Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2023, Vol. 59, No. 2, 774–794
https://doi.org/10.1214/22-AIHP1287

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES

Weak convergence of directed polymers to deterministic KPZ at high temperature

Sourav Chatterjee[a]

Departments of Mathematics and Statistics, Stanford University, Stanford, USA, [a]*souravc@stanford.edu*

Received 22 July 2021; revised 13 May 2022; accepted 30 May 2022

Abstract. It is shown that when $d \geq 3$, the growing random surface generated by the $(d+1)$-dimensional directed polymer model at sufficiently high temperature, after being smoothed by taking microscopic local averages, converges to a solution of the deterministic KPZ equation in a suitable scaling limit.

Résumé. On montre que quand $d \geq 3$, la surface aléatoire croissante engendrée par le modèle de polymère dirigé $(d+1)$-dimensionnel à une température suffisamment haute, après avoir été lissée en prenant des moyennes locales microscopiques, converge vers une solution de l'équation de KPZ déterministe dans une limite d'échelle appropriée.

MSC2020 subject classifications: 82C41; 60G60; 39A12

Keywords: Directed polymer; KPZ; Random surface; Scaling limit

1. The directed polymer model

Take any $d \geq 1$. Let β be a positive real number, let $\xi = (\xi_{t,x})_{t\in\mathbb{Z}_{\geq 1}, x\in\mathbb{Z}^d}$ be a collection of i.i.d. random variables and let $g : \mathbb{Z}^d \to \mathbb{R}$ be a function. The $(d+1)$-dimensional directed polymer model at inverse temperature β, in the random environment ξ, and initial condition g, defines a growing random surface $f : \mathbb{Z}_{\geq 0} \times \mathbb{Z}^d \to \mathbb{R}$ as follows. For any t and x, let $\mathcal{Q}(t,x)$ be the set of all nearest-neighbor paths in $\mathbb{Z}^d$ of length t that end at x. A typical element $q \in \mathcal{Q}(t,x)$ is a $(t+1)$-tuple $(q_0, q_1, \ldots, q_t)$ of elements of $\mathbb{Z}^d$, where $q_t = x$, and $|q_i - q_{i+1}| = 1$ for each i, where $|\cdot|$ denotes Euclidean norm. Let $f(0,x) = g(x)$ for each $x \in \mathbb{Z}^d$, and for $t \in \mathbb{Z}_{>0}$ and $x \in \mathbb{Z}^d$, let

$$f(t,x) = \frac{1}{\beta}\log\biggl[\sum_{q\in\mathcal{Q}(t,x)} \exp\biggl(\beta g(q_0) + \beta\sum_{i=1}^{t}\xi_{i,q_i}\biggr)\biggr]. \tag{1.1}$$

The directed polymer model has a long and storied history in probability and statistical physics. For recent surveys of this vast literature, see [1,9]. The model is supposed to converge to a scaling limit given by the Kardar–Parisi–Zhang (KPZ) equation [16]. Formally, the KPZ equation is given by

$$\partial_t h = \nu\Delta h + \frac{\lambda}{2}|\nabla h|^2 + \sqrt{D}\dot{W}, \tag{1.2}$$

where $h(t,x)$ is the height of a continuum random surface at time t and location x, $\dot{W}$ is a random field known as space-time white noise, and ν, λ, and D are nonnegative real-valued parameters. In dimension one, the meaning of the KPZ equation, as well as the convergence of the directed polymer model to a KPZ limit, are now well-understood – see [6] for a brief survey.

For $d \geq 3$, it has been shown recently in [18] that when β is sufficiently small, the discrete surface defined by (1.1) converges to a Gaussian field upon suitable centering and scaling. A similar result was proved for $d = 2$, when β decays as $t \to \infty$ like $(\log t)^{-1/2}$, in the earlier paper [2]. These Gaussian fields are the same as the ones coming from recent attempts at constructing solutions to the KPZ equation in $d \geq 2$ in the so-called *Edwards–Wilkinson regime* [3,8,10–14,

19,20]. In this regime, the gradient-squared term of the KPZ equation (1.2) does not appear (that is, $\lambda = 0$), which reduces the KPZ equation to the stochastic heat equation (SHE) with additive noise.

This paper proposes a different notion of KPZ convergence for directed polymers in $d \geq 3$ and sufficiently small β, where the limit is the *deterministic KPZ equation*

$$\partial_t h = \nu \Delta h + \frac{\lambda}{2}|\nabla h|^2, \tag{1.3}$$

which is just (1.2) with $D = 0$. The coefficients ν and λ are determined by β and d. The limit is obtained by smoothing the discrete surface (1.1) by locally averaging it over balls of small but growing radius. This notion of convergence to KPZ, by local smoothing of a discrete random surface, seems to be a new idea that has not appeared in the prior literature. One can say that this is convergence to KPZ 'on average', or just 'weak convergence'. It would be interesting to see if there are other models of surface growth that exhibit similar behavior (for example, those discussed in [7]).

2. Results

Fix $d \geq 3$, as before. Let $\xi = (\xi_{t,x})_{t \in \mathbb{Z}_{\geq 1}, x \in \mathbb{Z}^d}$ be a field of i.i.d. random variables, which we will refer to as the 'noise variables'. Let $g : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function. For each $\varepsilon > 0$, define $g_\varepsilon : \mathbb{Z}^d \to \mathbb{R}$ as $g_\varepsilon(x) := g(\varepsilon x)$. Let f_ε be the growing random surface generated by the directed polymer model with initial data $f_\varepsilon(0, \cdot) = g_\varepsilon(\cdot)$. Let $\mathcal{L}$ denote the set of all probability measures on $\mathbb{R}$ that are push-forwards of the standard Gaussian measure under some Lipschitz map. We assume that the law of the noise variables belongs to the class $\mathcal{L}$. Then, in particular, $\xi_{t,x}$ has a finite moment generating function $m(\beta) := \mathbb{E}(e^{\beta \xi_{t,x}})$. Define

$$\mu(\beta) := \frac{m(2\beta)}{m(\beta)^2}.$$

Let ρ_d be the probability that a simple symmetric random walk started at the origin in $\mathbb{Z}^d$ returns to the origin at least once. Since $d \geq 3$, $\rho_d < 1$. In particular, $\mu(0) = 1 < 1/\rho_d$. Let β_0 be the supremum of all nonnegative β such that $\mu(\beta) < 1/\rho_d$. It is easy to see that μ is a continuous function, which implies that $\beta_0 > 0$. The range $(0, \beta_0)$ is known as the "L^2 regime" of sufficiently high temperatures. For a justification of this nomenclature, as well as the importance of this regime and its various occurrences in the literature, see [18].

To state the main result about the scaling limit of f_ε, we first need to identify a constant related to the model, whose existence at high temperature is established by the following proposition.

Proposition 2.1. *When $\beta \in (0, \beta_0)$, the following limit exists and is finite*:

$$\eta(\beta) := \lim_{t \to \infty} \mathbb{E}\left[\log\left\{\frac{1}{(2dm(\beta))^t} \sum_{q \in \mathcal{Q}(t,0)} \exp\left(\beta \sum_{i=1}^{t} \xi_{i,q_i}\right)\right\}\right].$$

The quantity $\eta(\beta)$ is the 'second order term' in the asymptotic expression for the log partition function at high temperature as $t \to \infty$, the first order term being $t \log(2dm(\beta))$. It turns out that a multiple of this second order term has to be subtracted off when renormalizing the surface to obtain the deterministic KPZ limit below. It is not immediately clear if $\eta(\beta)$ admits an explicit calculation.

For $x \in \mathbb{Z}^d$ and $r > 0$, let $B(x, r)$ denote the set of all $y \in \mathbb{Z}^d$ such that $|x - y| \leq r$. Let $\{r_\varepsilon\}_{\varepsilon > 0}$ be any collection of positive real numbers tending to infinity as $\varepsilon \to 0$, such that $r_\varepsilon = o(\varepsilon^{-1})$. For $t \in \mathbb{R}$, let $[t]$ denote the greatest integer $\leq t$. For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, let $[x]$ denote the vector $([x_1], \ldots, [x_d])$. For $t \in \mathbb{R}_{\geq 0}$, $x \in \mathbb{R}^d$, and $\varepsilon > 0$, define the rescaled values

$$t_\varepsilon := \lfloor \varepsilon^{-2} t \rfloor, \qquad x_\varepsilon := \lfloor \varepsilon^{-1} x \rfloor. \tag{2.1}$$

Define the smoothed, rescaled and renormalized surface $\widetilde{f}^{(\varepsilon)} : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}$ as

$$\widetilde{f}^{(\varepsilon)}(t, x) := \frac{1}{|B(x_\varepsilon, r_\varepsilon)|} \sum_{y \in B(x_\varepsilon, r_\varepsilon)} f_\varepsilon(t_\varepsilon, y) - \frac{t_\varepsilon}{\beta} \log\bigl(2dm(\beta)\bigr) - \frac{\eta(\beta)}{\beta}. \tag{2.2}$$

The following theorem is the main result of this paper.

Theorem 2.2. *Let all notations and conditions be as above, with $d \geq 3$. Suppose that $\beta \in (0, \beta_0)$. Let h be the Cole–Hopf solution of the deterministic KPZ equation (1.3) with $\nu = 1/2d$, $\lambda = \beta/d$, and initial condition g. This is defined as*

$$h(t,x) = \frac{1}{\beta}\log \mathbb{E}\big[\exp\big(\beta g(\sqrt{t/d}\,Z + x)\big)\big],$$

where Z is a d-dimensional standard Gaussian random vector. Then for any $t > 0$ and $x \in \mathbb{R}^d$, the random variable $\widetilde{f}^{(\varepsilon)}(t,x)$ converges in L^2 to $h(t,x)$ as $\varepsilon \to 0$.

Theorem 2.2 proves convergence of $\widetilde{f}^{(\varepsilon)}(t,x)$ to the deterministic KPZ limit $h(t,x)$ for each fixed (t,x). This distinguishes it from much of the recent literature, which deal with convergence towards Edwards–Wilkinson limits on average – that is, after integrating the field against some test function [3,12–14,18]. From this perspective, the pointwise statement of Theorem 2.2 has some resemblance to the results of [10] and [11] which give pointwise laws of large numbers and central limit theorems for the difference between a solution of the KPZ equation with given initial data and a stationary solution of the KPZ equation.

The main idea behind the proof of Theorem 2.2 is the following. We show that $f_\varepsilon(t_\varepsilon, y)$ splits as the sum of a random quantity that has no dependence on the initial condition g, and an almost deterministic (i.e., highly concentrated) quantity that approximately equals $h(t,x)$ with high probability. The random part, moreover, is a function only of noise variables located very close to (t_ε, y), so that averaging the random parts over $y \in B(x_\varepsilon, r_\varepsilon)$ results in an approximately deterministic value that has no dependence on g. The renormalization term in (2.2) is this deterministic value.

One can also consider the rescaled and renormalized surface without smoothing, defined as

$$f^{(\varepsilon)}(t,x) := f_\varepsilon(t_\varepsilon, x_\varepsilon) - \frac{t_\varepsilon}{\beta}\log\big(2dm(\beta)\big) - \frac{\eta(\beta)}{\beta}.$$

It is a simple consequence of Theorem 2.2 that $f^{(\varepsilon)}$ converges to h in the a certain weak sense as $\varepsilon \to 0$.

Corollary 2.3. *Take any continuous function $\phi : \mathbb{R}^d \to \mathbb{R}$ with compact support. Then for any $t > 0$, as $\varepsilon \to 0$, the random variable $\int f^{(\varepsilon)}(t,x)\phi(x)\,dx$ converges in L^2 to $\int h(t,x)\phi(x)\,dx$.*

Results analogous to Corollary 2.3 have been proved for a model of continuous directed polymers in [12,20]. The weak limit in that model turns out to be the heat equation rather than the deterministic KPZ equation, because limit is obtained for the partition function rather than the log partition function. It is probable that the weak limit of the log partition function in the continuous model is also deterministic KPZ. It would be interesting to see if some analogue of Theorem 2.2 can be proved in the continuous setting. A natural open problem is to extend the above results to the entire high temperature regime $(0, \beta_c)$, where β_c is the smallest critical inverse temperature of the model (see [1] for the precise definition; it is known that $\beta_c \geq \beta_0$). It is possible that Theorem 2.2 and Corollary 2.3 are valid for all $\beta \in (0, \beta_c)$, especially in view of [1, Theorem 1.1], which shows that the endpoint distribution of the polymer is not localized when $\beta \leq \beta_c$. It is possible that delocalization of the endpoint is sufficient for the proof of Theorem 2.2 instead of the stronger condition $\beta \leq \beta_0$. This possibility is bolstered by the fact that a result similar to Corollary 2.3 for the partition function (instead of the log partition function) in the entire high temperature regime in a related continuous model has recently been established in [12].

Another interesting direction is to extend Theorem 2.2 to dimension two. It is unclear what the appropriate analogue of Theorem 2.2 would be in $d = 2$.

Investigating the behavior of second order fluctuations is a natural question. For example, in Theorem 2.2, one can ask if for some explicit $\gamma > 0$, $\varepsilon^{-\gamma}(\widetilde{f}^{(\varepsilon)}(t,x) - h(t,x))$ converges in law to a non-degenerate limiting distribution as $\varepsilon \to 0$. It is unclear whether this might hold for fixed (t,x), or after averaging against a smooth test function. A similar question can be asked in the context of Corollary 2.3, about the rescaled difference $\varepsilon^{-\gamma}\int(f^{(\varepsilon)}(t,x) - h(t,x))\phi(x)\,dx$. For this, the results of [12,13,18,20] indicate that the exponent γ could be $(d-2)/2$, because these papers show that $\varepsilon^{-(d-2)/2}\int(f^{(\varepsilon)}(t,x) - \mathbb{E}(f^{(\varepsilon)}(t,x)))\phi(x)\,dx$ has a non-degenerate limiting distribution in a number of related settings.

Finally, as pointed out by one of the referees (and in analogy with the results of [12,20]), it is likely that in Theorem 2.2 and Corollary 2.3, if one takes a suitable weak limit of the partition function instead of the log partition function, one would obtain the deterministic heat equation instead of the deterministic KPZ equation. It would be interesting to see if this is true, under a similar local averaging framework.

The rest of the paper is devoted to the proofs of the above results. Many technical ideas in the proofs are similar to ideas from various recent papers, such as [2,3,18]. The main new idea is the formulation of the scaling limit, presented above, that leads to the deterministic KPZ limit.

3. Concentration inequalities

We will use a technique from the classical theory of concentration of measure to control the lower tail probabilities of the partition functions of the polymer model. This technique was first used in [4]. For a recent application, see [3]. A complete account is given below for the reader's convenience.

First, we need the following classical result, known as the Gaussian concentration inequality (see [5, Appendix A]).

Theorem 3.1 (Gaussian concentration inequality). *Let X be an n-dimensional standard Gaussian random vector and let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function with Lipschitz constant L. Then for any $t \geq 0$,*

$$\mathbb{P}\big(|f(X) - \mathbb{E}(f(X))| \geq t\big) \leq 2e^{-t^2/2L^2}.$$

Let $\mathcal{L}(K)$ be the set of all probability measures on $\mathbb{R}$ that are obtained as push-forwards of the standard Gaussian measure under a Lipschitz map with Lipschitz constant $\leq K$. Let $\mathcal{L}_n(K)$ denote the set of the product measures on $\mathbb{R}^n$ whose marginal distributions are all in $\mathcal{L}(K)$. We will say that a random vector is in class $\mathcal{L}_n(K)$ if its law is in $\mathcal{L}_n(K)$. The following result is an easy corollary of the Gaussian concentration inequality.

Corollary 3.2. *Let X be an n-dimensional random vector in class $\mathcal{L}_n(K)$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function with Lipschitz constant L. Then for any $t \geq 0$,*

$$\mathbb{P}\big(|f(X) - \mathbb{E}(f(X))| \geq t\big) \leq 2e^{-t^2/2L^2K^2}.$$

Proof. If $X = (X_1, \ldots, X_n)$, we can write $X_i = g_i(Y_i)$ where $Y_1, \ldots, Y_n$ are i.i.d. standard Gaussian random variables and $g_1, \ldots, g_n$ are Lipschitz functions with Lipschitz constant K. Then $f(X)$, as a function of $Y = (Y_1, \ldots, Y_n)$, has Lipschitz constant $\leq LK$. The claimed inequality now follows from the Gaussian concentration inequality. □

A corollary of the above corollary is the following result, which is a standard idea from concentration of measure [17].

Corollary 3.3. *Let X be an n-dimensional random vector in class $\mathcal{L}_n(K)$. Let A be a closed subset of $\mathbb{R}^n$. Let $D := \inf\{\|X - x\| : x \in A\}$ and $p := \mathbb{P}(X \in A)$. Then for any $t \geq 0$,*

$$\mathbb{P}\big(D \geq K\sqrt{2\log(2/p)} + t\big) \leq 2e^{-t^2/2K^2}.$$

Proof. It is not hard to see that D is a Lipschitz function of X with Lipschitz constant 1. Therefore by Corollary 3.2,

$$\mathbb{P}\big(|D - \mathbb{E}(D)| \geq t\big) \leq 2e^{-t^2/2K^2}. \tag{3.1}$$

Since A is a closed set, the above inequality implies that

$$\begin{aligned} p &= \mathbb{P}(X \in A) = \mathbb{P}(D = 0) \\ &\leq \mathbb{P}\big(|D - \mathbb{E}(D)| \geq \mathbb{E}(D)\big) \leq 2e^{-\mathbb{E}(D)^2/2K^2}, \end{aligned}$$

which gives

$$\mathbb{E}(D) \leq K\sqrt{2\log(2/p)}.$$

Thus, again appealing to (3.1), we get the desired result. □

Using Corollary 3.3, we obtain the following variant of [17, Proposition 1.6].

Proposition 3.4. *Let X be an n-dimensional random vector in class $\mathcal{L}_n(K)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable convex function. Take any $a \in \mathbb{R}$ and $b > 0$, and let $p := \mathbb{P}(f(X) \geq a, |\nabla f(X)| \leq b)$. Then for any $t \geq 0$,*

$$\mathbb{P}\big(f(X) \leq a - Kb\sqrt{2\log(2/p)} - t\big) \leq 2e^{-t^2/2b^2K^2}.$$

Proof. By convexity,

$$f(y) - f(x) \le \nabla f(y) \cdot (y - x)$$

for any $x, y \in \mathbb{R}^n$. Thus, if $f(y) \ge a$ and $|\nabla f(y)| \le b$, we have that for any x,

$$f(x) \ge a - b|x - y|.$$

Let A be the set of all y such that $f(y) \ge a$ and $|\nabla f(y)| \le b$. Since f is continuously differentiable, A is a closed set. Let $p := \mathbb{P}(X \in A)$ and let D be the distance of X from A. Then by the above inequality,

$$f(X) \ge a - bD.$$

Thus, for any $t \ge 0$,

$$\begin{aligned}\mathbb{P}\bigl(f(X) \le a - Kb\sqrt{2\log(2/p)} - t\bigr) &\le \mathbb{P}\bigl(a - bD \le a - Kb\sqrt{2\log(2/p)} - t\bigr)\\ &= \mathbb{P}\bigl(D \ge K\sqrt{2\log(2/p)} + t/b\bigr).\end{aligned}$$

The claimed result now follows by Corollary 3.3. □

4. Random walk estimates

In this section, $C, C_1, C_2, \ldots$ will denote positive constants depending only on d, whose values may change from line to line. (Here d is the dimension fixed in Section 2. Recall that $d \ge 3$.) If a constant depends on some additional parameter θ, we will denote it by $C(\theta)$.

Let $\{S_n\}_{n\ge 0}$ be a simple symmetric random walk on $\mathbb{Z}^d$, starting at the origin. We will now collect several estimates for this walk that will be useful later. First, recall the well-known fact that for each n,

$$\sup_{x\in\mathbb{Z}^d} \mathbb{P}(S_n = x) \le \frac{C}{n^{d/2}}. \tag{4.1}$$

(The simplest way to show this is by Fourier transforms. For example, see [6, Lemma 18.3].) On the other hand, it is a simple consequence of Hoeffding's inequality [15] for sums of independent and uniformly bounded random variables that for any x and n,

$$\mathbb{P}(S_n = x) \le C_1 e^{-C_2|x|^2/n}. \tag{4.2}$$

Another easy consequence of Hoeffding's inequality that we will use later is that for any $\theta \ge 0$ and $n \ge 0$,

$$\mathbb{E}\bigl(e^{\theta|S_n|}\bigr) \le C_1 e^{C_2\theta^2 n}. \tag{4.3}$$

By Chebychev's inequality,

$$\mathbb{P}\bigl(|S_n| \le 2\sqrt{n}\bigr) \ge 1 - \frac{\mathbb{E}|S_n|^2}{4n} = \frac{3}{4}.$$

Thus, for any $\theta \ge 0$ and any n,

$$\mathbb{E}\bigl(e^{-\theta|S_n|}\bigr) \ge e^{-2\theta\sqrt{n}}\mathbb{P}\bigl(|S_n| \le 2\sqrt{n}\bigr) \ge \frac{3e^{-2\theta\sqrt{n}}}{4}. \tag{4.4}$$

Combining (4.1) and (4.2), and letting $n_0 := a(2 + |x|)^2/\log(2 + |x|)$ for a sufficiently small number a, we get

$$\begin{aligned}\mathbb{P}(S_n = x \text{ for some } n) &\le C_1 \sum_{n\le n_0} e^{-C_2|x|^2/n} + C_1 \sum_{n>n_0} n^{-d/2}\\ &\le C_1 n_0 e^{-C_2|x|^2/n_0} + C_1 n_0^{-(d-2)/2}\\ &\le C_1\bigl(\log(2 + |x|)\bigr)^{C_2}\bigl(2 + |x|\bigr)^{2-d}. \qquad (4.5)\end{aligned}$$

Recall that ρ_d is the probability that S_n returns to the origin at least once. Let $\{S'_n\}_{n\geq 0}$ be another d-dimensional simple symmetric random walks started at the origin, independent of $\{S_n\}_{n\geq 0}$. It is easy to see that $\{S_n - S'_n\}_{n\geq 0}$ is a Markov chain with the same law as the Markov chain $\{S_{2n}\}_{n\geq 0}$. Since a simple symmetric random walk can return to its starting point only at even times, this shows that if

$$N_n := \big|\big\{0 \leq k \leq n : S_k = S'_k\big\}\big|, \qquad N_\infty := \big|\big\{k \geq 0 : S_k = S'_k\big\}\big|, \tag{4.6}$$

then we have

$$\mathbb{P}(N_\infty = k) = \rho_d^{k-1}(1-\rho_d) \tag{4.7}$$

for each $k \geq 1$. In particular, $\mathbb{E}(e^{aN_\infty})$ is finite if $0 < a < \log(1/\rho_d)$.

Next, suppose that the two random walks are started from two different locations x and y. Let $N_n^{x,y}$ and $N_\infty^{x,y}$ be the analogues of N_n and N_∞ in this situation. Let $\tau^{x,y}$ be the smallest n such that $S_n = S'_n$. If there is no such n, let $\tau^{x,y} = \infty$. Let

$$\kappa(x,y) := \mathbb{P}\big(\tau^{x,y} < \infty\big). \tag{4.8}$$

Note that by (4.5), and the fact that $\{S_n - S'_n\}_{n\geq 0}$ behaves like $\{S_{2n}\}_{n\geq 0}$ started at $x-y$, we get

$$\begin{aligned}\kappa(x,y) &= \mathbb{P}(S_{2n} = 0 \text{ for some } n)\\ &\leq C_1\big(\log\big(2+|x-y|\big)\big)^{C_2}\big(2+|x-y|\big)^{2-d}.\end{aligned} \tag{4.9}$$

If $\tau^{x,y} = \infty$, then $N_\infty^{x,y} = 0$ and $N_n^{x,y} = 0$ for all n. On the other hand, if $\tau^{x,y}$ is finite, the strong Markov property of $\{S_n - S'_n\}_{n\geq 0}$ implies that $N_n^{x,y}$ has the same law as N_n and $N_\infty^{x,y}$ has the same law as N_∞. Thus, by (4.5), we get that for any function $f : \mathbb{Z}_{\geq 0} \to [0,\infty)$ with $f(0) = 0$,

$$\mathbb{E}\big(f\big(N_\infty^{x,y}\big)\big) = \mathbb{E}\big(f(N_\infty)\big)\kappa(x,y). \tag{4.10}$$

We will use the above facts on several occasions. We will also use the following lemma.

Lemma 4.1. *Let $\{S_n\}_{n\geq 0}$ and $\{S'_n\}_{n\geq 0}$ be independent simple symmetric random walks on $\mathbb{Z}^d$ starting at the origin. Take any $\alpha > 0$. Then for all n,*

$$\mathbb{E}\big[\kappa\big(S_n, S'_n\big)^\alpha\big] \leq C_1(\alpha)(\log n)^{C_2(\alpha)} n^{-(d-2)\alpha'/2},$$

where $\alpha' = \min\{\alpha, d/(d-2)\}$.

Proof. Since κ is uniformly bounded by 1, we have $\kappa(x,y)^\alpha \leq \kappa(x,y)^{\alpha'}$ for all x, y. So let us assume without loss of generality that $\alpha \leq d/(d-2)$. By the facts that $\kappa(x,y) \leq 1$ and $\kappa(x,y) = \kappa(0, x-y)$ for all x, y, and that $S_n - S'_n$ has the same law as S_{2n}, we get

$$\mathbb{E}\big[\kappa\big(S_n, S'_n\big)^\alpha\big] \leq \sum_{z:|z|\leq\sqrt{n}\log n} \kappa(0,z)^\alpha \mathbb{P}(S_{2n} = z) + \mathbb{P}\big(|S_{2n}| \geq \sqrt{n}\log n\big).$$

By Hoeffding's inequality,

$$\mathbb{P}\big(|S_{2n}| \geq \sqrt{n}\log n\big) \leq 2e^{-C(\log n)^2}.$$

On the other hand, by (4.1), $\mathbb{P}(S_{2n} = z) \leq Cn^{-d/2}$ for all z. Therefore, by (4.9),

$$\begin{aligned}&\sum_{z:|z|\leq\sqrt{n}\log n} \kappa(0,z)^\alpha \mathbb{P}(S_{2n} = z)\\ &\leq C_1(\alpha)(\log n)^{C_2(\alpha)} n^{-d/2} \sum_{z:|z|\leq\sqrt{n}\log n} \big(1+|z|\big)^{(2-d)\alpha}\\ &\leq C_1(\alpha)(\log n)^{C_2(\alpha)} n^{-d/2} \sum_{0\leq r\leq\sqrt{n}\log n} r^{d-1}(1+r)^{(2-d)\alpha}\\ &\leq C_1(\alpha)(\log n)^{C_2(\alpha)} n^{-(d-2)\alpha/2}.\end{aligned}$$

(The assumption that $\alpha \le d/(d-2)$ was used in the last step.) The proof is now completed by combining the three displays above. □

5. Discrete approximation of the Cole–Hopf solution

In this section, $C, C_1, C_2, \ldots$ will denote positive constants whose values may depend only on d, β and g. The values may change from line to line.

Let h be the Cole–Hopf solution to the deterministic KPZ equation, defined in the statement of Theorem 2.2. For a proof that this is indeed a solution, see [6]. We will need to work with a discrete approximation of h, defined as follows. For any $\varepsilon > 0$, define $G_\varepsilon : \mathbb{Z}_{\ge 0} \times \mathbb{Z}^d$ as

$$G_\varepsilon(t,x) := \mathbb{E}\big(e^{\beta g_\varepsilon(S_t + x)}\big),$$

where, as in Section 4, $\{S_n\}_{n\ge 0}$ denotes a simple symmetric random walk in $\mathbb{Z}^d$, started at the origin. Note that the definition of G_ε depends on β and g, but we are suppressing that for notational simplicity. The following lemmas about G_ε will be useful.

Lemma 5.1. *As $\varepsilon \to 0$, $\beta^{-1}\log G_\varepsilon(t_\varepsilon, x_\varepsilon) \to h(t,x)$ for any $(t,x) \in \mathbb{R}_{\ge 0}\times\mathbb{R}^d$.*

Proof. Note that

$$G_\varepsilon(t_\varepsilon, x_\varepsilon) = \mathbb{E}\big(e^{\beta g_\varepsilon(S_{t_\varepsilon} + x_\varepsilon)}\big) = \mathbb{E}\big(e^{\beta g(\varepsilon S_{t_\varepsilon} + \varepsilon x_\varepsilon)}\big).$$

By the multivariate central limit theorem, we know that as $\varepsilon \to 0$, $t_\varepsilon^{-1/2} S_{t_\varepsilon}$ converges in distribution to $d^{-1/2}Z$, where Z is a d-dimensional standard Gaussian random vector. Also, as $\varepsilon \to 0$, $\varepsilon x_\varepsilon \to x$ and $\varepsilon t_\varepsilon^{1/2} \to \sqrt{t}$. Thus, as $\varepsilon \to 0$, $\varepsilon S_{t_\varepsilon} + \varepsilon x$ converges in distribution to $\sqrt{t/d}\,Z + x$. Consequently, $e^{\beta g(\varepsilon S_{t_\varepsilon} + \varepsilon x_\varepsilon)}$ converges in distribution to $e^{\beta g(\sqrt{t/d}Z + x)}$.

Next, note that by the Lipschitz property of g, the inequality (4.3), and the facts that $\varepsilon|x_\varepsilon| \le C_1 + C_2|x|$ and $\varepsilon^2 t_\varepsilon \le C_1 + C_2 t$,

$$\begin{aligned}\mathbb{E}\big(e^{2\beta g_\varepsilon(S_{t_\varepsilon} + x_\varepsilon)}\big) &\le C_1 \mathbb{E}\big(e^{C_2\varepsilon(|S_{t_\varepsilon}| + |x_\varepsilon|)}\big)\\ &\le C_1 e^{C_2(\varepsilon^2 t_\varepsilon + \varepsilon|x_\varepsilon|)} \le C_3 e^{C_4(t+|x|)}.\end{aligned}$$

From this uniform boundedness of the second moment of $e^{\beta g(\varepsilon S_{t_\varepsilon} + \varepsilon x_\varepsilon)}$, and the distributional convergence proved in the previous paragraph, it is now easy to see that as $\varepsilon \to 0$,

$$G_\varepsilon(t_\varepsilon, x_\varepsilon) = \mathbb{E}\big(e^{\beta g(\varepsilon S_{t_\varepsilon} + \varepsilon x_\varepsilon)}\big) \to \mathbb{E}\big(e^{\beta g(\sqrt{t/d}Z + x)}\big),$$

which completes the proof of the lemma. □

Lemma 5.2. *For any t, x and y,*

$$\big|G_\varepsilon(t,x) - G_\varepsilon(t,y)\big| \le C_1\varepsilon|x-y|e^{C_2(\varepsilon|x| + \varepsilon|y| + \varepsilon^2 t)}.$$

Proof. By the Lipschitz property of g,

$$\begin{aligned}\big|G_\varepsilon(t,x) - G_\varepsilon(t,y)\big| &\le \mathbb{E}\big|e^{\beta g_\varepsilon(S_t + x)} - e^{\beta g_\varepsilon(S_t + y)}\big|\\ &\le C\varepsilon|x-y|\mathbb{E}\big(e^{\beta g_\varepsilon(S_t + x)} + e^{\beta g_\varepsilon(S_t + y)}\big)\\ &\le C_1\varepsilon|x-y|e^{C_2\varepsilon(|x| + |y|)}\mathbb{E}\big(e^{C_2\varepsilon|S_t|}\big).\end{aligned}$$

By (4.3), this completes the proof. □

Lemma 5.3. *For any s, t and x,*

$$\big|G_\varepsilon(s,x) - G_\varepsilon(t,x)\big| \le C_1\varepsilon|t-s|^{1/2}e^{C_2(\varepsilon|x| + \varepsilon^2 s + \varepsilon^2 t)}.$$

Proof. Note that by the Lipschitz property of g,

$$\begin{aligned}
&\bigl|G_\varepsilon(s,x) - G_\varepsilon(t,x)\bigr| \\
&\quad\le \mathbb{E}\bigl|e^{\beta g_\varepsilon(S_s+x)} - e^{\beta g_\varepsilon(S_t+x)}\bigr| \\
&\quad\le \beta\mathbb{E}\bigl[\bigl|g_\varepsilon(S_s+x) - g_\varepsilon(S_t+x)\bigr|\bigl(e^{\beta g_\varepsilon(S_s+x)} + e^{\beta g_\varepsilon(S_t+x)}\bigr)\bigr] \\
&\quad\le C\varepsilon\mathbb{E}\bigl[|S_s - S_t|\bigl(e^{\beta g_\varepsilon(S_s+x)} + e^{\beta g_\varepsilon(S_t+x)}\bigr)\bigr].
\end{aligned}$$

Now applying the Cauchy–Schwarz inequality, the fact that $\mathbb{E}[(S_s - S_t)^2] = |t-s|$, and the same technique as in the proof of Lemma 5.2 to obtain upper bounds on the expected values of the exponential terms, we get the desired result. □

Lemma 5.4. *We have*

$$C_1 e^{-C_2\varepsilon(|x|+\sqrt{t})} \le G_\varepsilon(t,x) \le C_3 e^{C_4(\varepsilon|x|+\varepsilon^2 t)}.$$

Proof. By the Lipschitz property of g,

$$\mathbb{E}\bigl(e^{-C_1 - C_2\varepsilon(|S_t|+|x|)}\bigr) \le G_\varepsilon(t,x) \le \mathbb{E}\bigl(e^{C_3 + C_4\varepsilon(|S_t|+|x|)}\bigr).$$

The proof is now easily completed by invoking (4.3) and (4.4). □

6. Concentration of the height function

Let σ be the law of the noise variables. In this section, $C, C_1, C_2, \ldots$ will denote positive constants depending only on d, β, σ and g, whose values may change from line to line. If a constant depends on some additional parameter θ, we will denote it by $C(\theta)$.

Fix some $\varepsilon > 0$. Recall the function $\mu(\beta) = m(2\beta)/m(\beta)^2$, where m is the moment generating function of the noise variables. Note that $\mu(\beta) \ge 1$ for any β. Take any $t \in \mathbb{Z}_{\ge 1}$ and $x \in \mathbb{Z}^d$. For $q \in \mathcal{Q}(t,x)$, let

$$Y_q(t,x) := \frac{\exp(\beta\sum_{i=1}^t \xi_{i,q_i})}{m(\beta)^t},$$

and let

$$Y(t,x) := \frac{1}{(2d)^t}\sum_{q\in\mathcal{Q}(t,x)} Y_q(t,x). \tag{6.1}$$

Also, define

$$Z_\varepsilon(t,x) := \frac{1}{(2d)^t}\sum_{q\in\mathcal{Q}(t,x)} e^{\beta g_\varepsilon(q_0)} Y_q(t,x).$$

Finally, let $F(t,x) := \log Y(t,x)$ and $F_\varepsilon(t,x) := \log Z_\varepsilon(t,x)$. In this section we will prove two lemmas that will show that F and F_ε have fluctuations of order 1, with exponentially decaying tails. These technical results will be used later in the proof.

Lemma 6.1. *If $\mu(\beta) < 1/\rho_d$, then for any $\theta > 0$, $\mathbb{E}(e^{-\theta F(t,x)}) \le C(\theta)$.*

Proof. Note that $\mathbb{E}(Y_q(t,x)) = 1$ for any t, x and q, and so $\mathbb{E}(Y(t,x)) = 1$. Next, note that for any t, x and y, and any $q \in \mathcal{Q}(t,x)$ and $r \in \mathcal{Q}(t,y)$,

$$\begin{aligned}
\mathbb{E}\bigl(Y_q(t,x)Y_r(t,y)\bigr) &= \frac{1}{m(\beta)^{2t}}\prod_{i=1}^t \mathbb{E}\bigl(e^{\beta(\xi_{i,q_i}+\xi_{i,r_i})}\bigr) \\
&= \mu(\beta)^{|q\cap r|},
\end{aligned} \tag{6.2}$$

where $q \cap r$ denotes the set of i such that $q_i = r_i$. Thus,

$$\begin{aligned}\mathbb{E}\big(Y(t,x)^2\big) &= \frac{1}{(2d)^{2t}} \sum_{q,r \in \mathcal{Q}(t,x)} \mathbb{E}\big(Y_q(t,x)Y_r(t,y)\big)\\ &= \frac{1}{(2d)^{2t}} \sum_{q,r \in \mathcal{Q}(t,x)} \mu(\beta)^{|q\cap r|}\\ &= \mathbb{E}\big(\mu(\beta)^{N_t}\big),\end{aligned}$$

where N_t is defined as in (4.6). By (4.7), the assumption that $\mu(\beta) < 1/\rho_d$, and the observation that $\mu(\beta) \ge 1$, this shows that

$$\mathbb{E}\big(Y(t,x)^2\big) \le \mathbb{E}\big(\mu(\beta)^{N_\infty}\big) \le C. \tag{6.3}$$

Also, as noted above, $\mathbb{E}(Y(t,x)) = 1$. Thus, by (6.3) and the Paley–Zygmund second moment inequality,

$$\mathbb{P}\big(Y(t,x) \ge 1/2\big) \ge C > 0. \tag{6.4}$$

Next, note that for any $1 \le s \le t$ and $y \in \mathbb{Z}^d$,

$$\begin{aligned}\frac{\partial}{\partial \xi_{s,y}} F(t,x) &= \frac{1}{Y(t,x)} \frac{\partial}{\partial \xi_{s,y}} Y(t,x)\\ &= \frac{1}{Y(t,x)(2d)^t} \sum_{q \in \mathcal{Q}(t,x)} \frac{\partial}{\partial \xi_{s,y}} Y_q(t,x)\\ &= \frac{\beta}{Y(t,x)(2d)^t} \sum_{q \in \mathcal{Q}(t,x), q(s)=y} Y_q(t,x).\end{aligned}$$

Let ∇F denote the gradient of F, considering F as a function of $(\xi_{s,y})_{1\le s\le t, y\in\mathbb{Z}^d}$. The above calculation shows that

$$\begin{aligned}\big|\nabla F(t,x)\big|^2 &= \frac{\beta^2}{Y(t,x)^2(2d)^{2t}} \sum_{1\le s\le t} \sum_{y\in\mathbb{Z}^d} \sum_{\substack{q,r\in\mathcal{Q}(t,x),\\ q(s)=r(s)=y}} Y_q(t,x)Y_r(t,x)\\ &= \frac{\beta^2}{Y(t,x)^2(2d)^{2t}} \sum_{q,r\in\mathcal{Q}(t,x)} |q\cap r| Y_q(t,x)Y_r(t,x).\end{aligned}$$

Thus, by (6.2), the assumption that $\mu(\beta) < 1/\rho_d$, and the fact that $\mu(\beta) \ge 1$, we get

$$\begin{aligned}\mathbb{E}\big|\nabla Y(t,x)\big|^2 &= \mathbb{E}\big|Y(t,x)\nabla F(t,x)\big|^2\\ &= \beta^2 \mathbb{E}\big(N_t \mu(\beta)^{N_t}\big) \le \beta^2 \mathbb{E}\big(N_\infty \mu(\beta)^{N_\infty}\big) \le C.\end{aligned}$$

Combining this with (6.4), we see that there are constants $a \in \mathbb{R}$ and $b > 0$, depending only on d, β and σ, such that

$$\begin{aligned}&\mathbb{P}\big(F(t,x) \ge a, \big|\nabla F(t,x)\big| \le b\big)\\ &\ge \mathbb{P}\big(F(t,x) \ge a\big) - \mathbb{P}\big(F(t,x) \ge a, \big|\nabla F(t,x)\big| > b\big)\\ &= \mathbb{P}\big(F(t,x) \ge a\big) - \mathbb{P}\big(Y(t,x) \ge e^a, \big|\nabla Y(t,x)\big| > bY(t,x)\big)\\ &\ge \mathbb{P}\big(F(t,x) \ge a\big) - \mathbb{P}\big(\big|\nabla Y(t,x)\big| > be^a\big)\\ &\ge C > 0.\end{aligned}$$

It is a standard fact, easily verifiable by computing second derivatives, that $F(t,x)$ is a convex function of the noise variables. Thus, by Proposition 3.4 and the above inequality, we have

$$\mathbb{P}\big(F(t,x) \le -C_1 - u\big) \le 2e^{-C_2 u^2}. \tag{6.5}$$

This shows that for any $\theta > 0$,

$$\mathbb{E}\big(e^{-\theta F(t,x)}\big) = \int_0^\infty \mathbb{P}\big(e^{-\theta F(t,x)} \geq v\big)\, dv \leq \int_0^\infty \mathbb{P}\big(F(t,x) \leq -\theta^{-1}\log v\big)\, dv \leq C(\theta).$$

This completes the proof of the lemma. □

Lemma 6.2. *If $\mu(\beta) < 1/\rho_d$, then for any $\varepsilon \in (0,1)$ and $\theta > 0$,*

$$\mathbb{E}\big(e^{-\theta F_\varepsilon(t,x)}\big) \leq \psi\big(\varepsilon|x| + \varepsilon\sqrt{t} + \varepsilon^2 t\big),$$

where $\psi : (0,\infty) \to (0,\infty)$ is an increasing continuous function that depends only on β, d, σ, g and θ.

Proof. The proof is similar to the proof of Lemma 6.1, with minor adjustments. Throughout, we will denote by $\psi_1, \psi_2, \ldots$ increasing, positive, continuous functions of $\varepsilon|x| + \varepsilon\sqrt{t} + \varepsilon^2 t$, which depend only on β, d, σ, g, and θ.

First, note that by the Lipschitz property of g, we have $|g_\varepsilon(x) - g_\varepsilon(y)| \leq C\varepsilon|x-y|$ for all x, y, and $|g_\varepsilon(x)| \leq C_1 + C_2\varepsilon|x|$ for all x. Thus,

$$\begin{aligned}\mathbb{E}\big(Z_\varepsilon(t,x)^2\big) &= \frac{1}{(2d)^{2t}} \sum_{q,r\in\mathcal{Q}(t,x)} e^{\beta g_\varepsilon(q_0)+\beta g_\varepsilon(r_0)}\mathbb{E}\big(Y_q(t,x)Y_r(t,y)\big)\\ &\leq \frac{C_1 e^{C_2\varepsilon|x|}}{(2d)^{2t}} \sum_{q,r\in\mathcal{Q}(t,x)} e^{C_3\varepsilon(|q_0-x|+|r_0-x|)}\mu(\beta)^{|q\cap r|}\\ &= C_1 e^{C_2\varepsilon|x|}\mathbb{E}\big(e^{C_3\varepsilon(|S_t|+|S'_t|)}\mu(\beta)^{N_t}\big), \qquad (6.6)\end{aligned}$$

where $\{S_n\}_{n\geq 0}$ and $\{S'_n\}_{n\geq 0}$ are independent simple symmetric random walks started at the origin in $\mathbb{Z}^d$, and N_t is the number of times they intersect up to time t. Since $\mu(\beta) < 1/\rho_d$, we can choose $\gamma > 1$, depending only on β, σ, and d, such that $\mu(\beta)^\gamma < 1/\rho_d$. Let $\gamma' := 2\gamma/(\gamma-1)$, so that $2/\gamma' + 1/\gamma = 1$. Applying Hölder's inequality, and the displays (4.3) and (4.7), we get

$$\begin{aligned}&\mathbb{E}\big(e^{C_3\varepsilon(|S_t|+|S'_t|)}\mu(\beta)^{N_t}\big)\\ &\quad\leq \big[\mathbb{E}\big(e^{C_3\varepsilon\gamma'|S_t|}\big)\big]^{1/\gamma'}\big[\mathbb{E}\big(e^{C_3\varepsilon\gamma'|S'_t|}\big)\big]^{1/\gamma'}\big[\mathbb{E}\big(\mu(\beta)^{\gamma N_t}\big)\big]^{1/\gamma}\\ &\quad\leq C_4 e^{C_5\varepsilon^2 t}.\end{aligned}$$

Plugging this into (6.6), we get

$$\mathbb{E}\big(Z_\varepsilon(t,x)^2\big) \leq \psi_1. \qquad (6.7)$$

On the other hand, since $\mathbb{E}(Y_q(t,x)) = 1$ and $g_\varepsilon(q_0) - g_\varepsilon(x) \geq -C\varepsilon|q_0 - x|$ for any $q \in \mathcal{Q}(t,x)$, we have

$$\begin{aligned}\mathbb{E}\big(Z_\varepsilon(t,x)\big) &= \frac{1}{(2d)^t}\sum_{q\in\mathcal{Q}(t,x)} e^{\beta g_\varepsilon(q_0)}\\ &\geq C_1 e^{-C_2\varepsilon|x|}\mathbb{E}\big(e^{-C_3\varepsilon|S_t|}\big). \qquad (6.8)\end{aligned}$$

Using (4.4) in (6.8) gives the lower bound

$$\mathbb{E}\big(Z_\varepsilon(t,x)\big) \geq 1/\psi_2. \qquad (6.9)$$

Using (6.7) and (6.9) and the Paley–Zygmund second moment inequality, we get

$$\mathbb{P}\big(Z_\varepsilon(t,x) \geq 1/\psi_3\big) \geq 1/\psi_4. \qquad (6.10)$$

Like in the proof of Lemma 6.1, we get

$$\bigl|\nabla Z_\varepsilon(t,x)\bigr|^2 = \frac{\beta^2}{(2d)^{2t}} \sum_{q,r\in\mathcal{Q}(t,x)} e^{\beta(g_\varepsilon(q_0)+g_\varepsilon(r_0))}|q\cap r|Y_q(t,x)Y_r(t,x),$$

which, using (6.2), gives

$$\mathbb{E}\bigl|\nabla Z_\varepsilon(t,x)\bigr|^2 = \beta^2\mathbb{E}\bigl(e^{\beta(g_\varepsilon(S_t+x)+g_\varepsilon(S'_t+x))}N_t\mu(\beta)^{N_t}\bigr).$$

Now proceeding as in the proof of (6.7) above, we get

$$\mathbb{E}\bigl|\nabla Z_\varepsilon(t,x)\bigr|^2 \le \psi_5. \tag{6.11}$$

Therefore, as in the proof of Lemma 6.1, we have that for any $a\in\mathbb{R}$ and $b>0$,

$$\begin{aligned}
&\mathbb{P}\bigl(F_\varepsilon(t,x)\ge a, \bigl|\nabla F_\varepsilon(t,x)\bigr|\le b\bigr)\\
&\quad\ge \mathbb{P}\bigl(Z_\varepsilon(t,x)\ge e^a\bigr) - \mathbb{P}\bigl(\bigl|\nabla Z_\varepsilon(t,x)\bigr| > be^a\bigr)\\
&\quad\ge \mathbb{P}\bigl(Z_\varepsilon(t,x)\ge e^a\bigr) - \frac{\psi_5}{b^2e^{2a}}.
\end{aligned}$$

Thus, choosing $a=-\log\psi_3$ and $b=\sqrt{2\psi_4\psi_5}\psi_3$, and using (6.10) and (6.11) we get

$$\mathbb{P}\bigl(F_\varepsilon(t,x)\ge a, \bigl|\nabla F_\varepsilon(t,x)\bigr|\le b\bigr)\ge 1/2\psi_4.$$

Again, by computing second derivatives, it is easy to verify that $F_\varepsilon(t,x)$ is a convex function of the noise variables. By Proposition 3.4, this shows that for all $t>0$,

$$\mathbb{P}\bigl(F_\varepsilon(t,x)\le -\psi_6 - t\bigr)\le 2e^{-t^2/\psi_7}.$$

As in the proof of Lemma 6.1, this implies that $\mathbb{E}(e^{-\theta F_\varepsilon(t,x)})\le\psi_8$. This completes the proof of the lemma. □

7. The main argument

This section contains the main body of the proof of Theorem 2.2. We will continue using the notations introduced in the previous sections. For the reader's convenience, the section is divided into small subsections. Choose two integers $1\le s<t$. These integers will be fixed throughout this section. We will also fix $x\in\mathbb{R}^d$, $\varepsilon\in(0,1)$, and $\beta>0$ such that $\mu(\beta)<1/\rho_d$.

In this section, $C, C_1, C_2, \ldots$ will denote positive constants whose values may depend only on β, d, σ, g, and the sum $\varepsilon|x|+\varepsilon(t-s)+\varepsilon\sqrt{t}+\varepsilon^2 t$. Furthermore, we will require that the dependence on the last item is increasing and continuous in nature. The reason is that we will later send ε to 0, varying s, t and x such that this quantity remains bounded. If a constant depends on some additional parameter θ, we will denote it by $C(\theta)$.

7.1. *The renormalized partition function*

Let $\mathcal{Q}(s,t,x)$ be the set of all nearest-neighbor paths $q=(q_s,q_{s+1},\ldots,q_t)$ with $q_t=x$. For $q\in\mathcal{Q}(s,t,x)$, let

$$Y_q(s,t,x) := \frac{\exp(\beta\sum_{i=s+1}^t \xi_{i,q_i})}{m(\beta)^{t-s}},$$

and define

$$Y(s,t,x) := \frac{1}{(2d)^{t-s}}\sum_{q\in\mathcal{Q}(s,t,x)} Y_q(s,t,x).$$

Let $D(s,t,x)$ be the set of possible values of q_s for $q\in\mathcal{Q}(s,t,x)$. For each $y\in D(s,t,x)$, let $\mathcal{Q}(s,t,x,y)$ for the set of $q\in\mathcal{Q}(s,t,x)$ with $q_s=y$, and define

$$Y(s,t,x,y) := \frac{1}{(2d)^{t-s}}\sum_{q\in\mathcal{Q}(s,t,x,y)} Y_q(s,t,x).$$

Define

$$\zeta(s,t,x,y) := \frac{Y(s,t,x,y)}{Y(s,t,x)}.$$

Note that $\zeta(s,t,x,\cdot)$ is a probability mass function on $D(s,t,x)$. Next, define

$$W_\varepsilon(s,t,x) = \frac{Z_\varepsilon(t,x)}{Y(s,t,x)}.$$

We will refer to $W_\varepsilon(s,t,x)$ as the *renormalized partition function* (as opposed to $Z_\varepsilon(t,x)$, which is the unrenormalized partition function). Our goal in this section is to show that this random variable is approximately equal to the deterministic quantity $G_\varepsilon(s,x)$ with high probability. The first step is the following lemma, which shows that $W_\varepsilon(s,t,x)$ is a weighted average of $Z_\varepsilon(s,y)$ over $y \in D(s,t,x)$.

Lemma 7.1. *We have*

$$W_\varepsilon(s,t,x) = \sum_{y \in D(s,t,x)} \zeta(s,t,x,y) Z_\varepsilon(s,y).$$

Proof. Note that any $q = (q_0,\dots,q_t) \in \mathcal{Q}(t,x)$ can be broken into two parts $(q_s,\dots,q_t) \in \mathcal{Q}(s,t,x)$ and $(q_0,\dots,q_s) \in \mathcal{Q}(s,q_s)$. Thus,

$$\begin{aligned}
Z_\varepsilon(t,x) &= \frac{1}{(2d)^t} \sum_{q \in \mathcal{Q}(t,x)} e^{\beta g_\varepsilon(q_0)} \frac{\exp(\beta \sum_{i=1}^t \xi_{i,q_i})}{m(\beta)^t} \\
&= \frac{1}{(2d)^t} \sum_{q \in \mathcal{Q}(s,t,x)} \sum_{r \in \mathcal{Q}(s,q_s)} e^{\beta g_\varepsilon(r_0)} \frac{\exp(\beta \sum_{i=s+1}^t \xi_{i,q_i})}{m(\beta)^{t-s}} \frac{\exp(\beta \sum_{i=1}^s \xi_{i,r_i})}{m(\beta)^s} \\
&= \frac{1}{(2d)^{t-s}} \sum_{q \in \mathcal{Q}(s,t,x)} \frac{\exp(\beta \sum_{i=s+1}^t \xi_{i,q_i})}{m(\beta)^{t-s}} Z_\varepsilon(s,q_s) \\
&= \sum_{y \in D(s,t,x)} Y(s,t,x,y) Z_\varepsilon(s,y).
\end{aligned}$$

Dividing both sides by $Y(s,t,x)$ gives the desired identity. □

7.2. *Expectation of the renormalized partition function*

Let $\mathcal{F}$ be the σ-algebra generated by the random variables $\{\xi_{u,y} : s+1 \le u \le t, y \in \mathbb{Z}^d\}$. Let $\mathbb{E}'$, $\mathbb{P}'$, Var' and Cov' denote conditional expectation, conditional probability, conditional variance and conditional covariance given $\mathcal{F}$. The following lemma shows that $\mathbb{E}'(W_\varepsilon(s,t,x))$ is a weighted average of $G_\varepsilon(s,y)$ over $y \in D(s,t,x)$.

Lemma 7.2. *We have*

$$\mathbb{E}'\big(W_\varepsilon(s,t,x)\big) = \sum_{y \in D(s,t,x)} \zeta(s,t,x,y) G_\varepsilon(s,y).$$

Proof. Since $\zeta(s,t,x,y)$ is $\mathcal{F}$-measurable and $Z_\varepsilon(s,y)$ is independent of $\mathcal{F}$ for any y, we have

$$\begin{aligned}
\mathbb{E}'\big(W_\varepsilon(s,t,x)\big) &= \sum_{y \in D(s,t,x)} \zeta(s,t,x,y) \mathbb{E}\big(Z_\varepsilon(s,y)\big) \\
&= \sum_{y \in D(s,t,x)} \zeta(s,t,x,y) \Big(\frac{1}{(2d)^s} \sum_{q \in \mathcal{Q}(s,y)} e^{\beta g_\varepsilon(q_0)} \Big).
\end{aligned}$$

To complete the proof, just note that the term inside the bracket on the right side is nothing but $G_\varepsilon(s,y)$. □

As a corollary of the above lemma and Lemma 5.2, we obtain the following result, which shows that $\mathbb{E}'(W_\varepsilon(s,t,x))$ is close to $G_\varepsilon(s,x)$ if $t-s = o(1/\varepsilon)$.

Corollary 7.3. *We have*

$$\bigl|\mathbb{E}'\bigl(W_\varepsilon(s,t,x)\bigr) - G_\varepsilon(s,x)\bigr| \le C\varepsilon(t-s)$$

Proof. Since $\zeta(s,t,x,\cdot)$ is a probability mass function on $D(s,t,x)$, we have

$$\begin{aligned}\bigl|\mathbb{E}'\bigl(W_\varepsilon(s,t,x)\bigr) - G_\varepsilon(s,x)\bigr| &= \Bigl|\sum_{y\in D(s,t,x)} \zeta(s,t,x,y)\bigl(G_\varepsilon(s,y) - G_\varepsilon(s,x)\bigr)\Bigr| \\ &\le \sum_{y\in D(s,t,x)} \zeta(s,t,x,y)\bigl|G_\varepsilon(s,y) - G_\varepsilon(s,x)\bigr|.\end{aligned}$$

To complete the proof, apply Lemma 5.2 to bound $|G_\varepsilon(s,y) - G_\varepsilon(s,x)|$ and use the fact that $|x-y| \le t-s$ for $y \in D(s,t,x)$. □

7.3. *Variance of the renormalized partition function*

We will now show that the renormalized partition function has small conditional variance if ε is small. The first step is to get a bound for covariances.

Lemma 7.4. *For any $y, y' \in D(s,t,x)$,*

$$\mathrm{Cov}\bigl(Z_\varepsilon(s,y), Z_\varepsilon\bigl(s,y'\bigr)\bigr) \le C_1\kappa\bigl(y,y'\bigr)^{C_2}.$$

Proof. By the Lipschitz property of g and the fact that $|y-x| \le t-s$ for any $y \in D(s,t,x)$, we have that for any $y, y' \in D(s,t,x)$,

$$\begin{aligned}&\mathrm{Cov}\bigl(Z_\varepsilon(s,y), Z_\varepsilon\bigl(s,y'\bigr)\bigr) \\ &= \frac{1}{(2d)^{2s}} \sum_{q\in\mathcal{Q}(s,y), r\in\mathcal{Q}(s,y')} e^{\beta(g_\varepsilon(q_0)+g_\varepsilon(r_0))}\,\mathrm{Cov}\bigl(Y_q(s,y), Y_r\bigl(s,y'\bigr)\bigr) \\ &\le \frac{C_1}{(2d)^{2s}} \sum_{q\in\mathcal{Q}(t,y), r\in\mathcal{Q}(t,y')} e^{C_2\varepsilon(|q_0-y|+|r_0-y'|)}\bigl(\mu(\beta)^{|q\cap r|}-1\bigr) \\ &= C_1\mathbb{E}\bigl(e^{C_2\varepsilon(|S_t-y|+|S'_t-y'|)}\bigl(\mu(\beta)^{N_t^{y,y'}}-1\bigr)\bigr),\end{aligned}$$

where $\{S_n\}_{n\ge0}$ and $\{S'_n\}_{n\ge0}$ denote independent simple symmetric random walks started from y and y', and $N_t^{y,y'}$ is the number of times they intersect up to time t. Since $\mu(\beta) < 1/\rho_d$, we can choose $\gamma > 1$, depending only on β, σ, and d, such that $\mu(\beta)^\gamma < 1/\rho_d$. Let $\gamma' := 2\gamma/(\gamma-1)$, so that $2/\gamma' + 1/\gamma = 1$. Applying Hölder's inequality, (4.3), and (4.10), we get

$$\begin{aligned}&\mathbb{E}\bigl(e^{C_2\varepsilon(|S_t-y|+|S'_t-y'|)}\bigl(\mu(\beta)^{N_t^{y,y'}}-1\bigr)\bigr) \\ &\le \bigl[\mathbb{E}\bigl(e^{C_2\gamma'\varepsilon|S_t-y|}\bigr)\bigr]^{1/\gamma'}\bigl[\mathbb{E}\bigl(e^{C_2\gamma'\varepsilon|S'_t-y'|}\bigr)\bigr]^{1/\gamma'}\bigl[\mathbb{E}\bigl|\mu(\beta)^{N_t^{y,y'}}-1\bigr|^\gamma\bigr]^{1/\gamma} \\ &\le C_3\bigl[\mathbb{E}\bigl(\mu(\beta)^{\gamma N_\infty}\bigr)\bigr]^{1/\gamma}\kappa\bigl(y,y'\bigr)^{1/\gamma},\end{aligned}$$

where N_∞ is the number of intersections of two independent simple symmetric random walks in $\mathbb{Z}^d$, started at the origin. Applying (4.7) completes the proof. □

The next lemma gives an upper bound for the $\mathrm{Var}'(W_\varepsilon(s,t,x))$, obtained using the bound on covariances from Lemma 7.4.

Lemma 7.5. *We have*

$$\mathrm{Var}'\bigl(W_\varepsilon(s,t,x)\bigr) \le C_1 \sum_{y,y'\in D(s,t,x)} \zeta(s,t,x,y)\zeta\bigl(s,t,x,y'\bigr)\kappa\bigl(y,y'\bigr)^{C_2}.$$

Proof. Since $\zeta(s,t,x,y)$ is $\mathcal{F}$-measurable for each y, and $\{Z_\varepsilon(s,y)\}_{y\in\mathbb{Z}^d}$ is independent of $\mathcal{F}$, we have

$$\begin{aligned}&\operatorname{Var}'\bigl(W_\varepsilon(s,t,x)\bigr)\\&\quad=\sum_{y,y'\in D(s,t,x)}\zeta(s,t,x,y)\zeta\bigl(s,t,x,y'\bigr)\operatorname{Cov}\bigl(Z_\varepsilon(s,y),Z_\varepsilon\bigl(s,y'\bigr)\bigr).\end{aligned}$$

The proof is now completed by invoking Lemma 7.4. □

Next, we use the above lemma, and Lemma 4.1, to obtain a bound on the expected value of $\operatorname{Var}'(W_\varepsilon(s,t,x))$.

Lemma 7.6. *We have*

$$\mathbb{E}\bigl[\operatorname{Var}'\bigl(W_\varepsilon(s,t,x)\bigr)\bigr]\le C_1(t-s)^{-C_2}.$$

Proof. Let C_2 be as in the statement of Lemma 7.5. First, suppose that in a particular realization of the noise field, we have $Y(s,t,x)\ge(t-s)^{-\alpha}$, where $\alpha>0$ is a constant that we will choose later. Then we have

$$\begin{aligned}&\sum_{y,y'\in D(s,t,x)}\zeta(s,t,x,y)\zeta\bigl(s,t,x,y'\bigr)\kappa\bigl(y,y'\bigr)^{C_2}\\&\quad\le(t-s)^{2\alpha}\sum_{y,y'\in D(s,t,x)}Y(s,t,x,y)Y\bigl(s,t,x,y'\bigr)\kappa\bigl(y,y'\bigr)^{C_2}.\end{aligned}$$

Now observe that

$$\begin{aligned}&\mathbb{E}\biggl(\sum_{y,y'\in D(s,t,x)}Y(s,t,x,y)Y\bigl(s,t,x,y'\bigr)\kappa\bigl(y,y'\bigr)^{C_2}\biggr)\\&\quad=\frac{1}{(2d)^{2(t-s)}}\sum_{y,y'\in D(s,t,x)}\sum_{\substack{q\in\mathcal{Q}(s,t,x,y),\\q'\in\mathcal{Q}(s,t,x,y')}}\mathbb{E}\bigl(Y_q(s,t,x,y)Y_{q'}\bigl(s,t,x,y'\bigr)\bigr)\kappa\bigl(y,y'\bigr)^{C_2}\\&\quad=\frac{1}{(2d)^{2(t-s)}}\sum_{y,y'\in D(s,t,x)}\sum_{\substack{q\in\mathcal{Q}(s,t,x,y),\\q'\in\mathcal{Q}(s,t,x,y')}}\mu(\beta)^{|q\cap q'|}\kappa\bigl(y,y'\bigr)^{C_2}\\&\quad=\frac{1}{(2d)^{2(t-s)}}\sum_{q,q'\in\mathcal{Q}(s,t,x)}\mu(\beta)^{|q\cap q'|}\kappa\bigl(q_s,q'_s\bigr)^{C_2}.\end{aligned}$$

The expression in the last line equals

$$\mathbb{E}\bigl(\mu(\beta)^{N_{t-s}}\kappa\bigl(S_{t-s},S'_{t-s}\bigr)^{C_2}\bigr),$$

where $\{S_n\}_{n\ge0}$ and $\{S'_n\}_{n\ge0}$ are independent simple symmetric random walks started at the origin in $\mathbb{Z}^d$, and N_{t-s} is the number of times they intersect up to time t. Since $\mu(\beta)<1/\rho_d$, we can choose $\delta>1$, depending only on β, σ, and d, such that $\mu(\beta)^\delta<1/\rho_d$. Let $\delta':=\delta/(\delta-1)$, so that $1/\delta'+1/\delta=1$. Applying Hölder's inequality, we get

$$\begin{aligned}\mathbb{E}\bigl(\mu(\beta)^{N_{t-s}}\kappa\bigl(S_{t-s},S'_{t-s}\bigr)^{C_2}\bigr)&\le\bigl[\mathbb{E}\bigl(\mu(\beta)^{\delta N_{t-s}}\bigr)\bigr]^{1/\delta}\bigl[\mathbb{E}\bigl(\kappa\bigl(S_{t-s},S'_{t-s}\bigr)^{C_2\delta'}\bigr)\bigr]^{1/\delta'}\\&\le\bigl[\mathbb{E}\bigl(\mu(\beta)^{\delta N_\infty}\bigr)\bigr]^{1/\delta}\bigl[\mathbb{E}\bigl(\kappa\bigl(S_{t-s},S'_{t-s}\bigr)^{C_2\delta'}\bigr)\bigr]^{1/\delta'}\end{aligned}$$

Applying (4.7) and Lemma 4.1 to the right side gives

$$\mathbb{E}\bigl(\mu(\beta)^{N_{t-s}}\kappa\bigl(S_{t-s},S'_{t-s}\bigr)^{C_2}\bigr)\le C_3(t-s)^{-C_4}.$$

Combining all of the above observations, and using Lemma 7.5, we get that

$$\mathbb{E}\bigl[\operatorname{Var}'\bigl(W_\varepsilon(s,t,x)\bigr);Y(s,t,x)\ge(t-s)^{-\alpha}\bigr]\le C_5(t-s)^{-C_6+2\alpha}.$$

On the other hand, since κ is uniformly bounded by 1, and $\zeta(s,t,x,\cdot)$ is a probability mass function, Lemma 7.5 implies that $\mathrm{Var}'(W_\varepsilon(s,t,x)) \le C_7$. Combining this with the above bound, we get

$$\mathbb{E}\big[\mathrm{Var}'\big(W_\varepsilon(s,t,x)\big)\big] \le C_5(t-s)^{-C_6+2\alpha} + C_7\mathbb{P}\big(Y(s,t,x) < (t-s)^{-\alpha}\big).$$

Now note that $Y(s,t,x)$ has the same distribution as $Y(t-s,x)$. Therefore by Lemma 6.1,

$$\begin{aligned}\mathbb{P}\big(Y(s,t,x) < (t-s)^{-\alpha}\big) &= \mathbb{P}\big(e^{-F(t-s,x)} > (t-s)^{\alpha}\big)\\ &\le (t-s)^{-\alpha}\mathbb{E}\big(e^{-F(t-s,x)}\big)\\ &\le C_8(t-s)^{-\alpha}.\end{aligned}$$

Plugging this bound into the preceding display, and choosing $\alpha = C_6/4$, we get the desired bound. □

7.4. *Concentration of the renormalized height function*

The goal of this subsection is to show that the random variable $\log W_\varepsilon(s,t,x)$, which we call the *renormalized height* of our random surface at (t,x), is close to the deterministic quantity $\log G_\varepsilon(s,x)$ in L^2 distance, if $1 \ll t-s \ll \varepsilon^{-1}$. The proof has several steps. The first step is the following, which shows that $W_\varepsilon(s,t,x)$ is close to $G_\varepsilon(s,x)$ in L^2 distance.

Lemma 7.7. *We have*

$$\mathbb{E}\big[\big(W_\varepsilon(s,t,x) - G_\varepsilon(s,x)\big)^2\big] \le C_1\big(\varepsilon^2(t-s)^2 + (t-s)^{-C_2}\big).$$

Proof. Simply observe that

$$\begin{aligned}&\mathbb{E}\big[\big(W_\varepsilon(s,t,x) - G_\varepsilon(s,x)\big)^2\big]\\ &\quad= \mathbb{E}\big[\mathrm{Var}'(W_\varepsilon(s,t,x)\big] + \mathbb{E}\big[\big(\mathbb{E}'\big(W_\varepsilon(s,t,x)\big) - G_\varepsilon(s,x)\big)^2\big],\end{aligned}$$

and apply Lemma 7.6 and Corollary 7.3 to bound the two terms on the right. □

Next, we show that $\log W_\varepsilon(s,t,x)$ is close to $\log G_\varepsilon(s,x)$ in $L^{1/2}$ distance.

Lemma 7.8. *We have*

$$\mathbb{E}\big|\log W_\varepsilon(s,t,x) - \log G_\varepsilon(s,x)\big|^{1/2} \le C_1\big(\sqrt{\varepsilon(t-s)} + (t-s)^{-C_2}\big).$$

Proof. First, note that

$$\begin{aligned}&\big|\log W_\varepsilon(s,t,x) - \log G_\varepsilon(s,x)\big|^{1/2}\\ &\quad\le \big|W_\varepsilon(s,t,x) - G_\varepsilon(s,x)\big|^{1/2}\big(W_\varepsilon(s,t,x)^{-1/2} + G_\varepsilon(s,x)^{-1/2}\big).\end{aligned}$$

By Lemma 6.2, the observation that $Y(s,t,x)$ has the same law as $Y(t-s,x)$, and the bound (6.3), we get

$$\begin{aligned}\mathbb{E}\big(W_\varepsilon(s,t,x)^{-1}\big) &= \mathbb{E}\left(\frac{Y(s,t,x)}{Z_\varepsilon(t,x)}\right)\\ &\le \big[\mathbb{E}\big(Y(t-s,x)^2\big)\big]^{1/2}\big[\mathbb{E}\big(Z_\varepsilon(t,x)^{-2}\big)\big]^{1/2} \le C.\end{aligned}$$

Also, by Lemma 5.4, $G_\varepsilon(s,x)^{-1/2} \le C$, and by Lemma 7.7,

$$\mathbb{E}\big|W_\varepsilon(s,t,x) - G_\varepsilon(s,x))\big| \le C_1\big(\varepsilon(t-s) + (t-s)^{-C_2}\big).$$

It is now easy to complete the proof by combining all of the above information and applying the Cauchy–Schwarz inequality. □

We want to improve Lemma 7.8 to an L^2 bound. For that, we need the following moment bound.

Lemma 7.9. *For any $\alpha \geq 1$,*

$$\mathbb{E}\big|\log W_\varepsilon(s,t,x)\big|^\alpha \leq C(\alpha).$$

Proof. First, note that since $Y(s,t,x)$ has the same distribution as $Y(t-s,x)$,

$$\begin{aligned}\mathbb{E}\big|\log W_\varepsilon(s,t,x)\big|^\alpha &\leq C(\alpha)\big(\mathbb{E}\big|\log Z_\varepsilon(t,x)\big|^\alpha + \mathbb{E}\big|\log Y(s,t,x)\big|^\alpha\big)\\ &= C(\alpha)\big(\mathbb{E}\big|F_\varepsilon(t,x)\big|^\alpha + \mathbb{E}\big|F(t-s,x)\big|^\alpha\big).\end{aligned}$$

Now, by Lemma 6.1,

$$\begin{aligned}\mathbb{E}\big|F(t-s,x)\big|^\alpha &\leq C(\alpha)\mathbb{E}\big(e^{|F(t-s,x)|}\big)\\ &\leq C(\alpha)\big(\mathbb{E}\big(e^{F(t-s,x)}\big) + \mathbb{E}\big(e^{-F(t-s,x)}\big)\big)\\ &\leq C(\alpha)\big(\mathbb{E}\big(Y(t-s,x)\big) + C\big).\end{aligned}$$

But, as already observed in the proof of Lemma 6.1, $\mathbb{E}(Y(t-s,x)) = 1$. Thus,

$$\mathbb{E}\big|F(t-s,x)\big|^\alpha \leq C(\alpha).$$

Similarly, by Lemma 6.2,

$$\begin{aligned}\mathbb{E}\big|F_\varepsilon(t,x)\big|^\alpha &\leq C(\alpha)\mathbb{E}\big(e^{|F_\varepsilon(t,x)|}\big)\\ &\leq C(\alpha)\big(\mathbb{E}\big(e^{F_\varepsilon(t,x)}\big) + \mathbb{E}\big(e^{-F_\varepsilon(t,x)}\big)\big)\\ &\leq C(\alpha)\big(\mathbb{E}\big(Z_\varepsilon(t,x)\big) + C\big).\end{aligned}$$

By (4.3) and the Lipschitz property of g,

$$\begin{aligned}\mathbb{E}\big(Z_\varepsilon(t,x)\big) &= \frac{1}{(2d)^t}\sum_{q\in\mathcal{Q}(t,x)} e^{\beta g_\varepsilon(q_0)}\\ &\leq C_1 e^{C_2\varepsilon|x|}\mathbb{E}\big(e^{C_3\varepsilon|S_t|}\big) \leq C_4.\end{aligned}$$

Combining the above observations, the proof is complete. □

We are now ready to prove the main result of this subsection, which is the following.

Lemma 7.10. *We have*

$$\mathbb{E}\big[\big(\log W_\varepsilon(s,t,x) - \log G_\varepsilon(s,x)\big)^2\big] \leq C_1\big(\big(\varepsilon(t-s)\big)^{1/4} + (t-s)^{-C_2}\big).$$

Proof. For a nonnegative random variable X, note that by the Cauchy–Schwarz inequality,

$$\mathbb{E}\big(X^2\big) = \mathbb{E}\big(X^{1/4}X^{7/4}\big) \leq \sqrt{\mathbb{E}\big(X^{1/2}\big)\mathbb{E}\big(X^{7/2}\big)}.$$

Applying this to the random variable $|\log W_\varepsilon(s,t,x) - \log G_\varepsilon(s,x)|$, and using Lemma 7.8, Lemma 7.9, and Lemma 5.4, we get the desired result. □

7.5. *Concentration of local averages*

Let $1 \leq r \leq \varepsilon^{-1}$ be a real number. Let $B(x,r)$ be the set of all points in $\mathbb{Z}^d$ that are within Euclidean distance r from x. Define

$$X := \frac{1}{|B(x,r)|}\sum_{y\in B(x,r)} F_\varepsilon(t,y). \tag{7.1}$$

Let $\alpha := \log G_\varepsilon(t,x) + \mathbb{E}(\log Y(t-s,0))$. The following result shows that X is close to α in L^2 distance if $1 \ll t-s \ll r \ll \varepsilon^{-1}$.

Lemma 7.11. *Let r, X and α be as above. Then*

$$\mathbb{E}\bigl[(X-\alpha)^2\bigr] \le C_1 r^{-d}(t-s)^d + C_1\bigl(\varepsilon(t-s)\bigr)^{1/4} + C_1(t-s)^{-C_2} + C_1\varepsilon^2(t-s) + C_1\varepsilon^2 r^2.$$

Proof. Note that paths in $\mathcal{Q}(s,t,x)$ and $\mathcal{Q}(s,t,y)$ do not intersect if $|x-y| > t-s$. Thus, $Y(s,t,x)$ and $Y(s,t,y)$ are independent random variables if $|x-y| > t-s$. On the other hand, from the proof of Lemma 7.9, we know that $\mathbb{E}|\log Y(s,t,x)|^2 \le C$. Combining these two observations, we get that

$$\begin{aligned}
&\operatorname{Var}\biggl(\frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\log Y(s,t,y)\biggr)\\
&= \frac{1}{|B(x,r)|^2}\sum_{y,y'\in B(x,r)}\operatorname{Cov}\bigl(\log Y(s,t,y), \log Y\bigl(s,t,y'\bigr)\bigr)\\
&= \frac{1}{|B(x,r)|^2}\sum_{y,y'\in B(x,r),\,|y-y'|\le t-s}\operatorname{Cov}\bigl(\log Y(s,t,y), \log Y\bigl(s,t,y'\bigr)\bigr)\\
&\le Cr^{-d}(t-s)^d.
\end{aligned}$$

Define

$$\gamma := \mathbb{E}\bigl(\log Y(t-s,0)\bigr) + \frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\log G_\varepsilon(s,y).$$

Recall that $F_\varepsilon(t,y) = \log Y(s,t,y) + \log W_\varepsilon(s,t,y)$. Since $Y(s,t,x)$ has the same distribution as $Y(t-s,x)$, we have $\mathbb{E}(\log Y(s,t,y)) = \mathbb{E}(\log Y(t-s,0))$ for all y. Thus,

$$\begin{aligned}
\mathbb{E}\bigl[(X-\gamma)^2\bigr] &= \mathbb{E}\biggl[\biggl(\frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\bigl(\log W_\varepsilon(s,t,y) - \log G_\varepsilon(s,y)\bigr)\\
&\qquad + \frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\bigl(\log Y(s,t,y) - \mathbb{E}\bigl(\log Y(s,t,y)\bigr)\bigr)\biggr)^2\biggr]\\
&\le 2\mathbb{E}\biggl[\biggl(\frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\bigl(\log W_\varepsilon(s,t,y) - \log G_\varepsilon(s,y)\bigr)\biggr)^2\biggr]\\
&\qquad + 2\operatorname{Var}\biggl(\frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\log Y(s,t,y)\biggr).
\end{aligned}$$

By Lemma 7.10,

$$\begin{aligned}
&\mathbb{E}\biggl[\biggl(\frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\bigl(\log W_\varepsilon(s,t,y) - \log G_\varepsilon(s,y)\bigr)\biggr)^2\biggr]\\
&\le \frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\mathbb{E}\bigl[\bigl(\log W_\varepsilon(s,t,y) - \log G_\varepsilon(s,y)\bigr)^2\bigr]\\
&\le C_1\bigl(\bigl(\varepsilon(t-s)\bigr)^{1/4} + (t-s)^{-C_2}\bigr).
\end{aligned}$$

Next, note that

$$\begin{aligned}
|\alpha-\gamma| &\le \bigl|\log G_\varepsilon(s,x) - \log G_\varepsilon(t,x)\bigr|\\
&\qquad + \frac{1}{|B(x,r)|}\sum_{y\in B(x,r)}\bigl|\log G_\varepsilon(s,y) - \log G_\varepsilon(s,x)\bigr|.
\end{aligned}$$

By Lemma 5.3 and Lemma 5.4, we have

$$\begin{aligned}&\bigl|\log G_\varepsilon(s,x) - \log G_\varepsilon(t,x)\bigr| \\ &\quad\le \bigl|G_\varepsilon(s,x) - G_\varepsilon(t,x)\bigr|\bigl(G_\varepsilon(s,x)^{-1} + G_\varepsilon(t,x)^{-1}\bigr) \\ &\quad\le C\varepsilon\sqrt{t-s}.\end{aligned}$$

Similarly, by Lemma 5.2 and Lemma 5.4, we have that for any $y \in B(x,r)$,

$$\begin{aligned}&\bigl|\log G_\varepsilon(s,y) - \log G_\varepsilon(s,x)\bigr| \\ &\quad\le \bigl|G_\varepsilon(s,y) - G_\varepsilon(s,x)\bigr|\bigl(G_\varepsilon(s,y)^{-1} + G_\varepsilon(s,x)^{-1}\bigr) \\ &\quad\le C\varepsilon|x-y| \le C\varepsilon r.\end{aligned}$$

Combining all of the above observations, we get the desired bound. □

8. Proof of Proposition 2.1

Proposition 2.1 can be proved by a martingale argument, but we already have all the ingredients to give a proof that yields a rate of convergence. The following proposition is a better version of Proposition 2.1.

Proposition 8.1. *Take any $\beta > 0$ such that $\mu(\beta) < 1/\rho_d$. Let $Y(t,x)$ be defined as in equation (6.1). For each positive integer t, let $\eta(\beta,t) := \mathbb{E}(\log Y(t,0))$. Then $\eta(\beta) := \lim_{t\to\infty}\eta(\beta,t)$ exists and is finite. Moreover, for any t,*

$$\bigl|\eta(\beta,t) - \eta(\beta)\bigr| \le C_1 t^{-C_2},$$

where C_1 and C_2 are positive constants that depend only on β, d and σ.

Proof. Take any $1 \le u < t$, and let $s := t-u$. Take any $\varepsilon \in (0,1)$. We will use Lemma 7.10 with $g \equiv 0$. When $g \equiv 0$, we have $G_\varepsilon(t,x) = 1$ for any ε, t and x. Also, we have $Z_\varepsilon(t,x) = Y(t,x)$. Thus, Lemma 7.10 gives

$$\begin{aligned}\bigl|\eta(\beta,t) - \eta(\beta,u)\bigr| &= \bigl|\mathbb{E}\bigl(\log Y(t,0)\bigr) - \mathbb{E}\bigl(\log Y(t-s,0)\bigr)\bigr| \\ &= \bigl|\mathbb{E}\bigl(\log Y(t,0)\bigr) - \mathbb{E}\bigl(\log Y(s,t,0)\bigr)\bigr| \\ &\le \mathbb{E}\bigl|\log Y(t,0) - \log Y(s,t,0)\bigr| \\ &\le \sqrt{\mathbb{E}\bigl[\bigl(\log Y(t,0) - \log Y(s,t,0)\bigr)^2\bigr]} \\ &\le C_1\bigl(\bigl(\varepsilon(t-s)\bigr)^{1/8} + (t-s)^{-C_2}\bigr),\end{aligned}$$

where C_1 and C_2 are positive constants that depend only on β, d, σ, and the sum $\varepsilon|x| + \varepsilon(t-s) + \varepsilon\sqrt{t} + \varepsilon^2 t$. Moreover the dependence on the sum is increasing and continuous. But the left side does not depend on ε. So we can take $\varepsilon \to 0$ and get the bound $|\eta(\beta,t) - \eta(\beta,u)| \le C_3 u^{-C_4}$, where C_3 and C_4 depend only on β, d, and σ. This suffices to prove both claims of the lemma. □

9. Proof of Theorem 2.2

Since the logarithm of any moment generating function is convex,

$$\frac{m'(\beta)}{m(\beta)} = \frac{d}{d\beta}\log m(\beta)$$

is an increasing function of β. Thus,

$$\frac{d}{d\beta}\log\mu(\beta) = 2\left(\frac{m'(2\beta)}{m(2\beta)} - \frac{m'(\beta)}{m(\beta)}\right) \ge 0.$$

This shows that μ is a non-decreasing function, which implies that $\mu(\beta) < 1/\rho_d$ for all $\beta \in (0, \beta_0)$. Thus, all the lemmas proved in the preceding sections are applicable when $\beta \in (0, \beta_0)$.

Take any $(t, x) \in \mathbb{R}_{>0} \times \mathbb{R}^d$. Let t_ε and x_ε be defined as in equation (2.1) and let $\widetilde{f}^{(\varepsilon)}$ be defined as in equation (2.2). Let $\{r_\varepsilon\}_{\varepsilon>0}$ be any collection of positive real numbers such that $1 \ll r_\varepsilon \ll \varepsilon^{-1}$ as $\varepsilon \to 0$. Define $s_\varepsilon := t_\varepsilon - [\sqrt{r_\varepsilon}]$. Define F_ε as in Section 6, so that for any $y \in \mathbb{Z}^d$,

$$F_\varepsilon(t_\varepsilon, y) = \beta f_\varepsilon(t_\varepsilon, y) - t_\varepsilon \log\bigl(2dm(\beta)\bigr). \tag{9.1}$$

In analogy with (7.1), let

$$X_\varepsilon := \frac{1}{|B(x_\varepsilon, r_\varepsilon)|} \sum_{y \in B(x_\varepsilon, r_\varepsilon)} F_\varepsilon(t_\varepsilon, y),$$

so that, by (9.1),

$$\widetilde{f}^{(\varepsilon)}(t, x) = \beta^{-1} X_\varepsilon - \beta^{-1}\eta(\beta). \tag{9.2}$$

Similarly, let $\alpha_\varepsilon := \log G_\varepsilon(t_\varepsilon, x_\varepsilon) + \mathbb{E}(\log Y(t_\varepsilon - s_\varepsilon, 0))$. Now observe that, as $\varepsilon \to 0$,

- $\varepsilon|x_\varepsilon| + \varepsilon(t_\varepsilon - s_\varepsilon) + \varepsilon\sqrt{t_\varepsilon} + \varepsilon^2 t_\varepsilon$ remains uniformly bounded above,
- $r_\varepsilon^{-d}(t_\varepsilon - s_\varepsilon)^d \to 0$,
- $\varepsilon(t_\varepsilon - s_\varepsilon) \to 0$,
- $t_\varepsilon - s_\varepsilon \to \infty$, and
- $\varepsilon r_\varepsilon \to 0$.

Using all of the above observations, and the bound from Lemma 7.11, we get that $X_\varepsilon - \alpha_\varepsilon \to 0$ in L^2 as $\varepsilon \to 0$. But by Lemma 5.1 and Proposition 8.1,

$$\lim_{\varepsilon \to 0} \alpha_\varepsilon = \beta h(t, x) + \eta(\beta).$$

Therefore, by (9.2), $\widetilde{f}^{(\varepsilon)}(t, x) \to h(t, x)$ in L^2 as $\varepsilon \to 0$.

10. Proof of Corollary 2.3

In this proof, $C, C_1, C_2, \ldots$ will denote positive constants that may depend only on β, d, σ, and g, whose values may change from line to line. As in the proof of Theorem 2.2, choose $\{r_\varepsilon\}_{\varepsilon>0}$ such that, as $\varepsilon \to 0$, $1 \ll r_\varepsilon \ll \varepsilon^{-1}$. Define $\widetilde{f}^{(\varepsilon)}$ as in (2.2). By the Cauchy–Schwarz inequality, for any $t > 0$,

$$\begin{aligned}
&\mathbb{E}\biggl[\biggl(\int \widetilde{f}^{(\varepsilon)}(t, x)\phi(x)\,dx - \int h(t, x)\phi(x)\,dx\biggr)^2\biggr] \\
&\quad\le \int \mathbb{E}\bigl[\bigl(\widetilde{f}^{(\varepsilon)}(t, x) - h(t, x)\bigr)^2\bigr]\bigl|\phi(x)\bigr|\,dx \int \bigl|\phi(x)\bigr|\,dx.
\end{aligned}$$

By Theorem 2.2, $\mathbb{E}[(\widetilde{f}^{(\varepsilon)}(t, x) - h(t, x))^2] \to 0$ for any x as $\varepsilon \to 0$. Moreover, by the bound from Lemma 7.11, we see that this convergence is uniform over any compact set. Since ϕ has compact support and is bounded, this proves that as $\varepsilon \to 0$,

$$\int \widetilde{f}^{(\varepsilon)}(t, x)\phi(x)\,dx \xrightarrow{L^2} \int h(t, x)\phi(x)\,dx.$$

Thus, it suffices to show that

$$\int \bigl(f^{(\varepsilon)}(t, x) - \widetilde{f}^{(\varepsilon)}(t, x)\bigr)\phi(x)\,dx \xrightarrow{L^2} 0. \tag{10.1}$$

To prove this, first note that

$$\widetilde{f}^{(\varepsilon)}(t, x) - f^{(\varepsilon)}(t, x) = \frac{1}{|B(x_\varepsilon, r_\varepsilon)|} \sum_{y \in B(x_\varepsilon, r_\varepsilon)} \bigl(f_\varepsilon(t_\varepsilon, y) - f_\varepsilon(t_\varepsilon, x_\varepsilon)\bigr)$$

$$= \frac{1}{|B(0,r_\varepsilon)|} \sum_{y\in B(0,r_\varepsilon)} \bigl(f_\varepsilon(t_\varepsilon, x_\varepsilon + y) - f_\varepsilon(t_\varepsilon, x_\varepsilon)\bigr).$$

Recall that $x_\varepsilon = [\varepsilon^{-1}x]$. Thus, by the change of variable $z = x + \varepsilon y$,

$$\begin{aligned}\int f_\varepsilon(t_\varepsilon, x_\varepsilon + y)\phi(x)\,dx &= \int f_\varepsilon\bigl(t_\varepsilon, [\varepsilon^{-1}x] + y\bigr)\phi(x)\,dx \\ &= \int f_\varepsilon\bigl(t_\varepsilon, [\varepsilon^{-1}z - y] + y\bigr)\phi(z - \varepsilon y)\,dz.\end{aligned}$$

Since $y \in \mathbb{Z}^d$, $[\varepsilon^{-1}z - y] = [\varepsilon^{-1}z] - y$. Thus,

$$\int f_\varepsilon(t_\varepsilon, x_\varepsilon + y)\phi(x)\,dx = \int f_\varepsilon(t_\varepsilon, x_\varepsilon)\phi(x - \varepsilon y)\,dx.$$

Recall the function F_ε defined in Section 6. The above calculations show that

$$\begin{aligned}&\int \bigl(f^{(\varepsilon)}(t,x) - \widetilde{f}^{(\varepsilon)}(t,x)\bigr)\phi(x)\,dx \\ &= \frac{1}{|B(0,r_\varepsilon)|} \sum_{y\in B(0,r_\varepsilon)} \int f_\varepsilon(t_\varepsilon, x_\varepsilon)\bigl(\phi(x) - \phi(x - \varepsilon y)\bigr)\,dx \\ &= \frac{1}{|B(0,r_\varepsilon)|} \sum_{y\in B(0,r_\varepsilon)} \int \beta^{-1}F_\varepsilon(t_\varepsilon, x_\varepsilon)\bigl(\phi(x) - \phi(x - \varepsilon y)\bigr)\,dx,\end{aligned}$$

where the last identity holds because $f_\varepsilon(t_\varepsilon, x_\varepsilon)$ and $\beta^{-1}F_\varepsilon(t_\varepsilon, x_\varepsilon)$ differ by a constant. Thus,

$$\begin{aligned}&\mathbb{E}\biggl[\biggl(\int \bigl(f^{(\varepsilon)}(t,x) - \widetilde{f}^{(\varepsilon)}(t,x)\bigr)\phi(x)\,dx\biggr)^2\biggr] \\ &\le \frac{1}{|B(0,r_\varepsilon)|} \sum_{y\in B(0,r_\varepsilon)} \mathbb{E}\biggl[\biggl(\int \beta^{-1}F_\varepsilon(t_\varepsilon, x_\varepsilon)\bigl(\phi(x) - \phi(x - \varepsilon y)\bigr)\,dx\biggr)^2\biggr].\end{aligned}$$

Let R be so large that $\phi(x) = 0$ when $|x| \ge R - 1$. For each $\delta > 0$, let

$$\omega(\delta) := \sup\bigl\{\bigl|\phi(y) - \phi(z)\bigr| : |y - z| \le \delta\bigr\}.$$

If ε is so small that $\varepsilon r_\varepsilon \le 1$, then for any $y \in B(0, r_\varepsilon)$, $|x - \varepsilon y| \ge |x| - 1$. Therefore, by the Cauchy–Schwarz inequality,

$$\begin{aligned}&\mathbb{E}\biggl[\biggl(\int \beta^{-1}F_\varepsilon(t_\varepsilon, x_\varepsilon)\bigl(\phi(x) - \phi(x - \varepsilon y)\bigr)\,dx\biggr)^2\biggr] \\ &= \mathbb{E}\biggl[\biggl(\int_{|x|\le R} \beta^{-1}F_\varepsilon(t_\varepsilon, x_\varepsilon)\bigl(\phi(x) - \phi(x - \varepsilon y)\bigr)\,dx\biggr)^2\biggr] \\ &\le C R^d \omega(\varepsilon r_\varepsilon) \int_{|x|\le R} \mathbb{E}\bigl[F_\varepsilon(t_\varepsilon, x_\varepsilon)^2\bigr]\,dx.\end{aligned}$$

From the proof of Lemma 7.9, it is easy to see that $\mathbb{E}[F_\varepsilon(t_\varepsilon, x_\varepsilon)^2] \le C$ uniformly over $|x| \le R$. Plugging this into the above, we get

$$\mathbb{E}\biggl[\biggl(\int \bigl(f^{(\varepsilon)}(t,x) - \widetilde{f}^{(\varepsilon)}(t,x)\bigr)\phi(x)\,dx\biggr)^2\biggr] \le C R^{2d}\omega(\varepsilon r_\varepsilon).$$

By the uniform continuity of ϕ, the right side tends to zero as $\varepsilon \to 0$. This proves equation (10.1), and hence completes the proof of the corollary.

Acknowledgements

I thank Chiranjib Mukherjee, Nikos Zygouras, and the anonymous referees for a number of helpful comments and references.

Funding

Research partially supported by NSF grant DMS-1855484.

References

[1] E. Bates and S. Chatterjee. The endpoint distribution of directed polymers. *Ann. Probab.* **48** (2) (2020) 817–871. MR4089496 https://doi.org/10.1214/19-AOP1376
[2] F. Caravenna, R. Sun and N. Zygouras. Universality in marginally relevant disordered systems. *Ann. Appl. Probab.* **27** (5) (2017) 3050–3112. MR3719953 https://doi.org/10.1214/17-AAP1276
[3] F. Caravenna, R. Sun and N. Zygouras. The two-dimensional KPZ equation in the entire subcritical regime. *Ann. Probab.* **48** (3) (2020) 1086–1127. MR4112709 https://doi.org/10.1214/19-AOP1383
[4] P. Carmona and Y. Hu. On the partition function of a directed polymer in a Gaussian random environment. *Probab. Theory Related Fields* **124** (3) (2002) 431–457. MR1939654 https://doi.org/10.1007/s004400200213
[5] S. Chatterjee. *Superconcentration and Related Topics*. Springer, Cham, 2014. MR3157205 https://doi.org/10.1007/978-3-319-03886-5
[6] S. Chatterjee Universality of deterministic KPZ, 2021a. ArXiv preprint. Available at arXiv:2102.13131.
[7] S. Chatterjee Superconcentration in surface growth, 2021b. ArXiv preprint. Available at arXiv:2103.09199.
[8] S. Chatterjee and A. Dunlap. Constructing a solution of the $(2+1)$-dimensional KPZ equation. *Ann. Probab.* **48** (2) (2020) 1014–1055. MR4089501 https://doi.org/10.1214/19-AOP1382
[9] F. Comets. *Directed Polymers in Random Environments. Lecture Notes from the 46th Probability Summer School Held in Saint-Flour, 2016*. Springer, Cham, 2017. MR3444835 https://doi.org/10.1007/978-3-319-50487-2
[10] F. Comets, C. Cosco and C. Mukherjee. Space-time fluctuation of the Kardar–Parisi–Zhang equation in $d \geq 3$ and the Gaussian free field, 2019. ArXiv preprint. Available at arXiv:1905.03200. MR4262971 https://doi.org/10.1103/physreve.103.042102
[11] F. Comets, C. Cosco and C. Mukherjee. Renormalizing the Kardar–Parisi–Zhang equation in $d \geq 3$ in weak disorder. *J. Stat. Phys.* **179** (3) (2020) 713–728. MR4099995 https://doi.org/10.1007/s10955-020-02539-7
[12] C. Cosco, S. Nakajima and M. Nakashima. Law of large numbers and fluctuations in the sub-critical and L^2 regions for SHE and KPZ equation in dimension $d \geq 3$, 2020. ArXiv preprint. Available at arXiv:2005.12689. MR4441505 https://doi.org/10.1016/j.spa.2022.05.010
[13] A. Dunlap, Y. Gu, L. Ryzhik and O. Zeitouni. Fluctuations of the solutions to the KPZ equation in dimensions three and higher. *Probab. Theory Related Fields* **176** (3) (2020) 1217–1258. MR4087492 https://doi.org/10.1007/s00440-019-00938-w
[14] Y. Gu. Gaussian fluctuations from the 2D KPZ equation. *Stoch. Partial Differ. Equ. Anal. Comput.* **8** (1) (2020) 150–185. MR4058958 https://doi.org/10.1007/s40072-019-00144-8
[15] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** (1963) 13–30. MR0144363
[16] M. Kardar, G. Parisi and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** (9) (1986) 889–892.
[17] M. Ledoux. *The Concentration of Measure Phenomenon*. American Mathematical Society, Providence, RI, 2001. MR1849347 https://doi.org/10.1090/surv/089
[18] D. Lygkonis and N. Zygouras. Edwards–Wilkinson fluctuations for the directed polymer in the full L^2-regime for dimensions $d \geq 3$. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (1) (2022) 65–104. MR4374673 https://doi.org/10.1214/21-aihp1173
[19] J. Magnen and J. Unterberger. The scaling limit of the KPZ equation in space dimension 3 and higher. *J. Stat. Phys.* **171** (4) (2018) 543–598. MR3790153 https://doi.org/10.1007/s10955-018-2014-0
[20] C. Mukherjee, A. Shamov and O. Zeitouni. Weak and strong disorder for the stochastic heat equation and continuous directed polymers in $d \geq 3$. *Electron. Commun. Probab.* **21** (2016) 12. MR3548773 https://doi.org/10.1214/16-ECP18