \tilde{h} , qui à priori comme nous l'avons déjà souligné, n'a aucune raison d'être dans $L^1(\mu)$, si seule la condition $h \in L^4(\mu \otimes \mu)$ (et donc $h \in L^2(\mu \otimes \mu)$) est spécifiée. A l'inverse, la suite $\widehat{U}_{[nt]}$,

$$\widehat{U}_{[nt]} = \sum_{0 \leq i, j \leq [nt]} h(\xi_{S_i}, \xi_{S_j}) \mathbf{1}_{\{S_i \neq S_j\}},$$

qui est définie en dehors des auto-intersections de la marche ($\{S_i \neq S_j\}$) et qui ne nécessite donc aucune condition supplémentaire sur \tilde{h} , apparaît comme le candidat idéal à la généralisation de la suite de processus associée aux U -statistiques classiques, $\Theta_n(t)$, qui, on le rappelle, est définie sur l'ensemble $\{1 \leq i \neq j \leq [nt]\}$:

$$\Theta_n(t) = \frac{1}{n} \sum_{\substack{i \neq j \\ i, j \in [nt]}} h(\xi_i, \xi_j).$$

- Dans le cas où la marche aléatoire $(S_n)_{n \in \mathbb{N}}$ est à valeurs dans \mathbb{Z} et vérifie les hypothèses de Kesten et Spitzer [44] (cf chapitre 2), les techniques de preuves, utilisées tout au long de ce chapitre, basées essentiellement sur la convergence presque-sûre des temps locaux d'auto-intersections V_n , ne nous permettent pas de conclure. En effet, dans ce dernier cas, les résultats concernant les temps locaux de la marche sont beaucoup moins forts puisqu'on ne dispose plus, pour les $(V_n)_{n \in \mathbb{N}}$, que d'une convergence en loi.

Conjecture de Cabus et Guillotin [18]

Bien qu'il ne soit pas possible d'adapter directement leur preuve dans le cas de la marche simple symétrique sur \mathbb{Z} , Cabus et Guillotin [18] conjecturent que dans ce contexte, sous des hypothèses appropriées à préciser,

$$n^{-\frac{3}{2}} \sum_{i, j=0}^{[nt]} h(\xi_{S_i}, \xi_{S_j})$$

converge faiblement, dans l'espace de Skorohod, vers le processus nondégénéré suivant :

$$W_t = \sum_{\nu=1}^{+\infty} \lambda_\nu (\Delta_t^{(\nu)})^2$$

où les $\Delta_t^{(\nu)}$, $\nu \geq 1$ sont des *mouvements browniens en scène aléatoire brownienne* (cf. Kesten et Spitzer [44]) définis ainsi :

$$\Delta_t^{(\nu)} = \int_{-\infty}^{+\infty} L_t(x) dB^{(\nu)}(x),$$

où $L_t(x)$ désigne le temps local au point $x \in \mathbb{R}$ d'un mouvement brownien $(W_t)_{t \geq 0}$ et où, $\{(B^{(\nu)}(t))_{t \in \mathbb{R}}, \nu = 1, 2, \dots\}$ désigne une suite de mouvements browniens bilatères, indépendants de $(W_t)_{t \geq 0}$.

CHAPITRE 4

U-statistiques indexées par une marche aléatoire récurrente sur \mathbb{Z}

Nous nous intéressons ici au cas des U -statistiques, basées sur un noyau de degré deux, indexées par une marche aléatoire $(S_n)_{n \in \mathbb{N}}$, récurrente, à valeurs dans \mathbb{Z} . Nous donnons deux théorèmes limite fonctionnels décrivant le comportement asymptotique des processus associés à ces U -statistiques échantillonnées et nous prouvons la conjecture de Cabus et Guillin [18].

1. Description du Modèle

Nous supposons tout-au-long de de cette partie, que toutes les variables aléatoires sont définies sur un même espace de probabilité. Notre problème concerne le comportement asymptotique, quand $n \rightarrow +\infty$, des sommes partielles de la forme :

$$U_n = \sum_{i,j=0}^n h(\xi_{S_i}, \xi_{S_j}), n \in \mathbb{N},$$

dans le cas où le noyau $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ désigne comme au chapitre précédent, une fonction mesurable, symétrique, telle que $h \in L^2(\mu \otimes \mu)$ et où la marche aléatoire $(S_n)_{n \in \mathbb{N}}$ ainsi que la scène aléatoire $(\xi_x)_{x \in \mathbb{Z}}$ vérifient les mêmes que hypothèses que dans l'article de Kesten et Spitzer [44]. C'est-à-dire que :

- la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, qui commence en 0, est associée par la relation

$$S_n = X_1 + \dots + X_n,$$

à une suite de variables aléatoires indépendantes identiquement distribuées à valeurs dans \mathbb{Z} , $(X_i)_{i \geq 1}$, qui appartiennent au domaine d'attraction de Z_α , une variable aléatoire de loi stable d'indice $1 < \alpha \leq 2$:

$$\frac{1}{n^{1/\alpha}} S_n \xrightarrow{\mathcal{D}} Z_\alpha.$$

- la scène aléatoire $(\xi_x)_{x \in \mathbb{Z}}$ désigne une suite de variables aléatoires réelles indépendantes, de même loi de probabilité μ , indépendante de la marche aléatoire $(S_n)_{n \in \mathbb{N}}$.

2. Résultats

Nous nous plaçons d'abord, dans le cadre d'un noyau h dégénéré, c'est-à-dire (cf chap. 2)

$$E(h(\xi_0, x)) = 0, \quad \text{pour } (\mu)\text{-presque tout } x \in \mathbb{R}.$$

Comme nous l'avons déjà vu dans les chapitres précédents, si l'on note T_h l'opérateur intégral défini à partir du noyau h (cf 106) et que l'on note $(\lambda_\nu)_{\nu \geq 1}$ ses valeurs propres, il existe dans $L^2(\mu)$, une suite orthonormale $(\phi_\nu)_{\nu \geq 1}$ de fonctions propres de T_h telles que

$$h(x, y) = \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(x) \phi_\nu(y), \quad \text{dans } L^2(\mu \otimes \mu).$$

Les $(\phi_\nu)_{\nu \geq 1}$, en vertu de la dégénérescence du noyau et de leur $L^2(\mu)$ -orthonormalité, vérifient pour tout $\nu, \sigma \geq 1$,

$$(50) \quad E(\phi_\nu(\xi_0)) = 0, \quad E(\phi_\nu(\xi_0)\phi_\sigma(\xi_0)) = \delta_{\nu,\sigma} \quad \text{où } \delta_{\nu,\sigma} = \begin{cases} 1 & \text{si } \nu = \sigma \\ 0 & \text{si } \nu \neq \sigma \end{cases}$$

En ce qui concerne h , dans le cas dégénéré, nous ferons l'hypothèse suivante :

$$h \in L^4(\mu \otimes \mu).$$

Cette hypothèse et le fait que les $(\phi_\nu)_{\nu \geq 1}$ soient des fonctions propres de T_h associées à des valeurs propres non nulles, entraîne, par simple application de l'inégalité de Cauchy-Schwartz que pour tout entier $\nu \geq 1$:

$$(51) \quad E((\phi_\nu(\xi_0))^4) < \infty.$$

Par ailleurs, en notant comme au chapitre précédent $\tilde{h} : \mathbb{R} \rightarrow \mathbb{R}$, la restriction de h à la droite $\{(x, y) \in \mathbb{R}^2 \mid y = x\}$, i.e. $\tilde{h}(x) = h(x, x)$, nous imposerons les conditions supplémentaires suivantes:

$$\tilde{h} \in L^4(\mu)$$

et

$$(52) \quad \tilde{h}(x) = \sum_{\nu=1}^{+\infty} \lambda_\nu \phi_\nu^2(x) \quad \text{dans } L^1(\mu).$$

On appellera δ_0 , la même constante que celle définie au chapitre 1, i.e.

$$\delta_0 = 1 - \frac{1}{2\alpha}.$$

Soit $(Y(t))_{t \geq 0}$ un processus de Lévy α -stable à trajectoires continues à droite, tel que $Y(1)$ ait la même loi que Z_α et soient $(B_+^{(\nu)}(t))_{t \geq 0}$, $(B_-^{(\nu)}(t))_{t \geq 0}$, $\nu = 1, 2, \dots$ une suite de paires de deux mouvements browniens indépendants, deux-à-deux indépendants et indépendants de $(Y(t))_{t \geq 0}$. Nous pouvons ainsi définir pour tout $\nu = 1, 2, \dots$, comme dans [44] (cf. chapitre 2), les processus :

$$\Delta_t^{(\nu)} = \int_0^\infty L_t(x) dB_+^{(\nu)}(x) + \int_0^\infty L_t(-x) dB_-^{(\nu)}(x),$$

encore notés

$$(53) \quad \Delta_t^{(\nu)} = \int_{\mathbb{R}} L_t(x) dB^{(\nu)}(x).$$

Par ailleurs on désignera par $(\Delta_t)_{t \in [0,1]}$ un processus égal en loi à l'un de ces processus.

Posons pour tout $t \in [0, 1]$ et tout $n \in \mathbb{N}$:

$$\mathcal{U}_n(t) := \frac{1}{n^{2\delta_0}} U_{[nt]}.$$

Nous définirons également le processus $(W_t)_{t \in [0,1]}$ tel que

$$(54) \quad W_t = \sum_{\nu=1}^{\infty} \lambda_\nu (\Delta_t^{(\nu)})^2.$$

En ce qui concerne la distribution asymptotique de $(\mathcal{U}_n(t))_{t \in [0,1]}$, quand n tend vers l'infini, dans le cas dégénéré, notre résultat est le suivant :

Théorème 2.1. *Dans le cas où le noyau h est dégénéré et sous les hypothèses précédentes, la suite de processus $\{(\mathcal{U}_n(t))_{t \in [0,1]}, n = 1, 2, \dots\}$ converge faiblement quand n tend vers l'infini, dans $(D[0, 1], \mathcal{S})$ vers le processus $(W_t)_{t \in [0,1]}$:*

$$\left(\mathcal{U}_n(t) := \frac{1}{n^{2\delta_0}} U_{[nt]} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (W_t)_{t \in [0,1]}.$$

Dans le cas d'un noyau h non dégénéré, nous supposons seulement que $h \in L^4(\mu \otimes \mu)$ et $\tilde{h} \in L^4(\mu)$, sans l'hypothèse supplémentaire (52) sur la décomposition de \tilde{h} dans $L^1(\mu)$. Notre second résultat est le suivant :

Théorème 2.2. *Soit h un noyau non dégénéré. Avec la notation $m = E(h(\xi_0, \xi_1))$ et $\sigma^2 = E(h(\xi_0, \xi_1)h(\xi_0, \xi_{-1})) - m^2$, la suite de processus $\{(U_{[nt]} - m[nt]^2)/(2t\sigma n^{2-1/2\alpha}), = 1, 2, \dots\}$ converge faiblement, quand n tend vers l'infini, dans $(D[0, 1], \mathcal{S})$ vers le processus $(\Delta_t)_{t \in [0,1]}$,*

$$\left(\frac{U_{[nt]} - m[nt]^2}{2t\sigma n^{2-1/2\alpha}} \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (\Delta_t)_{t \in [0,1]}.$$

Remarques :

- Comme pour le Théorème de Cabus et Guillin [18] (cf. Théorème 2.2, chapitre 3), les conditions de L^4 -intégrabilité de h et \tilde{h} ne servent que dans la preuve de la tension de la suite $\mathcal{U}_n(t)$. Pour la convergence des lois de dimension finie, seule la L_2 -intégrabilité de h et l'hypothèse (52) concernant la décomposition de \tilde{h} dans $L^1(\mu)$ sont nécessaires.
- Nous verrons par contre que cette dernière condition (52), qui n'était pas nécessaire dans le cas d'une marche aléatoire récurrente sur \mathbb{Z}^2 , est importante dans la technique de preuve utilisée ici. En effet, les résultats concernant la convergence des temps locaux d'auto-intersection de la marche aléatoire $(S_n)_{n \in \mathbb{N}}$ étant relativement faibles (convergence faible), nous ne pouvons pas appliquer la technique des preuves des théorèmes 2.1 et 2.2. Dans la démonstration que nous donnons, nous avons eu besoin de renforcer les conditions sur \tilde{h} .
- On sait (et c'est toujours vrai (cf. (31), chapitre 2)) que les valeurs propres $(\lambda_\nu)_{\nu \geq 1}$ d'un noyau $h \in L^2(\mu \otimes \mu)$ sont dans $\ell^2(\mathbb{N}^*)$ et vérifient

$$\sum_{\nu=1}^{\infty} \lambda_\nu^2 = E(h^2(\xi_1, \xi_2)).$$

Ici, l'hypothèse (52) implique de plus que

$$\left| \sum_{\nu=1}^{\infty} \lambda_\nu \right| < +\infty, \text{ avec } \sum_{\nu=1}^{\infty} \lambda_\nu = E(\tilde{h}(\xi_x)).$$

3. Propriétés des temps d'occupation de la marche aléatoire

Afin de prouver le Théorème 2.1, nous aurons besoin de résultats préliminaires concernant le comportement asymptotique des temps d'occupation $N_n(x)$ et des temps locaux d'auto-intersections V_n de la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, où, comme dans les chapitres précédents,

$N_n(x)$ désigne le nombre de visites de la marche aléatoire au point x dans l'intervalle de temps $[0, n]$, i.e.

$$N_n(x) = \sum_{i=0}^n 1_{\{S_i=x\}}$$

et où V_n désigne le nombre d'auto-intersections de $(S_n)_{n \in \mathbb{N}}$ jusqu'au temps n ,

$$V_n = \sum_{x \in \mathbb{Z}} N_n(x)^2.$$

Nous noterons ici, ainsi que tout au long de ce chapitre,

$$Q_n^{(p)} = \sum_{x \in \mathbb{Z}} N_n(x)^p.$$

D'après [44] (cf Lemme 3.2, chapitre 1), nous savons déjà que :
pour tout $x \in \mathbb{Z}$,

$$(55) \quad E(N_n(x)^p) = \mathcal{O}(n^{p(1-\frac{1}{\alpha})}), \quad \text{quand } n \rightarrow \infty,$$

et que pour une certaine constante $C > 0$,

$$(56) \quad E(V_n) \sim Cn^{2-\frac{1}{\alpha}}, \quad \text{quand } n \rightarrow +\infty.$$

Nous allons maintenant prouver le résultat suivant,

Lemme 3.1. *Pour tout $p \geq 1$ et tout $k \geq 1$,*

$$E(|Q_n^{(p)}|^k) = \mathcal{O}(n^{kp(1-\frac{1}{\alpha})+\frac{k}{\alpha}}), \quad \text{quand } n \rightarrow +\infty.$$

En particulier ($p = 2$), pour tout $k \geq 1$,

$$E(|V_n|^k) = \mathcal{O}(n^{2k\delta_0}) \quad \text{quand } n \rightarrow +\infty.$$

Preuve : Soit (Ω, Σ, P) un espace de probabilité sur lequel toutes les variables aléatoires sont définies. On notera $\|\cdot\|_k$ la norme définie sur l'espace $L^k(P)$ par

$$\|X\|_k = E(|X|^k)^{\frac{1}{k}},$$

pour tout X dans $L^k(P)$.

On voit que

$$\begin{aligned} Q_n^{(p)} &= \sum_{0 \leq i_1, i_2, \dots, i_p \leq n} 1_{\{S_{i_1} = S_{i_2} = \dots = S_{i_p}\}} \\ &\leq p! \sum_{0 \leq i_1 \leq i_2 \leq \dots \leq i_p \leq n} 1_{\{S_{i_1} = S_{i_2} = \dots = S_{i_p}\}}. \end{aligned}$$

Ainsi en utilisant le fait que $\|\cdot\|_k$ est une norme, il suit

$$(57) \quad \|Q_n^{(p)}\|_k \leq p! \sum_{0 \leq i_1 \leq n} \left\| \sum_{i_1 \leq i_2 \leq \dots \leq i_p \leq n} 1_{\{S_{i_1} = S_{i_2} = \dots = S_{i_p}\}} \right\|_k.$$

Maintenant, à i_1 fixé, on peut écrire

$$E\left(\left(\sum_{i_1 \leq i_2 \leq \dots \leq i_p \leq n} 1_{\{S_{i_1} = S_{i_2} = \dots = S_{i_p}\}}\right)^k\right) \leq E\left(\sum_{i_1 \leq i_2, \dots, i_{k(p-1)+1} \leq n} 1_{\{S_{i_1} = S_{i_2} = \dots = S_{i_{k(p-1)+1}}\}}\right)$$

Ici la propriété de Markov s'applique, ce qui conduit à

$$E\left(\left(\sum_{i_1 \leq i_2, \dots, \leq i_p \leq n} 1_{\{S_{i_1} = S_{i_2} = \dots = S_{i_p}\}}\right)^k\right) \leq E\left(N_n(0)^{k(p-1)}\right).$$

Il suit alors de cette inégalité et de (55) que

$$(58) \quad \left\| \sum_{i_1 \leq i_2 \leq \dots \leq i_p \leq n} 1_{\{S_{i_1} = S_{i_2} = \dots = S_{i_p}\}} \right\|_k = \mathcal{O}\left(n^{(p-1)(1-\frac{1}{\alpha})}\right).$$

On peut finalement conclure la preuve du Lemme en combinant (57) avec (58).

□

4. Preuve du Théorème 2.1

Nous allons d'abord montrer la convergence des lois de dimension finie puis, nous prouverons la tension de la suite de processus $(\mathcal{U}_n(t))_{t \in [0,1]}$.

• **Convergence des lois de dimension finie.**

Proposition 4.1. *Quand n tend vers l'infini, les lois de dimension finie de $(n^{-2\delta_0}U_{[nt]})_{t \in [0,1]}$ convergent vers les lois de dimension finie de $(W_t)_{t \in [0,1]}$.*

Preuve : Fixons $\theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^m$ et $0 \leq t_1 < \dots < t_m \leq 1$.

Nous allons d'abord commencer par choisir un entier arbitraire, $K \geq 1$. Pour cet entier K , nous définirons $h_K(x, y)$, le noyau défini par

$$(59) \quad h_K(x, y) = \sum_{\nu=1}^K \lambda_\nu \phi_\nu(x) \phi_\nu(y).$$

Nous noterons $U_{n,K}$ et $\mathcal{U}_{n,K}$ les analogues de U_n et \mathcal{U}_n , associés à h_K , i.e.

$$U_{n,K} = \sum_{i,j=0}^n h_K(\xi_{S_i}, \xi_{S_j}),$$

et

$$\mathcal{U}_{n,K}(t) = \frac{1}{n^{2\delta_0}} U_{[nt],K}.$$

La preuve de la Proposition 4.1 se décompose en trois étapes :

Étape 1 : En utilisant la définition de h_K (59), on voit que $\mathcal{U}_{n,K}$ peut s'écrire :

$$(60) \quad \mathcal{U}_{n,K}(t) = \sum_{\nu=1}^K \lambda_\nu \left(\frac{1}{n^{\delta_0}} \sum_{i=1}^{[nt]} \phi_\nu(\xi_{S_i}) \right)^2.$$

Pour tout $x \in \mathbb{Z}$, on définit le vecteur aléatoire $\tilde{\xi}_x^{(K)}$ par

$$\tilde{\xi}_x^{(K)} = (\phi_1(\xi_x), \dots, \phi_K(\xi_x)) \in \mathbb{R}^K.$$

Notons que pour tout $\nu \in \{1, \dots, K\}$, la suite $(\phi_\nu(\xi_x))_{x \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes identiquement distribuées. Par ailleurs, il découle de (50) que, pour tout $x \in \mathbb{Z}$, les composantes de $\tilde{\xi}_x^{(K)}$ sont des variables aléatoires décorréllées, d'espérance nulle et de variance 1. On peut donc appliquer le théorème central limite, i.e.

$$\frac{1}{\sqrt{n}} \sum_{x=1}^n \tilde{\xi}_x^{(K)} \xrightarrow{\mathcal{L}} Z_{1/2},$$

où $Z_{1/2}$ est un vecteur gaussien standard K -dimensionnel.

On définit pour tout $n \geq 1$,

$$\mathcal{Z}_{n,K}(t) = \frac{1}{n^{\delta_0}} \sum_{i=0}^{[nt]} \tilde{\xi}_{S_i}^{(K)}$$

Nous allons pouvoir appliquer ici, le Théorème de Maejima [48] (cf. Théorème 3.2 du chapitre 1) qui concerne le cas gaussien, i.e. quand $n \rightarrow +\infty$,

$$(61) \quad (\mathcal{Z}_{n,K}(t))_{t \in [0,1]} \text{ converge faiblement dans } D[0,1], \text{ vers } (\Delta_t^{(1)}, \dots, \Delta_t^{(K)})_{t \in [0,1]},$$

où les $\Delta_t^{(\nu)}$, $\nu = 1, \dots, K$ sont les processus définis dans la première section (cf. (53)).

Nous poserons ici,

$$\begin{aligned} L : \mathbb{R}^K &\rightarrow \mathbb{R} \\ (x_1, \dots, x_K) &\mapsto \sum_{\nu=1}^K \lambda_\nu x_\nu^2. \end{aligned}$$

Alors, $\mathcal{U}_{n,K}$ s'écrit,

$$(62) \quad \mathcal{U}_{n,K} = L(\mathcal{Z}_{n,K}).$$

Ainsi comme L est une fonction continue, il suit que, quand n tend vers l'infini :

$$(\mathcal{U}_{n,K}(t))_{t \in [0,1]} \text{ converge dans } (D[0,1], \mathcal{S}) \text{ vers } (W_K(t) := \sum_{\nu=1}^K \lambda_\nu (\Delta_t^{(\nu)})^2)_{t \in [0,1]}.$$

Par conséquent, quand $n \rightarrow \infty$, les lois de dimension finie de $(\mathcal{U}_{n,K}(t))_{t \in [0,1]}$ convergent vers celles de $(W_K(t))_{t \in [0,1]}$. En notant ϕ_K et $\phi_{n,K}$ les fonctions caractéristiques respectives de $(W_K(t_1), \dots, W_K(t_m))$ et $(\mathcal{U}_{n,K}(t_1), \dots, \mathcal{U}_{n,K}(t_m))$, nous voyons que, K et θ étant fixés :

$$(63) \quad |\phi_{n,K}(\theta) - \phi_K(\theta)| \longrightarrow 0, \quad \text{quand } n \rightarrow \infty.$$

Étape 2 : Nous allons maintenant nous intéresser à la suite $R_{n,K}$ définie par

$$\begin{aligned} R_{[nt],K} &= U_{[nt]} - U_{[nt],K} \\ &= \sum_{x,y \in \mathbb{Z}} N_{[nt]}(x) N_{[nt]}(y) (h - h_K)(\xi_x, \xi_y) \\ &= R_{[nt],K}^{(1)} + R_{[nt],K}^{(2)} \end{aligned}$$

où

$$R_{[nt],K}^{(1)} = \sum_{x \neq y} N_{[nt]}(x) N_{[nt]}(y) (h - h_K)(\xi_x, \xi_y)$$

et

$$R_{[nt],K}^{(2)} = \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) (h - h_K)(\xi_x, \xi_x).$$

En notant $\tilde{h}_K(x) = h_K(x, x)$, on peut réécrire $R_{[nt],K}^{(2)}$ sous la forme

$$R_{[nt],K}^{(2)} = \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) (\tilde{h} - \tilde{h}_K)(\xi_x).$$

Nous allons dans un premier temps, nous intéresser à $E\left(|R_{[nt],K}^{(1)}|^2\right)$. En utilisant l'indépendance de $(S_n)_{n \geq 0}$ et $(\xi_x)_{x \in \mathbb{Z}}$ ainsi que la dégénérescence de h , on voit que

$$\begin{aligned} E\left(|R_{[nt],K}^{(1)}|^2\right) &= \sum_{x \neq y} E\left(N_{[nt]}^2(x) N_{[nt]}^2(y)\right) \|h - h_K\|_{L^2(\mu \otimes \mu)}^2 \\ (64) \qquad \qquad \qquad &\leq E(V_{[nt]}^2) \|h - h_K\|_{L^2(\mu \otimes \mu)}^2 \end{aligned}$$

Par conséquent, il découle du Lemme 3.1, l'existence d'une constante $C_1 > 0$ telle que pour tout $n \geq 1$ et $t \in [0, 1]$,

$$(65) \qquad E\left(\left|\frac{R_{[nt],K}^{(1)}}{n^{2\delta_0}}\right|^2\right) \leq C_1 \|h - h_K\|_{L^2(\mu \otimes \mu)}^2.$$

Concentrons-nous maintenant sur $R_{n,K}^{(2)}$.

$$E\left(|R_{[nt],K}^{(2)}|\right) \leq E\left(\sum_x N_{[nt]}^2(x) |(\tilde{h} - \tilde{h}_K)(\xi_x)|\right).$$

L'indépendance de $(S_n)_{n \geq 0}$ et $(\xi_x)_{x \in \mathbb{Z}}$ implique que

$$E(|R_{[nt],K}^{(2)}(t)|) \leq E(V_{[nt]}) \|\tilde{h} - \tilde{h}_K\|_{L^1(\mu)}.$$

Enfin, par le Lemme 3.1, il existe une constante $C_2 > 0$ telle que pour tout $n \geq 1$ et $t \in [0, 1]$,

$$(66) \qquad E\left(\frac{|R_{[nt],K}^{(2)}|}{n^{2\delta_0}}\right) \leq C_2 \|\tilde{h} - \tilde{h}_K\|_{L^1(\mu)}.$$

Posons alors, $C = \max(C_1^{1/2}, C_2)$ et

$$\mathcal{R}_{n,K}(t) = \frac{1}{n^{2\delta_0}} R_{[nt],K}.$$

En vertu de (65) et (66), il suit que pour tout $n \geq 1$ et tout $t \in [0, 1]$,

$$(67) \qquad E\left(|\mathcal{R}_{n,K}(t)|\right) \leq C \left(\|h - h_K\|_{L^2(\mu \otimes \mu)} + \|\tilde{h} - \tilde{h}_K\|_{L^1(\mu)}\right).$$

En posant

$$\Theta_n = \sum_{j=1}^m \theta_j \mathcal{U}_n(t_j), \quad \Theta_{n,K} = \sum_{j=1}^m \theta_j \mathcal{U}_{n,K}(t_j),$$

et en notant, pour tout entier n , ϕ_n la fonction caractéristique de $(\mathcal{U}_n(t_1), \dots, \mathcal{U}_n(t_m))$, on voit alors que

$$\begin{aligned} |\phi_n(\theta) - \phi_{n,K}(\theta)| &= |E(e^{i\Theta_n} - e^{i\Theta_{n,K}})| \\ &\leq E|e^{i(\Theta_n - \Theta_{n,K})} - 1| \\ &\leq E|\Theta_n - \Theta_{n,K}| \\ &= E\left|\sum_{j=1}^m \theta_j \mathcal{R}_{n,K}(t_j)\right|. \end{aligned}$$

Ainsi, il vient, par (67), que pour tout $n \geq 1$,

$$|\phi_n(\theta) - \phi_{n,K}(\theta)| \leq C \sum_{j=1}^m |\theta_j| \left(\|h - h_K\|_{L^2(\mu \otimes \mu)} + \|\tilde{h} - \tilde{h}_K\|_{L^1(\mu)} \right).$$

Par conséquent, pour tout $\varepsilon > 0$ il existe un $K_0 > 0$ tel que pour tout $K > K_0$,

$$(68) \quad \sup_{n \geq 1} |\phi_n(\theta) - \phi_{n,K}(\theta)| < \varepsilon.$$

Étape 3 : Nous noterons ϕ la fonction caractéristique du vecteur aléatoire $(W_{t_1}, \dots, W_{t_m})$.

Nous allons d'abord montrer que $E(\Delta_1^4) < \infty$. On commence par appliquer l'inégalité de Burkholder-Davis-Gundy (cf Revuz et Yor [59]) à la composante $E((\int_0^{+\infty} L_1(x) dB_+(x))^4)$:

$$\begin{aligned} E\left(\left(\int_0^{+\infty} L_1(x) dB_+(x)\right)^4\right) &\leq KE\left(\left\langle \int_0^{+\infty} L_1(x) dB_+(x), \int_0^{+\infty} L_1(x) dB_+(x) \right\rangle^{4/2}\right) \\ &\leq KE\left(\left(\int_0^{+\infty} L_1^2(x) dx\right)^2\right), \end{aligned}$$

pour une certaine constante $K > 0$.

Nous allons maintenant voir que $E\left(\left(\int_{\mathbb{R}} L_1^2(x) dx\right)^2\right) < \infty$.

D'après Kesten et Spitzer [44], la suite $\frac{V_n}{n^{2\delta_0}}$ converge faiblement vers $\int_{\mathbb{R}} L_1(x)^2 dx$, et

donc, par conséquent, $(\frac{V_n}{n^{2\delta_0}})^2$ converge faiblement vers $(\int_{\mathbb{R}} L_1(x)^2 dx)^2$.

$$(69) \quad \left(\frac{V_n}{n^{2\delta_0}}\right)^2 \xrightarrow{\mathcal{D}} \left(\int_{\mathbb{R}} L_1^2(x) dx\right)^2.$$

D'autre part, d'après le Lemme 3.1, il existe une constante $C > 0$ telle que pour tout $n \geq 1$,

$$E\left(\left(\frac{V_n}{n^{2\delta_0}}\right)^4\right) \leq C.$$

Ainsi, la suite de variables aléatoires $(V_n^2/n^{4\delta_0})_{n \geq 1}$ est uniformément intégrable. On peut donc appliquer le Théorème 25.12 de Billingsley [10] :

$$E\left(\left(\int_{\mathbb{R}} L_1^2(x) dx\right)^2\right) \leq C.$$

Par l'auto-similarité de Δ_t ($\Delta_t \stackrel{d}{=} t^{\delta_0} \Delta_1$), on a aussi, pour tout $t \in [0, 1]$:

$$E\left(\left(\int_0^{+\infty} L_1(x) dB_+(x)\right)^4\right) \leq C < \infty.$$

On procède ensuite de la même manière avec la composante $E\left(\left(\int_0^{+\infty} L_1(x) dB_-(x)\right)^4\right)$.

Ainsi,

$$E(\Delta_t^4) \leq C < \infty,$$

pour une certaine constante C .

Revenons à la preuve de la convergence des lois de dimension finie de $W_{K,t}$ vers celles de W_t .

$$(70) \quad \left\| \sum_{j=1}^m \theta_j (W_{t_j} - W_{K,t_j}) \right\|_2 \leq \sum_{j=1}^m |\theta_j| \cdot \left\| W_{t_j} - W_{K,t_j} \right\|_2$$

De Cauchy-Schwartz et du fait que pour tout $\nu \neq \mu$, $(\Delta_t^{(\nu)}, \Delta_t^{(\mu)})_{t \geq 0}$ et $(\Delta_t^{(1)}, \Delta_t^{(2)})_{t \geq 0}$ sont égaux en lois, il découle :

$$\begin{aligned}
\|W_{t_j} - W_{K,t_j}\|_2^2 &= \sum_{\nu=K+1}^{\infty} \lambda_\nu^2 E((\Delta_{t_j}^{(\nu)})^4) + \sum_{\nu \neq \mu=K+1}^{\infty} \lambda_\nu \lambda_\mu E((\Delta_{t_j}^{(\nu)})^2 (\Delta_{t_j}^{(\mu)})^2) \\
&\leq t_j^{4\delta_0} E((\Delta_1)^4) \left(\sum_{\nu=K+1}^{\infty} \lambda_\nu^2 + \left(\sum_{\nu=K+1}^{\infty} \lambda_\nu \right)^2 \right) \\
&\leq C \left(\sum_{\nu=K+1}^{\infty} \lambda_\nu^2 + \left(\sum_{\nu=K+1}^{\infty} \lambda_\nu \right)^2 \right).
\end{aligned}$$

La convergence des séries $\sum_{\nu=1}^{\infty} \lambda_\nu^2$ et $\sum_{\nu=1}^{\infty} \lambda_\nu$ nous permet de conclure (cf (70)) que $\sum_{j=1}^m \theta_j W_{K,t_j}$ converge dans $L^2(P)$ vers $\sum_{j=1}^m \theta_j W_{t_j}$ et donc, que $\sum_{j=1}^m \theta_j W_{K,t_j}$ converge en loi vers $\sum_{j=1}^m \theta_j W_{t_j}$. Ainsi,

$$(71) \quad |\phi(\theta) - \phi_K(\theta)| \rightarrow 0, \quad K \rightarrow +\infty.$$

Finalement, il découle de (68) et (71), qu'étant donné $\varepsilon > 0$, on peut trouver un entier K_1 suffisamment grand pour qu'on ait simultanément

$$\sup_{n \geq 1} |\phi_n(\theta) - \phi_{n,K_1}(\theta)| < \varepsilon/3 \quad \text{et} \quad |\phi(\theta) - \phi_{K_1}(\theta)| < \varepsilon/3.$$

Par ailleurs, K_1 étant fixé, d'après (63), on sait qu'il existe un entier N tel que pour tout $n > N$

$$|\phi_{K_1}(\theta) - \phi_{n,K_1}(\theta)| < \varepsilon/3.$$

Ainsi, pour tout $n > N$,

$$|\phi(\theta) - \phi_n(\theta)| \leq |\phi(\theta) - \phi_{K_1}(\theta)| + |\phi_{K_1}(\theta) - \phi_{n,K_1}(\theta)| + |\phi_n(\theta) - \phi_{n,K_1}(\theta)| < \varepsilon,$$

ce qui termine la preuve de la convergence des lois de dimension finie. □

- **Tension de la famille** $(\mathcal{U}_n(t))_{t \in [0,1]}$.

Par le Théorème 15.6 de [11], afin de montrer la tension de $(\mathcal{U}_n(t))_{t \in [0,1]}$, il nous suffira de prouver le résultat suivant :

Proposition 4.2. *Il existe une constante strictement positive K et une constante $\gamma = 4\delta_0 > 1$, telles que pour tout $0 \leq t_1 < t < t_2 \leq 1$,*

$$E\left((\mathcal{U}_n(t_2) - \mathcal{U}_n(t))^2(\mathcal{U}_n(t) - \mathcal{U}_n(t_1))^2\right) \leq K(t_2 - t_1)^{4\delta_0}.$$

Nous allons commencer par introduire quelques notations. Dans la suite de ce chapitre, nous noterons pour tout s, t dans $[0, 1]$ et tout x, y dans \mathbb{Z} , $D_{s,t}(x, y)$ et $A_{s,t}(x)$ les quantités définies de la manière suivante

$$(72) \quad D_{s,t}(x, y) = N_{[nt]}(x)N_{[nt]}(y) - N_{[ns]}(x)N_{[ns]}(y),$$

$$(73) \quad A_{s,t}(x) = N_{[nt]}(x) - N_{[ns]}(x), \quad A_t(x) = A_{0,t}(x) = N_{[nt]}(x).$$

Preuve de la Proposition 4.2 : Les suites $(S_n)_{n \geq 0}$ et $(\xi_x)_{x \in \mathbb{Z}}$ sont indépendantes, ainsi,

$$(74) \quad \begin{aligned} E\left((\mathcal{U}_n(t_2) - \mathcal{U}_n(t))^2(\mathcal{U}_n(t) - \mathcal{U}_n(t_1))^2\right) &= \frac{1}{n^{8\delta_0}} \sum_{(x_i, y_i)_{1 \leq i \leq 4}} E\left(\prod_{i=1}^2 D_{t, t_2}(x_i, y_i) \prod_{i=3}^4 D_{t_1, t}(x_i, y_i)\right) \\ &\times E\left(\prod_{i=1}^4 h(\xi_{x_i}, \xi_{y_i})\right) \\ &\leq \frac{1}{n^{8\delta_0}} \sum_{(x_i, y_i)_{1 \leq i \leq 4}} E\left(\prod_{i=1}^4 D_{t_1, t_2}(x_i, y_i)\right) E\left(\prod_{i=1}^4 h(\xi_{x_i}, \xi_{y_i})\right). \end{aligned}$$

Posons

$$\alpha(x_1, y_1, \dots, x_4, y_4) = E\left(\prod_{i=1}^4 h(\xi_{x_i}, \xi_{y_i})\right)$$

et notons \mathcal{S} le support de α ,

$$\mathcal{S} = \{(x_1, \dots, y_4) \mid \alpha(x_1, \dots, y_4) = 0\}.$$

Notations : Afin de simplifier les notations, nous écrirons de manière équivalente

- $(x_i, y_i)_{1 \leq i \leq 4}$ et $(x_1, y_1)(x_2, y_2)(x_3, y_3)(x_4, y_4)$ pour le 8-uplet $(x_1, y_1, \dots, x_4, y_4)$.
- Pour tout $k \in \{2, 3, 4\}$, $(x, y)^k$ désignera $\overbrace{(x, y) \cdots (x, y)}^{k \text{ fois}}$.

Le fait que h soit dégénérée implique que α s'annule sur les 8-uplets (x_1, \dots, y_4) pour lesquels il existe une coordonnée qui n'apparaît qu'une seule fois. Par conséquent, tous les éléments de \mathcal{S} contiennent chaque coordonnée au moins deux fois. Nous allons maintenant

nous intéresser aux sous-ensembles suivants de \mathcal{S} :

Pour tout $i \in \{3, \dots, 8\}$,

$$\mathcal{S}_i = \{(x_j, y_j)_{1 \leq j \leq 4} \in \mathcal{S} \mid \text{une des coordonnées apparaît exactement } i \text{ fois}\}.$$

et

$$\mathcal{S}_2 = \{(x_i, y_i)_{1 \leq i \leq 4} \in \mathcal{S} \mid \text{toutes les coordonnées apparaissent exactement deux fois}\}.$$

On notera que $\mathcal{S} = \cup_{i=2}^8 \mathcal{S}_i$. Ici, comme précédemment, l'argument de la dégénérescence de h s'applique. Quitte à réordonner les coordonnées, cela implique que

$$\mathcal{S}_8 = \{(x_1, x_1)^4\},$$

$$(75) \quad \mathcal{S}_2 = \{(x_1, x_1)(x_2, x_2)(x_3, x_3)(x_4, x_4)\} \cup \{(x_1, x_1)(x_2, x_2)(x_3, y_3)^2\} \\ \cup \{(x_1, y_1)^2(x_2, y_2)^2\},$$

$$(76) \quad \mathcal{S}_3 = \mathcal{S}_5 = \{(x_1, y_1)^3(x_4, x_4) \text{ avec } x_1 \notin \{x_4, y_1\}\},$$

$$(77) \quad \mathcal{S}_4 = \{(x_1, x_1)^2(x_2, x_2)(x_3, x_3), (x_1, x_1)^2(x_2, x_3)^2, (x_1, x_1)(x_1, x_2)^2(x_3, x_3), \\ (x_1, x_2)^2(x_1, x_3)^2, \text{ avec } x_1 \notin \{x_2, x_3\}\},$$

$$(78) \quad \mathcal{S}_6 = \{(x_1, x_1)^3(x_2, x_2) \text{ avec } x_1 \neq x_2\} \cup \{(x_1, x_1)^2(x_1, x_2)^2 \text{ avec } x_1 \neq x_2\}.$$

Notons

$$C = \max_{\{(x_i, y_i)_{1 \leq i \leq 4} \in \mathcal{S}\}} E\left(\prod_{i=1}^4 h(\xi_{x_i}, \xi_{y_i})\right).$$

Comme $h \in L^4(\mu \otimes \mu)$ et $\tilde{h} \in L^4(\mu)$, $C < +\infty$.

Par (74), il suit alors que

$$(79) \quad E\left((\mathcal{U}_n(t_2) - \mathcal{U}_n(t))^2 (\mathcal{U}_n(t) - \mathcal{U}_n(t_1))^2\right) \leq \frac{C}{n^{8\delta_0}} \sum_{(x_i, y_i)_{i \leq 4} \in \mathcal{S}} E\left(\prod_{i=1}^4 D_{t_1, t_2}(x_i, y_i)\right).$$

Maintenant il nous reste à calculer les termes

$$\sum_{(x_i, y_i)_{i \leq 4} \in \mathcal{S}_k} E\left(\prod_{i=1}^4 D_{t_1, t_2}(x_i, y_i)\right), \quad k = 2, 3, \dots, 8.$$

Étant donné que le calcul de ces quantités repose sur les mêmes arguments, nous allons nous concentrer sur le cas où $k = 3$.

Il découle de l'inégalité de Cauchy-Schwartz, de (76) et de la positivité D_{t_1, t_2} que

$$(80) \quad \sum_{(x_i, y_i)_{i \leq 4} \in \mathcal{S}_3} E \left(\prod_{i=1}^4 D_{t_1, t_2}(x_i, y_i) \right) \leq E \left(\sum_{x_1, y_1} D_{t_1, t_2}^3(x_1, y_1) \sum_{x_4, y_4} D_{t_1, t_2}(x_4, y_4) \right) \\ \leq E \left(\left(\sum_{x_1, y_1} D_{t_1, t_2}^3(x_1, y_1) \right)^2 \right)^{1/2} E \left(\left(\sum_{x_4, y_4} D_{t_1, t_2}(x_4, y_4) \right)^2 \right)^{1/2}$$

Nous allons d'abord considérer le deuxième terme $E \left(\left(\sum_{x_4, y_4} D_{t_1, t_2}(x_4, y_4) \right)^2 \right)^{1/2}$.

$$(81) \quad \sum_{x_4, y_4} D_{t_1, t_2}(x_4, y_4) = \left(\sum_{x_4} N_{[nt_2]}(x_4) \right)^2 - \left(\sum_{x_4} N_{[nt_1]}(x_4) \right)^2 = [nt_2]^2 - [nt_1]^2$$

Comme $0 \leq t_1 < t_2 \leq 1$,

$$[nt_2]^2 - [nt_1]^2 \leq 2n([nt_2] - [nt_1]).$$

De deux choses l'une :

- Soit $1 \leq n(t_2 - t_1)$ et alors, $[nt_2] - [nt_1] \leq n(t_2 - t_1) + 1 \leq 2n(t_2 - t_1)$.
- soit $n(t_2 - t_1) \leq 1$ et dans ce cas $[nt_2] - [nt_1] = 0$.

On en déduit que

$$(82) \quad [nt_2]^2 - [nt_1]^2 \leq 4n^2(t_2 - t_1).$$

De (81) et (82), il découle donc que

$$(83) \quad E \left(\left(\sum_{x_4, y_4} D_{t_1, t_2}(x_4, y_4) \right)^2 \right)^{1/2} \leq 4n^2(t_2 - t_1).$$

Nous allons maintenant nous intéresser à l'intégrale

$$(84) \quad E \left(\left(\sum_{x_1, y_1} D_{t_1, t_2}^3(x_1, y_1) \right)^2 \right)^{1/2}.$$

Rappelons, à partir des définitions de (72) et (73) de D et A que

$$D_{t_1, t_2}(x, y) = A_{t_1, t_2}(x)A_{t_1, t_2}(y) + A_{t_1}(x)A_{t_1, t_2}(y) + A_{t_1}(y)A_{t_1, t_2}(x).$$

Afin d'estimer (84), nous allons regarder la norme L^2 (notée $\|\cdot\|_2$) des termes suivants

$$\alpha_1 = \sum_{x, y} A_{t_1, t_2}^3(x)A_{t_1, t_2}^3(y), \quad \alpha_2 = \sum_{x, y} A_{t_1}^3(x)A_{t_1, t_2}^3(y), \quad \alpha_3 = \sum_{x, y} A_{t_1}(x)A_{t_1, t_2}^2(x)A_{t_1, t_2}^3(y),$$

$$\alpha_4 = \sum_{x,y} A_{t_1}^2(x)A_{t_1,t_2}(x)A_{t_1,t_2}^3(y) \text{ et } \alpha_5 = \sum_{x,y} A_{t_1}^2(x)A_{t_1,t_2}(x)A_{t_1}(y)A_{t_1,t_2}^2(y).$$

Commençons par évaluer $\|\alpha_1\|_2$:

$$\|\alpha_1\|_2 = \left\| \left(\sum_x (N_{[nt_2]}(x) - N_{[nt_1]}(x))^3 \right)^2 \right\|_2$$

De la propriété de Markov, il suit que

$$\sum_x (N_{[nt_2]}(x) - N_{[nt_1]}(x))^3$$

a la même distribution que $Q_{[nt_2]-[nt_1]}^{(3)}$. Par conséquent, par le Lemme 3.1,

$$(85) \quad \|\alpha_1\|_2 = \left\| \left(Q_{[nt_2]-[nt_1]}^{(3)} \right)^2 \right\|_2 = \mathcal{O}\left(\{n(t_2 - t_1)\}^{6-4/\alpha}\right).$$

Nous allons nous intéresser maintenant à $\|\alpha_2\|_2$:

$$\|\alpha_2\|_2 = \left(E \left(\sum_x N_{[nt_1]}^3(x) \sum_y A_{t_1,t_2}^3(y) \right)^2 \right)^{1/2}.$$

La propriété de Markov s'applique encore ici : $\sum_x N_{[nt_1]}^3(x)$ et $\sum_y A_{t_1,t_2}^3(y)$ sont indépendants et respectivement distribués selon $Q_{[nt_1]}^{(3)}$ et $Q_{[nt_2]-[nt_1]}^{(3)}$. Ainsi,

$$(86) \quad \begin{aligned} \|\alpha_2\|_2 &= \|Q_{[nt_1]}^{(3)}\|_2 \|Q_{[nt_2]-[nt_1]}^{(3)}\|_2 \\ &= \mathcal{O}(n^{6-4/\alpha}(t_2 - t_1)^{3-2/\alpha}). \end{aligned}$$

Dans le calcul du carré de $\|\alpha_3\|_2$, la propriété de Markov conduit à

$$(87) \quad \begin{aligned} \|\alpha_3\|_2^2 &= E \left(\sum_{x,y} A_{t_1}(x)A_{t_1,t_2}^2(x)A_{t_1,t_2}^3(y) \right)^2 \\ &= \sum_{x_1,y_1,x_2,y_2} E \left(A_{t_1}(x_1)A_{t_1,t_2}(x_2) \right) E \left((A_{t_1,t_2}^2(x_1)A_{t_1,t_2}^2(x_2)A_{t_1,t_2}^3(y_1)A_{t_1,t_2}^3(y_2)) \right) \end{aligned}$$

Par l'inégalité de Cauchy-Schwartz et (55) il suit que pour tout x_1 et x_2 dans \mathbb{Z} ,

$$(88) \quad E \left(A_{t_1}(x_1)A_{t_1,t_2}(x_2) \right) = \mathcal{O}(n^{2-2/\alpha}).$$

L'inégalité de Cauchy-Schwartz s'applique encore dans l'estimation du terme de gauche

$$(89) \quad \begin{aligned} E \left(\left(\sum_{x_1,y_1} A_{t_1,t_2}^2(x_1)A_{t_1,t_2}^3(y_1) \right)^2 \right) &= E \left(\left(\sum_{x_1} A_{t_1,t_2}^2(x_1) \right)^2 \left(\sum_{y_1} A_{t_1,t_2}^3(y_1) \right)^2 \right) \\ &\leq \|Q_{t_1,t_2}^{(2)}\|_4^2 \|Q_{t_1,t_2}^{(3)}\|_4^2 \end{aligned}$$

où $\|\cdot\|_4$ désigne la norme L_4 . Maintenant, par (87), (88), (89) et le Lemme 3.1, il suit que

$$(90) \quad \|\alpha_3\|_2 = \mathcal{O}\left(n^{6-4/\alpha}(t_2 - t_1)^{5-3/\alpha}\right).$$

On applique les mêmes arguments pour le calcul des deux termes de gauche $\|\alpha_4\|_2$ et $\|\alpha_5\|_2$. On montre ainsi que

$$(91) \quad \|\alpha_4\|_2 = \mathcal{O}\left(n^{6-4/\alpha}(t_2 - t_1)^{4-2/\alpha}\right), \quad \|\alpha_5\|_2 = \mathcal{O}\left(n^{6-4/\alpha}(t_2 - t_1)^{3-1/\alpha}\right).$$

On déduit de (85), (86), (90) (91) et de $t_1, t_2 \in [0, 1]$ que

$$(92) \quad \left(E\left(\sum_{x,y} D_{t_1,t_2}^3(x,y)\right)^2\right)^{1/2} = \mathcal{O}\left(n^{6-4/\alpha}(t_2 - t_1)^{3-2/\alpha}\right).$$

Pour finir (92) et (83) s'appliquent dans l'estimation de (80), ce qui conduit à

$$\sum_{(x_i, y_i)_{i \leq 4} \in \mathcal{S}_3} E\left(\prod_{i=1}^4 D_{t_1, t_2}(x_i, y_i)\right) = \mathcal{O}\left(n^{8\delta_0}(t_2 - t_1)^{4-2/\alpha}\right).$$

Si l'on remplace dans les calculs précédents \mathcal{S}_3 par \mathcal{S}_i pour $i = 2, 4, \dots, 8$, les mêmes techniques s'appliquent et l'on peut vérifier que

$$(93) \quad \sum_{(x_i, y_i)_{i \leq 4} \in \mathcal{S}_k} E\left(\prod_{i=1}^4 D_{t_1, t_2}(x_i, y_i)\right) = \mathcal{O}\left(n^{8\delta_0}(t_2 - t_1)^{4-2/\alpha}\right), \quad k = 2, \dots, 6$$

et

$$(94) \quad \sum_{(x_i, y_i)_{i \leq 4} \in \mathcal{S}_8} E\left(\prod_{i=1}^4 D_{t_1, t_2}(x_i, y_i)\right) = \mathcal{O}\left(n^{8\delta_0-2/\alpha}(t_2 - t_1)^{4-2/\alpha}\right).$$

Finalement, il suit de (79), (93) et (94), qu'il existe une constante positive K et un $\gamma = 4 - 2/\alpha = 4\delta_0$, $\gamma > 1$, tels que, pour tout $0 \leq t_1 \leq t_2 \leq 1$

$$E\left((\mathcal{U}_n(t_2) - \mathcal{U}_n(t))^2(\mathcal{U}_n(t) - \mathcal{U}_n(t_1))^2\right) \leq K(t_2 - t_1)^{4\delta_0}.$$

□

5. Propriétés du processus limite dans le cas dégénéré.

Le processus $(W_t)_{t \in [0,1]} \in \mathcal{D}([0, \infty[, \mathbb{R})$, vérifie les propriétés suivantes :

Proposition 5.1. *Le processus $(W_t)_{t \in [0,1]}$ est auto-similaire d'indice $2\delta_0$. D'autre part, il admet une version continue.*

Preuve : Pour tout $K > 0$, le processus K -dimensionnel, $(\Delta_t^{(1)}, \dots, \Delta_t^{(K)})$ est auto-similaire d'indice δ_0 (cf. [48]) par conséquent, $(W_t)_{t \in [0,1]}$, qui est limite faible de la suite de processus $\{(W_{K,t})_{t \in [0,1]}, K = 1, 2, \dots\}$ est, lui, auto-similaire d'indice $2\delta_0$.

En utilisant les propriétés du processus $(\Delta_t)_{t \geq 0}$ (auto-similaire d'indice $2\delta_0$, à accroissements stationnaires), on montre, de la même façon que dans l'étape 3 de la section précédente, que pour tout $t, s \in [0, 1]$,

$$\begin{aligned} E\left((W_t - W_s)^2\right) &\leq (t - s)^{4\delta_0} \left(\sum_{\nu=1}^{+\infty} \lambda_\nu^2 + \left(\sum_{\nu=1}^{+\infty} \lambda_\nu \right)^2 \right) E\left((\Delta_1^{(1)})^4\right) \\ &\leq C(t - s)^{4\delta_0}. \end{aligned}$$

Comme $4\delta_0 > 1$, on conclut par le Théorème de Kolmogorov-Čentsov (cf. [43] p.53).
□

Remarque :

- Contrairement à ce qui se passait, dans le cadre des U -statistiques classiques (ainsi que dans le cadre des U -statistiques échantillonnées par une marche aléatoire $(S_n)_{n \in \mathbb{N}}$, sous les hypothèses décrites au chapitre 3), où le processus limite, $(U(t))_{t \in [0,1]}$, s'écrivait (cf. (37)) comme une somme infinie, pondérée, de processus indépendants (les $\{(B_t^{(\nu)})^2\}_{t \in [0,1]}, \nu = 1, 2, \dots\}$) nous allons voir qu'ici, les processus $\{(\Delta_t^{(\nu)})_{t \in [0,1]}, \nu = 1, 2, \dots\}$ qui apparaissent dans l'écriture (54) de $(W_t)_{t \in [0,1]}$, ne sont pas indépendants.

Preuve : Si les $\{(\Delta_t^{(\nu)})_{t \in [0,1]}, \nu = 1, 2, \dots\}$ étaient indépendants, ce serait en particulier le cas des deux variables aléatoires $\Delta_1^{(1)}$ et $\Delta_1^{(2)}$. D'après l'expression de la fonction caractéristique du vecteur aléatoire bi-dimensionnel $(\Delta_1^{(1)}, \Delta_1^{(2)})$ calculée en $\theta = (\theta_1, \theta_2)$ (voir Lemme 3.3 du chapitre 1 et [48]), si l'on note ϕ la fonction caractéristique de chacune des variables aléatoires $\Delta_1^{(1)}$ et $\Delta_1^{(2)}$ (équidistribuées), cela impliquerait que

$$(95) \quad \phi\left(\sqrt{\theta_1^2 + \theta_2^2}\right) = \phi(\theta_1)\phi(\theta_2), \quad \forall \theta_1, \theta_2, \in \mathbb{R},$$

On montre alors, (en posant $f(x) = \log(\phi(\sqrt{x}))$) et en se ramenant à l'ensemble des fonctions f continues sur \mathbb{R} vérifiant $f(x + y) = f(x) + f(y), \forall x, y \in \mathbb{R}$,) que les seules fonctions caractéristiques vérifiant (95) sont celles des v.a. gaussiennes. Or, les $(\Delta_t)_{t \in [0,1]}$ ne sont pas des processus gaussiens. □

6. Preuve du Théorème 2.2

Nous nous intéressons ici au cas où le noyau h est dégénéré. Comme au chapitre 3, nous allons voir que, dans ce cas de figure, $U_{[nt]}$ se comporte asymptotiquement, à une renormalisation près, comme une marche aléatoire en scène aléatoire, grâce à la h -décomposition de Hoeffding.

En effet, par cette décomposition, on peut écrire que pour tout $x \neq y$,

$$h(\xi_x, \xi_y) = m + (g(\xi_x) - m) + (g(\xi_y) - m) + \phi(\xi_x, \xi_y)$$

où $g(\xi_x) = E(h(\xi_x, \xi_y)|\xi_x)$ et ϕ est un noyau dégénéré.

En utilisant l'écriture précédente :

$$\begin{aligned} U_{[nt]} &= m([nt] + 1)^2 - mV_{[nt]} + \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m)) \\ &+ 2([nt] + 1) \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)(g(\xi_x) - m) + \sum_{x \neq y} N_{[nt]}(x)N_{[nt]}(y)\phi(\xi_x, \xi_y) \\ &= m([nt] + 1)^2 - mV_{[nt]} + \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m)) \\ &+ 2([nt] + 1) \sum_{k=0}^{[nt]} (g(\xi_{S_k}) - m) + \sum_{x \neq y} N_{[nt]}(x)N_{[nt]}(y)\phi(\xi_x, \xi_y) \end{aligned}$$

Par le Lemme 2.1,

$$E(V_n) = \mathcal{O}(n^{2\delta_0}) = \mathcal{O}(n^{2-\frac{1}{\alpha}})$$

ainsi $V_n/n^{2-1/2\alpha}$ converge en probabilité vers 0 quand n tend vers l'infini,

$$\frac{V_n}{n^{2-1/2\alpha}} \xrightarrow{P} 0.$$

En utilisant les mêmes arguments que pour le Théorème 2.3, chapitre 3, c'est-à-dire en calculant la norme L_2 (respectivement la norme L_1) de $\sum_{x \neq y} N_{[nt]}(x)N_{[nt]}(y)\phi(\xi_x, \xi_y)$ (resp. $\sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m))$) et en utilisant le Lemme 3.1, on montre que

$$\frac{1}{n^{2-\frac{1}{2\alpha}}} \left[\sum_{x \neq y} N_{[nt]}(x)N_{[nt]}(y)\phi(\xi_x, \xi_y) + \sum_{x \in \mathbb{Z}^2} N_{[nt]}(x)^2 (h(\xi_x, \xi_x) - 2(g(\xi_x) - m)) \right]$$

converge en probabilité vers 0 quand n tend vers l'infini.

Finalement, on applique le Théorème de Kesten et Spitzer (voir [44]),

$$\left(\frac{1}{\sigma n^{1-\frac{1}{2\alpha}}} \sum_{k=0}^{[nt]} (g(\xi_{S_k}) - m) \right)_{t \in [0,1]} \xrightarrow{\mathcal{D}} (\Delta_t)_{t \in [0,1]}, \quad n \rightarrow \infty$$

et notre résultat suit. □

7. Commentaires

• Dans le cas où le noyau est dégénéré, dans la preuve de la convergence des lois de dimensions finies, nous avons été amenés à nous restreindre à la classe \mathcal{L} des fonctions h de $L^2(\mu \otimes \mu)$ dont la restriction à la droite $\{(x, y) \in \mathbb{R}^2 \mid y = x\}$, \tilde{h} , définie par $\tilde{h}(x) = h(x, x)$ pour tout réel x , vérifie (52) :

$$\tilde{h}(x) = \sum_{\nu \in \mathbb{N}} \lambda_\nu \phi_\nu^2(x), \quad \text{dans } L^1(\mu).$$

Il existe de nombreux exemples de fonctions de $L^2(\mu \otimes \mu)$ qui ne vérifient pas la condition ci-dessus. C'est notamment le cas des fonctions dont la série des valeurs propres diverge, puisque (52) entraîne la convergence de $\sum_{\nu=1}^{\infty} \lambda_\nu$.

Contre-exemple : Notons μ_0 , la mesure uniforme sur $[0, 2\pi]$, i.e.

$$\mu_0 = \frac{1}{2\pi} \mathbf{1}_{[0, 2\pi]}(x) dx,$$

et T , le cercle unité du plan complexe. On peut identifier les fonctions définies sur T avec les fonctions définies sur \mathbb{R} de période 2π . Avec cette convention on définit $L^2(T)$ comme l'ensemble des fonctions complexes 2π -périodiques de $L^2(\mu_0)$. On notera $(. \mid .)$ le produit scalaire sur $L^2(T)$. A partir de toute fonction

$$k \in L^2(T) \text{ réelle, \quad telle que \quad } k(x) = k(-x), \quad \forall x \in \mathbb{R},$$

on peut construire un noyau $\mathcal{K} \in L^2(\mu_0 \otimes \mu_0)$ hermitien réel, en posant :

$$(96) \quad \mathcal{K}(x, y) = k(x - y), \quad \forall x, y \in \mathbb{R}.$$

La théorie des développements en séries de Fourier va nous permettre de décomposer \mathcal{K} sous la forme (32). Notons, pour tout $n \in \mathbb{Z}$,

$$e_n(x) = e^{inx}, \quad \forall x \in \mathbb{R}, \quad \text{et} \quad c_n = (k, e_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} k(t) e^{-int} dt.$$

Comme k est symétrique réelle (ce qui entraîne, pour tout $n \in \mathbb{N}$, que $c_n = c_{-n} \in \mathbb{R}$), son développement en série de Fourier s'écrit,

$$\begin{aligned} k(x) &= \sum_{n \in \mathbb{Z}} c_n e^{inx}, & \text{dans } L^2(T) \\ &= c_0 + \sum_{n \in \mathbb{N}^*} 2c_n \cos(nx), & \text{dans } L^2(T). \end{aligned}$$

On peut remarquer que les valeurs propres de l'opérateur $T_{\mathcal{K}}$, ne sont autres que les coefficients de Fourier de k , $(c_n)_{n \in \mathbb{Z}}$, et que la suite $(e_n)_{n \in \mathbb{Z}}$, de fonctions de $L^2(T)$, constitue une base orthonormale, de fonctions propres de $T_{\mathcal{K}}$, associées respectivement aux $(c_n)_{n \in \mathbb{Z}}$.

Nous pouvons donc décomposer le noyau \mathcal{K} , grâce au Théorème 3.3, chapitre 1, à partir des $(e_n)_{n \in \mathbb{Z}}$ et des $(c_n)_{n \in \mathbb{Z}}$, et l'on constate alors que l'on retombe exactement sur le développement en série de Fourier de k au point $x - y$,

$$\mathcal{K}(x, y) = \sum_{n \in \mathbb{Z}} c_n e_n(x) \overline{e_n(y)} = \sum_{n \in \mathbb{Z}} c_n e^{in(x-y)}, \quad \text{dans } L^2(\mu_0 \otimes \mu_0)$$

Dans le cadre des U -statistiques, la décomposition se fait selon des fonctions propres réelles. Posons pour tout x réel :

$$\phi_0(x) = 1, \quad \text{et pour tout } n \in \mathbb{N}^*, \quad \phi_{n,1}(x) = \sqrt{2} \cos(nx) \quad \text{et} \quad \phi_{n,2}(x) = \sqrt{2} \sin(nx).$$

La suite de fonctions $\{\phi_0, \phi_{n,1}, \phi_{n,2}, n \in \mathbb{N}^*\}$ constitue une base orthonormale de fonctions propres réelles de l'opérateur $T_{\mathcal{K}}$, associées aux valeurs propres de $T_{\mathcal{K}}, (c_n)_{n \in \mathbb{Z}}$, (respectivement de valeurs propres $\{c_0, c_n, c_n, n \in \mathbb{N}^*\}$). Ainsi, quitte à ré-indexer les (ϕ) , la décomposition,

$$\mathcal{K}(x, y) = c_0 + \sum_{n \in \mathbb{N}^*} c_n \phi_{n,1}(x) \phi_{n,2}(y) + \sum_{n \in \mathbb{N}^*} c_n \phi_{n,1}(x) \phi_{n,2}(y) \quad \text{dans } L^2(\mu_0 \otimes \mu_0),$$

correspond à une décomposition de \mathcal{K} , de la forme (32).

Venons-en aux contre-exemples.

Notons, pour tout $k \in L^2(T)$, \widehat{k} la suite des coefficients de Fourier de k . L'application

$$\begin{aligned} L^2(T) &\rightarrow \ell^2(\mathbb{Z}) \\ k &\mapsto \widehat{k} = ((k | e_n))_{n \in \mathbb{Z}}, \end{aligned}$$

est un isomorphisme d'espaces de Hilbert (par les Théorèmes de *Riesz-Fisher* et de *Parseval* [61]). Ainsi, à toute suite, σ , de l'ensemble

$$\mathcal{P} = \left\{ s = (s_n)_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \quad \mid \quad \left| \sum_{n \in \mathbb{N}} s_n \right| = +\infty \right\}.$$

est associée une fonction k_σ de $L^2(T)$ telle que $\widehat{k_\sigma} = \sigma$. On notera $\mathcal{K}_\sigma \in L^2(\mu_0 \otimes \mu_0)$ le noyau construit à partir de k_σ selon (96). Eh bien, tous les noyaux de l'ensemble $\{\mathcal{K}_\sigma \mid \sigma \in \mathcal{P}\}$ constituent des contre-exemples de fonction de $L^2(\mu_0 \otimes \mu_0)$ qui ne sont pas dans \mathcal{L} .

• Cependant, de nombreuses fonctions de $L^2(\mu \otimes \mu)$ appartiennent à la classe \mathcal{L} . C'est en particulier le cas de la classe de fonctions qui vérifient les hypothèses du *Théorème de Mercer* [42]. Nous allons citer ce théorème dans le cas où μ est une mesure uniforme sur un intervalle $[a, b]$,

$$\mu = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x) dx.$$

Signalons qu'il existe des versions de ce résultat énoncées dans des contextes un peu plus généraux, et nous renvoyons, par exemple, à l'ouvrage de Jörgens [42], pour plus de détails à ce sujet.

Théorème de Mercer 7.1. [42] *Soit h une fonction continue sur $[a, b] \times [a, b]$ dont l'opérateur T_h associé est positif, i.e.*

$$(97) \quad \int_a^b \int_a^b h(t, s) f(s) f(t) ds dt \geq 0, \quad \forall f \in L^2([a, b]) \text{ réelle},$$

alors (i) la représentation (32) est vérifiée en tous points de $[a, b] \times [a, b]$, i.e.

$$h(x, y) = \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(x) \phi_\nu(y), \quad \forall (x, y) \in [a, b] \times [a, b].$$

(ii) La série $\sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(x) \phi_\nu(y)$ converge absolument et uniformément sur $[a, b] \times [a, b]$.

(iii) (Formule de la trace) $\sum_{\nu=1}^{+\infty} \lambda_\nu = \frac{1}{b-a} \int_{-a}^{+b} h(x, x) dx$.

Remarque : Notons au passage qu'il ne suffit pas de supposer h continue pour avoir

$$h(x, y) = \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(x) \phi_\nu(y), \quad \text{pour tout } x, y \in [a, b].$$

Contre-exemple : Reprenons l'exemple, ci-dessus, dans lequel le noyau h était associé à une fonction continue 2π -périodique k par la relation, $h(x, y) = k(x - y)$ et où $[a, b] = [0, 2\pi]$. Si le Théorème de Mercer s'appliquait, sans qu'on ait besoin de supposer la positivité de T_h , cela impliquerait (en prenant $y = 0$) que toute fonction continue 2π -périodique, k , est la limite uniforme de sa série de Fourier. Or, il existe de nombreux exemples bien connus, pour lesquels ce n'est pas le cas (cf Rudin [61]).

• Nous allons voir maintenant en quoi il est nécessaire, dans le cadre du Théorème 2.1, de rajouter une condition concernant la restriction de h à la droite $\{(x, y) \in \mathbb{R}^2 \mid y = x\}$ (hypothèse (52)). Nous allons montrer, à l'aide d'un exemple que l'on peut construire deux versions, dans $L^4(\mu \otimes \mu)$, d'un même noyau h , notées h_1 et h_2 , telles que h_2 ne vérifie pas (52) et telles que les deux suites de processus attachés aux U -statistiques échantillonnées par la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, de noyaux respectifs h_1 et h_2 , ne possèdent pas la même limite faible.

Exemple : Notons μ la mesure de Lebesgue sur l'intervalle $[0, 1]$ et considérons les deux fonctions suivantes :

$$h_1(x, y) = 0, \quad \forall x, y \text{ dans } [0, 1],$$

et

$$h_2(x, y) = \begin{cases} 1 & \text{si } x = y \\ 0 & \text{sinon.} \end{cases}$$

Les deux noyaux h_1 et h_2 sont dégénérés et égaux dans $L^4(\mu \otimes \mu)$ puisque la mesure de Lebesgue de $\{(x, y) \in \mathbb{R}^2 \mid y = x\}$ est nulle.

On note

$$\mathcal{U}_{1,n}(t) = \frac{1}{n^{2\delta_0}} \sum_{i,j=0}^{[nt]} h_1(\xi_{S_i}, \xi_{S_j}), \quad \text{et} \quad \mathcal{U}_{2,n}(t) = \frac{1}{n^{2\delta_0}} \sum_{i,j=0}^{[nt]} h_2(\xi_{S_i}, \xi_{S_j}).$$

Alors,

$$\mathcal{U}_{1,n}(t) = 0, \quad \text{et} \quad \mathcal{U}_{2,n}(t) = \frac{1}{n^{2\delta_0}} \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x).$$

Nous voyons que les deux suites de processus $\{(\mathcal{U}_{1,n}(t))_{t \in [0,1]}, n = 1, 2, \dots\}$ et $\{(\mathcal{U}_{2,n}(t))_{t \in [0,1]}, n = 1, 2, \dots\}$, bien que basées sur deux noyaux égaux dans $L^4(\mu \otimes \mu)$, possèdent des limites faibles, dans $\mathcal{D}([0, 1])$, distinctes puisque d'après le Lemme 3.4 du chapitre 1 (cf. [44]),

quand $n \rightarrow \infty$,

$$(\mathcal{U}_{2,n}(t))_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(\int_{-\infty}^{\infty} L_t^2(x) dx \right)_{t \in [0,1]},$$

alors que

$$(\mathcal{U}_{1,n}(t))_{t \in [0,1]} \xrightarrow{\mathcal{D}} 0.$$

8. Problème ouvert

Pour éviter d'avoir à se restreindre à la sous-classe des fonctions h de $L^4(\mu \otimes \mu)$ qui vérifient l'hypothèse (52) ainsi que $\tilde{h} \in L^4(\mu)$, il semble relativement naturel d'introduire, comme au chapitre précédent, la suite des processus $(\widehat{\mathcal{U}}_n(t))_{t \in [0,1]}$, $n = 1, 2, \dots$ associés aux U -statistiques indexées par la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, définies en dehors des auto-intersections de $(S_n)_{n \in \mathbb{N}}$:

$$\widehat{\mathcal{U}}_{[nt]} = \frac{1}{n^{2\delta_0}} \sum_{0 \leq i, j \leq [nt]} h(\xi_{S_i}, \xi_{S_j}) \mathbf{1}_{\{S_i \neq S_j\}}.$$

Cela dit, nous allons voir que, dans le cadre considéré au cours de ce chapitre (celui d'une marche aléatoire récurrente uni-dimensionnelle), la technique de la preuve utilisée pour $(\mathcal{U}_n(t))_{t \in [0,1]}$ ne va pas nous permettre de conclure quant au comportement asymptotique de $(\widehat{\mathcal{U}}_{[nt]})_{t \in [0,1]}$. En effet, la méthode de la démonstration du Théorème 2.1, basée sur le Théorème 3.2 du chapitre 1 (cf. [48]), fait apparaître les auto-intersections de $(S_n)_{n \in \mathbb{N}}$ à travers le carré de la marche aléatoire en scène aléatoire vectorielle suivante, associée au noyau tronqué h_K :

$$\left(\sum_{i=0}^{[nt]} \tilde{\xi}_{S_i}^{(K)} \right)^2.$$

Pour pouvoir utiliser et appliquer ce résultat concernant les marches aléatoires en scène aléatoire vectorielle (le Théorème 3.2 du chapitre 1), à $(\widehat{\mathcal{U}}_{[nt]})_{t \in [0,1]}$, il faut réintroduire, à un moment ou à un autre, les auto-intersections de $(S_n)_{n \in \mathbb{N}}$ et nous allons brièvement tenter d'expliquer pourquoi, ici, les choses se compliquent.

Méthode 1 :

Nous noterons avec un "chapeau" les analogues, pris en dehors des auto-intersections de $(S_n)_{n \in \mathbb{N}}$, des quantités utilisées dans la preuve du Théorème 2.1 :

$$\begin{aligned}
\widehat{\mathcal{U}}_{n,K}(t) &= \frac{1}{n^{2\delta_0}} \sum_{\nu=1}^K \lambda_\nu \sum_{0 \leq i,j \leq [nt]} \phi_\nu(\xi_{S_i}) \phi_\nu(\xi_{S_j}) \mathbf{1}_{\{S_i \neq S_j\}} \\
&= \frac{1}{n^{2\delta_0}} \sum_{\nu=1}^K \lambda_\nu \left\{ \left(\sum_{i=0}^{[nt]} \phi_\nu(\xi_{S_i}) \right)^2 - \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) \phi_\nu^2(\xi_x) \right\} \\
&= \mathcal{U}_{n,K}(t) - \widetilde{\mathcal{U}}_{n,K}(t)
\end{aligned}$$

où $\mathcal{U}_{[nt],K}$ est le processus défini dans la section 4 par (60) et où

$$\widetilde{\mathcal{U}}_{[nt],K} = \frac{1}{n^{2\delta_0}} \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) \left(\sum_{\nu=1}^K \lambda_\nu \phi_\nu^2(\xi_x) \right).$$

Nous pouvons décrire le comportement asymptotique, quand $n \rightarrow \infty$, des deux suites de processus $(\mathcal{U}_{[nt],K})_{t \in [0,1]}$ et $(\widetilde{\mathcal{U}}_{[nt],K})_{t \in [0,1]}$:

- d'après l'étape 1 de la section 4, quand n tend vers l'infini,

$$(\mathcal{U}_{n,K}(t))_{t \in [0,1]} \xrightarrow{\mathcal{D}} (W_K(t) := \sum_{\nu=1}^K \lambda_\nu (\Delta_t^{(\nu)})^2)_{t \in [0,1]},$$

- d'autre part, nous allons voir que :

$$(\widetilde{\mathcal{U}}_{[nt],K})_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(\sum_{\nu=1}^K \lambda_\nu \int_{\mathbb{R}} L_t^2(x) dx \right)_{t \in [0,1]}, \quad \text{quand } n \rightarrow \infty.$$

On écrit,

$$(98) \quad \widetilde{\mathcal{U}}_{n,K} = \frac{1}{n^{2\delta_0}} \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) \left(\sum_{\nu=1}^K \lambda_\nu (\phi_\nu^2(\xi_x) - 1) \right) + \frac{1}{n^{2\delta_0}} \sum_{\nu=1}^K \lambda_\nu \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x).$$

Remarquons que, d'après (50) et (51), la variable aléatoire $\sum_{\nu=1}^K \lambda_\nu (\phi_\nu^2(\xi_x) - 1)$ est de carré intégrable. Nous pouvons donc, en conditionnant le terme de gauche, de l'écriture de $\widetilde{\mathcal{U}}_{n,K}$ (98), par rapport à la σ -algèbre \mathcal{A} , engendrée par X_1, X_2, \dots , calculer l'espérance de son carré. On trouve :

$$E\left(\left(\frac{1}{n^{2\delta_0}} \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) \left(\sum_{\nu=1}^K \lambda_\nu (\phi_\nu^2(\xi_x) - 1) \right) \right)^2 \middle| \mathcal{A} \right) \leq \frac{C_K}{n^{4\delta_0}} \sum_{x \in \mathbb{Z}} N_{[nt]}^4(x)$$

où C_K désigne une constante > 0 . On conclut en utilisant le Lemme 3.1 que, quand $n \rightarrow \infty$, le terme de gauche dans l'écriture de $\widetilde{\mathcal{U}}_{n,K}$ (98), converge en probabilité vers 0.

Par ailleurs, d'après le Lemme 3.4 du chapitre 1, le terme de droite converge faiblement, quand n tend vers l'infini, dans l'espace $(D([0, 1]), \mathcal{S})$ vers

$$\int_{\mathbb{R}} L_t^2(x) dx \sum_{\nu=1}^K \lambda_\nu.$$

Quoi qu'il en soit, dans tous les cas, il n'est pas possible de conclure au niveau de l'étape 1, qui concerne la limite quand $n \rightarrow \infty$ de la suite de processus $\{(\widehat{\mathcal{U}}_{n,K}(t))_{t \in [0,1]}, n = 1, 2, \dots\}$, puisqu'on ne sait rien dire, de manière générale, quant-au comportement asymptotique de la somme de deux processus qui convergent faiblement.

Méthode 2 :

Maintenant, abandonnons l'idée de supprimer l'hypothèse (52) (ainsi que le fait que $\tilde{h} \in L^4(\mu)$) et essayons de déduire directement le comportement asymptotique de $(\widehat{\mathcal{U}}_{[nt]})_{t \in [0,1]}$, de celui de $(\mathcal{U}_n(t))_{t \in [0,1]}$.

Soit

$$I_1 = E(\tilde{h}(\xi_x)) = \sum_{\nu=1}^{\infty} \lambda_\nu.$$

En écrivant :

$$\begin{aligned} \widehat{\mathcal{U}}_{[nt]} &= \mathcal{U}_n(t) - \frac{1}{n^{2\delta_0}} \sum_{0 \leq i, j \leq [nt]} h(\xi_{S_i}, \xi_{S_j}) \mathbf{1}_{\{S_i = S_j\}} \\ &= \mathcal{U}_n(t) - \frac{1}{n^{2\delta_0}} \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) h(\xi_x, \xi_x) \end{aligned}$$

on voit que se pose le même genre de problème que dans la méthode 1. En effet, ici comme \tilde{h} est supposée être dans $L^4(\mu)$ (donc dans $L^2(\mu)$), pour les mêmes raisons que celle évoquées ci-dessus :

$$\frac{1}{n^{2\delta_0}} \sum_{x \in \mathbb{Z}} N_{[nt]}^2(x) h(\xi_x, \xi_x) \xrightarrow{\mathcal{D}} I_1 \int_{\mathbb{R}} L_t^2(x) dx, \quad \text{quand } n \rightarrow \infty.$$

Même si l'on sait que

$$(\mathcal{U}_n(t))_{t \in [0,1]} \xrightarrow{\mathcal{D}} \left(\sum_{\nu=1}^{\infty} \lambda_\nu (\Delta_t^{(\nu)})^2 \right)_{t \in [0,1]}, \quad \text{quand } n \rightarrow \infty,$$

à ce stade, on ne peut pas vraiment plus conclure qu'avec la méthode précédente.

Dans tous les cas, on serait tenté de dire que la limite faible $(\widehat{\mathcal{U}}_n(t))_{t \in [0,1]}$ quand n tend vers l'infini, si elle existe, pourrait être de la forme informelle suivante

$$\left\langle \sum_{\nu=1}^{\infty} \lambda_\nu \left((\Delta_t^{(\nu)})^2 - \int_{\mathbb{R}} L_t^2(x) dx \right) \right\rangle.$$

CHAPITRE 5

Une nouvelle preuve des résultats du chapitre 3

Nous donnons ici une nouvelle preuve des Théorèmes 2.1 2.2, 3.1 et 3.2 du chapitre 3 (cf. [18]), basée sur la même méthode que celle utilisée dans le cas des marches aléatoires récurrentes en dimension 1.

Dans ce chapitre, nous allons brièvement voir comment la méthode de la preuve du Théorème 2.1, chapitre 4, s'applique aussi dans le cas où l'on considère, comme dans ([18]), une U -statistique échantillonnée :

- soit par une marche aléatoire récurrente sur \mathbb{Z}^2 ,
- soit par une marche aléatoire transiente sur $\mathbb{Z}^d, d \geq 1$.

Tout au long de ce chapitre, nous garderons les mêmes notations et les mêmes hypothèses qu'au chapitre 3.

1. Une nouvelle preuve des théorèmes 2.1 et 2.2 du chapitre 3

Nous traiterons d'abord du cas où $(S_n)_{n \in \mathbb{N}}$ est une marche aléatoire récurrente sur \mathbb{Z}^2 et nous nous intéresserons en premier lieu à la suite des processus $(\widehat{U}_n(t))_{t \in [0,1]}$, $n = 1, 2, \dots$, définis en dehors des auto-intersections de la marche aléatoire $(S_n)_{n \in \mathbb{N}}$ par l'équation (39).

1.1. Une nouvelle preuve du théorème 2.1 du chapitre 3.

La démonstration que nous allons donner diffère de celle de Cabus et Guillotin [18] exposée au chapitre 3, seulement au niveau de la preuve de la convergence des lois de dimension finie. L'adaptation de la technique de preuve du chapitre précédent nous permet d'obtenir une démonstration plus simple que celle de [18] et qui ne nécessite pas, ici, de conditions supplémentaires sur une éventuelle décomposition du noyau \tilde{h} .

Convergence des lois de dimension finie:

Comme au chapitre 3, on fixe $\theta_1, \dots, \theta_m \in \mathbb{R}$ et $0 = t_0 < t_1 < \dots < t_m \leq 1$ et l'on pose

$$\mathcal{W}_n = \sum_{j=1}^m \theta_j \widehat{\mathcal{U}}_n(t_j).$$

Pour tout entier $K > 0$, on appellera $\mathcal{W}_{n,K}$ et $(\widehat{\mathcal{U}}_{n,K}(t))_{t \in [0,1]}$ les équivalents basés sur le noyau tronqué h_K , défini, comme au chapitre 4, à partir de la décomposition de h par l'équation (59), respectivement de la suite de variables aléatoires $(\mathcal{W}_n)_{n \geq 1}$ et (resp.) de la suite des processus $(\widehat{\mathcal{U}}_n(t))_{t \in [0,1]}$, $n = 1, 2, \dots$:

$$(99) \quad \mathcal{W}_{n,K} = \sum_{j=1}^m \theta_j \widehat{\mathcal{U}}_{n,K}(t_j),$$

avec

$$\widehat{\mathcal{U}}_{n,K}(t) = \frac{1}{cn \log n} \sum_{0 \leq i, j \leq [nt]} h_K(\xi_{S_i}, \xi_{S_j}) \mathbf{1}_{\{S_i \neq S_j\}}.$$

D'autre part, on désignera par $(U_K(t))_{t \in [0,1]}$ le processus défini à partir de $(U(t))_{t \in [0,1]}$ par :

$$U_K(t) = \sum_{\nu=1}^K \lambda_\nu ((B_t^{(\nu)})^2 - t).$$

La méthode de preuve s'appuie essentiellement :

- sur le Théorème 4.2 du chapitre 1 (qui est l'équivalent vectoriel du Théorème de Bolthausen [13] et qui joue le même rôle que jouait le Théorème de Maejima [48] dans la preuve du Théorème 2.1, chapitre 4),
- ainsi que sur la loi des grands nombres suivante :

Lemme 1.1. *Quand n tend vers l'infini, pour tout $t \in [0, 1]$ et pour tout $\nu \geq 1$, $\frac{1}{cn \log n} \sum_{x \in \mathbb{Z}^2} N_{[nt]}^2(x) \phi_\nu^2(\xi_x)$ converge en probabilité vers t :*

$$\frac{1}{cn \log n} \sum_{x \in \mathbb{Z}^2} N_{[nt]}^2(x) \phi_\nu^2(\xi_x) \xrightarrow{P} t, \quad \text{quand } n \rightarrow \infty.$$

Preuve du Lemme 1.1 : On écrit,

$$(100) \quad \frac{1}{cn \log n} \sum_{x \in \mathbb{Z}^2} N_{[nt]}^2(x) \phi_\nu^2(\xi_x) = \frac{1}{cn \log n} \sum_{x \in \mathbb{Z}^2} N_{[nt]}^2(x) (\phi_\nu^2(\xi_x) - 1) + \frac{V_{[nt]}}{cn \log n}.$$

Par le Lemme 2.1, chapitre 3, pour tout $t \in [0, 1]$ $\frac{V_{[nt]}}{cn \log n}$ converge presque sûrement vers t , quand n tend vers l'infini,

$$(101) \quad \forall t \in [0, 1] \quad \lim_{n \rightarrow \infty} \frac{V_{[nt]}}{cn \log n} = t, \quad \text{p.s.}$$

D'autre part, en conditionnant le terme de gauche par rapport à la marche et en calculant l'espérance de son carré,

$$E\left(\left(\frac{1}{cn \log n} \sum_{x \in \mathbb{Z}^2} N_{[nt]}^2(x) (\phi_\nu^2(\xi_x) - 1)\right)^2 \middle| \mathcal{A}\right) \leq \frac{C_\nu}{n^2 \log^2 n} \sum_{x \in \mathbb{Z}^2} N_{[nt]}^4(x),$$

on voit en utilisant le lemme 2.2 du chapitre 3 que pour tout $\nu \geq 1$ et tout $t \in [0, 1]$

$$(102) \quad \frac{1}{cn \log n} \sum_{x \in \mathbb{Z}^2} N_{[nt]}^2(x) (\phi_\nu^2(\xi_x) - 1) \xrightarrow{P} 0, \quad \text{quand } n \rightarrow \infty.$$

On conclut le lemme grâce à (101) et (102) appliqués (100).

□

Revenons à la démonstration de la convergence des lois de dimension finie. L'idée ici, comme au chapitre précédent, est de s'intéresser et de se ramener au comment asymptotique de la suite, $(\widehat{\mathcal{U}}_{n,K}(t))_{t \in [0,1]}$, $n = 1, 2, \dots$, des U -statistiques échantillonnées par la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, basées sur le noyau tronqué, h_K . Le plan de la preuve est le même que celui du Théorème 2.1, chapitre 4. Il se décompose en trois étapes :

Étape 1 :

On ré-écrit la variable aléatoire $\widehat{\mathcal{W}}_{n,K}$ définie par (99) :

$$\begin{aligned} \widehat{\mathcal{W}}_{n,K} &= \frac{1}{cn \log n} \sum_{k=1}^m \theta_k \sum_{\nu=1}^K \lambda_\nu \left(\sum_{0 \leq i, j \leq [nt_k]} \phi_\nu(\xi_{S_i}) \phi_\nu(\xi_{S_j}) \mathbf{1}_{\{S_i \neq S_j\}} \right) \\ &= \sum_{\nu=1}^K \lambda_\nu \left(\frac{1}{cn \log n} \sum_{k=1}^m \theta_k \left(\sum_{i=0}^{[nt_k]} \phi_\nu(\xi_{S_i}) \right)^2 - \frac{1}{cn \log n} \sum_{k=1}^m \theta_k \sum_{x \in \mathbb{Z}^2} N_{[nt_k]}^2(x) \phi_\nu^2(\xi_x) \right) \end{aligned}$$

Pour tout $x \in \mathbb{Z}^2$, on définit comme au chapitre précédent, le vecteur aléatoire $\widetilde{\xi}_x^{(K)}$:

$$\widetilde{\xi}_x^{(K)} = (\phi_1(\xi_x), \dots, \phi_K(\xi_x)) \in \mathbb{R}^K.$$

Pour tout $\nu \in \{1, \dots, K\}$, la suite $(\phi_\nu(\xi_x))_{x \in \mathbb{Z}^2}$ est une suite de variables aléatoires indépendantes identiquement distribuées. Par ailleurs, il découle de (51) que, pour tout $x \in \mathbb{Z}^2$, les composantes de $\tilde{\xi}_x^{(K)}$ sont des variables aléatoires décorréées, d'espérance nulle et de variance 1. Nous allons pouvoir appliquer ici, le Théorème 4.2 du chapitre 1 qui est l'équivalent vectoriel du Théorème de Bolthausen [13] : quand $n \rightarrow +\infty$,

$$\left(\frac{1}{cn \log n} \sum_{i=0}^{[nt]} \tilde{\xi}_{S_i}^{(K)} \right)_{t \in [0,1]} \text{ converge faiblement dans } D[0,1], \text{ vers } (B_t^{(1)}, \dots, B_t^{(K)})_{t \in [0,1]},$$

où $(B_t^{(1)}, \dots, B_t^{(K)})_{t \in [0,1]}$ désigne un mouvement brownien standard K -dimensionnel.

De la continuité de l'application,

$$\begin{aligned} L : \mathbb{R}^K &\rightarrow \mathbb{R} \\ (x_1, \dots, x_K) &\mapsto \sum_{\nu=1}^K \lambda_\nu x_\nu^2. \end{aligned}$$

il découle que, quand n tend vers l'infini,

$$(103) \quad \sum_{k=1}^m \theta_k \sum_{\nu=1}^K \lambda_\nu \left(\frac{1}{cn \log n} \left(\sum_{i=0}^{[nt_k]} \phi_\nu(\xi_{S_i}) \right)^2 \right) \text{ converge en loi vers } \sum_{k=1}^m \theta_k \sum_{\nu=1}^K \lambda_\nu (B_{t_k}^{(\nu)})^2.$$

Par ailleurs, par le Lemme 1.1 : quand n tend vers l'infini,

$$(104) \quad \sum_{k=1}^m \theta_k \sum_{\nu=1}^K \lambda_\nu \left(\frac{1}{cn \log n} \sum_{x \in \mathbb{Z}^2} N_{[nt_k]}^2(x) \phi_\nu^2(\xi_x) \right) \xrightarrow{P} \sum_{k=1}^m \theta_k \sum_{\nu=1}^K \lambda_\nu t_k.$$

Reprenons l'écriture de $\widehat{\mathcal{W}}_{n,K}$ et appliquons les résultats (103) et (104). Il vient,

$$\widehat{\mathcal{W}}_{n,K} = \sum_{j=1}^m \theta_j \widehat{\mathcal{U}}_{n,K}(t_j) \text{ converge en loi vers } \sum_{k=1}^m \theta_k \sum_{\nu=1}^K \lambda_\nu ((B_{t_k}^{(\nu)})^2 - t_k), \text{ quand } n \rightarrow \infty.$$

Par conséquent, si l'on désigne par ψ_K et $\psi_{n,K}$ les fonctions caractéristiques respectives de $(U_K(t_1), \dots, U_K(t_m))$ et $(\widehat{\mathcal{U}}_{n,K}(t_1), \dots, \widehat{\mathcal{U}}_{n,K}(t_m))$, à K fixé,

$$(105) \quad |\psi_{n,K}(\theta) - \psi_K(\theta)| \rightarrow 0, \quad \text{quand } n \rightarrow \infty.$$

Étape 2 :

Définissons la suite de processus $(\widehat{R}_{n,K}(t))_{t \in [0,1]}$, $n = 1, 2, \dots$, respectivement $(\widehat{\mathcal{R}}_{n,K}(t))_{t \in [0,1]}$, $n = 1, 2, \dots$, qui correspond à la suite des restes entre les deux processus $(\widehat{U}_n(t))_{t \in [0,1]}$ et $(\widehat{U}_{n,K}(t))_{t \in [0,1]}$, respectivement entre $(\widehat{U}_n(t))_{t \in [0,1]}$ et $(\widehat{U}_{n,K}(t))_{t \in [0,1]}$:

$$\begin{aligned} \widehat{R}_{n,K}(t) &= \widehat{U}_n(t) - \widehat{U}_{n,K}(t), \\ &\text{et} \\ \widehat{\mathcal{R}}_{n,K}(t) &= \frac{\widehat{R}_{n,K}(t)}{cn \log n} \\ &= \widehat{U}_n(t) - \widehat{U}_{n,K}(t). \end{aligned}$$

Nous remarquerons que comme, ici, les processus sont définis en dehors des auto-intersections de la marche aléatoire $(S_n)_{n \in \mathbb{N}}$, $\widehat{R}_{n,K}$ correspond à l'équivalent, dans le cas où $(S_n)_{n \in \mathbb{N}}$ est une marche aléatoire récurrente sur \mathbb{Z}^2 , de $R_{n,K}^{(1)}$ défini dans l'étape 2 du chapitre 4. Comme on l'avait fait pour la suite de processus $(R_{n,K}^{(1)}(t))_{t \in [0,1]}$, on montre, en utilisant le lemme 2.4 du chapitre 3 (cf. [18]) qu'il existe une constante $C > 0$ telle que

$$E(|\mathcal{R}_{n,K}(t)|^2) \leq C \|h - h_K\|_{L^2(\mu \otimes \mu)}^2, \quad \forall n \geq 1 \text{ et } \forall t \in [0, 1].$$

En notant, pour tout entier n , ψ_n la fonction caractéristique de $(\mathcal{U}_n(t_1), \dots, \mathcal{U}_n(t_m))$, on montre alors, exactement de la même façon que dans l'étape 2 du chapitre 4 et en utilisant le résultat ci-dessus, que pour tout $\epsilon > 0$, il existe un $K_0 > 0$ tel que

$$(106) \quad \sup_{n \geq 1} |\psi_n(\theta) - \psi_{n,K}(\theta)| < \epsilon, \quad \forall K > K_0.$$

Étape 3 :

Calculons la moyenne quadratique de $\sum_{k=1}^m \theta_k (U(t_k) - U_K(t_k))$:

$$\begin{aligned} \left(E \left(\left| \sum_{k=1}^m \theta_k (U(t_k) - U_K(t_k)) \right|^2 \right) \right)^{1/2} &\leq \sum_{k=1}^m |\theta_k| \left(E(|U(t_k) - U_K(t_k)|)^2 \right)^{1/2} \\ &\leq \sum_{k=1}^m |\theta_k| \left(\sum_{\nu=K+1}^{\infty} \lambda_{\nu}^2 \right)^{1/2}. \end{aligned}$$

Comme la série $\sum_{\nu \geq 1} \lambda_{\nu}^2$ converge, il suit :

$$\sum_{k=1}^m \theta_k U_K(t_k) \rightarrow \sum_{k=1}^m \theta_k U(t_k), \text{ en moyenne quadratique, quand } K \rightarrow \infty,$$

d'où la convergence en loi de $(U_K(t_1), \dots, U_K(t_m))$ vers $(U(t_1), \dots, U(t_m))$.

Notons ψ la fonction caractéristique du vecteur aléatoire $(U(t_1), \dots, U(t_m))$, alors

$$(107) \quad \lim_{K \rightarrow \infty} |\psi(\theta) - \psi_K(\theta)| = 0.$$

Finalement, on conclut à la convergence des lois de dimension finie de la suite de processus $(\hat{U}_n(t))_{t \in [0,1]}$ vers les lois de dimension finie du processus $(U(t))_{t \in [0,1]}$, en combinant (105), (106) et (107).

□

1.2. Une nouvelle preuve du Théorème 2.2 du chapitre 3.

On écrit comme au chapitre 3 (cf. [18]) :

$$\mathcal{U}_n(t) = \hat{\mathcal{U}}_n(t) + \mathcal{T}_{[nt]},$$

où

$$\mathcal{T}_{[nt]} = \frac{T_{[nt]}}{cn \log n}$$

La preuve de la convergence des lois de dimension finie de la suite de processus $(\mathcal{U}_n(t))_{t \in [0,1]}$, $n = 1, 2, \dots$ vers celles de $(U(t) + I_1 t)_{t \in [0,1]}$ découle directement de la section précédente et du fait que pour tout $t \in [0, 1]$, $\mathcal{T}_{[nt]}$ converge en probabilité vers $I_1 t$, quand n tend vers

l'infini (ce que l'on montre de la même façon que pour le lemme 1.1).

2. Une nouvelle preuve des théorèmes 3.1 et 3.2 du chapitre 3

Dans cette sous-section les U -statistiques sont échantillonnées par $(S_n)_{n \in \mathbb{N}}$, une marche aléatoire transiente sur \mathbb{Z}^m , $m \geq 1$.

Ici encore, nous pouvons adapter la méthode de preuve du chapitre 4, respectivement à la convergence faible dans l'espace $(\mathcal{D}([0, 1]), \mathcal{S})$ de la suite de processus $(\frac{1}{\gamma n} \widehat{U}_{[nt]})_{t \in [0, 1]}$ vers $(U(t))_{t \in [0, 1]}$ (resp. à la convergence faible dans l'espace $(\mathcal{D}([0, 1]), \mathcal{S})$ de la suite de processus $(\frac{1}{\gamma n} U_{[nt]})_{t \in [0, 1]}$ vers $(U(t) + I_1 t)_{t \in [0, 1]}$).

Comme dans la sous-section précédente, seule la partie concernant la convergence des lois de dimension finie change par rapport à la démonstration de Cabus et Guillinot [18] exposée au chapitre 3. Il suffit donc remplacer dans le schéma de la preuve de la sous-section 1.1, le théorème de Bolthausen vectoriel et les résultats sur les temps locaux d'une marche aléatoire, par leurs équivalents dans le cas où $(S_n)_{n \in \mathbb{N}}$ est transiente sur \mathbb{Z}^m , $m \geq 1$ (cf. chapitre 1, section 5), afin de montrer que, si l'on fixe $\theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^m$ et $0 < t_1 < \dots < t_m < 1$,

$$\frac{1}{\gamma n} \sum_{k=1}^m \theta_k \widehat{U}_{[nt_k]} \text{ converge en loi vers } \sum_{k=1}^m \theta_k U(t_k), \quad \text{quand } n \rightarrow \infty.$$

□

Problème ouvert et conjecture

Une possibilité de prolongement de cette étude serait, par exemple, de regarder la distribution asymptotique des U -statistiques échantillonnées par une marche aléatoire dans le cas des noyaux d'ordre $k \geq 3$. Il semblerait que dans le cas des marches aléatoires soit récurrentes dans \mathbb{Z}^2 , soit transientes dans \mathbb{Z}^d , $d \geq 3$, le processus limite de la suite des processus associés aux U -statistiques échantillonnées par ces marches, pris en dehors de leurs auto-intersections, correctement renormalisés, soit le même que dans le cas classique (cf Théorème 3.6 du chapitre 2).

Partie 2

Étude de modèles stochastiques de repliement des protéines

CHAPITRE 6

Algorithmes $(1 + 1)$ et modèles de repliement des protéines

On considère deux versions d’un algorithme d’évolution simple modélisant le repliement d’une protéine, à température nulle : le $(1+1)$ -EA, sur le problème Leading Ones. Dans ce modèle schématique, la structure de la protéine est codée par une chaîne de bits de longueur n . Elle évolue vers sa configuration “native” en empruntant un chemin aléatoire de repliements successifs qui s’effectuent le long d’une suite de points de contacts. Nous étudions le comportement asymptotique du temps d’atteinte, dans le cas moyen, sous deux types de mutations différentes : le scénario “one flip” qui change un unique bit à la fois dans la chaîne de longueur n et le scénario “Bernoulli flip” qui change chaque bit de la chaîne indépendamment les uns des autres avec probabilité c/n . Pour chaque algorithme, nous donnons une loi des grands nombres, un théorème central limite et nous comparons la performance des deux modèles.

1. Introduction

Les algorithmes d’évolution (EA), introduits par J. Holland [40], sont des algorithmes d’optimisation stochastique, basés sur une analogie avec les mécanismes biologiques de l’évolution. Ils sont largement utilisés dans une grande variété de problèmes allant de la génétique des populations, à l’apprentissage en passant par l’optimisation. La tâche de l’EA est d’explorer un espace d’états, E , afin de maximiser une certaine fonction *fitness*, $f : E \rightarrow \mathbb{R}$, encore appelée *fonction d’adaptation*, en faisant évoluer une population d’individus, considérés comme des solutions potentielles au problème d’optimisation, sous l’action répétées d’opérations de mutations et de sélections. Chaque individu est évalué numériquement selon son fitness et les individus de fitness maximum sont recherchés. La dynamique des EA simule, à l’image de ce qui se passe dans les systèmes naturels, la survie des individus les mieux adaptés.

En dépit du grand nombre de résultats heuristiques existants dans ce domaine, les résultats mathématiques décrivant le comportement des EA restent relativement clairsemés. Parmi les exceptions se trouvent les travaux de R. Cerf [20, 19], Y. Rabinovich and A. Wigderson [56], G. Rudolph [63], C. Mazza and D. Piau [49], P. Del Moral and A. Guionnet [26], J. Bérard [5] and J. Bérard and A. Bienvenüe [6, 7].

D’autre part, en raison de la dynamique généralement complexe de ces algorithmes, les résultats de complexité sont relativement difficiles à obtenir et une approche commune

consiste à considérer des cas simplifiés. Parmi les EA les plus simples se trouvent les algorithmes $(1 + 1)$ -EA. Ils ont été étudiés par H. Muhlenbein [51], T. Bäck [2], G. Rudolph [63], J. Garnier et al. [34], and S. Droste et al. [30, 31].

Nous allons étudier ici, un algorithme $(1 + 1)$ spécifique, le $(1 + 1)$ -EA sur le problème LeadingOnes, qui permet de modéliser la dynamique du repliement d’une protéine vers sa structure tridimensionnelle active, ou état natif. Il s’agit là, de la version à température nulle, d’un algorithme décrit par les biophysiciens A. Bakk et al. [3], basé sur l’un des principaux modèles de repliement utilisés en biophysique, le modèle “fermeture éclair” ou “zipper model” (voir par exemple [4] pour une revue détaillée des différentes théories existantes). Nous nous intéresserons, plus particulièrement, à la vitesse de convergence de deux versions de cet algorithme.

1.1. Le modèle physique.

Les protéines se replient en une unique conformation *native*, ou biologiquement active, selon des échelles de temps variant de 10^{-3} s. à 1s. Or, si la dynamique du processus de repliement suivait une simple recherche aléatoire dans l’espace des conformations, cela impliquerait des échelles de temps astronomiques. Ce paradoxe est connu sous le nom de paradoxe de Levinthal [47]. Comment les protéines se replient-elles vers leur structure native? Voilà l’une des questions les plus intrigantes de la biophysique. Selon Anfinsen [1], l’état natif est aussi bien génétiquement que thermodynamiquement contrôlé, c.à.d. qu’il correspond à la configuration, dans laquelle l’énergie libre de Gibbs du système est la plus basse. En suggérant l’existence d’un contrôle cinétique, Levinthal propose la formation simultanée de petits noyaux structurés à plusieurs endroits de la chaîne polypeptidique qui initieraient et accéléreraient le repliement.

Il existe de nombreux points de vue concernant l’état de transition (TS). L’une des hypothèses est que la dynamique de (TS) consiste en un chemin aléatoire, conduisant la protéine vers son état natif, le long d’une descente guidée du paysage d’énergie libre (voir J. A. Schellman [64], K. A. Dill et al. [29]).

Le modèle de repliement de A. Bakk et al. [3] peut être décrit comme suit. La chaîne polypeptidique est équipée de n points de contacts c_1, \dots, c_n , appelés aussi noyaux, qui divisent la protéine en n sous-structures, ou sous-unités, distinctes. Pour i allant de 1 à

n , on attribue à c_i une variable de contact binaire, ϕ_i , qui indique si la sous structure se trouvant au niveau du i -ème point de contact, c_i , est pliée ($\phi_i = 1$) ou dépliée ($\phi_i = 0$). Par conséquent, la configuration de la protéine est entièrement modélisée par la chaîne de bits de longueur n , $\phi = (\phi_1, \phi_2, \dots, \phi_n)$. L'état natif correspond à la chaîne pour laquelle $\phi_i = 1$, pour tout $1 \leq i \leq n$, c.à.d. $(1, \dots, 1)$. Ainsi, il existe une bijection entre l'espace des configurations et l'ensemble $\{0, 1\}^n$.

Notons i_0 le plus petit indice $i \in \{1, \dots, n\}$ pour lequel c_i est déplié, c.à.d. pour lequel $\phi_i = 0$. On appellera partie ouverte de la protéine l'ensemble des points de contact $\{c_{i_0}, c_{i_0+1}, \dots, c_n\}$.

L'hypothèse concernant la dynamique du processus de repliement est la suivante : à chaque noyau est associé une énergie de $-\varepsilon_0$ si $i < i_0$, zéro sinon. Elle peut être implémenté à travers l'hamiltonien

$$\mathcal{H} = -\varepsilon_0(\phi_1 + \phi_1\phi_2 + \dots + \phi_1\dots\phi_n),$$

que l'on peut également réécrire à l'aide de la fonction LeadingOnes, L , définie sur l'espace des configurations $\{0, 1\}^n$, qui compte la longueur du plus long préfixe constitué uniquement de uns dans la chaîne de bits :

$$L(x) = \max\{k \geq 1 : \forall 1 \leq i \leq k, x_i = 1\} \cup \{0\}.$$

En effet,

$$(108) \quad \mathcal{H} = -\varepsilon_0 L(\phi) = -\varepsilon_0(i_0 - 1).$$

Ce modèle, dans lequel il n'y a pas d'énergie libre associée à la partie ouverte de la protéine, est également connu sous le nom de "modèle fermeture éclair" ou "zipper model". Remarquons qu'une descente le long du paysage d'énergie libre signifie à la fois le repliement de la sous-structure la plus à gauche qui n'est pas encore correctement repliée, c_{i_0} , et à la fois le "statu-quo" pour les sous-structures correctement repliées qui la précèdent (situées à sa gauche), c.à.d. les noyaux c_i pour $i < i_0$. Les étapes successives du processus de repliement s'effectuent dans un ordre bien spécifique et la dynamique de ce processus s'apparente à celle d'une fermeture éclair. D'autre part, si l'on regarde ce phénomène du point de vue de la chaîne de bit, baisser l'énergie libre de Gibbs revient exactement à augmenter la taille du plus long préfixe de uns.

L'algorithme proposé par A. Bakk et al. [3] afin d'explorer l'espace d'états $\{0, 1\}^n$ en vue de trouver les configurations d'énergie minimum, c.à.d. l'état natif, est basé sur la méthode de Monte Carlo Métropolis par chaîne de Markov (MCM) (cf. Binder [12]). Notons T la température du système, k le facteur usuel de Boltzmann et posons $\beta = 1/kT$. A tout instant k , l'état de la protéine est représenté par l'élément de $\{0, 1\}^n$, X_k . L'algorithme procède itérativement comme suit :

1. Mutation : $X_k \rightarrow X'_k$

La chaîne de bits, X_k , présente au temps k , effectue un pas d'une marche aléatoire sur $\{0, 1\}^n$ pour former une nouvelle configuration X'_k .

2. Sélection : $X'_k \rightarrow X_{k+1}$

La configuration, X'_k , est proposée pour la sélection : elle est acceptée pour former l'état de la protéine au temps $k + 1$, X_{k+1} , avec probabilité

$$P_{\text{accept}} = \min(1, \exp(-\beta\Delta\mathcal{H})), \text{ où}$$
$$\Delta\mathcal{H} = \mathcal{H}(X'_k) - \mathcal{H}(X_k),$$

et dans le cas où X'_k est rejetée, la configuration de la protéine reste la même : $X_{k+1} = X_k$.

La température T , à travers le facteur β , contrôle l'intensité de la sélection. Plus la température est basse (plus β est élevé), plus les configurations ayant de grandes valeurs de L sont favorisés par l'étape de sélection.

Nous nous concentrerons essentiellement sur le modèle MCM, à température $T = 0$, que nous noterons MCM₀. Cet algorithme est directement relié à un algorithme d'évolution spécifique qui appartient à la classe des *Stratégies d'évolution* (1 + 1) : le (1 + 1)-EA, encore appelé (1 + 1)-ES, sur le problème LeadingOnes.

1.2. L'approche (1 + 1)-EA.

Dans le modèle de repliement des protéines que nous considérons, minimiser l'hamiltonien \mathcal{H} , revient à maximiser la fonction LeadingOnes. De manière générale, le rôle d'un (1 + 1)-EA est d'optimiser une certaine fonction fitness, $f : \{0, 1\}^n \rightarrow \mathbb{R}$, en faisant évoluer une population réduite à un unique individu (ou chaîne de bits), sous l'action répétée de mutations et de sélections. La dynamique de ces algorithmes d'évolution (1 + 1), ou (1 + 1)-EA, peut-être formalisée à travers une chaîne de Markov à temps discret, de la façon suivante :

Nous nous plaçons ici, dans le *cas moyen* (“*mean case*”), dans lequel le premier individu, X_0 , est choisi uniformément au hasard dans $\{0, 1\}^n$.

1. Mutation : $X_k \rightarrow X'_k$

Comme dans la méthode de Monte Carlo Métropolis, à chaque étape, encore appelée génération, l'individu de la population au temps k , X_k , effectue un pas d'une marche aléatoire sur $\{0, 1\}^n$. Le nouvel individu ainsi créé est noté X'_k .

2. Sélection (version 1): $X'_k \rightarrow X_{k+1}$

Les deux individus X_k et X'_k proposés pour la sélection sont ensuite évalués en fonction de leur fitness. C'est l'individu ayant le plus grand fitness qui est ensuite sélectionné pour former la génération au temps $k + 1$, X_{k+1} :

$$(109) \quad X_{k+1} = \begin{cases} X'_k & \text{si } f(X'_k) > f(X_k), \\ X_k & \text{sinon.} \end{cases}$$

La sélection est déterministe.

Dans la littérature, il n'y a pas de véritable consensus, dans la définition des $(1 + 1)$ -EA, en ce qui concerne l'étape de sélection. Il arrive qu'elle soit remplacée par cette seconde version, légèrement différente :

2 bis. Sélection (version 2) : $X'_k \rightarrow X_{k+1}$

$$(110) \quad X_{k+1} = \begin{cases} X'_k & \text{si } f(X'_k) \geq f(X_k), \\ X_k & \text{sinon.} \end{cases}$$

Par exemple, Garnier et al. [34] considèrent la première version de la loi de sélection (109), alors que Droste et al. [30], [31] se concentrent sur la seconde version (110).

Quand le fitness est précisément la fonction LeadingOnes, L , on appellera cet algorithme, le $(1 + 1)_L$ -EA. Remarquons que dans le cadre LeadingOnes, la variante du $(1 + 1)_L$ -EA, dont le processus de sélection est défini par (110) n'est autre que MCM_0 . Afin de différencier ces deux algorithmes, nous garderons le nom MCM_0 , pour la version 2. Leur seule différence réside dans le fait que la procédure de sélection de MCM_0 accepte les candidats X'_k dont l'énergie est la même que celle de X_k , alors ces mêmes candidats sont refusés par $(1 + 1)_L$ -EA.

Remarques :

- La notation $(1 + 1)$ -EA rend compte du fait que l'on sélectionne parmi un parent et un enfant.
- On notera que dans le cas où le paysage de fitness présente des minima locaux la méthode $(1 + 1)$ -EA pourrait aboutir en une recherche sous-optimale, la chaîne de Markov restant prise au piège de minima locaux. Cependant, dans le cas de la fonction fitness LeadingOnes, nous sommes à l'abri de ce genre de problème.

1.3. Résultats.

Désignons par T_n (respectivement par \widehat{T}_n) le temps mis par le $(1 + 1)$ -EA (resp. par MCM_0) pour atteindre la configuration optimale (en terme de fitness score). Nous allons nous intéresser aux algorithmes MCM_0 et $(1 + 1)_L$ -EA dans le *cas moyen*, sous deux types de mutations : le cas one flip qui change un unique bit à la fois, choisi uniformément au hasard dans la chaîne de bits, et le Bernoulli flip, qui change chaque bit de la chaîne indépendamment avec la probabilité c/n . Comme nous l'avons brièvement rappelé dans l'introduction, la complexité de certains $(1 + 1)$ -EA a déjà été étudiée :

D'après Droste, Jansen et Wegener [30], $E(\widehat{T}_n) = \Theta(n \ln n)$ pour le Bernoulli flip dans le cas où la fonction fitness est une fonction linéaire sur $\{0, 1\}^n$. D'autre part, Droste et al. [31] ont montré, toujours dans le scénario Bernoulli flip, que la fonction LeadingOnes, est résoluble en temps moyen $\Theta(n^2)$.

Le cas où la fonction fitness correspond à la fonction OneMax $|\cdot|$, qui compte le nombre de uns dans la chaîne de bits,

$$|x| = \sum_{i=1}^n x_i,$$

a été étudié par Garnier et al. [34], à la fois dans le scénario one flip et dans le scénario Bernoulli flip. Dans le cas one flip, $(T_n - n \ln n)/n$ converge en distribution vers $-\ln 2 - \ln Z$. Dans le cas Bernoulli flip, $(T_n - c^{-1}e^c n \ln n)/n$ converge en distribution vers $-c^{-1}e^c \ln Z + C(c)$, où la loi de Z est exponentielle de paramètre 1 et $C(c)$ est une constante qui dépend de c .

Nous prouvons ici l'analogie pour le problème LeadingOnes du résultat de Garnier et al. [34]. Nous améliorons ainsi le résultat de Droste et al. [31]. Plus précisément, nous prouvons une loi des grands nombres, un théorème central limite, et nous comparons la performance des deux modèles. Pour finir, nous montrons que, dans les deux scénari (one et Bernoulli flip), les temps d'atteinte de MCM_0 et de $(1 + 1)_L$ -EA sont égaux en loi.

Théorème 1.1 (Cas one flip). (i) Pour $n \geq 1$, $E(T_n) = n^2/2$;

(ii) Quand $n \rightarrow \infty$, $T_n/E(T_n)$ converge en probabilité vers 1.

(iii) Quand $n \rightarrow \infty$, $(T_n - E(T_n))/n^{3/2}$ converge en distribution vers une variable aléatoire gaussienne, centrée, de variance $3/4$.

Théorème 1.2 (Cas Bernoulli flip). *i) Quand $n \rightarrow \infty$, $E(T_n) \sim m(c) n^2$, avec*

$$m(c) := (e^c - 1)/(2c^2).$$

ii) Quand $n \rightarrow \infty$, $T_n/E(T_n)$ converge en probabilité vers 1.

iii) De plus, $(T_n - m(c) n^2)/n^{3/2}$ converge en distribution vers une variable aléatoire gaussienne, centrée, de variance $\sigma^2(c)$, avec

$$\sigma^2(c) := 3(e^{2c} - 1)/(8c^3).$$

Notons que $m(c) > 1/2$ pour tout $c > 0$.

Corollaire 1.1. *Quand $n \rightarrow \infty$, $E(T_n)$ dans le cas Bernoulli flip est supérieur à $E(T_n)$ dans le cas one flip, pour n'importe quelle valeur de c .*

Théorème 1.3. *Pour tout $n \in \mathbb{N}$, dans le cas one flip comme dans le cas Bernoulli flip, \widehat{T}_n possède la même distribution que T_n .*

2. Preuve du Théorème 1.1

La loi de T_n , conditionnée par $|X_0|$, est la loi d'une somme de variables aléatoires géométriques, voir le Lemme 2.1. Ceci entraîne la partie (i) du Théorème. Comme le TCL implique la loi des grands nombres, nous prouverons donc le TCL de la partie (iii).

La distribution de T_n découle d'une simple observation. Dans le $(1 + 1)$ -EA associé au problème LeadingOnes, une mutation est acceptée si et seulement si elle ajoute 1 au nombre de premiers uns. Par conséquent, dans le cas one flip, la chaîne effectue un saut quand le zéro le plus à gauche est changé en un. Tous les autres changements, laissent la chaîne telle quelle. Ainsi, les zéros dans X_0 sont successivement changés, de gauche à droite, jusqu'à ce que l'on ait atteint l'individu (la configuration) optimal(e), $(1, 1, \dots, 1)$.

Ici, comme tout au long de cette section, on posera $\varepsilon = 1/n$ quand nous aurons à faire à des algorithmes de longueur n , $|x|$ désignera le nombre de uns dans x , la loi géométrique $\mathcal{G}(p)$ de paramètre p sera définie par

$$\mathcal{G}(p) = \sum_{n \geq 1} p(1-p)^{n-1} \delta_n,$$

et la loi binomiale négative de paramètre (k_0, p) , $\mathcal{NB}(k_0, p)$, attribuera le poids suivant à $k \geq k_0$:

$$\binom{k-1}{k-k_0} p^{k_0} (1-p)^{k-k_0}.$$

Lemme 2.1. *Si $|X_0| = n - k_0$, T_n est la somme de k_0 variables aléatoires i.i.d. $\mathcal{G}(\varepsilon)$. Ainsi, la loi de T_n est binomiale négative de paramètre (k_0, ε) .*

Preuve du Lemme 2.1 : Soit $\tau_0 = 0$ et, pour tout $k \geq 0$,

$$(111) \quad \tau_{k+1} = \inf\{i \geq \tau_k; X_i \neq X_{\tau_k}\}, \quad \sigma_{k+1} = \tau_{k+1} - \tau_k, \quad \tilde{X}_k = X_{\tau_k}.$$

En d'autres termes, \tilde{X}_k désigne la position de la chaîne après son k -ième saut, et $|\tilde{X}_k| = n - k_0 + k$. Le zéro le plus à gauche de \tilde{X}_k est changé en un après un séjour en \tilde{X}_k de durée σ_{k+1} . Ainsi, $(\sigma_k)_k$ est i.i.d. de loi $\mathcal{G}(\varepsilon)$. Il reste à remarquer que

$$T_n = \sigma_1 + \cdots + \sigma_{k_0}.$$

□

Preuve de la partie (iii) : On notera E_k le conditionnement par rapport à $\{|X_0| = n - k\}$. Comme X_0 est uniforme, la loi μ de $|X_0|$ est binomiale $(n, 1/2)$. D'après le Lemme 2.1, sous P_k , T_n est la somme de k variables aléatoires géométriques i.i.d. de paramètre ε . Ainsi,

$$(112) \quad E_k(e^{-\alpha T_n}) = e^{-\alpha k} \varepsilon^k [1 - (1 - \varepsilon) e^{-\alpha}]^{-k}.$$

Posons $\Theta_n = (T_n - n^2/2)/n^{3/2}$. La décomposition de la transformée de Laplace de Θ_n le long des valeurs de $|X_0|$, la forme explicite de μ , et l'équation (112) réunies, entraînent que

$$E(e^{-\alpha \Theta_n}) = e^{\sqrt{n}\alpha/2} 2^{-n} \sum_{k=0}^n \binom{n}{k} e^{-\alpha k/n^{3/2}} \varepsilon^k [1 - (1 - \varepsilon) e^{-\alpha/n^{3/2}}]^{-k}.$$

Ce que l'on peut ré-écrire,

$$E(e^{-\alpha \Theta_n}) = e^{\sqrt{n}\alpha/2} 2^{-n} (1 + \beta_n)^n,$$

avec

$$\beta_n = \varepsilon e^{-\alpha/n^{3/2}} [1 - (1 - \varepsilon) e^{-\alpha/n^{3/2}}]^{-1}.$$

Rappelons que $\varepsilon = 1/n$. Le développement limité de β_n donne

$$\beta_n = 1 - \alpha/\sqrt{n} + \alpha^2/n + o(1/n).$$

Ce qui implique que

$$E(e^{-\alpha \Theta_n}) \rightarrow e^{3\alpha^2/8},$$

quand $n \rightarrow \infty$. Cela conclut la preuve. \square

3. Preuve du Théorème 1.2

Nous allons d'abord décrire la loi de T_n conditionnellement aux valeurs prises par L le long du chemin emprunté par $(X_k)_{k \in \mathbb{N}}$ jusqu'au temps T_n , c'est-à-dire jusqu'à ce que l'individu optimal $(1, 1, \dots, 1)$ soit atteint. Nous allons voir que cette loi est la loi d'une somme de variables aléatoires géométrique indépendantes (voir le Lemme 3.3). Nous en déduirons la loi de T_n , voir Proposition 3.1. Ceci entraînera la partie (i) du théorème.

Pour les mêmes raisons que dans la section précédente, étant donné que le TCL implique la loi des grands nombres, nous prouverons donc le TCL de la partie (iii).

Rappelons que le $(1 + 1)$ EA, dans le contexte LeadingOnes, accepte une mutation si et seulement si le nombre de premiers uns est augmenté. Ainsi la dynamique de l'algorithme Bernoulli flip procède comme suit : la chaîne effectue un saut pour former un nouvel individu, au temps $k + 1$, si et seulement si les premiers uns de X_k restent inchangés en même temps que le zéro, le plus à gauche, est changé en un, quelles que soient les valeurs prises par les autres bits.

Ici, comme tout au long de cette section, on prendra $\varepsilon = c/n$ quand on aura à faire à des algorithmes sur des chaînes de longueur n . Pour tout $i \geq 0$, notons $p(n, i) = \varepsilon(1 - \varepsilon)^i$. Comme dans le contexte one flip, \tilde{X}_k désignera la position de la chaîne après son k -ième saut. Nous garderons également les mêmes définitions pour σ_k et τ_k .

Pour tout $k \geq 0$, posons

$$\ell_k = L(\tilde{X}_k).$$

Soit Y_0 tel que

$$X_0 = (1^{\ell_0}, 0, Y_0).$$

Pour tout $k \geq 1$, définissons Y_k et W_k par

$$\tilde{X}_k = (1^{\ell_{k-1}}, 1, W_k) = (1^{\ell_k}, 0, Y_k).$$

Afin de calculer la loi de T_n conditionnellement à (ℓ_j) dans le Lemme 3.3, nous aurons besoin des Lemmes 3.1 et 3.2 ci-dessous

Lemme 3.1. (i) Pour tout $k \geq 0$, σ_k dépend du passé seulement à travers le dernier score de ℓ_{k-1} . En d'autres termes, la loi de σ_k , conditionnellement à $(X_t)_{\{t < \tau_k\}}$, est la loi de

σ_k , conditionnellement à ℓ_{k-1} .

(ii) Pour tout $k \geq 0$, sachant $\{\ell_{k-1} = i\}$, la loi de σ_k est $\mathcal{G}(p(n, i))$.

Preuve : Le temps de séjour σ_k est le temps mis par l'algorithme pour sauter de \tilde{X}_{k-1} à \tilde{X}_k . Comme le zéro le plus à gauche de \tilde{X}_{k-1} est en position $\ell_{k-1} + 1$, on a besoin de changer le $(\ell_{k-1} + 1)$ -ième bit, tout en laissant les ℓ_{k-1} premiers bits tels quels. Ainsi,

$$P(\sigma_k = t | \tilde{X}_0, \dots, \tilde{X}_{k-1}) = \varepsilon(1 - \varepsilon)^{\ell_{k-1}} [1 - \varepsilon(1 - \varepsilon)^{\ell_{k-1}}]^{(t-1)}.$$

□

Lemme 3.2. Pour tout $k \geq 1$ soit $\hat{\mathcal{F}}_k = \sigma\{\sigma_i, \ell_j : i \leq k, j \leq k-1\}$, alors la loi de ℓ_k conditionnellement à $\hat{\mathcal{F}}_k$ est la loi de ℓ_k conditionnellement à ℓ_{k-1} .

Preuve : Notons P_i la probabilité sachant $\{\ell_0 = i\}$. D'après la propriété de Markov :
(113)

$$P(\ell_k = i_k | \sigma_k = t_k, \ell_{k-1} = i_{k-1}, \dots, \sigma_1 = i_1, \ell_0 = i_0) = P(\ell_1 = i_k | \sigma_1 = t_k, \ell_0 = i_{k-1})$$

Comme $\{\ell_1 = i_1, \sigma_1 = t, \ell_0 = i_0\} = \{L(X_t) = i_1, L(X_{t-1}) = \dots = L(X_0) = i_0\}$, comme $\{\sigma_1 = t, \ell_0 = i_0\} = \{L(X_t) \neq i_0, L(X_{t-1}) = \dots = L(X_0) = i_0\}$ et en utilisant la propriété de Markov sur $(L(X_t))_t$ on en déduit l'expression suivante :

$$P(\ell_1 = i_k | \sigma_1 = t_k, \ell_0 = i_{k-1}) = \frac{P(L(X_1) = i_k | L(X_0) = i_{k-1})}{P(L(X_1) \neq i_{k-1} | L(X_0) = i_{k-1})}$$

Cette quantité est indépendante de t_k , ainsi si on reconsidère (113):

$$P(\ell_k = i_k | \sigma_k = t_k, \ell_{k-1} = i_{k-1}, \dots, \sigma_1 = i_1, \ell_0 = i_0) = P(\ell_k = i_k | \ell_{k-1} = i_{k-1})$$

□

Lemme 3.3. Conditionnellement à $\{\ell_0 = i_0, \dots, \ell_{J-1} = i_{J-1}, \ell_J = n\}$, T_n est la somme de J variables aléatoires géométriques indépendantes de paramètres respectifs $p(n, i_0), \dots, p(n, i_{J-1})$.

Preuve : Sachant les différentes valeurs successives prises par la fonction LeadingOnes $\{\ell_0 = i_0, \dots, \ell_J = n\}$ jusqu'à l'obtention de l'individu optimal, $T_n = \sum_{k=1}^J \sigma_k$. Ainsi,
(114)

$$P(T_n = t | \ell_J = n, \dots, \ell_0 = i_0) = \sum_{t_1 + \dots + t_k = t} P(\sigma_J = t_J, \dots, \sigma_1 = t_1 | \ell_J = n, \dots, \ell_0 = i_0)$$

En utilisant les Lemmes (3.2) et (3.1) on peut en déduire par récurrence que :

$$P(\ell_J = n, \sigma_J = t_J, \dots, \sigma_1 = t_1, \ell_0 = i_0) = \prod_{k=1}^J P(\ell_k = i_k | \ell_{k-1} = i_{k-1}) P(\sigma_k = t_k | \ell_{k-1} = i_{k-1})$$

Par conséquent, comme $\prod_{k=1}^J P(\ell_k = i_k | \ell_{k-1} = i_{k-1}) = P(\ell_J = n, \dots, \ell_0 = i_0)$,

$$P(\sigma_1 = t_1, \dots, \sigma_J = t_J | \ell_J = n, \dots, \ell_0 = i_0) = \prod_{k=1}^J P(\sigma_k = t_k | \ell_{k-1} = i_{k-1})$$

Nous pouvons maintenant remplacer cette dernière équation dans (114),

$$P(T_n = t | \ell_J = n, \dots, \ell_0 = i_0) = \sum_{t_1 + \dots + t_k = t} \prod_{k=1}^J P(\sigma_k = t_k | \ell_{k-1} = i_{k-1})$$

Soit $q(n, i_k)$ la loi de probabilité de $\mathcal{G}(p(n, i_k))$.

Alors d'après le Lemme 3.1, nous pouvons écrire, en reconnaissant un produit de convolution que :

$$P(T_n = t | \ell_J = n, \dots, \ell_0 = i_0) = q(n, i_{J-1}) * \dots * q(n, i_0)(t)$$

Ainsi, sachant que notre exploration va sauter J fois avant d'atteindre sa cible $(1, 1, \dots, 1)$ et sachant $\{\ell_0 = i_0, \dots, \ell_J = n\}$, T_n suit la même loi que la somme de J variables aléatoires indépendantes $\mathcal{G}(p(n, i_0)), \dots, \mathcal{G}(p(n, i_{J-1}))$.

Concentrons-nous maintenant sur $P(\ell_0 = i_0, \dots, \ell_{J-1} = i_{J-1}, \ell_J = n)$. Afin de calculer cette quantité, nous aurons besoin du Lemme suivant :

Lemme 3.4. *Soit $k \geq 1$. Si X_0 est choisi uniformément au hasard dans $\{0, 1\}^n$, alors, sachant $\{\ell_{k-1} = i\}$, W_k suit la loi uniforme $\{0, 1\}^{n-i-1}$:*

$$\mathcal{L}(W_k | \ell_{k-1} = i) = \mathcal{U}(\{0, 1\}^{n-i-1}).$$

Preuve : Intéressons-nous d'abord au cas où $k = 1$. Soit P_i la probabilité conditionnelle à $\{\ell_0 = i\}$. Soit μ la loi de probabilité de Y_0 sachant $\{\ell_0 = i\}$. Comme X_0 est choisi uniformément au hasard dans $\{0, 1\}^n$, μ est la loi uniforme sur $\{0, 1\}^{n-i-1}$.

$$(115) \quad P_i(W_1 = w) = \sum_{t \geq 1} P_i(\tilde{X}_1 = (1^i, 1, w), \sigma_1 = t)$$

Comme $\{X_{\sigma_1} = (1^i, 1, w), \sigma_1 = t, L(X_0) = i\} = \{X_t = (1^i, 1, w), L(X_{t-1}) = \dots = L(X_0) = i\}$ et par la propriété de Markov, l'équation (115) peut-être ré-écrite :

$$P_i(W_1 = w) = \sum_{t \geq 1} P_i(X_1 = (1^i, 1, w)) P_i(L(X_1) = L(X_0))^{t-1}.$$

D'où,

$$(116) \quad P_i(W_1 = w) = \frac{P_i(X_1 = (1^i, 1, w))}{P_i(L(X_1) > L(X_0))}.$$

Rappelons que la chaîne de Markov saute de X_0 vers une configuration ayant une plus grande valeur de fitness, au temps 1, si, à la fois, aucun des ℓ_0 premiers uns de X_0 ne change et si à la fois, le zéro de X_0 le plus à gauche est changé en un. Ainsi,

$$(117) \quad P_i(L(X_1) > L(X_0)) = \varepsilon(1 - \varepsilon)^i$$

D'un autre côté, comme $P_i(Y_0 = u) = \mu(u) = 1/2^{n-i-1}$,

$$(118) \quad P_i(X_1 = (1^i, 1, w)) = 1/2^{n-i-1} \sum_{u \in \{0,1\}^{n-i-1}} P(X_1 = (1^i, 1, w) | X_0 = (1^i, 0, u))$$

Désignons par $d(w, u)$ la distance de Hamming entre w et u , il vient alors

$$(119) \quad P(X_1 = (1^i, 1, w) | X_0 = (1^i, 0, u)) = \varepsilon(1 - \varepsilon)^i \varepsilon^{d(w,u)} (1 - \varepsilon)^{n-i-1-d(w,u)}$$

Pour finir les équations (116), (117), (118) et (119) ainsi que

$$\sum_{u \in \{0,1\}^{n-i-1}} \varepsilon^{d(w,u)} (1 - \varepsilon)^{n-i-1-d(w,u)} = 1,$$

entraînent que :

$$(120) \quad P_i(W_1 = w) = 1/2^{n-i-1}$$

Nous pouvons maintenant, en utilisant la propriété de Markov forte, en déduire la preuve pour tout $k \geq 2$. □

Lemme 3.5. *Si X_0 est choisi uniformément au hasard dans l'espace d'état, pour tout $k \geq 1$, la probabilité conditionnelle de ℓ_k , sachant $\{\ell_{k-1} = i_{k-1}\}$, satisfait:*

$$\begin{aligned} P(\ell_k = j_k | \ell_{k-1} = j_{k-1}) &= 2^{-(j_k - j_{k-1})} & \text{si } j_{k-1} + 1 \leq j_k < n \\ &= 2^{-(n - j_{k-1} - 1)} & \text{si } j_k = n \end{aligned}$$

Preuve : Ceci est une conséquence directe du Lemme 3.4. □

Maintenant que nous connaissons les probabilités de transition de la suite des valeurs successives prises par la fonction `LeadingOnes` jusqu'à ce que l'individu cible $(1, 1, \dots, 1)$ soit atteint, ainsi que la loi de T_n conditionnellement à ces valeurs, nous pouvons calculer la loi de probabilité de T_n :

Proposition 3.1. *Si X_0 est choisi uniformément au hasard dans $\{0, 1\}^n$, alors, la loi de probabilité de T_n vérifie :*

$$(121) \quad P(T_n = t) = \frac{1}{2^n} \sum_J \sum_{\{0 \leq i_0 \leq i_1 \leq \dots \leq i_J = n\}} q(n, i_{J-1}) * \dots * q(n, i_0)(t)$$

Preuve : X_0 étant choisi uniformément au hasard dans l'espace d'état, $P(\ell_0 = i_0) = \frac{1}{2^{i_0+1}}$. Ainsi, par le Lemme 3.5,

$$P(\ell_0 = i_0, \dots, \ell_J = n) = \frac{1}{2^n}$$

Le résultat est alors une conséquence directe du Lemme 3.3. □

Preuve de la partie (iii) : Posons $\Theta_n = (T_n - \frac{n^2(e^c-1)}{2c^2})|n^{3/2}$.

Alors,

$$(122) \quad E(\exp(-\alpha\Theta_n)) = \exp\left(\alpha\frac{\sqrt{n}}{2c^2}(e^c - 1)\right) E\left(\exp(-\alpha\frac{T_n}{n^{3/2}})\right)$$

En vertu de (121), qui donne la distribution de T_n :

$$E\left(\exp(-\alpha\frac{T_n}{n^{3/2}})\right) = \frac{1}{2^n} \sum_J \sum_{i_0 < \dots < i_{J-1}} E\left(\exp(-\frac{\alpha}{n^{3/2}}(\mathcal{G}(p(n, i_0)) + \dots + \mathcal{G}(p(n, i_{J-1}))))\right)$$

Par l'indépendance des variables aléatoires $(\mathcal{G}(p(n, i_k)))_{i_k}$,

$$\begin{aligned} E\left(\exp(-\alpha\frac{T_n}{n^{3/2}})\right) &= \frac{1}{2^n} \sum_J \sum_{i_0 < \dots < i_{J-1}} \prod_{k=0}^{J-1} E\left(-\frac{\alpha}{n^{3/2}}\mathcal{G}(p(n, i_k))\right) \\ &= \frac{1}{2^n} \prod_{k=0}^{n-1} \left(1 + \phi_k\left(\frac{\alpha}{n^{3/2}}\right)\right) \end{aligned}$$

où $\phi_k(\alpha)$ désigne le transformée de Laplace de $\mathcal{G}(p(n, k))$.

$$\phi_k(\alpha) = \frac{e^{-\alpha}p(n, k)}{1 - (1 - p(n, k))e^{-\alpha}}.$$

En se souvenant que $p(n, i) = \varepsilon(1 - \varepsilon)^i$ et $\varepsilon = c/n$, on déduit que :

$$\prod_{k=0}^{n-1} \left(1 + \phi_k\left(\frac{\alpha}{n^{3/2}}\right)\right) \simeq_{n \rightarrow +\infty} 2^n \exp\left(-\alpha \frac{\sqrt{n}}{2c^2}(e^c - 1)\right) \exp\left(\frac{3\alpha^2(e^{2c} - 1)}{8c^3}\right)$$

Ainsi, en remplaçant ceci dans (122), quand n tend vers $+\infty$:

$$E(\exp(-\alpha\Theta_n)) \simeq \exp\left(\frac{3\alpha^2(e^{2c} - 1)}{8c^3}\right)$$

Nous reconnaissons ici, la transformée de Laplace d'une variable aléatoire gaussienne, centrée de variance $\frac{3(e^{2c} - 1)}{8c^3}$. Ce qui conclut la preuve.

4. Preuve du Théorème 1.3

Dans le contexte MCM_0 , à chaque étape de son exploration, l'algorithme a la possibilité de visiter plusieurs configurations distinctes de la protéine ayant toutes même valeur de fitness, avant de véritablement sauter (en fitness) vers une nouvelle configuration de fitness strictement supérieur. Ce qui n'était pas autorisé dans le contexte $(1 + 1)_L\text{-EA}$, où la configuration de la protéine si elle ne restait pas sur place, ne pouvait qu'effectuer des sauts vers des configurations de fitness strictement supérieur.

Comme dans la section précédente, on désignera par $(\tilde{X}_k)_{k \in \mathbb{N}}$ la chaîne définie par les configurations de la protéine prises aux instants de sauts $(\tau_k)_k$. Nous allons succinctement décrire la preuve du Théorème 1.3.

4.1. Preuve du cas one flip.

Ici, comme dans la section 1, nous poserons à nouveau $\varepsilon = 1/n$. Dans le contexte $(1 + 1)_L\text{-EA}$, une fois que l'on connaît, configuration initiale, X_0 , le chemin $(\tilde{X}_k)_{k \in \mathbb{N}}$ devient totalement déterministe : le descendre signifie alors exactement changer, un par un, de gauche à droite, les zéros en uns. Ce n'est plus le cas dans le contexte MCM_0 et nous ne pouvons pas directement déduire la preuve en adaptant directement celle du Théorème 1.1.

Cependant, l'idée de la démonstration est très proche de celle du Théorème 1.2. Considérons d'abord la loi de \hat{T}_n conditionnement aux valeurs prises par L le long du chemin

$(X_k)_{k \in \mathbb{N}}$ jusqu'à ce que l'algorithme ait atteint l'état natif de la protéine $(1, 1, \dots, 1)$. Les mêmes arguments simples s'appliquent. Le Lemme 3.2 reste vrai et nous montrons que, comme dans le cas one flip dans le cadre $(1+1)_L$ -EA, $(\sigma)_k$ est une suite de variables aléatoires i.i.d. de loi $\mathcal{G}(\varepsilon)$. Nous pouvons maintenant aisément adapter le Lemme 3.3 :

Lemme 4.1. *Conditionnellement à $\{\ell_0 = i_0, \dots, \ell_{J-1} = i_{J-1}, \ell_J = n\}$, \widehat{T}_n est la somme de J variables aléatoires géométriques i.i.d. de paramètre ε , i.e. \widehat{T}_n est binomiale négative de paramètre (J, ε) .*

Le Lemme 3.4 est toujours vérifié :

Preuve : La démonstration est une copie de celle du Lemme 3.4 jusqu'à l'équation (116) :

$$P_i(W_1 = w) = \frac{P_i(X_1 = (1^i, 1, w))}{P_i(L(X_1) > L(X_0))}$$

Dans le scénario one flip, un unique bit est changé à la fois, durant l'étape de mutation. Par conséquent, afin d'obtenir la configuration $(1^i, 1, w)$ au temps 1, nous avons besoin qu'au temps 0, la protéine soit dans l'état $(1^i, 0, w)$. Ainsi,

$$(123) \quad P_i(X_1 = (1^i, 1, w)) = P(X_1 = (1^i, 1, w) | X_0 = (1^i, 0, w)) P_i(Y_0 = w) = 1/(n2^{n-i-1})$$

D'un autre côté,

$$(124) \quad P_i(L(X_1) > L(X_0)) = 1/n.$$

Les égalités (123) et (124) appliquées à l'équation (116) donnent le résultat pour $k = 1$. Pour finir, la propriété de Markov forte s'applique ce qui termine la preuve pour $k > 1$.

□

Maintenant, nous allons pouvoir revenir à la preuve du Théorème 1.3, dans le cas one flip :

A partir de ce qui précède, nous pouvons déduire la loi de probabilité de \widehat{T}_n :

$$\begin{aligned} P(\widehat{T}_n = t) &= \frac{1}{2^n} \sum_J \sum_{\{o \leq i_0 \leq i_1 \leq \dots \leq i_J = n\}} \mathcal{NB}(J, \varepsilon)(t) \\ &= \sum_J \frac{1}{2^n} \binom{n}{J} \mathcal{NB}(J, \varepsilon)(t) \end{aligned}$$

Rappelons par le Lemme 2.1 que conditionnellement à $|X_0| = n - J$, T_n suit une loi binomiale négative de paramètre (J, ε) . Comme X_0 est choisi uniformément au hasard dans $\{0, 1\}^n$, cela entraîne que pour tout $t \geq 0$:

$$P(T_n = t) = P(\widehat{T}_n = t)$$

□

4.2. Preuve dans le cas Bernoulli flip .

Les dynamiques de MCM_0 et de $(1+1)_L$ -EA diffèrent légèrement. Cependant, on remarque que dans le contexte MCM_0 , les Lemmes 3.1, 3.2, 3.4, 3.5 sont encore vrais. Maintenant, comme la preuve de la Proposition 3.1 est entièrement basée sur ces lemmes, on conclut que la loi de probabilité du temps d'atteinte est la même, dans les scenari MCM_0 et $(1+1)_L$ -EA.

□

5. Conclusion

Après avoir examiné les deux versions (one et Bernoulli flip) des algorithmes d'évolution auxquels nous nous sommes intéressés ici, nous arrivons à la conclusion suivante : comme $(e^c - 1)/c^2 > 1$ pour tout $c \in \mathbb{R}_+$, l'espérance du temps d'atteinte dans le cas Bernoulli flip est plus grande que celle du one flip. Ainsi on peut conclure que le one flip est plus efficace que le Bernoulli flip, en terme de temps d'atteinte moyen. La même conclusion avait été trouvée par Garnier et al. [34] pour le problème OneMax.

Cette meilleure performance du one flip suggère que, malgré la capacité du Bernoulli flip à passer de n'importe quelle région de l'espace d'état à n'importe quelle autre, en une seule itération du processus de recherche, le Bernoulli flip résulte en une convergence plus lente vers un individu donné, dans le paysage d'énergie LeadingOnes.

Afin d'expliquer ce phénomène, comme la chaîne de Markov qui modélise notre recherche $(1+1)$ accepte une mutation seulement dans le cas d'une augmentation du nombre de premiers uns, nous remarquons les faits suivants : dans le contexte Bernoulli flip, plus l'algorithme se rapproche de l'individu cible, plus il met de temps avant de sauter; d'un autre côté, dans le cas one flip, le nombre de premiers uns présents dans la chaîne de bits n'interfère pas dans la loi de probabilité du temps mis, par l'algorithme de recherche, pour

sauter. Aussi, cette loi de probabilité reste stable à mesure que la recherche s’approche de l’individu optimal.

6. Problèmes ouverts

Question 1 : Convergence presque sûre et grandes déviations.

Une première possibilité de prolongement, à l’étude de ces algorithmes d’évolution simplifiés, serait de regarder des résultats de type grandes déviations, relatifs au temps d’atteinte T_n . Par ailleurs, un résultat de concentration de la mesure sur les probabilités $P(|T_n/E(T_n) - 1| > \varepsilon), \varepsilon > 0$ pourrait éventuellement nous permettre de déduire la convergence presque sûre de T_n/n^2 .

Question 2 : Comment le fitness croît-il en fonction du temps?

Notons $\gamma(n) = n^2/2$ (resp. $m(c)n^2$) pour le one (resp. Bernoulli) flip. En effectuant un changement d’échelle, pour mieux comprendre la dynamique du repliement, on pourrait regarder le comportement asymptotique, quand $n \rightarrow \infty$, du processus associé à la proportion de partie repliée, de la protéine au temps $[t\gamma(n)]$:

$$\left(\frac{L(X_{[t\gamma(n)]})}{n} \right)_{t \geq 0}$$

Question 3 : Qu’en est-il du *pire cas*?

Notons que le *pire cas* (“worst case” dans la littérature anglo-saxonne), pour lequel la configuration initiale de la protéine correspond au cas où toutes les sous-structures sont dépliées, c.à.d. $X_0 = (0, 0, \dots, 0)$, n’a pas été traité. Dans ce cadre, en ce qui concerne le Bernoulli flip, les choses sont légèrement plus compliquées. Reprenons les notations de la section 1.1. L’un des arguments clé, qui fait marcher la preuve du *cas moyen*, est contenu dans le Lemme 3.4. A savoir : quelle que soit $i_0 - 1$, la taille du nombre de premiers uns de la chaîne, la partie *ouverte* de la protéine prise au delà du premier 0 ($\phi_{i_0} = 0$) et qui est codée par $(\phi_{i_0+1}, \dots, \phi_n)$ reste, toujours, uniformément distribué sur $\{0, 1\}^{n-i_0}$. Ainsi, la dynamique *cas moyen*, répète un certain motif, vérifié à chaque itération de l’algorithme : l’uniforme distribution, dans l’espace des configurations, de la partie *ouverte* de la protéine, privée de la première sous-structure dénaturée (le premier zéro). Ceci découle du fait que la loi uniforme sur $\{0, 1\}^n$ est la mesure invariante de la matrice de transition associée aux mutations-Bernoulli flip. Par contre, il n’est plus vrai dans le *pire cas* du scénario Bernoulli flip, qu’à chaque itération de l’algorithme, la partie *ouverte* de la chaîne polypeptidique $(\phi_{i_0+1}, \dots, \phi_n)$ est uniformément distribuée. Dans tous les cas,

la structure de cette partie *ouverte*, n'est pas distribuée selon un même type de loi qui se répéterait à chaque étape. Notons, en revanche, que quand le nombre d'itérations tend vers l'infini, la loi de $(\phi_{i_0}, \dots, \phi_n)$ converge vers la loi uniforme. Aussi, il serait intéressant d'étudier la dynamique, *pire cas*-Bernoulli flip, d'un peu plus près.

Remarques : • Par contre, le *pire cas* sous l'hypothèse one flip de $(1+1)_L$ -EA se traite aisément. En effet, dans ce cadre, la partie ouverte est toujours dans l'état $(\phi_{i_0+1} = 0, \dots, \phi_n = 0)$. On peut alors montrer que :

- (i) Pour tout $n \geq 1$, $E(T_n) = n^2$;
- (ii) Quand $n \rightarrow \infty$, $T_n/E(T_n)$ converge en probabilité vers 1.
- (iii) Quand $n \rightarrow \infty$, $(T_n - E(T_n))/n^{3/2}$ converge en loi vers une variable aléatoire gaussienne centrée, de variance 1.

• Si l'on se place, d'un point de vue thermodynamique, il n'y a pas de différence entre l'état $(0, 0, \dots, 0)$ et les états pour lesquels $\phi_1 = 0$, puisque la partie ouverte de la protéine n'intervient pas dans le calcul de l'hamiltonien. On pourrait définir le *pire cas thermodynamique*, simplement comme étant l'état *dénaturé* (c.à.d. pour lequel $\phi_1 = 0$). Dans ce cas, la partie ouverte, (ϕ_2, \dots, ϕ_n) , est, à nouveau, uniformément distribuée sur $\{0, 1\}^{n-1}$. Il s'agit alors d'une simple adaptation des scénari déjà traités.

Question 4 : Comment accélérer la convergence?

En faisant référence, aux remarques faites dans la section précédente, qui soulignaient la meilleure performance du one flip, dans les dernières étapes de la recherche aléatoire, pour cibler précisément un individu donné, une fois que la recherche aléatoire s'est suffisamment rapproché, en opposition à la capacité du Bernoulli flip de passer d'une région à l'autre en une seule itération, une extension naturelle afin d'accélérer la convergence, serait de considérer une mutation inhomogène dans le temps, qui commencerait par un fort taux de mutation et se terminerait par un one flip. On pourrait également imaginer un modèle où les sous-unités ne changeraient plus de manière indépendante et où le taux de mutation ne serait plus déterministe en c/n mais de la forme X/n , pour une certaine variable aléatoire X à déterminer. Une autre possibilité pour accélérer la vitesse de convergence, serait d'augmenter la taille de la population en considérant des algorithmes élitistes de la forme $(\lambda + \mu)$ -EA.

Question 5 : Comment se comporte le modèle MCM, à température $T > 0$?

Revenons au modèle MCM de A. Bakk et al.[3] de la section 1.1. (dans le cas où $T = 1/\beta > 0$). Écrivons la fonction de partition du système : pour tout entier $i_0 \in [0, n]$, il

y a 2^{n-i_0} configurations possibles dans lesquelles la première sous-unité dépliée est $\phi_{i_0} = 0$, ainsi,

$$Z_n(\beta) = \sum_{i_0=1}^n 2^{n-i_0} e^{\beta \varepsilon_0 (i_0-1)} + e^{\beta n \varepsilon_0} = 2^{n-1} \frac{1 - e^{n(\beta \varepsilon_0 - \ln 2)}}{1 - e^{\beta \varepsilon_0 - \ln 2}} + e^{\beta n \varepsilon_0}.$$

On voit alors apparaître une température de transition $T_c = \ln 2 / \varepsilon_0$ au delà de laquelle, c'est l'état dénaturé (de fitness nul, c.à.d. $\phi_1 = 0$) qui est thermodynamiquement favorisé, contrairement à ce qui se passe dans le cas $0 < T < T_c$, où l'algorithme converge vers l'état natif $(1, 1, \dots, 1)$. Il serait intéressant d'étudier la dynamique du modèle *MCM* de plus près.

Questions 6 et 7 : Modélisations non uni-dimensionnelles.

La schématisation linéaire de la protéine, sous forme de chaîne de bits uni-dimensionnelle, semble peu conforme à la réalité. Un développement possible serait de se pencher sur des modélisations spatiales, plus complexes. Les deux types de modèles suivants, permettraient une représentation plus appropriée, de la structure des chaînes polypeptidiques :

1. **Le Modèle-arbre.** Dans un premier modèle, on peut représenter la protéine par un *arbre* binaire, ou encore *d*-aire, \tilde{T}_0 , blanc. A chaque instant k , la partie repliée de la protéine, correspond à un sous-arbre de T_0 , colorié en noir, noté T_k . Une augmentation du fitness de la protéine se traduit par le coloriage, en noir, d'une arête à la frontière entre \tilde{T}_0 et T_k .
2. **Le Modèle-tableau.** Une autre possibilité est de représenter la protéine par une boîte de taille $N \times N$ dans \mathbb{Z}^2 . Dans cette boîte, pour modéliser la dynamique du processus de repliement, on considère des *tableaux de Young* (la partie repliée), munis d'un ordre partiel.

Deux types de questions se posent :

- (i) A quelle vitesse arrive-t-on à l'équilibre?
- (ii) Peut-on décrire la géométrie asymptotique de la frontière partie repliée/partie dépliée, à l'équilibre, en particulier quand la taille de la protéine tend vers l'infini?

Remarque : Ce dernier axe de recherche (le modèle-tableau) fait notamment penser aux travaux de Cerf [22, 21] sur le cristal de Wulff. Cependant ici, le hamiltonien contient des termes d'interactions entre spins non voisins.

Bibliographie

- [1] C. B. Anfinsen. Principles that govern the folding of a protein chain (Nobel lecture). *Science*, 191:223–230, 1973.
- [2] T. Bäck. Optimal mutation rates in genetic search. In S. Forrest, editor, *Proceedings of the 5th International Conference on Genetic Algorithms*, pages 2–8. Morgan Kaufmann, 1993.
- [3] A. Bakk, J. S. Høye, A. Hansen, S. Sneppen, and M. H. Jensen. Pathways in two-state protein folding. *Biophys. J.*, 79:2722–2727, 2000.
- [4] N. Benhabiles, A. Thomas, and R. Brasseur. Les mécanismes de repliement des protéines solubles. *Biotechnol. Agron. Soc. Environ.*, 4(2):71–81, 2000.
- [5] J. Bérard. Genetic algorithms in random environments: two examples. Preprint.
- [6] J. Bérard and A. Bienvenüe. Convergence of a genetic algorithm with finite population. In *Mathematics and computer science (Versailles, 2000)*, pages 155–163. Birkhäuser, Basel, 2000.
- [7] J. Bérard and A. Bienvenüe. Sharp asymptotic results for simplified mutation-selection algorithms. *Ann. Appl. Probab.*, 13(4):1534–1568, 2003.
- [8] H.-G. Beyer. A alternative explanation for the manner in which genetic algorithms operate. *Biosystems*, 41:1–15, 1997.
- [9] H.-G. Beyer and H. P. Schwefel. Evolution strategies a comprehensive introduction. *Natural Computing*, 1:3–52, 2002.
- [10] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.
- [11] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [12] K. Binder. *Applications of the Monte Carlo Method in Statistical Physics*. Springer Verlag, Berlin, 1987.
- [13] E. Bolthausen. A central limit theorem for two-dimensional random walks in random sceneries. *Ann. Probab.*, 17(1):108–115, 1989.
- [14] A. N. Borodin. A limit theorem for sums of independent random variables defined on a recurrent random walk. *Dokl. Akad. Nauk SSSR*, 246(4):786–787, 1979.
- [15] A. N. Borodin. Limit theorems for sums of independent random variables defined on a transient random walk. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 85:17–29, 237, 244, 1979. Investigations in the theory of probability distributions, IV.
- [16] E. Boylan. Local times for a class of Markoff processes. *Illinois J. Math.*, 8:19–39, 1964.
- [17] H. Buchwalter. *Variation sur l'analyse*. ellipse, 1992.
- [18] P. Cabus and N. Guillotin-Plantard. Functional limit theorems for U -statistics indexed by a random walk. *Stochastic Process. Appl.*, 101(1):143–160, 2002.
- [19] R. Cerf. The dynamics of mutation-selection algorithms with large population sizes. *Ann. Inst. H. Poincaré Probab. Statist.*, 32(4):455–508, 1996.
- [20] R. Cerf. Asymptotic convergence of genetic algorithms. *Adv. in Appl. Probab.*, 30(2):521–550, 1998.

- [21] R. Cerf and R. Kenyon. The low-temperature expansion of the Wulff crystal in the 3D Ising model. *Comm. Math. Phys.*, 222(1):147–179, 2001.
- [22] R. Cerf and Á. Pisztora. On the Wulff crystal in the Ising model. *Ann. Probab.*, 28(3):947–1017, 2000.
- [23] E. Csáki, W. König, and Z. Shi. An embedding for the Kesten-Spitzer random walk in random scenery. *Stochastic Process. Appl.*, 82(2):283–292, 1999.
- [24] E. Csáki, P. Révész, and Z. Shi. A strong invariance principle for two-dimensional random walk in random scenery. *Stochastic Process. Appl.*, 92(2):181–200, 2001.
- [25] D. Dacunha-Castelle and M. Duflo. *Probabilités et statistiques. Tome 2*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master’s Degree]. Masson, Paris, 1983. Problèmes à temps mobile. [Movable-time problems].
- [26] P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(2):155–194, 2001.
- [27] F. den Hollander, J. Naudts, and P. Scheunders. A long-time tail for random walk in random scenery. *J. Statist. Phys.*, 66(5-6):1527–1555, 1992.
- [28] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.
- [29] K.A. Dill, K. M. Feibig, and H. S. Chan. Cooperativity in protein folding kinetics. *Proc. Natl. Acad. Sci. USA*, 90:1942–1946, 1993.
- [30] S. Droste, T. Jansen, and I. Wegener. A rigorous complexity analysis of the $(1+1)$ evolutionary algorithm for linear functions with boolean inputs. In *Proceedings of the IEEE International Conference on Evolutionary Computation ICEC’98*, pages 499–504. IEEE Press, 1998.
- [31] S. Droste, T. Jansen, and I. Wegener. On the analysis of the $(1+1)$ evolutionary algorithm. *Theoret. Comput. Sci.*, 276(1-2):51–81, 2002.
- [32] E. B. Dynkin and A. Mandelbaum. Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.*, 11(3):739–745, 1983.
- [33] D.B. Fogel, A.J. Owens, and M.J. Walsh. *Artificial Intelligence through Simulated Evolution*. Wiley, New York, 1966.
- [34] J. Garnier, L. Kallel, and M. Schoenauer. Rigorous hitting times for binary mutations. *Evolutionary Computation*, 7(2):173–203, 1999.
- [35] R. K. Gettoor and H. Kesten. Continuity of local times for Markov processes. *Compositio Math.*, 24:277–303, 1972.
- [36] I. I. Gikhman and A. V. Skorokhod. *Introduction to the theory of random processes*. Dover Publications Inc., Mineola, NY, 1996. Translated from the 1965 Russian original, Reprint of the 1969 English translation, With a preface by Warren M. Hirsch.
- [37] D.E. Goldberg. *Genetic algorithms in search, optimization and machine learning*. Addison Wesley, 1989.
- [38] G. G. Gregory. Large sample theory for U -statistics and tests of fit. *Ann. Statist.*, 5(1):110–123, 1977.
- [39] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19:293–325, 1948.
- [40] J. H. Holland. *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, Mich., 1975. An introductory analysis with applications to biology, control, and artificial intelligence.
- [41] W. N. Hudson. Operator-stable distributions and stable marginals. *J. Multivariate Anal.*, 10(1):26–37, 1980.

- [42] K. Jörgens. *Linear integral operators*, volume 7 of *Surveys and Reference Works in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass., 1982. Translated from the German by G. F. Roach.
- [43] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.
- [44] H. Kesten and F. Spitzer. A limit theorem related to a new class of self-similar processes. *Z. Wahrsch. Verw. Gebiete*, 50(1):5–25, 1979.
- [45] D. Khoshnevisan and T. M. Lewis. A law of the iterated logarithm for stable processes in random scenery. *Stochastic Process. Appl.*, 74(1):89–121, 1998.
- [46] A.J. Lee. *U-statistics. Theory and practice*. Marcel Dekker, Inc., New York, 1990.
- [47] C. Levinthal. Principles that govern the folding of a protein chain (Nobel lecture). *J. Chem. Phys.*, 65:44–45, 1968.
- [48] M. Maejima. Limit theorems related to a class of operator-self-similar processes. *Nagoya Math. J.*, 142:161–181, 1996.
- [49] C. Mazza and D. Piau. On the effect of selection in genetic algorithms. *Random Structures Algorithms*, 18(2):185–200, 2001.
- [50] P. A. Meyer. Un cours sur les intégrales stochastiques. In *Séminaire de Probabilités, X (Seconde partie: Théorie des intégrales stochastiques, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 245–400. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.
- [51] H. Mühlenbein. How genetic algorithms really work: I. Mutation and hill-climbing. In R. Männer and R. Manderick, editors, *Parallel Problem Solving from Nature PPSN II*, pages 15–25. North Holland, Amsterdam, 1992.
- [52] G. Neuhaus. Functional limit theorems for U -statistics in the degenerate case. *J. Multivariate Anal.*, 7(3):424–439, 1977.
- [53] D. Piau. Scaling exponents of random walks in random sceneries. *Stochastic Process. Appl.*, 100:3–25, 2002.
- [54] D. Piau. Further scaling exponents of random walks in random sceneries. To appear in *Stochastic Process. Appl.*, 2004.
- [55] A. Prugel-Bennett and Shapiro J.L. An analysis of genetic algorithms using statistical mechanics. *Phys. Rev. Lett.*, 72(9):1305–1309, 1994.
- [56] Y. Rabinovich and A. Wigderson. Techniques for bounding the convergence rate of genetic algorithms. *Random Structures Algorithms*, 14(2):111–138, 1999.
- [57] I. Rechenberg. *Evolutionstrategie: Optimierung Technischer Systeme nach Prinzipien des Biologischen Evolution*. Fromman-Holzboog Verlag, Stuttgart, 1973.
- [58] P. Révész and Z. Shi. Strong approximation of spatial random walk in random scenery. *Stochastic Process. Appl.*, 88(2):329–345, 2000.
- [59] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [60] H. Rubin and R. A. Vitale. Asymptotic distribution of symmetric statistics. *Ann. Statist.*, 8(1):165–170, 1980.
- [61] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
- [62] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.
- [63] G. Rudolph. *Convergence Properties of Evolutionary Algorithms*. Kovac, Hamburg, 1997.

- [64] J. A. Schellman. The factors affecting the stability of hydrogen polypeptide structures in solution. *J. Phys. Chem.*, 62, 1485-1494.
- [65] H.-P. Schwefel. *Numerical Optimization of Computer Models*. Chichester: Wiley & Sons, 1981.
- [66] R. J. Serfling. *Approximation theorems of mathematical statistics*. John Wiley & Sons Inc., New York, 1980. Wiley Series in Probability and Mathematical Statistics.
- [67] F. Spitzer. *Principles of random walks*. Springer-Verlag, New York, second edition, 1976. Graduate Texts in Mathematics, Vol. 34.
- [68] M. D. Vose. *The simple genetic algorithm*. Complex Adaptive Systems. MIT Press, Cambridge, MA, 1999. Foundations and theory, A Bradford Book.