THE $2d$ KPZ AS A MARGINALLY RELEVANT DISORDERED SYSTEM

NIKOS ZYGOURAS

Abstract. We describe a series of works, joint with Francesco Caravenna and Rongfeng Sun, which make some first steps towards the understanding of scaling limits of disordered systems in a suitable weak disorder regime, where disorder has a so-called marginally relevant effect. This includes some first understanding of the KPZ equation in two space dimensions, which in the language of SPDEs is the "critical dimension". Among the results that we will describe is a phase transition and the identification of the KPZ solution below a critical temperature, which falls into the Edwards-Wilkinson universality class. Emphasis will be given on conveying the main ideas, stripped off the technical parts, as well as describing the methods which include Lindeberg principles, fourth moment theorems, analysis on Wiener spaces, multiscale analysis etc.

Contents

1. Introduction.

The KPZ equation is a stochastic PDE, formally written as

$$\partial_t h(t,x) = \frac{1}{2}\Delta h(t,x) + \frac{1}{2}|\nabla h(t,x)|^2 + \beta\, \dot{W}(t,x), \qquad t>0, x\in\mathbb{R}^d, \tag{1.1}$$

Date: September 22, 2019.

1991 *Mathematics Subject Classification.* Primary: 82B44.

Key words and phrases. Lindeberg theorem, Fourth moment theorem, Wiener chaos, multilinear chaos, disordered systems.

where $\dot{W}(t,x)$ is the space-time white noise, defined as a gaussian field, which is delta-correlated in space and time as

$$\mathbb{E}[\dot{W}(t,x)\,\dot{W}(s,y)] = \delta(t-s)\delta(x-y).$$

The parameter $\beta > 0$ modulates the strength of the noise. The KPZ equation was introduced by Kardar, Parisi and Zhang [KPZ86] as a model for random interface growth obeying three basic principles:

- growth is proportional to the normal vector of the interface $\sqrt{1+|\nabla h|^2}$, which is "approximated" via a Taylor expansion by $1+\frac{1}{2}|\nabla h|^2$,
- there is local smoothing, represented by the Laplacian Δh,
- there is local randomness, represented by the white noise $\xi(t,x)$.

The well posedness of equation (1.1) is subject to questioning as one expects the solution to be "rough" in space, thus leading to ambiguities in the definition of the term $|\nabla h|^2$, since ∇h should be a distribution rather than a function.

A vast amount of work has been carried out in spatial dimension one, which has by now put into firm grounds (and has significantly extended) the original predictions of Kardar, Parisi and Zhang that the fluctuation of the "solution" $h(t,x)$ should be of order $t^{1/3}$, contrary to the usual $t^{1/2}$ dictated by the central limit theorem. This fact was established by first discovering and analysing an "integrable" structure via the study of discrete models, such as the *asymmetric exclusion process*, the *directed polymer model*, *last passage percolation* etc., and then approximating via the discrete models the solution to (1.1) in a suitable sense. We refer to the surveys [C12, QS15, BP14, BG16, Z18] for some relevant reviews. Thus, devising a way to provide a meaning to the notion of solution to the KPZ equation (and therefore a means of approximating it) is an important task. In dimension one the first result of this type was achieved by Bertini and Giacomin [BG97], who approximated the KPZ via a particle system called the asymmetric exclusion process; this is a one dimensional particle system where each particle jumps independently at an exponentially distributed time with probability p to the right and $1-p$ to the left, with the constraint that no two or more particles occupy the same site.

More robust approaches emerged in recent years, giving rise to new revolutionary theories that further allowed to treat a wide class of singular stochastic PDEs. These are the theories of Regulariy Structures by Hairer [H14], of Paractontrolled Distributions by Gubinelli-Imkeller-Perkowski [GIP15] and of Energy Solutions by Goncalves and Jara [GJ10, GJ14]. A renormalisation approach to the one dimensional KPZ equation has also been successful through the work of Kupiainen [K14].

In higher dimensions the situation is much less understood. A firm prediction about the exponents that govern the fluctuations of the solution to the KPZ is missing and so does a solution theory. Dimension two is characterised as a *critical dimension* and marks the limitation of the above solution theories. The goal of these notes is to review some first steps that have taken place more recently into understanding the two dimensional KPZ equation. Moreover, we will see that some of the mechanisms governing the KPZ equation in the critical dimension two also govern a wider class of models of statistical mechanics where disorder is present and has a so called *marginal* effect. Our focus here will be mainly to summarise a series of works [CSZ17a, CSZ17b, CSZ18a, CSZ18b, CSZ18c, CSZ18+]. At the end of this introduction we will briefly describe some other interesting works around this topic.

One may see the qualitative difference between dimension one and higher dimensions via a *renormalisation* procedure. Let us describe this on the closely (as we will explain) related model of the stochastic heat equation (SHE), which will also link nicely to the theme of disordered systems and disorder relevance. The stochastic heat equation writes as

$$\partial_t u(t,x) = \frac{1}{2}\Delta u(t,x) + \dot{W}(t,x)u(t,x), \qquad \text{with} \quad t>0, x\in\mathbb{R}^d. \tag{1.2}$$

and its solution is formally related to the solution of the KPZ via the Hopf-Cole transformation $h = \log u$ (we will be more precise about this and some subtleties later on). The renormalisation we alluded to earlier amounts, in its simplest formulation, to a scaling of the variables as

$$(t,x) \mapsto (\varepsilon^2 t, \varepsilon x)\,. \tag{1.3}$$

Using the gaussian scaling property of the white noise, which says that

$$\dot{W}(\varepsilon^2 t, \varepsilon x) := \frac{W(\mathrm{d}(\varepsilon^2\, t)\,,\, \mathrm{d}(\varepsilon x))}{\mathrm{d}(\varepsilon^2\, t)\, \mathrm{d}(\varepsilon x)} \overset{\mathrm{d}}{=} \varepsilon^{-1-\frac{d}{2}}\, \frac{W(\mathrm{d}t\,,\, \mathrm{d}x)}{\mathrm{d}t\,\mathrm{d}x} = \varepsilon^{-1-\frac{d}{2}}\dot{W}(t,x),$$

it is not difficult to see that $u^\varepsilon(t,x) := u(\varepsilon^2 t, \varepsilon x)$ formally solves the SPDE

$$\partial_t u^\varepsilon = \frac{1}{2}\Delta u^\varepsilon + \beta\varepsilon^{1-\frac{d}{2}}\dot{\widetilde{W}}\,u^\varepsilon, \tag{1.4}$$

where $\dot{\widetilde{W}}$ is a new space-time White noise obtained from $\dot{W}$ via the above scaling. Therefore, space-time renormalisation has the effect of changing the strength of the noise to $\varepsilon^{1-\frac{d}{2}}\beta$.

We now see that if $d<2$, then, as $\varepsilon\to 0$, the strength of the noise in the renormalized equation goes to zero, which means that the noise will have a gradually decreasing effect on the regularity of the solution to the SHE and thus a solution can be suitably defined. On the other hand, for $d>2$ the noise should crucially affect the solution as its strength after renormalisation increases. One also sees that $d=2$ is a critical dimension as the renormalisation leaves the noise invariant and thus no conclusion can be drawn on the effect of noise to the existence and regularity of a solution.

In this notes we will be viewing the KPZ and SHE equations within the framework of *disordered systems*. In the realm of disordered systems, one is interested in large scale effects, as these would be captured by the reciprocal change of variables $(t,x)\mapsto(\varepsilon^{-2}t, \varepsilon^{-1}x)$. This will result to a renormalized equation where now the strength of the noise is $\varepsilon^{\frac{d}{2}-1}\beta$. Then for $d<2$, then noise (disorder) has a prevailing effect (amounting to disorder relevance), while for $d>2$ the effect of the noise vanishes, amounting to disorder irrelevance. However, again, when $d=2$ the renormalisation leaves the noise invariant and no conclusion can be drawn on the effect of noise. This is the marginal case. We will provide more details on this framework in Section 2.

The rigorous procedure that is followed in order to make sense of singular SPDEs like (1.2) is to first look at an equation with some sorts of smoothed out or mollified noise and show that the solution to this has some sort of limit when the mollification is taken away. In the case of SHE one typically proceeds by looking at the equation

$$\partial_t u^\varepsilon = \frac{1}{2}\Delta u^\varepsilon + \beta u^\varepsilon \dot{W}^\varepsilon, \tag{1.5}$$

where $\dot W^\varepsilon$ is the mollified *in space* noise $\dot W^\varepsilon(t,x) := \int_{\mathbb{R}^2} j_\varepsilon(x-y)\dot W(t,y)\mathrm{d}y$ with $j_\varepsilon(x) := \varepsilon^{-d} j(x/\varepsilon)$ and with $j \in C_c^\infty(\mathbb{R}^d)$ a symmetric, probability density on $\mathbb{R}^d$. $\dot W^\varepsilon$ is still a white noise in time. In particular, $\dot W^\varepsilon(t,\cdot)\mathrm{d}t =: \dot W^\varepsilon(\mathrm{d}t,\cdot)$ is a Brownian motion with quadratic variation equal to $\varepsilon^{-d}\|j\|^2_{L^2(\mathbb{R}^d)}$.

Using an adaptation of the Feynman-Kac formula for Brownian-like potentials such as $\dot W^\varepsilon(\mathrm{d}t,\cdot)$ one can write a path integral representation for the solution to (1.5) as [BC95, Sec. 3 and (3.22)]

$$u^\varepsilon(t,x) = \mathrm{E}_x\Big[\exp\Big\{\beta\int_0^t \dot W^\varepsilon(t-s,B_s)\mathrm{d}s - \frac12\beta^2\,\mathbb{E}\Big[\Big(\int_0^t \dot W^\varepsilon(t-s,B_s)\mathrm{d}s\Big)^2\Big]\Big\}\Big], \tag{1.6}$$

where E_x is expectation with respect to a standard Brownian motion $(B_s)_{s\geqslant 0}$ on $\mathbb{R}^d$ with $B_0 = x$. $\mathbb{E}$ denotes the expectation with respect to the white noise. By time reversal, we note that $u^\varepsilon(t,x)$ has the same distribution (for fixed (t,x)) as

$$\begin{aligned}\widetilde u^\varepsilon(t,x) &:= \mathrm{E}_x\Big[\exp\Big\{\beta\int_0^t \dot W^\varepsilon(s,B_s)\mathrm{d}s - \frac12\beta^2\,\mathbb{E}\Big[\Big(\int_0^t \dot W^\varepsilon(s,B_s)\mathrm{d}s\Big)^2\Big]\Big\}\Big] \\ &= \mathrm{E}_x\Big[\exp\Big\{\beta\int_0^t\int_{\mathbb{R}^2} j_\varepsilon(B_s-y)\dot W(s,y)\mathrm{d}s\mathrm{d}y - \frac12\beta^2 t\|j_\varepsilon\|_2^2\Big\}\Big] \\ &= \mathrm{E}_{\varepsilon^{-1}x}\Big[\exp\Big\{\beta\,\varepsilon^{1-\frac d2}\int_0^{\varepsilon^{-2}t}\int_{\mathbb{R}^2} j(B_{\widetilde s}-\widetilde y)\dot{\widetilde W}(\widetilde s,\widetilde y)\mathrm{d}\widetilde s\mathrm{d}\widetilde y - \frac{\beta^2\varepsilon^{2(1-\frac d2)}}{2}(\varepsilon^{-2}t)\|j\|_2^2\Big\}\Big],\end{aligned} \tag{1.7}$$

where in the last step we made the change of variables $(\varepsilon\widetilde y, \varepsilon^2\widetilde s) := (y,s)$ using that, by scaling, $\dot{\widetilde W}$ defined by $\dot{\widetilde W}(\widetilde s,\widetilde y)\mathrm{d}\widetilde s\mathrm{d}\widetilde y := \varepsilon^{-1-\frac d2}\dot W(\varepsilon^2\widetilde s,\varepsilon\widetilde y)\mathrm{d}(\varepsilon^2\widetilde s)\mathrm{d}(\varepsilon\widetilde y)$ is also a two-dimensional space-time white noise.

Formula (1.7) is what is known as the (partition function of the) *continuum directed polymer in a random medium* or simply the *continuum directed polymer model.* The role of the polymer is played by the Brownian path $(B_s)_{s>0}$ and the role of the random medium is played by the white noise. One approach to the SHE (and, thus, to the KPZ equation) will be to use the polymer representation and justify that a limit exists as $\varepsilon \to 0$. In dimension one this was successfully dealt with by Bertini and Cancrini in [BC95]. In dimension two we have that $1-\frac d2 = 0$ and thus the regularity provided by the factor $\varepsilon^{1-\frac d2}$ in $d=1$ disappears and things become more subtle. It turns out that, in order to have a hope to obtain a non trivial limit, one needs to consider a renormalised version of the mollified SHE (1.5) where β is chosen to be $\beta_\varepsilon = \hat\beta\sqrt{2\pi/\log\varepsilon^{-1}}$. The reason for this choice will be clarified later. The goal then is to establish a limit for

$$\mathrm{E}_{\varepsilon^{-1}x}\Big[\exp\Big\{\beta_\varepsilon\int_0^{\varepsilon^{-2}t}\int_{\mathbb{R}^2} j(B_s-y)\dot W(s,y)\mathrm{d}s\mathrm{d}y - \frac{\beta_\varepsilon^2}{2}(\varepsilon^{-2}t)\|j\|_2^2\Big\}\Big]. \tag{1.8}$$

It is worth highlighting the following. In dimension two, choosing $\beta_\varepsilon = \hat\beta\sqrt{2\pi/\log\varepsilon^{-1}}$, the term containing the noise in

$$\partial_t u^\varepsilon = \frac12\Delta u^\varepsilon + \frac{\hat\beta\sqrt{2\pi}}{\sqrt{\log\varepsilon^{-1}}}u^\varepsilon\dot W^\varepsilon,$$

appears to be reduced to zero. Thus the naive guess would be that the limit is just the solution of the standard heat equation, which, assuming initial condition equal to one, would be also equal to one. However, as we will see, this is certainly not the case. For any fixed space-time point $(t,x) \in \mathbb{R}_+\times\mathbb{R}^2$, $u^\varepsilon(t,x)$ will converge to a log-normal variable, as $\varepsilon \downarrow 0$, if

$\hat{\beta} < 1$. As a field, $u^\varepsilon(t,\cdot)$ is uncorrelated as captured by a central limit theorem in the form that for every test function $\phi \in C_c^\infty(\mathbb{R}^2)$

$$\sqrt{\frac{\log \varepsilon^{-1}}{2\pi}} \int_{\mathbb{R}^2} (u^\varepsilon(t,x) - 1)\, \phi(x)\, \mathrm{d}x \xrightarrow[\varepsilon \to 0]{d} \int_{\mathbb{R}^2} v^{(c_{\hat{\beta}})}(t,x) \phi(x)\, \mathrm{d}x \tag{1.9}$$

where the centering 1 is due to the assumption that the initial condition is 1 and where $v^{(c)}(t,x)$ is the solution of the two-dimensional *additive* stochastic heat equation

$$\begin{cases} \partial_t v^{(c)}(t,x) = \frac{1}{2}\Delta v^{(c)}(t,x) + c\,\dot{W}(t,x) \\ \quad v(0,x) \equiv 0, \end{cases} \tag{1.10}$$

The value of c_β in (1.9) equals $c_{\hat{\beta}} := \sqrt{\frac{\hat{\beta}^2}{1-\hat{\beta}^2}}$. The solution to this linear equation is well defined and it is a mean zero gaussian field. This field is what governs the so-called *Edward-Wilkinson universality* class, which is characterised by gaussian fluctuations.

Interesting correlations also appear at scales ε^χ with $\chi \in (0,1)$ (still for $\beta_\varepsilon = \hat{\beta}\sqrt{2\pi/\log \varepsilon^{-1}}$ with $\hat{\beta} < 1$). In this case $u^\varepsilon(t, \varepsilon^\chi \cdot)$ converges to a field of correlated log-normals with the correlations depending on the exponent χ. We will not expand in this direction and we will refer to [CSZ17b, Theorems 2.12 and 2.15] for details.

More interesting is the situation at or above the critical temperature $\hat{\beta}_{\text{critical}} = 1$ where, on the one hand, $u^\varepsilon(t,x)$ turns out to converge to zero, for every fixed (t,x) but, on the other hand, the field is strongly correlated. In particular, at $\hat{\beta} = \hat{\beta}_{\text{critical}} = 1$ one expects a non-trivial limit for

$$\int_{\mathbb{R}^2} (u^\varepsilon(t,x) - 1)\, \phi(x)\, \mathrm{d}x$$

where comparing to (1.9) we notice the absence of the factor $\sqrt{\log \varepsilon^{-1}}$. However, the scaling limit in this case is still not understood. In fact, even the existence of a unique limit has not been proved, yet. The only results that exists are moment computations [BC98, CSZ18a, GQT19], which provide existence of non-trivial (i.e. not constant) subsequential limits but however do not provide uniqueness or do not determine the limit, except the fact that limits are log-correlated and non gaussian fields. Let us mention, though, some interesting recents works of Clark [C19a, C19b, C19c] where he studies polymers on the so-called *hierarchical diamond lattice* in a weak disorder scaling, similar as we consider here, at the critical temperature and proves existence and some characterisation of the scaling limit. It is possible that there exists some relation between these scaling limits and the ones for the $2d$ polymer / SHE. So far, there have not been any works above the critical value, i.e. $\hat{\beta} > \hat{\beta}_{\text{critical}} = 1$.

Coming back to the KPZ equation, let us see how we can transfer some of the understanding described around the SHE. First, let us relate the two solutions via the Hopf-Cole transformation $h^\varepsilon := \log u^\varepsilon$. One needs to be careful in deriving the equation for h^ε as u^ε is the solution of a stochastic PDE and thus one needs to employ Itô calculus when differentiating with respect to time. Doing this carefully, we have that the equation for h^ε is

$$\partial_t h^\varepsilon(t,x) = \frac{1}{2}\Delta h^\varepsilon(t,x) + \frac{1}{2}|\nabla h^\varepsilon(t,x)|^2 + \beta \dot{W}^\varepsilon(t,x) - \frac{\beta^2}{2}\varepsilon^{-d}\|j\|^2_{L^2(\mathbb{R}^d)}. \tag{1.11}$$

where one notices the correction term $-\frac{\beta^2}{2}\varepsilon^{-d}\|j\|^2_{L^2(\mathbb{R}^d)}$, which converges to infinity as $\varepsilon \to 0$.

Restricting attention to dimension two, we choose, as in the SHE, $\beta = \beta_\varepsilon = \hat\beta\sqrt{2\pi/\log\varepsilon^{-1}}$ and the equation we deal with is

$$\partial_t h^\varepsilon(t,x) = \frac{1}{2}\Delta h^\varepsilon(t,x) + \frac{1}{2}|\nabla h^\varepsilon(t,x)|^2 + \beta_\varepsilon^2 \dot W^\varepsilon(t,x) - \frac{\beta_\varepsilon^2}{2}\varepsilon^{-2}\|j\|^2_{L^2(\mathbb{R}^d)}. \tag{1.12}$$

It turns out that equation (1.12) has certain invariance (see, for example, [CSZ18b, Appendix A]), which leads to the equivalent formulation

$$\partial_t \widetilde h^\varepsilon(t,x) = \frac{1}{2}\Delta \widetilde h^\varepsilon(t,x) + \frac{\beta_\varepsilon}{2}|\nabla \widetilde h^\varepsilon(t,x)|^2 + \dot W^\varepsilon(t,x). \tag{1.13}$$

The relation between the solutions to the two equations is that

$$h^\varepsilon(t,x) = \beta_\varepsilon\, \widetilde h^\varepsilon(t,x).$$

Again, looking at the form of equation (1.13) we see that the nonlinearity is gradually reduced to zero and one might naively expect that the limit satisfies the additive SHE (5.4) with $c = 1$, as would be by just dropping out the nonlinearity in (1.13). This is not completely true, though, as we will see that for $\hat\beta < 1$

$$\frac{1}{\beta_\varepsilon}\int_{\mathbb{R}^2} (h^\varepsilon(t,x) - 1)\,\phi(x)\,\mathrm{d}x \overset{d}{=} \int_{\mathbb{R}^2} (\widetilde h^\varepsilon(t,x) - 1)\,\phi(x)\,\mathrm{d}x \xrightarrow[\varepsilon\to 0]{d} \int_{\mathbb{R}^2} v^{(\widetilde c_{\hat\beta})}(t,x)\phi(x)\,\mathrm{d}x$$

with $\widetilde c_{\hat\beta} = \sqrt{\frac{1}{1-\hat\beta^2}}$, which is strictly larger than 1. This suggests that the term $\frac{\beta_\varepsilon}{2}|\nabla\widetilde h^\varepsilon|^2$ produces in the limit a noise term. We remark that the situation for $\hat\beta \geqslant 1$ has not yet been settled for the KPZ equation.

In these notes we will mostly work in a discrete setting, with the discrete version of the partition function of the continuum polymer. This is not only a matter of convenience. It fits into a more general framework of studying scaling limits of so called *disordered systems*.

The discrete version of the polymer partition function is

$$Z_{N,\beta} := \mathrm{E}\Big[e^{\sum_{n=1}^N (\beta\omega(n,S_n) - \lambda(\beta))}\Big].$$

Here $(\omega(n,x))_{n\geqslant 1, x\in\mathbb{Z}^d}$ is a collection of i.i.d. random variables, which we assume to have mean zero, variance one and log-moment generating function $\lambda(\beta) := \log\mathbb{E}\big[e^{\beta\omega}\big] < \infty$ for all $\beta > 0$. This family of variables plays the role of the white noise. The role of the Brownian path $(B_s)_{s>0}$ is played by a simple, symmetric random walk $(S_n)_{n\geqslant 1}$ on $\mathbb{Z}^d$.

In dimension one, choosing $\varepsilon = 1/\sqrt{N}$, we have that the discrete analogue of (1.7) is the partition function Z_{N,β_N} with $\beta_N = \hat\beta N^{-1/4}$. The fact that in this case Z_{N,β_N} has a well defined limit was shown by Alberts-Khanin-Quastel [AKQ14] (see also [CSZ17a] for a more general framework). In dimension two we will choose $\beta_N = \hat\beta/\sqrt{\log N}$ in analogy with what was described above within the SHE and KPZ framework.

The directed polymer model is an example of a *disordered system*, where a "pure" statistical model (in this case the simple random walk) is perturbed by disorder / noise. The interest then is to understand the effect that disorder has on the pure systems and whether *arbitrarily small amount of disorder* is sufficient to change its statistical properties. If it does, then disorder is called *relevant* while if a sufficiently large amount of disorder is required, then disorder is called *irrelevant*. In [CSZ17a, CSZ18+] we formulated the question of disorder relevance in the form of whether a partition function, such as that of the directed polymer, has a non trivial. i.e. random, limit for suitable choice of the parameter β going to zero as the size of the system increases. In [CSZ17b] we extended this framework to include systems where disorder has a *marginally relevant* effect. This included the two dimensional directed polymer and SHE. Our treatment of the two dimensional directed polymer and SHE follows the methods which were motivated by the study of more general disordered systems and the phenomenon of disorder relevance.

Before closing this introduction let us make a quick review of the recent literature around higher dimensional KPZ and SHE equations, which has been marked by a surge of activity. In dimension two, which we will be mainly interested here, Chatterjee and Dunlap, using different methods, proved tightness of the solution to the KPZ equation (1.13). Gu [G18], using methods around Malliavin calculus, proved Edwards-Wilkinson universality. Neither of these works covers the whole subcritical regime $\hat{\beta} < \hat{\beta}_{\text{critical}}$.

In $d \geqslant 3$ the strength of the noise in the SHE (1.5) is modulated as $\hat{\beta}\varepsilon^{\frac{2-d}{2}}$ and there also exists a critical value $\hat{\beta}_{\text{critical}}$, which marks a transition between weak and strong disorder. The existence of this phase transition was shown in [MSZ16, CCM18] and it is in the spirit of the transition that has been known for directed polymers in $d \geqslant 3$ from the works of Comets, Shiga, Yoshida [CSY04]. In particular, if $\hat{\beta} < \hat{\beta}_{\text{critical}}$ then, for fixed (t, x), the solution $u^\varepsilon(t, x)$ converges to an a.s. positive limit, while for $\hat{\beta} > \hat{\beta}_{\text{critical}}$ it converges to zero. Contrary to dimension two, where we have identified precisely the critical value of $\hat{\beta}$, in $d \geqslant 3$ the understanding of the weak-to-strong transition and the critical temperature that marks this transition is rather poor . For example there is no characterisation of the critical value of $\hat{\beta}$.

The Malliavin calculus approach of [G18] was used earlier in [DGRZ18] to prove Edwards-Wilkinson universality for the KPZ equation in $d \geqslant 3$. Earlier works on the $d \geqslant 3$ KPZ via renormalisation techniques are those of Magnen and Unterberger [MU18]. These works also do not cover the whole subcritical regime. The advantage of the methods exposed in these notes compared to those in the above works, which allow to cover the whole subcritical regime in $d = 2$, is that they make a detailed analysis of the polynomial / Wiener chaos expansion of the polymer model / SHE. There is currently work in progress [L19] to extend the Edwards-Wilkinson universality for $d \geqslant 3$ in the whole subcritical regime.

Let us remark that in space dimension $d = 1$, the Cole-Hopf solution $h(t, x) := \log u(t, x)$ of the KPZ equation (1.1) is well-defined as a random function, for all $\beta \in (0, \infty)$. Moreover, there is no phase transition in the one-point distribution as β varies. Thus, Edwards-Wilkinson fluctuations for $h(t, x)$ and $u(t, x)$ can be easily established as $\beta \downarrow 0$, combining Wiener chaos and Taylor expansion (because $u(t, x) \to 1$).

It is also interesting to consider a variation of the KPZ equation, which is called *anisotropic KPZ* and where the nonlinearity $|\nabla h|^2$ is replaced by $\langle \nabla h, A\nabla h\rangle$ for some matrix A with $\det(A) \leqslant 0$. There is some belief and some evidence that the anisotropic KPZ falls into the Edwards-Wilkinson universality class. We refer to the review [T17] for some details on the current understanding. Let us just mention some recent work by Cannizzaro, Erhard and Schönbauer [CES19], where A is the diagonal matrix with diagonal elements $1, -1$ and the nonlinearity is modulate in a similar fashion as the $2d$ KPZ equation by $\beta/\sqrt{\log \varepsilon^{-1}}$ †. They prove tightness of the fields and non triviality of the limit points. Unlike the usual KPZ equation there is currently no indication of a phase transition in $\hat{\beta}$ in the anisotropic KPZ.

†the approximation of the anisotropic KPZ that is performed in [CES19] is a bit different than the mollification procedure we have been working with. It goes into working in the Fourier space and performing a truncation in the Fourier modes

2. Disorder relevance

Let us start with a description of the notion of a disordered system. Consider an open set $\Omega \subseteq \mathbb{R}^d$ and define the lattice $\Omega_\delta := (\delta\mathbb{Z})^d \cap \Omega$, for $\delta > 0$, which is the support of a *"random field"* $\sigma = (\sigma_x)_{x\in\Omega_\delta}$ whose law is determined by a probability measure denoted by $\mathrm{P}^{\mathrm{ref}}_{\Omega_\delta}$. Typically, the field takes values $\sigma_x \in \{0,1\}$ or $\{\pm 1\}$. Even though it also sensible to consider fields that take non binary values, currently the treatment of such fields is not covered by the methods we will describe.

Some examples of such fields can be:

- *Ising models.* In this case, $\Omega_\delta := (\delta\mathbb{Z})^d \cap \Omega$ with δ representing the mesh of the grid on $\Omega \subset \mathbb{R}^d$ and $\sigma_x \in \{\pm 1\}$. The measure $\mathrm{P}^{\mathrm{ref}}_{\Omega_\delta}$ is the Ising measure given by
$$\mathrm{P}^{\mathrm{ref}}_{\Omega_\delta}(\sigma) = \frac{1}{Z^{\mathrm{ref}}_{\Omega_\delta}} e^{\,J\sum_{x\sim y}\sigma_x\sigma_y},$$
where $x \sim y$ means that sites x, y are nearest neighbour, i.e. connected by an edge of $\mathbb{Z}^d$. J is a coupling constant, which represents the strength of interaction between neighbouring values of the field σ and
$$Z^{\mathrm{ref}}_{\Omega_\delta} := \sum_{\sigma} e^{\,J\sum_{x\sim y}\sigma_x\sigma_y},$$
is the partition function.
- *Random walks.* In this case, Ω is typically $\mathbb{Z}^d \times \{0,1,...,N\}$ for $N \geqslant 1$ and Ω_δ is the scaled version $N^{-1/2}\mathbb{Z}^d \times N^{-1}(\mathbb{N} \cap [0,1])$, We notice that in this case the lattice is given its more natural parabolic scaling. The field $\sigma = (\sigma_{n,x})_{n\leqslant N, x\in\mathbb{Z}^d}$ is $\sigma_{n,x} = \mathbb{1}_{\{S_n = x\}}$, where $(S_n)_{n\geqslant 1}$ is the trajectory of a random walk.

A disordered system arises when on the lattice Ω_δ, on top of the reference field σ, there exists an additional randomness, $\omega := (\omega_x)_{x\in\Omega_\delta}$ modelled in the form of an i.i.d. collection, which is typically assumed to be of mean zero, variance one and having exponential moments (although it is sensible to relax the exponential moment assumption and consider heavy tailed fields, in which case new phenomena often arise, see for example [AL11, HM07, DZ16, BT18, BT19]). We call the randomness ω *disorder* and denote its law by $\mathbb{P}$ and its expectation with respect to it by $\mathbb{E}$.

Given a realisation of the disorder ω, the *disordered model* is defined as the following probability measure $\mathrm{P}^{\omega}_{\Omega_\delta;\lambda,h}$ for the field $\sigma = (\sigma_x)_{x\in\Omega_\delta}$:

$$\mathrm{P}^{\omega}_{\Omega_\delta;\beta,h}(\sigma) := \frac{e^{\sum_{x\in\Omega_\delta}(\beta\omega_x + h)\sigma_x}}{Z^{\omega}_{\Omega_\delta;\beta,h}} \mathrm{P}^{\mathrm{ref}}_{\Omega_\delta}(\sigma)\,, \tag{2.1}$$

where now the partition function is defined by

$$Z^{\omega}_{\Omega_\delta;\beta,h} := \mathrm{E}^{\mathrm{ref}}_{\Omega_\delta}\big[e^{\sum_{x\in\Omega_\delta}(\beta\omega_x + h)\sigma_x}\big]\,. \tag{2.2}$$

and we remark that in this case it is itself a random variable, depending on the realisation ω.

A question of central interest in statistical physics but often very poorly understood is

Q. *" does an arbitrarily small amount of disorder change the statistical mechanics properties of the reference field ? "*

In the 70s A.B. Harris [H74] proposed the following criterion, which is known as *Harris criterion*:

Harris criterion: *If d is the (effective) dimension and γ is the correlation length exponent (we will be more precise about this below) of the reference model, then if $\gamma > \frac{d}{2}$, the model is* **disorder irrelevant***, meaning that small enough amount of disorder is not sufficient to change its statistical properties. If $\gamma < \frac{d}{2}$, then the model is* **disorder relevant***, meaning that any arbitrarily small amount of disorder does change its statistical properties.*

Let us first define what we mean by a *correlation length exponent* here: Consider the (what is called) *k-point correlation function* of the field σ to be:

$$\psi_\delta^{(k)}(x_1, ..., x_k) := \mathrm{E}^{\mathrm{ref}}_{\Omega_\delta}[\sigma_{x_1} \cdots \sigma_{x_k}].$$

Then the correlation length exponent can be defined as the exponent γ such that

$$(\delta^{-\gamma})^k \, \psi_\delta^{(k)}(x_1, \dots, x_k) \xrightarrow[\delta \downarrow 0]{} \boldsymbol{\psi}_\Omega^{(k)}(x_1, \dots, x_k). \tag{2.3}$$

where the limit is to be thought of as pointwise, although stronger forms such as in $L^2(\Omega^k)$ will be needed for the framework we will develop.

Even though very simple, actually rigorously verifying the Harris criterion in concrete examples is often difficult and requires a careful case by case analysis (although one can transfer some intuition and a set of "general principles" between different problems). An overview of the features of disorder relevance and the challenges verifying the Harris criterion can be found in [G11].

In [CSZ17a] we proposed a different point of view for the Harris criterion focusing on the existence of *non-trivial* (i.e. random) scaling limits of the partition functions when β, h are suitably scaled to zero as $\delta \to 0$. The question can be phrased as:

Q. *Consider the partition function of a disordered model as defined in* (2.2)*. Can we choose $\beta = \beta_\delta$ and $h = h_\delta$, converging to zero as $\delta \to 0$, such that $Z^\omega_{\Omega_\delta;\beta_\delta,h_\delta}$ converges in distribution to a random (i.e. finite and not constant) random variable $\boldsymbol{Z}^W_{\Omega;\hat\beta,\hat h}$* ?

Here W denotes white noise on $\mathbb{R}^d$ and we request that the limit should be a non trivial function of an underlying white noise.

We will now describe this point of view, whose core is *multilinear* and *Wiener chaos expansion* and *Lindeberg principles for multilinear polynomials*, which we will describe in detail in the following section.

Although it makes sense to consider a general value of h, we will, for simplicity, restrict ourselves to the choice of $h = -\lambda(\beta)$, where $\lambda(\beta) := \log \mathbb{E}[e^{\beta\omega_x}]$. We denote the partition function associated to this choice by $Z^\omega_{\Omega_\delta;\beta}$.

Let us assume that the field (σ_x) takes values in $\{0,1\}$. The starting point is to write the partition function in the form of a multilinear polynomial. We do this via what is called in statistical mechanics *high temperature* or *Mayer expansion*, which goes by writing

$$Z^\omega_{\Omega_\delta;\beta} = \mathrm{E}^{\mathrm{ref}}_{\Omega_\delta}\Bigg[\prod_{x \in \Omega_\delta} \big(1 + \beta\sigma_x\zeta_x\big)\Bigg], \qquad \text{where} \qquad \zeta_x := \frac{e^{(\beta\omega_x - \lambda(\beta))} - 1}{\beta}. \tag{2.4}$$

Here we used the fact that

$$e^{(\beta\omega_x - \lambda(\beta))\sigma_x} - 1 = \Big(e^{(\beta\omega_x - \lambda(\beta))} - 1\Big)\sigma_x, \qquad \text{if } \sigma_x \in \{0,1\}.$$

Expanding the product and interchanging the (finite) summation with the expectation $\mathrm{E}^{\mathrm{ref}}_{\Omega_\delta}$, we write

$$Z^{\omega}_{\Omega_\delta;\beta} = 1 + \sum_{k=1}^{\infty} \beta^k \sum_{x_1,\dots,x_k \in \Omega_\delta} \mathrm{E}^{\mathrm{ref}}_{\Omega_\delta}\big[\sigma_{x_1}\cdots\sigma_{x_k}\big] \prod_{i=1}^{k} \zeta_{x_i},$$

where the inner sum is taken over k-tuples of *distinct* $x_1, \dots, x_k \in \Omega_\delta$ (and so the sum over k even though written as an infinite sum it is in fact finite). Denoting $\psi^{(k)}_\delta(x_1,\dots,x_k) := \mathrm{E}^{\mathrm{ref}}_{\Omega_\delta}[\sigma_{x_1}\cdots\sigma_{x_k}]$ we write

$$Z^{\omega}_{\Omega_\delta;\beta} = 1 + \sum_{k=1}^{\infty} (\beta\delta^{\gamma})^k \sum_{x_1,\dots,x_k \in \Omega_\delta} (\delta^{-\gamma})^k\, \psi^{(k)}_\delta(x_1,\dots,x_k) \prod_{i=1}^{k} \zeta_{x_i}, \tag{2.5}$$

where we have inserted the assumed scaling of the k-point correlation function. Note that the random variables (ζ_x) are mean zero precisely due to the choice of the parameter h to be equal to $-\lambda(\beta)$. Moreover, for β small, they have asymptotically unit variance.

At this point we need of a *Lindeberg principle* : suppose that we can replace the random variables (ζ_x), from (2.4), by standard normal variables, which we denote by (ξ_x). If so, then we could model this new collection of i.i.d. normal via a White noise $W(\cdot)$ on $\mathbb{R}^d$ as

$$\xi_x = |C_{x,\delta}|^{-1/2}\, W\big(C_{x,\delta}\big),$$

where $C_{x,\delta}$ is the cube in $(\delta\mathbb{Z})^d$ with side length δ, "bottom-left" corner equal to x and volume $|C_{x,\delta}| = \delta^d$. Consider now the partition function

$$Z^{W}_{\Omega_\delta;\beta} = 1 + \sum_{k=1}^{\infty} \big(\beta\delta^{\gamma-\frac{d}{2}}\big)^k \sum_{x_1,\dots,x_k \in \Omega_\delta} (\delta^{-\gamma})^k\, \psi^{(k)}_\delta(x_1,\dots,x_k) \prod_{i=1}^{k} W(C_{x_i,\delta}), \tag{2.6}$$

which can also be written as an iterated Wiener-Itô integral in the form

$$Z^{W}_{\Omega_\delta;\beta} = 1 + \sum_{k=1}^{\infty} \big(\beta\delta^{\gamma-\frac{d}{2}}\big)^k \int\cdots\int_{\Omega^k} (\delta^{-\gamma})^k\, \psi^{(k,\mathrm{ext})}_\delta(x_1,\dots,x_k) \prod_{i=1}^{k} W(\mathrm{d}x_i),$$

where $\psi^{(k,\mathrm{ext})}_\delta$ is the piecewise constant function on Ω^k, which takes the constant value $\psi^{(k)}_\delta$ on the cubes $C_{x_1,\delta} \times \cdots \times C_{x_k,\delta}$.

Choosing now

$$\beta = \beta_\delta = \hat\beta\, \delta^{\frac{d}{2}-\gamma}, \tag{2.7}$$

ones sees via an easy $L^2(\mathbb{P})$ estimate and using assumption (2.3) (strengthened to hold in an $L^2(\Omega^k)$ sense) that

$$Z^{W}_{\Omega_\delta;\hat\beta} \xrightarrow[\delta\downarrow 0]{L^2(\mathbb{P})} 1 + \sum_{k=1}^{\infty} \hat\beta^k \int\cdots\int_{\Omega^k} \boldsymbol{\psi}^{(k)}(x_1,\dots,x_k) \prod_{i=1}^{k} W(\mathrm{d}x_i). \tag{2.8}$$

We should here remark at the **consistency with the Harris criterion**: the scaling of β in (2.7) is consistent with the requirement that $\beta_\delta \to 0$ for $\delta \to 0$ (thus disorder is gradually weaker) if $\gamma < d/2$. In the case that $\gamma > d/2$ it turns out that any scaling of β tending to zero as $\delta \to 0$ will always lead to be a trivial, i.e. non random and in fact constant, limit.

As we see, the main point in obtaining the scaling limit of the disordered partition function is justifying the passage (in the limit $\delta \to 0$) from (2.5) to (2.6). This step is precisely achieved with the Lindeberg principle, which will be described in the next section, see Theorem 3.2. Let us note that if h in (2.2) is taken to be different than $-\lambda(\beta)$, then the random variables (ζ_x) are not mean zero and in this case one needs to be more careful as issues of convergence of the series in (2.5) arise. Moreover, one needs an extension of Lindeberg Theorem 3.2 that will cover the situation of non-mean-zero variables. These issues have been settled and suitable extensions of Theorem 3.2 have been achieved in [CSZ17a].

At the marginal case $\frac{d}{2} = \gamma$ the above procedure has two problems. The first one is that the scaling of β is not well defined as the exponent $\frac{d}{2} - \gamma$ vanishes. This is rectified by typically choosing a logarithmic (or more general slowly varying) scaling. However, another more serious issue typically arises, that the limiting correlation kernel $\boldsymbol{\psi}^k_\Omega$ in (2.3) is not in $L^2(\Omega^k)$ and thus the candidate Wiener integrals in (2.8) are not well defined. This is the situation that one also faces when attempting to define the limit of the directed polymer model or of the SHE and KPZ. We will see in Section 4 that a different structure takes place in this situation.

3. Some general tools: Chaos expansions, Lindeberg principles, Fourth moment theorems, Hypercontractivity

3.1. Multilinear polynomials and Lindeberg principle. Let us define multilinear polynomials as follows. Consider a family of i.i.d. random variables $\xi := (\xi_x)_{x\in\mathsf{S}}$ indexed by a countable set S Let $\mathcal{P}^{\text{fin}}(\mathsf{S}) := \{I \subset \mathsf{S}\colon |I| < \infty\}$, the set of all finite subsets of S. Consider also a function $\psi\colon \mathcal{P}^{\text{fin}}(\mathsf{S}) \to \mathbb{R}$. Then a multilinear polynomial of disorder ξ, associated to ψ is defined as

$$\Psi(\xi) := \sum_{I\in\mathcal{P}^{\text{fin}}(\mathsf{S})} \psi(I)\xi^I, \quad \text{where} \quad \xi^I := \prod_{a\in I} \xi_a, \quad \text{with} \quad \xi^{\varnothing} := 1. \tag{3.1}$$

Assuming that $\mathbb{E}[\xi_a] = 0$ and $\mathbb{V}\text{ar}(\xi_a) = 1$, it is easy to compute the variance of $\Psi(\xi)$ as

$$\mathbb{V}\text{ar}(\Psi(\xi)) = \boldsymbol{\sigma}^2_\Psi := \sum_{I\in\mathcal{P}^{\text{fin}}(\mathsf{S}), I\neq\varnothing} \psi(I)^2. \tag{3.2}$$

An important feature, that we would like to quantify, is the "*influence*" that a single variable has on the overall random function. In other words, "how much" does the random function change if we change, e.g. by resampling, one of its random variables.

This motivates putting the notion of *influence* in a mathematical context. We define

Definition 3.1. *Let $(\omega_x)_{x\in\mathsf{S}}$ be a family of i.i.d., mean zero and variance one real valued variables indexed by a countable set S. Let $f\colon \mathbb{R}^{\mathsf{S}} \to \mathbb{R}$ be a function of this family of variables. The influence of entry $x \in \mathsf{S}$ is defined as*

$$\text{Inf}_x(f) := \mathbb{E}\Big[\, \mathbb{V}\text{ar}(f(\omega) \,\big|\, \{\omega_y\}_{y\neq x})\,\Big].$$

In the case of multilinear polynomials the influence of entry $x \in \mathsf{S}$ equals

$$\text{Inf}_x\big(\Psi\big) = \mathbb{E}\left[\Big(\frac{\partial\Psi}{\partial\omega_x}\Big)^2\right] = \sum_{I\ni x} \psi(I)^2.$$

The notion of influence plays an important role in the "replacement principle" that we discussed earlier (see discussion around (2.6)) and which goes under the name of Lindeberg principle. We have the following theorem

Theorem 3.2. *Let $\zeta = (\zeta_a)_{a\in\mathsf{S}}$ and $\xi = (\xi_a)_{a\in\mathsf{S}}$ be two families of independent random variables with mean zero, variance one and uniformly integrable second moment. Let $\Psi(\xi), \Psi(\zeta)$ be the associated multilinear polynomials as defined in* (3.1) *and assume that $\boldsymbol{\sigma}_\Psi^2 := \sum_{\varnothing\neq I\in\mathcal{P}^{\mathrm{fin}}(\mathsf{S})} \psi(I)^2$ is finite.*

Then for every $f \in C_b^3(\mathbb{R})$ and any $\varepsilon > 0$, there exists C_ε depending not only on ε but also on $\|f'\|_\infty, \|f''\|_\infty, \|f'''\|_\infty$ and $\boldsymbol{\sigma}_\Psi^2$, such that

$$\Big|\mathbb{E}\big[f\big(\Psi(\xi)\big)\big] - \mathbb{E}\big[f\big(\Psi(\zeta)\big)\big]\Big| \leqslant \varepsilon + C_\varepsilon \sqrt{\max_{a\in\mathsf{S}} \mathrm{Inf}_a(\Psi)}. \tag{3.3}$$

The above theorem was proved in [CSZ17a] and it is an improvement of a theorem in [MOO10] (see also [R74]), where the Lindeberg principle for multilinear polynomials was proved under the assumption of finite third moment. The above theorem captures an optimal, in terms of moments, condition. In [CSZ17a] a more quantitative expression of the right hand side on (3.3) was provided. Moreover, in [CSZ17a] a statement of the Lindeberg principle for non mean zero variables was proved.

A direct consequence of the above theorem is that if one has a sequence of multi-linear functionals Ψ_n for which it holds that

$$\max_{a\in\mathsf{S}} \mathrm{Inf}_a(\Psi_n) \xrightarrow[n\to\infty]{} 0, \tag{3.4}$$

then the asymptotic distribution of $\Psi_n(\xi)$ and $\Psi_n(\zeta)$ are the same assuming that the families ξ and ζ have matching first and second moment (e.g. mean zero and variance one) and uniformly integrable second moments. Assumption (3.4) is typically satisfied when one considers multilinear polynomials corresponding to partition functions of disorder relevant systems and thus the Lindeberg principle of Theorem 3.2 facilitates the passage between representations (2.5) and (2.6).

We will provide the proof of Theorem 3.2 in the Appendix.

3.2. Fourth Moment Theorem.

A main tool that we will use to handle the asymptotic limits of the two dimensional SHE and KPZ as well as of general marginally relevant disordered systems (in the subcritical regime $\hat\beta < \beta_{\mathrm{critical}}$) is the so-called *fourth moment theorem.* In the form of multilinear polynomials this remarkable type of theorem asserts that a sequence of multilinear polynomials of mean zero and variance one random variables ξ with sufficient moments (recall also notation from (3.1))

$$\Psi_n(\zeta) = \sum_{I\in\mathcal{P}^{\mathrm{fin}}(\mathsf{S})} \psi_n(I)\xi^I,$$

converges to a standard normal variable if and only if its variance converges to 1 and its fourth moment converges to 3.

Fourth moment theorems were (re)discovered and popularised by Nualart and Peccati [NP05] in the context of Wiener chaoses. Versions of the fourth moment theorem in the setting of discrete chaoses were discovered earlier in the study of statistics by Sevastyanov [S61] (for bilinear forms) and later by de Jong [dJ87, dJ90]. Since the work of Nualart and Peccati there has been an explosion of fourth moment theorems in various contexts and with many applications. Some of the most sharp proof techniques make use of Malliavin calculus (see for example [NPR10]).

Let us state a version of the fourth moment theorem from [CSZ17b], which is an extension (using the Lindeberg Theorem 3.2) of a theorem of Nourdin, Peccati and Reinert [NPR10] to random variables with just uniformly integrable second moments.

Theorem 3.3. *For each $N \in \mathbb{N}$, let $(\xi_x)_{x\in\mathsf{S}}$ be independent random variables with mean 0 and variance 1, indexed by a countable set S. Assume that $(\xi_x^2)_{x\in\mathsf{S}}$ are uniformly integrable. Fix $k \in \mathbb{N}$ and $d_1, \ldots, d_k \in \mathbb{N}$. For each $1 \leqslant i \leqslant k$, let $\Psi_N^{(i)}(\xi)$ be a multi-linear polynomial in $(\xi_x)_{x\in\mathsf{S}}$ of degree d_i, i.e.,*

$$\Psi_N^{(i)}(\xi) = \sum_{I\subset \mathsf{S},\, |I|=d_i} \psi_N^{(i)}(I)\xi^I \qquad \text{for some real-valued } \psi_N^{(i)}(\cdot).$$

Assume further that:

(i) *For all $1 \leqslant i, j \leqslant k$, $\mathbb{E}[\Psi_N^{(i)}(\xi)\Psi_N^{(j)}(\xi)] \to V(i,j)$ for some matrix V as $N \to \infty$;*

(ii) *For each $1 \leqslant i \leqslant k$, $\mathbb{E}[\Psi_N^{(i)}(\omega)^4] \to 3\,V(i,i)^2$ as $N \to \infty$, where we have replaced $(\xi_x)_{x\in\mathsf{S}}$ by i.i.d. standard normal random variables $(\omega_x)_{x\in\mathsf{S}}$;*

(iii) *The maximal influence of each variable ξ_x on the polynomials of degree one among $(\Psi_N^{(i)}(\xi_x))_{1\leqslant i\leqslant k}$ is asymptotically negligible, i.e., for each $1 \leqslant i \leqslant k$,*

$$\max_{x\in\mathsf{S}} |\psi_N^{(i)}(x)| \to 0 \qquad \text{as} \quad N \to \infty. \tag{3.5}$$

Then $(\Psi_N^{(i)}(\xi)_{1\leqslant i\leqslant k}$ converge jointly in law to a centered Gaussian vector with covariance V.

3.3. Hypercontractivity. Let us now discuss the notion of hypercontractivity. Some references on hypercontractivity are [S98] for a discrete setting and [J97] for hypercontractivity on Gaussian spaces.

The notion of hypercontractivity is very useful when one needs to control higher than two moments via second moments. The significance of this is that when dealing with multilinear expansions their second moments can be easily computed (see for example (3.2)), while higher moments, which do appear in our KPZ estimates, are not easily computable, in particular when these are non integers. Thus having a tool that will allow this reduction in a *sharp* way is very important and we will make use of this in Section 5 when proving Edwards-Wilkinson universality of the 2d KPZ.

Let us give the following definition of hypercontractivity for multi-linear polynomials.

Definition 3.4. *Let $\Psi(\xi) := \sum_{I\in\mathcal{P}^{\mathrm{fin}}(\mathsf{S})} \psi(I)\xi^I$ be a multi-linear polynomial of the family of random variables $\xi = (\xi_a)_{a\in\mathsf{S}}$. For $\varrho > 0$, define the operator T_ϱ acting on the multilinear polynomial as*

$$\big(T_\varrho\Psi\big)(\xi) := \sum_{I\in\mathcal{P}^{\mathrm{fin}}(\mathsf{S})} \varrho^{|I|}\,\psi(I)\,\xi^I,$$

where $|I|$ denotes the cardinality of the set $|I|$. For $\varrho \geqslant 1$ and $1 \leqslant p \leqslant q < \infty$, we will say that the family ξ is $(p, q, \varrho)-$hypercontractive if

$$\|\Psi\|_q \leqslant \|\,T_\varrho\Psi\,\|_p.$$

for **all** *multi-linear polynomials Ψ.*

The following theorem shows that families of random variables with more than two moments are hypercontractive.

Theorem 3.5. *Let $\xi = (\xi_x)_{x\in\mathsf{S}}$ be a family of i.i.d. random variables such that*

$$\mathbb{E}\big[\xi_x\big] = 0, \quad \mathbb{E}\big[\xi_x^2\big] = 1 \quad \text{and} \quad \exists p_0 \in (2,\infty)\colon \mathbb{E}\big[\xi_x^{p_0}\big] < \infty$$

Then, for every $p \in (2, p_0)$ the family ξ is $(2, p, \varrho_p)$-hypercontractive with

$$\lim_{p \downarrow 2} \varrho_p = 1. \tag{3.6}$$

In particular, we have that, for every every multilinear polynomial $\Psi(\xi) = \sum_{I \in \mathcal{P}^{\mathrm{fin}}(\mathsf{S})} \psi(I) \xi^I$ and for $p \in (2, p_0)$,

$$\mathbb{E}\big[|\Psi(\xi)|^p\big] \leqslant \Big(\sum_{I \in \mathcal{P}^{\mathrm{fin}}(\mathsf{S})} \varrho_p^{|I|} \, \psi(I) \, \xi^I \Big)^{p/2}.$$

Except for the sharp asymptotic (3.6), this theorem was proved in [MOO10] as an extension of the corresponding result in the Gaussian framework, see [J97, S98]. The sharp asymptotic (3.6) was proved in [CSZ18b, Theorem B.1] and is important for proving the Edwards-Wilkinson fluctuations for the KPZ in the *entire* subcritical regime $\hat{\beta} < 1$ in Theorem 5.1. The estimate on the hypercontractivity constant given in [MOO10, Proposition 3.16] was

$$2\sqrt{p-1} \sup_{N \in \mathbb{N}} \frac{\mathbb{E}[|\xi_x|^p]^{1/p}}{\mathbb{E}[|\xi_x|^2]^{1/2}} = 2\sqrt{p-1} \sup_{N \in \mathbb{N}} \mathbb{E}[|\xi_x|^p]^{1/p},$$

which when $p \downarrow 2$ it converges to 2, instead of the natural value 1. This extra factor 2 is the byproduct of a non-optimal symmetrization argument in the proof in [MOO10].

4. Marginal relevance via the $2d$ directed polymer and SHE

We will now study the case of marginal relevance that we touched upon at the end of Section 2. The two dimensional polymer and SHE fall in this category and so we will use this as the main example in this section. For other marginal models that fall under the scope of the methods described here we refer to [CSZ17b].

Let $(S_n)_{n \geqslant 1}$ be a simple, two dimensional random walk and let the disorder $(\omega_{n,x})_{n \in \mathbb{N}, x \in \mathbb{Z}^d}$ satisfy the usual conditions of mean zero, variance one and exponential moments $\lambda(\beta) := \log \mathbb{E}[e^{\beta\omega}] < \infty$. We will denote the law and expectation of a random walk starting at location x by P_x and E_x, respectively, while the law and expectation with respect to the disorder will be denoted by $\mathbb{P}$ and $\mathbb{E}$. We set

$$q_n(x) := \mathrm{P}(S_n = x). \tag{4.1}$$

The local limit theorem asserts that

$$q_n(x) = 2 g_{n/2}(x) \mathbb{1}_{\{n + x_1 + x_2 \in 2\mathbb{Z}\}} + o(\tfrac{1}{n}), \qquad \text{uniformly for} \quad x \in \mathbb{Z}^2, \tag{4.2}$$

where the factor 2 is due to periodicity and $g_t(x) = \frac{1}{2\pi t} e^{-|x|^2/2t}$ is the two dimensional heat kernel. Denoting by $\widetilde{S}$ an independent copy of S, we define the expected *overlap*, which will play an important role in the normalisation below, by

$$R_N := \sum_{n=1}^N \mathrm{P}(S_n = \widetilde{S}_n) = \sum_{n=1}^N \sum_{x \in \mathbb{Z}^2} q_n(x)^2 = \sum_{n=1}^N q_{2n}(0) = \frac{\log N}{\pi} + O(1), \tag{4.3}$$

where we used the convolution property of the random walk $\sum_x q_n(x)^2 = q_{2n}(0)$ and the local limit theorem asymptotics. The partition function of the directed polymer model, for a random walk starting at location x, is given by

$$Z_{N,\beta}(x) := \mathrm{E}_x\Big[e^{\sum_{n=1}^N (\beta\,\omega(n,S_n) - \lambda(\beta))} \Big] = \mathrm{E}_x\Big[e^{\sum_{1 \leqslant n \leqslant N, y \in \mathbb{Z}^2} (\beta\,\omega(n,y) - \lambda(\beta))\, \mathbb{1}_{\{S_n = y\}}} \Big]. \tag{4.4}$$

Performing a similar expansion as in (2.4) we can write the partition function as

$$Z_{N,\beta}(x) = 1 + \sum_{k=1}^{N} \sigma(\beta)^k \sum_{\substack{0=n_0<n_1<\ldots<n_k\leqslant N \\ x_0=x,\ x_1,\ldots,x_k\in\mathbb{Z}^2}} \prod_{i=1}^{k} q_{n_i-n_{i-1}}(x_i - x_{i-1})\, \xi_{n_i,x_i}, \tag{4.5}$$

where

$$\xi_{n,x} := \sigma(\beta)^{-1}\Big(e^{\beta\omega(n,x)-\lambda(\beta)} - 1\Big) \qquad \text{with} \qquad \sigma = \sigma(\beta) := \sqrt{e^{\lambda(2\beta)-2\lambda(\beta)} - 1}. \tag{4.6}$$

$\sigma(\beta)$ is chosen so that the random variables $\xi_{n,x}$ are normalized to have variance one. But since $\sigma(\beta) \sim \beta$, for β small, we will usually replace it by β without extra reference.

The determination of the scaling of β in terms of N comes from a variance computation

$$\begin{aligned}\mathbb{E}[Z_{N,\beta}(x)^2] &= 1 + \sum_{k=1}^{N} \sigma(\beta)^{2k} \sum_{\substack{0=:n_0<n_1<\ldots<n_k\leqslant N \\ x_0=:x,\ x_1,\ldots,x_k\in\mathbb{Z}^2}} \prod_{i=1}^{k} q_{n_i-n_{i-1}}(x_i - x_{i-1})^2 \\ &= 1 + \sum_{k=1}^{N} \sigma(\beta)^{2k} \sum_{0=:n_0<n_1<\ldots<n_k\leqslant N} q_{2(n_i-n_{i-1})}(0)\,.\end{aligned} \tag{4.7}$$

Looking, for example, at the first term of this expansion and using (4.3) we see that to keep this term of order one, we should choose $\beta = \beta_N$ as

$$\beta_N := \frac{\hat\beta}{\sqrt{R_N}} = \frac{\sqrt{\pi}\hat\beta}{\sqrt{\log N}}(1 + o(1))\,, \qquad \text{for } \hat\beta \in (0,\infty). \tag{4.8}$$

This turns out to also be the right choice in order to keep the variance of the rest of the terms to be of order one. The logarithmic scaling, here, should be contrasted to the power law scaling in (2.7). This reflects the fact that we are at the marginal case and so the exponent in (2.7) vanishes. Indeed, under the parabolic scaling and the local limit theorem (4.2), it holds that $Nq_{tN}(x\sqrt{N})$ has a non trivial limit and so the correlation exponent is $\gamma = 1$, reflected by the power of N multiplying the kernel q. On the other hand, the effective dimension is d_{eff} = one time scaling + twice space scaling $= 1 + 2 \cdot \frac{1}{2} = 2$. Thus $d_{\text{eff}}/2 = \gamma$.

The first guess for the continuum limit of $Z_{N,\beta_N}(x)$ under this choice of β_N would be

$$1 + \sum_{k\geqslant 1} \hat\beta^k \int\!\!\cdots\!\!\int_{\substack{0<t_1<\cdots<t_{k-1}<\varepsilon^{-2} \\ x_1,\ldots,x_k\in\mathbb{R}^2,\, x_0=x}} \prod_{i=1}^{k} g_{t_i-t_{i-1}}(x_i - x_{i-1})\, W(\mathrm{d}t_i, \mathrm{d}x_i).$$

Notice that this would be the obvious form for the solution to the stochastic heat equation via the standard Picard iteration. However, there is a problem as the above integrals are not properly defined in the Itô sense. This can be easily checked by computing the L^2 norm of, for example, the first integral which is

$$\hat\beta^2 \int_0^1 \int_{\mathbb{R}^2} g_t(x)^2 \mathrm{d}t\, \mathrm{d}x = \hat\beta^2 \int_0^1 \frac{1}{4\pi t}\, \mathrm{d}t,$$

which blows up logarithmically.

The scaling limit is indeed more subtle and less obvious. Remarkably, the situation is universal among models which can be characterised as *marginally relevant* in the sense described in the following theorem

Theorem 4.1 ([CSZ17b]). *Let $Z^{\mathrm{marginal}}_{N,\beta_N}$ be a multilinear polynomial (typically a partition function) of the form*

$$Z^{\mathrm{marginal}}_{N,\beta_N} = 1 + \sum_{k=1}^{N} \beta_N^k \sum_{\substack{1\leqslant n_1<\cdots<n_k\leqslant N\\ x_1,\dots,x_k\in\mathbb{Z}^d}} \prod_{i=1}^{k} q_{n_i-n_{i-1}}(x_i-x_{i-1})\,\xi_{n_i,x_i},$$

where we may assume that x_0 in the the above summation is equal to 0. $(\xi_{n,x})_{n\in\mathbb{N},x\in\mathbb{Z}^d}$ is a collection of i.i.d. mean zero, variance one random variables with exponential moments and the kernel $(q_n(x))_{n\in\mathbb{N},x\in\mathbb{Z}^d}$ satisfies that

$$R_N := \sum_{n=1}^{N}\sum_{x\in\mathbb{Z}^d} q_n(x)^2 \quad \text{grows to infinity as a slowly varying function.} \tag{4.9}$$

We also assume that the kernel $q_n(x)$ satisfies a type of local limit theorem, i.e.

$$\sup_{x\in\mathbb{Z}^d}\Big\{n^\gamma\, q_n(x) - g\big(\frac{x}{n^a}\big)\Big\} \xrightarrow[n\to\infty]{} 0\,, \tag{4.10}$$

for a sufficiently smooth density $g(\cdot)$ and exponents $a,\gamma>0$ †. Then, if $\beta_N := \hat\beta/\sqrt{R_N}$, it holds that

$$Z^{\mathrm{marginal}}_{N,\beta_N} \xrightarrow[N\to\infty]{d} \begin{cases} \exp\big(\sigma_{\hat\beta}\mathsf{X} - \frac12\sigma^2_{\hat\beta}\big) & \text{if } \hat\beta<1\\ 0 & \text{if } \hat\beta\geqslant 1\end{cases}. \tag{4.11}$$

where X is a standard normal variable and $\sigma^2_{\hat\beta} = \log(1-\hat\beta^2)^{-1}$.

We will outline the main ideas of the above theorem below. Before, let us describe the analogue of this theorem for the stochastic heat equation in the following

Theorem 4.2. *Let $j\in C_c^\infty(\mathbb{R}^2)$ be a probability density on $\mathbb{R}^2$ with $j(x)=j(-x)$, and let $J := j*j$. For $\varepsilon>0$, let $j_\varepsilon(x) := \varepsilon^{-2}j(x/\varepsilon)$ and define the mollified noise $\dot W^\varepsilon$ by $\dot W^\varepsilon(t,x) := \int_{\mathbb{R}^2} j_\varepsilon(x-y)\dot W(t,y)\mathrm{d}y$. Then the solution to the regularised SHE*

$$\partial_t u^\varepsilon = \frac12\Delta u^\varepsilon + \beta_\varepsilon u^\varepsilon \dot W^\varepsilon, \qquad u^\varepsilon(0,\cdot)\equiv 1, \tag{4.12}$$

with $\beta_\varepsilon = \hat\beta\sqrt{2\pi/\log\varepsilon^{-1}}$ satisfies the pointwise (i.e. for fixed $t\in\mathbb{R}, x\in\mathbb{R}^2$) distributional limit

$$u^\varepsilon(t,x) \xrightarrow[\varepsilon\to 0]{(d)} \begin{cases} \exp\big(\sigma_{\hat\beta}\mathsf{X} - \frac12\sigma^2_{\hat\beta}\big) & \text{if } \hat\beta<1\\ 0 & \text{if } \hat\beta\geqslant 1\end{cases}. \tag{4.13}$$

where X is a standard normal variable and $\sigma^2_{\hat\beta} = \log(1-\hat\beta^2)^{-1}$.

We remark that the difference between the factor 2π in the above theorem and the factor π in (4.3) is due to the periodicity of the walk. The proof of Theorem 4.2 is given in Section

---

†we see that in order for (4.9) to hold, the exponents a,γ need to satisfy the relation $\gamma = \frac{1+ad}{2}$, which is consistent with Harris criterion for effective dimension $\mathrm{d_{eff}} := 1+ad$. But also notice that in order for the limit of $q_n(\cdot)$ to be a probability density, it is required that $\gamma = ad$, which combined with the relation $\gamma = \frac{1+ad}{2}$, leads to the interpretation of marginality, in terms of the physical dimension as $d = 1/a$

9 of [CSZ17b] and it is based on an approximation of the solution to the SHE, in the form of a Wiener chaos expansion

$$u^{\varepsilon}(t,z) \overset{d}{=} 1+\sum_{k\geqslant 1} \beta_{\varepsilon}^{k} \underset{\substack{0<t_1<\cdots<t_k<\varepsilon^{-2}t \\ \vec{x}\in(\mathbb{R}^2)^k}}{\int\cdots\int} \Bigg(\int_{(\mathbb{R}^2)^k} \prod_{i=1}^{k} g_{t_i-t_{i-1}}(y_i-y_{i-1})\, j(y_i-x_i)\,\mathrm{d}y_i \Bigg) \prod_{i=1}^{k} \xi(t_i,x_i)\mathrm{d}t_i\mathrm{d}x_i\,, \tag{4.14}$$

where $y_0 = \varepsilon^{-1}z$, by the partition function of the directed polymer model. Since this approximation is mostly technical and long but does not bring in any news insights, we will not expose it here.

Before sketching the proof Theorem 4.1, let us remark on the significance of the critical value $\hat{\beta}_{\text{critical}} = 1$. This lies on the fact that for $\hat{\beta} < 1$, the $L^2(\mathbb{P})$ norm of $Z^{\text{marginal}}_{N,\beta_N}$ is uniformly bounded in N, while for $\hat{\beta} \geqslant 1$ it increases to infinity as $N \to \infty$. For $\hat{\beta} = 1$ precise estimates [CSZ18c] show that this $L^2(\mathbb{P})$ norm grows as $\log N$.

Sketch of the proof of Theorem 4.1. We will outline the proof having in mind the case of a directed two dimensional polymer, in which case $q_n(x)$ is the transition probability kernel of a two dimensional simple random and, therefore from (4.3), $R_N \sim \frac{1}{\pi}\log N$.

An important first observation has to do with the correct **time scale** upon which one observes a change in the fluctuations. To determine this time scale, one may look at the partition function of a system of length tN for arbitrary $t > 0$. Computing the variance of Z_{tN,β_N} with β_N as in (4.8), one obtains that it is asymptotically independent of t when $N \to \infty$. One can be easily convinced about this by looking, for example, at the variance of the first term in the chaos expansion (4.5), which behaves as (we denote by $\sigma_N := \sigma(\beta_N)$)

$$\sigma_N^2\, \mathbb{V}\mathrm{ar}\Big(\sum_{\substack{1\leqslant n\leqslant tN \\ x\in\mathbb{Z}^2}} q_n(x)\xi_{n,x} \Big) = \frac{\hat{\beta}^2}{R_N} \sum_{\substack{1\leqslant n\leqslant tN \\ x\in\mathbb{Z}^2}} q_n(x)^2 = \hat{\beta}^2 \frac{R_{tN}}{R_N} \xrightarrow[N\to\infty]{} \hat{\beta}^2,$$

which is independent of t, since R_N is a slowly varying function. Moreover, a similar computation shows that the contribution to the fluctuations from disorder $\xi_{n,x}$ sampled in the time interval $[tN, N]$ is negligible, for any $t > 0$ fixed. On the other hand, one starts seeing a change in the fluctuations when looking at time scales N^t, with $t > 0$. These facts dictate that the meaningful time scale is not tN but N^t and that the partition function Z_{N,β_N} essentially depends only on disorder $\xi_{n,x}$ with $n/N \to 0$, as $N \to \infty$. We remark that this observation will also be important later on when we will try to approximate the KPZ equation via the SHE.

To quantify the observation on the time scale, we decompose the summations over $n_1, ..., n_k$ in the multilinear expansion (4.5), over intervals $n_j - n_{j-1} \in I_{i_j}$, with $I_{i_j} = \big(N^{\frac{i_j-1}{M}}, N^{\frac{i_j}{M}}\big]$, $i_j \in \{1, ..., M\}$ with M being a coarse graining parameter (which will eventually tend to infinity). We can then rewrite the k-th term in the expansion (4.5) as

$$\frac{\hat{\beta}^k}{M^{k/2}} \sum_{1\leqslant i_1,\ldots,i_k\leqslant M} \Theta^{N,M}_{i_1,\ldots,i_k} \qquad \text{where} \tag{4.15}$$

$$\Theta^{N,M}_{i_1,\ldots,i_k} := \Big(\frac{M}{R_N}\Big)^{k/2} \sum_{\substack{n_j-n_{j-1}\in I_{i_j} \text{ for } j=1,\ldots,k \\ x_1,\ldots,x_k\in Z^d}} \prod_{j=1}^{k} q_{n_j-n_{j-1}}(x_j - x_{j-1})\,\xi_{n_j,x_j}.$$

For technical reasons, that should become obvious below, we are led to restrict the summation in (4.15) to the subset

$$
\begin{aligned}
&i_1, \dots, i_k \in \{1, \dots, M\}_\sharp \qquad \text{with} \\
&\{1, \dots, M\}_\sharp := \{\boldsymbol{i} = (i_1, \dots, i_k) \colon |i_j - i_{j'}| \geqslant 2, \text{ for all } j \neq j'\}.
\end{aligned}
$$

It is not difficult to justify this restriction via an $L^2(\mathbb{P})$ but we will omit the details.

We now observe that if an index i_j is a running maximum for the k-tuple $\boldsymbol{i} := (i_1, \dots, i_k)$, i.e. $i_j > \max\{i_1, \dots, i_{j-1}\}$ then $\big(N^{\frac{i_j-1}{M}}, N^{\frac{i_j}{M}}\big] \ni n_j \gg n_r \in \big(N^{\frac{i_r-1}{M}}, N^{\frac{i_r}{M}}\big]$, for all $r < j$, when $N \to \infty$. This is the point where we also use the restriction into $\{1, \dots, M\}_\sharp$. This implies that $q_{n_j - n_{j-1}}(x_j - x_{j-1}) \approx q_{n_j}(x_j)$ for $n_j \in I_{i_j}$ and $n_{j-1} \in I_{i_{j-1}}$, where the drop out of the spatial term x_{j-1} makes use of the diffusive properties of the random walk. Thus, decomposing the sequence $\boldsymbol{i} := (i_1, \dots, i_k)$ according to its running maxima, i.e. $\boldsymbol{i} = (\boldsymbol{i}^{(1)}, \dots, \boldsymbol{i}^{(\mathtt{m}(\boldsymbol{i}))})$ with $\boldsymbol{i}^{(r)} := (i_{\ell_r}, \dots, i_{\ell_{r+1}-1})$ and with $i_1 = i_{\ell_1} < i_{\ell_2} < \cdots < i_{\ell_\mathtt{m}}$ being the successive running maxima, it can be shown that (4.15) asymptotically factorizes for large N as

$$
\frac{\hat\beta^k}{M^{\frac{k}{2}}} \sum_{\boldsymbol{i} \in \{1, \dots, M\}_\sharp^k} \Theta^{N;M}_{\boldsymbol{i}^{(1)}} \Theta^{N;M}_{\boldsymbol{i}^{(2)}} \cdots \Theta^{N;M}_{\boldsymbol{i}^{(\mathtt{m})}} . \tag{4.16}
$$

The heart of the matter is to show that all the $\Theta^{N;M}_{\boldsymbol{i}^{(j)}}$ converge jointly, when $N \to \infty$ to standard normal variables. This is where the fourth moment theorem is used and we will show how this is done in Proposition 4.3 that follows. Assuming this, let us see how we can obtain the convergence to the log-normal distribution in (4.11) when $\hat\beta < 1$.

We can start by replacing, using the Lindeberg principle, the $\Theta^{N;M}_{\boldsymbol{i}}$ variables in (4.16) by standard normals, which we denote by $\zeta_{\boldsymbol{i}}$. Then denoting by

$$
\zeta_r(a) := \sum_{(a, a_2, \dots, a_r) \in \{1, \dots, a-1\}^{r-1}} \zeta_{(a, a_2, \dots, a_r)}
$$

we have that $Z^{\text{marginal}}_{N,\beta_N}$ is approximately (in the large N limit)

$$
Z^{\text{marginal}}_{N,\beta_N} \approx 1 + \sum_{k=1}^{\infty} \sum_{m=1}^{k} \frac{\beta^k}{M^{\frac{k}{2}}} \sum_{\substack{1 \leqslant \ell_1 < \cdots < \ell_m \leqslant k \\ 1 \leqslant a_1 < a_2 < \cdots < a_m \leqslant M}} \prod_{j=1}^{m} \zeta_{\ell_{j+1} - \ell_j}(a_j),
$$

where m denotes the number of running maxima in the sequence $\boldsymbol{i} = (i_1, \dots, i_k)$ in (4.15) (thus determining the number of dominated sequences), $\ell_1, \dots, \ell_m$ denotes the location of the running maxima in $\boldsymbol{i}$ and $a_1, \dots, a_m$ denote the values of $i_{\ell_1}, \dots, i_{\ell_m}$. We can continue by rewriting the above as

$$
\begin{aligned}
&1+\sum_{k=1}^{\infty}\sum_{m=1}^{k}\frac{\hat\beta^k}{M^{\frac{k}{2}}}\sum_{\substack{r_1,\dots r_m\in\mathbb{N}\\ r_1+\cdots+r_m=k}}\;\sum_{1\leqslant a_1<a_2<\cdots<a_m\leqslant M}\;\prod_{j=1}^{m}\zeta_{r_j}(a_j)\\
&= 1+\sum_{m=1}^{\infty}\sum_{r_1,\dots,r_m\in\mathbb{N}}\;\sum_{1\leqslant a_1<a_2<\dots<a_m\leqslant M}\;\prod_{j=1}^{m}\frac{\hat\beta^{r_j}}{M^{\frac{r_j}{2}}}\,\zeta_{r_j}(a_j)\\
&= 1+\sum_{m=1}^{\infty}\sum_{r_1,\dots,r_m\in\mathbb{N}}\;\sum_{\substack{0<t_1<t_2<\dots<t_m\leqslant 1\\ t_1,\dots,t_m\in\frac{1}{M}\mathbb{N}}}\;\prod_{j=1}^{m}\frac{\hat\beta^{r_j}}{M^{\frac{r_j}{2}}}\,\zeta_{r_j}(Mt_j)\\
&= 1+\sum_{m=1}^{\infty}\sum_{\substack{0<t_1<t_2<\dots<t_m\leqslant 1\\ t_1,\dots,t_m\in\frac{1}{M}\mathbb{N}}}\;\prod_{j=1}^{m}\Big\{\sum_{r\in\mathbb{N}}\frac{\hat\beta^{r}}{M^{\frac{r}{2}}}\,\zeta_{r}(Mt_j)\Big\},
\end{aligned}
\tag{4.17}
$$

In the first equality we interchanged the summations over m and k and for this the assumption $\hat\beta<1$, that ensures convergence in $L^2(\mathbb{P})$ is crucially used.

Since $(\hat\beta/\sqrt{M})^r\zeta_r(Mt)$ are normal random variables, independent for different values of $r\in\mathbb{N}$ and $t\in M^{-1}\mathbb{N}$, we have that the random variables

$$
\Xi_{M,t}:=\sum_{r\in\mathbb{N}}\frac{\hat\beta^{r}}{M^{\frac{r}{2}}}\,\zeta_r(Mt),\qquad t\in(0,1]\cap\frac{1}{M}\mathbb{N},
$$

are also independent normal with mean zero and variance

$$
\mathbb{V}\mathrm{ar}(\Xi_{M,t})=\sum_{r\in\mathbb{N}}\frac{\hat\beta^{2r}}{M^r}\mathbb{V}\mathrm{ar}(\xi_r(Mt))=\sum_{r\in\mathbb{N}}\frac{\hat\beta^{2r}}{M^r}(Mt-1)^{r-1}=\frac{\hat\beta^2}{M}\cdot\frac{1+\varepsilon_M(t)}{1-\hat\beta^2 t},
$$

with the error $\varepsilon_M(t)$ easily seen to converge to 0, uniformly in $t\in[0,1]$, as $M\to\infty$ for $\hat\beta<1$. We can, therefore, represent $\Xi_{M,t}$ in terms of a standard, one dimensional Wiener process W:

$$
\Xi_{M,t}=\frac{\hat\beta(1+\varepsilon_M(t))}{\sqrt{1-\hat\beta^2\,t}}\int_{t-\frac{1}{M}}^{t}\mathrm{d}W_s,\qquad t\in[0,1]\cap\frac{1}{M}\mathbb{N}.
\tag{4.18}
$$

and we can rewrite (4.17) as

$$
1+\sum_{m=1}^{\infty}\sum_{\substack{0<t_1<t_2<\dots<t_m\leqslant 1\\ t_1,\dots,t_m\in\frac{1}{M}\mathbb{N}}}\;\prod_{j=1}^{m}\frac{\hat\beta(1+\varepsilon_M(t))}{\sqrt{1-\hat\beta^2\,t_j}}\int_{t_j-\frac{1}{M}}^{t_j}\mathrm{d}W_s.
\tag{4.19}
$$

So, for $\hat\beta<1$, we have that (4.18) converges in $L^2(\mathbb{P})$, for $M\to\infty$, to

$$
1+\sum_{m=1}^{\infty}\int\!\cdots\!\int_{0<t_1<\dots<t_m<1}\prod_{j=1}^{m}\frac{\hat\beta}{\sqrt{1-\hat\beta^2\,t_j}}\mathrm{d}W_{t_j}=\exp\Big\{\int_0^1\frac{\hat\beta}{\sqrt{1-\hat\beta^2\,t}}\mathrm{d}W(t)-\frac{1}{2}\mathbb{E}\Big[\Big(\int_0^1\frac{\hat\beta}{\sqrt{1-\hat\beta^2\,t}}\mathrm{d}W(t)\Big)^2\Big]\Big\},
$$

where the last equality holds by the properties of the Wick exponential [J97, §3.2]. Since $\int_0^1\frac{\hat\beta}{\sqrt{1-\hat\beta^2\,t}}\mathrm{d}W(t)$ is a gaussian variable with variance $\int_0^1\frac{\hat\beta^2\,\mathrm{d}t}{1-\hat\beta^2}=\log(1-\hat\beta^2)$, the result follows.

This concludes the proof of the log-normality in the subcritical regime $\hat\beta < 1$. The proof that at the critical temperature $\hat\beta = 1$ the limit of $Z^{\rm marginal}_{N,\beta_N}$ is zero makes use of the convergence for $\hat\beta < 1$ and goes via a fractional moment computation. For $\vartheta < 1$,

$$\lim_{N\to\infty} \mathbb{E}[(Z^{\rm marginal}_{N,\beta_N})^\vartheta] = \mathbb{E}\Big[\exp\Big(\vartheta\int_0^1 \frac{\hat\beta}{\sqrt{1-\hat\beta^2\,t}}{\rm d}W(t) - \frac{\vartheta}{2}\int_0^1 \frac{\hat\beta^2}{1-\hat\beta^2\,t}{\rm d}t\Big)\Big]$$
$$= \exp\Big(\frac{\vartheta(\vartheta-1)}{2}\int_0^1 \frac{\hat\beta^2}{1-\hat\beta^2\,t}{\rm d}t\Big) = \big(1-\hat\beta^2\big)^{-\frac{\vartheta(\vartheta-1)}{2}},$$

which, since $\vartheta < 1$, goes to zero as $\hat\beta \nearrow 1$. The proof is completed by a monotonicity property of $\beta \mapsto \mathbb{E}\big[\big(Z^{\rm marginal}_{N,\beta}\big)^\vartheta\big]$, which allows for an interchange of the limits in N and $\hat\beta$ in the above computation; for details see [CSZ17b, Theorem 2.8 and its proof]. ☐

We will now prove the core proposition, which outlines how $\Theta^{N;M}_{\boldsymbol{i}}$ converges to a normal variable for just a single $\boldsymbol{i}$. The joint convergence, that is required in the previous theorem, follows easily from the computation that we will outline and the general statement of Theorem 3.3.

Proposition 4.3 (Dominated sequences). *Let $\Theta^{N,M}_{\boldsymbol{i}}$ be defined as in (4.15) with the kernel $q_n(x)$ satisfying the assumption of marginal relevance as in Theorem 4.1 and $\boldsymbol{i} = (i_1,...,i_k)$, for some $k \geqslant 1$, being a dominated sequence, that is $i_1 > i_2,...,i_k$. Then, for every fixed M, $\Theta^{N,M}_{\boldsymbol{i}}$ converges to a standard normal as $N \to \infty$.*

Proof. The proof uses the Fourth Moment Theorem. Let $\boldsymbol{i} = (i_1,...,i_k)$ be a dominated sequence and compute

$$\mathbb{E}\Big[\big(\Theta^{N,M}_{\boldsymbol{i}}\big)^4\Big] = \Big(\frac{M}{R_N}\Big)^{2k} \sum_{\substack{a_j-a_{j-1}\in I_{i_j}\\ x_1,...,x_k}} \sum_{\substack{b_j-b_{j-1}\in I_{i_j}\\ y_1,...,y_k}} \sum_{\substack{c_j-c_{j-1}\in I_{i_j}\\ z_1,...,z_k}} \sum_{\substack{d_j-d_{j-1}\in I_{i_j}\\ w_1,...,w_k}} \mathbb{E}\Big[\prod_{j=1}^k \xi_{a_j,x_j}\xi_{b_j,y_j}\xi_{c_j,z_j}\xi_{d_j,w_j}\Big]$$
$$\times \prod_{j=1}^k q_{a_j-a_{j-1}}(x_j-x_{j-1})\, q_{b_j-b_{j-1}}(y_j-y_{j-1})\, q_{c_j-c_{j-1}}(z_j-z_{j-1})\, q_{d_j-d_{j-1}}(w_j-w_{j-1}). \tag{4.20}$$

By the Lindeberg principle, Theorem 3.2, we may assume that the random variables $\xi_{n,x}$ are standard normals and therefore the expectation that appears inside the sum above will be zero unless the variables $\big(\xi_{a_j,x_j},\xi_{b_j,y_j},\xi_{c_j,z_j},\xi_{d_j,w_j} : j = 1,...,k\big)$ match in pairs or in quadruples.

We will show that there is only one case that contributes to the asymptotic behaviour when $N \to \infty$, which is when either $(a_j,x_j) = (b_j,y_j)$ and $(c_j,z_j) = (d_j,w_j)$ for all j or $(a_j,x_j) = (c_j,z_j)$ and $(b_j,y_j) = (d_j,w_j)$ for all j or $(a_j,x_j) = (d_j,w_j)$ and $(c_j,z_j) = (b_j,y_j)$ for all j. The restriction to these *three* possibilities is what gives that the limit of the fourth moment of $\Theta^{N,M}_{\boldsymbol{i}}$ converges to three - we also need to notice that the contribution of each term is 1. To see this last point, look, for example, at the case $(a_j,x_j) = (b_j,y_j)$. The corresponding term in the right hand side of (4.20) equals (we also use that $\mathbb{E}[\xi^2] = 1$)

$$\Big(\frac{M}{R_N}\Big)^{2k} \sum_{\substack{a_j - a_{j-1} \in I_{i_j} \\ x_1,\dots,x_k}} \sum_{\substack{c_j - c_{j-1} \in I_{i_j} \\ z_1,\dots,z_k}} \prod_{j=1}^{k} q_{a_j - a_{j-1}}(x_j - x_{j-1})^2 \, q_{c_j - c_{j-1}}(z_j - z_{j-1})^2$$

$$= \Big(\frac{M}{R_N}\Big)^{2k} \Big(\sum_{\substack{a_j - a_{j-1} \in I_{i_j} \\ x_1,\dots,x_k}} \prod_{j=1}^{k} q_{a_j - a_{j-1}}(x_j - x_{j-1})^2 \Big)^2. \tag{4.21}$$

Having in mind the case of the two dimensional directed polymer, that we are concerned with here, in which case $R_N \sim \frac{1}{\pi} \log N$, we notice that

$$\frac{M}{R_N} \sum_{\substack{a_j - a_{j-1} \in I_{i_j} \\ x_k \in \mathbb{Z}^2}} q_{a_j - a_{j-1}}(x_j - x_{j-1})^2 = \frac{M}{R_N} (R_{N^{i_j/M}} - R_{N^{(i_j-1)/M}}) \approx 1. \tag{4.22}$$

Let us now describe how the rest of the possible matching cases lead to negligible contribution. For this, let us label the elements of the set $\{(a_j, x_j), (b_j, y_j), (c_j, z_j), (d_j, w_j) \colon j = 1, \dots, k\}$ as $\{(f_1, \chi_1), \dots, (f_p, \chi_p)\}$ with $p \leqslant 2k$ denoting the cardinality of the set. The first case to exclude is the case where a quadruple matching exists.

Quadruple matchings. In this case $p < 2k$. Every time we sum a double matching, we will have a sum of the form

$$\sum_{\substack{a_r \in a_{r-1} + I_{i_r} \\ b_m \in b_{m-1} + I_{i_m} \\ x_r = y_m \in \mathbb{Z}^2 \text{ and } a_r = b_m}} q_{a_r - a_{r-1}}(x_r - x_{r-1}) q_{b_m - b_{m-1}}(y_m - y_{m-1}), \tag{4.23}$$

which, by Cauchy-Schwarz and the computations in (4.21), (4.22), is easily seen to be bounded by $(R_N/M)^2$.

On the other hand, when a quadruple matching occurs, we have a sum of the form

$$\sum_{\substack{a_r \in a_{r-1} + I_{i_r}, \, b_m \in b_{m-1} + I_{i_m} \\ c_u \in c_{u-1} + I_{i_u}, \, d_v \in d_{v-1} + I_{i_v} \\ x_r = y_m = z_u = w_v \in \mathbb{Z}^2 \\ a_r = b_m = c_u = d_v}} q_{a_r - a_{r-1}}(x_r - x_{r-1}) q_{b_m - b_{m-1}}(y_m - y_{m-1}) \times q_{c_u - c_{u-1}}(z_u - z_{u-1}) q_{d_v - d_{v-1}}(w_v - w_{v-1}),$$

and bounding the last two kernels by 1 we come back to the same sum as in (4.23) and, thus, the contribution of this quadruple summation is also bounded by R_N/M. Performing successively all summations of $\{(f_1, \chi_1), \dots, (f_p, \chi_p)\}$ we obtain a bound of order $(R_N/M)^p$, which, since $p < 2k$, is dominated by the factor $(M/R_N)^{2k}$ in (4.20) (recall that $R_N \to \infty$).

Mixed pairwise matchings. This case is the central one and here is where one sees the significance of the logarithmic (or in general slowly varying) growth of the overlap R_N, i.e. *marginal relevance*, as well as the role of the *dominated sequence*.

The situation here is that we only have pairwise matchings but the labels mix. For example, we may have that $(f_1, \chi_1) = (a_1, x_1) = (b_1, y_1)$ but then the (a, x) label does not continue to match with the (b, y) label but, for example, $(f_2, \chi_2) = (a_2, x_2) = (c_1, z_1)$. In this case, let us look at the normalised sum

$$\frac{M}{R_N} \sum_{\substack{a_2 \in a_1+I_{i_2}\,,\, c_1 \in I_{i_1} \\ x_2=z_1 \in \mathbb{Z}^2 \text{ and } a_2=c_1}} q_{a_2-a_1}(x_2-x_1) q_{c_1}(z_1).$$

We notice that the matching $a_2 = c_1$ constraints the range of c_1 from its original set I_{i_1} to $a_1 + I_{i_2}$. Thus, via Cauchy-Schwarz we bound this by

$$\frac{M}{R_N}\Big(\sum_{a_2 \in a_1+I_{i_2}\,,\, x_2 \in \mathbb{Z}^2} q_{a_2-a_1}(x_2-x_1)^2\Big)^{1/2}\Big(\sum_{c_1 \in a_1+I_{i_2}\,,\, z_1 \in \mathbb{Z}^2} q_{c_1}(z_1)^2\Big)^{1/2}.$$

The first sum is bounded by R_N/M but the second one is

$$\sum_{c_1 \in a_1+I_{i_2}\,,\, z_1 \in \mathbb{Z}^2} q_{c_1}(z_1)^2 = \sum_{c_1 \in a_1+I_{i_2}} q_{2c_1}(0)^2 \approx \log(a_1 + N^{i_2/M}) - \log(a_1 + N^{(i_2-1)/M})$$

and since $a_1 \in I_{i_1}$ it is of order $N^{i_1/M} >> N^{i_2/M}$, the above difference converges to zero as N tends to infinity.

The more general mix-and-match-labels case follows the same route. □

Theorem 4.1 describes the asymptotics of a *single* partition function of a marginally relevant model when the starting point of the polymer path is fixed. We can also ask about the asymptotics of the joint laws of $\big(Z^{\text{marginal}}_{N,\beta_N}(x)\big)_{x\in\mathbb{Z}^d}$. This will be described by the next theorem, which, for simplicity in terms of notation, we only state in the case of the standard two dimensional directed polymer model.

Theorem 4.4 (Partition function of $2d$ polymer as a field). *Let $Z_{N,\beta_N}(x)$ be the partition function of a directed polymer corresponding to a two dimensional, simple random walk. Let β_N be chosen as*

$$\beta_N := \frac{\hat\beta}{\sqrt{R_N}}, \quad \text{with} \quad \hat\beta < 1 \quad \text{and} \quad R_N := \sum_{n\leqslant N\,,\, x\in Z^2} q_n(x)^2 = \frac{\log N}{\pi}(1+o(1)). \tag{4.24}$$

Let also $\phi \in C_b(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ be a test function. Then

$$Z_{N,\beta_N}(\phi) := \frac{\sqrt{R_N}}{N} \sum_{x\in\mathbb{Z}^d} \big(Z_{N,\beta_N}(x) - 1\big)\phi\big(\tfrac{x}{\sqrt{N}}\big)$$

converges to a gaussian variable with mean zero and variance

$$\sigma^2_{\hat\beta,\phi} = \frac{\hat\beta^2}{1-\hat\beta^2}\int_{R^2}\int_{R^2} \phi(x)K(x,y)\phi(y)\,\mathrm{d}x\mathrm{d}y, \qquad \text{with} \qquad K(x,y) = \int_0^1 \frac{1}{2\pi t} e^{-\frac{|x|^2}{2t}}\,\mathrm{d}t.$$

Proof. The proof of this theorem follows similar lines as that of Theorem 4.1 and also makes crucial use of Proposition 4.3 on the asymptotic normality of dominated sequences. However, there is a key difference with Theorem 4.1, where we saw, via variance estimates, that the main contribution to fluctuations of the partition function with fixed starting point comes from disorder $\xi_{n,x}$ with $n = N^t$ for $t < 1$. Here the fact that we average at spatial scales $\sqrt{N}$ will make those contributions to the fluctuations of $Z_{N,\beta_N}(\phi)$ negligible. Thus, here, the noise that will drive the fluctuations is $\xi_{n,x}$ with n of order N †. This can be seen

†This qualitative difference will also play a crucial role in the KPZ fluctuations and the approximation of the KPZ by the field of partition functions

again via a variance computation: starting from the chaos expansion

$$Z_{N,\beta_N}(\phi) = \frac{\sqrt{R_N}}{N} \sum_{k=1}^{N} \beta_N^k \sum_{\substack{1 \leqslant n_1 < \cdots < n_k \leqslant N \\ x_0, x_1, \ldots, x_k \in \mathbb{Z}^d}} \phi\big(\tfrac{x_0}{\sqrt{N}}\big) \prod_{i=1}^{k} q_{n_i - n_{i-1}}(x_i - x_{i-1})\ \xi_{n_i, x_i},$$

the variance is computed as

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}\,\big(Z_{N,\beta_N}(\phi)\big) \\
&= \frac{R_N}{N^2} \sum_{k=1}^{N} \beta_N^{2k} \sum_{\substack{1 \leqslant n_1 < \cdots < n_k \leqslant N \\ x_0, \tilde{x}_0, x_1, \ldots, x_k \in \mathbb{Z}^d}} \phi\big(\tfrac{x_0}{\sqrt{N}}\big)\phi\big(\tfrac{\tilde{x}_0}{\sqrt{N}}\big)\, q_{n_1}(x_1 - x_0) q_{n_1}(x_1 - \tilde{x}_0) \prod_{i=2}^{k} q_{n_i - n_{i-1}}(x_i - x_{i-1})^2.
\end{aligned}$$

Summing successively over the variables $(n_k, x_k), (n_{k-1}, x_{k-1}), \ldots, (n_2, x_2)$ and using (4.24) will produce a factor of R_N^{k-1}, which will be cancelled by $k-1$ powers of β_N^2. There remains one more power of β_N^2 which will then cancel the prefactor R_N (recall that $\beta_N = \hat{\beta}/\sqrt{R_N}$). Thus, we have that the variance is approximately, for large N, equal to

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\,\big(Z_{N,\beta_N}(\phi)\big) &\approx \frac{1}{N^2} \sum_{k=1}^{\infty} \hat{\beta}^{2k} \sum_{1 \leqslant n_1 \leqslant N} \sum_{x_0, \tilde{x}_0, x_1} \phi\big(\tfrac{x_0}{\sqrt{N}}\big)\phi\big(\tfrac{\tilde{x}_0}{\sqrt{N}}\big) q_{n_1}(x_1 - x_0) q_{n_1}(x_1 - \tilde{x}_0) \\
&= \frac{\hat{\beta}^2}{1 - \hat{\beta}^2} \frac{1}{N^2} \sum_{1 \leqslant n_1 \leqslant N} \sum_{x_0, \tilde{x}_0} \phi\big(\tfrac{x_0}{\sqrt{N}}\big)\phi\big(\tfrac{\tilde{x}_0}{\sqrt{N}}\big) q_{2n_1}(x_0 - \tilde{x}_0),
\end{aligned}$$

where in the second equality we performed the geometric summation over the $\hat{\beta}^{2k}$ and we summed over x_1 using the convolution property of the random walk. Using the local limit theorem, we see that the variance is approximately

$$\mathbb{V}\mathrm{ar}\,\big(Z_{N,\beta_N}(\phi)\big) \approx \frac{\hat{\beta}^2}{1 - \hat{\beta}^2} \frac{1}{N} \sum_{1 \leqslant n_1 \leqslant N} \frac{1}{N^2} \sum_{x_0, \tilde{x}_0} \phi\big(\tfrac{x_0}{\sqrt{N}}\big)\phi\big(\tfrac{\tilde{x}_0}{\sqrt{N}}\big) \cdot \frac{1}{\pi \frac{n_1}{N}} \exp\Big(-\frac{|x_0 - \tilde{x}_0|^2/N}{2n_1/N}\Big).$$

The Riemann sum approximation shows immediately that contributions from $n_1 = o(N)$ are negligible. Moreover, it also shows that the limiting variance is

$$\sigma^2_{\hat{\beta},\phi} = \frac{\hat{\beta}^2}{1 - \hat{\beta}^2} \int_{R^2} \int_{R^2} \phi(x) K(x,y) \phi(y)\, \mathrm{d}x \mathrm{d}y, \qquad \text{with} \qquad K(x,y) = \int_0^1 \frac{1}{2\pi t} e^{-\frac{|x|^2}{2t}}\, \mathrm{d}t.$$

as claimed in the statement of the theorem.

Having made this crucial observation, the proof of the theorem proceeds as the proof of Theorem 4.1 by first coarse graining the temporal variables so that $n_j - n_{j-1} \in I_{i_j}$ with $I_j = \big(N^{\frac{i_j - 1}{M}}, N^{\frac{i_j - 1}{M}}\big]$ as

$$Z_{N,\beta_N}(\phi) = \sum_{k \geqslant 1} \frac{\hat{\beta}^k}{M^{k/2}} \sum_{1 \leqslant i_1, \ldots, i_k \leqslant M} \Theta^{N,M,\phi}_{i_1,\ldots,i_k} \qquad \text{where} \tag{4.25}$$

$$\Theta^{N,M,\phi}_{i_1,\ldots,i_k} := \Big(\frac{M}{R_N}\Big)^{k/2} \sum_{\substack{n_j - n_{j-1} \in I_{i_j} \text{ for } j = 1, \ldots, k \\ x_0, x_1, \ldots, x_k \in \mathbb{Z}^2}} \phi\big(\tfrac{x_0}{\sqrt{N}}\big) \prod_{j=1}^{k} q_{n_j - n_{j-1}}(x_j - x_{j-1})\, \xi_{n_j, x_j}.$$

But now the crucial observation on the time scale comes into place, imposing that the main contribution is when n_1 is of order N, thus forcing $i_1 = 1$. Therefore, the decomposition of $Z_{N,\beta_N}(\phi)$ into dominated subsequence will consist of only *one* dominated subsequence, that of (4.25) with $i_1 = 1$ as (opposed to the decomposition of Z_{N,β_N} for fixed initial point as in (4.16)). A slight modification of Proposition 4.3 shows that $\Theta^{N,M,\phi}_{i_1,\dots,i_k}$ with $i_1 = 1$ is asymptotically normal variable and thus $Z_{N,\beta_N}(\phi)$ as in (4.25) is a sum of asymptotically uncorrelated, gaussian variables. This now easily leads to the conclusion.

□

5. The two dimensional KPZ in the subcritical regime

With respect to the KPZ equation, Theorem 4.1 translates, via the Hopf-Cole transformation $h^\varepsilon = \log u^\varepsilon$, to

$$h^\varepsilon(t,x) \xrightarrow[\varepsilon\downarrow 0]{d} \begin{cases} \sigma_{\hat\beta}\mathsf{X} - \frac12\sigma^2_{\hat\beta} & \text{if } \hat\beta < 1 \\ -\infty & \text{if } \hat\beta \geqslant 1 \end{cases} \qquad \text{with} \quad \sigma^2_{\hat\beta} := \log\tfrac{1}{1-\hat\beta^2}\,, \quad \mathsf{X} \sim N(0,1)\,, \tag{5.1}$$

which indicates a phase transition at $\hat\beta = 1$. In this section we will describe that, when viewed as a field, the solution to the KPZ in dimension two falls into the Edwards-Wikinson class in the subcritical regime $\hat\beta < 1$. In particular, we have the following theorem

Theorem 5.1 (Edwards-Wilkinson fluctuations for 2-dimensional KPZ - [CSZ18b]**).** *Let h^ε be the solution of the mollified KPZ equation* (1.11) *with initial condition $h^\varepsilon(0,x) \equiv 1$ and with $\beta_\varepsilon = \hat\beta\sqrt{2\pi/\log\varepsilon^{-1}}$ and $\hat\beta \in (0,1)$. Denote*

$$\mathfrak{h}^\varepsilon(t,x) := \frac{h^\varepsilon(t,x) - \mathbb{E}[h^\varepsilon(t,x)]}{\beta_\varepsilon} = \frac{\sqrt{\log\varepsilon^{-1}}}{\sqrt{2\pi}\,\hat\beta}\big(h^\varepsilon(t,x) - \mathbb{E}[h^\varepsilon(t,x)]\big)\,, \tag{5.2}$$

where the centering satisfies $\mathbb{E}[h^\varepsilon(t,x)] = -\frac12\sigma^2_{\hat\beta} + o(1)$ as $\varepsilon \downarrow 0$, see (5.1)*.*

For any $t > 0$ and $\phi \in C_c(\mathbb{R}^2)$, the following convergence in law holds:

$$\langle \mathfrak{h}^\varepsilon(t,\cdot), \phi(\cdot)\rangle = \int_{\mathbb{R}^2} \mathfrak{h}^\varepsilon(t,x)\phi(x)\mathrm{d}x \xrightarrow[\varepsilon\downarrow 0]{d} \langle v^{(c_{\hat\beta})}(t,\cdot), \phi(\cdot)\rangle, \tag{5.3}$$

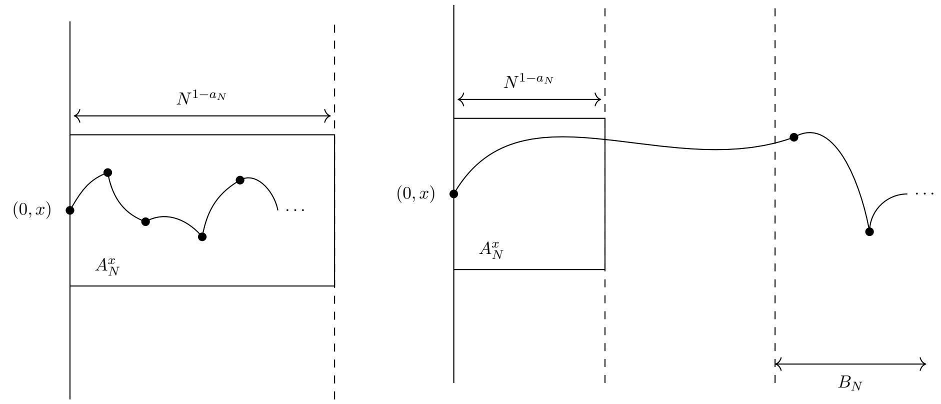
where $v^{(c)}(t,x)$ is the solution of the two-dimensional additive stochastic heat equation

$$\begin{cases} \partial_t v^{(c)}(t,x) = \dfrac12 \Delta v^{(c)}(t,x) + c\,\xi(t,x) \\ \;\; v^{(c)}(0,x) \equiv 0 \end{cases}, \qquad \textit{where} \;\; c := c_{\hat\beta} := \sqrt{\tfrac{1}{1-\hat\beta^2}}\,. \tag{5.4}$$

In terms of the two dimensional polymer model, the above theorem writes as in the following theorem. This will be the theorem whose proof we will outline. The proof of Theorem 5.1 follows exactly the same lines if instead of working with the polynomial chaos expansion of the partition function we work with the Wiener chaos expansion of u^ε as in (4.14). We refer to [CSZ18b, Section 5] for details.

We have

Theorem 5.2 (Edwards-Wilkinson fluctuations for directed polymer - [CSZ18b]**).** *Let $Z_{N,\beta_N}(x)$ be the family of partition functions defined as in* (4.4) *with β_N as in* (4.8) *and $\hat\beta \in (0,1)$. The disorder ω satisfies the usual assumptions of mean zero, variance one and finite exponential moments and in addition we require it to satisfy a concentration property*



(A) Partition fuction $Z^A_{N,\beta_N}(x)$. (B) Partition function $Z^B_{N,\beta_N}(x)$.

FIGURE 1. The above figures depict the chaos expansions of $Z^A_{N,\beta_N}(x)$ and $Z^B_{N,\beta_N}(x)$. The disorder sampled by $Z^A_{N,\beta_N}(x)$ is restricted to the set A^x_N, while that of $Z^B_{N,\beta_N}(x)$ is restricted to B_N

†:

$$\exists \gamma > 1, C_1, C_2 \in (0,\infty) : \text{ for all } n \in \mathbb{N} \text{ and } f : \mathbb{R}^n \to \mathbb{R} \text{ convex and 1-Lipschitz}$$
$$\mathbb{P}\Big(\big|f(\omega_1,\ldots,\omega_N) - M_f\big| \geqslant t\Big) \leqslant C_1 \exp\Big(-\frac{t^\gamma}{C_2}\Big), \tag{5.5}$$

where M_f denotes a median of $f(\omega_1,\ldots,\omega_N)$. Denote

$$\mathfrak{h}_N(t,x) := \frac{\log Z_{tN}(x\sqrt{N}) - \mathbb{E}[\log Z_{tN}]}{\beta_N} = \frac{\sqrt{\log N}}{\sqrt{\pi}\,\hat\beta}\big(\log Z_{tN}(x\sqrt{N}) - \mathbb{E}[\log Z_{tN}]\big). \tag{5.6}$$

For any $t > 0$ and $\phi \in C_c^\infty(\mathbb{R}^2)$, the following convergence in law holds, with $c_{\hat\beta}$ as in (5.4)*:*

$$\langle \mathfrak{h}_N(t,\cdot), \phi(\cdot)\rangle = \int_{\mathbb{R}^2} \mathfrak{h}_N(t,x)\,\phi(x)\,\mathrm{d}x \xrightarrow[N\to\infty]{d} \langle v^{(\sqrt{2}c_{\hat\beta})}(t/2,\cdot), \phi(\cdot)\rangle, \tag{5.7}$$

where $v^{(c)}(s,x)$ is the solution of the two-dimensional additive SHE as in (5.4).

Outline of the proof of Theorem 5.2. The main idea is to try to "linearize" the logarithm of the partition function. The way to achieve this is guided by the observation (see the discussion at the beginning of the proof of Theorem 4.1) that the main contribution to the fluctuations of the partition function $Z_{N,\beta_N}(x)$ comes from disorder $\xi_{n,x}$ with $n = o(N)$ and in particular with $n = N^t$ for $t < 1$. This leads us to define the set

$$A^x_N := \Big\{(n,z) \in \mathbb{N} \times \mathbb{Z}^2 \;:\; n \leqslant N^{1-a_N}\,,\, |z - x| < N^{\frac{1}{2}-\frac{a_N}{4}}\Big\}, \tag{5.8}$$

where

$$a_N = \frac{1}{(\log N)^{1-\gamma}} \qquad \text{with} \qquad \gamma \in (0,\gamma^*), \tag{5.9}$$

†Condition (5.5) is satisfied if ω are bounded, Gaussian, or if they have a density $\exp(-V(\cdot) + U(\cdot))$, with V uniformly strictly convex and U bounded. We refer to [Led01] for more details.

for some $\gamma^* > 0$ depending only on $\hat{\beta}$. The precise choice of γ^* is more of a technical nature and we will not bother with it here; one can refer for details in [CSZ18b]. The spatial coordinates of the set A_N^x are essentially restricted to a slightly superdiffusive window in order to make sure that the random walk path stays, with high probability, within this box during the corresponding time scale.

We define now the partition function $Z_{N,\beta}^A(x)$ which only samples disorder in A_N^x, i.e.

$$Z_{N,\beta_N}^A(x) := \mathrm{E}_x\big[e^{H_{A_N^x,\beta_N}}\big], \qquad \text{where} \qquad H_{A_N^x,\beta_N} := \sum_{(n,x)\in A_N^x} (\beta_N \omega_{n,x} - \lambda(\beta_N))\mathbb{1}_{\{S_n=x\}}\,. \tag{5.10}$$

This allows to decompose the original partition function $Z_{N,\beta_N}(x)$ as follows:

$$Z_{N,\beta_N}(x) = Z_{N,\beta_N}^A(x) + \hat{Z}_{N,\beta_N}^A(x), \tag{5.11}$$

where $\hat{Z}_{N,\beta_N}^A(x)$ is defined via this relation as the "remainder". Since, as we mentioned, $Z_{N,\beta_N}^A(x)$ captures the main contribution in Z_{N,β_N}, we expect that for any *fixed* x, $\hat{Z}_{N,\beta_N}^A(x) \ll Z_{N,\beta_N}^A(x)$ in a suitable sense. In particular, an $L^2(\mathbb{P})$ estimate shows that

$$\forall \hat{\beta} \in (0,1)\ \exists\, C_{\hat{\beta}} < \infty \ \text{ such that } \ \forall N \in \mathbb{N}: \qquad \mathbb{E}[\hat{Z}_{N,\beta_N}^A(x)^2] \leqslant C_{\hat{\beta}}\, a_N\,. \tag{5.12}$$

with a_n as defined in (5.9). The proof of this estimate is not difficult but it is, nevertheless, a bit technical and it can be found in Section 3.4 of [CSZ18b]. We now have the approximation

$$\log Z_{N,\beta_N}(x) = \log Z_{N,\beta_N}^A(x) + \log\Big(1 + \frac{\hat{Z}_{N,\beta_N}^A(x)}{Z_{N,\beta_N}^A(x)}\Big) \approx \log Z_{N,\beta_N}^A(x) + \frac{\hat{Z}_{N,\beta_N}^A(x)}{Z_{N,\beta_N}^A(x)}. \tag{5.13}$$

This approximation is quantified via the following estimate:

Estimate 1. Define the error $O_N(x)$ via

$$\log Z_{N,\beta_N}(x) = \log Z_{N,\beta_N}^A(x) + \frac{\hat{Z}_{N,\beta_N}^A(x)}{Z_{N,\beta_N}^A(x)} + O_N(x). \tag{5.14}$$

Then for every suitable test function $\phi(\cdot)$ we have that

$$\sqrt{\log N}\,\frac{1}{N}\sum_{x\in\mathbb{Z}^2}\big(O_N(x) - \mathbb{E}[O_N(x)]\big)\,\phi(\tfrac{x}{\sqrt{N}}) \xrightarrow[N\to\infty]{L^2(\mathbb{P})} 0\,. \tag{5.15}$$

The proof of this estimate uses a simple Taylor expansion estimate, which says that, essentially, the error term $O_N(z)$ is bounded by $\big(\frac{\hat{Z}_{N,\beta_N}^A(x)}{Z_{N,\beta_N}^A(x)}\big)^2$. In order to estimate this error, one needs to use Hölder inequality, in order to separate the numerator and denominator as $\mathbb{E}\Big[\big(\hat{Z}_{N,\beta_N}^A(x)\big)^{2p}\Big]^{1/p}\cdot\mathbb{E}\Big[\big(Z_{N,\beta_N}^A(x)\big)^{-2q}\Big]^{1/q}$. The estimate on the first expectation makes use of the Hypercontractivity, Theorem 3.5:

$$\begin{aligned}\mathbb{E}\Big[\big(\hat{Z}_{N,\beta_N}^A(x)\big)^{2p}\Big] &\leqslant 1 + \sum_{k\geqslant 1}^{N}\big(\varrho_{2p}\,\beta_N^2\big)^k \sum_{\substack{1\leqslant n_1<\cdots<n_k\leqslant N\\ x_1,\ldots,x_k\in\mathbb{Z}^d\\ \exists j\in\{1,\ldots,k\}:\,(n_j,x_j)\notin A}} \prod_{i=1}^{k} q_{n_i-n_{i-1}}(x_i-x_{i-1})^2 \qquad (5.16)\\ &\leqslant 1 + \sum_{k\geqslant 1}^{N}\big(\varrho_{2p}\,\beta_N^2\big)^k \sum_{\substack{1\leqslant n_1<\cdots<n_k\leqslant N\\ x_1,\ldots,x_k\in\mathbb{Z}^d}} \prod_{i=1}^{k} q_{n_i-n_{i-1}}(x_i-x_{i-1})^2,\end{aligned}$$

where ϱ_{2p} is the hypercontractivity constant. The significance of the estimate $\lim_{p\to 1} \varrho_{2p} = 1$, that was proved in Theorem 3.5, is that by choosing p sufficiently close to 1, we have that $\varrho_p \beta_N^2 = \varrho_p \hat{\beta}^2 \sqrt{\pi/\log N}$, and $\varrho_p \hat{\beta}^2$ is still less than 1. Thus the right hand side is finite. Moreover, feeding the first line of (5.16) into (5.12) will eventually show that $\mathbb{E}\Big[\big(\hat{Z}^A_{N,\beta_N}(x)\big)^{2p}\Big]$ is sufficiently small when p is sufficiently close to 1, so that (5.15) holds.

We should remark that choosing p close to 1 has the consequence that q is made very large But this is still fine in terms of estimating $\mathbb{E}\Big[\big(Z^A_{N,\beta_N}(x)\big)^{-2q}\Big]$ since all negative moments of the partition function can be shown to exist. In particular, we have (see [CSZ18b, Poposition 3.1]) that

Negative tails. For any $\hat{\beta} \in (0,1)$, there exists $c_{\hat{\beta}} \in (0,\infty)$ with the following property: for every $N \in \mathbb{N}$ and for every choice of $\Lambda \subseteq \{1,\dots,N\} \times \mathbb{Z}^2$, one has

$$\forall t \geqslant 0: \qquad \mathbb{P}(\log Z_{\Lambda,\beta_N} \leqslant -t) \leqslant c_{\hat{\beta}}\, e^{-t^\gamma / c_{\hat{\beta}}}, \tag{5.17}$$

where $\gamma > 1$ is the same exponent appearing in assumption (5.5).

The proof of the negative moment tails makes use of an interesting concentration estimate, which is of general interest. This was proved in [CTT17, Proposition 3.4], inspired by [Led01, Proposition 1.6].

Proposition 5.3 (Concentration estimate, [CTT17, Led01]**).** *Assume that disorder ω has the concentration property* (5.5)*. There exist constants $c_1, c_2 \in (0,\infty)$ such that, for every $n \in \mathbb{N}$ and for every differentiable convex function $f\colon \mathbb{R}^n \to \mathbb{R}$, the following bound holds for all $a \in \mathbb{R}$ and $t, c \in (0,\infty)$,*

$$\mathbb{P}\big(f(\omega) \leqslant a - t\big)\ \mathbb{P}\big(f(\omega) \geqslant a, |\nabla f(\omega)| \leqslant c\big) \leqslant c_1 \exp\Big(-\frac{(t/c)^\gamma}{c_2}\Big), \tag{5.18}$$

where $\omega = (\omega_1, \dots, \omega_n)$ and $|\nabla f(\omega)| := \sqrt{\sum_{i=1}^n (\partial_i f(\omega))^2}$ is the norm of the gradient.

The second step, after **Estimate 1**, is to use the other important observation, already discussed at the beginning of the proof of Theorem 4.4, that, when averaged at spatial scales of order $\sqrt{N}$, the contributions to the averaged field from disorder $\xi_{n,x}$ with $n = o(N)$ actually become negligible. In particular,

Estimate 2. For $Z^A_{N,\beta_N}(\cdot)$ defined as in (5.10) and any suitable test function ϕ

$$\sqrt{\log N}\, \frac{1}{N} \sum_{x\in\mathbb{Z}^2} \big(\log Z^A_{N,\beta_N}(x) - \mathbb{E}[\log Z^A_{N,\beta_N}(x)]\big)\, \phi(\tfrac{x}{\sqrt{N}}) \xrightarrow[N\to\infty]{L^2(\mathbb{P})} 0\,. \tag{5.19}$$

The proof of this is a fairly simple $L^2(\mathbb{P})$ estimate where we only use the boundedness of moments of $\log Z^A_{N,\beta_N}$, which follows from the negative tail estimate 5.17 as well as that for $\hat{\beta} < 1$ it holds that Z^A_{N,β_N} has bounded second moment.

Estimates 1. and **2.** imply that the fluctuations of the average of $Z_{N,\beta_N}(\cdot)$ will be governed by the fluctuations of the average of the fraction $\big(\frac{\hat{Z}^A_{N,\beta_N}(\cdot)}{Z^A_{N,\beta_N}(\cdot)}\big)$. The crucial point here is that the numerator of this fraction *approximately factorises* in a way that cancels the denominator and what remains is a sort of a restricted polymer partition function for which we can apply a variation of Theorem 4.4.

To see this, we notice that, for any fixed x, $\hat{Z}^A_{N,\beta_N}(x)$ is by definition a "partition function" where disorder outside the box A^x_N is necessarily sampled. This can then be either disorder $\xi_{n,z}$ with $n < N^{1-a_N}$ and with z such that $|z-x| > N^{\frac{1}{2}-\frac{a_N}{4}}$ or disorder $\xi_{n,z}$ with $n > N^{1-a_N}$. The first possibility is negligible as then the random walk will have to travel superdiffusively (this is the reason why the width of the box A^x_N was chosen to be slightly larger than the diffusive scale). In the second situation, the main contribution will actually come from the sampling of disorder in the set

$$B_N := \big((N^{1-9a_N/40}, N] \cap \mathbb{N} \big) \times \mathbb{Z}^2, \tag{5.20}$$

(the choice of the exponent 9/40 is mostly arbitrary and rather technical). This is simply because the slice $\big((N^{1-a_N/4}, N^{1-9a_N/40}] \cap \mathbb{N} \big) \times \mathbb{Z}^2$ is "thin", i.e. its volume is negligible compared to that of B_N. Defining now the corresponding partition function

$$Z^B_{N,\beta_N}(x) := \mathrm{E}_x\big[e^{H_{B_N,\beta_N}}\big] \quad \text{where} \quad H_{B_N,\beta_N} := \sum_{(n,x)\in B_N} (\beta_N \omega_{n,x} - \lambda(\beta_N)) \mathbb{1}_{\{S_n = x\}}$$

we will have that

$$\hat{Z}^A_{N,\beta_N}(x) \approx Z^A_{N,\beta_N}(x) \left(Z^B_{N,\beta_N}(x) - 1 \right), \tag{5.21}$$

The quantitative estimate related to this is

> **Estimate 3.** For $Z^A_{N,\beta_N}(\cdot)$, $\hat{Z}^A_{N,\beta_N}(\cdot)$, $Z^B_{N,\beta_N}(\cdot)$ defined as above and any suitable test function ϕ, we have that
>
> $$\sqrt{\log N}\, \frac{1}{N} \sum_{x\in\mathbb{Z}^2} \left(\frac{\hat{Z}^A_{N,\beta_N}(x)}{Z^A_{N,\beta_N}(x)} - \left(Z^B_{N,\beta_N}(x) - 1 \right) \right) \phi\big(\tfrac{x}{\sqrt{N}}\big) \xrightarrow[N\to\infty]{L^1(\mathbb{P})} 0 .$$

To understand the reason behind this last estimate and (5.21), let us decompose the chaos expansion of $\hat{Z}^A_{N,\beta_N}(x)$ according to the last disorder $\xi_{t,w}$ sampled with $t < N^{1-a_N}$ (see also the Figure above). As we already said, we assume that there is no disorder $\xi_{n,z}$ sampled with $n < N^{1-a_N}$ and $|z-x| > N^{\frac{1}{2}-\frac{a_N}{4}}$ (such contributions are negligible) and moreover we assume that the first disorder $\xi_{n,z}$ that is sampled after time N^{1-a_N} will be such that $n > N^{1-9a_n/40}$ (again, as already mentioned contributions, from sampling disorder between times N^{1-a_N} and $N^{1-9a_n/40}$ are negligible). Thus,

$$\hat{Z}^A_{N,\beta_N}(x) \approx \sum_{\substack{(t,w)\,\in\,\{(0,x)\}\cup A^x_N \\ (r,z)\,\in\, B_N}} Z^A_{0,t,\beta_N}(x,w) \cdot q_{r-t}(z-w) \cdot \sigma_N\, \xi_{r,z} \cdot Z_{r,N,\beta_N}(z), \tag{5.22}$$

where $Z^A_{0,t,\beta_N}(x,w)$ denotes the "point-to-point" partition function where the random walk starts from $(0,x)$ and ends at (t,w) (with the convention that $Z^A_{0,t,\beta_N}(x,w) := 1$ if $(t,w) = (0,x)$) and restricted to sample disorder only in the set A^x_N. Moreover, $Z_{r,N,\beta_N}(z)$ denotes the partition function where the walk starts at time r from position z and runs until time N without any constraint at its end point. The main observation now is that

$$q_{r-t}(z-w) \approx q_r(z-x) \qquad \text{for} \qquad r > N^{1-9a_N/40} >> N^{1-a_N} \geqslant t,$$

using also the diffusive properties of the random walk to say that in these time scales $z - x \approx \sqrt{r-t} \approx \sqrt{r} \approx z - w$. This leads to an (asymptotic) factorisation of (5.22) as

$$\sum_{(t,w)\in\{(0,x)\}\cup A_N^x} Z^A_{0,t,\beta_N}(x,w)\cdot \sum_{(r,z)\in B_N} q_r(z-x)\cdot\sigma_N\,\xi_{r,z}\cdot\, Z_{r,N,\beta_N}(z) = Z^A_{N,\beta_N}(Z^B_{N,\beta_N}(x)-1), \tag{5.23}$$

which is the desired factorisation.

The above estimates reduce the study of the fluctuation to those of $Z^B_{N,\beta_N}(\cdot)$:

Final step. Let $v^{(c)}(s,x)$ be the solution of the two-dimensional additive SHE. Then

$$\frac{\sqrt{\log N}}{\sqrt{\pi}\,\hat\beta}\,\frac{1}{N}\sum_{x\in\mathbb{Z}^2}\big(Z^B_{N,\beta_N}(x)-1\big)\,\phi(\tfrac{x}{\sqrt N}) \xrightarrow[N\to\infty]{d} \langle v^{(\sqrt2 c_{\hat\beta})}(1/2,\cdot),\phi\rangle, \tag{5.24}$$

which is essentially Theorem 4.4. □

6. Criticality and moment estimates

The critical case for the two dimensional SHE corresponds to temperature scaling as $\beta_\varepsilon = \sqrt{2\pi/\log\varepsilon^{-1}}$. In fact, it turns out that there is a critical window of the form

$$\beta_\varepsilon^2 = \frac{2\pi}{\log\varepsilon^{-1}} + \frac{\vartheta}{(\log\varepsilon^{-1})^2}, \qquad \text{for} \quad \vartheta\in\mathbb{R},$$

where one observes a non trivial behaviour depending on the tuning parameter ϑ. In the case of the two dimensional polymer the critical case corresponds to temperature scaling β_N so that

$$\sigma_N^2 := \sigma(\beta_N)^2 = \frac{1}{R_N}\Big(1+\frac{\vartheta+o(1)}{\log N}\Big), \qquad \vartheta\in\mathbb{R}, \tag{6.1}$$

where we recall that $\sigma(\beta)^2 := e^{\lambda(2\beta)-2\lambda(\beta)} - 1$ and the asymptotic $R_N = \frac{(1+o(1))}{\pi}\log N$.

For the sake of brevity of exposition we will only discuss the polymer case. The details for the SHE can be found in [CSZ18b]. Let us denote by $Z_N^{\text{crit.}}$ the partition function of the two dimensional directed polymer Z_{N,β_N} with β_N as in (6.1). As we have already seen in Theorem 4.1, $Z_N^{\text{crit.}}$ converges in distribution to 0. However, we will see that when averaged over the starting point it exhibits a nontrivial behaviour.

Let us highlight that the critical temperature is marked by the fact that its second moment grows to infinity as $N\to\infty$. In fact, we have the more precise asymptotic information, which we will explain below, that it grows as $\log N$.

Proposition 6.1. *Let $Z_N^{\text{crit.}}$ be the partition function of the two dimensional polymer with β_N as in* (6.1)*. Then*

$$\mathbb{V}\text{ar}\Big(Z_N^{\text{crit.}}\Big) \approx \log N\int_0^1 G_\vartheta(t)\mathrm{d}t, \quad \textit{with} \quad G_\vartheta(t) = \int_0^\infty e^{(\vartheta-\gamma)s}\frac{s t^{s-1}}{\Gamma(s+1)}\mathrm{d}s, \tag{6.2}$$

where γ is the Euler-Mascheroni constant $\gamma\approx 0.577$ and $\Gamma(\cdot)$ is the gamma function.

The proof of this theorem goes via creating a link to renewal theory, which further allows for refinements that are useful towards higher moment estimates. We will provide the proof later. For the moment let us mark a distinction between the behaviour of the partition function of a random polymer starting from a fixed point, which without loss of generality we assume to be zero, and the behaviour when the partition function is averaged over its starting point against suitable test functions. Assuming $\phi \in C_c^\infty(\mathbb{R}^2)$ it turns out that the variance of

$$Z_N^{\text{crit.}}(\phi) := \frac{1}{N} \sum_{x \in \mathbb{Z}^2} \Big(Z_N^{\text{crit.}}(x) - 1 \Big)\, \phi(\frac{x}{\sqrt{N}})$$

remains bounded as $N \to \infty$. In fact it turns out that,

$$\lim_{N\to\infty} \mathbb{V}\text{ar}\Big[Z_N^{\text{crit.}}(\phi) \Big] \;=\; \int_{\mathbb{R}^2\times\mathbb{R}^2} \phi(z)\, \phi(z')\, K_\vartheta(z-z')\, \mathrm{d}z\, \mathrm{d}z' \,, \tag{6.3}$$

where the covariance kernel $K_\vartheta(\cdot)$ is given by

$$K_\vartheta(x) := \pi \int\limits_{0<u<v<1} g_u(x)\, G_\vartheta(v-u)\, \mathrm{d}u\, \mathrm{d}v \,. \tag{6.4}$$

where G_ϑ is defined as above and $g_u(x)$ is the heat kernel. It is worth to remark that the kernel $K_\vartheta(x) \sim C \log \frac{1}{|x|}$ as $x \sim 0$, which means that any (conjecturally unique) limit of the field is log-correlated.

The boundedness of $Z_N^{\text{crit.}}(\phi)$ shows the existence of limits of the field at criticality *without any rescaling* (as was the case below the critical temperature, see Theorem 4.4). However, to ensure that the / any limiting field is non trivial, i.e. not just "flat" (Lebesque), requires boundedness of higher moments. This is because in order to say that $\mathbb{V}\text{ar}\big(\lim_{N\to\infty} Z_N^{\text{crit.}}(\phi)\big) = \lim_{N\to\infty} \mathbb{V}\text{ar}\big(Z_N^{\text{crit.}}(\phi)\big)$ (the latter being non zero as we remarked above), we need control of higher moments that will allow to interchange the limits via uniform integrability. The first such estimate was achieved in [CSZ18b]:

Theorem 6.2 (Third moment). *Let $\phi \in C_c(\mathbb{R}^2)$, $\vartheta \in \mathbb{R}$. Let $Z_N^{\text{crit.}}$ be the partition function corresponding to the choice of critical β_N (6.1). Then*

$$\lim_{N\to\infty} \mathbb{E}\Big[\big(Z_N^{\text{crit.}}(\phi) - \mathbb{E}\big[Z_N^{\text{crit.}}(\phi)\big] \big)^3 \Big] \;=\; \int\limits_{(\mathbb{R}^2)^3} \phi(z)\, \phi(z')\, \phi(z'')\, M_\vartheta(z,z',z'')\, \mathrm{d}z\, \mathrm{d}z'\, \mathrm{d}z'' \;<\; \infty \,, \tag{6.5}$$

where the kernel $M_\vartheta(\cdot)$ is given by

$$M_\vartheta(z,z',z'') \;:=\; \sum_{m=2}^{\infty} 2^{m-1}\, \pi^m \left\{ \mathcal{I}_\vartheta^{(m)}(z,z',z'') + \mathcal{I}_\vartheta^{(m)}(z',z'',z) + \mathcal{I}_\vartheta^{(m)}(z'',z,z') \right\}, \tag{6.6}$$

with $\mathcal{I}_\vartheta^{(m)}(\cdot)$ defined as follows:

$$\begin{aligned} \mathcal{I}_\vartheta^{(m)}(z,z',z'') := \int\!\!\cdots\!\!\int\limits_{\substack{0<a_1<b_1<\ldots<a_m<b_m<t \\ x_1,y_1,\ldots,x_m,y_m\in\mathbb{R}^2}} & \mathrm{d}\vec{a}\, \mathrm{d}\vec{b}\, \mathrm{d}\vec{x}\, \mathrm{d}\vec{y}\;\; g_{\frac{a_1}{2}}(x_1-z)\, g_{\frac{a_1}{2}}(x_1-z')\, g_{\frac{a_2}{2}}(x_2-z'') \\ & \cdot G_\vartheta(b_1-a_1, y_1-x_1)\, g_{\frac{a_2-b_1}{2}}(x_2-y_1)\, G_\vartheta(b_2-a_2, y_2-x_2) \\ & \cdot \prod_{i=3}^{m} g_{\frac{a_i-b_{i-2}}{2}}(x_i-y_{i-2})\, g_{\frac{a_i-b_{i-1}}{2}}(x_i-y_{i-1})\, G_\vartheta(b_i-a_i, y_i-x_i)\,, \end{aligned} \tag{6.7}$$

where $G_\vartheta(t,x) = G_\vartheta(t) g_{t/4}(x)$ with $G_\vartheta(t)$ as in (6.2) and $g_t(x)$ the two dimensional heat kernel.

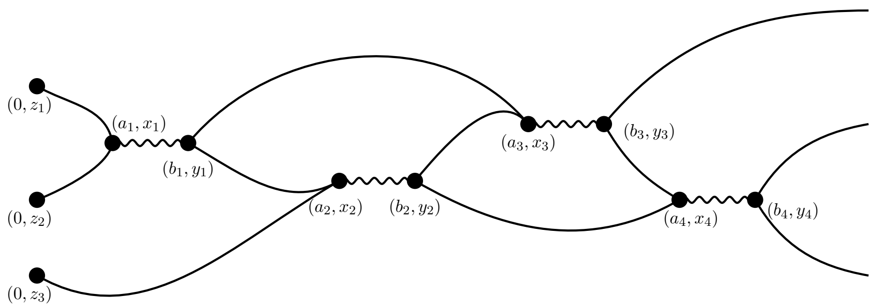


FIGURE 2. Diagramatic representation of the expansion (6.10) of the third moment. Curly lines between nodes (a_i, x_i) and (b_i, y_i) have weight $U_N(b_i - x_i, y_i - x_i)$, coming for pairwise matchings between a single pair of copies AB, BC or CA, while solid, curved lines between nodes (a_i, x_i) and (b_{i-1}, y_{i-1}) or between (a_i, x_i) and (b_{i-2}, y_{i-2}) indicate a weight $q_{b_{i-1},a_i}(y_{i-1}, x_i)$ and $q_{b_{i-2},a_i}(y_{i-2}, x_i)$, respectively.

The analogue of this theorem for the two-dimensional SHE was also established in [CSZ17b]. More recently Gu-Quastel-Tsai [GQT19] established the analogue of the above theorem for higher than three moments for the SHE. We will not expose it here as it requires a different set of notation but we will give an informal description of their result and compare it with the formulation of Theorem 6.2 below.

Let us give a very brief sketch of the framework of the proof of Theorem 6.2.

Sketch of the proof of Theorem 6.2. Our framework involves again the polynomial chaos expansion of the partition function (4.5). For conciseness we will introduce the notation

$$q_t^N(\phi, x) := \sum_{y \in \mathbb{Z}^2} q_t(x-y)\phi\big(\frac{y}{\sqrt{N}}\big),$$

which incorporates the averaging over the initial condition combined with the first transition kernel in the chaos expansion. Replacing $Z_N^{\text{crit.}}$ with is chaos expansion and using Fubini to develop the third power we have

$$\begin{aligned}
&\mathrm{E}\Big[\big(Z_N^{\text{crit.}}(\phi) - \mathbb{E}[Z_N^{\text{crit.}}(\phi)]\big)^3\Big] \\
&\quad = \sum_{\substack{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C} \subseteq \{1,\dots,tN\}\times\mathbb{Z}^2 \\ |\boldsymbol{A}|\geqslant 1,\, |\boldsymbol{B}|\geqslant 1,\, |\boldsymbol{C}|\geqslant 1}} \frac{\sigma_N^{|\boldsymbol{A}|+|\boldsymbol{B}|+|\boldsymbol{C}|}}{N^3}\, q^N_{s,a_1}(\phi, x_1)\, q^N_{s,b_1}(\phi, y_1) \cdot q^N_{s,c_1}(\phi, z_1) \cdot \\
&\qquad \cdot \mathbb{E}\left[\xi_{A_1} \prod_{i=2}^{|\boldsymbol{A}|} \xi_{A_i}\, q(A_{i-1}, A_i) \cdot \xi_{B_1} \prod_{j=2}^{|\boldsymbol{B}|} \xi_{B_j}\, q(B_{j-1}, B_j) \cdot \xi_{C_1} \prod_{k=2}^{|\boldsymbol{C}|} \xi_{C_k}\, q(C_{k-1}, C_k)\right]
\end{aligned} \tag{6.8}$$

where we have used the shorthand notation $\boldsymbol{A} = (A_1, \dots, A_{|\boldsymbol{A}|})$ with $A_i = (a_i, x_i) \in \mathbb{Z}^3_{\text{even}}$, and $\boldsymbol{B}$, $\boldsymbol{C}$ defined similarly, with $B_j = (b_j, y_j)$, $C_k = (c_k, z_k)$, and we have set

$$q(A_{i-1}, A_i) := q_{a_i - a_{i-1}}(x_i - x_{i-1}).$$

When $|\boldsymbol{A}| = 1$, we use the convention that $\prod_{i=2}^{|\boldsymbol{A}|} \dots$ equals 1 and similarly for $\boldsymbol{B}$ and $\boldsymbol{C}$.

Denote $\boldsymbol{D} := \boldsymbol{A} \cup \boldsymbol{B} \cup \boldsymbol{C} \subset \{1,\ldots,N\} \times \mathbb{Z}^2$, with $\boldsymbol{D} = (D_1, \ldots, D_{|\boldsymbol{D}|})$ and $D_i = (d_i, w_i)$. Since $\mathbb{E}[\xi_z] = 0$, the contributions to $M^{N,\mathrm{NT}}_{s,t}(\phi,\psi)$ come only from $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ where the points in $\boldsymbol{A} \cup \boldsymbol{B} \cup \boldsymbol{C}$ pair up - for the sake of exposition we ignore here a triple matching. That is, we ignore the case that $A_i = B_j = C_k$ for some i, k, j. In particular, we assume that $k := |\boldsymbol{D}| = \frac{1}{2}(|\boldsymbol{A}| + |\boldsymbol{B}| + |\boldsymbol{C}|)$ and that each point D_j belongs to exactly two of the three sets $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$. Hence we can associate a vector $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_k)$ of labels $\ell_j \in \{AB,\, BC,\, AC\}$. Note that there is a one to one correspondence between $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ and $(\boldsymbol{D}, \boldsymbol{\ell})$. So we can write (6.8) as

$$\begin{aligned} &\frac{1}{N^3} \sum_{k=2}^{\infty} \sigma_N^{2k} \sum_{\substack{\boldsymbol{D} \subseteq \{1,\ldots,N\} \times \mathbb{Z}^2 \\ |\boldsymbol{D}| = k \geqslant 2}} \; \sum_{\boldsymbol{\ell} \in \{AB,\, BC,\, AC\}^k} q^N_{s,a_1}(\phi, x_1)\, q^N_{s,b_1}(\phi, y_1)\, q^N_{s,c_1}(\phi, z_1) \,\cdot \\ &\qquad \cdot \prod_{i=2}^{|\boldsymbol{A}|} q(A_{i-1}, A_i) \prod_{j=2}^{|\boldsymbol{B}|} q(B_{j-1}, B_j) \prod_{m=2}^{|\boldsymbol{C}|} q(C_{m-1}, C_m)\,, \end{aligned} \tag{6.9}$$

with the $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ in the above expression being implicitly determined by $(\boldsymbol{D}, \boldsymbol{\ell})$.

We now make a combinatorial observation, see also Figure 2. The sequence $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_k)$ consists of consecutive *stretches* $(\ell_1, \ldots, \ell_i)$, $(\ell_{i+1}, \ldots, \ell_j)$, etc., such that the labels are constant in each stretch and change from one stretch to the next. Any stretch, say $(\ell_p, \ldots, \ell_q)$, has a first point $D_p = (a, x)$ and a last point $D_q = (b, y)$. Let m denote the number of stretches and let (a_i, x_i) and (b_i, y_i), with $a_i \leqslant b_i$, be the first and last points of the i-th stretch.

We now rewrite (6.9) by summing over $m \in \mathbb{N}$, $(a_1, b_1, \ldots, a_m, b_m)$, and $(x_1, y_1, \ldots, x_m, y_m)$. The sum over the labels of $\boldsymbol{\ell}$ leads to a combinatorial factor $3 \cdot 2^{m-1}$, because there are 3 choices for the label of the first stretch and two choices for the label of the following stretches. Once we fix (a_1, x_1) and (b_1, y_1), summing over all possible configurations inside the first stretch gives the factor

$$\sum_{r=1}^{\infty} \sigma_N^{2(r+1)} \sum_{\substack{a_1 = t_0 < t_1 < \ldots < t_r = b_1 \\ z_0 = x_1,\, z_1, z_2, \ldots, z_{r-1} \in \mathbb{Z}^2,\, z_r = y_1}} \prod_{i=1}^{r} q_{t_{i-1}, t_i}(z_{i-1}, z_i)^2 =: \sigma_N^2\, U_N(b_1 - a_1,\, y_1 - x_1)\,,$$

The quantity U_N in the right hand side is defined via this relation and it is closely related to the (point-to-point) variance of the polymer partition function. A similar factor arises from each stretch and this leads to the following expression for the centred third moment (assuming we have ignored the case of triple matchings of the ξ variables, hence the quotation marks in the equality below; this is a technical point that can be dealt with some extra work)

$$\mathrm{E}\Big[\big(Z_N^{\text{crit.}}(\phi)-\mathbb{E}[Z_N^{\text{crit.}}(\phi)]\big)^3\Big] \text{``}=\text{''} \sum_{m=2}^{\infty} 3\cdot 2^{m-1} I_\vartheta^{(N,m)}(\phi), \qquad \text{where}$$

$$\begin{aligned} I_\vartheta^{(N,m)}(\phi) := \frac{\sigma_N^{2m}}{N^3} \sum_{\substack{s<a_1\leqslant b_1<a_2\leqslant b_2<\ldots<a_m\leqslant b_m<t\\ x_1,y_1,x_2,y_2,\ldots,x_m,y_m\in\mathbb{Z}^2}} & q_{s,a_1}^N(\phi,x_1)^2\, q_{s,a_2}^N(\phi,x_2)\,\cdot \\ & \cdot\, U_N(b_1-a_1,y_1-x_1)\, q_{b_1,a_2}(y_1,x_2)\, U_N(b_2-a_2,y_2-x_2)\,\cdot \\ & \cdot \prod_{i=3}^{m}\Big\{q_{b_{i-2},a_i}(y_{i-2},x_i)\, q_{b_{i-1},a_i}(y_{i-1},x_i)\, U_N(b_i-a_i,y_i-x_i)\Big\}\,, \end{aligned} \tag{6.10}$$

with the convention that $\prod_{i=3}^m\{\ldots\}=1$ for $m=2$. Note that the sum starts with $m=2$ because in (6.9), we have $|\boldsymbol{A}|,|\boldsymbol{B}|,|\boldsymbol{C}|\geqslant 1$.

Passing from (6.10) to (6.5)-(6.6) amounts to a Riemann sum approximation after scaling the time variables proportionally to N and the space variables proportionally to $\sqrt{N}$. Crucial to this limiting procedure, as well as ensuring that the resulting series converge, is the asymptotic behaviour of $U_N(tN,x\sqrt{N})$ for large N and $t\in\mathbb{R}, x\in\mathbb{R}^2$. In particular, for $x\in\mathbb{Z}^2, n\in\mathbb{N}$, it holds that

$$U_N(n,x)\approx \frac{\log N}{N^2} G_\vartheta\Big(\frac{n}{N},\frac{x}{\sqrt{N}}\Big)\, 2\mathbb{1}_{\{n+x_1+x_2\ \text{even}\}}, \tag{6.11}$$

with $G_\vartheta(t,x)=G_\vartheta(t)g_{t/4}(x)$ and $G_\vartheta(t)$ defined in (6.4). The factors of $\log N$ will cancel with the factors of σ_N^2 in (6.10) and the factors N^{-2} will be absorbed by the Riemann sum approximations. Moreover, for every fixed $\vartheta\in\mathbb{R}$, we have the asymptotic behaviour

$$G_\vartheta(t)=\frac{1}{t\,(\log\frac{1}{t})^2}+\frac{2\vartheta+o(1)}{t\,(\log\frac{1}{t})^3} \qquad \text{as} \quad t\to 0\,. \tag{6.12}$$

These asymptotic behaviours are based on the renewal theory framework, same as the one that underlies the variance asymptotics in Proposition 6.1. Even though we will not discuss the details, which can be found in [CSZ18c], the underlying framework will become clear when we sketch the proof of Proposition 6.1 below. One thing that should be remarked is the bare integrability of $G_\vartheta(t)$ which shows how marginal is the integrability of the moments of the averaged field of partition functions in two dimensions.

Having these estimates, what remains to conclude is to ensure that the series $\sum_{m=2}^\infty 3\cdot 2^{m-1}I_\vartheta^{(N,m)}(\phi)$ in (6.10) converge, uniformly in N. This point is quite technical due to the interlacing structure as shown in Figure 2 and we refer for the details to [CSZ18b]. □

Sketch of the proof of Proposition 6.1. As we have already seen a few times, using the polynomial chaos expansion of the partition function its variance can be written as

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\left(Z_N^{\mathrm{crit.}}\right) &= \sum_{k\geqslant 1} \sigma_N^{2k} \sum_{\substack{1\leqslant n_1<\cdots<n_k\leqslant N \\ x_1,\ldots,x_k\in\mathbb{Z}^2}} \prod_{i=1}^{k} q_{n_i-n_{i-1}}(x_i-x_{i-1})^2 \\
&= \sum_{k\geqslant 1} \sigma_N^{2k} \sum_{1\leqslant n_1<\cdots<n_k\leqslant N} \prod_{i=1}^{k} q_{2(n_i-n_{i-1})}(0)^2 \\
&= \sum_{k\geqslant 1} \Big(1+\frac{\vartheta+o(1)}{\log N}\Big)^k \frac{1}{R_N^k} \sum_{1\leqslant n_1<\cdots<n_k\leqslant N} \prod_{i=1}^{k} q_{2(n_i-n_{i-1})}(0)^2,
\end{aligned}$$

where in the second equality we just used the convolution property of the random walk and in the third the definition of the choice of σ_N^2. Now, we will write the last convolution as a renewal probability. In particular, we define the i.i.d. random variables $T_1^{(N)},\ldots,T_k^{(N)}$ with

$$\mathrm{P}\big(T_1^{(N)}=n\big) = \frac{1}{R_N} q_{2n}(0)\mathbb{1}_{n\leqslant N} \approx \frac{1}{\log N}\frac{\mathbb{1}_{n\leqslant N}}{n},$$

(with the last due to the local limit theorem and the asymptotics of R_N) and

$$\tau_k^{(N)} := T_1^{(N)}+\cdots+T_k^{(N)}.$$

We can then write

$$\mathbb{V}\mathrm{ar}\left(Z_N^{\mathrm{crit.}}\right) = \sum_{k\geqslant 1} \Big(1+\frac{\vartheta+o(1)}{\log N}\Big)^k \mathrm{P}\big(\tau_k^{(N)}\leqslant N\big). \tag{6.13}$$

The point now is that $\big(\frac{1}{N}\tau^{(N)}_{s\log N}\big)_{s>0}$ converges to a process $(Y_s)_{s>0}$ with a density which can be explicitly computed. We will see this in a moment, but let us now use this fact to conclude the asymptotics of the variance from (6.13). This boils down to a Riemann sum approximation as

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\left(Z_N^{\mathrm{crit.}}\right) &= \sum_{k\geqslant 1} \Big(1+\frac{\vartheta+o(1)}{\log N}\Big)^{2k} \mathrm{P}\big(\tau_k^{(N)}\leqslant N\big) \\
&= \log N \,\frac{1}{\log N} \sum_{s\in\frac{1}{\log N}\mathbb{N}} \Big(1+\frac{\vartheta+o(1)}{\log N}\Big)^{s\log N} \mathrm{P}\big(\tau^{(N)}_{s\log N}\leqslant N\big) \\
&\approx \log N \int_0^\infty e^{\vartheta s}\mathrm{P}(Y_s\leqslant 1)\,\mathrm{d}s = \log N \int_0^\infty e^{\vartheta s}\int_0^1 f_s(t)\mathrm{d}t\,\mathrm{d}s,
\end{aligned}$$

where $f_s(t)$ is the density of Y_s, which can be computed exactly (see [CSZ18c]) as

$$f_s(t) = \begin{cases} \dfrac{s\,t^{s-1}\,e^{-\gamma\,s}}{\Gamma(s+1)} & \text{for } t\in(0,1], \\[2ex] \dfrac{s\,t^{s-1}e^{-\gamma s}}{\Gamma(s+1)} - st^{s-1}\displaystyle\int_0^{t-1}\frac{f_s(a)}{(1+a)^s}\mathrm{d}a & \text{for } t\in(1,\infty), \end{cases} \tag{6.14}$$

This leads to the form of the asymptotic variance as in (6.2).

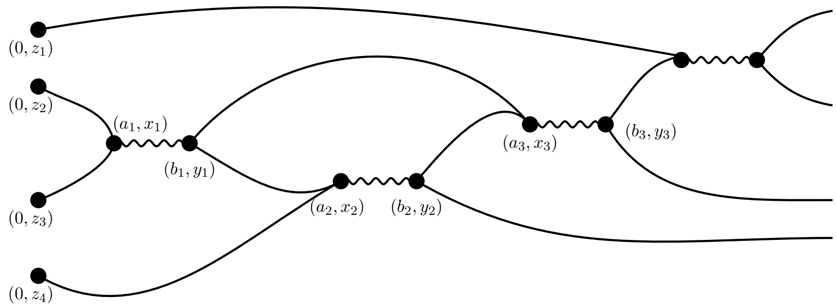


FIGURE 3. Diagramatic representation of the expansion of in the case of the fourth moment in analogy with the diagram of Figure 2. Curly and curved lines bare the same weights as in Figure 2. Here we notice the non locality of the topmost lace.

The convergence of $\big(\frac{1}{N}\,\tau^{(N)}_{s\log N}\big)_{s>0}$ can be easily seen via a Fourier transform computation as

$$\mathrm{E}\Big[e^{\frac{\lambda}{N}\,\tau^{(N)}_{s\log N}}\Big]=\Big(\mathrm{E}\Big[e^{\frac{\lambda}{N}\,T^{(N)}_1}\Big]\Big)^{s\log N}\approx\Big(1+\frac{1}{\log N}\sum_{k=1}^{N}\big(e^{\frac{\lambda}{N}n}-1\big)\frac{1}{n}\Big)^{s\log N},$$

which again by a Riemann sum approximation converges to

$$\exp\Big(s\int_0^1\big(e^{\lambda x}-1\big)\frac{\mathrm{d}x}{x}\Big).$$

This expression plus the independence of the increments, inherited by the independence to $\{T^{(N)}_k\}_{k\geqslant 1}$, shows that $(Y_s)_{s>0}$ is a Lévy type process with Lévy measure $\mathbb{1}_{(0,1)}(x)\,\mathrm{d}x/x$. The fact that the density of Y_s can be computed explicitly as in (6.14) is a non trivial fact and was done in [CSZ18c, Appendix B]. Its computability is related to an invariance of the process Y, which amounts to the fact that conditionally on all the jumps up to time s being smaller than t, the law of Y_s/t is the same as the law of Y_s (see [CSZ18c, Proposition B.1]).

It is worth remarking that the density $f_s(\cdot)$ is related to what is called the *Dickman function*, which is a very distinguished function in analytic number theory. In particular, if we define (the Dickman function)

$$\varrho(t)=e^{\gamma}f_1(t),$$

then $\varrho(1/t)$ equals the asymptotic probability that the largest prime factor of a number chosen uniformly from $\{1,...,N\}$ is less than N^t, see [Ten95]. □

Before closing this section let us comment on the higher moments of the averaged field and the work of Gu-Quastel-Tsai [GQT19]. If we wanted to adapt the approach we described for the third moment, then we would need to deal with (further) non local interactions. For example (see Figure 2), if we wanted to compute the fourth moment, then we would need to deal with four copies of polymer and consider the pairwise matchings, as was done earlier. The non locality in this approach would consist of the possible scenario (among others) that three of the copies match pairwise for some time, until, only much later, the copy that was left alone (in the case of Figure 2 this would correspond to the topmost line) starts matching

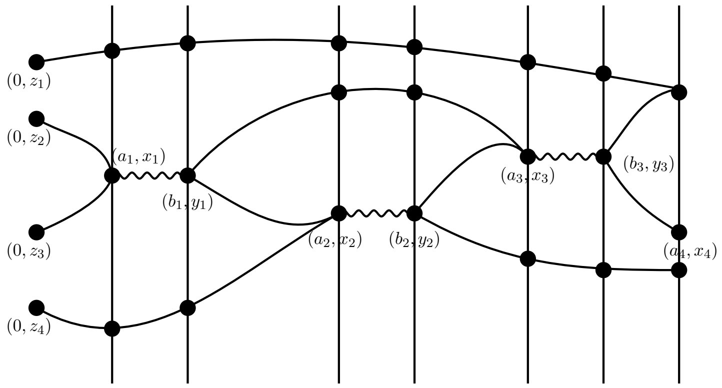


FIGURE 4. Diagramatic representation of the expansion of the fourth moment (in a discrete format) followed by [GQT19]. Curly and curved lines bare the same (or rather continuous analogues of the) weights as in Figures 2 and 3. Additionally, this diagram keeps track of all the marked points on the vertical lines and not only the beginning and end points of the curly lines as in the previous two figures.

with one of the other three copies. This connection is non local and keeping track of the starting points of non local laces is complicated.

The approach of [GQT19], was to introduce additional space-time points, see Figure 4, and consider these in the decomposition of the summation. Notice that we could remove these additional points by summing over them and this would bring us back to the previous decomposition as in Figure 3. However, keeping track of the leads to a Markovian structure, which allows to handle the combinatorics easier. This approach was inspired by previous works on Hamiltonians with point interactions[DR04]

$$-\Delta + \sum_{i<j} \delta(x_i - x_j), \qquad \text{on } \mathbb{R}^2.$$

It also used some crucial estimates on suitable norms of the operator corresponding to propagation between points in strips without curly lines from [DFT94]. The corresponding estimates on the operators corresponding to propagation between points in strips with curly lines is close in spirit to our estimates around the function $U_N(n,x)$ as derived from (6.11).

7. APPENDIX

Proof of Theorem 3.2. Without loss of generality we will assume that the index set S is finite and for notational simplicity we identify it with $\{1,...,n\}$. More crucially, we will assume that Ψ has degree ℓ which stays bounded in n, that is

$$\Psi(\xi) = \sum_{I \subset \mathsf{S}\,,\, |I| \leqslant \ell} \psi(I)\, \xi^I.$$

This assumption can be justified by a simple truncation argument, we refer to [CSZ17a] for details. For a function $f \in C_b^3(\mathbb{R})$ we denote

$$g(x_1,...,x_n) := f(\Psi(x)), \tag{7.1}$$

and

$$h_{n,j}^{X^j}(y) := g(\zeta_1,...,\zeta_{j-1}, y, \xi_{j+1},...,\xi_n), \quad \text{with} \quad X^j := (\zeta_1,...,\zeta_{j-1},\xi_{j+1},...,\xi_n), \tag{7.2}$$

and we have that

$$f(\Psi(\xi)) - f(\Psi(\zeta)) = \sum_{j=1}^{n} \big(h_{n,j}^{X^j}(\xi_j) - h_{n,j}^{X^j}(\zeta_j)\big). \tag{7.3}$$

We now Taylor expand the function $h_{n,j}^{X^j}(\cdot)$ around zero as

$$h_{n,j}^{X^j}(y) = h_{n,j}^{X^j}(0) + \big(\partial_y\, h_{n,j}^{X^j}(0)\big)y + \frac{1}{2}\big(\partial_y^2\, h_{n,j}^{X^j}(0)\big)y^2 + R_{n,j}^{X^j}(y),$$

where the error term

$$R_{n,i}^{X^j}(y) = \frac{1}{2}\int_0^y \Big(\partial_y^3 h_{n,i}^{X^j}(t)\Big)\,(y-t)^2 \mathrm{d}t, \tag{7.4}$$

and the following two bounds hold:

$$\big|\, R_{n,i}^{X^j}(y) \,\big| \leqslant \frac{1}{6}\|\, \partial_y^3 h_{n,i}^{X^j}\,\|_\infty\, |y|^3 = \frac{1}{6}\|f'''\|_\infty |y|^3 \tag{7.5}$$

$$\big|\, R_{n,i}^{X^j}(y) \,\big| \leqslant \|\, \partial_y^2 h_{n,i}^{X^j}\,\|_\infty\, y^2 = \|f''\|_\infty y^2. \tag{7.6}$$

The first bound follows by bounding $\partial_y^3 h_{n,i}^{X^j}$ in (7.4) by its supremum norm, while for the second bound we first perform an integration by parts and write the remainder as

$$R_{n,i}^{X^j}(y) = -\frac{1}{2}\partial_y^2 h_{n,i}^{X^j}(0)y^2 + \int_0^y \partial_y^2 h_{n,i}^{X^j}(t)(y-t)\,\mathrm{d}t,$$

and then bound $\partial_y^2 h_{n,i}^{X^j}$ by its supremum norm.

Inserting this Taylor expansion into (7.3) for $h_{n,j}^{X^j}(\xi_j)$ and $h_{n,j}^{X^j}(\zeta_j)$ and using the fact that the first and second moments of the ξ and ζ variables match, we have the estimate

$$\begin{aligned}\Big|\mathbb{E}[f(\Psi(\xi))] - \mathbb{E}[f(\Psi(\zeta))]\Big| &= \Big|\sum_{j=1}^{N} \mathbb{E}\Big[R_{n,j}^{X^j}(\xi_j) - R_{n,j}^{X^j}(\zeta_j)\Big]\Big| \\ &\leqslant \sum_{j=1}^{N} \mathbb{E}\Big[|R_{n,j}^{X^j}(\zeta_j)|\Big] + \sum_{j=1}^{N} \mathbb{E}\Big[|R_{n,j}^{X^j}(\xi_j)|\Big].\end{aligned}$$

The derivatives of $h_j^x(\cdot) := g(x_1, ..., x_{j-1}, y, x_{j+1}, ..., x_n)$ with g defined as in (7.1) are computed as:

$$\begin{aligned}\big(\partial_y^m h_j^x\big)\,(y) &= f^{(m)}\Big(\Psi(x_1,\ldots,x_{j-1},y,x_{j+1},\ldots,x_N)\Big)\Big(\,\partial_y\Psi(x_1,\ldots,x_{j-1},y,x_{j+1},\ldots,x_N)\,\Big)^m \\ &= f^{(m)}\Big(\Psi(x_1,\ldots,x_{j-1},y,x_{j+1},\ldots,x_N)\Big)\,\Big(\sum_{I\ni j}\psi(I)x^{I\setminus\{j\}}\Big)^m.\end{aligned}$$

Defining

$$L_j(x) := \sum_{I\ni j}\psi(I)x^I,$$

we obtain that bounds (7.5) and (7.6) on $R_{n,j}^{X^j}(y)$ give

$$\sum_{j=1}^{N} \mathbb{E}\Big[|R_{n,j}^{X^j}(\zeta_j)|\Big] \leqslant C_f \sum_{j=1}^{N} \mathbb{E}\left[\varphi(L_j(X^j))\right], \tag{7.7}$$

with $C_f = \max\{\|f'\|_\infty, \|f^{(2)}\|_\infty, \|f^{(3)}\|_\infty\}$ and $\varphi(x) := \min\{\frac{|x|^3}{6}, |x|^2\}$. To proceed with a sharp estimate on (7.7) under only the assumption of uniformly integrable second moments, we need to truncate the random variables in a way that also respects an orthogonality. The general truncation is described as follows:

Truncation procedure : Fix $M \in (0,\infty)$. We can decompose any real-valued random variable Y with zero mean and finite variance as

$$Y = Y^- + Y^+ , \tag{7.8}$$

where Y^-, Y^+ are functions of Y and possibly of some extra randomness, such that

$$\begin{aligned} &\mathbb{E}[Y^-] = \mathbb{E}[Y^+] = 0\,, && Y^-Y^+ = 0\,, \\ &|Y^-| \leqslant |Y|\,\mathbb{1}_{\{|Y|\leqslant M\}}\,, && \mathbb{E}[(Y^+)^2] \leqslant 2\,\mathbb{E}[Y^2\mathbb{1}_{\{|Y|>M\}}]\,. \end{aligned} \tag{7.9}$$

We postpone the proof of the truncation properties (7.9) until the end of the proof of this theorem. Assuming these properties, we proceed by denoting by X^{j-} the vector X^j from (7.2) with all its entries truncated as above and also $X^{j+} := X^j - X^{j-}$. Noting the elementary inequality

$$\varphi(a+b) \leqslant 2a^2 + \frac{4}{3}|b|^3, \qquad \text{for real } a, b,$$

we have that the bound in (7.7) can be extended to

$$\mathbb{E}\big[\varphi(L_j(X^j))\big] \leqslant 4\,\mathbb{E}\Big[\big(L_j(X^j) - L_j(X^{j-})\big)^2\Big] + \frac{4}{3}\,\mathbb{E}\big[|L_j(X^{j-})|^3\big]. \tag{7.10}$$

Estimate on the first term in (7.10): To estimate the first term in (7.10) we write

$$L_j(X^j) - L_j(X^{j-}) = \sum_{I\ni j}\psi(I)\sum_{\Gamma\subseteq I,\,|\Gamma|\geqslant 1}(X^{j+})^\Gamma(X^{j-})^{I\setminus\Gamma}\,.$$

By (7.9) the random variables $X_1^{j-}, X_1^{j+}, X_2^{j-}, X_2^{j+}, \ldots$ are orthogonal. Setting $\sigma^2_{\pm,i} := \mathbb{E}[(X_i^{j\pm})^2]$ and observing that $\sigma^2_{-,i} + \sigma^2_{+,i} = \mathbb{V}\mathrm{ar}(X_i^j) = 1$, we obtain

$$\begin{aligned} \mathbb{E}\big[\big(L_j(X^j) - L_j(X^{j-})\big)^2\big] &= \sum_{I\ni j}\psi(I)^2\sum_{\Gamma\subseteq I,\,|\Gamma|\geqslant 1}(\sigma_+^2)^\Gamma(\sigma_-^2)^{I\setminus\Gamma} \\ &= \sum_{I\ni j}\psi(I)^2\big(1-(\sigma_-^2)^I\big) \leqslant \sum_{I\ni j}\psi(I)^2\big(1-(1-\overline{\sigma}_+^2)^{|I|}\big)\,, \end{aligned} \tag{7.11}$$

where

$$\overline{\sigma}_+^2 := \max_{i=1,\ldots,N}\sigma^2_{+,i} = \max_{i=1,\ldots,n}\mathbb{E}[(X_i^{j+})^2] \leqslant 2\max_{i=1,\ldots,n}\mathbb{E}[(X_i^j)^2\,\mathbb{1}_{\{|X_i^j|>M\}}] \leqslant 2\,\boldsymbol{m}_2^{>M}\,,$$

having used (7.9) and having defined

$$\boldsymbol{m}_2^{>M} := \sup_{X\in\{\zeta_i,\xi_i\}_{i\geqslant 1}}\mathbb{E}\big[X^2\mathbb{1}_{|X|\geqslant M}\big].$$

Using the estimate $\big(1-(1-\overline{\sigma}_+^2)^{|I|} \leqslant |I|\,\overline{\sigma}_+^2$ in (7.11) we have that

$$\begin{aligned}\sum_{j=1}^{n} \mathbb{E}\Big[\big(L_j(X^j)-L_j(X^{j-})\big)^2\Big] \;&\leqslant\; 2\,\boldsymbol{m}_2^{>M} \sum_j \bigg(\sum_{I\ni j} |I|\,\psi(I)^2\bigg) \\ &\leqslant 2\,\boldsymbol{m}_2^{>M}\,\ell^2 \sum_I \psi(I)^2\,, \end{aligned} \tag{7.12}$$

where we recall that ℓ is the degree of Ψ. Given the uniform integrability of the second moment, the last bound can be made arbitrarily small, say less than ε, by choosing M large enough.

Estimate on the second term in (7.10)**:** For the second term we will use hypercontractivity bound Theorem 3.5 with an non optimal constant (as provided in [MOO10]) $\widetilde{\varrho}_3 := 2\sqrt{2}\,\max_{i\leqslant n} \frac{\|X_i^{j-}\|_3}{\|X_i^{j-}\|_2}$. In particular,

$$\|L_j(X^{j-})\|_3 \leqslant \widetilde{\varrho}_3^{\,\ell}\,\|L_j(X^{j-})\|_2\,, \tag{7.13}$$

Since for every i we have that $|X_i^{j-}| \leqslant |X_i^j|\,\mathbb{1}_{|X_i^{j-}|\leqslant M}$, by (7.9), we have

$$\|X_i^{j-}\|_3 \leqslant \mathbb{E}\big[\,|X_i^j|^3\,\mathbb{1}_{\{|X_i^j|\leqslant M\}}\,\big]^{1/3} \leqslant \big(\boldsymbol{m}_3^{\leqslant M}\big)^{1/3}\,,$$

with $\boldsymbol{m}_3^{\leqslant M}$ being the maximum truncated third moment of variables $\xi_i, \zeta_i,\ i\geqslant 1$. On the other hand, again by (7.9), we have that for every i

$$\begin{aligned}\|X_i^{j-}\|_2^2 &= \|X_i^j\|_2^2 - \|X_i^{j+}\|_2^2 = \mathbb{E}[(X_i^j)^2] - \mathbb{E}[(X_i^{j+})^2] \geqslant \mathbb{E}[(X_i^j)^2] - 2\mathbb{E}[(X_i^j)^2\mathbb{1}_{\{|X_i^j|>M\}}] \\ &= 1 - 2\,\mathbb{E}[(X_i^j)^2\mathbb{1}_{\{|X_i^j|>M\}}] \geqslant 1 - 2\boldsymbol{m}_2^{>M}\,,\end{aligned}$$

hence

$$\widetilde{\varrho}_3 \leqslant 2\sqrt{2}\frac{\big(\boldsymbol{m}_3^{\leqslant M}\big)^{1/3}}{\sqrt{1-2\boldsymbol{m}_2^{>M}}} \leqslant 4\big(\boldsymbol{m}_3^{\leqslant M}\big)^{1/3}\,,$$

provided $\boldsymbol{m}_{>M}^{(2)} \leqslant \frac{1}{4}$, which can be achieved by choosing M large enough, thanks to the uniform integrability of the second moment. Therefore, (7.13) yields

$$\mathbb{E}\left[|L_j(X^{j-})|^3\right] \leqslant 64^\ell\,\big(\boldsymbol{m}_3^{\leqslant M}\big)^\ell\,\mathbb{E}\left[L_j(X^{j-})^2\right]^{3/2}\,.$$

Note that, since $\mathbb{E}[(X_i^{j-})^2] \leqslant \mathbb{E}[(X_i^j)^2] = 1$, we have

$$\mathbb{E}\left[L_j(X^{j-})^2\right] = \sum_{I\ni j}\psi(I)^2 \prod_{i\in I}\mathbb{E}[(X_i^{j-})^2] \leqslant \sum_{I\ni j}\psi(I)^2 = \mathrm{Inf}_j[\Psi]\,.$$

Therefore

$$\begin{aligned}\sum_{j=1}^{N}\mathbb{E}\left[|L_j(X^{j-})|^3\right] &\leqslant 64^\ell\;\big(\boldsymbol{m}_3^{\leqslant M}\big)^\ell\Big(\max_i\sqrt{\mathrm{Inf}_i[\Psi]}\Big)\;\sum_j\sum_{I\ni j}\psi(I)^2 \\ &\leqslant \ell\,64^\ell\,\big(\boldsymbol{m}_3^{\leqslant M}\big)^\ell\;\Big(\max_i\sqrt{\mathrm{Inf}_i[\Psi]}\Big)\;\sum_{|I|\leqslant\ell}\psi(I)^2\,.\end{aligned} \tag{7.14}$$

Together with bound (7.12), this shows the desired bound (3.3).

Proof of truncation properties (7.9). Let $M > 0$. If $\mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y\leqslant M\}}] = 0$ we are done: just choose $Y^- := Y\mathbb{1}_{\{-M\leqslant Y\leqslant M\}}$ and $Y^+ := Y - Y^-$. If, on the other hand, $\mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y\leqslant M\}}] > 0$ (the strictly negative case is analogous), we set

$$\overline{T} := \sup\{T \in [0, M] : \ \mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y\leqslant T\}}] \leqslant 0\} \in [0, M] .$$

Note that $\mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y\leqslant \overline{T}\}}] \geqslant 0$, because $T \mapsto \mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y\leqslant T\}}]$ is right-continuous. If $\mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y\leqslant \overline{T}\}}] = 0$, defining $Y^- := Y\mathbb{1}_{\{-M\leqslant Y\leqslant \overline{T}\}}$ and $Y^+ := Y - Y^-$, all the properties in (7.9) are clearly satisfied, except the last one that will be checked below. Finally, we consider the case $\mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y\leqslant \overline{T}\}}] > 0$ (then necessarily $\overline{T} > 0$). Since $\mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y< \overline{T}\}}] \leqslant 0$ by definition of $\overline{T}$, we must have $\mathbb{P}(Y = \overline{T}) > 0$. Then take a random variable U uniformly distributed in $(0, 1)$ and independent of Y, and define

$$Y^- := Y\big(\mathbb{1}_{\{-M\leqslant Y<\overline{T}\}} + \mathbb{1}_{\{Y=\overline{T},\, U\leqslant\varrho\}}\big) , \qquad \text{where} \qquad \varrho := \frac{-\mathbb{E}[Y\mathbb{1}_{\{-M\leqslant Y<\overline{T}\}}]}{\overline{T}\,\mathbb{P}(Y = \overline{T})} \in (0, 1) .$$

Setting $Y^+ := Y - Y^-$, all the properties (7.9) but the last one are clearly satisfied.

For the last property, we write

$$\mathbb{E}[(Y^+)^2] = \mathbb{E}[(Y^+)^2\mathbb{1}_{\{|Y|>M\}}] + \mathbb{E}[(Y^+)^2\mathbb{1}_{\{|Y|\leqslant M\}}] = \mathbb{E}[Y^2\mathbb{1}_{\{|Y|>M\}}] + \mathbb{E}[(Y^+)^2\mathbb{1}_{\{|Y|\leqslant M\}}] ,$$

because $Y^+ = Y$ on the event $\{|Y| > M\}$. For the second term, since $0 \leqslant Y^+ \leqslant M$ on the event $\{|Y| \leqslant M\}$, we can write $(Y^+)^2 \leqslant MY^+$ (no absolute value needed). Since $Y^- = Y^-\mathbb{1}_{\{|Y|\leqslant M\}}$ has zero mean by (7.9), we obtain

$$\begin{aligned}\mathbb{E}[(Y^+)^2\mathbb{1}_{\{|Y|\leqslant M\}}] &\leqslant M\,\mathbb{E}[Y^+\mathbb{1}_{\{|Y|\leqslant M\}}] = M\,\mathbb{E}[(Y^+ + Y^-)\mathbb{1}_{\{|Y|\leqslant M\}}]\\ &= M\,\mathbb{E}[Y\mathbb{1}_{\{|Y|\leqslant M\}}] = M\,(-\mathbb{E}[Y\mathbb{1}_{\{|Y|>M\}}]) \leqslant \mathbb{E}[Y^2\mathbb{1}_{\{|Y|>M\}}] ,\end{aligned}$$

where we have used the fact that $\mathbb{E}[Y] = 0$ by assumption, and Markov's inequality. The last relation in (7.9) is proved. □

Acknowledgements

I would like to thank the Hausdorff Center, Bonn and the organisers of the programme "*New Frontiers in Singular SPDEs and Scaling Limits*", G. Cannizzaro, A. Chandra, D. Erhard, Hao Shen and B.E. Ugurcan for the opportunity to deliver the lecture series based on these notes. I am mostly grateful to my collaborators on these works, Francesco Caravenna and Rongfeng Sun, for both a wonderful collaboration and great friendship, that has made this work both possible and fun.

References

[AKQ14] T. Alberts, K. Khanin, J. Quastel. Intermediate Disorder for $1 + 1$ Dimensional Directed Polymers. *Ann. Probab.* 42, 1212–1256, 2014.

[AL11] A. Auffinger, O. Louidor, Directed polymers in a random environment with heavy tails. *Comm. Pure and Appl. Math.*, 64(2), 183-204, (2011).

[BT18] Q. Berger, N. Torri, Directed polymers in heavy-tail random environment, *Ann. Prob.*, to appear, arXiv:1802.03355, (2018).

[BT19] Q. Berger, N. Torri, Entropy-controlled last-passage percolation, *Ann. Appl. Prob.*, 29(3), 1878-1903, (2019).

[BC95] L. Bertini and N. Cancrini. The stochastic heat equation: Feynman-Kac formula and intermittence. *J. Stat. Phys.* 78, 1377–1401, 1995.

[BC98] L. Bertini and N. Cancrini. The two-dimensional stochastic heat equation: renormalizing a multiplicative noise. *J. Phys. A: Math. Gen.* 31, 615, 1998.

[BG97] L. Bertini, G. Giacomin, Stochastic Burgers and KPZ equations from particle systems, *Comm. Math. Phys.*, 183(3), 571-607, (1997).

[BG16] A. Borodin, V. Gorin, Lectures on integrable probability, *Probability and statistical physics in St. Petersburg,* 91(155-214), (2016).

[BP14] A. Borodin, L. Petrov, Integrable probability: From representation theory to Macdonald processes, *Probability Surveys*, 11, 1-58, (2014).

[CES19] G. Cannizzaro, D. Erhard, P. Schönbauer, 2D Anisotropic KPZ at stationarity: scaling, tightness and non triviality. arXiv preprint arXiv:1907.01530, (2019).

[CSZ17a] F. Caravenna, R. Sun, N. Zygouras, Polynomial chaos and scaling limits of disordered systems. *J. Eur. Math. Soc.* 19 (2017), 1–65.

[CSZ17b] F. Caravenna, R. Sun, N. Zygouras, Universality in marginally relevant disordered systems. *Ann. Appl. Probab.* 27 (2017), 3050–3112.

[CSZ18a] F. Caravenna, R. Sun, N. Zygouras, On the moments of the (2+1)-dimensional directed polymer and stochastic heat equation in the critical window. *Comm. Math. Phys.*, to appear arXiv:1808.03586 (2018).

[CSZ18b] F. Caravenna, R. Sun, N. Zygouras, The two-dimensional KPZ equation in the entire subcritical regime, *Ann. Prob.* to appear,

[CSZ18c] F. Caravenna, R. Sun, N. Zygouras, The Dickman subordinator, renewal theorems, and disordered systems, *Elect. J. Prob., to appear* arXiv:1805.01465, (2018)

[CSZ18+] F. Caravenna, R. Sun, N. Zygouras, Sacling limits of disordered systems and disorder relevance *Proceedings of the XVIII International Congress of Mathematical Physics,* arxiv.1602.05825

[CTT17] F. Caravenna, F. L. Toninelli, N. Torri. Universality for the pinning model in the weak coupling regime. *Ann. Probab.* 45 (2017), 2154–2209.

[C09] S. Chatterjee, Fluctuations of eigenvalues and second order Poincaré inequalities, *Prob. Th. Rel. Fields*, 143, 1-40, (2009)

[CD18] S. Chatterjee, A. Dunlap, Constructing a solution of the $(2 + 1)$-dimensional KPZ equation, *Ann. Prob.*, to appear, arXiv:1809.00803 (2018).

[C19a] J. T. Clark, Continuum models of directed polymers on disordered diamond fractals in the critical case. arXiv preprint arXiv:1908.07120, (2019).

[C19b] J. T. Clark, Weak-disorder limit at criticality for directed polymers on hierarchical graphs, arXiv preprint arXiv:1908.06555, (2019).

[C19c] J. T. Clark, The conditional Gaussian multiplicative chaos structure underlying a critical continuum random polymer model on a diamond fractal. arXiv preprint arXiv:1908.08192, (2019)

[CSY04] F. Comets, T. Shiga, N. Yoshida, Probabilistic analysis of directed polymers in a random environment: a review. *Stochastic analysis on large scale interacting systems Mathematical Society of Japan*, 115-142, (2004).

[CCM19] F. Comets, C. Cosco, C. Mukherjee, Renormalizing the Kardar-Parisi-Zhang equation in $d \geqslant 3$ in weak disorder. arXiv preprint arXiv:1902.04104, (2019).

[CCM18] F. Comets, C. Cosco, C. Mukherjee. Fluctuation and rate of convergence for the stochastic heat equation in weak disorder. arXiv:1807.03902 (2018).

[C12] I. Corwin. The Kardar-Parisi-Zhang equation and universality class. *Random Matrices: Theory and Applications*, 1 (2012), 1130001.

[DFT94] G.F. Dell'Antonio, R. Figari, A. Teta, Hamiltonians for systems of N particles interacting through point interactions, *Annales de l'IHP Physique théorique*, Vol. 60, No. 3, pp. 253-290, (1994).

[DZ16] P.S. Dey, N.Zygouras; High temperature limits of directed polymers with heavy tail disorder, *Annals of Prob.*, 2016, Vol. 44, No. 6, 4006-4048,

[DGLT09] B. Derrida, G. Giacomin, H. Lacoin, F. L. Toninelli, Fractional moment bounds and disorder relevance for pinning models, *Comm. Math. Phys.*, 287(3), 867-887, (2009).

[DR04] J. Dimock, S. G. Rajeev, Multi-particle Schrödinger operators with point interactions in the plane, *J. of Physics A: Mathematical and General*, 37(39), 9157, (2004)

[DGRZ18] A. Dunlap, Y. Gu, L. Ryzhik, O. Zeitouni, Fluctuations of the solutions to the KPZ equation in dimensions three and higher. arXiv preprint arXiv:1812.05768 (2018).

[G11] G. Giacomin, Disorder and Critical Phenomena Through Basic Probability Models: École d'Été de Probabilitś de Saint-Flour XL-2010 (Vol. 2025), *Springer*,(2011).

[GJ14] P. Goncalves, M. Jara, Nonlinear fluctuations of weakly asymmetric interacting particle systems, *Archive for Rational Mechanics and Analysis*, 212(2), 597-644, (2014).

[GJ10] P. Goncalves, M. Jara, Universality of KPZ equation. arXiv preprint arXiv:1003.4478, (2010).

[G18] Y. Gu, Gaussian fluctuations from the 2D KPZ equation, *Stoch. and Partial Differential Equations: Analysis and Computations*, 1-36, (2018).

[GQT19] Y. Gu, J. Quastel, L.C. Tsai, Moments of the 2D SHE at criticality. arXiv preprint arXiv:1905.11310, (2019).

[GIP15] M. Gubinelli, P. Imkeller, N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics*, Pi (2015) Vol. 3, no. 6.

[H14] M. Hairer. A theory of regularity structures. *Inventiones Math.* 198, 269–504, 2014.

[HM07] B. Hambly, J. B. Martin, Heavy tails in last-passage percolation, *Prob. Theory and Rel. Fields*, 137:227-275, (2007)

[H74] A.B. Harris, Effect of random defects on the critical behaviour of Ising models. *Journal of Physics C: Solid State Physics*, 1974.

[dJ87] P. de Jong. A central limit theorem for generalized quadratic forms. *Prob. Th. Rel. Fields* 75, 261–277, 1987.

[dJ90] P. de Jong. A central limit theorem for generalized multilinear forms. *J. Multivariate Anal.* 34, 275–289, 1990.

[J97] S. Janson, Gaussian hilbert spaces, *Cambridge University Press*, Vol. 129, (1997).

[KPZ86] M. Kardar, G. Parisi, Y.-C. Zhang. Dynamic Scaling of Growing Interfaces. *Physical Review Letters* 56 (1986), 889–892.

[K14] A. Kupiainen. Renormalisation Group and Stochastic PDE's. arXiv:1410.3094, 2014.

[Led96] M. Ledoux. Isoperimetry and Gaussian analysis. Chapter 3 (pp. 165-294) in *Lectures on Probability Theory and Statistics, Ecole d'Eté de Probabilités de Saint-Flour XXIV–1994*, Springer (1996).

[Led01] M. Ledoux. The concentration of measure phenomenon. Mathematical Surveys and Monographs, Vol. 89, *American Mathematical Society* (2001).

[L19] D. Lygkonis, Fluctuations of the directed polymer partition function in dimensions $d \geqslant 2$, MSc thesis, University of Warwick, (2019).

[MU18] J. Magnen, J. Unterberger. The scaling limit of the KPZ equation in space dimension 3 and higher. *J. Stat. Phys.* 171 (2018), Volume 171, 543–598.

[MOO10] E. Mossel, R. O'Donnell, K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. Math* 171 (2010), 295–341.

[MSZ16] C. Mukherjee, A. Shamov, O. Zeitouni, Weak and strong disorder for the stochastic heat equation and continuous directed polymers in $d \geqslant 3$. *Elect.. Comm. Probab.* 21, no. 61, (2016),

[NP09] I. Nourdin, G. Peccati, Stein's method on Wiener chaos, *Prob. Th. Rel. Fields*, 145(1-2), 75-118, (2009).

[NPR10] I. Nourdin, G. Peccati, G. Reinert. Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *Ann. Probab.* 38, 1947-1985, (2010).

[NP05] D. Nualart, G. Peccati, Central limit theorems for sequences of multiple stochastic integrals, Ann. Prob., 33(1), 177-193, (2005).

[QS15] J. Quastel, H. Spohn. The one-dimensional KPZ equation and its universality class. *J. Stat. Phys.* 160 (2015), 965–984.

[R74] V. I. Rotar', Some limit theorems for polynomials of second degree, *Theory of Probability & Its Applications*, 18(3), 499-507, (1974).

[S61] B. A. Sevastyanov, A class of limit distributions for quadratic forms of normal stochastic variables *Theor. Probab. Appl,* 6, 337-340, (1961).

[S98] J. Szulga, Introduction to random chaos, *Chapman & Hall, London*, (1998).

[Ten95] G. Tenenbaum. Introduction to Analytic and Probabilistic Number Theory. *Cambridge University Press* (1995).

[T17] F. L. Toninelli. (2 + 1)-dimensional interface dynamics: mixing time, hydrodynamic limit and Anisotropic KPZ growth. To appear in *Proceedings of the International Congress of Mathematicians 2018*, arXiv:1711.0557.

[Z18] N. Zygouras Some algebraic structures in the KPZ universality, arXiv:1812.07204, (2018)