A PDE hierarchy for directed polymers in random environments

Yu Gu1,* and Christopher Henderson2

1 Department of Mathematics, University of Maryland, College Park, MD 20742, United States of America
2 Department of Mathematics, University of Arizona, Tucson, AZ 85721, United States of America

E-mail: yugull05@gmail.com

Received 12 November 2020, revised 22 July 2021
Accepted for publication 3 September 2021
Published 16 September 2021

Abstract
For a Brownian directed polymer in a Gaussian random environment, with \( q(t, \cdot) \) denoting the quenched endpoint density and \( Q_n(t, x_1, \ldots, x_n) = \mathbb{E}[q(t, x_1) \ldots q(t, x_n)] \), we derive a hierarchical PDE system satisfied by \( \{Q_n\} \). We present two applications of the system: (i) we compute the generator of \( \{\mu_t(dx) = q(t, x)dx\} \) for some special functionals, where \( \{\mu_t(dx)\} \) is viewed as a Markov process taking values in the space of probability measures; (ii) in the high temperature regime with \( d \geq 3 \), we prove a quantitative central limit theorem for the annealed endpoint distribution of the diffusively rescaled polymer path. We also study a nonlocal diffusion-reaction equation motivated by the generator and establish a super-diffusive \( O(t^{2/3}) \) scaling.

Keywords: directed polymer, stochastic heat equation, reaction-diffusion equation

Mathematics Subject Classification numbers: 60H15, 35K57, 82D60.

1. Introduction

The study of directed polymers in random environments has witnessed important progress in recent years. A common feature of the models is an interaction between a reference path measure, which is typically given by a random walk or a Brownian motion, and a background random environment. The polymer measure is then formulated as the Gibbs measure with a Hamiltonian describing the accumulated energy collected along the path in the random environment. The physically interesting quantities include the fluctuations of the free energy, the typical behaviors of the paths, and so on, see e.g. the books [18, 24, 28] and the references therein.

*Author to whom any correspondence should be addressed.
Recommended by Dr Sandra Cerrai.
While the random walk/Brownian motion is diffusive, the random environment can change the polymer’s behavior drastically. Indeed, when the temperature is low, the typical path is expected to be super-diffusive, and the transverse displacement of the path of length $T$ is expected to be of order $T^\xi$ with an exponent $\xi > \frac{1}{2}$. In $d = 1$, it has been conjectured that $\xi = \frac{2}{3}$ and the directed polymer model falls into the KPZ universality class. The proofs of this conjecture in several settings can be found, for example, in [3, 5, 6, 11–13, 23, 26, 33, 36, 39, 41]. For a more complete list, we refer to the review articles [22, 38]. Another feature of the polymer paths in low temperatures is that they localise and concentrate in small regions, see e.g. [7, 15, 19, 43] and the references therein.

In this paper, we consider the Wiener measure as the reference path measure and a generalised spacetime Gaussian random field as the environment, and our focus will be on the endpoint distribution of the polymer path. The results presented in the sequel offer a possibly new perspective of this problem. We develop a deterministic hierarchy governing the evolution of the endpoint distribution. From the hierarchy, we compute the generator of the Markov process associated with the endpoint distribution, for linear functionals. In $d \geq 3$ and for high temperatures, we make use of the hierarchy to quantify the diffusive behavior of the polymer path. We also study a nonlocal reaction-diffusion equation motivated by the generator, and establish a super-diffusive behavior with exponent $\frac{2}{3}$; that is, we establish spreading on spatial scales $O(t^{2/3})$. Our arguments are based on a mixture of tools from both probability and partial differential equations.

1.1. Main result

Let $\eta$ be a spacetime white noise built on the probability space $(\Omega, \mathcal{F}, P)$ and the expectation with respect to $\eta$ is denoted by $\mathbb{E}$. Fix a mollifier $0 \leq \phi \in C_\infty^c(\mathbb{R}^d)$ with $\int \phi = 1$. We smooth $\eta$ in the $x$-variable and define

$$\eta_\phi(t, x) = \int_{\mathbb{R}^d} \phi(x - y)\eta(t, y)dy,$$

which is a generalised Gaussian random field. The covariance function of $\eta_\phi$ is then given by

$$\mathbb{E}[\eta_\phi(t, x)\eta_\phi(s, y)] = \delta(t - s)R(x - y),$$

with the spatial covariance function

$$R(x) = \int_{\mathbb{R}^d} \phi(x + y)\phi(y)dy \in C_\infty^c(\mathbb{R}^d).$$

Let $(\Sigma, \mathcal{A}, \mathbb{P})$ be another probability space and $w$ be a Brownian motion built on it. The initial location $w_0$ is distributed according to $\mu_0(dx)$. Throughout the paper, we assume

$$\mu_0(dx) = q_0(x)dx, \quad q_0 \geq 0, \quad \int_{\mathbb{R}^d} q_0(x)dx = 1,$$

and consider the two cases (i) $q_0 \in C_c(\mathbb{R}^d)$, (ii) $q_0(x) = \delta(x)$. While other initial distributions can be considered as well, our main focus here is on $q_0(x) = \delta(x)$. We denote the expectation with respect to $w$ by $\mathbb{E}_{\mu_0}$, and define the energy of any Brownian path $w : [0, T] \to \mathbb{R}^d$ in the Gaussian environment $\eta_\phi$ by

$$H(T, w) = \int_0^T \eta_\phi(t, w_t)dt. \quad (1.1)$$
In this paper, we study the endpoint distribution of the random polymer, obtained by tilting the Brownian motion by a factor of $e^{i\theta H(T, w)}$, where $\beta > 0$ is the inverse temperature; that is, for any $T \geq 0$, we are interested in:

\[
\mu_T(dx) = q(T, x)dx,
\]
\[
q(T, x) = Z(T)^{-1}\mathbb{E}_{\mu_0}[\delta(w_T - x) \exp(\beta H(T, w) - \frac{1}{2}\beta^2 R(0) T)],
\]
\[
and Z(T) = \mathbb{E}_{\mu_0}[\exp(\beta H(T, w) - \frac{1}{2}\beta^2 R(0) T)].
\]

(1.2)

In $d = 1$, we also consider the random environment given by the spacetime white noise $\eta$, without any mollification in the spatial variable. To unify the notation, we allow $\phi$ to be the Dirac function $\phi(x) = \delta(x)$ in $d = 1$, and the spatial covariance function in this case is

\[
R(x) = \int_{\mathbb{R}} \phi(x + y)\phi(y)dy = \delta(x).
\]

We define

\[
q(T, x) = Z(T)^{-1}\mathbb{E}_{\mu_0}\left[\delta(w_T - x) \exp\left(\beta \int_0^T \eta(s, w_s)ds\right)\right],
\]
\[
Z(T) = \mathbb{E}_{\mu_0}\left[\exp\left(\beta \int_0^T \eta(s, w_s)ds\right)\right].
\]

(1.3)

Here: $\exp$: is the Wick-ordered exponential, see e.g. [37]. In both (1.2) and (1.3), the endpoint density $q$ is related to a stochastic heat equation with a multiplicative noise, which we will make more precise in section 2. We emphasise that the case of $\phi(x) = R(x) = \delta(x)$ is restricted to $d = 1$.

For any $t \geq 0, n \geq 1$ and $x_{1:n} := (x_1, \ldots, x_n) \in \mathbb{R}^{nd}$, define the $n$-point density by

\[
Q_n(t, x_{1:n}) = \mathbb{E}[q(t, x_1) \ldots q(t, x_n)].
\]

(1.4)

For any two functions $f, g$, we define $\langle f, g \rangle := \int f(x)g(x)dx$ as long as the integral is well-defined in a standard way. If $f$ and $g$ also depend on the $t$ variable, then we write $\langle f(t), g(t) \rangle = \int f(t, x)g(t, x)dx$; see section 1.5 for more details on the notations and conventions. We are now able to state the first main result.

**Theorem 1.1.** For any $n \geq 1$ and $T > 0$, if $f \in C_b^{1,2}([0, T] \times \mathbb{R}^{nd})$, we have

\[
\langle f(T), Q_n(T) \rangle = \langle f(0), q_0^{\otimes n} \rangle + \int_0^T \langle (\partial_t + \frac{1}{2}\Delta)f(t), Q_n(t) \rangle dt
\]
\[
+ \beta^2 \sum_{k=0}^{n} \int_0^T \langle f_{k,R}(t), Q_{n+k}(t) \rangle dt,
\]

(1.5)

where the functions $f_{k,R} : [0, T] \times \mathbb{R}^{(n+k)d} \to \mathbb{R}$ are given by
Let \( \mathcal{M}_1(\mathbb{R}^d) \) be the space of probability measures on \( \mathbb{R}^d \). Due to white-in-time correlation of \( \eta \), \( \{\mu_T\}_{T \geq 0} \) is a Markov process taking values in \( \mathcal{M}_1(\mathbb{R}^d) \). For any \( f \in C_0(\mathbb{R}^d) \), we associate it in the natural way with a functional \( F_f : \mathcal{M}_1(\mathbb{R}^d) \to \mathbb{R} \) given by

\[
F_f(\mu) = \langle f, \mu \rangle = \int_{\mathbb{R}^d} f(x) \mu(dx),
\]

where we abused the notation to also let \( \langle \cdot, \cdot \rangle \) denote the pairing between \( C_0(\mathbb{R}^d) \) and \( \mathcal{M}_1(\mathbb{R}^d) \). Denote the generator of \( \{\mu_T\}_{T \geq 0} \) by \( \mathcal{L} \), and let \(*\) denote the convolution. An immediate consequence of theorem 1.1 is

**Corollary 1.2.** Assume \( \mu_0(dx) = q_0(x)dx \) with \( q_0 \in C_c(\mathbb{R}^d) \). For any \( f \in C^2_b(\mathbb{R}^d) \),

\[
\mathcal{L} F_f(\mu_0) = \left\langle \frac{1}{2} \Delta f, q_0 \right\rangle + \beta^2 \langle f, T q_0 \rangle,
\]

with the operator \( T \) defined as

\[
T q_0(x) = \langle R \ast q_0, q_0 \rangle q_0(x) - q_0(x) R \ast q_0(x).
\]

Another application of theorem 1.1 is to study the diffusive behavior of the polymer endpoint in a high temperature regime when \( d \geq 3 \). It is well-known from the classical work \([2, 10, 20, 32, 42]\) that in this case and under a diffusive rescaling, the polymer endpoint converges to a standard normal distribution in the quenched sense. In our notation, as \( q(T, \cdot) \) denotes the quenched density of the endpoint of the polymer of length \( T \), the result says that for almost every realisation of the random environment, and any \( h \in C_0(\mathbb{R}^d) \), we have

\[
\int_{\mathbb{R}^d} h(x) T^{d/2} q(T, \sqrt{T} x) dx \to \int_{\mathbb{R}^d} h(x) G_h(x) dx, \quad \text{as } T \to \infty.
\]
Here $G_t(\cdot)$ is the density of $N(0, L_0)$. The results in [10, 20, 32] are for a discrete i.i.d. random environment. In the setting of the continuous Gaussian environment considered in this paper, the same result was proved in [14]. The above quenched central limit theorem (1.10) immediately implies the annealed one (recall that $Q_1 = \mathbb{E}[q]$)

$$
\int_{\mathbb{R}^d} h(x)T^{d/2}Q_1(T, \sqrt{T}x)dx \to \int_{\mathbb{R}^d} h(x)G_1(x)dx, \quad \text{as } T \to \infty. \quad (1.11)
$$

As \(\{Q_n\}_{n \geq 1}\) solves the PDE hierarchy (1.7), which can be viewed as a ‘perturbation’ of the heat equation, it is natural to ask if we can analyse the system of equations and show that the ‘perturbation’ is indeed small in this asymptotic regime. It turns out that the hierarchy provides a nice analytic framework for us to give a simple proof of (1.11) and to also quantify the convergence rate.

Let $X_T$ denote a random variable with the density $T^{d/2}Q_1(T, \sqrt{T}x)$, then the Wasserstein distance between $X_T$ and $N(0, L_0)$ is defined as

$$
d_W(X_T, N(0, L_0)) := \sup_{h \in \text{Lip}(1)} \left| \int_{\mathbb{R}^d} h(x)T^{d/2}Q_1(T, \sqrt{T}x)dx - \int_{\mathbb{R}^d} h(x)G_1(x)dx \right|,
$$

where Lip(1) = \{h \in C(\mathbb{R}^d) : |h(x) - h(y)| \leq |x - y|\}.

**Theorem 1.3.** Assume $q_0(\cdot) = \delta(\cdot)$. In $d \geq 3$, there exist positive constants $\beta_0(d, R), C(d, R, \beta)$ such that if $\beta < \beta_0$, we have for all $T > 0$ that

$$
d_W(X_T, N(0, L_0)) \leq C \left( \frac{\log T}{\sqrt{T}} 1_{d=3} + \frac{1}{\sqrt{T}} 1_{d \geq 4} \right). \quad (1.12)
$$

Similar results were obtained in [9, theorem 4], and these suggest that the error estimates above are sharp when $d \geq 5$.

Of particular interest is the mean square displacement of the polymer endpoint. We have the following error bound.

**Theorem 1.4.** Under the same assumption of theorem 1.3, it holds for all $T > 0$ that

$$
\left| \int_{\mathbb{R}^d} x^2Q_1(T, x)dx - d \right| \leq C \left( \frac{1}{\sqrt{T}} 1_{d=3} + \frac{\log T}{T} 1_{d=4} + \frac{1}{T} 1_{d \geq 5} \right). \quad (1.13)
$$

There are several results on error estimates for the mean square displacement [2, 32, 42], and theorem 1.4 seems to provide the best rate. We discuss the relation of theorem 1.4 to the previous results in more details in remark 1.14 below.

### 1.3. A nonlocal reaction-diffusion equation motivated by $\mathcal{L}$

Since $\{\mu_T\}_{T \geq 0}$ is a Markov process, we can consider the forward Kolmogorov equation associated with it. Recall that $\mathcal{L}$ denotes its generator. For any $f \in C_b(\mathbb{R}^d)$, let

$$
\mathcal{U} : [0, \infty) \times \mathcal{M}_1(\mathbb{R}^d) \to \mathbb{R}
$$

be the solution to

$$
\partial_t \mathcal{U}(t, \mu) = \mathcal{L}\mathcal{U}(t, \mu), \quad t > 0,
$$

$$
\mathcal{U}(0, \mu) = (f, \mu). \quad (1.14)
$$
With \( \mu_0(dx) = q_0(x)dx \), the solution can be written as

\[
\mathcal{H}(t, \mu_0) = \mathcal{E}[\langle f, \mu_t \rangle] = \int_{\mathbb{R}} f(x) Q_1(t, x) dx,
\]

where we used the fact that \( \mu_t(dx) = q(t, x)dx \) and \( \mathcal{E}[q(t, x)] = Q_1(t, x) \). Thus, one can try to study the asymptotic behavior of \( Q_1(t, \cdot) \) as \( t \to \infty \) by considering the equation (1.14), rather than the system (1.16).

Nevertheless, computing the generator \( \mathcal{L} \) for general functionals is difficult, and, in addition, the obtained formula might not be easy to manipulate. In corollary 1.2, we considered the simplest possible linear functionals, and the equation (1.8) shows that, for these functionals on \( \mathcal{M}_1(\mathbb{R}^d) \), the action of \( \mathcal{L} \) is equivalent to that of the differential operator \( \frac{1}{2} \Delta + \beta^2 T \) acting on the corresponding density. This motivates us to consider the following deterministic PDE associated with the operator \( T \). Take the case of \( R(\cdot) = \delta(\cdot) \) in \( d = 1 \), and denote the \( L^2(\mathbb{R}^d) \) norm by \( \| \cdot \| \). For any \( f \in L^2(\mathbb{R}) \), we have

\[
Tf(x) = \| f \|^2 f(x) - f(x)^2.
\]

Consider the following equation

\[
\begin{align*}
\partial_t g(t, x) &= \frac{1}{2} \Delta g(t, x) + \beta^2 T g(t, \cdot) \\
&= \frac{1}{2} \Delta g(t, x) + \beta^2 \| g(t, \cdot) \|^2 g(t, x) - \beta^2 g(t, x)^2, \\
g(0, x) &= q_0(x).
\end{align*}
\]

We see in the sequel that this describes the evolution of a probability density.

It is unclear whether (1.15) has anything to do with (1.14), since we only know that the two evolutions match at \( t = 0 \), which is due to the fact that \( \mathcal{H}(0, \mu) = \langle f, \mu \rangle \) is a linear functional on \( \mathcal{M}_1(\mathbb{R}) \). Nevertheless, the following result shows a super-diffusive behavior of \( g \) with the exponent \( \frac{2}{3} \).

**Theorem 1.5.** In \( d = 1 \), assume \( 0 \leq q_0 \in C_c(\mathbb{R}) \) and \( \int_{\mathbb{R}} q_0(x) dx = 1 \). For any \( p \geq 1 \), there exists a constant \( C = C(p, \beta, q_0) > 0 \) such that

\[
C^{-1} T^{\frac{2}{3}} \leq \int_{\mathbb{R}} |x|^p g(T, x) dx \leq C T^{\frac{2}{3}}, \quad \text{for all } T \geq 1.
\]

It is unclear to us whether the above super-diffusion with the ‘right’ exponent \( \frac{2}{3} \) is a coincidence or not. In the case of the spacetime white noise environment in \( d = 1 \), the function \( Q_1 \) is the annealed density of the endpoint of the continuum directed random polymer [1], and \( \{ Q_n \}_{n \geq 1} \) is a weak solution of the infinite system:
\[\partial_t Q_n(t, x_1, \ldots, x_n) = \frac{1}{2} \Delta Q_n(t, x_1, \ldots, x_n) + \beta^2 Q_n(t, x_1, \ldots, x_n) \sum_{1 \leq i < j \leq n} \delta(x_i - x_j) - \beta^2 n \sum_{i=1}^n Q_{n+1}(x_1, \ldots, x_n, x_i) + \beta^2 n(n+1) \int_{\mathbb{R}} Q_{n+2}(x_1, \ldots, x_n, \tilde{x}, \tilde{x}) d\tilde{x}, \quad n \geq 1. \quad (1.16)\]

For this model, the super-diffusion with exponent \(\frac{5}{3}\) was shown in [23, theorem 1.11]. Taking \(n = 1\), the above system yields
\[\partial_t Q_1(t, x) = \frac{1}{2} \Delta Q_1(t, x) + \beta^2 \int_{\mathbb{R}} Q_3(t, x, \tilde{x}, \tilde{x}) d\tilde{x} - \beta^2 Q_2(t, x, x). \quad (1.17)\]

Thus, the question reduces to justifying if (1.15) is a reasonable approximation of (1.17) or not. Incidentally, if we make the assumption of a factorised joint density to close the hierarchy
\[Q_2(t, x_1, x_2) \approx Q_1(t, x_1) Q_1(t, x_2), \quad Q_3(t, x_1, x_2, x_3) \approx \prod_{j=1}^3 Q_1(t, x_j), \quad (1.18)\]

which is similar in spirit to the molecular chaos assumption in the BBGKY hierarchy of kinetic theory [16], then (1.17) reduces to (1.15):
\[\partial_t Q_1(t, x) \approx \frac{1}{2} \Delta Q_1(t, x) + \beta^2 \|Q_1(t, \cdot)\|^2 Q_1(t, x) - \beta^2 Q_2^2(t, x). \quad (1.19)\]

However, we do not claim either (1.18) or (1.19) in this paper.

### 1.4. Discussions

As the endpoint density of the Brownian motion solves the standard heat equation, we look for a counterpart when the Brownian motion is weighted by a random environment. For each fixed realisation of the random environment, it is known that the polymer model is equivalent to a diffusion in a (different) random environment [4, theorem 2]. Thus, in the quenched setting, the analogue of the standard heat equation we are looking for is a Fokker–Planck equation with a random coefficient, describing the evolution of the density of the aforementioned diffusion. However, studying either the solution to the Fokker–Planck equation or its ensemble average seems to be as complicated as the polymer model itself; hence, the main message we wish to convey here is the following: rather than studying the single point distribution, one could instead look at the multipoint distributions defined in (1.4). By definition, for each \(T \geq 0\), \(Q_n(T, \cdot)\) is a probability density on \(\mathbb{R}^n\). While we do not have an underlying dynamics that reproduces the evolution of \(Q_n\), heuristically, it can be viewed as the joint density of \(n\) particles, interacting indirectly through their separate individual interaction with the common random environment, similar to ‘independent walkers in the same environment’. Theorem 1.1, which comes from a straightforward application of Itô’s formula, shows that \(\{Q_n\}_{n \geq 1}\) solves a hierarchical PDE system. In this way, the study of the endpoint distribution of the random
polymer, in the annealed setting, may be reduced to the study of $Q_1$ and the analysis of the deterministic PDE system satisfied by $\{Q_n\}_{n \geq 1}$.

We make a few remarks.

**Remark 1.6.** A result similar to theorem 1.1 was proved in [15, theorem 3.1] for a different polymer model (albeit not formulated as a PDE hierarchy), with a random walk reference path measure and an environment on $\mathbb{R}_+ \times \mathbb{Z}^d$ made of i.i.d. copies of Brownian motions.

**Remark 1.7.** We show in the sequel (see lemma A.2) that

$$Q_n(t, x_1, \ldots, x_n) = \mathbb{E}\left[ u(t, x_1) \ldots u(t, x_n) \right], \quad (1.20)$$

with $u$ solving the stochastic heat equation

$$\partial_t u = \frac{1}{2} \Delta u + \beta u \eta, \quad u(0, \cdot) = q_0(\cdot).$$

If we only consider the numerator of the rhs of (1.20), and define

$$\tilde{Q}_n(t, x_1, \ldots, x_n) = \mathbb{E}[u(t, x_1) \ldots u(t, x_n)],$$

it is well-known that $\tilde{Q}_n$ solves

$$\partial_t \tilde{Q}_n = \frac{1}{2} \Delta \tilde{Q}_n + \beta^2 \tilde{Q}_n \sum_{1 \leq i < j \leq n} R(x_i - x_j) =: \mathcal{H}_n \tilde{Q}_n. \quad (1.21)$$

In this case there is no coupling between $\tilde{Q}_n$ for different values of $n$, and the Hamiltonian $\mathcal{H}_n$ is the so-called delta-Bose gas if we have a contact interaction $R(\cdot) = \delta(\cdot)$ in $d = 1$. There are many studies on the moments of the stochastic heat equation, either relying on the Feynman–Kac representation of the solution to (1.21) or the spectral property of $\mathcal{H}_n$, and we refer to the monograph [35] and the references therein. The equation (1.21) should be compared with (1.7), in which we have additional terms related to $Q_{n+1}$ and $Q_{n+2}$.

**Remark 1.8.** For the **annealed** endpoint distribution considered in the present paper, with the density given by $\mathbb{E}[q(T, \cdot)]$, it is conjectured that in $d = 1$, the rescaled density $\frac{T}{2} \mathbb{E}[q(T, T^+ \cdot)]$ converges weakly in space to some universal limit as $T \to \infty$, see the result on a related last passage percolation model [34]. The limit was further identified in [27, 40].

**Remark 1.9.** The asymptotics of the solution of (1.15) are not obvious. In order to understand the exponent $\frac{2}{3}$ in theorem 1.5, one can make the following back-of-the-envelope computation. If we assume spreading at spatial scales $O(t^p)$, then to preserve the fact that $g$ is a probability measure, we must have $\sup g \sim O(t^{-p})$. This yields $\|g\|^2 \sim O(t^{-p})$. In order to use this, we linearise the equation around zero to obtain

$$\partial_t g \approx \frac{1}{2} \Delta g + \beta^2 \|g\|^2 g \approx \frac{1}{2} \Delta g + O(t^{-p})g.$$

Then, using the large $x$ asymptotics of the heat equation and the fact that the last term yields an integrating factor, we find

$$g(t, x) \approx e^{\int_0^t \frac{x^2}{4} \partial_x^2 + O(\log x)} \approx e^{\alpha^{-p}e^{-\frac{x^2}{4}} \partial_x^2 - \frac{x^2}{4}}.$$
For consistency with our assumption of spreading in $x$ like $O(t^p)$, we require that $g$ is ‘large’ to the left of $O(t^p)$ and ‘small’ to the right of $O(t^p)$. This means that the two terms in the exponent should cancel at $x \sim O(t^p)$. In other words, we require $O(t^{1-p}) = (t^p)^2/2t$. Solving this yields $p = \frac{2}{3}$. Unfortunately, this argument is far from rigorous. Instead, as with the heat equation, in order to establish the spreading behavior of $g$, the key estimate is an upper bound on the $L^\infty$ norm that yields decay to zero at the sharp rate, which is $O(t^{-2/3})$. Since the Laplacian (diffusion) can only cause decay like $O(t^{-1/2})$ in $d = 1$, then nonlinear terms have to provide the mechanism for this decay. Our proof proceeds by establishing a functional inequality relating the two nonlinear terms at any maximum of $g$. This, combined with a differential inequality satisfied by the maximum, shows that $g(t) \lesssim t^{-2/3}$. From there we obtain the upper bound in theorem 1.5 via the construction of a supersolution and the lower bound via a simple variational argument.

**Remark 1.10.** It is natural to investigate further questions on the PDE (1.15), including the asymptotics of $T^{2/3}g(T, T^{2/3} \cdot)$ as $T \to \infty$ and the behavior of $g$ in high dimensions. One can study the equation corresponding to the spatially correlated noise:

$$
\partial_t g(t, x) = \frac{1}{2} \Delta g(t, x) + \beta^2 \langle R * g(t, \cdot), g(t, \cdot) \rangle g(t, x) - \beta^2 g(t, x) R * g(t, x),
$$

where $R \in C^\infty_c(\mathbb{R}^d)$ is the spatial covariance function of the noise. Compared to (1.15) which is the case of $R(\cdot) = \delta(\cdot)$, the above equation is ‘more nonlocal’, and the analytic tools used in this paper do not seem to apply. Indeed, the functional inequalities and delicate identities used in the proof of theorem 1.5 either are not true or do not make sense and do not have obvious analogues in this more general setting. These are presented in [29].

**Remark 1.11.** We note that the proof of theorem 1.5 does not use any regularity of $q_0$, and would apply equally well to $q_0$ that is a localised probability measure such as $\delta$; however, it is not immediately obvious that (1.15) is well-posed with measure initial data. Hence, in this work, we impose the condition that $q_0 \in C_c(\mathbb{R})$ in order to avoid technical issues. In a future work, we show that this condition may be relaxed; that is, the estimates established here are sufficient to establish such a well-posedness result for localised probability measures.

**Remark 1.12.** In $d \geq 3$ and the high temperature regime, theorems 1.3 and 1.4 concern the annealed distribution $Q_1$, which is obtained after taking an average with respect to the random environment. It is natural to ask about the extra error induced by the random fluctuations of the environment. As the focus of the paper is on the PDE hierarchy, which involves $Q_1$ rather than $q$, we do not study this problem here, but note that it is not very hard to extract error bounds on the random fluctuations from the proof of theorem 1.1. We refer to remark 3.1 below for more details.

**Remark 1.13.** In theorems 1.3 and 1.4, we are deep in the high temperature regime $\beta \ll 1$ to obtain the error estimate in the annealed central limit theorem. The quenched central limit theorem actually holds in the full weak disorder regime $\beta < \beta_c$ for some critical value of $\beta_c$ [20]. As we strive for a more quantitative estimate here, it is unclear to us whether similar error estimates can be proved for all $\beta < \beta_c$.

**Remark 1.14.** The first error bound on the convergence of the mean square displacement was proved in [32, equation (1.7)], and their result in the discrete setting translates to our case as
\[ \left| \frac{1}{T} \int_{\mathbb{R}^d} |x|^2 Q_t(T, x) \, dx - d \right| \leq C \frac{1}{T^{\theta - \delta}}. \]

Here \( \theta = \min\left(\frac{d-2}{d}, \frac{1}{2}\right) \) and \( \delta > 0 \) can be arbitrarily small. In [2, 42], similar results were shown which corresponds to the following in our setting

\[ \left| \frac{1}{T} \int_{\mathbb{R}^d} |x|^2 Q_t(T, x) \, dx - d \right| \leq C \left( \frac{1}{T^{(d-2)/2}} \mathbb{1}_{2 < d < 6} + \frac{\sqrt{\log T}}{T} \mathbb{1}_{d = 6} + \frac{1}{T} \mathbb{1}_{d > 6} \right). \]

It is worth mentioning that in these works, the quenched error estimates were also proved, while we only focus on the annealed case here.

**Organisation of the paper.** In section 2, we briefly discuss the connection between the directed polymer and the stochastic heat equation, which will be used later in our proof. Sections 3–5 are devoted to the proofs of theorems 1.1, 1.3 and 1.5 respectively. In appendix A, we review some basics about stochastic heat equations for the convenience of readers. The proofs of some technical lemmas are presented in appendix B.

### 1.5. Notation and conventions

We recall and define some notation.

(a) The expectation with respect to the Gaussian random environment is denoted by \( E \), and the expectation with respect to Brownian motions is \( \mathbb{E} \).

(b) We consider two cases of spatial covariance functions of the Gaussian environment (a) \( R(\cdot) \in C^\infty_c(\mathbb{R}^d), d \geq 1 \) and (b) \( R(\cdot) = \delta(\cdot), d = 1 \).

(c) The initial distribution \( \mu_0(dx) = q_0(x) \, dx \) is fixed, and we include the two cases (a) \( q_0(\cdot) \in C_c(\mathbb{R}^d) \) and (b) \( q_0(\cdot) = \delta(\cdot) \).

(d) For functions \( f, g \) and measure \( \mu \), we write \( \langle f, g \rangle = \int fg, \langle f, \mu \rangle = \int f(x) \mu(dx) \), and \( \| f \|^2 = \langle f, f \rangle \).

(e) We use * to denote convolution in the spatial variable \( x \), and the standard heat kernel of \( \partial_t - \frac{1}{2} \Delta \) is \( G_t(x) = (2\pi t)^{-d/2} \exp(-|x|^2/(2t)) \).

### 2. Directed polymer and stochastic heat equation

In this section, we briefly discuss the relationship between directed polymers and the stochastic heat equation with a multiplicative noise.

First, we define the time reversal of \( \eta \) and \( \eta_0 \):

\[ \xi(t, x) = \eta(-t, x), \quad \xi_0(t, x) = \eta_0(-t, x) = \int_{\mathbb{R}^d} \phi(x - y) \xi(t, y) \, dy. \quad (2.1) \]

Fix the inverse temperature \( \beta > 0 \). For any \( s \in \mathbb{R} \) and \( x \in \mathbb{R}^d \), define \( U(s, x; t, y) \) as the solution of

\[ \begin{align*}
\partial_t U(s, x; t, y) &= \frac{1}{2} \Delta_x U(s, x; t, y) + \beta \, U(s, x; t, y) \xi_0(t, y), \quad t > s, y \in \mathbb{R}^d, \\
U(s, x; s, y) &= \delta(y - x).
\end{align*} \quad (2.2) \]
Here the product between $U$ and $\xi \phi$ is interpreted in the Itô–Walsh sense [44], and we have included the spacetime white noise case of $\phi(\cdot) = \delta(\cdot)$ in $d = 1$. Then the quenched endpoint density $q(T, x)$, defined in (1.2) and (1.3), is also given by

$$q(T, x) = \frac{\int_{\mathbb{R}^d} U(-T, x; 0, y)q_0(y)dy}{\int_{\mathbb{R}^d} U(-T, x; 0, y)q_0(y)dy}.$$  (2.3)

To see this, consider the spatially correlated case with $\phi \in C^\infty_c(\mathbb{R}^d)$. We only need to use the Feynman–Kac formula [8] to rewrite

$$\int_{\mathbb{R}^d} U(-T, x; 0, y)q_0(y)dy = \mathbb{E}_{\mu_0} \left[ \delta(w_T - x)e^{\beta \int_0^T \xi(\cdot, w_s)yds - \frac{1}{2} \beta^2 R(0)T} \right].$$  (2.4)

Here we recall that $w$ is a standard Brownian motion that is independent from $\xi$, and $\mathbb{E}_{\mu_0}$ denotes the expectation with respect to $w$, with the starting point

$$w_0 \sim \mu_0(dx) = q_0(x)dx.$$  

For any $y \in \mathbb{R}^d$, let $\mathbb{E}_y$ denote the expectation with respect to the Brownian motion starting at $w_0 = y$. Then we can rewrite (2.4) as

$$\int_{\mathbb{R}^d} U(-T, x; 0, y)q_0(y)dy = \int_{\mathbb{R}^d} q_0(y)\mathbb{E}_y \left[ \delta(w_T - x)e^{\beta \int_0^T \xi(\cdot, w_s)yds - \frac{1}{2} \beta^2 R(0)T} \right] dy.$$

For the case of $\phi(\cdot) = \delta(\cdot)$ in $d = 1$, by the definition of the Wick-ordered exponential, we have

$$\mathbb{E}_{\mu_0} \left[ \delta(w_T - x) : \exp \left( \beta \int_0^T \eta(s, w_s)ds \right) \right] = \mathbb{E}_{\mu_0} \left[ \delta(w_T - x) : \exp \left( \beta \int_0^T \xi(-s, w_s)ds \right) \right] = \int_{\mathbb{R}^d} U(-T, x; 0, y)q_0(y)dy.$$

3. Proof of theorem 1.1

We make use of the Feynman–Kac representation in (1.2) to study the case of spatially correlated noise, i.e. when the spatial covariance function $R(\cdot) \in C^\infty_c(\mathbb{R}^d)$. Through an approximation argument, we derive the corresponding result for the case of spacetime white noise.
3.1. Colored noise environment: $R(\cdot) \in C_c^\infty(\mathbb{R}^d), d \geq 1$

We first introduce some notation. Fix $n \geq 1$, $T > 0$ and a $C_b^{1,2}$ function $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$, we define

$$M(t, w) = \exp(\beta H(t, w) - \frac{1}{2} \beta^2 R(0)t),$$

$$Y_f(t) = f(t, w_1^n, \ldots, w_n^n) \prod_{j=1}^n M(t, w^j), \quad \text{and}$$

$$X_f(t) = \mathbb{E}_{\omega_0}[Y_f(t)],$$

where $\{w^j\}_{j=1,\ldots,n}$ are independent copies of Brownian motions built on $(\Sigma, A, \mathbb{P})$. Thus, we have

$$X_1(t) = Z(t)^n,$$

where $Z(t)$ is the partition function defined in $(1.2)$ and $1$ stands for the constant function $1(x) \equiv 1$. With the new notation, we define

$$X_f(t) := \langle f(t, \cdot), \mu_t^{\omega_n} \rangle = \langle f(t, \cdot), q(t, \cdot)^{\omega_n} \rangle = \frac{X_f(t)}{X_1(t)} = \frac{X_f(t)}{Z(t)^n}.$$

**Proof of 1.5** In the following, the differential $d$ is the full stochastic differential with respect to both the Gaussian environment and the Brownian motions.

We first note that for each fixed $w$,

$$H(t, w) = \int_0^t \eta_0(s, w_s)ds$$

is a Brownian motion with variance $R(0)$, and for $w^l$, $w^l$, we have the bracket process

$$\langle H(\cdot, w^l), H(\cdot, w^l) \rangle_t = \int_0^t R(w^l_t - w^l_t)ds.$$  

This implies

$$dM(t, w) = \beta M(t, w)dH(t, w),$$

$$d(M(\cdot, w^l), M(\cdot, w^l))_t = \beta^2 M(t, w^l)M(t, w^l)R(w^l_t - w^l_t)dt.$$  

(3.2)

We also know that $H(t, w)$ is independent of $w$.

Now we apply Itô’s formula to $Y_f(t)$:

$$dY_f(t) = d \left[ f(t, w_1^n, \ldots, w_n^n) \prod_{j=1}^n M(t, w^j) \right]$$

$$= f(t, w_1^n, \ldots, w_n^n) d \left[ \prod_{j=1}^n M(t, w^j) \right] + \prod_{j=1}^n M(t, w^j) d[f(t, w_1^n, \ldots, w_n^n)]$$

$$+ d(f(c, w_1^n, \ldots, w_n^n), \prod_{j=1}^n M(c, w^j)).$$
By (3.2), we have
\[
\begin{align*}
\frac{d}{dt} & \prod_{j=1}^{n} M(t, w^j) = \beta \sum_{k=1}^{n} \prod_{j=1}^{n} M(t, w^j) dH(t, w^k) \\
& + \beta^2 \sum_{1 \leq k < l \leq n} \prod_{j=1}^{n} M(t, w^j) R(w^k_t - w^l_t) dt.
\end{align*}
\]
We also have
\[
\begin{align*}
\frac{df}{dt}(t, w_1, \ldots, w_n) = & \sum_{j=1}^{n} \nabla_j f(t, w_1, \ldots, w_n) dw_j + \left( \partial_t + \frac{1}{2} \Delta \right) f(t, w_1, \ldots, w_n) dt,
\end{align*}
\]
where \( \nabla_j \) denotes the gradient with respect to the \( j \)-th variable and \( \Delta = \sum_{j=1}^{n} \nabla_j \cdot \nabla_j \). Since \( H(t, w) \) is independent of \( w \), by (3.2), we have
\[
\langle f(w_1, \ldots, w_n), \prod_{j=1}^{n} M(\cdot, w^j) \rangle_t \equiv 0.
\]
Thus, we have
\[
\begin{align*}
Y_f(T) &= f(0, w_1^0, \ldots, w_n^0) + \beta \sum_{k=1}^{n} \int_0^T Y_f(t) dH(t, w^k) \\
& + \beta^2 \sum_{1 \leq k < l \leq n} \int_0^T Y_f(t) R(w^k_t - w^l_t) dt \\
& + \sum_{k=1}^{n} \int_0^T \prod_{j=1}^{n} M(t, w^j) \nabla_k f(t, w_1^1, \ldots, w_n^k) dw_k \\
& + \int_0^T \prod_{j=1}^{n} M(t, w^j) \left( \partial_t + \frac{1}{2} \Delta \right) f(t, w_1^1, \ldots, w_n^k) dt.
\end{align*}
\]
Taking expectation with respect to the Brownian motions, we have
\[
\begin{align*}
X_f(T) &= \mathbb{E}_\mu [Y_f(T)] = \mathbb{E}_\mu [f(0, w_1^0, \ldots, w_n^0)] + \beta \sum_{k=1}^{n} \mathbb{E}_\mu \left[ \int_0^T Y_f(t) dH(t, w^k) \right] \\
& + \beta^2 \sum_{1 \leq k < l \leq n} \int_0^T \mathbb{E}_\mu [Y_f(t) R(w^k_t - w^l_t)] dt \\
& + \int_0^T \mathbb{E}_\mu \left[ \prod_{j=1}^{n} M(t, w^j) \left( \partial_t + \frac{1}{2} \Delta \right) f(t, w_1^1, \ldots, w_n^k) \right] dt.
\end{align*}
\] (3.3)
For the second term on the rhs, recall that $\eta_b(t, x) = \int_0^x \phi(x - y) \eta(t, y) dy$ with $\eta$ a spacetime white noise, using (1.1), we can rewrite it as
\[
\beta \sum_{k=1}^n \mathbb{E}_{\mu_0} \left[ \int_0^T Y_f(t) dH(t, w^k) \right] = \beta \sum_{k=1}^n \int_0^T \int_{\mathbb{R}^d} \mathbb{E}_{\mu_0} [Y_f(t) \phi(u^k_t - y)] \eta(t, y) dy \, dt.
\] (3.4)

Now we apply Itô’s formula to $X_f(t) = X_f(t)Z(t)^{-n}$:
\[
dX_f(t) = \frac{1}{Z(t)^n} dX_f(t) - \frac{nX_f(t)}{Z(t)^{n+1}} dZ(t) + \frac{n(n + 1)X_f(t)}{2Z(t)^{n+2}} d\langle Z, Z \rangle_t - \frac{n}{Z(t)^{n+1}} d\langle X_f, Z \rangle_t.
\] (3.5)

By (3.3), the martingale component of $X_f$ is given by (3.4). A simpler version of (3.3) also gives
\[
Z(T) = 1 + \beta \int_0^T \int_{\mathbb{R}^d} \mathbb{E}_{\mu_0} [M(t, w) \phi(w_t - y)] \eta(t, y) dy \, dt,
\]
which implies
\[
d\langle X_f, Z \rangle_t = \beta^2 \sum_{k=1}^n \int_0^T \mathbb{E}_{\mu_0} \left[ f(t, w_t^1, \ldots, w_t^n) \prod_{j=1}^{n+1} M(t, w^j) \phi(u_t^{k+1} - y) \phi(u_t^{n+1} - y) \right] dy \, dt
\]
\[
= \beta^2 \sum_{k=1}^n \mathbb{E}_{\mu_0} \left[ f(t, w_t^1, \ldots, w_t^n) \prod_{j=1}^{n+1} M(t, w^j) R(u_t^{k+1} - u_t^{n+1}) \right] dt.
\] (3.6)

Combining (3.3)–(3.6), we have
\[
X_f(T) = X_f(0) + \sum_{i=1}^4 I_i(T),
\] (3.7)
with
\[
I_i(T) = \int_0^T \mathbb{E}_{\mu_0} \left[ \prod_{j=1}^n M(t, w^j) \left( \partial_t + \frac{1}{2} \Delta \right) f(t, w_t^1, \ldots, w_t^n) \right] \, dt,
\]
\[ I_3(T) = -n \int_0^T \frac{X_f(t)}{Z(t)^{n+1}} dZ(t) \]

\[ = -\frac{n}{\beta} \int_0^T \frac{X_f(t)}{Z(t)^{n+2}} \mathbb{E}^\nu_{0} [M(t, w) \phi(w_1 - y)] \eta(t, y) dy \, dt, \]

\[ I_3(T) = \frac{n(n + 1)}{2} \int_0^T \frac{X_f(t)}{Z(t)^{n+2}} \mathbb{E}^\nu_{0} \left[ \prod_{j=1}^{n+1} M(t, w^j) \cdot R(w_1^j - w_2^j) \right] \, dt, \]

and

\[ I_4(T) = -n \int_0^T \frac{1}{Z(t)^{n+1}} d(X_f, Z), \]

\[ = -\frac{n^2 \beta}{2} \sum_{k=1}^{n} \int_0^T \frac{1}{Z(t)^{n+2}} \mathbb{E}^\nu_{0} \left[ f(t, w_1^1, \ldots, w_2^k) \prod_{j=1}^{n+1} M(t, w^j) R(w_1^j - w_2^j) \right] \, dt. \]

Taking the expectation with respect to \( \eta \), we have

\[ \mathbb{E}[I_1(T)] = \int_0^T \langle \beta^2 f_{0, \rho}(t) + \left( \partial_t + \frac{\beta}{2} \Delta \right) f(t), Q_\rho(t) \rangle \, dt, \quad \mathbb{E}[I_2(T)] = 0, \]

\[ \mathbb{E}[I_3(T)] = \int_0^T \langle \beta^2 f \rho(t), Q_{n+2}(t) \rangle \, dt, \]

\[ \mathbb{E}[I_4(T)] = \int_0^T \langle \beta^2 f_{1, \rho}(t), Q_{n+1}(t) \rangle \, dt. \]

It suffices to note that \( \mathbb{E}[X_f(T)] = \langle f(T), Q_\rho(T) \rangle \) and \( X_f(0) = \langle f(0), q_0^{(\rho)} \rangle \) to complete the proof of (1.5). \( \square \)

**Remark 3.1.** The equation (1.5) results from taking expectation with respect to the noise on both sides of (3.7), in which the martingale terms disappear. If we are interested in the size of the random fluctuations, it suffices to study the martingale terms in \( I_1(T), I_2(T) \). Actually, an SPDE satisfied by \( q(t, x) \) was derived in [30, proposition 3.2].

### 3.2. White noise environment: \( R(t) = \delta(t), d = 1 \)

The proof in this case is through an approximation of the spacetime white noise by colored noise. Recall that we only consider \( d = 1 \) in this case.

For each \( \varepsilon > 0 \), define

\[ \xi_\varepsilon(t, x) = \int_{\mathbb{R}} \phi_\varepsilon(x - y) \xi(t, y) dy, \quad \phi_\varepsilon(x) = \frac{1}{\varepsilon} \phi\left( \frac{x}{\varepsilon} \right) \]

and the spatial covariance function

\[ R_\varepsilon(x) = \frac{1}{\varepsilon} R\left( \frac{x}{\varepsilon} \right) = \int_{\mathbb{R}} \phi_\varepsilon(x + y) \phi_\varepsilon(y) dy. \]
For any \( s \in \mathbb{R}, x \in \mathbb{R} \), let \( U_\varepsilon \) be the solution to
\[
\partial_t U_\varepsilon (s, x; t, y) = \frac{1}{2} \Delta_y U_\varepsilon (s, x; t, y) + \beta U_\varepsilon (s, x; t, y) \xi(t, y), \quad t > s, \quad y \in \mathbb{R},
\]
\[
U_\varepsilon (s, x; s, y) = \delta(y - x).
\]
(3.10)

In other words, \( U_\varepsilon \) solves (2.2) with \( \xi_0 \) replaced by \( \xi_\varepsilon \). Similarly, the quenched end-point distribution of the polymer in the environment \( \xi_\varepsilon \), with the starting point distributed as \( \mu_{0,\varepsilon}(dx) = q_0(x)dx \), is defined as
\[
\mu_{T,\varepsilon}(dx) = q_{\varepsilon}(T, x)dx, \quad q_{\varepsilon}(T, x) = \frac{\int_0^T U_\varepsilon(-T, x; 0, y)q_0(y)dy}{\int_\mathbb{R} U_\varepsilon(-T, x; 0, y)q_0(y)dy}.
\]
(3.11)

For any \( n \geq 1, T > 0 \) and \( x_1, n = (x_1, \ldots, x_n) \in \mathbb{R}^n \), define
\[
Q_{n,\varepsilon}(T, x_1, n) = \mathbb{E}[q_{\varepsilon}(T, x_1) \ldots q_{\varepsilon}(T, x_n)].
\]

By (1.5) (for the case of \( \phi \in C_c^\infty(\mathbb{R}^d) \)), we have, for any \( f \in C_b^{1,2}([0, T] \times \mathbb{R}^n) \),
\[
\langle f(T), Q_{n,\varepsilon}(T) \rangle = \langle f(0), q_0^n \rangle + \int_0^T \left( \partial_t + \frac{1}{2} \Delta_t \right) f(t, \mathbb{E}[Q_{n,\varepsilon}(t)])dt
+ \beta^2 \sum_{k=0}^{n-1} \int_0^T \langle f_{k,\varepsilon}(t), Q_{n,\varepsilon+k,\varepsilon}(t) \rangle dt,
\]
(3.12)

with the shorthand notation \( f_{k,\varepsilon} := f_{k,\varepsilon, \varepsilon} \). Recall that for any \( R \in C_c^\infty(\mathbb{R}^d) \), the functions \( f_{k,\varepsilon, \varepsilon} \) were defined in (1.6), for \( k = 0, 1, 2 \).

To prove (1.5) for the case of \( \phi(\cdot) = \delta(\cdot) \), we need the following two technical lemmas:

**Lemma 3.2.** For any \( p \geq 1, t > 0, x \in \mathbb{R} \), \( q_{\varepsilon}(t, x) \rightarrow q(t, x) \) in \( L^p(\Omega, \mathcal{F}, \mathbb{P}) \), as \( \varepsilon \rightarrow 0 \). The convergence is uniform for \( x \in \mathbb{R} \) and \( t \) in compact subsets of \( (0, \infty) \). In addition, there exists \( C = C(p, \beta, T) > 0 \) such that for all \( \varepsilon \in (0, 1), t \in (0, T], x \in \mathbb{R} \),
\[
\mathbb{E}[|q_{\varepsilon}(t, x)|^p] + \mathbb{E}[|q(t, x)|^p] \leq C(G_t * q_0(x))^p.
\]
(3.13)

**Lemma 3.3.** \( Q_{n} \) is continuous on \( (0, \infty) \times \mathbb{R}^n \).

The proofs of lemmas 3.2 and 3.3 is given in appendix A.

**Corollary 3.4.** There exist \( C = C(\beta, T) > 0 \) such that for all \( t \in (0, T], (x_1, \ldots, x_n) \in \mathbb{R}^n \) and \( \varepsilon \in (0, 1) \),
\[
Q_{n,\varepsilon}(t, x_1, \ldots, x_n) + Q_{\varepsilon}(t, x_1, \ldots, x_n) \leq C \prod_{j=1}^n G_t * q_0(x_j).
\]

In addition, \( Q_{n,\varepsilon}(t, x_1, \ldots, x_n) \rightarrow Q_{\varepsilon}(t, x_1, \ldots, x_n) \) as \( \varepsilon \rightarrow 0 \), and the convergence is uniform for \( (x_1, \ldots, x_n) \in \mathbb{R}^n \) and \( t \) in compact subsets of \( (0, \infty) \).
Now we can finish the proof of Theorem 1.1.

**Proof.** We start from (3.12) and pass to the limit of \(\varepsilon \to 0\) for each term. First, by Corollary 3.4, we have

\[
\langle f(T), Q_{n;\varepsilon}(T) \rangle \to \langle f(T), Q_n(T) \rangle,
\]

and

\[
\int_0^T \left( \partial_t + \frac{1}{2} \Delta \right) f(t), Q_{n;\varepsilon}(t) \, dt \to \int_0^T \left( \partial_t + \frac{1}{2} \Delta \right) f(t), Q_n(t) \, dt.
\]

The rest of the \(\varepsilon\)-dependent terms in (3.12) are treated in the same way, so we take \(\int_0^T \langle f_{0,\varepsilon}(t), Q_{n;\varepsilon}(t) \rangle \, dt\) as an example: for any \(i, j\) from the summation. By the change of variable \(x_i \mapsto x_i, x_j \mapsto x_i - \varepsilon x_j\), the integral equals to

\[
\int_{\mathbb{R}^n} f(t, x_1, \ldots, x_i, \ldots, x_i - \varepsilon x_j, \ldots, x_n) \, dx_i = \sum_{1 \leq i < j \leq n} \int_{\mathbb{R}^n} f(t, x_1, \ldots, x_i, \ldots, x_i - \varepsilon x_j, \ldots, x_n) \, dx_i.
\]

By Corollary 3.4, we have

\[
Q_{n;\varepsilon}(t, x_1, \ldots, x_i, \ldots, x_i - \varepsilon x_j, \ldots, x_n) \leq C T^{\frac{1}{2}} \prod_{\ell \neq j} G_t * \delta_0(x_\ell),
\]

where we also used the elementary estimate \(G_t * \delta_0(x_i - \varepsilon x_j) \leq C T^{\frac{1}{2}}\), which clearly holds for the two cases of \(\delta_0\) we considered in the paper: \(\delta_0 \in C_c(\mathbb{R})\) or \(\delta_0(x) = \delta(x)\). For fixed \(t \in (0, T)\) and \((x_1, \ldots, x_n) \in \mathbb{R}^n\), by the continuity of \(f\), Lemma 3.3 and Corollary 3.4, we obtain

\[
f(t, x_1, \ldots, x_i, \ldots, x_i - \varepsilon x_j, \ldots, x_n) Q_{n;\varepsilon}(t, x_1, \ldots, x_i, \ldots, x_i - \varepsilon x_j, \ldots, x_n)
\]

\[
\to f(t, x_1, \ldots, x_i, \ldots, x_i - \varepsilon x_j, \ldots, x_n) Q_n(t, x_1, \ldots, x_i, \ldots, x_i - \varepsilon x_j, \ldots, x_n), \quad \text{as } \varepsilon \to 0.
\]

Note that \(\int R = 1\), we can apply dominated convergence theorem to conclude that

\[
\int_0^T \langle f_{0,\varepsilon}(t), Q_{n;\varepsilon}(t) \rangle \, dt \to \int_0^T \langle f_{0,\varepsilon}(t), Q_n(t) \rangle \, dt, \quad \text{as } \varepsilon \to 0.
\]

Here we recall from (1.6) that

\[
f_{0,\varepsilon}(t, x_{1:n}) = f(t, x_{1:n}) \sum_{1 \leq i < j \leq n} \delta(x_i - x_j).
\]

The proof is complete. \(\square\)
3.3. Proof of corollary 1.2

The proof of the cases \( R(\cdot) \in C^\infty_c(\mathbb{R}^d) \) and \( R(\cdot) = \delta(\cdot) \) are similar, and we only deal with the latter. Fix \( n = 1, q_0 \in C_c(\mathbb{R}^d) \) and \( f \in C^2_c(\mathbb{R}^d) \), by theorem 1.1, we have

\[
\langle f, Q(T) \rangle = \langle f, q_0 \rangle + \int_0^T \left( \frac{1}{2} \Delta f, Q_1(t) \right) dt + \beta^2 \sum_{k=0}^2 \int_0^T \langle f_{k(\cdot)}, Q_{1+k}(t) \rangle dt.
\]

Recall that \( F_f(\mu_T) = \langle f, \mu_T \rangle \), so we have

\[
\frac{E[F_f(\mu_T)] - F_f(\mu_0)}{T} = \frac{\langle f, Q(T) \rangle - \langle f, q_0 \rangle}{T} = \frac{1}{T} \int_0^T \left( \frac{1}{2} \Delta f, Q_1(t) \right) dt + \beta^2 \sum_{k=0}^2 \frac{1}{T} \int_0^T \langle f_{k(\cdot)}, Q_{1+k}(t) \rangle dt.
\]

By definition \( f_{0,1(\cdot)} = 0 \) when \( n = 1 \). For \( k = 1 \), we have

\[
\frac{1}{T} \int_0^T \langle f_{k(\cdot)}, Q_{1+k}(t) \rangle dt = -\frac{1}{T} \int_0^T \int_{\mathbb{R}^2} f(x_1) \delta(x_1 - x_2) Q_2(t, x_1, x_2) dx_1 dx_2 dt = -\frac{1}{T} \int_0^T \int_{\mathbb{R}} f(x_1) Q_2(t, x_1, x_1) dx_1 dt = -\int_0^1 \int_{\mathbb{R}} f(x_1) Q_2(1, x_1, x_1) dx_1 dt.
\]

Similarly, when \( k = 2 \), we have

\[
\frac{1}{T} \int_0^T \langle f_{k(\cdot)}, Q_{1+k}(t) \rangle dt = \frac{1}{T} \int_0^T \int_{\mathbb{R}^3} f(x_1) \delta(x_2 - x_3) Q_3(t, x_1, x_2, x_3) dx_1 dx_2 dx_3 dt = \frac{1}{T} \int_0^T \int_{\mathbb{R}^2} f(x_1) Q_3(t, x_1, x_2, x_2) dx_1 dx_2 dt = \int_0^1 \int_{\mathbb{R}^2} f(x_1) Q_3(1, x_1, x_2, x_2) dx_1 dx_2 dt.
\]

By applying corollary 3.4 and lemma 3.5 below, we have

\[
\frac{E[F_f(\mu_T)] - F_f(\mu_0)}{T} \to \left( \frac{1}{2} \Delta f, q_0 \right) - \beta^2 \int_{\mathbb{R}} f(x_1) q_0(x_1) x_1 dx_1
\]

\[
+ \beta^2 \int_{\mathbb{R}^2} f(x_1) q_0(x_1) q_0(x_2) x_1 dx_1 dx_2
\]

as \( T \to 0 \). The rhs equals to

\[
\left( \frac{1}{2} \Delta f, q_0 \right) + \beta^2 (f_{1,1(\cdot)}, q_0^3) + \beta^2 (f_{2,1(\cdot)}, q_0^3) = \left( \frac{1}{2} \Delta f, q_0 \right) + \beta^2 (f, T q_0),
\]

7352
which completes the proof of (1.8).

**Lemma 3.5.** Assume \( q_0 \in C_c(\mathbb{R}^d) \). For any \( n \geq 1 \) and \( (x_1, \ldots, x_n) \in \mathbb{R}^{nd} \), as \( t \to 0 \),

\[
Q_n(t, x_1, \ldots, x_n) \to \prod_{j=1}^{n} q_0(x_j).
\]

The proof of lemma 3.5 is given in appendix A.

### 4. Quantitative central limit theorem: proofs of theorems 1.3 and 1.4

In this section, we consider the high dimensions \( d \geq 3 \) and a high temperature regime with \( \beta \ll 1 \). The goal is to prove theorems 1.3 and 1.4. With a change of variable, (1.12) and (1.13) are equivalent with, for any \( h \in \text{Lip}(1) \),

\[
\left| \int_{\mathbb{R}^d} h(\epsilon x) Q_1 \left( \frac{1}{\epsilon^2}, x \right) dx - \int_{\mathbb{R}^d} h(x) G_1(x) dx \right| \leq C \left( \epsilon \right) |\log \epsilon| \mathbb{I}_{d=3} + \epsilon \mathbb{I}_{d \geq 4} ,
\]

(4.1)

and

\[
\left| \int_{\mathbb{R}^d} |\epsilon x|^2 Q_1 \left( \frac{1}{\epsilon^2}, x \right) dx - d \right| \leq C \left( \epsilon \right) \mathbb{I}_{d=3} + \epsilon^2 |\log \epsilon| \mathbb{I}_{d=4} + \epsilon^2 \mathbb{I}_{d \geq 5} .
\]

(4.2)

To unify the notation, we view (4.2) as a special case of (4.1) with \( h(x) = |x|^2 \) even though this choice of \( h \) is not an element of Lip(1). Note that, although the function \( h \) here is not necessarily bounded, it grows at most polynomially at infinity. Thus, by corollary 3.4, it is easy to see that the two integrals in (4.1) are both well-defined. Our proof below is based on selecting appropriate test functions in the equation satisfied by \( Q_t \) in order to quantify the cancellation between the \( Q_2 \) and \( Q_3 \) terms.

Throughout the section, we assume that \( q_0(\cdot) = \delta(\cdot) \); that is, the starting point of the polymer path is at the origin. Recall that in high dimensions we assumed the random environment is smooth in the spatial variable and \( R(\cdot) \in C_c^\infty(\mathbb{R}^d) \) is the spatial covariance function.

#### 4.1. Error form

The first step is to derive an exact error expression in (4.1) using the hierarchical PDE system. We define an auxiliary test function as follows: for any \( \epsilon > 0 \) and a function \( h \), let \( f_\epsilon(t,x) \) be the solution to the backward heat equation

\[
\partial_t f_\epsilon(t,x) + \frac{1}{2} \Delta f_\epsilon(t,x) = 0, \quad t < \frac{1}{\epsilon^2}, \quad x \in \mathbb{R}^d,
\]

\[
f_\epsilon \left( \frac{1}{\epsilon^2}, x \right) = h(\epsilon x).
\]

(4.3)
Then we have

**Lemma 4.1.** For a continuous function $h$ with at most polynomial growth at infinity, we have

$$
\mathcal{E}_\epsilon(h) := \int_{\mathbb{R}^d} h(\epsilon x) Q_1 \left( \frac{1}{\epsilon^2}, x \right) dx - \int_{\mathbb{R}^d} h(x) G_1(x) dx
$$

$$= \beta^2 \int_0^{\epsilon^{-2}} \int_{\mathbb{R}^d} [f(\epsilon t, x) - f_\epsilon(t, y)] R(y - z) Q_3(t, x, y, z) dy dz dt.
$$

(4.4)

**Proof.** We first assume in addition that $h \in C^1_b(\mathbb{R}^d)$. As $f_\epsilon$ solves the backward heat equation, it holds that $f_\epsilon \in C^1_b((0, \epsilon^{-2}] \times \mathbb{R}^d)$. In the hierarchical PDE system (1.5), we take $n = 1$, $T = \epsilon^{-2}$, and the test function to be $f_\epsilon$ to obtain

$$
\int_{\mathbb{R}^d} f_\epsilon \left( \frac{1}{\epsilon^2}, x \right) Q_1 \left( \frac{1}{\epsilon^2}, x \right) dx = \int_{\mathbb{R}^d} f_\epsilon(0, x) Q_1(0, x) dx
$$

$$- \beta^2 \int_0^{\epsilon^{-2}} \int_{\mathbb{R}^d} f_\epsilon(t, x) R(x - y) Q_3(t, x, y, y) dx dy dt
$$

$$+ \beta^2 \int_0^{\epsilon^{-2}} \int_{\mathbb{R}^d} f_\epsilon(t, x) R(y - z) Q_3(t, x, y, z) dy dz dt.
$$

(4.5)

As $Q_1(0, x) = q_0(x) = \delta(x)$, the first term on the rhs of (4.5) is

$$
\int_{\mathbb{R}^d} f_\epsilon(0, x) Q_1(0, x) dx = f_\epsilon(0, 0) = \int_{\mathbb{R}^d} h(\epsilon x) G_{\epsilon^{-2}}(x) dx = \int_{\mathbb{R}^d} h(x) G_1(x) dx,
$$

(4.6)

where the last step is through a change of variable and using the scaling property of the heat kernel.

By definition, $Q_3(t, x_1, \ldots, x_n)$ is symmetric in the $x$-variables and

$$
Q_3(t, x, y) = \int_{\mathbb{R}^d} Q_3(t, x, y, z) dz.
$$

Thus, (4.5) can be rewritten as

$$
\mathcal{E}_\epsilon(h) = \beta^2 \int_0^{\epsilon^{-2}} \int_{\mathbb{R}^d} [f_\epsilon(t, x) - f_\epsilon(t, y)] R(y - z) Q_3(t, x, y, z) dy dz dt.
$$

where we also used the fact that $R(\cdot)$ is even. Through an approximation and the bound on $Q_n$, given in corollary 3.4, the above identity extends to the case of $h$ having at most polynomial growth at infinity, which completes the proof.

**Remark 4.2.** For the case of $q_0(\cdot) \in C_c(\mathbb{R}^d)$, a similar error decomposition as (4.4) can be derived. The only change to make in the proof is in (4.6), where an extra error term comes out of the weak convergence of $\frac{1}{\epsilon^2} q_0(\frac{x}{\epsilon}) \rightarrow \delta(x)$. 

7354
4.2. Estimating $\mathcal{E}(h)$

The proofs of theorems 1.3 and 1.4 reduce to the estimate of $\mathcal{E}(h)$ for $h \in \text{Lip}(1)$ and $h(x) = |x|^2$ respectively. By using a probabilistic representation, the following bounds on $Q_n$ hold in the high temperature regime in $d \geq 3$:

**Lemma 4.3.** For any $d \geq 3$ and $n \geq 1$, there exists constants $\beta_0(d,n,R) > 0$ and $C(d,n,R,\beta)$ such that if $\beta < \beta_0(d,n,R)$, we have

$$Q_n(t,x_1,\ldots,x_n) \leq C \prod_{j=1}^{n} G_t(x_j), \quad \text{for all } t > 0, \quad x_1,\ldots,x_n \in \mathbb{R}^d.$$  

The proof of lemma 4.3 is in appendix B.

Before undertaking the proofs of theorems 1.3 and 1.4, we provide a heuristic argument that shows how the diffusive behavior of the polymer endpoint behavior follows from the convergence of the error $\mathcal{E}_c(h)$ to zero. Recall that

$$\mathcal{E}_c(h) = \beta^2 \int_0^\infty \int_{\mathbb{R}^d} [f_x(t,x) - f_x(t,y)] R(y-z) Q_3(t,x,y,z) dx \, dy \, dz \, dt.$$  

Consider the simple case of $h \in C_b(\mathbb{R}^d)$, so $|f_x| \leq \sup_x |h(x)|$. The key point here is that, with the assumption of $\beta \ll 1$ and $d \geq 3$,

$$\int_{\mathbb{R}^d} R(y-z) Q_3(t,x,y,z) dx \, dy \, dz = \int_{\mathbb{R}^d} R(y-z) Q_2(t,y,z) dy \, dz \leq C t^{-d/2}$$  

and, hence, is integrable for $t \in [1, \infty)$. This is ultimately related to the fast decay of the heat kernel in high dimensions $d \geq 3$, with the smallness of $\beta$ ensuring that the effect of the random environment is ‘summable’ in the limit (hidden in the proof of lemma 4.3). Therefore, the main contribution to $\mathcal{E}_c(h)$ actually comes from the time integration in a microscopically large domain $[0,M]$ for $1 \ll M \ll \varepsilon^{-2}$; that is

$$\mathcal{E}_c(h) \approx \beta^2 \int_0^M \int_{\mathbb{R}^d} [f_x(t,x) - f_x(t,y)] R(y-z) Q_3(t,x,y,z) dx \, dy \, dz \, dt.$$  

On the other hand, when $t \leq M \ll \varepsilon^{-2}$, the $f$ terms cancel for the following reason: for any fixed $(t,x)$, it is straightforward to check that

$$f_x(t,x) = \int_{\mathbb{R}^d} h(\varepsilon z) G_{\varepsilon^{-2}}(x-z) \, dz$$  

$$= \int_{\mathbb{R}^d} h(z) G_1 (\varepsilon z - x) \, dz \rightarrow \int_{\mathbb{R}^d} h(z) G_1(z) \, dz,$$

which is independent of $x$. Thus, by the dominated convergence theorem, we obtain that $\mathcal{E}_c(h) \rightarrow 0$. The proofs of theorems 1.3 and 1.4 then rely on quantifying this argument, which we do now.

In the proofs below, $C$ is a constant independent of $\varepsilon$ which may change from line to line.

**Proof of theorem 1.3.** Fix $h \in \text{Lip}(1)$, we have

$$|f_x(t,x) - f_x(t,y)| \leq \int_{\mathbb{R}^d} G_{\varepsilon^{-2}}(z)[h(\varepsilon(x-z)) - h(\varepsilon(y-z))] \, dz \leq \varepsilon |x-y|.$$
which implies
\[
|E_\varepsilon(h)| \leq \beta^2 \varepsilon \int_0^{e^{-2}} \int_{\mathbb{R}^d} |x - y| R(y - z) Q_3(t, x, y, z) dx dy dz \quad (4.7)
\]
\[
\leq C \varepsilon \int_0^{e^{-2}} \int_{\mathbb{R}^d} |x - y| R(y - z) G(x) G(y) G(z) dx dy dz.
\]

While the above integral can be estimated directly (as in the proof of theorem 1.4 below), we present a simple probabilistic argument here. Let \(B^1, B^2, B^3\) be three independent Brownian motions, then the integral can be written as
\[
\int_{\mathbb{R}^d} |x - y| R(y - z) G_t(x) G_t(y) G_t(z) dx dy dz = \mathbb{E}[|B_t^2 - B_t^1| \cdot R(B_t^3 - B_t^1)].
\]

Since \(B^2 - B^1\) and \(B^3 - B^1\) are correlated Brownian motions with variance 2 and covariance 1, we can rewrite the expectation as
\[
\mathbb{E}[|B_t^2 - B_t^1| \cdot R(B_t^3 - B_t^1)] = \mathbb{E} \left[ \sqrt{\frac{1}{2} W_t^1} - \sqrt{\frac{3}{2} W_t^2} \cdot R(\sqrt{2} W_t^1) \right],
\]
with \(W^1, W^2\) independent Brownian motions. Since \(R \in C^\infty_c\), we estimate the above expectation by
\[
\mathbb{E} \left[ \sqrt{\frac{1}{2} W_t^1} - \sqrt{\frac{3}{2} W_t^2} \cdot R(\sqrt{2} W_t^1) \right] \leq C \mathbb{E} \left[ \sqrt{\frac{1}{2} W_t^1} - \sqrt{\frac{3}{2} W_t^2} \cdot 1_{|[W^2]| \leq M} \right]
\]
for some constant \(M > 0\). For \(t \leq 1\), we have the obvious bound
\[
\mathbb{E} \left[ \sqrt{\frac{1}{2} W_t^1} - \sqrt{\frac{3}{2} W_t^2} \cdot 1_{|[W^2]| \leq M} \right] \leq C.
\]

For \(t > 1\), by first averaging \(W^2\), we have
\[
\mathbb{E} \left[ \sqrt{\frac{1}{2} W_t^1} - \sqrt{\frac{3}{2} W_t^2} \cdot 1_{|[W^2]| \leq M} \right] \leq C \sqrt{7} \mathbb{P}(|W_t^1| \leq M) \leq C t^{-\frac{1}{4}}.
\]

Combining the two cases and plugging into (4.7), we derive
\[
|E_\varepsilon(h)| \leq C \varepsilon \left( 1 + \int_1^{e^{-2}} t^{-\frac{1}{4}} dt \right) \leq C (\varepsilon |\log \varepsilon| 1_{d=3} + \varepsilon 1_{d \geq 4}).
\]

The proof is complete. \(\square\)

**Proof of theorem 1.4.** Let \(h(x) = |x|^2\), then
\[
f_\varepsilon(t, x) = \int_{\mathbb{R}^d} G_{e^{-2}t}(y)|\varepsilon(x - y)|^2 dy = \varepsilon^2 |x|^2 + (1 - \varepsilon^2 t)d.
\]

7356
Applying lemmas 4.1 and 4.3, we have
\[|E_ε(h)| = \left| \beta^2 ε^2 \int_{0}^{\infty} \int_{\mathbb{R}^d} (|x|^2 - |y|^2) R(y-z) Q_3(t, x, y, z) dx \, dy \, dz \, dr \right| \]
\[\leq C ε^2 \int_{0}^{\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|x|^2 - |y|^2) R(y-z) G_t(x) G_t(y) G_t(z) dx \, dy \, dz \, dr. \tag{4.8} \]

To estimate the above integral, as $R$ is compactly supported (suppose its support has a radius $M > 0$), we have
\[\int_{\mathbb{R}^d} R(y-z) G_t(z) dz \leq C \int_{\mathbb{R}^d} I_{|z| \leq M} G_t(z) dz \]
\[\leq C \left( I_{t \leq 1} + I_{t > 1} \sup_{z \in \mathbb{R}^d} G_t(z) \right). \]

For the case of $t > 1$, by considering $|y| \leq 2M$ and $|y| > 2M$ separately, we derive
\[\sup_{z \in \mathbb{R}^d} G_t(z) \leq CG_{c_1 t}(z), \text{ for some } c_1 > 0. \]

From (4.8) and the above estimate, the following bound holds
\[|E_ε(h)| \leq C ε^2 \int_{0}^{\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|x|^2 - |y|^2) G_t(x) G_t(y) dx \, dy \, dz \, dr \]
\[+ C ε^2 \int_{0}^{\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|x|^2 - |y|^2) G_t(x) G_t(y) G_{c_1 t}(y) dx \, dy \, dz \, dr = A_1 + A_2. \]

We first get $A_1 \leq C ε^2$. For $A_2$, by the fact that
\[G_t(y) G_{c_1 t}(y) \leq C t^{-d/2} G_{c_2 t}(y), \text{ for some } c_2 > 0, \]
we have
\[A_2 \leq C ε^2 \int_{1}^{\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} t^{-d/2} (|x|^2 - |y|^2) G_t(x) G_{c_1 t}(y) dx \, dy \, dz \, dr \]
\[\leq C ε^2 \int_{1}^{\infty} t^{-d/2} dt \leq C \left( ε^2 I_{d=3} + ε^2 \log ε I_{d=4} + ε^2 I_{d>5} \right). \]

This completes the proof. \(\square\)

5. Growth of moments: proof of theorem 1.5

In the interest of the simplest presentation, we remove the parameter $β$ by scaling. Indeed, let
\[\overline{g}(t, x) = \beta^{-2} g(t\beta^{-4}, x\beta^{-2}), \text{ and observe that} \]
\[\overline{g}_t = \overline{g}_{xx} + \overline{g} \|\overline{g}\|^2 - \overline{g}. \]
Hence, for the remainder of the section, we set $\beta = 1$; that is, we are interested in
\begin{equation}
\begin{aligned}
\partial_t g(t, x) &= \frac{1}{2} \Delta g(t, x) + \|g(t, \cdot)\|^2 g(t, x) - g(t, x)^2, \\
g(0, x) &= q_0(x).
\end{aligned}
\tag{5.1}
\end{equation}

Here we abused notation by reverting to $g$ as opposed to using $\overline{g}$. Undoing this simple scaling reveals the dependence on $\beta$ of our results.

In order to control the moments of $g$, it is necessary to understand the asymptotic behavior of $\|g\|^2$ as $t \to \infty$. By interpolation and the fact that $g$ is a probability density (noted below), it is enough to control the maximum of $g$. In the following section, we state the main estimate on the decay of the maximum of $g$. After, we show how to use this to obtain upper and lower bounds on the moments of $g$ by constructing sharp sub and supersolutions of $g$. In section 5.2, we show how to obtain the correct asymptotics on the maximum of $g$.

5.1. Statement of the main inequality and its application to the moments of $g$

In order to streamline the argument, we define a few quantities that play key roles in the proof. For any $t \geq 0$, let
\begin{equation}
M(t) = \max_{x \in \mathbb{R}} g(t, x), \quad E(t) = \int_\mathbb{R} g(t, x)^2 dx, \quad D(t) = \int_\mathbb{R} |g_x(t, x)|^2 dx.
\end{equation}

The key inequality that we require is stated in the following proposition, proved in section 5.2.

**Proposition 5.1.** There is a universal constant $C_0$, independent of the initial data, such that
\[ M(t) \leq \frac{C_0}{t^{1/3}} \] for all $t > 0$.

Two more useful facts are the following. Integrating (5.1), we see that
\[ \frac{d}{dt} \int_\mathbb{R} g(t, x) dx = E(t) \left( \int_\mathbb{R} g(t, x) dx - 1 \right). \]

Since $\int_\mathbb{R} g(0, x) dx = \int_\mathbb{R} q_0(x) dx = 1$, by assumption, a simple ODE argument yields, for any $t \geq 0$,
\[ \int_\mathbb{R} g(t, x) dx = \int_\mathbb{R} q_0(x) dx = 1. \]

Thus $g(t, \cdot)$ is a probability density (note that $g \geq 0$ by comparison principle and the fact that $q_0 \geq 0$). This is unsurprising given the derivation of the model (5.1); however, it is crucial in our analysis. Indeed, we immediately deduce the following useful inequality:
\[ E(t) \leq M(t) \int_\mathbb{R} g(t, x) dx = M(t). \]
We now show how to conclude theorem 1.5 assuming proposition 5.1. We begin with the upper bound.

**Proof of the upper bound in theorem 1.5.** The first step is to replace the $\|g\|^2 = E$ term in (5.1). From proposition 5.1 and (5.4), we see that

$$E(t) \leq \frac{C_0}{t^{2/3}}.$$  

This, along with (5.1), implies that

$$\partial_t g - \frac{1}{2} \Delta g - \frac{C_0}{t^{2/3}} g \leq 0.$$  

The comparison principle implies that $g \leq \overline{g}$, where $\overline{g}$ solves

$$\begin{cases}
\partial_t \overline{g} - \frac{1}{2} \Delta \overline{g} - \frac{C_0}{t^{2/3}} \overline{g} = 0 & \text{in } (0, \infty) \times \mathbb{R}, \\
\overline{g} = q_0 & \text{on } 0 \times \mathbb{R}.
\end{cases} \quad (5.5)$$

The second step is to obtain a bound on $\overline{g}$, and hence, on $g$, for large $x$. The first thing to notice is that

$$\overline{g}(t, x) = \exp \left\{ - \int_0^t \frac{C_0}{s^{2/3}} ds \right\} \overline{g}(t, x)$$

solves the heat equation, $\partial_t \overline{g} = \frac{1}{2} \Delta \overline{g}$. It follows that

$$\overline{g}(t, x) = \exp \left\{ \int_0^t \frac{C_0}{s^{2/3}} ds \right\} \overline{g}(t, x) = \exp \left\{ \int_0^t \frac{C_0}{s^{2/3}} ds \right\} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2}} q_0(x-y) dy.$$

By assumption, $q_0$ is compactly supported. A straightforward estimate of the convolution, as well as a simple evaluation of the time integral, yields, for any $t \geq 1$ and any $x$,

$$g(t, x) \leq \overline{g}(t, x) \leq \frac{C}{\sqrt{t}} e^{3C_0 t^{1/3} - \frac{x^2}{2}}$$

for some positive constant $C$ depending only on the initial data.

We now conclude the bound on the moments of $g$. Pairing the above arguments with (5.1), we have established that, for all $t \geq 1$,

$$g(t, x) \leq C \min \left\{ \frac{1}{t^{2/3}}, \frac{1}{t^{1/2}} \exp \left\{ 3C_0 t^{1/3} - \frac{x^2}{2t} \right\} \right\}.$$

We now use this to conclude the proof. Indeed, for any $p \geq 1$, we find

$$\int_{\mathbb{R}} |x|^p g(t, x) dx \leq \int_{|x| \leq 6C_0 t^{2/3}} |x|^p \frac{C}{t^{2/3}} dx + \int_{|x| > 6C_0 t^{2/3}} |x|^p \frac{C}{t^{1/2}} e^{3C_0 t^{1/3} - \frac{x^2}{2}} dx$$

$$\leq Ct^2 + C t^{p/2} \int_{|x| > 6C_0 t^{2/3}} |y|^p e^{3C_0 t^{1/3} - \frac{x^2}{2}} dy$$

$$\leq Ct^2 + C t^{p/2} e^{(p-1)/6} e^{-2(3C_0 t^{1/3} - (3C_0 t^{1/3})).$$
where $C$ is a constant depending only on $q_0$ and $p$ that changes line-by-line. The second term clearly tends to zero. This completes the proof.

It is now possible to deduce the lower bound using the upper bound given in proposition 5.1 and (5.3). We require one lemma.

**Lemma 5.2.** Assume that $\lambda > 0$, $d \geq 1$, and $w : [0, \infty) \to \mathbb{R}$ is an increasing function. Then

$$
\min_{0 \leq g \leq \lambda} \int_{B_r} w(|x|)g(x)dx = \lambda \int_{B_{\lambda}^{1/d} - 1/d} w(|x|)dx.
$$

where $B_r$ denotes the ball centered at the origin with radius $r$ and $\omega_d$ is the volume of the unit ball in $\mathbb{R}^d$.

This lemma is elementary and follows from the fact that the minimiser is clearly $\lambda 1_{B_{\lambda}^{1/d} - 1/d}(x)$. Hence, we omit its proof. We now conclude the proof of theorem 1.5.

**Proof of the lower bound in theorem 1.5.** From proposition 5.1, we know that $g(t, x) \leq \frac{C_0}{(1 + t)^{2/3}} = : \lambda$ for all $t$. Hence

$$
\int_{\mathbb{R}} \left| x \right|^p g(t, x)dx \geq \min_{0 \leq g \leq \lambda} \int_{\mathbb{R}} \left| x \right|^p g(x)dx.
$$

Applying lemma 5.2, we have

$$
\min_{0 \leq g \leq \lambda} \int_{\mathbb{R}} \left| x \right|^p g(x)dx \geq \lambda \int_{\mathbb{R}^{1/2\lambda}} \left| x \right|^p dx = \int_{-1/2\lambda}^{1/2\lambda} \left| x \right|^p dx = \frac{2^{-p+1}}{p+1} \left( \frac{1+1^{2/3}}{C_0} \right)^p.
$$

which concludes the proof.

5.2. Decay of the maximum of $g$

Classical techniques for decay of parabolic equations are often based on Nash’s inequality, which relates the $L^2$ norm of the gradient of $g$ with the $L^2$ norm of $g$. Such an estimate necessarily gives decay like $O(t^{-1/2})$ in $d = 1$, which is slower than the rate of decay we prove below. Hence, such a strategy is not useful here.

In other words, the Laplacian term (and the related Dirichlet energy $D$) is not sufficient to obtain decay like $O(t^{-2/3})$. The only other term in the equation is $g(E - g)$, and, hence, our proof must be based on this term. The key observation is that near the maximum of $g$, we expect $g(E - g) \approx M(E - M) < 0$. As such, we require an estimate that quantifies how negative this term is.

In fact, our argument is more subtle than this. We use the decay induced by both terms $-D$ and $-M(M - E)$. Indeed, if $M - E$ is large, then the nonlinear term $-M(M - E)$ is a large negative number. On the other hand, if $M \approx E$, it must be that $g$ ‘flattens’ quickly after reaching the maximum, making $D$ large (recall that $g$ is a probability measure so if $M \approx E$ then $g$ is near the optimal case in Hölder’s inequality, which, in turn, implies that $g$ is nearly an indicator
function). In both cases, we get a large decay term. The key estimate quantifying this heuristic is the following, which is proved at the end of the section.

**Lemma 5.3.** There is a universal constant $C_1 > 0$ such that

$$M - E \geq \frac{M^4}{C_1 D}.$$ 

Before beginning the proof of proposition 5.1, we collect two more inequalities. The first is that, for any $0 < t_1 < t_2$,

$$M(t_2) \leq M(t_1) + \int_{t_1}^{t_2} (E(s)M(s) - M(s)^2)ds. \quad (5.6)$$

Informally, this can be seen by noting that, at a maximum, $\Delta g \leq 0$, so that (5.1) reads

$$\dot{M} \leq EM - M^2,$$

where we used the physics notation $\cdot$ to denote the time derivative. This differential inequality has to be interpreted in the suitable weak sense, but this purely technical issue is standard in parabolic theory and, hence, we omit the details.

The second inequality is, for all $t_1 < t_2$,

$$E(t_2) \leq E(t_1) - \int_{t_1}^{t_2} D(s)ds. \quad (5.7)$$

In order to see this, simply multiply (5.1) by $g$ and integrate in $x$ in order to obtain

$$\frac{1}{2} \dot{E} + \frac{1}{2} D \leq g^2 - \int g^3.$$

Since

$$E(t) = \int_{\mathbb{R}} g^{3/2}(t, x)g^{1/2}(t, x)dx \leq \left( \int_{\mathbb{R}} g^2(t, x)dx \right)^{1/2} \left( \int_{\mathbb{R}} g(t, x)dx \right)^{1/2},$$

then $\frac{1}{2} \dot{E} + \frac{1}{2} D \leq 0$. Integrating this in time yields (5.7).

**Proof of proposition 5.1.** Let

$$t_0 = \sup \left\{ t > 0 : \sup_{s \in [0,t]} s^{2/3}M(s) < A \right\}, \quad A = 2C_1^{1/3},$$

with the $C_1$ from lemma 5.3. It is clear that if $t_0 = \infty$ then the proof is finished. We proceed by contradiction assuming that $t_0$ is finite. By continuity, it is also clear that $M(t_0) = A_{t_0}^{2/3}$. There are two cases to consider.

**Case one:** $M(t_0/2) > 2M(t_0)$. Since $t_0/2 < t_0$, then, using the definition of $t_0$, we have

$$\frac{2A}{t_0^{2/3}} = 2M(t_0) < M(t_0/2) < \frac{A}{(t_0/2)^{2/3}} = \frac{2^{2/3}A}{t_0^{2/3}}.$$ 

This is a contradiction since $2 > 2^{2/3}$. Hence, this case cannot occur.

**Case two:** $M(t_0/2) \leq 2M(t_0)$. We first combine lemma 5.3 and (5.6) to find

$$M(t) \leq M(t_0/2) - \frac{1}{C_1} \int_{t_0/2}^{t} \frac{M(s)^2}{D(s)}ds.$$
Since this is true for all $t$, it follows that

$$M(t_0) \leq \overline{M}(t_0),$$

where $\overline{M} = -C_1 M^4 D^{-1}$ and $\overline{M}(t_0/2) = M(t_0/2)$. Elementary calculus yields

$$M(t_0) \leq \overline{M}(t_0) = \left( M(t_0/2)^4 + \frac{4}{C_1} \int_{t_0/2}^{t_0} D(s)^{-1} \, ds \right)^{-1/4}. \tag{5.8}$$

Then, using (in order) Cauchy–Schwarz, (5.7), (5.4), and the assumption that $M(t_0/2) \leq 2M(t_0)$, we find

$$\frac{1}{4} t_0^2 = \left( \int_{t_0/2}^{t_0} \sqrt{D(s)} \, ds \right)^2 \leq (E(t_0/2) - E(t_0)) \int_{t_0/2}^{t_0} D(s)^{-1} \, ds \leq E(t_0/2) \int_{t_0/2}^{t_0} D(s)^{-1} \, ds \leq 2M(t_0) \int_{t_0/2}^{t_0} D(s)^{-1} \, ds. \tag{5.9}$$

Using this inequality in (5.8), we obtain

$$M(t_0) \leq \left( \frac{1}{M(t_0/2)^4} + \frac{1}{2C_1 M(t_0)} \right)^{-1/4} \leq \left( \frac{1}{2C_1 M(t_0)} \right)^{1/2} = \left( \frac{2C_1 M(t_0)}{t_0^2} \right)^{1/2}. \tag{5.10}$$

Re-arranging this yields

$$M(t_0) < \frac{(2C_1)^{1/3}}{t_0^{1/3}}. \tag{5.11}$$

However, by the construction of $t_0$, we have that $M(t_0) = At_0^{-2/3} = 2C_1 t_0^{-1/3} t_0^{-2/3}$. Hence, we have reached a contradiction, and we conclude that case two cannot occur either.

Since both cases yield a contradiction, it follows that $t_0 = \infty$, which completes the proof. \qed

It only remains to establish lemma 5.3. We do this now. The idea of the proof is to re-write $M - E$ in terms of a single integral term and then use the proof of a lemma of Constantin, Kiselev, Oberman, and Ryzhik [21, lemma 2]. This lemma was originally used to establish the key inequality in a proof of lower bounds on the speed of Fisher–KPP fronts in the presence of shear flows in a cylinder.

**Proof of lemma 5.3.** As time plays no role in this lemma, we omit it notationally. First, observe that, due to (5.3), we have

$$M - E = M \int_R g(x) \, dx - \int_R g(x)^2 \, dx = \int_R g(x)(M - g(x)) \, dx.$$ 

Notice that the integrand $g(M - g)$ is nonnegative.
Since $M$ is the maximum of $g$ and $\lim_{x \to -\infty} g(x) = 0$, we can find $x_1 < x_2$ such that
\[
g(x_1) = \frac{M}{3}, \quad g(x_2) = \frac{2M}{3}, \quad \text{and} \quad \frac{M}{3} \leq g(x) \leq \frac{2M}{3} \quad \text{for all} \quad x \in (x_1, x_2).
\]
\begin{equation}
(5.12)
\end{equation}

Then we have that
\[
\frac{M}{3} = \int_{x_1}^{x_2} g_x \, dx \leq \sqrt{x_2 - x_1} \left( \int_{\mathbb{R}} |g_x|^2 \, dx \right)^{1/2}.
\]
\begin{equation}
(5.13)
\end{equation}

On the other hand, since $M/3 \leq g(x) \leq 2M/3$ for all $x \in (x_1, x_2)$, then $g(M - g) \geq M^2/9$ on $(x_1, x_2)$. It follows that
\[
\int_{\mathbb{R}} g(x)(M - g(x)) \, dx \geq \int_{x_1}^{x_2} g(x)(M - g(x)) \, dx \geq \int_{x_1}^{x_2} \frac{M^2}{9} \, dx = \frac{M^2}{9} |x_2 - x_1|.
\]
\begin{equation}
(5.14)
\end{equation}

After squaring (5.13) and inserting (5.14) into it, we find
\[
\frac{M^2}{9} \leq \frac{9}{M^2} \int_{\mathbb{R}} g(x)(M - g(x)) \, dx \int_{\mathbb{R}} |g_x|^2 \, dx,
\]
\begin{equation}
(5.15)
\end{equation}

which yields the claim. □

Acknowledgments

YG was partially supported by the NSF through DMS-1907928 and the Center for Nonlinear Analysis of CMU. CH was partially supported by NSF Grant DMS-2003110. YG would like to thank Xi Geng for several discussions.

Appendix A. Basics about stochastic heat equation

For the convenience of the reader, we present some standard facts about the stochastic heat equation. We will first discuss the case of the spacetime white noise in $d = 1$, then that of colored noise in $d \geq 1$.

A.1. Spacetime white noise in $d = 1$

For readability, we define the following notation. For any $\varepsilon > 0$ and for any $f$ such that either $0 \leq f \in C_0(\mathbb{R})$ or $f(\cdot) = \delta(\cdot)$, we define $u_{\varepsilon, f}$ and $u_f$ as the solutions to the following equations in $d = 1$:
\[
\begin{align*}
\partial_t u_{\varepsilon, f}(t, x) &= \frac{1}{2} \Delta u_{\varepsilon, f}(t, x) + \beta u_{\varepsilon, f}(t, x)\eta(t, x), & u_{\varepsilon, f}(0, x) = f(x), \\
\partial_t u_f(t, x) &= \frac{1}{2} \Delta u_f(t, x) + \beta u_f(t, x)\eta(t, x), & u_f(0, x) = f(x),
\end{align*}
\]
where $\eta$ is the spacetime white noise, $\eta_\varepsilon$ is the spatial mollification of $\eta$, and they are time reversals of $\xi, \xi_\varepsilon$:

$$\eta(t,x) = \xi(-t,x), \quad \eta_\varepsilon(t,x) = \xi_\varepsilon(-t,x).$$

**Lemma A.1.**

(a) For any $p \geq 1, T > 0$, there exists $C = C(p, \beta, T) > 0$ such that

$$E[u_{\varepsilon, f}(t, x)^p] + E[u_f(t, x)^p] \leq C(G_t * f(x))^p,$$

for all $\varepsilon \in (0,1), t \in (0, T], x \in \mathbb{R}$;

(b) $u_{\varepsilon, f}(t, x) \to u_f(t, x)$ in $L^p(\Omega)$ as $\varepsilon \to 0$, uniformly in $x \in \mathbb{R}$ and $t$ in compact subsets of $(0, \infty)$;

(c) For $f \in C_0(\mathbb{R})$, we have $u_f(t, x) \to f(x)$ in $L^p(\Omega)$ as $t \to 0$, for each $x \in \mathbb{R}$;

(d) $u_f \in C(0, \infty) \times \mathbb{R}, L^p(\Omega)$).

**Proof.** The moments bounds in (a) can be found$^3$ in [17, theorem 1.7]. The convergence in (b) was proved in [8, theorem 2.2]. For (c), we write

$$u_f(t, x) = G_t * f(x) + \beta \int_0^t \int_\mathbb{R} G_{t-s}(x-y) u_f(s, y) \eta(s, y) dy ds.$$ 

By the moments bounds in (a) and the BDG inequality, we can show the term of the stochastic integral $\int_0^t \int_\mathbb{R} G_{t-s}(x-y) u_f(s, y) \eta(s, y) dy ds$ goes to zero in $L^p(\Omega)$, as $t \to 0$. Finally, (d) comes from the standard moment estimate: for any $n \in \mathbb{N}$ and $s, t \in [n^{-1}, n], x, y \in \mathbb{R}$, there exists $C = C(n, p)$ such that

$$E[u_f(t, x) - u_f(s, y)]^p \leq C \left( |t-s|^p + |x-y|^p \right).$$

The proof is complete. \hfill $\Box$

Recall that

$$q_\varepsilon(t, x) = \frac{\int_{\mathbb{R}} U_\varepsilon(-t, x; 0, y) q_0(y) dy}{\int_{\mathbb{R}} U_\varepsilon(-t, x; 0, y) q_0(y) dy}, \quad q(t, x) = \frac{\int_{\mathbb{R}} U(-t, x; 0, y) q_0(y) dy}{\int_{\mathbb{R}} U(-t, x; 0, y) q_0(y) dy}$$

with $U_\varepsilon, U$ solving (3.10) and (2.2) respectively. Define

$$\tilde{q}_\varepsilon(t, x) = \frac{u_{\varepsilon, q_0}(t, x)}{\int_{\mathbb{R}} u_{\varepsilon, q_0}(t, x) dx}, \quad \tilde{q}(t, x) = \frac{u_{q_0}(t, x)}{\int_{\mathbb{R}} u_{q_0}(t, x) dx}. \quad \text{(A.1)}$$

**Lemma A.2.** For any $\varepsilon > 0$ and $t > 0, x \in \mathbb{R}$, we have

$$q_\varepsilon(t, x) = \tilde{q}_\varepsilon(t, x), \quad q(t, x) = \tilde{q}(t, x).$$

$^3$In [17], the bounds are stated only for $p \geq 2$. However, it is also shown that, in our setting, $u_f, u_{\varepsilon, f} > 0$. The bound for $p \in (1, 2)$ follows by Hölder’s inequality.
Proof. By the Feynman–Kac formula, we have
\[
\int_R U_c(-t, x; 0, y)q_0(y)dy = \int_R q_0(y)\mathbb{E}_x\left[\delta(u_t - x)e^{\beta \int_0^t \xi(s)ds - \frac{1}{2}\beta^2 R_c(0)t}\right] dy
\]
\[
= \int_R q_0(y)G_t(x - y)\mathbb{E}_y\left[e^{\beta \int_0^t \xi(s)ds - \frac{1}{2}\beta^2 R_c(0)t} \mid w_t = x\right] dy,
\]
where we recall that \(\mathbb{E}_x\) is the expectation with respect to \(w\) starting from \(w_0 = y\) and \(G_t(\cdot)\) is the standard heat kernel. Changing variables in the exponent \(s \mapsto t - s\) and using the time reversal \(\eta_\varepsilon(t, \cdot) = \xi(t - \cdot)\), we find
\[
\int_R U_c(-t, x; 0, y)q_0(y)dy = \int_R q_0(y)G_t(x - y)\mathbb{E}_y\left[e^{\beta \int_0^t \xi(s)ds - \frac{1}{2}\beta^2 R_c(0)t} \mid w_t = x\right] dy
\]
\[
= \int_R q_0(y)G_t(x - y)\mathbb{E}_y\left[e^{\beta \int_0^t \xi(s)ds - \frac{1}{2}\beta^2 R_c(0)t} \mid w_t = y\right] dy.
\]
On the other hand, we write
\[
u_{c, y_0}(t, x) = \mathbb{E}_x\left[\int_0^t q_0(w_s)\mathbb{E}_y\left[e^{\beta \int_0^t \xi(s)ds - \frac{1}{2}\beta^2 R_c(0)t} \mid w_t = y\right] ds\right]
\]
\[
= \int_R q_0(y)G_t(y - x)\mathbb{E}_y\left[e^{\beta \int_0^t \xi(s)ds - \frac{1}{2}\beta^2 R_c(0)t} \mid w_t = y\right] dy.
\]
The two expressions are equal to each other, so we have \(q_c(t, x) = \tilde{q}_c(t, x)\). By sending \(\varepsilon \to 0\), we also have \(q(t, x) = \tilde{q}(t, x)\), which completes the proof. \(\square\)

Lemma A.3. For any \(p \geq 1, T > 0\), there exists \(C = C(p, T)\) such that, for \(\varepsilon \in (0, 1), t \in [0, T]\), we have
\[
\mathbb{E}\left[\left(\int_{\mathbb{R}^2} U_c(-t, \tilde{x}; 0, y)q_0(y)dy d\tilde{x}\right)^{-p}\right] + \mathbb{E}\left[\left(\int_{\mathbb{R}^2} U(-t, \tilde{x}; 0, y)q_0(y)dy d\tilde{x}\right)^{-p}\right] \leq C.
\]

Proof. Define \(U_{c, 1}(-t; s, y) := \int_{\mathbb{R}^2} U_c(-t, \tilde{x}; s, y)d\tilde{x}\), which solves
\[
\partial_t U_{c, 1} = \frac{1}{2} \Delta U_{c, 1} + \beta U_{c, 1} \xi(s, y), \quad U_{c, 1}(-t; -t, y) = 1.
\]
By the statistical shift-invariance of the noise, we have
\[
U_{c, 1}(-t; 0, \cdot) \overset{\text{law}}{=} u_{c, 1}(t, \cdot).
\]
for each $t > 0$. Combining this with lemma A.2, a simple observation is that, for each $t > 0$,
\[
\int_{\mathbb{R}} u_{\varepsilon,t}(y) q_0(y) dy \overset{\text{law}}{=} \int_{\mathbb{R}} u_{\varepsilon,0}(t, \tilde{x}) d\tilde{x}.
\]
(A.2)

By Jensen’s inequality, we have
\[
\left( \int_{\mathbb{R}} U_{\varepsilon,t}(-t; 0, y) q_0(y) dy \right)^{-p} \leq \int_{\mathbb{R}} U_{\varepsilon,t}(-t; 0, y)^{-p} q_0(y) dy,
\]
which implies
\[
E \left[ \left( \int_{\mathbb{R}^d} U_{\varepsilon,t}(-t; \tilde{x}, 0, y) q_0(y) dy d\tilde{x} \right)^{-p} \right] \leq \int_{\mathbb{R}} E[U_{\varepsilon,t}(-t; 0, y)^{-p}] q_0(y) dy \tag{A.3}
\]
where we used the stationarity of $U_{\varepsilon,t}(-t; 0, y)$ in the $y$ variable in the last step. The negative moment bound
\[
\sup_{\varepsilon \in (0,1), t \in (0,T]} E[u_{\varepsilon,t}(t, \cdot)^{-p}] \leq C(p, \beta, T)
\]
can be found, e.g. in [31, corollary 4.8], and hence the proof is complete. \hfill \square

A.2. Colored noise in $d \geq 1$

For colored noise $\eta(t, x)$, we consider the equation
\[
\partial_t u_f(t, x) = \frac{1}{2} \Delta u_f(t, x) + \beta u_f(t, x) \eta_0(t, x), \quad t > 0, x \in \mathbb{R}^d,
\]
\[
u_f(0, x) = f(x),
\]
where $d \geq 1$ and either $0 \leq f \in C_b(\mathbb{R}^d)$ or $f(x) = \delta(x)$. The follow lemma holds:

**Lemma A.4.** For any $p \geq 1$, $T > 0$, there exists $C = C(d, p, \beta, T) > 0$ such that
\[
E[u_f(t, x)^p] \leq C(G_t \ast (f(x))^p),
\]
\[
E[u_f(t, x)^{-p}] \leq C,
\]
for all $t \in (0, T], x \in \mathbb{R}^d$.

The positive moment bound can be found in [17, theorem 1.7], and the negative moment bound in [31, corollary 4.8].

**Appendix B. Proofs of lemmas 3.2, 3.3, 3.5 and 4.3**

**Proof of lemma 3.2.** By lemma A.2, we know that
\[
q_\varepsilon(t, x) = \frac{u_{\varepsilon,0}(t, x)}{\int_{\mathbb{R}} u_{\varepsilon,0}(t, \tilde{x}) d\tilde{x}} \quad \text{and} \quad q(t, x) = \frac{u_{0}(t, x)}{\int_{\mathbb{R}} u_{0}(t, \tilde{x}) d\tilde{x}}.
\]
Applying lemma A.1(a), lemma A.3 and Hölder’s inequality, we derive (3.13). To show the convergence of \( q(t, x) \rightarrow q(t, x) \), we write

\[
q(t, x) - q(t, x) = \frac{u_{c,q_0}(t, x) - u_{q_0}(t, x)}{\int_{\mathbb{R}} u_{c,q_0}(t, \tilde{x})dx} + \frac{u_{q_0}(t, x)}{\int_{\mathbb{R}} u_{q_0}(t, \tilde{x})dx} \left( \int_{\mathbb{R}} u_{q_0}(t, \tilde{x})d\tilde{x} - \int_{\mathbb{R}} u_{c,q_0}(t, \tilde{x})d\tilde{x} \right).
\]

The first term on the rhs goes to zero in \( L^p(\Omega) \) by lemma A.1(b) and lemma A.3. Together with (A.2), we can show the second term goes to zero similarly. \( \square \)

**Proof of lemma 3.3.** The continuity of \( Q_n \) follows from (3.13), Hölder’s inequality and the convergence of

\[
E[|q(t_n, x_n) - q(t, x)|^p] \rightarrow 0 \tag{B.1}
\]

if \((t_n, x_n) \rightarrow (t, x) \in (0, \infty) \times \mathbb{R} \) as \( n \rightarrow \infty \). To prove (B.1), it suffices to apply lemmas A.1(d), A.2 and A.3. We omit the details here. \( \square \)

**Proof of lemma 3.5.** By the assumption of \( q_0 \in C_c(\mathbb{R}) \) and lemma A.1(a), we know that \( E[|q(t, x)|^p] \leq C \) uniformly in \( t \in (0, T], x \in \mathbb{R} \). Then it suffices to show that for \( x \in \mathbb{R} \), \( q(t, x) \rightarrow q_0(x) \) in \( L^p(\Omega) \) as \( t \rightarrow 0 \). We write

\[
q(t, x) - q_0(x) = \frac{u_{q_0}(t, x) - q_0(x)}{\int_{\mathbb{R}} u_{q_0}(t, \tilde{x})d\tilde{x}} + \frac{q_0(x)}{\int_{\mathbb{R}} u_{q_0}(t, \tilde{x})d\tilde{x}} \left( 1 - \int_{\mathbb{R}} u_{q_0}(t, \tilde{x})d\tilde{x} \right),
\]

and the first term on the rhs goes to zero in \( L^p(\Omega) \) by lemmas A.1(c) and A.3. For the second term, the proof is similar with another use of (A.2) and the fact that \( \int q_0 = 1 \). \( \square \)

**Proof of lemma 4.3.** The goal is to show

\[
Q_n(t, x_1, \ldots, x_n) = E[q(t, x_1) \ldots q(t, x_n)] \leq C \prod_{j=1}^{n} G_j(x_j)
\]

in \( d \geq 3 \) when \( \beta \) is small. By lemma A.2, we have

\[
q(t, x) = \frac{u_{q_0}(t, x)}{\int_{\mathbb{R}} u_{q_0}(t, \tilde{x})d\tilde{x}},
\]

with \( u_{q_0} \) the solution to the stochastic heat equation starting from \( q_0 \):

\[
\partial_t u_{q_0} = \frac{1}{2} \Delta u_{q_0} + \beta u_{q_0} w_0, \quad u_{q_0}(0, x) = q_0(x).
\]

Recall that here we consider the initial data \( q_0(x) = \delta(x) \).

By the proof of lemma A.3 (see (A.3)), we have

\[
E \left[ \left( \int_{\mathbb{R}^d} u_{q_0}(t, \tilde{x})d\tilde{x} \right)^{-p} \right] \leq E[|u_1(t, \cdot)^{-p}|],
\]

with \( u_1 \) solving the equation with constant initial data

\[
\partial_t u_1 = \frac{1}{2} \Delta u_1 + \beta u_1 w_0, \quad u_1(0, x) = 1.
\]
By [25, proposition 2.3], we know that for any $d \geq 3$ and $p \geq 1$, there exists positive constants $\beta_0$, $C > 0$ such that if $\beta < \beta_0$, then

$$\sup_{t \geq 0} \mathbb{E}[u_t(t, \cdot)^{-p}] \leq C.$$ 

By Hölder’s inequality, this implies

$$Q_n(t, x_1, \ldots, x_n) \leq C \prod_{j=1}^n \mathbb{E}[u_0(t, x_j)^{2n}]^{1/n}.$$ 

Thus, the proof of the lemma reduces to the following moment estimate of $u_0(t, x)$: for any $n \geq 1$, there exists $\beta_0$, $C > 0$ such that if $\beta < \beta_0$, then

$$\mathbb{E}[u_0(t, x)^n]^{1/n} \leq C G_t(x), \quad \text{for all } t > 0, x \in \mathbb{R}^d. \quad (B.2)$$

To prove (B.2), we first express $u_0$ by the Feynman–Kac formula as

$$u_0(t, x) = \mathbb{E}_x \left[ q_0(w_t) e^{\beta \int_0^t \eta \varphi(t-s, w_s) \, ds - \frac{1}{2} \beta^2 R(0)} t \right],$$

with $\mathbb{E}_x$ the expectation with respect to the Brownian motion $w$ starting at $x$. By the replica method, the $n$-th moment equals

$$\mathbb{E}[u_0(t, x)^n] = \mathbb{E}_x \left[ \prod_{j=1}^n q_0(w_{t_j}) e^{\beta \sum_{1 \leq i < j \leq n} R(w_{t_i} - w_{t_j})} \right],$$

with $\{w_j\}_{j=1}^n$ independent Brownian motions starting at $x$. By conditioning on the positions of $w_j$ at time $t$, we further write the expectation as an integral

$$\mathbb{E}[u_0(t, x)^n] = \int_{\mathbb{R}^d} \prod_{j=1}^n q_0(y_j) G_t(x - y_j)$$

$$\times \mathbb{E}_x \left[ e^{\beta \sum_{1 \leq i < j \leq n} R(w_{t_i} - w_{t_j})} \mid w_{t_j} = y_j, j = 1, \ldots, n \right] dy_1 \ldots dy_n.$$

By [25, lemma A.2], the following estimate holds for $\beta \ll 1$:

$$\sup_{t > 0, x, y_1, \ldots, y_n \in \mathbb{R}^d} \mathbb{E}_x \left[ e^{\beta \sum_{1 \leq i < j \leq n} R(w_{t_i} - w_{t_j})} \mid w_{t_j} = y_j, j = 1, \ldots, n \right] \leq C,$$

which yields

$$\mathbb{E}[u_0(t, x)^n] \leq C \left( \int_{\mathbb{R}^d} q_0(y) G_t(x - y) \, dy \right)^n = C G_t(x)^n,$$

and completes the proof. □

References

[1] Alberts T, Khanin K and Quastel J 2014 The continuum directed random polymer J. Stat. Phys. 154 305–26

[2] Albeverio S and Zhou X 1996 A martingale approach to directed polymers in a random environment J. Theor. Probab. 9 171189
[3] Amir G, Corwin I and Quastel J 2011 Probability distribution of the free energy of the continuum directed random polymer in 1 + 1 dimensions Commun. Pure Appl. Math. 64 466-537
[4] Baktihin Y and Khanin K 2018 On global solutions of the random Hamilton–Jacobi equations and the KPZ problem Nonlinearity 31 R93–R121
[5] Balázs M, Quastel J and Seppäläinen T 2011 Fluctuation exponent of the KPZ/stochastic Burgers equation J. Am. Math. Soc. 24 683–708
[6] Barraquand G and Corwin I 2017 Random-walk in Beta-distributed random environment Probab. Theor. Relat. Fields 167 1057–116
[7] Bates E and Chatterjee S 2016 The endpoint distribution of directed polymers (arXiv:1612.03443)
[8] Bertini L and Cancrini N 1995 The stochastic heat equation: Feynman–Kac formula and intermittence J. Stat. Phys. 78 1377–401
[9] Boldrighini C, Minlos R A and Pellegrinotti A 1997 Almost-sure central limit theorem for directed polymers and random corrections Commun. Math. Phys. 189 533–57
[10] Bolthausen E 1989 A note on the diffusion of directed polymers in a random environment Commun. Math. Phys. 123 529–34
[11] Borodin A and Corwin I 2014 Macdonald processes Probab. Theor. Relat. Fields 158 225–400
[12] Borodin A, Corwin I and Ferrari P 2014 Free energy fluctuations for directed polymers in random media in 1 + 1 dimension Commun. Pure Appl. Math. 67 1129–214
[13] Borodin A, Corwin I and Remyen D 2013 Log-gamma polymer free energy fluctuations via a Fredholm determinant identity Commun. Math. Phys. 324 215–32
[14] Bröker Y and Mukherjee C 2019 Quenched central limit theorem for the stochastic heat equation in weak disorder Probability and Analysis in Interacting Physical Systems, in Honor of SRS Varadhan (Berlin: Springer) pp 173–89
[15] Carmona P and Hu Y 2006 Strong disorder implies strong localization for directed polymers in a random environment ALEA 2 217–29 http://alea.impa.br/articles/v2/02-09.pdf
[16] Cerignani C, Illner R and Pulvirenti M 2013 The Mathematical Theory of Dilute Gases vol 106 (Berlin: Springer)
[17] Chen L and Huang J 2019 Comparison principle for stochastic heat equation on \(\mathbb{R}^d\) Ann. Probab. 47 989–1035
[18] Comets F 2017 Directed polymers in random environments Probability in Saint-Flour, Lecture Notes in Mathematics vol 2175 (Berlin: Springer)
[19] Comets F, Shiga T and Yoshida N 2003 Directed polymers in a random environment: path localization and strong disorder Bernoulli 9 705–23
[20] Comets F and Yoshida N 2006 Directed polymers in random environment are diffusive at weak disorder Ann. Probab. 34 1746–70
[21] Constantin P, Kiselev A, Oberman A and Ryzhik L 2000 Bulk burning rate in passive-reactive diffusion Arch. Rational Mech. Anal. 154 53–91
[22] Corwin I 2012 The Kardar–Parisi–Zhang equation and universality class Random Matrices: Theory Appl. 01 1130001
[23] Corwin I and Hammond A 2016 KPZ line ensemble Probab. Theor. Relat. Fields 166 67–185
[24] Den Hollander F 2009 Random Polymers: École d’Été de Probabilités de Saint-Flour XXXVII–2007 (Berlin: Springer)
[25] Dunlap A, Gu Y, Ryzhik L and Zeitouni O 2020 Fluctuations of the solutions to the KPZ equation in dimensions three and higher Probab. Theor. Relat. Fields 176 1217–58
[26] Ferrari P L and Spohn H 2006 Scaling limit for the space-time covariance of the stationary totally asymmetric simple exclusion process Commun. Math. Phys. 265 1–44
[27] Flores G M, Quastel J and Remenik D 2013 Endpoint distribution of directed polymers in 1 + 1 dimensions Commun. Math. Phys. 317 363–80
[28] Giacomin G 2007 Random Polymer Models (London: Imperial College Press, World Scientific)
[29] Gu Y and Henderson C 2020 Long-time behavior for a nonlocal model from directed polymers (arXiv:2010.13874)
[30] Gu Y and Komorowski T 2021 KPZ on torus: Gaussian fluctuations (arXiv:2104.13540)
[31] Hu Y and Lê K 2018 Asymptotics of the density of parabolic Anderson random fields (arXiv:1801.03386)
[32] Imbrie J Z and Spencer T 1988 Diffusion of directed polymers in a random environment J. Stat. Phys. 52 609–26
[33] Johansson K 2000 Shape fluctuations and random matrices Commun. Math. Phys. 209 437–76
[34] Johansson K 2003 Discrete polynuclear growth and determinantal processes Commun. Math. Phys. 242 277–329

[35] Khoshnevisan D 2014 Analysis of Stochastic Partial Differential Equations vol 119 (Providence, RI: American Mathematical Society)

[36] Landim C, Quastel J, Salmhofer M and Yau H-T 2004 Superdiffusivity of asymmetric exclusion process in dimensions one and two Commun. Math. Phys. 244 455–81

[37] Nualart D and Rozovskii B 1997 Weighted stochastic Sobolev spaces and bilinear SPDEs driven by space–time white noise J. Funct. Anal. 149 200–25

[38] Quastel J and Spohn H 2015 The one-dimensional KPZ equation and its universality class J. Stat. Phys. 160 965–84

[39] Quastel J and Valko B 2007 $t^{1/3}$ superdiffusivity of finite-range asymmetric exclusion processes on $\mathbb{Z}$ Commun. Math. Phys. 273 379–94

[40] Schehr G 2012 Extremes of $n$ vicious walkers for large $n$: application to the directed polymer and KPZ interfaces J. Stat. Phys. 149 385–410

[41] Seppäläinen T 2012 Scaling for a one-dimensional directed polymer with boundary conditions Ann. Probab. 40 19–73

[42] Song R and Zhou X 1996 A remark on diffusion on directed polymers in random environment J. Stat. Phys. 85 277–89

[43] Vargas V 2007 Strong localization and macroscopic atoms for directed polymers Probab. Theor. Relat. Fields 138 391–410

[44] Walsh J B 1986 An introduction to stochastic partial differential equations École d’Été de Probabilités de Saint Flour XIV-1984 (Berlin: Springer) pp 265–439