EFFECTIVE DIFFUSIVITIES IN PERIODIC KPZ

YU GU, TOMASZ KOMOROWSKI

ABSTRACT. For the KPZ equation on a torus with a $1 + 1$ spacetime white noise, it was shown in [34, 28] that the height function satisfies a central limit theorem, and the variance can be written as the expectation of an exponential functional of Brownian bridges. In this paper, we consider another physically relevant quantity, the winding number of the directed polymer on a cylinder, or equivalently, the displacement of the directed polymer endpoint in a spatially periodic random environment. It was shown in [33] that the polymer endpoint satisfies a central limit theorem on diffusive scales. The main result of this paper is an explicit expression of the effective diffusivity, in terms of the expectation of another exponential functional of Brownian bridges. Our argument is based on a combination of tools from Malliavin calculus, homogenization, and diffusion in distribution-valued random environments.

KEYWORDS: KPZ equation, directed polymer, diffusion in random environment, homogenization.

1. INTRODUCTION

1.1. **Main result.** We are interested in the stochastic heat equation (SHE) with a spatially periodic space-time white noise

$$
\begin{aligned}
(1.1) \qquad \partial_t Z &= \tfrac{1}{2}\Delta Z + \beta Z \xi, \qquad t > 0, x \in \mathbb{R}, \\
Z(0, x) &= \delta(x).
\end{aligned}
$$

Here ξ is a generalized Gaussian random field over $\mathbb{R} \times \mathbb{R}$ with the covariance function

$$
\mathbf{E}[\xi(t, x)\xi(s, y)] = \delta(t - s) \sum_{n \in \mathbb{Z}} \delta(x - y + n).
$$

We assume that ξ is built on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and view it as a spacetime white noise on $\mathbb{R} \times \mathbb{T}$ that is periodically extended to $\mathbb{R} \times \mathbb{R}$. Here $\mathbb{T}$ is the unit torus, defined as the interval $[0, 1]$ with identified endpoints. The product between Z and ξ is interpreted in the Itô-Walsh sense.

The random function Z is associated with the model of a directed polymer in a random environment and the parameter $\beta > 0$ plays the role of the inverse temperature. There are two important physical quantities: the free energy and the endpoint displacement of the polymer. Define $\bar{Z}_t = \int_{\mathbb{R}} Z(t, x)dx$ as the point-to-line partition function. It was known that $\log \bar{Z}_t$ satisfies a central limit theorem, with the drift and the variance described explicitly in terms of some auxiliary Brownian bridges, see [34, Theorem 1.1], [15, Proposition 4.1] and [28, Eq. (2.10), (4.2)]:

$$
\frac{\log \bar{Z}_t + \gamma(\beta)t}{\sqrt{t}} \Rightarrow N(0, \Sigma^2(\beta)), \quad \text{as } t \to \infty,
$$

where $\gamma(\beta), \Sigma^2(\beta) > 0$ are constants given by

(1.2)
$$\gamma(\beta) = \frac{\beta^2}{2} \mathbb{E}_{W_1}\left[\frac{1}{\left(\int_0^1 e^{\beta W_1(y)}dy\right)^2}\right] = \frac{\beta^2}{2} + \frac{\beta^4}{24},$$

$$\Sigma^2(\beta) = \beta^2 \mathbb{E}_{W_1}\left[\left(\mathbb{E}_{W_2}\frac{1}{\int_0^1 e^{\beta(W_1(y)+W_2(y))}dy}\right)^2\right].$$

Here W_1, W_2 are independent Brownian bridges connecting $(0,0)$ and $(1,0)$, and throughout the paper, we use $\mathbb{E}_{W_i}$ as the expectation on W_i.

Define

(1.3)
$$\rho(t,x) = \frac{Z(t,x)}{\bar{Z}_t}, \quad x \in \mathbb{R},$$

which is the quenched endpoint density of the associated directed polymer. It was shown in [33] that there exists some $\sigma^2(\beta) > 0$ such that for any f belonging to $C_b(\mathbb{R})$ - the space of bounded continuous functions on $\mathbb{R}$ - we have

(1.4)
$$\mathbf{E}\int_{\mathbb{R}} f\left(\frac{x}{\sqrt{t}}\right)\rho(t,x)dx \to \int_{\mathbb{R}} f(x)\frac{1}{\sqrt{2\pi\sigma^2(\beta)}}\exp\left(-\frac{x^2}{2\sigma^2(\beta)}\right)dx,$$

as $t \to \infty$. In other words, the annealed endpoint distribution converges under the diffusive scaling to a centered Gaussian with variance $\sigma^2(\beta)$.

The main result of the paper is to derive an explicit expression of $\sigma^2(\beta)$:

Theorem 1.1. *Suppose that $\{W_i\}_{i=1,2,3}$ are three independent Brownian bridges connecting $(0,0)$ and $(1,0)$. Then, the effective diffusivity $\sigma^2(\beta)$ admits the following expression:*

(1.5)
$$\sigma^2(\beta) = 1 + \beta^2 \mathbb{E}_{W_1}\mathbb{E}_{W_3}\left[\mathcal{A}(\beta, W_1, W_3)^2\right],$$

where

$$\mathcal{A}(\beta, W_1, W_3) = \int_{\mathbb{T}^2} \Xi(\beta, y, W_1)\left(\frac{e^{\beta W_1(z)+\beta W_3(z)}}{\int_{\mathbb{T}} e^{\beta W_1(z')+\beta W_3(z')}dz'} - 1\right)1_{[0,y]}(z)dydz,$$

with

$$\Xi(\beta, y, W_1) = \mathbb{E}_{W_2}\left[\frac{e^{\beta W_2(y)-\beta W_1(y)}}{\left(\int_{\mathbb{T}} e^{\beta W_2(y')}e^{-\beta W_1(y')}dy'\right)^2}\right].$$

1.2. **Motivation.** There is an intriguing relation between the two diffusion constants $\Sigma^2(\beta)$ and $\sigma^2(\beta)$, predicted by Éric Brunet through the replica method, see [12, Eq. (20)]. We translate it in our notations[1]:

(1.6)
$$\sigma^2(\beta) = \frac{2}{\beta^2}\Sigma^2(\beta) - \frac{1}{2\beta}\frac{d}{d\beta}\Sigma^2(\beta) = -\frac{\beta^3}{2}\frac{d}{d\beta}\left(\frac{\Sigma^2(\beta)}{\beta^4}\right), \qquad \beta > 0.$$

In order to apply the replica method to study the fluctuations of the free energy $\log \bar{Z}_t$, it reduces to computing the ground state energy of the n–particle Delta Bose gas

$$\mathcal{H}_n = \frac{1}{2}\sum_{i=1}^n \nabla_{x_i}^2 + \beta^2 \sum_{1 \le i < j \le n} \delta(x_i - x_j)$$

on the torus (with periodic boundary condition), for all $n \ge 1$. This was done in [13, 14] by solving the Bethe Ansatz equation. The quenched limiting free energy $\gamma(\beta)$ in (1.2) matches the prediction given by the replica method, see [12, Eq. (13)]. There is also a

---

[1]The γ parameter in [12, Eq. (20)] is β^2 in our notation.

predicted integral form for $\Sigma^2(\beta)$, see [12, Eq. (14)], but we do not know how to compute the expectation in the expression of $\Sigma^2(\beta)$ in (1.2).

Meanwhile, it is more complicated to compute the effective diffusivity of the polymer endpoint using the replica method, and it turns out that one needs to find the ground state energy of the same Hamiltonian but with a different boundary condition, see [12, Section 3.1]. Through solving a new Bethe Ansatz equation, Brunet found a relation between the ground state energies of the two Hamiltonians with different boundary conditions, through which he derived (1.6).

It was our goal to justify (1.6) rigorously, and to understand why there exists such a relation between the two (non-universal) diffusion constants. With Theorem 1.1, it seems that we are halfway there, since both $\sigma^2(\beta)$ and $\Sigma^2(\beta)$ are now written as the expectations of exponential functionals of independent Brownian bridges. Nevertheless, to prove (1.6) using the expressions in (1.2) and (1.5) seems to be a highly nontrivial problem on Brownian bridges, which we choose not to pursue here. Using (1.2) and (1.5), one can perform a small β expansion, which indeed matches the predictions in [12], and suggests the validity of (1.6):

Corollary 1.2. *For $\beta \ll 1$, we have*

(1.7)
$$\sigma^2(\beta) = 1 + \frac{\beta^4}{360} + O(\beta^6),$$
$$\Sigma^2(\beta) = \beta^2 + \frac{\beta^4}{12} - \frac{\beta^6}{360} + O(\beta^8).$$

Another reason why one may be interested in explicit expressions of the diffusion constants comes from the attempt to understand the $1:2:3$ scaling in the $1+1$ KPZ universality class. In this paper, the length of the torus was fixed to be $L = 1$. If we denote the diffusion constants by $\Sigma^2(\beta, L)$ and $\sigma^2(\beta, L)$ for general $L > 0$, it is straightforward to derive the following relations through a rescaling:

$$\Sigma^2(\beta, 1) = L^2 \Sigma^2(\frac{\beta}{\sqrt{L}}, L),$$
$$\sigma^2(\beta, 1) = \sigma^2(\frac{\beta}{\sqrt{L}}, L).$$

Therefore, the asymptotic behaviors of $\Sigma^2(\beta, 1)$ and $\sigma^2(\beta, 1)$ as $\beta \to \infty$ translate into the behaviors of $\Sigma^2(1, L)$ and $\sigma^2(1, L)$ as $L \to \infty$. For $L = \infty$ with the problem posed on the whole line, it is known that the free energy and the polymer endpoint behave sub- and super-diffusively with the $1/3$ and $2/3$ exponents, so one expects that $\Sigma^2(1, \infty) = 0$ and $\sigma^2(1, \infty) = \infty$. As a matter of fact, the decaying rate of $\Sigma^2(1, L) \downarrow 0$ and the growing rate of $\sigma^2(1, L) \uparrow \infty$ as $L \to \infty$ is related to the $1:2:3$ scaling:

(i) For the height function (free energy), the critical scale comes from the balance of the one-point variance $t\Sigma^2(1, L)$ and the transversal roughness L. The transversal roughness here refers to the fluctuations of $h(t, x) - h(t, 0)$ at stationarity, where h solves the KPZ equation on the torus. It was shown in [28, Theorem 4.1] that, as $L \to \infty$, the variance $\Sigma^2(1, L)$ decays like $L^{-1/2}$, which suggests the $1:2:3$ scaling and also leads to the optimal variance of the height function in certain regimes with $t, L \to \infty$ together, see [28, Theorem 1.1]. The heuristic is that, suppose $t\Sigma^2(1, L) \asymp tL^{-1/2} \gg L$, then the transversal roughness is much smaller than the one-point fluctuation, and if we write $h(t, x) = h(t, 0) + (h(t, x) - h(t, 0))$, then the second term on the r.h.s. has a much smaller fluctuation than the first term, so that $h(t, x)$ and $h(t, 0)$ are almost perfectly correlated for $t \gg 1$. Only when $tL^{-1/2}$ is of order $O(L)$, we have the nontrivial correlation kicks in, and this gives $L \asymp t^{2/3}$.

(ii) For the polymer endpoint, the critical scale comes from the balance of the diffusive standard deviation $\sqrt{t\sigma^2(1,L)}$ and the length of the cell L. It is expected and also predicted by the replica method that $\sigma^2(1,L) \asymp L^{1/2}$, which suggests the $1:2:3$ scaling. Roughly speaking, in the region of $\sqrt{t\sigma^2(1,L)} \asymp \sqrt{tL^{1/2}} \gg L$, the polymer path visits many cells and one still expects to see the homogenization phenomenon and the central limit theorem, thus, it is natural to guess the critical length scale comes from the relation $\sqrt{tL^{1/2}} \asymp L$, which leads to $L \asymp t^{2/3}$. Starting from the expression (1.5), one may try to show the equivalent statement of $\sigma^2(\beta,1) \asymp \beta$ as $\beta \to \infty$, although we do not pursue it here.

In the end, let us point out that, in order to prove the asymptotic behaviors of the diffusion constants, it simplifies greatly if one could derive some explicit formulas first. For general homogenization type problems, the effective diffusivity is typically written in an abstract form, in some cases through the solution to a cell problem, the so-called corrector, in some other cases in terms of the Green-Kubo formula or the solution to a certain variational problem. Therefore, it was actually a surprise to us that the $\sigma^2(\beta)$ (and $\Sigma^2(\beta)$) may be expressed explicitly in terms of the underlying invariant measure for the KPZ equation on the unit torus, which is the Brownian bridge. In a sense, the corrector is explicit in our problem, which we will discuss in more detail later. As mentioned previously, the central limit theorem was already derived in [33], with $\sigma^2(\beta)$ written as the integral of an abstract covariance function, but it seems hopeless to us to do any quantitative analysis from that expression.

1.3. **Context.** The study of the KPZ equation and the $1+1$ KPZ universality class has witnessed remarkable progress during the past years. We will not attempt to survey the huge literature here, and refer the readers to the reviews [21, 46, 47] and the references therein. For a nice introduction to the SHE, we refer to the monograph [40]. Our problem is more related to those in a bounded region since the compactness makes a huge difference concerning the large scale random fluctuations. For (totally) asymmetric simple exclusion processes on a finite line segment with periodic or open boundaries, exact diffusion constants for the particle current were derived in [25, 26, 27]. More recently, there is a series of work [3, 4, 5, 6, 7, 44] on the regimes where the size of the torus and the time go to infinity simultaneously, and the fluctuations of the height function were derived in different cases. For the open KPZ equation, the one on an interval with Neumann boundary conditions, a lot of recent progress has been on the construction of explicit invariant measures [9, 10, 16, 22, 23].

The problem we consider here can also be considered as the homogenization of a diffusion in a random environment. Although the directed polymer is almost always formulated in the form of Gibbs measures of paths in a random environment, it can also be viewed as a passive scalar problem, with the drift given by the solution to a stochastic Burgers equation. This perspective plays a very important role in our analysis: for the polymer endpoint, one can decompose its total variance into two parts, the variance of the quenched mean and the mean of the quenched variance, which, roughly speaking, correspond to the disorder variance and the thermal variance respectively. It turns out that the main difficulty comes from analyzing the quenched mean, which is actually the corrector if we put it in the context of homogenization, see Section 1.4 for more discussion. For our case of a spacetime white noise, the solution to the stochastic Burgers equation is not function-valued, so one needs to deal with the so-called singular diffusion. Luckily for us, there has been a lot of recent progress on the study of diffusion with (very singular) distributional drifts [17, 24, 36, 38, 42], which was inspired by the breakthrough in the study of singular SPDEs, using tools such as rough paths [36], regularity structures [37] and paracontrolled calculus [35], and we will borrow tools from there.

Another important ingredient we rely on to derive the explicit expressions of the diffusion constants is the so-called Clark-Ocone formula from Malliavin calculus. It involves computing the Malliavin derivative of the free energy or the quenched mean of the polymer endpoint with respect to the underlying white noise, and rewriting their fluctuations in the form of Itô integrals. The key here is the *Itô integral*, which can be viewed as the sum of martingale differences, so that the second moment can be computed in a rather straightforward way. To compare with the homogenization argument, one should note that, from a probabilistic perspective, the whole point of constructing the corrector through solving a certain Poisson equation is to extract the martingale part from the drift, which contributes to the effective diffusivity. In a certain sense, the Clark-Ocone formula provides a different way of performing the semi-martingale decomposition, and in some cases it leads to "new" formulas like (1.2) and (1.5). For the directed polymer, there is a simple yet crucial property that the derivative of the free energy with respect to the noise is given by the quenched density of the polymer path. Compared to other homogenization problems, it is this special property that makes the Clark-Ocone formula useful in this case.

In order to study the super-diffusive behaviors in turbulent transport, it is natural to analyze how the effective diffusivity diverges as the domain in consideration becomes larger and larger. In [30], the so-called box diffusivity was defined, and its asymptotics was analyzed to derive the relevant scaling exponents of several models. More recently, for the model of a Brownian particle in the curl of a two-dimensional Gaussian free field, the same approach was adopted in [19, 2], and through a "progressive" homogenization, the precise order of the mean square displacement and a quenched central limit theorem were derived. Similar results were obtained in [18, 32], taking inspirations from [43, 49].

1.4. **Sketch of the argument and main challenges.** In this section, we will formulate the problem as a diffusion in a random environment and explain on a formal level the main steps of the proof. Some proofs will be written in different ways later (as always), but we are hoping that a bird's-eye view at this stage will help the readers to understand the ideas in the argument.

To study the polymer endpoint, through a Girsanov transformation, it is equivalent with analyzing the following diffusion with a random drift:

$$(1.8) \qquad d\mathcal{X}_s = u(t - s, \mathcal{X}_s)ds + dB_s, \qquad \mathcal{X}_0 = 0,$$

where B is a standard Brownian motion and u is the stationary solution to the stochastic Burgers equation

$$(1.9) \qquad \partial_t u = \tfrac{1}{2}\Delta u + \tfrac{1}{2}\nabla u^2 + \beta\nabla\xi.$$

Here the driving force ξ of the evolving random environment is independent of the thermal noise B. Then $\mathcal{X}_t$ can be written as

$$\mathcal{X}_t = \int_0^t u(t - s, \mathcal{X}_s)ds + B_t,$$

and, as in standard homogenization, the first step is to extract the martingale part from the drift, see e.g. [41, Chapters 9 and 11]. To do that, we consider the following equation

$$(1.10) \qquad \partial_t \phi = \frac{1}{2}\Delta\phi + u\nabla\phi + u, \quad \phi(0, x) = 0,$$

and call ϕ the corrector, which could be obtained through a formal two-scale expansion of the backward Kolmogorov equation with drift u. More precisely, consider the equation

$$\partial_t f_\varepsilon = \tfrac{1}{2}\Delta f_\varepsilon + \tfrac{1}{\varepsilon}u\big(\tfrac{t}{\varepsilon^2}, \tfrac{x}{\varepsilon}\big)\nabla f_\varepsilon,$$

if we plug in the ansatz $f_\varepsilon(t, x) = \bar{f}(t, x) + \varepsilon\mathfrak{f}_1(t, x, s, y) + \varepsilon^2\mathfrak{f}_2(t, x, s, y) + \ldots$, with the fast variables $s = \tfrac{t}{\varepsilon^2}, y = \tfrac{x}{\varepsilon}$, and identity the terms of the same order in ε, then we obtain $\mathfrak{f}_1(t, x, s, y) = \partial_x\bar{f}(t, x)\phi(s, y)$, where ϕ solves an equation of the form (1.10).

There are different ways of formulating the corrector equation which may or may not incorporate the evolution of the underlying environment. For our purpose, it is more convenient to consider the above one, where the dynamics of u is not taken into account. One can show that (with $\mathbb{E}_B$ denoting the expectation only on B)

$$(1.11) \qquad \phi(t,0) = \mathbb{E}_B \int_0^t u(t-s, \mathcal{X}_s) ds,$$

and

$$(1.12) \qquad \mathcal{X}_t = \phi(t,0) + \int_0^t (1 + \nabla\phi(t-s, \mathcal{X}_s)) dB_s.$$

In this way, one can view the corrector $\phi(t,0)$ as the quenched mean: $\phi(t,0) = \mathbb{E}_B \mathcal{X}_t$. Since u is periodic in space, one can also write $\phi(t,0) = \mathbb{E}_B \int_0^t u(t-s, \dot{\mathcal{X}}_s) ds$, with $\dot{\mathcal{X}}$ representing the fractional part of $\mathcal{X}$. It is not hard to imagine that both the solution to the Burgers equation and the process $\dot{\mathcal{X}}_s$ are fast mixing in time, so $\phi(t,0)$ can be viewed as a sum of weakly dependent random variables with $\phi(t,0)/\sqrt{t}$ satisfying a central limit theorem. Therefore, both terms on the r.h.s. of (1.12) contribute to the limiting effective diffusivity. This is in sharp contrast to the case of a divergence form operator or a divergence free drift, where the corrector grows sublinearly hence does not contribute to the limiting diffusivity.

It turns out the martingale part $\int_0^t (1 + \nabla\phi(t-s, \mathcal{X}_s)) dB_s$ is easy to study, because, for directed polymer in a random environment that is statistically shear-invariant (which is our case), through the Gibbs measure formulation, one can show that the mean of the quenched variance equals to t:

$$\mathbf{E}\mathbb{E}_B \int_0^t (1 + \nabla\phi(t-s, \mathcal{X}_s))^2 ds = t.$$

This is where the "1" factor in (1.5) comes from.

The main difficulty of the paper then reduces to deriving an explicit variance of $\frac{\phi(t,0)}{\sqrt{t}}$. At this point, it is worth emphasizing that $\mathcal{X}_t$ does *not* satisfy a quenched central limit theorem. More precisely, for the polymer endpoint $\mathcal{X}_t$, one can view the $\phi(t,0)$ on the r.h.s. of (1.12) as describing the fluctuations resulting from the disorders, and the Itô integral with respect to B as the thermal fluctuations. In our periodic setting, we expect a quenched central limit theorem to hold for the Itô integral term, while the $\phi(t,0)$ term is "deterministic" if we quench the random environment, and $\frac{\phi(t,0)}{\sqrt{t}}$ does not vanish as $t \gg 1$. A similar case was discussed in [31], with a non-quenched central limit theorem proved for a diffusion in a random environment. On a different note, one should view the $\phi(t,0)$ term as a measurement of how "localized" the polymer endpoint is: suppose we sample two different polymer paths from the quenched Gibbs measure, they "share" the $\phi(t,0)$ term. To compare to the non-periodic setting, where $\mathcal{X}_t$ is expected to be of order $t^{2/3}$ and to localize around some favorite point, we suspect that, as the size of the torus goes to infinity, this favorite point comes from the $\phi(t,0)$ term, which is a functional of the random environment. Meanwhile, the Itô integral term becomes of order $O(1)$ as $t \to \infty$, so $\mathcal{X}_t$ stays $O(1)$ distance away from $\phi(t,0)$. The second moment of the Itô integral term blows up as $t \to \infty$, but that is due to the heavy tail [29].

To study the quenched mean, it is more convenient to consider the Gibbs formulation. We study $X_t = \int_{\mathbb{R}} x\rho(t,x) dx$, which has the same law as $\phi(t,0)$. Applying the Clark-Ocone formula, we can write X_t as an Itô integral, with respect to the underlying white noise ξ:

$$(1.13) \qquad X_t = \int_0^t \int_{\mathbb{T}} f(s,y) \xi(s,y) dy ds,$$

where $f(s,y) = \mathbf{E}[\mathcal{D}_{s,y} X_t | \mathcal{F}_s]$ with $\mathcal{D}_{s,y}$ denoting the Malliavin derivative operator, and $\mathcal{F}_s$ is the natural filtration generated by ξ. The calculation of $\mathcal{D}_{s,y} X_t$ is rather technical,

but in the end we obtain $f(s, \cdot)$ as a functional of $\big(u(s, \cdot), \nabla \phi(s, \cdot)\big)$ (which is a surprise to us). Thus, to derive the formula for the variance of X_t, one needs to understand the joint distribution of $u(s, \cdot)$ and $\nabla \phi(s, 0)$.

Define $g = 1 + \nabla \phi$, we know from (1.10) that g satisfies the Fokker-Planck equation

$$(1.14) \qquad \partial_t g = \frac{1}{2} \Delta g + \nabla(ug).$$

In this way, the problem reduces to understanding the joint distribution of $u(s, \cdot)$ and $g(s, \cdot)$. One should note that, for general drifts, this is not possible. Here is another key point in the argument: g is the density of another diffusion in random environment:

$$(1.15) \qquad d\mathcal{Y}_s = -u(s, \mathcal{Y}_s)ds + dB_s.$$

Compare the two SDEs satisfied by $\mathcal{X}$ and $\mathcal{Y}$, we note that the random drifts are time reversal of each other, up to the opposite sign. Roughly speaking, since $\mathcal{Y}$ follows approximately the characteristics of the Burgers flow, one may suspect that it corresponds to the so-called second class particle in the context of exclusion processes. For the drift given by (1.9), there is a time reversal anti-symmetry

$$(1.16) \qquad \{u(t-s, y)\}_{s \in [0,t], y \in \mathbb{T}} \stackrel{\text{law}}{=} \{-u(s, y)\}_{s \in [0,t], y \in \mathbb{T}}$$

that holds for any $t > 0$. Using the above equality in law, one can re-connect $\mathcal{Y}$ to $\mathcal{X}$ and derive the joint distribution of $u(s, \cdot)$ and $g(s, \cdot)$, through the Gibbs formulation of the directed polymers. This, together with the Clark-Ocone representation (1.13), eventually helps us to compute the asymptotic variance of $X_t/\sqrt{t}$, which corresponds to the second term on the r.h.s. of (1.5).

Overall, the fact that $\sigma^2(\beta)$ can be written explicitly in terms of the underlying invariant measure comes as a surprise to us, and the situation here seems to be more subtle than the study of the free energy $\log \bar{Z}_t$. For a noise that is white in time but smooth in space, we have obtained a formula for the variance of $\log \bar{Z}_t$, similar to the one for $\Sigma^2(\beta)$ given in (1.2), see [28, Eq. (5.6)]. For the variance of the polymer endpoint, our argument relies crucially on the time reversal anti-symmetry in (1.16), which does not hold for a colored noise, and this is actually the only place in the paper where the noise is required to be white. It is also unclear to us whether there exists a similar result for a colored noise. On a heuristic level, the time reversal anti-symmetry allows us to go back and forth between the Gibbsian and the Markovian settings, connecting the polymer measure, which is the diffusion with a non-Markovian drift described by (1.8), to the "second class particle" diffusion with a Markovian drift, given in (1.15). The zero-temperature version of the polymer model is the last passage percolation, and one may want to draw the connection between the two diffusions and the duality between the geodesics and the competing interfaces [48]. In the end, the Gibbs measure formulation allows us to express the joint distribution of the endpoint density "g" and the drift "u" explicitly.

1.5. **Organization of the paper.** From the above discussion, we immediately identify a few difficulties in implementing the argument. (i) To study the quenched mean $\phi(t, 0)$, it is actually more convenient to use the Gibbs measure formulation, where one can write the quenched mean as $\int_{\mathbb{R}} x\rho(t, x)dx$, with the endpoint density ρ defined in (1.3). To apply the Clark-Ocone formula, we need to compute $\mathcal{D}_{s,y}\rho(t, x)$, which involves cancellations and is considerably more complicated than computing the Malliavin derivative of the free energy. This is the main task of Section 2. (ii) To use the time reversal anti-symmetry, we need to initiate the Burgers equation (1.9) at stationarity, but this is not the case for the point-to-line directed polymer. Therefore, in Section 3, we will derive an error estimate which will help us to approximate the free boundary condition by the stationary boundary condition. (iii) The SDEs (1.8) and (1.15) are only symbolic since the drift u

is distribution-valued, and this is where the recently developed theories for singular diffusions play a role. In Section 4, we will study the two related passive scalar problems and make use of the time reversal anti-symmetry to derive the joint distribution of $u(s, \cdot)$ and $\nabla\phi(s, \cdot)$. (iv) In Section 5, we compute the variance of $\phi(t, 0)$, using the fast mixing of $u(s, \cdot)$, or equivalently, the polymer endpoint density on the torus, and complete the proof of Theorem 1.1.

1.6. **Notations.** (i) We write $q_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$ as the standard heat kernel on $\mathbb{R}$, and $G_t(x) = \sum_{n\in\mathbb{Z}} q_t(x+n)$ is the heat kernel on $\mathbb{T}$.

(ii) We use W_i, with the index i, to represent independent Brownian bridges connecting $(0, 0)$ and $(1, 0)$, periodically extended to $\mathbb{R}$. We will use B to represent standard Brownian motion on $\mathbb{R}$.

(iii) Since there are different sources of randomness, we use $\mathbf{E}$ to denote the expectation on ξ, and $\mathbb{E}_{W_i}, \mathbb{E}_B$ to denote the expectations on W_i, B respectively.

2. QUENCHED MEAN AND CLARK-OCONE REPRESENTATION

The goal of this section is to present the first key step in the proof of Theorem 1.1. The total variance of the polymer endpoint is decomposed into two parts: the variance of the quenched mean and the mean of the quenched variance. It turns out that for our model with the noise satisfying the shear-invariance property, the mean of the quenched variance is easy to deal with, see (2.5) below, which is sometimes regarded as a physics folklore, see e.g. [12, Eq. (42)]. The key difficulty is then to study the fluctuations of the quenched mean. Our idea is similar to that of [28], where tools from Malliavin calculus were borrowed, and the quenched mean will be written as an explicit Itô integral through the Clark-Ocone formula, see (2.18) below.

2.1. **Variance decomposition and quenched mean.** It was known from [33, Eq. (3.21), Proposition 5.1] that the effective diffusivity in (1.4) can be obtained through

$$\sigma^2(\beta) = \lim_{t\to\infty} \frac{1}{t} V_t, \tag{2.1}$$

where

$$V_t := \mathbf{E} \int_{\mathbb{R}} x^2 \rho(t, x) dx \tag{2.2}$$

is the annealed variance. Let

$$X_t := \int_{\mathbb{R}} x\rho(t, x)dx, \qquad \mathcal{V}_t := \int_{\mathbb{R}} x^2 \rho(t, x)dx - X_t^2 \tag{2.3}$$

be the quenched mean and the quenched variance of the polymer endpoint, respectively. Then, we can obviously write

$$V_t = \mathbf{E}X_t^2 + \mathbf{E}\mathcal{V}_t. \tag{2.4}$$

It is folklore, e.g. see a proof in [33, Lemma 5.3], that the average of the quenched variance is t, i.e.,

$$\mathbf{E}\mathcal{V}_t = t. \tag{2.5}$$

Thus, to derive an expression for $\sigma^2(\beta)$, it suffices to study the first term on the r.h.s. of (2.4), which is actually the variance of the quenched mean.

Let us introduce some notations. Let $\mathcal{Z}_{t,s}(x,y)$ be the propagator of the SHE and $\mathcal{G}_{t,s}(x,y)$ be its periodic counterpart. Namely, for any fixed (s,y), $\mathcal{Z}_{t,s}(x,y)$ solves

(2.6)
$$\partial_t \mathcal{Z}_{t,s}(x,y) = \frac{1}{2}\Delta_x \mathcal{Z}_{t,s}(x,y) + \beta \mathcal{Z}_{t,s}(x,y)\xi(t,x), \qquad t > s, x \in \mathbb{R},$$
$$\mathcal{Z}_{s,s}(x,y) = \delta(x-y),$$

and

$$\mathcal{G}_{t,s}(x,y) = \sum_{n \in \mathbb{Z}} \mathcal{Z}_{t,s}(x, y+n) = \sum_{n \in \mathbb{Z}} \mathcal{Z}_{t,s}(x+n, y).$$

So we know that the solution to (1.1) is actually $Z(t,x) = \mathcal{Z}_{t,0}(x,0)$.

It is worth emphasizing that one should view $\mathcal{Z}$ as the propagator of SHE on $\mathbb{R}_+ \times \mathbb{R}$ while $\mathcal{G}$ as the propagator of the same equation on $\mathbb{R}_+ \times \mathbb{T}$. This leads to the fact that, if we view the polymer path as lying on the half plane $\mathbb{R}_+ \times \mathbb{R}$, then the quenched density should be expressed in terms of $\mathcal{Z}$; while if we view the polymer path as lying on the half cylinder $\mathbb{R}_+ \times \mathbb{T}$, the quenched density should be expressed in terms of $\mathcal{G}$.

Next, we introduce the following notations on the endpoint densities of the "forward" and "backward" polymer path on a cylinder. Let $\mathcal{M}_1(\mathbb{T})$ be the set of all Borel probability measures on $\mathbb{T}$. For any $\nu \in \mathcal{M}_1(\mathbb{T})$ and $t > s$, we let

(2.7)
$$\rho_{\mathrm{f}}(t,x;s,\nu) = \frac{\int_{\mathbb{T}} \mathcal{G}_{t,s}(x,y)\nu(dy)}{\int_{\mathbb{T}^2} \mathcal{G}_{t,s}(x',y)\nu(dy)dx'}, \quad x \in \mathbb{T},$$
$$\rho_{\mathrm{b}}(t,\nu;s,y) = \frac{\int_{\mathbb{T}} \mathcal{G}_{t,s}(x,y)\nu(dx)}{\int_{\mathbb{T}^2} \mathcal{G}_{t,s}(x,y')\nu(dx)dy'}, \quad y \in \mathbb{T}.$$

Here the subscripts "f, b" represent "forward" and "backward" respectively. When ν is a Dirac measure on $\mathbb{T}$, we write

$$\rho_{\mathrm{f}}(t,x;s,y) = \rho_{\mathrm{f}}(t,x;s,\delta_y) \quad \text{and} \quad \rho_{\mathrm{b}}(t,x;s,y) = \rho_{\mathrm{b}}(t,\delta_x;s,y).$$

We also need the simplified notation

(2.8)
$$\mathcal{G}_{t,s}(-,y) = \int_{\mathbb{T}} \mathcal{G}_{t,s}(x,y)dx = \int_{\mathbb{R}} \mathcal{Z}_{t,s}(x,y)dx,$$

which can be viewed as the point-to-line partition function with the starting point (s,y).

For any $\theta \in \mathbb{R}$, define

(2.9)
$$\mathcal{E}_\theta(t) = \int_{\mathbb{R}} e^{\theta x} \mathcal{Z}_{t,0}(x,0)dx,$$

so one can write

(2.10)
$$X_t = \partial_\theta \log \mathcal{E}_\theta(t)\Big|_{\theta=0}.$$

It is clear that X_t is a square integrable random variable that is measurable with respect to $\mathcal{F}_t$ - the natural filtration corresponding to the noise $\{\xi(s,y)\}_{(s,y) \in \mathbb{R} \times \mathbb{T}}$. The idea is to apply the Clark-Ocone formula to write X_t as an Itô integral with respect to ξ, through which we compute the second moment.

2.2. **Clark-Ocone representation.** We will first write $\log \mathcal{E}_\theta(t) - \mathbf{E}\log \mathcal{E}_\theta(t)$ as an Itô integral, then take derivative on θ to obtain the representation for X_t.

Recall that we view ξ as the spacetime white noise on $\mathbb{R} \times \mathbb{T}$. Let $\mathcal{D}_{s,y}$ denote the Malliavin derivative [45, Chapter 1]. Namely, for "smooth" random variables X, $(\mathcal{D}_{s,y}X)_{(s,y) \in \mathbb{R} \times \mathbb{T}}$ is an $L^2(\mathbb{R} \times \mathbb{T})$–valued random variable.

By the Clark-Ocone formula, see e.g. [20, Proposition 6.3], we have

(2.11)
$$\log \mathcal{E}_\theta(t) - \mathbf{E}\log \mathcal{E}_\theta(t) = \int_0^t \int_{\mathbb{T}} \mathbf{E}[\mathcal{D}_{s,y}\log \mathcal{E}_\theta(t)|\mathcal{F}_s]\xi(s,y)dyds.$$

First, let us consider the mean. Fix $t > 0$. Since

$$(2.12) \qquad \{\mathcal{Z}_{t,0}(x,0)\}_{x \in \mathbb{R}} \overset{\text{law}}{=} \{\mathcal{Z}_{t,0}(0,x)\}_{x \in \mathbb{R}},$$

we have

$$\mathbf{E}\log\mathcal{E}_\theta(t) = \mathbf{E}\log\left(\int_{\mathbb{R}} e^{\theta x}\mathcal{Z}_{t,0}(x,0)dx\right) = \mathbf{E}\log\left(\int_{\mathbb{R}}\mathcal{Z}_{t,0}(0,x)e^{\theta x}dx\right)$$

$$= \mathbf{E}\log\left(\int_{\mathbb{R}}\mathcal{Z}_{t,0}(0,x)dx\right) + \frac{\theta^2 t}{2}.$$

For the last "=", see the proof of Lemma 5.3 in [33]. Thus, we conclude that

$$(2.13) \qquad \mathbf{E}\log\mathcal{E}_\theta(t) = \mathbf{E}\log\mathcal{E}_0(t) + \frac{1}{2}\theta^2 t.$$

Here is the lemma on the Malliavin derivative:

Lemma 2.1. *For any $s < t$ and $y \in \mathbb{T}$, we have*

$$(2.14) \qquad \mathcal{D}_{s,y}\mathcal{E}_\theta(t) = \beta \sum_{n \in \mathbb{Z}} \mathcal{Z}_{s,0}(y+n,0)\int_{\mathbb{R}} e^{\theta x}\mathcal{Z}_{t,s}(x,y+n)dx.$$

The proof of the lemma is standard and follows from an approximation argument and application of the Feynman-Kac formula (see [11]), so we only sketch it here. To consider some approximation, let us introduce a few notations that will be used throughout the paper. For a given $\varepsilon > 0$, let ξ^ε be a spatial mollification of ξ:

$$\xi^\varepsilon(t,x) = \int_{\mathbb{R}} \eta^\varepsilon(x-y)\xi(t,y)dy, \qquad t,x \in \mathbb{R}$$

where η^ε is a symmetric approximation of the Dirac function, as $\varepsilon \to 0$, that is supported in $[-1/4, 1/4]$. Let R^ε be the spatial covariance function of ξ^ε, which is the periodic extension of $\eta^\varepsilon \star \eta^\varepsilon$:

$$R^\varepsilon(x) = \sum_{n \in \mathbb{Z}} \eta^\varepsilon \star \eta^\varepsilon(x+n),$$

and let $\mathcal{Z}^\varepsilon$ be the propagator of the SHE with the mollified noise ξ^ε, defined in the same way as (2.6). Finally, let $\mathcal{E}_\theta^\varepsilon(t) = \int_{\mathbb{R}} e^{\theta x}\mathcal{Z}_{t,0}^\varepsilon(x,0)dx$.

Proof of Lemma 2.1. By the Feynman-Kac formula, see [11, Theorem 2.2], we can write

$$\mathcal{E}_\theta^\varepsilon(t) = \mathbb{E}_B\left[e^{\beta\int_0^t \xi^\varepsilon(\ell,B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t}e^{\theta B_t}\right],$$

where B is a standard Brownian motion starting from the origin and the expectation $\mathbb{E}_B$ is taken only with respect to B. Since $\mathcal{E}_\theta^\varepsilon(t) \to \mathcal{E}_\theta(t)$ in $L^2(\Omega)$, it suffices to show that the convergence of $\beta^{-1}\mathcal{D}_{s,y}\mathcal{E}_\theta^\varepsilon(t)$ to the r.h.s. of (2.14) holds in $L^2(\Omega, L^2([0,t]\times\mathbb{T}))$.

By the definition of ξ^ε, we have

$$\int_0^t \xi^\varepsilon(\ell,B_\ell)d\ell = \int_0^t\int_{\mathbb{T}}\sum_{n \in \mathbb{Z}}\eta^\varepsilon(B_\ell - y - n)\xi(\ell,y)dyd\ell,$$

so for each fixed realization B, it holds that

$$\mathcal{D}_{s,y}\left(\int_0^t \xi^\varepsilon(\ell,B_\ell)d\ell\right) = \sum_{n \in \mathbb{Z}}\eta^\varepsilon(B_s - y - n).$$

This leads to

$$\beta^{-1}\mathcal{D}_{s,y}\mathcal{E}_\theta^\varepsilon(t) = \sum_{n \in \mathbb{Z}}\mathbb{E}_B\left[e^{\beta\int_0^t\xi^\varepsilon(\ell,B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t}\eta^\varepsilon(B_s - y - n)e^{\theta B_t}\right]$$

$$= \sum_{n \in \mathbb{Z}}\int_{\mathbb{R}} q_s(w)\eta^\varepsilon(w-y-n)\mathbb{E}_B[e^{\beta\int_0^t\xi^\varepsilon(\ell,B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t}e^{\theta B_t}|B_s = w]dw$$

$$= \sum_{n \in \mathbb{Z}}\int_{\mathbb{R}}\eta^\varepsilon(w-y-n)\left(\int_{\mathbb{R}} e^{\theta x}\mathcal{Z}_{t,s}^\varepsilon(x,w)dx\right)\mathcal{Z}_{s,0}^\varepsilon(w,0)dw.$$

Sending $\varepsilon \to 0$, we complete the proof. $\square$

Now we can state the following proposition:

Proposition 2.2. *For any $0 \le s < t$ and $y \in \mathbb{T}$, we have*

$$\mathbf{E}[\mathcal{D}_{s,y} \log \mathcal{E}_\theta(t)|\mathcal{F}_s]$$

$$= \beta \mathbf{E}\left[\frac{\mathcal{G}_{t,s}(-,y) \sum_{n \in \mathbb{Z}} e^{\theta(y+n)} \mathcal{Z}_{s,0}(y+n,0)}{\int_{\mathbb{R}} \mathcal{G}_{t,s}(-,y') e^{\theta y'} \mathcal{Z}_{s,0}(y',0) dy'}\bigg|\mathcal{F}_s\right].$$

Proof. We proceed by approximation. Following the proof of Lemma 2.1, we have

$$\frac{1}{\beta} \frac{\mathcal{D}_{s,y} \mathcal{E}_\theta^\varepsilon(t)}{\mathcal{E}_\theta^\varepsilon(t)} = \frac{\sum_{n \in \mathbb{Z}} \mathbb{E}_B\left[e^{\beta \int_0^t \xi^\varepsilon(\ell, B_\ell) d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t} \eta^\varepsilon(B_s - y - n) e^{\theta B_t}\right]}{\mathbb{E}_B\left[e^{\beta \int_0^t \xi^\varepsilon(\ell, B_\ell) d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t} e^{\theta B_t}\right]}.$$

The main challenge is hence to compute the conditional expectation given $\mathcal{F}_s$. To continue, first, for both the numerator and the denominator, we write

$$e^{\theta B_t} = e^{\theta B_s} e^{\frac{1}{2}\theta^2(t-s)} \cdot e^{\theta(B_t - B_s) - \frac{1}{2}\theta^2(t-s)},$$

and view the last exponential factor on the right as a change of measure. Applying the Girsanov theorem to the Brownian motion in $[s,t]$, we have

$$(2.15) \qquad \frac{1}{\beta} \frac{\mathcal{D}_{s,y} \mathcal{E}_\theta^\varepsilon(t)}{\mathcal{E}_\theta^\varepsilon(t)} = \frac{\sum_{n \in \mathbb{Z}} \mathbb{E}_B\left[e^{\beta \int_0^t \xi^\varepsilon(\ell, \tilde{B}_\ell^\theta) d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t} \eta^\varepsilon(B_s - y - n) e^{\theta B_s}\right]}{\mathbb{E}_B\left[e^{\beta \int_0^t \xi^\varepsilon(\ell, \tilde{B}_\ell^\theta) d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t} e^{\theta B_s}\right]},$$

where $\{\tilde{B}_\ell^\theta\}_\ell$ is obtained from $\{B_\ell\}_\ell$ by adding the drift θ in the interval $[s,t]$:

$$\tilde{B}_\ell^\theta = B_\ell 1_{\ell \in [0,s]} + [B_\ell + \theta(\ell - s)] 1_{\ell \in (s,t]}.$$

In this way, we can rewrite

$$(2.16) \qquad \int_0^t \xi^\varepsilon(\ell, \tilde{B}_\ell^\theta) d\ell = \int_0^s \xi^\varepsilon(\ell, B_\ell) d\ell + \int_0^{t-s} \xi^\varepsilon(s + \ell, B_{s+\ell} + \theta\ell) d\ell.$$

Here is the key of the argument: to compute $\mathbf{E}[\frac{\mathcal{D}_{s,y} \mathcal{E}_\theta^\varepsilon(t)}{\mathcal{E}_\theta^\varepsilon(t)}|\mathcal{F}_s]$, we need to average out the noise in the domain $[s,t] \times \mathbb{R}$, but by (2.15) and (2.16), we see that the dependence on the noise beyond s is only through the factor $\int_0^{t-s} \xi^\varepsilon(s + \ell, B_{s+\ell} + \theta\ell) d\ell$. Since ξ^ε is white in time and stationary in space, we have, for fixed s,

$$\{\xi^\varepsilon(s + \ell, x + \theta\ell)\}_{\ell \ge 0, x \in \mathbb{R}} \overset{\text{law}}{=} \{\xi^\varepsilon(s + \ell, x)\}_{\ell \ge 0, x \in \mathbb{R}}.$$

In other words, when computing the conditional expectation, one can replace $\int_0^t \xi^\varepsilon(\ell, \tilde{B}_\ell^\theta) d\ell$ by

$$\int_0^s \xi^\varepsilon(\ell, B_\ell) d\ell + \int_0^{t-s} \xi^\varepsilon(s + \ell, B_{s+\ell}) d\ell = \int_0^t \xi^\varepsilon(\ell, B_\ell) d\ell.$$

As a result

$$(2.17) \qquad \begin{aligned} &\frac{1}{\beta} \mathbf{E}\left[\frac{\mathcal{D}_{s,y} \mathcal{E}_\theta^\varepsilon(t)}{\mathcal{E}_\theta^\varepsilon(t)}\bigg|\mathcal{F}_s\right] \\ &= \mathbf{E}\left[\frac{\sum_{n \in \mathbb{Z}} \mathbb{E}_B\left[e^{\beta \int_0^t \xi^\varepsilon(\ell, B_\ell) d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t} \eta^\varepsilon(B_s - y - n) e^{\theta B_s}\right]}{\mathbb{E}_B\left[e^{\beta \int_0^t \xi^\varepsilon(\ell, B_\ell) d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t} e^{\theta B_s}\right]}\bigg|\mathcal{F}_s\right]. \end{aligned}$$

Now we can pass to the limit, as $\varepsilon \to 0$, and arguing as in the proof of Lemma 2.1, we conclude that the right hand side of (2.17) tends to

$$\mathbf{E}\left[\frac{\sum_{n \in \mathbb{Z}} \mathcal{G}_{t,s}(-,y+n) \cdot e^{\theta(y+n)} \mathcal{Z}_{s,0}(y+n,0)}{\int_{\mathbb{R}} \mathcal{G}_{t,s}(-,y') e^{\theta y'} \mathcal{Z}_{s,0}(y',0) dy'}\bigg|\mathcal{F}_s\right],$$

where we recall that $\mathcal{G}_{t,s}(-,y) = \int_{\mathbb{R}} \mathcal{Z}_{t,s}(x,y) dx$. It is easy to see that $y \mapsto \mathcal{G}_{t,s}(-,y)$ is 1-periodic in y, so the proof is complete. $\square$

By (2.10), computing derivative on θ on both sides of (2.11) and using (2.13), we obtain

$$X_t = \frac{d}{d\theta}_{|\theta=0} \log \mathcal{E}_\theta(t) = \frac{d}{d\theta}_{|\theta=0} \int_0^t \int_\mathbb{T} \mathbf{E}[\mathcal{D}_{s,y} \log \mathcal{E}_\theta(t)|\mathcal{F}_s]\xi(s,y)dyds.$$

Applying Proposition 2.2, we further derive

$$(2.18) \qquad X_t = \beta \int_0^t \int_\mathbb{T} \mathbf{E}[I_{1,t}(s,y) - I_{2,t}(s,y)|\mathcal{F}_s]\xi(s,y)dyds,$$

with

$$(2.19) \qquad I_{1,t}(s,y) = \frac{\mathcal{G}_{t,s}(-,y)\sum_{n\in\mathbb{Z}}(y+n)\mathcal{Z}_{s,0}(y+n,0)}{\mathcal{G}_{t,0}(-,0)},$$

and

$$(2.20) \qquad I_{2,t}(s,y) = \rho_\mathrm{m}(t,-|s,y)\frac{\int_\mathbb{R} \mathcal{G}_{t,s}(-,y')y'\mathcal{Z}_{s,0}(y',0)dy'}{\mathcal{G}_{t,0}(-,0)},$$

where

$$(2.21) \qquad \rho_\mathrm{m}(t,-|s,y) = \frac{\mathcal{G}_{t,s}(-,y)\mathcal{G}_{s,0}(y,0)}{\mathcal{G}_{t,0}(-,0)},$$

is the midpoint density of the polymer, i.e. it is the quenched density at time s of the point-to-line polymer, starting at $(0,0)$. To simplify the notation, we omitted the dependence on the starting point in the expression of ρ_m. In particular, we have $\int_\mathbb{T} \rho_\mathrm{m}(t,-|s,y)dy = 1$. Here the subscript "m" represents "middle".

For any $s > 0, y \in \mathbb{T}$, define

$$(2.22) \qquad \psi(s,y) = -\frac{\sum_{n\in\mathbb{Z}}(y+n)\mathcal{Z}_{s,0}(y+n,0)}{\mathcal{G}_{s,0}(y,0)} = \frac{\sum_{n\in\mathbb{Z}}(n-y)\mathcal{Z}_{s,0}(y,n)}{\mathcal{G}_{s,0}(y,0)}.$$

Here in the second "=", we changed $n \to -n$ and used the fact that $\mathcal{Z}_{s,0}(y-n,0) = \mathcal{Z}_{s,0}(y,n)$. One can interpret $\psi(s,y)$ as the "winding number" of the point-to-point polymer path on the cylinder, starting at (s,y) and ending at $(0,0)$, see more discussion in Section 3. The last equality follows from the fact that in the case of a 1-periodic noise, we have

$$\mathcal{Z}_{s,0}(y-n,0) = \mathcal{Z}_{s,0}(y,n), \quad n \in \mathbb{Z}.$$

We have the following lemma, which expresses $I_{1,t} - I_{2,t}$ in terms of ρ_m and ψ and will serve as the starting point of the rest of the analysis.

Lemma 2.3. *We have*

$$(2.23) \quad I_{1,t}(s,y) - I_{2,t}(s,y) = \rho_\mathrm{m}(t,-|s,y)\int_\mathbb{T}[\psi(s,y')-\psi(s,y)]\rho_\mathrm{m}(t,-|s,y')dy'.$$

Proof. First, using the definition (2.21), we can rewrite (2.19) in the form

$$I_{1,t}(s,y) = -\rho_\mathrm{m}(t,-|s,y)\psi(s,y).$$

Then, we rewrite the numerator of the second factor on the right hand side of (2.20) as

$$\int_\mathbb{R} \mathcal{G}_{t,s}(-,y')y'\mathcal{Z}_{s,0}(y',0)dy'$$

$$= \int_\mathbb{T} \mathcal{G}_{t,s}(-,y')\mathcal{G}_{s,0}(y',0)\frac{\sum_{n\in\mathbb{Z}}(y'+n)\mathcal{Z}_{s,0}(y'+n,0)}{\mathcal{G}_{s,0}(y',0)}dy',$$

where we used the fact that $\mathcal{G}_{t,s}(-,y)$ is 1−periodic in y. Thus, invoking the definitions (2.21) and (2.22), we can rewrite (2.20) in the form

$$I_{2,t}(s,y) = \rho_\mathrm{m}(t,-|s,y)\int_\mathbb{T} \rho_\mathrm{m}(t,-|s,y')\frac{\sum_{n\in\mathbb{Z}}(y'+n)\mathcal{Z}_{s,0}(y'+n,0)}{\mathcal{G}_{s,0}(y',0)}dy'$$

$$= -\rho_\mathrm{m}(t,-|s,y)\int_\mathbb{T} \rho_\mathrm{m}(t,-|s,y')\psi(s,y')dy'.$$

The proof is complete. □

To summarize, the main results of this section are (2.18) and (2.23). They allow to express the quenched mean $X_t = \int_{\mathbb{R}} x \rho(t, x) dx$ as a more or less explicit Itô integral. The integrand involves the quenched midpoint density $\rho_{\mathrm{m}}(t, -|s, y)$ (see (2.23)), which we understand well, and the difference between "winding numbers" of the polymer paths starting at (s, y') and (s, y), i.e. the factor $\psi(s, y') - \psi(s, y)$, which we do not understand so well so far and presents itself the main challenge. In a sense, to derive an explicit variance formula for X_t, one needs to understand the joint distribution of $\rho_{\mathrm{m}}(t, -|s, y)$ and $\psi(s, y') - \psi(s, y)$. This will be the content of the next two sections.

3. Winding number and stabilization

In this section, let us look more closely at the function $\psi(s, y)$ defined in (2.22). Recall that a key quantity in the Clark-Ocone representation (2.18) is $\psi(s, y') - \psi(s, y)$. The goal of this section is to reframe the problem so that we are in a stationary setting, and the main result, Proposition 3.2 below, proves a stationary approximation of $\psi(s, y') - \psi(s, y)$.

To better understand it, we define an ε-approximation of $\psi(s, y)$. Recall that $\mathcal{Z}^\varepsilon$ is the propagator of SHE with noise ξ^ε. Let also $\mathcal{G}^\varepsilon$ be the ε-counterpart of $\mathcal{G}$, and define

$$\psi^\varepsilon(s, y) = \frac{\sum_{n \in \mathbb{Z}}(n - y)\mathcal{Z}_{s,0}^\varepsilon(y, n)}{\mathcal{G}_{s,0}^\varepsilon(y, 0)}.$$

By the Feynman-Kac formula, we can rewrite ψ^ε and interpret its integer part (floor, or ceiling depending on whether it is positive or negative respectively) as the winding number of a directed polymer on a cylinder:

Lemma 3.1. *For any $s > 0$ and $y \in \mathbb{T}$, we have*

$$(3.1) \qquad \psi^\varepsilon(s, y) = \frac{\mathbb{E}_B\left[e^{\beta \int_0^s \xi^\varepsilon(s-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)s}(B_s - y) \mid B_0 = y, \dot{B}_s = 0\right]}{\mathbb{E}_B\left[e^{\beta \int_0^s \xi^\varepsilon(s-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)s} \mid B_0 = y, \dot{B}_s = 0\right]},$$

where $\dot{B}$ denotes the fractional part of B (in other words, $\dot{B}$ is a Brownian motion on $\mathbb{T}$).

Proof. To verify the above formula, we rewrite the numerator of the expression on the right hand side of (3.1):

$$\mathbb{E}_B\left[e^{\beta \int_0^s \xi^\varepsilon(s-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)s}(B_s - y) \mid B_0 = y, \dot{B}_s = 0\right]$$

$$= \frac{\sum_{n \in \mathbb{Z}} \mathbb{E}_B\left[e^{\beta \int_0^s \xi^\varepsilon(s-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)s}(B_s - y)\delta(B_s - n) \mid B_0 = y\right]}{\sum_{n \in \mathbb{Z}} q_s(n - y)}$$

$$= \frac{\sum_{n \in \mathbb{Z}}(n - y)\mathcal{Z}_{s,0}^\varepsilon(y, n)}{G_s(y)}.$$

The denominator can be rewritten in the same way. □

The above lemma interprets $\psi^\varepsilon(s, y)$ as the winding number, up to a fractional part, of the polymer path around the cylinder, with the starting point (s, y) and the ending point $(0, 0)$. Due to the fast mixing of the polymer endpoint distribution on the torus, the prescribed distribution of the ending point plays no role in the long time, and it is actually more convenient to consider the ending point at stationarity. For this purpose, we define the key object

$$(3.2) \qquad \phi(s, y) = \frac{\int_{\mathbb{R}} \mathcal{Z}_{s,0}(y, y')(y' - y)e^{\beta W(y')}dy'}{\int_{\mathbb{R}} \mathcal{Z}_{s,0}(y, y')e^{\beta W(y')}dy'},$$

where W is a Brownian bridge on $\mathbb{T}$, periodically extended to $\mathbb{R}$. One should view ϕ as a "stationary" approximation of ψ, and the main result of the section is the following estimate, which will help us to eventually replace the ψ factors in (2.23) by the corresponding ϕ factors.

Proposition 3.2. *For any $p \in [1, \infty)$, there exists $C, \lambda > 0$ depending on p so that*

$$
(3.3) \qquad \sup_{y, y' \in \mathbb{T}} \mathbb{E}_W \mathbf{E} |[\psi(s, y') - \psi(s, y)] - [\phi(s, y') - \phi(s, y)]|^p \leq C e^{-\lambda s}, \quad s \geq 0.
$$

Before commenting on the proof of Proposition 3.2, let us introduce some notation: for two probability measures μ, ν on $\mathbb{T}$, we periodically extend ν to $\mathbb{R}$, and define

$$
(3.4) \qquad
\begin{aligned}
\mathcal{W}(s; \mu, \nu) &= \frac{\int_{\mathbb{T}} \left(\int_{\mathbb{R}} \mathcal{Z}_{s,0}(y, y')(y' - y) \nu(dy') \right) \mu(dy)}{\int_{\mathbb{T}} \left(\int_{\mathbb{R}} \mathcal{Z}_{s,0}(y, y') \nu(dy') \right) \mu(dy)} \\
&= \frac{\int_{\mathbb{T}} \left(\int_{\mathbb{R}} \mathcal{Z}_{s,0}(y, y')(y' - y) \nu(dy') \right) \mu(dy)}{\int_{\mathbb{T}} \left(\int_{\mathbb{T}} \mathcal{G}_{s,0}(y, y') \nu(dy') \right) \mu(dy)}.
\end{aligned}
$$

With some abuse of terminology ($\mathcal{W}$ needs not be an integer), we shall refer to $\mathcal{W}(s; \mu, \nu)$ as the winding number of the polymer path with initial distribution μ (at time s) and ending distribution ν (at time 0).

Using this notation we can write

$$
(3.5) \qquad \psi(s, y) = \mathcal{W}(s; \delta_y, \delta_0), \qquad \phi(s, y) = \mathcal{W}(s; \delta_y, \rho_W),
$$

where (with some abuse of notation) we have used ρ_W to denote the (random) probability measure on $\mathbb{T}$ with density

$$
\rho_W(x) = \frac{e^{\beta W(x)}}{\int_{\mathbb{T}} e^{\beta W(x')} dx'}.
$$

To see on a heuristic level why Proposition 3.2 holds, we note that, for the same ending point distribution ν but different initial distributions μ_1, μ_2, the quantity $\mathcal{W}(s, \mu_1, \nu) - \mathcal{W}(s, \mu_2, \nu)$ is of order $O(1)$ for $s \gg 1$, see Corollary 3.9 below, and it mostly depends on the random environment $\xi(\ell, \cdot)$ with $s - \ell \sim O(1)$, by the fast mixing of the polymer endpoint density on the torus, which will become more clear later. Therefore, for $s \gg 1$, its dependence on ν is exponentially small in s, which explains (3.3).

In light of (3.5), Proposition 3.2 is a conclusion of the following more general result: for any $p \in [1, \infty)$, there exists $C, \lambda > 0$ depending on p so that

$$
(3.6) \qquad
\begin{aligned}
\sup_{\mu_i, \nu_i, i=1,2} &\mathbf{E} \big| [\mathcal{W}(s; \mu_1, \nu_1) - \mathcal{W}(s, \mu_2, \nu_1)] - [\mathcal{W}(s; \mu_1, \nu_2) - \mathcal{W}(s; \mu_2, \nu_2)] \big|^p \\
&\leq C e^{-\lambda s}, \quad s \geq 0,
\end{aligned}
$$

where the supremum is taken with respect to all $\mu_1, \mu_2, \nu_1, \nu_2 \in \mathcal{M}_1(\mathbb{T})$.

The proof of (3.6) is given in Section 3.3 below. Roughly speaking, we will write the winding number $\mathcal{W}(s; \mu, \nu)$ as the running sum of a Markov chain, with the chain representing the winding number accumulated by the polymer path during each time interval of length 1. The fast mixing of the chain drives the exponential decay in (3.6). Before presenting the argument, we need some auxiliary results contained in the following sections.

3.1. **Rewriting the winding number in terms of a Markov chain.** Recall that $\mathcal{W}(s; \mu, \nu)$ was defined in (3.4). For notational convenience in this section, it is more convenient to view μ as the ending distribution and ν as the starting distribution. We extend μ periodically to $\mathbb{R}$. The following lemma holds:

Lemma 3.3. *We have*

$$(3.7) \qquad \mathcal{W}(s; \mu, \nu) = -\frac{\int_{\mathbb{R}} \left(\int_{\mathbb{T}} \mathcal{Z}_{s,0}(y, y')(y - y')\nu(dy') \right) \mu(dy)}{\int_{\mathbb{R}} \left(\int_{\mathbb{T}} \mathcal{Z}_{s,0}(y, y')\nu(dy') \right) \mu(dy)}.$$

Proof. To see why (3.7) holds, consider the numerator of $\mathcal{W}(s; \mu, \nu)$ in (3.4) first. By the periodicity of ν and $\mathcal{Z}$, we can write

$$\int_{\mathbb{T}} \left(\int_{\mathbb{R}} \mathcal{Z}_{s,0}(y, y')(y' - y)\nu(dy') \right) \mu(dy)$$

$$= \sum_{n \in \mathbb{Z}} \int_{\mathbb{T}^2} \mathcal{Z}_{s,0}(y, y' + n)(y' + n - y)\nu(dy')\mu(dy)$$

$$= \sum_{n \in \mathbb{Z}} \int_{\mathbb{T}^2} \mathcal{Z}_{s,0}(y - n, y')(y' + n - y)\nu(dy')\mu(dy)$$

$$= -\sum_{n \in \mathbb{Z}} \int_{\mathbb{T}^2} \mathcal{Z}_{s,0}(y + n, y')(y + n - y')\nu(dy')\mu(dy)$$

$$= -\int_{\mathbb{R}} \left(\int_{\mathbb{T}} \mathcal{Z}_{s,0}(y, y')(y - y')\nu(dy') \right) \mu(dy),$$

since μ is also periodically extended. The denominator is treated in the same way, which completes the proof of (3.7). □

For given $\mu, \nu \in \mathcal{M}_1(\mathbb{T})$ and each realization of the noise, we introduce a Markov chain so that the winding number of the polymer path is written as the running sum of the chain. This has also been used in [33] to prove the central limit theorem of the winding number, and we repeat it here for the convenience of readers.

Fix any $s > 0$ and denote $N = \lfloor s \rfloor$. Fix $y \in \mathbb{R}, y' \in \mathbb{T}$. We first write it as $y = j_{N+1} + x_{N+1}$ for some $j_{N+1} \in \mathbb{Z}$ and $x_{N+1} \in [0, 1)$. Since $\mathcal{Z}$ is the propagator of SHE, we have

$$\mathcal{Z}_{s,0}(y, y') = \int_{\mathbb{R}} \mathcal{Z}_{s,N}(j_{N+1} + x_{N+1}, x)\mathcal{Z}_{N,0}(x, y')dx$$

$$= \sum_{j_N \in \mathbb{Z}} \int_{\mathbb{T}} \mathcal{Z}_{s,N}(j_{N+1} + x_{N+1}, j_N + x_N)\mathcal{Z}_{N,0}(j_N + x_N, y')dx_N.$$

Iterating the above relation, we obtain

$$\mathcal{Z}_{s,0}(y, y') = \sum_{j_1, \dots, j_N \in \mathbb{Z}} \int_{\mathbb{T}^N} \mathcal{Z}_{s,N}(j_{N+1} + x_{N+1}, j_N + x_N)$$

$$\times \prod_{k=1}^N \mathcal{Z}_{k,k-1}(j_k + x_k, j_{k-1} + x_{k-1})d\mathbf{x}_{1,N},$$

where we used the simplified notation $d\mathbf{x}_{1,N} = dx_1 \dots dx_N$ and the convention $j_0 = 0, x_0 = y'$. In other words, in the above integration, we have decomposed $\mathbb{R} = \cup_{j \in \mathbb{Z}}[j, j + 1)$, then integrate in each interval and sum them up. One should think of the variable $j_k + x_k$ as representing the location of the polymer path at time k, with j_k the integer part and x_k the fractional part.

Now we make use of the periodicity:

$$(3.8) \qquad \begin{aligned} \sum_{j_k \in \mathbb{Z}} \mathcal{Z}_{k,k-1}(j_k + x_k, j_{k-1} + x_{k-1}) &= \sum_{j_k \in \mathbb{Z}} \mathcal{Z}_{k,k-1}(j_k - j_{k-1} + x_k, x_{k-1}) \\ &= \sum_{j_k \in \mathbb{Z}} \mathcal{Z}_{k,k-1}(j_k + x_k, x_{k-1}) = \mathcal{G}_{k,k-1}(x_k, x_{k-1}), \end{aligned}$$

where $\mathcal{G}$ is the periodic propagator of SHE. Then we can write

(3.9)

$$\mathcal{Z}_{s,0}(y,y') = \int_{\mathbb{T}^N} \left(\sum_{j_1,\dots,j_N \in \mathbb{Z}} \frac{\mathcal{Z}_{s,N}(j_{N+1}+x_{N+1}, j_N+x_N) \prod_{k=1}^N \mathcal{Z}_{k,k-1}(j_k+x_k, j_{k-1}+x_{k-1})}{\mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^N \mathcal{G}_{k,k-1}(x_k, x_{k-1})} \right)$$
$$\times \mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^N \mathcal{G}_{k,k-1}(x_k, x_{k-1}) dx_{1,N}.$$

Fix the realization of the random noise and the vector

(3.10) $\mathbf{x} := (x_0, \dots, x_{N+1}) \in \mathbb{T}^{N+2}.$

We construct an integer-valued, time inhomogeneous Markov chain $\{Y_j\}_{j=1}^N$, with

$$\mathbb{P}_{\mathbf{x}}[Y_1 = j_1] = \frac{\mathcal{Z}_{1,0}(j_1+x_1, x_0)}{\mathcal{G}_{1,0}(x_1, x_0)},$$

$$\mathbb{P}_{\mathbf{x}}[Y_2 = j_2 | Y_1 = j_1] = \frac{\mathcal{Z}_{2,1}(j_2+x_2, j_1+x_1)}{\mathcal{G}_{2,1}(x_2, x_1)},$$

(3.11) $\cdots$

$$\mathbb{P}_{\mathbf{x}}[Y_N = j_N | Y_{N-1} = j_{N-1}] = \frac{\mathcal{Z}_{N,N-1}(j_N+x_N, j_{N-1}+x_{N-1})}{\mathcal{G}_{N,N-1}(x_N, x_{N-1})},$$

$$\mathbb{P}_{\mathbf{x}}[Y_{N+1} = j_{N+1} | Y_N = j_N] = \frac{\mathcal{Z}_{s,N}(j_{N+1}+x_{N+1}, j_N+x_N)}{\mathcal{G}_{s,N}(x_{N+1}, x_N)}$$

By (3.8) and (3.11), it is clear that Y_{N+1} is a sum of independent random variables, for each fixed realization of the noise and $\mathbf{x}$. We rewrite it as

(3.12) $Y_{N+1} = \sum_{k=1}^{N+1} \eta_k,$ with $Y_0 = 0,$ $\eta_k = Y_k - Y_{k-1},$

and one can interpret η_k as the winding number accumulated during the time interval $[k-1, k]$ for $k = 1, \dots, N$, and η_{N+1} corresponds to the winding number in $[N, s]$. We have

$$\mathbb{P}_{\mathbf{x}}[\eta_k = j] = \frac{\mathcal{Z}_{k,k-1}(x_k+j, x_{k-1})}{\mathcal{G}_{k,k-1}(x_k, x_{k-1})}, \qquad j \in \mathbb{Z}, \quad k = 1, \dots, N,$$

(3.13)

$$\mathbb{P}_{\mathbf{x}}[\eta_{N+1} = j] = \frac{\mathcal{Z}_{s,N}(x_{N+1}+j, x_N)}{\mathcal{G}_{s,N}(x_{N+1}, x_N)}, \qquad j \in \mathbb{Z}.$$

With the Markov chain, one can write the summation in (3.9) as

$$\sum_{j_1,\dots,j_N} \frac{\mathcal{Z}_{s,N}(j_{N+1}+x_{N+1}, j_N+x_N) \prod_{k=1}^N \mathcal{Z}_{k,k-1}(j_k+x_k, j_{k-1}+x_{k-1})}{\mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^N \mathcal{G}_{k,k-1}(x_k, x_{k-1})}$$
$$= \mathbb{P}_{\mathbf{x}}[Y_{N+1} = j_{N+1}],$$

where, to emphasize the dependence of the Markov chain on $x_0, \dots, x_{N+1}$, we have denoted the probability by $\mathbb{P}_{\mathbf{x}}$.

Therefore, (3.9) is rewritten as

(3.14)

$$\mathcal{Z}_{s,0}(y,y') = \int_{\mathbb{T}^N} \mathbb{P}_{\mathbf{x}}[Y_{N+1} = j_{N+1}] \mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^N \mathcal{G}_{k,k-1}(x_k, x_{k-1}) d\mathbf{x}_{1,N}.$$

Now we introduce an additional notation: suppose that $f, g \in \mathbb{D}_c(\mathbb{T})$ - the set of continuous densities on the torus $\mathbb{T}$. Define

(3.15) $\mathscr{G}_N(\mathbf{x}; f, g) := \dfrac{f(x_{N+1}) \mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^N \mathcal{G}_{k,k-1}(x_k, x_{k-1}) g(x_0)}{C_{N,0}(f, g)},$

with $\mathbf{x} = (x_0, \ldots, x_{N+1})$ and the normalization factor
$$(3.16)$$
$$C_{N,0}(f, g) := \int_{\mathbb{T}^{N+2}} f(x_{N+1}) \mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^{N} \mathcal{G}_{k,k-1}(x_k, x_{k-1}) g(x_0) d\mathbf{x}_{0,N+1}$$
$$= \int_{\mathbb{T}^2} f(x_{N+1}) \mathcal{G}_{s,0}(x_{N+1}, x_0) g(x_0) dx_0 dx_{N+1}.$$

For each realization of the random environment, one should view $\mathscr{G}_N(\mathbf{x}; f, g)$ as the joint density of the polymer on the cylinder, evaluated at

$$(0, x_0), (1, x_1), \ldots, (N, x_N), (s, x_{N+1}),$$

with x_0 and x_{N+1} sampled from the densities g, f respectively. For any $\mu, \nu \in \mathcal{M}_1(\mathbb{T})$ and $(x_1, \ldots, x_N) \in \mathbb{T}^N$, we abuse the notation and write $\mathscr{G}_N(\mathbf{x}; \mu, \nu)$ as a measure

$$(3.17) \qquad \mathscr{G}_N(\mathbf{x}; \mu, \nu) := \frac{\mu(dx_{N+1}) \mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^{N} \mathcal{G}_{k,k-1}(x_k, x_{k-1}) \nu(dx_0)}{C_{N,0}(\mu, \nu)}.$$

In this case, $C_{N,0}(\mu, \nu)$ equals to

$$C_{N,0}(\mu, \nu) := \int_{\mathbb{T}^{N+2}} \mathcal{G}_{s,N}(x_{N+1}, x_N) \prod_{k=1}^{N} \mathcal{G}_{k,k-1}(x_k, x_{k-1}) \nu(dx_0) d\mathbf{x}_{1,N} \mu(dx_{N+1}).$$

Recall that the winding number $\mathscr{W}(s; \mu, \nu)$ has the representation given by (3.7). With the Markov chain $\{Y_j\}_{j=1}^{N}$ and the above notation, one can express $\mathcal{Z}_{s,0}(y, y')$ in terms of Y_{N+1} and $\mathscr{G}_N$. This leads to

Lemma 3.4. *We have*

$$(3.18) \qquad \mathcal{W}(s; \mu, \nu) = -\int_{\mathbb{T}^{N+2}} (x_{N+1} + \mathbb{E}_{\mathbf{x}} Y_{N+1} - x_0) \mathscr{G}_N(\mathbf{x}; \mu, \nu) d\mathbf{x}_{1,N},$$

where $\mathbb{E}_{\mathbf{x}}$ is the expectation w.r.t. $\mathbb{P}_{\mathbf{x}}$.

Remark 3.5. On the r.h.s. of (3.18), the domain of integration is $\mathbb{T}^{N+2}$, and this is because we have defined $\mathscr{G}_N(\mathbf{x}; \mu, \nu)$ as in (3.17) which includes $\mu(dx_{N+1})$ and $\nu(dx_0)$.

Proof. Recall the expression for $\mathcal{W}(s; \mu, \nu)$ obtained in (3.7):

$$\mathcal{W}(s; \mu, \nu) = -\frac{\int_{\mathbb{R}} \left(\int_{\mathbb{T}} \mathcal{Z}_{s,0}(y, y')(y - y') \nu(dy') \right) \mu(dy)}{\int_{\mathbb{R}} \left(\int_{\mathbb{T}} \mathcal{Z}_{s,0}(y, y') \nu(dy') \right) \mu(dy)}.$$

For the numerator, using (3.14) and (3.17), we rewrite it as (with $y = j_{N+1} + x_{N+1}, y' = x_0$)

$$\int_{\mathbb{R}} \left(\int_{\mathbb{T}} \mathcal{Z}_{s,0}(y, y')(y - y') \nu(dy') \right) \mu(dy)$$
$$= C_{N,0}(\mu, \nu) \sum_{j_{N+1} \in \mathbb{Z}} \int_{\mathbb{T}^2} \mathbb{P}_{\mathbf{x}}[Y_{N+1} = j_{N+1}](j_{N+1} + x_{N+1} - x_0) \mathscr{G}_N(\mathbf{x}; \mu, \nu) dx_{1,N}$$
$$= C_{N,0}(\mu, \nu) \int_{\mathbb{T}^2} (\mathbb{E}_{\mathbf{x}}[Y_{N+1}] + x_{N+1} - x_0) \mathscr{G}_N(\mathbf{x}; \mu, \nu) dx_{1,N}.$$

Similarly, it is easy to check that the denominator equals to $C_{N,0}(\mu, \nu)$. The proof is complete. □

3.2. Exponential mixing.
In this section, we start from the expression in (3.18) and make use of the exponential mixing of the endpoint distribution of directed polymer on the cylinder to prove (3.6).

First, we need the following moment estimate

Lemma 3.6. *For any $p \in [1, \infty)$, we have*

$$(3.19) \qquad \sup_{s \in (0,2]} \sup_{x,y \in \mathbb{T}} \mathbf{E} \sum_{j \in \mathbb{Z}} |j|^p \frac{\mathcal{Z}_{s,0}(x+j,y)}{\mathcal{G}_{s,0}(x,y)} \le C_p.$$

As a result, we conclude that

$$(3.20) \qquad \sup_{N \ge 1} \sup_{\mathbf{x}} \max_{k=1,\dots,N+1} \mathbf{E} \mathbb{E}_{\mathbf{x}} |\eta_k|^p \le C_p.$$

Similar estimates can be found in [33, Lemma 3.2, Lemma 4.1]. The proof here covers more general cases.

Proof. First we state the following estimate, which can be proved using rather standard techniques (e.g. see [39, Theorem 2.6], the proof in case of spatially periodic noise can be repeated verbatim): there exists $C > 0$, only depending on $p \ge 1$, such that for $s \in (0, 2]$, $x, y \in \mathbb{T}$ and $j \in \mathbb{Z}$, we have

$$(3.21) \qquad \mathbf{E} \mathcal{Z}_{s,0}(x+j,y)^p \le C q_s(x+j-y)^p.$$

From (3.21) and the triangle inequality, we have (with $\| \cdot \|_p$ denoting the $L^p(\Omega)$ norm)

$$(3.22) \qquad \| \mathcal{G}_{s,0}(x,y) \|_p = \| \sum_{j \in \mathbb{Z}} \mathcal{Z}_{s,0}(x+j,y) \|_p \le C \sum_{j \in \mathbb{Z}} q_s(x+j-y) = C G_s(x-y)$$

We also have the negative moment estimate by [39, Corollary 4.8][2]: for any $p \in [1, \infty)$ there exists $C > 0$ such that

$$(3.23) \qquad \mathbf{E} \mathcal{G}_{s,0}(x,y)^{-p} \le C G_s(x-y)^{-p}, \quad x, y \in \mathbb{T}, \, s \in (0,2].$$

Applying the Cauchy-Schwarz inequality and the estimate (3.21) together with (3.23), we derive

$$\mathbf{E} \sum_{j \in \mathbb{Z}} |j|^p \frac{\mathcal{Z}_{s,0}(x+j,y)}{\mathcal{G}_{s,0}(x,y)} \le C \sum_{j \in \mathbb{Z}} |j|^p \frac{q_s(x+j-y)}{G_s(x-y)}.$$

Now we write

$$\sum_{j \in \mathbb{Z}} |j|^p \frac{q_s(x+j-y)}{G_s(x-y)} \le C + \sum_{|j|>100} |j|^p \frac{q_s(x+j-y)}{q_s(x-y)}$$

and note that the r.h.s. is bounded, uniformly in $s \in (0,2]$ and $x, y \in \mathbb{T}$. This completes the proof of (3.19). To see why (3.20) holds, we only need to recall the definition of η_k in (3.13) and apply (3.19). □

Next, we need the following lemma, which, roughly speaking, states that the dependence on η_k on the two distributions μ and ν is exponentially small if $k \gg 1$ and $N - k \gg 1$ (recall that $N = \lfloor s \rfloor$).

Lemma 3.7. *For any $p \in [1, \infty)$, there exist $C, \lambda > 0$ such that*

$$(3.24) \qquad \begin{aligned} \mathbf{E} \left| \int_{\mathbb{T}^{N+2}} \mathbb{E}_{\mathbf{x}} \eta_k [\mathscr{G}_N(\mathbf{x}; \mu_1, \nu_1) - \mathscr{G}_N(\mathbf{x}; \mu_1, \nu_2)] dx_{1,N} \right|^p &\le C e^{-\lambda k}, \\ \mathbf{E} \left| \int_{\mathbb{T}^{N+2}} \mathbb{E}_{\mathbf{x}} \eta_k [\mathscr{G}_N(\mathbf{x}; \mu_1, \nu_1) - \mathscr{G}_N(\mathbf{x}; \mu_2, \nu_1)] dx_{1,N} \right|^p &\le C e^{-\lambda(N-k)}, \end{aligned}$$

for all $\mu_i, \nu_i \in \mathcal{M}_1(\mathbb{T})$, with $i = 1, 2$, and $k = 1, \dots, N+1$.

Proof. We will only consider the first inequality as the second one is proved in the same way. By Lemma 3.6, we only need to consider the case of $N \gg 1$ and $k \gg 1$.

The case of $1 \le k \le N$. By definition (3.13), we know that

$$\mathbb{E}_{\mathbf{x}} \eta_k = \sum_j j \mathbb{P}_{\mathbf{x}} [\eta_k = j] = \sum_j j \frac{\mathcal{Z}_{k,k-1}(x_k+j, x_{k-1})}{\mathcal{G}_{k,k-1}(x_k, x_{k-1})}$$

---

[2]The result in [39] is for the equation on the whole space, but the proof applies verbatim to the torus case.

which only depends on x_{k-1}, x_k and the noise in the time interval $[k-1, k]$. Using the forward and backward polymer endpoint densities, defined in (2.7), we can integrate out all x−variables except x_{k-1}, x_k and obtain

$$(3.25) \quad \begin{aligned} &\int_{\mathbb{T}^{N+2}} \mathbb{E}_{\mathbf{x}} \eta_k \, \mathscr{G}_N(\mathbf{x}; \mu_1, \nu_i) dx_{1,N} \\ &= a_i^{-1} \int_{\mathbb{T}^2} \mathbb{E}_{\mathbf{x}} \eta_k \, \rho_{\mathrm{b}}(s, \mu_1; k, x_k) \mathcal{G}_{k,k-1}(x_k, x_{k-1}) \rho_{\mathrm{f}}(k-1, x_{k-1}; 0, \nu_i) dx_{k-1} dx_k, \end{aligned}$$

with

$$a_i := \int_{\mathbb{T}^2} \rho_{\mathrm{b}}(s, \mu_1; k, x_k) \mathcal{G}_{k,k-1}(x_k, x_{k-1}) \rho_{\mathrm{f}}(k-1, x_{k-1}; 0, \nu_i) dx_{k-1} dx_k, \quad i = 1, 2.$$

Then one can write the difference as

$$\int_{\mathbb{T}^{N+2}} \mathbb{E}_{\mathbf{x}} \eta_k \left[\mathscr{G}_N(\mathbf{x}; \mu_1, \nu_1) - \mathscr{G}_N(\mathbf{x}; \mu_1, \nu_2) \right] dx_{1,N} = E_1 + E_2$$

with

$$\begin{aligned} E_1 = a_1^{-1} \int_{\mathbb{T}^2} \mathbb{E}_{\mathbf{x}} \eta_k \, \rho_{\mathrm{b}}(s, \mu_1; k, x_k) \mathcal{G}_{k,k-1}(x_k, x_{k-1}) \\ \times \left[\rho_{\mathrm{f}}(k-1, x_{k-1}; 0, \nu_1) - \rho_{\mathrm{f}}(k-1, x_{k-1}; 0, \nu_2) \right] dx_{k-1} dx_k, \end{aligned}$$

and

$$\begin{aligned} E_2 = \int_{\mathbb{T}^2} \mathbb{E}_{\mathbf{x}} \eta_k \, \rho_{\mathrm{b}}(s, \mu_1; k, x_k) \mathcal{G}_{k,k-1}(x_k, x_{k-1}) \rho_{\mathrm{f}}(k-1, x_{k-1}; 0, \nu_2) dx_{k-1} dx_k \\ \times (a_1^{-1} - a_2^{-1}). \end{aligned}$$

It suffices to show that

$$(3.26) \qquad \mathbf{E}|E_1 + E_2|^p \le Ce^{-\lambda k},$$

for some $C, \lambda > 0$ depending only on p.

We collect three facts which will be used in the proof:

(i) there exists C, depending only on $p \in [1, \infty)$ such that
$$\mathbf{E} \sup_{x,y \in \mathbb{T}} \mathcal{G}_{k,k-1}(x,y)^p + \mathbf{E} \sup_{x,y \in \mathbb{T}} \mathcal{G}_{k,k-1}(x,y)^{-p} \le C,$$
see [34, Lemma 4.1].

(ii) by Proposition A.1, there exists C, depending only on $p \in [1, \infty)$, such that
$$\mathbf{E} \sup_{x \in \mathbb{T}} |\rho_{\mathrm{f}}(k, x; 0, \nu_1) - \rho_{\mathrm{f}}(k, x; 0, \nu_2)|^p \le Ce^{-\lambda k}, \qquad k \ge 10.$$

(iii) there exists C, depending only on $p \in [1, \infty)$, such that
$$\mathbf{E} \sup_{x \in \mathbb{T}, \nu \in \mathcal{M}_1(\mathbb{T})} |\rho_{\mathrm{f}}(k, x; 0, \nu)|^p \le C, \qquad k \ge 10.$$

Using (i) and Jensen's inequality, we conclude that for any $p \ge 1$,
$$\mathbf{E}\left[a_i^p + a_i^{-p} \right] \le C, \quad i = 1, 2.$$

Now applying the Hölder inequality and (ii), we have

$$\mathbf{E}|E_1|^p \le \sqrt{\mathbf{E} a_1^{-2p}} \sqrt{\mathbf{E} \int_{\mathbb{T}^2} \rho_{\mathrm{b}}(s, \mu_1; k, x_k) |E_3|^{2p} dx_{k-1} dx_k},$$

with

$$E_3 = \mathbb{E}_{\mathbf{x}} \eta_k \, \mathcal{G}_{k,k-1}(x_k, x_{k-1}) \left[\rho_{\mathrm{f}}(k-1, x_{k-1}; 0, \nu_1) - \rho_{\mathrm{f}}(k-1, x_{k-1}; 0, \nu_2) \right].$$

Note that $\rho_{\mathrm{b}}(s, \mu_1; k, \cdot)$ is independent of E_3 and we have the bound

$$\begin{aligned} &\mathbf{E} \int_{\mathbb{T}^2} \rho_{\mathrm{b}}(s, \mu_1; k, x_k) |E_3|^{2p} dx_{k-1} dx_k \\ &= \int_{\mathbb{T}^2} \mathbf{E} \rho_{\mathrm{b}}(s, \mu_1; k, x_k) \mathbf{E}|E_3|^{2p} dx_{k-1} dx_k \le Ce^{-\lambda k}, \end{aligned}$$

thanks to (i), (ii) and Lemma 3.6. This implies that $\mathbf{E}|E_1|^p \le Ce^{-\lambda k}$. The analysis of E_2 is similar, so we do not repeat it here. Combine them together, we conclude the proof of (3.26).

The case of $k = N + 1$. In this case we know that

$$\mathbb{E}_{\mathbf{x}}\eta_{N+1} = \sum_j j \frac{\mathcal{Z}_{s,N}(x_{N+1}+j, x_N)}{\mathcal{G}_{s,N}(x_{N+1}, x_N)},$$

and for $i = 1, 2$, we have

$$(3.27) \quad \begin{aligned} &\int_{\mathbb{T}^{N+2}} \mathbb{E}_{\mathbf{x}}\eta_{N+1} \mathscr{G}_N(\mathbf{x}; \mu_1, \nu_i)dx_{0,N+1} \\ &= a_i^{-1} \int_{\mathbb{T}^2} \mathbb{E}_{\mathbf{x}}\eta_{N+1} \mathcal{G}_{s,N}(x_{N+1}, x_N)\rho_{\mathsf{f}}(N, x_N; 0, \nu_i)dx_N \mu_1(dx_{N+1}), \end{aligned}$$

with

$$a_i = \int_{\mathbb{T}^2} \mathcal{G}_{s,N}(x_{N+1}, x_N)\rho_{\mathsf{f}}(N, x_N; 0, \nu_i)dx_N\mu_1(dx_{N+1}).$$

Compared to (3.25), the difference is we do not have the factor ρ_{b} in (3.27). The rest of the proof is the same once we use (iii) and the following estimate (see e.g. [34, Lemma B.2])

$$\sup_{x_{N+1}\in\mathbb{T}} \left\{ \mathbf{E}\left|\int_{\mathbb{T}} \mathcal{G}_{s,N}(x_{N+1}, x_N)dx_N\right|^p + \mathbf{E}\left|\int_{\mathbb{T}} \mathcal{G}_{s,N}(x_{N+1}, x_N)dx_N\right|^{-p} \right\} \le C.$$

The proof is complete. □

A similar proof of the previous lemma leads to

Lemma 3.8. *There exist $C, \lambda > 0$ such that*

$$(3.28) \quad \mathbf{E}\left|\int_{\mathbb{T}^{N+2}} x_{N+1}[\mathscr{G}_N(\mathbf{x}; \mu_1, \nu_1) - \mathscr{G}_N(\mathbf{x}; \mu_1, \nu_2)]d\mathbf{x}_{1,N}\right|^p \le Ce^{-\lambda N},$$

$$(3.29) \quad \mathbf{E}\left|\int_{\mathbb{T}^{N+2}} x_0[\mathscr{G}_N(\mathbf{x}; \mu_1, \nu_1) - \mathscr{G}_N(\mathbf{x}; \mu_2, \nu_1)]d\mathbf{x}_{1,N}\right|^p \le Ce^{-\lambda N},$$

for any $\mu_i, \nu_i \in \mathcal{M}_1(\mathbb{T})$, $i = 1, 2$.

Using (3.12), (3.18) and Lemma 3.7, we obtain

Corollary 3.9. *For any $p \in [1, \infty)$, there exists $C > 0$ such that for all $s \ge 0$,*

$$(3.30) \quad \sup_{\mu_1, \mu_2, \nu_1, \nu_2 \in \mathcal{M}_1(\mathbb{T})} \mathbf{E}\big|\mathcal{W}(s; \mu_1, \nu_1) - \mathcal{W}(s, \mu_2, \nu_2)\big|^p \le C.$$

Remark 3.10. Recall that $\psi(s, y) = \mathcal{W}(s; \delta_y, \delta_0)$ and $\phi(s, y) = \mathcal{W}(s; \delta_y, \rho_W)$. It is straightforward to check that another consequence of Lemmas 3.7 and 3.8 is the convergence in distribution of $\psi(s, y') - \psi(s, y)$ and $\phi(s, y') - \phi(s, y)$ as $s \to \infty$, for any $y', y \in \mathbb{T}$. As a matter of fact, there is a one-force-one-solution principle: suppose we abuse the notation and let $\psi(s, y)$ (or $\phi(s, y)$) be the winding number of the polymer path in the time interval $[-K, s]$, then $\psi(s, y') - \psi(s, y)$ (or $\phi(s, y') - \phi(s, y)$) converges in $L^p(\Omega)$ as $K \to \infty$.

3.3. **Proof of Proposition 3.2.** With Lemmas 3.7 and 3.8, we can complete the proof of (3.6) and conclude the proof of Proposition 3.2.

Proof of (3.6). By Lemma 3.6, it suffices to consider those $N \gg 1$. By (3.18) and (3.12), we can first bound the difference as

$$\begin{aligned} &\big|[\mathcal{W}(s; \mu_1, \nu_1) - \mathcal{W}(s, \mu_2, \nu_1)] - [\mathcal{W}(s; \mu_1, \nu_2) - \mathcal{W}(s; \mu_2, \nu_2)]\big| \\ &\le E_1 + E_2 + E_3, \end{aligned}$$

with

$$E_1 = \sum_{i=1}^{2} \left| \int_{\mathbb{T}^{N+2}} x_{N+1} [\mathscr{G}_N(\mathbf{x}; \mu_i, \nu_1) - \mathscr{G}_N(\mathbf{x}; \mu_i, \nu_2)] d\mathbf{x}_{1,N} \right|,$$

$$E_2 = \sum_{i=1}^{2} \left| \int_{\mathbb{T}^{N+2}} x_0 [\mathscr{G}_N(\mathbf{x}; \mu_1, \nu_i) - \mathscr{G}_N(\mathbf{x}; \mu_2, \nu_i)] d\mathbf{x}_{1,N} \right|,$$

$$E_3 = \sum_{i=1}^{2} \left| \sum_{N+1 \geq k \geq N/2} \int_{\mathbb{T}^{N+2}} \mathbb{E}_{\mathbf{x}} \eta_k [\mathscr{G}_N(\mathbf{x}; \mu_i, \nu_1) - \mathscr{G}_N(\mathbf{x}; \mu_i, \nu_2)] dx_{1,N} \right|$$

$$+ \sum_{i=1}^{2} \left| \sum_{1 \leq k < N/2} \int_{\mathbb{T}^{N+2}} \mathbb{E}_{\mathbf{x}} \eta_k [\mathscr{G}_N(\mathbf{x}; \mu_1, \nu_i) - \mathscr{G}_N(\mathbf{x}; \mu_2, \nu_i)] dx_{1,N} \right|.$$

Then it suffices to apply the estimates in Lemmas 3.7 and 3.8 to complete the proof. □

4. Passive scalar in Burgers flow and time reversal

Recall that in Section 2, we have used the Clark-Ocone formula to write the quantity of interest, the quenched mean X_t, as an Itô integral

$$(4.1) \qquad X_t = \int_{\mathbb{R}} x\rho(t,x)dx = \beta \int_0^t \int_{\mathbb{T}} \mathbf{E}[I_{1,t}(s,y) - I_{2,t}(s,y) | \mathcal{F}_s] \xi(s,y) dy ds,$$

with

$$(4.2) \qquad I_{1,t}(s,y) - I_{2,t}(s,y) = \rho_{\mathrm{m}}(t,-|s,y) \int_{\mathbb{T}} [\psi(s,y') - \psi(s,y)] \rho_{\mathrm{m}}(t,-|s,y') dy'.$$

In Section 3, we have derived the estimate in Proposition 3.2 which will eventually enable us to replace $\psi(s,y') - \psi(s,y)$ by its "stationary" version $\phi(s,y') - \phi(s,y)$. Therefore, to have an explicit expression for the variance of X_t, it remains yet to understand the joint distribution of $\rho_{\mathrm{m}}(t,-|s,\cdot)$ and $\nabla\phi(s,\cdot)$. It turns out to be the major challenge and is the goal of this section.

The rest of this section is divided into three parts. In Section 4.1, we use the Girsanov transformation to reduce the study of the directed polymer to a diffusion in a *non-Markovian* drift. By a standard homogenization argument, the quenched mean of the polymer endpoint turns out to be the corrector from a formal two-scale expansion. In Section 4.2, we further study the gradient of the corrector and draw the connection to another diffusion with a *Markovian* drift. Using the time-reversal anti-symmetry between the aforementioned two drifts, we form a link between the Gibbsian and the Markovian settings, which eventually enable us to derive the joint law of the gradient of the corrector and the drift, see Proposition 4.8 below. This may be compared to the study of geodesics and competing interfaces in the context of last passage percolation [48]. On the technical side, all these will be done on the level of approximations, since the drifts are distribution valued. In Section 4.3 we borrow the tools from singular diffusion to pass to the limit.

4.1. **Winding number, diffusion in random environment, corrector.** The goal of this section is to utilize the well-known connection between the directed polymer and the passive scalar in a random drift, which solves the stochastic Burgers equation, and to rewrite the difference $\phi(s,y') - \phi(s,y)$ in terms of the solution of a Fokker-Planck equation with another, different but related, drift. One technical difficulty is that, since the random environment for the directed polymer is the spacetime white noise, the random drift for the passive scalar is distribution valued. For this reason we start with an approximation.

Recall that ϕ was defined in (3.2). For any $t > 0, x \in \mathbb{T}$, we define the ε−approximation of ϕ as

$$
\begin{aligned}
\phi^\varepsilon(t, x) &= \frac{\int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x')(x' - x)e^{\beta W(x')}dx'}{\int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x')e^{\beta W(x')}dx'} \\
&= \frac{\mathbb{E}_B\big[e^{\beta \int_0^t \xi^\varepsilon(t-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t}e^{\beta W(B_t)}(B_t - x)|B_0 = x\big]}{\mathbb{E}_B\big[e^{\beta \int_0^t \xi^\varepsilon(t-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t}e^{\beta W(B_t)}|B_0 = x\big]}.
\end{aligned}
\tag{4.3}
$$

For any $s < t, x, y \in \mathbb{T}$, define the midpoint density of the "stationary" polymer at (s, y) by

$$
\rho_{\mathrm{m}}^\varepsilon(t, x|s, y) = \frac{\mathbb{E}_B\big[e^{\beta \int_0^t \xi^\varepsilon(t-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t}e^{\beta W(B_t)}\delta(\dot{B}_{t-s} - y)|B_0 = x\big]}{\mathbb{E}_B\big[e^{\beta \int_0^t \xi^\varepsilon(t-\ell, B_\ell)d\ell - \frac{1}{2}\beta^2 R^\varepsilon(0)t}e^{\beta W(B_t)}|B_0 = x\big]},
\tag{4.4}
$$

where we recall that $\dot{B}_s$ denotes the fractional part of B_s. It is clear that

$$
\int_{\mathbb{T}} \rho_{\mathrm{m}}^\varepsilon(t, x|s, y)dy = 1.
$$

If we view the polymer path as lying on the cylinder, starting from (t, x) and going backward in time with the terminal potential given by βW, then $\rho_{\mathrm{m}}^\varepsilon(t, x|s, \cdot)$ is precisely the quenched density at time s. Using the propagator of SHE (with mollified noise ξ^ε), we can rewrite $\rho_{\mathrm{m}}^\varepsilon$ as

$$
\begin{aligned}
\rho_{\mathrm{m}}^\varepsilon(t, x|s, y) &= \frac{\sum_{n \in \mathbb{Z}} \mathcal{Z}_{t,s}^\varepsilon(x, y+n)\int_{\mathbb{R}} \mathcal{Z}_{s,0}^\varepsilon(y+n, x')e^{\beta W(x')}dx'}{\int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x')e^{\beta W(x')}dx'} \\
&= \frac{\mathcal{G}_{t,s}^\varepsilon(x, y)\int_{\mathbb{T}} \mathcal{G}_{s,0}^\varepsilon(y, x')e^{\beta W(x')}dx'}{\int_{\mathbb{T}} \mathcal{G}_{t,0}^\varepsilon(x, x')e^{\beta W(x')}dx'}.
\end{aligned}
\tag{4.5}
$$

Remark 4.1. At this point, it is worth pointing out the difference between the two midpoint densities, $\rho_{\mathrm{m}}(t, -|s, y)$ and $\rho_{\mathrm{m}}^\varepsilon(t, x|s, y)$, defined in (2.21) and (4.4) respectively – one should not confuse them with each other. Apart from the different noise ξ and ξ^ε, the prescribed distribution of two endpoints are different: for $\rho_{\mathrm{m}}(t, -|s, y)$, the endpoint at t is distributed according to Lebesgue measure on $\mathbb{T}$ and the endpoint at 0 is fixed at 0; for $\rho_{\mathrm{m}}^\varepsilon(t, x|s, y)$, the endpoint at t is fixed at x and the endpoint at 0 is distributed according to the density $\rho_W(x) = e^{\beta W(x)}/\int_{\mathbb{T}} e^{\beta W(x')}dx'$.

Define $Z_W^\varepsilon(t, x) := \int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, y)e^{\beta W(y)}dy$, which solves

$$
\partial_t Z_W^\varepsilon = \frac{1}{2}\Delta Z_W^\varepsilon + \beta Z_W^\varepsilon \xi^\varepsilon, \qquad Z_W^\varepsilon(0, x) = e^{\beta W(x)},
\tag{4.6}
$$

and define

$$
h^\varepsilon = \log Z_W^\varepsilon, \qquad u^\varepsilon = \nabla h^\varepsilon.
\tag{4.7}
$$

Here one should keep in mind that both h^ε and u^ε depend on W, and we only have kept it implicit to make the notation less burdensome.

From now on, we fix some $t > 0$ and $x \in \mathbb{T}$. Consider the SDE

$$
d\mathcal{X}_s^\varepsilon = u^\varepsilon(t - s, \mathcal{X}_s^\varepsilon)ds + dB_s, \qquad \mathcal{X}_0^\varepsilon = x.
\tag{4.8}
$$

We emphasize that the above is a simplified notation where the dependence on (t, x) is omitted.

Lemma 4.2. *The quenched law of $\{\mathcal{X}_s^\varepsilon\}_{0 \le s \le t}$ coincides with that of the path $\{B_s\}_{0 \le s \le t}$ of the random polymer under the measure*

$$
\frac{1}{Z_W^\varepsilon(t, x)} \exp\left\{\beta \int_0^t \xi^\varepsilon(t - s, B_s)ds - \frac{\beta^2}{2}R^\varepsilon(0)t + \beta W(B_t)\right\} \mathbb{W}_t(dB),
$$

where $\mathbb{W}_t$ is the Wiener measure on $C[0, t]$ with $B_0 = x$.

Proof. The result is well-known, and the proof is rather standard, so we only present it here for the convenience of readers. Recall that $h^\varepsilon = \log Z_W^\varepsilon$. Consider $Y_s = h^\varepsilon(t-s, B_s)$, with B a standard Brownian motion starting at x. Applying Ito formula, we have

$$dY_s = (-\partial_t h^\varepsilon + \frac{1}{2}\Delta h^\varepsilon)ds + \nabla h^\varepsilon dB_s$$
$$= \left(\nabla h^\varepsilon(t-s,B_s)dB_s - \frac{1}{2}|\nabla h^\varepsilon(t-s,B_s)|^2 ds\right) - \beta\xi^\varepsilon(t-s,B_s)ds + \frac{1}{2}\beta^2 R^\varepsilon(0)ds,$$

which implies that (recall that $h^\varepsilon(0,x) = \beta W(x)$)

$$\beta W(B_t) = h^\varepsilon(t,x) + \int_0^t \left(\nabla h^\varepsilon(t-s,B_s)dB_s - \frac{1}{2}|\nabla h^\varepsilon(t-s,B_s)|^2 ds\right)$$
$$- \beta\int_0^t \xi^\varepsilon(t-s,B_s)ds + \frac{1}{2}\beta^2 R^\varepsilon(0)t.$$

So one can write

$$\frac{1}{Z_W^\varepsilon(t,x)} \exp\left\{\beta\int_0^t \xi^\varepsilon(t-s,B_s)ds - \frac{1}{2}\beta^2 R^\varepsilon(0)t + \beta W(B_t)\right\}$$
$$= \exp\left\{\int_0^t \nabla h^\varepsilon(t-s,B_s)dB_s - \frac{1}{2}\int_0^t |\nabla h^\varepsilon(t-s,B_s)|^2 ds\right\}.$$

It remains to apply the Girsanov theorem to complete the proof. $\qquad\square$

Using the previous lemma, we can express the winding number ϕ^ε in terms of the diffusion $\mathcal{X}^\varepsilon$:

Lemma 4.3. *For any $t > 0, x \in \mathbb{T}$, we have*

$$(4.9) \qquad \begin{aligned} \phi^\varepsilon(t,x) = \mathbb{E}_B[\mathcal{X}_t^\varepsilon] - x &= \int_0^t \mathbb{E}_B u^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)ds \\ &= \int_0^t \int_\mathbb{T} u^\varepsilon(t-s,y)\rho_\mathrm{m}^\varepsilon(t,x|t-s,y)dyds. \end{aligned}$$

Proof. First, from (4.3) and Lemma 4.2, we know that $\phi^\varepsilon(t,x) = \mathbb{E}_B \mathcal{X}_t^\varepsilon - x$. Then, from the SDE (4.8), we have

$$\mathcal{X}_t^\varepsilon = x + \int_0^t u^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)ds + B_t.$$

Taking expectation on B yields the second equality in (4.9).

Thanks to Lemma 4.2, the quenched density of $\mathcal{X}_s^\varepsilon$ on the torus is $\rho_\mathrm{m}^\varepsilon(t,x|t-s,\cdot)$, so the third equality in (4.9) follows, which completes the proof. $\square$

Define

$$(4.10) \qquad g^\varepsilon(t,x) = \nabla(x + \phi^\varepsilon(t,x)) = 1 + \nabla\phi^\varepsilon(t,x),$$

which will play a crucial role later. Next we show that g^ε solve a Fokker-Planck equation with the random drift given by u^ε:

Lemma 4.4. *We have: ϕ^ε solves*

$$(4.11) \qquad \partial_t\phi^\varepsilon = \tfrac{1}{2}\Delta\phi^\varepsilon + u^\varepsilon\nabla\phi^\varepsilon + u^\varepsilon, \qquad \phi^\varepsilon(0,x) = 0,$$

and g^ε solves

$$(4.12) \qquad \partial_t g^\varepsilon = \tfrac{1}{2}\Delta g^\varepsilon + \nabla(u^\varepsilon g^\varepsilon), \qquad g^\varepsilon(0,x) = 1.$$

Proof. Consider the solution to (4.11), with a bit of an abuse of notation we still denote it by ϕ^ε. Applying the Itô formula for the process $Y_s = \phi^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)$, we obtain

$$dY_s = (-\partial_t\phi^\varepsilon + \frac{1}{2}\Delta\phi^\varepsilon + u^\varepsilon\nabla\phi^\varepsilon)ds + \nabla\phi^\varepsilon dB_s$$
$$= -u^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)ds + \nabla\phi^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)dB_s,$$

which implies that (recall that $\mathcal{X}_0^\varepsilon = x$)

$$(4.13) \qquad 0 = \phi^\varepsilon(t,x) - \int_0^t u^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)ds + \int_0^t \nabla\phi^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)dB_s.$$

Taking expectation with respect to B on both sides shows that

$$\phi^\varepsilon(t,x) = \mathbb{E}_B \int_0^t u^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)ds.$$

Applying Lemma 4.3, we complete the proof of (4.11). Since $\phi^\varepsilon(t,x)$ is smooth in the x−variable and by definition $g^\varepsilon = 1 + \nabla\phi^\varepsilon$, we take derivatives in x on both sides of (4.11) to obtain (4.12). □

The equation (4.11) for ϕ^ε can be viewed as the corrector equation for the diffusion in random environment $\mathcal{X}_s^\varepsilon$ (it is the ε−version of (1.10)). Namely, using ϕ^ε (see (4.13)), we can decompose the drift term in (4.8) as

$$\int_0^t u^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)ds = \phi^\varepsilon(t,x) + \int_0^t \nabla\phi^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)dB_s,$$

so the process itself can be written as

$$(4.14) \qquad \begin{aligned} \mathcal{X}_t^\varepsilon &= x + \int_0^t u^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)ds + B_t \\ &= x + \phi^\varepsilon(t,x) + \int_0^t g^\varepsilon(t-s,\mathcal{X}_s^\varepsilon)dB_s. \end{aligned}$$

In other words, to study the fluctuations of the directed polymer endpoint, by Lemma 4.2, it suffices to consider the process $\mathcal{X}^\varepsilon$, and the decomposition in (4.14) is the start of the standard homogenization argument. Nevertheless, unlike the "static" homogenization where the corrector (constituting the so called co-boundary term) does not contribute on the diffusive scale, see e.g. the models in [41, Chapter 9] and the references cited there, in our dynamic setting with a compressible drift, the ϕ^ε term contributes to the limiting Gaussian distribution. Actually, the main challenge of the paper is to quantify this contribution.

4.2. **Fokker-Planck equation and time reversal.** Recall that ϕ^ε was defined in (4.3), and as $\varepsilon \to 0$ it converges to

$$(4.15) \qquad \phi(t,x) = \frac{\int_\mathbb{R} \mathcal{Z}_{t,0}(x,x')(x'-x)e^{\beta W(x')}dx'}{\int_\mathbb{R} \mathcal{Z}_{t,0}(x,x')e^{\beta W(x')}dx'}$$

In addition, we have $Z_W^\varepsilon, h^\varepsilon, u^\varepsilon$ converge to Z_W, h, u respectively, with

$$(4.16) \qquad \begin{aligned} Z_W(t,x) &= \int_\mathbb{R} \mathcal{Z}_{t,0}(x,y)e^{\beta W(y)}dy, \\ h &= \log Z_W, \qquad u = \nabla h. \end{aligned}$$

Thus, h, u solve the KPZ, stochastic Burgers equation, respectively, that start from stationarity. Similarly, the dependence of h, u on W was kept implicit.

Our goal is to derive the joint distribution of $(u(s,\cdot),\nabla\phi(s,\cdot))$ for large $s \gg 1$, which was needed in the Clark-Ocone representation (see (4.1) and (4.2)). Here $u(s,\cdot)$ will appear in the term $\rho_\mathrm{m}(t,-|s,y)$, and, once we approximate ψ by ϕ, the other term can be written as $\phi(s,y') - \phi(s,y) = \int_y^{y'} \nabla\phi(s,z)dz$. It is worth pointing out that $\nabla\phi(s,\cdot)$ and $u(s,\cdot)$ are distribution-valued processes and what we actually need is the function-valued "increment processes" $\phi(s,y) - \phi(s,0)$ and $h(s,y) - h(s,0) = \int_0^y u(s,z)dz$.

Since ϕ can be approximated by ϕ^ε and the increments of ϕ^ε is related to g^ε through

$$\phi^\varepsilon(s,y') - \phi^\varepsilon(s,y) = \int_y^{y'} \nabla\phi^\varepsilon(s,z)dz = \int_y^{y'} (g^\varepsilon(s,z)-1)dz,$$

it suffices to study the joint distribution of $u^\varepsilon(s, \cdot)$ and $g^\varepsilon(s, \cdot)$. On the other hand, we know from Lemma 4.4 that g^ε solves the Fokker-Planck equation with the drift given by u^ε:

$$(4.17) \qquad \partial_t g^\varepsilon = \frac{1}{2}\Delta g^\varepsilon + \nabla(u^\varepsilon g^\varepsilon), \qquad g^\varepsilon(0, x) = 1.$$

Thus, the problem reduces to studying the above equation and the joint distribution of the coefficient and the solution. For a Fokker-Planck equation driven by a general random drift, it does not seem likely that one could obtain such explicit descriptions. However, in our case of the Burgers drift, there is an amazing time reversal anti-symmetry (see [8, Proposition 1.1]):

Proposition 4.5. *Fix any $t > 0$. Then*

$$(4.18) \qquad \{u(s, \cdot)\}_{s \in [0,t]} \overset{law}{=} \{-u(t - s, \cdot)\}_{s \in [0,t]}.$$

The above relation (4.18) only holds for u, but not for u^ε. Nevertheless, this inspires us to define another diffusion driven by a time reversed drift and connect g^ε in (4.17) to its Fokker-Planck equation.

Before that, we first note that the solution to (4.17) is periodic in space, so we view it as the Fokker-Planck equation satisfied by the density of the following diffusion in random environment, with $\mathbb{T}$ being the state space:

$$(4.19) \qquad d\mathcal{Y}_s^\varepsilon = -u^\varepsilon(s, \mathcal{Y}_s^\varepsilon)ds + dB_s, \qquad \mathcal{Y}_0^\varepsilon \sim m_\mathbb{T},$$

where "$\mathcal{Y}_0^\varepsilon \sim m_\mathbb{T}$" means that $\mathcal{Y}_0^\varepsilon$ is sampled from the Lebesgue measure on $\mathbb{T}$, denoted by $m_\mathbb{T}$. It is clear that for any $s > 0$, $g^\varepsilon(s, \cdot)$ is the quenched density of $\mathcal{Y}_s^\varepsilon$ on $\mathbb{T}$,

$$g^\varepsilon(s, y) \geq 0 \quad \text{and} \quad \int_\mathbb{T} g^\varepsilon(s, y)dy = 1.$$

Inspired by the time reversal anti-symmetry in (4.18), we consider another diffusion in random environment, which is related to the directed polymer in light of Lemma 4.2: for fixed $t > 0$, let $\tilde{\mathcal{Y}}^\varepsilon$ solve

$$(4.20) \qquad d\tilde{\mathcal{Y}}_s^\varepsilon = u^\varepsilon(t - s, \tilde{\mathcal{Y}}_s^\varepsilon)ds + dB_s, \qquad \tilde{\mathcal{Y}}_0^\varepsilon \sim m_\mathbb{T}.$$

Let $\tilde{g}^\varepsilon(t; s, \cdot)$ be the quenched density of $\tilde{\mathcal{Y}}_s^\varepsilon$ (again on $\mathbb{T}$ rather than $\mathbb{R}$). Comparing $\tilde{\mathcal{Y}}_s^\varepsilon$ to $\mathcal{X}_s^\varepsilon$ defined in (4.8), there are two differences: (i) the distribution of the starting point, (ii) we have chosen $\mathbb{R}$ as the state space for $\mathcal{X}_s^\varepsilon$, while the state space for $\tilde{\mathcal{Y}}_s^\varepsilon$ is $\mathbb{T}$. Here $t > 0$ is fixed, and we have used the simplified notations $\mathcal{X}_s^\varepsilon, \tilde{\mathcal{Y}}_s^\varepsilon$, where the dependence on t was again omitted.

We have the following lemma which relates $\tilde{g}^\varepsilon$ to the midpoint density of a directed polymer:

Lemma 4.6. *For any $s \in (0, t]$ and $y \in \mathbb{T}$, we have*

$$\begin{aligned} \tilde{g}^\varepsilon(t; s, y) &= \int_\mathbb{T} \rho_\mathrm{m}^\varepsilon(t, x | t - s, y)dx \\ &= \int_\mathbb{T} \left(\frac{\mathcal{G}_{t,t-s}^\varepsilon(x, y)\int_\mathbb{T}\mathcal{G}_{t-s,0}^\varepsilon(y, x')e^{\beta W(x')}dx'}{\int_\mathbb{T}\mathcal{G}_{t,0}^\varepsilon(x, x')e^{\beta W(x')}dx'} \right)dx, \end{aligned}$$

where $\rho_\mathrm{m}^\varepsilon$ was given in (4.5).

Proof. By Lemma 4.2, we know that if $\tilde{\mathcal{Y}}_0^\varepsilon = x$, then the quenched density of $\tilde{\mathcal{Y}}_s^\varepsilon$ is given by $\rho_\mathrm{m}^\varepsilon(t, x | t - s, y)$ and this actually completes the proof since $\tilde{\mathcal{Y}}_0^\varepsilon$ is sampled from the Lebesgue measure on $\mathbb{T}$. □

We are particularly interested in $\tilde{g}^\varepsilon(t; t, \cdot)$, which takes the form

$$\tilde{g}^\varepsilon(t; t, y) = \int_{\mathbb{T}} \left(\frac{\mathcal{G}_{t,0}^\varepsilon(x, y) e^{\beta W(y)}}{\int_{\mathbb{T}} \mathcal{G}_{t,0}^\varepsilon(x, x') e^{\beta W(x')} dx'} \right) dx, \qquad y \in \mathbb{T}.$$

Define the random density

$$(4.21) \qquad \tilde{g}(t, y) = \int_{\mathbb{T}} \left(\frac{\mathcal{G}_{t,0}(x, y) e^{\beta W(y)}}{\int_{\mathbb{T}} \mathcal{G}_{t,0}(x, x') e^{\beta W(x')} dx'} \right) dx, \qquad y \in \mathbb{T}.$$

The following result is rather standard, and the proof follows an approximation argument similar to that of [11, Theorem 2.2].

Lemma 4.7. *For any $t > 0$, we have*

$$\tilde{g}^\varepsilon(t; t, \cdot) \to \tilde{g}(t, \cdot), \quad as \quad \varepsilon \to 0$$

in $C(\mathbb{T})$ in probability.

Define

$$(4.22) \qquad \tilde{\phi}(t, y) = \int_0^y [\tilde{g}(t, y') - 1] dy', \qquad y \in \mathbb{T}.$$

Here is the main result of this section (recall that $\nabla h^\varepsilon = u^\varepsilon$ and $\nabla \phi^\varepsilon = g^\varepsilon - 1$).

Proposition 4.8. *Fix any $t > 0$, as $\varepsilon \to 0$, we have the following weak convergence on $C(\mathbb{T}) \times C(\mathbb{T})$:*

$$(4.23) \qquad (h^\varepsilon(t, \cdot) - h^\varepsilon(t, 0), \phi^\varepsilon(t, \cdot) - \phi^\varepsilon(t, 0)) \Rightarrow (-\beta W(\cdot), \tilde{\phi}(t, \cdot)).$$

As a result,

$$(4.24) \qquad (h(t, \cdot) - h(t, 0), \phi(t, \cdot) - \phi(t, 0)) \overset{law}{=} (-\beta W(\cdot), \tilde{\phi}(t, \cdot)).$$

The above result comes from the time reversal anti-symmetry (4.18), which only holds in the limiting case of $\varepsilon = 0$. To understand where (4.24) comes from, let us pretend the anti-symmetry also holds for $\varepsilon > 0$: if $\{u^\varepsilon(s, \cdot)\}_{s \in [0,t]} \overset{law}{=} \{-u^\varepsilon(t - s, \cdot)\}_{s \in [0,t]}$, then by comparing the two diffusions described by (4.19) and (4.20) respectively, it is straightforward to conclude that $(u^\varepsilon(t, \cdot), g^\varepsilon(t, \cdot)) \overset{law}{=} (-u^\varepsilon(0, \cdot), \tilde{g}^\varepsilon(t; t, \cdot))$, which can be seen as an ε-version of (4.24). Although this argument does not apply to $\varepsilon > 0$, as the time reversal anti-symmetry for the Burgers solution holds only in the limiting case of $\varepsilon = 0$, it is natural to expect (4.23) to hold. The proof of Proposition 4.8 uses the well-developed results for singular diffusions with distribution-valued drifts, which we prove in the next section.

Remark 4.9. The relation (4.24) gives the explicit joint law of $(u(t, \cdot), \nabla \phi(t, \cdot))$. On the formal level, $g = 1 + \nabla \phi$ solves the Fokker-Planck equation $\partial_t g = \frac{1}{2} \Delta g + \nabla(ug)$, so if we consider the singular diffusion $d\mathcal{Y}_s = -u(s, \mathcal{Y}_s) ds + dB_s$, it actually leads to the explicit invariant measure of the process of "the environment seen from the particle".

4.3. **Singular diffusion and proof of Proposition 4.8.** Fix $t > 0$ in this section. Recall that

$$\phi^\varepsilon(t, x) = \frac{\int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x')(x' - x) e^{\beta W(x')} dx'}{\int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x') e^{\beta W(x')} dx'}$$

$$= \frac{\int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x') x' e^{\beta W(x')} dx'}{\int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x') e^{\beta W(x')} dx'} - x,$$

$$h^\varepsilon(t, x) = \log \int_{\mathbb{R}} \mathcal{Z}_{t,0}^\varepsilon(x, x') e^{\beta W(x')} dx',$$

and ϕ, h are of the same form as $\phi^\varepsilon, h^\varepsilon$, with $\mathcal{Z}^\varepsilon$ replaced by $\mathcal{Z}$. Therefore, $\phi^\varepsilon, h^\varepsilon$ are both written as smooth functionals of functions of the form $\int_\mathbb{R} \mathcal{Z}^\varepsilon_{t,0}(x, x') f(x') dx'$ for some continuous function f, namely, the solution to SHE driven by ξ^ε started from f. Thus, by a standard argument (see e.g. [11, Theorem 2.2]) we know that, for each $t > 0$,

$$(h^\varepsilon(t, \cdot), \phi^\varepsilon(t, \cdot)) \Rightarrow (h(t, \cdot), \phi(t, \cdot)),$$

weakly in $C(\mathbb{T}) \times C(\mathbb{T})$.

With the above weak convergence, to prove Proposition 4.8, it suffices to show the convergence of finite dimensional distributions in (4.23). In other words, take any $n \geq 1$ and $x_1, \ldots, x_n \in \mathbb{T}$, we aim at proving the convergence in distribution of

$$(4.25) \qquad (X_1^\varepsilon, \ldots, X_n^\varepsilon, Y_1^\varepsilon, \ldots, Y_n^\varepsilon) \Rightarrow (X_1, \ldots, X_n, Y_1, \ldots, Y_n)$$

as $\varepsilon \to 0$, with

$$(4.26) \qquad \begin{aligned} X_i^\varepsilon &= h^\varepsilon(t, x_i) - h^\varepsilon(t, 0), & X_i &= -\beta W(x_i), \\ Y_i^\varepsilon &= \phi^\varepsilon(t, x_i) - \phi^\varepsilon(t, 0), & Y_i &= \tilde{\phi}(t, x_i). \end{aligned}$$

Before going to the proof of (4.25), we first introduce a few more notations. Denote $\zeta = (\xi, W)$ which represents the underlying random realization. For each realization of ζ (hence each realization of the drift u^ε), let $\mathbb{Q}_t^\varepsilon, \tilde{\mathbb{Q}}_t^\varepsilon$ be the probability measures on $C([0, t], \mathbb{T})$ induced by $\{\mathcal{Y}_s^\varepsilon\}_{s \in [0,t]}, \{\tilde{\mathcal{Y}}_s^\varepsilon\}_{s \in [0,t]}$ respectively. Namely, $\mathbb{Q}_t^\varepsilon, \tilde{\mathbb{Q}}_t^\varepsilon$ are the quenched probability measures for the diffusions in random environments, and they depend on $\varepsilon > 0$ and also on the random realization ζ. Let $\mathcal{M}_1(C([0, t], \mathbb{T}))$ denote the set of probability measures on $C([0, t], \mathbb{T})$, equipped with the topology of weak convergence. One can view $\mathbb{Q}_t^\varepsilon, \tilde{\mathbb{Q}}_t^\varepsilon$ as random variables taking values in $\mathcal{M}_1(C([0, t], \mathbb{T}))$. Define

$$(4.27) \qquad \mathscr{U}_t = \{-u(s, x)\}_{s \in [0,t], x \in \mathbb{T}}, \qquad \tilde{\mathscr{U}}_t = \{u(t - s, x)\}_{s \in [0,t], x \in \mathbb{T}},$$

which represent the drift for the two diffusions in random environment (in the limit of $\varepsilon = 0$), and we view them as random variables taking values $C([0, t], C^{-\alpha}(\mathbb{T}))$ for some $\alpha > 1/2$.

Next, we state a result which is borrowed from the singular diffusion literature and will play a crucial role in our analysis.

Proposition 4.10. *As $\varepsilon \to 0$, $\mathbb{Q}_t^\varepsilon$ and $\tilde{\mathbb{Q}}_t^\varepsilon$ converge weakly almost surely. In other words, there exist $\mathbb{Q}_t, \tilde{\mathbb{Q}}_t$, which are random variables taking values in $\mathcal{M}_1(C([0, t], \mathbb{T}))$, such that for almost every ζ, we have $\mathbb{Q}_t^\varepsilon \Rightarrow \mathbb{Q}_t$ and $\tilde{\mathbb{Q}}_t^\varepsilon \Rightarrow \tilde{\mathbb{Q}}_t$ as $\varepsilon \to 0$, where "$\Rightarrow$" represents the weak convergence of measures. In addition, we have*

$$(4.28) \qquad (\mathscr{U}_t, \mathbb{Q}_t) \overset{law}{=} (\tilde{\mathscr{U}}_t, \tilde{\mathbb{Q}}_t).$$

The proof of Proposition 4.10 will be presented in Section B of the Appendix. We will first use it to complete the proof of (4.25), which is divided into several steps.

Proof of (4.25). Throughout the proof, fix $n \geq 1$ and $x_1, \ldots, x_n \in \mathbb{T}$. Let $\{\omega_s\}_{s \in [0,t]}$ be the canonical process of $C([0, t], \mathbb{T})$.

Step 1. The goal is to show that under $\tilde{\mathbb{Q}}_t$, the random variable ω_t has the density given by $\tilde{g}(t, \cdot)$ defined in (4.21): for any $a < b \in \mathbb{T}$, we have

$$(4.29) \qquad \tilde{\mathbb{Q}}_t(\omega_t \in [a, b]) = \int_a^b \tilde{g}(t, x) dx.$$

First, by definition we have $\tilde{\mathbb{Q}}_t^\varepsilon(\omega_t \in [a, b]) = \int_a^b \tilde{g}^\varepsilon(t; t, x) dx$. Then, by Lemma 4.7, $\tilde{g}^\varepsilon(t; t, \cdot) \to \tilde{g}(t, \cdot)$ in $C(\mathbb{T})$ in probability. Thus, upon extracting a subsequence, we have $\tilde{g}^\varepsilon(t; t, \cdot) \to \tilde{g}(t, \cdot)$ in $C(\mathbb{T})$ almost surely. By Proposition 4.10, we have $\tilde{\mathbb{Q}}_t^\varepsilon \Rightarrow \tilde{\mathbb{Q}}_t$ almost surely, which implies that, ω_t under $\tilde{\mathbb{Q}}_t^\varepsilon$ converges in distribution to ω_t under $\tilde{\mathbb{Q}}_t$ (again, for

almost every realization of ζ). By the convergence of the density $\tilde{g}^\varepsilon(t; t, \cdot)$ to the density $\tilde{g}(t, \cdot)$, this further implies that ω_t under $\tilde{\mathbb{Q}}_t$ has density given by $\tilde{g}(t, \cdot)$, hence completing the proof of (4.29).

Step 2. The goal is to show that for any $i = 1, \ldots, n$,

$$
\begin{aligned}
(X_i^\varepsilon, Y_i^\varepsilon) &= (h^\varepsilon(t, x_i) - h^\varepsilon(t, 0), \phi^\varepsilon(t, x_i) - \phi^\varepsilon(t, 0)) \\
&\to (h(t, x_i) - h(t, 0), \mathbb{Q}_t(\omega_t \in [0, x_i]) - x_i) =: (\bar{X}_i, \bar{Y}_i)
\end{aligned}
\tag{4.30}
$$

in probability. Since $g^\varepsilon = 1 + \nabla \phi^\varepsilon$, we first write

$$
\phi^\varepsilon(t, x_i) - \phi^\varepsilon(t, 0) = \int_0^{x_i} g^\varepsilon(t, y) dy - x_i.
$$

Recall that g^ε solves the Fokker-Planck equation (4.17), with $\mathcal{Y}_s^\varepsilon$ the underlying diffusion given by (4.19). By the definition of $\mathbb{Q}_t^\varepsilon$, we can write the integral on the r.h.s. of the above equation as

$$
\int_0^{x_i} g^\varepsilon(t, y) dy = \mathbb{Q}_t^\varepsilon(\omega_t \in [0, x_i]).
$$

From (4.29), we know that, for any $x \in \mathbb{T}$,

$$
\tilde{\mathbb{Q}}_t(\omega_t = x) = 0,
\tag{4.31}
$$

which implies that $\mathbb{Q}_t(\omega_t = x) = 0$, because we have $\mathbb{Q}_t \overset{\text{law}}{=} \tilde{\mathbb{Q}}_t$ by Proposition 4.10. Thus, by the convergence of $\mathbb{Q}_t^\varepsilon \Rightarrow \mathbb{Q}_t$, we can pass to the limit and obtain that, almost surely,

$$
\begin{aligned}
\phi^\varepsilon(t, x_i) - \phi^\varepsilon(t, 0) &= \mathbb{Q}_t^\varepsilon(\omega_t \in [0, x_i]) - x_i \\
&\to \mathbb{Q}_t(\omega_t \in [0, x_i]) - x_i.
\end{aligned}
$$

This completes the proof of (4.30).

Step 3. The goal is to show that

$$
(\bar{X}_1, \ldots, \bar{X}_n, \bar{Y}_1, \ldots, \bar{Y}_n) \overset{\text{law}}{=} (X_1, \ldots, X_n, Y_1, \ldots, Y_n),
\tag{4.32}
$$

which is the last piece we need to complete the proof of (4.25). To show (4.32), we first rewrite X_i, Y_i as (by (4.26) and (4.22))

$$
\begin{aligned}
X_i &= -\beta W(x_i) = -\int_0^{x_i} u(0, y) dy, \\
Y_i &= \tilde{\phi}(t, x_i) = \int_0^{x_i} \tilde{g}(t, y) dy - x_i = \tilde{\mathbb{Q}}_t(\omega_t \in [0, x_i]) - x_i,
\end{aligned}
\tag{4.33}
$$

where the last "=" comes from (4.29). Then we recall the definition of $\bar{X}_i, \bar{Y}_i$:

$$
\begin{aligned}
\bar{X}_i &= h(t, x_i) - h(t, 0) = \int_0^{x_i} u(t, y) dy, \\
\bar{Y}_i &= \mathbb{Q}_t(\omega_t \in [0, x_i]) - x_i.
\end{aligned}
\tag{4.34}
$$

If we compare (4.33) and (4.34), then (4.32) is a direct consequence of (4.28). □

5. A VARIANCE FORMULA

In this section, we combine the results obtained in the previous sections to derive a formula of $\sigma^2(\beta)$. Recall that in Section 2, we have obtained the following expression

$$
X_t = \int_{\mathbb{R}} x\rho(t, x) dx = \beta \int_0^t \int_{\mathbb{T}} \mathbf{E}[I_{1,t}(s, y) - I_{2,t}(s, y) | \mathcal{F}_s] \xi(s, y) dy ds,
\tag{5.1}
$$

with

$$
I_{1,t}(s, y) - I_{2,t}(s, y) = \rho_{\mathrm{m}}(t, -|s, y) \int_{\mathbb{T}} [\psi(s, y') - \psi(s, y)] \rho_{\mathrm{m}}(t, -|s, y') dy',
\tag{5.2}
$$

and $\rho_{\mathrm{m}}(t, -|s, y)$ and $\psi(s, y)$ given by, cf (2.21) and (2.22),

$$
(5.3) \quad
\begin{aligned}
\rho_{\mathrm{m}}(t, -|s, y) &= \frac{\mathcal{G}_{t,s}(-, y)\mathcal{G}_{s,0}(y, 0)}{\mathcal{G}_{t,0}(-, 0)} = \frac{\rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y)\rho_{\mathrm{f}}(s, y; 0, 0)}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y')\rho_{\mathrm{f}}(s, y'; 0, 0)dy'}, \\
\psi(s, y) &= \frac{\sum_{n \in \mathbb{Z}}(n - y)\mathcal{Z}_{s,0}(y, n)}{\mathcal{G}_{s,0}(y, 0)}.
\end{aligned}
$$

Here we recall that $m_{\mathbb{T}}$ denotes the Lebesgue measure on $\mathbb{T}$. Obviously, the random function $\psi(s, \cdot)$ is $\mathcal{F}_s$–measurable. The factor $\rho_{\mathrm{m}}(t, -|s, y)$ depends on both the forward and backward polymer endpoint densities, and it is clear that $\rho_{\mathrm{f}}(s, y; 0, 0)$ is $\mathcal{F}_s$–measurable and $\rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y)$ is independent of $\mathcal{F}_s$. In other words, to compute the conditional expectation on the r.h.s. of (5.1), we only need to average out the factor ρ_{b}.

This section contains a series of (rather standard) technical estimates which enable us to replace $\rho_{\mathrm{m}}(t, -|s, y)$ and $\psi(s, y)$ by their "equilibrium" versions. We briefly sketch them below so that the readers could have a big picture before going to the details. First, for $s \gg 1$ and $t - s \gg 1$, by the exponential mixing of the polymer endpoint density, one can replace the ρ_{f} and ρ_{b} in (5.3) by their stationary counterparts, and the ρ_{f} factor will only depends on $h(s, \cdot) - h(s, 0)$, where we recall that h solves the KPZ equation started at stationarity. Secondly, by Proposition 3.2, one can replace $\psi(s, y') - \psi(s, y)$ by $\phi(s, y') - \phi(s, y)$. After these approximations, what will appear in the Clark-Ocone representation only involves $\nabla h(s, \cdot)$ and $\nabla \phi(s, \cdot)$, since we will first average out the ρ_{b} factor. Now we make use of the key identity (4.24) to compute the expectation of the resulting functional of $(\nabla h(s, \cdot), \nabla \phi(s, \cdot))$, and this leads to an explicit expression of $\sigma^2(\beta)$.

5.1. Stationary approximation.

Recall that h was defined in (4.16), which is the solution to the KPZ equation started from stationary initial data βW. As we will have multiple independent Brownian bridges later, for notational convenience, we let $W_1 = W$. Define the "approximate" midpoint density (as an approximation of $\rho_{\mathrm{m}}(t, -|s, y)$)

$$
(5.4) \quad
\begin{aligned}
\rho_{\mathrm{app,m}}(t, -|s, y) &= \frac{\rho_{\mathrm{b}}(t, \rho_2; s, y)e^{h(s, y)}}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(t, \rho_2; s, y')e^{h(s, y')}dy'} \\
&= \frac{\rho_{\mathrm{b}}(t, \rho_2; s, y)\rho_{\mathrm{f}}(s, y; 0, \rho_1)}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(t, \rho_2; s, y')\rho_{\mathrm{f}}(s, y'; 0, \rho_1)dy'},
\end{aligned}
$$

with

$$
(5.5) \quad \rho_i(y) = \frac{e^{\beta W_i(y)}}{\int_{\mathbb{T}} e^{\beta W_i(y')}dy'}, \quad i = 1, 2,
$$

and W_2 is a Brownian bridge independent of ξ and W_1. Compare the expressions of $\rho_{\mathrm{app,m}}$ and ρ_{m}, the difference lies in the prescribed distribution of the starting and ending points. By the invariance of the law of ρ_1, ρ_2 under the forward/backward polymer endpoint evolution, we actually have that for each $t \geq s > 0$,

$$
(5.6) \quad \{\rho_{\mathrm{app,m}}(t, -|s, y)\}_{y \in \mathbb{T}} \overset{\text{law}}{=} \left\{ \frac{e^{\beta(W_1(y)+W_2(y))}}{\int_{\mathbb{T}} e^{\beta(W_1(y')+W_2(y'))}dy'} \right\}_{y \in \mathbb{T}},
$$

where, by the elementary properties of the Brownian bridge, the r.h.s. is statistically invariant under rotation on a torus (see e.g. [28, Lemma 4.2]).

For $s \gg 1$ and $t - s \gg 1$, by the exponential mixing of the endpoint distribution of directed polymers, see [34, Theorem 2.3], one may expect that $\rho_{\mathrm{m}}(t, -|s, y)$, see (5.3), is close to $\rho_{\mathrm{app,m}}(t, -|s, y)$. This is the contents of the following lemma.

Lemma 5.1. *For any $p \in [1, \infty)$, there exist $C, \lambda > 0$ such that*

$$
(5.7) \quad \sup_{y \in \mathbb{T}} \mathbb{E}_{W_1}\mathbb{E}_{W_2}\mathbf{E}|\rho_{\mathrm{m}}(t, -|s, y) - \rho_{\mathrm{app,m}}(t, -|s, y)|^p \leq C(e^{-\lambda(t-s)} + e^{-\lambda s})
$$

for all $t \geq 100$ and $s \in [1, t-1]$.

The proof is rather standard given the exponential mixing property formulated in Proposition A.1 below. We postpone it till the Appendix A.2.

Define

$$(5.8) \qquad J_t(s, y) := \rho_{\mathrm{app,m}}(t, -|s, y) \int_{\mathbb{T}} [\phi(s, y') - \phi(s, y)] \rho_{\mathrm{app,m}}(t, -|s, y') dy',$$

and

$$(5.9) \qquad Y_t := \beta \int_0^t \int_{\mathbb{T}} \mathbb{E}_{W_2} \mathbf{E}[J_t(s, y) | \mathcal{F}_s] \xi(s, y) dy ds.$$

The $J_t(s, y)$ should be viewed as a "stationary" approximation of $I_{1,t}(s, y) - I_{2,t}(s, y)$, with ρ_{m} replaced by $\rho_{\mathrm{app,m}}$ and ψ by ϕ. We emphasize that Y_t only depends on the noise ξ and the Brownian bridge W_1. The Itô integral with respect to ξ in (5.9) is performed for a fixed realization of W_1.

The main result of the present section is the following lemma, which implies that $\frac{X_t - Y_t}{\sqrt{t}} \to 0$ as $t \to \infty$:

Lemma 5.2. *We have* $\sup_{t \geq 100} \mathbb{E}_{W_1} \mathbf{E} |X_t - Y_t|^2 < \infty$.

Proof. First, we note that X_t does not depend on W_1. Fix any $t \geq 100$. The proof consists of several steps.

Step 1, uniform bounds. By (3.5) and Corollary 3.9, we have for any $p \geq 1$, there exits a constant $C > 0$ such that

$$(5.10) \qquad \sup_{s \geq 0, y, y' \in \mathbb{T}} \{\mathbf{E} |\psi(s, y') - \psi(s, y)|^p + \mathbb{E}_{W_1} \mathbf{E} |\phi(s, y') - \phi(s, y)|^p\} \leq C.$$

All constants C appearing below may depend on $p \in [1, +\infty)$, but are independent of other parameters such as t, s etc. In addition, by (5.6), we have

$$(5.11) \qquad \sup_{s \in [0, t], y \in \mathbb{T}} \mathbb{E}_{W_1} \mathbb{E}_{W_2} \mathbf{E} \rho_{\mathrm{app,m}}(t, -|s, y)^p \leq C.$$

Combining estimates (5.10) and (5.11), we conclude that

$$(5.12) \qquad \sup_{s \in (0, t], y \in \mathbb{T}} \mathbb{E}_{W_1} \mathbb{E}_{W_2} \mathbf{E} |J_t(s, y)|^p \leq C.$$

Next, we derive an estimate on $I_{1,t} - I_{2,t}$. First, for $\rho_{\mathrm{m}}(t, -|s, y)$, starting from the expression in (5.3), by the Hölder estimate together with (A.4) and (A.1), we conclude that

$$\mathbf{E} \rho_{\mathrm{m}}(t, -|s, y)^p \leq C, \qquad \text{for } s \in [1, t-1], y \in \mathbb{T}.$$

For other values of s, we will use the following estimates: (i) for $s \in (t-1, t]$, we have

$$(5.13) \qquad \sup_{s \in (t-1, t], y \in \mathbb{T}} \mathbf{E}[\rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y)^p + \rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y)^{-p}] \leq C,$$

and (ii) for $s \in (0, 1)$, we have

$$(5.14) \qquad \mathbf{E}[\rho_{\mathrm{f}}(s, y; 0, 0)^p] \leq C G_s(y)^p, \qquad \mathbf{E}[\rho_{\mathrm{f}}(s, y; 0, 0)^{-p}] \leq C G_s(y)^{-p}.$$

Here we recall that G is the heat kernel on torus.

Suppose now that $s \in [0, 1)$. Then we have

$$\rho_{\mathrm{m}}(t, -|s, y) \leq \rho_{\mathrm{f}}(s, y; 0, 0) \sup_{y' \in \mathbb{T}} \rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y') \sup_{y' \in \mathbb{T}} \rho_{\mathrm{b}}^{-1}(t, m_{\mathbb{T}}; s, y').$$

On the other hand, when $s \in (t-1, t]$, then

$$\rho_{\mathrm{m}}(t, -|s, y) \leq \rho_{\mathrm{f}}(s, y; 0, 0) \sup_{y' \in \mathbb{T}} \rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y') \sup_{y' \in \mathbb{T}} \rho_{\mathrm{f}}^{-1}(s, y'; 0, 0).$$

Using again Hölder inequality together with (5.13) and (5.14), we derive that

$$(5.15) \qquad \mathbf{E}\rho_{\mathrm{m}}(t,-|s,y)^p \le C\big(G_s(y)^p 1_{s\in(0,1)} + 1_{s\in(1,t]}\big), \qquad 0 < s \le t, \, y \in \mathbb{T}.$$

Hence, from (5.2), (5.10) and (5.15), we conclude that, for all $s \in (0,t], y \in \mathbb{T}$,

$$(5.16) \qquad \mathbf{E}|I_{1,t}(s,y) - I_{2,t}(s,y)|^p \le C\big(G_s(y)^p 1_{s\in(0,1)} + 1_{s\in(1,t]}\big).$$

Step 2, approximation. We claim that there exist $C, \lambda > 0$ such that

$$(5.17) \qquad \sup_{y\in\mathbb{T}} \mathbb{E}_{W_1}\mathbb{E}_{W_2}\mathbf{E}|I_{1,t}(s,y) - I_{2,t}(s,y) - J_t(s,y)|^p \le C\big(e^{-\lambda s} + e^{-\lambda(t-s)}\big)$$

for all $s \in [1, t-1]$.

Indeed, using (5.2) and (5.8), we can decompose the difference as

$$I_{1,t}(s,y) - I_{2,t}(s,y) - J_t(s,y) = \sum_{i=1}^3 E_i,$$

with

$$E_1 = \rho_{\mathrm{m}}(t,-|s,y) \int_{\mathbb{T}} [\psi(s,y') - \psi(s,y) - \phi(s,y') + \phi(s,y)]\rho_{\mathrm{m}}(t,-|s,y')dy',$$

$$E_2 = (\rho_{\mathrm{m}}(t,-|s,y) - \rho_{\mathrm{app,m}}(t,-;s,y)) \int_{\mathbb{T}} [\phi(s,y') - \phi(s,y)]\rho_{\mathrm{m}}(t,-|s,y')dy',$$

$$E_3 = \rho_{\mathrm{app,m}}(t,-|s,y) \int_{\mathbb{T}} [\phi(s,y') - \phi(s,y)](\rho_{\mathrm{m}}(t,-|s,y') - \rho_{\mathrm{app,m}}(t,-|s,y'))dy'.$$

Now we apply Proposition 3.2 and (5.15) to conclude that for $s \in [1, t-1]$,

$$\mathbb{E}_{W_1}\mathbf{E}|E_1|^p \le Ce^{-\lambda s},$$

From Lemma 5.1, (5.10) and (5.15), we obtain that

$$\mathbb{E}_{W_1}\mathbb{E}_{W_2}\mathbf{E}|E_2|^p \le C(e^{-\lambda(t-s)} + e^{-\lambda s}),$$

Finally from Lemma 5.1, (5.10) and (5.11), we get

$$\mathbb{E}_{W_1}\mathbb{E}_{W_2}\mathbf{E}|E_3|^p \le C(e^{-\lambda(t-s)} + e^{-\lambda s}).$$

This completes the proof of (5.17).

Step 3, conclusion. We have (see (5.1) and (5.9))

$$X_t - Y_t = \beta \int_0^t \int_{\mathbb{T}} \mathbb{E}_{W_2}\mathbf{E}[I_{1,t}(s,y) - I_{2,t}(s,y) - J_t(s,y)|\mathcal{F}_s]\xi(s,y)dyds.$$

For each realization of W_1, applying Itô isometry, we have

$$\mathbf{E}|X_t - Y_t|^2 = \beta^2 \int_0^t \int_{\mathbb{T}} \mathbf{E}\Big[\big(\mathbb{E}_{W_2}\mathbf{E}[I_{1,t}(s,y) - I_{2,t}(s,y) - J_t(s,y)|\mathcal{F}_s]\big)^2\Big] dyds.$$

Taking expectation on W_1 and using the Jensen inequality, we obtain

$$\mathbb{E}_{W_1}\mathbf{E}|X_t - Y_t|^2 \le \beta^2 \int_0^t \int_{\mathbb{T}} \mathbb{E}_{W_1}\mathbb{E}_{W_2}\mathbf{E}\big[|I_{1,t}(s,y) - I_{2,t}(s,y) - J_t(s,y)|^2\big]dyds.$$

Applying (5.12) and (5.16) to the integral in $s \in (0,1] \cup [t-1,t]$ and (5.17) to the integral in $s \in [1, t-1]$, we complete the proof. $\square$

5.2. **Itô isometry and time reversal.** By Lemma 5.2, to derive the limit of $t^{-1}\mathbf{E}X_t^2$, it suffices to do the same for $\mathbb{E}_{W_1}\mathbf{E}Y_t^2/t$. Recall that Y_t and $J_t(s,y)$ are given by (5.9) and (5.8) respectively and W_1, W_2 are two independent Brownian bridges satisfying $W_j(0) = W_j(1) = 0$, $j = 1, 2$.

Before applying Itô isometry to compute the variance, there is one more simplification we need to deal with the conditional expectation. Define

$$(5.18) \qquad \rho_W(s,y) := \frac{e^{\beta W_2(y)+h(s,y)}}{\int_{\mathbb{T}} e^{\beta W_2(y')+h(s,y')}dy'} = \frac{e^{\beta W_2(y)+h(s,y)-h(s,0)}}{\int_{\mathbb{T}} e^{\beta W_2(y')+h(s,y')-h(s,0)}dy'},$$

where we recall that $h(s,y) = \log Z_{W_1}(s,y)$ (cf (4.16)), and define

$$(5.19) \qquad J_W(s,y) := \rho_W(s,y) \int_{\mathbb{T}}[\phi(s,y') - \phi(s,y)]\rho_W(s,y')dy',$$

where ϕ is defined by (4.15):

$$\phi(t,x) = \frac{\int_{\mathbb{R}} \mathscr{Z}_{t,0}(x,x')(x'-x)e^{\beta W_1(x')}dx'}{\int_{\mathbb{R}} \mathscr{Z}_{t,0}(x,x')e^{\beta W_1(x')}dx'}.$$

We claim that

$$(5.20) \qquad \mathbb{E}_{W_2}\mathbf{E}[J_t(s,y)|\mathcal{F}_s] = \mathbb{E}_{W_2}J_W(s,y).$$

First, since both $h(s,\cdot)$ and $\phi(s,\cdot)$ are $\mathcal{F}_s$−measurable, we note that the right hand side is also $\mathcal{F}_s$−measurable. Secondly, recall that

$$\rho_{\mathrm{app,m}}(t,-|s,y) = \frac{\rho_{\mathrm{b}}(t,\rho_2;s,y)e^{h(s,y)}}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(t,\rho_2;s,y')e^{h(s,y')}dy'},$$

and we have, for any $t > s$ fixed,

$$\{\rho_{\mathrm{b}}(t,\rho_2;s,y)\}_{y\in\mathbb{T}} \stackrel{\mathrm{law}}{=} \left\{\frac{e^{\beta W_2(y)}}{\int_{\mathbb{T}} e^{\beta W_2(y')}dy'}\right\}_{y\in\mathbb{T}}.$$

Since $\rho_{\mathrm{b}}(t,\rho_2;s,y)$ depends only on the noise $\{\xi(\ell,x)\}_{(\ell,x)\in[s,t]\times\mathbb{R}}$ and the bridge W_2, we can write that

$$\mathbb{E}_{W_2}\mathbf{E}[\rho_{\mathrm{app,m}}(t,-|s,y)\rho_{\mathrm{app,m}}(t,-|s,y')|\mathcal{F}_s]$$
$$= \mathbb{E}_{W_2}[\rho_W(s,y)\rho_W(s,y')], \quad y,y' \in \mathbb{T}.$$

Due to $\mathcal{F}_s$−measurability of both $h(s,\cdot)$ and $\phi(s,\cdot)$, the above yields (5.20), which further implies

$$Y_t = \beta \int_0^t \int_{\mathbb{T}} \mathbb{E}_{W_2}[J_W(s,y)]\xi(s,y)dyds.$$

By Itô isometry (applied for each realization of W_1), we have

$$(5.21) \qquad \mathbb{E}_{W_1}\mathbf{E}Y_t^2 = \beta^2 \int_0^t \int_{\mathbb{T}} \mathbb{E}_{W_1}\mathbf{E}\left[\left(\mathbb{E}_{W_2}[J_W(s,y)]\right)^2\right]dyds.$$

From the expression of $J_W(s,y)$ in (5.19), we see that in order to compute the expectation on the r.h.s., one needs to understand the distribution of

$$\left(h(s,\cdot) - h(s,0), \phi(s,\cdot) - \phi(s,0)\right).$$

This is where the time reversal anti-symmetry (see (4.18)) and Proposition 4.8 come into play. Define

$$(5.22) \qquad \tilde{J}_W(s,y) := \tilde{\rho}_W(y) \int_{\mathbb{T}}[\tilde{\phi}(s,y') - \tilde{\phi}(s,y)]\tilde{\rho}_W(y')dy',$$

where $\tilde{\phi}$ was defined in (4.22) and

$$(5.23) \qquad \tilde{\rho}_W(y) := \frac{e^{\beta W_2(y)}e^{-\beta W_1(y)}}{\int_{\mathbb{T}} e^{\beta W_2(y')}e^{-\beta W_1(y')}dy'}.$$

We have the following key proposition:

Proposition 5.3. *For any $s > 0, y \in \mathbb{T}$, we have*

$$(5.24) \qquad J_W(s, y) \overset{\text{law}}{=} \tilde{J}_W(s, y).$$

In consequence,

$$(5.25) \qquad \mathbb{E}_{W_1} \mathbf{E}\big[\big(\mathbb{E}_{W_2}[J_W(s, y)] \big)^2 \big] = \mathbb{E}_{W_1} \mathbf{E}\big[\big(\mathbb{E}_{W_2}[\tilde{J}_W(s, y)] \big)^2 \big].$$

Proof. By (5.18) and (5.19), one can write

$$
\begin{aligned}
(5.26) \qquad J_W(s, y) =\, & \frac{e^{\beta W_2(y) + h(s,y) - h(s,0)}}{\int_{\mathbb{T}} e^{\beta W_2(y') + h(s,y') - h(s,0)} dy'} \\
& \times \int_{\mathbb{T}} [\phi(s, y') - \phi(s, y)] \frac{e^{\beta W_2(y') + h(s,y') - h(s,0)}}{\int_{\mathbb{T}} e^{\beta W_2(y'') + h(s,y'') - h(s,0)} dy''} dy'.
\end{aligned}
$$

By (4.24), we know that

$$(h(s, \cdot) - h(s, 0), \phi(s, \cdot) - \phi(s, 0)) \overset{\text{law}}{=} (-\beta W_1(\cdot), \tilde{\phi}(s, \cdot)).$$

If we replace $h(s, y) - h(s, 0)$ by $-\beta W_1(\cdot)$ and ϕ by $\tilde{\phi}$ in (5.26), we have $J_W(s, y)$ becomes $\tilde{J}_W(s, y)$. This completes the proof. □

5.3. **Stationary approximation again.**

In the expression of $\tilde{J}_W$ defined in (5.22), there is still an s−dependence through the function $\tilde{\phi}(s, \cdot)$. The last step is to use a stationary approximation of $\tilde{\phi}(s, \cdot)$ for large $s \gg 1$. By the definition of $\tilde{\phi}$ in (4.22), we have

$$\tilde{\phi}(s, y') - \tilde{\phi}(s, y) = \int_y^{y'} [\tilde{g}(s, z) - 1] dz,$$

with (cf (4.21))

$$\tilde{g}(s, z) = \int_{\mathbb{T}} \left(\frac{\rho_{\mathrm{b}}(s, x; 0, z) e^{\beta W_1(z)}}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(s, x; 0, z') e^{\beta W_1(z')} dz'} \right) dx, \qquad z \in \mathbb{T}.$$

For $s \gg 1$ and any $x \in \mathbb{T}$, we expect the backward polymer endpoint density $\rho_{\mathrm{b}}(s, x; 0, \cdot)$ to stabilize and reach stationarity.

This inspires us to define

$$(5.27) \qquad \tilde{g}_{\mathrm{app}}(s, z) := \frac{\rho_{\mathrm{b}}(s, \rho_3; 0, z) e^{\beta W_1(z)}}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(s, \rho_3; 0, z') e^{\beta W_1(z')} dz'}, \qquad s \geq 0, z \in \mathbb{T},$$

where

$$\rho_3(z) = \frac{e^{\beta W_3(z)}}{\int_{\mathbb{T}} e^{\beta W_3(z')} dz'},$$

and W_3 is another Brownian bridge independent of W_1, W_2, ξ.

By the same argument as used in the proof of Lemma 5.1, see Appendix A.2, we have: for any $p \in [1, \infty)$, there exist $C, \lambda > 0$ such that

$$(5.28) \qquad \sup_{z \in \mathbb{T}} \mathbb{E}_{W_1} \mathbb{E}_{W_3} \mathbf{E} |\tilde{g}(s, z) - \tilde{g}_{\mathrm{app}}(s, z)|^p \leq C e^{-\lambda s}, \qquad s \geq 1.$$

Again, by the invariance of the law of ρ_3 under the backward polymer endpoint evolution, it is clear that for each $s \geq 0$, we have

$$
\begin{aligned}
(5.29) \qquad & \left\{ \left(\tilde{g}_{\mathrm{app}}(s, z), \tilde{\rho}_W(z') \right) \right\}_{(z,z') \in \mathbb{T}^2} \overset{\text{law}}{=} \left\{ \left(\tilde{g}_{\mathrm{app}}(0, z), \tilde{\rho}_W(z') \right) \right\}_{(z,z') \in \mathbb{T}^2} \\
& = \left\{ \left(\frac{e^{\beta W_3(z) + \beta W_1(z)}}{\int_{\mathbb{T}} e^{\beta W_3(z'') + \beta W_1(z'')} dz''}, \frac{e^{\beta W_2(z')} e^{-\beta W_1(z')}}{\int_{\mathbb{T}} e^{\beta W_2(z'')} e^{-\beta W_1(z'')} dz''} \right) \right\}_{(z,z') \in \mathbb{T}^2}.
\end{aligned}
$$

Define

$$(5.30) \qquad \tilde{J}_{W,\mathrm{app}}(s,y) := \tilde{\rho}_W(y) \int_{\mathbb{T}} \left(\int_y^{y'} [\tilde{g}_{\mathrm{app}}(s,z)-1]dz \right) \tilde{\rho}_W(y')dy'.$$

By (5.29), we know that

$$(5.31) \qquad \left\{ \tilde{J}_{W,\mathrm{app}}(s,y) \right\}_{y \in \mathbb{T}} \overset{\mathrm{law}}{=} \left\{ \tilde{J}_{W,\mathrm{app}}(0,y) \right\}_{y \in \mathbb{T}},$$

and from the definitions of $\tilde{\rho}_W$ and $\tilde{g}_{\mathrm{app}}$, it is clear that $\tilde{J}_{W,\mathrm{app}}(0,\cdot)$ only involves the three independent Brownian bridges $\{W_i\}_{i=1,2,3}$.

The following lemma is the last piece we need to compute the variance:

Lemma 5.4. *There exist $C, \lambda > 0$ such that*

$$\sup_{y \in \mathbb{T}} \mathbb{E}_W \mathbf{E}\big[|\tilde{J}_W(s,y) - \tilde{J}_{W,\mathrm{app}}(s,y)|^2\big] \leq Ce^{-\lambda s}, \quad s \geq 1.$$

Proof. From the expressions of $\tilde{J}_W(s,y)$ and $\tilde{J}_{W,\mathrm{app}}(s,y)$ in (5.22) and (5.30), we have

$$\tilde{J}_W(s,y) - \tilde{J}_{W,\mathrm{app}}(s,y)$$
$$= \tilde{\rho}_W(y) \int_{\mathbb{T}} \left(\int_y^{y'} [\tilde{g}(s,z) - \tilde{g}_{\mathrm{app}}(s,z)]dz \right) \tilde{\rho}_W(y')dy'.$$

It suffices to apply (5.28) together with the Hölder inequality to complete the proof. □

A corollary of the above lemma is

Corollary 5.5. *We have*

$$(5.32) \qquad \lim_{t \to \infty} \frac{\mathbf{E}X_t^2}{t} = \lim_{t \to \infty} \frac{\mathbb{E}_{W_1}\mathbf{E}Y_t^2}{t} = \beta^2 \int_{\mathbb{T}} \mathbb{E}_{W_1} \mathbb{E}_{W_3} \Big[\big(\mathbb{E}_{W_2}[\tilde{J}_{W,\mathrm{app}}(0,y)] \big)^2 \Big] dy.$$

Proof. Recall that we already know from [33, Section 5] that $\mathbf{E}X_t^2/t$ converges as $t \to \infty$, so, the first equality in (5.32) is a consequence of Lemma 5.2. For the second equality, we have from (5.21) and (5.25)

$$\mathbb{E}_{W_1}\mathbf{E}Y_t^2 = \beta^2 \int_0^t \int_{\mathbb{T}} \mathbb{E}_{W_1}\mathbf{E}\big[\big(\mathbb{E}_{W_2}[J_W(s,y)] \big)^2 \big]dyds$$
$$= \beta^2 \int_0^t \int_{\mathbb{T}} \mathbb{E}_{W_1}\mathbf{E}\big[\big(\mathbb{E}_{W_2}[\tilde{J}_W(s,y)] \big)^2 \big]dyds.$$

Applying Lemma 5.4 and (5.31), we complete the proof. □

By (2.1), (2.5) and the above corollary, we conclude that for any $\beta > 0$ we have

$$(5.33) \qquad \sigma^2(\beta) = 1 + \beta^2 \int_{\mathbb{T}} \mathbb{E}_{W_1} \mathbb{E}_{W_3} \Big[\big(\mathbb{E}_{W_2}[\tilde{J}_{W,\mathrm{app}}(0,y)] \big)^2 \Big] dy.$$

To write the effective diffusivity more explicitly, we have the following lemma, using which we complete the proof of Theorem 1.1.

Lemma 5.6. *The process $\{\tilde{J}_{W,\mathrm{app}}(0,y)\}_{y \in \mathbb{T}}$ is stationary. As a result, we have*

$$(5.34) \qquad \sigma^2(\beta) = 1 + \beta^2 \mathbb{E}_{W_1} \mathbb{E}_{W_3} \Big[\big(\mathbb{E}_{W_2}[\tilde{J}_{W,\mathrm{app}}(0,0)] \big)^2 \Big],$$

where

$$\mathbb{E}_{W_2}\tilde{J}_{W,\mathrm{app}}(0,0) = \int_{\mathbb{T}^2} \Xi(\beta,y,W_1) \left(\frac{e^{\beta W_1(z)+\beta W_3(z)}}{\int_{\mathbb{T}} e^{\beta W_1(z')+\beta W_3(z')}dz'} - 1 \right) 1_{[0,y]}(z)dydz,$$

with

$$\Xi(\beta,y,W_1) = \mathbb{E}_{W_2} \left[\frac{e^{\beta W_2(y)-\beta W_1(y)}}{\left(\int_{\mathbb{T}} e^{\beta W_2(y')}e^{-\beta W_1(y')}dy' \right)^2} \right].$$

Proof. By (5.27), we know that $\tilde{g}_{\mathrm{app}}(s, \cdot)$ is a continuous density on $\mathbb{T}$, so

$$y' \mapsto \int_y^{y'} [\tilde{g}_{\mathrm{app}}(s, z) - 1] dz$$

is a 1-periodic function for each $y \in \mathbb{T}$. For any $x \in \mathbb{T}$, we can write

$$\tilde{J}_{W,\mathrm{app}}(0, y + x) = \tilde{\rho}_W(y + x) \int_{\mathbb{T}} \left(\int_{y+x}^{y'} [\tilde{g}_{\mathrm{app}}(0, z) - 1] dz \right) \tilde{\rho}_W(y') dy'$$

$$= \tilde{\rho}_W(y + x) \int_{\mathbb{T}} \left(\int_{y+x}^{y'+x} [\tilde{g}_{\mathrm{app}}(0, z) - 1] dz \right) \tilde{\rho}_W(y' + x) dy'$$

$$= \tilde{\rho}_W(y + x) \int_{\mathbb{T}} \left(\int_y^{y'} [\tilde{g}_{\mathrm{app}}(0, z + x) - 1] dz \right) \tilde{\rho}_W(y' + x) dy',$$

and the conclusion of the lemma follows from the joint stationarity of

$$\{(\tilde{g}_{\mathrm{app}}(0, z), \tilde{\rho}_W(z'))\}_{(z,z') \in \mathbb{T}^2},$$

see (5.29). $\qquad\square$

APPENDIX A. ENDPOINT DISTRIBUTION OF THE DIRECTED POLYMER ON A CYLINDER

A.1. **Exponential mixing of endpoint distribution.** Here we summarize a few results on the exponential mixing of the endpoint distribution of the directed polymer on a torus, which will be used throughout the paper. Recall that $\rho_{\mathrm{f}}, \rho_{\mathrm{b}}$ were defined in (2.7). By the time reversal invariance of the spacetime white noise, we have

(A.1) $$\{\rho_{\mathrm{f}}(t, x; s, \nu)\}_{x \in \mathbb{T}} \stackrel{\mathrm{law}}{=} \{\rho_{\mathrm{b}}(t, \nu; s, x)\}_{x \in \mathbb{T}}.$$

Here is the main result on ρ_{f}. By the above identity, the same conclusion applies to ρ_{b}.

Proposition A.1. *The Markov process $\{\rho_{\mathrm{f}}(t, \cdot; s, \nu)\}_{t \geq 0}$, taking values in $\mathcal{M}_1(\mathbb{T})$, has a unique invariant measure given by the law of the $\mathcal{M}_1(\mathbb{T})$-valued random variable $\rho(y) dy$, where*

(A.2) $$\rho(y) = \frac{e^{\beta W(y)}}{\int_{\mathbb{T}} e^{\beta W(y')} dy'},$$

and W is a Brownian bridge satisfying $W(0) = W(1) = 0$.

Furthermore, for any $p \geq 1$, there exist $C, \lambda > 0$ such that for all $t \geq 1$,

(A.3) $$\mathbf{E} \sup_{\nu, \nu' \in \mathcal{M}_1(\mathbb{T})} \sup_{x \in \mathbb{T}} |\rho_{\mathrm{f}}(t, x; 0, \nu) - \rho_{\mathrm{f}}(t, x; 0, \nu')|^p \leq C e^{-\lambda t},$$

and

(A.4) $$\mathbf{E} \sup_{\nu \in \mathcal{M}_1(\mathbb{T})} \sup_{x \in \mathbb{T}} \{\rho_{\mathrm{f}}(t, x; 0, \nu)^p + \rho_{\mathrm{f}}(t, x; 0, \nu)^{-p}\} \leq C.$$

Proof. The fact that $\{\rho_{\mathrm{f}}(t, \cdot; s, \nu)\}_{t \geq 0}$ is Markovian follows from [34, Lemma 2.2]. Estimate (A.3) is an immediate consequence of [34, Proposition 4.7].

To show (A.4), we note that

$$\rho_{\mathrm{f}}(t, x; 0, \nu) = \frac{\int_{\mathbb{T}} \mathcal{G}_{t, t-1/2}(x, x') \rho_{\mathrm{f}}(t - 1/2, x'; 0, \nu) dx'}{\int_{\mathbb{T}^2} \mathcal{G}_{t, t-1/2}(x'', x') \rho_{\mathrm{f}}(t - 1/2, x'; 0, \nu) dx' dx''}$$

$$\leq \sup_{x, x' \in \mathbb{T}} \mathcal{G}_{t, t-1/2}(x, x') \sup_{x', x'' \in \mathbb{T}} \mathcal{G}_{t, t-1/2}^{-1}(x'', x').$$

Estimate (A.4) then follows from an application of [34, Lemma 4.1] and the Hölder inequality. $\qquad\square$

A.2. **Proof of Lemma 5.1.** Recall that the goal was to show that for $s \in [1, t-1]$,

(A.5) $\sup_{y \in \mathbb{T}} \mathbb{E}_{W_1} \mathbb{E}_{W_2} \mathbf{E} |\rho_{\mathrm{m}}(t, -|s, y) - \rho_{\mathrm{app,m}}(t, -|s, y)|^p \leq C(e^{-\lambda(t-s)} + e^{-\lambda s})$,

with

$$\rho_{\mathrm{m}}(t, -|s, y) = \frac{\rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y)\rho_{\mathrm{f}}(s, y; 0, 0)}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y')\rho_{\mathrm{f}}(s, y'; 0, 0)dy'},$$

$$\rho_{\mathrm{app,m}}(t, -|s, y) = \frac{\rho_{\mathrm{b}}(t, \rho_2; s, y)\rho_{\mathrm{f}}(s, y; 0, \rho_1)}{\int_{\mathbb{T}} \rho_{\mathrm{b}}(t, \rho_2; s, y')\rho_{\mathrm{f}}(s, y'; 0, \rho_1)dy'}.$$

By (A.1) and (A.3), we have

$$\mathbb{E}_{W_2} \mathbf{E}\left[\sup_{y \in \mathbb{T}} |\rho_{\mathrm{b}}(t, m_{\mathbb{T}}; s, y) - \rho_{\mathrm{b}}(t, \rho_2; s, y)|^p\right]$$

$$+ \mathbb{E}_{W_1} \mathbf{E}\left[\sup_{y \in \mathbb{T}} |\rho_{\mathrm{f}}(s, y; 0, 0) - \rho_{\mathrm{f}}(s, y; 0, \rho_1)|^p\right] \leq C(e^{-\lambda(t-s)} + e^{-\lambda s}).$$

With the above estimate, the rest of the proof is rather standard, with several uses of Hölder inequality together with (A.4). We do not repeat it here.

APPENDIX B. SINGULAR DIFFUSION: PROOF OF PROPOSITION 4.10

Let us first recall the main ideas in Section 4 to see the role played by singular diffusion. The main argument in Section 4 was done on the level of approximation for each $\varepsilon > 0$ and only passed to the limit through Proposition 4.10 – in the following we will sketch formally the argument for $\varepsilon = 0$.

Recall that the goal was to describe the joint distribution of $h(t, \cdot) - h(t, 0)$ and $\phi(t, \cdot) - \phi(t, 0)$, where h solves the KPZ equation at stationarity, and $\phi(t, x)$, defined in (3.2), is the displacement of the polymer started at (t, x), running backward in time, with the terminal potential $e^{\beta W(\cdot)}$. By the Girsanov theorem and Feynman-Kac representation, one can interpret $\phi(t, x)$ as the quenched mean displacement of the diffusion in the Burgers drift:

$$\phi(t, x) = \mathbb{E}_B \mathcal{X}_t - x,$$

where $\{\mathcal{X}_s\}_{s \in [0, t]}$ solves the following formal SDE

(B.1) $d\mathcal{X}_s = u(t-s, \mathcal{X}_s)ds + dB_s$, $\mathcal{X}_0 = x$.

Note that (B.1) is only symbolic, since $u = \nabla h$ is the distributional-valued solution to the stochastic Burgers equation. In addition, one can show that ϕ solves the following PDE with distribution-valued coefficients

(B.2) $\partial_t \phi = \frac{1}{2}\Delta\phi + u\nabla\phi + u$, $\phi(0, x) = 0$,

which is also symbolic. To study the increment process of ϕ, it suffices to consider $g = 1 + \nabla\phi$, which solves the Fokker-Planck equation with a random distribution-valued coefficient

(B.3) $\partial_s g(s, y) = \frac{1}{2}\Delta_y g(s, y) + \nabla_y(u(s, y)g(s, y))$, $g(0, x) = 1$.

Since $\nabla h = u$, to study the joint distribution of $h(t, \cdot) - h(t, 0)$ and $\phi(t, \cdot) - \phi(t, 0)$, the problem reduces to studying the joint distribution of $u(t, \cdot)$ and $g(t, \cdot)$. The above Fokker-Planck equation is related to another diffusion in the Burgers flow:

(B.4) $d\mathcal{Y}_s = -u(s, \mathcal{Y}_s)ds + dB_s$, $\mathcal{Y}_0 \sim m_{\mathbb{T}}$.

Inspired by the time-reversal anti-symmetry of the Burgers flow (4.18), we consider $\tilde{\mathcal{Y}}_s$ given by

(B.5) $d\tilde{\mathcal{Y}}_s = u(t-s, \tilde{\mathcal{Y}}_s)ds + dB_s$, $\tilde{\mathcal{Y}}_0 \sim m_{\mathbb{T}}$,

and its density, denoted by $\tilde{g}(t; s, \cdot)$, evolves according to

(B.6) $\qquad \partial_s \tilde{g}(t; s, y) = \frac{1}{2}\Delta_y \tilde{g}(t; s, y) - \nabla_y(u(t - s, y)\tilde{g}(t; s, y)), \qquad \tilde{g}(t; 0, y) = 1.$

Compare the two (formal) Fokker-Planck equations (B.3) and (B.6), since the coefficients have the same law, it is natural to expect the solutions to have the same law, and, the joint distributions of the coefficient and the solution are the same as well. This turns out to be precisely the (4.28) in the statement of Proposition 4.10:

$$(\mathscr{U}_t, \mathbb{Q}_t) \overset{\text{law}}{=} (\tilde{\mathscr{U}}_t, \tilde{\mathbb{Q}}_t),$$

where $\mathscr{U}_t, \tilde{\mathscr{U}}_t$ are the corresponding coefficient processes, and $\mathbb{Q}_t, \tilde{\mathbb{Q}}_t$ are the probability measures corresponding to g and $\tilde{g}$.

The singular diffusions described by the formal SDE (B.1), (B.4) and (B.5) can all be made sense pathwisely, that is, for each realization of u, as the solutions to the corresponding martingale problems [17, 24], using the tools of rough path, regularity structures and paracontrolled calculus. The singular Fokker-Planck equations (B.3) and (B.6) can also be made sense pathwisely. As these are not the focus of the paper, we refrain from going to the details and only refer the readers to the aforementioned papers. Our proof of Proposition 4.10 is based on [24], which was the first one to give a rigorous meaning of the so-called continuum directed random polymer, introduced in [1], as a singular diffusion.

Proof of Proposition 4.10. Fix $t > 0$, recall that $\mathbb{Q}_t^\varepsilon, \tilde{\mathbb{Q}}_t^\varepsilon$ are the quenched probability measures on $C([0, t], \mathbb{T})$ of

$$d\mathcal{Y}_s^\varepsilon = -u^\varepsilon(s, \mathcal{Y}_s^\varepsilon)ds + dB_s, \qquad \mathcal{Y}_0^\varepsilon \sim m_\mathbb{T},$$
$$d\tilde{\mathcal{Y}}_s^\varepsilon = u^\varepsilon(t - s, \tilde{\mathcal{Y}}_s^\varepsilon)ds + dB_s, \qquad \tilde{\mathcal{Y}}_0^\varepsilon \sim m_\mathbb{T}.$$

For any $x \in \mathbb{T}$, we denote $\mathbb{Q}_{t,x}^\varepsilon, \tilde{\mathbb{Q}}_{t,x}^\varepsilon$ as the corresponding measures when the starting point is $\mathcal{Y}_0^\varepsilon = \tilde{\mathcal{Y}}_0^\varepsilon = x$. Applying [24, Theorem 31], we know that, for almost every realization ζ, $\tilde{\mathbb{Q}}_{t,x}^\varepsilon$ converges weakly[3], with the limit denoted by $\tilde{\mathbb{Q}}_{t,x}$. Furthermore, this can be done for all $x \in \mathbb{T}$ with the same ζ. As a result, for almost every realization ζ, we have the convergence of $\tilde{\mathbb{Q}}_t^\varepsilon = \int_\mathbb{T} \tilde{\mathbb{Q}}_{t,x}^\varepsilon dx$.

To show the convergence of $\mathbb{Q}_t^\varepsilon$, one could either follow the proof in [24], which is essentially a repetition, or follow the paracontrolled approach outlined in [17]. Since the proofs are almost the same, we do not do it here.

In the end, one needs to show the identity-in-law:

$$(\mathscr{U}_t, \mathbb{Q}_t) \overset{\text{law}}{=} (\tilde{\mathscr{U}}_t, \tilde{\mathbb{Q}}_t).$$

This immediately comes from the pathwise construction: $\mathbb{Q}_t$ and $\tilde{\mathbb{Q}}_t$ are constructed for almost every realization of $\mathscr{U}_t = \{-u(s, x)\}_{s \in [0,t], x \in \mathbb{T}}$ and $\tilde{\mathscr{U}}_t = \{u(t - s, x)\}_{s \in [0,t], x \in \mathbb{T}}$ respectively. In other words, they are deterministic functionals of $\mathscr{U}_t$ and $\tilde{\mathscr{U}}_t$ respectively. By the fact that $\mathscr{U}_t \overset{\text{law}}{=} \tilde{\mathscr{U}}_t$, we complete the proof. □

References

[1] T. ALBERTS, K. KHANIN AND J. QUASTEL, *The continuum directed random polymer*, Journal of Statistical Physics 154.1-2 (2014): 305-326. 37

[2] S. ARMSTRONG, A. BOU-RABEE AND T. KUUSI, *Superdiffusive central limit theorem for a Brownian particle in a critically-correlated incompressible random drift*, arXiv preprint arXiv:2404.01115 (2024). 5

[3] J. BAIK AND Z. LIU, *TASEP on a ring in sub-relaxation time scale*, J. Stat. Phys., 165 (2016), 1051–1085. 4

---

[3]The result in [24, Theorem 31] is only for the solution to the KPZ equation with constant initial data (rather than at stationarity), but the same proof applies verbatim to our setting.

[4] J. BAIK AND Z. LIU, *Fluctuations of TASEP on a ring in relaxation time scale*, Comm. Pure Appl. Math., 71 (2018), 747–813. 4

[5] J. BAIK AND Z. LIU, *Multipoint distribution of periodic TASEP*, J. Amer. Math. Soc., 32 (2019), 609–674. 4

[6] J. BAIK AND Z. LIU, *Periodic TASEP with general initial conditions*, Probab. Theory Related Fields, 179 (2021), 1047–1144. 4

[7] J. BAIK, Z. LIU, AND G. L. F. SILVA, *Limiting one-point distribution of periodic TASEP*, Aug. 2020, arXiv preprint 2008.07024v1. 4

[8] M. BALÁZS, J. QUASTEL, AND T. SEPPÄLÄINEN, *Fluctuation exponent of the KPZ/stochastic Burgers equation*, J. Amer. Math. Soc., 24 (2011), 683–708. 25

[9] G. BARRAQUAND, I. CORWIN AND Z. YANG, *Stationary measures for integrable polymers on a strip*, arXiv preprint arXiv:2306.05983 (2023). 4

[10] G. BARRAQUAND AND P. L. DOUSSAL, *Steady state of the KPZ equation on an interval and Liouville quantum mechanics*, Europhysics Letters 137.6 (2022): 61003. 4

[11] L. BERTINI AND N. CANCRINI, *The Stochastic Heat Equation: Feynman-Kac Formula and Intermittence*, J. Stat. Phys., 78 (1995), 1377–1401. 10, 26, 27

[12] É. BRUNET, *Fluctuations of the winding number of a directed polymer in a random medium*, Physical Review E 68.4 (2003): 041101. 2, 3, 8

[13] É. BRUNET AND B. DERRIDA, *Ground state energy of a non-integer number of particles with δ attractive interactions*, Phys. A, 279 (2000), 398–407. 2

[14] É. BRUNET AND B. DERRIDA, *Probability distribution of the free energy of a directed polymer in a random medium*, Physical Review E 61.6 (2000): 6789. 2

[15] É. BRUNET, Y. GU AND T. KOMOROWSKI, *High temperature behaviors of the directed polymer on a cylinder*, arxiv preprint arXiv:2110.07368v3. 1

[16] W. BRYC, A. KUZNETSOV, Y. WANG, AND J. WESOLOWSKI, *Markov processes related to the stationary measure for the open KPZ equation*, July 2021, arXiv preprint 2105.03946v2. 4

[17] G. CANNIZZARO AND K. CHOUK, *Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential*, The Annals of Probability 46.3 (2018): 1710-1763. 4, 37

[18] G. CANNIZZARO, L. HAUNSCHMID-SIBITZ AND F. TONINELLI, $\sqrt{\log t}$-*Superdiffusivity for a Brownian particle in the curl of the 2D GFF*, The Annals of Probability 50.6 (2022): 2475-2498. 5

[19] G. CHATZIGEORGIOU, P. MORFE, F. OTTO AND L. WANG, *The Gaussian free-field as a stream function: asymptotics of effective diffusivity in infra-red cut-off*, arXiv preprint arXiv:2212.14244. 5

[20] L. CHEN, D. KHOSHNEVISAN, D. NUALART, AND F. PU, *Spatial ergodicity for SPDEs via Poincaré-type inequalities* July 2019, arXiv preprint 1907.11553v1. 9

[21] I. CORWIN, *The Kardar-Parisi-Zhang equation and universality class*, Random Matrices Theory Appl., 1 (2012), 1130001, 76. 4

[22] I. CORWIN, *Some recent progress on the stationary measure for the open KPZ equation*, Toeplitz Operators and Random Matrices: In Memory of Harold Widom (2022): 321-360. 4

[23] I. CORWIN AND A. KNIZEL, *Stationary measure for the open KPZ equation*, Apr. 2021, arXiv preprint 2103.12253v2. 4

[24] F. DELARUE AND R. DIEL, *Rough paths and 1d SDE with a time dependent distributional drift: application to polymers*, Probability Theory and Related Fields 165.1-2 (2016): 1-63. 4, 37

[25] B. DERRIDA, M. R. EVANS AND D. MUKAMEL, *Exact diffusion constant for one-dimensional asymmetric exclusion models*, J. Phys. A, 26 (1993), 4911–4918. 4

[26] B. DERRIDA, M. R. EVANS AND K. MALLICK, *Exact diffusion constant of a one-dimensional asymmetric exclusion model with open boundaries*, Journal of statistical physics 79 (1995): 833-874. 4

[27] B. DERRIDA AND K. MALLICK, *Exact diffusion constant for the one-dimensional partially asymmetric exclusion model*, Journal of Physics A: mathematical and general 30.4 (1997): 1031. 4

[28] A. DUNLAP, Y. GU AND T. KOMOROWSKI, *Fluctuation exponents of the KPZ equation on a large torus*, *Comm. on Pure and Applied Math.*, 2023, https://doi.org/10.1002/cpa.22110 1, 3, 7, 8, 29

[29] A. DUNLAP, Y. GU AND L. LI, *Localization length of the 1+ 1 continuum directed random polymer*, Annales Henri Poincaré. Vol. 24. No. 7. Cham: Springer International Publishing, 2023. 6

[30] A. FANNJIANG, *Anomalous diffusion in random flows*, Mathematics of Multiscale Materials (1998): 81-99. 5

[31] B. FEHRMAN, *Stochastic homogenization with space-time ergodic divergence-free drift*, arXiv preprint arXiv:2207.14555 (2022). 6

[32] G. FELTES AND H. WEBER, *Brownian particle in the curl of 2-d stochastic heat equations*, arXiv preprint arXiv:2211.02194 (2022). 5

[33] Y. GU AND T. KOMOROWSKI, *Fluctuations of the winding number of a directed polymer on a cylinder*, to appear in *SIAM Journ. Math. Anal.* available at https://arxiv.org/abs/2207.14091. 1, 2, 4, 8, 10, 15, 18, 34

[34] Y. GU AND T. KOMOROWSKI, *KPZ on torus: Gaussian fluctuations*, June 2021, arXiv preprint 2104.13540v2, to appear in *Ann. Inst. H. Poincare, Prob. and Stat.* 1, 19, 20, 29, 35

[35] M. GUBINELLI, P. IMKELLER AND N. PERKOWSKI, *Paracontrolled distributions and singular PDEs*, Forum of Mathematics, Pi. Vol. 3. Cambridge University Press, 2015. 4

[36] M. HAIRER, *Solving the KPZ equation*, Annals of mathematics (2013): 559-664. 4

[37] M. HAIRER, *A theory of regularity structures*, Inventiones mathematicae 198.2 (2014): 269-504. 4

[38] Z. HAO, X. ZHANG, R. ZHU AND X. ZHU, *Singular kinetic equations and applications*, arXiv preprint arXiv:2108.05042. 4

[39] Y. HU AND K. LÊ, *Asymptotics of the density of parabolic Anderson random fields*, Annales de l'Institut Henri Poincare (B) Probabilites et statistiques. Vol. 58. No. 1. Institut Henri Poincaré, 2022. 18

[40] D. KHOSHNEVISAN, *Analysis of stochastic partial differential equations*, Vol. 119. American Mathematical Soc., 2014. 4

[41] T. KOMOROWSKI, C. LANDIM AND S. OLLA, *Fluctuations in Markov Processes. Time Symmetry and Martingale Approximation*, Springer Ser.: *Grundlehren der mathematischen Wissenschaften*, Vol. **345**, 2012. 5, 24

[42] H. KREMP AND N. PERKOWSKI, *Multidimensional SDE with distributional drift and Lévy noise*, Bernoulli 28.3 (2022): 1757-1783. 4

[43] C. LANDIM, J. QUASTEL, M. SALMHOFER AND H. T. YAU, *Superdiffusivity of asymmetric exclusion process in dimensions one and two*, Communications in mathematical physics, 244, (2004) 455-481. 5

[44] Z. LIU, *Height fluctuations of stationary TASEP on a ring in relaxation time scale*, Ann. Inst. H. Poincaré B, 54 (2018), 1031–1057. 4

[45] D. NUALART, *The Malliavin calculus and related topics*, Vol. 1995. Berlin: Springer, 2006. 9

[46] J. QUASTEL, *Introduction to KPZ*, in Current Developments in Mathematics, 2011, Int. Press, Somerville, MA, 2012, 125–194. 4

[47] J. QUASTEL AND H. SPOHN, *The one-dimensional KPZ equation and its universality class*, J. Stat. Phys., 160 (2015), 965–984. 4

[48] T. SEPPÄLÄINEN, *Existence, uniqueness and coalescence of directed planar geodesics: Proof via the increment-stationary growth process*, Ann. Inst. H. Poincare Probab. Statist. 56(3): 1775-1791, 2020. 7, 21

[49] H. T. YAU, $(\log t)^{2/3}$ *Law of the Two Dimensional Asymmetric Simple Exclusion Process*, Annals of mathematics (2004): 377-405. 5

(Yu Gu) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742, USA.

Email address: yugu105@gmail.com

(Tomasz Komorowski) INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES, UL. ŚNIADECKICH 8, 00-636 WARSAW, POLAND.

Email address: tkomorowski@impan.pl