Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 188, 1–22.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1240

Multi-point Lyapunov exponents of the Stochastic Heat Equation*

Yier Lin†

Abstract

We obtain the multi-point positive integer Lyapunov exponents of the Stochastic Heat Equation (SHE) and provide three expressions for them. We prove the result by matching the upper and lower bounds for the Lyapunov exponents. The upper bound is obtained by analyzing the contour integral formula in [4]. For the lower bound, we apply an induction argument, relying on a tree recently appeared in [71]. The tree is related to the optimal trajectories of the Brownian motions in the Feynman-Kac formula.

Keywords: KPZ; SHE; Lyapunov exponents; multi-point.
MSC2020 subject classifications: 60H15; 60F10; 82B23.
Submitted to EJP on March 14, 2024, final version accepted on November 4, 2024.

1 Introduction

In this paper, we consider the Stochastic Heat Equation (SHE) in one spatial dimension,

$$\partial_t Z(t,x) = \frac{1}{2}\partial_{xx} Z(t,x) + \xi(t,x) Z(t,x), \qquad (t,x) \in \mathbb{R}_{>0} \times \mathbb{R},$$

where ξ is the spacetime white noise.

The SHE is closely related to the Kardar-Parisi-Zhang (KPZ) equation [35]

$$\partial_t h(t,x) = \frac{1}{2}\partial_{xx} h(t,x) + \frac{1}{2}(\partial_x h(t,x))^2 + \xi(t,x)$$

via the Hopf-Cole transform $Z(t,x) = \exp(h(t,x))$. The KPZ equation is a paradigm for modeling the random interface growth [64, 13], a universal scaling limit of the weakly asymmetric interacting particle systems and a testing ground for the study of nonlinear stochastic PDEs.

*We acknowledge the support of a research fund from Department of Statistics of the University of Chicago.
†University of Chicago, United States of America. E-mail: ylin10@uchicago.edu

We focus on the SHE starting from the Dirac delta initial data $Z(0,\cdot)=\delta(\cdot)$. The SHE starting from the Dirac delta initial data has a unique mild solution that satisfies

$$Z(t,x)=\mathsf{q}(t,x)+\int_0^t\int_{\mathbb{R}}\mathsf{q}(t-s,x-y)Z(s,y)\xi(s,y)dsdy,\qquad (t,x)\in\mathbb{R}_{>0}\times\mathbb{R},$$

where $\mathsf{q}(t,x):=\frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}$ is the standard heat kernel. The stochastic integral against the spacetime white noise is interpreted in the Itô's sense. In addition, Z is strictly positive on $(t,x)\in\mathbb{R}_{>0}\times\mathbb{R}$ [62, 23].

1.1 Main result

Throughout the paper, we consider the *hyperbolic scaling* $Z_T(t,x):=Z(Tt,Tx)$. We compute the multi-point Lyapunov exponents of the SHE that are defined as the following limit

$$\lim_{T\to\infty}\frac{1}{T}\log\mathbb{E}\Big[\prod_{i=1}^n Z_T(t,x_i)^{m_i}\Big].$$

Theorem 1.1. *For fixed* $n\in\mathbb{Z}_{\geq 1}$, $t>0$, $\vec{x}=(x_1<\ldots<x_n)\in\mathbb{R}^n$ *and* $\vec{m}=(m_1,\ldots,m_n)\in\mathbb{Z}_{\geq 1}^n$, *we have*

$$\lim_{T\to\infty}\frac{1}{T}\log\mathbb{E}\Big[\prod_{i=1}^n Z_T(t,x_i)^{m_i}\Big]=\gamma(t,\vec{x},\vec{m}),$$

where $\gamma:=\gamma_1=\gamma_2=\gamma_3$, *the expressions of the* $\{\gamma_i\}_{i=1}^3$ *are given in* (1.1) - (1.3).

We set

$$\mathfrak{m}:=\sum_{i=1}^n m_i \text{ and } u_k:=x_j \text{ if } \sum_{i=1}^{j-1}m_i<k\leq\sum_{i=1}^{j}m_i,$$

see Figure 1.1 for an visualization.

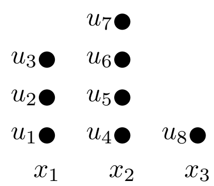

Figure 1: Take $n=3$, $\vec{x}=(x_1<x_2<x_3)$ and $m_1=3$, $m_2=4$ and $m_3=1$, we have $(u_1,\ldots,u_8)=(x_1,x_1,x_1,x_2,x_2,x_2,x_2,x_3)$.

We define

$$\gamma_1(t,\vec{x},\vec{m}):=\inf\Big\{\sum_{i=1}^{\mathfrak{m}}\frac{t}{2}a_i^2+\sum_{i=1}^{\mathfrak{m}}u_ia_i : a_i-a_{i+1}\geq 1,\, i=1,\ldots,\mathfrak{m}-1\Big\}\tag{1.1}$$

$$\gamma_2(t,\vec{x},\vec{m}):=\inf\Big\{\sum_{i=1}^{n}\frac{m_it}{2}(b_i+\frac{x_i}{t})^2+\frac{(m_i^3-m_i)t}{24}-\frac{m_ix_i^2}{2t}:\; b_i-b_{i+1}\geq\frac{m_i+m_{i+1}}{2}, i=1,\ldots,n-1\Big\}.\tag{1.2}$$

The a_i and b_i in the infimums above are real numbers. We define $\mathfrak{m}_{\mathfrak{b}_j}:=\sum_{i\in\mathfrak{b}_j}m_i$ and

$$\gamma_3(t,\vec{x},\vec{m}):=\sum_{j=1}^{\mathfrak{n}}\Big(\frac{(\mathfrak{m}_{\mathfrak{b}_j}^3-\mathfrak{m}_{\mathfrak{b}_j})t}{24}-\sum_{k,\ell\in\mathfrak{b}_j,k<\ell}\frac{m_km_\ell}{2}|x_k-x_\ell|-\frac{(\sum_{k\in\mathfrak{b}_j}m_kx_k)^2}{2t\mathfrak{m}_{\mathfrak{b}_j}}\Big).\tag{1.3}$$

The positive integer $\mathfrak{n}$ and sets $\{\mathfrak{b}_j\}_{j=1}^{\mathfrak{n}}$ (which form a partition of $\{1,\dots,n\}$) will be defined in Section 3.1.

When n is small, we have more explicit expressions for the Lyapunov exponents.

Corollary 1.2. *Take $n=1$ in Theorem 1.1, we obtain the one-point Lyapunov exponents of the SHE*

$$\lim_{T\to\infty}\frac{1}{T}\log\mathbb{E}[Z_T(t,x_1)^{m_1}] = \frac{(m_1^3-m_1)t}{24}-\frac{m_1x_1^2}{2t}.$$

Take $n=2$ in Theorem 1.1, we have

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{T}\log\mathbb{E}\Big[\prod_{i=1}^{2}Z_T(t,x_i)^{m_i}\Big]\\ &=\begin{cases}\frac{((m_1+m_2)^3-(m_1+m_2))t}{24}-\frac{m_1m_2(x_2-x_1)}{2}-\frac{(m_1x_1+m_2x_2)^2}{2(m_1+m_2)t} & \textit{if } 0<\frac{x_2-x_1}{t}\le\frac{m_1+m_2}{2},\\ \frac{(m_1^3-m_1)t}{24}+\frac{(m_2^3-m_2)t}{24}-\frac{m_1x_1^2}{2t}-\frac{m_2x_2^2}{2t}, & \textit{if } \frac{x_2-x_1}{t}>\frac{m_1+m_2}{2}.\end{cases}\end{aligned}$$

As a byproduct, Theorem 1.1 shows that $\gamma_1=\gamma_2=\gamma_3$, which is surprising since the expressions of them are quite different. In the following, we characterize the minimizer of the variational expression γ_1. We define a map $f:\{1,\dots,\mathfrak{m}\}\to\{1,\dots,n\}$ such that $u_i=x_{f(i)}$.

Corollary 1.3. *The infimum in γ_1 has a unique minimizer $(a_1^*,\dots,a_{\mathfrak{m}}^*)$. In addition, we have $a_i^*-a_{i+1}^*=1$ if there exists $j\in\{1,\dots,\mathfrak{n}\}$ such that $f(i),f(i+1)\in\mathfrak{b}_j$. We have $a_i^*-a_{i+1}^*>1$ if there does not exist such j.*

1.2 Proof idea

Let us explain the idea for proving Theorem 1.1. We respectively show the upper bound

$$\limsup_{T\to\infty}\frac{1}{T}\log\mathbb{E}\Big[\prod_{i=1}^{n}Z_T(t,x_i)^{m_i}\Big]\le\gamma_1(t,\vec{x},\vec{m}), \tag{1.4}$$

and the lower bound

$$\liminf_{T\to\infty}\frac{1}{T}\log\mathbb{E}\Big[\prod_{i=1}^{n}Z_T(t,x_i)^{m_i}\Big]\ge\gamma_3(t,\vec{x},\vec{m}). \tag{1.5}$$

After obtaining these bounds, we conclude Theorem 1.1 by showing $\gamma_1\le\gamma_2\le\gamma_3$. In the following, we focus on explaining the idea for obtaining the bounds (1.4) and (1.5).

1.2.1 Upper bound

It is known that the moments of the SHE $\mathbb{E}[\prod_{i=1}^n Z(t,x_i)]$ solve a PDE called the *delta Bose gas* [4, Section 6.2]. More precisely, let $U(t;x_1,\dots,x_n):=\mathbb{E}[\prod_{i=1}^n Z(t,x_i)]$, we have

$$\partial_t U=\frac{1}{2}\Delta U+\frac{1}{2}\sum_{i\neq j}\delta(x_i-x_j)U.$$

The solution to the above PDE (starting from certain initial data) admits a contour integral expression [4, Proposition 6.2.3]. By a straightforward analysis of the contour integral, we obtain the upper bound.

1.2.2 Lower bound

Apply the Feynman-Kac formula to solve the delta Bose gas, we can express the moments of the SHE as expectations of the Brownian local times, namely, $\mathbb{E}[\prod_{i=1}^{n} Z_T(t,x_i)^{m_i}]$ is equal to

$$\mathbb{E}\Big[\exp\Big(\sum_{1\le i<j\le \mathfrak{m}}\int_0^{Tt}\delta(W_s^i-W_s^j)ds\Big)\prod_{i=1}^{\mathfrak{m}}\delta(W_{Tt}^i)\Big], \tag{1.6}$$

the $\{W^i\}_{i=1}^{\mathfrak{m}}$ are independent Brownian motions. Note that $W_0^j = Tu_j$, namely $W_0^j = Tx_i$ if $\sum_{k=1}^{i-1} m_k < j \le \sum_{k=1}^{i} m_k$. Hence, there are m_i Brownian motions starting from the location Tx_i. The first Dirac function in (1.6) is understood as local time. The second Dirac function here is understood as the distributional limit of the heat kernel $\mathsf{q}(t,\cdot)$ as $t\to 0$. As a consequence, one can interpret (1.6) as

$$\mathbb{E}\Big[\exp\Big(\sum_{1\le i<j\le \mathfrak{m}}\int_0^{Tt}\delta(\mathsf{W}_s^i-\mathsf{W}_s^j)ds\Big)\prod_{i=1}^{\mathfrak{m}}\mathsf{q}(Tt,Tu_i)\Big],$$

where $\{\mathsf{W}^j\}_{j=1}^{\mathfrak{m}}$ are independent Brownian bridges with $\mathsf{W}_0^j = Tu_j$ and $\mathsf{W}_{Tt}^j = 0$.

We want to understand the optimal trajectories of the Brownian motions, i.e. the deterministic trajectories where the Brownian motions stay around contribute most to the expectation (1.6).

The Dirac function at the terminal time in (1.6) forces the Brownian motions to end up at 0. From the perspective of large deviations, to compute the asymptotics of expectations, one needs to figure out the optimal product of the value of the random variable and the its probability to take such value.. For the expectation (1.6), in order to contribute more to the Brownian local times on the exponential, the Brownian motions tend to move close to one another before the terminal time to gain more local time. Once they are close, they no longer want to be apart. On the other hand, they avoid traveling too fast to become close since this will make the transition probabilities too small. Consider the macroscopic picture by scaling both space and time by T, the above discussion suggests that the optimal trajectories of the Brownian motions form a tree as shown in Figure 2. In Section 3.1, we will characterize precisely this tree using the attractive Brownian particles. Following [71, Section 2.3], we call this tree the *optimal clusters*.

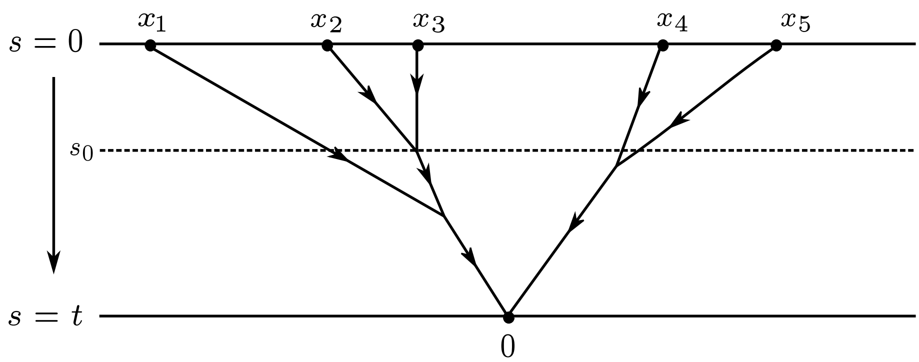

Figure 2: The optimal trajectories form a tree.

We proceed to explain how to obtain the lower bound (1.5). The idea is to apply an induction argument over n, which is the number of different locations that the Brownian motions start from. Let s_0 be the first time that the number of different points in the optimal clusters is smaller than n, i.e. the first time that two branches in the tree merge. We break the integral in (1.6) into integrals over the time interval $[0, s_0]$ and $[s_0, t]$ (remember the time and space have been scaled by T) and restrict on the event that the Brownian motions at time s_0 stay around the optimal clusters. Since the number of different points in the optimal clusters at time s_0 is less than n, we can lower bound the contribution of the Brownian motions in time $[s_0, t]$ using the induction hypothesis. In addition, we can lower bound the contribution in time $[0, s_0]$ by dropping the Brownian local times if the two Brownian motions start from different x_i, this is fine since the optimal trajectories of the Brownian motions with different starting locations do not overlap before time s_0. Put together the lower bounds for the integrals over $[0, s_0]$ and $[s_0, t]$, we obtain the desired lower bound (1.5).

We remark that the actual proof in Section 3.2 is slightly different from what has been explained above, although the idea behind is quite similar. Instead of using (1.6), we rely on the semigroup identity of the SHE to carry out the proof.

1.3 Discussion

Let us discuss three related work. In [14], the authors obtained the one-point Lyapunov exponents of the SHE under Dirac delta initial data by applying a residue expansion to the contour integral (2.1). Since there are many poles in the contour integral, book-keeping the residues when deforming the contour is not easy. The advantage in the one-point situation is that since $u_1 = \ldots = u_{\mathrm{m}}$, the exponential function in (2.1) is symmetric in $z_1, \ldots, z_{\mathrm{m}}$, and we have a simple residue expansion thanks to [4, Proposition 6.2.7]. For the multi-point $(n > 1)$ situation, the residue expansion seems much harder to be analyzed due to the lack of symmetry. Instead of relying on the contour integral formula to derive the full asymptotic of the Lyapunov exponents, we only use it to obtain an (optimal) upper bound.

[24] obtained the sharp bounds for the two-point upper tail probability of the KPZ equation in finite large time using the Gibbsian line ensembles [17, 18]. The method therein should work for obtaining the multi-point upper tail bounds and could be applied to obtain the multi-point (even for non-integer valued) Lyapunov exponents of the SHE. This approach, however, seems a bit indirect and we are not sure whether it will lead to a simple expression of the Lyapunov exponents.

The work [71] studied the Lyapunov exponents of the SHE

$$\frac{1}{N^3 T} \log \mathbb{E}\Big[\prod_{i=1}^{n} Z_T(t, N x_i)^{N m_i}\Big]$$

under the *high-moment regime* $N^2 T \to \infty$ and $N \to \infty$ and obtained the multi-point Lyapunov exponents by studying the large deviations of the attractive Brownian particles. In this paper, we consider the hyperbolic scaling $N = 1$ and $T \to \infty$, which is different from the high-moment regime. Our method relies on the exact formula and an induction argument, which is new. We have not seen the multi-point Lyapunov exponents of the SHE being studied in the context of hyperbolic scaling before.

1.4 Literature review

The one-point Lyapunov exponents of the SHE with different initial conditions have been studied in [3, 10, 14, 21, 30]. The moments and Lyapunov exponents are useful for studying the property of *intermittency* [25, 26, 36, 10], the density function and large deviations of the SHE (and its variant), see [12, 29, 5, 8, 31, 33, 37, 14, 9, 21, 55, 22, 27, 32, 30].

The Lyapunov exponents of the SHE are closely related to the upper tail bounds and Large Deviation Principle (LDP) of the KPZ equation. The one-point tail bounds and LDP of the KPZ equation with different boundary and initial conditions have been intensively studied recently in the physics work [53, 54, 44, 65, 16, 45, 46, 51, 42, 47, 52] and mathematics work [15, 14, 7, 21, 38, 55, 6, 69, 20, 30]. The two (and potentially multi)-point upper tail bounds and the terminal-time limit shapes of the KPZ equation have been studied in [24]. The Freidlin-Wentzell LDP/weak noise theory has been used to study the one-point LDP and most probable shapes of the KPZ equation, see the physics work [39, 40, 41, 59, 34, 60, 61, 66, 67, 2, 68] and mathematics work [56, 56, 57, 28]. The connection between the Freidlin-Wentzell LDP/weak noise theory and the integrable PDE has been studied in the physics work [43, 48, 49, 50] and mathematics work [70]. The LDP and spacetime limit shapes of the KPZ equation in the deep upper tails have been studied recently in [71, 58].

Outline

In Section 2, we prove the upper bound (1.4). In Section 3, we prove the lower bound (1.5). In Section 4, we establish a continuity result, which is a technical input for proving the lower bound. In Section 5, we prove Theorem 1.1 and its corollaries.

2 The upper bound

In this section, we will prove the upper bound (1.4), which is stated as the following proposition. The major input is a contour integral expression for the moments of the SHE.

Proposition 2.1. *Under the same setting as Theorem 1.1, we have*

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{E}\Big[\prod_{i=1}^{n} Z_T(t,x_i)^{m_i}\Big] \leq \gamma_1(t,\vec{x},\vec{m}).$$

Proof. Recall that $u_k := x_j$ if $\sum_{i=1}^{j-1} m_i < k \leq \sum_{i=1}^{j} m_i$. It is straightforward that

$$\mathbb{E}\Big[\prod_{i=1}^{n} Z_T(t,x_i)^{m_i}\Big] = \mathbb{E}\Big[\prod_{i=1}^{\mathfrak{m}} Z_T(t,u_i)\Big].$$

By [4, Proposition 6.2.3], we know that

$$\mathbb{E}\Big[\prod_{i=1}^{\mathfrak{m}} Z_T(t,u_i)\Big] = \frac{1}{(2\pi\mathbf{i})^{\mathfrak{m}}} \int \cdots \int \prod_{1\leq A<B\leq \mathfrak{m}} \frac{z_A - z_B}{z_A - z_B - 1} \exp\Big(\sum_{j=1}^{\mathfrak{m}} Ttz_j^2/2 + \sum_{j=1}^{\mathfrak{m}} Tu_j z_j\Big) \prod_{j=1}^{\mathfrak{m}} dz_j, x \qquad (2.1)$$

the z_j-contour is given by $a_j + \mathbf{i}\mathbb{R}$ and $\{a_k\}_{k=1}^{\mathfrak{m}}$ can be any real numbers satisfying $a_k - a_{k+1} > 1$, $k = 1, \dots, \mathfrak{m}-1$. We remark that the proof for the identity above in [4, Proposition 6.2.3.] is not fully rigorous. However, one can combine [4, Proposition 5.4.8] together with [63, Corollary 1.7] to obtain a rigorous proof.

Apply a change of variable $z_j = a_j + \mathbf{i}y_j$ for $j = 1, \dots, \mathfrak{m}$, the triangle inequality

$\int \ldots \leq \int |\ldots|$, and $|\frac{z_A - z_B}{z_A - z_B - 1}| \leq \frac{a_A - a_B}{a_A - a_B - 1}$ for $A < B$ (noting that $a_A - a_B > 1$), we have

$$\begin{aligned}
&\mathbb{E}\Big[\prod_{i=1}^{\mathfrak{m}} Z_T(t, u_i)\Big] \\
&\leq \frac{1}{(2\pi)^{\mathfrak{m}}} \prod_{1 \leq A < B \leq \mathfrak{m}} \Big|\frac{a_A - a_B}{a_A - a_B - 1}\Big| \exp\left(\sum_{j=1}^{\mathfrak{m}} Tta_j^2/2 + \sum_{j=1}^{\mathfrak{m}} Tu_j a_j\right) \int_{\mathbb{R}^{\mathfrak{m}}} \prod_{j=1}^{\mathfrak{m}} e^{-Tty_j^2/2} dy_j \\
&= (2\pi Tt)^{-\mathfrak{m}/2} \prod_{1 \leq A < B \leq \mathfrak{m}} \Big|\frac{a_A - a_B}{a_A - a_B - 1}\Big| \exp\left(\sum_{j=1}^{\mathfrak{m}} Tta_j^2/2 + \sum_{j=1}^{\mathfrak{m}} Tu_j a_j\right).
\end{aligned}$$

Take the logarithm of both sides above, divide by T and let $T \to \infty$, we know that

$$\limsup_{T \to \infty} \frac{1}{T} \log \mathbb{E}\Big[\prod_{i=1}^{\mathfrak{m}} Z_T(t, u_i)\Big] \leq \sum_{j=1}^{\mathfrak{m}} \frac{ta_j^2}{2} + \sum_{j=1}^{\mathfrak{m}} u_j a_j.$$

The inequality above holds for $a_1, \ldots, a_{\mathfrak{m}}$ satisfying $a_k - a_{k+1} > 1$, $k = 1, \ldots, \mathfrak{m} - 1$. By the continuity of the right hand side above in $a_1, \ldots, a_{\mathfrak{m}}$, the inequality also holds for any $a_1, \ldots, a_{\mathfrak{m}}$ satisfying $a_k - a_{k+1} \geq 1$, $k = 1, \ldots, \mathfrak{m} - 1$. Take the infimum of the right hand side over $a_1, \ldots, a_{\mathfrak{m}}$ under this constraint, we conclude the proposition. □

3 The lower bound

In this section, we will prove the lower bound (1.5), which is stated as the following proposition.

Proposition 3.1. *Under the same setting as Theorem 1.1, we have*

$$\liminf_{T \to \infty} \frac{1}{T} \log \mathbb{E}\Big[\prod_{i=1}^{n} Z_T(t, x_i)^{m_i}\Big] \geq \gamma_3(t, \vec{x}, \vec{m}).$$

3.1 The inertia clusters and optimal clusters

We recall the *inertia clusters* $\boldsymbol{\zeta} = \{\boldsymbol{\zeta}_i\}_{i=1}^n$ and *optimal clusters* $\boldsymbol{\xi} = \{\boldsymbol{\xi}_i\}_{i=1}^n$ from [71, Section 2.3], see Figure 3 for a visualization. Let us first define the inertia clusters, then we use them to define the optimal clusters.

The inertia clusters are the trajectories of the point masses which run backward in time (compared with the time unit used in Z_T) from $s = 0$ to $s = t$. We say the inertia clusters start from $(\vec{x}, \vec{m})$ if at time $s = 0$, they start from the point masses with weight m_i and location x_i, $i = 1, \ldots, n$. As time evolves, the point mass with weight m_i will travel with a constant speed $\phi_i := \frac{1}{2}(m_{i+1} + \ldots + m_n) - \frac{1}{2}(m_1 + \ldots + m_{i-1})$. When the point masses with weights m_i and m_{i+1} and speeds ϕ_i and ϕ_{i+1} collide, they merge into a single point mass with weight $m_i + m_{i+1}$ and travel with speed $\frac{m_i \phi_i + m_{i+1} \phi_{i+1}}{m_i + m_{i+1}}$, this follows the conservation of momentum. Let $\boldsymbol{\zeta}_i : [0, t] \to \mathbb{R}, i = 1, \ldots, n$ denote the trajectories of the inertia clusters. Examine the point masses which have merged between time $[0, t]$, we obtain a partition of $\{1, \ldots, n\}$. Denote the partition to be $\mathfrak{B} = \{\mathfrak{b}_i\}_{i=1}^{\mathfrak{n}}$, where

$$\mathfrak{n} := \text{the number of different clusters at time } t.$$

We can order $\mathfrak{b}_1, \ldots, \mathfrak{b}_{\mathfrak{n}}$ such that for $i < j$, the elements of $\mathfrak{b}_i$ are smaller than those in $\mathfrak{b}_j$. Note that we have $\boldsymbol{\zeta}_i(t) = \boldsymbol{\zeta}_j(t)$ if and only if i, j belong to the same $\mathfrak{b}_k$ for some k. We define $\boldsymbol{\zeta}_{\mathfrak{b}_k}(t) := \boldsymbol{\zeta}_i(t) = \boldsymbol{\zeta}_j(t)$.

We proceed to define the optimal clusters $\boldsymbol{\xi}$. Let $v_j := \boldsymbol{\zeta}_{\mathfrak{b}_j}(t)/t$ for $j = 1, \ldots, \mathfrak{n}$. We define the trajectories of the optimal clusters $\{\boldsymbol{\xi}_i\}_{i=1}^n$ by applying a constant drift to the inertia clusters: $\boldsymbol{\xi}_i(s) := \boldsymbol{\zeta}_i(s) - v_j s$ for $i \in \mathfrak{b}_j$. Note that we have $\boldsymbol{\xi}_i(t) = 0$ for $i = 1, \ldots, n$.

Let us discuss non-rigorously why the optimal clusters should be the optimal trajectories of the Brownian motions for the expectation in (1.6). Following [71, Appendix A], we use the Tanaka's formula to write $\int_0^{Tt} \delta(W_s^i - W_s^j)ds = -\frac{1}{2}\int_0^{Tt} \operatorname{sgn}(W_s^i - W_s^j)d(W_s^i - W_s^j) + \frac{1}{2}|W_s^i - W_s^j|\big|_{s=0}^{s=Tt}$ with $\operatorname{sgn}(x) = \mathbf{1}_{\{x>0\}} - \mathbf{1}_{\{x<0\}}$. The quadratic variation of the stochastic integral on the resulting exponential is equal to $\frac{(\mathfrak{m}^3-\mathfrak{m})t}{12}$. By applying Girsanov theorem, (1.6) equals

$$\exp\Big(\frac{(\mathfrak{m}^3-\mathfrak{m})Tt}{24} - \sum_{1\le k<\ell\le n} \frac{Tm_k m_\ell}{2}|x_k - x_\ell|\Big)\mathbb{E}\Big[\exp\Big(\frac{1}{2}\sum_{1\le i<j\le \mathfrak{m}} |X_{Tt}^i - X_{Tt}^j|\Big)\prod_{i=1}^{\mathfrak{m}} \delta(X_{Tt}^i)\Big], \tag{3.1}$$

where $dX_s^i = \frac{1}{2}\sum_{j=1}^n \operatorname{sgn}(X_s^j - X_s^i) + dW_s^i$ with $X_0^i = Tu_i$. Note that the time is of order T and the diffusion has an order of $\sqrt{T}$, which is negligible compared with the drift as $T \to \infty$. Drop the diffusion and solve the deterministic equations $dX_s^i = \frac{1}{2}\sum_{j=1}^n \operatorname{sgn}(X_s^j - X_s^i)$. We refer to the resulting $\{X^i\}_{i=1}^{\mathfrak{m}}$ as the *attractive Brownian particles*. The i-th Brownian particle has the drift $\frac{1}{2}\sum_{j=1}^n \operatorname{sgn}(X_s^j - X_s^i)$. In addition, when two Brownian particles meet, they stay together afterward to contribute to the local time. Hence, the weight of the point masses in the inertia clusters can be viewed as the number of Brownian particles staying together. Scale the space and time by T, the trajectories of $\{X^i\}_{i=1}^{\mathfrak{m}}$ are given by the inertia clusters $\boldsymbol{\zeta}$. The Dirac delta function in (3.1) forces the clusters to end at 0. The most economic way to fulfill this (in terms of transition probability) is to apply a constant drift to each group in the inertia clusters. This leads to the optimal clusters $\boldsymbol{\xi}$.

For our proof in the paper, we *only* need the definition of $\boldsymbol{\xi}$. The discussion in the previous paragraph is only to explain how $\boldsymbol{\xi}$ appears and will not be used.

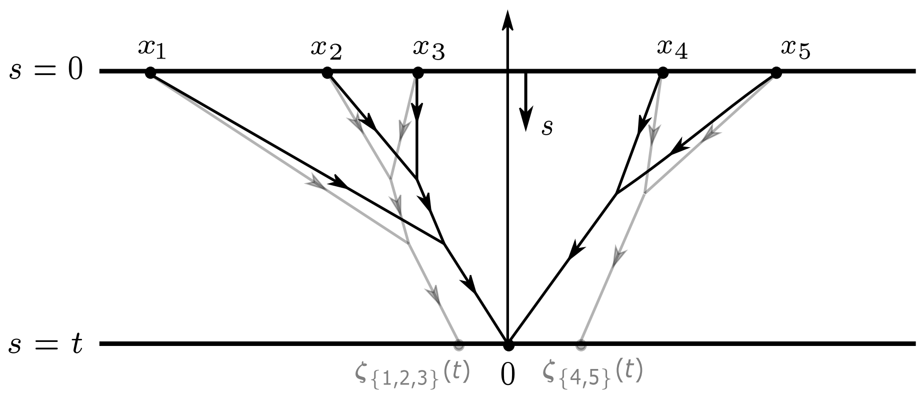

Figure 3: An illustration of the inertia clusters $\boldsymbol{\zeta}$ (gray) and the optimal clusters $\boldsymbol{\xi}$ (black) when $n = 5$. In the figure, we have $\mathfrak{n} = 2$. The partition of $\{1, 2, 3, 4, 5\}$ is given by $\mathfrak{B} = \{\mathfrak{b}_1, \mathfrak{b}_2\}$ where $\mathfrak{b}_1 = \{1, 2, 3\}$ and $\mathfrak{b}_2 = \{4, 5\}$.

3.2 Proof of Proposition 3.1

We use the Feynman-Kac formula to introduce a four parameter version of Z_T. For $r < t$, we set

$$Z_T(r, y; t, x) := \mathbb{E}\Big[\exp\Big(\int_0^{T(t-r)} \xi(Tt - s, W_s)ds\Big)\delta(W_{T(t-r)} - Ty)\Big]. \tag{3.2}$$

In particular, we have $Z_T(t, x) = Z_T(0, 0; t, x)$.

The exponential above is the Wick exponential and W is a standard Brownian motion starting from $W_0 = Tx$. One correct way to interpret the Wick exponential is to Taylor expand it, time-order the multiple Itô integrals, and then switch the order of integration with $\mathbb{E}$ (see (see [13, Section 4.1.1])). Doing this results in a series of multiple stochastic integrals

$$Z_T(r, y; t, x) = \mathfrak{q}(T(t-r), T(x-y)) \sum_{n=0}^{\infty} \int_{0 \le t_1 \le \dots \le t_n \le T(t-r)} \int_{\mathbb{R}^n} p_{BB}(t_1, \dots t_n; x_1, \dots x_n) \times \\ \xi(t_1, x_1) \dots \xi(t_n, x_n) dt_1 dx_1 \dots dt_n dx_n.$$

where $(t_1, ..., t_n; x_1 ..., x_n)$ represents the n-step transition probability of a Brownian bridge started at Ty at time 0 and ended at Tx and time $T(t-r)$ to go through positions x_i at times t_i for $i = 1, \dots, n$. The multiple stochastic integrals are those of Itô.

By (3.2), we have the semigroup identity: For $r < s < t$,

$$Z_T(r, y; t, x) = \int_{\mathbb{R}} T Z_T(r, y; s, z) Z_T(s, z; t, x) dz. \tag{3.3}$$

In the following,

$$A \gtrsim B \text{ means } \liminf_{T \to \infty} \frac{1}{T} \log(A/B) \ge 0.$$

Under this notation, showing Proposition 3.1 is the same as showing

$$\mathbb{E}\Big[\prod_{i=1}^{n} Z_T(t, x_i)^{m_i}\Big] \gtrsim e^{T\gamma_3(t, \vec{x}, \vec{m})}. \tag{3.4}$$

Proof of Proposition 3.1. We apply an induction over n for proving (3.4). When $n = 1$, We apply [14, Lemma 4.1], which states that for $T > \pi$,

$$\mathbb{E}\Big[Z(2T, 0)^k e^{\frac{kT}{12}}\Big] \le \frac{69 k! e^{\frac{Tk^3}{12}}}{2\sqrt{\pi T} k^{\frac{3}{2}}} \le C(k) e^{\frac{Tk^3}{12}}.$$

We then use the fact that for any fixed t, the stochastic process $\{\frac{Z(t,x)}{\mathfrak{q}(t,x)}\}_{x \in \mathbb{R}}$ is stationary (which follows from [1, Proposition 2.3]). Recall that $\mathfrak{q}$ represents the heat kernel. This implies that for $Tt > 2\pi$,

$$\mathbb{E}\Big[Z_T(t, x_1)^{m_1}\Big] = \mathbb{E}\Big[Z(Tt, Tx_1)^{m_1}\Big] = \mathbb{E}\Big[Z(Tt, 0)^{m_1}\Big] \mathfrak{q}(Tt, Tx)^{m_1} \le C(m_1) e^{\frac{Ttm_1^3}{24}}.$$

When $n > 1$, we only need to prove (3.4) under the *induction hypothesis* that

$$\mathbb{E}[\prod_{i=1}^{n'} Z_T(t', x'_i)^{m'_i}] \gtrsim e^{T\gamma_3(t, \vec{x}', \vec{m}')}$$

holds for any $t' > 0$, $\vec{x}' = (x'_1 < \dots < x'_{n'}) \in \mathbb{R}^{n'}$, $\vec{m}' = (m'_1, \dots, m'_{n'}) \in \mathbb{Z}_{\ge 1}^{n'}$ with $n' < n$.

Recall the number $\mathfrak{n}$ and the partition $\{\mathfrak{b}_j\}_{j=1}^{\mathfrak{n}}$ from Section 3.1. We divide the proof into two cases: $\mathfrak{n} > 1$ and $\mathfrak{n} = 1$. When $\mathfrak{n} > 1$, we have $|\mathfrak{b}_j| < n$ for each $j \in \{1, \dots, \mathfrak{n}\}$. By the induction hypothesis, we know that

$$\mathbb{E}\Big[\prod_{k \in \mathfrak{b}_j} Z_T(t, x_k)^{m_k}\Big] \gtrsim e^{T\gamma_3(t, \vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j})}, \tag{3.5}$$

where $\vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j}$ are the vectors for $\{x_i\}_{i \in \mathfrak{b}_j}, \{m_i\}_{i \in \mathfrak{b}_j}$. Note that the trajectories of the inertia clusters starting from $(\vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j})$ are still given by $\{\boldsymbol{\zeta}_i\}_{i \in \mathfrak{b}_j}$. Hence, if we start the Brownian particles from $(\vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j})$, there is only one cluster at time $s = t$. We have

$$\gamma_3(t, \vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j}) = \frac{(\mathfrak{m}_{\mathfrak{b}_j}^3 - \mathfrak{m}_{\mathfrak{b}_j})t}{24} - \sum_{k, \ell \in \mathfrak{b}_j, k < \ell} \frac{m_k m_\ell}{2} |x_k - x_\ell| - \frac{(\sum_{k \in \mathfrak{b}_j} m_k x_k)^2}{2t\mathfrak{m}_{\mathfrak{b}_j}}.$$

Apply Lemma A.1 and then (3.5) for $j = 1, \dots, \mathfrak{n}$, we obtain the desired inequality

$$\mathbb{E}\Big[\prod_{i=1}^{n} Z_T(t,x_i)^{m_i}\Big] \geq \prod_{j=1}^{\mathfrak{n}} \mathbb{E}\Big[\prod_{k\in\mathfrak{b}_j} Z_T(t,x_k)^{m_k}\Big] \gtrsim \exp\Big(\sum_{j=1}^{\mathfrak{n}} T\gamma_3(t,\vec{x}_{\mathfrak{b}_j},\vec{m}_{\mathfrak{b}_j})\Big) = e^{T\gamma_3(t,\vec{x},\vec{m})}.$$

We proceed to prove Proposition 3.1 when $\mathfrak{n}=1$. Define

$$s_0 = \inf\{s : |\{\boldsymbol{\xi}_1(s),\dots,\boldsymbol{\xi}_n(s)\}| < n\}, \tag{3.6}$$

which is the first time that a merging occurs among point masses in the optimal clusters. Fix $\delta > 0$, apply the semigroup identity (3.3), and write $Z_T(t,x_k) = \int_{\mathbb{R}} TZ_T(t-s_0,y_{k,\ell})Z_T(t-s_0,y_{k,\ell};t,x_k)dy_{k,\ell}$ for $\ell = 1,\dots,m_k$. Use Fubini's theorem to exchange the expectation and integral, then apply the independence between $Z_T(t-s_0,\cdot)$ and $Z_T(t-s_0,\cdot;t,\cdot)$, and finally restrict the domain of integral for $y_{k,\ell}$ to $[\boldsymbol{\xi}_k(s_0)-\delta,\boldsymbol{\xi}_k(s_0)+\delta]$. By the semigroup property, we know that $\mathbb{E}[\prod_{k=1}^n Z_T(t,x_k)^{m_k}]$ is lower bounded by

$$\int_{\mathbb{R}^{\mathfrak{m}}} T^{\mathfrak{m}}\mathbb{E}\Big[\prod_{k=1}^{n}\prod_{\ell=1}^{m_k} Z_T(t-s_0,y_{k,\ell})\Big]\mathbb{E}\Big[\prod_{k=1}^{n}\prod_{\ell=1}^{m_k} Z_T(t-s_0,y_{k,\ell};t,x_k)\Big]\cdot \prod_{k=1}^{n}\prod_{\ell=1}^{m_k}\mathbf{1}_{\{|y_{k,\ell}-\boldsymbol{\xi}_k(s_0)|\leq\delta\}}\,dy_{k,\ell}. \tag{3.7}$$

We lower bound the first and second expectation in the integral of (3.7), assuming that $|y_{k,\ell}-x_k|\leq\delta$. For the second expectation, apply Lemma A.1, we get $\mathbb{E}[\prod_{k=1}^n\prod_{\ell=1}^{m_k} Z_T(t-s_0,y_{k,\ell};t,x_k)] \geq \prod_{k=1}^n \mathbb{E}[\prod_{\ell=1}^{m_k} Z_T(t-s_0,y_{k,\ell};t,x_k)]$. By applying Proposition 4.2 to the preceding right hand side (noting that $|y_{k,\ell}-\boldsymbol{\xi}_k(s_0)|\leq\delta$) together with [14, Lemma 4.1], we get

$$\mathbb{E}\Big[\prod_{k=1}^{n}\prod_{\ell=1}^{m_k} Z_T(t-s_0,y_{k,\ell};t,x_k)\Big] \gtrsim \exp\bigg(T\Big(\sum_{k=1}^{n}\frac{(m_k^3-m_k)s_0}{24}-\frac{m_k(x_k-\boldsymbol{\xi}_k(s_0))^2}{2s_0}+f_1(\delta)\Big)\bigg), \tag{3.8}$$

where $\lim_{\delta\to 0} f_1(\delta) = 0$.

We examine the point masses that have merged at time $s=s_0$ and let $\vec{x}' = (x_1',\dots,x_{n'}')$ and $\vec{m}' = (m_1',\dots,m_{n'}')$ denote the locations and weights of them. Since a merging happens at time $s=s_0$, we know that $n'<n$. For the first expectation in (3.7), use Proposition 4.2, we get $\mathbb{E}[\prod_{k=1}^n\prod_{\ell=1}^{m_k} Z_T(t-s_0,y_{k,\ell})] \geq \mathbb{E}[\prod_{k=1}^n Z_T(t-s_0,\boldsymbol{\xi}_k(s_0))^{m_k}]e^{Tf_2(\delta)}$. Rewrite $\prod_{k=1}^n Z_T(t-s_0,\boldsymbol{\xi}_k(s_0))^{m_k}$ as $\prod_{k=1}^{n'} Z_T(t-s_0,x_k')^{m_k'}$ and apply the induction hypothesis, we get a lower bound

$$\begin{aligned}\mathbb{E}\Big[\prod_{k=1}^{n}\prod_{\ell=1}^{m_k} Z_T(t-s_0,y_{k,\ell})\Big] &\geq \mathbb{E}\Big[\prod_{k=1}^{n'} Z_T(t-s_0,x_k')^{m_k'}\Big]e^{Tf_2(\delta)}\\ &\gtrsim \exp\big(T\gamma_3(t-s_0,\vec{x}',\vec{m}')+Tf_2(\delta)\big),\end{aligned} \tag{3.9}$$

where $\lim_{\delta\to 0} f_2(\delta)=0$. By applying the lower bounds (3.8)-(3.9) to the right hand side of (3.7) and then letting $\delta\to 0$, we conclude that

$$\begin{aligned}\mathbb{E}\Big[\prod_{i=1}^{n} Z_T(t,x_i)^{m_i}\Big] &\gtrsim \exp\bigg(T\Big(\sum_{k=1}^{n}\frac{(m_k^3-m_k)s_0}{24}-\frac{m_k(x_k-\boldsymbol{\xi}_k(s_0))^2}{2s_0}+\gamma_3(t-s_0,\vec{x}',\vec{m}')\Big)\bigg)\\ &= e^{T\gamma_3(t,\vec{x},\vec{m})}.\end{aligned}$$

The last equality is due to Lemma B.1. □

4 A continuity result for the moments

The main result in this section is Proposition 4.2, which proves a continuity result for the moments of the SHE. We prove the following lemma as a preparation.

Lemma 4.1. *Recall that* $\mathsf{q}(t,x) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$. *For* $w_i, v_i \in \mathbb{R}$, $i = 1, \dots, n$ *and* $T(t-r) \geq 2\pi$, *we have*

$$\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big] \leq 69 n! \exp\Big(\frac{n^3 T(t-r)}{24}\Big) \prod_{i=1}^n \mathsf{q}(T(t-r), T(w_i - v_i)). \tag{4.1}$$

$$\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big] \geq \prod_{i=1}^n \mathsf{q}(T(t-r), T(w_i - v_i)). \tag{4.2}$$

Proof. To prove (4.1), we apply Hölder's inequality and get

$$\begin{aligned}\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big] &\leq \prod_{i=1}^n \big(\mathbb{E}[Z_T(r, v_i; t, w_i)^n]\big)^{1/n} \\ &\leq 69 n! \exp\Big(\frac{n^3 T(t-r)}{24}\Big) \prod_{i=1}^n \mathsf{q}(T(t-r), T(w_i - v_i)),\end{aligned}$$

where the last equality follows from [14, Lemma 4.1].

To prove (4.2), by Lemma A.1 and $\mathbb{E}[Z_T(r, v_i; t, w_i)] = \mathsf{q}(T(t-r), T(w_i - v_i))$, we have

$$\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big] \geq \prod_{i=1}^n \mathbb{E}[Z_T(r, v_i; t, w_i)] = \prod_{i=1}^n \mathsf{q}(T(t-r), T(w_i - v_i)).$$

This concludes the lemma. □

Let $\vec{w} = (w_1, \dots, w_n)$. Define $\|\vec{w}\|_\infty = \max_{i=1,\dots,n} |w_i|$ and do it similarly for $\vec{w}', \vec{v}, \vec{v}'$. The following proposition is the main result of this section.

Proposition 4.2. *For fixed* $R, \epsilon > 0$, $r < t$ *and* $n \in \mathbb{Z}_{\geq 1}$, *there exists* $\delta > 0$ *such that*

$$-\epsilon \leq \liminf_{T\to\infty} T^{-1} \log \frac{\mathbb{E}[\prod_{i=1}^n Z_T(r, v_i; t, w_i)]}{\mathbb{E}[\prod_{i=1}^n Z_T(r, v_i'; t, w_i')]} \leq \limsup_{T\to\infty} T^{-1} \log \frac{\mathbb{E}[\prod_{i=1}^n Z_T(r, v_i; t, w_i)]}{\mathbb{E}[\prod_{i=1}^n Z_T(r, v_i'; t, w_i')]} \leq \epsilon$$

The inequality holds uniformly for $\limsup$ *and* $\liminf$ *in* $\vec{w}, \vec{w}', \vec{v}, \vec{v}'$ *satisfying* $\|\vec{w} - \vec{w}'\|_\infty, \|\vec{v} - \vec{v}'\|_\infty \leq \delta$ *and* $\|\vec{w}\|_\infty, \|\vec{w}'\|_\infty, \|\vec{v}\|_\infty, \|\vec{v}'\|_\infty \leq R$.

Proof. By the Feynman-Kac formula (3.2), the stochastic process $\{Z_T(r, v; t, w)\}_{(w,v)\in\mathbb{R}}$ has the same probability distribution as $\{Z_T(r, w; t, v)\}_{(w,v)\in\mathbb{R}}$. Hence, for the upper bounds, we can assume $\vec{w}' = \vec{w}$ in the proposition. It suffices to prove that

$$\limsup_{T\to\infty} T^{-1} \log \frac{\mathbb{E}[\prod_{i=1}^n Z_T(r, v_i; t, w_i)]}{\mathbb{E}[\prod_{i=1}^n Z_T(r, v_i'; t, w_i)]} \leq \epsilon$$

holds uniformly for $\vec{w}, \vec{v}$ and $\vec{v}'$. The proof of the lower bound for $\liminf$ can be obtained by swapping $\vec{v}, \vec{v}'$.

Write the constant $C = C(n, R, r, t)$ to simplify the notation. It is enough to prove that

$$\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big] \leq C e^{\frac{1}{2}T\epsilon} \mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i'; t, w_i)\Big] + \frac{1}{2} \mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big]. \tag{4.3}$$

We fix $t_0 \in (r, t)$ which will be specified later. Apply the semigroup identity (3.3) and then Fubini's theorem, we have

$$\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big] = \int_{\mathbb{R}^n} T^n \mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t_0, y_i)\Big] \mathbb{E}\Big[\prod_{i=1}^n Z_T(t_0, y_i; t, w_i)\Big] \prod_{i=1}^n dy_i. \tag{4.4}$$

Fix a constant K that will be specified later. Let $A := \{\vec{y} = (y_1, \dots, y_n) \in \mathbb{R}^n : \|\vec{y}\|_\infty \le K\}$. Write the above integral $\int_{\mathbb{R}^n} \dots$ as $\mathsf{E}_1 + \mathsf{E}_2$ where $\mathsf{E}_1 = \int_A \dots$ and $\mathsf{E}_2 = \int_{A^c} \dots$. To prove (4.3), we need to show that

$$\mathsf{E}_1 \le Ce^{\frac{1}{2}T\epsilon}\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i'; t, w_i)\Big], \quad \mathsf{E}_2 \le \frac{1}{2}\mathbb{E}\Big[\prod_{i=1}^n Z_T(r, v_i; t, w_i)\Big]. \tag{4.5}$$

We first prove the upper bound for E_2. Apply (4.1) to upper bound the first and second expectations in the integrand of E_2, we get

$$\mathsf{E}_2 \le 69n!e^{\frac{n^3T(t-r)}{24}} \int_{A^c} T^n \prod_{i=1}^n \mathsf{q}(T(t_0 - r), T(y_i - v_i))\mathsf{q}(T(t - t_0), T(w_i - y_i))dy_i. \tag{4.6}$$

Take a large enough $K = (1 + t - r)(10R + 69n!)$, it is straightforward to see that for all $\|\vec{w}\|_\infty, \|\vec{v}\|_\infty \le R$ and $t_0 \in (t, r)$, we have,

$$69n!e^{\frac{n^3T(t-r)}{24}} \int_{A^c} T^n \frac{\prod_{i=1}^n \mathsf{q}(T(t_0 - r), T(y_i - v_i))\mathsf{q}(T(t - t_0), T(w_i - y_i))dy_i}{\prod_{i=1}^n \mathsf{q}(T(t - r), T(v_i - w_i))} \le \frac{1}{2}. \tag{4.7}$$

The inequality (4.7) can be proved by interpreting the fraction above as the probability density function of n independent Brownian bridges starting from $T\vec{v}$ at time 0 and ending at $T\vec{w}$ at time $T(t - r)$. The integral is the probability of the event that at least one of the Brownian bridges go beyond $[-TK, TK]$ at time $T(t_0 - r)$.

Continue our proof and multiply both sides of (4.7) by $\prod_{i=1}^n \mathsf{q}(T(t - r), T(v_i - w_i))$, apply the resulting inequality to upper bound the right hand side of (4.6) and finally apply (4.2), we get $\mathsf{E}_2 \le \frac{1}{2}\mathbb{E}[\prod_{i=1}^n Z_T(r, v_i; t, w_i)]$.

We proceed to upper bound E_1. Recall that $\mathsf{E}_1 = \int_A \dots$ where $\dots$ is given by the integrand on the right hand side of (4.4). By applying (4.1) to upper bound the first expectation in the integrand, we have

$$\mathsf{E}_1 \le Ce^{\frac{1}{24}n^3T(t_0-r)} \int_A T^n \mathbb{E}\Big[\prod_{i=1}^n Z_T(t_0, y_i; t, w_i)\Big] \prod_{i=1}^n \mathsf{q}(T(t_0 - r), T(y_i - v_i))dy_i. \tag{4.8}$$

Define $\mathsf{p}(t, x) := -\frac{x^2}{2t}$. It is straightforward to check that for $\|\vec{y}\|_\infty \le K$ and $\|\vec{v}\|_\infty$, $\|\vec{v}'\|_\infty \le R$, we have

$$\Big|\sum_{i=1}^n (\mathsf{p}(T(t_0 - r), T(y_i - v_i)) - \sum_{i=1}^n \mathsf{p}(T(t_0 - r), T(y_i - v_i'))\Big| \le \frac{CT\|\vec{v} - \vec{v}'\|_\infty}{t_0 - r}.$$

Take the exponential, the above inequality implies that

$$\prod_{i=1}^n \mathsf{q}(T(t_0 - r), T(y_i - v_i)) \le \exp\Big(\frac{CT\|\vec{v} - \vec{v}'\|_\infty}{t_0 - r}\Big) \prod_{i=1}^n \mathsf{q}(T(t_0 - r), T(y_i - v_i')). \tag{4.9}$$

Apply (4.9) to upper bound the right hand side of (4.8), and then release the domain of integral to $\mathbb{R}^n$, we have

$$\mathsf{E}_1 \le Ce^{\frac{1}{24}n^3T(t_0-r)} \exp\Big(\frac{CT\|\vec{v} - \vec{v}'\|_\infty}{t_0 - r}\Big) \cdot \int_{\mathbb{R}^n} T^n \mathbb{E}\Big[\prod_{i=1}^n Z_T(t_0, y_i; t, w_i)\Big] \prod_{i=1}^n \mathsf{q}(T(t_0 - r), T(y_i - v_i'))dy_i.$$

Apply (4.2) to upper bound $\prod_{i=1}^{n} \mathsf{q}(T(t_0 - r), T(y_i - v_i'))$ by $\mathbb{E}[\prod_{i=1}^{n} Z_T(r, v_i'; t_0, y_i)]$ and then use the Fubini's theorem together with the semigroup identity (3.3), we get

$$\mathsf{E}_1 \le Ce^{\frac{1}{24}n^3 T(t_0-r)} \exp\Big(\frac{CT\|\vec{v}-\vec{v}'\|_\infty}{t_0-r}\Big)\mathbb{E}\Big[\prod_{i=1}^{n} Z_T(r, v_i'; t, w_i)\Big].$$

Take $\delta = \min(\frac{\epsilon^2}{16C^2}, \frac{36\epsilon^2}{n^6})$, $t_0 = r + \delta^{\frac{1}{2}}$. Using $\|\vec{v}-\vec{v}'\|_\infty \le \delta$, we obtain the upper bound (4.5) and thus (4.3). ☐

5 Proof of Theorem 1.1 and its corollaries

In this section, we will prove the following lemma and conclude the proof of Theorem 1.1.

Lemma 5.1. *We have* $\gamma_1 \le \gamma_2 \le \gamma_3$.

Proof. We first prove $\gamma_1 \le \gamma_2$ by finding $a_1, \ldots, a_{\mathfrak{m}}$ which satisfy $a_i - a_{i+1} \ge 1$ for $i = 1, \ldots, \mathfrak{m} - 1$ and $\sum_{i=1}^{\mathfrak{m}} \frac{t}{2} a_i^2 + u_i a_i = \gamma_2(t, \vec{x}, \vec{m})$. Let $(b_1, \ldots, b_n)$ be a minimizer of γ_2. For $j = 1, \ldots, n$, set $S_j := \sum_{i=1}^{j} m_i$ and

$$a_k := b_j + \frac{m_j+1}{2} - k + \sum_{i=1}^{j-1} m_i, \qquad \text{if } S_{j-1} < k \le S_j. \tag{5.1}$$

Note that we have $a_k - a_{k+1} = 1$ unless $k = S_j$ for some $j \in \{1, \ldots, n-1\}$. Since $b_j - b_{j+1} \ge \frac{m_j + m_{j+1}}{2}$ for $j = 1, \ldots, n-1$, we know that $a_k - a_{k+1} \ge 1$ for $k \in \{S_1, \ldots, S_{n-1}\}$. This implies that $a_k - a_{k+1} \ge 1$ for all $k = 1, \ldots, \mathfrak{m}-1$. Recall that $u_k = x_j$ if $S_{j-1} < k \le S_j$, one can directly check that

$$\sum_{j=1}^{\mathfrak{m}} \frac{t}{2} a_j^2 + \sum_{j=1}^{\mathfrak{m}} u_j a_j = \sum_{i=1}^{n} \sum_{k=S_{i-1}+1}^{S_i} \frac{t}{2} a_k^2 + x_i a_k. \tag{5.2}$$

In addition, we have for each $i \in \{1, \ldots, n\}$,

$$\begin{aligned}
&\sum_{k=S_{i-1}+1}^{S_i} \frac{t}{2} a_k^2 + x_i a_k \\
&= \sum_{k=S_{i-1}+1}^{S_i} \frac{t}{2}\Big(b_i + \frac{m_i+1}{2} - k + \sum_{j=1}^{i-1} m_i\Big)^2 + x_i\Big(b_i + \frac{m_i+1}{2} - k + \sum_{j=1}^{i-1} m_i\Big) \\
&= \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 + \frac{(m_i^3 - m_i)t}{24} - \frac{m_i x_i^2}{2t}
\end{aligned} \tag{5.3}$$

In the second equality, the term $\frac{m_i^3 - m_i}{24}$ comes from the sum $\sum_{k=Si-1+1}^{S_i} (\frac{m_i+1}{2} - k + \sum_{j=1}^{i-1} m_i)^2$. By (5.2) and (5.3), we have

$$\sum_{j=1}^{\mathfrak{m}} \frac{t}{2} a_j^2 + \sum_{j=1}^{\mathfrak{m}} u_j a_j = \sum_{i=1}^{n} \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 + \frac{(m_i^3 - m_i)t}{24} - \frac{m_i x_i^2}{2t} = \gamma_2(t, \vec{x}, \vec{m}). \tag{5.4}$$

This implies that $\gamma_1 \le \gamma_2$.

We proceed to show $\gamma_2 \le \gamma_3$ by finding $b_1, \ldots, b_n$ which satisfy $b_i - b_{i+1} \ge \frac{m_i + m_{i+1}}{2}$ for $i = 1, \ldots, n-1$ and $\sum_{i=1}^{n} \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 + \frac{(m_i^3 - m_i)t}{24} - \frac{m_i x_i^2}{2t} = \gamma_3(t, \vec{x}, \vec{m})$. Recall the notation for the inertia clusters from Section 3.1. Let $n_i := |\mathfrak{b}_i|$ for $i = 1, \dots, \mathfrak{n}$ and $N_k := \sum_{i=1}^{k} n_i$. Then, we have $\mathfrak{b}_k = \{N_{k-1}+1, \dots, N_k\}$. Let

$$\mathfrak{c}_k(s) := \frac{\sum_{i=N_{k-1}+1}^{N_k} m_i \boldsymbol{\zeta}_i(s)}{\mathfrak{m}_{\mathfrak{b}_k}}.$$

Note that $\mathfrak{c}_k(s)$ is the location of the center of mass for $\{\boldsymbol{\zeta}_i(s)\}_{i\in\mathfrak{b}_k}$. An important observation is that $\mathfrak{c}_k$ travels with a constant speed $\varphi_k := \frac{1}{2}(\sum_{i=N_k+1}^{n} m_i - \sum_{i=1}^{N_{k-1}} m_i)$. Since $\boldsymbol{\zeta}_i(0) = x_i$, we have $\mathfrak{c}_k(0) = \frac{\sum_{i=N_{k-1}+1}^{N_k} m_i x_i}{\mathfrak{m}_{\mathfrak{b}_k}}$. Moreover, we have $\mathfrak{c}_k(t) = \boldsymbol{\zeta}_{\mathfrak{b}_k}(t)$, which implies that $\mathfrak{c}_k(t) < \mathfrak{c}_{k+1}(t)$ for all $k = 1, \dots, \mathfrak{n}-1$. Using this together with $\mathfrak{c}_j(t) = \varphi_j t + \mathfrak{c}_j(0)$, we have $\mathfrak{c}_{k+1}(0) - \mathfrak{c}_k(0) > -(\varphi_{k+1} - \varphi_k)t$, which is equivalent to

$$\frac{\sum_{i=N_k+1}^{N_{k+1}} m_i x_i}{\mathfrak{m}_{\mathfrak{b}_{k+1}}} - \frac{\sum_{i=N_{k-1}+1}^{N_k} m_i x_i}{\mathfrak{m}_{\mathfrak{b}_k}} > \frac{(\sum_{i=N_{k-1}+1}^{N_{k+1}} m_i)t}{2}. \tag{5.5}$$

We set

$$b_i := -\frac{\sum_{j=N_{k-1}+1}^{N_k} m_j x_j}{\mathfrak{m}_{\mathfrak{b}_k} t} + \frac{\sum_{j=i+1}^{N_k} m_j - \sum_{j=N_{k-1}+1}^{i-1} m_j}{2}, \qquad \text{for } N_{k-1} < i \le N_k. \tag{5.6}$$

By (5.5), one can verify that $b_i - b_{i+1} \ge \frac{m_i + m_{i+1}}{2}$ for $i = 1, \dots, n-1$. Recall that

$$\gamma_3(t, \vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j}) = \frac{(\mathfrak{m}_{\mathfrak{b}_j}^3 - \mathfrak{m}_{\mathfrak{b}_j})t}{24} - \sum_{k,\ell\in\mathfrak{b}_j, k<\ell} \frac{m_k m_\ell}{2}|x_k - x_\ell| - \frac{(\sum_{k\in\mathfrak{b}_j} m_k x_k)^2}{2t\mathfrak{m}_{\mathfrak{b}_j}}.$$

A straightforward (although tedious) computation implies that

$$\sum_{i=N_{j-1}+1}^{N_j} \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 + \frac{(m_i^3 - m_i)t}{24} - \frac{m_i x_i^2}{2t} = \gamma_3(t, \vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j}). \tag{5.7}$$

See Appendix C for detail. Summing both sides above over $j = 1, \dots, \mathfrak{n}$, we have

$$\sum_{i=1}^{n} \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 + \frac{(m_i^3 - m_i)t}{24} - \frac{m_i x_i^2}{2t} = \gamma_3(t, \vec{x}, \vec{m}). \tag{5.8}$$

This shows that $\gamma_2 \le \gamma_3$. □

Proof of Theorem 1.1. Apply Propositions 2.1 and 3.1, we know that $\gamma_1 \ge \gamma_3$. Using this together Lemma 5.1, we have $\gamma_1 = \gamma_2 = \gamma_3$. Using this together with Propositions 2.1 and 3.1, we conclude Theorem 1.1. □

Proof of Corollary 1.2. Use the expression of γ_2, the result is straightforward for $n = 1$. When $n = 2$, we have

$$\gamma_2(t, \vec{x}, \vec{m}) := \inf\Big\{ \sum_{i=1}^{2} \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 + \frac{(m_i^3 - m_i)t}{24} - \frac{m_i x_i^2}{2t} : b_1 - b_2 \ge \frac{m_1 + m_2}{2} \Big\}.$$

When $\frac{x_2 - x_1}{t} \ge \frac{m_1 + m_2}{2}$, the target function is minimized at $b_1 = -\frac{x_1}{t}$ and $b_2 = -\frac{x_2}{t}$. When $0 < \frac{x_2 - x_1}{t} \le \frac{m_1 + m_2}{2}$, the target function in the infimum is minimized at the boundary $b_1 - b_2 = \frac{m_1 + m_2}{2}$ with $b_1 = -\frac{m_1 x_1 + m_2 x_2}{(m_1 + m_2)t} + \frac{m_2}{2}$. Insert the minimizers into the target function, we obtain the desired result. □

Proof of Corollary 1.3. Since the target function that we want to minimize in γ_1 is continuous and goes to infinity when we send $\sum_{i=1}^{\mathfrak{m}} |a_i|$ to infinity, we know that the infimum in γ_1 has a minimizer. Moreover, the domain $\{(a_1, \dots, a_{\mathfrak{m}}) : a_i - a_{i+1} \geq 1, \forall\, i \in \{1, \dots, \mathfrak{m}-1\}\}$ is convex and the target function is strictly convex, thus the infimum in γ_1 has a unique minimizer $(a_1^*, \dots, a_{\mathfrak{m}}^*)$. By a similar argument, we know that γ_2 has a unique minimizer $(b_1^*, \dots, b_n^*)$. Using (5.1) - (5.4) and $\gamma_1 = \gamma_2$, we have

$$a_k^* = b_j^* + \frac{m_j + 1}{2} - k + \sum_{i=1}^{j-1} m_i, \qquad \text{if } S_{j-1} < k \leq S_j. \tag{5.9}$$

Moreover, use (5.6) - (5.7) and $\gamma_2 = \gamma_3$, we know that

$$b_i^* := -\frac{\sum_{j=N_{k-1}+1}^{N_k} m_j x_j}{\mathfrak{m}_{\mathfrak{b}_k} t} + \frac{\sum_{j=i+1}^{N_k} m_j - \sum_{j=N_{k-1}+1}^{i-1} m_j}{2}, \qquad \text{for } N_{k-1} < i \leq N_k.$$

By (5.5), we have $b_i^* - b_{i+1}^* = \frac{m_i + m_{i+1}}{2}$ if $i, i+1$ belong to the same $\mathfrak{b}_j$ and $b_i^* - b_{i+1}^* > \frac{m_i + m_{i+1}}{2}$ if not. Use this together with (5.9), we conclude that $a_i^* - a_{i+1}^* = 1$ if and only if $f(i)$ and $f(i+1)$ belong to the same $\mathfrak{b}_j$. ☐

A A correlation inequality

In this section, we prove the following inequality.

Lemma A.1. *For any $T > 0$ and $t > r$, integers $1 \leq k \leq n$, and real numbers $w_1, \dots, w_n$, $v_1, \dots, v_n$, we have*

$$\mathbb{E}\Big[\prod_{i=1}^{n} Z_T(r, v_i; t, w_i)\Big] \geq \mathbb{E}\Big[\prod_{i=1}^{k} Z_T(r, v_i; t, w_i)\Big]\mathbb{E}\Big[\prod_{i=k+1}^{n} Z_T(r, v_i; t, w_i)\Big].$$

Proof. Without loss of generality, we can take $T = 1$ and write Z_1 as Z. We claim that for any positive real numbers $s_1, \dots, s_n$, we have

$$\begin{aligned} &\mathbb{P}\Big(\bigcap_{i=1}^{n}\{Z(r, v_i; t, w_i) \geq s_i\}\Big) \\ &\geq \mathbb{P}\Big(\bigcap_{i=1}^{k}\{Z(r, v_i; t, w_i) \geq s_i\}\Big)\mathbb{P}\Big(\bigcap_{i=k+1}^{n}\{Z(r, v_i; t, w_i) \geq s_i\}\Big). \end{aligned} \tag{A.1}$$

The proof of the claim follows the idea of [19, Proposition 1] and uses the FKG-Harris inequality (see [11, Proposition A.1]) at the level of the discrete polymer model. By [1, Theorem 2.7], we can approximate the four-parameter process $Z(r, y; t, x)$ in terms of the partition function of discrete polymer models, which is denoted as $\mathcal{Z}_\epsilon(r, y; t, x)$. More precisely, at the process level, $\mathcal{Z}_\epsilon(r, y; t, x)$ converges in distribution to $Z(r, y; t, x)$ as $\epsilon \to 0$. It is straightforward that $\mathcal{Z}_\epsilon(r, y; t, x)$ is an increasing function of the i.i.d. random variables that we put on the lattice $\mathbb{Z}^2$ in the discrete polymer models. Hence, the events $A_{1,\epsilon} = \cap_{i=1}^{k}\{\mathcal{Z}_\epsilon(r, v_i; t, w_i) \geq s_i\}$ and $A_{2,\epsilon} = \cap_{i=k+1}^{n}\{\mathcal{Z}_\epsilon(r, v_i; t, w_i) \geq s_i\}$ are increasing events. Note that although the lattice $\mathbb{Z}^2$ is infinite, however, only a finite number of random variables on the lattice affects the value of $\mathcal{Z}_\epsilon(r, v_i; t, w_i)$ for $i = 1, \dots, n$. Note that these random variables are all independent due to the definition of the discrete polymer. By the FKG-Harris inequality, $\mathbb{P}(A_{1,\epsilon} \cap A_{2,\epsilon}) \geq \mathbb{P}(A_{1,\epsilon})\mathbb{P}(A_{2,\epsilon})$. Send $\epsilon \to 0$, we conclude (A.1).

We proceed to conclude Lemma A.1. By applying Fubini's theorem, we have

$$\mathbb{E}\Big[\prod_{i=1}^{n} Z(r, v_i; t, w_i)\Big] = \int_{\mathbb{R}_{>0}^n} \mathbb{P}\Big(\bigcap_{i=1}^{n}\{Z(r, v_i; t, w_i) \geq s_i\}\Big) ds_1 \dots ds_n.$$

Apply (A.1) to the right hand side and then use Fubini's theorem, we have

$$\begin{aligned}
&\mathbb{E}\Big[\prod_{i=1}^{n} Z(r,v_i;t,w_i)\Big] \\
&\geq \int_{\mathbb{R}^k_{>0}} \mathbb{P}\Big(\bigcap_{i=1}^{k}\{Z(r,v_i;t,w_i)\geq s_i\}\Big)\int_{\mathbb{R}^{n-k}_{>0}} \mathbb{P}\Big(\bigcap_{i=k+1}^{n}\{Z(r,v_i;t,w_i)\geq s_i\}\Big)\prod_{i=1}^{n} ds_i \\
&= \mathbb{E}\Big[\prod_{i=1}^{k} Z(r,v_i;t,w_i)\Big]\mathbb{E}\Big[\prod_{i=k+1}^{n} Z(r,v_i;t,w_i)\Big]. \qquad \square
\end{aligned}$$

B An identity

In this section, we will prove the last equality in the proof of Proposition 3.1. Recall that for the optimal clusters starting from $(\vec{x},\vec{m})$, we have assumed $\mathfrak{n}=1$ and set $s_0=\inf\{s:|\{\boldsymbol{\xi}_1(s),\dots,\boldsymbol{\xi}_n(s)\}|<n\}$. We let $(\vec{x}',\vec{m}')$ be the (different) locations and weights of the point masses at time $s=s_0$. Note that $(\vec{x}',\vec{m}')$ is obtained from $(\vec{\boldsymbol{\xi}}(s_0),\vec{m})$ by examining the point masses that have merged.

Lemma B.1. *We have*

$$\sum_{k=1}^{n}\frac{(m_k^3-m_k)s_0}{24}-\frac{m_k(x_k-\boldsymbol{\xi}_k(s_0))^2}{2s_0}+\gamma_3(t-s_0,\vec{x}',\vec{m}')=\gamma_3(t,\vec{x},\vec{m}). \tag{B.1}$$

Proof. Recall from (3.6) that s_0 is the first time that a merging occurs among point masses in the optimal clusters. Set $k\in\{1,\dots,n-1\}$ to be a fixed number such that

$$\frac{x_{k+1}-x_k}{(m_k+m_{k+1})/2}=\min_{j\in\{1,\dots,n-1\}}\frac{x_{j+1}-x_j}{(m_j+m_{j+1})/2}. \tag{B.2}$$

By (B.2), we know that the point masses starting from x_k and x_{k+1} will merge first. Moreover, the speed of particle x_k and x_{k+1} are, respectively, $\frac{1}{2}(\sum_{i=k+1}^{n}m_i-\sum_{i=1}^{k-1}m_i)$ and $\frac{1}{2}(\sum_{i=k+2}^{n}m_i-\sum_{i=1}^{k}m_i)$. This implies the time for the k-th and $k+1$-th point masses to merge is $s_0=\frac{x_{k+1}-x_k}{(m_k+m_{k+1})/2}$.

It is not hard to check that $\boldsymbol{\xi}_k(s_0)=x_k+(\phi_k-v)s_0$ with

$$\phi_k=\frac{\sum_{j=k+1}^{n}m_j-\sum_{j=1}^{k-1}m_j}{2}\ \text{and}\ v=\frac{\sum_{j=1}^{n}m_jx_j}{\mathfrak{m}t}. \tag{B.3}$$

Since $\mathfrak{n}=1$, we have

$$\gamma_3(t,\vec{x},\vec{m})=\frac{(\mathfrak{m}^3-\mathfrak{m})t}{24}-\sum_{1\leq j<k\leq n}\frac{m_jm_k(x_k-x_j)}{2}-\frac{(\sum_{j=1}^{n}m_jx_j)^2}{2\mathfrak{m}t},$$

$$\gamma_3(t-s_0,\vec{x}',\vec{m}')=\frac{(\mathfrak{m}^3-\mathfrak{m})(t-s_0)}{24}-\sum_{1\leq j<k\leq n}\frac{m_jm_k(\boldsymbol{\xi}_k(s_0)-\boldsymbol{\xi}_j(s_0))}{2}-\frac{(\sum_{j=1}^{n}m_j\boldsymbol{\xi}_j(s_0))^2}{2\mathfrak{m}(t-s_0)}.$$

This implies that

$$\begin{aligned}
&\gamma_3(t,\vec{x},\vec{m})-\gamma_3(t-s_0,\vec{x}',\vec{m}') \\
&=\frac{(\mathfrak{m}^3-\mathfrak{m})s_0}{24}-\sum_{1\leq j<k\leq n}\frac{m_jm_k(x_k-x_j)}{2}-\frac{(\sum_{j=1}^{n}m_jx_j)^2}{2\mathfrak{m}t} \\
&\quad+\sum_{1\leq j<k\leq n}\frac{m_jm_k(\boldsymbol{\xi}_k(s_0)-\boldsymbol{\xi}_j(s_0))}{2}+\frac{(\sum_{j=1}^{n}m_j\boldsymbol{\xi}_j(s_0))^2}{2\mathfrak{m}(t-s_0)}.
\end{aligned} \tag{B.4}$$

By the definition of $\boldsymbol{\xi}$ and (B.3), we have

$$\frac{m_j m_k(\boldsymbol{\xi}_k(s_0) - \boldsymbol{\xi}_j(s_0))}{2} = \frac{m_j m_k(x_k - x_j)}{2} + \frac{m_j m_k(\phi_k - \phi_j)}{2}.$$

Inserting this into (B.4), we get

$$\begin{aligned}&\gamma_3(t, \vec{x}, \vec{m}) - \gamma_3(t - s_0, \vec{x}', \vec{m}') \\ &= \frac{(\mathfrak{m}^3 - \mathfrak{m})s_0}{24} - \frac{(\sum_{j=1}^n m_j x_j)^2}{2\mathfrak{m}t} + \sum_{1 \le j < k \le n} \frac{m_j m_k(\phi_k - \phi_j)s_0}{2} + \frac{(\sum_{j=1}^n m_j \boldsymbol{\xi}_j(s_0))^2}{2\mathfrak{m}(t - s_0)}.\end{aligned} \tag{B.5}$$

Next, we observe that

$$\frac{(\sum_{j=1}^n m_j x_j)^2}{2\mathfrak{m}t} = \frac{\mathfrak{m}v^2 s_0}{2}, \qquad \frac{(\sum_{j=1}^n m_j \boldsymbol{\xi}_j(s_0))^2}{2\mathfrak{m}(t - s_0)} = \frac{\mathfrak{m}v^2(t - s_0)}{2}.$$

Inserting these into (B.5), we get

$$\gamma_3(t, \vec{x}, \vec{m}) - \gamma_3(t - s_0, \vec{x}', \vec{m}') = \frac{(\mathfrak{m}^3 - \mathfrak{m})s_0}{24} + \sum_{1 \le j < k \le n} \frac{m_j m_k(\phi_k - \phi_j)s_0}{2} - \frac{\mathfrak{m}v^2 s_0}{2}. \tag{B.6}$$

On the other hand, we have

$$\begin{aligned}&\sum_{k=1}^n \frac{(m_k^3 - m_k)s_0}{24} - \frac{m_k(x_k - \boldsymbol{\xi}_k(s_0))^2}{2s_0} \\ &= \sum_{k=1}^n \frac{(m_k^3 - m_k)s_0}{24} - \frac{m_k(\phi_k + v)^2 s_0}{2} \\ &= \sum_{k=1}^n \Big(\frac{(m_k^3 - m_k)s_0}{24} - \frac{m_k \phi_k^2 s_0}{2}\Big) - \sum_{k=1}^n m_k \phi_k v s_0 - \frac{\mathfrak{m}v^2 s_0}{2}.\end{aligned}$$

One can verify that $\sum_{k=1}^n m_k \phi_k = 0$. As a result,

$$\sum_{k=1}^n \frac{(m_k^3 - m_k)s_0}{24} - \frac{m_k(x_k - \boldsymbol{\xi}_k(s_0))^2}{2s_0} = \sum_{k=1}^n \Big(\frac{(m_k^3 - m_k)s_0}{24} - \frac{m_k \phi_k^2 s_0}{2}\Big) - \frac{\mathfrak{m}v^2 s_0}{2}. \tag{B.7}$$

One can check that

$$\frac{(\mathfrak{m}^3 - \mathfrak{m})s_0}{24} + \sum_{1 \le j < k \le n} \frac{m_j m_k(\phi_k - \phi_j)s_0}{2} = \sum_{k=1}^n \Big(\frac{(m_k^3 - m_k)s_0}{24} - \frac{m_k \phi_k^2 s_0}{2}\Big).$$

Using this together with (B.6) and (B.7), we have

$$\frac{(\mathfrak{m}^3 - \mathfrak{m})s_0}{24} + \sum_{1 \le j < k \le n} \frac{m_j m_k(\phi_k - \phi_j)s_0}{2} = \gamma_3(t, \vec{x}, \vec{m}) - \gamma_3(t - s_0, \vec{x}', \vec{m}').$$

This concludes the lemma. □

C Detailed computation for (5.7)

We present detailed computation for (5.7). It is clear that (5.7) is equivalent to

$$\sum_{i=N_{j-1}+1}^{N_j} \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 - \frac{m_i x_i^2}{2t} = \gamma_3(t, \vec{x}_{\mathfrak{b}_j}, \vec{m}_{\mathfrak{b}_j}) - \sum_{i=N_{j-1}+1}^{N_j} \frac{(m_i^3 - m_i)t}{24}. \tag{C.1}$$

We first look at the left hand side of (C.1):

$$\sum_{i=N_{j-1}+1}^{N_j} \frac{m_i t}{2}(b_i + \frac{x_i}{t})^2 - \frac{m_i x_i^2}{2t} = \sum_{i=N_{j-1}+1}^{N_j} \frac{m_i b_i^2 t}{2} + m_i x_i b_i. \tag{C.2}$$

Let $d_j := \frac{\sum_{k=N_{j-1}+1}^{N_j} m_k x_k}{\mathfrak{m}_{\mathfrak{b}_j} t}$. Recall from (5.6) that for $N_{j-1} < i \le N_j$, we have

$$\begin{aligned} b_i &= -\frac{\sum_{k=N_{j-1}+1}^{N_j} m_k x_k}{\mathfrak{m}_{\mathfrak{b}_j} t} + \frac{\sum_{k=i+1}^{N_j} m_k - \sum_{k=N_{j-1}+1}^{i-1} m_k}{2} \\ &= -d_j + \frac{\sum_{k=i+1}^{N_j} m_k - \sum_{k=N_{j-1}+1}^{i-1} m_k}{2}. \end{aligned}$$

Plugging this into (C.2), we have

$$\begin{aligned} \sum_{i=N_{j-1}+1}^{N_j} \frac{m_i t b_i^2}{2} + m_i x_i b_i &= \sum_{i=N_{j-1}+1}^{N_j} \frac{m_i t}{2}\Big(- d_j + \frac{\sum_{k=i+1}^{N_j} m_k - \sum_{k=N_{j-1}+1}^{i-1} m_k}{2}\Big)^2 \\ &\quad + m_i x_i \Big(- d_j + \frac{\sum_{k=i+1}^{N_j} m_k - \sum_{k=N_{j-1}+1}^{i-1} m_k}{2}\Big) \\ &= A_1 + A_2 + A_3 + A_4 \end{aligned} \tag{C.3}$$

where

$$\begin{aligned} A_1 &:= \frac{m_i t}{2} d_j^2 \Big(\sum_{k=N_{j-1}+1}^{N_j} m_k \Big) - \Big(\sum_{k=N_{j-1}+1}^{N_j} m_k x_k \Big) d_j, \\ A_2 &:= \sum_{i=N_{j-1}+1}^{N_j} m_i x_i \frac{\sum_{k=i+1}^{N_j} m_k - \sum_{k=N_{j-1}+1}^{i-1} m_k}{2}, \\ A_3 &:= -\sum_{i=N_{j-1}+1}^{N_j} \frac{m_i t}{2} d_j \frac{\sum_{k=i+1}^{N_j} m_k - \sum_{k=N_{j-1}+1}^{i-1} m_k}{2}, \\ A_4 &:= \sum_{i=N_{j-1}+1}^{N_j} \frac{1}{8} t m_i \Big(\sum_{k=i+1}^{N_j} m_k - \sum_{k=N_{j-1}+1}^{i-1} m_k \Big)^2. \end{aligned}$$

It is straightforward to verify that

$$\begin{aligned} A_1 &= -\frac{(\sum_{k\in\mathfrak{b}_j} m_k x_k)^2}{2t\mathfrak{m}_{\mathfrak{b}_j}}, \\ A_2 &= -\sum_{N_{j-1}+1\le k<\ell\le N_j} \frac{m_k m_\ell}{2}(x_\ell - x_k), \\ A_3 &= 0, \\ A_4 &= \frac{(\mathfrak{m}_{\mathfrak{b}_i}^3 - \mathfrak{m}_{\mathfrak{b}_i})t}{24} - \sum_{i=N_{j-1}+1}^{N_j} \frac{(m_i^3 - m_i)t}{24}. \end{aligned}$$

Plugging these into (C.3), we conclude with (C.1), thus establishing (5.7).

References

[1] Tom Alberts, Konstantin Khanin, and Jeremy Quastel. The intermediate disorder regime for directed polymers in dimension $1+1$. *Ann. Probab.*, 42(3):1212–1256, 2014. MR3189070

[2] Tomer Asida, Eli Livne, and Baruch Meerson. Large fluctuations of a Kardar-Parisi-Zhang interface on a half line: The height statistics at a shifted point. *Physical Review E*, 99(4):042132, 2019.

[3] Lorenzo Bertini and Nicoletta Cancrini. The stochastic heat equation: Feynman-Kac formula and intermittence. *Journal of statistical Physics*, 78:1377–1401, 1995. MR1316109

[4] Alexei Borodin and Ivan Corwin. Macdonald processes. *Probability Theory and Related Fields*, 158(1-2):225–400, 2014. MR3152785

[5] Alexei Borodin and Ivan Corwin. Moments and Lyapunov exponents for the parabolic Anderson model. *Ann. Appl. Probab.*, 24(3):1172–1198, 2014. MR3199983

[6] Mattia Cafasso and Tom Claeys. A Riemann-Hilbert approach to the lower tail of the Kardar-Parisi-Zhang equation. *Comm. Pure Appl. Math.*, 75(3):493–540, 2022. MR4373176

[7] Mattia Cafasso, Tom Claeys, and Giulio Ruzza. Airy kernel determinant solutions to the KdV equation and integro-differential Painlevé equations. *Comm. Math. Phys.*, 386(2):1107–1153, 2021. MR4294287

[8] Le Chen and Robert C. Dalang. Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions. *Ann. Probab.*, 43(6):3006–3051, 2015. MR3433576

[9] Le Chen, Yaozhong Hu, and David Nualart. Regularity and strict positivity of densities for the nonlinear stochastic heat equation. *Mem. Amer. Math. Soc.*, 273(1340):v+102, 2021. MR4334477

[10] Xia Chen. Precise intermittency for the parabolic Anderson equation with an $(1+1)$-dimensional time–space white noise. *Annales de l'I.H.P. Probabilités et statistiques*, 51(4):1486–1499, 2015. MR3414455

[11] Francis Comets. *Directed polymers in random environments*, volume 2175 of *Lecture Notes in Mathematics*. Springer, Cham, 2017. Lecture notes from the 46th Probability Summer School held in Saint-Flour, 2016. MR3444835

[12] Daniel Conus, Mathew Joseph, and Davar Khoshnevisan. On the chaotic character of the stochastic heat equation, before the onset of intermitttency. *Ann. Probab.*, 41(3B):2225–2260, 2013. MR3098071

[13] Ivan Corwin. The Kardar–Parisi–Zhang equation and universality class. *Random matrices: Theory and applications*, 1(01):1130001, 2012. MR2930377

[14] Ivan Corwin and Promit Ghosal. KPZ equation tails for general initial data. *Electronic Journal of Probability*, 25(none):1 – 38, 2020. MR4115735

[15] Ivan Corwin and Promit Ghosal. Lower tail of the KPZ equation. *Duke Math. J.*, 169(7):1329–1395, 2020. MR4094738

[16] Ivan Corwin, Promit Ghosal, Alexandre Krajenbrink, Pierre Le Doussal, and Li-Cheng Tsai. Coulomb-gas electrostatics controls large fluctuations of the Kardar-Parisi-Zhang equation. *Physical review letters*, 121(6):060201, 2018.

[17] Ivan Corwin and Alan Hammond. Brownian Gibbs property for Airy line ensembles. *Invent. Math.*, 195(2):441–508, 2014. MR3152753

[18] Ivan Corwin and Alan Hammond. KPZ line ensemble. *Probab. Theory Related Fields*, 166(1-2):67–185, 2016. MR3547737

[19] Ivan Corwin and Jeremy Quastel. Crossover distributions at the edge of the rarefaction fan. *Ann. Probab.*, 41(3A):1243–1314, 2013. MR3098678

[20] Sayan Das and Promit Ghosal. Law of iterated logarithms and fractal properties of the KPZ equation. *Ann. Probab.*, 51(3):930–986, 2023. MR4583059

[21] Sayan Das and Li-Cheng Tsai. Fractional moments of the stochastic heat equation. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(2):778 – 799, 2021. MR4260483

[22] Sayan Das and Weitao Zhu. Upper-tail large deviation principle for the ASEP. *Electron. J. Probab.*, 27:Paper No. 11, 34, 2022. MR4366818

[23] Gregorio R. Moreno Flores. On the (strict) positivity of solutions of the stochastic heat equation. *The Annals of Probability*, 42(4):1635–1643, 2014. MR3262487

[24] Shirshendu Ganguly and Milind Hegde. Sharp upper tail estimates and limit shapes for the KPZ equation via the tangent method. *arXiv preprint arXiv:2208.08922*, 2022.

[25] J. Gärtner and S. A. Molchanov. Parabolic problems for the Anderson model. I. Intermittency and related topics. *Comm. Math. Phys.*, 132(3):613–655, 1990. MR1069840

[26] Jürgen Gärtner, Wolfgang König, and Stanislav Molchanov. Geometric characterization of intermittency in the parabolic Anderson model. *Ann. Probab.*, 35(2):439–499, 2007. MR2308585

[27] Pierre Yves Gaudreau Lamarre, Promit Ghosal, and Yuchen Liao. Moment Intermittency in the PAM with Asymptotically Singular Noise. *arXiv preprint arXiv:2206.13622*, 2022. MR4636689

[28] Pierre Yves Gaudreau Lamarre, Yier Lin, and Li-Cheng Tsai. KPZ equation with a small noise, deep upper tail and limit shape. *Probab. Theory Related Fields*, 185(3-4):885–920, 2023. MR4556284

[29] Nicos Georgiou and Timo Seppäläinen. Large deviation rate functions for the partition function in a log-gamma distributed random potential. *Ann. Probab.*, 41(6):4248–4286, 2013. MR3161474

[30] Promit Ghosal and Yier Lin. Lyapunov exponents of the SHE under general initial data. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 59(1):476 – 502, 2023. MR4533737

[31] Yaozhong Hu, Jingyu Huang, David Nualart, and Samy Tindel. Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency. *Electron. J. Probab.*, 20:no. 55, 50, 2015. MR3354615

[32] Yaozhong Hu and Khoa Lê. Asymptotics of the density of parabolic Anderson random fields. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(1):105–133, 2022. MR4374674

[33] Chris Janjigian. Large deviations of the free energy in the O'Connell-Yor polymer. *J. Stat. Phys.*, 160(4):1054–1080, 2015. MR3373651

[34] Alex Kamenev, Baruch Meerson, and Pavel V Sasorov. Short-time height distribution in the one-dimensional Kardar-Parisi-Zhang equation: Starting from a parabola. *Physical Review E*, 94(3):032108, 2016. MR3731817

[35] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889, 1986.

[36] Davar Khoshnevisan. *Analysis of stochastic partial differential equations*, volume 119 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2014. MR3222416

[37] Davar Khoshnevisan, Kunwoo Kim, and Yimin Xiao. Intermittency and multifractality: a case study via parabolic stochastic PDEs. *Ann. Probab.*, 45(6A):3697–3751, 2017. MR3729613

[38] Yujin H. Kim. The lower tail of the half-space KPZ equation. *Stochastic Process. Appl.*, 142:365–406, 2021. MR4318370

[39] IV Kolokolov and SE Korshunov. Optimal fluctuation approach to a directed polymer in a random medium. *Physical Review B*, 75(14):140201, 2007. MR2546056

[40] IV Kolokolov and SE Korshunov. Universal and nonuniversal tails of distribution functions in the directed polymer and Kardar-Parisi-Zhang problems. *Physical Review B*, 78(2):024206, 2008.

[41] IV Kolokolov and SE Korshunov. Explicit solution of the optimal fluctuation problem for an elastic string in a random medium. *Physical Review E*, 80(3):031107, 2009.

[42] Alexandre Krajenbrink. *Beyond the typical fluctuations: a journey to the large deviations in the Kardar-Parisi-Zhang growth model*. PhD thesis, Université Paris sciences et lettres, 2019.

[43] Alexandre Krajenbrink. From painlevé to zakharov–shabat and beyond: Fredholm determinants and integro-differential hierarchies. *Journal of Physics A: Mathematical and Theoretical*, 54(3):035001, 2020. MR4209129

[44] Alexandre Krajenbrink and Pierre Le Doussal. Exact short-time height distribution in the one-dimensional Kardar-Parisi-Zhang equation with Brownian initial condition. *Physical Review E*, 96(2):020102, 2017.

[45] Alexandre Krajenbrink and Pierre Le Doussal. Large fluctuations of the KPZ equation in a half-space. *SciPost Physics*, 5(4):032, 2018.

[46] Alexandre Krajenbrink and Pierre Le Doussal. Simple derivation of the $(-\lambda H)^{5/2}$ tail for the 1D KPZ equation. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(6):063210, 2018. MR3832292

[47] Alexandre Krajenbrink and Pierre Le Doussal. Linear statistics and pushed coulomb gas at the edge of β-random matrices: Four paths to large deviations. *Europhysics Letters*, 125(2):20009, 2019.

[48] Alexandre Krajenbrink and Pierre Le Doussal. Inverse scattering of the Zakharov-Shabat system solves the weak noise theory of the Kardar-Parisi-Zhang equation. *Physical Review Letters*, 127(6):064101, 2021. MR4312166

[49] Alexandre Krajenbrink and Pierre Le Doussal. Inverse scattering solution of the weak noise theory of the Kardar-Parisi-Zhang equation with flat and Brownian initial conditions. *Physical Review E*, 105(5):054142, 2022. MR4445107

[50] Alexandre Krajenbrink and Pierre Le Doussal. Crossover from the macroscopic fluctuation theory to the Kardar-Parisi-Zhang equation controls the large deviations beyond Einstein's diffusion. *Physical Review E*, 107(1):014137, 2023. MR4548902

[51] Alexandre Krajenbrink, Pierre Le Doussal, and Sylvain Prolhac. Systematic time expansion for the Kardar–Parisi–Zhang equation, linear statistics of the GUE at the edge and trapped fermions. *Nuclear Physics B*, 936:239–305, 2018. MR3869768

[52] Pierre Le Doussal. Large deviations for the Kardar–Parisi–Zhang equation from the Kadomtsev–Petviashvili equation. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(4):043201, 2020. MR4148618

[53] Pierre Le Doussal, Satya N Majumdar, Alberto Rosso, and Grégory Schehr. Exact short-time height distribution in the one-dimensional Kardar-Parisi-Zhang equation and edge fermions at high temperature. *Physical review letters*, 117(7):070403, 2016.

[54] Pierre Le Doussal, Satya N Majumdar, and Grégory Schehr. Large deviations for the height in 1D Kardar-Parisi-Zhang growth at late times. *Europhysics Letters*, 113(6):60004, 2016.

[55] Yier Lin. Lyapunov exponents of the half-line SHE. *J. Stat. Phys.*, 183(3):Paper No. 37, 34, 2021. MR4261708

[56] Yier Lin and Li-Cheng Tsai. Short time large deviations of the KPZ equation. *Comm. Math. Phys.*, 386(1):359–393, 2021. MR4287189

[57] Yier Lin and Li-Cheng Tsai. A lower-tail limit in the weak noise theory. *arXiv preprint arXiv:2210.05629*, 2022. MR4630784

[58] Yier Lin and Li-Cheng Tsai. Spacetime limit shapes of the KPZ equation in the upper tails. *arXiv preprint arXiv:2304.14380*, 2023. MR4556284

[59] Baruch Meerson, Eytan Katzav, and Arkady Vilenkin. Large deviations of surface height in the Kardar-Parisi-Zhang equation. *Physical review letters*, 116(7):070601, 2016. MR3582125

[60] Baruch Meerson and Johannes Schmidt. Height distribution tails in the Kardar–Parisi–Zhang equation with Brownian initial conditions. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(10):103207, 2017. MR3722588

[61] Baruch Meerson and Arkady Vilenkin. Large fluctuations of a Kardar-Parisi-Zhang interface on a half line. *Physical Review E*, 98(3):032145, 2018.

[62] Carl Mueller. On the support of solutions to the heat equation with noise. *Stochastics: An International Journal of Probability and Stochastic Processes*, 37(4):225–245, 1991. MR1149348

[63] Mihai Nica. Intermediate disorder limits for multi-layer semi-discrete directed polymers. *Electron. J. Probab.*, 26:Paper No. 62, 50, 2021. MR4254804

[64] Jeremy Quastel. Introduction to KPZ. *Current developments in mathematics*, 2011(1), 2011. MR3098078

[65] Pavel Sasorov, Baruch Meerson, and Sylvain Prolhac. Large deviations of surface height in the $1+1$-dimensional Kardar–Parisi–Zhang equation: exact long-time results for $\lambda H < 0$. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(6):063203, 2017. MR3673439

[66] Naftali R Smith and Baruch Meerson. Exact short-time height distribution for the flat Kardar-Parisi-Zhang interface. *Physical Review E*, 97(5):052110, 2018.

[67] Naftali R Smith, Baruch Meerson, and Pavel Sasorov. Finite-size effects in the short-time height distribution of the Kardar–Parisi–Zhang equation. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(2):023202, 2018. MR3772417

[68] Naftali R Smith, Baruch Meerson, and Arkady Vilenkin. Time-averaged height distribution of the Kardar–Parisi–Zhang interface. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(5):053207, 2019. MR3998606

[69] Li-Cheng Tsai. Exact lower-tail large deviations of the KPZ equation. *Duke Math. J.*, 171(9):1879–1922, 2022. MR4484218

[70] Li-Cheng Tsai. Integrability in the weak noise theory. *arXiv preprint arXiv:2204.00614*, 2022. MR4630784

[71] Li-Cheng Tsai. High moments of the SHE in the clustering regimes. *arXiv preprint arXiv:2304.14375*, 2023. MR4799315

Acknowledgments. We thank Ivan Corwin, Greg Lawler, Russel Lyons and Li-Cheng Tsai for the helpful discussion. We thank Guillaume Barraquand for telling us a rigorous approach for proving the exact formula (2.1).