



Scaling limit of the collision measures of multiple random walks

Dinh-Toan Nguyen

► To cite this version:

| Dinh-Toan Nguyen. Scaling limit of the collision measures of multiple random walks. 2022. <hal-03603284>

HAL Id: hal-03603284

<https://hal.science/hal-03603284v1>

Preprint submitted on 9 Mar 2022

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.



HAL Authorization

Scaling limit of the collision measures of multiple random walks

Dinh-Toan Nguyen^{* †}

Abstract

For an integer $k \geq 2$, let $S^{(1)}, S^{(2)}, \dots, S^{(k)}$ be k independent simple symmetric random walks on \mathbb{Z} . A pair (n, z) is called a collision event if there are at least two distinct random walks, namely, $S^{(i)}, S^{(j)}$ satisfying $S_n^{(i)} = S_n^{(j)} = z$. We show that under the same scaling as in Donsker's theorem, the sequence of random measures representing these collision events converges to a non-trivial random measure on $[0, 1] \times \mathbb{R}$. Moreover, the limit random measure can be characterized using Wiener chaos. The proof is inspired by methods from statistical mechanics, especially, by a partition function that has been developed for the study of directed polymers in random environments.

1 Introduction

For an integer $k \geq 2$, let $S^{(1)}, S^{(2)}, \dots, S^{(k)}$ be k independent simple symmetric random walks (SSRWs) on \mathbb{Z} , defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ (see [20, p.3]). A pair $(n, z) \in \mathbb{N} \times \mathbb{Z}$ is called a *collision event* if there are at least two random walks that collide (occupy the same position at the same time) at the time n and the location z (see Figure 1), and n is then called a *collision time*.

First mentioned in Pólya's note [18], the collision of random walks has since then been a classic topic in probability theory. Recently, this topic has gained more attention from researchers working on the random walks on graphs [4, 12] and random environments [11, 9, 3].

When consider only two random walks, collision problems are strongly related to Brownian local time [15, 19, 23]. The convergence of collision times

MSC2020 subject classifications: 60F05, 60G57, 82C05.

Key words and phrases: Random Walks, Collisions, Scaling Limits, Wiener Chaos, Partition Functions, Random Measures, U -statistics, Random Environment

^{*}LAMA, Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS, F-77454 Marne-la-Vallée, France.

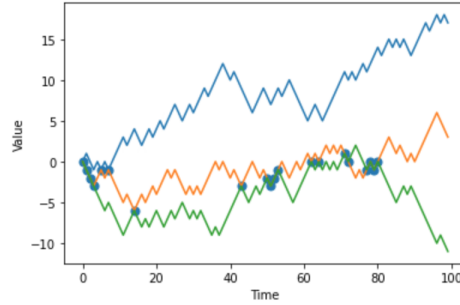
[†]Département de Mathématiques, Université du Québec à Montréal (UQAM), Montréal, QC H2X 3Y7, Canada.

can be achieved by coupling a new SSRW, the difference between two given random walks, with a Brownian motion using Skorokhod's embedding [20, p.52][10, Theorem 8.6.1]. However, these methods cannot be easily adapted to give convergence results for the collisions of more than two SSRWs because the couplings rely heavily the choices of stopping times which are proper for each SSRW.

To the best of our knowledge, scaling limit results for collisions of $k > 2$ random walks are still limited.

In this paper, we investigate a relatively uncommon aspect of random walk collisions, concerning their duality with the partition function of a directed polymer model in statistical mechanics [8]. By following the ideas developed in [8, 2], we obtain new results on when and where the collisions of these random walks occur after long observation, or more precisely, on the scaling limit of the empirical measures of the collision events.

Figure 1: An example of collision events when $k = 3, N = 100$. The collision events are represented by blue dots.



Our study objects are as follows:

Definition 1.1. For each $N \in \mathbb{N}$, we define the collision measures of k random walks $S^{(1)}, S^{(2)}, \dots, S^{(k)}$ until time N to be:

$$\Pi_N := \sum_{n=1}^N \sum_{z \in \mathbb{Z}} \sum_{\substack{1 \leq i < j \leq k: \\ S_n^{(i)} = S_n^{(j)} = z}} \delta_{\left(\frac{n}{N}, \frac{z}{\sqrt{N}}\right)}, \text{ and}$$

$$\Pi'_N := \sum_{n=1}^N \sum_{z \in \mathbb{Z}} \delta_{\left(\frac{n}{N}, \frac{z}{\sqrt{N}}\right)} \mathbb{1}_{\{(n,z) \text{ is an collision event}\}},$$

where δ is the Dirac measure.

For each $N \in \mathbb{N}$, the main difference between Π'_N and Π_N is that Π_N takes into account the multiplicity of collision events. For example, if the number of considered random walks is 3 and it happens that $S_n^{(1)} = S_n^{(2)} = S_n^{(3)} = z$ for some (n, z) , then the Dirac measure $\delta_{\left(\frac{n}{N}, \frac{z}{\sqrt{N}}\right)}$ will appear 3 times in the

summation in Π_N while the number of its appearance in Π'_N is still one. Concerning the scaling choice, one can observe that this is the same scaling as in Donker's theorem, which also suggests that the distribution of $\hat{\Pi}_N$ is closely related to Brownian local time. Indeed, when $k = 2$, the total measure of Π_N is equal to

$$\#\{n \in \llbracket 1, N \rrbracket : S_n^{(1)} - S_n^{(2)} = 0\} \stackrel{(d)}{=} \#\{n \in \llbracket 1, N \rrbracket : S_{2n}^{(1)} = 0\} = L_{S^{(1)}}^0(2N), \quad (1)$$

where $\stackrel{(d)}{=}$ is the equality in law, and $L_S^0(t)$ is the local time at position 0 during the period $[0, t]$ of some walk S . So by the convergence of local time of simple random walks, the equation (1) implies $\frac{1}{\sqrt{N}} \|\Pi_N\| = O_{\mathbf{P}}(1)$, where $\|\Pi_N\|$ denotes the total measure of Π_N .

In this work, we not only bound the sequence $(\frac{1}{\sqrt{N}}\Pi_N; N \in \mathbb{N})$, but also prove that this sequence of random measures converges to a non-trivial random measure \mathcal{N} on $[0, 1] \times \mathbb{R}$. Before giving our main result, let us recall the convergence of random measures.

Definition 1.2. (*Convergence of random measures*) Suppose ξ, ξ_1, ξ_2, \dots are random finite measures on $([0, 1] \times \mathbb{R}, \mathcal{B}([0, 1] \times \mathbb{R}))$, we say that $\xi_n \xrightarrow[n \rightarrow \infty]{wd} \xi$ if the sequence of real random variables $(\xi_n(f), n \in \mathbb{N})$ converges in distribution to $\xi(f)$ when n goes to infinity for all bounded continuous function $f \in \mathcal{C}_b([0, 1] \times \mathbb{R})$. Here, $\mu(f)$ denotes the integral $\int f d\mu$ for any (random) measure μ and bounded measurable function f .

Here are our main results:

Theorem 1.3. (*Convergence of collision measures and characterization of the limit random measure*)

- There is a random finite positive measure \mathcal{N} on the measurable space $([0, 1] \times \mathbb{R}, \mathcal{B}([0, 1] \times \mathbb{R}))$ such that:

$$\frac{1}{\sqrt{N}}\Pi_N \xrightarrow[N \rightarrow +\infty]{wd} \mathcal{N} \quad \text{and} \quad \frac{1}{\sqrt{N}}\Pi'_N \xrightarrow[N \rightarrow +\infty]{wd} \mathcal{N}.$$

- Furthermore, for all nonnegative bounded continuous function $f \in \mathcal{C}_{b,+}([0, 1] \times \mathbb{R})$, the exponential moment of \mathcal{N} with respect to f is equal to k -th moment of a positive random variable $\mathcal{Z}_{\sqrt{2}f}$:

$$\mathbf{E} \left[e^{\mathcal{N}(f)} \right] = \mathbf{E} \left[\left(\mathcal{Z}_{\sqrt{2}f} \right)^k \right].$$

where for each $a \in \mathcal{C}_{b,+}([0, 1] \times \mathbb{R})$, the random variable \mathcal{Z}_a is identified as the sum of multiple stochastic integrals given by:

$$\mathcal{Z}_a := 1 + \sum_{n=1}^{\infty} \int_{\Delta_n} \int_{\mathbb{R}^n} \prod_{i=1}^n \left[a(\mathbf{t}_i, \mathbf{x}_i) \varrho(\mathbf{t}_i - \mathbf{t}_{i-1}, \mathbf{x}_i - \mathbf{x}_{i-1}) W(dt_i d\mathbf{x}_i) \right], \quad (2)$$

where, W is the white noise based on the Lebesgue measure on $[0, 1] \times \mathbb{R}$, $\mathbf{x}_0 = 0$, ϱ is the standard Gaussian heat kernel

$$\varrho(t, x) = \frac{e^{-x^2/2t}}{\sqrt{2\pi t}},$$

and Δ_n is the n -dimensional simplex

$$\Delta_n := \{\mathbf{t} \in [0, 1]^n : 0 = \mathbf{t}_0 < \mathbf{t}_1 < \mathbf{t}_2 < \dots < \mathbf{t}_n \leq 1\}. \quad (3)$$

We refer to [25, Chapter 1] for an introduction on white noise (see also Section 3.2 of this article).

As mentioned earlier, we will prove this theorem by investigating the connection between the collision measures of random walks with a model in statistical mechanics; hence, we will introduce many auxiliary notions in our paper, such as *random environment* ω , *partition function* \mathfrak{Z}_N and *U-statistics* $\mathcal{S}_n^N(\cdot)$.

The general idea is that by associating each point (n, z) on the grid $\mathbb{N} \times \mathbb{Z}$ with a random variable, we can change the underlying framework from studying a deterministic grid to studying a collection of random variables indexed by $\mathbb{N} \times \mathbb{Z}$. For a such collection, the range of possible tools from statistical mechanics is large. Indeed, the partition function we will use is developed to study a directed polymer model [2, ?, 5].

The organization of this paper is as follows: Section 2 introduces some basic notions that we will use in the sequel to explain our main ideas, especially, the relation between the concept of partition functions and the collision measures Π_N . Section 3 gives a brief review on U -statistics and Wiener chaos. At the end of this section, we prove some new theorems on the convergence of U -statistics, on which our asymptotic result on partition functions (Theorem 2.3) is based. Section 4 presents a short study on the random variable \mathcal{Z}_a defined in (2), and our proof of Theorem 2.3. Finally, Section 5 combines all proved results to show the *weak tightness* of $(\frac{1}{\sqrt{N}}\Pi_N, N \in \mathbb{N})$, and prove Theorem 1.3.

Some auxiliary results are presented in the Appendix at the end of this article.

2 Partition functions and main ideas of proof

2.1 Partition functions

We introduce a collection $\omega := (\omega(i, z) : i \in \mathbb{N}, z \in \mathbb{Z})$ of independent Rademacher variables indexed by $\mathbb{N} \times \mathbb{Z}$, i.e., for all $(n, z) \in \mathbb{N} \times \mathbb{Z}$,

$$\mathbf{P}(\omega(n, z) = -1) = \mathbf{P}(\omega(n, z) = 1) = \frac{1}{2}.$$

These random variables are created by extending our existing probability space $(\Omega, \mathcal{A}, \mathbf{P})$ so that $\omega, S^{(1)}, S^{(2)}, \dots, S^{(k)}$ are independent. In the sequel, for a real number β and a real function A on $\mathbb{N} \times \mathbb{Z}$, $\beta\omega$ and $A\omega$ are defined as:

$$\begin{aligned} A\omega &:= (A(i, z)\omega(i, z) : i \in \mathbb{N}, z \in \mathbb{Z}) . \\ \beta\omega &:= (\beta\omega(i, z) : i \in \mathbb{N}, z \in \mathbb{Z}) . \end{aligned}$$

As briefly explained in Section 1, the role ω is to add new degrees of freedom to the existing model, by which we have more flexibility to create more objects. The partition function \mathfrak{Z} is one of such objects:

Definition 2.1. *For any positive integer N and any real function A on $\mathbb{N} \times \mathbb{Z}$, the **partition function** $\mathfrak{Z}_N(A)$ is defined as the conditional expectation:*

$$\mathfrak{Z}_N(A) := \mathbf{E} \left[\prod_{n=1}^N (1 + A(n, S_n^{(1)})\omega(n, S_n^{(1)})) \mid \omega \right] .$$

Note that $\mathfrak{Z}_N(A)$ is a random variable depending only on the value of ω .

2.2 Main ideas

The starting point of our paper and the proof of our main results is a heuristic relation between the partition functions \mathfrak{Z}_N and the random measure Π_N : Given a nonnegative bounded function A on $\mathbb{N} \times \mathbb{Z}$, since $S^{(1)}, \dots, S^{(k)}$ are i.i.d., we have:

$$\begin{aligned} \mathbf{E} \left[\mathfrak{Z}_N \left(\frac{1}{N^{1/4}} A \right)^k \right] &= \mathbf{E} \left[\mathbf{E} \left[\prod_{n=1}^N \left(1 + \frac{1}{N^{1/4}} A(n, S_n^{(1)})\omega(n, S_n^{(1)}) \right) \mid \omega \right]^k \right] \\ &= \mathbf{E} \left[\mathbf{E} \left[\prod_{i=1}^k \prod_{n=1}^N \left(1 + \frac{1}{N^{1/4}} A(n, S_n^{(i)})\omega(n, S_n^{(i)}) \right) \mid \omega \right] \right] \\ &= \mathbf{E} \left[\prod_{n=1}^N \prod_{i=1}^k \left(1 + \frac{1}{N^{1/4}} A(n, S_n^{(i)})\omega(n, S_n^{(i)}) \right) \right] \\ &= \mathbf{E} \left[\mathbf{E} \left[\prod_{n=1}^N \prod_{i=1}^k \left(1 + \frac{1}{N^{1/4}} A(n, S_n^{(i)})\omega(n, S_n^{(i)}) \right) \mid S^{(1)}, \dots, S^{(k)} \right] \right] \\ &= \mathbf{E} \left[\prod_{n=1}^N \mathbf{E} \left[\prod_{i=1}^k \left(1 + \frac{1}{N^{1/4}} A(n, S_n^{(i)})\omega(n, S_n^{(i)}) \right) \mid S^{(1)}, \dots, S^{(k)} \right] \right] \\ &= \mathbf{E} \left[\prod_{n=1}^N \left[1 + \frac{1}{N^{1/2}} \left(\sum_{\substack{1 \leq i < j \leq k: \\ S_n^{(i)} = S_n^{(j)} = z}} A^2(n, z) \right) + \frac{1}{N^{3/4}} (\dots) + \dots \right] \right] \end{aligned}$$

Then since $1 + x \approx e^x$, heuristically, we deduce:

$$\begin{aligned} \mathbf{E} \left[\mathfrak{Z}_N \left(\frac{1}{N^{1/4}} A \right)^k \right] &\approx \mathbf{E} \left[\prod_{n=1}^N \left[\exp \left(\frac{1}{N^{1/2}} \sum_{\substack{1 \leq i < j \leq k: \\ S_n^{(i)} = S_n^{(j)} = z}} A^2(n, z) \right) \right] \right] \\ &= \mathbf{E} \left[\exp \left(\frac{1}{N^{1/2}} \sum_{n=1}^N \sum_{\substack{1 \leq i < j \leq k: \\ S_n^{(i)} = S_n^{(j)} = z}} A^2(n, z) \right) \right] = \mathbf{E} \left[\exp \left(\frac{1}{N^{1/2}} \Pi_N(f_N) \right) \right], \end{aligned}$$

where f_N is a measurable function such that $f_N \left(\frac{n}{N}, \frac{z}{\sqrt{N}} \right) = A(n, z)$ for all $n \in \mathbb{N}, z \in \mathbb{Z}$.

In short, by abuse of notation, the above observation suggests that:

$$\mathbf{E} \left[e^{N^{-1/2} \Pi_N} \right] \approx \mathbf{E} \left[(\mathfrak{Z}_N)^k \right]. \quad (4)$$

In other words, if we have a good understanding of \mathfrak{Z} , we will have good information on Π_N, Π'_N .

Then to study the partition function \mathfrak{Z}_N , we base our study on the paper [2], in which Alberts et al. studied the scaling limit of \mathfrak{Z} when the function A is constant. In our study, we generalize their results for a sequence of functions $(A_N, N \in \mathbb{N})$ satisfying certain conditions. An expansion of Wiener chaoses emerges naturally in our limit objects because, as we will see, each term in the algebraic expansion of \mathfrak{Z}_N (cf. Proposition 4.6) converges to a Wiener chaos. To this aim, we will have to introduce some U -statistics and study their asymptotic behavior in Section 3.

2.3 Main results on partition functions

We terminate this section by presenting our results on the asymptotic behavior of \mathfrak{Z}_N . The proofs will be presented later in Section 4.

Notation 2.2. For $(t, x) \in [0, 1] \times \mathbb{R}$, $[t, x]_N$ denotes the unique pair of integer (i, z) such that :

- $(t, x) \in \left(\frac{i-1}{N}, \frac{i}{N} \right] \times \left(\frac{z-1}{\sqrt{N}}, \frac{z+1}{\sqrt{N}} \right]$,
- i and z have same parity.

Theorem 2.3. Let $(A_n, n \in \mathbb{N})$ be a sequence of real functions whose domain is $\mathbb{N} \times \mathbb{Z}$ such that:

- i. $\sup_N \|A_N\|_\infty < +\infty$,
- ii. there is a measurable function $a \in L^\infty([0, 1] \times \mathbb{R})$ such that:

$$\lim_{N \rightarrow +\infty} A_N([t, x]_N) = a(t, x) \quad a.e.$$

Then as N converges to infinity, we have:

$$\mathfrak{Z}_N(N^{-1/4}A_N) \xrightarrow{(d)} \mathcal{Z}_{\sqrt{2}a}$$

This theorem is a generalization of Proposition 5.3 in [2] where the sequence $(A_N; N \in \mathbb{N})$ is replaced by a fixed constant $\beta \geq 0$.

We will also prove that under some conditions, the partition functions are uniformly bounded in L^k :

Theorem 2.4. *For a sequence of real functions $(A_n, n \in \mathbb{N})$ on $\mathbb{N} \times \mathbb{Z}$ such that $\sup_N \|A_N\|_\infty < +\infty$, we have:*

$$\limsup_N \mathbf{E} \left[\left(\mathfrak{Z}_N(N^{-1/4}A_N) \right)^k \right] < +\infty.$$

Notice that even though k is fixed in our study, the definition of \mathfrak{Z} does not depend on k . So, the above sequence $(\mathfrak{Z}_N(N^{-1/4}A_N), N \in \mathbb{N})$ is also uniformly bounded in L^p for all $p \in \mathbb{N}$, which implies directly the following corollary:

Corollary 2.5. *For a sequence of real functions A_1, A_2, \dots on $\mathbb{N} \times \mathbb{Z}$ such that $\sup_N \|A_N\|_\infty < +\infty$, the sequence of random variables $(\mathfrak{Z}_N(N^{-1/4}A_N), N \in \mathbb{N})$ is uniformly L^k -integrable.*

3 U -Statistics: related notions and limit theorems

Let $E_n^N := \{\mathbf{i} \in \llbracket 1, N \rrbracket^n : \mathbf{i}_j \neq \mathbf{i}_l \text{ for } j \neq l\}$. In this paper, we are interested in sums of the form:

$$\sum_{\mathbf{i} \in E_n^N} \sum_{\substack{\mathbf{z} \in \mathbb{Z}^n \\ \mathbf{i} \leftrightarrow \mathbf{z}}} \overline{g_N}(\mathbf{i}, \mathbf{z}) A_N(\mathbf{i}, \mathbf{z}) \omega(\mathbf{i}, \mathbf{z}), \quad (5)$$

for some weight functions $\overline{g_N}$ specified later. The notation $\mathbf{i} \leftrightarrow \mathbf{z}$ means that:

Notation 3.1. $\mathbf{i} \leftrightarrow \mathbf{z}$ means that for all $j \in \llbracket 1, n \rrbracket$, the corresponding j -th coordinates of \mathbf{i} and \mathbf{z} , namely \mathbf{i}_j and \mathbf{z}_j , have the same parity.

We will see that sums of this type appear naturally when we expand the partition functions \mathfrak{Z}_N (see (19)).

The organization of this section is as follows: Sections 3.1 and 3.2 introduce the framework of Theorem 3.12 which is our main result on the convergence of sums of the form (5). Section 3.3 presents the proof of Theorem 3.12 and other related results.

The approach we used in this section is standard in the theory of U -statistics. Interested readers can consult the book [16] of Korolyuk and Borovskikh for a more rigorous introduction of this theory.

3.1 Introduction of U -statistics S_n^N

We first make precise the definition of the weight functions $(\overline{g}_N, N \in \mathbb{N})$ in the above sum.

Let g be a function in $L^2([0, 1]^n \times \mathbb{R}^n)$. For each N , the weight functions $\overline{g_N}$ associated to g is defined by the following procedure:

First, we partition the space $(0, 1]^n \times \mathbb{R}^n$ in rectangles of the form:

$$\mathcal{R}_n^N := \left\{ \left(\frac{\mathbf{i} - \mathbf{1}}{N}, \frac{\mathbf{i}}{N} \right) \times \left(\frac{\mathbf{z} - \mathbf{1}}{\sqrt{N}}, \frac{\mathbf{z} + \mathbf{1}}{\sqrt{N}} \right) : \mathbf{i} \in D_n^N, \mathbf{z} \in \mathbb{Z}^n, \mathbf{i} \leftrightarrow \mathbf{z} \right\}.$$

with $\mathbf{1}$ being the vector of ones and D_n^N being the *integer simplex*:

$$D_n^N := \{\mathbf{i} \in \llbracket 1, N \rrbracket^n : 1 \leq \mathbf{i}_1 < \mathbf{i}_2 \cdots < \mathbf{i}_n \leq N\}. \quad (6)$$

Visually, \mathcal{R}_n^N is a collection of nonoverlapping translations of the base rectangle:

$$\left(\frac{0}{N}, \frac{1}{N}\right)^n \times \left(\frac{-1}{\sqrt{N}}, \frac{1}{\sqrt{N}}\right)^n.$$

Then, the function \overline{g}_N is defined as the average of g on each rectangles above. More precisely, for any $(\mathbf{t}, \mathbf{x}) \in (0, 1]^n \times \mathbb{R}^n$, $\overline{g}_N(\mathbf{t}, \mathbf{x})$ is defined as the mean:

$$\overline{g_N}(\mathbf{t}, \mathbf{x}) := \frac{1}{|R|} \int_R g(\mathbf{s}, \mathbf{y}) d\mathbf{s} d\mathbf{y},$$

where R is the unique rectangle in \mathcal{R}_n^N that contains (\mathbf{t}, \mathbf{x}) , and $|R|$ denotes the Lebesgue measure of R . In probabilistic terms, $\overline{g_N}$ is simply the conditional expectation of g onto the rectangles of \mathcal{R}_n^N . We note that $|R| = 2^n N^{-3n/2}$. This term will appear recurrently in most of our computations.

Suppose $(A_N, N \in \mathbb{N})$ is a sequence of real-valued functions on $\mathbb{N} \times \mathbb{Z}$.

Notation 3.2. For any n -tuple $\mathbf{i} \in E_n^N$ and n -tuple $\mathbf{z} \in \mathbb{Z}^n$, $A_N(\mathbf{i}, \mathbf{z})$ and $\omega(\mathbf{i}, \mathbf{z})$ denote

$$A_N(\mathbf{i}, \mathbf{z}) := A_N(\mathbf{i}_1, \mathbf{z}_1)A_N(\mathbf{i}_2, \mathbf{z}_2)...A_N(\mathbf{i}_n, \mathbf{z}_n),$$

$$\omega(\mathbf{i}, \mathbf{z}) := \omega(\mathbf{i}_1, \mathbf{z}_1)\omega(\mathbf{i}_2, \mathbf{z}_2)...\omega(\mathbf{i}_n, \mathbf{z}_n),$$

with \mathbf{i}_j being the j -th coordinate of \mathbf{i} as defined previously.

Now, we define the weighted U -statistics \mathcal{S}_n^N .

Definition 3.3. Suppose $(A_n, n \in \mathbb{N})$ is a sequence of bounded real-valued functions on $\mathbb{N} \times \mathbb{Z}$. For any function $g \in L^2([0, 1]^n \times \mathbb{R}^n)$, the U -statistics \mathcal{S}_n^N is defined as:

$$\mathcal{S}_n^N(g) := 2^{n/2} \sum_{\mathbf{i} \in E_n^N} \sum_{\substack{\mathbf{z} \in \mathbb{Z}^N: \\ \mathbf{i} \leftrightarrow \mathbf{z}}} \overline{g_N} \left(\frac{\mathbf{i}}{N}, \frac{\mathbf{z}}{\sqrt{N}} \right) A_N(\mathbf{i}, \mathbf{z}) \omega(\mathbf{i}, \mathbf{z}). \quad (7)$$

We first give some basic properties of the U -statistics S_n^N .

Proposition 3.4. *Suppose $(A_N, N \in \mathbb{N})$ is a sequence of bounded real-valued functions on $\mathbb{N} \times \mathbb{Z}$. For all positive integers n and N , we have:*

i. (Well-posedness) $\mathcal{S}_n^N(g)$ is well-defined and has zero mean for all $g \in L^2([0, 1]^n \times \mathbb{R}^n)$.

ii. (Linearity) For all $f, g \in L^2([0, 1]^n \times \mathbb{R}^n)$, $\alpha, \beta \in \mathbb{R}$

$$\mathcal{S}_n^N(\alpha f + \beta g) = \alpha \mathcal{S}_n^N(f) + \beta \mathcal{S}_n^N(g).$$

iii. (L^2 -boundedness) If $c > 0$ is a number such that $\|A_N\|_\infty \leq c$, then for all $g \in L^2([0, 1]^n \times \mathbb{R}^n)$:

$$\mathbf{E}[\mathcal{S}_n^N(g)^2] \leq c^{2n} N^{3n/2} \|g\|_2^2.$$

iv. (Uncorrelatedness of U -statistics of different orders) If n_1, n_2 are two different positive integers, then $\forall g_i \in L^2([0, 1]^{n_i} \times \mathbb{R}^{n_i}) \quad i = 1, 2$,

$$\mathbf{E}[\mathcal{S}_{n_1}^N(g_1) \mathcal{S}_{n_2}^N(g_2)] = 0.$$

Proof. Assume that f and g have compact supports, then the sums in $S_n^N(f)$ and $S_n^N(g)$ have a finite number of terms; thus, point *i* is trivial. Point *ii* is also trivial by recalling that ω is a collection of centered random variables. Now, for point *iii*, observe that for any $\mathbf{i}, \mathbf{i}' \in E_n^N, \mathbf{x}, \mathbf{x}' \in \mathbb{Z}^n$:

$$\mathbf{E} \left[\prod_{l=1}^n \omega(\mathbf{i}_l, \mathbf{x}_l) \prod_{l=1}^n \omega(\mathbf{i}'_l, \mathbf{x}'_l) \right] = \mathbb{1}_{\{\mathbf{i}=\mathbf{i}', \mathbf{x}=\mathbf{x}'\}}.$$

Hence

$$\begin{aligned} \mathbf{E} \left[\mathcal{S}_n^N(g)^2 \right] &= 2^n \sum_{\mathbf{i} \in E_n^N} \sum_{\substack{\mathbf{z} \in \mathbb{Z}^n \\ \mathbf{i} \leftrightarrow \mathbf{z}}} A_N(\mathbf{i}, \mathbf{z})^2 \overline{g_N} \left(\frac{\mathbf{i}}{N}, \frac{\mathbf{z}}{\sqrt{N}} \right)^2 \\ &\leq 2^n \sum_{\mathbf{i} \in [1, N]^n} \sum_{\substack{\mathbf{z} \in \mathbb{Z}^n \\ \mathbf{i} \leftrightarrow \mathbf{z}}} c^{2n} \frac{1}{|\mathcal{R}|} \int_{\mathcal{R}} g(\mathbf{t}, \mathbf{y})^2 d\mathbf{t} d\mathbf{y} \\ &= N^{3n/2} c^{2n} \int_{[0, 1]^n} \int_{\mathbb{R}^n} g(\mathbf{t}, \mathbf{y})^2 d\mathbf{t} d\mathbf{y}. \end{aligned}$$

The last inequality is simply an application of the Cauchy-Schwarz lemma. So our theorem is valid for compactly supported functions. In other words, $g \mapsto S_n^N(g)$ is a linear Lipschitz continuous mapping that maps the space $L^2_{\text{compact}}([0, 1]^n \times \mathbb{R}^n)$ into $L^2(\mathbf{P})$. Hence, all the properties *i, ii, iii* can be extended naturally to all $L^2([0, 1]^n \times \mathbb{R}^n)$ by the density of $L^2_{\text{compact}}([0, 1]^n \times \mathbb{R}^n)$.

\mathbb{R}^n) in $L^2([0, 1]^n \times \mathbb{R}^n)$.

For the covariance relation in point iv , one can observe that if $\mathbf{i} \in E_{n_1}^N, \mathbf{x} \in \mathbb{Z}^{n_1}, \mathbf{i}' \in E_{n_2}^N, \mathbf{x}' \in \mathbb{Z}^{n_2}$, then

$$\mathbf{E} \left[\prod_{l=1}^{n_1} \omega(\mathbf{i}_l, \mathbf{x}_l) \prod_{l=1}^{n_2} \omega(\mathbf{i}'_l, \mathbf{x}'_l) \right] = 0.$$

because there is necessarily one ω term that is distinct from the others, and its independence from the rest implies zero expectation.

Hence, iv is clearly true if g_1, g_2 have compact supports. The extension to non-compactly-supported functions can also be obtained by a density argument as above. \square

Now, to characterize rigorously the limit of the U -statistics $(S_n^N, N \geq 1)$, we need to introduce the Wiener chaos.

3.2 Wiener chaos

3.2.1 White noise and stochastic integration on $[0, 1] \times \mathbb{R}$

This section recalls the elementary theory of white noise and stochastic integration on the measure space $([0, 1] \times \mathbb{R}, \mathcal{B}, dt \otimes dx)$. Here \mathcal{B} is the Borel σ -algebra, and $dt \otimes dx$ denotes Lebesgue measure on $[0, 1] \times \mathbb{R}$. For more details on Wiener chaos, we invite readers to read [17, Chapter 1] or [13, Chapter 11].

Let \mathcal{B}_f be the collection of all Borel sets of $[0, 1] \times \mathbb{R}$ with finite Lebesgue measure. Observe that $\mathcal{B} = \sigma(\mathcal{B}_f)$.

Definition 3.5. A *white noise* on $[0, 1] \times \mathbb{R}$ is a collection of mean zero Gaussian random variables indexed by \mathcal{B}_f

$$W = \{W(A) : A \in \mathcal{B}_f\}$$

such that for any $h \in \mathbb{N}$ and every finite collection (A_1, A_2, \dots, A_h) of elements of \mathcal{B}_f , the tuple $(W(A_1), \dots, W(A_h))$ is a h -dimensional Gaussian vector, with mean zero and covariance structure:

$$\mathbf{E}[W(A)W(B)] = |A \cap B|.$$

So in particular, if A and B are disjoint then $W(A)$ and $W(B)$ are independent.

For any $g \in L^2([0, 1] \times \mathbb{R}, \mathcal{B}, dt \otimes dx)$, the stochastic integral

$$I_1(g) := \int_0^1 \int_{\mathbb{R}} g(t, x) W(dt \otimes dx)$$

is constructed by first defining I_1 on simple functions then extending I_1 via density arguments [13, p.210]. In the end, for each $g \in L^2([0, 1] \times \mathbb{R})$, we

have that $I_1(g) \sim N(0, \|g\|_2^2)$, so in particular, I_1 preserves the Hilbert space structure of $L^2([0, 1] \times \mathbb{R})$,

$$\mathbf{E}(I_1(g)I_1(h)) = \int_0^1 \int g(t, x)h(t, x)dt dx$$

This construction idea can be extended to higher dimensions (see [17, p. 9,10]) to give a sense of the following notation of multiple stochastic integrals for any $n > 1$ and function $g \in L^2([0, 1]^n \times \mathbb{R}^n)$:

$$I_n(g) := \int_{[0,1]^n} \int_{\mathbb{R}^n} g(\mathbf{t}, \mathbf{x}) W^{\otimes n}(d\mathbf{t}d\mathbf{x}),$$

where $W^{\otimes n}(d\mathbf{t}d\mathbf{x}) := W(dt_1dx_1)W(dt_2dx_2) \cdots W(dt_ndx_n)$.

However, the mapping $I_n : L^2([0, 1]^n \times \mathbb{R}^n) \rightarrow L^2(\mathbf{P})$ is no longer injective. For example, if A, B are two disjoint compacts of $[0, 1] \times \mathbb{R}$, we see that $I_2(\mathbb{1}_{A \times B}) = W(A)W(B) = I_2(\mathbb{1}_{B \times A})$ even though $\mathbb{1}_{A \times B} \neq \mathbb{1}_{B \times A}$. Nonetheless, we observe that the restriction of I_n on the subspace $L_{\text{sym}}^2([0, 1]^n \times \mathbb{R}^n)$ (see Definition 3.6) of $L^2([0, 1]^n \times \mathbb{R}^n)$ is an isometry [17, p. 9,10].

Definition 3.6. A function $g \in L^2([0, 1]^n \times \mathbb{R}^n)$ is said to be symmetric if $g(\mathbf{t}, \mathbf{x}) = g(\pi\mathbf{t}, \pi\mathbf{x})$ for all $(\mathbf{t}, \mathbf{x}) \in [0, 1]^n \times \mathbb{R}^n$, permutation π on $\{1, \dots, n\}$, where $\pi\mathbf{t} := \mathbf{t}_{\pi(1)}, \dots, \mathbf{t}_{\pi(n)}$, $\pi\mathbf{x} := \mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(n)}$.

The set $L_{\text{sym}}^2([0, 1]^n \times \mathbb{R}^n)$ is then defined as the subspace of all symmetric functions of $L^2([0, 1]^n \times \mathbb{R}^n)$.

Notation 3.7. For any $n \in \mathbb{N}$, $g \in L^2([0, 1]^n \times \mathbb{R}^n)$, we denote by $\text{Sym } g(\mathbf{t}, \mathbf{x})$ the symmetrization of g defined by:

$$\text{Sym } g(\mathbf{t}, \mathbf{x}) := \frac{1}{n!} \sum_{\pi \text{ is a permutation of } \{1, \dots, n\}} g(\pi\mathbf{t}, \pi\mathbf{x}). \quad (8)$$

In summary, we have the following theorem which is a standard result in the theory of stochastic integration:

Theorem 3.8. There exists a continuous linear mapping $I_n : L^2([0, 1]^n \times \mathbb{R}^n) \rightarrow L^2(\mathbf{P})$ such that for any n -tuple of disjoint finite measurable sets A_1, A_2, \dots, A_n in $\mathcal{B}([0, 1] \times \mathbb{R})$:

$$I_n(\mathbb{1}_{A_1 \times A_2 \cdots \times A_n}) = W(A_1)W(A_2) \cdots W(A_n)$$

Furthermore, for all $g \in L^2([0, 1]^n \times \mathbb{R}^n)$,

$$\mathbf{E} [I_n(g)^2] \leq \|g\|_2^2, \quad (9)$$

and the equality occurs if and only if g is symmetric.

Proof. The first part is a summary of the construction of multiple stochastic integration in [17, p.8,9]. For the inequality, observe that:

$$\begin{aligned} \mathbf{E} [I_n(g)^2] &= \mathbf{E} [I_n(\text{Sym } g)^2] \stackrel{\text{Isometry}}{=} \|\text{Sym } (g)\|_2^2 \\ &= \left\| \frac{1}{n!} \sum_{\pi \text{ is a permutation}} g \circ \pi \right\|_2^2 \\ &\leq \frac{1}{n!} \sum_{\pi \text{ is a permutation}} \|g \circ \pi\|_2^2 = \|g\|_2^2 \end{aligned}$$

by Cauchy-Schwarz's inequality. Here, by abuse of notation, $(g \circ \pi)(\mathbf{t}, \mathbf{x})$ denotes $g(\pi\mathbf{t}, \pi\mathbf{x})$. \square

3.2.2 Wiener chaos on $[0, 1] \times \mathbb{R}$

This section provides a short introduction to the Wiener chaos's theory.

Wiener chaos may be regarded as a way of representing random variables as infinite sums of multiple stochastic integrals.

For a white noise W , we denote by \mathcal{F}_W the complete σ -algebra generated by random variables $(W(A), A \in \mathcal{B}_f)$, the Wiener chaos decomposition theorem states (see [17, Theorem 1.1.2]):

Proposition 3.9. (*Wiener chaos decomposition*) *For every random variable $X \in L^2(\Omega, \mathcal{F}_W, \mathbf{P})$, there is a unique sequence of symmetric functions $g_n \in L^2_{\text{Sym}}([0, 1]^n \times \mathbb{R}^n)$, $n \geq 1$, such that:*

$$X = \sum_{n=0}^{\infty} I_n(g_n).$$

Here g_0 is simply a constant and I_0 is the identity mapping on the constants.

In fact, for $n \geq 1$, the terms of the chaos series are all mean zero, so g_0 must be the mean of X . Moreover, by the orthogonality of $I_{n_1}(g_1)$ and $I_{n_2}(g_2)$ for $n_1 \neq n_2$ (see [17, p.9]), we have the relation:

$$\mathbf{E}[X^2] = \sum_{n=0}^{\infty} \|g_n\|_2^2.$$

Now, we define two important spaces of collections of functions:

Definition 3.10. *The Fock space over $L^2([0, 1] \times \mathbb{R})$ is defined to be the Hilbert space:*

$$F := \left\{ \mathbf{g} = (g_0, g_1, \dots) \in \bigoplus_{n=0}^{\infty} L^2([0, 1]^n \times \mathbb{R}^n) : \sum_{n=0}^{\infty} \|g_n\|_2^2 < \infty \right\} \quad (10)$$

equipped with the inner product $\langle \mathbf{g}, \mathbf{f} \rangle_F = \sum_{n=0}^{\infty} \langle \mathbf{g}_n, \mathbf{f}_n \rangle_{L^2([0,1]^n \times \mathbb{R}^n)}$. Then, the symmetric Fock space F_{sym} is defined as the Hilbert subspace of F that contains only collections of symmetric functions, i.e.,

$$F_{\text{sym}} := F \cap \left(\bigoplus_{n=0}^{\infty} L^2_{\text{sym}}([0,1]^n \times \mathbb{R}^n) \right).$$

The result in Proposition 3.9 works also in reverse, that is, the mapping

$$\begin{aligned} \mathbf{I}: F_{\text{sym}} &\longrightarrow L^2(\Omega, \mathcal{F}_W, \mathbf{P}) \\ (g_0, g_1, \dots) &\longmapsto \sum_{n \geq 0} I_n(g_n) \end{aligned}$$

is an isometry. This fact will be useful for the justification for the well-posedness of \mathcal{Z}_a in Section 4.

3.3 Limit theorems for U -statistics

In this section, we prove two limit theorems (Theorem 3.11 and Theorem 3.12) for our U -statistics S_n^N defined by (7). They are extensions of Theorem 4.3 and Lemma 4.4 in [2] with non-constant A_n . The result of the second theorem will be useful for the rest of this paper while the first is crucial for the proof of the second.

Theorem 3.11. (*Convergence of U -statistics to Stochastic Integrals*)

Suppose the functions A_1, A_2, \dots in the definition 3.3 of the U -statistics satisfy the following conditions:

i. $\sup_N \|A_N\|_{\infty} < +\infty$.

ii. There is a measurable function $a \in L^{\infty}([0,1] \times \mathbb{R})$ such that:

$$\lim_{N \rightarrow +\infty} A_N([t, x]_N) = a(t, x) \quad \text{a.e.}$$

Then, for any positive integer n and function $g \in L^2([0,1]^n \times \mathbb{R}^n)$, we have:

$$N^{-3n/4} \mathcal{S}_n^N(g) \xrightarrow[N \rightarrow +\infty]{(d)} \int_{[0,1]^n} \int_{\mathbb{R}^n} g(\mathbf{t}, \mathbf{x}) a^{\otimes n}(\mathbf{t}, \mathbf{x}) W^{\otimes n}(d\mathbf{t} d\mathbf{x}).$$

Moreover, for any finite collection of $n_1, \dots, n_m \in \mathbb{N}_0$ and g_1, \dots, g_m with $g_i \in L^2([0,1]^{n_i} \times \mathbb{R}^{n_i})$, one has the joint convergence

$$(N^{-3n_1/4} \mathcal{S}_{n_1}^N(g_1), \dots, N^{-3n_m/4} \mathcal{S}_{n_m}^N(g_m)) \xrightarrow[N \rightarrow +\infty]{(d)} (\tilde{I}_{n_1}(g_1), \dots, (\tilde{I}_{n_m}(g_m)).$$

where, for $n \geq 1$,

$$\tilde{I}_n(g) := \int_{[0,1]^n} \int_{\mathbb{R}^n} g(\mathbf{t}, \mathbf{x}) a^{\otimes n}(\mathbf{t}, \mathbf{x}) W^{\otimes n}(d\mathbf{t} d\mathbf{x}), \quad (11)$$

and

$$a^{\otimes n}(\mathbf{t}, \mathbf{x}) = a(\mathbf{t}_1, \mathbf{x}_1) \cdots a(\mathbf{t}_n, \mathbf{x}_n). \quad (12)$$

Theorem 3.12. Suppose the functions A_1, A_2, \dots in the definition 3.3 of the U -statistics satisfy the following conditions:

- i. $\sup_N \|A_N\|_\infty < c$ for some $c > 0$.
- ii. There is a measurable function $a \in L^\infty([0, 1] \times \mathbb{R})$ such that:

$$\lim_{N \rightarrow +\infty} A_N([t, x]_N) = a(t, x) \quad a.e.$$

Then if $(g_n, n \in \mathbb{N}_0)$ is a sequence of functions such that $(c^n g_n, n \in \mathbb{N}_0)$ belongs to the Fock space F , we have:

$$\sum_{n=0}^{\infty} N^{-3n/4} \mathcal{S}_n^N(g_n) \xrightarrow[N \rightarrow \infty]{(d)} \sum_{n=0}^{\infty} \int_{[0,1]^n \times \mathbb{R}^n} g_n(\mathbf{t}, \mathbf{x}) a^{\otimes n}(\mathbf{t}, \mathbf{x}) W^{\otimes n}(d\mathbf{t} d\mathbf{x}).$$

Remark 3.13. Notice that by the definition, $(\mathbf{i}, \mathbf{z}) \mapsto A_n(\mathbf{i}, \mathbf{z})$ is symmetric, hence $\mathcal{S}_n^N(g_n) = \mathcal{S}_n^N(\text{Sym}(g_n))$. This allows us to only consider symmetric functions in the proof of Theorem 3.12.

Besides, to prove the above two theorems, we will repeatedly use the following lemma in Billingsley [6, Theorem 3.2]:

Lemma 3.14. For $n, N \in \mathbb{N}$, let Y_n^N, Y_n, Y^N, Y be real-valued random variables defined on a common probability space such that $Y_n \xrightarrow[n \rightarrow +\infty]{(d)} Y$ and that for all n , $Y_n^N \xrightarrow[N \rightarrow +\infty]{(d)} Y_n$. If for each $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbf{P}(|Y_n^N - Y^N| \geq \epsilon) = 0$$

Then $Y^N \xrightarrow[N \rightarrow \infty]{(d)} Y$.

Proof of Theorem 3.11. Let $c := \sup_N \|A_N\|_\infty$.

Step 1: Let $n = 1$ and assume that g is a continuous and compactly supported function.

Rewrite $N^{-3/4} \mathcal{S}_1^N(g)$ as a weighted sum of elements in ω :

$$N^{-3/4} \mathcal{S}_1^N(g) = \sum_{i \in [1, N]} \sum_{z \in \mathbb{Z}} \mathbb{1}_{i \leftrightarrow z} \cdot 2^{1/2} N^{-3/4} \bar{g}_N \left(\frac{i}{N}, \frac{z}{\sqrt{N}} \right) A_N(i, z) \omega(i, z). \quad (13)$$

Because, g has compact support, the number of nonzero terms in the above sum is finite. Besides, recall that ω is a collection of independent random variables having zero mean and variance 1, one has:

$$\begin{aligned} \mathbf{E} \left[N^{-3/2} \mathcal{S}_1^N(g)^2 \right] &= 2N^{-3/2} \sum_{i \in [1, N]} \sum_{\substack{z \in \mathbb{Z} \\ i \leftrightarrow x}} \overline{g_N} \left(\frac{i}{N}, \frac{z}{\sqrt{N}} \right)^2 A_N(i, z)^2 \\ &= \int_{[0,1]} \int_{\mathbb{R}} \overline{g_N}^2(t, x) A_N([t, x]_N)^2 dt dx \xrightarrow{N \rightarrow +\infty} \int_{[0,1]} \int_{\mathbb{R}} g^2 a^2(t, x) dt dx. \end{aligned}$$

where the last convergence follows from the dominated convergence theorem and the fact that g is continuous and compactly supported.

Hence, by Lindeberg-Feller's central limit theorem (see [10, Theorem 3.4.5]),

$$N^{-3/4} \mathcal{S}_1^N(g) \xrightarrow[N \rightarrow +\infty]{(d)} \mathcal{N}(0, \int_{[0,1] \times \mathbb{R}} g^2 a^2(t, x) dt dx),$$

Note that the second condition of Lindeberg-Feller is satisfied because the supremum of all the terms in Equation (13) is smaller or equal to

$$2^{1/2} N^{-3/4} \|g\|_{\infty} \|A_N\|_{\infty},$$

which converges to 0 when $N \rightarrow \infty$. Thus, by the isometry of the stochastic integration I_1 , we have proved that:

$$N^{-3/4} \mathcal{S}_1^N(g) \xrightarrow[N \rightarrow +\infty]{(d)} \int_{[0,1]} \int_{\mathbb{R}} g(t, x) a(t, x) W(dt dx) = \tilde{I}_1(g).$$

Step 2: Let $n = 1$ and g be any function in $L^2([0, 1] \times \mathbb{R})$.

Because the space of continuous and compactly supported functions $C_c([0, 1] \times \mathbb{R})$ is dense in $L^2([0, 1] \times \mathbb{R})$ [7, Theorem 4.12], there exists a sequence of continuous and compactly supported functions $(g_m, m \in \mathbb{N}_0)$ converging to g in L^2 . Thus, by combining with the fact that $\|a\|_{\infty} \leq c$, this implies

$$\tilde{I}_1(g_m) \xrightarrow[m \rightarrow \infty]{L^2} \tilde{I}_1(g).$$

Besides, for all N , observe that:

$$\mathbf{E} \left[N^{-3/2} (\mathcal{S}_1^N(g - g_m))^2 \right] \leq c^2 \|g - g_m\|_2^2.$$

So the above observations and Step 1 give the following diagram:

$$\begin{array}{ccc} N^{-3/4} \mathcal{S}_1^N(g_m) & \xrightarrow[N \rightarrow +\infty]{(d)} & \tilde{I}_1(g_m) \\ \text{in } L^2(\mathbf{P}), \text{ uniformly in } N \downarrow m \rightarrow +\infty & & m \rightarrow +\infty \downarrow L^2 \\ N^{-3/4} \mathcal{S}_1^N(g) & & \tilde{I}_1(g) \end{array}$$

Thus thanks to Lemma 3.14, we duce that $N^{-3/4}\mathcal{S}_1^N(g) \xrightarrow[N \rightarrow +\infty]{(d)} \tilde{I}_1(g)$.

Step 3: Now we will prove that for all $m \in \mathbb{N}$ and m functions $g_1, g_2, \dots, g_m \in L^2([0, 1] \times \mathbb{R})$, we have:

$$\left(N^{-3/4}\mathcal{S}_1^N(g_1), \dots, N^{-3/4}\mathcal{S}_1^N(g_m)\right) \xrightarrow[N \rightarrow \infty]{(d)} (\tilde{I}_1(g_1), \dots, \tilde{I}_1(g_m)). \quad (14)$$

Indeed, for any m real numbers $\alpha_1, \dots, \alpha_m$, our result in *Step 1* shows that:

$$\alpha_1 N^{-3/4}\mathcal{S}_1^N(g_1) + \dots \alpha_m N^{-3/4}\mathcal{S}_1^N(g_m) \xrightarrow[N \rightarrow \infty]{(d)} \alpha_1 \tilde{I}_1(g_1) + \dots \alpha_m \tilde{I}_1(g_m).$$

Thus, by Cramer-Wold Theorem [13, Corollary 4.5], the convergence (14) is valid.

Step 4: For $n > 1$ and g of the form:

$$g(\mathbf{t}, \mathbf{x}) = g^{(1)}(\mathbf{t}_1, \mathbf{x}_1)g^{(2)}(\mathbf{t}_2, \mathbf{x}_2) \cdots g^{(n)}(\mathbf{t}_n, \mathbf{x}_n). \quad (15)$$

where $g^{(j)}$ are functions of $L^2([0, 1] \times \mathbb{R})$ with disjoint compact supports. As the supports of $g^{(j)}$ are disjoint compacts, if N is large enough, the supports of $\overline{g_N^{(j)}}$ are also disjoint. Thus we have the first equality in the following argument:

$$N^{-3n/4}\mathcal{S}_n^N(g) = \prod_{j=1}^n N^{-3/4}\mathcal{S}_1^N(g^{(j)}) \xrightarrow[N \rightarrow \infty]{(d)} \prod_{j=1}^n \tilde{I}_1(g^{(j)}) = \tilde{I}_n(g).$$

The latter limit in law is obtained by using the convergence of the joint random variables $(N^{-3/4}\mathcal{S}_1^N(g^{(1)}), \dots, N^{-3/4}\mathcal{S}_1^N(g^{(n)}))$.

Step 5: Now, for any m -tuple of functions g_1, g_2, \dots, g_m of the form (15), by a similar argument, one can show the joint convergence:

$$\left(N^{-3n/4}\mathcal{S}_n^N(g_1), \dots, N^{-3n/4}\mathcal{S}_n^N(g_m)\right) \xrightarrow[N \rightarrow +\infty]{(d)} (\tilde{I}_n(g_1), \dots, \tilde{I}_n(g_m)).$$

Hence, by the linearity of \mathcal{S} and I_n , for any linear combination g of functions of the form (15), one has the convergence $N^{-3n/4}\mathcal{S}_n^N(g) \xrightarrow[N \rightarrow +\infty]{(d)} \tilde{I}_n(g)$.

Besides, the space of such linear combinations is dense in $L^2([0, 1]^n \times \mathbb{R}^n)$ (the space of step functions is dense in $L^2([0, 1]^n \times \mathbb{R}^n)$ [7, Proof of Theorem 4.13], then we shrink the support of each step function in an appropriate way), by a same density argument as in *Step 1*, one can conclude that for any $g \in L^2([0, 1]^n \times \mathbb{R}^n)$:

$$N^{-3n/4}\mathcal{S}_n^N(g) \xrightarrow[N \rightarrow +\infty]{(d)} \tilde{I}_n(g).$$

The proof for the desired joint convergence for different n is just a repeat of *Step 4* and *Step 5*. \square

Proof of Theorem 3.12. Without loss of generality, we assume $c = 1$. Using Cauchy-Schwarz's inequality as in the proof of Theorem 3.8, one can show that:

$$\text{Sym } \mathbf{g} := (\text{Sym } g_0, \text{Sym } g_1, \text{Sym } g_2, \dots) \in F_{\text{Sym}}.$$

By the symmetry of \mathcal{S}_n^N and the stochastic integrations $(\tilde{I}_n, n \in \mathbb{N}_0)$, without loss of generality, we can suppose that $g \in F_{\text{Sym}}$ (see Remark 3.13). Since \mathbf{I} is an isometry from the symmetric Fock space into $L^2(\mathbf{P})$, we automatically get that $\sum_{n=0}^M \tilde{I}_n(g_n) \rightarrow \sum_{n \geq 0} \tilde{I}_n(g_n)$ in $L^2(\mathbf{P})$, as M goes to infinity. Since $\text{Var}(N^{-3n/4} \mathcal{S}_n^N(g_n)) \leq \|g_n\|_2^2$ (see Proposition 3.4), this also implies

$$\sum_{n=0}^M N^{-3n/4} \mathcal{S}_n^N(g_n) \xrightarrow{M \rightarrow +\infty} \sum_{n=0}^{\infty} N^{-3n/4} \mathcal{S}_n^N(g_n)$$

in $L^2(\mathbf{P})$, uniformly in N . Theorem 3.11 implies that:

$$\sum_{n=0}^M N^{-3n/4} \mathcal{S}_n^N(g_n) \xrightarrow[N \rightarrow \infty]{(d)} \sum_{n=0}^M \tilde{I}_n(g_n).$$

Combining these asymptotic results, we have by Lemma 3.14 the following diagram:

$$\begin{array}{ccc} \sum_{n=0}^M N^{-3n/4} \mathcal{S}_n^N(g_n) & \xrightarrow[N \rightarrow +\infty]{(d)} & \sum_{n=0}^M \tilde{I}_n(g_n) \\ \text{in } L^2, \text{ uniformly in } N \downarrow M \rightarrow +\infty & & M \rightarrow +\infty \downarrow \text{in } L^2 \\ \sum_{n=0}^{\infty} N^{-3n/4} \mathcal{S}_n^N(g_n) & \xrightarrow[N \rightarrow +\infty]{(d)} & \sum_{n=0}^{\infty} \tilde{I}_n(g_n). \end{array}$$

□

4 Limit theorems for partition functions

In this section, we study the convergence of partition functions \mathfrak{Z}_N (Definition 2.1). First, we verify the well-posedness of the limit value \mathcal{Z}_a given in Theorem 1.3, for all $a \in L^\infty([0, 1] \times \mathbb{R})$ by

$$\mathcal{Z}_a := 1 + \sum_{n=1}^{\infty} \int_{\Delta_n} \int_{\mathbb{R}^n} \prod_{i=1}^n \left[a(\mathbf{t}_i, \mathbf{x}_i) \varrho(\mathbf{t}_i - \mathbf{t}_{i-1}, \mathbf{x}_i - \mathbf{x}_{i-1}) W(d\mathbf{t}_i, d\mathbf{x}_i) \right].$$

where ϱ is the gaussian kernel, and W is the white noise based on the Lebesgue measure on $[0, 1] \times \mathbb{R}$.

4.1 Study of \mathcal{Z}_a

4.1.1 Brownian motion and simple random walk

Let $(S_n, n \in \mathbb{N}_0)$ denote a simple random walk on \mathbb{Z} and $(B_t, t \in \mathbb{R}_{\geq 0})$ denote a Brownian motion on \mathbb{R} .

For $i \in \mathbb{N}$, $t \geq 0$ and $x \in \mathbb{R}$, we define:

$$p(i, x) := \mathbf{P}(S_i = x) \quad \varrho(t, x) := \frac{e^{-x^2/2t}}{\sqrt{2\pi t}} \quad (16)$$

We will make heavy use of the finite dimensional distributions of both simple random walk and Brownian motion. For notations, we introduce for $n \in \mathbb{N}$, $\mathbf{i} \in D_n^N$ (D_n^N being the integer simplex (6)), $\mathbf{z} \in \mathbb{Z}^n$, $\mathbf{t} \in \Delta_n$ (Δ_n being the real simplex (3)), $\mathbf{x} \in \mathbb{R}^n$:

$$p_n(\mathbf{i}, \mathbf{z}) := \prod_{j=1}^n p(\mathbf{i}_j - \mathbf{i}_{j-1}, \mathbf{z}_j - \mathbf{z}_{j-1}) = \mathbf{P}(S_{\mathbf{i}_1} = \mathbf{z}_1, \dots, S_{\mathbf{i}_n} = \mathbf{z}_n), \quad (17)$$

and

$$\varrho_n(\mathbf{t}, \mathbf{x}) := \prod_{j=1}^n \varrho(\mathbf{t}_j - \mathbf{t}_{j-1}, \mathbf{x}_j - \mathbf{x}_{j-1}). \quad (18)$$

For convenience, we respectively extend the domains of p_n and ϱ to $\llbracket 1, N \rrbracket^n \times \mathbb{Z}^n$ and to $[0, 1]^n \times \mathbb{R}^n$ by letting p_n and ϱ_n to be zero outside $D_n^N \times \mathbb{Z}^n$ and $\Delta_n \times \mathbb{R}^n$.

4.1.2 Wiener chaos for Brownian transition probabilities

The Brownian transition probabilities can generate many elements in the Fock space F (see Definition 3.10). Let us recall here Notation (12) of a^{\otimes} .

Proposition 4.1. *For every measurable bounded function $a \in L^\infty([0, 1] \times \mathbb{R})$, let $\varrho(a) := (1, a\varrho_1, a^{\otimes 2}\varrho_2, a^{\otimes 3}\varrho_3, \dots)$. be a weighted ordered collection (indexed by \mathbb{N}_0) of Brownian transition probabilities ϱ_n that depends on a . Then, $\varrho(a)$ is an element in the Fock space F , i.e., $\sum_{n \geq 0} \|a^{\otimes n}\varrho_n\|_2^2 < \infty$.*

Remark 4.2. *In particular, if a is a constant function, that is, a is equal to some constant β then $\varrho(\beta) = (1, \beta\varrho_1, \beta^2\varrho_2, \beta^3\varrho_3, \dots)$.*

Proof of Proposition 4.1. Recall that a is a bounded function, then there is a positive number β such that $\|a\|_\infty \leq \beta$. Hence, $\sum_{n \geq 0} \|a^{\otimes n}\varrho_n\|_2^2 \leq \sum_{n \geq 0} \beta^{2n} \|\varrho_n\|_2^2$. Thus, it suffices to prove that $\varrho(\beta)$ belongs to the Fock space. Indeed, observe that when $\mathbf{t} \in \Delta_n$,

$$\varrho_n(\mathbf{t}, \mathbf{x})^2 = \varrho_n(\mathbf{t}, \mathbf{x}\sqrt{2}) \prod_{j=1}^n \frac{1}{\sqrt{2\pi(\mathbf{t}_j - \mathbf{t}_{j-1})}}.$$

Hence,

$$\begin{aligned}
\int_{[0,1]^n} \int_{\mathbb{R}^n} \varrho_n(\mathbf{t}, \mathbf{x})^2 d\mathbf{x} d\mathbf{t} &= \int_{\Delta_n} \int_{\mathbb{R}^n} \varrho_n(\mathbf{t}, \mathbf{x})^2 d\mathbf{x} d\mathbf{t} \\
&= (4\pi)^{-n/2} \int_{\Delta_n} \prod_{j=1}^n \frac{1}{\sqrt{\mathbf{t}_j - \mathbf{t}_{j-1}}} d\mathbf{t} \\
&= (4\pi)^{-n/2} B\left(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}, 1\right) = \frac{1}{2^n \Gamma((n/2) + 1)}.
\end{aligned}$$

where B is the Beta function and Γ is the Gamma function.

The second to last equality comes from recognizing that the integrand is the density of the Dirichlet distribution, for which the beta function B is the normalizing constant.

Besides, Gamma function converges extremely rapid to infinity, faster than any exponential functions [1, Sterling's formula, 6.1.37]. Hence, the decay of the above expression shows that $\|\varrho(\beta)\|_F^2 < \infty$ for all $\beta \in \mathbb{R}_+$. \square

So naturally, we have the following corollary on the well-posedness of \mathcal{Z}_a .

Proposition 4.3. *For all measurable bounded function a on $[0, 1] \times \mathbb{R}$, the Wiener chaos \mathcal{Z}_a is well-defined and has the representation $\mathcal{Z}_a = \mathbf{I}(\varrho(a))$.*

4.2 Relation between \mathfrak{Z} and U -statistics

We begin with establishing the relation between partition functions \mathfrak{Z} and U -statistics, then we will prove Theorem 2.3.

For convenience, we extend Notation 2.2 $[t, x]_N$ for a pair $(t, x) \in [0, 1] \times \mathbb{R}$ to higher dimensions:

Notation 4.4. *For any pair $(\mathbf{t}, \mathbf{x}) \in (0, 1]^n \times \mathbb{R}^n$, we let $[\mathbf{t}, \mathbf{x}]_N$ denote the unique pair $(\mathbf{i}, \mathbf{z}) \in \llbracket 1, N \rrbracket^n \times \mathbb{Z}^n$ such that:*

$$i. (\mathbf{t}, \mathbf{x}) \in \left(\frac{\mathbf{i}-1}{N}, \frac{\mathbf{i}}{N}\right] \times \left(\frac{\mathbf{z}-1}{\sqrt{N}}, \frac{\mathbf{z}+1}{\sqrt{N}}\right],$$

ii. \mathbf{i} and \mathbf{z} have the same parity.

Definition 4.5. *For $n, N \geq 1$, define $p_n^N : [0, 1]^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ by*

$$p_n^N(\mathbf{t}, \mathbf{x}) = 2^{-n} p_n([\mathbf{t}, \mathbf{x}]_N) \mathbb{1}_{\lceil N\mathbf{t} \rceil \in D_n^N},$$

where $\lceil N\mathbf{t} \rceil$ is the usual ceiling function, that is, for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{Z}^n$, $\lceil \mathbf{x} \rceil = \mathbf{z}$ if and only if for all i , \mathbf{z}_i is the smallest integer bigger than or equal to \mathbf{x}_i .

We observe that the condition $\lceil N\mathbf{t} \rceil \in D_n^N$ implies that p_n^N is identically zero if $n > N$. Besides, we also see that p_n^N is constant on each rectangle

in \mathcal{R}_n^N , so the average $\overline{p_n^N} = p_n^N$ and in particular, for $\mathbf{i} \in E_n^N, \mathbf{z} \in \mathbb{Z}^n$ such that $\mathbf{i} \leftrightarrow \mathbf{z}$, we have:

$$p_n^N \left(\frac{\mathbf{i}}{N}, \frac{\mathbf{z}}{\sqrt{N}} \right) \mathbf{1}_{\mathbf{i} \in D_n^N} = 2^{-n} p_n(\mathbf{i}, \mathbf{z}) \mathbf{1}_{\mathbf{i} \in D_n^N}.$$

Thus, by definition of \mathcal{S}_n^N (see Definition 3.3) ,

$$\mathcal{S}_n^N(p_n^N) = 2^{-n/2} \sum_{\mathbf{i} \in D_n^N} \sum_{\mathbf{z} \in \mathbb{Z}^n} p_n(\mathbf{i}, \mathbf{z}) \omega(\mathbf{i}, \mathbf{z}) A_N(\mathbf{i}, \mathbf{z}).$$

Note that the condition $\mathbf{i} \leftrightarrow \mathbf{z}$ is already handled by p_n . This leads to the following relation:

Proposition 4.6. *For all number real $\beta \in \mathbb{R}$ and positive integer $N \in \mathbb{N}$, the partition functions \mathfrak{Z}_N can be rewritten as:*

$$\mathfrak{Z}_N(\beta A_N) = \sum_{n=0}^N 2^{n/2} \beta^n \mathcal{S}_n^N(p_n^N).$$

Remark 4.7. *So in particular,*

$$\mathfrak{Z}_N(N^{-1/4} A_N) = \sum_{n=0}^N 2^{n/2} N^{-3n/4} \mathcal{S}_n^N(N^{n/2} p_n^N).$$

This equality is useful for our proof of Theorem 2.3.

Proof of Proposition 4.6. By definition,

$$\begin{aligned} \mathfrak{Z}_N(\beta A_N) &= \mathbf{E} \left[\prod_{n=1}^N (1 + \beta A_N(n, S_n) \omega(n, S_n)) \middle| \omega \right] \\ &= \mathbf{E} \left[1 + \sum_{n=1}^N \sum_{\mathbf{i} \in D_n^N} \beta^n \left(\prod_{j=1}^n A_N(\mathbf{i}_j, S_{\mathbf{i}_j}) \right) \left(\prod_{j=1}^n \omega(\mathbf{i}_j, S_{\mathbf{i}_j}) \right) \middle| \omega \right] \\ &= \mathbf{E} \left[1 + \sum_{n=1}^N \sum_{\mathbf{i} \in D_n^N} \sum_{\mathbf{z} \in \mathbb{Z}^n} \beta^n \left(\prod_{j=1}^n \mathbb{1}_{S_{\mathbf{i}} = \mathbf{z}_j} \right) \left(\prod_{j=1}^n A_N(\mathbf{i}_j, \mathbf{z}_j) \right) \left(\prod_{j=1}^n \omega(\mathbf{i}_j, \mathbf{z}_j) \right) \middle| \omega \right] \\ &= 1 + \sum_{n=1}^N \sum_{\mathbf{i} \in D_n^N} \sum_{\mathbf{z} \in \mathbb{Z}^n} \beta^n \mathbf{E} \left[\prod_{j=1}^n \mathbb{1}_{S_{\mathbf{i}} = \mathbf{z}_j} \right] A_N(\mathbf{i}, \mathbf{z}) \omega(\mathbf{i}, \mathbf{z}) \\ &= 1 + \sum_{n=1}^N \sum_{\mathbf{i} \in D_n^N} \sum_{\mathbf{z} \in \mathbb{Z}^n} \beta^n p_n(\mathbf{i}, \mathbf{z}) A_N(\mathbf{i}, \mathbf{z}) \omega(\mathbf{i}, \mathbf{z}) = 1 + \sum_{n=1}^N \beta^n 2^{n/2} \mathcal{S}_n^N(p_n^N). \end{aligned} \tag{19}$$

Thus, $\mathfrak{Z}_N(\beta A_N) = 1 + \sum_{n=1}^N 2^{n/2} \beta^n \mathcal{S}_n^N(p_n^N)$. \square

Now, we are ready to give a proof of Theorem 2.3.

Proof of Theorem 2.3. First observe that Theorem 3.12 and Proposition 4.1 imply that :

$$\sum_{n=0}^{\infty} N^{-3n/4} S_n^N(2^{n/2} \varrho_n) \xrightarrow{(d)} \sum_{n=0}^{\infty} \tilde{I}_n(\varrho_n 2^{n/2}) = \mathbf{I}(\varrho(\sqrt{2}a)) = \mathcal{Z}_{\sqrt{2}a}.$$

as N converges to infinity. Now we show that the difference between this term and $\mathfrak{Z}_N(N^{-1/4} A_N \omega)$ goes to zero as N converges to infinity. Observe that:

$$\begin{aligned} & \sum_{n=0}^{\infty} N^{-3n/4} S_n^N(2^{n/2} \varrho_n) - \mathfrak{Z}_N(N^{-1/4} A_N \omega) \\ &= \sum_{n=0}^N 2^{n/2} N^{-3n/4} S_n^N(\varrho_n - N^{n/2} p_n^N) + \sum_{n=N+1}^{\infty} N^{-3n/4} S_n^N(2^{n/2} \varrho_n). \end{aligned}$$

By Proposition 3.4, the second term is bounded in L^2 by the square root of

$$\sum_{n=N+1}^{\infty} 2^n c^{2n} \|\varrho_n\|_2^2.$$

which goes to zero as $N \rightarrow \infty$ by Proposition 4.1.

For the first term, using again Proposition 3.4, we note that its L^2 -norm is bounded above by the square root of

$$\sum_{n=0}^N 2^n c^{2n} \|\varrho_n - N^{n/2} p_n^N\|_2^2.$$

From Lemma 4.8 below, we know that there is a constant $C > 0$ such that for all $n \in \mathbb{N}$,

$$\sup_N \|N^{n/2} p_n^N\|_2 \leq C^n \|\varrho_n\|_2 \quad \text{and} \quad \lim_{N \rightarrow \infty} \|\varrho_k - N^{n/2} p_n^N\|_2^2 = 0.$$

Since by Proposition 4.1, the sequence $2^{n+1} c^n (1 + C^n) \|\varrho_n\|_2^2$ is summable, by the dominated convergence theorem, we can easily deduce that:

$$\lim_{N \rightarrow +\infty} \sum_{n=0}^N 2^n c^{2n} \|\varrho_n - N^{n/2} p_n^N\|_2^2 = \sum_{n=0}^{\infty} \lim_{N \rightarrow +\infty} 2^n c^{2n} \|\varrho_n - N^{n/2} p_n^N\|_2^2 = 0.$$

Theorem 2.3 is therefore proved. \square

Lemma 4.8. *For all n , we have the L^2 -convergence:*

$$\lim_{N \rightarrow +\infty} \|\varrho_n - N^{n/2} p_n^N\|_2 = 0,$$

and moreover, there exists a constant C such that for all $n \in \mathbb{N}$,

$$\sup_N \|N^{n/2} p_n^N\|_2 \leq C^n \|\varrho_n\|_2.$$

Proof. From the local central limit theorem, we deduce that for any fixed $n \in \mathbb{N}$, $N^{n/2}p_n^N$ converges almost surely to ϱ_n as N goes to infinity. So by the general Lebesgue dominated convergence theorem [21, Theorem 19], to prove our L^2 convergence, it suffices to find a function $g \in L^2([0, 1]^n \times \mathbb{R}^n)$ and a sequence $(g_N, N \in \mathbb{N})$ of functions in $L^2([0, 1]^n \times \mathbb{R}^n)$ such that:

- i. $(N^{n/2}p_n^N)^2 \leq g_N$ for all N .
- ii. g_N converges pointwise to g when N converges to infinity.
- iii. $\lim_{N \rightarrow \infty} \int_{[0, 1]^n \times \mathbb{R}^n} g_N = \int_{[0, 1]^n \times \mathbb{R}^n} g < \infty$.

By Definition (16) of p and Sterling's formula (see [1, Sterling's formula, 6.1.37]), we observe that there exists a constant C such that $\sqrt{ip(i, x)} \leq C$ for all i and x , therefore:

$$\sup_{\mathbf{z} \in \mathbb{Z}^n} p_n(\mathbf{i}, \mathbf{z}) \leq C^n \prod_{j=1}^n \frac{1}{\sqrt{\mathbf{i}_j - \mathbf{i}_{j-1}}}.$$

From this and by Definition 4.5 of p_n^N , we have:

$$\left(N^{n/2} p_n^N(\mathbf{t}, \mathbf{x}) \right)^2 \leq (C/2)^n h \left(\frac{\lceil N\mathbf{t} \rceil}{N} \right) N^{n/2} p_n^N(\mathbf{t}, \mathbf{x}).$$

where $h(\mathbf{t}) = \prod_{j=1}^n \frac{1}{\sqrt{\mathbf{t}_j - \mathbf{t}_{j-1}}} \mathbb{1}_{\{\mathbf{t} \in \Delta_n\}}$.

Let us choose for all N the function

$$g_N(\mathbf{t}, \mathbf{x}) := (C/2)^n h \left(\frac{\lceil N\mathbf{t} \rceil}{N} \right) N^{n/2} p_n^N(\mathbf{t}, \mathbf{x}),$$

and let

$$g(\mathbf{t}, \mathbf{x}) := (C/2)^n h(\mathbf{t}) \varrho_n(\mathbf{t}, \mathbf{x}).$$

Clearly, the conditions i. and ii. for the generalized dominated convergence Theorem are satisfied. For the last condition, we first notice that:

$$\int_{[0, 1]^n \times \mathbb{R}^n} g(\mathbf{t}, \mathbf{x}) d\mathbf{t} d\mathbf{x} = (C/2)^n \int_{[0, 1]^n} h(\mathbf{t}) d\mathbf{t}.$$

Then by definition of p_n^N , we have the following equalities:

$$\begin{aligned}
& \int_{[0,1]^n \times \mathbb{R}^n} g_N(\mathbf{t}, \mathbf{x}) d\mathbf{t} d\mathbf{x} \\
&= \sum_{\substack{\mathbf{i} \in [1, N]^n, \mathbf{z} \in \mathbb{Z}^n: \\ \mathbf{i} \text{ and } \mathbf{z} \text{ have the same parity}}} \int_{\left(\frac{\mathbf{i}-1}{N}, \frac{\mathbf{i}}{N}\right] \times \left(\frac{\mathbf{z}-1}{\sqrt{N}}, \frac{\mathbf{z}+1}{\sqrt{N}}\right]} g_N(\mathbf{t}, \mathbf{x}) d\mathbf{t} d\mathbf{x} \\
&= \sum_{\substack{\mathbf{i} \in [1, N]^n, \mathbf{z} \in \mathbb{Z}^n: \\ \mathbf{i} \text{ and } \mathbf{z} \text{ have the same parity}}} \left(N^{-3n/2} 2^n\right) \left[(C/2)^n h\left(\frac{\mathbf{i}}{N}\right) 2^{-n} N^{n/2} p_n(\mathbf{i}, \mathbf{z}) \mathbb{1}_{\mathbf{i} \in D_n^N} \right] \\
&= (C/2)^n N^{-n} \sum_{\mathbf{i} \in D_n^N} \sum_{\mathbf{z} \in \mathbb{Z}^n} h\left(\frac{\mathbf{i}}{N}\right) p_n(\mathbf{i}, \mathbf{z}) \mathbb{1}_{\{\mathbf{i} \text{ and } \mathbf{z} \text{ have the same parity}\}} \\
&= (C/2)^n N^{-n} \sum_{\mathbf{i} \in D_n^N} h\left(\frac{\mathbf{i}}{N}\right) = (C/2)^n \int_{[0,1]^n} h\left(\frac{[N\mathbf{t}]}{N}\right) d\mathbf{t}.
\end{aligned}$$

So, what is left to do is prove that

$$\lim_{N \rightarrow \infty} \int_{[0,1]^n} h\left(\frac{[N\mathbf{t}]}{N}\right) d\mathbf{t} = \int_{[0,1]^n} h(\mathbf{t}) d\mathbf{t} \quad \text{and} \quad \int_{[0,1]^n} h(\mathbf{t}) d\mathbf{t} < \infty.$$

which is true because $h\left(\frac{[N\mathbf{t}]}{N}\right)$ converges pointwise to $h(\mathbf{t})$ for all \mathbf{t} and they form a uniformly integrable sequence of functions in $L^2([0,1] \times \mathbb{R})$. Indeed, the uniform integrability is due to the fact that:

$$\begin{aligned}
& \int_{\Delta_n} \left[h\left(\frac{[N\mathbf{t}]}{N}\right) \right]^{3/2} d\mathbf{t} = N^{-n} \sum_{\mathbf{i} \in D_n^N} \prod_{j=1}^n \left(\frac{\mathbf{i}_j}{N} - \frac{\mathbf{i}_{j-1}}{N} \right)^{-3/4} \\
&= \sum_{\mathbf{i} \in D_n^N} 2^{3n/4} \prod_{j=1}^n \left(\frac{2\mathbf{i}_j - 2\mathbf{i}_{j-1}}{N} \right)^{-3/4} \\
&\leq \sum_{\mathbf{i} \in D_n^N} \int_{[0,1]^n} 2^{3n/4} \prod_{j=1}^n \left(\frac{\mathbf{i}_j - \mathbf{s}_j}{N} - \frac{\mathbf{i}_{j-1} - \mathbf{s}_{j-1}}{N} \right)^{-3/4} d\mathbf{s} \quad (20) \\
&\leq 2^{3n/4} \int_{\Delta_n} \prod_{j=1}^n (\mathbf{t}_j - \mathbf{t}_{j-1})^{-3/4} d\mathbf{t} < \infty.
\end{aligned}$$

Note that in the inequality (20), we have used the fact that for all positive integer $m \geq 1$ and real numbers $a, b \in [0, 1]$: $2m \geq m + a - b > 0$.

For the inequality in the latter part of our lemma, by what we have proved

so far, we observe that:

$$\begin{aligned}
\|N^{n/2}p_n^N\|_2^2 &\leq (C/2)^n \int_{\Delta_n} h\left(\frac{[N\mathbf{t}]}{N}\right) d\mathbf{t} \\
&\leq (C/2)^n 2^{n/2} \int_{\Delta_n} \prod_{j=1}^n (\mathbf{t}_j - \mathbf{t}_{j-1})^{-1/2} d\mathbf{t} \\
&= C^n 2^{-n/2} \int_{\Delta_n} \int_{\mathbb{R}^n} (4\pi)^{n/2} [\varrho_n(\mathbf{t}, \mathbf{x})]^2 d\mathbf{t} d\mathbf{x} = C^n (2\pi)^{n/2} \|\varrho_n\|_2^2,
\end{aligned}$$

where the second inequality is obtained similarly as (20).
Hence, we have our desired conclusion. □

5 Asymptotics of collision measures

5.1 Convergence of exponential moments

We first prove Theorem 2.4 on the uniform boundedness of moments of the partition functions. Then we will study the convergence of the exponential moments of $(\frac{1}{\sqrt{N}}\Pi_N; N \in \mathbb{N})$.

Proof of Theorem 2.4. Let c be a positive number such that $c \geq \sup_N \|A_N\|_\infty$. Without loss of generality, assume N is sufficiently large (i.e. $N > c^4$) such that the partition function $\mathfrak{Z}_N\left(\frac{1}{N^{1/4}}A_N\right)$ is a positive random variable. Recall that in page 6, we have shown that:

$$\begin{aligned}
&\mathbf{E} \left[\mathfrak{Z}_N \left(\frac{1}{N^{1/4}} A_N \right)^k \right] = \\
&\mathbf{E} \left[\prod_{n=1}^N \mathbf{E} \left[\prod_{i=1}^k \left(1 + \frac{1}{N^{1/4}} A_N(n, S_n^{(i)}) \omega(n, S_n^{(i)}) \right) \middle| S^{(1)}, S^{(2)}, \dots, S^{(k)} \right] \right].
\end{aligned}$$

Now, define for $n \geq 1$:

$$X_{N,n} := \mathbf{E} \left[\prod_{i=1}^k \left(1 + \frac{1}{N^{1/4}} A_N(n, S_n^{(i)}) \omega(n, S_n^{(i)}) \right) \middle| S^{(1)}, S^{(2)}, \dots, S^{(k)} \right] - 1, \tag{21}$$

and

$$T_N := \sum_{n=1}^N X_{N,n}. \tag{22}$$

Because ω is a collection of independent Rademacher random variables, we easily notice that $X_{N,n} \geq 0$ $\mathbf{P} - a.s.$, since $X_{N,n}$ can be represented by:

$$X_{N,n} = \sum_{l=2}^k \sum_{1 \leq i_1 < i_2 < \dots < i_l \leq k} N^{-l/4} \prod_{h=1}^l A_N(n, S_n^{(i_h)}) \underbrace{\mathbf{E} \left[\prod_{h=1}^l \omega(n, S_n^{(i_h)}) \right]}_{\text{is either 0 or 1}}. \quad (23)$$

Consequently,

$$\mathbf{E} \left[3_N \left(\frac{1}{N^{1/4}} A_N \right)^k \right] = \mathbf{E} \left[\prod_{n=1}^N (1 + X_{N,n}) \right] \leq \mathbf{E} [e^{T_N}]. \quad (24)$$

where we have used the classical inequality that $\forall x \in \mathbb{R} : 1 + x \leq e^x$ and $X_{N,n} \geq 0$.

Then, for each n , let us introduce the number $U^{(n)}$ of pairs (i, j) such that $S_n^{(i)} = S_n^{(j)}$, i.e.,

$$U^{(n)} := \sum_{1 \leq i < j \leq k} \mathbf{1}_{S_n^{(i)} = S_n^{(j)}}.$$

We observe that on the event $\{U^{(n)} = 0\}$, $X_{N,n}$ is equal to zero, and on the event $\{U^{(n)} \geq 1\}$,

$$\sum_{1 \leq i_1 < i_2 < \dots < i_l \leq k} \mathbf{E} \left[\omega(n, S_n^{(i_1)}) \dots \omega(n, S_n^{(i_l)}) \right] \leq \binom{k}{l} \leq \binom{k}{l} U^{(n)}. \quad (25)$$

Thus,

$$X_{N,n} \leq \sum_{l=2}^k N^{-l/4} c^l \binom{k}{l} U^{(n)} \leq (c+1)^k N^{-1/2} U^{(n)}. \quad (26)$$

So by combining the inequalities (24) and (26), one sees that:

$$\begin{aligned} & \mathbf{E} \left[3_N \left(\frac{1}{N^{1/4}} A_N \omega \right)^k \right] \leq \mathbf{E} [e^{T_N}] \\ & \leq \mathbf{E} \left[\exp \left((c+1)^k N^{-1/2} \sum_{1 \leq i < j \leq k} \sum_{n=1}^N \mathbf{1}_{S_n^{(i)} = S_n^{(j)}} \right) \right] \\ & = \mathbf{E} \left[\prod_{(i,j): 1 \leq i < j \leq k} \exp \left((c+1)^k N^{-1/2} \sum_{n=1}^N \mathbf{1}_{S_n^{(i)} = S_n^{(j)}} \right) \right] \\ & \leq \mathbf{E} \left[\exp \left(\frac{k(k-1)}{2} (c+1)^k N^{-1/2} \sum_{n=1}^N \mathbf{1}_{S_n^{(1)} = S_n^{(2)}} \right) \right], \end{aligned}$$

by Hölder's inequality. Besides, using Theorem B.2 in Appendices, we can prove that for all $\beta \geq 0$:

$$\sup_N \mathbf{E} \left[\exp \left(\beta N^{-1/2} \sum_{n=1}^N \mathbf{1}_{S_n^{(1)}=S_n^{(2)}} \right) \right] < +\infty.$$

Thus, we imply the desired conclusion. \square

Remark 5.1. *Using the same argument as in the above proof, one can see that:*

$$\sup_N \mathbf{E} \left(e^{\beta T_N} \right) < \infty \quad \forall \beta \geq 0.$$

Hence, in particular, $(e^{\beta T_N}, N \in \mathbb{N})$ is uniformly integrable.

If we do not care about $U^{(n)}$, we can just have $X_{N,n} \leq (c+1)^k N^{-1/2}$. This remark will be useful in our proof for Theorem 5.2.

We now give result on the convergence of the exponential moments of $(\frac{1}{\sqrt{N}}\Pi_N, N \in \mathbb{N})$.

Theorem 5.2. *For any bounded positive continuous function $f \in \mathcal{C}_{b,+}([0, 1] \times \mathbb{R})$, we have:*

$$\mathbf{E} \left[\exp \left(\frac{1}{\sqrt{N}} \Pi_N(f) \right) \right] \xrightarrow{N \rightarrow +\infty} \mathbf{E} \left[(\mathcal{Z}_{\sqrt{2f}})^k \right].$$

Proof of Theorem 5.2. For any bounded nonnegative continuous function $f \in \mathcal{C}_{b,+}([0, 1] \times \mathbb{R})$, let

- A_1, A_2, \dots be a sequence of real functions defined on $\mathbb{N} \times \mathbb{Z}$ such that:

$$A_N(n, z) := \sqrt{f} \left(\frac{n}{N}, \frac{z}{\sqrt{N}} \right) \quad \forall n \in \mathbb{N}, z \in \mathbb{Z}.$$

- $a := \sqrt{f}$ and $c := \|a\|_\infty$.

Notice that due to the continuity of f , $\lim_{N \rightarrow \infty} A([t, x]_N) = a(t, x)$ for all $(t, x) \in [0, 1] \times \mathbb{R}$. Thus, $(A_N, N \in \mathbb{N})$ satisfies the condition of Theorem 2.3 and therefore:

$$\mathfrak{Z}_N(N^{-1/4}A_N) \xrightarrow[N \rightarrow \infty]{(d)} \mathcal{Z}_{\sqrt{2f}}.$$

Hence, from the uniform integrability in Corollary 2.5, we deduce that:

$$\mathbf{E} \left[\left(\mathfrak{Z}_N(N^{-1/4}A_N) \right)^k \right] \xrightarrow{N \rightarrow +\infty} \mathbf{E} \left[(\mathcal{Z}_{\sqrt{2f}})^k \right]. \quad (27)$$

Using again the quantity $X_{N,n}$ defined by (21), we have shown in (24) that:

$$\mathbf{E} \left[\left(3_N(N^{-1/4} A_N) \right)^k \right] = \mathbf{E} \left[\prod_{n=1}^N (1 + X_{N,n}) \right]$$

So the convergence (27) can be rewritten as:

$$\mathbf{E} \left[\prod_{n=1}^N (1 + X_{N,n}) \right] \xrightarrow{N \rightarrow +\infty} \mathbf{E} \left[(\mathcal{Z}_{\sqrt{2f}})^k \right]. \quad (28)$$

From Remark 5.1, we know that the sequence $(T_N, N \in \mathbb{N})$ with $T_N = \sum_{n=1}^N X_{N,n}$ satisfies that $(e^{\beta T_N}, N \in \mathbb{N})$ is uniformly integrable and that:

$$0 \leq X_{N,n} \leq (c+1)^k N^{-1/2}.$$

Hence, using Theorem A.3 in Appendix A and the convergence (28), we deduce that:

$$\mathbf{E} \left[e^{T_N} \right] \xrightarrow{N \rightarrow +\infty} \mathbf{E} \left[(\mathcal{Z}_{\sqrt{2f}})^k \right]. \quad (29)$$

We now investigate the relation between T_N and $\frac{1}{\sqrt{N}} \Pi_N(f)$. Observe that:

$$0 \leq \frac{1}{\sqrt{N}} \Pi_N(f) \leq T_N.$$

Indeed, from the expansion (23), we have:

$$\begin{aligned} T_N - \frac{1}{\sqrt{N}} \Pi_N(f) &= \\ &= \sum_{n=1}^N \sum_{l=3}^k \sum_{1 \leq i_1 < i_2 < \dots < i_l \leq k} N^{-l/4} \prod_{h=1}^l A_N(n, S_n^{(i_h)}) \mathbf{E} \left[\prod_{h=1}^l \omega(n, S_n^{(i_h)}) \right] \geq 0. \end{aligned}$$

Then following the same arguments that have been used to bound T_N in (25) and (26), one can show that:

$$T_N - \frac{1}{\sqrt{N}} \Pi_N(f) \leq (c+1)^k N^{-3/4} \sum_{1 \leq i < j \leq k} \sum_{n=1}^N \mathbb{1}_{S_n^{(i)} = S_n^{(j)}}.$$

where the upper bound converges in distribution to 0 when N converges to infinity.

Thus, by applying Lemma C.1 in Appendices to two sequences $\left(e^{\frac{1}{\sqrt{N}} \Pi_N(f)}, N \in \mathbb{N} \right)$ and $(e^{T_N}, N \in \mathbb{N})$, one can conclude that:

$$\mathbf{E} \left[\exp \left(\frac{1}{\sqrt{N}} \Pi_N(f) \right) \right] \xrightarrow{N \rightarrow +\infty} \mathbf{E} \left[(\mathcal{Z}_{\sqrt{2f}})^k \right] < \infty.$$

□

5.2 Convergence of collision measures

We begin by proving the weak tightness of $(\frac{1}{\sqrt{N}}\Pi_N, N \in \mathbb{N})$, then giving the proof of Theorem 1.3. We refer to [14, p.118,119] for the weak tightness. The weak tightness is crucial as it allows us to take convergent subsequences of $(\frac{1}{\sqrt{N}}\Pi_N; N \in \mathbb{N})$ (cf. Theorem 5.3).

Theorem 5.3. (*Prokhorov's theorem, [6, Theorems 5.1 and 5.2, p.59-60]*)

Let E be a Polish space and Θ be a family of probability measures on E , then Θ is tight if and only if Θ is a relatively compact subset of \mathcal{P} , where \mathcal{P} is the topological space of all probability measures on E , equipped with the weak convergence topology (see [14, Chapter 4] for more details).

Remark 5.4. *In our framework of random measures, E is taken to be the space of all positive finite measures on $[0, 1] \times \mathbb{R}$, equipped with weak convergence topology.*

Theorem 5.5. *The sequence of random measures $(\frac{1}{\sqrt{N}}\Pi_N, N \in \mathbb{N})$ is weakly tight.*

Proof of Theorem 5.5. Let \mathcal{M} denote the set of all finite positive measures on the Polish space $[0, 1] \times \mathbb{R}$, and

$$\begin{aligned} L_m &:= \{\mu \in \mathcal{M} : \|\mu\| \leq m\}, \\ M_m &:= \{\mu \in \mathcal{M} : \text{supp}\mu \subset [0, 1] \times [-m, m]\}, \\ K_m &:= L_m \cap M_m. \end{aligned}$$

So K_m is a collection of some measures that are uniformly bounded and contained within the same compact set. Thus, by Lemma 4.4 in [14], K_m is a weakly relatively compact subset of \mathcal{M} . So, by the definition of tightness, it suffices to prove that

$$\lim_{m \rightarrow +\infty} \sup_N \mathbf{P}(N^{-1/2}\Pi_N \notin K_m) = 0,$$

which is true because

$$\lim_{m \rightarrow +\infty} \sup_N \mathbf{P}(N^{-1/2}\Pi_N \notin M_m) = 0 \text{ and } \lim_{m \rightarrow +\infty} \sup_N \mathbf{P}(N^{-1/2}\Pi_N \notin L_m) = 0.$$

Indeed, for M_m , we observe that:

$$\begin{aligned} &\mathbf{P}(N^{-1/2}\Pi_N \notin M_m) \\ &\leq \mathbf{P}\left(\sup_{\substack{1 \leq n \leq N \\ 1 \leq i \leq k}} |S_n^{(i)}| > m\sqrt{N}\right) \leq k\mathbf{P}\left(\frac{\sup_{1 \leq n \leq N} |S_n|}{\sqrt{N}} > m\right). \end{aligned}$$

Since the sequence $\left(\frac{1}{\sqrt{N}} \sup_{1 \leq n \leq N} |S_n|, N \in \mathbb{N}\right)$ converges in distribution to a real random variable (by Donsker's theorem), this sequence is tight by Prokhorov's theorem [6, Theorem 5.2, p. 60]. Thus,

$$\lim_{m \rightarrow +\infty} \sup_N \mathbf{P}(N^{-1/2} \Pi_N \notin M_m) \leq k \lim_{m \rightarrow +\infty} \sup_N \mathbf{P}\left(\frac{\sup_{1 \leq n \leq N} |S_n|}{\sqrt{N}} > m\right) = 0.$$

For L_m , we have:

$$\begin{aligned} \mathbf{P}(N^{-1/2} \Pi_N \notin L_m) &\leq \mathbf{P}\left(\sum_{n=1}^N \sum_{1 \leq i < j \leq k} \mathbb{1}_{S_n^{(i)} = S_n^{(j)}} > m\sqrt{N}\right) \\ &\leq \frac{k(k-1)}{2} \mathbf{P}\left(\frac{1}{\sqrt{N}} \sum_{n=1}^N \mathbb{1}_{S_n^{(1)} = S_n^{(2)}} > \frac{2m}{k(k-1)}\right) \\ &= \frac{k(k-1)}{2} \mathbf{P}\left(\frac{1}{\sqrt{N}} \sum_{n=1}^{2N} \mathbb{1}_{S_n=0} > \frac{2m}{k(k-1)}\right). \end{aligned}$$

Similarly, because $\left(\frac{1}{\sqrt{N}} \sum_{n=1}^{2N} \mathbb{1}_{S_n=0}, N \in \mathbb{N}\right)$ also converges in distribution [20, Theorem 10.1], we have

$$\lim_{m \rightarrow +\infty} \sup_N \mathbf{P}(N^{-1/2} \Pi_N \notin L_m) = 0.$$

Hence the conclusion. \square

Now, by combining all results we have shown so far, we can give the proof of Theorem 1.3.

Proof of Theorem 1.3. By Theorem 5.5 and Prokhorov's theorem [6, Theorem 5.1], there exists a random finite positive measure \mathcal{N}' on $[0, 1] \times \mathbb{R}$ such that there is a subsequence of $(\frac{1}{\sqrt{N}} \Pi_N, N \in \mathbb{N})$ that converges in distribution to \mathcal{N}' . For convenience, assume that \mathcal{N}' is defined on the existing probability space $(\Omega, \mathcal{A}, \mathbf{P})$.

Besides, for any $f \in C_{b,+}([0, 1] \times \mathbb{R})$, by the proof of Theorem 5.2, it is known that: $(e^{\frac{1}{\sqrt{N}} \Pi_N(f)}, N \in \mathbb{N})$ is uniformly integrable. Thus, $\mathbf{E}[e^{\mathcal{N}'(f)}]$ is finite and equal to $\mathbf{E}\left[\left(\mathcal{Z}_{\sqrt{2}f}\right)^k\right]$.

We see that to show $\frac{1}{\sqrt{N}} \Pi_N \xrightarrow[N \rightarrow \infty]{wd} \mathcal{N}'$, it suffices to prove that \mathcal{N}' is uniquely defined in distribution.

Indeed, let \mathcal{N}'' be another random bounded measure on $[0, 1] \times \mathbb{R}$ such that there is a subsequence of $(\frac{1}{\sqrt{N}} \Pi_N, N \in \mathbb{N})$ that converges in distribution to it. Assume \mathcal{N}'' is also defined on $(\Omega, \mathcal{A}, \mathbf{P})$.

In the following, we will prove that $\mathcal{N}'(h) \stackrel{(d)}{=} \mathcal{N}''(h)$ for all $h \in C_b([0, 1] \times \mathbb{R})$,

then the uniqueness of \mathcal{N}' follows immediately from Lemma 4.7 in [14]. Let f, g be two continuous nonnegative bounded functions on $[0, 1] \times \mathbb{R}$. For any two nonnegative numbers a and b , $af + bg$ is also a continuous bounded nonnegative function. Hence,

$$\mathbf{E} \left[e^{\mathcal{N}'(af+bg)} \right] = \mathbf{P} \left[\left(\mathcal{Z}_{\sqrt{2(af+bg)}} \right)^k \right] = \mathbf{E} \left[e^{\mathcal{N}''(af+bg)} \right].$$

Either, for all $a, b \geq 0$,

$$\mathbf{E} \left[e^{a\mathcal{N}'(f)+b\mathcal{N}'(g)} \right] = \mathbf{E} \left[e^{a\mathcal{N}''(f)+b\mathcal{N}''(g)} \right].$$

Or, for all $a, b \geq 0$,

$$a\mathcal{N}'(f) + b\mathcal{N}'(g) \stackrel{(d)}{=} a\mathcal{N}''(f) + b\mathcal{N}''(g).$$

So by Cramer-Wold theorem [13, Corollary 4.5], we have:

$$(\mathcal{N}'(f), \mathcal{N}'(g)) \stackrel{(d)}{=} (\mathcal{N}''(f), \mathcal{N}''(g)).$$

Then using Cramer-Wold Theorem again, we deduce that $\mathcal{N}'(f - g) \stackrel{(d)}{=} \mathcal{N}''(f - g)$ for all $f, g \in \mathcal{C}_{b,+}([0, 1] \times \mathbb{R})$. So, $\mathcal{N}'(h) \stackrel{(d)}{=} \mathcal{N}''(h)$ for all $h \in \mathcal{C}_b([0, 1] \times \mathbb{R})$ because any bounded continuous function h can be written as the difference of two continuous bounded nonnegative functions.

Thus, we proved that $\frac{1}{\sqrt{N}}\Pi_N \xrightarrow[N \rightarrow \infty]{wd} \mathcal{N}$, where \mathcal{N} is a positive random measure $[0, 1] \times \mathbb{N}$ that is uniquely defined in distribution by the following equation for all $f \in \mathcal{C}_{b,+}([0, 1] \times \mathbb{R})$:

$$\mathbf{E}(e^{\mathcal{N}(f)}) = \mathbf{E} \left[\left(\mathcal{Z}_{\sqrt{2f}} \right)^k \right].$$

Finally, the convergence of $(\frac{1}{\sqrt{N}}\Pi'_N, N \in \mathbb{N})$ follows directly from the convergence of $(\frac{1}{\sqrt{N}}\Pi_N, N \in \mathbb{N})$ and Lemma C.1 by noticing that $\Pi_N(f) \geq \Pi'_N(f) \geq 0$ for all $f \in \mathcal{C}_{b,+}([0, 1] \times \mathbb{R})$, and

$$\begin{aligned} & \mathbf{E} \left(\frac{1}{\sqrt{N}} \|\Pi_N - \Pi'_N\| \right) \\ & \leq \frac{1}{\sqrt{N}} \mathbf{E} \left[\sum_{n=1}^N \sum_{z \in \mathbb{Z}} \binom{k}{2} \sum_{1 \leq i_1 \leq i_2 \leq i_3 \leq k} \mathbf{1}_{\{S_n^{(i_1)} = S_n^{(i_2)} = S_n^{(i_3)} = z\}} \right] \\ & \leq \frac{k^5}{\sqrt{N}} \sum_{n=1}^N \mathbf{P}(S_n^{(1)} = S_n^{(2)} = S_n^{(3)}) \leq \frac{k^5}{\sqrt{N}} \sum_{n=1}^N \max_{z \in \mathbb{Z}} (\mathbf{P}(S_n^{(3)} = z)) \mathbf{P}(S_n^{(1)} = S_n^{(2)}) \\ & = \frac{k^5}{\sqrt{N}} \sum_{n=1}^N \frac{1}{2^n} \binom{n}{\lceil n/2 \rceil} \frac{1}{2^{2n}} \binom{2n}{n} \leq \frac{k^5}{\sqrt{N}} C^2 \sum_{n=1}^N \frac{1}{n} \xrightarrow[N \rightarrow \infty]{} 0, \end{aligned}$$

for some constant C such that $\frac{1}{2^n} \binom{n}{\lceil n/2 \rceil} \leq C \frac{1}{\sqrt{n}}$ for all $n \in \mathbb{N}$. Note that such C exists thanks to Sterling's formula. Hence, our theorem is proved. \square

Acknowledgement

I am indebted to my Master thesis supervisor Quentin Berger for his invaluable help during my master internship and for introducing me to the techniques of U -statistics for solving problems in Statistical Mechanics. My sincere gratitude is reserved for Nicolas Fournier for many crucial discussions. Also, I'm fortunate to have Viet-Chi Tran and H  l  ne Gu  rin as my current supervisors, without their reviews and their push, I could not have finished this paper. Finally, this project is supported by Mathematics for Public Health (MfPH) program at the Fields Institute for Research in Mathematical Sciences, Canada, and partly funded by the B  zout Labex, funded by ANR, reference ANR-10-LABX-58.

Appendix A On the asymptotic relation between products and sums of independent random variables

We consider a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. For any N , let $X_N = (X_{N,n}, n \in \mathbb{N})$ be a sequence of nonnegative random variables such that the sum $S_N = \sum_{n \geq 1} X_{N,n}$ is almost surely finite.

Suppose that there exists a sequence of numbers $(c_N, N \in \mathbb{N})$ converging to 0 such that for all N ,

$$c_N \geq \|X_N\|_\infty = \sup_n |X_{N,n}|.$$

Let

$$P_N := \prod_{n \geq 1} (1 + X_{N,n}).$$

In this Appendix, we establish two relations between the sum S_N and the product P_N when N converges to infinity. Note that we do not assume $(X_{N,n}; n \in \mathbb{N}, N \in \mathbb{N})$ to be independent nor identically distributed.

Theorem A.1. (First relation) For any real random variable Y , the following two assertions are equivalent:

$$1) \quad S_N \xrightarrow[N \rightarrow +\infty]{(d)} Y \qquad 2) \quad P_N \xrightarrow[N \rightarrow +\infty]{(d)} e^Y.$$

Remark A.2. *There is no moment assumption on Y .*

Theorem A.3. (Second relation) Assume that the sequence $(\exp(S_N), N \in \mathbb{N})$ is uniformly integrable. Then for any real constant C , the following two assertions are equivalent:

$$1) \quad \mathbf{E}[e^{S_N}] \xrightarrow[N \rightarrow +\infty]{} C \qquad 2) \quad \mathbf{E}[P_N] \xrightarrow[N \rightarrow +\infty]{} C.$$

Proof of Theorem A.1. Let us first prove that 1) \Rightarrow 2). The inequality $x - \frac{x^2}{2} \leq \ln(1+x) \leq x$ and the assumption imply that:

$$0 \leq S_N - \ln(P_N) \leq \frac{1}{2} \sum_{n \geq 1} X_{N,n}^2 \leq c_N S_N \xrightarrow[N \rightarrow +\infty]{(d)} 0.$$

Hence, by Slutsky's lemma [24, Lemma 2.8], $\ln(P_N) \xrightarrow[N \rightarrow +\infty]{(d)} Y$.

Let us now prove that 2) \Rightarrow 1), we see that for all $x > 0$, $0 \leq x - \ln(1+x) \leq x \ln(1+x)$. We deduce

$$0 \leq S_N - \ln(P_N) \leq \sum_{n \geq 1} X_{N,n} \ln(1 + X_{N,n}) \leq c_N \ln(P_N) \xrightarrow[N \rightarrow +\infty]{(d)} 0.$$

Thus, $S_N \xrightarrow[N \rightarrow +\infty]{(d)} Y$. The equivalence is proved. \square

Proof of Theorem A.3. For the 1) \Rightarrow 2) direction:

The sequence $(\exp(S_n), n \in \mathbb{N})$ being uniformly integrable, thus there is a subsequence $(n_k, k \in \mathbb{N})$ of \mathbb{N} and a random variable $Z \in L^1$ such that:

$$\exp S_{n_k} \xrightarrow[n \rightarrow \infty]{(d)} Z \quad \text{and} \quad \mathbf{E}[\exp(S_{n_k})] \xrightarrow[k \rightarrow \infty]{} \mathbf{E}[Z].$$

We deduce that $\mathbf{E}[Z] = C$ and by Theorem A.1, we have $P_{n_k} \xrightarrow[n \rightarrow \infty]{(d)} Z$.

Besides, the uniform integrability of $(\exp(S_n), n \in \mathbb{N})$ implies the uniform integrability of $(P_N, N \in \mathbb{N})$ ($0 \leq P_N \leq e^{S_N}$). So,

$$\mathbf{E}[P_{n_k}] \xrightarrow[k \rightarrow \infty]{} \mathbf{E}[Z] = C.$$

Notice that the uniform integrability and the convergence $\mathbf{E}(S_N) \xrightarrow[N \rightarrow \infty]{} C$ are still valid if we take any subsequence $(S_{m_i}, i \in \mathbb{N})$ of $(S_N, N \in \mathbb{N})$.

Thus, the result so far implies that for every subsequence $(m_i, i \in \mathbb{N})$ of \mathbb{N} , there is a subsequence $(m_{i_k}, k \in \mathbb{N})$ of (m_i) such that:

$$\mathbf{E}[P_{m_{i_k}}] \xrightarrow[k \rightarrow \infty]{} C.$$

The first implication is proved. The reciprocal is similar. \square

Remark A.4. Note that uniform integrability implies tightness.

Appendix B Some auxiliary results on random walks

Let $(S_n, n \in \mathbb{N}_0)$ be a simple symmetric random walks on \mathbb{Z} and :

- i. $(X_k, k \in \mathbb{N}_0)$ be a sequence of random variables such that $X_0 = 0$ and $X_k := \inf\{N > X_{k-1} : S_N = 0\}$ for all positive integer k ,

- ii. $\mathcal{F} := (\mathcal{F}_k, k \in \mathbb{N}_0)$ be the canonical filtration of the process $(X_k, k \in \mathbb{N}_0)$,
- iii. $T_k := X_k - X_{k-1}$ for all positive integer k ,
- iv. $\tau_N := \inf\{k \geq 0 : X_k \geq N\}$.

Clearly, by definition, for each N , τ_N is a stopping time with respect to the filtration \mathcal{F} and by Markov's property of S , $(T_k, k \in \mathbb{N})$ is a sequence of independent identically distributed random variables.

Notice that T_1 is the first time after 0 at which the random walk S returns to the position 0. Clearly, this stopping time is well-known. One of its properties is that

Lemma B.1. *There is a positive constant C such that for all $k \in \mathbb{N}$,*

$$\mathbf{P}(T_1 = 2k) = 2^{-2k+1} \frac{1}{k} \binom{2k-2}{k-1} \geq \frac{C}{k^{3/2}}.$$

Indeed, this lemma is just a combination of Theorem 9.2 in [20] and Sterling's formula.

Concerning τ_N , by its definition, we have the following equality which will be useful for our later analysis:

$$\tau_N - 1 = \sup\{k \geq 0 : X_k \leq N - 1\} = \sum_{n=1}^{N-1} \mathbb{1}_{S_n=0}.$$

In the following is the main theorem of this Section.

Theorem B.2. *(Boundedness of exponential moments of local times)*

Let S be a random simple walk on \mathbb{Z} starting from 0, then for any constant $\beta \geq 0$, we have:

$$\sup_N \mathbf{E} \left[\exp \left(\beta N^{-1/2} \sum_{n=1}^N \mathbb{1}_{S_n=0} \right) \right] < +\infty.$$

This is a corollary of Lemma 4.2 in [22]. Here, we give an alternative proof.

Proof. The main idea to prove this theorem is to construct many appropriate martingales to estimate the underlying exponential moment. The construction is as follows, for each $N \in \mathbb{N}$, define:

- i. $X_n^N := \sum_{i=1}^n \min(T_i, N)$.
- ii. $\gamma_N := \inf\{n \geq 1 : X_n^N \geq N\}$.
- iii. $\lambda_N(\beta) := -\log \mathbf{E}(e^{-\beta \min(T_1, N)}) > 0 \quad \forall N \in \mathbb{N}, \beta > 0$.

iii. $M_n^N := \exp(-\beta X_n^N + n\lambda_N(\beta))$.

Then by noticing that the random variables T_1, T_2, \dots are i.i.d, we see that for each N , $(M_n^N, n \in \mathbb{N})$ is a martingale with respect to the filtration \mathcal{F} . In addition, because $\forall n, N : X_n^N \geq n$, $\forall N : \tau_N \leq N$. Hence by the optional sampling theorem, $\forall N \in \mathbb{N}, \beta > 0$,

$$\mathbf{E} [\exp(-\beta X_{\gamma_N}^N + \gamma_N \lambda_N(\beta))] = 1.$$

Besides, by definition of γ_N and X^N , we have:

$$X_{\gamma_N}^N = X_{\gamma_N-1}^N + \min(T_{\gamma_N}, N) \leq N + N = 2N.$$

Thus, $e^{2\beta} \geq \mathbf{E}(e^{\gamma_N \lambda_N(\beta/N)})$.

Hence, Lemma B.3 implies that for all $\beta > 0$,

$$\sup_N \mathbf{E} \left[\exp\left(\frac{1}{2}c(\beta)\gamma_N/\sqrt{N}\right) \right] < \infty,$$

where $c(\beta) := C \int_0^{1/2} \frac{1}{t^{3/2}}(1 - e^{-2t\beta})dt$ and C is the constant defined in Lemma B.1.

By noticing that $\lim_{\beta \rightarrow \infty} c(\beta) = \infty$ and $\forall N : \tau_N = \gamma_N$, we conclude that for all $\beta > 0$:

$$\sup_N \mathbf{E} \left[\exp(\beta \tau_N / \sqrt{N}) \right] < \infty,$$

which is essentially our desired conclusion because $\tau_N - 1 = \sum_{n=1}^{N-1} \mathbb{1}_{S_n=0}$. \square

Lemma B.3. *The sequence of functions $(\lambda_N, N \in \mathbb{N})$ given in the proof of Theorem B.2 satisfies the following inequality:*

$$\liminf_{N \rightarrow \infty} \sqrt{N} \lambda_N(\beta/N) \geq c(\beta),$$

with $c(\beta) := C \int_0^{1/2} \frac{1}{t^{3/2}}(1 - e^{-2t\beta})dt$, where C is the constant defined in the Lemma B.1.

Proof. For any $\beta > 0$ and $N \geq 2$, we have:

$$\begin{aligned} & 1 - \mathbf{E} \left[e^{-\beta \min(T_1, N)/N} \right] \\ &= \sum_{k=1}^{\lfloor N/2 \rfloor} \mathbf{P}(T_1 = 2k)(1 - e^{-2k\beta/N}) + \mathbf{P}(T_1 \geq 2\lfloor N/2 \rfloor + 2)(1 - e^{-\beta}) \\ &\geq \sum_{k=1}^{\lfloor N/2 \rfloor} \frac{C}{k^{3/2}}(1 - e^{-2k\beta/N}) \end{aligned}$$

Thus,

$$\liminf_{N \rightarrow \infty} \sqrt{N} \left(1 - \mathbf{E} \left[e^{-\beta \min(T_1, N)/N} \right] \right) \geq C \int_0^{1/2} \frac{1}{t^{3/2}}(1 - e^{-2t\beta})dt = c(\beta) > 0.$$

From which, we conclude $\liminf_{N \rightarrow \infty} \sqrt{N} \lambda_N(\beta/N) \geq c(\beta)$. \square

Appendix C A useful lemma

Lemma C.1. *Let $(U_n), (V_n)$ be two sequences of positive random variables such that $0 \leq U_n \leq V_n$ for all n , and V_1, V_2, \dots are uniformly integrable. Then if $\frac{V_n}{U_n} \xrightarrow[n \rightarrow +\infty]{(d)} 1$ and $\lim_{n \rightarrow \infty} \mathbf{E}(V_n) = C$, then $\lim_{n \rightarrow \infty} \mathbf{E}(U_n) = C$.*

Proof. The uniform integrability of (V_n) implies the uniform integrability of (U_n) . The uniform integrability of (U_n) implies that for every subsequence (n_k) of \mathbb{N} , there exists a subsequence (n_{k_l}) of (n_k) such that $(U_{n_{k_l}}, l \in \mathbb{N})$ converges in distribution to a random variable Z . The convergence of $(\frac{V_n}{U_n}, n \in \mathbb{N})$ implies that $(V_{n_{k_l}}, l \in \mathbb{N})$ also converges in distribution to Z . Then, the uniform integrability implies that $\lim_l \mathbf{E}(U_{n_{k_l}}) = C = \lim_l \mathbf{E}(V_{n_{k_l}})$. Hence the conclusion. \square

References

- [1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables. 10th printing, with corrections.* National Bureau of Standards. A Wiley-Interscience Publication. New York etc.: John Wiley & Sons. xiv, 1046 pp. (1972)., 1972.
- [2] T. Alberts, K. Khanin, and J. Quastel. The intermediate disorder regime for directed polymers in dimension $1 + 1$. *Ann. Probab.*, 42(3):1212–1256, 2014.
- [3] L. Avena, O. Blondel, and A. Faggionato. Analysis of random walks in dynamic random environments via L^2 -perturbations. *Stochastic Process. Appl.*, 128(10):3490–3530, 2018.
- [4] M. T. Barlow, Y. Peres, and P. Sousi. Collisions of random walks. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(4):922–946, 2012.
- [5] Q. Berger and H. Lacoin. The scaling limit of the directed polymer with power-law tail disorder. *Comm. Math. Phys.*, 386(2):1051–1105, 2021.
- [6] P. Billingsley. *Convergence of probability measures.* Chichester: Wiley, 1999.
- [7] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* New York, NY: Springer, 2011.
- [8] P. Carmona and H. Yueyun. On the partition function of a directed polymer in a Gaussian random environment. *Probab. Theory Related Fields*, 124(3):431–457, 2002.

- [9] X. Chen. Gaussian bounds and collisions of variable speed random walks on lattices with power law conductances. *Stochastic Process. Appl.*, 126(10):3041–3064, 2016.
- [10] R. Durrett. *Probability. Theory and examples.*, volume 31. Cambridge: Cambridge University Press, 2010.
- [11] N. Halberstam and T. Hutchcroft. Collisions of random walks in dynamic random environments. *Electron. J. Probab.*, 27:–, 2022.
- [12] T. Hutchcroft and Y. Peres. Collisions of random walks in reversible random graphs. *Electron. Commun. Probab.*, 20:no. 63, 6, 2015.
- [13] O. Kallenberg. *Foundations of modern probability.* New York, NY: Springer, 1997.
- [14] O. Kallenberg. *Random measures, theory and applications*, volume 77. Cham: Springer, 2017.
- [15] F. B. Knight. Random walks and a sojourn density process of Brownian motion. *Trans. Am. Math. Soc.*, 109:56–86, 1963.
- [16] V. S. Korolyuk and Yu. V. Borovskikh. *Theory of U-statistics. Updated and transl. from the Russian by P. V. Malyshev and D. V. Malyshev*, volume 273. Dordrecht: Kluwer Academic Publishers, 1994.
- [17] D. Nualart. *The Malliavin calculus and related topics.* Berlin: Springer, 2006.
- [18] G. Pólya. *Collected papers. Vol. IV*, volume 22 of *Mathematicians of Our Time.* MIT Press, Cambridge, MA, 1984. Probability; combinatorics; teaching and learning in mathematics, Edited by Gian-Carlo Rota, M. C. Reynolds and R. M. Shortt.
- [19] P. Révész. Local time and invariance. Analytical methods in probability theory, Proc. Conf., Oberwolfach 1980, Lect. Notes Math. 861, 128-145 (1981)., 1981.
- [20] P. Révész. *Random walk in random and non-random environments.* Hackensack, NJ: World Scientific, 2005.
- [21] H. Royden and P. M. Fitzpatrick. *Real analysis.* New York, NY: Prentice Hall, 2010.
- [22] J. Sohler. Finite size scaling for homogeneous pinning models. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 6:163–177, 2009.
- [23] T. Szabados and B. Székely. An elementary approach to Brownian local time based on simple, symmetric random walks. *Period. Math. Hung.*, 51(1):79–98, 2005.

- [24] A. W. van der Vaart. *Asymptotic statistics*, volume 3. Cambridge: Cambridge Univ. Press, 1998.
- [25] J. B. Walsh. An introduction to stochastic partial differential equations. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 265–439. Springer, Berlin, 1986.