APPETIZERS IN SHE/KPZ

YU GU

Abstract. This is the lecture note I prepared for the SPDE online summer school, based on the joint works with Guillaume Bal, Alexander Dunlap, Tomasz Komorowski, Jiawei Li, Lenya Ryzhik, Li-Cheng Tsai, and Ofer Zeitouni. The central theme is to study the long time behaviors of a heat equation with a spacetime random potential that decorrelates sufficiently rapidly.

We start with a random potential of size ε. The central limit theorem gives the first time scale ε^{-2} on which one observes nontrivial effects, which is a result of stochastic homogenization. A natural question is how far one can go beyond the time scale of ε^{-2}. In $d = 1$ we discuss the Wong-Zakai theorem of Hairer-Pardoux extending it to ε^{-4}. In $d \geq 3$, we present a work proving the Edwards-Wilkinson limit on a "very long" time scale, with a weak yet order $O(1)$ potential. The results are then extended to the Kardar-Parisi-Zhang (KPZ) equation. In the end we discuss one of the intermittent cases, and prove a central limit theorem for the KPZ equation on the torus.

Contents

1. Introduction

We are interested in the heat equation with a spacetime random potential

$$\partial_t u = \tfrac{1}{2}\Delta u + uV(t,x), \qquad t>0, x\in\mathbb{R}^d. \tag{1.1}$$

Here V is a smooth stationary Gaussian process that decorrelates sufficiently rapidly. Without loss of generality we assume $\mathbb{E}V(t,x)=0$ and denote its covariance function by

$$\mathbb{E}V(t,x)V(s,y) = R(t-s,x-y), \tag{1.2}$$

where R is a smooth function that decays fast.

Assuming $u(0,x)\geq 0$ and is not identically zero, then define $h(t,x)=\log u(t,x)$, which solves

$$\partial_t h = \tfrac{1}{2}\Delta h + \tfrac{1}{2}|\nabla h|^2 + V(t,x), \qquad t>0, x\in\mathbb{R}^d. \tag{1.3}$$

When V is a spacetime white noise, i.e., $R(t-s,x-y)=\delta(t-s)\delta(x-y)$, this is the so-called KPZ equation, a default model of surface growth subjected to random perturbations. The conjecture in $d=1$ is

$$\varepsilon h(\tfrac{t}{\varepsilon^3},\tfrac{x}{\varepsilon^2}) + C_\varepsilon t \Rightarrow \text{KPZ fixed point}, \tag{1.4}$$

which has actually been proved recently when V is a spacetime white noise [48, 51].

1.1. Weak noise.

We will be less ambitious here and first study the case of a weak potential. Let $\varepsilon\ll 1$, and consider

$$\partial_t u = \tfrac{1}{2}\Delta u + \varepsilon uV(t,x). \tag{1.5}$$

One should think of (1.5) as the microscopic model with ε being a fixed parameter which describes the strength of the disorder. The question is on which scale one can observe the nontrivial effects from the random potential V. It is clear that we should wait for a long time for the weak randomness to accumulate to become something of order $O(1)$: for fixed (t,x), $u(t,x)\to\bar{u}(t,x)$ with $\bar{u}$ solving the unperturbed heat equation $\partial_t\bar{u}=\frac{1}{2}\Delta\bar{u}$.

If we consider the time integral of the potential

$$X(t,x) = \int_0^t \varepsilon V(s,x)ds,$$

a simple calculation shows that

$$\begin{aligned}\mathrm{Var}[X(\tfrac{t}{\varepsilon^2},x)] = \mathbb{E}[X(\tfrac{t}{\varepsilon^2},x)^2] &= \varepsilon^2\int_{[0,t/\varepsilon^2]^2} R(s_1-s_2,0)ds_1ds_2\\ &\to t\int_{\mathbb{R}} R(s,0)ds,\end{aligned} \tag{1.6}$$

which implies that the first time scale should be ε^{-2}. After all, the above calculation shows that if $t\ll\varepsilon^{-2}$, we have $\mathbb{E}[X(t,x)^2]\ll 1$, while if $t\gg\varepsilon^{-2}$, $\mathbb{E}[X(t,x)^2]\gg 1$. Note that we do not consider the degenerate case of $\int R(s,0)ds=0$. The scale ε^{-2} is by no means surprising, and is merely a result of the classical central limit theorem (CLT). Since we considered Gaussian potential, the calculation already shows the convergence of $X(\frac{t}{\varepsilon^2},x)$ to some Gaussian distribution. The same conclusion holds provided that we assume V satisfies some mixing property.

The above simpleminded discussion indicates that one can consider the diffusively rescaled solution $u_\varepsilon(t,x)=u(\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})$, which satisfies

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon}V(\tfrac{t}{\varepsilon^2},\tfrac{x}{\varepsilon})u_\varepsilon, \tag{1.7}$$

and ask about the limit of u_ε and the role played by the large highly-oscillatory random potential $\frac{1}{\varepsilon}V(\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})$. A result of Pardoux-Piatnitski [46] (for more general V which is not necessarily Gaussian) shows that

$$u_\varepsilon(t,x) \to u_{\hom}(t,x) \tag{1.8}$$

with

$$\partial_t u_{\hom}(t,x) = \tfrac{1}{2}\Delta u_{\hom}(t,x) + c_1 u_{\hom}(t,x), \tag{1.9}$$

where $c_1 > 0$ is a positive constant that can be expressed explicitly in terms of R and the Gaussian kernel. In other words, the heterogeneous coefficient $\frac{1}{\varepsilon}V(\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})$ can be approximated by the homogenized coefficient c_1 when $\varepsilon \ll 1$. This is a homogenization type result which can be either interpreted from the probabilistic or analytic perspective, and we will discuss it in Section 2.

One can ask what happens beyond the homogenization/CLT regime. Given any parameter $\alpha > 0$, define $u_\varepsilon(t,x) = u(\frac{t}{\varepsilon^\alpha},\frac{x}{\varepsilon^{\alpha/2}})$, which solves

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon^{\alpha-1}}V(\tfrac{t}{\varepsilon^\alpha},\tfrac{x}{\varepsilon^{\alpha/2}})u_\varepsilon. \tag{1.10}$$

Again we considered the diffusive scaling which is natural here since we are working on the heat equation and viewing the randomness as a perturbation, but it is worth emphasizing that this is only the case in the perturbative regime, as can be seen from (1.4), where $x \sim t^{2/3}$ is the right scaling. When $\alpha = 2$, the Pardoux-Piatnitski result shows that the random potential is being replaced by a positive constant in the limit. Now the question is, what happens for $\alpha > 2$? We first write the above equation as

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon^{\alpha/2-1}}\left(\tfrac{1}{\varepsilon^{\alpha/2}}V(\tfrac{t}{\varepsilon^\alpha},\tfrac{x}{\varepsilon^{\alpha/2}})\right)u_\varepsilon.$$

In light of the homogenization result, the large highly-oscillatory random potential $\frac{1}{\varepsilon^{\alpha/2}}V(\frac{t}{\varepsilon^\alpha},\frac{x}{\varepsilon^{\alpha/2}})$ should be replaced by c_1 as $\varepsilon \to 0$ (compared to (1.7), it is only a notational change of $\varepsilon \to \varepsilon^{\alpha/2}$). Because $\alpha > 2$, there is a diverging factor $\frac{1}{\varepsilon^{\alpha/2-1}}$, so it is tempting to write the approximate equation as

$$\partial_t u_\varepsilon \approx \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon^{\alpha/2-1}}c_1 u_\varepsilon,$$

which is wrong, as we will see later, but in any case it indicates that one needs to renormalize in this case, or in other words, to subtract a diverging constant.

A different way (from the calculation in (1.6)) to guess a *canonial* scale is to look at the limit of the driving force. In (1.10), let $V_\varepsilon(t,x) = \frac{1}{\varepsilon^{\alpha-1}}V(\frac{t}{\varepsilon^\alpha},\frac{x}{\varepsilon^{\alpha/2}})$, which is again a zero-mean stationary Gaussian process. To determine its possible limit as $\varepsilon \to 0$, we only need to compute the covariance function:

$$R_\varepsilon(t-s,x-y) = \mathbb{E}[V_\varepsilon(t,x)V_\varepsilon(s,y)] = \tfrac{1}{\varepsilon^{2\alpha-2}}R(\tfrac{t-s}{\varepsilon^\alpha},\tfrac{x-y}{\varepsilon^{\alpha/2}}).$$

When $\alpha = \frac{4}{2-d}$, we have

$$2\alpha - 2 = \alpha + \tfrac{\alpha d}{2} \tag{1.11}$$

and by further imposing that $\alpha > 0$ (which is natural in our problem since we are considering the long time behavior) so that we stay in $d = 1$, it leads to $\alpha = 4$ and

$$R_\varepsilon(t-s,x-y) = \tfrac{1}{\varepsilon^6}R(\tfrac{t-s}{\varepsilon^4},\tfrac{x-y}{\varepsilon^2}) \to \sigma^2\delta(t-s)\delta(x-y),$$

with $\sigma^2 = \int R(t,x)dxdt$. Again we will consider the non-degenerate case of $\sigma^2 > 0$. In other words, in $d = 1$ and if we consider the time scale ε^{-4}, V_ε scales to a spacetime white noise (with coefficient σ). This indicates that it might be possible to consider the problem

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon^3}V(\tfrac{t}{\varepsilon^4},\tfrac{x}{\varepsilon^2})u_\varepsilon \tag{1.12}$$

and send $\varepsilon \to 0$ to obtain the SHE with a multiplicative spacetime white noise ξ

$$\partial_t \mathcal{U} = \tfrac{1}{2}\Delta \mathcal{U} + \sigma \mathcal{U} \xi. \tag{1.13}$$

From the discussion that leads to the need for renormalization, we know that it can not be true that $u(\frac{t}{\varepsilon^4}, \frac{x}{\varepsilon^2}) \Rightarrow \mathcal{U}(t,x)$. Indeed, what has been proved by Hairer-Pardoux [38] (for more general semilinear equations) is, there exists $c_1, c_2 > 0$ such that as $\varepsilon \to 0$,

$$u(\tfrac{t}{\varepsilon^4}, \tfrac{x}{\varepsilon^2}) \exp(-\tfrac{c_1 t}{\varepsilon^2} - c_2 t) \Rightarrow \mathcal{U}(t,x), \tag{1.14}$$

and the c_1 constant is the same homogenization constant appearing in (1.9). To see the renormalization more clearly, one can write down the equation satisfied by $u_\varepsilon(t,x) = u(\frac{t}{\varepsilon^4}, \frac{x}{\varepsilon^2}) \exp(-\frac{c_1 t}{\varepsilon^2} - c_2 t)$

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon^3} V(\tfrac{t}{\varepsilon^4}, \tfrac{x}{\varepsilon^2}) u_\varepsilon - (\tfrac{c_1}{\varepsilon^2} + c_2) u_\varepsilon. \tag{1.15}$$

This is viewed as a Wong-Zakai type result, since in the limiting SPDE (1.13), the product between $\mathcal{U}$ and ξ is interpreted in the Itô sense, which is actually the only way to interpret it, and before passing to the limit, we have the "physical" product $\frac{1}{\varepsilon^3} V(\frac{t}{\varepsilon^4}, \frac{x}{\varepsilon^2}) u_\varepsilon$. The renormalization term $(\frac{c_1}{\varepsilon^2} + c_2) u_\varepsilon$ can be viewed as the Wong-Zakai correction. We will discuss the proof of (1.14) in more detail in Section 3.

Given (1.14), one can ask two further questions in $d = 1$. (i) What happens when $\alpha \in (2,4)$? The two boundary cases of $\alpha = 2, 4$ correspond to homogenization and convergence to SHE respectively, so one might expect a transition in between. Here is a heuristic picture: on the time scale ε^{-2}, although the overall effects from the random potential is being averaged into a deterministic constant c_1, there are still small random fluctuations (which goes to zero as $\varepsilon \to 0$). As time goes by, those random fluctuations accumulate and finally become of order $O(1)$ when we reach the time scale ε^{-4}, and they combine together to build up the solution to (1.13). Between ε^{-2} and ε^{-4}, the random fluctuations are still small so we are unable to observe any random limit, however, the deterministic effects are overwhelming and requires a renormalization. After one subtracts a diverging constant, what remains is simply the standard heat equation with no perturbation, and it turns out that the small random fluctuations satisfy a central limit theorem after a rescaling. (ii) Another question is on the regime $\alpha > 4$, which is in some sense related to studying the long time behavior of (1.13) (provided one takes a two-step limit), and one would expect results of the type of (1.4).

The relation (1.11) and the need for $\alpha > 0$ restricts our study to $d = 1$, and this is related to the fact that we started from a potential of size ε and considered the scaling depending on ε. To obtain the SHE (1.13) in $d \geq 2$ through a formal scaling, one needs to put $\varepsilon^0 = 1$ in $d = 2$ or ε to some negative power in $d \geq 3$ in front of V before the scaling. It indicates that in high dimensions maybe one can consider "longer" time scales. This is not a very precise statement, and we emphasize here that the length of the time scale itself is relative to the size of the potential. For example, if the potential is of size $O(1)$, there is no reason for differentiating which time scale is longer. The ε^{-2} or ε^{-4} in the previous discussion is only because we have a potential of size ε.

1.2. **Strong noise.** Without the ε prefactor, one is left with the equation (1.1) and immediately faces the problem of intermittency, which usually comes with a dynamic system with a noisy perturbation injected in a multiplicative way. Consider the simplest one dimensional example of geometric Brownian motion:

$$dZ_t = Z_t dB_t, \qquad Z_0 = 1,$$

where B is a standard Brownian motion. The solution is

$$Z_t = e^{B_t - \frac{1}{2}t},$$

which decays to zero exponentially fast, almost surely as $t \to \infty$, so one do not observe "interesting" long time behaviors of Z_t. However, the moments of the solution grows exponentially fast in t: for any $n \in \mathbb{Z}_+$,

$$\mathbb{E}[Z_t^n] = \mathbb{E}[e^{nB_t - \frac{nt}{2}}] = e^{\frac{1}{2}n(n-1)t} \tag{1.16}$$

and we define $\gamma_n = \frac{1}{2}n(n-1)$, so $\frac{\gamma_n}{n} = \frac{1}{2}(n-1)$ increases with respect to n, fitting for a notion of *intermittency*, see a beautiful introduction in [40]. In many cases, the moments are supposed to describe the statistical properties of the random variable – if they do not grow very fast, it determines the distribution. This is not the case here: we know that $Z_t \to 0$ almost surely but its moments grow exponentially fast, and the growth is getting faster and faster as the order increases. It indicates that in most of the probability space, the solution is close to zero, and only a few rare events under which the solution is unreasonably large contribute to the moments. In some sense, the moment calculation as in (1.16), and more complicated cases of SHE, only gives the (upper) tail information of the distribution we are interested in. If we are concerned with the "bulk", for example in this simple example one can write

$$\frac{\log Z_t + \frac{1}{2}t}{\sqrt{t}} = \frac{B_t}{\sqrt{t}} = N(0,1)$$

and pretend that we do not know the explicit solution of geometric Brownian motion, then the above expression provides a more detailed description of Z_t when t is large: rather than simply stating that $Z_t \to 0$ exponentially fast, we know that the exponent satisfies a central limit theorem after a centering. This is the prototype of example one should keep in mind concerning intermittency in the context of SPDE. For more complicated problems such as

$$\partial_t u = \tfrac{1}{2}\Delta u + uV(t,x), \tag{1.17}$$

we would like to do the same and ask

$$\frac{\log u(t,x) + ?t}{t^?} \Rightarrow ?$$

The first "?" is on how fast the solution decays/grows to infinity and the constant is sometimes referred to as the quenched Lyapunov exponent; the second is on the size of the fluctuation and one expects that in $d = 1$ the exponent is $1/3$; the third is on the asymptotic limiting distribution and for certain special initial data it is related to the largest eigenvalue of special random matrices (again in $d = 1$). The fact that the solution decays or grows exponentially indicates that all the interesting and meaningful evolutions happen on the level of the exponent. After taking the logarithm, one is left with the KPZ equation.

Strong noise: the only non-intermittent case. To study $\log u$ is a challenging problem in the intermittent setting, which we will not touch in this note (except when the equation is posed on the torus). There is one non-intermittent case when the equation is posed on the whole space. To see it clearly, we introduce a parameter β into (1.17) and consider

$$\partial_t u = \tfrac{1}{2}\Delta u + \beta uV(t,x). \tag{1.18}$$

We emphasize that β is not the ε in (1.5), and we will not rescale the equation according to β. It is just a fixed parameter and sometimes we interpret it as the inverse temperature if the solution to (1.18) is viewed as the partition function of a directed polymer in a random environment. For the sake of discussion, we assume V is white in time with a smooth spatial covariance function $R(\cdot)$, and the

product between u and V is interpreted in the Itô sense. Studying the white-in-time potential is technically simpler than the case when V is colored in time, but one expect similar *qualitative* behaviors. For the problem to be non-intermittent, one may require the moments do not grow exponentially fast. We consider flat initial data $u(0,x) \equiv 1$ for simplicity, and the second moment of u admits a Feynman-Kac representation

$$\mathbb{E}[u(t,x)^2] = \mathbb{E}_B[\exp(\beta^2 \int_0^t R(B_s^1 - B_s^2)ds)],$$

where B^1, B^2 are two independent standard Brownian motions and $\mathbb{E}_B$ is the expectation with respect to the Brownian motions. Note that the r.h.s. does not depend on x because we start from flat initial data. The above expression is related to the heat equation with a deterministic potential

$$\partial_t f = \Delta f + \beta^2 f R(x). \tag{1.19}$$

Actually by the Feynman-Kac formula, one easily checks that $\mathbb{E}[u(t,x)^2] = f(t,x)$ (if we start from $f(0,x) \equiv 1$). Therefore, the long time behavior of $\mathbb{E}[u(t,x)^2]$ is associated with the spectral property of the operator $\Delta + \beta^2 R(\cdot)$. From the probabilistic point of view, if we assume R is compactly supported (which we will do throughout the note), the term $\int_0^t R(B_s^1 - B_s^2)ds$ measures the amount of "intersection time" of two independent Brownian motions (or the amount of time a Brownian motion with variance 2 spending near the origin). It is a simple exercise to check that in $d \geq 3$, because the Brownian motion is transient, the random variable $\int_0^\infty R(B_s^1 - B_s^2)ds$ has an exponential tail, and

$$\sup_{t\geq 0} \mathbb{E}[u(t,x)^2] = \mathbb{E}_B[\exp(\beta^2 \int_0^\infty R(B_s^1 - B_s^2)ds)] < \infty, \qquad \text{if } \beta < \beta_c$$

for some critical value β_c. From the spectral point of view, the operator $\Delta + \beta^2 R(\cdot)$ has no positive eigenvalue in this case.

In this case ($d \geq 3$ and $\beta < \beta_c$), because the second moment $\mathbb{E}[u(t,x)^2]$ stabilizes as $t \to \infty$, one may expect that the solution $u(t)$, viewed as a Markov process, converges to a nontrivial equilibrium ($u = 0$ is always an equilibrium, which we consider as trivial). This is the only case in which the SHE (1.18) stabilizes to a nontrivial state. In all other cases, one expects u to go to zero. The local long time behavior of u is then clear: $u(t,\cdot)$ converges to some (non-explicit) stationary distribution as $t \to \infty$. Then one is interested in the more "global" behavior, as a random Schwartz distribution, i.e., when tested against smooth functions, what is the long time behavior of u?

To study the "global" behavior of u in the non-intermittent case, a first question is on the spatial correlation scale, i.e., for $t \gg 1$, what is the critical length scale ℓ_c in the x variable so that

$$\mathrm{Cor}[u(t,x), u(t,y)] \approx 0, \text{ if } |x-y| \gg \ell_c$$
$$\mathrm{Cor}[u(t,x), u(t,y)] \approx 1, \text{ if } |x-y| \ll \ell_c$$

It turns out that the spatial correlation length is $\ell_c = \sqrt{t}$, which can be seen from either the probabilistic representation through the intersection of Brownian motions or the PDE (1.19). Then one introduce a parameter $\varepsilon \ll 1$ and define

$$u_\varepsilon(t,x) = u(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}).$$

The problem reduces to studying the asymptotics of u_ε as a random Schwartz distribution. The ε here is different from the previous one which is used to describe the strength of the potential, and it is just a scaling parameter with no physical

meaning. In this case it makes sense to consider the initial data that varies on the scale of ε^{-1}: $u(0,x) = u_0(\varepsilon x)$ so $u_\varepsilon(0,x) = u_0(x)$.

Since we are interested in the random fluctuations, the problem reduces to studying the asymptotics of the random variable

$$\int [u_\varepsilon(t,x) - \mathbb{E}u_\varepsilon(t,x)]g(x)dx.$$

We consider a spatial test function g here, but one can also take a spacetime test function. Since the equation (1.18) is interpreted in the Itô sense, we have $\mathbb{E}u_\varepsilon(t,x) = \bar{u}(t,x)$, with $\bar{u}$ solving the unperturbed heat equation

$$\partial_t \bar{u} = \tfrac{1}{2}\Delta\bar{u}, \qquad \bar{u}(0,x) = u_0(x)$$

and the random fluctuation satisfies the equation

$$\partial_t(u_\varepsilon - \bar{u}) = \tfrac{1}{2}\Delta(u_\varepsilon - \bar{u}) + \tfrac{\beta}{\varepsilon^2}V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon})u_\varepsilon$$

By the same discussion as in (1.11), we can write

$$\tfrac{1}{\varepsilon^2}V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}) = \varepsilon^{d/2-1}\left(\tfrac{1}{\varepsilon^{1+d/2}}V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon})\right)$$

with the term in the bracket scales to a spacetime white noise, so we have

$$\partial_t\left(\tfrac{u_\varepsilon - \bar{u}}{\varepsilon^{d/2-1}}\right) = \tfrac{1}{2}\Delta\left(\tfrac{u_\varepsilon - \bar{u}}{\varepsilon^{d/2-1}}\right) + \tfrac{\beta}{\varepsilon^{1+d/2}}V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon})u_\varepsilon,$$

and it remains to study the limit of the second term on the r.h.s., which we rewrite as

$$\tfrac{1}{\varepsilon^{1+d/2}}U(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}) := \tfrac{1}{\varepsilon^{1+d/2}}V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon})u(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}).$$

Assume for the moment that $u_0 \equiv 1$ so $\bar{u} \equiv 1$, then by the previous discussion on the local behavior of u, we know that $u(t,x)$ converges in distribution as $t \to \infty$ to some spacetime stationary solution to (1.18). Heuristically, the amount of time it takes for u to reach the (local) equilibrium is only *microscopically* large, therefore, we can approximate the $u(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$ in the above expression by the spacetime stationary solution evaluated at $(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$ when ε is small, then it is not hard to imagine that $\frac{1}{\varepsilon^{1+d/2}}U(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$ scales to a spacetime white noise, whose covariance structure would encode the information from the spacetime stationary solution. In the general case when u_0 and $\bar{u}$ are not constants, we have a similar approximation of the form

$$\tfrac{1}{\varepsilon^{1+d/2}}U(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}) \approx \text{ spacetime white noise } \times \bar{u}(t,x).$$

The main result says that, as Schwartz distributions,

$$\tfrac{u_\varepsilon - \bar{u}}{\varepsilon^{d/2-1}} \Rightarrow \mathcal{U} \tag{1.20}$$

with

$$\partial_t\mathcal{U} = \tfrac{1}{2}\Delta\mathcal{U} + \beta\nu_{\text{eff}}\xi\bar{u} \tag{1.21}$$

where ξ is a standard spacetime white noise and ν_{eff}^2 is the effective variance which results from the resonance between u_ε and the driving force and incorporates the information of the local equilibrium. The SPDE satisfied by $\mathcal{U}$ is a stochastic heat equation with an additive noise, and the solution is an explicit convolution of the heat kernel with the noise, which is a Gaussian process. It is the so-called Edwards-Wilkinson model, and is an infinite dimensional Ornstein-Uhlenbeck process when written in the Fourier domain. It is one of the default models of Gaussian fluctuations.

The takeaway message from the above discussion is that the local equilibrium statistics plays a crucial role in determine the large scale behavior of the field. The reason why there exists a local equilibrium in this particular example is that we are in a non-intermittent regime ($d \geq 3$ and $\beta < \beta_c$) and the solution to the microscopic model (1.18) (or the corresponding KPZ equation) stabilizes. To prove

something like (1.4) in the intermittent case, one needs to find and make use of other stationary quantities, and in many cases it is the solution to the Burgers equation or its discrete counterpart. Expressed using the solution to SHE, it is

$$\nabla h = \nabla \log u = \tfrac{\nabla u}{u}.$$

In other words, in those intermittent case, u itself goes to zero almost surely but $\frac{\nabla u}{u}$ converges to some stationary random field. Incidentally, when the equation is posed on the torus, another quantity that stabilizes is the endpoint distribution of a directed polymer. In terms of SHE, it is

$$\rho = \tfrac{u}{\int u}.$$

By utilizing the mixing property of ρ, one can prove a Gaussian version of (1.4) for the equation on the torus. This will be discussed in detail in Section 5.

Another interesting question in the non-intermittent case is what happens to $\log u$. Define $h = \log u$ which solves the KPZ type equation

$$\partial_t h = \tfrac{1}{2}\Delta h + \tfrac{1}{2}|\nabla h|^2 + \beta V \tag{1.22}$$

We assumed that V is smooth in the spatial variable so there is no problem with applying Itô formula to rigorously derive the above equation, and there is a constant Itô-Stratonovitch term which we omitted to simplify the presentation. The local statistical behavior of h as $t \to \infty$ is clear: since u converges to some stationary solution, after taking the log which is a continuous operation, we know that h must also converge to a stationary solution. What is more difficult is the global behavior. In other words, if we view h as a random Schwartz distribution, what is the limit? The previous discussion shows that for any test function g,

$$\int \left(\tfrac{u_\varepsilon - \bar{u}}{\varepsilon^{d/2-1}}\right) g \Rightarrow \int \mathcal{U} g \tag{1.23}$$

in distribution. Define $h_\varepsilon(t,x) = h(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}) = \log u_\varepsilon(t,x)$, what we would like to understand is the fluctuation of

$$\int h_\varepsilon g = \int (\log u_\varepsilon) g \tag{1.24}$$

Clearly it is not a trivial exercise to directly use (1.23) to deal with (1.24). After all, we are not supposed to interchange the nonlinear function "log" with the integral sign.

Here is a question that can be asked in a very general context and appears frequently in the context of SPDE. Given a stationary random field W whose properties we know very well, e.g., how fast it decorrelates/mixes etc. Applying a smooth function f so that we obtain another stationary random field $f(W)$, what can we say about the property of this new random field? If W is a Gaussian process, it is clear what one should do: expanding f in Hermite polynomials

$$f(W) = \sum_{n \geq 0} f_n H_n(W) \tag{1.25}$$

so we have by orthogonality that

$$\mathbb{E}[f(W(x))f(W(y))] = \sum_{n \geq 0} f_n^2 \mathbb{E}[H_n(W(x))H_n(W(y))]$$

Then the decorrelation property of W is transferred to $f(W)$ through the above formula. For the non-Gaussian case, which is what we have (u is a non-Gaussian field), it is unclear how one should proceed.

It turns out that under the same condition $\beta < \beta_c$ and $d \geq 3$, the same convergence holds for h_ε

$$\frac{h_\varepsilon - \mathbb{E}h_\varepsilon}{\varepsilon^{d/2-1}} \Rightarrow \mathcal{U}. \tag{1.26}$$

As indicated by the above discussion, the key in proving this result is to study the correlation property of h_ε. This will be done using tools of Malliavin calculus. The idea is to express $h_\varepsilon - \mathbb{E}[h_\varepsilon] = \log u_\varepsilon - \mathbb{E}[\log u_\varepsilon]$ as an Ito integral, which is done using the Clark-Ocone formula and in some sense plays the role of a mild formulation. The integrand is related to the density of a directed polymer, which, in the non-intermittent case ($d \geq 3$ and $\beta < \beta_c$), behaves like a Brownian motion on large scales. We will discuss the proof in detail in Section 4.

The problem of studying h_ε as a random Schwartz distribution is related to the attempt of making sense of the KPZ equation in high dimensions. Suppose we write the formal equation

$$\partial_t H = \tfrac{1}{2}\Delta H + \tfrac{1}{2}\kappa|\nabla H|^2 + \xi$$

where ξ is a spacetime white noise and we put a parameter κ in front of the nonlinear term which will play a role later. In $d = 1$, what has been done is to consider a mollification of ξ on the scale of ε, denoted by ξ_ε, and the corresponding equation on the level of ε with a renormalization

$$\partial_t H_\varepsilon = \tfrac{1}{2}\Delta H_\varepsilon + \tfrac{1}{2}\kappa|\nabla H_\varepsilon|^2 + \xi_\varepsilon - C_\varepsilon$$

If C_ε is chosen properly, then removing the mollification, one can show that H_ε converges with the limit defined as the solution to 1d KPZ which also coincides with the Hopf-Cole solution, which is $\log \mathcal{U}$ with $\mathcal{U}$ solving (1.13). Another way of seeing this is through the Wong-Zakai theorem: taking logarithm on (1.15), the $\frac{1}{\varepsilon^3}V(\frac{t}{\varepsilon^4}, \frac{x}{\varepsilon^2})$ corresponds to ξ_ε and $\frac{c_1}{\varepsilon^2} + c_2$ is the C_ε here. In high dimensions, this is a dead end – no matter how one chooses the renormalization constant, the H_ε does not converge. It turns out that the nonlinear term $|\nabla H_\varepsilon|^2$ is so "large" that a simple centering does not remedy it. One can do a simple calculation to see that the variance of $|\nabla \mathcal{G}\xi_\varepsilon|^2$ blows up as $\varepsilon \to 0$, where we used $\mathcal{G}$ to denote $(\partial_t - \frac{1}{2}\Delta)^{-1}$. This indicates that maybe one should adjust the coupling constant κ: consider

$$\partial_t H_\varepsilon = \tfrac{1}{2}\Delta H_\varepsilon + \tfrac{1}{2}\kappa_\varepsilon|\nabla H_\varepsilon|^2 + \xi_\varepsilon - C_\varepsilon \tag{1.27}$$

and tune the parameter κ_ε until one has some nontrivial limit. Again by plugging the first chaos $\mathcal{G}\xi_\varepsilon$ and requiring the variance to converge, we see that (in $d \geq 3$)

$$\kappa_\varepsilon = \beta\varepsilon^{d/2-1}$$

is a possible choice. Here we kept a (fixed) parameter β.

The connection between H_ε, which solves (1.27), and $h_\varepsilon = \log u_\varepsilon$ is simple: define

$$\tilde{h}_\varepsilon = \tfrac{h_\varepsilon}{\beta\varepsilon^{d/2-1}},$$

then $\tilde{h}_\varepsilon$ solves the same equation as H_ε (modulo the renormalization constant C_ε). This can be seen through rescaling (1.22). By (1.26), we have

$$H_\varepsilon - \mathbb{E}H_\varepsilon \Rightarrow \beta^{-1}\mathcal{U} \tag{1.28}$$

In other words, in $d \geq 3$ and for $\beta < \beta_c$, the fluctuations of (1.27) is given by the Edwards-Wilkinson equation.

What happens on the level of (1.27) is that the ξ_ε converges to ξ and the nonlinear term, after a centering, converges to an independent spacetime white noise. The two noises combine together to produce the effective noise in (1.21).

The convergence to a Gaussian limit is probably boring from a physical point of view. Although the nonlinear term contributes to the limit, it only becomes part of

the noise, and the high order term in the dynamics does not change so we still have the heat operator $\partial_t - \frac{1}{2}\Delta$ in the limit (in the case of a noise that is smooth in time as well, the coefficient in front of Δ is homogenized in the limit, so in this sense it modifies the high order term, but again only its coefficient). What one would like to eventually understand is what happens in one of those intermittent cases, and a slightly more modest goal is to study the behavior of h_ε when $\beta = \beta_c$. There are very few results in this direction and we will not touch them in the note.

Summary. The SHE and KPZ are connected to many objects in probability, PDE, statistical physics, etc. During the past decade, there have been many exciting developments advancing our understanding of the two fundamental random PDEs. What is presented here is only a very small fraction, served as an "appetizer" in a sense. There are excellent monographs and reviews on the subject, from different perspectives and based on various tools, and we refer the interested readers to [40, 21, 22, 47, 49, 34, 52].

Acknowledgement. The work was partially supported by the NSF through DMS-1613301/1807748/1907928. I would like to thank Guillaume Bal and Lenya Ryzhik for their guidances and inspirations.

2. Homogenization

The goal in this section is to study the homogenization of the equation

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon} V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}) u_\varepsilon \tag{2.1}$$

with $u_\varepsilon(0,x) = u_0(x)$. Here V is assumed to be a smooth Gaussian process that decorrelates fast and has a covariance function R. To simplify the notation, sometimes we will write $V_\varepsilon(t,x) = \frac{1}{\varepsilon} V(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$. The main result is that

Theorem 2.1. *For each (t,x), $u_\varepsilon(t,x) \to u_{\rm hom}(t,x)$ in probability, with*

$$\partial_t u_{\rm hom} = \tfrac{1}{2}\Delta u_{\rm hom} + c_1 u_{\rm hom}, \qquad u_{\rm hom}(0,x) = u_0(x)$$

and $c_1 = \int_0^\infty \int q_t(x) R(t,x) dx dt$.

We will sketch two different proofs, based on diagram expansions and the Feynman-Kac formula respectively. The theory of regularity structures [36] and paracontrolled calculus [33] certainly apply to this problem as well, but we will try to avoid introducing new languages.

2.1. Diagramm.

The (classical) solution to (2.1) satisfies

$$u_\varepsilon(t,x) = q_t \star u_0(x) + \int_0^t \int q_{t-s}(x-y) u_\varepsilon(s,y) V_\varepsilon(s,y) dy ds \tag{2.2}$$

where we used $q_t(x)$ to denote the standard heat kernel, i.e., it is the density of $N(0,t)$. We can keep iterating the above integral equation to obtain an infinite series expansion:

$$u_\varepsilon(t,x) = \sum_{n\geq 0} u_{n,\varepsilon}(t,x)$$

with

$$u_{n,\varepsilon}(t,x) = \int_{[0,t]^n_<} \int \left(\prod_{j=0}^{n-1} q_{s_j - s_{j+1}}(y_j - y_{j+1}) V_\varepsilon(s_{j+1}, y_{j+1}) \right) q_{s_n}(y_n - y_{n+1}) u_0(y_{n+1}) dy ds. \tag{2.3}$$

where we used the convention $s_0 = t, y_0 = x$. One can check that the above series converges to the solution u_ε. So in a way, the solution is written explicitly. The problem reduces to passing to the limit term by term (maybe jointly).

Remark 2.2. Here we assumed V is a smooth Gaussian process. If V is white in time and the equation (2.1) is interpreted in the Ito sense, then the same expansion gives the Wiener chaos decomposition of the random variable $u_\varepsilon(t,x)$, provided that we interpreted the integral in (2.3) as the iterated Ito integral. As we will see immediately, the white in time nature could simplify a lot of calculations. A natural question is if we can come up with a different expansion from (2.3) which is more similar to iterated Ito integrals. This will be our approach to the Wong-Zakai theorem.

To show $u_\varepsilon \to u_{\text{hom}}$ in probability, it suffices to prove the stronger result

$$\mathbb{E}u_\varepsilon \to u_{\text{hom}}, \quad \text{and} \quad \mathbb{E}u_\varepsilon^2 \to u_{\text{hom}}^2.$$

We describe the proof of the convergence of the expectation. For each n,

$$\begin{aligned}&\mathbb{E}[u_{n,\varepsilon}(t,x)]\\ &= \int_{[0,t]^n_<} \int \mathbb{E}\left[\prod_{j=0}^{n-1} q_{s_j - s_{j+1}}(y_j - y_{j+1}) V_\varepsilon(s_{j+1}, y_{j+1})\right] q_{s_n}(y_n - y_{n+1}) u_0(y_{n+1}) dy ds.\end{aligned}$$

For the expectation to be nonzero, n needs to be even, because $\{V_\varepsilon(s_j,y_j)\}_{j=1}^n$ are correlated Gaussian random variables of zero mean. The calculation of the moments invokes the Wick formula

$$\mathbb{E}[\prod_{j=1}^n V_\varepsilon(s_j,y_j)] = \sum_\tau \prod_j R_\varepsilon(s_j - s_{\tau(j)}, y_j - y_{\tau(j)})$$

Here we use R_ε to denote the covariance function of V_ε, and the summation is over all pairings of $\{1,\dots,n\}$, denoted by τ. The product in j is over a given point of the pairing (by some convention, e.g., the smaller index). For example, when $n=4$, the r.h.s. of the above expression is

$$\begin{aligned}&R_\varepsilon(s_1-s_2, y_1-y_2) R_\varepsilon(s_3-s_4, y_3-y_4)\\ &+R_\varepsilon(s_1-s_3, y_1-y_3) R_\varepsilon(s_2-s_4, y_2-y_4)\\ &+R_\varepsilon(s_1-s_4, y_1-y_4) R_\varepsilon(s_2-s_3, y_2-y_3)\end{aligned} \tag{2.4}$$

Therefore, the expectation of $u_{n,\varepsilon}$ can be written as

$$\begin{aligned}\mathbb{E}u_{n,\varepsilon}(t,x) = \sum_\tau \int_{[0,t]^n_<} \int \prod_{j=0}^{n-1} q_{s_j-s_{j+1}}(y_j - y_{j+1}) \prod_j R_\varepsilon(s_j - s_{\tau(j)}, y_j - y_{\tau(j)})\\ \times q_{s_n}(y_n - y_{n+1}) u_0(y_{n+1}) dy ds\end{aligned} \tag{2.5}$$

It remains to pass to the limit of the many terms above. It is important to remember that, for the time variable, we are integrating in the simplex $t > s_1 > \ldots > s_n > 0$, so for example we know that $s_1 - s_2$ is smaller than $s_1 - s_3$. As the covariance function R_ε is assumed to be fast decaying, it is not hard to convince ourselves that maybe the main contribution in (2.4) comes from the term $R_\varepsilon(s_1-s_2, y_1-y_2)R_\varepsilon(s_3-s_4, y_3-y_4)$. Similarly, for any even n, the main contribution in (2.5) should come from the pairing $(1,2),(3,4),\dots,(n-1,n)$. This is called the simple diagram, or sometimes the ladder diagram, which in many cases consisting of the main contribution. In many problems where one has no better approach other than a brutal force expansion (e.g. the Schrödinger equation), the main technical efforts are devoted to proving the non-simple diagrams do not contribute in the limit. In our setting, one could make use of the mixing in time variable, i.e., when $t-s$ is large, $V(t,x)$ is independent of $V(s,y)$ (if R is compactly supported), which simplifies a lot of the analysis. When V only depends on the spatial variable, the diagram analysis becomes more involved.

Remark 2.3. When V is assumed to be white in time with a covariance function $\delta(t-s)R(x-y)$, we have $\mathbb{E}u_{n,\varepsilon} = 0$ for $n \geq 1$, which is just a simple property of the Wiener chaos. From the above discussion, it is also clear as we are integrating in the simplex so two time variables can not coincide, and integrating the Dirac function only gives zero.

Let us not go into the technical details of analyzing different diagrams and simply study the ladder one. Take $n = 2$ for example (in this case the ladder diagram is the only one), we have

$$\mathbb{E}u_{2,\varepsilon}(t,x) = \int_0^t \int q_{t-s_1}(x-y_1)q_{s_1-s_2}(y_1-y_2)R_\varepsilon(s_1-s_2,y_1-y_2)q_{s_2}(y_2-y_3)u_0(y_3)dyds$$

Recall that R_ε is the covariance function of $\varepsilon^{-1}V(\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})$, so

$$R_\varepsilon(s_1-s_2,y_1-y_2) = \tfrac{1}{\varepsilon^2}R(\tfrac{s_1-s_2}{\varepsilon^2},\tfrac{y_1-y_2}{\varepsilon})$$

After a change of variable the above integral can be written as

$$\mathbb{E}u_{2,\varepsilon}(t,x) = \int_{\ell_1+\varepsilon^2\ell_2+\ell_3=t} \int q_{\ell_1}(x-y_1)q_{\ell_2}(z)R(\ell_2,z)q_{\ell_3}(y_1-\varepsilon z-y_3)u_0(y_3)dydzd\ell$$

which converges to

$$\begin{aligned}&\int_0^\infty \int q_{\ell_2}(z)R(\ell_2,z)d\ell_2dz \times \int_{\ell_1+\ell_3=t} \int q_{\ell_1}(x-y_1)q_{\ell_3}(y_1-y_3)u_0(y_3)dyd\ell\\ &= c_1\int_0^t \int q_{t-\ell_3}(x-y_1)q_{\ell_3}(y_1-y_3)u_0(y_3)dyd\ell = c_1 t\int q_t(x-y)u_0(y)dy\end{aligned}$$

Roughly speaking, what happens is the factor

$$V_\varepsilon(s_1,y_1)q_{s_1-s_2}(y_1-y_2)V_\varepsilon(s_2,y_2)$$

being averaged into the constant c_1.

By almost exactly the same calculation, we can show the convergence of $\mathbb{E}u_{n,\varepsilon}$. Summing all terms up (and ignoring the non-ladder diagrams which requires a few pages of analysis), we conclude that

$$\sum_{n\geq 0}\mathbb{E}u_{n,\varepsilon}(t,x) \to \sum_{n\geq 0}\tfrac{(ct)^n}{n!}\int q_t(x-y)u_0(y)dy = u_{\mathrm{hom}}(t,x).$$

Similarly, one can study the second moment and show that $\mathbb{E}u_\varepsilon(t,x)^2 \to u_{\mathrm{hom}}(t,x)^2$, which completes the proof of Theorem 2.1.

Remark 2.4. Theorem 2.1 can also be easily proved using the theory of regularity structures or paracontrolled calculus (see in particular [34, Chapter 4]). Rather than writing the solution explicitly as an infinite series, the idea is to explore the continuous dependence of the "solution" on the "coefficient" in a bigger space. In some sense, the two approaches can be compared through the example of solving an ODE: the diagram expansion here corresponds to finding all derivatives of the solution then write the solution of the ODE as an infinite Taylor series, while the theory of regularity structures and paracontrolled calculus aims at building a solution theory through the Picard iteration of some fixed point equation which also leads to a continuous dependence of the solution on the coefficient. On the other hand, it is not hard to imagine that the diagram approach can be adapted to study the Schrödinger equation with a random potential and prove similar results [30], which falls out of the scope of the other theories.

2.2. **Feynman-Kac.** Another instructive approach to prove Theorem 2.1 is through the Feynman-Kac formula, which brings in probabilistic tools and connects to the subject of directed polymers. The solution to (2.1) can be written as

$$u_\varepsilon(t,x) = \mathbb{E}_B[u_0(x+B_t)e^{\int_0^t V_\varepsilon(t-s,x+B_s)ds}]$$

where we used $\mathbb{E}_B$ to represent the expectation with respect to the Brownian motion B that is independent of V. Then the problem reduces to studying the weak limit of $(B_t, Y_\varepsilon(t))$ with

$$Y_\varepsilon(t) = \int_0^t V_\varepsilon(t-s, x+B_s)ds.$$

One needs to be careful here, as the weak convergence does not directly lead to the convergence of u_ε due to the fact that $\mathbb{E}_B$ is only a partial expectation.

The interested quantity is a Brownian motion in a random environment, which should be distinguished from some other random walk/diffusion in random environment where the environment actually affects the walker, e.g. as a drift – here the Brownian motion is independent from V. For this reason, sometimes it is called random walk/diffusion in random scenery. By a time reversal and a scaling, it is equivalent with studying

$$X_\varepsilon(t) = \varepsilon \int_0^{t/\varepsilon^2} V(s, B_s)ds.$$

The main result is the following:

Proposition 2.5. *For any $t > 0$, $(\varepsilon B_{t/\varepsilon^2}, X_\varepsilon(t)) \Rightarrow (Z_1, Z_2)$ with Z_1 independent of Z_2 and $Z_1 \sim N(0,t), Z_2 \sim N(0,\sigma^2 t)$. The variance $\sigma^2 = 2c_1$.*

Remark 2.6. We are not being very precise when stating the weak convergence in Proposition 2.5. The proof we will give below leads to a result that is stronger than annealed convergence (when considering the product probability space) but weaker than the quenched convergence (for almost every realization of V).

Remark 2.7. After passing to the limit, the fact that Z_1 being independent of Z_2 shows that the Brownian motion is lost in X_ε, which is typical in homogenization. In some other cases, the Brownian path is memorized in the limit, showing localization behaviors. In some sense, if we interpret the behaviors of random PDE in terms of the underlying Brownian motion in random environment or more precisely, the directed polymer, then intermittency is related to localization and non-intermittency corresponds to homogenization.

Ignoring the uniform integrability issue, we have

$$\mathbb{E}u_\varepsilon(t,x) = \mathbb{E}\mathbb{E}_B[u_0(x+B_t)e^{\int_0^t V_\varepsilon(t-s,x+B_s)ds}] \to \mathbb{E}[u_0(x+Z_1)e^{Z_2}] = u_{\text{hom}}(t,x).$$

The proof of $\mathbb{E}u_\varepsilon^2 \to u_{\text{hom}}^2$ requires introducing another Brownian motion independent of B and proving a weak convergent result that is similar to Proposition 2.5. We will not go into the detail since it is similar. The rest of the section is to explain and describe the proof of Proposition 2.5.

Roughly speaking, to prove the weak convergence of $X_\varepsilon(t)$, one can either utilize the mixing of V in time or the mixing induced by the Brownian motion. They will lead to slightly different convergences. The approach we use here is through a martingale decomposition and a martingale central limit theorem. A different proof is given in [46] by using the mixing of V. Rather than the mixing assumption, the martingale approach only requires the basic stationarity and ergodicity of the environment, together with the necessary integrability of R so that $\sigma^2 < \infty$.

The idea is to solve a Poisson equation and treat $X_\varepsilon(t)$ as the Ito correction term. Suppose we solve the equation

$$-(\partial_t \Phi + \tfrac{1}{2}\Delta\Phi) = V \tag{2.6}$$

then applying Ito formula leads to

$$\varepsilon\Phi(\tfrac{t}{\varepsilon^2}, B_{t/\varepsilon^2}) - \varepsilon\Phi(0, B_0) = \varepsilon \int_0^{t/\varepsilon^2} \nabla\Phi(s, B_s) dB_s - \varepsilon \int_0^{t/\varepsilon^2} V(s, B_s) ds$$

so we write

$$X_\varepsilon(t) = \varepsilon\Phi(0, B_0) - \varepsilon\Phi(\tfrac{t}{\varepsilon^2}, B_{t/\varepsilon^2}) + \varepsilon \int_0^{t/\varepsilon^2} \nabla\Phi(s, B_s) dB_s$$

If one can show that the two remainder terms $\varepsilon\Phi(\frac{t}{\varepsilon^2}, B_{t/\varepsilon^2}), \varepsilon\Phi(0, B_0)$ vanish as $\varepsilon \to 0$, it might be possible to show that the martingale term converges to our desired limit. To apply martingale CLT and show the convergence of quadratic variation, it is sometimes more convenient to work in a stationary setting in which one can make use of ergodic theorem. The equation (2.6) does not always have a stationary solution, i.e., a solution that is a spacetime stationary random field (it does in $d \geq 3$). For this reason, we add a regularization parameter and consider the regularized Poisson equation

$$\lambda\Phi_\lambda - (\partial_t + \tfrac{1}{2}\Delta)\Phi_\lambda = V \tag{2.7}$$

By a spectral representation, the solution can be written explicitly, and one has the martingale decomposition

$$X_\varepsilon(t) = \varepsilon\Phi_\lambda(0, B_0) - \varepsilon\Phi_\lambda(\tfrac{t}{\varepsilon^2}, B_{t/\varepsilon^2}) + \varepsilon\lambda \int_0^{t/\varepsilon^2} \Phi_\lambda(s, B_s) ds + \varepsilon \int_0^{t/\varepsilon^2} \nabla\Phi_\lambda(s, B_s) dB_s \tag{2.8}$$

A simple calculation shows that

$$\lambda\mathbb{E}[\Phi_\lambda^2] \to 0, \qquad \text{as } \lambda \to 0$$

using which one can prove the three remainders terms on the r.h.s. of (2.8) goes to zero, upon choosing $\lambda = \varepsilon^2$. For the martingale term, the key is that $\nabla\Phi_\lambda$ converges as $\lambda \to 0$ to some stationary random field, denoted by η. It is worth emphasizing that the limit of Φ_λ is not necessarily stationary – in $d = 1, 2$ the Poisson equation (2.6) has no stationary solution. However, it has a stationary gradient in the limit, which is sort of a folklore in homogenization. Thus, we have the martingale approximation

$$X_\varepsilon(t) \approx M_\varepsilon(t) := \varepsilon \int_0^{t/\varepsilon^2} \eta(s, B_s) dB_s$$

with the quadratic variation

$$\langle M_\varepsilon \rangle_t = \varepsilon^2 \int_0^{t/\varepsilon^2} |\eta(s, B_s)|^2 ds.$$

Applying ergodic theorem, we derive that $\langle M_\varepsilon \rangle_t \to t\mathbb{E}[|\eta|^2]$, which implies that $M_\varepsilon(t) \Rightarrow \sqrt{\mathbb{E}[|\eta|^2]} W_t$, where W is a standard Brownian motion. Actually, consider the two-component martingale $(\varepsilon B_{t/\varepsilon^2}, M_\varepsilon(t))$, the same proof shows that the limit are two independent Brownian motions.

Remark 2.8. To apply ergodic theorem, one need to write the quadratic variation $\langle M_\varepsilon \rangle$ in a different way. As η is a stationary random field, there exists a random variable $\tilde{\eta}$ such that $\eta(t, x; \omega) = \tilde{\eta}(\tau_{(t,x)}\omega)$, where τ is a group of measure-preserving transformation associated with the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Here we used ω to

represent the random realization sampled from Ω. Using $\tilde{\eta}$ and τ, the quadratic variation is written as

$$\langle M_\varepsilon \rangle_t = \varepsilon^2 \int_0^{t/\varepsilon^2} \tilde{\eta}(\tau_{(s,B_s)}\omega) ds$$

A canonical object to define here is the "environment seen from particle"

$$\omega_s := \tau_{(s,B_s)}\omega$$

and one should think of ω as representing the environment, so ω_s is the environment viewed in the moving frame of the Brownian motion and time. In this case, one can show that $\{\omega_s\}$ is a Markov process with $\mathbb{P}$ (the original probability measure) being an ergodic invariant measure, so $\langle M_\varepsilon \rangle_t$ converges almost surely.

Polymer. In this section we introduce the model of directed polymer, and explain its connection to heat equation with a random potential and phrase the homogenization result in this context. For an excellent introduction to directed polymers, we refer to the book [19].

The directed polymer is a fundamental statistical physics model with the Gibbs measure given by the weight $\exp(\beta \int_0^t V(s, B_s; \omega) ds)$. Here t is the length of the polymer path, V is the random environment, β is the inverse temperature, and we have specified the dependence of V on the random realization ω.

For each $t > 0$ and random realization ω, we define a probability measure on $C[0,t]$ by tilting the Wiener measure using the aforementioned weight. Denote the expectation under the new measure by $\hat{\mathbb{E}}_{t,\omega}$, then for any bounded continuous function $F : C[0,t] \to \mathbb{R}$,

$$\hat{\mathbb{E}}_{t,\omega} F(B) = \frac{\mathbb{E}_B F(B) e^{\beta \int_0^t V(s,B_s;\omega) ds}}{\mathbb{E}_B e^{\beta \int_0^t V(s,B_s;\omega) ds}} \tag{2.9}$$

Recall that we use $\mathbb{E}$ to denote the expectation with respect to V, $\mathbb{E}_B$ the expectation with respect to B. We emphasize that here the polymer measure is defined as a random measure which depends on the length of the polymer path. One of the most basic question is to ask whether the polymer path is diffusive or not, in the thermodynamic limit (as $t \to \infty$). We can look at the endpoint distribution, formally written as

$$\rho(t,x) = \frac{\mathbb{E}_B \delta(B_t - x) e^{\beta \int_0^t V(s,B_s;\omega) ds}}{\mathbb{E}_B e^{\beta \int_0^t V(s,B_s;\omega) ds}}.$$

A simple calculation shows that

$$\rho(t,x) = \frac{u(t,x;0)}{\int u(t,x';0) dx'}$$

where we use $u(t,x;0)$ to denote the solution to

$$\partial_t u = \tfrac{1}{2}\Delta u + uV, \qquad u(0,x) = \delta(x).$$

Then the question reduces to studying how the density $\rho(t,\cdot)$ spreads in large time.

By Proposition 2.5 (and the sketched proof), we derive that for any bounded continuous $f : \mathbb{R}^d \to \mathbb{R}$,

$$\frac{\mathbb{E}_B f(\varepsilon B_{t/\varepsilon^2}) e^{\beta\varepsilon \int_0^{t/\varepsilon^2} V(s,B_s;\omega) ds}}{\mathbb{E}_B e^{\beta\varepsilon \int_0^{t/\varepsilon^2} V(s,B_s;\omega) ds}} \to \int f(x) q_t(x) dx$$

with q_t the density of $N(0,t)$. In other words, in this homogenization regime where we have a weak potential of size ε and consider the length of the polymer depending on the weakness $t \sim \varepsilon^{-2}$, the polymer path is still diffusive and the law of the rescaled

endpoint, under the polymer measure, converges to $N(0,t)$. This is also related to the discussion in Remark 2.7.

One can ask what happens to the polymer measure if we consider a longer path $t \gg \varepsilon^{-2}$ – eventually we wish to study the measure defined in (2.9) and send $t \to \infty$. This is clearly related to studying the longer time scale of the random PDE. As we will see later, in the non-intermittent regime, the polymer path is always diffusive, the law of the rescaled path being not necessarily a Brownian motion though.

References. For the heat equation with a random potential, homogenization results of the form (2.1) can be found in [46, 28, 27, 8, 7, 42, 39], for time-dependent, time-constant, Gaussian, and non-Gaussian potentials. The error estimates in the homogenization and corresponding central limit theorems are proved in [28, 8, 7, 26]. The corrector and martingale method of study diffusion in random environment can be found in the excellent monograph [41].

3. Wong-Zakai

The goal in this section is to study

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon^3} V(\tfrac{t}{\varepsilon^4}, \tfrac{x}{\varepsilon^2}) u_\varepsilon - C_\varepsilon u_\varepsilon \tag{3.1}$$

with $u_\varepsilon(0,x) = u_0(x)$. Throughout the section, we assume $d = 1$. Without loss of generality, we assume $\int R(t,x)dtdx = 1$. Here is the main result

Theorem 3.1. *Let $C_\varepsilon = \frac{c_1}{\varepsilon^2} + c_2$ with $c_1, c_2 > 0$ defined in Proposition 3.2. For each (t,x), we have $u_\varepsilon(t,x) \Rightarrow \mathcal{U}(t,x)$ in distribution, where $\mathcal{U}$ solves*

$$\partial_t \mathcal{U} = \tfrac{1}{2}\Delta \mathcal{U} + \mathcal{U}\xi, \qquad \mathcal{U}(0,x) = u_0(x) \tag{3.2}$$

where ξ is a $1+1$ spacetime white noise.

The equation (3.2) is the only SHE that can be made sense of with a space-time white noise. The solution is a continuous random field that satisfies the mild formulation

$$\mathcal{U}(t,x) = q_t \star u_0(x) + \int_0^t \int q_{t-s}(x-y)\mathcal{U}(s,y)\xi(s,y)dyds \tag{3.3}$$

where the stochastic integral is interpreted in the Ito-Walsh sense. A good introduction to the solution theory can be found e.g. in [40]. To see why it is impossible to have a function-valued solution in $d \geq 2$, it suffices to consider the formal stochastic integral $\int_0^t \int q_{t-s}(x-y)\xi(s,y)dyds$. For each (t,x) fixed,

$$\int_0^t \int q_{t-s}(x-y)^2 dyds = \int_0^t \tfrac{1}{(4\pi(t-s))^{d/2}} ds = \infty \qquad \text{in } d \geq 2 \tag{3.4}$$

therefore, the first chaos is not defined for fixed (t,x), although it still makes sense as a Schwartz distribution: for any test function g, the random variable

$$\int_0^t \int g(x) q_{t-s}(x-y)\xi(s,y)dydsdx$$

is well-defined. This indicates that to study SHE in $d \geq 2$ (as an equation), we may view it as a random Schwartz distribution. This is a bit different from the study of directed polymer, where the interested quantity is the partition function which is typically considering the polymer path starting at a fixed point. On the other hand, as we are interested in the large scale behavior and the divergence in (3.4) only results from small scale singularity, in high dimensions one can study the equation with a noise that is smooth and decorrelates fast and expect the large scale behavior is universal.

Iterating the mild formulation (3.3), we can write

$$\mathcal{U}(t,x) = \sum_{n\geq 0} \mathcal{U}_n(t,x) = \sum_{n\geq 0} \mathcal{I}_n(f_n(t,x)),$$

which is the so-called Wiener chaos expansion:

(3.5)
$$\mathcal{I}_n(f_n(t,x)) = \int_{[0,t]^n_<} \int \left(\prod_{j=0}^{n-1} q_{s_j - s_{j+1}}(y_j - y_{j+1}) \xi(s_{j+1}, y_{j+1}) \right) q_{s_n}(y_n - y_{n+1}) u_0(y_{n+1}) dy ds.$$

Here we used $f_n(t,x)$ to denote the integrand in the iterated Ito integral:

$$f_n(t,x,\cdot) = \prod_{j=0}^{n-1} q_{s_j - s_{j+1}}(y_j - y_{j+1}) \int q_{s_n}(y_n - y_{n+1}) u_0(y_{n+1}) dy_{n+1}$$

with the convention $s_0 = t, y_0 = x$. One should compare (3.5) with (2.3), which are almost the same expression except that V_ε is replaced by ξ. However, the integral are interpreted in completely different ways. The one in (2.3) is a deterministic Lebesgue integral and the V_ε there is function-valued. Thus, the integral in (2.3) is computed for each realization of the smooth Gaussian field. In (3.5), we first integrate $q_{s_{n-1}-s_n}(y_{n-1} - y_n) \int q_{s_n}(y_n - y_{n+1}) u_0(y_{n+1}) dy_{n+1}$ (as a function of (s_n, y_n)) against ξ, in the domain $[0, s_{n-1}) \times \mathbb{R}$. The resulting random process, indexed by (s_{n-1}, y_{n-1}), is adapted to the filtration generated by ξ. Then we multiply it by $q_{s_{n-2}-s_{n-1}}(y_{n-2} - y_{n-1})$ and integrate against ξ again, in $[0, s_{n-2}) \times \mathbb{R}$, to obtain a process indexed by (s_{n-2}, y_{n-2}). Keep performing the Ito integration, eventually we obtain a process indexed by (t,x), which is the n–th order chaos $\mathcal{I}_n(f_n(t,x))$.

We use the following elementary example to showcase the difference between the two expansions. Consider the Ito SDE

$$dZ_t = Z_t dB_t, \qquad Z_0 = 1 \tag{3.6}$$

the solution is a geometric Brownian motion which satisfies

$$\begin{aligned} Z_t =& 1 + \int_0^t Z_s dB_s = 1 + \int_0^t (1 + \int_0^s Z_\ell dB_\ell) dB_s \\ =& 1 + B_t + \ldots \int_{[0,t]^n_<} dB_{s_n} dB_{s_{n-1}} \ldots dB_{s_1} + \ldots \end{aligned} \tag{3.7}$$

The above series is the Wiener chaos expansion of the random variable $Z_t = e^{B_t - \frac{1}{2}t}$, and the n–th term is essentially the n–th order Hermite polynomial evaluated at B_t. One should think of the expansion in (3.5) as nothing else but an infinite dimensional version of (3.7), where the "1-d" integration operator $\int^t$ is being replaced by the heat semigroup $e^{\Delta t/2}$.

Now we smooth the Brownian motion and define

$$B_\varepsilon(t) = \int \phi_\varepsilon(t-s) B_s ds \tag{3.8}$$

where ϕ_ε is an approximation of identity. Thus, for each $\varepsilon > 0$, B_ε has smooth sample paths, and one can consider the (classical) ODE corresponding to (3.6)

$$Z'_\varepsilon(t) = Z_\varepsilon(t) B'_\varepsilon(t), \qquad Z_\varepsilon(0) = 1. \tag{3.9}$$

We can again iterate the integral formulation of the above ODE and obtain

$$Z_\varepsilon(t) = 1 + B_\varepsilon(t) + \ldots + \int_{[0,t]^n_<} \prod_{j=1}^n B'_\varepsilon(s_j) ds + \ldots \tag{3.10}$$

It is clear that the expansion in (3.10) should correspond to the one in (2.3).

Obviously we do not have $Z_\varepsilon(t) \to Z_t$ as $\varepsilon \to 0$, despite the convergence of $B_\varepsilon(t) \to B(t)$:

$$Z_\varepsilon(t) = e^{B_\varepsilon(t)} \to e^{B_t} \neq e^{B_t - \frac{1}{2}t} = Z_t$$

The extra factor of $\frac{1}{2}t$ is the Ito-Stratonovitch correction. Instead, if we consider the "renormalized" ODE

$$Z'_\varepsilon(t) = Z_\varepsilon(t)B'_\varepsilon(t) - \tfrac{1}{2}Z_\varepsilon(t), \qquad Z_\varepsilon(0) = 1$$

then we have the desired convergence $Z_\varepsilon(t) \to Z_t$.

One should think of Theorem 3.1 as an infinite dimensional version of the above ODE example, in which case the $\frac{1}{2}$ factor is being replaced by the C_ε in (3.1).

3.1. **Wiener Chaos and Stroock formula.** There are only a few ways of proving convergence to SHE. The moments of $\mathcal{U}$ grows fast which does not necessarily determine the distribution. Three methods are (i) the martingale approach [11]; (ii) the chaos expansion [2]; (iii) the theory of regularity structures/paracontrolled calculus. We will not go into the third one. By assuming that V is Markov in time, there could be a martingale approach based on perturbed test function expansions, which we will not go into detail either. So we are left with the chaos expansion.

Before sketching the proof, let us emphasize that the main difficulty in proving Theorem 3.1 lies in the fact that V is colored in time. If we assume V is white in time and smooth in space, then one can interpret the equation (3.1) in the Ito sense, write down the chaos expansion easily and pass to the limit term by term. This is the approach in [10]. In a sense, the approach in [2] is similar: in the discrete setting when the random environment are "i.i.d.", it is "white" in time. When the noise is colored in time, it is more involved to write down the chaos expansion, which is not the expansion obtained in (2.3) or (3.10).

To compare (3.7) with (3.10), the main difference is that the terms in (3.7) are orthogonal to each other, therefore when passing to the limit in $L^2(\Omega)$, they do not interact with each other. In (3.10), this is not the case. For example, the third term $\int_{[0,t]^3_<} B'_\varepsilon(s_3)B'_\varepsilon(s_2)B'_\varepsilon(s_1)ds$ already contains a component in the first chaos. With different terms interacting with each other, it seems highly nontrivial to put the renormalization constants in the right place.

Therefore, we need to come up with a different expansion from (2.3) or (3.10). There is a general result in Malliavin calculus which helps with it, the so-called Stroock formula.

Suppose we have some underlying spacetime white noise ξ, and a random variable X that is measurable with respect to ξ, and is infinitely Malliavin-differentiable. Then the coefficient of the n–th order chaos of X is given by

$$f_n = \tfrac{1}{n!}\mathbb{E}[\mathcal{D}^n X] \tag{3.11}$$

Here we used $\mathcal{D}$ to denote the Malliavin derivative. Using the above expression, we can write $X = \sum_{n\geq 0}\mathcal{I}_n(f_n)$, where $\mathcal{I}_n(f_n)$ is the Ito-Wiener integral of f_n with respect to ξ. For more details, we refer to [45].

Take the ODE for example, suppose we want to write down the chaos expansion of $Z_\varepsilon(t) = e^{B_\varepsilon(t)}$, for some fixed t. If we write $B(t) = \int_0^t \eta(s)ds$ with η the underlying white noise, then

$$B_\varepsilon(t) = \int \phi_\varepsilon(t-s)B_s ds = \int \psi_\varepsilon(t,s)\eta(s)ds,$$

for the explicit deterministic function $\psi_\varepsilon(t,s) = \int \phi_\varepsilon(t-\ell)1_{\ell>s}d\ell$. Then $\mathcal{D}_s B_\varepsilon(t) = \psi_\varepsilon(t,s)$ and we have

$$\mathcal{D}^n_{s_1,\dots,s_n} Z_\varepsilon(t) = e^{B_\varepsilon(t)} \prod_{j=1}^n \psi_\varepsilon(t,s_j)$$

The chaos expansion coefficient is given by

$$f_n(s_1,\dots,s_n) = \tfrac{1}{n!}\mathbb{E}[\mathcal{D}^n_{s_1,\dots,s_n} Z_\varepsilon(t)] = \tfrac{1}{n!} e^{\frac{1}{2}\mathrm{Var}[B_\varepsilon(t)]} \prod_{j=1}^n \psi_\varepsilon(t,s_j).$$

Sending $\varepsilon \to 0$, we obtain the r.h.s. converges to

$$\tfrac{1}{n!} e^{\frac{1}{2}t} \prod_{j=1}^n 1_{[0,t]}(s_j)$$

which, after a symmetrization, becomes $e^{\frac{1}{2}t} 1_{0<s_n<\dots<s_1<t}$, the expansion coefficient of e^{B_t}. Of course this is a toy example in which one can get the coefficient more efficiently, but the point here is that the almost exact same approach works for Theorem 3.1.

Now we can sketch the proof of Theorem 3.1. To simplify the presentation, we consider the flat initial data $u_0 \equiv 1$

Step 1. We first need to figure out what the renormalization constant C_ε is. Since we expect that $u_\varepsilon \Rightarrow \mathcal{U}$, with a possible uniform integrability, this would imply $\mathbb{E}[u_\varepsilon] \to \mathbb{E}[\mathcal{U}]$. For the limiting SHE, the expectation is simple: $\mathbb{E}\mathcal{U}$ just solves the standard heat equation (in the case of $u_0 \equiv 1$, we have $\mathbb{E}\mathcal{U} \equiv 1$). Therefore, in some sense we only need to tune the constant C_ε so that $\mathbb{E}u_\varepsilon$ converges to the right limit. We emphasize that this does not always work. A priori it is unclear whether the limiting SPDE has finite moments – it is not the case for some other problems!

The expectation can be expressed using the Feynman-Kac formula:

$$\begin{aligned}\mathbb{E}u_\varepsilon(t,x) =& \mathbb{E}u_\varepsilon(t,0) = \mathbb{E}\mathbb{E}_B e^{\varepsilon \int_0^{t/\varepsilon^4} V(t/\varepsilon^4 - s, B_s)ds - C_\varepsilon t} \\ =& \mathbb{E}_B e^{\frac{1}{2}\varepsilon^2 \int_{[0,t/\varepsilon^4]^2} R(s_1-s_2, B_{s_1}-B_{s_2})ds - C_\varepsilon t}\end{aligned}$$

Consider the process

$$Y_\varepsilon(t) := \tfrac{1}{2}\varepsilon^2 \int_{[0,t/\varepsilon^4]^2} R(s_1-s_2, B_{s_1}-B_{s_2})ds$$

which in the case when V is white in time is simply $Y_\varepsilon(t) = \frac{1}{2}R(0)\frac{t}{\varepsilon^2}$ (with $R(0)$ here the spatial covariance function evaluated at 0). The colored time makes the analysis more involved. By a change of variable, we rewrite

$$Y_\varepsilon(t) = \varepsilon^2 \int_0^{t/\varepsilon^4} ds_1 \left(\int_0^{s_1} R(s_2, B_{s_1} - B_{s_1-s_2})ds_2 \right).$$

Assuming R is compactly supported, so for s_1 large, the domain of integration for the bracket term is fixed. WLOG assume $R(s,\cdot) = 0$ for $|s| \geq 1$, then

$$Y_\varepsilon(t) = \varepsilon^2 \int_1^{t/\varepsilon^4} ds_1 \left(\int_0^1 R(s_2, B_{s_1} - B_{s_1-s_2})ds_2 \right) + O(\varepsilon^2)$$

By the property of Brownian motion, the process $\chi(s_1) := \int_0^1 R(s_2, B_{s_1} - B_{s_1-s_2})ds_2$ is stationary and has a finite range of dependence. We can assume B is a two-sided Brownian motion so χ is defined on $\mathbb{R}$. Since the integral $\varepsilon^2 \int_1^{t/\varepsilon^4}$ is of CLT scaling, it is an exercise to show that

Proposition 3.2. *$Y_\varepsilon(t) - c_1\frac{t}{\varepsilon^2} \Rightarrow \sqrt{2c_2}W(t)$ in $C[0,\infty)$, where W is a standard Brownian motion. Here*

$$c_1 = \mathbb{E}\chi(s), \qquad c_2 = \int_0^\infty \mathbb{E}[\chi(s)\chi(0)]ds$$

After checking the uniform integrability, we pass to the limit to derive that

$$\mathbb{E}u_\varepsilon(t,x) = \mathbb{E}_B e^{Y_\varepsilon(t)-C_\varepsilon t} \to \mathbb{E}e^{\sqrt{2c_2}W(t)-c_2 t} = 1$$

Step 2. Now we use the Stroock formula to write down the chaos expansion for $u_\varepsilon(t,x)$ and show that it converges to the chaos expansion of $\mathcal{U}(t,x)$. Again we rely on the Feynman-Kac formula to compute the Malliavin derivative. Before going to the detail, let us put the problem in a slightly different setup. By the scaling property, we know that $V_\varepsilon(t,x) = \varepsilon^{-3}V(\frac{t}{\varepsilon^4},\frac{x}{\varepsilon^2})$ has the same law as

$$\xi_\delta(t,x) = \int \phi_\delta(t-s,x-y)\xi(s,y)dyds$$

where $\phi_\delta(t,x) = \frac{1}{\delta^3}\phi(\frac{t}{\delta^2},\frac{x}{\delta})$ is assumed to be a symmetric mollifier and $\delta = \varepsilon^2$. The relation between R and ϕ is given by $R = \phi \star \phi$. This is convenient because in the original setup, we only expect convergence in distribution, but now with ξ_δ we expect convergence in probability (since $\xi_\delta \to \xi$ in probability)– sometimes proving strong convergence is easier than proving weak convergence.

Remark 3.3. It seems to be a philosophical question whether one should view the problem as a large scale problem (starting with the potential of size ε and consider the scaling $(\frac{t}{\varepsilon^4},\frac{x}{\varepsilon^2})$) or a small scale problem (starting with the white noise ξ and consider its approximation ξ_δ), and we will certainly not go into that.

With $V_\varepsilon(t,x)$ replaced by ξ_δ, we write

$$u_\varepsilon(t,x) = \mathbb{E}_B e^{\int_0^t \xi_\delta(t-s,x+B_s)ds - C_\varepsilon t}$$

The exponent can be written as

$$\begin{aligned}\int_0^t \xi_\delta(t-s,x+B_s)ds &= \int_0^t \left(\int \phi_\delta(t-s-\ell,x+B_s-y)\xi(\ell,y)d\ell dy\right)ds\\ &= \int\left(\int_0^t \phi_\delta(t-s-\ell,x+B_s-y)ds\right)\xi(\ell,y)d\ell dy\end{aligned}$$

Since ϕ_δ approximates the Dirac function as $\delta \to 0$, on a purely formal level, we can write

(3.12)
$$\Phi^\varepsilon_{t,x}(\ell,y) := \int_0^t \phi_\delta(t-s-\ell,x+B_s-y)ds \to \int_0^t \delta(t-s-\ell,x+B_s-y)ds = \delta(x+B_{t-\ell}-y).$$

To compute the Malliavin derivative, we interchange it with the expectation

$$\begin{aligned}\mathcal{D}_{\ell,y}u_\varepsilon(t,x) =& \mathbb{E}_B[\mathcal{D}_{\ell,y}e^{\int_0^t \xi_\delta(t-s,x+B_s)ds - C_\varepsilon t}]\\ =& \mathbb{E}_B\left[e^{\int_0^t \xi_\delta(t-s,x+B_s)ds - C_\varepsilon t}\Phi^\varepsilon_{t,x}(\ell,y)\right]\end{aligned} \tag{3.13}$$

Taking the expectation, we obtain the coefficient of the first order chaos

$$\mathbb{E}[\mathcal{D}_{\ell,y}u_\varepsilon(t,x)] = \mathbb{E}_B\left[e^{\tilde{Y}_\delta(t)-C_\varepsilon t}\Phi^\varepsilon_{t,x}(\ell,y)\right]$$

We expect that $\mathbb{E}[\mathcal{D}_{\ell,y}u_\varepsilon(t,x)] \to q_{t-\ell}(x-y)$, which is the coefficient of the first chaos of $\mathcal{U}(t,x)$. Here

$$\tilde{Y}_\delta(t) = \frac{1}{2}\int_{[0,t]^2} R_\delta(s_1-s_2,B_{s_1}-B_{s_2})$$

with $R_\delta = \phi_\delta \star \phi_\delta$ is the covariance function of ξ_δ and a scaling argument shows that $\tilde{Y}_\delta$ has the same law as Y_ε. With a bit more effort, we can extend Proposition 3.2 to the joint convergence of

$$(\tilde{Y}_\delta(t) - c_1 \tfrac{t}{\delta}, B_t) \Rightarrow (\sqrt{2c_2} W(t), B_t) \tag{3.14}$$

in $C[0,\infty)$, with B independent of W. Therefore, again on a purely formal level, we have (invoking (3.12))

$$\begin{aligned} &\mathbb{E}[\mathcal{D}_{\ell,y} u_\varepsilon(t,x)] = \mathbb{E}_B\left[e^{\tilde{Y}_\delta(t) - C_\varepsilon t} \Phi^\varepsilon_{t,x}(\ell, y)\right] \\ &\to \mathbb{E}[e^{\sqrt{2c_2} W(t) - c_2 t} \delta(x + B_{t-\ell} - y)] = q_{t-\ell}(x-y). \end{aligned}$$

Unfortunately, the joint convergence in (3.14) is not enough to justify the previous passing to the limit. The convergence in (3.12) is just symbolic and it is about the property of the Brownian motion on small scales, not compatible with the weak convergence we were able to establish in (3.14). To pass to the limit, we need a type of local central limit theorem – after all, the limit we get in the end is the heat kernel $q_{t-\ell}(x-y)$, which is a density rather than a probability. This probably can be done, but is complicated in any case. Here is a brief sketch of how we avoid dealing with it.

Step 2.1. We first show that $u_\varepsilon(t,x)$ converges in $L^2(\Omega)$ by proving

$$\mathbb{E}[u_{\varepsilon_1}(t,x) u_{\varepsilon_2}(t,x)]$$

converges as $\varepsilon_1, \varepsilon_2 \to 0$. It shows the convergence by proving it is a Cauchy sequence, without identifying the limit. This relies on a similar calculation using the Feynman-Kac formula and some estimates on the intersection local time of two independent Brownian motion.

Step 2.2. It remains to identify the limit as $\mathcal{U}(t,x)$. The fact that $\{u_\varepsilon(t,x)\}_\varepsilon$ is a Cauchy sequence in $L^2(\Omega)$ implies that all chaos are Cauchy in $L^2(\Omega)$ (due to orthogonality), then as a (deterministic) function of $(\ell, y) \in (0,\infty)\times\mathbb{R}$, $\mathbb{E}[\mathcal{D}_{\ell,y} u_\varepsilon(t,x)]$ must converge in $L^2(0,\infty)\times\mathbb{R}$. Given the strong convergence, to identify its limit, it suffices to proceed weakly in space, i.e., we take any test function g and pass to the limit of the integral

$$\int g(\ell, y) \mathbb{E}[\mathcal{D}_{\ell,y} u_\varepsilon(t,x)] dy d\ell$$

which leads to the factor

$$\int g(\ell, y) \Phi^\varepsilon_{t,x}(\ell, y) dy d\ell$$

and becomes sort of a probability rather than a density. Then it is enough to pass to the limit using (3.14). This also completes the proof of Theorem 3.1.

Remark 3.4. Through the Stroock formula, we also see the connection between the KPZ equation and the directed polymers. From (3.13) we derive

$$\mathcal{D}_{\ell,y} \log u_\varepsilon(t,x) = \frac{\mathcal{D}_{\ell,y} u_\varepsilon(t,x)}{u_\varepsilon(t,x)} = \frac{\mathbb{E}_B\left[e^{\int_0^t \xi_\delta(t-s, x+B_s) ds} \Phi^\varepsilon_{t,x}(\ell, y)\right]}{\mathbb{E}_B\left[e^{\int_0^t \xi_\delta(t-s, x+B_s) ds}\right]}$$

the expectation of which is actually the coefficient of the first order chaos of the random variable $h_\varepsilon(t,x) = \log u_\varepsilon(t,x)$. It is clear that

$$\mathcal{D}_{\ell,y} \log u_\varepsilon(t,x) > 0, \qquad \int \mathcal{D}_{\ell,y} \log u_\varepsilon(t,x) dy = 1$$

so it is a random probability density. From the expression, we see that it is almost exactly the density of the polymer path at ℓ, with the starting point (t,x) and the path running backward in time.

Remark 3.5 (Transition from homogenization). Compare Theorems 3.1 and 2.1, the difference is that a longer time scale $\varepsilon \mapsto \frac{t}{\varepsilon^4}$ in Theorem 3.1. As one can imagine, the chaos expansion approach can also be used to study the (simpler) cases of $\varepsilon \mapsto \frac{t}{\varepsilon^\alpha}$ with $\alpha \in [2,4)$, where $\alpha = 2$ corresponds to the homogenization regime. Assuming u_ε solves

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \tfrac{1}{\varepsilon^{\alpha-1}} V(\tfrac{t}{\varepsilon^\alpha}, \tfrac{x}{\varepsilon^{\alpha/2}}) u_\varepsilon$$

the result says the following: for any $\alpha \in [2,4)$ and each (t,x),

$$u_\varepsilon(t,x) \exp(-\tfrac{c_1 t}{\varepsilon^{\alpha-2}}) \to \bar{u}(t,x) \text{ in } L^2(\Omega) \tag{3.15}$$

with $\bar{u}$ solving the unperturbed heat equation. It covers Theorem 2.1 as a special case. As expected, as soon as $\alpha > 2$, one needs to renormalize. What happens in the above convergence is that only the zeroth order chaos, i.e., the mean, survives in the limit, and all other terms vanish as $\varepsilon \to 0$.

If we view the homogenization result as a law of large numbers, it is natural to further study the error and derive a possible central limit theorem. Here is the result which can be obtained as a corollary of Theorem 3.1: for $\alpha \in [2,4)$ and each (t,x),

$$\tfrac{1}{\varepsilon^{1-\alpha/4}} (u_\varepsilon(t,x) - \mathbb{E}u_\varepsilon(t,x)) \exp(-\tfrac{c_1 t}{\varepsilon^{\alpha-2}}) \Rightarrow \mathscr{U}(t,x) \tag{3.16}$$

with

$$\partial_t \mathscr{U}(t,x) = \tfrac{1}{2}\Delta \mathscr{U}(t,x) + \bar{u}(t,x)\xi(t,x)$$

The limit $\mathscr{U}(t,x)$ is a Gaussian process, so the fluctuations in homogenization fall into the Edwards-Wilkinson class. The CLT in (3.16) can be explained as follows. Suppose we write $u_\varepsilon(t,x)$ in the chaos expansion

$$u_\varepsilon(t,x) = \mathbb{E}[u_\varepsilon(t,x)] + \sum_{n\geq 1} u_{\varepsilon,n}(t,x)$$

where $u_{\varepsilon,n}(t,x)$ denotes the n–th order chaos. By what is shown in Theorem 3.1, one easily sees that

$$\tfrac{1}{\varepsilon^{(1-\alpha/4)n}} \exp(-\tfrac{c_1 t}{\varepsilon^{\alpha-2}}) u_{\varepsilon,n}(t,x)$$

should converge to the n–th chaos of SHE, denoted by $\mathcal{U}_n(t,x)$. The factor $\frac{1}{\varepsilon^{(1-\alpha/4)n}}$ comes from the fact that

$$\tfrac{1}{\varepsilon^{\alpha-1}} V(\tfrac{t}{\varepsilon^\alpha}, \tfrac{x}{\varepsilon^{\alpha/2}}) = \varepsilon^{1-\alpha/4} \left(\tfrac{1}{\varepsilon^{3\alpha/4}} V(\tfrac{t}{\varepsilon^\alpha}, \tfrac{x}{\varepsilon^{\alpha/2}}) \right)$$

with the term in the bracket scales to a spacetime white noise. Thus, the renormalized random fluctuation can be roughly written as

$$[u_\varepsilon(t,x) - \mathbb{E}u_\varepsilon(t,x)] \exp(-\tfrac{c_1 t}{\varepsilon^{\alpha-2}}) \approx \sum_{n\geq 1} \varepsilon^{(1-\alpha/4)n} \mathcal{U}_n(t,x)$$

For $\alpha \in [2,4)$, the main contribution to the r.h.s. is the first chaos, which is of Gaussian distribution, and all high order terms are much smaller. This leads to the Gaussian fluctuation in (3.16). The above approximation also shows the transition: as we increase α to 4, all chaos becomes order $O(1)$, building up the solution to SHE.

Remark 3.6. The chaos expansion approach seems to require the least machinery to prove something like Theorem 3.1, but it is also pretty restrictive. In a sense, one needs to be able to compute the chaos expansion more or less explicitly. For the linear equation, this was done through the Feynman-Kac formula, as in (3.13). For nonlinear problems or the case of a non-Gaussian noise, it is unclear how one should proceed in this way.

References. The Wong-Zakai theorem was first proved in [38] for general semi-linear equations. The result was later extended to non-Gaussian noise in [15], and the problem on the whole line was studied in [37]. The approach presented here is based on [32].

4. KPZ to Edwards-Wilkinson

The goal in this section is to study the following equation

$$\partial_t H_\varepsilon = \tfrac{1}{2}\Delta H_\varepsilon + \tfrac{1}{2}\kappa_\varepsilon |\nabla H_\varepsilon|^2 + \xi_\varepsilon, \qquad x \in \mathbb{R}^d, d \geq 3, \tag{4.1}$$

with $\kappa_\varepsilon = \beta\varepsilon^{d/2-1}$, and ξ_ε is a spatial mollification of the spacetime white noise ξ on the scale of ε, i.e.,

$$\xi_\varepsilon(t,x) = \int \phi_\varepsilon(x-y)\xi(t,y)dy, \qquad \phi_\varepsilon(\cdot) = \tfrac{1}{\varepsilon^d}\phi(\tfrac{\cdot}{\varepsilon}) \tag{4.2}$$

We start from the flat initial data $H_\varepsilon(0,x) \equiv 0$. Here is the main result:

Theorem 4.1. *There exists $\beta_c > 0$ such that if $\beta < \beta_c$, then for any test function g,*

$$\int [H_\varepsilon(t,x) - \mathbb{E}H_\varepsilon(t,x)]g(x)dx \Rightarrow \int \mathcal{U}(t,x)g(x)dx$$

in distribution, with $\mathcal{U}$ solving the Edwards-Wilkinson equation

$$\partial_t \mathcal{U} = \tfrac{1}{2}\Delta\mathcal{U} + \nu_{\text{eff}}\xi, \qquad \mathcal{U}(0,x) = 0 \tag{4.3}$$

where ν^2_{eff} is given in (4.17).

From (4.1), we can write down the equation for the centered fluctuations

$$\partial_t[H_\varepsilon - \mathbb{E}H_\varepsilon] = \tfrac{1}{2}\Delta[H_\varepsilon - \mathbb{E}H_\varepsilon] + \tfrac{1}{2}\kappa_\varepsilon[|\nabla H_\varepsilon|^2 - \mathbb{E}|\nabla H_\varepsilon|^2] + \xi_\varepsilon \tag{4.4}$$

What happens as $\varepsilon \to 0$ is that $\xi_\varepsilon \to \xi$, and

$$\tfrac{1}{2}\kappa_\varepsilon[|\nabla H_\varepsilon|^2 - \mathbb{E}|\nabla H_\varepsilon|^2] \Rightarrow \sigma_{\text{eff}}\eta$$

with η an independent spacetime white noise. Thus, the effective noise in (4.3) is actually $\xi + \sigma_{\text{eff}}\eta$ and the effective variance is $\nu^2_{\text{eff}} = 1 + \sigma^2_{\text{eff}} > 1$. We have abused the notation to continue use ξ (we only care about distribution after all), but one should not think of the ξ in (4.3) as the one in (4.1). The nonlinear term becoming asymptotically independent in the limit seems to be a rather general fact, which also applies to other problems such as the one considered in [17]. We will discuss it in more detail in Remark 4.3.

It is helpful to recall the microscopic model corresponding to (4.1). Let u solve

$$\partial_t u = \tfrac{1}{2}\Delta u + \beta u V(t,x) \tag{4.5}$$

with $V(t,x) = \int \phi(x-y)\xi(t,y)dy$. Consider $h_\varepsilon(t,x) = h(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}) = \log u(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$ which solves

$$\partial_t h_\varepsilon = \tfrac{1}{2}\Delta h_\varepsilon + \tfrac{1}{2}|\nabla h_\varepsilon|^2 + \beta\tfrac{1}{\varepsilon^2}V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon})$$

then $\tilde{h}_\varepsilon = \frac{h_\varepsilon}{\beta\varepsilon^{d/2-1}}$ solves

$$\partial_t \tilde{h}_\varepsilon = \tfrac{1}{2}\Delta\tilde{h}_\varepsilon + \tfrac{1}{2}\beta\varepsilon^{d/2-1}|\nabla h_\varepsilon|^2 + \tfrac{1}{\varepsilon^{1+d/2}}V(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}).$$

It is easy to check that $\frac{1}{\varepsilon^{1+d/2}}V(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$ has the same law as $\xi_\varepsilon(t,x)$ (as a process). Thus, $\tilde{h}_\varepsilon$ has the same law as H_ε.

We will not directly study the nonlinear equation (4.1), although it would be nice if one could. Instead, we will study the linear equation (4.5) and how its solution behaves under the nonlinear transformation. As discussed in the introduction, the key difficulty is to transfer the correlation property of u into that of h. It is clear

the approach here does not work for other nonlinear equations when there is no Hopf-Cole anymore.

4.1. **Clark-Ocone.** The key difficulty lies in studying the asymptotics of the two point correlation function of H_ε. In other words, we would like to have a handy formula for

$$\mathbb{E}\big[(H_\varepsilon(t,x)-\mathbb{E}H_\varepsilon(t,x))(H_\varepsilon(t,y)-\mathbb{E}H_\varepsilon(t,y))\big],$$

so that one can do asymptotics. The idea is to write $H_\varepsilon-\mathbb{E}H_\varepsilon$ as an Ito integral so one can apply Ito isometry, and this will be done through the Clark-Ocone formula. In a sense, we are looking for a mild formulation of (4.1), and the Clark-Ocone formula does (part of) the job.

On the Gaussian space associated with the noise ξ, suppose a random variable Z is in $D^{1,2}$ which is the Sobolev space with the Malliavin derivative, then the Clark-Ocone formula tells that

$$Z-\mathbb{E}Z=\int_0^\infty\int\mathbb{E}[\mathcal{D}_{s,y}Z|\mathcal{F}_s]\xi(s,y)dyds \tag{4.6}$$

Here $\mathcal{D}$ denotes the Malliavin derivative, and $\{\mathcal{F}_s\}_{s\geq0}$ is the filtration generated by the noise ξ. For the KPZ equation, as we explained in Remark 3.4, the Malliavin derivative $\mathcal{D}_{s,y}h(t,x)$ can be interpreted as the density of the directed polymer at (s,y), with the starting point at (t,x). Thus, applying (4.6) to $h(t,x)$, the r.h.s. becomes the density of the directed polymer (which is itself random) Ito-integrating against the noise, and a key here is that, before performing the Ito integration, the randomness in $[s,t]\times\mathbb{R}^d$ should be first averaged out, which is the role by the conditional expectation in (4.6). One should compare (4.6) with the mild formulation of SHE, in which the Ito integral takes the form $\int_0^t\int q_{t-s}(x-y)u(s,y)\xi(s,y)dyds$ (this mild formulation can actually be easily obtained through (4.6) as we will see). In some sense, a difference is the heat kernel $q_{t-s}(x-y)$ being replaced by the (averaged) density of the directed polymer.

By a similar scaling as before, define

$$u_\varepsilon(t,x)=e^{\kappa_\varepsilon H_\varepsilon(t,x)-\frac12\kappa_\varepsilon^2R_\varepsilon(0)t}$$

where $R_\varepsilon(\cdot)=\frac{1}{\varepsilon^d}R(\frac{\cdot}{\varepsilon})$ is the spatial covariance function of ξ_ε and $R=\phi\star\phi$, we have

$$\partial_tu_\varepsilon=\tfrac12\Delta u_\varepsilon+\kappa_\varepsilon u_\varepsilon\xi_\varepsilon,\qquad u_\varepsilon(0,x)=1. \tag{4.7}$$

We write the solution to (4.7) by Feynman-Kac formula

$$u_\varepsilon(t,x)=\mathbb{E}_B\exp(\kappa_\varepsilon\int_0^t\xi_\varepsilon(t-s,x+B_s)ds-\tfrac12\kappa_\varepsilon^2R_\varepsilon(0)t)$$

(Here (4.7) is interpreted in the Ito sense, which only differs from the Stratonovitch sense by a multiple constant.) Then

$$\mathcal{D}_{s,y}H_\varepsilon(t,x)=\tfrac{1}{\kappa_\varepsilon}\mathcal{D}_{s,y}\log u_\varepsilon(t,x)=\tfrac{1}{\kappa_\varepsilon}\tfrac{\mathcal{D}_{s,y}u_\varepsilon(t,x)}{u_\varepsilon(t,x)}$$

By (4.2), we write the exponent as

$$\int_0^t\xi_\varepsilon(t-s,x+B_s)ds=\int_0^t\int\phi_\varepsilon(x+B_{t-s}-y)\xi(s,y)dyds$$

which implies that

$$\mathcal{D}_{s,y}u_\varepsilon(t,x)=\kappa_\varepsilon\mathbb{1}_{[0,t]}(s)\mathbb{E}_B\Big[e^{\kappa_\varepsilon\int_0^t\xi_\varepsilon(\ell,x+B_{t-\ell})d\ell-\frac12\kappa_\varepsilon^2R_\varepsilon(0)t}\phi_\varepsilon(x+B_{t-s}-y)\Big]$$

Applying (4.6), we further derive

(4.8)
$$H_\varepsilon(t,x) - \mathbb{E}H_\varepsilon(t,x)$$
$$= \int_0^t \int \mathbb{E}\left[\frac{\mathbb{E}_B\left[e^{\kappa_\varepsilon \int_0^t \xi_\varepsilon(\ell, x+B_{t-\ell})d\ell - \frac{1}{2}\kappa_\varepsilon^2 R_\varepsilon(0)t} \phi_\varepsilon(x+B_{t-s}-y)\right]}{\mathbb{E}_B\left[e^{\kappa_\varepsilon \int_0^t \xi_\varepsilon(\ell, x+B_{t-\ell})d\ell - \frac{1}{2}\kappa_\varepsilon^2 R_\varepsilon(0)t}\right]} \,\middle|\, \mathcal{F}_s \right] \xi(s,y)dyds.$$

Note that the factor $e^{-\frac{1}{2}\kappa_\varepsilon^2 R_\varepsilon(0)t}$ appears in both the numerator and the denominator, and we kept it there for future purpose. The difficulty in dealing with the above expression is that we need to compute the conditional expectation of some variable with a random denominator.

For the numerator in (4.8), we condition on the value B_{t-s} and write

$$\mathbb{E}_B\left[e^{\kappa_\varepsilon \int_0^t \xi_\varepsilon(\ell, x+B_{t-\ell})d\ell - \kappa_\varepsilon^2 R_\varepsilon(0)t/2} \phi_\varepsilon(x+B_{t-s}-y)\right]$$
$$= \int \mathbb{E}_B\left[e^{\kappa_\varepsilon \int_0^t \xi_\varepsilon(\ell, x+B_{t-\ell})d\ell - \kappa_\varepsilon^2 R_\varepsilon(0)t/2} \,\middle|\, B_{t-s}=w\right] \phi_\varepsilon(x+w-y) q_{t-s}(w)dw$$

For the exponential factor, we have

$$\mathbb{E}_B\left[e^{\kappa_\varepsilon \int_0^t \xi_\varepsilon(\ell, x+B_{t-\ell})d\ell - \kappa_\varepsilon^2 R_\varepsilon(0)t/2} \,\middle|\, B_{t-s}=w\right]$$
$$= \mathcal{E}_{t,s}(x, x+w)\mathcal{E}_{s,0}(x+w)$$

where

$$\mathcal{E}_{t,s}(x,x+w) = \mathbb{E}_B[e^{\kappa_\varepsilon \int_s^t \xi_\varepsilon(\ell, x+B_{t-\ell})d\ell - \kappa_\varepsilon^2 R_\varepsilon(0)(t-s)/2} \,|\, B_{t-s}=w]$$
$$\mathcal{E}_{s,0}(x+w) = \mathbb{E}_B[e^{\kappa_\varepsilon \int_0^s \xi_\varepsilon(\ell, x+B_{t-\ell})d\ell - \kappa_\varepsilon^2 R_\varepsilon(0)s/2} \,|\, B_{t-s}=w]$$

The convention here is that $\mathcal{E}_{s_1,s_2}(x_1,x_2)$ is the point-to-point partition function of the directed polymer from (s_1,x_1) to (s_2,x_2), with $s_1 > s_2$ and we run backward in time, and $\mathcal{E}_s(x)$ is the point-to-line partition function starting from (s,x) and running backward in time. Therefore, the integrand in (4.8) is rewritten as

$$\begin{aligned} &\int \mathbb{E}\left[\frac{\mathcal{E}_{t,s}(x,x+w)\mathcal{E}_{s,0}(x+w)}{\mathcal{E}_{t,0}(x)} \,|\, \mathcal{F}_s\right] \phi_\varepsilon(x+w-y) q_{t-s}(w)dw \\ &= \int \mathbb{E}\left[\frac{\mathcal{E}_{t,s}(x,w)}{\mathcal{E}_{t,0}(x)} \,|\, \mathcal{F}_s\right] \mathcal{E}_{s,0}(w) \phi_\varepsilon(w-y) q_{t-s}(x-w)dw \qquad (4.9) \\ &= \int \mathbb{E}\left[\frac{\mathcal{E}_{t,s}(x,w)}{\mathcal{E}_{t,0}(x)} \,|\, \mathcal{F}_s\right] u_\varepsilon(s,w) \phi_\varepsilon(w-y) q_{t-s}(x-w)dw \end{aligned}$$

Without the random denominator in (4.9), which is actually what happens when the Clark-Ocone formula is applied to u_ε, we use the simple fact that $\mathcal{E}_{t,s}(x,w)$ is independent of $\mathcal{F}_s$ and $\mathbb{E}[\mathcal{E}_{t,s}(x,w)] = 1$ to conclude that the above expression equals to

$$\int u_\varepsilon(s,w)\phi_\varepsilon(w-y)q_{t-s}(x-w)dw$$

which implies

$$u_\varepsilon - \mathbb{E}u_\varepsilon = \kappa_\varepsilon \int_0^t \int \left(\int u_\varepsilon(s,w)\phi_\varepsilon(w-y)q_{t-s}(x-w)dw \right) \xi(s,y)dyds \qquad (4.10)$$

Integrating y first, this is precisely the mild formulation of (4.7).

In (4.9), the situation is of course more complicated, but in our non-intermittent regime where the underlying microscopic dynamics quickly reaches the local equilibrium, a similar strategy inspired by the above calculation can be implemented.

First, by a scaling, we know that $u_\varepsilon(t,x)$ has the same law as $u(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$ with u solving (4.5). From the discussion in the introduction, under the assumption that $d \geq 3$ and $\beta < \beta_c$, the solution to (4.5) converges to some stationary distribution. Therefore, the random variable $\mathcal{E}_{t,0}(x) = u_\varepsilon(t,x)$ essentially depends on a thin layer

of the random environment near t. This is because the convergence to stationarity only requires a microscopically large amount of time, which could be much smaller than ε^{-2}. In other words, although the random variable $u(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon})$ is a measurable function of $\{V(s,y) : (s,y) \in [0, \frac{t}{\varepsilon^2}] \times \mathbb{R}^d\}$, we can approximate it by a random variable that only depends on $(s,y) \in [\frac{t}{\varepsilon^2} - M_\varepsilon, \frac{t}{\varepsilon^2}] \times \mathbb{R}^d$, with $1 \ll M_\varepsilon \ll \varepsilon^{-2}$. In the macroscopic variable, it means that we can approximate $u_\varepsilon(t,x)$ by a random variable that is measurable with respect to the environment in $(s,y) \in [t - o(1), t] \times \mathbb{R}^d$.

We rewrite (4.8) as

$$\begin{aligned} &H_\varepsilon(t,x) - \mathbb{E}H_\varepsilon(t,x) \\ &= \int_0^t \int \left(\int \mathbb{E}\left[\frac{\mathcal{E}_{t,s}(x,w)}{\mathcal{E}_{t,0}(x)} \,|\, \mathcal{F}_s \right] u_\varepsilon(s,w) \phi_\varepsilon(w-y) q_{t-s}(x-w) dw \right) \xi(s,y) dy ds. \end{aligned} \tag{4.11}$$

and the proof consists of three steps.

Step 1. By the previous discussion, we know that $\mathcal{E}_{t,0}(x) = u_\varepsilon(t,x)$ can be approximated by

$$\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x) := \mathbb{E}_B e^{\kappa_\varepsilon \int_0^{\tau_\varepsilon} \xi_\varepsilon(t-\ell, x+B_\ell) d\ell - \kappa_\varepsilon^2 R_\varepsilon(0)\tau_\varepsilon/2}$$

with some $\varepsilon^2 \ll \tau_\varepsilon \ll 1$ (τ_ε is the M_ε in the previous paragraph multiplied by ε^2). Suppose we can control the error in the approximation (which needs a negative moment estimate of $u_\varepsilon(t,x)$), then

$$\begin{aligned} &H_\varepsilon(t,x) - \mathbb{E}H_\varepsilon(t,x) \\ &\approx \int_0^t \int \left(\int \mathbb{E}\left[\frac{\mathcal{E}_{t,s}(x,w)}{\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)} \,|\, \mathcal{F}_s \right] u_\varepsilon(s,w) \phi_\varepsilon(w-y) q_{t-s}(x-w) dw \right) \xi(s,y) dy ds. \end{aligned} \tag{4.12}$$

Step 2. We know that $\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)$ is measurable with respect to the noise in $[t-\tau_\varepsilon, t]$, which is independent of $\mathcal{F}_s$ for those $s < t - \tau_\varepsilon$. Suppose we can control the integral in the thin layer, then

(4.13)

$$\begin{aligned} &H_\varepsilon(t,x) - \mathbb{E}H_\varepsilon(t,x) \\ &\approx \int_0^{t-\tau_\varepsilon} \int \left(\int \mathbb{E}\left[\frac{\mathcal{E}_{t,s}(x,w)}{\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)} \,|\, \mathcal{F}_s \right] u_\varepsilon(s,w) \phi_\varepsilon(w-y) q_{t-s}(x-w) dw \right) \xi(s,y) dy ds \\ &= \int_0^{t-\tau_\varepsilon} \int \left(\int \mathbb{E}\left[\frac{\mathcal{E}_{t,s}(x,w)}{\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)} \right] u_\varepsilon(s,w) \phi_\varepsilon(w-y) q_{t-s}(x-w) dw \right) \xi(s,y) dy ds \end{aligned}$$

In this way, we got rid of the conditional expectation.

Step 3. It remains to study $\mathbb{E}[\frac{\mathcal{E}_{t,s}(x,w)}{\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)}]$. Recall that $\mathcal{E}_{t,s}(x,w)$ is the point-to-point partition function of the directed polymer, an average with respect to a Brownian bridge connection x and w. By the time reversal property of the bridge, we immediately see that the two endpoints are symmetric. Therefore, for any fixed $s < t$, by the same discussion as before, the random variable $\mathcal{E}_{t,s}(x,w)$ essentially depends on the random noise in two thin layers: $\ell \in [t-\tau_\varepsilon, t]$ and $[s, s+\tau_\varepsilon]$. Define

$$\bar{\mathcal{E}}_{s+\tau_\varepsilon, s}(w) = \mathbb{E}_B\left[e^{\kappa_\varepsilon \int_0^{\tau_\varepsilon} \xi_\varepsilon(s+\ell, w+B_\ell) d\ell - \kappa_\varepsilon^2 R_\varepsilon(0)\tau_\varepsilon/2} \right],$$

one should check that under the assumption of $d \geq 3, \beta < \beta_c$, we have the approximation

$$\mathcal{E}_{t,s}(x,w) \approx \tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x) \bar{\mathcal{E}}_{s+\tau_\varepsilon, s}(w) \tag{4.14}$$

Thus, in this non-intermittent regime, the point-to-point partition function is factorized into two point-to-line partition functions. With the approximation in (4.14),

we conclude that

$$\mathbb{E}\big[\tfrac{\mathcal{E}_{t,s}(x,w)}{\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)}\big] \approx \mathbb{E}\big[\tfrac{\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)\bar{\mathcal{E}}_{s+\tau_\varepsilon,s}(w)}{\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)}\big] = \mathbb{E}\bar{\mathcal{E}}_{s+\tau_\varepsilon,s}(w) = 1$$

which implies

$$\begin{aligned} &H_\varepsilon(t,x) - \mathbb{E}H_\varepsilon(t,x) \\ &\approx \int_0^{t-\tau_\varepsilon} \int \Big(\int u_\varepsilon(s,w)\phi_\varepsilon(w-y)q_{t-s}(x-w)dw \Big) \xi(s,y)dyds \end{aligned} \tag{4.15}$$

By (4.10), the r.h.s. of (4.15) is approximately $\frac{1}{\kappa_\varepsilon}(u_\varepsilon - \mathbb{E}u_\varepsilon)$, then the result follows from (1.20). The proof of Theorem 4.1 is complete modulo the verification of the approximations in (4.12), (4.13) and (4.14).

It might have appeared that the above discussion applies for a fixed (t,x) – this is not true and one can not pass to the limit in (4.15) for any fixed (t,x), because the limiting Edwards-Wilkinson solution does not make sense for any (t,x). We should interpret the "$\approx$" in (4.12), (4.13) and (4.15) as being close in $L^2(\Omega)$ after integrating against a smooth test function.

We rewrite (4.15) as

$$\begin{aligned} &\tfrac{1}{\kappa_\varepsilon}[\log u_\varepsilon(t,x) - \mathbb{E}\log u_\varepsilon(t,x)] \\ &\approx \int_0^{t-\tau_\varepsilon} \int \Big(\int u_\varepsilon(s,w)\phi_\varepsilon(w-y)q_{t-s}(x-w)dw \Big) \xi(s,y)dyds, \end{aligned}$$

so that one can view the role of the Clark-Ocone formula as a linearization. Nothing is special about the logarithm when it comes to linearization, thus, we may consider any smooth function f, and for $f(u_\varepsilon(t,x))$, ask the same question, as the one for (1.25). The above discussion applies verbatim and leads to

$$\begin{aligned} &\tfrac{1}{\kappa_\varepsilon}[f(u_\varepsilon(t,x)) - \mathbb{E}f(u_\varepsilon(t,x))] \\ &\approx \int_0^{t-\tau_\varepsilon} \int \Big(\int \mathbb{E}[f'(\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x))\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)]u_\varepsilon(s,w)\phi_\varepsilon(w-y)q_{t-s}(x-w)dw \Big) \xi(s,y)dyds \end{aligned} \tag{4.16}$$

Compare to (4.15), the only difference in (4.16) is the extra factor

$$\mathbb{E}[f'(\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x))\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)].$$

This comes from Step 3 where we have the approximation

$$\begin{aligned} \mathbb{E}[f'(\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x))\mathcal{E}_{t,s}(x,w)] &\approx \mathbb{E}[f'(\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x))\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)\bar{\mathcal{E}}_{s+\tau_\varepsilon,s}(w)] \\ &= \mathbb{E}[f'(\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x))\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)]\mathbb{E}[\bar{\mathcal{E}}_{s+\tau_\varepsilon,s}(w)] \\ &= \mathbb{E}[f'(\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x))\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)]. \end{aligned}$$

It happens to be 1 when $f = \log$. As ε goes to zero, $\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)$ is close in distribution to the stationary distribution, which we denote by Z_∞, then

$$\mathbb{E}[f'(\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x))\tilde{\mathcal{E}}_{t,t-\tau_\varepsilon}(x)] \to \sigma_f := \mathbb{E}[f'(Z_\infty)Z_\infty].$$

It further leads to the convergence of

$$\tfrac{1}{\kappa_\varepsilon}[f(u_\varepsilon(t,x)) - \mathbb{E}f(u_\varepsilon(t,x))] \Rightarrow \sigma_f \mathcal{U}.$$

In this way, the results for u_ε and H_ε can be viewed as special cases of the above convergence, with $f(x) = x$ and $\log x$ respectively.

Remark 4.2 (Effective variance). Let us explain heuristically where the effective variance ν_{eff}^2 comes from. We will work with the SHE

$$\partial_t u_\varepsilon = \tfrac{1}{2}\Delta u_\varepsilon + \kappa_\varepsilon u_\varepsilon \xi_\varepsilon, \qquad u_\varepsilon(0,x) = 1$$

The mild formulation gives

$$\tfrac{u_\varepsilon(t,x)-1}{\kappa_\varepsilon} = \int_0^t \int q_{t-s}(x-y)u_\varepsilon(s,y)\xi_\varepsilon(s,y)dyds$$

As explained in the introduction, the factor $u_\varepsilon\xi_\varepsilon$ becomes an effective spacetime white noise as $\varepsilon \to 0$. The fact that we obtained a white limit is easy to see: by the convergence to local stationarity, $u_\varepsilon(s,y)$ only depends on the noise near (s,y), so $u_\varepsilon\xi_\varepsilon$ can be viewed as a local function of the underlying noise. In addition, we know that $u_\varepsilon(s,y)$ has approximately the same law as $Z(\frac{s}{\varepsilon^2},\frac{y}{\varepsilon})$, where Z is the stationary solution to (4.5). Since the product between u_ε and ξ_ε is in the Ito sense, we can view them as being independent, therefore, the spatial covariance function of $u_\varepsilon(s,y)\xi_\varepsilon(s,y)$ is approximately $\frac{1}{\varepsilon^d}R_Z(\frac{\cdot}{\varepsilon})R(\frac{\cdot}{\varepsilon})$, where R_Z is the spatial covariance function of Z. Thus, as $\varepsilon \to 0$, $u_\varepsilon(s,y)\xi_\varepsilon(s,y)$ becomes a spacetime white noise with the effective variance given by

$$\nu_{\text{eff}}^2 = \int R_Z(x)R(x)dx. \tag{4.17}$$

Remark 4.3 (Nonlinear term). It does not seem easy to directly prove the convergence of

$$\mathcal{N}_\varepsilon := \tfrac{1}{2}\kappa_\varepsilon[|\nabla H_\varepsilon|^2 - \mathbb{E}|\nabla H_\varepsilon|^2] \Rightarrow \sigma_{\text{eff}}\eta$$

with η independent of ξ. By (4.4) we can write

$$H_\varepsilon - \mathbb{E}H_\varepsilon = \mathcal{G}\mathcal{N}_\varepsilon + \mathcal{G}\xi_\varepsilon$$

with $\mathcal{G} = (\partial_t - \frac{1}{2}\Delta)^{-1}$. To see that $\mathcal{G}\mathcal{N}_\varepsilon$ and $\mathcal{G}\xi_\varepsilon$ are asymptotically independent, at least uncorrelated, we only need to note that, by (4.15), $H_\varepsilon - \mathbb{E}H_\varepsilon \approx \frac{u_\varepsilon - 1}{\kappa_\varepsilon}$, whose first chaos is exactly $\mathcal{G}\xi_\varepsilon$. This shows that, up to some small error, $\mathcal{G}\mathcal{N}_\varepsilon$ lives in higher order chaos. Thus, passing to the limit, $\mathcal{G}\mathcal{N}_\varepsilon$ must be asymptotically uncorrelated with $\mathcal{G}\xi_\varepsilon$. The independence comes from the joint convergence to Gaussian, which we will not discuss here.

References. The Edwards-Wilkinson limit of the KPZ equation in $d \geq 3$ was first proved in [44], for a spacetime mollified potential. A similar result for SHE in $d \geq 3$ was proved in [31]. The $d = 2$ case was studied in [16, 14, 25]. The approach presented here is based on [24]. Further extensions can be found in [20, 23, 43].

5. KPZ on torus

In this section, we study the KPZ equation on the unit torus, which is formally written as

$$\partial_t h = \tfrac{1}{2}\Delta h + \tfrac{1}{2}|\nabla h|^2 + \xi, \qquad t > 0, x \in \mathbb{T}^d \tag{5.1}$$

We assume the noise ξ is white in time and smooth in space, with a spatial covariance function $R(\cdot)$. Two cases are considered: (i) $d \geq 1$ and $R(\cdot) : \mathbb{T}^d \to \mathbb{R}$ is a smooth function; (ii) $d = 1$ and $R(\cdot) = \delta(\cdot)$. In case (i), the equation (5.1) makes sense in its form since h is smooth in the x−variable. The initial data $h(0,x)$ can be essentially anything.

Here is the main result

Theorem 5.1. *There exists $\gamma, \sigma > 0$ depending on $d, R(\cdot)$ so that for any $x \in \mathbb{T}^d$, we have*

$$\frac{h(t,x) + \gamma t}{\sqrt{t}} \Rightarrow N(0,\sigma^2) \tag{5.2}$$

in distribution as $t \to \infty$.

We will not strive for the sharpest condition on $R(\cdot)$ under which Theorem 5.1 holds, but only consider what we think are the most interesting cases here. Throughout the section, we will study $u(t,x) = e^{h(t,x)}$ which solves the SHE

$$\partial_t u = \tfrac{1}{2}\Delta u + u\xi. \tag{5.3}$$

We interpret the above SHE in the Ito sense, so from (5.1) to (5.3), there is actually an Ito-Stratonovitch correction term which in case (i) is $-\frac{1}{2}R(0)u$ and in case (ii) is $-\infty$. Since they are simply constants, we do not bother to put them into the equation.

Compare (5.2) to (1.4), we see the $t^{1/3}$ fluctuation becomes $t^{1/2}$ when the whole line is replaced by the torus in $d = 1$. So in some sense, the super-concentration turns into concentration, and it would be interesting to study the cross-over regime. We refer to [4, 5, 6] for the relevant works in periodic TASEP.

The solution to (5.3) is intermittent. If we compute the moments, say with a flat initial data $u(0,x) \equiv 1$, then

$$\mathbb{E}u(t,x)^n = \mathbb{E}_B \exp\Big(\sum_{i,j=1}^{n} \mathbb{1}_{i\neq j} \int_0^t \delta(B_s^i - B_s^j)ds\Big) \tag{5.4}$$

Here $B^i, i = 1,\dots,n$ are independent Brownian motions on the torus. Equivalently, one can view them as Brownian motions on $\mathbb{R}^d$ and perform a periodization of the Dirac function. Therefore, $\mathbb{E}u(t,x)^n$ will be larger than the case when the equation is posed on the whole space: to contribute to the exponent in (5.4), it is easier for the Brownian motions to intersect on the torus than the whole space. Since the moment growth is at least the same, we know that on the torus, the SHE is also intermittent. For the case of a $1+1$ spacetime white noise on the whole line, it was shown in [18] that

$$\mathbb{E}u(t,x)^n \sim e^{\gamma_n t} \qquad \text{with } \gamma_n = \tfrac{1}{24}n(n^2-1)$$

It seems to be an open problem to compute γ_n when the equation is on the torus.

In any case, we are not interested in the moments of u here, but the long time behavior of $\log u$. While we only proved the convergence of the marginal, it is not hard to convince ourselves that it should also be true that

$$\varepsilon h\big(\tfrac{t}{\varepsilon^2}, x\big) + \tfrac{\gamma t}{\varepsilon} \Rightarrow \sigma W_t$$

as a process in t, where W is a standard Brownian motion. Thus, one can roughly write that

$$u(t,x) = e^{h(t,x)} \approx e^{\sigma W_t - \gamma t}$$

with the above "$\approx$" interpreted carefully (certainly not in the sense of the moments being close). The above approximation is actually exact when $R(\cdot)$ is constant – in this case, the noise only depends on the time variable, so the solution to (5.3), when started with a constant initial data, is a geometric Brownian motion. As soon as $R(\cdot)$ is non-constant, the problem becomes much more complicated.

Unlike the non-intermittent case we studied in Section 4, the result here shows that $u \to 0$ exponentially fast and $h \to -\infty$ with a linear speed, so we need to find other quantities that stabilize and give us some type of local stationarity. One such quantity is $\phi = \nabla h = \frac{\nabla u}{u}$, the solution to the Burgers equation

$$\partial_t \phi = \tfrac{1}{2}\Delta\phi + \phi\cdot\nabla\phi + \nabla\xi.$$

Taking the derivative of h removes its zero mode, which makes the convergence to stationarity possible. Using ϕ, one can write h (at least formally) as

$$h(t,x) = \int_0^t \int q_{t-s}(x-y)|\phi(s,y)|^2 dyds + \int_0^t \int q_{t-s}(x-y)\xi(s,y)dyds$$

when we start from $h(0,x) \equiv 0$. The second term on the r.h.s. is easy to analyze, and on the torus, only the zero mode contributes as $t \to \infty$. The key is to study the nonlinear term. For $t - s \gg 1$, the heat kernel $q_{t-s}(x-y)$ is almost a constant (on the torus), so one can view the first term on the r.h.s. as a time average of the $|\phi(s,y)|^2$. The fast mixing of ϕ then drives a central limit theorem.

One of the difficulties in implementing the above argument is that in the case of a $1+1$ spacetime white noise, ϕ is not function-valued, so the study of terms of the form $\int_0^t \int q_{t-s}(x-y)|\phi(s,y)|^2 dyds$ presumably requires tools from singular SPDE [35].

We study a different quantity:

$$\rho(t,x) = \frac{u(t,x)}{\int u(t,x')dx'} \tag{5.5}$$

which is a density function on $\mathbb{T}^d$. The fact that we are in a compact domain ensures no mass is leaking as $t \to \infty$, so one may expect that $\rho(t,\cdot)$ converges to some stationary distribution. When we are on the whole line, as $t \to \infty$, the density $\rho(t,\cdot)$ spreads out and goes to 0 as one can imagine. One may view ρ as the endpoint distribution of the directed polymer.

To be more precise, we introduce the following notations: let $u(t,x;f)$ be the solution to (5.3) with initial data f and $u(t,x;y) = u(t,x;\delta_y)$. Similarly, $\rho(t,x;f) = \frac{u(t,x;f)}{\int u(t,x';f)dx'}$.

When ξ is smooth in space, by Feynman-Kac, we can write

$$\begin{aligned} u(t,x;f) &= \int u(t,x;y)f(y)dy = \int \mathbb{E}_x[\delta(B_t - y)e^{\int_0^t \xi(t-\ell,B_\ell)d\ell - R(0)t/2}]f(y)dy \\ &= \int \mathbb{E}_y[\delta(B_t - x)e^{\int_0^t \xi(\ell,B_\ell)d\ell - R(0)t/2}]f(y)dy \end{aligned}$$

where we used $\mathbb{E}_x$ to denote the expectation with respect to the Brownian motion starting at $B_0 = x$. Therefore,

$$\rho(t,x;f) = \frac{\int \mathbb{E}_y[\delta(B_t - x)e^{\int_0^t \xi(\ell,B_\ell)d\ell}]f(y)dy}{\int \mathbb{E}_y[e^{\int_0^t \xi(\ell,B_\ell)d\ell}]f(y)dy}$$

which can be viewed as the density of endpoint distribution of the directed polymer with the initial distribution given by f. When ξ is a spacetime white noise, with an appropriate definition of the corresponding directed polymer [1], this is also true.

Suppose we can show that $\rho(t,\cdot)$ mixes very fast and converges to some stationary distribution, it remains to connect $h = \log u$ with ρ, and prove that ρ derives the CLT in (5.2). First, we write

$$\log \rho(t,x) = \log u(t,x) - \log Z_t = h(t,x) - \log Z_t$$

with

$$Z_t = \int u(t,x')dx' \tag{5.6}$$

Since ρ is expected to stabilize, the difference between $h(t,x)$ and $\log Z_t$ is of order $O(1)$, which vanishes after the $\sqrt{t}$ scaling. Then it remains to prove Theorem 5.1 for $\log Z_t$.

The idea is to use the semimartingale decomposition. It is well-known that Z_t is a positive martingale, which can be seen from either the Feynman-Kac or the mild formulation of (5.3): for any f that is a density, we have

$$u(t,x;f) = \int q_t(x-y)f(y)dy + \int_0^t \int q_{t-s}(x-y)u(s,y;f)\xi(s,y)dyds$$

which implies

$$Z_t = 1 + \int_0^t \int u(s,y;f)\xi(s,y)dyds$$

Define M_t through the SDE

$$dM_t = Z_t^{-1}dZ_t, \qquad M_0 = 0$$

we have

$$Z_t = e^{M_t - \frac{1}{2}\langle M\rangle_t}$$

so

$$\log Z_t = M_t - \tfrac{1}{2}\langle M\rangle_t \tag{5.7}$$

is the semimartingale we were looking for. By the definition of ρ, we have

$$M_t = \int_0^t \int \rho(s,y;f)\xi(s,y)dyds \tag{5.8}$$

and

$$\langle M\rangle_t = \int_0^t \int \rho(s,y_1;f)\rho(s,y_2;f)R(y_1-y_2)dy_1dy_2ds. \tag{5.9}$$

To summarize, the quantity we are interested in, $\log Z_t$, is written as a linear combination of M_t and $\langle M\rangle_t$, both expressed in terms of ρ. The mixing of ρ derives central limit theorems for M_t and $\langle M\rangle_t - \mathbb{E}\langle M\rangle_t$, respectively.

Remark 5.2. It is worth comparing (5.7) with the Clark-Ocone representation:

$$\log Z_t - \mathbb{E}\log Z_t = \int_0^t \int \mathbb{E}[\tfrac{\mathcal{D}_{s,y}Z_t}{Z_t} \,|\, \mathcal{F}_s]\xi(s,y)dyds \tag{5.10}$$

In the above formula, $\frac{\mathcal{D}_{s,y}Z_t}{Z_t}$ is also related to the directed polymer: take the case of a smooth noise as an example, $Z_t = \mathbb{E}_B e^{\int_0^t \xi(\ell,B_\ell)d\ell - \frac{1}{2}R(0)t}$, then

$$\frac{\mathcal{D}_{s,y}Z_t}{Z_t} = \frac{\mathbb{E}_B \exp(\int_0^t \xi(\ell,B_\ell)d\ell)\phi(B_s-y)}{\mathbb{E}_B \exp(\int_0^t \xi(\ell,B_\ell)d\ell)}.$$

A key difference is that, the polymer measure in (5.10) is fixed, i.e., the length $t>0$ is given, while in the expression of M_t and $\langle M\rangle_t$, the measure is changing with respect to s.

We rewrite the quadratic variation as

$$\langle M\rangle_t = \int_0^t \mathcal{R}(\rho(s))ds$$

with the functional $\mathcal{R}(v) = \int R(x-y)v(x)v(y)dxdy$ (for a spacetime white noise, $\mathcal{R}(v) = \|v\|_2^2$). So it is an additive functional of $\rho(s)$, and if ρ mixes fast, $\mathcal{R}(\rho(s))$ also decorrelates fast, which presumably leads to Gaussian behaviors of $\langle M\rangle_t$ after a centering. To make the heuristics precise, the idea is to solve a Poisson equation and extract the martingale part from $\langle M\rangle_t$ and apply a martingale CLT (it is similar to the homogenization approach in Section 2.2).

Denote the generator of $\{\rho(s)\}_{s\geq 0}$ by $\mathcal{L}$ and the invariant measure of $\{\rho(s)\}_{s\geq 0}$ by π_∞, we consider the Poisson equation

$$-\mathcal{L}\chi = \tilde{\mathcal{R}} = \mathcal{R} - 2\gamma$$

with

$$2\gamma = \int \mathcal{R}(\rho)\pi_\infty(d\rho) \tag{5.11}$$

Then,

$$\chi(\rho(t)) - \chi(\rho(0)) = \mathbf{M}_t + \int_0^t \mathcal{L}\chi(\rho(s))ds$$

for another martingale $\mathbf{M}_t$, and since we can write

$$\langle M\rangle_t - 2\gamma t = \int_0^t \tilde{\mathcal{R}}(\rho(s))ds = -\int_0^t \mathcal{L}\chi(\rho(s))ds$$

we have

$$\langle M\rangle_t - 2\gamma t = \mathbf{M}_t + \chi(\rho(0)) - \chi(\rho(t))$$

and

$$\begin{aligned}\log Z_t + \gamma t =& M_t - \tfrac{1}{2}(\langle M\rangle_t - 2\gamma t)\\ =& M_t - \tfrac{1}{2}\mathbf{M}_t + \tfrac{1}{2}\chi(\rho(t)) - \tfrac{1}{2}\chi(\rho(0))\\ =& \mathscr{M}_t + \varepsilon_t\end{aligned} \tag{5.12}$$

Here $\mathscr{M}_t = M_t - \frac{1}{2}\mathbf{M}_t$ is the martingale term and $\varepsilon_t = \frac{1}{2}\chi(\rho(t)) - \frac{1}{2}\chi(\rho(0))$ is the remainder. To prove the Gaussian fluctuations of $\log Z_t + \gamma_t$, it suffices to show that

Proposition 5.3. *As $t \to \infty$, we have $\frac{\mathscr{M}_t}{\sqrt{t}} \Rightarrow N(0,\sigma^2)$ and $\frac{\varepsilon_t}{\sqrt{t}} \to 0$.*

Now everything reduces to proving a geometric ergodicity of the Markov process $\{\rho(s)\}_{s\geq 0}$. First, it will imply the boundedness of the corrector χ, hence verifying that $\frac{\varepsilon_t}{\sqrt{t}} \to 0$ as $t \to \infty$. Then for the martingale term $\mathscr{M}_t$, once we write its quadratic variation explicitly, it is again an additive functional of $\{\rho(s)\}_{s\geq 0}$, so applying (a version of) ergodic theorem, one can show the convergence of quadratic variation, which implies the convergence of $\frac{\mathscr{M}_t}{\sqrt{t}} \Rightarrow N(0,\sigma^2)$.

We will not go into the detail of proving Proposition 5.3, which is a rather standard homogenization argument once we have the fast mixing of the underlying process. In the next section, we sketch the proof of the geometric ergodicity of $\{\rho(s)\}_{s\geq 0}$, which is stated in Proposition 5.4 below.

5.1. **Sinai's chain.** For any $t > s$ and $x, y \in \mathbb{T}^d$, define $\mathcal{Z}_{t,s}(x,y)$ as the solution to

$$\begin{aligned}&\partial_t \mathcal{Z}_{t,s}(x,y) = \tfrac{1}{2}\Delta_x \mathcal{Z}_{t,s}(x,y) + \mathcal{Z}_{t,s}(x,y)\xi(t,x), \qquad t > s,\\ &\mathcal{Z}_{s,s}(x,y) = \delta(x-y).\end{aligned} \tag{5.13}$$

In other words, $\mathcal{Z}_{t,s}(x,y)$ is the propagator of the SHE from (s,y) to (t,x).

Fix $t > 1$ and for simplicity assume $t = N \in \mathbb{Z}$, we have

$$u(t,x) = \int_{\mathbb{T}^d} \mathcal{Z}_{t,t-1}(x,y_1)u(t-1,y_1)dy_1. \tag{5.14}$$

Iterating the above formula, we derive

$$u(t,x) = \int_{\mathbb{T}^{Nd}} \prod_{j=0}^{N-1} \mathcal{Z}_{t-j,t-j-1}(y_j,y_{j+1})u_0(y_N)dy_1\dots dy_{N-1}dy_N. \tag{5.15}$$

where $y_0 = x$ and u_0 is the initial data for (5.3). It is worth noting that the above expression resembles the product of random matrices

Following Sinai [50], it seems to be rather standard to introduce a Markov chain to rewrite $u(t,x)$. The Markov chain $\{Y_n\}_{n=1}^N$, which takes values in $\mathbb{T}^d$, is constructed as follows. We run the chain backward in time. Let $\pi_N(y_N)$ be the density of X_N and $\pi_k(y_k\,|\,y_{k+1})$ be the transition density from $Y_{k+1} = y_{k+1}$ to $Y_k = y_k$:

(5.16)

$$\pi_1(y_1\,|\,y_2) = \frac{\mathcal{Z}_{t-1,t-2}(y_1,y_2)}{\int_{\mathbb{T}^d}\mathcal{Z}_{t-1,t-2}(y_1,y_2)dy_1}$$

$$\pi_k(y_k\,|\,y_{k+1}) = \frac{\int_{\mathbb{T}^{(k-1)d}}\mathcal{Z}_{t-1,t-2}(y_1,y_2)\dots\mathcal{Z}_{t-k,t-k-1}(y_k,y_{k+1})dy_1\dots dy_{k-1}}{\int_{\mathbb{T}^{kd}}\mathcal{Z}_{t-1,t-2}(y_1,y_2)\dots\mathcal{Z}_{t-k,t-k-1}(y_k,y_{k+1})dy_1\dots dy_{k-1}dy_k}, \qquad k \leq N-1.$$

and

$$\pi_N(y_N) = \frac{\int_{\mathbb{T}^{(N-1)d}} \left(\prod_{j=1}^{N-1} \mathcal{Z}_{t-j,t-j-1}(y_j, y_{j+1})\right) u_0(y_N) dy_1 \dots dy_{N-1}}{\int_{\mathbb{T}^{Nd}} \left(\prod_{j=1}^{N-1} \mathcal{Z}_{t-j,t-j-1}(y_j, y_{j+1})\right) u_0(y_N) dy_1 \dots dy_N}$$

We first note that the joint density of $Y_1, \dots, Y_N$ is

$$\pi_1(y_1 \,|\, y_2) \dots \pi_{N-1}(y_{N-1} \,|\, y_N) \pi_N(y_N) = A \left(\prod_{j=1}^{N-1} \mathcal{Z}_{t-j,t-j-1}(y_j, y_{j+1}) \right) u_0(y_N)$$

with the (random) constant A given by

$$A = \left(\int_{\mathbb{T}^{Nd}} \left(\prod_{j=1}^{N-1} \mathcal{Z}_{t-j,t-j-1}(y_j, y_{j+1}) \right) u_0(y_N) dy_1 \dots dy_N \right)^{-1}.$$

Let $\mathbb{E}_\pi$ denote the expectation of the Markov chain, then (5.15) can be written as

(5.17)

$$\begin{aligned} u(t,x) =& \int_{\mathbb{T}^{Nd}} \mathcal{Z}_{t,t-1}(x, y_1) \pi_1(y_1 \,|\, y_2) \dots \pi_{N-1}(y_{N-1} \,|\, y_N) \pi_N(y_N) A^{-1} dy_1 \dots dy_N \\ =& A^{-1} \mathbb{E}_\pi[\mathcal{Z}_{t,t-1}(x, Y_1)] \end{aligned}$$

thus,

$$\rho(t,x) = \frac{u(t,x)}{\int_{\mathbb{T}^d} u(t,y) dy} = \frac{\mathbb{E}_\pi[\mathcal{Z}_{t,t-1}(x, Y_1)]}{\int_{\mathbb{T}^d} \mathbb{E}_\pi[\mathcal{Z}_{t,t-1}(y, Y_1)] dy} \tag{5.18}$$

The key here is that the random constant A disappears when we take the ratio (the same happens to the Burgers equation). Note that the dependence of $\rho(t,x)$ on the random environment $\{\xi(s,\cdot) : s \le t-1\}$ is only through the random variable Y_1, hence it suffices to study the mixing property of the Markov chain.

We note that

$$\inf_{y_k, y_{k+1} \in \mathbb{T}^d} \pi_k(y_k \,|\, y_{k+1}) \ge \frac{\inf_{y_{k-1}, y_k, y_{k+1} \in \mathbb{T}^d} \mathcal{Z}_{t-k+1,t-k}(y_{k-1}, y_k) \mathcal{Z}_{t-k,t-k-1}(y_k, y_{k-1})}{\sup_{y_{k-1}, y_{k+1}} \int_{\mathbb{T}^d} \mathcal{Z}_{t-k+1,t-k}(y_{k-1}, y_k) \mathcal{Z}_{t-k,t-k-1}(y_k, y_{k+1}) dy_k}$$

and the r.h.s. only depends on the noise in the time interval $[t-k-1, t-k+1]$. For any s, the random variable $\mathcal{Z}_{s,s-1}(x,y)$ is roughly the heat kernel multiplying a random factor that can be statistically bounded away from zero and infinity (for this we need the positive and negative moments estimate of the Green's function of SHE). Since the heat kernel (at time 1) is also bounded away from zero and infinity (for this we use the fact that the spatial domain is compact), it is not hard to imagine that, with a positive probability, the transition density π_k can be bounded from below by a positive constant. More precisely, one can show that , there exists a $\delta > 0$ depending on $d, R(\cdot)$ and events $B_{t-k+1,t-k-1}(\delta)$ only involving the random environment $\{\xi(s,\cdot) : s \in [t-k-1, t-k+1]\}$ such that

$$\mathbb{P}[B_{t-k+1,t-k-1}(\delta)] > \delta \tag{5.19}$$

and under $B_{t-k+1,t-k-1}(\delta)$, the transition density

$$\inf_{y_k, y_{k+1} \in \mathbb{T}^d} \pi_k(y_k \,|\, y_{k+1}) > \delta.$$

With the above Doeblin condition, one can easily construct a coupling to show that the Markov chain mixes fast. Suppose that $s < t$, we consider the subintervals of of the form $[t-k-1, t-k+1]$ in $[s,t]$, and if one of those events $B_{t-k+1,t-k-1}(\delta)$ occurs, then the transition kernel π_k is bounded from below by δ, so we can write

$$\pi_k(y_k \,|\, y_{k+1}) = \delta + (1-\delta) \tfrac{\pi_k(y_k \,|\, y_{k+1}) - \delta}{1-\delta}.$$

Introducing a Bernoulli random variable τ_k such that $\mathbb{P}[\tau_k = 1] = \delta$, to sample from $\pi_k(y_k \,|\, y_{k+1})$, it is equivalent with sampling from the uniform distribution if $\tau_k = 1$

and from $\frac{\pi_k(y_k \mid y_{k+1})-\delta}{1-\delta}$ if $\tau_k = 0$. Thus, with probability δ, the distribution of Y_k given $Y_{k+1} = y_{k+1}$ is uniform, which renews the chain. We also easily see that the probability of not renewing is exponentially small with respect to the length of $t-s$. This leads to the fast mixing of $\rho(t)$:

Proposition 5.4. *For any $s < t$, there exists a random function $\rho_s(t,\cdot)$ that is a function of the noise in $[s,t]\times\mathbb{R}^d$, such that for any $p \geq 1$,*

$$\mathbb{E}\|\rho(t) - \rho_s(t)\|_\infty^p \leq Ce^{-\lambda(t-s)}$$

with the constants $\lambda, C > 0$ only depending on $d, R(\cdot), p$.

References. The approach presented here is based on [29]. Results on the fluctuations of the KPZ equation on the whole line can be found e.g. in [3, 9, 12, 13].

References

[1] T. Alberts, K. Khanin, and J. Quastel, *The continuum directed random polymer*, Journal of Statistical Physics, 154 (2014), pp. 305–326. 30

[2] ———, *The intermediate disorder regime for directed polymers in dimension 1+1*, Annals of Probability, 42 (2014), pp. 1212–1256. 18

[3] G. Amir, I. Corwin, and J. Quastel, *Probability distribution of the free energy of the continuum directed random polymer in 1+1 dimensions*, Communications on pure and applied mathematics, 64 (2011), pp. 466–537. 34

[4] J. Baik and Z. Liu, *Tasep on a ring in sub-relaxation time scale*, Journal of Statistical Physics, 165 (2016), pp. 1051–1085. 29

[5] ———, *Fluctuations of tasep on a ring in relaxation time scale*, Communications on Pure and Applied Mathematics, 71 (2018), pp. 747–813. 29

[6] J. Baik, Z. Liu, and G. L. Silva, *Limiting one-point distribution of periodic tasep*, arXiv preprint arXiv:2008.07024, (2020). 29

[7] G. Bal, *Homogenization with large spatial random potential*, Multiscale Modeling & Simulation, 8 (2010), pp. 1484–1510. 16

[8] ———, *Convergence to homogenized or stochastic partial differential equations*, Applied Mathematics Research eXpress, 2011 (2011), pp. 215–241. 16

[9] M. Balázs, J. Quastel, and T. Seppäläinen, *Fluctuation exponent of the kpz/stochastic burgers equation*, Journal of the American Mathematical Society, 24 (2011), pp. 683–708. 34

[10] L. Bertini and N. Cancrini, *The stochastic heat equation: Feynman-kac formula and intermittence*, Journal of statistical Physics, 78 (1995), pp. 1377–1401. 18

[11] L. Bertini and G. Giacomin, *Stochastic burgers and kpz equations from particle systems*, Communications in mathematical physics, 183 (1997), pp. 571–607. 18

[12] A. Borodin, I. Corwin, and P. Ferrari, *Free energy fluctuations for directed polymers in random media in $1+1$ dimension*, Communications on Pure and Applied Mathematics, 67 (2014), pp. 1129–1214. 34

[13] A. Borodin, I. Corwin, P. Ferrari, and B. Veto, *Height fluctuations for the stationary kpz equation*, Mathematical Physics, Analysis and Geometry, 18 (2015), pp. 1–95. 34

[14] F. Caravenna, R. Sun, and N. Zygouras, *The two-dimensional kpz equation in the entire subcritical regime*, Annals of Probability, 48 (2020), pp. 1086–1127. 28

[15] A. Chandra and H. Shen, *Moment bounds for spdes with non-gaussian fields and application to the wong-zakai problem*, Electronic Journal of Probability, 22 (2017). 23

[16] S. Chatterjee, A. Dunlap, et al., *Constructing a solution of the $(2+1)$-dimensional kpz equation*, Annals of Probability, 48 (2020), pp. 1014–1055. 28

[17] L. Chen, D. Khoshnevisan, D. Nualart, and F. Pu, *Central limit theorems for spatial averages of the stochastic heat equation via malliavin-stein's method*, arXiv preprint arXiv:2008.02408, (2020). 23

[18] X. Chen, *Precise intermittency for the parabolic anderson equation with an $(1+1)$-dimensional time–space white noise*, in Annales de l'IHP Probabilités et statistiques, vol. 51, 2015, pp. 1486–1499. 29

[19] F. Comets, *Directed polymers in random environments*, Springer, 2017. 15

[20] F. Comets, C. Cosco, and C. Mukherjee, *Space-time fluctuation of the kardar-parisi-zhang equation in $d \geq 3$ and the gaussian free field*, arXiv preprint arXiv:1905.03200, (2019). 28

[21] I. Corwin, *The kardar–parisi–zhang equation and universality class*, Random matrices: Theory and applications, 1 (2012), p. 1130001. 10

[22] I. Corwin and H. Shen, *Some recent progress in singular stochastic pdes*, arXiv preprint arXiv:1904.00334, (2019). 10
[23] C. Cosco, S. Nakajima, and M. Nakashima, *Law of large numbers and fluctuations in the sub-critical and l^2 regions for she and kpz equation in dimension $d \geq 3$*, arXiv preprint arXiv:2005.12689, (2020). 28
[24] A. Dunlap, Y. Gu, L. Ryzhik, and O. Zeitouni, *Fluctuations of the solutions to the kpz equation in dimensions three and higher*, Probability Theory and Related Fields, 176 (2020), pp. 1217–1258. 28
[25] Y. Gu, *Gaussian fluctuations from the 2d kpz equation*, Stochastics and Partial Differential Equations: Analysis and Computations, 8 (2020), pp. 150–185. 28
[26] Y. Gu and G. Bal, *Fluctuations of parabolic equations with large random potentials*, Stochastic Partial Differential Equations: Analysis and Computations, 3 (2015), pp. 1–51. 16
[27] ———, *Homogenization of parabolic equations with large time-dependent random potential*, Stochastic Processes and their Applications, 125 (2015), pp. 91–115. 16
[28] ———, *Weak convergence approach for parabolic equations with large, highly oscillatory, random potential*, in Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, vol. 52, Institut Henri Poincaré, 2016, pp. 261–285. 16
[29] Y. Gu and T. Komorowski, *Kpz on torus: Gaussian fluctuations*, arXiv preprint arXiv:2104.13540, (2021). 34
[30] Y. Gu and L. Ryzhik, *The random schrödinger equation: Homogenization in time-dependent potentials*, Multiscale Modeling & Simulation, 14 (2016), pp. 323–363. 12
[31] Y. Gu, L. Ryzhik, and O. Zeitouni, *The edwards–wilkinson limit of the random heat equation in dimensions three and higher*, Communications in Mathematical Physics, 363 (2018), pp. 351–388. 28
[32] Y. Gu and L.-C. Tsai, *Another look into the wong–zakai theorem for stochastic heat equation*, Annals of Applied Probability, 29 (2019), pp. 3037–3061. 23
[33] M. Gubinelli, P. Imkeller, and N. Perkowski, *Paracontrolled distributions and singular pdes*, in Forum of Mathematics, Pi, vol. 3, Cambridge University Press, 2015. 10
[34] M. Gubinelli and N. Perkowski, *Lectures on singular stochastic pdes*, arXiv preprint arXiv:1502.00157, (2015). 10, 12
[35] ———, *The infinitesimal generator of the stochastic burgers equation*, Probability Theory and Related Fields, 178 (2020), pp. 1067–1124. 30
[36] M. Hairer, *A theory of regularity structures*, Inventiones mathematicae, 198 (2014), pp. 269–504. 10
[37] M. Hairer and C. Labbé, *Multiplicative stochastic heat equations on the whole space*, Journal of the European Mathematical Society, 20 (2018), pp. 1005–1054. 23
[38] M. Hairer and É. Pardoux, *A wong-zakai theorem for stochastic pdes*, Journal of the Mathematical Society of Japan, 67 (2015), pp. 1551–1604. 4, 23
[39] M. Hairer, E. Pardoux, and A. Piatnitski, *Random homogenisation of a highly oscillatory singular potential*, Stochastic Partial Differential Equations: Analysis and Computations, 1 (2013), pp. 571–605. 16
[40] D. Khoshnevisan, *Analysis of stochastic partial differential equations*, vol. 119, American Mathematical Soc., 2014. 5, 10, 16
[41] T. Komorowski, C. Landim, and S. Olla, *Fluctuations in Markov processes: time symmetry and martingale approximation*, vol. 345, Springer Science & Business Media, 2012. 16
[42] A. Lejay, *Homogenization of divergence-form operators with lower-order terms in random media*, Probability theory and related fields, 120 (2001), pp. 255–276. 16
[43] D. Lygkonis and N. Zygouras, *Edwards-wilkinson fluctuations for the directed polymer in the full l^2-regime for dimensions $d \geq 3$*, arXiv preprint arXiv:2005.12706, (2020). 28
[44] J. Magnen and J. Unterberger, *The scaling limit of the kpz equation in space dimension 3 and higher*, Journal of Statistical Physics, 171 (2018), pp. 543–598. 28
[45] D. Nualart, *The Malliavin calculus and related topics*, vol. 1995, Springer, 2006. 18
[46] É. Pardoux and A. Piatnitski, *Homogenization of a singular random one-dimensional pde with time-varying coefficients*, The Annals of Probability, 40 (2012), pp. 1316–1356. 3, 13, 16
[47] J. Quastel, *Introduction to kpz*, Current developments in mathematics, 2011 (2011). 10
[48] J. Quastel and S. Sarkar, *Convergence of exclusion processes and kpz equation to the kpz fixed point*, arXiv e-prints, (2020), pp. arXiv–2008. 2
[49] J. Quastel and H. Spohn, *The one-dimensional kpz equation and its universality class*, Journal of Statistical Physics, 160 (2015), pp. 965–984. 10
[50] Y. G. Sinai, *Two results concerning asymptotic behavior of solutions of the burgers equation with force*, Journal of statistical physics, 64 (1991), pp. 1–12. 32

[51] B. Virag, *The heat and the landscape i*, arXiv preprint arXiv:2008.07241, (2020). 2
[52] N. Zygouras, *The 2d kpz as a marginally relevant disordered system*, (2019). 10