Università degli studi di Padova

Université Paris Dauphine - PSL

Double Degree Program MAPPA

# A Central Limit Theorem for two-dimentional directed polymers in a random environment


Supervisor:
**Prof. Clément Cosco**
Cosupervisor:
**Prof. Alberto Chiarini**

Graduate:
**Anna Donadini**
Student number UniPD:
**2056068**
Student number Dauphine:
**22200157**


13 September 2023
Academic Year 2022-2023

# Contents

# Introduction

## General presentation of the model

The problem of directed polymers in a random environment is a well known subject in the statistical physics and probability communities. It was first studied in [10] and received its first mathematical treatment in [11]. In the setting of the $d$-dimensional integer lattice, the polymer measure is a random probability measure on paths of $d$-dimensional nearest-neighbour lattice walks. The randomness of the polymer measure comes from an i.i.d. collection of random variables $\omega(n,x)$, placed on the sites of $\mathbb{N}\times\mathbb{Z}^d$, which are called the random environment. More specifically, the random environment is a product probability measure on the space of all $\omega : \mathbb{N}\times\mathbb{Z}^d \to \mathbb{R}$, so that the variables $\omega(n,x)$ are independent and have moments of all orders. The directed polymer model in random environment can then be defined as a random walk in a random potential. Indeed, considering the simple random walk $(S_n)_{n\in\mathbb{N}}$ starting from $x \in \mathbb{Z}^d$ and given a fixed environment $\omega(n,x)$, $n \in \mathbb{N}$, $x \in \mathbb{Z}^d$, the energy of an $n$-step path is defined as

$$H_n^\omega(S) = \sum_{i=1}^{n} \omega(n, S_n), \tag{H}$$

and the polymer measure is defined in the Gibbsian way as

$$P_{n,\beta}^\omega(S) = \frac{e^{\beta H_n^\omega(S)} P(S)}{Z_n^\omega(\beta)}, \tag{$\mathbb{P}$}$$

where $\beta > 0$ is the inverse temperature, $P$ is the probability measure associated to the simple symmetric random walk started at the origin and $Z_n^\omega(\beta) = E[e^{\beta H_n^\omega(S)}]$ is the partition function, i.e. the normalizing constant to make $P_{n,\beta}^\omega$ a probability measure. This way, the polymer is attracted to sites where the random environment is positive and repelled by sites where the environment is negative. The overall goal of the subject is to study the behaviour of the polymer when the inverse temperature $\beta$ and the dimension $d$ vary and $n$ gets large.

To picture intuitively how this model works, let's consider the example of a hydrophilic polymer chain (i.e. a long chain of monomers) floating in water. Due to the thermal fluctuation, the shape of the polymer should be understood as a random object. We now suppose that the water contains randomly placed hydrophobic molecules as impurities, which repel the hydrophilic monomers of the polymer. The question is: how do the impurities affect the global shape of the polymer chain? Let's consider the model defined in (H) and ($\mathbb{P}$), where we let the random environment $\omega(n, x)$ describe the presence of an impurity at site $(n, x)$. Informally, we can guess two scenarios: if the space dimension $d$ is large and the temperature $(1/\beta)$ is high, the impurities should not affect the global shape of the polymer; on the other hand, if $d$ is small or the environment is strong, the polymer will not be able to avoid the impurities if it stays at reasonable distances. Indeed, in the first case the polymer has much space to move and the impurities are weak, so it has no interest in travelling far and only wants to pick a specific portion of rewards, using a local strategy. In the second case, the polymer will have an advantage in travelling far in order to find more favourable areas. These atypical areas in the medium correspond to a higher density of rewards, and a precise geometry making them feasible for the walk ([13]). The first scenario falls in the so called *weak disorder regime*, the second in the *strong disorder regime*.

Let's be more precise on the definition of strong and weak disorder. Consider the partition function of the directed polymer model

$$W_N(\beta, x) = E_x\left[e^{\sum_{n=1}^{N}(\beta\omega(n,S_n)-\lambda(\beta))}\right], \tag{W}$$

where $(S_n)_{n\in\mathbb{N}}$ is the simple random walk on $\mathbb{Z}^d$ starting at $x \in \mathbb{Z}^d$, $\beta > 0$ is the inverse temperature and $\lambda(\beta)$ is a normalizing constant, so that the partition function has mean one. Then, setting $\mathcal{G}_N := \sigma\left\{\omega(i, x) : i \le N, x \in \mathbb{Z}^d\right\}$, $W_N(\beta, x)$ is a $\mathcal{G}_N$-measurable positive martingale and therefore it converges almost surely to a limit $W_\infty(\beta, x)$. We say that if $W_\infty(\beta, 0) > 0$ then weak disorder holds, whereas if $W_\infty(\beta, 0) = 0$ strong disorder holds. Weak disorder leads to a diffusive behavior of the polymer, while strong disorder forces the polymer to localize. We refer to [13] for a quantitative description of these phenomena and a more complete history of the model.

Moreover, it is known that a phase transition occurs between the diffusive behaviour in the weak disorder and the localized behaviour in the strong disorder. In fact, Comets and Yoshida showed in [12] that in every dimension $d$ there exists a critical value $\beta_c$ that determines a phase transition: for $0 \le \beta < \beta_c$ the polymer is in a weak disorder regime (diffusive-type behaviour), while for $\beta > \beta_c$ we find a strong disorder regime (localization). Note that they also proved that for $d = 1, 2$ the critical value is $\beta_c = 0$.

For $d \ge 3$ the critical value $\beta_c$ is only known to exist and to be positive, but its precise value remains largely unknown. In particular, almost nothing is known about the behavior of the polymer at the critical point. An important step forward in this direction was made by Caravenna, Sun and Zygouras [1] in the marginal dimension $d = 2$, where the polymer is always localized. They showed that one can actually zoom close to the critical value $\beta_c = 0$ and still find a phase transition from a diffusive behaviour to a localized one. More precisely, defining the *replica overlap*

$$R_N := E_0^{\otimes 2}\left[\sum_{n=1}^{N} \mathbb{1}_{S_n^1 = S_n^2}\right],$$

which takes into account the number of intersections between two copies of the random walk, and considering for $\hat{\beta} > 0$ the parameter

$$\beta_N := \frac{\hat{\beta}}{\sqrt{R_N}} \xrightarrow[N\to\infty]{} 0,$$

we find a critical value $\hat{\beta}_c = 1$, for which we can state a *Central Limit Theorem* as follows. Let $x \in \mathbb{Z}^2$ and let

$$W_N(\hat{\beta}, x) := \mathbb{E}_x\left[e^{\sum_{n=1}^N (\beta_N \omega(n,S_n) - \lambda(\beta_N))}\right]$$

be the partition function of the rescaled two-dimensional directed polymer. Then, setting $\lambda^2 := \log\left(\frac{1}{1-\hat{\beta}}\right)$, we have that

$$W_N \xrightarrow[N\to\infty]{\mathcal{L}} e^{\mathcal{N}(-\frac{\lambda^2}{2},\lambda^2)} \qquad \text{for } \hat{\beta} < 1 \tag{i}$$

$$W_N \xrightarrow[N\to\infty]{\mathbb{P}} 0 \qquad \text{for } \hat{\beta} \geq 1. \tag{ii}$$

This result has been first proved in [1] and another proof was later proposed in [5]. Both proofs rely on rather sophisticated chaos expansion techniques, and what we aim in this project is to give an alternative proof to (i) using direct and accessible tools. The main idea of our proof is to decompose $\log W_N$ as a sum of independent random variables in order to justify the convergence of $\log W_N$ to a Gaussian distribution via classical central limit theorem arguments. The heuristics of the proof are presented in Section 2.1. Indeed, our approach is motivated both by the desire to give a simpler proof of the Central Limit theorem and by the hope of establishing a connection between the partition function of the 2D polymer model and branching processes, cf. Remark 2.1.2.

## Connections to other models

The study and development of the directed polymer model is relevant in many fields. One of the main applications is that the partition function provides a discretization of the solution of the multiplicative Stochastic Heat Equation (SHE) and that the logarithm of the partition function corresponds to the solution of the Kardar-Parisi-Zhang (KPZ) equation. See [1], [5] and [14] for examples and insights of such connection.

Another connection of the directed polymer model is with branching random walk in random environment. Consider the directed polymer model. Let the environment take only non negative values. One can interpret the model in terms of a branching population (without deaths) moving in a random potential: the polymer measure with time horizon $n$ is equal to the law of the ancestral path of an individual selected uniformly at random in the full population at time $n$. Moreover, the partition function of the polymer model is equal to the expected population size at time $n$ in a fixed environment. In the case of general environment, one can define a branching random walk in random environment such that, conditionally on survival at time n, the law of the ancestral path of an individual selected uniformly at random in the full population at that time is the polymer measure. We refer to [13] for a broader and more detailed discussion of the topic.

# Chapter 1

# Preliminaries

## 1.1 Notation and definition of the model

In this section, we present all the objects defining our model and set all the necessary notation.
Let $(S_n)_{n\in\mathbb{N}}$ denote the simple random walk on $\mathbb{Z}^2$ and $(S_n^1)_{n\in\mathbb{N}}, (S_n^2)_{n\in\mathbb{N}}$ be two independent copies of $(S_n)_{n\in\mathbb{N}}$ starting from the same point. Let $\omega(n,x)$, for $n \in \mathbb{N}$ and $x \in \mathbb{Z}^2$, be i.i.d. random variables constituting the environment. In the following, we will denote with $P_x(\cdot)$ the probability with respect to the random walk started at $x \in \mathbb{Z}^2$ and with $\mathbb{P}(\cdot)$ the probability with respect to the environment. Analogously, $E_x[\cdot]$ will be the expectation over the simple random walk started at $x \in \mathbb{Z}^2$ and $\mathbb{E}[\cdot]$ the expectation over the environment. We use the notation $E$ without any subscript when considering $E_0$.

Suppose that $\mathbb{E}[\omega(n,x)] = 0$, $\mathbb{E}[\omega(n,x)^2] = 1$ and $\mathbb{E}[e^{\beta\omega(n,x)}] < \infty$ for all $\beta > 0$. We define, for $(n,x) \in \mathbb{N}\times\mathbb{Z}^2$, the parameter $\lambda(\beta_N)$ as

$$\lambda(\beta_N) := \log\left(\mathbb{E}[e^{\beta_N\omega(n,x)}]\right), \tag{1.1}$$

where

$$\beta_N = \frac{\hat{\beta}}{\sqrt{R_N}}, \qquad R_N = E_0^{\otimes 2}\left[\sum_{n=1}^{N} \mathbb{1}_{S_n^1 = S_n^2}\right]. \tag{1.2}$$

Note that $\lambda(\beta_N)$ does not depend on $(n,x)$ as the variables $\omega$ that account for the random environment are independent and identically distributed. Note also that, by the Local Central Limit Theorem ([6, Sec. 1.2]), we have that as $N \to \infty$

$$R_N \sim \frac{\log N}{\pi}, \tag{1.3}$$

thus $\beta_N$ goes to 0 for $N \to \infty$ like $\hat{\beta}\sqrt{\frac{\pi}{\log N}}$.

In this setting, our main object of study will be the partition function of the $2d$ random polymer starting from a point $x \in \mathbb{Z}^2$, which we define as

$$W_N(\hat{\beta}, x) := E_x\left[e^{\sum_{n=1}^{N}(\beta_N\omega(n,S_n) - \lambda(\beta_N))}\right]. \tag{1.4}$$

Note that the expectation is taken only with respect to the simple random walk, thus $W_N$ is a random variable with respect to the environment $\omega$; notice that it also has mean one. To analyze the behaviour of this object, we define some approximations as follows.

Let $a, b \in [0,1]$ with $a \leq b$ and $M > 0$ be a positive integer. Then, consider

$$Z_{a,b,N}(\hat{\beta}, x) = E_x \left[ e^{\sum_{n=\lceil N^a \rceil+1}^{\lceil N^b \rceil} (\beta_N \omega(n,S_n) - \lambda(\beta_N))} \right], \tag{1.5}$$

and more specifically

$$Z_{k,M,N}(\hat{\beta}, x) := Z_{\frac{k}{M}, \frac{k+1}{M}, N} = E_x \left[ e^{\sum_{n=t_k+1}^{t_{k+1}} (\beta_N \omega(n,S_n) - \lambda(\beta_N))} \right], \tag{1.6}$$

where we set $t_0 = 1$ and $t_k = \lceil N^{\frac{k}{M}} \rceil$, so that the variables are defined on disjoint time intervals and are thus independent. Notice that all the objects defined are random variables of expectation equal to 1. For simplicity, in the following pages we will drop the indexes $N$ and $M$. We also use the short notations $W_N = W_N(\hat{\beta}, 0)$ and $Z_{a,b} = Z_{a,b,N}(\hat{\beta}, 0)$.

**Remark.** Since the starting point $x$ does not play any particular role by space-shift invariance, we can consider the walk starting from 0, without loss of generality.

## 1.2 A theorem of convergence

The purpose of this section is to present some preliminary results, which will later turn out to be key steps in the proof of (i). Recall the definition of $Z_{a,b}$ given in (1.5). The first lemma we prove exhibits an alternative expression of the second moment of $Z_{a,b}$ in terms of the number of times the simple random walk crosses zero at even steps.

**Lemma 1.2.1** (Second moment of $Z_{a,b}$). *Let $\lambda_2(\beta_N) := \lambda(2\beta_N) - 2\lambda(\beta_N)$. Then for any $a, b \in [0,1]$ with $a \leq b$, we have*

$$\mathbb{E}[Z_{a,b}^2] = E\left[ e^{\lambda_2(\beta_N) \sum_{n=\lceil N^a \rceil+1}^{\lceil N^b \rceil} \mathbb{1}_{S_{2n}=0}} \right] \tag{1.7}$$

*Proof.* We have

$$\begin{aligned}
\mathbb{E}[Z_{a,b}^2] &= \mathbb{E}\left[ \left( E\left[ e^{\sum_{n=\lceil N^a \rceil+1}^{\lceil N^b \rceil} (\beta_N \omega(n,S_n) - \lambda(\beta_N))} \right] \right)^2 \right] \\
&= \mathbb{E}\left[ E^{\otimes 2} \left[ e^{\sum_{n=\lceil N^a \rceil+1}^{\lceil N^b \rceil} (\beta_N \omega(n,S_n^1) - \lambda(\beta_N))} e^{\sum_{n=\lceil N^a \rceil+1}^{\lceil N^b \rceil} (\beta_N \omega(n,S_n^2) - \lambda(\beta_N))} \right] \right] \\
&= E^{\otimes 2} \left[ \mathbb{E}\left[ e^{\sum_{n=\lceil N^a \rceil+1}^{\lceil N^b \rceil} (\beta_N \omega(n,S_n^1) + \beta_N \omega(n,S_n^2) - 2\lambda(\beta_N))} \right] \right] \\
&= E^{\otimes 2} \left[ \prod_{n=\lceil N^a \rceil+1}^{\lceil N^b \rceil} \left( \mathbb{E}\left[ e^{\beta_N \omega(n,S_n^1) + \beta_N \omega(n,S_n^2)} \right] e^{-2\lambda(\beta_N)} \right) \right],
\end{aligned}$$

where the last equality holds by independence of the disorder $\omega$. At this point we notice that

$$\mathbb{E}\left[e^{\beta_N\omega(n,S_n^1)+\beta_N\omega(n,S_n^2)}\right] = \begin{cases} e^{\lambda(2\beta_N)} & \text{if} \quad S_n^1 = S_n^2 \\ e^{2\lambda(\beta_N)} & \text{if} \quad S_n^1 \neq S_n^2, \end{cases}$$

thus deducing that

$$\mathbb{E}\left[e^{\beta_N\omega(n,S_n^1)+\beta_N\omega(n,S_n^2)}\right] e^{-2\lambda(\beta_N)} = \begin{cases} e^{\lambda(2\beta_N)-2\lambda(\beta_N)} & \text{if} \quad S_n^1 = S_n^2 \\ 1 & \text{if} \quad S_n^1 \neq S_n^2. \end{cases}$$

Now, as they are starting from the same point, $(S_n^1)_{n\in\mathbb{N}}$ and $(S_n^2)_{n\in\mathbb{N}}$ have the same parity, therefore we can conclude:

$$\mathbb{E}[Z_{a,b}^2] = E^{\otimes 2}\left[\prod_{n=\lceil N^a\rceil+1}^{\lceil N^b\rceil} e^{\lambda_2(\beta_N)\mathbb{1}_{S_n^1=S_n^2}}\right]$$
$$= E^{\otimes 2}\left[e^{\lambda_2(\beta_N)\sum_{n=\lceil N^a\rceil+1}^{\lceil N^b\rceil}\mathbb{1}_{S_n^1=S_n^2}}\right] = E\left[e^{\lambda_2(\beta_N)\sum_{n=\lceil N^a\rceil+1}^{\lceil N^b\rceil}\mathbb{1}_{S_{2n}=0}}\right].$$

□

**Remark 1.2.2.** *If the $\omega$'s are Gaussian, then $\lambda(\beta_N) = \frac{\beta_N^2}{2}$ and $\lambda_2(\beta_N) = \beta_N^2$.*

The second result we want to present in this section is a theorem of convergence for the second moment of the variables $Z_{a,b}$. To do so, we start by recalling a chaos expansion formula.

**Remark 1.2.3** (Expansion of a product)**.**

$$\prod_{k=1}^{T}(1+x_k) = 1 + \sum_{k=1}^{T}\sum_{1\le n_1<\cdots<n_k\le T}\prod_{i=1}^{k} x_{n_i} \tag{1.8}$$

We are now ready to prove convergence of the second moment of $Z_{a,b}$ and consequently deduce convergence of $\mathbb{E}[W_N^2]$ itself.

**Theorem 1.2.4** (Convergence of the second moment)**.** *For any $a, b \in [0,1]$ with $a \le b$, we have*

$$\mathbb{E}\left[Z_{a,b}^2\right] \xrightarrow[N\to\infty]{} e^{\lambda_{a,b}^2}, \tag{1.9}$$

*where $\lambda_{a,b}^2 := \log\left(\frac{1-a\hat{\beta}^2}{1-b\hat{\beta}^2}\right)$.*

**Remark 1.2.5.** *The previous theorem implies that*

$$\mathbb{E}[W_N^2] \xrightarrow[N\to\infty]{} e^{\lambda^2}, \tag{1.10}$$

*where we set* $\lambda^2 := \lambda_{0,1}^2 = \log\left(\frac{1}{1-\hat\beta^2}\right)$. *In fact,* (1.10) *follows from*

$$\begin{aligned}\mathbb{E}\left[Z_{0,1}^2\right] \le \mathbb{E}[W_N^2] &= \mathbb{E}\left[E^{\otimes 2}\left[e^{\sum_{n=1}^N(\beta_N\omega(n,S_n^1)+\beta_N\omega(n,S_n^2)-2\lambda(\beta_N))}\right]\right]\\ &= E^{\otimes 2}\left[\mathbb{E}\left[e^{\beta_N\omega(1,S_1^1)+\beta_N\omega(1,S_1^2)-2\lambda(\beta_N)}\right]\mathbb{E}\left[e^{\sum_{n=2}^N(\beta_N\omega(n,S_n^1)+\beta_N\omega(n,S_n^2)-2\lambda(\beta_N))}\right]\right]\\ &= E^{\otimes 2}\left[e^{\lambda_2(\beta_N)\mathbb{1}_{S_1^1=S_1^2}}\mathbb{E}\left[e^{\sum_{n=2}^N(\beta_N\omega(n,S_n^1)+\beta_N\omega(n,S_n^2)-2\lambda(\beta_N))}\right]\right] \le e^{\lambda_2(\beta_N)}\mathbb{E}[Z_{0,1}^2],\end{aligned}$$

*and* $e^{\lambda_2(\beta_N)} \xrightarrow[N\to\infty]{} 1$ *as* $\lambda_2(\beta_N) \sim \beta_N^2$.

*Proof.* To prove (1.9) we will provide an upper and a lower bound on the second moment of $Z_{a,b}$ (we refer to [1, Sec. 3] for similar computations).

Consider $\Lambda_N := e^{\lambda_2(\beta_N)} - 1$ and call $p_n(x) = \mathbb{P}_0(S_n = x)$ the probability of being in $x$ at time $n$ when starting from 0, $p_{n-m}(x,y)$ the probability of going from $x$ at time $m$ to $y$ at time $n$. Before starting with the computations, notice that for any $k \ge 1$:

$$e^{\lambda_2(\beta_N)\mathbb{1}_{S_{2n}=0}} - 1 = (e^{\lambda_2(\beta_N)} - 1)\mathbb{1}_{S_{2n}=0} = \Lambda_N\mathbb{1}_{S_{2n}=0} \tag{1}$$

$$\begin{aligned}\Big\{(n_1,\dots,n_k) \text{ s.t. } \lceil N^a\rceil + 1 \le n_1 < \dots < n_k \le \lceil N^b\rceil\Big\} \subset \Omega_N := \Big\{(n_1,\dots,n_k) \text{ s.t.}\\ n_i - n_{i-1} \in [\![1, \lceil N^b\rceil]\!] \text{ for } i = 2,\dots,k \text{ and } n_1 \in [\![\lceil N^a\rceil, \lceil N^b\rceil]\!]\Big\}\end{aligned} \tag{2}$$

$$\begin{aligned}\Big\{(n_1,\dots,n_k) \text{ s.t. } \lceil N^a\rceil + 1 \le n_1 < \dots < n_k \le \lceil N^b\rceil\Big\} \supset \Gamma_N := \Big\{(n_1,\dots,n_k) \text{ s.t.}\\ n_i - n_{i-1} \in [\![1, \frac{\lceil N^b\rceil - \lceil N^a\rceil}{k}]\!] \text{ for } i = 2,\dots,k\\ \text{and } n_1 \in [\![\lceil N^a\rceil + 1, \lceil N^a\rceil + 1 + \frac{\lceil N^b\rceil - \lceil N^a\rceil}{k}]\!]\Big\},\end{aligned} \tag{3}$$

where we use the notation $[\![A,B]\!] := [A,B] \cap \mathbb{Z}$.
Let now $n_0 = 0$ and recall that $R_h := \sum_{n=1}^h p_{2n}(0)$, $h \ge 1$. Then, using Lemma 1.2.1, we compute an upper bound:

$$\begin{aligned}\mathbb{E}\left[Z_{a,b}^2\right] &\overset{(1.7)}{=} E\left[e^{\sum_{n=\lceil N^a\rceil+1}^{\lceil N^b\rceil}\lambda_2(\beta_N)\mathbb{1}_{S_{2n}=0}}\right]\\ &= E\left[\prod_{n=\lceil N^a\rceil+1}^{\lceil N^b\rceil} e^{\lambda_2(\beta_N)\mathbb{1}_{S_{2n}=0}}\right] \overset{1}{=} E\left[\prod_{n=\lceil N^a\rceil+1}^{\lceil N^b\rceil}(\Lambda_N\mathbb{1}_{S_{2n}=0}+1)\right]\\ &\overset{(1.8)}{=} E\left[1 + \sum_{k=1}^\infty \sum_{\lceil N^a\rceil+1\le n_1<\dots<n_k\le\lceil N^b\rceil}\prod_{i=1}^k \Lambda_N\mathbb{1}_{S_{2n_i}=0}\right]\\ &= 1 + \sum_{k=1}^\infty \sum_{\lceil N^a\rceil+1\le n_1<\dots<n_k\le\lceil N^b\rceil}\prod_{i=1}^k \Lambda_N p_{2(n_i-n_{i-1})}(0).\end{aligned}$$

At this point, introducing $\Omega_N$ from (2), we obtain using (2) that

$$\mathbb{E}\left[Z_{a,b}^2\right] \overset{(2)}{\leq} 1 + \sum_{k=1}^{\infty} \Lambda_N^k \sum_{n_i \in \Omega_N} \prod_{i=1}^{k} p_{2(n_i - n_{i-1})}(0),$$

so, if for $i \geq 2$ we call $\tilde{n} := n_i - n_{i-1}$, by definition of $\Omega_N$ and of $R_h$ we recover that for every $k \geq 1$

$$\begin{aligned} \sum_{n_i \in \Omega_N} \prod_{i=1}^{k} p_{2(n_i - n_{i-1})}(0) &= \sum_{n_i \in \Omega_N} p_{2n_1}(0) p_{2(n_i - n_{i-1})}(0) \cdots p_{2(n_k - n_{k-1})}(0) \\ = \sum_{n_1 \in [\![N^a, N^b]\!]} p_{2n_1}(0) \left( \sum_{\tilde{n} \in [\![1, N^b]\!]} p_{2\tilde{n}}(0) \right)^{k-1} &= \left( R_{\lceil N^b \rceil} - R_{\lceil N^a \rceil} \right) \left( R_{\lceil N^b \rceil} \right)^{k-1}. \end{aligned} \tag{1.11}$$

Now we can conclude with an upper bound:

$$\begin{aligned} \mathbb{E}\left[Z_{a,b}^2\right] &\leq 1 + \sum_{k=1}^{\infty} (\Lambda_N)^k (R_{\lceil N^b \rceil} - R_{\lceil N^a \rceil})(R_{\lceil N^b \rceil})^{k-1} \\ &= 1 + \frac{\Lambda_N (R_{\lceil N^b \rceil} - R_{\lceil N^a \rceil})}{1 - \Lambda_N R_{\lceil N^b \rceil}} \xrightarrow[N \to \infty]{} 1 + \frac{(b-a)\hat{\beta}^2}{1 - b\hat{\beta}^2} \\ &= \frac{1 - a\hat{\beta}^2}{1 - b\hat{\beta}^2} = e^{\lambda_{a,b}^2}, \end{aligned} \tag{1.12}$$

where the convergence holds thanks to (1.3), as it also implies that

$$\Lambda_N \sim \frac{\hat{\beta}^2 \pi}{\log N}, \tag{1.13}$$

because, for $N \to \infty$, $e^{\beta_N^2} - 1 = e^{\hat{\beta}^2 / R_N} - 1 \sim \frac{\hat{\beta}^2}{R_N} \overset{(1.3)}{\sim} \frac{\pi \hat{\beta}^2}{\log N}$.

With similar computations, we can also determine a lower bound.
In fact, we have that

$$\begin{aligned} \mathbb{E}\left[Z_{a,b}^2\right] &= \sum_{k=0}^{\infty} (\Lambda_N)^k \sum_{\lceil N^a \rceil + 1 \leq n_1 < \cdots < n_k \leq \lceil N^b \rceil} \prod_{i=1}^{k} p_{2(n_i - n_{i-1})}(0) \\ &= 1 + \sum_{k=1}^{\infty} (\Lambda_N)^k \sum_{\lceil N^a \rceil + 1 \leq n_1 < \cdots < n_k \leq \lceil N^b \rceil} \prod_{i=1}^{k} p_{2(n_i - n_{i-1})}(0), \end{aligned}$$

thus, setting $\Gamma_N$ as in (3), we get that

$$\mathbb{E}\left[Z_{a,b}^2\right] \overset{(3)}{\geq} 1 + \sum_{k=1}^{\infty} (\Lambda_N)^k \sum_{n_i \in \Gamma_n} \prod_{i=1}^{k} p_{2(n_i - n_{i-1})}(0),$$

so as similarly seen in (1.11), if for $i \geq 2$ we call $\tilde{n} := n_i - n_{i-1}$, by definition of $\Gamma_N$ and of $R_h$ we recover that for every $k \geq 1$

$$\mathbb{E}\left[Z_{a,b}^2\right] \geq 1 + \sum_{k=1}^{\infty} (\Lambda_N)^k \left( \sum_{n_1=\lceil N^a \rceil+1}^{\lceil N^a \rceil+1+\frac{\lceil N^b \rceil - \lceil N^a \rceil}{k}} p_{2n_1}(0) \right) \left( \sum_{\tilde{n}=1}^{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{k}} p_{2\tilde{n}}(0) \right)^{k-1},$$

so that

$$\begin{aligned}
\mathbb{E}\left[Z_{a,b}^2\right] &\geq 1 + \sum_{k=1}^{\infty} (\Lambda_N)^k \left( R_{\lceil N^a \rceil+1+\frac{\lceil N^b \rceil - \lceil N^a \rceil}{k}} - R_{\lceil N^a \rceil+1} \right) \left( R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{k}} \right)^{k-1} \\
&\overset{\forall M_0 \geq 0}{\geq} 1 + \sum_{k=1}^{M_0} (\Lambda_N)^k \left( R_{\lceil N^a \rceil+1+\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}} - R_{\lceil N^a \rceil+1} \right) \left( R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}} \right)^{k-1} \\
&= 1 + \left( \frac{R_{\lceil N^a \rceil+1+\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}} - R_{\lceil N^a \rceil+1}}{R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}}} \right) \left( \frac{1 - \left( \Lambda_N R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}} \right)^{M_0+1}}{1 - \Lambda_N R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}}} - 1 \right).
\end{aligned}$$

So, by (1.3) and (1.13), we can conclude that $\forall M_0 > 0$

$$\begin{aligned}
&\liminf_{N\to\infty} \left( 1 + \left( \frac{R_{\lceil N^a \rceil+1+\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}} - R_{\lceil N^a \rceil+1}}{R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}}} \right) \left( \frac{1 - \left( \Lambda_N R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}} \right)^{M_0+1}}{1 - \Lambda_N R_{\frac{\lceil N^b \rceil - \lceil N^a \rceil}{M_0}}} - 1 \right) \right) \\
&\qquad = 1 + \left( \frac{(b-a)\hat{\beta}^2}{b\hat{\beta}^2} \right) \left( \frac{1 - (b\hat{\beta}^2)^{M_0+1}}{1 - b\hat{\beta}^2} - 1 \right) \\
&\qquad \xrightarrow[M_0\to\infty]{} 1 + \left( \frac{(b-a)\hat{\beta}^2}{b\hat{\beta}^2} \right) \left( \frac{1}{1 - b\hat{\beta}^2} - 1 \right) = \frac{1 - a\hat{\beta}^2}{1 - b\hat{\beta}^2} = e^{\lambda_{a,b}^2},
\end{aligned}$$

and in particular

$$\liminf_{N\to\infty} \mathbb{E}\left[Z_{a,b}^2\right] \geq e^{\lambda_{a,b}^2}.$$

□

# Chapter 2

# Central Limit Theorem

The aim of this chapter is to give an alternative proof of the Central Limit Theorem for $2d$ directed polymers to the already existing proof of [1] and [5]. In particular, we will focus on proving (i). We begin by explaining the high-level structure of the proof.

## 2.1 Heuristics of the proof

**Theorem 2.1.1** (Central Limit Theorem)**.**

$$W_N \xrightarrow[N\to\infty]{\mathcal{L}} e^{\mathcal{N}(-\frac{\lambda^2}{2},\lambda^2)} \qquad \textit{for } \hat{\beta} < 1 \tag{i}$$

$$W_N \xrightarrow[N\to\infty]{\mathbb{P}} 0 \qquad \textit{for } \hat{\beta} \geq 1. \tag{ii}$$

The proof we aim to put forward is divided in multiple steps.
First, as done in (1.6), for $M > 0$ and $k \in \mathbb{N}$, we define the independent variables

$$Z_k(\hat{\beta}, x) := E\left[e^{\sum_{n=t_k+1}^{t_{k+1}}(\beta_N\omega(n,S_n)-\lambda(\beta_N))}\right] \text{with } t_0 = 1,\ t_k = \lceil N^{\frac{k}{M}} \rceil. \tag{2.1}$$

Recall that in Theorem 1.2.4 we proved convergence to $e^{\lambda^2_{a,b}}$ of

$$Z_{a,b}(\hat{\beta}, x) := E\left[e^{\sum_{n=\lceil N^a\rceil+1}^{\lceil N^b\rceil}(\beta_N\omega(n,S_n)-\lambda(\beta_N))}\right],$$

for any $a, b \in [0, 1]$, $a \leq b$. This also proves that the second moment of the partition function $W_N$ and the second moment of the product of the $Z_k$'s as $N$ tends to infinity converge to $e^{\lambda^2}$.

Then, the first step in the proof of (ii) is to show that the product of the $Z_k$'s approximates the partition function of the $2d$ polymer in $L^2$, namely that for every $M > 0$ we have

$$W_N - \prod_{k=0}^{M-1} Z_k \xrightarrow[N\to\infty]{L^2} 0. \tag{2.2}$$

One can prove this statement by considering three copies of the random walk starting from the same point and proving some asymptotic independence.

Then, we define the centered random variable $U_k := Z_k - 1$ and consequently prove that

$$\log \prod_{k=0}^{M-1} Z_k = \sum_{k=0}^{M-1} \log(1+U_k)$$

is close in probability to

$$\sum_{k=0}^{M-1} (U_k - \frac{1}{2}U_k^2).$$

The key step in the proof of this *Taylor expansion* is a quantitative bound on the $2+\varepsilon$ moment of the $U_k$, which can be recovered thanks to *hypercontractivity* of polynomial chaos ([2], eq. (3.10)). In particular, we are able to state that

$$\mathbb{E}[U_k^{2+\varepsilon}] \leq \frac{c}{M^{1+\frac{\varepsilon}{2}}}, \tag{2.3}$$

where $c > 0$ denotes a constant depending only on $\hat{\beta}$.

We continue the proof of the theorem by using a bound on the Wasserstein distance between the law of the sum of the $U_k$'s and the law of a Gaussian with mean 0 and variance $\lambda^2$ ([4], eq. (1.3b)). In particular:

$$\mathcal{W}_1\left(\mathcal{L}\left(\sum_{k=0}^{M-1} U_k\right), \mathcal{N}(0,\lambda^2)\right) \leq c \sum_{k=0}^{M-1} \mathbb{E}[U_k^{2+\varepsilon}] \overset{(2.3)}{\leq} \frac{c}{M^{\frac{\varepsilon}{2}}}.$$

Furthermore, by a truncation argument, (2.3) enables us to argue that $\sum_{k=0}^{M-1} \frac{U_k^2}{2}$ concentrates around its mean, which we prove to be asymptotically close to $\frac{\lambda^2}{2}$.

At this point we can conclude: combining (2.2) and the fact that $\log \prod_{k=0}^{M-1} Z_k$ is close in probability to $\sum_{k=0}^{M-1} \left(U_k - \frac{U_k^2}{2}\right)$, which is itself close to $\mathcal{N}\left(-\frac{\lambda^2}{2}, \lambda^2\right)$, the conclusion follows

$$W_N \xrightarrow[N\to\infty]{\mathcal{L}} e^{\mathcal{N}(-\frac{\lambda^2}{2},\lambda^2)}.$$

**Remark 2.1.2** (Connection to branching random walks)**.** *As a byproduct of the proof, it holds that pointwise* $\log W_N(x) - \mathbb{E}[\log W_N(x)] \approx M_n(x)$*, where* $M_n(x) = \sum_{k=0}^{M-1} Z_k(x) - 1$ *is a sum of independent random variables (with non-constant variance). It turns out that* $(M_N(x))$*,* $|x| \leq \sqrt{N}$ *can be interpreted as a branching random walk process (very roughly, the points* $|x| \leq \sqrt{N}$ *stand for the leaves of the binary tree of the corresponding branching random walk). We refer to [7] for an introduction to Branching random walks and their relation to the maximum of log-correlated fields. We note that in general, studying the maximum of log-correlated fields requires a comparison to some underlying branching random walk. Here, we believe that* $M_n(x)$ *is a good candidate for such a comparison, which may help to obtain sharp estimates on the maximum of* $\log W_N(x)$ *for* $|x| \leq \sqrt{N}$ *(see [3] for a more extended discussion about this question).*

## 2.2 Step 1

**Proposition 2.2.1.** *For all integers $M > 0$ we have*

$$W_N - \prod_{k=0}^{M-1} Z_k \xrightarrow[N\to\infty]{L^2} 0. \tag{2.4}$$

*Proof.* Recall that $t_1 = \lceil N^{1/M} \rceil$. Fix $M > 0$ and consider

$$Z_0(\hat{\beta}, x) = E\left[e^{\sum_{n=1}^{t_1} h_n}\right] \qquad \tilde{Z}_0(\hat{\beta}, x) = E\left[e^{\sum_{n=t_1+1}^{N} h_n}\right],$$

where we defined

$$h_n := \beta_N \omega(n, S_n) - \lambda(\beta_N).$$

In particular, $Z_0 \perp\!\!\!\perp \tilde{Z}_0$ as

$$Z_0 \in \sigma\left(\left\{\omega(n,\cdot) \text{ for } n \le \lceil N^{1/M} \rceil\right\}\right), \qquad \tilde{Z}_0 \in \sigma\left(\left\{\omega(n,\cdot) \text{ for } n \ge \lceil N^{1/M} \rceil + 1\right\}\right).$$

We first prove that

$$\mathbb{E}[|W_N - Z_0\tilde{Z}_0|^2] \xrightarrow[N\to\infty]{} 0. \tag{2.5}$$

We have $\mathbb{E}[|W_N - Z_0\tilde{Z}_0|^2] = \mathbb{E}[W_N^2] - 2\mathbb{E}[W_N Z_0 \tilde{Z}_0] + \mathbb{E}[Z_0^2\tilde{Z}_0^2]$ and by Theorem 1.2.4 we know that

$$\mathbb{E}[W_N^2] \xrightarrow[N\to\infty]{} e^{\lambda^2} \tag{2.6}$$

$$\mathbb{E}[Z_0^2\tilde{Z}_0^2] = \mathbb{E}[Z_0^2]\mathbb{E}[\tilde{Z}_0^2] \xrightarrow[N\to\infty]{} \frac{1}{1 - \frac{1}{M}\hat{\beta}^2} \cdot \frac{1 - \frac{1}{M}\hat{\beta}^2}{1 - \hat{\beta}^2} = e^{\lambda^2}; \tag{2.7}$$

thus, it suffices to prove that $\mathbb{E}[W_N Z_0 \tilde{Z}_0] \xrightarrow[N\to\infty]{} e^{\lambda^2}$.

Consider $S^1, S^2$ and $S^3$ three independent copies of the random walk starting all from the origin. We can write

$$\begin{aligned} W_N Z_0 \tilde{Z}_0 &= E\left[e^{\sum_{n=1}^{N} h_n}\right] E\left[e^{\sum_{n=1}^{t_1} h_n}\right] E\left[e^{\sum_{n=t_1+1}^{N} h_n}\right] \\ &= E^{\otimes 3}\left[e^{\sum_{n=1}^{N} h_n^1} e^{\sum_{n=1}^{t_1} h_n^2} e^{\sum_{n=t_1+1}^{N} h_n^3}\right] \\ &= E^{\otimes 3}\left[e^{\sum_{n=1}^{t_1} h_n^{1,2}} e^{\sum_{n=t_1+1}^{N} h_n^{1,3}}\right], \end{aligned}$$

where we set

$$\begin{aligned} h_n^i &:= \beta_N \omega(n, S_n^i) - \lambda(\beta_N) \text{ for } i \in \{1,2,3\} \\ h_n^{1,j} &:= \beta_N \omega(n, S_n^1) + \beta_N \omega(n, S_n^j) - 2\lambda(\beta_N) \text{ for } j \in \{2,3\}. \end{aligned}$$

We then notice that

$$e_1 = e^{\sum_{n=1}^{t_1} h_n^{1,2}} \in \sigma\left(\left\{\omega_n \text{ for } n \le \lceil N^{1/M} \rceil\right\}\right)$$
$$e_2 = e^{\sum_{n=t_1+1}^{N} h_n^{1,3}} \in \sigma\left(\left\{\omega_n \text{ for } n \ge \lceil N^{1/M} \rceil + 1\right\}\right),$$

thus $e_1 \perp\!\!\!\perp e_2$ with respect to $\mathbb{P}$ and we can write

$$\mathbb{E}[W_N Z_0 \tilde{Z}_0] = \mathbb{E}\left[E^{\otimes 3}[e_1 e_2]\right] = E^{\otimes 3}\left[\mathbb{E}[e_1 e_2]\right] = E^{\otimes 3}\left[\mathbb{E}[e_1]\mathbb{E}[e_2]\right].$$

Defining now

$$l_n^{1,j} := \lambda_2(\beta_N) \mathbb{1}_{S_n^1 = S_n^j} \text{ for } j \in \{2,3\},$$

analogously seen in the proof of Lemma 1.2.1, we have

$$\mathbb{E}\left[e^{\sum_{n=1}^{T} h_n^{1,j}}\right] = e^{\sum_{n=1}^{T} l_n^{1,j}},$$

which implies

$$\mathbb{E}[e_1] = e^{\sum_{n=1}^{t_1} l_n^{1,2}}, \quad \mathbb{E}[e_2] = e^{\sum_{n=t_1+1}^{N} l_n^{1,3}}$$

and thus

$$\mathbb{E}[W_N Z_0 \tilde{Z}_0] = E^{\otimes 3}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} e^{\sum_{n=t_1+1}^{N} l_n^{1,3}}\right].$$

We would now like to split the triple mean into the product of two means: such thing cannot be done right away as the two factors inside the mean are not independent ($S_n^1$ is present in both). However, the number of times in which $S_n^1$ and $S_n^2$ meet should not interfere on the long run with the number of times $S_n^1$ and $S_n^3$ meet, therefore some splitting should be allowed.

Before going on with the computations, fix $\alpha > 0$ and define the set

$$A_N = \{x \in \mathbb{Z}^2 : |x| \le \alpha^{-1}\sqrt{t_1}, |x|_1 = 2k \text{ for some } k \in \mathbb{N}\},$$

where with $|\cdot|$ we denote the $L^2$ norm and with $|\cdot|_1$ the $L^1$ norm. This will be useful when considering the difference $S_n^1 - S_n^3$, which is in law equal to $S_{2n}$: indeed, as the copies of the random walk we are considering start at the same point, they have the same parity and thus their difference only proceeds with even steps. Therefore, calling $l_{2n} := \lambda_2(\beta_N)\mathbb{1}_{S_{2n}=0}$, we get

$$E_{x,y}^{\otimes 2}\left[e^{\sum_{n=1}^{N-t_1} l_n^{1,3}}\right] = E_{x-y}\left[e^{\sum_{n=1}^{N-t_1} l_{2n}}\right]. \tag{2.8}$$

We are now in the position to continue:

$$\begin{aligned}
\mathbb{E}[W_N Z_0 \tilde{Z}_0] &= E^{\otimes 3}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} e^{\sum_{n=t_1+1}^{N} l_n^{1,3}}\right] \\
&\overset{\text{Markov Property}}{=} E^{\otimes 3}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} E_{S_{t_1}^1, S_{t_1}^3}^{\otimes 2}\left[e^{\sum_{n=1}^{N-t_1} l_n^{1,3}}\right]\right] \\
&\ge E^{\otimes 3}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} \mathbb{1}_{S_{t_1}^1 - S_{t_1}^3 \in A_N} E_{S_{t_1}^1, S_{t_1}^3}^{\otimes 2}\left[e^{\sum_{n=1}^{N-t_1} l_n^{1,3}}\right]\right] \\
&\overset{(2.8)}{=} E^{\otimes 3}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} \mathbb{1}_{S_{t_1}^1 - S_{t_1}^3 \in A_N} E_{S_{t_1}^1 - S_{t_1}^3}\left[e^{\sum_{n=1}^{N-t_1} l_{2n}}\right]\right] \\
&\ge E^{\otimes 2}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} \mathbb{1}_{S_{t_1}^1 - S_{t_1}^3 \in A_N}\right] \inf_{x \in A_N} E_x\left[e^{\sum_{n=1}^{N-t_1} l_{2n}}\right].
\end{aligned}$$

Define now

$$\psi_N = E^{\otimes 2}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} \mathbb{1}_{S^1_{t_1}-S^3_{t_1}\in A_N}\right], \qquad \varphi_N = \inf_{x\in A_N} E_x\left[e^{\sum_{n=1}^{N-t_1} l_{2n}}\right],$$

so that $\mathbb{E}[W_N Z_0 \tilde{Z}_0] \geq \psi_N \varphi_N$. Hence, it suffices to show that

$$\liminf_{\alpha\to 0} \liminf_{N\to\infty} \psi_N \geq e^{\lambda^2_{0,1/M}} \tag{2.9}$$

and that

$$\lim_{\alpha\to 0} \liminf_{N\to\infty} \varphi_N \geq e^{\lambda^2_{1/M,1}}. \tag{2.10}$$

So, for $\alpha > 0$ arbitrary, we would obtain that

$$\liminf_{N\to\infty} \mathbb{E}[W_N Z_0 \tilde{Z}_0] \geq e^{\lambda^2_{0,1/M}} e^{\lambda^2_{1/M,1}} = e^{\lambda^2}. \tag{2.11}$$

First, we estimate $\psi_N$: as $\mathbb{1}_A = 1 - \mathbb{1}^c_A$, we have

$$\psi_N = E^{\otimes 2}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}}\right] - E^{\otimes 2}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} \mathbb{1}_{S^1_{t_1}-S^3_{t_1}\notin A_N}\right] \tag{2.12}$$

but we know from Theorem 1.2.4 that, recalling $t_1 = \lceil N^{1/M} \rceil$,

$$E^{\otimes 2}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}}\right] \xrightarrow[N\to\infty]{} e^{\lambda^2_{0,1/M}} = \frac{1}{1-\frac{1}{M}\hat{\beta}}, \tag{2.13}$$

so it suffices to estimate the second term of (2.12). By Hölder Inequality with $p, q \geq 1$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\begin{aligned} E^{\otimes 2}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} \mathbb{1}_{S^1_{t_1}-S^3_{t_1}\notin A_N}\right] &\leq E^{\otimes 2}\left[\left(e^{\sum_{n=1}^{t_1} l_n^{1,2}}\right)^p\right]^{\frac{1}{p}} P^{\otimes 2}\left(S^1_{t_1} - S^3_{t_1} \notin A_N\right)^{\frac{1}{q}} \\ &= E^{\otimes 2}\left[\left(e^{\sum_{n=1}^{t_1} l_n^{1,2}}\right)^p\right]^{\frac{1}{p}} P\left(|S_{2t_1}| > \alpha^{-1} N^{\frac{1}{2M}}\right)^{\frac{1}{q}}, \end{aligned}$$

so that

$$\limsup_{\alpha\to 0} \limsup_{N\to\infty} E^{\otimes 2}\left[e^{\sum_{n=1}^{t_1} l_n^{1,2}} \mathbb{1}_{S^1_{t_1}-S^3_{t_1}\notin A_N}\right] = 0. \tag{2.14}$$

Notice that (2.14) holds true because, for $p = p(\hat{\beta}) \geq 1$ small enough, we have

$$E^{\otimes 2}\left[e^{p\lambda_2(\beta_N)\sum_{n=1}^{t_1} \mathbb{1}_{S^1_n = S^2_n}}\right]^{\frac{1}{p}} \xrightarrow[N\to\infty]{} e^{\lambda(p\hat{\beta})^2_{0,\frac{1}{M}}} = \frac{1}{1-\frac{1}{M}p\hat{\beta}} \leq \frac{1}{1-p\hat{\beta}} < \infty$$

and, by Hoeffding's inequality, we also have

$$P\left(|S_{2t_1}| > \alpha^{-1} N^{\frac{1}{2M}}\right)^{\frac{1}{q}} \leq e^{-c\alpha^{-2}} \xrightarrow[\alpha\to 0]{} 0.$$

So by (2.12), (2.13) and (2.14) we have proven (2.9).

Now we want to prove (2.10). Take $x \in A_N$. Then:

$$
\begin{aligned}
E_x\left[e^{\sum_{n=1}^{N-t_1} l_{2n}}\right] &= E_x\left[\prod_{n=1}^{N-t_1} e^{l_{2n}}\right] = E_x\left[\prod_{n=1}^{N-t_1}\left(e^{l_{2n}}-1+1\right)\right] \\
&= E_x\left[\prod_{n=1}^{N-t_1}(1+\Lambda_N \mathbb{1}_{S_{2n}=0})\right] \overset{(1.8)}{=} 1+\sum_{k=1}^{N-t_1}\Lambda_N^k \sum_{1\le n_1<\dots<n_k\le N-t_1} E_x\left[\prod_{i=1}^{k}\mathbb{1}_{S_{2n_i}=0}\right] \\
&= \sum_{k=0}^{N-t_1}\Lambda_N^k \sum_{1\le n_1<\dots<n_k\le N-t_1} p_{2n_1}(x)\prod_{i=2}^{k} p_{2(n_i-n_{i-1})}(0) \\
&\ge \sum_{k=0}^{N-t_1}\Lambda_N^k \sum_{\alpha^{-3}t_1\le n_1<\dots<n_k\le N-t_1} p_{2n_1}(x)\prod_{i=2}^{k} p_{2(n_i-n_{i-1})}(0) \\
&\overset{\text{LLT}}{\ge} e^{-c\alpha}\left(\sum_{k=0}^{N-t_1}\Lambda_N^k \sum_{\alpha^{-3}t_1\le n_1<\dots<n_k\le N-t_1} \prod_{i=1}^{k} p_{2(n_i-n_{i-1})}(0)\right)(1+o(1)) \qquad (2.15)\\
&\overset{(1.8)}{=} e^{-c\alpha}E_0\left[e^{\sum_{\alpha^{-3}t_1}^{N-t_1}\lambda_2(\beta_N)\mathbb{1}_{S_{2n}=0}}\right](1+o(1)), \qquad (2.16)
\end{aligned}
$$

where (2.15) holds because, by the Local Limit Theorem ([6, Sec. 1.2]), we have for $x \in A_N$ and $n_1 \ge \alpha^{-3}t_1$ that

$$p_{2n_1}(x) \ge e^{-c\frac{|x|^2}{n_1}} p_{2n_1}(0)(1+o(1)) \ge e^{-c\frac{\alpha^{-2}}{\alpha^{-3}}} p_{2n_1}(0)(1+o(1)) = e^{-c\alpha}p_{2n_1}(0)(1+o(1)).$$

Now, for all fixed $\alpha > 0$, the proof of Theorem 1.2.4 entails that:

$$\lim_{N\to\infty} E_0\left[e^{\sum_{n=\alpha^{-3}t_1}^{N-t_1} l_{2n}}\right] = e^{\lambda^2_{1/M,1}} = \frac{1-\frac{1}{M}\hat{\beta}}{1-\hat{\beta}}. \qquad (2.17)$$

At this point, (2.10) is proven thanks to (2.16) and (2.17). Therefore, putting together (2.6), (2.7), (2.9) and (2.10) we deduce (2.5), as desired.

At this point, we notice that by repeating a similar argument we have

$$\tilde{Z}_0 - Z_1\tilde{Z}_1 \xrightarrow[N\to\infty]{L^2} 0, \qquad (2.18)$$

where

$$Z_1 = E\left[e^{\sum_{n=t_1+1}^{t_2} h_n}\right], \qquad \tilde{Z}_1 = E\left[e^{\sum_{n=t_2+1}^{N} h_n}\right].$$

Combining what we know so far, we get

$$W_N - Z_0Z_1\tilde{Z}_1 = W_N - Z_0\tilde{Z}_0 + Z_0\tilde{Z}_0 - Z_0Z_1\tilde{Z}_1 = W_N - Z_0\tilde{Z}_0 + Z_0(\tilde{Z}_0 - Z_1\tilde{Z}_1) \quad (2.19)$$

but $Z_0 \perp\!\!\!\perp \tilde{Z}_0 - Z_1\tilde{Z}_1$, so by (2.18)

$$\mathbb{E}[|Z_0(\tilde{Z}_0 - Z_1\tilde{Z}_1)|^2] = \mathbb{E}[|Z_0\|^2]\mathbb{E}[|\tilde{Z}_0 - Z_1\tilde{Z}_1|^2] \xrightarrow[N\to\infty]{} 0$$

and therefore, by (2.5) and (2.19), we have

$$W_N - Z_0 Z_1 \tilde{Z}_1 \xrightarrow[N\to\infty]{L^2} 0.$$

In a similar way, one can prove that $\forall k \leq M-1$ we have

$$\tilde{Z}_{k-1} - Z_k \tilde{Z}_k \xrightarrow[N\to\infty]{L^2} 0, \tag{2.20}$$

where we defined

$$Z_k = E\left[e^{\sum_{n=t_k+1}^{t_{k+1}} h_n}\right] \qquad \tilde{Z}_k = E\left[e^{\sum_{n=t_{k+1}+1}^{N} h_n}\right].$$

Then, by (2.20) and independence of $\tilde{Z}_{k-1} - Z_k \tilde{Z}_k$ and $(Z_i)_{i\leq k-1}$, we have

$$W_N - \prod_{k=0}^{M-1} Z_k = W_N - Z_0\tilde{Z}_0 + Z_0\tilde{Z}_0 - Z_0 Z_1 \tilde{Z}_1 + \cdots + \prod_{i=0}^{M-2} Z_i \tilde{Z}_{M-2} - \prod_{i=0}^{M-2} Z_i Z_{M-1}\tilde{Z}_{M-1}$$

$$= W_N - Z_0\tilde{Z}_0 + Z_0(\tilde{Z}_0 - Z_1\tilde{Z}_1) + \cdots + \prod_{i=0}^{M-2} Z_i(\tilde{Z}_{M-2} - Z_{M-1}\tilde{Z}_{M-1}) \xrightarrow[N\to\infty]{L^2} 0$$

and the proof is thus concluded.

□

## 2.3 Step 2

Define $U_k := Z_k - 1$. Notice that $U_k$ is, for every value of $k$, a centered random variable and that $U_0, \ldots, U_{M-1}$ are independent.

**Lemma 2.3.1** (Variance of $U_k$). *There exists $C_{\hat{\beta}} > 0$ such that*

$$\limsup_{N\to\infty} \sup_{k\leq M-1} \mathbb{E}\left[U_k^2\right] \leq \frac{C_{\hat{\beta}}}{M} \tag{2.21}$$

*Proof.* By the same computations as in (1.12), we can see that for any $M > 0$ fixed and for any $k \leq M-1$

$$\mathbb{E}[U_k^2] = \mathbb{E}[Z_k^2] - 1 \overset{(1.12)}{\leq} \frac{\Lambda_N(R_{t_{k+1}} - R_{t_k})}{1 - \Lambda_N R_{t_{k+1}}}. \tag{2.22}$$

Moreover, by (1.3) and (1.13), we have that, for some constant $C_{\hat{\beta}}$ depending on $\hat{\beta}$

$$\limsup_{N\to\infty} \frac{1}{1-\Lambda_N R_N} \leq C_{\hat{\beta}} < \infty, \tag{2.23}$$

and as $N \to \infty$,

$$\Lambda_N(R_{t_{k+1}} - R_{t_k}) \sim \frac{\hat{\beta}^2}{M}. \tag{2.24}$$

So, by (2.22), (2.23) and (2.24) there exists $C_{\hat{\beta}} > 0$ such that (2.21) is verified. □

In fact, by hypercontractivity of polynomial chaos ([2, (3.10)]), we can push this estimate to $2+\varepsilon_0$ moments of $U_k$, for some $\varepsilon_0$, as explained in the next lemma.

**Lemma 2.3.2** (Moment estimate)**.** *There exist $\varepsilon_0 = \varepsilon_0(\hat{\beta}) \in (0,1)$ so that $\forall M > 0$ we have*

$$\limsup_{N\to\infty} \sup_{k\le M-1} \mathbb{E}[|U_k|^{2+\varepsilon_0}] \le \frac{c}{M^{1+\frac{\varepsilon_0}{2}}} \tag{2.25}$$

*with $c = c_{\hat{\beta}}$.*

*Proof.* Set $X_0^{(N)} := 0$ and

$$X_k^{(N)} := \sum_{t_k+1\le n_1<\cdots<n_k\le t_{k+1}} E\left[\prod_{i=1}^{k}\left(e^{\beta_N\omega(n_i,S_{n_i})-\lambda(\beta_N)}-1\right)\right].$$

We can then express the variables $U_k$ in terms of the variables $X_k$ by chaos expansion as follows:

$$\begin{aligned} U_k = Z_k - 1 &= E\left[e^{\sum_{n=t_k+1}^{t_{k+1}}(\beta_N\omega(n,S_n)-\lambda(\beta_N))}\right]-1 \\ &= E\left[\prod_{n=t_k+1}^{t_{k+1}} e^{\beta_N\omega(n,S_n)-\lambda(\beta_N)}\right]-1. \end{aligned}$$

Then, adding and subtracting 1 inside the product, we obtain

$$\begin{aligned} U_k &= E\left[\prod_{n=t_k+1}^{t_{k+1}}\left(1+\left(e^{\beta_N\omega(n,S_n)-\lambda(\beta_N)}-1\right)\right)-1\right] \\ &\overset{(1.8)}{=} \sum_{k=1}^{\infty}\sum_{t_k+1\le n_1<\cdots<n_k\le t_{k+1}} E\left[\prod_{i=1}^{k}\left(e^{\beta_N\omega(n_i,S_{n_i})-\lambda(\beta_N)}-1\right)\right]. \end{aligned}$$

In particular, we have

$$U_k = \sum_{k=1}^{\infty} X_k^{(N)}.$$

Now we can apply hypercontractivity (see [2, 3.10]) and deduce that for every $\varepsilon > 0$ there exists a constant $c_\varepsilon$ uniform in $N$ such that $c_\varepsilon \to 1$ as $\varepsilon \to 0$ and

$$\begin{aligned} \mathbb{E}\left[|U_k|^{2+\varepsilon}\right] &= \mathbb{E}\left[\left|\sum_{k=1}^{\infty} X_k^{(N)}\right|^{2+\varepsilon}\right] \le \left(\sum_{k=1}^{\infty} c_\varepsilon^{2k}\mathbb{E}\left[(X_k^{(N)})^2\right]\right)^{1+\frac{\varepsilon}{2}} \\ &= \left(\sum_{k=1}^{\infty} c_\varepsilon^{2k}\sum_{t_k+1\le n_1<\cdots<n_k\le t_{k+1}} \mathbb{E}\left[\prod_{i=1}^{k}\left(\Lambda_N \mathbb{1}_{S_{2(n_i-n_{i-1})}=0}\right)\right]\right)^{1+\frac{\varepsilon}{2}}. \end{aligned}$$

At this point we can choose $\varepsilon_0 = \varepsilon_0(\hat\beta) \in (0,1)$ such that $c_{\varepsilon_0}\hat\beta < 1$ and conclude, similarly to (1.12), that for $N$ large enough

$$\mathbb{E}\left[|U_k|^{2+\varepsilon_0}\right] \leq \left(\frac{c_{\varepsilon_0}^2 \Lambda_N (R_{t_{k+1}} - R_{t_k})}{1 - c_{\varepsilon_0}^2 \Lambda_N R_N}\right)^{1+\frac{\varepsilon_0}{2}} \leq \left(\frac{(c_{\varepsilon_0}\hat\beta)^2}{1-(c_{\varepsilon_0}\hat\beta)^2}\frac{1}{M}\right)^{1+\frac{\varepsilon_0}{2}} = \frac{c_{\hat\beta}}{M^{1+\frac{\varepsilon_0}{2}}},$$

where the second inequality follows from (1.3), (1.13) and the Local Limit Theorem ([6, Sec. 1.2]), as similarly seen in the proof of (2.21). □

From this point on, we fix $\varepsilon_0 \in (0,1)$ such that (2.25) holds.

**Proposition 2.3.3.** *For all $M > 0$, for all $\varepsilon > 0$ and*

$$\limsup_{N\to\infty} \mathbb{P}\left(\left|\sum_{k=0}^{M-1} \log(1+U_k) - \sum_{k=0}^{M-1}\left(U_k - \frac{U_k^2}{2}\right)\right| \geq \varepsilon\right) \leq \frac{2c_{\hat\beta}}{\varepsilon^{2+\varepsilon_0} M^{\frac{\varepsilon_0}{2}}}. \tag{2.26}$$

*Proof.* For $\varepsilon > 0$, define the events

$$F_{M,\varepsilon} := \bigcap_{k\leq M-1} \{|U_k| < \varepsilon\} \qquad F_{M,\varepsilon}^c := \bigcup_{k\leq M-1} \{|U_k| \geq \varepsilon\}$$

and write $f(U) := \sum_{k=0}^{M-1} \log(1+U_k) - \sum_{k=0}^{M-1}(U_k - \frac{1}{2}U_k^2)$. Then:

$$\mathbb{P}(|f(U)| \geq \varepsilon) = \mathbb{P}(|f(U)| \geq \varepsilon, F_{M,\varepsilon}) + \mathbb{P}(|f(U)| \geq \varepsilon, F_{M,\varepsilon}^c).$$

Let's estimate the two terms on the right hand side separately. For the first term, recalling that by Taylor expansion, there exists $\varepsilon > 0$ small enough such that $\forall |x| \leq \varepsilon$

$$\left|\log(1+x) - \left(x - \frac{x^2}{2}\right)\right| \leq |x|^{2+\varepsilon_0}, \tag{2.27}$$

we have that

$$\begin{aligned}\mathbb{P}(|f(U)| \geq \varepsilon, F_{M,\varepsilon}) &\overset{(2.27)}{\leq} \mathbb{P}\left(\sum_{k=0}^{M-1} |U_k|^{2+\varepsilon_0} \geq \varepsilon, F_{M,\varepsilon}\right) \leq \mathbb{P}\left(\sum_{k=0}^{M-1} |U_k|^{2+\varepsilon_0} \geq \varepsilon\right) \\ &\overset{\text{Markov}}{\leq} \frac{\mathbb{E}\left[\sum_{k=0}^{M-1} |U_k|^{2+\varepsilon_0}\right]}{\varepsilon} \leq \frac{M}{\varepsilon} \sup_{k\leq M-1} \mathbb{E}\left[|U_k|^{2+\varepsilon_0}\right] \overset{(2.25)}{\leq} \frac{c_{\hat\beta}}{\varepsilon M^{\frac{\varepsilon_0}{2}}}.\end{aligned} \tag{2.28}$$

On the other hand, for the second term we have

$$\begin{aligned}\mathbb{P}(|f(U)| \geq \varepsilon, F_{M,\varepsilon}^c) &\leq \mathbb{P}(F_{M,\varepsilon}^c) = \mathbb{P}\left(\cup_{k\leq M}\{|U_k| \geq \varepsilon\}\right) \leq M \sup_{k\leq M} \mathbb{P}(|U_k| \geq \varepsilon) \\ &\overset{\text{Markov}}{\leq} M \sup_{k\leq M} \frac{\mathbb{E}\left[|U_k|^{2+\varepsilon_0}\right]}{\varepsilon^{2+\varepsilon_0}} \overset{(2.25)}{\leq} \frac{c_{\hat\beta}}{\varepsilon^{2+\varepsilon_0} M^{\frac{\varepsilon_0}{2}}}.\end{aligned} \tag{2.29}$$

Thus, noticing that $\frac{1}{\varepsilon} < \frac{1}{\varepsilon^{2+\varepsilon_0}}$, we conclude using (2.28) and (2.29)

$$\mathbb{P}(|f(U)| \geq \varepsilon) \leq \frac{c_{\hat\beta}}{\varepsilon M^{\frac{\varepsilon_0}{2}}} + \frac{c_{\hat\beta}}{\varepsilon^{2+\varepsilon_0} M^{\frac{\varepsilon_0}{2}}} \leq \frac{2c_{\hat\beta}}{\varepsilon^{2+\varepsilon_0} M^{\frac{\varepsilon_0}{2}}}.$$

□

## 2.4 Step 3

As last step of this proof, we will find estimates on the distance between $\sum_{k=0}^{M-1}\left(U_k - \frac{U_k^2}{2}\right)$ and $\mathcal{N}(-\frac{\lambda^2}{2}, \lambda^2)$ uniformly in $N$. First of all, keeping in mind (2), we have the following lemma.

**Lemma 2.4.1.** *We have that*

$$\limsup_{N\to\infty} \mathcal{W}_1\left(\mathcal{L}\left(\sum_{k=0}^{M-1} U_k\right), \mathcal{N}(0,\lambda^2)\right) \leq \frac{c'_{\hat{\beta}}}{M^{\frac{\varepsilon_0}{2}}}. \tag{2.30}$$

*Proof.*

$$\mathcal{W}_1\left(\mathcal{L}\left(\sum_{k=0}^{M-1} U_k\right), \mathcal{N}(0,\lambda^2)\right) \leq C\sum_{k=0}^{M-1} \mathbb{E}[U_k^{2+\varepsilon_0}] \overset{(2.25)}{\leq} \frac{c'_{\hat{\beta}}}{M^{\frac{\varepsilon_0}{2}}},$$

where the first inequality holds by the uniform estimate on the Wasserstein distance between the law of the sum of independent random variables and the law of a Gaussian that we find in [4, (1.3b)]. □

Now, the only thing left to show is that $\sum_{k=0}^{M-1}\frac{U_k^2}{2} \xrightarrow[M\to\infty]{} \frac{\lambda^2}{2}$ uniformly in $N$.

**Lemma 2.4.2.** *For every $\varepsilon > 0$,*

$$\limsup_{M\to\infty}\limsup_{N\to\infty} \mathbb{P}\left(\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2} - \frac{\lambda^2}{2}\right| > \varepsilon\right) = 0. \tag{2.31}$$

*Proof.* Take $\alpha > 0$ and $\varepsilon > 0$. First, we want to show that we can truncate the variables $U_k$ as follows:

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{k=0}^{M-1} U_k^2 - \sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right|\right] &= \mathbb{E}\left[\sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2>\frac{\alpha}{M}}\right] \leq \mathbb{E}\left[\sum_{k=0}^{M-1} |U_k|^{2+\varepsilon_0}\frac{M^{\frac{\varepsilon_0}{2}}}{\alpha^{\frac{\varepsilon_0}{2}}}\right] \\ &= \frac{M^{\frac{\varepsilon_0}{2}}\sum_{k=0}^{M-1}\mathbb{E}\left[|U_k|^{2+\varepsilon_0}\right]}{\alpha^{\frac{\varepsilon_0}{2}}} \overset{(2.25)}{\leq} \frac{c_{\hat{\beta}}}{\alpha^{\frac{\varepsilon_0}{2}}},\end{aligned} \tag{2.32}$$

where the last inequality holds for $N \geq N_0(M)$ large enough. This also implies that

$$\left|\mathbb{E}\left[\sum_{k=0}^{M-1} U_k^2\right] - \mathbb{E}\left[\sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right]\right| \leq \mathbb{E}\left[\left|\sum_{k=0}^{M-1} U_k^2 - \sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right|\right] \overset{(2.32)}{\leq} \frac{c_{\hat{\beta}}}{\alpha^{\frac{\varepsilon_0}{2}}}. \tag{2.33}$$

Moreover, by Theorem 1.2.4, we have that

$$\begin{aligned}\mathbb{E}\left[\sum_{k=0}^{M-1} U_k^2\right] &= \sum_{k=0}^{M-1} \mathbb{E}\left[U_k^2\right] = \sum_{k=0}^{M-1} \left(\mathbb{E}\left[Z_k^2\right]-1\right) \xrightarrow[N\to\infty]{\text{Theorem 1.2.4}} \sum_{k=0}^{M-1}\left(\frac{1-\frac{k}{M}\hat\beta^2}{1-\frac{k+1}{M}\hat\beta^2}-1\right)\\ &= \sum_{k=0}^{M-1}\frac{1}{M}\left(\frac{\hat\beta^2}{1-\hat\beta^2\frac{k+1}{M}}\right) \xrightarrow[M\to\infty]{\text{Riemann sum}} \int_0^1 \frac{\hat\beta^2}{1-\hat\beta^2 x}dx\\ &= -\log\left(\frac{1}{\hat\beta^2}-1\right)+\log\left(\frac{1}{\hat\beta^2}\right) = \log\left(\frac{1}{1-\hat\beta^2}\right) = \lambda^2. \end{aligned} \tag{2.34}$$

At this point, equation (2.34) together with (2.33) implies that, for $N \geq N_0(M)$ large enough

$$\left|\mathbb{E}\left[\sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}\right]-\lambda^2\right| \leq \frac{c_{\hat\beta}}{\alpha^{\frac{\varepsilon_0}{2}}}, \tag{2.35}$$

Therefore, the only thing left to prove is that $\sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}$ is close to its mean. Indeed, as the random variables $U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}} - \mathbb{E}\left[U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}\right]$ are independent and centered, we have:

$$\begin{aligned}&\mathbb{P}\left(\left|\sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}} - \mathbb{E}\left[\sum_{k=0}^{M-1} U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}\right]\right| > \varepsilon\right)\\ &\leq \mathbb{P}\left(\sum_{k=0}^{M-1}\left|U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}} - \mathbb{E}\left[U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}\right]\right| > \varepsilon\right)\\ &\overset{\text{Chebyshev}}{\leq} \varepsilon^{-1}\sum_{k=0}^{M-1}\mathbb{E}\left[\left(U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}} - \mathbb{E}\left[U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}\right]\right)^2\right]\\ &\leq \varepsilon^{-1}\sum_{k=0}^{M-1}\left(\mathbb{E}\left[\left(U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}\right)^2\right] + \mathbb{E}\left[U_k^2 \mathbb{1}_{U_k^2\leq \frac{\alpha}{M}}\right]^2\right)\\ &\leq 2\frac{M}{\varepsilon}\frac{\alpha^2}{M^2} = 2\frac{\alpha^2}{\varepsilon M}. \end{aligned} \tag{2.36}$$

All these computations imply that, for $N \geq N_0(M)$ large enough and for $\varepsilon > 0$

$$\begin{aligned}
\mathbb{P}\left(\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}-\frac{\lambda^2}{2}\right|>\varepsilon\right) &\leq \mathbb{P}\left(\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}-\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right|>\varepsilon\right)+\\
&+\mathbb{P}\left(\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}-\mathbb{E}\left[\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right]\right|>\varepsilon\right)+\\
&+\mathbb{P}\left(\left|\mathbb{E}\left[\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right]-\frac{\lambda^2}{2}\right|>\varepsilon\right)\\
&\overset{\text{Markov}}{\leq} \varepsilon^{-1}\mathbb{E}\left[\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}-\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right|\right]+\\
+\mathbb{P}\left(\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}-\mathbb{E}\left[\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right]\right|>\varepsilon\right)&+\varepsilon^{-1}\left|\mathbb{E}\left[\sum_{k=0}^{M-1}\frac{U_k^2}{2}\mathbb{1}_{U_k^2\leq\frac{\alpha}{M}}\right]-\frac{\lambda^2}{2}\right|\\
&\leq \frac{3c_{\hat{\beta}}}{\varepsilon\alpha^{\frac{\varepsilon_0}{2}}}+\frac{2\alpha^2}{\varepsilon M},
\end{aligned}$$

where in the last inequality we used (2.32), (2.36), (2.33) and (2.35). Indeed, we obtain (2.31) since $\alpha > 0$ is arbitrary. □

## 2.5 Proof of the Central Limit Theorem

We are now ready to prove Theorem 1.2.4 (i). Our estimates we have collected will enable us to control

$$\left|\mathbb{E}\left[\varphi\left(W_N\right)\right]-\mathbb{E}\left[\varphi\left(e^X\right)\right]\right|, \tag{2.37}$$

where $X \overset{\mathcal{L}}{\sim} \mathcal{N}\left(-\frac{\lambda^2}{2},\lambda^2\right)$ and $\varphi$ is a smooth function with bounded support, thus we aim to prove that (2.37) tends to 0 as $N$ tends to infinity. This would show vague convergence of $W_N$ to the exponential of a random variable distributed like a Gaussian $\mathcal{N}\left(-\frac{\lambda^2}{2},\lambda^2\right)$. However, as $W_N$ is positive and $\mathbb{E}[W_N] = 1$, we can deduce tightness of $(W_N)_N$ and finally infer the desired convergence in law. Let's proceed.

Let $M, \varepsilon > 0$ and $\varphi \in \mathcal{C}_c^\infty$. Let $X \overset{\mathcal{L}}{\sim} \mathcal{N}\left(-\frac{\lambda^2}{2},\lambda^2\right)$ be a gaussian random variable. Then, by the triangle inequality, we have

$$\begin{aligned}
\left|\mathbb{E}\left[\varphi\left(W_N\right)\right]-\mathbb{E}\left[\varphi\left(e^X\right)\right]\right| &\leq \left|\mathbb{E}\left[\varphi\left(W_N\right)\right]-\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1}Z_k\right)\right]\right|\\
&+\left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1}Z_k\right)\right]-\mathbb{E}\left[\varphi\left(e^X\right)\right]\right|.
\end{aligned} \tag{2.38}$$

Now, the first term in the right-hand side of (2.38) satisfies

$$\begin{aligned}\limsup_{N\to\infty}\left|\mathbb{E}\left[\varphi\left(W_N\right)-\varphi\left(\prod_{k=0}^{M-1}Z_k\right)\right]\right| &\leq \limsup_{N\to\infty}\|\varphi\|_{\mathrm{Lip}}\mathbb{E}\left[\left|W_N-\prod_{k=0}^{M-1}Z_k\right|\right]\\ &\overset{(2.4)}{\leq} \|\varphi\|_{\mathrm{Lip}}\limsup_{N\to\infty}\mathbb{E}\left[\left|W_N-\prod_{k=0}^{M-1}Z_k\right|^2\right]^{\frac{1}{2}}=0, \end{aligned}\tag{2.39}$$

by Proposition 2.2.1. To deal with the second term we make the following remarks.

**Remark 2.5.1.** *Let $\varphi\in\mathcal{C}_c^\infty(\mathbb{R})$. Then $\psi:=\varphi\circ\exp$ is $\mathcal{C}_c^\infty(\mathbb{R}_+)$ and Lipschitz continuous on all $\mathbb{R}$. Let's be more precise. It is clear that for $x\to\infty$ the function $\psi$ remains smooth with compact support, but we cannot state the same for $x\to-\infty$. However, considering $x,y\leq 0$, we have*

$$|\psi(x)-\psi(y)|=|\varphi(e^x)-\varphi(e^y)|\leq\|\varphi\|_{Lip}|e^x-e^y|\overset{x,y\leq 0}{\leq}\|\varphi\|_{Lip}|x-y|,\tag{2.40}$$

*so we recover that $\psi$ is lipschitz for negative values of $x$. At this point, as we knew that $\psi\in\mathcal{C}_c^\infty(\mathbb{R}_+)$, we can deduce lipschitzianity on all the domain $\mathbb{R}$.*

**Remark 2.5.2.** *Let $\psi$ as defined in Remark 2.5.1. Then $\phi:=\psi\circ(\cdot-\frac{\lambda^2}{2})$ is $\mathcal{C}^\infty(\mathbb{R}_+)$ and Lipschitz continuous on all $\mathbb{R}$. Indeed, the same argument detailed in Remark 2.5.1 holds in this case and in fact we have*

$$|\phi(x)-\phi(y)|=\left|\psi\left(x-\frac{\lambda^2}{2}\right)-\psi\left(y-\frac{\lambda^2}{2}\right)\right|\leq\|\psi\|_{Lip}|x-y|.\tag{2.41}$$

We then decompose the second term in the right-hand side of (2.38) as

$$\begin{aligned}\left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1}Z_k\right)-\varphi\left(e^X\right)\right]\right| &\leq \left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1}Z_k\right)-\varphi\left(\prod_{k=0}^{M-1}e^{U_k-\frac{U_k^2}{2}}\right)\right]\right|+\\ &+\left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1}e^{U_k-\frac{U_k^2}{2}}\right)-\varphi\left(e^X\right)\right]\right|\end{aligned}\tag{2.42}$$

so that by Remark 2.5.1, together with Proposition 2.3.3, Lemma 2.4.2 and Lemma 2.4.1, we can show that

$$\limsup_{M\to\infty}\limsup_{N\to\infty}\left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1}Z_k\right)-\varphi\left(e^X\right)\right]\right|\leq 2\varepsilon\|\psi\|_{\mathrm{Lip}}.\tag{2.43}$$

Indeed, setting

$$\Psi_1:=\psi\left(\sum_{k=0}^{M-1}\log Z_k\right)-\psi\left(\sum_{k=0}^{M-1}U_k-\frac{U_k^2}{2}\right)$$

$$\Psi_2:=\psi\left(\sum_{k=0}^{M-1}U_k-\frac{U_k^2}{2}\right)-\psi\left(\left(\sum_{k=0}^{M-1}U_k\right)-\frac{\lambda^2}{2}\right)$$

we make the following estimates:

$$
\begin{aligned}
\left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1} Z_k\right)-\varphi\left(\prod_{k=0}^{M-1} e^{U_k-\frac{U_k^2}{2}}\right)\right]\right| &\leq \mathbb{E}\left[\left|\psi\left(\sum_{k=0}^{M-1}\log Z_k\right)-\psi\left(\sum_{k=0}^{M-1} U_k-\frac{U_k^2}{2}\right)\right|\right] \\
=\mathbb{E}\left[|\Psi_1|\,\mathbb{1}_{\left|\sum_{k=0}^{M-1}\log Z_k-\sum_{k=0}^{M-1}U_k-\frac{U_k^2}{2}\right|\geq\varepsilon}\right] &+\mathbb{E}\left[|\Psi_1|\,\mathbb{1}_{\left|\sum_{k=0}^{M-1}\log Z_k-\sum_{k=0}^{M-1}U_k-\frac{U_k^2}{2}\right|<\varepsilon}\right] \\
\overset{(2.40)}{\leq} 2\|\psi\|_\infty\mathbb{P}\left(\left|\sum_{k=0}^{M-1}\log Z_k-\sum_{k=0}^{M-1}U_k-\frac{U_k^2}{2}\right|\geq\varepsilon\right) &+\varepsilon\|\psi\|_{\text{Lip}} \\
\overset{(2.26)}{\leq} 2\|\psi\|_\infty\left(\frac{2c_{\hat\beta}}{\varepsilon^{2+\varepsilon_0}M^{\frac{\varepsilon_0}{2}}}\right) &+\varepsilon\|\psi\|_{\text{Lip}},
\end{aligned}
\tag{2.44}
$$

where the last equality holds for $N \geq N_0(\hat\beta, M)$ large enough. Taking $N\to\infty$ and then $M\to\infty$ gives one $\varepsilon\|\psi\|_{\text{Lip}}$ term in (2.43). Turning to the second term in the right-hand side of (2.42),

$$
\begin{aligned}
\left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1} e^{U_k-\frac{U_k^2}{2}}\right)-\varphi\left(e^X\right)\right]\right| &= \left|\mathbb{E}\left[\psi\left(\sum_{k=0}^{M-1} U_k-\frac{U_k^2}{2}\right)-\psi\left(X\right)\right]\right| \\
&\leq |\mathbb{E}\left[\Psi_2(U_k)\right]|+\left|\mathbb{E}\left[\psi\left(\left(\sum_{k=0}^{M-1}U_k\right)-\frac{\lambda^2}{2}\right)-\psi\left(X\right)\right]\right|,
\end{aligned}
$$

so we find that

$$
\limsup_{M\to\infty}\limsup_{N\to\infty}\left|\mathbb{E}\left[\varphi\left(\prod_{k=0}^{M-1} e^{U_k-\frac{U_k^2}{2}}\right)-\varphi\left(e^X\right)\right]\right| \leq \varepsilon\|\psi\|_{\text{Lip}}, \tag{2.45}
$$

giving the second $\varepsilon\|\psi\|_{\text{Lip}}$ term in (2.43), where (2.45) follows from

$$
\begin{aligned}
|\mathbb{E}\left[\Psi_2(U_k)\right]| &= \mathbb{E}\left[|\Psi_2(U_k)|\,\mathbb{1}_{\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}-\frac{\lambda^2}{2}\right|>\varepsilon}\right]+\mathbb{E}\left[|\Psi_2(U_k)|\,\mathbb{1}_{\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}-\frac{\lambda^2}{2}\right|\leq\varepsilon}\right] \\
&\overset{(2.40)}{\leq} 2\|\psi\|_\infty\mathbb{P}\left(\left|\sum_{k=0}^{M-1}\frac{U_k^2}{2}-\frac{\lambda^2}{2}\right|>\varepsilon\right)+\varepsilon\|\psi\|_{\text{Lip}},
\end{aligned}
$$

and

$$
\begin{aligned}
\left|\mathbb{E}\left[\psi\left(\left(\sum_{k=0}^{M-1}U_k\right)-\frac{\lambda^2}{2}\right)-\psi\left(X\right)\right]\right| &= \left|\mathbb{E}\left[\phi\left(\sum_{k=0}^{M-1}U_k\right)-\phi\left(X+\frac{\lambda^2}{2}\right)\right]\right| \\
&\overset{(3)}{\leq} \|\phi\|_{\text{Lip}}\mathcal{W}_1\left(\mathcal{L}\left(\sum_{k=0}^{M-1}U_k\right),\mathcal{N}\left(0,\lambda^2\right)\right),
\end{aligned}
$$

together with (2.31) and (2.30).We emphasize that the latter computation is derived thanks to the Kantorovich-Rubenstein duality theorem recalled in (3).

Finally, putting together (2.39), (2.43) and taking, in order, $N \to \infty$, $M \to \infty$ and $\varepsilon \to 0$ we obtain that for any $\varphi \in \mathcal{C}_c^\infty$

$$\left|\mathbb{E}\left[\varphi\left(W_N\right) - \varphi\left(e^X\right)\right]\right| \xrightarrow[N\to\infty]{} 0,$$

which proves *vague convergence* of (the law of) $W_N$ to (the law of) $e^X$, where $X \sim \mathcal{N}\left(-\frac{\lambda^2}{2}, \lambda^2\right)$. However, in our setting, proving vague convergence is actually equivalent to proving convergence in law as the sequence $(W_N)_N$ is tight. In fact, for every $N \geq 1$ and for every $\varepsilon > 0$, considering $M > 1/\varepsilon$, we have that

$$\mathbb{P}(|W_N| > M) \overset{\text{Markov}}{\leq} \frac{\mathbb{E}[W_N]}{M} < \varepsilon \mathbb{E}[W_N] = \varepsilon,$$

which gives tightness of $(W_N)_N$. Therefore, we have proved the desired statement:

$$W_N \xrightarrow{\mathcal{L}} e^{\mathcal{N}\left(-\frac{\lambda^2}{2},\lambda^2\right)}.$$

# Appendix

In this appendix, we will recall some general facts and definitions that we used throughout the development of the proof of Theorem 2.1.1 (i). First of all, we recall the definition of Wasserstein distance ([8]), which is used in Lemma 2.4.1.

**Definition** (Wasserstein distances)**.** Let $(\mathcal{X}, d)$ be a Polish metric space and let $p \in [1, \infty)$. For any two probability measures $\mu$ and $\nu$ on $\mathcal{X}$, the Wasserstein distance of order $p$ between $\mu$ and $\nu$ is defined by the formula

$$\mathcal{W}_p(\mu, \nu) := \inf \left\{ \mathbb{E}\left[d(X,Y)^p\right]^{\frac{1}{p}}, \quad \mathcal{L}(X) = \mu \text{ and } \mathcal{L}(Y) = \nu \right\}, \tag{1}$$

where by $\mathbb{E}[\cdot]$ we are referring to an expectation with respect to a coupling of $X$ and $Y$.

**Remark.** Notice that, in our setting, we are working with $p = 1$ and with $d(x, y) = |x - y|$, which gives the formula

$$\mathcal{W}_1(\mu, \nu) := \inf \left\{ \mathbb{E}\left[|X - Y|\right], \quad \mathcal{L}(X) = \mu \text{ and } \mathcal{L}(Y) = \nu \right\}. \tag{2}$$

Recall that there also exists a duality formula for the 1-Wasserstein distance:

**Proposition** (Kantorovich-Rubenstein duality theorem)**.** *When $\mu$ and $\nu$ have bounded support, we have that*

$$\mathcal{W}_1(\mu, \nu) = \sup \left\{ \mathbb{E}_\mu\left[f(X)\right] - \mathbb{E}_\nu\left[f(Y)\right], \ f \textit{ continuous and } \|f\|_{Lip} \leq 1 \right\}. \tag{3}$$

# Bibliography


[1] F. Caravenna, R. Sun and N. Zygouras, *Universality in Marginally Relevant Disordered Systems*, The Annals of Applied Probability, 2017.

[2] F. Caravenna, R. Sun and N. Zygouras, *The Two-Dimensional KPZ Equation in the Entire Subcritical Regime*, The Annals of Probability, 2020.

[3] C. Cosco and Ofer Zeitouni, *Moments of Partition Functions of 2D Gaussian Polymers in the Weak Disorder Regime*, to appear in Communications in Mathematical Physics.

[4] E. Rio, *Upper Bounds for Minimal Distances in the Central Limit Theorem*, Annales de l'Institut Henri Poincaré-Probabilités et Statistiques, 2009.

[5] F. Caravenna and F. Cottini, *Gaussian Limits for Subcritical Chaos*, Electronic Journal of Probability, 2022.

[6] G. F. Lawler, *Random Walk: A Modern Introduction*, Cambridge University Press, 2010.

[7] O. Zeitouni, *Branching random walks and Gaussian Fields*, Notes for Lectures, 2012.

[8] C. Villani, *Optimal Transport*, Springer Berlin, Heidelberg, 2008.

[9] T. Alberts, K. Khanin and J. Quastel, *The intermidiate disorder regime for directed polymers in dimension 1+1*, The Annals of Probability, 2014.

[10] D. A. Huse and C. L. Henley, *Pinning and roughening of domain walls in ising systems due to random impurities*, Phys. Rev. Lett. , 1985.

[11] J. Z. Imbrie and T. Spencer, *Diffusion of directed polymers in a random environment*, J. Stat. Phys. , 1988.

[12] F. Comets and N. Yoshida, *Directed polymers in random environment are diffusive at weak disorder*, The Annals of Probability, 2006.

[13] F. Comets, *Directed polymers in random environment*, École d'Été de Probilités de Saint-Flour XLVI-2016, Springer, 2017.

[14] L. Bertini and N. Cancrini, *The two-dimensional stochastic heat equation: renormalizing a multiplicative noise*, Journal of Physics A: Mathematical and General, 31(2): 615-622, Jan. 1998.