



Contents lists available at ScienceDirect

Journal of Functional Analysis

journal homepage: [www.elsevier.com/locate/jfa](http://www.elsevier.com/locate/jfa)

Regular Article

## High moments of the SHE in the clustering regimes

Li-Cheng Tsai

Department of Mathematics, University of Utah, United States of America



## ARTICLE INFO

## Article history:

Received 9 May 2023

Accepted 2 September 2024

Available online 16 September 2024

Communicated by Ivan Corwin

## Keywords:

Moment Lyapunov exponent

Large deviation principle

Interacting diffusion

Stochastic heat equation

## ABSTRACT

We analyze the high moments of the Stochastic Heat Equation (SHE) via a transformation to the attractive Brownian Particles (BPs), which are Brownian motions interacting via pairwise attractive drift. In those scaling regimes where the particles tend to cluster, we prove a Large Deviation Principle (LDP) for the empirical measure of the attractive BPs. Under the delta(-like) initial condition, we characterize the unique minimizer of the rate function and relate the minimizer to the spacetime limit shapes of the Kardar–Parisi–Zhang (KPZ) equation in the upper tails. The results of this paper are used in the companion paper [75] to prove an  $n$ -point, upper-tail LDP for the KPZ equation and to characterize the corresponding spacetime limit shape.

© 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

## 1. Introduction

This paper is motivated by the study of the large deviations and spacetime limit shapes of the Kardar–Parisi–Zhang (KPZ) equation. Introduced in [64], the KPZ equation

$$\partial_t h = \frac{1}{2} \partial_{xx} h + \frac{1}{2} (\partial_x h)^2 + \eta, \quad h = h(t, x), \quad (t, x) \in (0, \infty) \times \mathbb{R} \quad (1.1)$$

---

E-mail address: [lctsai.math@gmail.com](mailto:lctsai.math@gmail.com).

describes the evolution of a randomly growing interface, where  $\eta = \eta(t, x)$  denotes the spacetime white noise. This equation plays a central role in nonequilibrium statistical mechanics and has been widely studied in mathematics and physics; we refer to [82,19,81,21,20] for reviews on the mathematical literature related to the KPZ equation.

This paper and the companion paper [75] use the moments of the Stochastic Heat Equation (SHE) to obtain a Large Deviation Principle (LDP) and the corresponding spacetime limit shape of the KPZ equation. Recall that the SHE

$$\partial_t Z = \frac{1}{2} \partial_{xx} Z + \eta Z \quad (1.2)$$

gives the solution of (1.1) via  $h := \log Z$ , so the moment generating function of the KPZ equation is related to the moments of the SHE. This relation serves as a gateway to studying the KPZ equation, since the moments of the SHE are accessible from a number of tools, including the delta Bose gas and Feynman–Kac formula. At the one-point level, this relation has led to fruitful results on the one-point upper tail of the KPZ equation [16,11,26,22,71,32].

The goal of this paper is to characterize the *multi-point* moment Lyapunov exponents

$$\lim_{N \rightarrow \infty} \frac{1}{N^3 T_N} \log \mathbf{E} \left[ \prod_{\mathbf{c}=1}^n Z(T\mathbf{t}, NT\mathbf{x}_{\mathbf{c}})^{Nm_{\mathbf{c}}} \right] \quad (1.3)$$

for all positive powers,  $Nm_{\mathbf{c}} \in (0, \infty)$ , and for fixed  $\mathbf{x}_1 < \dots < \mathbf{x}_n \in \mathbb{R}$ . Hereafter  $T = T_N$  is the scale of time, and  $N$  is the scale of the powers. The only conditions we impose on  $N$  and  $T = T_N$  are

$$N \rightarrow \infty, \quad N^2 T = N^2 T_N \rightarrow \infty. \quad (1.4)$$

Our results actually hold for any  $T = T_{N,A} := A/N^2$  with  $(N, A) \rightarrow (\infty, \infty)$ ; we take  $T = T_N$  to depend only on  $N$  for the convenience of notation. Note that (1.4) allows  $T_N \rightarrow 0$ ,  $T_N \rightarrow 1$ , and  $T_N \rightarrow \infty$ . The second condition in (1.4) underscores what we call the clustering behaviors of the attractive Brownian Particles, as explained in the paragraph after next. The first condition in (1.4) reduces the task of characterizing (1.3) for positive powers,  $Nm_{\mathbf{c}} \in (0, \infty)$ , to that of integer powers. Indeed,  $Nm_{\mathbf{c}} \in \mathbb{Z}_{>0}$  translates into  $m_{\mathbf{c}} \in \frac{1}{N}\mathbb{Z}_{>0}$ , which becomes a denser and denser subset of  $(0, \infty)$  as  $N \rightarrow \infty$ .

Let us briefly describe our methods and results; the full description will be given in Section 2. To analyze the moment Lyapunov exponents, we express the integer moments of the SHE in terms of a system of attractive Brownian Particles (BPs), which are Brownian Motions (BMs) interacting via pairwise attractive drift. The task of analyzing the moment Lyapunov exponents turns into proving the LDP for the attractive BPs. Theorem 2.3 gives the sample-path LDP for the empirical measure of the attractive BPs. Next, we specialize the initial condition into the delta-like initial condition (defined

in (2.14)). Theorem 2.4 explicitly characterizes the unique minimizer of the rate function (of the LDP for the attractive BPs) under the delta-like initial condition. We call the minimizer the **optimal deviation**. Corollary 2.5 gives the limit of (1.3) under the delta-like initial condition. Finally, in Theorem 2.6, we show how the moment Lyapunov exponents and optimal deviation are related to the corresponding rate function and limit shape of the KPZ equation. The results of this paper are used in the companion paper [75] to prove an  $n$ -point, upper-tail LDP for the KPZ equation and to characterize the corresponding spacetime limit shape.

Under (1.4), the attractive BPs exhibit what we call the **clustering behavior**. Two effects contribute to the evolution of the attractive BPs: the attractive drift and diffusive effect. The drift pulls the particles together, while the diffusive effect spreads them out. As will be explained in Section 2.2, under the second condition in (1.4), the attractive drift dominates the diffusive effect, so the BPs tend to cluster. Theorem 2.4 (also Theorem 2.6(b)) shows that, under the delta-like initial condition, the optimal deviation consists of a number of clusters and explicitly describes the spacetime trajectories of the clusters. In the context of the upper-tail LDPs for the KPZ equation, the clustering behavior (of the attractive BPs) corresponds to what is called the noise-corridor effect in [75, Sect. 1.3]. Those noise corridors are exactly the trajectories of the clusters.

Our proof does not rely on integrability or explicit formulas. Thanks to the integrability of the delta Bose gas, moments of the SHE enjoy explicit formulas. The formulas offer a potential path to obtain the limit in (1.3), but they do not seem to provide information for proving a localization result like Corollary 2.5(b), which is a crucial technical input for the proof in [75]. Still, it is interesting to see if one can obtain the limit in (1.3) also from the formulas.

Let us compare (1.4) with two commonly-considered scaling regimes. First,  $T_N = A/N^2$ , with  $A < \infty$  fixed and  $N \rightarrow \infty$ , corresponds to the Freidlin–Wentzell/weak-noise LDP for the KPZ equation. At the level of the attractive BPs, this is the diffusive regime, and the LDP is proven in [24] for a general class of rank-based diffusions that includes the attractive BPs. The LDP in [24] and the LDP proven here are very different in nature, as will be explained in Section 2.2. Next,  $N = 1$  and  $T \rightarrow \infty$  correspond to the hyperbolic scaling regime, in long time, of the KPZ equation, namely  $h_T(t, x) := T^{-1}h(Tt, Tx)$ . This is perhaps the most natural scaling regime in long time, while (1.4) allows us to probe any deviation much larger than those in the hyperbolic regime. For the hyperbolic scaling regime, since  $N^2T = T \rightarrow \infty$ , we also expect the clustering behaviors to happen. One may seek to generalize our approach to obtain the multipoint, positive-integer moment Lyapunov exponents in this regime. This has been achieved by [72], based on exact formulas as well as ideas about the optimal clusters from Section 2.3 of this paper.

We end the introduction with a brief literature review. The moments of the SHE and its variant have been used to study the intermittency property [34, 31, 47], large deviations, and the density function of the SHE (and its variants) in [18, 35, 4, 16, 9, 52, 45, 37, 11, 17, 26, 22, 71]. The connection between the delta Bose gas and attractive BPs has been used in the physics work [68] to study their relaxation properties and their hy-

drodynamic limits. The attractive BPs is a special case of diffusions with rank-based drifts, or rank-based diffusions for short. We refer to [28,6,80,29,42,5,87,43,83,24,25,44,84,85,93,94,10,23,86,2,3] and the references therein for the literature on rank-based diffusions. Recently, there has been much interest in the LDPs for the KPZ equation in mathematics and physics. Several strands of methods produce detailed information on the one-point tail probabilities and the one-point rate function. This includes the physics works [69,70,54,90,13,55,56,61,65,57,67], the simulation works [38,39,36,40], and the mathematics works [11,12,8,26,22,48,71,7,32,30,95]. For the Freidlin–Wentzell/weak-noise LDP, behaviors of the one-point rate function and the corresponding most probable shape(s) for various initial conditions and boundary conditions have been predicted [49–51,76,62,77,79,89,91,1,92], some of which recently proven [73,74,33]; an intriguing symmetry breaking and second-order phase transition has been discovered in [46,88] via numerical means and analytically derived in [54,59]; a connection to integrable PDEs is recently established and studied in the physics works [66,58–60] and the mathematically rigorous work [96].

### *Outline*

In Section 2, we state the results and present some discussions. In Section 3, we introduce some notation, definitions, and tools. Sections 4–6 make up the proof of Theorem 2.3, the LDP for the attractive BPs: We establish properties of the rate function in Section 4, prove the LDP upper bound in Section 5, and prove the LDP lower bound in Section 6. In Section 7, we specialize the setting into the delta-like initial condition and establish results on the moment Lyapunov exponents, stated as Theorem 2.4 and Corollary 2.5. Finally, in Section 8, we prove Theorem 2.6 that relates the moment Lyapunov exponent and optimal deviation to the rate function and spacetime limit shape of the KPZ equation.

### *Acknowledgments*

I thank Ivan Corwin and Jeremy Quastel for showing me the derivation described in Remark 2.1 and thank Ivan Corwin for suggesting the possibility of extracting tail information from the exact formulas of the integer moments of the SHE. The knowledge has inspired me to pursue this work. I thank Jeremy Quastel for many useful discussions about the large deviations of KPZ-related models and thank Yau-Yuan Mao for the help with some of the figures in this paper. I thank the anonymous reviewers for their careful reading of and useful comments on the manuscript. The research of Tsai was partially supported by the NSF through DMS-2243112 and the Alfred P. Sloan Foundation through the Sloan Research Fellowship FG-2022-19308.

## 2. Results and discussions

### 2.1. Moment Lyapunov exponents, attractive BPs

The first goal of this paper is to prove that the  $n$ -point Lyapunov exponent (1.3) exists and to characterize it, for suitable initial conditions for  $Z$ . In (1.3) and hereafter,  $t \in (0, \infty)$ ,  $\mathbf{x}_1 < \dots < \mathbf{x}_n \in \mathbb{R}$ , and  $\mathbf{m}_1, \dots, \mathbf{m}_n \in (0, \infty)$ . Throughout this paper we will only work with integer moments, so an integer part is implicitly taken whenever needed; for example,  $N\mathbf{m}_\epsilon := \lceil N\mathbf{m}_\epsilon \rceil$  in (1.3). As will be shown in Theorem 2.6(a), the limit of (1.3) is continuous in  $\vec{\mathbf{m}} \in [0, \infty)^n$ . Hence, once we obtain the limit for  $\lceil N\vec{\mathbf{m}} \rceil \in (\mathbb{Z}_{>0})^n$ , the result automatically extends to  $N\vec{\mathbf{m}} \in [0, \infty)^n$ .

To access (1.3), let us express the moments in terms of the attractive BPs. First, by [41, Theorem 5.3], the one-point moment of  $Z$  can be expressed as an expectation over independent BMs. The same proof there works for multipoint (in space) moments, giving, for  $y_1, \dots, y_{N\mathbf{m}} \in \mathbb{R}$ ,

$$\mathbb{E} \left[ \prod_{i=1}^{N\mathbf{m}} Z(t, y_i) \right] = \mathbb{E}_{\text{BM}} \left[ e^{\sum_{i < j} \int_0^t ds \delta_0(X_i(t-s) - X_j(t-s))} \prod_{i=1}^{N\mathbf{m}} Z(0, X_i(t)) \right], \quad (2.1)$$

where  $X_1(s) - y_1, \dots, X_{N\mathbf{m}}(s) - y_{N\mathbf{m}}$  are independent standard BMs under  $\mathbb{E}_{\text{BM}}$ , and the integrals are interpreted as localtimes. Hereafter,  $\mathbf{m} \in (0, \infty)$  is fixed and is taken to be  $\mathbf{m} = \mathbf{m}_1 + \dots + \mathbf{m}_n$  whenever we analyze (1.3). The localtimes in (2.1) can be removed by transforming  $\mathbb{E}_{\text{BM}}$  into a different law  $\mathbb{E}$ . Let  $\mathbb{P}$  and  $\mathbb{E}$  denote the law and expectation under which  $X_1, \dots, X_{N\mathbf{m}}$  evolve as a system of attractive BPs as

$$dX_i(s) = \sum_{j=1}^{N\mathbf{m}} \frac{1}{2} \text{sgn}(X_j(s) - X_i(s)) ds + dB_i(s), \quad X_i(0) = y_i, \quad (2.2)$$

where  $B_1, \dots, B_{N\mathbf{m}}$  are independent standard BMs, and  $\text{sgn}(x) := (x/|x|)\mathbf{1}_{\{x \neq 0\}}$ . We index the particles so that they are ordered at the start:  $X_1(0) \leq X_2(0) \leq \dots$ , and the particles can exchange orders as time evolves. By [43], the equations (2.2) have a unique strong (pathwise) solution. Start from  $\mathbb{P}_{\text{BM}}$  (under which  $X_1, \dots, X_{N\mathbf{m}}$  are independent BMs), apply Tanaka's formula to  $X_i - X_j$ , sum the result over  $i < j \in \{1, \dots, N\mathbf{m}\}$ , and exponentiate the result. Doing so gives

$$\text{under } \mathbb{P}_{\text{BM}}, \quad e^{\sum_{i < j} \int_0^t ds \delta_0(X_i - X_j)} = \exp \left( - \sum_{i < j} \int_0^t d(X_i - X_j) \frac{1}{2} \text{sgn}(X_i - X_j) \right) \quad (2.3a)$$

$$\cdot \exp \left( - \sum_{i < j} \frac{1}{2} |X_i(0) - X_j(0)| + \sum_{i < j} \frac{1}{2} |X_i(t) - X_j(t)| \right). \quad (2.3b)$$

Rewrite the sum in (2.3a) as  $-\frac{1}{2} \sum_{i=1}^{Nm} \int_0^t dX_i \sum_{j:j \neq i} \text{sgn}(X_i - X_j)$ . As a process in  $t$ , this term has quadratic variation  $tNm((Nm)^2 - 1)/12$ ; see [16, Eq. (2.1)]. Given this property, Girsanov's theorem asserts that

$$\frac{d\mathbb{P}_{\text{BM}}}{d\mathbb{P}}(\vec{X}) = (\text{right hand side of (2.3a)}) \cdot e^{Nm((Nm)^2 - 1)\frac{t}{24}}. \quad (2.4)$$

Let  $\psi_{\text{gd}}(\vec{x}) := \exp(-\frac{1}{2} \sum_{i < j} |x_i - x_j|) = \exp(-\frac{1}{4} \sum_{i,j=1}^{Nm} |x_i - x_j|)$ . Combining (2.3)–(2.4) gives

$$\mathbb{E}_{\text{BM}}[e^{\sum_{i < j} \int_0^t ds \delta_0(X_i(t-s) - X_j(t-s))}(\cdot)] = e^{Nm((Nm)^2 - 1)\frac{t}{24}} \mathbb{E}[\psi_{\text{gd}}(\vec{X}(0))(\cdot) / \psi_{\text{gd}}(\vec{X}(t))]. \quad (2.5)$$

We view both sides of (2.5) as operators that act on a measurable function  $F(\vec{X})$  of  $(X_i(\cdot))_{i=1,\dots,Nm}$  such that  $\psi_{\text{gd}}(\vec{X}(0))F(\vec{X})/\psi_{\text{gd}}(\vec{X}(t))$  is integrable under  $\mathbb{E}$ .

**Remark 2.1.** A similar application of Tanaka's formula was used in [16] to access the one-point integer moments of the SHE. The transformation (2.5) can also be derived by conjugating the Hamiltonian of the delta Bose gas by its ground state. This is done in [68, Eq. (5)–(6)] at a physics level of rigor.

To analyze the attractive BPs, we will mostly work with the (scaled) empirical measure

$$\mu_N(s) := \frac{1}{N} \sum_{i=1}^{Nm} \delta_{X_i(Ts)/(NT)}. \quad (2.6)$$

Hereafter, we write  $T = T_N$ , with the understanding that  $N \rightarrow \infty$  is always taken under (1.4). Writing  $\langle \lambda, f \rangle := \int_{\mathbb{R}} \lambda(dx) f(x)$  for the action of a Borel measure  $\lambda$  on  $f$ , we rewrite (2.5) as

$$\mathbb{E}_{\text{BM}}[e^{\sum_{i < j} \int_0^{Tt} ds \delta_0(X_i(Tt-s) - X_j(Tt-s))}(\cdot)] \quad (2.7a)$$

$$= \exp(N^3 T m^3 (1 - \frac{1}{(Nm)^2}) \frac{t}{24}) \mathbb{E}\left[\exp\left(N^3 T \langle \mu_N(s)^{\otimes 2}, \frac{1}{4} |x - x'| \rangle \Big|_{s=0}^{s=t}\right)(\cdot)\right], \quad (2.7b)$$

where  $\lambda^{\otimes 2} := \lambda \otimes \lambda$  stands for the product measure that acts on  $\mathbb{R}^2 = \{(x, x')\}$ .

**Convention 2.2.** We have and will continue to use  $t$  to denote the time variable of the SHE, and use  $s$  to denote the time variable of  $X_1, X_2, \dots$ . We call  $Z(t, \cdot)|_{t=0}$  and  $Z(t, \cdot)|_{t=Nt}$  respectively the **initial** and **terminal conditions** of the SHE, and call  $X_i(s)|_{s=0}$  and  $X_i(s)|_{s=Nt}$  respectively the **starting** and **ending conditions** of  $X_i$ . Since  $t = Nt - s$ , there is a *time reversal* between  $t$  and  $s$ .

## 2.2. Result: LDP for the attractive BPs

The task of analyzing the moment Lyapunov exponent (1.3) turns into studying the large deviations of the attractive BPs. Indeed, the quantities in (2.7) are of exponential scales, so the expectation is controlled by the large deviations of the attractive BPs.

Our result, Theorem 2.3, gives the sample-path LDP for  $\mu_N$ .

Let us prepare some notation for Theorem 2.3. Let  $(\mathbf{m}\mathcal{P}(\mathbb{R}))$  denote the space of positive Borel measures on  $\mathbb{R}$  with total mass  $\mathbf{m}$ , and endow the space with the weak\* topology, namely the topology induced by convergence in distribution, which is denoted by  $\Rightarrow$ . Consider  $\mathcal{C}([0, \mathfrak{t}], \mathbf{m}\mathcal{P}(\mathbb{R}))$ . A sequence  $\mu_k$  converging to  $\mu$  in this space means  $\sup_{s \in [0, \mathfrak{t}]} |\langle \mu_k(s) - \mu(s), \varphi \rangle| \rightarrow 0$  for any  $\varphi \in \mathcal{C}_b(\mathbb{R})$ . Endow  $\mathcal{C}([0, \mathfrak{t}], \mathbf{m}\mathcal{P}(\mathbb{R}))$  with the topology induced by the convergence sequences. Namely, a set  $A$  in  $\mathcal{C}([0, \mathfrak{t}], \mathbf{m}\mathcal{P}(\mathbb{R}))$  is closed if and only if  $A$  contains the limit of any sequence  $\{\mu_k\}_{k=1,2,\dots} \subset A$  that converges. In Section 3.2, we will introduce metrics for these topologies. Given a topological space  $\mathcal{X}$ , a sequence of  $\mathcal{X}$ -valued random variables  $R_N$  **satisfying the LDP with the speed  $s_N$  and rate function  $I$**  means that  $-\inf_O I \leq \liminf_{N \rightarrow \infty} \frac{1}{s_N} \log \mathbb{P}[R_N \in O]$  and  $\limsup_{N \rightarrow \infty} \frac{1}{s_N} \log \mathbb{P}[R_N \in C] \leq -\inf_C I$  for every open  $O \subset \mathcal{X}$  and closed  $C \subset \mathcal{X}$ . A function  $I : \mathcal{X} \rightarrow [0, \infty]$  being a **good rate function** means that its is lower-semicontinuous and  $\{I \leq r\}$  is compact for any  $r < \infty$ .

More notation for Theorem 2.3 is in place. For  $\lambda \in \mathbf{m}\mathcal{P}(\mathbb{R})$ , set

$$\Phi[\lambda](x) := \langle \lambda, \tfrac{1}{2} \operatorname{sgn}(\cdot - x) \rangle = -\tfrac{1}{2} \langle \lambda, \mathbf{1}_{(-\infty, x)} \rangle + \tfrac{1}{2} \langle \lambda, \mathbf{1}_{(x, \infty)} \rangle. \quad (2.8)$$

Recall that  $\operatorname{sgn}(0) := 0$ , so the right hand side excludes the mass of  $\lambda$  at  $x$ . This function gives the amount of drift a particle feels in the system of attractive BPs. More explicitly, the drift in (2.2) is equal to  $N\Phi[\lambda](x)$  with  $\lambda = \mu_N(s)$  and  $x = X_i(Ts)/(NT)$ . Next, let  $\mathcal{C}_b^{1,1}([0, \mathfrak{t}], \mathbb{R})$  be the space of functions  $\mathbf{h} = \mathbf{h}(s, x)$  with continuous and bounded first derivatives in time and space. Let

$$\Lambda(\mu, \mathbf{h}) := \langle \mu(s), \mathbf{h}(s) \rangle \Big|_{s=0}^{s=\mathfrak{t}} - \int_0^{\mathfrak{t}} ds \langle \mu, \partial_s \mathbf{h} \rangle - \int_0^{\mathfrak{t}} ds \langle \mu, \Phi[\mu] \partial_x \mathbf{h} \rangle, \quad (2.9)$$

$$\mathcal{I}(\mu) := \sup \left\{ \Lambda(\mu, \mathbf{h}) - \int_0^{\mathfrak{t}} ds \frac{1}{2} \langle \mu, (\partial_x \mathbf{h})^2 \rangle : \mathbf{h} \in \mathcal{C}_b^{1,1}([0, \mathfrak{t}], \mathbb{R}) \right\}. \quad (2.10)$$

Hereafter, we adopt shorthand notation such as  $\mathbf{h}(s) := \mathbf{h}(s, \cdot)$ ,  $\langle \mu, \partial_s \mathbf{h} \rangle := \langle \mu(s), \partial_s \mathbf{h}(s) \rangle$ , and  $\Phi[\mu] := \Phi[\mu(s)](x)$ . Expressions like (2.10) arise in the martingale method [63] for proving LDP upper bounds. To prove the LDP lower bound and to analyze  $\mathcal{I}$ , it will be convenient to have an expression more explicit than (2.10). Given any  $\lambda \in \mathbf{m}\mathcal{P}(\mathbb{R})$ , let

$$F[\lambda](x) := \langle \lambda, \mathbf{1}_{(-\infty, x]} \rangle, \quad \mathfrak{X}[\lambda](a) := \inf \{x \in \mathbb{R} : a \leq F[\lambda](x)\}, \quad (2.11)$$

be the Cumulative Distribution Function (CDF) and quantile function, aka inverse CDF. We refer to  $a \in [0, \mathfrak{m}]$  as the **quantile coordinate**. As will be shown in Section 4.1, the rate function  $\mathcal{I}$  permits a **quantile representation**  $\mathbb{I}$ . Namely,  $\mathcal{I} = \mathbb{I}$ , where

$$\mathbb{I}(\mu) := \int_0^{\mathfrak{t}} ds \int_0^{\mathfrak{m}} da \frac{1}{2} (\partial_s(\mathfrak{X}[\mu]) - \Phi[\mu](\mathfrak{X}[\mu]))^2, \quad (2.12)$$

where  $\mathfrak{X}[\mu] := \mathfrak{X}[\mu(s)](a)$ . Given that  $|\Phi[\mu]| \leq \mathfrak{m}/2$ , when  $\partial_s(\mathfrak{X}[\mu]) \notin \mathcal{L}^2([0, \mathfrak{t}] \times [0, \mathfrak{m}])$ , we define the integral in (2.12) to be  $+\infty$ .

**Theorem 2.3.** Fix any  $\mathfrak{t}, \mathfrak{m} \in (0, \infty)$  and a starting condition  $\mu_{\text{sc}} \in \mathfrak{m}\mathcal{P}(\mathbb{R})$ . Assume (1.4).

Started from a deterministic  $\mu_N(0)$  such that  $\mu_N(0) \Rightarrow \mu_{\text{sc}}$ , the empirical measure  $\mu_N$  satisfies the LDP on  $\mathcal{C}([0, \mathfrak{t}], \mathfrak{m}\mathcal{P}(\mathbb{R}))$  with the speed  $N^3T = N^3T_N$  and the rate function

$$\mathcal{I}_\star(\mu) := \begin{cases} \mathcal{I}(\mu) = \mathbb{I}(\mu) & \text{when } \mu(0) = \mu_{\text{sc}}, \\ +\infty & \text{when } \mu(0) \neq \mu_{\text{sc}}, \end{cases} \quad (2.13)$$

and  $\mathcal{I}_\star$  is a good rate function.

We will refer to  $\mathcal{I}_\star$ ,  $\mathcal{I}$ , and  $\mathbb{I}$  all as rate functions; as shown in Sections 4, they are lower semicontinuous, nonnegative, and not identically infinite.

As mentioned previously, the condition  $N^2T \rightarrow \infty$  underscores the clustering behaviors. In (2.2), the drift pulls the particles together, while the diffusive effect spreads them out. For  $O(N)$  particles evolving over time  $O(T)$ , the contribution of the drift is  $O(NT)$ , while the contribution of the diffusive effect is  $O(\sqrt{T})$ . The condition  $N^2T \rightarrow \infty$  amounts to saying that the drift dominates the diffusive effect. Under this condition, the attractive BPs tend to cluster, so we refer to the regimes with  $N^2T \rightarrow \infty$  as the **clustering regimes**. In these regimes, deviations  $\mu \in \mathcal{C}([0, \mathfrak{t}], \mathfrak{m}\mathcal{P}(\mathbb{R}))$  with *atoms* are relevant.

We emphasize that Theorem 2.3 differs from the LDP proven in [24]. The work [24] considered a general class of rank-based diffusions and proved their LDPs under the scaling regime that corresponds to  $N \rightarrow \infty$  and  $TN^2 = A < \infty$ , which we refer to as the **diffusive regime**. In the diffusive regime, the drift and diffusive effect compete at the same footing. We hence do not expect deviations with atoms (that persist over time) to be relevant. This distinction has a significant implication in the proof, which is explained in the next paragraph.

Let us describe the major challenge in proving Theorem 2.3. The challenge stems from the discontinuity of the drift, which is common in the study of rank-based diffusions. With the presence of atomic deviations, however, the issue of discontinuity becomes *much more severe*. To see why, take any deviation  $\mu \in \mathcal{C}([0, \mathfrak{t}], \mathfrak{m}\mathcal{P})$ . At the macroscopic level, when the attractive BPs follow  $\mu$ , the drift should approximate  $\Phi[\mu(s)](x)$ . For a



non-atomic  $\mu(s)$ , the macroscopic drift  $\Phi[\mu(s)](x)$  is continuous in  $x$ , but for an atomic  $\mu(s)$ , the macroscopic drift is not continuous: for example  $\Phi[\delta_0](x) = (\mathbf{1}_{(-\infty, 0)}(x) - \mathbf{1}_{(0, \infty)}(x))/2$ . Put differently, in the diffusive regime, the effect of the discontinuity in the drift diminishes in the  $N \rightarrow \infty$  limit, while in the clustering regimes, the effect of the discontinuity *persists*. This issue renders many commonly-used tools non-applicable and manifests itself in the proof of the LDP lower bound, where we (have to) use a cluster approximation with the help of the quantile representation (2.12) of the rate function.

### 2.3. Result: applications to the moment Lyapunov exponents [section rewritten]

The results in Sections 2.3–2.4 are for a particular initial condition, which we now define. For  $\mathbf{x}' \in \mathbb{R}$  and a small  $\alpha > 0$ , let  $Z_\alpha = Z_\alpha(t, x)$  be the solution of the SHE with the **delta-like initial condition**

$$Z_\alpha(0, NT \cdot) := \mathbf{1}_{[-\alpha + \mathbf{x}', \mathbf{x}' + \alpha]}. \quad (2.14)$$

We will send  $N \rightarrow \infty$  first and  $\alpha \rightarrow 0$  later. The solution of the SHE with this initial condition, scaled in time by  $T$  and in space by  $NT$ , can be written by the Feynman–Kac formula as

$$Z_{N, \alpha}(t, x) := Z_\alpha(Tt, NTx) := \mathbb{E}_{\text{BM}} \left[ e^{\int_0^{Tt} ds \, \eta(Tt-s, X(s))} \mathbf{1}_{[-\alpha + \mathbf{x}', \mathbf{x}' + \alpha]} \left( \frac{1}{NT} X(Tt) \right) \right], \quad (2.15)$$

where  $X = (\text{standard BM}) + NTx$ . For the purpose of studying the LDPs and limit shapes of the KPZ equation, the delta-like initial condition approximates the true delta initial condition  $\delta_{\mathbf{x}'}$  sufficiently well; this is seen in the analysis of the companion paper [75].

We now explain how to obtain the multipoint moment Lyapunov exponents of  $Z_{N, \alpha}$ . To set up the notation, fix any  $\mathbf{x}_1 < \dots < \mathbf{x}_n \in \mathbb{R}$ ,  $\mathbf{x}' \in \mathbb{R}$ , let  $\vec{\mathbf{m}} = (\mathbf{m}_1, \dots, \mathbf{m}_n) \in [0, \infty)^n$  and  $\mathbf{m} = \mathbf{m}_1 + \dots + \mathbf{m}_n$  be as before, and let

$$L_{\text{SHE}}(\mathbf{x}' \xrightarrow{t} (\vec{\mathbf{x}}, \vec{\mathbf{m}})) := \sup \left\{ \frac{t\mathbf{m}^3}{24} + \langle \mu(s)^{\otimes 2}, \frac{1}{4}|x - x'| \rangle \Big|_{s=0}^{s=t} - \mathbb{I}(\mu) : \right. \quad (2.16a)$$

$$\left. \mu \in \mathcal{C}([0, t], \mathbf{m} \mathcal{P}(\mathbb{R})), \mu(0) = \sum_{\mathbf{c}=1}^n \mathbf{m}_{\mathbf{c}} \delta_{\mathbf{x}_{\mathbf{c}}}, \mu(t) = \mathbf{m} \delta_{\mathbf{x}'} \right\}. \quad (2.16b)$$

Let  $\mathbf{c}_N(i) := \min\{\mathbf{c} : \mathbf{m}_1 + \dots + \mathbf{m}_{\mathbf{c}} \geq i/N\}$  and set  $y_i = \mathbf{x}_{\mathbf{c}_N(i)}$ . This way,  $\mathbf{E}[\prod_{i=1}^{N\mathbf{m}} Z(Tt, NTy_i)] = \mathbf{E}[\prod_{\mathbf{c}=1}^n Z(Tt, NT\mathbf{x}_{\mathbf{c}})^{N\mathbf{m}_{\mathbf{c}}}]$ . Specialize to the delta-like initial condition (2.14), use (2.1) for  $t = Tt$  and  $y_i \mapsto NTy_i$ , use (2.7), and apply  $\frac{1}{N^3 T} \log(\cdot)$  to the result. Doing so gives

$$\begin{aligned} \frac{1}{N^3 T} \log \mathbf{E} \left[ \prod_{\mathbf{c}=1}^n Z_{N,\alpha}(\mathbf{t}, \mathbf{x}_{\mathbf{c}})^{\mathbf{m}_{\mathbf{c}}} \right] &= \frac{\mathbf{t} \mathbf{m}^3}{24} + \frac{1}{4} \sum_{\mathbf{c}, \mathbf{c}'=1}^n \mathbf{m}_{\mathbf{c}} \mathbf{m}_{\mathbf{c}'} |\mathbf{x}_{\mathbf{c}} - \mathbf{x}_{\mathbf{c}'}| + o_N(1) \\ &+ \frac{1}{N^3 T} \log \mathbb{P} \left[ X_i(0) = NT \mathbf{x}_{\mathbf{c}_N(i)}, |X_i(T\mathbf{t}) - NT \mathbf{x}'| \leq NT\alpha, \forall i \right], \end{aligned} \quad (2.17)$$

where  $o_N(1)$  denotes a term that converges to zero as  $N \rightarrow \infty$ . Rewrite (2.17) as

$$\frac{1}{N^3 T} \log \mathbf{E} \left[ \prod_{\mathbf{c}=1}^n Z_{N,\alpha}(\mathbf{t}, \mathbf{x}_{\mathbf{c}})^{\mathbf{m}_{\mathbf{c}}} \right] = \frac{\mathbf{t} \mathbf{m}^3}{24} + \langle \boldsymbol{\mu}_N(s)^{\otimes 2}, \frac{1}{4} |x - x'| \rangle \Big|_{s=0}^{s=\mathbf{t}} + o(1) \quad (2.18a)$$

$$+ \frac{1}{N^3 T} \log \mathbb{P} \left[ \boldsymbol{\mu}_N(0) = \sum_{\mathbf{c}=1}^n \mathbf{m}_{\mathbf{c}} \delta_{\mathbf{x}_{\mathbf{c}}}, \text{supp}(\boldsymbol{\mu}_N(\mathbf{t})) \subset [-\delta + \mathbf{x}', \mathbf{x}' + \delta] \right]. \quad (2.18b)$$

It now suffices to obtain the limit of (2.18b) under  $\lim_{\alpha \rightarrow 0} \lim_{N \rightarrow \infty}$ . This will be carried out in Section 7.2 with the aid of Theorem 2.3. The result gives

$$\limsup_{\alpha \rightarrow 0} \limsup_{N \rightarrow \infty} \left| (2.18b) + \inf \left\{ \mathbb{I}(\mu) : \mu(0) = \sum_{\mathbf{c}=1}^n \mathbf{m}_{\mathbf{c}} \delta_{\mathbf{x}_{\mathbf{c}}}; \mu(\mathbf{t}) = \delta_{\mathbf{x}'} \right\} \right| = 0. \quad (2.19)$$

Combining (2.18)–(2.19) gives the multipoint moment Lyapunov exponents of  $Z_{N,\alpha}$ :

$$\limsup_{\alpha \rightarrow 0} \limsup_{N \rightarrow \infty} \left| \frac{1}{N^3 T} \log \mathbf{E} \left[ \prod_{\mathbf{c}=1}^n Z_{N,\alpha}(\mathbf{t}, \mathbf{x}_{\mathbf{c}})^{\mathbf{m}_{\mathbf{c}}} \right] - L_{\text{SHE}}(\mathbf{x}' \xrightarrow{\mathbf{t}} (\vec{\mathbf{x}}, \vec{\mathbf{m}})) \right| = 0. \quad (2.20)$$

For our subsequent analysis, the functional in the supremum in (2.16a) is not quite convenient, and we need an equivalent expression of it. Let

$$\mathbb{L}_{[s_1, s_2]}(\mu) := \int_{s_1}^{s_2} ds \left( \sum_x \frac{1}{24} \langle \mu, \mathbf{1}_{\{x\}} \rangle^3 - \int_0^{\mathbf{m}} da \frac{1}{2} (\partial_s \mathfrak{X}[\mu])^2 \right), \quad (2.21)$$

where the sum runs over atoms of  $\mu$ . In Appendix A, we prove that

$$\frac{\mathbf{t} \mathbf{m}^3}{24} + \langle \mu(s)^{\otimes 2}, \frac{1}{4} |x - x'| \rangle \Big|_{s=0}^{s=\mathbf{t}} - \mathbb{I}(\mu) = \mathbb{L}_{[0, \mathbf{t}]}(\mu). \quad (2.22)$$

Given this, let us rewrite (2.16) as

$$L_{\text{SHE}}(\mathbf{x}' \xrightarrow{\mathbf{t}} (\vec{\mathbf{x}}, \vec{\mathbf{m}})) := \sup \left\{ \mathbb{L}_{[0, \mathbf{t}]}(\mu) : \right. \quad (2.23a)$$

$$\left. \mu \in \mathcal{C}([0, \mathbf{t}], \mathbf{m} \mathcal{P}(\mathbb{R})), \mu(0) = \sum_{\mathbf{c}=1}^n \mathbf{m}_{\mathbf{c}} \delta_{\mathbf{x}_{\mathbf{c}}}, \mu(\mathbf{t}) = \mathbf{m} \delta_{\mathbf{x}'} \right\}, \quad (2.23b)$$

where  $\vec{\mathbf{m}} = (\mathbf{m}_1, \dots, \mathbf{m}_n)$  and  $\mathbf{m} := \mathbf{m}_1 + \dots + \mathbf{m}_n$ .

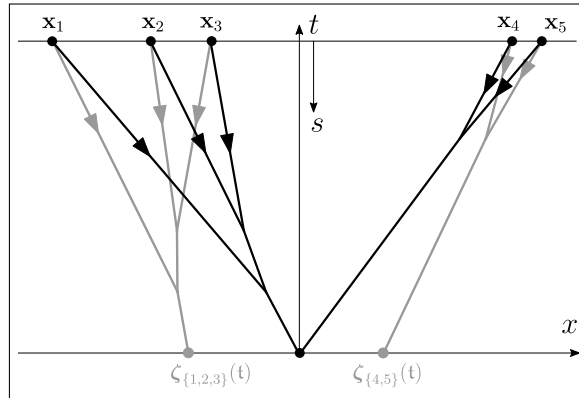


Fig. 1. The inertia clusters (gray) and optimal clusters (black). In this figure,  $\mathfrak{B} = \{\{1, 2, 3\}, \{4, 5\}\}$ , and  $\mathbf{x}' = 0$ .

Our result below, Theorem 2.4, gives the unique minimizer of (2.23). To state the result we need to construct a few objects. We begin by introducing what we call **inertia clusters**. The inertia clusters are point masses  $\mathbf{m}_1, \dots, \mathbf{m}_n$  that start at  $s = 0$  from  $\mathbf{x}_1 < \dots < \mathbf{x}_n$  with velocities  $\phi_c := \frac{1}{2}(-\mathbf{m}_1 \dots - \mathbf{m}_{c-1} + \mathbf{m}_{c+1} + \dots + \mathbf{m}_n)$ ,  $c = 1, \dots, n$ . They travel at constant velocities until they meet, and when they meet, they merge according to the conservation of momentum. For example, if  $\zeta_c$  and  $\zeta_{c+1}$  meet, they merged into a single cluster with mass  $(\mathbf{m}_c + \mathbf{m}_{c+1})$  and velocity  $(\mathbf{m}_c \phi_c + \mathbf{m}_{c+1} \phi_{c+1}) / (\mathbf{m}_c + \mathbf{m}_{c+1})$ . Let  $\zeta_c = \zeta_c(s) \in \mathcal{C}[0, t]$ ,  $c = 1, \dots, n$  denote the trajectories of the inertial clusters. Examine which inertia clusters have merged within  $(0, t)$  and lump the indices of those clusters together. Doing so gives a partition of  $\{1, \dots, n\}$  into intervals. We let  $\mathfrak{B}$  denote this partition, and call an element  $\mathfrak{b}$  in  $\mathfrak{B}$  a **branch**. Namely,  $c, c' \in \mathfrak{b}$  if and only if  $\zeta_c$  and  $\zeta_{c'}$  merged with  $s \in (0, t)$ . Note that branches depend on  $t$ . All the inertia clusters within a branch end up at the same position at  $s = t$ , namely  $\zeta_c(t) = \zeta_{c'}(t)$  for all  $c, c' \in \mathfrak{b}$ ; call this position  $\zeta_{\mathfrak{b}}(t)$ . In general,  $\zeta_{\mathfrak{b}}(t) \neq \mathbf{x}'$ . To *bring* the clusters to  $\mathbf{x}'$  at the ending time, apply a constant drift:

$$\xi_c(s) := \zeta_c(s) + v_{\mathfrak{b}} s, \quad v_{\mathfrak{b}} := (\mathbf{x}' - \zeta_{\mathfrak{b}}(t)) / t, \quad c \in \mathfrak{b}, \mathfrak{b} \in \mathfrak{B}. \quad (2.24)$$

The resulting deviation  $\xi = \sum_{c=1}^n \mathbf{m}_c \delta_{\xi_c}$  is the **optimal deviation** and we call  $\xi_1, \dots, \xi_n$  the **optimal clusters**. See Fig. 1. By construction,  $\xi_c$  and  $\xi_{c'}$  merge within  $(0, t)$  if and only if  $c, c' \in \mathfrak{b}$ .

**Theorem 2.4.** *The optimal deviation  $\xi$  is the unique minimizer of (2.23).*

For the analysis in the companion paper [75], the result (2.20) along does not suffice, and we need a “localized” version of it as well. To set up the notation, For  $f \in \mathcal{C}[0, Ts']$ , set

$$\text{Dist}_{N,s}(f, \xi) := \min_{c=1, \dots, n} \{|\frac{1}{NT} f(Ts) - \xi_c(s)|\}, \quad (2.25)$$

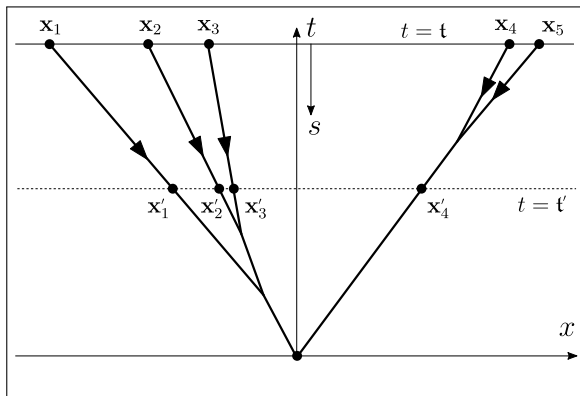


Fig. 2. Intermediate-time configuration. In this figure,  $\mathfrak{C}(1) = \{1\}$ ,  $\mathfrak{C}(2) = \{2\}$ ,  $\mathfrak{C}(3) = \{3\}$ ,  $\mathfrak{C}(4) = \{4, 5\}$ .

$$\text{Dist}_{N,[0,s']}(f, \xi) := \sup_{s \in [0,s']} \text{Dist}_{N,s}(f, \xi). \quad (2.26)$$

Define the localized version of  $Z_{N,\alpha}$ :

$$Z_{N,\alpha}^{\beta\text{-loc}}(t, x) := \mathbb{E}_{\text{BM}} \left[ e^{\int_0^{Tt} ds \, \eta(Tt-s, X(s))} \mathbf{1}_{[-\alpha + \mathbf{x}', \mathbf{x}' + \alpha]} \left( \frac{1}{NT} X(Tt) \right) \cdot \mathbf{1}_{\{\text{Dist}_{N,[0,t]}(X, \xi(\cdot + (t-t))) \leq \beta\}} \right], \quad (2.27)$$

where  $X = (\text{standard BM}) + NTx$ . The last indicator in (2.27) constrains the BM to stay close to  $\xi$ , in the post-scale units. By construction,  $Z_{N,\alpha}^{\beta\text{-loc}} \leq Z_{N,\alpha}$ . Below, Corollary 2.5(a) just restates (2.20), while Corollary 2.5(b) is the localized version of (2.20) that will be needed in [75].

**Corollary 2.5.**

- (a)  $\limsup_{\alpha \rightarrow 0} \limsup_{N \rightarrow \infty} \left| \frac{1}{N^3 T} \log \mathbf{E} \left[ \prod_{\mathfrak{c}=1}^n Z_{N,\alpha}(\mathfrak{t}, \mathbf{x}_{\mathfrak{c}})^{N\mathfrak{m}_{\mathfrak{c}}} \right] - L_{\text{SHE}}(\mathbf{x}' \xrightarrow{t} (\vec{\mathbf{x}}, \vec{\mathbf{m}})) \right| = 0.$   
 (b) For any nonempty  $A \subset \{1, \dots, n\}$  and  $\beta > 0$ ,

$$\limsup_{\alpha \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbf{E} \left[ \prod_{\mathfrak{c} \in A} (Z_{N,\alpha} - Z_{N,\alpha}^{\beta\text{-loc}})(\mathfrak{t}, \mathbf{x}_{\mathfrak{c}})^{N\mathfrak{m}_{\mathfrak{c}}} \cdot \prod_{\mathfrak{c} \notin A} Z_{N,\alpha}(\mathfrak{t}, \mathbf{x}_{\mathfrak{c}})^{N\mathfrak{m}_{\mathfrak{c}}} \right] < L_{\text{SHE}}(\mathbf{x}' \xrightarrow{t} (\vec{\mathbf{x}}, \vec{\mathbf{m}})). \quad (2.28)$$

As mentioned in Section 1, the analog of Corollary 2.5(a) in the hyperbolic scaling regime ( $T \rightarrow \infty$ ,  $N = 1$ ,  $\vec{\mathbf{m}} \in \mathbb{Z}_{>0}^n$ ) has been proven in [72].

The discussion that leads up to (2.20) holds for more general initial conditions. For example, one can consider the multi-delta-like initial condition

$$Z'_\alpha(0, NT \cdot) := e^{N^2 T \beta_1} \mathbf{1}_{[-\alpha + \mathbf{x}'_1, \mathbf{x}'_1 + \alpha]} + \dots + e^{N^2 T \beta_m} \mathbf{1}_{[-\alpha + \mathbf{x}'_m, \mathbf{x}'_m + \alpha]}, \quad (2.29)$$

for  $\mathbf{x}'_1 < \dots < \mathbf{x}'_m \in \mathbb{R}$ ,  $\beta_1, \dots, \beta_m \in \mathbb{R}$ , with  $N \rightarrow \infty$  first and  $\alpha \rightarrow 0$  later. On the other hand, the way we characterize the unique minimizer in Theorem 2.4 works only for the delta-like initial condition (2.14).

#### 2.4. Result: connection to $I_{\text{KPZ}}$ and the limit shape [section rewritten]

We begin by preparing some notation. Fix  $t \in (0, t]$  and  $\mathbf{x}_1 < \dots < \mathbf{x}_n$ , and let  $\mathbf{p}(t) = \mathbf{p}(t, x) := -x^2/(2t)$  and

$$\mathcal{R}(t, \vec{\mathbf{x}}) := \{\vec{\mathbf{r}} = (\mathbf{r}_c)_{c=1}^n : \mathbf{r}_c \geq \mathbf{p}(t, \mathbf{x}_c), \ c = 1, \dots, n\}. \quad (2.30)$$

Given any  $\vec{\mathbf{r}} \in \mathcal{R}(t, \vec{\mathbf{x}})$ , let  $\mathbf{f}_{\star, t, \vec{\mathbf{x}}, \vec{\mathbf{r}}} = \mathbf{f}_\star(x)$  be the piecewise  $\mathcal{C}^1$  function on  $\mathbb{R}$  characterized by the properties:  $\mathbf{f}_\star(\mathbf{x}_c) = \mathbf{r}_c$ , for all  $c$ ;  $\mathbf{f}_\star \geq \mathbf{p}(t)$ ;  $\mathbf{f}_\star(x) = \mathbf{p}(t, x)$  for all  $|x|$  large enough;  $\partial_x \mathbf{f}_\star$  is constant on  $\{\mathbf{f}_\star > \mathbf{p}(t)\} \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ ;  $\mathbf{f}_\star$  is  $\mathcal{C}^1$  except at  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . See Fig. 4 for an illustration. Set

$$I_{\text{KPZ}}(t, \vec{\mathbf{x}}, \vec{\mathbf{r}}) := \int_{\mathbb{R}} dx \left( \frac{1}{2} (\partial_x \mathbf{f}_{\star, t, \vec{\mathbf{x}}, \vec{\mathbf{r}}})^2 - \frac{1}{2} (\partial_x \mathbf{p}(t))^2 \right), \quad (2.31)$$

$$\mathbf{h}_\star(t, x) := \inf \left\{ \frac{(x - y)^2}{2(t - t)} + \mathbf{f}_\star(y) : y \in \mathbb{R} \right\}, \quad (t, x) \in (0, 1] \times \mathbb{R}. \quad (2.32)$$

See Fig. 3 for an illustration of  $\mathbf{h}_\star$ . The space  $\mathcal{R}(t, \vec{\mathbf{x}})$  is the space of deviations of the  $n$ -point upper-tail LDP for the KPZ equation at  $(t, \mathbf{x}_1), \dots, (t, \mathbf{x}_n)$ . Because our method relies on positive moments, the results in this paper and in [75] are restricted to the subspace

$$\mathcal{R}_{\text{conc}}(t, \vec{\mathbf{x}}) := \{\vec{\mathbf{r}} \in \mathcal{R}(t, \vec{\mathbf{x}}) : \mathbf{f}_{\star, t, \vec{\mathbf{x}}, \vec{\mathbf{r}}} \text{ is concave}\}. \quad (2.33)$$

Let us mention a related property. Recall that the hypograph of a function is  $\text{hypo}(\mathbf{f}) := \{(\mathbf{x}, \mathbf{r}) : \mathbf{r} \leq \mathbf{f}(\mathbf{x}), \mathbf{x} \in \mathbb{R}\}$ . When  $\vec{\mathbf{r}} \in \mathcal{R}_{\text{conc}}(t, \vec{\mathbf{x}})$ , the function  $\mathbf{f}_{\star, t, \vec{\mathbf{x}}, \vec{\mathbf{r}}} = \mathbf{f}_\star$  has its hypograph  $\text{hypo}(\mathbf{f}_\star)$  given by the convex hull of  $\text{hypo}(\mathbf{p}(1)) \cup \{(\mathbf{x}_c, \mathbf{r}_c) : c = 1, \dots, n\}$ , but this property fails when  $\vec{\mathbf{r}} \in \mathcal{R}(t, \vec{\mathbf{x}}) \setminus \mathcal{R}_{\text{conc}}(t, \vec{\mathbf{x}})$ .

The companion paper [75] shows that the KPZ equation satisfies the finite-dimensional LDP with the rate function  $I_{\text{KPZ}}$ , and that  $\mathbf{h}_\star$  gives the corresponding spacetime limit shape, under the same scaling regimes considered in this paper; see Theorem 1.1 in [73] for the precise statement. A similar LDP is proven in [30] based the Brownian Gibbs resampling property [14, 15]. The results in [30] cover a different set of scaling regimes, work for configurations in  $\mathcal{R}(t, \vec{\mathbf{x}})$ , and give very detailed probability bounds.

Theorem 2.6 gives the connection between  $I_{\text{KPZ}}$ ,  $L_{\text{SHE}}$ ,  $\mathbf{h}_\star$ , which is used in [75].

#### Notation for Theorem 2.6

- (a) Fix  $t$  and  $\vec{x}$ ; view  $I_{\text{KPZ}}(t, \vec{x}, \vec{r}) =: I_{\text{KPZ}}(\vec{r})$  and  $L_{\text{SHE}}(0 \xrightarrow{t} (\vec{x}, \vec{m})) =: L_{\text{SHE}}(\vec{m})$  as functions on  $\mathcal{R}_{\text{conc}}(t, \vec{x})$  and on  $[0, \infty)^n$ , respectively.
- (b) Fix any  $(\mathbf{x}_1 < \dots < \mathbf{x}_n)$ , consider the terminal time  $\mathbf{t}$ , take any pair  $(\vec{r}, \vec{m}) \in \mathcal{R}_{\text{conc}}(\mathbf{t}, \vec{x})^\circ \times (0, \infty)^n$  that satisfies  $(\nabla_{\vec{r}} I_{\text{KPZ}})(t, \vec{x}, \vec{r}) = \vec{m}$ , and let  $\mathbf{h}_\star$  and  $\xi$  be the corresponding limit shape and optimal deviation, respectively. As will be explained in Section 8.3, the limit shape  $\mathbf{h}_\star$  is a weak solution of the Hamilton–Jacobi equation of Burgers’ equation and has shocks.
- (c) Let  $\vec{x}, \vec{r}$ , and  $\vec{m}$  be as in Part (b) and take any intermediate time  $\mathbf{t}' \in (0, \mathbf{t}]$ . Traveling in backward time, some of the optimal clusters may have merged by time  $s = \mathbf{t} - \mathbf{t}'$ . Let  $\{\xi_{\mathbf{c}}(\mathbf{t} - \mathbf{t}')\}_{\mathbf{c}=1}^n = \{\mathbf{x}'_1 < \dots < \mathbf{x}'_{n'}\}$  denote the distinct positions of the clusters at that time. Accordingly, let  $\vec{r}' := (\mathbf{h}_\star(\mathbf{t}', \mathbf{x}'_a))_{a=1}^{n'}$ ,  $\mathfrak{C}(\mathbf{a}) := \{\mathbf{c} : \xi_{\mathbf{c}}(\mathbf{t} - \mathbf{t}') = \mathbf{x}'_a\}$ ,  $\mathbf{m}'_a := \sum_{\mathbf{c} \in \mathfrak{C}(\mathbf{a})} \mathbf{m}_{\mathbf{c}}$ , and  $\vec{m}' := (\mathbf{m}'_a)_{a=1}^{n'}$ ; see Fig. 2 for an illustration.

**Theorem 2.6.** *Notation as above.*

- (a) *The functions  $I_{\text{KPZ}}$  and  $L_{\text{SHE}}$  are continuous, strictly convex, and the Legendre transform of each other. Further,  $\nabla I_{\text{KPZ}} = \nabla_{\vec{r}} I_{\text{KPZ}} : \mathcal{R}_{\text{conc}}(t, \vec{x}) \rightarrow [0, \infty)^n$  is a homeomorphism.*
- (b) *The trajectories of the shocks in  $\mathbf{h}_\star$  and the trajectories of the optimal clusters in  $\xi$  coincide.*
- (c) *We have*

$$(\nabla_{\vec{r}} I_{\text{KPZ}})(\mathbf{t}', \vec{x}', \vec{r}') = \vec{m}', \quad (2.34)$$

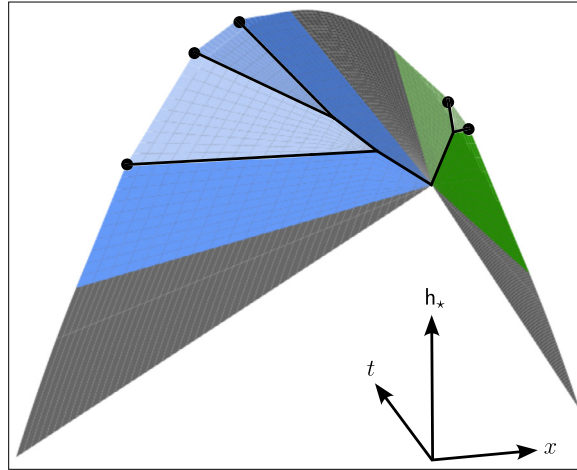
$$L_{\text{SHE}}(0 \xrightarrow{\mathbf{t}} (\vec{x}, \vec{m})) = L_{\text{SHE}}(0 \xrightarrow{\mathbf{t}'} (\vec{x}', \vec{m}')) + \sum_{a=1}^{n'} L_{\text{SHE}}(\mathbf{x}'_a \xrightarrow{\mathbf{t}-\mathbf{t}'} (\mathbf{x}_{\mathbf{c}}, \mathbf{m}_{\mathbf{c}})_{\mathbf{c} \in \mathfrak{C}(\mathbf{a})}). \quad (2.35)$$

### 3. Notation, definitions, tools

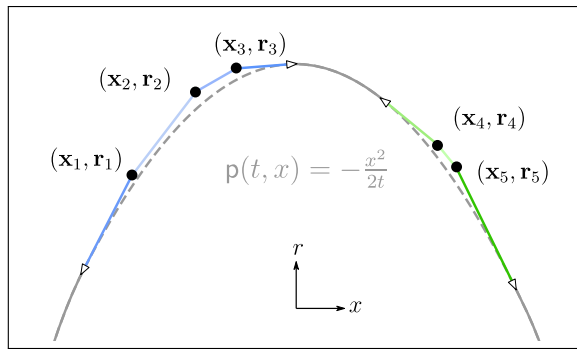
#### 3.1. Reduction to $\mathbf{m} = 1$

We begin by explaining how Theorem 2.3, Theorem 2.4, and Corollary 2.5 follow from the special case of  $\mathbf{m} = 1$ . The key lies in certain scaling relations. Consider the scaling operator  $\langle \frac{1}{\mathbf{m}} \mathfrak{S}_{\mathbf{m}} \lambda, \mathbf{f} \rangle := \langle \lambda, \mathbf{f}(\cdot/\mathbf{m}) \rangle$ . The map  $\frac{1}{\mathbf{m}} \mathfrak{S}_{\mathbf{m}} : \mathbf{m} \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$  is a homeomorphism. Applying  $\frac{1}{\mathbf{m}} \mathfrak{S}_{\mathbf{m}}$  to  $\mu_N(s)$  gives  $\frac{1}{\mathbf{m}} \mathfrak{S}_{\mathbf{m}} \mu_N(s) = \frac{1}{N\mathbf{m}} \sum_{i=1}^{N\mathbf{m}} \delta_{X_i(Ts)/(N\mathbf{m}T)}$ . This can be viewed as a unit-mass empirical measure with  $N' := N\mathbf{m}$  being the scaling parameter. We will verify in Appendix A the scaling identities

$$\mathcal{I}(\tfrac{1}{\mathbf{m}} \mathfrak{S}_{\mathbf{m}} \mu) = \tfrac{1}{\mathbf{m}^3} \mathcal{I}(\mu), \quad \mathbb{I}(\tfrac{1}{\mathbf{m}} \mathfrak{S}_{\mathbf{m}} \mu) = \tfrac{1}{\mathbf{m}^3} \mathbb{I}(\mu), \quad \text{for all } \mu \in \mathcal{C}([0, \mathbf{t}], \mathbf{m} \mathcal{P}(\mathbb{R})) \quad (3.1)$$



**Fig. 3.** The limit shape  $h_*$ . In the gray regions  $h_* = p$ ; in the colored regions  $h_*$  is piecewise linear.



**Fig. 4.** The function  $f_* = f_{*,t,\bar{x},\bar{r}}$ .

The  $1/m^3$  factor absorbs the change in the LDP speed by going from  $N$  to  $N' = Nm$ . Combining what said above shows that Theorem 2.3, Theorem 2.4, and Corollary 2.5 follow from the special case of  $m = 1$ .

### 3.2. The spaces $\mathcal{P}(\mathbb{R})$ and $\mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$

By Section 3.1, we consider  $m = 1$  only.

Let us introduce some metrics on  $\mathcal{P}(\mathbb{R})$  and  $\mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$ . Even though we endow  $\mathcal{P}(\mathbb{R})$  with the weak\* topology, it will be convenient to also consider the 1-Wasserstein metric:

$$\mathcal{W}(\lambda, \lambda') := \inf \{ \langle \pi, |x - x'| \rangle : \pi \in \mathcal{P}(\mathbb{R}^2), \langle \pi, \cdot \otimes 1 \rangle = \lambda, \langle \pi, 1 \otimes \cdot \rangle = \lambda' \}. \quad (3.2)$$

The 1-Wasserstein metric on  $\mathcal{P}(\mathbb{R})$  permits the inverse-CDF formula:

$$\mathcal{W}(\lambda, \lambda') = \int_{\mathbb{R}} dx |F[\lambda](x) - F[\lambda'](x)| = \int_0^1 da |\mathfrak{X}[\lambda](a) - \mathfrak{X}[\lambda'](a)|. \quad (3.3)$$

With  $\mathbb{R}$  being non-compact, the 1-Wasserstein metric produces a topology *stronger* than the weak\* topology on  $\mathcal{P}(\mathbb{R})$ . To metrize the weak\* topology, we introduce

$$\text{dist}(\lambda, \lambda') := \sum_{k=1}^{\infty} 2^{-k} \min \left\{ 1, \int_{-k}^k dx |F[\lambda](x) - F[\lambda'](x)| \right\}. \quad (3.4)$$

It is not hard to check that  $\text{dist}$  metrizes the weak\* topology on  $\mathcal{P}(\mathbb{R})$ . Accordingly,

$$\text{dist}_{[0, \mathfrak{t}]}(\mu, \mu') := \sup \left\{ \text{dist}(\mu(s), \mu'(s)) : s \in [0, \mathfrak{t}] \right\} \quad (3.5)$$

metrizes the topology on  $\mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  introduced in Section 2.2. Let us note two useful inequalities related to these metrics. First, by (3.3)–(3.4),  $\text{dist}(\lambda, \lambda') \leq \mathcal{W}(\lambda, \lambda')$ . Next, for any  $v_1, v_2, \dots \in (0, 1]$  that add to 1 and any  $y_1, y_2, \dots, y'_1, y'_2, \dots \in \mathbb{R}$ ,

$$\text{dist} \left( \sum_i v_i \delta_{y_i}, \sum_i v_i \delta_{y'_i} \right) \leq \mathcal{W} \left( \sum_i v_i \delta_{y_i}, \sum_i v_i \delta_{y'_i} \right) \leq \sum_i v_i |y_i - y'_i|, \quad (3.6)$$

which holds thanks to the coupling  $\pi(\{y_i, y'_i\}) = v_i, i = 1, 2, \dots$

We will need a criterion for a set  $S \subset \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  to be precompact. Recall the topology of  $\mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  from before Theorem 2.3 and recall that we endow  $\mathcal{P}(\mathbb{R})$  with the weak\* topology. First, by a generalized version of the Arzelà–Ascoli theorem, the set  $S$  is precompact if it is equi-continuous and if  $\{\mu(s) : \mu \in S, s \in [0, \mathfrak{t}]\}$  is precompact in  $\mathcal{P}(\mathbb{R})$ ; see [78, Thm. 47.1] for example. By the Banach–Alaoglu theorem, for any  $b < \infty$ , the set  $\{\lambda \in \mathcal{P}(\mathbb{R}) : \text{supp}(\lambda) \subset [-b, b]\}$  is compact. From this property, it is not hard to show that, for any  $b_1, b_2, \dots \rightarrow \infty$ , the set  $\cap_{k=1}^{\infty} \{\lambda \in \mathcal{P}(\mathbb{R}) : \lambda(\mathbb{R} \setminus [-b_k, b_k]) \leq 1/k\}$  is precompact in  $\mathcal{P}(\mathbb{R})$ . These properties give the following criterion.

**Lemma 3.1.** *A set  $S \subset \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  is precompact if*

- (i) *the set  $S$  is equicontinuous with respect to  $\text{dist}$ , and*
- (ii) *there exists  $b_k \rightarrow \infty$  such that  $\langle \mu(s), \mathbf{1}_{\mathbb{R} \setminus [-b_k, b_k]} \rangle \leq 1/k$ , for all  $k \in \mathbb{Z}_{>0}$ ,  $\mu \in S$ , and  $s \in [0, \mathfrak{t}]$ .*

Here is a list of useful properties.

**Lemma 3.2.**

- (a) *For any  $\lambda \in \mathcal{P}(\mathbb{R})$  and  $f \in \mathcal{L}^1(\mathbb{R}, \lambda)$ ,  $\langle \lambda, f \rangle = \int_0^1 da f(\mathfrak{X}[\lambda](a))$ .*



- (b) For any  $\mu \in \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$ , the tail mass  $\sup_{s \in [0, \mathfrak{t}]} \langle \mu(s), \mathbf{1}_{\mathbb{R} \setminus [-b, b]} \rangle$  tends to zero as  $b \rightarrow \infty$ .
- (c) For any  $\mu \in \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$ , for Lebesgue almost every  $a \in [0, 1]$ ,  $\mathfrak{X}[\mu(\cdot)](a) \in \mathcal{C}[0, \mathfrak{t}]$ .

Part (a) is the standard inverse-CDF formula; Parts (b)–(c) are proven in Appendix A.

### 3.3. Expressing $\Phi[\lambda]$ in the quantile coordinate

For any  $\lambda \in \mathcal{P}(\mathbb{R})$  and  $a \in [0, 1]$ , let

$$a_- = a_-[\lambda](a) := \inf\{a' \in [0, 1] : \mathfrak{X}[\lambda](a') = \mathfrak{X}[\lambda](a)\}, \quad (3.7a)$$

$$a_+ = a_+[\lambda](a) := \sup\{a' \in [0, 1] : \mathfrak{X}[\lambda](a') = \mathfrak{X}[\lambda](a)\}. \quad (3.7b)$$

Note that  $a_- < a_+$  if and only if  $\lambda$  has an atom at  $\mathfrak{X}[\lambda](a)$ . Refer to the second expression of  $\Phi[\lambda]$  in (2.8) and note that  $\lambda((-\infty, \mathfrak{X}[\lambda](a))) = a_-$  and  $\lambda((\mathfrak{X}[\lambda](a), \infty)) = (1 - a_+)$ . We have

$$\Phi[\lambda](\mathfrak{X}[\lambda](a)) = \begin{cases} \frac{1}{2} - a, & \text{when } a_- = a_+, \\ \frac{1}{2} - \frac{a_- + a_+}{2} = \frac{1}{a_+ - a_-} \int_{a_-}^{a_+} da \left(\frac{1}{2} - a\right), & \text{when } a_- < a_+. \end{cases} \quad (3.8)$$

Fix any  $f \in \mathcal{L}^1(\mathbb{R}, \lambda)$ . By (3.8), we have  $f(\mathfrak{X}[\lambda](a)) \cdot \Phi[\lambda](\mathfrak{X}[\lambda](a)) = f(\mathfrak{X}[\lambda](a)) \cdot (\frac{1}{2} - a)$  whenever  $a_- = a_+$ , and  $\int_{a_-}^{a_+} da f(\mathfrak{X}[\lambda](a)) \cdot \Phi[\lambda](\mathfrak{X}[\lambda](a)) = \int_{a_-}^{a_+} da f(\mathfrak{X}[\lambda](a)) \cdot (\frac{1}{2} - a)$  whenever  $a_- < a_+$ . Combining these properties gives

$$\int_0^1 da f(\mathfrak{X}[\lambda](a)) \cdot \Phi[\lambda](\mathfrak{X}[\lambda](a)) = \int_0^1 da f(\mathfrak{X}[\lambda](a)) \left(\frac{1}{2} - a\right). \quad (3.9)$$

### 3.4. Dividing a measure

Let us introduce a procedure of dividing a given  $\mu \in \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  into pieces with constant masses. Fix any  $\mathbf{m}_1, \dots, \mathbf{m}_n \in (0, 1]$  that add up to  $\mathbf{m} = 1$ . For any  $s \in [0, \mathfrak{t}]$  and  $\mathfrak{c} \in \{1, \dots, n\}$ , consider  $F_{\mathfrak{c}}(s, x) := \max\{\min\{F[\mu(s)](x) - (\mathbf{m}_1 + \dots + \mathbf{m}_{\mathfrak{c}-1}), \mathbf{m}_{\mathfrak{c}}\}, 0\}$ . Namely, we consider the graph of  $F[\mu(s)](\cdot)$  between the horizontal levels  $(\mathbf{m}_1 + \dots + \mathbf{m}_{\mathfrak{c}-1})$  and  $(\mathbf{m}_1 + \dots + \mathbf{m}_{\mathfrak{c}})$  and shift the graph down so that the lower level is at 0. The result  $F_{\mathfrak{c}}(s, x)$  is the CDF of a measure with total mass  $\mathbf{m}_{\mathfrak{c}}$ , and we let  $\mu_{\mathfrak{c}}(s) \in \mathbf{m}_{\mathfrak{c}}\mathcal{P}(\mathbb{R})$  denote that measure.

Here are a few properties of  $\mu_{\mathfrak{c}}$ . First,  $\mu_{\mathfrak{c}}(s)$  is continuous in  $s$ . To see why, note that, by construction,  $|F_{\mathfrak{c}}(s', x) - F_{\mathfrak{c}}(s, x)| \leq |F(s', x) - F(s, x)|$ . Using this property in conjunction with (3.4) and the continuity of  $\mu$  gives the continuity of  $\mu_{\mathfrak{c}}$ . Next, by construction, the measure  $\mu_{\mathfrak{c}}(s)$  is supported in  $[\mathfrak{X}[\mu(s)](\mathbf{m}_1 + \dots + \mathbf{m}_{\mathfrak{c}-1}), \mathfrak{X}[\mu(s)](\mathbf{m}_1 +$

$\dots + \mathbf{m}_c]$ . Finally, note that  $\mu_c(s)$  and  $\mu_{c+1}(s)$  can both have an atom at  $\mathfrak{X}[\mu(s)](\mathbf{m}_1 + \dots + \mathbf{m}_c)$ .

#### 4. LDP for the attractive BPs: properties of the rate functions

As explained in Section 3.1, we will only consider  $\mathbf{m} = 1$ , so all deviations take value in  $\mathcal{P}(\mathbb{R})$ .

##### 4.1. The quantile representation

Here, we show that  $\mathcal{I} = \mathbb{I}$  on  $\mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$ . To simplify notation, write  $\mathfrak{X}[\mu(s)](a) = \mathfrak{X}(s, a)$ ,  $\Phi[\mu(s)](x) = \Phi(s, x)$ , and  $F[\mu(s)](x) := F(s, x)$ .

**Case 1:**  $\partial_s \mathfrak{X}[\mu] \in \mathcal{L}^2([0, t] \times [0, 1])$ . We seek to express the terms in (2.10) in the quantile coordinate. Take any  $\mathbf{h} \in \mathcal{C}_b^{1,1}([0, t], \mathbb{R})$ . Under the assumption of Case 1,  $\mathfrak{X}(s, a)$  is differentiable in  $s$  Lebesgue almost everywhere (a.e.) on  $[0, t] \times [0, 1]$ , so  $\frac{d}{ds} \mathbf{h}(s, \mathfrak{X}(s, a)) - \partial_s \mathbf{h}(s, \mathfrak{X}(s, a)) = \partial_x \mathbf{h}(s, \mathfrak{X}(s, a)) \partial_s \mathfrak{X}(s, a)$  Lebesgue a.e. Apply  $\int_{[0, t]} ds \int_0^1 da$  to both sides. On the left hand side of the result, use Lemma 3.2(a) in reverse to turn the result into  $\langle \mu(s), \mathbf{h}(s) \rangle|_0^t - \int_{[0, t]} ds \langle \mu, \partial_s \mathbf{h} \rangle$ , which gives the first two terms in (2.9). Next, use Lemma 3.2(a) to express the last term in (2.9) and the integral term in (2.10) in the quantile coordinate. Collecting the preceding results gives

$$\Lambda(\mu, \mathbf{h}) - \int_0^t ds \frac{1}{2} \langle \mu, (\partial_x \mathbf{h})^2 \rangle = \int_0^t ds \int_0^1 da \left( (\partial_x \mathbf{h})_{\mathfrak{X}} (\partial_s \mathfrak{X} - \Phi_{\mathfrak{X}}) - \frac{1}{2} ((\partial_x \mathbf{h})_{\mathfrak{X}})^2 \right), \quad (4.1)$$

where the subscript means  $(\varphi_{\mathfrak{X}})(s, a) := \varphi(t, \mathfrak{X}(s, a))$ , transformation to the quantile coordinate. Taking the supremum over  $\mathbf{h} \in \mathcal{C}_b^{1,1}([0, t], \mathbb{R})$  gives

$$\mathcal{I}(\mu) = \sup_{\mathbf{h}} \int_0^t ds \int_0^1 da \left( (\partial_x \mathbf{h})_{\mathfrak{X}} (\partial_s \mathfrak{X} - \Phi_{\mathfrak{X}}) - \frac{1}{2} ((\partial_x \mathbf{h})_{\mathfrak{X}})^2 \right). \quad (4.2)$$

Let us simplify (4.2). Within the integral, write  $\partial_x \mathbf{h} = \mathbf{g}$  and complete the square to get  $-\frac{1}{2}(\mathbf{g}_{\mathfrak{X}} - \partial_s \mathfrak{X} + \Phi_{\mathfrak{X}})^2 + \frac{1}{2}(\partial_s \mathfrak{X} - \Phi_{\mathfrak{X}})^2$ , and recognize the contribution of  $\frac{1}{2}(\partial_s \mathfrak{X} - \Phi_{\mathfrak{X}})^2$  as  $\mathbb{I}(\mu)$ . The supremum is taken over  $\mathbf{h} \in \mathcal{C}_b^{1,1}([0, t], \mathbb{R})$ . Via an approximation argument, we can replace this supremum with the supremum over those  $\mathbf{g}$ s with  $\int_{[0, t]} ds \langle \mu, \mathbf{g}^2 \rangle < \infty$ . Doing so gives

$$\mathcal{I}(\mu) = \mathbb{I}(\mu) - \inf \left\{ \int_0^t ds \int_0^1 da \frac{1}{2} (\mathbf{g}_{\mathfrak{X}} - \partial_s \mathfrak{X} + \Phi_{\mathfrak{X}})^2 : \int_0^t ds \langle \mu, \mathbf{g}^2 \rangle < \infty \right\}. \quad (4.3)$$

It remains only to show that the infimum in (4.3) is zero. Take any partition  $\Gamma = \{0 = s_0 < s_1 < \dots < s_{|\Gamma|} = t\}$  of  $[0, t]$ , for any  $\varphi \in \mathcal{C}[0, t]$ , let

$$(\partial_\Gamma \varphi)(s) := \sum_{j=1}^{|\Gamma|} \frac{\varphi(s_j) - \varphi(s_{j-1})}{s_j - s_{j-1}} \mathbf{1}_{[s_{j-1}, s_j)}(s) \quad (4.4)$$

denote the discrete-time derivative with respect to  $\Gamma$ , and take  $\mathbf{g}(s, x) := (\partial_\Gamma \mathfrak{X}(\cdot, F(\cdot, x))(s) - \Phi(s, x))$ . Insert this  $\mathbf{g}$  into (4.3). The terms  $-\Phi_{\mathfrak{X}}$  and  $\Phi_{\mathfrak{X}}$  cancel each other. For the discrete-derivative term, note that  $\mathfrak{X}(s_j, F(s_j, x))_{\mathfrak{X}} := \mathfrak{X}(s_j, F(s_j, \mathfrak{X}(s_j, a)))$ . It is not hard to check that, for every  $s_j$ , the last expression is equal to  $\mathfrak{X}(s_j, a)$  Lebesgue a.e. on  $[0, 1]$ . The resulting integral hence reads  $\int_{[0, t]} ds \int_0^1 da \frac{1}{2} (\partial_\Gamma \mathfrak{X} - \partial_s \mathfrak{X})^2$ . Under the assumption of Case 1, this integral tends to zero as the mesh of  $\Gamma$  tends to zero. This completes the proof for Case 1.

**Case 2:**  $\partial_s \mathfrak{X}[\mu] \notin \mathcal{L}^2([0, t] \times [0, 1])$ . In this case, by definition,  $\mathbb{I}(\mu) := \infty$ . Take any partition  $\Gamma = \{0 = s_0 < s_1 < \dots < s_{|\Gamma|} = t\}$  of  $[0, t]$  and let  $\partial_\Gamma$  be as in (4.4). We will show that

$$\mathcal{I}(\mu) + \frac{t}{8} \geq \frac{1}{4} \int_0^t ds \int_0^1 da (\partial_\Gamma \mathfrak{X})^2. \quad (4.5)$$

By the assumption of Case 2 and Lemma 3.2(c), the right hand side of (4.5) tends to  $\infty$  as the mesh of  $\Gamma$  tends to zero. Hence proving (4.5) will give the desired result  $\mathcal{I}(\mu) = \mathbb{I}(\mu) = \infty$ .

The proof will invoke a variant of  $\mathcal{I}$ . Let us introduce this variant and its properties.

$$\mathcal{I}^\circ_{[s', s'']}(\mu) := \sup_{\mathbf{h} \in \mathcal{C}_b^{1,1}([s', s''], \mathbb{R})} \left\{ \langle \mu(s), \mathbf{h}(s) \rangle \Big|_{s'}^{s''} - \int_{s'}^{s''} ds \langle \mu, \partial_s \mathbf{h} \rangle - \int_{s'}^{s''} ds \langle \mu, (\partial_x \mathbf{h})^2 \rangle \right\}. \quad (4.6)$$

The following properties are not difficult to verify, which we do in Appendix A.

- (a) Additivity in time:  $\mathcal{I}^\circ_{[s_1, s_2]} + \mathcal{I}^\circ_{[s_2, s_3]} = \mathcal{I}^\circ_{[s_1, s_3]}$ .
- (b) Monotonicity in time: For all  $[s_2, s_3] \subset [s_1, s_4]$ ,  $\mathcal{I}^\circ_{[s_2, s_3]} \leq \mathcal{I}^\circ_{[s_1, s_4]}$ .
- (c) Convexity:  $\mu \mapsto \mathcal{I}^\circ_{[s', s'']}(\mu)$  is convex.
- (d) Space translation invariance:  $\mathcal{I}^\circ_{[s', s'']}(\mathsf{T}_{0,y}\mu) = \mathcal{I}^\circ_{[s', s'']}(\mu)$ , where  $\mathsf{T}_{0,y}$  translates  $\mu$  in space by  $y$ , namely  $\langle \mathsf{T}_{0,y}\mu(s), \mathbf{f} \rangle := \langle \mu(s), \mathbf{f}(\cdot - y) \rangle$  for  $\mathbf{f} \in \mathcal{C}_b(\mathbb{R})$ .
- (e) For any time-independent  $\lambda \in \mathcal{P}(\mathbb{R})$ ,  $\mathcal{I}^\circ_{[s', s'']}(\lambda) = 0$ .

We begin by bounding  $\mathcal{I}(\mu)$  from below. Young's inequality and the property  $|\Phi[\mu]| \leq 1/2$  together give  $-\Phi[\mu]\partial_x \mathbf{h} \geq -\frac{1}{8} - \frac{1}{2}(\partial_x \mathbf{h})^2$ . Insert this inequality into (2.9)–(2.10), and, within the result, move the term  $\int_{[0, t]} ds \langle \mu, -1/8 \rangle = -t/8$  to the left hand side. Doing so gives  $\mathcal{I}(\mu) + t/8 \geq \mathcal{I}^\circ_{[0, t]}(\mu)$ . Next, we mollify  $\mu$ . Define the time-space translation operator by  $\langle (\mathsf{T}_{s,y}\nu)(s'), \mathbf{f} \rangle := \langle \nu(s' + s), \mathbf{f}(\cdot - y) \rangle$ , take  $\varphi_\delta(x) := \exp(-x^2/(2\delta))/\sqrt{2\pi\delta}$ , and use  $\varphi_\delta$  to mollify  $\mu$  to get

$$\mu_\delta := \int_{\mathbb{R}} ds \int_{\mathbb{R}} dy \varphi_\delta(s) \varphi_\delta(y) \mathsf{T}_{s,y}\mu, \quad (4.7)$$

with the convention  $\mu(s)|_{s<0} := \mu(0)$  and  $\mu(s)|_{s>t} := \mu(t)$ . Apply  $\mathcal{I}^\circ_{[0,t]}(\cdot)$  to both sides of (4.7), use Property (c), and use Property (d). Doing so gives

$$\mathcal{I}^\circ_{[0,t]}(\mu_\delta) \leq \int_{\mathbb{R}} ds \int_{\mathbb{R}} dy \varphi_\delta(s) \varphi_\delta(y) \mathcal{I}^\circ_{[0,t]}(\mathbb{T}_{s,y}\mu) = \int_{\mathbb{R}} ds \varphi_\delta(s) \mathcal{I}^\circ_{[0,t]}(\mathbb{T}_{s,0}\mu). \quad (4.8)$$

Next, for  $s \geq 0$ , using Properties (a), (e), and (b) in order gives  $\mathcal{I}^\circ_{[0,t]}(\mathbb{T}_{s,0}\mu) = \mathcal{I}^\circ_{[0,t-s]}(\mu) + \mathcal{I}^\circ_{[t-s,t]}(\mu(t)) = \mathcal{I}^\circ_{[t-s,t]}(\mu) + 0 \leq \mathcal{I}^\circ_{[0,t]}(\mu) + 0$ , and similarly for  $s < 0$ . Hence

$$\mathcal{I}^\circ_{[0,t]}(\mu_\delta) \leq \left( \int_{\mathbb{R}} ds \varphi_\delta(s) \right) \mathcal{I}^\circ_{[0,t]}(\mu) = \mathcal{I}^\circ_{[0,t]}(\mu) \leq \mathcal{I}(\mu) + \frac{t}{8}. \quad (4.9)$$

Write  $\mathcal{I}^\circ_j := \mathcal{I}^\circ_{[s_{j-1}, s_j]}$ . The last bound and Property (a) give

$$\mathcal{I}(\mu) + \frac{t}{8} \geq \mathcal{I}^\circ_1(\mu_\delta) + \dots + \mathcal{I}^\circ_{|\Gamma|}(\mu_\delta). \quad (4.10)$$

We next bound  $\mathcal{I}^\circ_j(\mu_\delta)$  from below. Write  $F_\delta(s, x) := F[\mu_\delta(s)](x)$  and  $\mathfrak{X}_\delta(s, a) := \mathfrak{X}[\mu_\delta(s)](a)$  to simplify notation. By the construction of  $\mu_\delta$ , both functions are  $\mathcal{C}^\infty$ . Further, since the mollifier  $\varphi_\delta$  is strictly positive everywhere,  $F_\delta(s, x)$  strictly increases in  $x$ , whereby  $F_\delta(s, \mathfrak{X}_\delta(s, a)) = a$  for all  $s$  and  $a$ . Fix any  $v \in \mathcal{C}^\infty[0, 1]$  and let  $h(s, x) := \int_{-\infty}^x dy v(F_\delta(s, y))$ . For this test function,  $\frac{d}{ds} h(s, \mathfrak{X}_\delta(s, x)) = (\partial_s h)_{\mathfrak{X}_\delta} + (\partial_x h)_{\mathfrak{X}_\delta} \partial_s \mathfrak{X}_\delta$  and  $(\partial_x h)_{\mathfrak{X}_\delta} = (v \circ F_\delta)_{\mathfrak{X}_\delta}$ , which is equal to  $v$  thanks to the relation  $F_\delta(s, \mathfrak{X}_\delta(s, a)) = a$ . Hence  $\frac{d}{ds} h(s, \mathfrak{X}_\delta(s, x)) - (\partial_s h)_{\mathfrak{X}_\delta} - (\partial_x h)_{\mathfrak{X}_\delta}^2 = v \partial_s \mathfrak{X}_\delta - v^2$ . Apply  $\int_{[s_{j-1}, s_j]} ds \int_{[0, 1]} da$  to both sides and compare the result with (4.6). Doing so gives

$$\mathcal{I}^\circ_j(\mu_\delta) \geq \int_0^1 da \left( v(a) \mathfrak{X}_\delta(s, a) \Big|_{s_{j-1}}^{s_j} - v(a)^2 (s_j - s_{j-1}) \right). \quad (4.11)$$

Optimizing (4.11) over  $v \in \mathcal{C}^\infty[0, 1]$  gives

$$\mathcal{I}^\circ_j(\mu_\delta) \geq (s_j - s_{j-1}) \int_0^1 da \frac{1}{4} \left( \frac{\mathfrak{X}_\delta(s_j, a) - \mathfrak{X}_\delta(s_{j-1}, a)}{s_j - s_{j-1}} \right)^2. \quad (4.12)$$

We are now ready to conclude the desired result. Combining (4.10) and (4.12) gives

$$\mathcal{I}(\mu) + \frac{t}{8} \geq \frac{1}{4} \int_{[0,t]} ds \int_{[0,1]} da (\partial_\Gamma \mathfrak{X}_\delta)^2. \quad (4.13)$$

With the aid of Lemma 3.2(c), it is not hard to check that  $\mathfrak{X}_\delta \rightarrow \mathfrak{X}$  as  $\delta \rightarrow 0$  Lebesgue a.e. on  $[0, t] \times [0, 1]$ . Given this property, sending  $\delta \rightarrow 0$  in (4.13) with the aid of Fatou's lemma gives the desired result (4.5).

#### 4.2. The rate function $\mathcal{I}_\star$ is a good

Here we prove that  $\mathcal{I}_\star$  is good. Recall from (2.13) that  $\mathcal{I}_\star$  consists of  $\mathcal{I}$  and a dependence on the initial condition. It hence suffices to show that  $\mathcal{I} = \mathbb{I}$  is a good rate function when restricted to  $\{\mu : \mu(0) = \mu_{\text{sc}}\}$ . We will show that  $\mathcal{I}$  is lower-semicontinuous on  $\mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  and that  $S := \{\mu : \mathbb{I}(\mu) \leq r, \mu(0) = \mu_{\text{sc}}\}$  is precompact.

Let us show the lower semicontinuity of  $\mathcal{I}$ . We begin with some reductions. Recall from (2.10) that  $\mathcal{I}$  is defined as a supremum. Since the supremum of any set of continuous functions is lower semicontinuous, it suffices to check that, for any  $\mathfrak{h} \in \mathcal{C}_b^{1,1}([0, \mathfrak{t}], \mathbb{R})$ ,

$$\mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R})) \rightarrow \mathbb{R}, \quad \mu \mapsto \Lambda(\mu, \mathfrak{h}) - \int_0^{\mathfrak{t}} ds \frac{1}{2} \langle \mu, (\partial_x \mathfrak{h})^2 \rangle \quad \text{is continuous.} \quad (4.14)$$

Every term in (2.9) and (2.10) is readily seen to be continuous except for  $\int_{[0, \mathfrak{t}]} ds \langle \mu, \Phi[\mu](\partial_x \mathfrak{h}) \rangle$ . It hence suffices to show that, for any  $\mathfrak{f} \in \mathcal{C}_b(\mathbb{R})$ , the map  $\mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R}$ :  $\lambda \mapsto \langle \lambda, \Phi[\lambda] \mathfrak{f} \rangle$  is continuous. To this end, combine Lemma 3.2(a) and the identity (3.9) to get  $\langle \lambda, \Phi[\lambda] \mathfrak{f} \rangle = \int_0^1 da \left(\frac{1}{2} - a\right) \mathfrak{f}(\mathfrak{X}(a))$ . Take any sequence  $\lambda_k \Rightarrow \lambda$ . Note that  $\lambda_k \Rightarrow \lambda$  implies  $\mathfrak{X}[\lambda_k] \rightarrow \mathfrak{X}[\lambda]$  Lebesgue a.e. on  $[0, 1]$ . This property together with the bounded convergence theorem gives  $\langle \lambda_k, \Phi[\lambda_k] \mathfrak{f} \rangle = \int_0^1 da \left(\frac{1}{2} - a\right) \mathfrak{f}(\mathfrak{X}[\lambda_k](a)) \rightarrow \int_0^1 da \left(\frac{1}{2} - a\right) \mathfrak{f}(\mathfrak{X}(a)) = \langle \lambda, \Phi[\lambda] \mathfrak{f} \rangle$ .

To prepare for the proof of the precompactness, we derive a time-continuity estimate. Recall from (2.12) that  $\mathbb{I}$  is defined as an integral over time. Fix  $s' < s'' \in [0, \mathfrak{t}]$ , forgo the integral outside  $[s', s'']$ , factor out  $\frac{1}{2}(s'' - s')$  from the integral, and apply Jensen's inequality with respect to  $(s'' - s')^{-1} \int_{[s', s'']} ds \int_{[0, 1]} da$ . Doing so gives

$$\mathbb{I}(\mu) \geq \frac{1}{2}(s'' - s') \left( \frac{\int_0^1 da \mathfrak{X}[\mu(s)]|_{s'}^{s''}}{s'' - s'} - \frac{\int_{s'}^{s''} ds \int_0^1 da \Phi[\mu] \mathfrak{x}}{s'' - s'} \right)^2. \quad (4.15)$$

Call the first and second terms within the last square  $b$  and  $b'$  respectively. Use the inequality  $(b - b')^2 \geq \frac{1}{2}b^2 - b'^2$  and note that  $|b'| \leq \|\Phi[\mu]\|_\infty \leq 1/2$ . After being simplified, the result reads  $(s'' - s')\mathbb{I}(\mu) + \frac{(s'' - s')^2}{8} \geq \frac{1}{4} \int_0^1 da (\mathfrak{X}[\mu(s'')] - \mathfrak{X}[\mu(s')])^2$ . By the Cauchy-Schwarz inequality, the last integral is bounded from below by  $\int_0^1 da |\mathfrak{X}[\mu(s'')] - \mathfrak{X}[\mu(s')]|$ . We arrive at the time-continuity estimate:

$$\frac{1}{4} \int_0^1 da \left| \mathfrak{X}[\mu(s'')] - \mathfrak{X}[\mu(s')] \right| \leq (s'' - s')\mathbb{I}(\mu) + \frac{1}{8}(s'' - s')^2. \quad (4.16)$$

Based on (4.16), we fix  $r \in [0, \infty)$  and  $\mu_{\text{sc}} \in \mathcal{P}(\mathbb{R})$  and show the precompactness of  $S := \{\mu : \mathbb{I}(\mu) \leq r, \mu(0) = \mu_{\text{sc}}\}$ . We will do so by verifying the conditions in Lemma 3.1. Referring to (3.3)–(3.4), we see that  $\text{dist}(\mu(s''), \mu(s'))$  is bounded by

4 times the left hand side of (4.16). Hence the equicontinuity of  $S$ , which is required by Lemma 3.1(i), follows. To verify the condition in Lemma 3.1(ii), take any  $\mu \in S$  and write  $\langle \mu(s), \mathbf{1}_{\mathbb{R} \setminus [-b, b]} \rangle$  as  $\int da \mathbf{1}\{|\mathfrak{X}[\mu(s)]| > b\}$ . Bound the last integral by  $\int da \mathbf{1}\{|\mathfrak{X}[\mu(0)]| > b/2\} + \int da \mathbf{1}\{|\mathfrak{X}[\mu(s)] - \mathfrak{X}[\mu(0)]| > b/2\}$ . Recognize the former integral as  $\langle \mu(0), \mathbf{1}_{\mathbb{R} \setminus [-b/2, b/2]} \rangle$  and bound the second integral by using Markov's inequality and (4.16) for  $(s', s'') = (0, s)$ . Doing so gives the bound  $(8/b)(s\mathbb{I}(\mu) + s^2/8)$ , which is at most  $(8r\mathfrak{t} + \mathfrak{t}^2)/b$ . Hence,  $\langle \mu(s), \mathbf{1}_{\mathbb{R} \setminus [-b, b]} \rangle \leq \langle \mu_{sc}, \mathbf{1}_{\mathbb{R} \setminus [-b/2, b/2]} \rangle + (8r\mathfrak{t} + \mathfrak{t}^2)/b$ . From this, we see that the condition in Lemma 3.1(ii) is satisfied for a suitable choice of  $b_k \rightarrow \infty$ .

## 5. LDP for the attractive BPs: upper bound

Here we prove the LDP upper bound in Theorem 2.3. We achieve this by first establishing the exponential tightness of  $\mu_N$  (defined at the beginning of Section 5.1) and then proving the weak LDP upper bound (defined at the beginning of Section 5.2). As was explained in Section 3.1, we consider  $\mathfrak{m} = 1$  only.

### 5.1. Exponential tightness [section rewritten]

Here, we seek to prove that  $\mu_N$  is exponentially tight, which means, for any  $\varepsilon > 0$ , there exists a precompact  $S \subset \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  such that

$$\limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P}[\mu_N \notin S] \leq -\frac{1}{\varepsilon}. \quad (5.1)$$

The first step is to devise some events to control the BMs  $B_i$  in (2.2). Consider the events

$$\mathcal{U}_N([s_1, s_2], v) := \left\{ \sum_{i=1}^N \sup_{s \in [s_1, s_2]} |B_i(Ts) - B_i(Ts_1)| \leq N^2 T v |s_2 - s_1|^{1/3} \right\}, \quad (5.2)$$

$$\mathcal{U}_N(v) := \bigcap_{\ell \geq 1} \bigcap_{j=1}^{\ell} \mathcal{U}_N\left(\left[\frac{j-1}{\ell} \mathfrak{t}, \frac{j}{\ell} \mathfrak{t}\right], v\right). \quad (5.3)$$

To control the summand in (5.2), set  $\overline{B}_i[s_1, s_2] := \sup_{s \in [s_1, s_2]} (B_i(Ts) - B_i(Ts_1))$  and  $\underline{B}_i[s_1, s_2] := -\inf_{s \in [s_1, s_2]} (B_i(Ts) - B_i(Ts_1))$  and write

$$\sup_{s \in [s_1, s_2]} |B_i(Ts) - B_i(Ts_1)| = \max\{\overline{B}_i[s_1, s_2], \underline{B}_i[s_1, s_2]\} \leq \overline{B}_i[s_1, s_2] + \underline{B}_i[s_1, s_2]. \quad (5.4)$$

Sum (5.4) over  $i = 1, \dots, N$ . For  $\lambda \in \mathbb{R}$ , we have the inequality

$$\mathbb{E}\left[e^{\lambda \sum_{i=1}^N \sup_{s \in [s_1, s_2]} |B_i(Ts) - B_i(Ts_1)|}\right] \leq \mathbb{E}\left[e^{\lambda \sum_{i=1}^N \overline{B}_i[s_1, s_2]} \cdot e^{\lambda \sum_{i=1}^N \underline{B}_i[s_1, s_2]}\right]. \quad (5.5)$$

The random variables  $\overline{B}_i[s_1, s_2]$  and  $\underline{B}_i[s_1, s_2]$  each have the same law as  $\sqrt{T(s_2 - s_1)}|B_i(1)|$ . On the right hand side of (5.5), use the Cauchy–Schwarz inequality, use the just-mentioned property about the law, and use the property that  $B_1, \dots, B_N$  are independent. Doing so gives

$$\mathbb{E}[e^{\lambda \sum_{i=1}^N \sup_{s \in [s_1, s_2]} |B_i(Ts) - B_i(Ts_1)|}] \leq (\mathbb{E}[e^{2\lambda \sqrt{T(s_2 - s_1)}|B_1(1)|}])^N. \quad (5.6)$$

The inequality  $\exp(|r|) \leq \exp(r) + \exp(-r)$  and the identity  $\mathbb{E}[\exp(\pm r B_1(1))] = \exp(r^2/2)$  together give  $\mathbb{E}[e^{2\lambda \sqrt{T(s_2 - s_1)}|B_1(1)|}] \leq 2e^{2\lambda^2 T(s_2 - s_1)}$ . Using this in (5.6) gives

$$\mathbb{E}[e^{\lambda \sum_{i=1}^N \sup_{s \in [s_1, s_2]} |B_i(Ts) - B_i(Ts_1)|}] \leq 2^N e^{2\lambda^2 NT(s_2 - s_1)}. \quad (5.7)$$

Combining (5.7) and Chernoff's bound  $\mathbb{P}[A > a] \leq e^{-\lambda a} \mathbb{E}[e^{\lambda A}]$  with  $A = \sum_{i=1}^N \sup_{s \in [s_1, s_2]} |B_i(Ts) - B_i(Ts_1)|$ ,  $a = N^2 T v |s_2 - s_1|^{1/3}$ , and  $\lambda = N |s_2 - s_1|^{-2/3} v/4$  yields

$$\mathbb{P}[\mathcal{U}_N([s_1, s_2], v)^c] \leq 2^N \exp\left(-\frac{1}{8} N^3 T v^2 |s_2 - s_1|^{-1/3}\right). \quad (5.8)$$

Next, use (5.8) and the union bound to bound  $\mathbb{P}[\mathcal{U}_N(v)^c]$ . We have, for all  $v \geq 1$  and some universal constant  $c < \infty$ ,

$$\mathbb{P}[\mathcal{U}_N(v)^c] \leq 2^N \sum_{\ell=1}^{\infty} \ell \cdot e^{-\frac{1}{8} N^3 T v^2 \ell^{1/3}} \leq c 2^N e^{-\frac{1}{8} N^3 T v^2}. \quad (5.9)$$

Under (1.4), the factor  $c 2^N$  is negligible compared to  $\exp(O(N^3 T))$ . Use (5.9) to fix a large enough  $v$  such that

$$\limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P}[\mathcal{U}_N(v)^c] \leq -\frac{1}{\varepsilon}. \quad (5.10)$$

Having fixed  $v$ , we write  $\mathcal{U}_N(v) = \mathcal{U}_N$  hereafter.

We next prove that, under  $\mathcal{U}_N$ , all samples of  $\boldsymbol{\mu}_N$  are contained in a fixed precompact set. This will imply the desired exponential tightness. Proving this statement amounts to verifying the conditions in Lemma 3.1 — for all realizations of  $\boldsymbol{\mu}_N$  under  $\mathcal{U}_N$  — with a fixed choice of  $b_1, b_2, \dots \rightarrow \infty$ .

To verify the equicontinuity required by Lemma 3.1(i), apply (3.6) to get

$$\text{dist}(\boldsymbol{\mu}_N(s'), \boldsymbol{\mu}_N(s'')) \leq \frac{1}{N} \sum_{i=1}^N \frac{|X_i(Ts'') - X_i(Ts')|}{NT}, \quad \text{for all } s' < s'' \in [0, t]. \quad (5.11)$$

To bound the last sum, integrate (2.2) over  $[Ts', Ts'']$  and use  $|\text{sgn}| \leq 1$ . Doing so gives

$$\frac{1}{N} \sum_{i=1}^N \frac{|X_i(Ts'') - X_i(Ts')|}{NT} \leq \frac{|s'' - s'|}{2} + \frac{1}{N} \sum_{i=1}^N \frac{|B_i(Ts'') - B_i(Ts')|}{NT}. \quad (5.12)$$

Under  $\mathcal{U}_N$ , the last term in (5.12) is bounded by  $2v|s'' - s'|^{1/3}$ . This is because, for any given  $s' < s'' \in [0, t]$ , we can find  $\ell \in \mathbb{Z}_{>1}$  and  $j \in \{1, \dots, \ell - 1\}$  such that  $s', s'' \in [t(j-1)/\ell, t(j+1)/\ell]$  and that  $s'' - s' > t/\ell$ . Hence

$$\text{under } \mathcal{U}_N, \text{ for all } s' < s'' \in [0, t],$$

$$\frac{1}{N} \sum_{i=1}^N \frac{1}{NT} |X_i(Ts'') - X_i(Ts')| \leq \frac{1}{2}|s'' - s'| + 2v|s'' - s'|^{1/3}. \quad (5.13)$$

Combining (5.11) and (5.13) gives, under  $\mathcal{U}_N$ ,  $\text{dist}(\mu_N(s'), \mu_N(s'')) \leq |s'' - s'|/2 + 2v|s'' - s'|^{1/3}$ , so the desired equicontinuity holds.

To verify the condition in Lemma 3.1(ii), start by writing

$$\langle \mu_N(s), \mathbf{1}_{\mathbb{R} \setminus [-2b, 2b]} \rangle \leq \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{\frac{1}{NT} |X_i(Ts) - X_i(0)| > b\}} + \langle \mu_N(0), \mathbf{1}_{\mathbb{R} \setminus (-b, b)} \rangle. \quad (5.14)$$

We have  $\limsup_{N \rightarrow \infty} \langle \mu_N(0), \mathbf{1}_{\mathbb{R} \setminus (-b, b)} \rangle \leq \langle \mu_{\text{sc}}, \mathbf{1}_{\mathbb{R} \setminus (-b, b)} \rangle$ , because  $\mu_N(0) \Rightarrow \mu_{\text{sc}}$  and because  $\mathbb{R} \setminus (-b, b)$  is closed. Next, bound the summand in (5.14) by  $b^{-1}|X_i(Ts) - X_i(0)|/(NT)$  and apply (5.13). The result gives that, under  $\mathcal{U}_N$ , the first term on the right side of (5.14) is bounded by  $b^{-1}(s \cdot 2^{-1} + 2vs^{1/3}) \leq b^{-1}(t/2 + 2vt^{1/3})$ . These bounds together verify the condition in Lemma 3.1(ii) for  $b_k := (t/2 + 2vt^{1/3})2k$ .

## 5.2. The weak upper bound

We begin by stating the goal. First, given the exponential tightness, it suffices to prove a weak LDP upper bound, namely the LDP upper bound where the closed set is assumed to be compact. Fix any compact  $K \subset \mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$ . Recall  $\mathcal{I}_*$  from (2.13). If  $K \cap \{\mu : \mu(0) = \mu_{\text{sc}}\} = \emptyset$ , then  $\inf_K \mathcal{I}_* = +\infty$  and the desired upper bound follows trivially. We hence assume  $K \cap \{\mu : \mu(0) = \mu_{\text{sc}}\} \neq \emptyset$ , whereby  $\inf_K \mathcal{I}_* = \inf_K \mathcal{I}$ . Further, recall from (2.10) that  $\mathcal{I}$  is defined as a supremum over  $\mathbf{h} \in \mathcal{C}_b^{1,1}([0, t], \mathbb{R})$ . Via an approximation argument, the supremum can be replaced by the one over  $\mathbf{h} \in \mathcal{C}_c^\infty([0, t], \mathbb{R})$ . Hence our goal is to show

$$\limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P}[\mu_N \in K] \leq - \inf_{\mu \in K} \sup_{\mathbf{h} \in \mathcal{C}_c^\infty([0, t], \mathbb{R})} \left\{ \Lambda(\mu, \mathbf{h}) - \int_0^t ds \frac{1}{2} \langle \mu, (\partial_x \mathbf{h})^2 \rangle \right\}. \quad (5.15)$$

Let us use the martingale method [63] to prove (5.15). Take any  $\mathbf{h} \in \mathcal{C}_c^\infty([0, t], \mathbb{R})$  and apply Itô's calculus to  $N^3 T \langle \mu_N(s), \mathbf{h}(s) \rangle = N^2 T \sum_i \mathbf{h}(s, X_i(Ts)/(NT))$  with the aid of (2.2) to get

$$N^3 T \langle \mu_N(s), \mathbf{h}(s) \rangle \Big|_0^{s'} - N^3 T \int_0^{s'} ds \langle \mu_N, \partial_s \mathbf{h} + \Phi[\mu_N](\partial_x \mathbf{h}) + \frac{1}{2N^2 T} \partial_{xx} \mathbf{h} \rangle \quad (5.16a)$$



$$= NT^{1/2} \sum_{i=1}^N \int_0^{s'} dB_i(s) (\partial_x h)(s, \frac{1}{NT} X_i(Ts)). \quad (5.16b)$$

This result implies that the expression in (5.16a), when viewed as a process in  $s'$ , is a martingale with the quadratic variation  $N^3 T \int_0^{s'} ds \langle \mu_N, (\partial_x h)^2 \rangle$ . Hence  $\mathbb{E}[\exp(N^3 T U_N(\mu_N))] = 1$ , where

$$U_N(\mu) := \Lambda(\mu, h) - \int_0^t ds \frac{1}{2} \langle \mu, (\partial_x h)^2 \rangle - \frac{1}{2N^2 T} \int_0^t ds \langle \mu, \partial_{xx} h \rangle, \quad (5.17)$$

and  $\Lambda$  was defined in (2.9). Given any Borel  $A \subset \mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$ , write

$$\mathbb{P}[\mu_N \in A] = \mathbb{E}[\mathbf{1}_A(\mu_N) \exp(-N^3 T U_N(\mu_N)) \exp(N^3 T U_N(\mu_N))], \quad (5.18)$$

bound the first two factors together by  $\sup_{\mu \in A} \exp(-N^3 T U_N(\mu))$ , use  $\mathbb{E}[\exp(N^3 T U_N(\mu_N))] = 1$  for the remaining factor, and apply  $\limsup_N (N^3 T)^{-1} \log(\cdot)$  to both sides of the result. Doing so gives

$$\limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P}[\mu_N \in A] \leq \lim_{N \rightarrow \infty} \sup_{\mu \in A} U_N(\mu) \quad (5.19a)$$

$$= - \inf_{\mu \in A} \left\{ \Lambda(\mu, h) - \int_{[0, t]} ds \frac{1}{2} \langle \mu, (\partial_x h)^2 \rangle \right\}. \quad (5.19b)$$

Since this holds for all  $h \in \mathcal{C}_c^\infty([0, t], \mathbb{R})$ , we further obtain

$$\limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P}[\mu_N \in A] \leq - \sup_{h \in \mathcal{C}_c^\infty([0, t], \mathbb{R})} \inf_{\mu \in A} \left\{ \Lambda(\mu, h) - \int_0^t ds \frac{1}{2} \langle \mu, (\partial_x h)^2 \rangle \right\}. \quad (5.20)$$

We seek to swap the supremum and infimum on the right hand side. To this end, apply Lemmata 3.2–3.3 in Appendix 2 in [53], with  $J_h(\mu) = -\Lambda(\mu, h) + \int_{[0, t]} ds \frac{1}{2} \langle \mu, (\partial_x h)^2 \rangle$ . Under our notation, these Lemmata assert that, if  $J_h(\mu)$  is upper-semicontinuous in  $\mu$  for every  $h$ , and if (5.20) holds for every open  $A$ , then (5.15) holds for every compact  $K$ . As shown after (4.14),  $J_h(\mu)$  is continuous in  $\mu$  for every  $h$ . We have proved (5.20) for every Borel  $A$ . Hence the desired weak upper bound (5.15) follows.

## 6. LDP for the attractive BPs: lower bound

We begin by setting up the goal of the proof. As mentioned in Section 3.1, we consider  $m = 1$  only. Indeed, proving the LDP lower bound amounts to proving that, for any  $\mu \in \mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$  and  $\varepsilon > 0$ ,  $-\mathcal{I}_\star(\mu) \leq \liminf_{\varepsilon \rightarrow 0} \liminf_{N \rightarrow \infty} (N^3 T)^{-1} \log \mathbb{P}[\text{dist}_{[0, t]}(\mu_N, \mu) < \varepsilon]$ . Recall from (2.13) that, when  $\mathcal{I}_\star(\mu) < \infty$ , we have  $\mu(0) = \mu_{sc}$  and  $\mathcal{I}_\star(\mu) = \mathbb{I}(\mu)$ . Our goal is hence as follows.

**Proposition 6.1.** For any  $\mu \in \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  with  $\mu(0) = \mu_{\text{sc}}$  and  $\mathbb{I}(\mu) < \infty$ ,

$$\liminf_{r \rightarrow 0} \liminf_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P}[\text{dist}_{[0, \mathfrak{t}]}(\mu_N, \mu) < r] \geq -\mathbb{I}(\mu). \quad (6.1)$$

### 6.1. Proving a preliminary version of Proposition 6.1

Let us introduce some classes of deviations. We call a deviation  $\xi \in \mathcal{C}([0, \mathfrak{t}], \mathcal{P}(\mathbb{R}))$  **clustering** if it is of the form  $\xi(s) = \sum_{\mathfrak{c}=1}^n \mathfrak{m}_{\mathfrak{c}} \delta_{\xi_{\mathfrak{c}}(s)}$ , for some  $\xi_1 \leq \dots \leq \xi_n \in \mathcal{C}[0, \mathfrak{t}]$  and some  $\mathfrak{m}_1, \dots, \mathfrak{m}_n \in (0, 1]$  that add up to  $\mathfrak{m} = 1$ . We call  $\xi_1, \dots, \xi_n$  the **clusters** or the **trajectories of clusters**. We call a deviation **Piecewise-Linear(PL)-clustering** if it is clustering and the trajectories of its clusters are piecewise linear.

As a first step toward proving Proposition 6.1, here we prove a preliminary version of it where  $\mu$  is replaced by a PL-clustering deviation  $\xi$ . For a PL-clustering deviation, we will use a different way (than  $\text{dist}_{[0, \mathfrak{t}]}$ ) to measure how close the empirical measure is to the deviation. Recall that the attractive BPs are ordered at the start:  $X_1(0) \leq \dots \leq X_N(0)$ . For a given PL-clustering  $\xi$ , to each  $X_i$  assign a cluster  $\xi_{\mathfrak{c}}$  via the ordering of the *indices*:

$$\mathfrak{c}_{N, \xi}(i) := \min\{\mathfrak{c} : \mathfrak{m}_1 + \dots + \mathfrak{m}_{\mathfrak{c}} \geq i/N\}, \mathcal{I}_{N, \xi}(\mathfrak{c}) := \mathfrak{c}_{N, \xi}^{-1}(\{\mathfrak{c}\}) = \{i : \mathfrak{c}_{N, \xi}(i) = \mathfrak{c}\}. \quad (6.2)$$

For  $\vec{x} \in (\mathcal{C}[0, T\mathfrak{t}])^N$ , set

$$\mathfrak{D}\text{ist}_{N, s}(\vec{x}, \xi) := \max\left\{\left|\frac{1}{NT}x_i(Ts) - \xi_{\mathfrak{c}_{N, \xi}(i)}(s)\right| : i = 1, \dots, N\right\}, \quad (6.3)$$

$$\mathfrak{D}\text{ist}_{N, [s', s'']}(\vec{x}, \xi) := \sup_{s \in [s', s'']} \mathfrak{D}\text{ist}_{N, s}(\vec{x}, \xi). \quad (6.4)$$

We measure how close the system of attractive BPs is to  $\xi$  by  $\mathfrak{D}\text{ist}_{N, s}(\vec{X}, \xi)$ , which is a (much) finer measurement than  $\text{dist}(\mu_N(s), \xi(s))$ . By (3.6),

$$\text{dist}(\mu_N(s), \xi(s)) \leq \mathfrak{D}\text{ist}_{N, s}(\vec{X}, \xi). \quad (6.5)$$

Hereafter,  $c = c(v_1, v_2, \dots)$  denotes a constant that depends only on the designated variables  $v_1, v_2, \dots$ . We seek to prove the following preliminary version of Proposition 6.1.

**Proposition 6.2.** Given any PL-clustering  $\xi$ , there exists a  $c = c(\xi) \in [1, \infty)$  such that

$$\liminf_{r \rightarrow 0} \liminf_{N \rightarrow \infty} \inf_{\mathfrak{D}\text{ist}_{N, 0}(\vec{X}, \xi) \leq r} \left\{ \frac{1}{N^3 T} \log \mathbb{P}[\mathfrak{D}\text{ist}_{N, [0, \mathfrak{t}]}(\vec{X}, \xi) \leq cr \mid \vec{X}(0)] \right\} \geq -\mathbb{I}(\mu). \quad (6.6)$$

In (6.6) and similarly hereafter, the conditional probability  $\mathbb{P}[\dots \mid \vec{X}(0)]$  is viewed as a function of  $\vec{X}(0)$ , and the infimum is taken over those  $\vec{X}(0)$ s that satisfy  $\mathfrak{D}\text{ist}_{N, 0}(\vec{X}, \xi) \leq r$ .

**Proof.** We begin by setting up the notation. Let  $0 = s_0 < s_1 < \dots < s_\ell = \mathfrak{t}$  be the time between which the clusters of  $\xi$  are linear. For a small  $\delta > 0$  to be fixed shortly, we call  $[s_0, s_1 - \delta], [s_1 + \delta, s_2 - \delta], \dots, [s_{\ell-1} + \delta, s_\ell]$  the **linear segments**, and call  $[s_1 - \delta, s_1 + \delta], [s_2 - \delta, s_2 + \delta], \dots$ , the **transition segments**. To fix the value of  $\delta$ , consider the maximum speed of all clusters  $\dot{\xi}_{\max} := \sup_{[0, \mathfrak{t}], \mathfrak{c}} |\dot{\xi}_{\mathfrak{c}}|$ , and note that, by the definition of the  $s_j$ s, within each  $(s_{j-1}, s_j)$ , the clusters of  $\xi$  either never meet or completely coincide. We fix a small enough  $\delta$  such that

$$4\delta\dot{\xi}_{\max} \leq 2 \cdot 3^0 r, \text{ and} \quad (6.7)$$

for every linear segment, either

$$|\xi_{\mathfrak{c}} - \xi_{\mathfrak{c}'}| > 2 \cdot 3^{2\ell} r \text{ everywhere on that linear segment, or} \quad (6.8)$$

$$|\xi_{\mathfrak{c}} - \xi_{\mathfrak{c}'}| = 0 \text{ everywhere on that linear segment.}$$

**Step 1: analysis within a linear segment.** Let us set up the notation and state the goal of this step. Fix a linear segment  $[s', s'']$  and define the event

$$\mathcal{C}_{N, \beta}[s', s''] := \{\mathfrak{D}\text{ist}_{N, [s', s'']}(\vec{X}, \xi) < 3^\beta r\}. \quad (6.9)$$

Hereafter,  $\beta \in \{0, 1, \dots, 2\ell - 2\}$  is an auxiliary parameter. Let  $\mathbb{I}_{[s', s'']}$  be the analog of  $\mathbb{I}$  where the time integral (see (2.12)) is restricted to  $[s', s'']$  and let  $\dot{\varphi} := d\varphi/ds$ . It is readily checked that

$$\text{for a clustering } \nu = \sum_{\mathfrak{c}=1}^n \mathfrak{m}_{\mathfrak{c}} \delta_{\nu_{\mathfrak{c}}}, \quad \mathbb{I}_{[s', s'']}(\nu) = \int_{s'}^{s''} ds \sum_{\mathfrak{c}=1}^n \frac{\mathfrak{m}_{\mathfrak{c}}}{2} (\dot{\nu}_{\mathfrak{c}} - \Phi[\nu](\nu_{\mathfrak{c}}))^2. \quad (6.10)$$

The goal of this step is to prove that

$$\liminf_{N \rightarrow \infty} \inf_{\mathfrak{D}\text{ist}_{N, s'}(\vec{X}, \xi) \leq 3^\beta r} \left\{ \frac{1}{N^3 T} \log \mathbb{P}[\mathcal{C}_{N, \beta+1}[s', s''] \mid \vec{X}(Ts')] \right\} \geq -\mathbb{I}_{[s', s'']}(\xi). \quad (6.11)$$

Recall from (6.8) that, within  $[s', s'']$ , any pair of clusters either stay strictly apart or completely coincide. After combining those clusters that completely coincide, we have

$$|\xi_{\mathfrak{c}} - \xi_{\mathfrak{c}'}|_{[s', s'']} > 2 \cdot 3^{2\ell} r, \quad \forall \mathfrak{c} \neq \mathfrak{c}'. \quad (6.12)$$

The first step of proving (6.11) is to set up Girsanov's transform. Within  $[s', s'']$ , the cluster  $\xi_{\mathfrak{c}}$  travels at a constant velocity  $d\xi_{\mathfrak{c}}/ds := \dot{\xi}_{\mathfrak{c}}$ . Letting  $v_{\mathfrak{c}} := \dot{\xi}_{\mathfrak{c}} - \phi_{\mathfrak{c}}$ , where  $\phi_{\mathfrak{c}} := \frac{1}{2}(\dots - \mathfrak{m}_{\mathfrak{c}-1} + \mathfrak{m}_{\mathfrak{c}+1} + \dots)$ , we seek to apply Girsanov's transform to turn  $\mathbb{P}$  into another law  $\mathbb{Q}$  where

$$\text{under } \mathbb{Q}, \quad dX_i = \sum_{j=1}^N \frac{1}{2} \operatorname{sgn}(X_j - X_i) ds + N v_{\mathbf{c}_{N,\xi}(i)} ds + dB_i. \quad (6.13)$$

In plain words, we pursue a “strategy” where each  $X_i$  receives an additional drift  $N v_{\mathbf{c}_{N,\xi}(i)}$ . The term  $\dot{\xi}_{\mathbf{c}}$  in  $v_{\mathbf{c}}$  helps  $X_i(T \cdot)/(NT)$  follow the cluster  $\xi_{\mathbf{c}}$ , while the term  $-\phi_{\mathbf{c}}$  counters the “pulling” from those particles with a different assigned cluster, namely the effect of the drift  $\sum_{j \notin \mathcal{J}_{N,\xi}(\mathbf{c})} \operatorname{sgn}(X_j - X_i)/2$ . Hereafter, we will use the phrase “pulling” similarly, to refer to the effect of the drift coming from a set of particles.

The next step is to set up a stopping time and prepare the relevant properties. Let  $\sigma$  be the first time when the condition required by  $\mathcal{C}_{N,\beta+1}[s', s'']$  is violated, namely

$$\sigma := \inf \{s \in [s', s''] : |\frac{1}{NT} X_i(Ts) - \xi_{\mathbf{c}_{N,\xi}(i)}(s)| = 3^{\beta+1}r, \text{ for some } i\}. \quad (6.14)$$

Since  $\beta < 2\ell$ , by (6.12), particles with different assigned clusters stay strictly ordered within  $[s', \sigma]$ . This property implies, for any  $i \in \mathcal{J}_{N,\xi}(\mathbf{c})$ , the pulling from particles outside  $\mathcal{J}_{N,\xi}(\mathbf{c})$  is given by  $\sum_{j \notin \mathcal{J}_{N,\xi}(\mathbf{c})} \frac{1}{2} \operatorname{sgn}(j - i)$ . The last expression sums to  $N \frac{1}{2}(-\mathbf{m}_1 \dots - \mathbf{m}_{\mathbf{c}-1} + \mathbf{m}_{\mathbf{c}+1} + \dots + \mathbf{m}_n) = N\phi_{\mathbf{c}}$ . Hence, within  $s \in [s', \sigma]$  and for each  $i \in \mathcal{J}_{N,\xi}(\mathbf{c})$ ,

$$\sum_{j=1}^N \frac{1}{2} \operatorname{sgn}(X_j(Ts) - X_i(Ts)) = \sum_{j \in \mathcal{J}_{N,\xi}(\mathbf{c})} \frac{1}{2} \operatorname{sgn}(X_j(Ts) - X_i(Ts)) + N\phi_{\mathbf{c}}. \quad (6.15)$$

Having set up the stopping time  $\sigma$ , we next analyze the evolution of the particles within  $s \in [s', \sigma]$  and under the condition  $\mathfrak{Dist}_{N,s'}(\vec{X}, \xi) \leq 3^\beta r$ , which is required in (6.11). We will analyze separately the evolution of the center of mass (defined later) relative to their assigned clusters and the evolution of a particle relative to its center of mass. Hereafter, unless otherwise noted,  $s \in [s', \sigma]$ .

We now define the center of mass and analyze its evolution. Let  $X_{\mathbf{c}} := \sum_{i \in \mathcal{J}_{N,\xi}(\mathbf{c})} X_i / (N\mathbf{m}_{\mathbf{c}})$  be the center of mass. Take the average of (6.13) over  $i \in \mathcal{J}_{N,\xi}(\mathbf{c})$ . The pulling from within  $i \in \mathcal{J}_{N,\xi}(\mathbf{c})$  averages to zero; the pulling from outside of  $\mathcal{J}_{N,\xi}(\mathbf{c})$  averages to  $N\phi_{\mathbf{c}}$  by (6.15). Hence  $dX_{\mathbf{c}} = N\dot{\xi}_{\mathbf{c}} ds + dB_{\mathbf{c}}$ , where  $B_{\mathbf{c}} := \sum_{i \in \mathcal{J}_{N,\xi}(\mathbf{c})} B_i / (N\mathbf{m}_{\mathbf{c}})$ . Integrating this equation gives

$$|\frac{1}{NT} X_{\mathbf{c}}(Ts) - \xi_{\mathbf{c}}(s)| \leq |\frac{1}{NT} X_{\mathbf{c}}(Ts') - \xi_{\mathbf{c}}(s')| + \frac{1}{NT} |B_{\mathbf{c}}(Ts) - B_{\mathbf{c}}(Ts')|. \quad (6.16)$$

Under the condition  $\mathfrak{Dist}_{N,s'}(\vec{X}, \xi) \leq 3^\beta r$ , the first term on the right hand side is at most  $3^\beta r$ , so

$$\begin{aligned} |\frac{1}{NT} X_{\mathbf{c}}(Ts) - \xi_{\mathbf{c}}(s)| &\leq 3^\beta r + \frac{1}{NT} |B_{\mathbf{c}}(Ts) - B_{\mathbf{c}}(Ts')| \\ &\leq 3^\beta r + \frac{1}{NT} \max \{|B_i(Ts) - B_i(Ts')| : i = 1, \dots, N\}. \end{aligned} \quad (6.17)$$

We next analyze the motion of a particle relative to its center of mass. Fix any  $\mathbf{c}$ . As seen in (6.15), a particle with  $i \in \mathcal{J}_{N,\xi}(\mathbf{c})$  feels the pulling from within and outside of

$\mathfrak{I}_{N,\xi}(\mathbf{c})$ . Intuitively, the former should only pull particles in  $\mathfrak{I}_{N,\xi}(\mathbf{c})$  closer together. To make this statement precise, we appeal to a comparison result from [86]. Write  $\mathfrak{I}_{N,\xi}(\mathbf{c}) = \{i_1, \dots, i_2\}$ , rank  $\{X_i(Ts)\}_{i \in \mathfrak{I}_{N,\xi}}$  to get  $X_{(i_1)}(Ts) \leq \dots \leq X_{(i_2)}(Ts)$ , let  $Z_i(Ts) := X_{(i+1)}(Ts) - X_{(i)}(Ts)$  be the gap process, and let  $L_{i,i+1}$  denote the localtime of  $Z_i$  at 0. By [43, Proposition 3], the ranked processes satisfy

$$dX_{(i)} = \left( N\phi_{\mathbf{c}} + Nv_{\mathbf{c}} + \sum_{j \in \mathfrak{I}_{N,\xi}(\mathbf{c})} \frac{1}{2} \text{sgn}(j-i) \right) ds + dW_i + \frac{1}{2} dL_{i-1,i} - \frac{1}{2} dL_{i,i+1}, \quad (6.18)$$

where  $W_{i_1}, \dots, W_{i_2}$  are independent standard BMs, and we assume  $W_i(s') = 0$  without loss of generality. Next, consider the analog of  $X_{(i)}$  without the inner pulling. Namely, for  $i \in \mathfrak{I}_{N,\xi}(\mathbf{c})$ , consider

$$d\bar{X}_{(i)} = N(\phi_{\mathbf{c}} + v_{\mathbf{c}}) ds + dW_i + \frac{1}{2} d\bar{L}_{i-1,i} - \frac{1}{2} d\bar{L}_{i,i+1}, \quad \bar{X}_i(Ts') = X_{(i)}(Ts'), \quad (6.19)$$

where  $\bar{L}_{i,i+1}$  is the localtime of  $\bar{X}_{(i+1)} - \bar{X}_{(i)}$  at 0. Let  $\bar{Z}_i := \bar{X}_{(i+1)} - \bar{X}_{(i)}$  denote the gap process and let  $f_i(s) := W_{i+1}(s) - W_i(s) + X_{(i)}(Ts')$ . From (6.18)–(6.19), we infer the equations of the gap processes

$$dZ_{(i)} = df_i - ds - \frac{1}{2} dL_{i-1,i} + dL_{i,i+1} - \frac{1}{2} dL_{i+1,i+2}, \quad (6.20)$$

$$d\bar{Z}_{(i)} = df_i - \frac{1}{2} dL_{i-1,i} + dL_{i,i+1} - \frac{1}{2} dL_{i+1,i+2}, \quad (6.21)$$

for  $i = i_1, \dots, i_2 - 1$ , with the convention that  $L_{i-1,i} := 0$  and  $L_{i_2,i_2+1} := 0$ . Now apply [86, Theorem 3.1] to our  $Z = (Z_i)_{i \in [i_1, i_2)}$  and  $\bar{Z} = (\bar{Z}_i)_{i \in [i_1, i_2)}$ , with  $f_i(s) - (s - s')$  being the driving function for  $Z_i$  and with  $f_i$  being the driving function for  $\bar{Z}_i$ . The result gives  $Z_i(Ts) \leq \bar{Z}_i(Ts)$  for all  $i \in [i_1, i_2 - 1)$  and  $s \in [s', \sigma]$ . Further, letting  $\bar{X}_{\mathbf{c}} := \sum_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})} \bar{X}_{(i)} / |\mathfrak{I}_{N,\xi}(\mathbf{c})|$  denote the center of mass of  $(\bar{X}_{(i)})_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})}$  and averaging (6.19) over  $i \in \mathfrak{I}_{N,\xi}(\mathbf{c})$ , we see that  $X_{\mathbf{c}} = \bar{X}_{\mathbf{c}}$ . This property, together with the comparison result of the gap processes, gives

$$\max \{ |X_i - X_{\mathbf{c}}| : i \in \mathfrak{I}_{N,\xi}(\mathbf{c}) \} \leq \max \{ |\bar{X}_{(i)} - \bar{X}_{\mathbf{c}}| : i \in \mathfrak{I}_{N,\xi}(\mathbf{c}) \}. \quad (6.22)$$

Next, let  $\bar{B}_i$ , for  $i \in \mathfrak{I}_{N,\xi}(\mathbf{c})$ , denote independent standard BMs, and consider

$$Y_i(Ts) := \bar{B}_i(Ts) - \bar{B}_i(Ts') + N(\phi_{\mathbf{c}} + v_{\mathbf{c}})(Ts - Ts') + X_i(Ts'), \quad (6.23)$$

for  $s \in [s', \sigma]$  and  $i \in \mathfrak{I}_{N,\xi}(\mathbf{c})$ . If we rank  $\{Y_i\}_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})}$ , the resulting process has the same law as  $(\bar{X}_{(i)})_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})}$ . This property can be checked from (6.19) by using the generators for example, which we omit. Given this property, defining the event

$$\mathcal{F}_N(r') := \{ |X_i(Ts) - X_{\mathbf{c}}(Ts)| < NT(r' + 3^\beta r), \forall i \in \mathfrak{I}_{N,\xi}(\mathbf{c}), \forall s \in [s', \sigma] \}, \quad (6.24)$$

and letting  $Y_c := \sum_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})} Y_i / |\mathfrak{I}_{N,\xi}(\mathbf{c})|$  denote the center of mass of  $(Y_i)_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})}$ , we have

$$\begin{aligned} & \mathbb{Q}[\mathcal{F}_N(r') | \vec{X}(Ts')] \\ & \geq \mathbb{Q}[|Y_i(Ts) - Y_c(Ts)| < NT(r' + 3^\beta r), \forall i \in \mathfrak{I}_{N,\xi}(\mathbf{c}), \forall s \in [s', \sigma]]. \end{aligned} \quad (6.25)$$

Next, we seek to bound  $|Y_i(Ts) - Y_c(Ts)|$ . Recall the expression of  $Y_i(Ts)$  for (6.23). Averaging (6.23) over  $i \in \mathfrak{I}_{N,\xi}(\mathbf{c})$  gives the expression  $Y_c(Ts) = (\mathfrak{I}_{N,\xi}(\mathbf{c}))^{-1} \sum_{j \in \mathfrak{I}_{N,\xi}(\mathbf{c})} (\overline{B}_j(Ts) - \overline{B}_j(Ts')) + N(\phi_c + v_c)(Ts - Ts') + X_c(Ts')$ . Inserting these expressions into  $|Y_i(Ts) - Y_c(Ts)|$ , and bound the result using the triangle inequality. Doing so gives

$$|Y_i(Ts) - Y_c(Ts)| \leq 2 \max_{j \in \mathfrak{I}_{N,\xi}(\mathbf{c})} |\overline{B}_j(Ts) - \overline{B}_j(Ts')| + |X_i(Ts') - X_c(Ts')|. \quad (6.26)$$

Under the condition  $\mathfrak{D}\text{ist}_{N,s'}(\vec{X}, \xi) \leq 3^\beta r$ , the last term in (6.26) is bounded by  $NT3^\beta r$ . Hence, under the condition  $\mathfrak{D}\text{ist}_{N,s'}(\vec{X}, \xi) \leq 3^\beta r$ ,

$$\max_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})} |Y_i(Ts) - Y_c(Ts)| \leq 2 \max_{i \in \mathfrak{I}_{N,\xi}(\mathbf{c})} |\overline{B}_i(Ts) - \overline{B}_i(Ts')| + NT3^\beta r. \quad (6.27)$$

Using this property in (6.25) gives

$$\begin{aligned} & \inf_{\mathfrak{D}\text{ist}_{N,s'}(\vec{X}, \xi) \leq 3^\beta r} \mathbb{P}[\mathcal{F}_N(r') | \vec{X}(Ts')] \\ & \geq \mathbb{P}[|\overline{B}_i(Ts) - \overline{B}_i(Ts')| < \tfrac{1}{2}NT r', \forall i \in \mathfrak{I}_{N,\xi}(\mathbf{c}), \forall s \in [s', \sigma]]. \end{aligned} \quad (6.28)$$

Based on what we obtained so far, we now complete the proof of (6.11). Combining (6.17) and (6.24) gives, under the condition  $\mathfrak{D}\text{ist}_{N,s'}(\vec{X}, \xi) \leq 3^\beta r$  and within  $s \in [s', \sigma]$ ,

$$\left| \frac{1}{NT} X_i(Ts) - \xi_{c_{N,\xi}(i)}(s) \right| \leq 2 \cdot 3^\beta r + \frac{1}{NT} \max_{i=1, \dots, N} |B_i(Ts) - B_i(Ts')| + r' \quad \text{on } \mathcal{F}_N(r'). \quad (6.29)$$

We will take  $r' \rightarrow 0$  later, so let us assume  $r' < r/2$  hereafter. Consider the event

$$\mathcal{E}_N(r') := \{|B_i(Ts) - B_i(Ts')| < NT r', \forall s \in [s', s''], \forall i = 1, \dots, N\}. \quad (6.30)$$

Under  $\mathcal{F}_N(r') \cap \mathcal{E}_N(r')$  and given the assumption  $r' < r/2$ , the right hand side of (6.29) is strictly less than  $3^{\beta+1}r$ . Hence, under the condition  $\mathfrak{D}\text{ist}_{N,s'}(\vec{X}, \xi) \leq 3^\beta r$ , we have  $\mathcal{F}_N(r') \cap \mathcal{E}_N(r') \subset \{\sigma > s''\} \subset \mathcal{C}_{N,\beta}[s', s'']$ . Now apply Lemma C.1(a) with the  $v_i = v_{c_{N,\xi}(i)}$  defined above (6.13) and with  $\mathcal{E} = \mathcal{F}_N(r') \cap \mathcal{E}_N(r')$ . In the result recognize that  $(s'' - s')(2N)^{-1} \sum_{i=1}^N v_{c_{N,\xi}(i)}^2 = \mathbb{I}_{[s', s'']}(s)$  with the aid of (6.10), and bound  $N^{-1} \sum_{i=1}^N |v_{c_{N,\xi}(i)}| r'$  by  $c(\xi) r'$ . Doing so gives

$$\begin{aligned} & \inf_{\mathfrak{Dist}_{N,s'}(\vec{X},\xi) \leq 3^\beta r} \left\{ \frac{1}{N^3 T} \log \mathbb{P} \left[ \mathcal{F}_N(r') \cap \mathcal{E}_N(r') \mid \vec{X}(Ts') \right] \right\} \\ & \geq \inf_{\mathfrak{Dist}_{N,s'}(\vec{X},\xi) \leq 3^\beta r} \frac{1}{N^3 T} \log \mathbb{Q} \left[ \mathcal{F}_N(r') \cap \mathcal{E}_N(r') \mid \vec{X}(Ts') \right] - c(\xi)r' - \mathbb{I}_{[s',s'']}(\xi). \end{aligned} \quad (6.31)$$

On the left hand side of (6.31), bound the probability from above by  $\mathbb{P}[\mathcal{C}_{N,\beta}[s',s''] \mid \vec{X}(Ts')]$ , and send  $N \rightarrow \infty$  on both sides. From (6.28) and (6.30), it is not hard to check that the first term on the right hand side of (6.31) converges to 0 as  $N \rightarrow \infty$ . Finally sending  $r' \rightarrow 0$  gives the desired result (6.11).

**Step 2: analysis within a transition segment.** Let us set up the notation and state the goal of this step. Fix a transition segment  $[s',s'']$ , consider the line (in the  $(s,x)$  plane) that connects  $(s',\xi_{c_{N,\xi}(i)}(s'))$  and  $(s'',\xi_{c_{N,\xi}(i)}(s''))$ , and let

$$w_i = w_i[s',s''] := \frac{\xi_{c_{N,\xi}(i)}(s'') - \xi_{c_{N,\xi}(i)}(s')}{s'' - s'} = N \int_{(i-1)/N}^{i/N} da \frac{\mathfrak{X}[\xi(s'')] - \mathfrak{X}[\xi(s')]}{s'' - s'} \quad (6.32)$$

be the velocity when traveling along the line. Consider

$$\mathcal{C}'_{N,\beta}[s',s''] := \left\{ \left| \frac{1}{NT} X_i(Ts) - \xi_{c_{N,\xi}(i)}(s') - (s - s')w_i \right| \leq 3^\beta r, \forall s \in [s',s''], \forall i \right\}, \quad (6.33)$$

$$\text{Cost}_{[s',s'']}(\lambda,\lambda') := (s'' - s') \int_0^1 da \left( \frac{|\mathfrak{X}[\lambda'] - \mathfrak{X}[\lambda]|}{s'' - s'} + \frac{1}{2} \right)^2. \quad (6.34)$$

By (6.7), it is not hard to check from (6.33) that

$$\mathcal{C}'_{N,\beta}[s',s''] \subset \{ \mathfrak{Dist}_{N,[s',s'']}(\vec{X},\xi) \leq 3^{\beta+1}r \}. \quad (6.35)$$

The goal of this step is to prove

$$\liminf_{N \rightarrow \infty} \inf \left\{ \frac{1}{N^3 T} \log \mathbb{P} \left[ \mathcal{C}'_{N,\beta+1}[s',s''] \mid \vec{X}(Ts') \right] \right\} \geq -\text{Cost}_{[s',s'']}(\xi(s'),\xi(s'')), \quad (6.36)$$

where the infimum runs over  $\mathfrak{Dist}_{N,s'}(\vec{X},\xi(s')) \leq 3^\beta r$ , just like in Step 1.

To prove (6.36), consider the law  $\mathbb{Q}$  under which  $dX_i = Nw_i ds + dB_i$ . Apply Lemma C.1(b) with  $v_i = w_i$ ,  $J_i = \{1, \dots, N\}$ ,  $\mathcal{E} = \mathcal{C}'_{N,\beta}[s',s'']$ , and  $p = 2$ . Doing so gives

$$\begin{aligned} & \frac{1}{N^3 T} \log \mathbb{P} \left[ \mathcal{C}'_{N,\beta+1}[s',s''] \mid \vec{X}(Ts') \right] \\ & \geq \frac{2}{N^3 T} \log \mathbb{Q} \left[ \mathcal{C}'_{N,\beta+1}[s',s''] \mid \vec{X}(Ts') \right] - \frac{s'' - s'}{N} \sum_{i=1}^N \left( |w_i| + \frac{1}{2} \right)^2. \end{aligned} \quad (6.37)$$

It is readily checked that the first term on the right hand side of (6.37) tends to 0 as  $N \rightarrow \infty$ , uniformly over  $\mathfrak{Dist}_{N,s'}(\vec{X}, \xi(s')) \leq 3^\beta r$ . To bound the last term in (6.37), use (6.32) and Jensen's inequality to get

$$\left(|w_i| + \frac{1}{2}\right)^2 \leq N \int_{(i-1)/N}^{i/N} da \left( \frac{|\mathfrak{X}[\xi(s'')] - \mathfrak{X}[\xi(s')]|}{s'' - s'} + \frac{1}{2} \right)^2. \quad (6.38)$$

Applying  $(s'' - s')N^{-1} \sum_{i=1}^N$  on both sides of (6.38) gives  $(s'' - s')N^{-1} \sum_i (|w_i| + 1/2)^2 \leq \text{Cost}_{[s', s'']}(\xi(s'), \xi(s''))$ . The desired result (6.36) hence follows.

**Step 3: Combining Steps 1–2 to complete the proof.** Apply (6.11) and (6.36) in alternating order over the segments  $[s_0, s_1 - \delta]$ ,  $[s_1 - \delta, s_1 + \delta]$ ,  $\dots$ ,  $[s_{\ell-1} + \delta, s_\ell]$ , with  $\beta = 0, 1, \dots, 2\ell - 2$ . By (6.9) and (6.33),  $\mathcal{C}_{N,1}[s_0, s_1 - \delta] \subset \{\mathfrak{Dist}_{N,s_1-\delta}(\vec{X}, \xi) \leq 3^1\}$ ,  $\mathcal{C}'_{N,2}[s_1 - \delta, s_1 + \delta] \subset \{\mathfrak{Dist}_{N,s_1+\delta}(\vec{X}, \xi) \leq 3^2\}$ ,  $\dots$ , so the resulting bounds can be concatenated. The result gives

$$\begin{aligned} \liminf_{N \rightarrow \infty} \inf_{\mathfrak{Dist}_{N,0}(\vec{X}, \xi) \leq r} & \left\{ \frac{1}{N^3 T} \log \mathbb{P}[\mathfrak{Dist}_{N,[0,t]}(\vec{X}, \xi) \leq 3^{2\ell-1} r \mid \vec{X}(0)] \right\} \\ & \geq - \left( \sum \mathbb{I}_{[s', s'']}(\xi) + \sum \text{Cost}_{[s', s'']}(\xi(s'), \xi(s'')) \right), \end{aligned} \quad (6.39)$$

where the first sum runs over all linear segments, and the second runs over all transition segments. The first sum is at most  $\mathbb{I}(\xi)$ , while the second sum tends to zero as  $\delta$  tends to zero, as is readily checked from (6.34) and the assumption that  $\partial_s \mathfrak{X}[\mu] \in \mathcal{L}^2([0, t] \times [0, 1])$ . Recall that  $\delta$  is chosen so that (6.7)–(6.8) hold. With  $\xi$  being PL-clustering, we can choose  $\delta$  such that  $\delta \leq c(\xi)r$ . This way,  $\delta \rightarrow 0$  as  $r \rightarrow 0$ . Sending  $r \rightarrow 0$  in (6.39) gives the desired result (6.6).  $\square$

## 6.2. Proving Proposition 6.1 under Assumption 6.3

Based on the results in Section 6.1, here we prove Proposition 6.1 under an additional assumption. Identify  $\{\lambda \in \mathcal{P}(\mathbb{R}) : \text{supp}(\lambda) \subset [-b, b]\}$  with  $\mathcal{P}[-b, b]$ . Namely, view those measures supported in  $[-b, b]$  as measures on  $[-b, b]$ .

**Assumption 6.3.** There exists a  $b < \infty$  such that  $\{\mu_N(0)\}_{N \in \mathbb{Z}_{>0}}, \{\mu(s)\}_{s \in [0,t]} \subset \mathcal{P}[-b, b]$ .

Namely, we assume that  $\mu_N(0)$  and  $\mu(s)$  have a compact support, uniformly over  $N \in \mathbb{Z}_{>0}$  and  $s \in [0, t]$ . Fixing a  $\mu$  and  $\mu_N(0)$  that satisfy  $\mathbb{I}(\mu) < \infty$ ,  $\mu_N(0) \Rightarrow \mu(0)$ , and Assumption 6.3, we seek to prove (6.1) for this  $\mu$  and  $\mu_N(0)$ .

The proof requires an approximation tool. We state it here and put its proof in Appendix B: For any  $[s', s''] \subset [0, t]$ ,



$$\inf \left\{ \text{dist}_{[s', s'']}(\xi, \mu) + |\mathbb{I}_{[s', s'']}(\xi) - \mathbb{I}_{[s', s'']}(\mu)| : \right. \\ \left. \xi \text{ is PL-clustering, } \text{supp}(\xi(s)) \subset [-b, b], \forall s \in [s', s''] \right\} = 0. \quad (6.40)$$

Let us outline the proof. We will construct a small  $s_0$  and use (6.40) to construct a PL-clustering  $\xi$  that approximates  $\mu$ , and perform analysis over  $[0, s_0]$  and  $[s_0, t]$  separately.

Let us construct  $s_0$  and  $\xi$ . Fix an  $\varepsilon > 0$ . By the assumption  $\mathbb{I}(\mu) < \infty$ ,  $\partial_s \mathfrak{X}[\mu] \in \mathcal{L}^2([0, t] \times [0, 1])$ . Using this property and the time continuity of  $\mu$  to find an  $s_0$  such that  $\text{Cost}_{[0, s_0]}(\mu(0), \mu(s_0)) < \varepsilon$  and  $\sup_{[0, s_0]} \text{dist}(\mu(s), \mu(s_0)) < \varepsilon$ . Turning to the construction of  $\xi$ , we begin by noting that  $\text{Cost}_{[0, s_0]} : (\mathcal{P}[-b, b])^2 \rightarrow \mathbb{R}$  is continuous, which is not hard to verify from (3.4) and (6.34). Granted this continuity, use (6.40) for  $[s', s''] = [s_0, t]$  to obtain a PL-clustering  $\xi$  such that  $\text{dist}_{[s_0, t]}(\xi, \mu) < \varepsilon$ ,  $|\mathbb{I}_{[s_0, t]}(\xi) - \mathbb{I}_{[s_0, t]}(\mu)| < \varepsilon$ , and

$$\text{Cost}_{[0, s_0]}(\mu(0), \xi(s_0)) < \varepsilon. \quad (6.41)$$

Given that  $\text{dist}(\mu_N(0), \mu(0)) \rightarrow 0$ , for all  $N$  large enough,  $\text{Cost}_{[0, s_0]}(\mu_N(0), \xi(s_0)) < \varepsilon$ .

Let us perform analysis over  $[0, s_0]$  and  $[s_0, t]$ , beginning with the former. We seek to adapt the argument in Step 2 in Section 6.1. More precisely, we seek to apply the argument with  $w_i[s', s'']$  (defined in (6.32)) replaced by  $w_i(s_0) := (\xi_{\epsilon_N, \xi(i)}(s_0) - X_i(0)/(NT))/s_0$  and  $\mathcal{C}'_{N, \beta}[s', s'']$  (defined in (6.33)) replaced by

$$\mathcal{C}''_{N, r}[0, s_0] := \left\{ \left| \frac{1}{NT} X_i(Ts) - \frac{1}{NT} X_i(0) - w_i(s_0)s \right| < r, \forall s \in [0, s_0], \forall i \right\}. \quad (6.42)$$

Applying the argument in Step 2 in Section 6.1 with these adaptations gives

$$\liminf_{r \rightarrow 0} \liminf_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P}[\mathcal{C}''_{N, r}[0, s_0]] \geq -\limsup_{N \rightarrow \infty} \text{Cost}_{[0, s_0]}(\mu_N(0), \xi(s_0)). \quad (6.43)$$

Unlike in (6.36), we need not impose any constraint on  $\vec{X}(0)$  in (6.43). This is because when defining the event in (6.42), we use the reference point  $X_i(0)/(NT)$  instead of  $\xi_{\epsilon_N, \xi(i)}(0)$ ; compared with (6.33). Given that  $\text{dist}_{[0, s_0]}(\mu(s), \mu(0)) < \varepsilon$ ,  $\text{dist}(\xi(s_0), \mu(s_0)) < \varepsilon$ , and  $\text{dist}(\mu_N(0), \mu(0)) \rightarrow 0$ , for all  $N$  large enough we have  $\mathcal{C}''_{N, r}[0, s_0] \subset \{\text{dist}_{[0, s_0]}(\mu_N, \mu) < r + 2\varepsilon\}$ . Also, note that

$$\mathcal{C}''_{N, r}[0, s_0] \subset \{\mathfrak{D}\text{ist}_{N, s_0}(\vec{X}, \xi) \leq r\}. \quad (6.44)$$

Move on to  $[s_0, t]$ . Apply Proposition 6.2 with  $[0, t] \mapsto [s_0, t]$ . In the result, use (6.5) and  $\text{dist}_{[s_0, t]}(\xi, \mu) < \varepsilon$ . Doing so gives

$$\liminf_{r \rightarrow 0} \liminf_{N \rightarrow \infty} \inf \left\{ \frac{1}{N^3 T} \log \mathbb{P}[\text{dist}_{[s_0, t]}(\mu_N, \xi) < \varepsilon + cr \mid \vec{X}(Ts_0)] \right\} \geq -\mathbb{I}_{[s_0, t]}(\mu), \quad (6.45)$$

where the infimum runs over  $\mathfrak{D}\text{ist}_{N, s_0}(\vec{X}, \xi) \leq r$ . Further,

$$-\mathbb{I}_{[s_0, t]}(\mu) \geq -\mathbb{I}_{[0, t]}(\mu) = -\mathbb{I}(\mu). \quad (6.46)$$

Combine (6.43) and (6.45) with the aid of (6.44). In the result, use (6.41) and (6.46), and send  $\varepsilon \rightarrow 0$ . Doing so gives (6.1) under Assumption 6.3.

### 6.3. Removing Assumption 6.3

Having proven (6.1) under Assumption 6.3, we explain how the same result follows without the assumption. Taking any  $\mu \in \mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$  with  $\mathbb{I}(\mu) < \infty$  and  $\mu(0) = \mu_{sc}$ , we seek to show that (6.1) holds for  $\mu_N$  and this  $\mu$ .

We begin by truncating  $\mu$  and  $\mu_N$  so that the result satisfies Assumption 6.3. Fix a sequence  $1/2 > r_1 > r_2 > \dots \rightarrow 0$  such that  $\mathfrak{X}[\mu_{sc}]$  is continuous at every  $r_k$  and  $1 - r_k$ . Fix any  $k$ . Recall the scaling operator  $\mathfrak{S}_m$  from Section 3.1. Apply the procedure in Section 3.4 with  $(m_1, m_2, m_3) = (r_k, 1 - 2r_k, r_k)$  to get  $\mu_1, \mu_2, \mu_3$  and take  $\nu := \mu_2$  as the truncated deviation. By Lemma 3.2(b), the truncated deviation  $\nu$  satisfies Assumption 6.3. Next, truncate the empirical measure similarly. Recall that  $X_1(0) \leq \dots \leq X_N(0)$ . Set  $N_- := \lceil Nr_k \rceil$ ,  $N_+ := \lfloor N(1 - r_k) \rfloor$ ,  $N' := N_+ - N_-$ , and  $\nu_N(s) := \frac{1}{N} \sum_{N_- \leq i \leq N_+} \delta_{X_i(Ts)/(NT)}$ . Note that  $\mu_N(0) \Rightarrow \mu_{sc}$  is equivalent to  $\mathfrak{X}[\mu_N(0)]$  converging to  $\mathfrak{X}[\mu_{sc}]$  everywhere the latter is continuous. Recall that  $\mathfrak{X}[\mu_{sc}]$  is continuous at each  $r_k$  and  $1 - r_k$ . These properties imply, as  $N \rightarrow \infty$ ,  $\nu_N(0) \Rightarrow \nu(0)$ ,  $X_{N_-}(0)/(NT) \rightarrow \mathfrak{X}[\mu_{sc}](r_k)$ , and  $X_{N_+}(0)/(NT) \rightarrow \mathfrak{X}[\mu_{sc}](1 - r_k)$ . The last two convergences show that  $\nu_N(0)$  is supported in an  $N$ -independent bounded interval.

Next, we claim that

$$\text{dist}_{[0, t]}(\mu_N, \mu) \leq \text{dist}_{[0, t]}(\nu_N, \nu) \leq c \cdot (r_k + \frac{1}{N}). \quad (6.47)$$

Recall from Section 3.4 that the graph of  $F[\nu(s)] = F[\mu_2(s)]$  is obtained by taking the graph of  $F[\mu(s)]$  between the levels  $r_k$  and  $1 - r_k$  and shift the graph down so that the lower level is at 0. The same property holds for  $(F[\nu_N(s)], F[\mu_N(s)])$  with  $(r_k, 1 - r_k) \mapsto (N_-/N, 1 - N_+/N)$ . By construction,  $N_-/N$  and  $N_+/N$  differ from  $r_k$  by at most  $1/N$ . Using these properties gives

$$\int_{-k}^k da |F[\mu_N(s)] - F[\mu(s)]| \leq \int_{-k}^k da \left( |F[\nu_N(s)] - F[\nu(s)]| + c \cdot (r_k + \frac{1}{N}) \right). \quad (6.48)$$

Combining (6.48) and (3.4) gives (6.47).

Next, we would like to apply the result of Section 6.2 with  $\mu_N$  and  $\mu$  being replaced by  $\nu_N$  and  $\nu$ . However, note that the particles in  $\nu_N$  are not autonomous and feel the pulling from particles outside of  $\nu_N$ . Namely, after being restricted to  $i \in [N_-, N_+]$ , the system of equations (2.2) still contains terms  $\frac{1}{2} \text{sgn}(X_j - X_i)$  from outside of the system  $j \notin [N_-, N_+]$ . To resolve this issue, consider the law  $\mathbb{Q}$  under which

$$dX_i = \sum_{j \in [N_-, N_+]} \frac{1}{2} \operatorname{sgn}(X_j - X_i) ds + dB_i, \quad \text{for all } i \in [N_-, N_+], \quad (6.49)$$

$$dX_i = \sum_{j=1}^N \frac{1}{2} \operatorname{sgn}(X_j - X_i) ds + dB_i, \quad \text{for all } i \notin [N_-, N_+]. \quad (6.50)$$

Under  $\mathbb{Q}$ , the truncated empirical measure  $\nu_N$  evolves autonomously. Apply Lemma C.1(b) with  $[s', s''] = [0, t]$ ,  $v_i = 0$ ,  $J_i = \{1, \dots, N_- - 1, N_+ + 1, \dots, N\}$  for  $i \in [N_-, N_+]$ , and  $J_i = \emptyset$  for  $i \notin [N_-, N_+]$ . In the result, bound  $|J_i|/(2N)$  by  $cr_k$ . Doing so gives

$$\frac{1}{N^3 T} \log \mathbb{P} [\operatorname{dist}_{[0, t]}(\nu_N, \nu) < r] \geq \frac{p}{N^3 T} \log \mathbb{Q} [\operatorname{dist}_{[0, t]}(\nu_N, \nu) < r] - \frac{pt(cr_k)^2}{2(p-1)}. \quad (6.51)$$

Next, apply the result of Section 6.2 with  $(\mu_N, \mu, \mathbb{P}) \mapsto (\nu_N, \nu, \mathbb{Q})$ . Note that, even though  $\nu_N$  and  $\nu$  do not have unit mass, by the scaling arguments in Section 3.1, the result still applies. Doing so gives, for any  $p > 1$  and  $k \in \mathbb{Z}_{>0}$ ,

$$\liminf_{r \rightarrow 0} \liminf_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P} [\operatorname{dist}_{[0, t]}(\nu_N, \nu) < r] \geq -p\mathbb{I}(\nu) - \frac{pt(cr_k)^2}{2(p-1)}. \quad (6.52)$$

We are now ready to complete the proof. So far we have omitted most  $k$  dependence. Restore it by writing  $\nu = \nu_k$ ,  $\nu_N = \nu_{N,k}$ , and  $N' = N'_k$ . By the construction of  $\nu = \nu_k$  and the assumption  $\mathbb{I}(\mu) < \infty$ , it is not hard to check that  $\mathbb{I}(\nu_k) \rightarrow \mathbb{I}(\mu)$  as  $k \rightarrow \infty$ . In (6.52), send  $k \rightarrow \infty$  with the aid of this property and (6.47). Finally sending  $p \rightarrow 1$  completes the proof.

## 7. Applications to the moment Lyapunov exponents

As was explained in Section 3.1, we consider  $\mathfrak{m} = 1$  only.

### 7.1. Proof of Theorem 2.4

We begin with some notation. Let  $\mathbf{x}_c$ ,  $\mathbf{x}'$ ,  $\mathbf{m}_c$ , and the optimal deviation  $\xi = \sum_{c=1}^n \mathbf{m}_c \delta_{\xi_c}$  be as in Section 2.3. Recall branches from there and recall that  $\mathfrak{B}$  denotes the set of branches. Let  $\mathbf{m}_b := \sum_{c \in b} \mathbf{m}_c$  be the total mass within a given branch  $b$ . By definition, branches are disjoint intervals in  $\{1, \dots, n\}$ , so we can order them, and we use  $<$  and  $\leq$  to denote the ordering. For a given branch  $b$ , we write  $b \ominus 1$  and  $b \oplus 1$  for the respective branches that precedes and succeeds  $b$  in  $\mathfrak{B}$ .

To facilitate the proof, let us prepare a few properties of the optimal clusters  $\xi_1, \dots, \xi_n$ . From their definition given after Theorem 2.4, it is not hard to check that

$$\dot{\xi}_c = \Phi[\xi](\xi_c) + v_b, \quad \text{for all } c \in b \text{ and for all } s \in [0, t] \text{ except when a merge happens,} \quad (7.1)$$

where  $v_b$  is given in (2.24). Recall also that those  $\xi_c$  and  $\xi_{c'}$  belonging to different branches do not meet within  $(0, t)$ . Using this property to take the mass-weighted average of (7.1) over  $c \in b$  gives

$$\dot{\xi}_b = \phi_b + v_b, \quad \phi_b := \frac{1}{2}(\dots - m_{b \ominus 1} + m_{b \oplus 1} + \dots). \quad (7.2)$$

This equation shows that  $\dot{\xi}_b$  is a constant. Combining (7.1)–(7.2) gives  $\dot{\xi}_c - \Phi[\xi](\xi_c) = \dot{\xi}_b - \phi_b$ , for all  $c \in b$ . Inserting this identity into (6.10) for  $\nu = \xi$  and  $[s', s''] = [0, t]$  gives

$$\mathbb{I}(\xi) = \sum_{b \in \mathfrak{B}} t \frac{m_b}{2} (\dot{\xi}_b - \phi_b)^2. \quad (7.3)$$

We now begin the proof, starting with a reduction. Set the starting and ending conditions to be  $\mu_{sc} := \sum_c m_c \delta_{x_c}$  and  $\mu_{ec} := \delta_{x'}$ , with  $1 = m_1 + \dots + m_n$ . By (2.16),

$$\begin{aligned} L_{\text{SHE}}(x' \xrightarrow{t} (\bar{x}, \bar{m})) \\ = \frac{t}{24} + \langle \mu_{sc}^{\otimes 2}, \frac{1}{4}|x - x'| \rangle - \langle \mu_{ec}^{\otimes 2}, \frac{1}{4}|x - x'| \rangle - \inf \{ \mathbb{I}(\mu) : \mu(0) = \mu_{sc}, \mu(t) = \mu_{ec} \}. \end{aligned} \quad (7.4)$$

Take any  $\mu \in \mathcal{C}([0, t], \mathcal{P}(\mathbb{R}))$  with  $\mu(0) = \mu_{sc}$  and  $\mu(t) = \mu_{ec}$ . Given (7.4), it suffices to prove that  $\mathbb{I}(\mu) \geq \mathbb{I}(\xi)$  and that the equality holds only if  $\mu = \xi$ . Without loss of generality, assume  $\mathbb{I}(\mu) < \infty$ , which implies  $\partial_s \mathfrak{X}[\mu] \in \mathcal{L}^2([0, t] \times [0, 1])$ .

**Step 1: proving that  $\mathbb{I}(\mu) \geq \mathbb{I}(\xi)$ .** We begin with some notation. First, write  $\mathfrak{X}[\mu(s)](a) = \mathfrak{X}(s, a)$  and  $\Phi[\mu(s)](\mathfrak{X}(s, a)) = \Phi_{\mathfrak{X}}(s, a)$  to simplify notation. Next, apply the procedure in Section 3.4 with  $(m_1, m_2, \dots) = (m_b)_{b \in \mathfrak{B}}$  to get  $(\mu_b)_{b \in \mathfrak{B}}$ , where the branches are ordered as described previously. Let  $M_b := \sum_{b' \leq b} m_{b'}$  be the cumulative mass.

Let us derive a lower bound on  $\mathbb{I}(\mu)$ . Refer to (2.12) for the definition of  $\mathbb{I}(\mu)$  and divide the integral  $\int_0^1 da$  into  $\int_{M_{b \ominus 1}}^{M_b} da$ ,  $b \in \mathfrak{B}$ , with the convention that  $M_{(\text{first cluster}) \ominus 1} := 0$ . Note that this procedure is equivalent to the procedure of dividing  $\mu$  in the previous paragraph. In each of the resulting integral, multiply and divide by  $tm_b$  and apply Jensen's inequality to get

$$\mathbb{I}(\mu) = \sum_{b \in \mathfrak{B}} t \frac{m_b}{2} \int_0^t \frac{ds}{t} \int_{M_{b \ominus 1}}^{M_b} \frac{da}{m_b} \left( \partial_s \mathfrak{X} - \Phi_{\mathfrak{X}} \right)^2 \quad (7.5)$$

$$\geq \sum_{b \in \mathfrak{B}} t \frac{m_b}{2} \left( \int_0^t \frac{ds}{t} \int_{M_{b \ominus 1}}^{M_b} \frac{da}{m_b} \partial_s \mathfrak{X} - \int_0^t \frac{ds}{t} \int_{M_{b \ominus 1}}^{M_b} \frac{da}{m_b} \Phi_{\mathfrak{X}} \right)^2. \quad (7.6)$$

Next we derive an expression for the first integral in (7.6). Evaluate the integral over  $[0, t]$  by the fundamental theorem of calculus, recognize the result as  $\langle \mu_b(s), x \rangle|_0^t / (tm_b)$ ,

and use  $\mu(0) = \xi(0)$  and  $\mu(t) = \xi(t)$ . Recall the center of mass  $\xi_b := \sum_{c \in b} (m_c \xi_c) / m_b$ . The resulting expression is hence  $\xi_b(s)|_0^t / t$ . As shown in (7.2),  $\dot{\xi}_b$  is a constant, so

$$\int_0^t \frac{ds}{t} \int_{M_{b \ominus 1}}^{M_b} \frac{da}{m_b} \partial_s \mathfrak{X} = \dot{\xi}_b \quad (7.7)$$

Next we derive an expression for the last integral in (7.6). Recall from (3.8) that the integrand takes two different forms depending on whether  $a_- = a_+$  or  $a_- < a_+$ , where  $a_{\pm} = a_{\pm}[\mu(s)](a)$  is defined in (3.7). This dichotomy is particularly relevant at  $a = M_{b \ominus 1}$  and  $a = M_b$ , which are the boundaries of the last integral in (7.6). With this in mind, set  $M_b^{\pm} = M_b^{\pm}(s) := a_{\pm}[\mu(s)](M_b)$ , and note that

$$M_b^- \leq M_b \leq M_b^+. \quad (7.8)$$

Divide the integral over  $[M_{b \ominus 1}, M_b]$  into integrals over  $[M_{b \ominus 1}, M_{b \ominus 1}^+]$ ,  $[M_{b \ominus 1}^+, M_b^-]$ , and  $[M_b^-, M_b]$  and use (3.8) to evaluate the integrals:

$$\int_{M_{b \ominus 1}}^{M_b} \frac{da}{m_b} \Phi_{\mathfrak{X}} = \frac{1}{m_b} \left( A + \int_{M_{b \ominus 1}^+}^{M_b^-} da \left( \frac{1}{2} - a \right) + A' \right), \quad (7.9)$$

where  $A := \frac{1}{2}(M_{b \ominus 1}^+ - M_{b \ominus 1})(1 - M_{b \ominus 1}^+ - M_{b \ominus 1}^-)$  and  $A' := \frac{1}{2}(M_b - M_b^-)(1 - M_b^+ - M_b^-)$ . Recall  $\phi_b$  from (7.2) and write it as

$$\phi_b = \frac{1}{m_b} \int_{M_{b \ominus 1}}^{M_b} da \left( \frac{1}{2} - a \right) = \frac{1}{m_b} \left( A'' + \int_{M_{b \ominus 1}^+}^{M_b^-} da \left( \frac{1}{2} - a \right) + A''' \right), \quad (7.10)$$

where  $A'' := \frac{1}{2}(M_{b \ominus 1}^+ - M_{b \ominus 1})(1 - M_{b \ominus 1}^+ - M_{b \ominus 1})$  and  $A''' := \frac{1}{2}(M_b - M_b^-)(1 - M_b - M_b^-)$ . Subtract (7.10) from (7.9) and simplify  $(A - A'')$  and  $(A' - A''')$  in the result. Doing so gives

$$\int_{M_{b \ominus 1}}^{M_b} \frac{da}{m_b} \Phi_{\mathfrak{X}} = \phi_b + \frac{1}{m_b} D_{b \ominus 1} - \frac{1}{m_b} D_b, \quad (7.11)$$

where

$$D_b := \frac{1}{2}(M_{b \ominus 1}^+ - M_b)(M_b - M_b^-), \quad (7.12)$$

with the convention that  $D_{(\text{first branch}) \ominus 1} := 0$ .

We are now ready to complete the proof of Step 1. Insert (7.7) and (7.11) into (7.6), expand the resulting square, and compare the result with (7.3). Doing so gives

$$\mathbb{I}(\mu) \geq \mathbb{I}(\xi) + \sum_{\mathbf{b} \in \mathfrak{B}} \frac{\mathfrak{m}_{\mathbf{b}}}{2} \left( \int_0^{\mathfrak{t}} \frac{ds}{\mathfrak{t}} \frac{D_{\mathbf{b}} - D_{\mathbf{b} \ominus 1}}{\mathfrak{m}_{\mathbf{b}}} \right)^2 + \sum_{\mathbf{b} \in \mathfrak{B}} (\dot{\xi}_{\mathbf{b}} - \phi_{\mathbf{b}}) \int_0^{\mathfrak{t}} ds (D_{\mathbf{b}} - D_{\mathbf{b} \ominus 1}). \quad (7.13)$$

Rewrite the last sum as  $\sum_{\mathbf{b} \in \mathfrak{B}} \int_0^{\mathfrak{t}} ds (-\dot{\xi}_{\mathbf{b} \oplus 1} + \phi_{\mathbf{b} \oplus 1} + \dot{\xi}_{\mathbf{b}} - \phi_{\mathbf{b}}) D_{\mathbf{b}}$ , with the convention that  $\dot{\xi}_{(\text{last branch}) \oplus 1} := 0$  and  $\phi_{(\text{last branch}) \oplus 1} := 0$ . We arrive at

$$\mathbb{I}(\mu) \geq \mathbb{I}(\xi) + \sum_{\mathbf{b} \in \mathfrak{B}} \frac{1}{2\mathfrak{m}_{\mathbf{b}}} \left( \int_0^{\mathfrak{t}} ds (D_{\mathbf{b}} - D_{\mathbf{b} \ominus 1}) \right)^2 \quad (7.14a)$$

$$+ \sum_{\mathbf{b} \in \mathfrak{B}} \int_0^{\mathfrak{t}} ds (-\dot{\xi}_{\mathbf{b} \oplus 1} + \dot{\xi}_{\mathbf{b}} + \phi_{\mathbf{b} \oplus 1} - \phi_{\mathbf{b}}) D_{\mathbf{b}}. \quad (7.14b)$$

By (7.2),  $(-\dot{\xi}_{\mathbf{b} \oplus 1} - \dot{\xi}_{\mathbf{b}} + \phi_{\mathbf{b} \oplus 1} - \phi_{\mathbf{b}}) = -v_{\mathbf{b} \oplus 1} + v_{\mathbf{b}}$ . This quantity is strictly positive, as is readily checked from (2.24). Also,  $D_{\mathbf{b}}$  is nonnegative by (7.8) and (7.12). Hence  $\mathbb{I}(\mu) \geq \mathbb{I}(\xi)$ .

**Step 2: proving that  $\mathbb{I}(\mu) = \mathbb{I}(\xi)$  implies  $\mu = \xi$ .** The strategy is to extract information on  $\mu$  from the condition  $\mathbb{I}(\mu) = \mathbb{I}(\xi)$ . Set  $\mathfrak{X}_{\mathbf{b}}(s) := \mathfrak{X}[\mu(s)](M_{\mathbf{b}})$  and note that  $\mu_{\mathbf{b}}(s)$  is supported in  $[\mathfrak{X}_{\mathbf{b} \ominus 1}(s), \mathfrak{X}_{\mathbf{b}}(s)]$ . As explained after (7.14),  $(\dot{\xi}_{\mathbf{b} \oplus 1} - \dot{\xi}_{\mathbf{b}} - \phi_{\mathbf{b} \oplus 1} + \phi_{\mathbf{b}}) > 0$ ; recall that  $D_{\mathbf{b}} \geq 0$  by (7.8) and (7.12). Using these properties and the condition  $\mathbb{I}(\mu) = \mathbb{I}(\xi)$  in (7.14b) gives  $D_{\mathbf{b}} = 0$  for all  $\mathbf{b}$ . This property together with (7.8) and (7.12) forces  $M_{\mathbf{b}}^- = M_{\mathbf{b}}^+$ . This being true for all  $\mathbf{b}$  implies that  $\mu_{\mathbf{b}}(s)$  has no atoms at  $\mathfrak{X}_{\mathbf{b} \ominus 1}(s)$  and  $\mathfrak{X}_{\mathbf{b}}(s)$ .

We continue to extract information from the condition  $\mathbb{I}(\mu) = \mathbb{I}(\xi)$ . The condition forces the inequality in (7.6) to be an equality, which in turn forces  $\partial_s \mathfrak{X} - \Phi_{\mathfrak{X}}$  to be a constant Lebesgue a.e. on  $[0, \mathfrak{t}] \times [M_{\mathbf{b} \ominus 1}, M_{\mathbf{b}}]$ . By (7.7), (7.11), and  $D_{\mathbf{b}} = 0$ , the constant is  $\dot{\xi}_{\mathbf{b}} - \phi_{\mathbf{b}}$ . Hence

$$\partial_s \mathfrak{X} - \Phi_{\mathfrak{X}} = \dot{\xi}_{\mathbf{b}} - \phi_{\mathbf{b}}, \quad \text{Lebesgue a.e. on } [0, \mathfrak{t}] \times [M_{\mathbf{b} \ominus 1}, M_{\mathbf{b}}]. \quad (7.15)$$

We seek to “localize” (7.15) onto  $\mu_{\mathbf{b}}$ . More precisely, we seek to rewrite (7.15) in terms of  $\mathfrak{X}[\mu_{\mathbf{b}}]$  and  $\Phi[\mu_{\mathbf{b}}]$ . First, by the construction of  $\mu_{\mathbf{b}}$  in the first paragraph in Step 1,  $\mathfrak{X}[\mu_{\mathbf{b}}(s)](a + M_{\mathbf{b} \ominus 1}) = \mathfrak{X}(s, a)$  for all  $a \in [0, \mathfrak{m}_{\mathbf{b}}]$ , so  $(\partial_s \mathfrak{X})|_{(s, a + M_{\mathbf{b} \ominus 1})} = (\partial_s \mathfrak{X}[\mu_{\mathbf{b}}])|_{(s, a)}$ . Next, the definition (2.8) of  $\Phi[\cdot]$  gives

$$\Phi[\mu(s)](x) = \Phi[\mu_{\mathbf{b}}(s)](x) - \frac{1}{2} \sum_{\mathbf{b}' < \mathbf{b}} \langle \mu_{\mathbf{b}'}(s), \mathbf{1}_{(-\infty, x)} \rangle + \frac{1}{2} \sum_{\mathbf{b}' > \mathbf{b}} \langle \mu_{\mathbf{b}'}(s), \mathbf{1}_{(x, \infty)} \rangle. \quad (7.16)$$

Since  $\mu_b(s)$  is supported in  $[\mathfrak{X}_{b\ominus 1}(s), \mathfrak{X}_b(s)]$ , we consider those  $x$  in this interval only. As mentioned in the previous paragraph,  $\mu_{b'}(s)$  has not atoms at  $\mathfrak{X}_{b'\ominus 1}(s)$  and  $\mathfrak{X}_{b'}(s)$ , for all  $b'$ . Using this property in (7.16) simplifies the last two sums together into  $\phi_b$ , defined in (7.2). We now rewrite (7.15) as

$$\partial_s \mathfrak{X}[\mu_b] - \Phi[\mu_b](\mathfrak{X}[\mu_b]) = \dot{\xi}_b, \quad \text{Lebesgue a.e. on } [0, t] \times [0, \mathfrak{m}_b]. \quad (7.17)$$

Further using (7.2) gives

$$\partial_s \mathfrak{X}[\mu_b] - \Phi[\mu_b](\mathfrak{X}[\mu_b]) = \phi_b + v_b, \quad \text{Lebesgue a.e. on } [0, t] \times [0, \mathfrak{m}_b]. \quad (7.18)$$

Apply the procedure in Section 3.4 to divide  $\mu_b$  according to the masses  $(\mathfrak{m}_c)_{c \in b}$  to get  $(\mu_c)_{c \in b}$ .

Taking any  $c \in b$ , let us show that  $\mu_c(s)$  is supported at a single point. This amounts to showing that  $\mathfrak{X}[\mu_b(s)]|_{a \in U_c}$  is a constant, where  $U_c := (\sum_{c' \in b, c' < c} \mathfrak{m}_{c'}, \sum_{c' \in b, c' \leq c} \mathfrak{m}_{c'})$ . Given (7.17), consider any  $a_1 < a_2 \in U_c$  such that, for  $j = 1, 2$ ,

$$\partial_s \mathfrak{X}[\mu_b(s)](a_j) - \Phi[\mu_b(s)](\mathfrak{X}[\mu_b])(a_j) = \dot{\xi}_b \quad \text{holds for Lebesgue a.e. } s \in [0, t]. \quad (7.19)$$

It is readily checked from (2.8) and (2.11) that  $\Phi[\lambda](x)$  is nonincreasing in  $x$  and  $\mathfrak{X}[\lambda](a)$  is nondecreasing in  $a$ . Hence  $\Phi[\mu_b(s)](\mathfrak{X}[\mu_b])(a_1) \geq \Phi[\mu_b(s)](\mathfrak{X}[\mu_b])(a_2)$ . Take the difference of (7.19) for  $j = 1$  and  $j = 2$ , use the last inequality, integrate the result over  $[0, s]$ , and use  $\mathfrak{X}[\mu(0)](a_1) = \mathfrak{X}[\mu(0)](a_2) = \mathbf{x}_c$ . The result gives  $\mathfrak{X}[\mu(s)](a_1) \geq \mathfrak{X}[\mu(s)](a_2)$ . Since  $\mathfrak{X}[\lambda](a)$  is nondecreasing in  $a$ ,  $\mathfrak{X}[\mu(s)]|_{a \in [a_1, a_2]}$  is a constant. Sending  $[a_1, a_2] \rightarrow U_c$  shows that  $\mathfrak{X}[\mu_b(s)]|_{a \in U_c}$  is a constant. Hence  $\mu_c(s)$  is supported at a single point; let  $x_{\mu, c}(s) = x_{\mu, c}$  denote this point.

It remains only to show that  $x_{\mu, 1}, \dots, x_{\mu, n}$  coincidence with  $\xi_1, \dots, \xi_n$ . The measure  $\mu(s)$  being continuous in  $s$  implies that  $x_{\mu, c}(s)$  is too. Hence, there exists an  $s_0 > 0$  such that  $x_{\mu, 1}, \dots, x_{\mu, n}$  do not meet within  $[0, s_0)$ . Within  $s \in [0, s_0)$ , the equation (7.18) gives  $\dot{x}_{\mu, c} = \Phi[\mu_b](x_{\mu, c}) + \phi_b + v_b$ , for  $c \in b$ . The first two terms on the right hand side add to  $\frac{1}{2}(-\mathfrak{m}_1 - \dots - \mathfrak{m}_{c-1} + \mathfrak{m}_{c+1} + \dots + \mathfrak{m}_n) =: \phi_c$ , so  $x_{\mu, c}$  moves at the constant velocity  $\phi_c + v_b$  within  $[0, s_0)$ . This description of evolution matches that of the optimal clusters  $\xi_1, \dots, \xi_n$ , so  $x_{\mu, c}|_{[0, s_0]} = \xi_c|_{[0, s_0]}$ , for all  $c$ . By a time-continuity argument, this equality extends the first time when a merge happens in the optimal clusters. Take this first merge time as the new starting time and run the same argument. Continuing inductively completes the proof.

## 7.2. Proof of Corollary 2.5

Recall that we consider  $\mathfrak{m} = 1$  only.

(a) Given (2.18), the proof amounts to showing that

$$\frac{1}{N^3 T} \log \mathbb{P} \left[ \left| \frac{1}{NT} X_i(Tt) - \mathbf{x}' \right| \leq \alpha, \forall i \right] + \inf \left\{ \mathbb{I}(\mu) : \mu(0) = \sum_{c=1}^n \mathbf{m}_c \delta_{\mathbf{x}_c}, \mu(t) = \delta_{\mathbf{x}'} \right\} \quad (7.20)$$

tends to zero as  $N \rightarrow \infty$  first and  $\alpha \rightarrow 0$  later, where the attractive BPs start from  $X_i(0)/(NT) = \mathbf{x}_{c_{N,\xi}(i)}$  or equivalently  $\mu_N(0) = \sum_{c=1}^n \mathbf{m}_c \delta_{\mathbf{x}_c}$ . By (3.6), the event  $\{|X_i(Tt)/NT - \mathbf{x}'| \leq r, \forall i\}$  implies  $\text{dist}(\mu_N(t), \mu_{ec}) \leq r$ . Hence, by Theorem 2.3 and the contraction principle, the limit of (7.20) is  $\leq 0$ . To prove the reverse inequality, note that, by Theorem 2.4, the infimum in (7.20) is  $\mathbb{I}(\xi)$ . Applying Proposition 6.2 with  $\xi = \xi$  gives the reversed inequality  $\geq 0$ .

(b) Let us set up the notation and goal. Fix a nonempty  $A \subset \{1, \dots, n\}$ , recall  $c_{N,\xi}$  from (6.2), and let  $\mathfrak{I}_{N,\xi}(A) = \{i : c_{N,\xi}(i) \in A\}$ . Since  $A$  is nonempty,  $|\mathfrak{I}_{N,\xi}(A)|/N \geq \min\{\mathbf{m}_1, \dots, \mathbf{m}_n\} := \mathbf{m}_* > 0$ . Recall  $\text{Dist}_{N,[s',s'']}$  from (2.25)–(2.26), and, for  $\beta > 0$ , consider the event  $\mathcal{D}_N := \{\text{Dist}_{N,[0,t]}(X_i, \xi) > \beta, \forall i \in \mathfrak{I}_N(A)\}$ . Similar to (a), proving (b) amounts to proving that

$$\limsup_{\alpha \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P} \left[ \left\{ \left| \frac{1}{NT} X_i(Tt) - \mathbf{x}' \right| \leq \alpha, \forall i \right\} \cap \mathcal{D}_N \right] < -\mathbb{I}(\xi). \quad (7.21)$$

The strategy of proving (7.21) is to derive a lower bound on  $\text{dist}_{[0,t]}(\mu_N, \xi)$  from the condition imposed by  $\mathcal{D}_N$ . First, it is not hard to check from (3.4) that, for a large enough  $k_0 = k_0(\xi)$ ,

$$\text{dist}_{[0,t]}(\mu_N, \xi) \geq \text{dist}(\mu_N(s), \xi(s)) \geq 2^{-k_0} \min \left\{ 1, \frac{1}{N} \sum_{i=1}^N \text{Dist}_{N,s}(X_i, \xi) \right\}. \quad (7.22)$$

Under  $\mathcal{D}_N$ , for every  $i \in \mathfrak{I}_{N,\xi}(A)$ , there exists a random  $s(i) \in [0, t]$  such that  $\text{Dist}_{N,s(i)}(X_i, \xi) > \beta$ . If these  $s(i)$ s happen to be all the same, by (7.22),  $\text{dist}_{[0,t]}(\mu_N, \xi) \geq 2^{-k_0} \mathbf{m}_* \min\{1, \beta\}$ . In general, those  $s(i)$ s are different, so we need some time-continuity estimates. Recall the event  $\mathcal{U}_N(v)$  from (5.3). Use the bounds derived after (5.3) to fix a  $v$  so that

$$\limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P} [\mathcal{U}_N(v)^c] < -\mathbb{I}(\xi), \quad (7.23)$$

and write  $\mathcal{U}_N(v) = \mathcal{U}_N$  hereafter. We claim that, there exists a  $c_1 = c_1(v, \xi) < \infty$  such that

$$\begin{aligned} & \text{under } \mathcal{D}_N \cap \mathcal{U}_N, \quad \sup_{s \in [0,t]} \left\{ \left| \{i \in \mathfrak{I}_{N,\xi}(A) : \text{Dist}_{N,s}(X_i, \xi) > \beta/2\} \right| \right\} \\ & \geq N \mathbf{m}_* / c_1. \end{aligned} \quad (7.24)$$



Namely, under  $\mathcal{U}_N$ , a  $1/c_1$  fraction of  $i \in \mathfrak{I}_{N,\xi}(A)$  simultaneously satisfy  $\text{Dist}_{N,s}(X_i, \xi) > \beta/2$  at least once within  $s \in [0, \mathfrak{t}]$ . To see why, divide  $[0, \mathfrak{t}]$  into  $\ell$  equally spaced subintervals. Within each subinterval, apply the continuity estimate (5.13) and use the continuity of  $\xi_{\mathfrak{c}}(s)$ ,  $\mathfrak{c} = 1, \dots, n$ . By choosing  $\ell$  large enough (depending only on  $v$ ,  $\xi$ , and  $\beta$ ), we ensure that, under  $\mathcal{U}_N$ , the bound

$$\max_{k=1, \dots, \ell} \sup_{s, s' \in [\frac{(k-1)\mathfrak{t}}{\ell}, \frac{k\mathfrak{t}}{\ell}]} \left| \left| \frac{1}{NT} X_i(Ts) - \xi_{\mathfrak{c}_{N,\xi}(i)}(s) \right| - \left| \frac{1}{NT} X_i(Ts') - \xi_{\mathfrak{c}_{N,\xi}(i)}(s') \right| \right| \leq \frac{\beta}{2} \quad (7.25)$$

holds for at least  $N(1 - \mathfrak{m}_*/2)$  many  $i \in \{1, \dots, N\}$ . Since  $|\mathfrak{I}_{N,\xi}(A)| \geq N\mathfrak{m}_*$ , the bound (7.25) holds for at least  $N\mathfrak{m}_*/2$  many  $i \in \mathfrak{I}_{N,\xi}(A)$ . The claim (7.24) hence follows with  $c_1 = 2\ell$ . Set  $c_2 := 2^{-k_0} \min\{1, \beta\mathfrak{m}_*/(2c_1)\} > 0$ . Combining (7.22)–(7.24) gives

$$\mathcal{D}_N \cap \mathcal{U}_N \subset \{\text{dist}_{[0,\mathfrak{t}]}(\mu_N, \xi) \geq c_2\}. \quad (7.26)$$

Let us complete the proof of (7.21). First, by (7.26) and Theorem 2.3,

$$\begin{aligned} & \limsup_{\alpha \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N^3 T} \log \mathbb{P} \left[ \left\{ \left| \frac{1}{NT} X_i(T\mathfrak{t}) - \mathbf{x}' \right| \leq \alpha, \forall i \right\} \cap \mathcal{D}_N \cap \mathcal{U}_N \right] \\ & \leq \inf \left\{ \mathbb{I}(\mu) : \text{dist}_{[0,\mathfrak{t}]}(\mu, \xi) \geq c_2 : \mu(0) = \sum_{\mathfrak{c}} \mathfrak{m}_{\mathfrak{c}} \delta_{\mathbf{x}_{\mathfrak{c}}}, \mu(\mathfrak{t}) = \mathfrak{m} \delta_{\mathbf{x}'} \right\}. \end{aligned} \quad (7.27)$$

By Theorem 2.4, the right hand side of (7.27) is strictly smaller than  $-\mathbb{I}(\xi)$ . Given this and (7.23), the desired result (7.21) now follows.

## 8. Connection to $I_{\text{KPZ}}$ and the limit shape

Here, we establish a few properties of  $I_{\text{KPZ}}$  and the limit shape. In doing so, we prove Theorem 2.6: Part (a) is proven in Sections 8.1, 8.2, and 8.4; Parts (b)–(c) are proven in Section 8.3.

### 8.1. Basic properties of $I_{\text{KPZ}}$

Fix  $t$  and  $(\mathbf{x}_1 < \dots < \mathbf{x}_n)$  and writing  $I_{\text{KPZ}}(t, \vec{\mathbf{x}}, \vec{\mathbf{r}}) = I_{\text{KPZ}}(\vec{\mathbf{r}})$ ,  $\mathbf{f}_{\star, t, \vec{\mathbf{x}}, \vec{\mathbf{r}}} = \mathbf{f}_{\star, \vec{\mathbf{r}}}$ ,  $\mathcal{R}(t, \vec{\mathbf{x}}) = \mathcal{R}$ , and  $\mathcal{R}_{\text{conc}}(t, \vec{\mathbf{x}}) = \mathcal{R}_{\text{conc}}$  to simplify notation.

Let us show that  $I_{\text{KPZ}} : \mathcal{R} \rightarrow [0, \infty)$  is strictly convex, which implies that  $I_{\text{KPZ}}|_{\mathcal{R}_{\text{conc}}}$  is strictly convex. The key is to recognize  $\mathbf{f}_{\star, \vec{\mathbf{r}}}$  as the minimizer of a variational problem. Recall that  $\mathbf{p}(t, x) = -x^2/(2t)$ . For any piecewise  $\mathcal{C}^1$  function  $\mathbf{f} : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\mathbf{f} \geq \mathbf{p}(t)$  and  $\mathbf{f}(x) = \mathbf{p}(t, x)$  for all large enough  $|x|$ , consider  $I_{\text{BM}}(\mathbf{f}) := \int_{\mathbb{R}} dx \left( \frac{1}{2} (\partial_x \mathbf{f})^2 - \frac{1}{2} (\partial_x \mathbf{p}(t))^2 \right)$ . Note that the integral is well-defined and finite because  $\mathbf{f}(x) = \mathbf{p}(t, x)$  for all large enough  $|x|$ . It is not hard to check that

$$I_{\text{KPZ}}(\vec{\mathbf{r}}) := I_{\text{BM}}(\mathbf{f}_{\star, \vec{\mathbf{r}}}) = \inf \{ I_{\text{BM}}(\mathbf{f}) : \mathbf{f}(\mathbf{x}_{\mathbf{c}}) = \mathbf{r}_{\mathbf{c}}, \mathbf{c} = 1, \dots, n \}. \quad (8.1)$$

Namely,  $\mathbf{f}_{\star, \vec{\mathbf{r}}}$  is the minimizer of the infimum in (8.1). For any  $\gamma \in (0, 1)$ , use the convexity of  $u \mapsto u^2/2$  to write

$$\frac{1}{2}(\gamma \partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_1} + (1 - \gamma) \partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_2})^2 \leq \gamma \frac{1}{2}(\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_1})^2 + (1 - \gamma) \frac{1}{2}(\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_2})^2. \quad (8.2)$$

Subtract  $(\partial_x(-\mathbf{p}(t)))^2/2$  from both sides and integrate both sides over  $x \in \mathbb{R}$ . The right hand side is  $\gamma I_{\text{KPZ}}(\vec{\mathbf{r}}_1) + (1 - \gamma) I_{\text{KPZ}}(\vec{\mathbf{r}}_2)$ . The left hand side is at least  $I_{\text{KPZ}}(\gamma \vec{\mathbf{r}}_1 + (1 - \gamma) \vec{\mathbf{r}}_2)$  by the variational characterization (8.1). This proves the convexity. To prove the strictness, note that the equality in (8.2) holds only if  $\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_1} = \partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_2}$ . This being true for Lebesgue a.e.  $x \in \mathbb{R}$  forces  $\vec{\mathbf{r}}_1 = \vec{\mathbf{r}}_2$ .

Next, we show that, for every  $\vec{\mathbf{r}} \in \mathcal{R}$ ,

$$\begin{aligned} \partial_{\mathbf{r}_{\mathbf{c}}} I_{\text{KPZ}}(\vec{\mathbf{r}}) &:= \partial_{\mathbf{r}_{\mathbf{c}}} I_{\text{KPZ}}(t, \vec{\mathbf{x}}, \vec{\mathbf{r}}) = (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}})(\mathbf{x}_{\mathbf{c}}^-) - (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}})(\mathbf{x}_{\mathbf{c}}^+), \\ &\text{for all } \mathbf{c} = 1, \dots, n. \end{aligned} \quad (8.3)$$

Let  $L_{\mathbf{c}, \mathbf{c}+1} = L_{\mathbf{c}, \mathbf{c}+1}(x)$  denote the linear function such that  $L_{\mathbf{c}, \mathbf{c}+1}(\mathbf{x}_{\mathbf{c}}) = \mathbf{r}_{\mathbf{c}}$  and that  $L_{\mathbf{c}, \mathbf{c}+1}(\mathbf{x}_{\mathbf{c}+1}) = \mathbf{r}_{\mathbf{c}+1}$ , and let  $\mathbf{x}_{\mathbf{c}, -} := \inf \{ x \geq \mathbf{x}_{\mathbf{c}-1} : \partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}}(x) = \partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}}(\mathbf{x}_{\mathbf{c}}^-) \}$ , with the convention  $\mathbf{x}_0 := -\infty$ . In words,  $\mathbf{x}_{\mathbf{c}, -}$  is  $\mathbf{x}_{\mathbf{c}-1}$  when  $L_{\mathbf{c}, \mathbf{c}+1}|_{[\mathbf{x}_{\mathbf{c}}, \mathbf{x}_{\mathbf{c}+1}]} \geq \mathbf{p}(t)|_{[\mathbf{x}_{\mathbf{c}}, \mathbf{x}_{\mathbf{c}+1}]}$ , otherwise  $\mathbf{x}_{\mathbf{c}, -}$  is the “tangent point” to the left of  $\mathbf{x}_{\mathbf{c}}$ ; the tangent points are those labeled by triangles in Fig. 4. Define  $\mathbf{x}_{\mathbf{c}, +}$  similarly. To prove (8.3), it suffices to consider those  $\vec{\mathbf{r}} \in \mathcal{R}$  such that

$$\text{either } L_{\mathbf{c}-1, \mathbf{c}}|_{[\mathbf{x}_{\mathbf{c}-1}, \mathbf{x}_{\mathbf{c}}]} > \mathbf{p}(t)|_{[\mathbf{x}_{\mathbf{c}-1}, \mathbf{x}_{\mathbf{c}}]} \text{ or } \mathbf{x}_{\mathbf{c}-1, +} < \mathbf{x}_{\mathbf{c}, -} \text{ are both tangent points,} \quad (8.4a)$$

$$\text{either } L_{\mathbf{c}, \mathbf{c}+1}|_{[\mathbf{x}_{\mathbf{c}}, \mathbf{x}_{\mathbf{c}+1}]} > \mathbf{p}(t)|_{[\mathbf{x}_{\mathbf{c}}, \mathbf{x}_{\mathbf{c}+1}]} \text{ or } \mathbf{x}_{\mathbf{c}, +} < \mathbf{x}_{\mathbf{c}+1, -} \text{ are both tangent points.} \quad (8.4b)$$

Indeed, for fixed  $\mathbf{r}_1, \dots, \mathbf{r}_{\mathbf{c}-1}, \mathbf{r}_{\mathbf{c}+1}, \dots, \mathbf{r}_n$ , there exist at most two values of  $\mathbf{r}_{\mathbf{c}}$  for which (8.4) fails. Also, as is readily checked, the right hand side (8.3) is continuous on  $\mathcal{R}$ . Hence, once (8.3) is proven for those  $\vec{\mathbf{r}} \in \mathcal{R}$  satisfying (8.4), the result extends to all  $\vec{\mathbf{r}} \in \mathcal{R}$ .

Let us prove (8.3) under the assumption (8.4). Under this assumption, perturbing  $\mathbf{r}_{\mathbf{c}}$  changes  $\mathbf{f}_{\star, \vec{\mathbf{r}}}$  only within  $x \in [\mathbf{x}_{\mathbf{c}, -}, \mathbf{x}_{\mathbf{c}}] \cup [\mathbf{x}_{\mathbf{c}}, \mathbf{x}_{\mathbf{c}, +}]$ , so

$$\begin{aligned} \partial_{\mathbf{r}_{\mathbf{c}}} I_{\text{KPZ}}(\vec{\mathbf{r}}) &= \partial_{\mathbf{r}_{\mathbf{c}}} \int_{\mathbf{x}_{\mathbf{c}, -}}^{\mathbf{x}_{\mathbf{c}}} dx \left( \frac{1}{2}(\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}})^2 - \frac{1}{2}(\partial_x \mathbf{p}(t))^2 \right) \\ &\quad + \partial_{\mathbf{r}_{\mathbf{c}}} \int_{\mathbf{x}_{\mathbf{c}}}^{\mathbf{x}_{\mathbf{c}, +}} dx \left( \frac{1}{2}(\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}})^2 - \frac{1}{2}(\partial_x \mathbf{p}(t))^2 \right). \end{aligned} \quad (8.5)$$

In the integrals in (8.1), only  $\mathbf{x}_{\mathbf{c},-}, \mathbf{x}_{\mathbf{c},+}, \mathbf{f}_{\star, \bar{\mathbf{r}}}$  may depend on  $\mathbf{r}_{\mathbf{c}}$ . Under (8.4), either  $\partial \mathbf{x}_{\mathbf{c},-} / \partial \mathbf{r}_{\mathbf{c}} = 0$  or  $(\partial_x \mathbf{f}_{\star, \bar{\mathbf{r}}})(\mathbf{x}_{\mathbf{c},-}) = (\partial_x \mathbf{p})(t, \mathbf{x}_{\mathbf{c},-})$ . Hence, in (8.5), the contribution of differentiating  $\mathbf{x}_{\mathbf{c},-}$  is zero. The same holds for  $\mathbf{x}_{\mathbf{c},+}$ . What remains is the contribution of differentiating the integrands:

$$\partial_{\mathbf{c}} I_{\text{KPZ}}(\bar{\mathbf{r}}) = \int_{[\mathbf{x}_{\mathbf{c},-}, \mathbf{x}_{\mathbf{c}}]} dx \partial_x \mathbf{f}_{\star, \bar{\mathbf{r}}} \frac{1}{\mathbf{x}_{\mathbf{c}} - \mathbf{x}_{\mathbf{c},-}} + \int_{[\mathbf{x}_{\mathbf{c}}, \mathbf{x}_{\mathbf{c},+}]} dx \partial_x \mathbf{f}_{\star, \bar{\mathbf{r}}} \left( \frac{-1}{\mathbf{x}_{\mathbf{c},+} - \mathbf{x}_{\mathbf{c}}} \right). \quad (8.6)$$

Note that the integrands are constant, so the last expression evaluates to the right hand side of (8.3).

8.2. The map  $\nabla I_{\text{KPZ}} : \mathcal{R}_{\text{conc}} \rightarrow [0, \infty)^n$  is a homeomorphism

Notation as in Section 8.1.

First, by the strict convexity from Section 8.1,  $\nabla I_{\text{KPZ}}$  is injective, and from (8.3), it is not hard to check that  $\nabla I_{\text{KPZ}}$  is continuous.

To prepare for the rest of the proof, for any  $\vec{\mathbf{m}} \in [0, \infty)^n$ , we consider  $S(\vec{\mathbf{m}}) := \{\bar{\mathbf{r}} \in \mathcal{R}_{\text{conc}} : \partial_{\mathbf{c}} I_{\text{KPZ}}(\bar{\mathbf{r}}) \leq \mathbf{m}_{\mathbf{c}}, \mathbf{c} = 1, \dots, n\}$  and establish a few properties of it. First, the set  $S(\vec{\mathbf{m}})$  is nonempty because it contains  $(-\mathbf{x}_{\mathbf{c}}^2/(2t))_{\mathbf{c}=1}^n$ . Next, we claim that, for any compact  $K \subset \mathbb{R}^n$ ,  $S(K) := \cup_{\vec{\mathbf{m}} \in K} S(\vec{\mathbf{m}})$  is compact. It is not hard to check that  $S(K)$  is closed. By definition, every  $\bar{\mathbf{r}} \in S(K)$  satisfies  $\mathbf{r}_1 \geq -\mathbf{x}_1^2/(2t), \dots, \mathbf{r}_n \geq -\mathbf{x}_n^2/(2t)$ , so we only need upper bounds on  $\mathbf{r}_1, \dots, \mathbf{r}_n$ . Sum the formula (8.3) over  $\mathbf{c}, \dots, n$  to get

$$(\partial_x \mathbf{f}_{\star, \bar{\mathbf{r}}})(\mathbf{x}_{\mathbf{c}}^-) - (\partial_x \mathbf{f}_{\star, \bar{\mathbf{r}}})(\mathbf{x}_{\mathbf{n}}^+) = \partial_{\mathbf{c}} I_{\text{KPZ}}(\bar{\mathbf{r}}) + \dots + \partial_n I_{\text{KPZ}}(\bar{\mathbf{r}}) \leq \mathbf{m}_{\mathbf{c}} + \dots + \mathbf{m}_n. \quad (8.7)$$

Recall  $\mathbf{x}_{\mathbf{c}, \pm}$  from after (8.3), and note that  $(\partial_x \mathbf{f}_{\star, \bar{\mathbf{r}}})(\mathbf{x}_{\mathbf{n}}^+) = (\partial_x \mathbf{p})(t, \mathbf{x}_{\mathbf{n},+}) \leq (\partial_x \mathbf{p})(t, \mathbf{x}_{\mathbf{n}})$ . Hence  $\partial_x (\mathbf{f}_{\star, \bar{\mathbf{r}}})(\mathbf{x}_{\mathbf{c}}^-)$  is bounded from above, and the bound can be chosen uniformly over  $\vec{\mathbf{m}} \in K$ . Using this property inductively for  $\mathbf{c} = 1, \dots, n$  shows that  $\mathbf{r}_1, \dots, \mathbf{r}_n$  are bounded from above, uniformly over  $\vec{\mathbf{m}} \in K$ .

Let us prove that  $\nabla I_{\text{KPZ}}$  is surjective. Let  $\pi_{\mathbf{c}} : \mathbb{R}^n \rightarrow \mathbb{R}$  be the projection onto the  $\mathbf{c}$ th coordinate, take any  $\vec{\mathbf{m}} \in [0, \infty)^n$ , consider  $\mathbf{r}_{*, \mathbf{c}'} := \sup \pi_{\mathbf{c}'}(S(\vec{\mathbf{m}}))$ , and set  $\vec{\mathbf{r}}_* := (\mathbf{r}_{*, \mathbf{c}'}^n)_{\mathbf{c}'=1}$ . Fix any  $\mathbf{c}$ . By the construction of  $\vec{\mathbf{r}}_*$  and the compactness of  $S(\vec{\mathbf{m}}) = S(\{\vec{\mathbf{m}}\})$ , there exists a convergent sequence  $\vec{\mathbf{r}}_{(k)} \rightarrow \vec{\mathbf{r}}_{(\infty)}$  in  $S(\vec{\mathbf{m}})$  such that  $\mathbf{r}_{(\infty), \mathbf{c}} = \mathbf{r}_{*, \mathbf{c}}$  and  $\mathbf{r}_{(\infty), \mathbf{c}'} \leq \mathbf{r}_{*, \mathbf{c}'}$  for all  $\mathbf{c}'$ . These properties together with the property that  $\mathbf{f}_{\star, \vec{\mathbf{r}}_{(\infty)}}$  is concave gives that

$$(\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_*})(\mathbf{x}_{\mathbf{c}}^-) - (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_*})(\mathbf{x}_{\mathbf{c}}^+) \leq (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_{(\infty)}})(\mathbf{x}_{\mathbf{c}}^-) - (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_{(\infty)}})(\mathbf{x}_{\mathbf{c}}^+). \quad (8.8)$$

The right hand side is at most  $\mathbf{m}_{\mathbf{c}}$  because  $\vec{\mathbf{r}}_{(\infty)} \in S(\vec{\mathbf{m}})$ . We arrive at the inequality

$$\partial_{\mathbf{c}} I_{\text{KPZ}}(\vec{\mathbf{r}}_*) = (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_*})(\mathbf{x}_{\mathbf{c}}^-) - (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}_*})(\mathbf{x}_{\mathbf{c}}^+) \leq \mathbf{m}_{\mathbf{c}}. \quad (8.9)$$

Next, we have the following property that can be checked by referring to Fig. 4: For any  $\vec{\mathbf{r}} \in \mathcal{R}_{\text{conc}}$  (and in fact for any  $\vec{\mathbf{r}} \in \mathcal{R}$ ), when  $\mathbf{r}_{\mathbf{c}}$  increases while other components remain fixed, the quantity  $\partial_{\mathbf{c}} I_{\text{KPZ}} = (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}})(\mathbf{x}_{\mathbf{c}}^-) - (\partial_x \mathbf{f}_{\star, \vec{\mathbf{r}}})(\mathbf{x}_{\mathbf{c}}^+)$  increases while other  $(\partial_{\mathbf{c}'} I_{\text{KPZ}})$ s decrease or stay the same. This property forces the inequality in (8.9) to be an equality: Otherwise, by increasing  $\mathbf{r}_{\mathbf{c}}$ , we can increase  $\partial_{\mathbf{c}} I_{\text{KPZ}}$  while maintaining the inequality  $\partial_{\mathbf{c}'} I_{\text{KPZ}} \leq \mathbf{m}_{\mathbf{c}'}$  for all  $\mathbf{c}'$ , which contradicts the construction of  $\vec{\mathbf{r}}_{\star}$ . Since  $\mathbf{c}$  was arbitrary,  $\partial_{\mathbf{c}} I_{\text{KPZ}}(\vec{\mathbf{r}}_{\star}) = \mathbf{m}_{\mathbf{c}}$ , for  $\mathbf{c} = 1, \dots, n$ , and this gives the desired surjectivity.

Finally, we prove that  $(\nabla I_{\text{KPZ}})^{-1}$  is continuous. The previous paragraph gives  $(\nabla I_{\text{KPZ}})(S(\vec{\mathbf{m}})) \ni \vec{\mathbf{m}}$ . Take any  $b < \infty$  and consider  $S' := S([0, b]^n)$ . We have  $(\nabla I_{\text{KPZ}})(S') \supset [0, b]^n$ . Since  $S'$  is compact, the continuity of  $\nabla I_{\text{KPZ}}$  implies the continuity of  $(\nabla I_{\text{KPZ}})^{-1}$  on  $[0, b]^n$ .

### 8.3. Limit shape and its shocks

Fix  $\vec{\mathbf{r}} \in \mathcal{R}_{\text{conc}} = \mathcal{R}_{\text{conc}}(\mathbf{t}, \vec{\mathbf{x}})^{\circ}$ . Hereafter, we will often drop the  $\mathbf{t}, \vec{\mathbf{x}}, \vec{\mathbf{r}}$  dependence to simplify notation.

We begin by recalling some PDE background related to  $\mathbf{h}_{\star}$ ; we refer to [27, Ch. 3] for an introduction on this topic. Consider Burgers' equation and its Hamilton–Jacobi equation

$$\partial_t \mathbf{u} = \frac{1}{2} \partial_x (\mathbf{u}^2), \quad \partial_t \mathbf{h} = \frac{1}{2} (\partial_x \mathbf{h})^2, \quad (t, x) \in (0, \mathbf{t}] \times \mathbb{R}. \quad (8.10)$$

The two equations are related by  $\partial_x \mathbf{h} = \mathbf{u}$ . These equations can have multiple weak solutions under a given initial condition, but have a unique entropy solution given by the Hopf–Lax operator:

$$\mathbf{h}(t) = \text{HL}_t(\mathbf{h}(0)), \quad \text{HL}_t(\mathbf{f})(x) := \sup \left\{ -\frac{(x-y)^2}{2t} + \mathbf{f}(y) : y \in \mathbb{R} \right\}. \quad (8.11)$$

Now, *time reverse* what is described above and consider the backward version of (8.10)–(8.11):

$$\partial_s \mathbf{h}(\mathbf{t} - s, x) = -\frac{1}{2} (\partial_x \mathbf{h}(\mathbf{t} - s, x))^2, \quad \partial_s \mathbf{u}(\mathbf{t} - s, x) = -\frac{1}{2} \partial_x (\mathbf{u}(\mathbf{t} - s, x))^2, \quad (8.12)$$

$$(\text{HL}_s^{\text{bk}}(\mathbf{f}))(x) := \inf_{y \in \mathbb{R}} \left\{ \frac{(x-y)^2}{2s} + \mathbf{f}(y) \right\}. \quad (8.13)$$

Under such notation, the limit shape  $\mathbf{h}_{\star}$ , defined in (2.32), is obtained by taking  $\mathbf{f}_{\star} = \mathbf{f}_{\star, \mathbf{t}, \vec{\mathbf{x}}, \vec{\mathbf{r}}}$  as the terminal condition and evolving it backward by  $\text{HL}^{\text{bk}}$ . As is readily checked, a weak solution of (8.12) is also a weak solution (8.10). On the other hand, an entropy solution of (8.12) (which we call backward entropy) is in general *not* an entropy solution of (8.10) (which we call forward entropy). Hence  $\mathbf{h}_{\star}$  is a weak solution of (8.10) and (8.12), is backward entropy, but is not forward entropy. Put  $\mathbf{u}_{\star} := \partial_x \mathbf{h}_{\star}$ . The **characteristics** are

linear trajectories in spacetime along which  $u_*$  is constant. The **shocks** are trajectories in spacetime across which  $u_*$  is not continuous.

Let us give a geometric description of  $h_*$ . Write  $h_*(t) = f_*$  as an infimum of lines:

$$f_* = \min \left\{ \min \left\{ L_{\triangleleft}^b, L_{c, c+1}^b, L_{\triangleright}^b : b \in \mathfrak{B}_* \right\}, \right. \quad (8.14a)$$

$$\left. \inf \left\{ L : L \text{ is a tangent line of } p(t) \text{ at places where } p(t) = f_* \right\} \right\}. \quad (8.14b)$$

Here,  $\mathfrak{B}_*$  is a partition of  $\{1, \dots, n\}$  into intervals, where each  $b \in \mathfrak{B}_*$  is a maximal set of  $c$ s such that  $(f_* - p(t))|_{x \in [-x_c, x_{c'}]} > 0$  for all  $c, c' \in b$ ; we will show later that  $\mathfrak{B}_* = \mathfrak{B}$ . Let  $y_{\triangleleft}^b$  and  $y_{\triangleright}^b$  be the left and right tangent points as depicted in Fig. 5. The minimum in (8.14a) runs over all  $b \in \mathfrak{B}_*$  and  $c, c+1 \in b$ , and accounts for the piecewise linear part of  $f_*$ , as depicted in Fig. 5. The infimum in (8.14b) accounts for the parabolic part of  $f_*$ , as depicted in Fig. 6. Next, apply  $HL_s^{bk}$  to both sides of (8.14) and exchange the infimum over  $y$  (in (8.13)) with other minimums and infimums to get

$$h_*(t-s) = \min \left\{ \min \left\{ HL_s^{bk}(L_{\triangleleft}^b), HL_s^{bk}(L_{c, c+1}^b), HL_s^{bk}(L_{\triangleright}^b) : b \in \mathfrak{B}_* \right\}, \right. \quad (8.15a)$$

$$\left. \inf \left\{ HL_s^{bk}(L) : L \text{ is a tangent line of } p(t) \text{ at places where } p(t) = f_* \right\} \right\}. \quad (8.15b)$$

For a linear function  $L(x) = v_1x + v_2$ , it is readily checked that  $HL_s^{bk}(L) = L - (s/2)(\partial_x L)^2$ . It is also readily checked that  $HL_s^{bk}(p(t)) = p(t-s)$ . Hence, for  $s \in (0, t)$ , the limit shape  $h_*(t-s)$  is obtained by vertically shifting the lines in Figs. 5–6 by  $-(s/2) \cdot (\text{slope})^2$  and taking the infimum of the result. Note that every shifted line stays above  $p(t-s)$ , since  $HL_s^{bk}$  preserves orders: Namely,  $f_1 \geq f_2$  implies  $HL_s^{bk}(f_1) \geq HL_s^{bk}(f_2)$ . Within each  $b \in \mathfrak{B}_*$ , the leftmost and rightmost lines (those indexed by  $\triangleleft$  and  $\triangleright$ ) touch the parabola  $p(t-s)$  at tangent. At the tangent points,  $\partial_x p(t-s) = -x/(t-s)$  is equal to the slopes of those lines, so the  $x$  coordinates of the tangent points are  $(t-s)y_{\triangleleft}^b/t$  and  $(t-s)y_{\triangleright}^b/t$ , which trace out the dashed lines in Fig. 7.

Based on the preceding description of  $h_*$ , we infer some properties of  $h_*$  and its shocks. Let  $\mathcal{C}_b := \{(t, x) : t \in [0, t], x \in [(t/t)y_{\triangleleft}^b, (t/t)y_{\triangleright}^b]\}$  be the spacetime region bounded by the tangent points. They are the colored regions in Fig. 7.

- (I) Outside  $\cup_{b \in \mathfrak{B}_*} \mathcal{C}_b$ ,  $h_*(t, x) = p(t, x) = -x^2/(2t)$ , so the characteristics are straight lines that connect  $(t, x) = (0, 0)$  and  $\{t\} \times \mathbb{R}$ ; see Fig. 7.
- (II) Within each  $\mathcal{C}_b$ ,  $u_*$  is piecewise constant, with values given by the slopes of  $L_{\triangleleft}^b, L_{\triangleright}^b$ , and a subset of  $\{L_{c, c+1}^b : c, c+1 \in b\}$ ; the jumps of  $u_*$  occur exactly along the shocks; see Fig. 7.

Let  $\sigma_c = \sigma_c(s)$  denote the shocks, parameter in the backward time  $s$ . Let  $t', x'_a$  be as in the notation of Theorem 2.6(c). By the Rankine–Hugoniot relation (see [27, Ch. 3] for example)

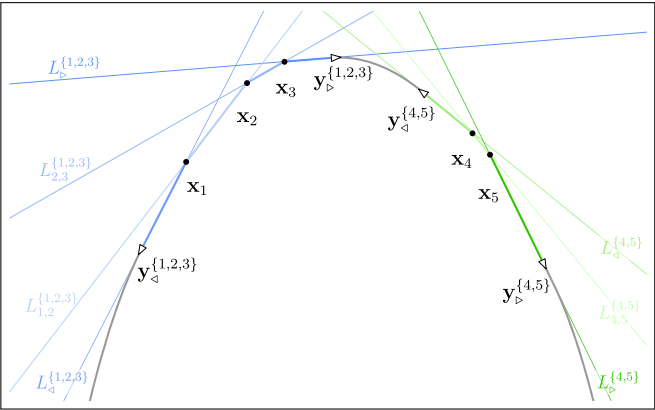


Fig. 5.  $f_*$  as an infimum of lines, the piecewise linear part. In this Figure,  $\mathfrak{B}_* = \{\{1, 2, 3\}, \{4, 5\}\}$ .

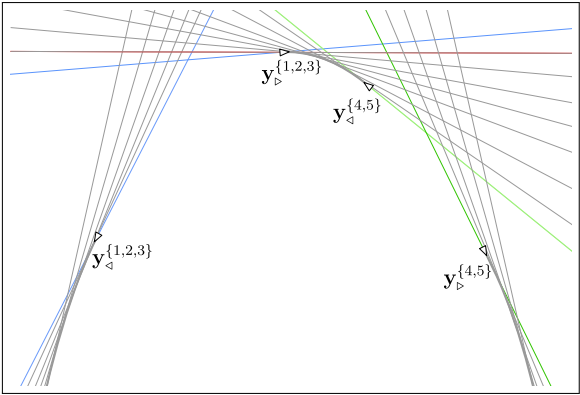


Fig. 6.  $f_*$  as an infimum of lines, the parabolic part.

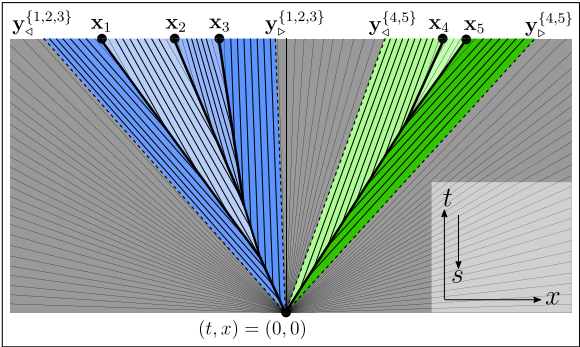


Fig. 7. The shocks (thick solid lines), characteristics (thin solid lines), and tangent points (dashed lines).

$$\dot{\sigma}_a|_{s=t-t'} := (\frac{d}{ds}\sigma_a)(t-t') = \frac{1}{2}(u_{*,a}^- + u_{*,a}^+), \quad \text{where } u_{*,a}^\pm := u_*(t', \sigma_a(t-t')^\pm).$$

(8.16)

This relation and Properties (I)–(II) together give the following property.

- (III) For each  $\mathfrak{b} \in \mathfrak{B}_*$ , those shocks  $\sigma_c$  with  $c \in \mathfrak{b}$  stay within  $\mathcal{C}_{\mathfrak{b}}$  and travel at constant velocities except when they meet. Further, shocks  $\sigma_c$  and  $\sigma_{c'}$  merge within  $s \in (0, t)$  if and only if they belong to the same  $\mathfrak{b} \in \mathfrak{B}_*$ , so  $\mathfrak{B}_* = \mathfrak{B}$ .

By Property (II),  $h_*$  is linear in a neighborhood on either side of a shock and hence solves (8.10) *classically* there. Using this property gives the following.

- (IV) Let  $t', \mathbf{x}'_a$  be as in the notation of Theorem 2.6(c). For all  $t' \in (0, t)$  except when shocks merge,  $\frac{d}{dt} h_*(t, \sigma_a(t-t))|_{t=t'} = \frac{1}{2}(u_{*,a}^-)^2 - u_{*,a}^- \dot{\sigma}_a$ , where  $u_{*,a}^- := u_*(t, \sigma_a(t-t'))^-$ .

Let us prove Theorem 2.6(c). To prove (2.34), recall that  $\nabla_{\vec{r}} I_{\text{KPZ}}(t, \vec{x}, \vec{r}) = \vec{m}$  and combine this relation with the formula (8.3) for  $t = t$  to get  $u_{*,c}^- - u_{*,c}^+ = m_c$ . Next, use the formula (8.3) for  $t = t'$  to get  $\nabla_{\vec{r}} I_{\text{KPZ}}(t', \vec{x}', \vec{r}') = u_{*,a}^- - u_{*,a}^+$ , telescope the right hand side as  $\sum_{c \in \mathfrak{C}(a)} (u_{*,c}^- - u_{*,c}^+)$ , use  $u_{*,c}^- - u_{*,c}^+ = m_c$ , and recognize the resulting sum as  $m'_a$ . Doing so concludes (2.34). To prove (2.35), let  $\xi$  be the optimal deviation for (2.23) with  $\mathbf{x}' = 0$ . From the definition of the optimal deviation (in Section 2.3), it is readily checked that  $\xi(\cdot + (t-t'))|_{[0,t]}$  is the optimal deviation for (2.23) with  $[0, t] \mapsto [0, t']$ ,  $\mathbf{x}' = 0$ ,  $(\vec{x}, \vec{m}) \mapsto (\vec{m}', \vec{x}')$ . Similarly, for each  $a = 1, \dots, n'$ , the deviation  $\sum_{c \in \mathfrak{C}(a)} m_c \delta \xi_c|_{[0,t-t']}$  is the optimal deviation for (2.23) with  $[0, t] \mapsto [0, t-t']$ ,  $\mathbf{x}' = \mathbf{x}'_a$ ,  $(\vec{x}, \vec{m}) \mapsto (\mathbf{x}_c, m_c)_{c \in \mathfrak{C}(a)}$ . These properties together with Theorem 2.4 give (2.35).

Let us prove Theorem 2.6(b). With  $\vec{x}, \vec{r}$  having been fixed, set  $\vec{m} := (\nabla I_{\text{KPZ}})(t, \vec{x}, \vec{r})$ . For this  $\vec{m}$ , let  $(\xi_c)_{c=1}^n$  be the optimal clusters. Take any intermediate time  $t' \in (0, t]$  and let  $a \in \{1, \dots, n'\}$ ,  $\mathfrak{C}(a)$ ,  $m'_a$  be as in Notation for Theorem 2.6(c), and recall  $v_b$  from before Theorem 2.4. We seek to prove that, for any  $a \in \{1, \dots, n'\}$  with  $\mathfrak{C}(a) \subset \mathfrak{b}$ ,

$$\dot{\sigma}_a = \frac{1}{2} \sum_{a'=1}^{n'} \text{sgn}(\sigma_{a'} - \sigma_a) m'_{a'} + v_b. \quad (8.17)$$

Referring to the definition of the optimal clusters (before Theorem 2.6(c)), we see that (8.17) means that the shocks evolve the same way (in the backward time  $s$ ) as the optimal clusters. Since the shocks and optimal clusters share the same starting condition, namely  $(\sigma_c(0))_{c=1}^n = (\mathbf{x}_c)_{c=1}^n = (\xi_c(0))_{c=1}^n$ , once (8.17) is proven, the desired result  $(\sigma_c)_{c=1}^n = (\xi_c)_{c=1}^n$  follows. To prove (8.17), fix  $a, b$  with  $\mathfrak{C}(a) \subset \mathfrak{b}$ . Consider the set  $\{a' : \mathfrak{C}(a') = \mathfrak{b}\}$  and write it at  $\{a' : \mathfrak{C}(a') = \mathfrak{b}\} = [a^\triangleleft, \dots, a^\triangleright] \cap \mathbb{Z}$ . Let  $u_{*,\triangleleft}^b$  and  $u_{*,\triangleright}^b$  denote the respective slopes of  $L_{\triangleleft}^b$  and  $L_{\triangleright}^b$ . By Property (II),  $u_{*,\triangleleft}^b = u_{*,a^\triangleleft}^-$ ,  $u_{*,\triangleright}^b = u_{*,a^\triangleright}^+$ , and  $u_{*,a'}^+ = u_{*,a'+1}^-$  for all  $a' \in [a^\triangleleft, a^\triangleright - 1]$ . Use these properties to telescope the right hand side of (8.16) as

$$\dot{\sigma}_{\mathbf{a}} = \frac{1}{2} \left( u_{\star, \triangleleft}^{\mathbf{b}} + \sum_{\mathbf{a}' \in [\mathbf{a}^{\triangleleft}, \mathbf{a})} (-u_{\star, \mathbf{a}'}^{-} + u_{\star, \mathbf{a}'}^{+}) + \sum_{\mathbf{a}' \in (\mathbf{a}, \mathbf{a}^{\triangleright}]} (u_{\star, \mathbf{a}'}^{-} - u_{\star, \mathbf{a}'}^{+}) + u_{\star, \triangleright}^{\mathbf{b}} \right). \quad (8.18)$$

Next, combine (2.34) and (8.3) to get

$$\mathbf{m}'_{\mathbf{a}} = u_{\star, \mathbf{a}}^{-} - u_{\star, \mathbf{a}}^{+}, \quad \mathbf{a} = 1, \dots, n'. \quad (8.19)$$

Inserting (8.19) into (8.18) gives

$$\dot{\sigma}_{\mathbf{a}} = \frac{1}{2} (u_{\star, \triangleleft}^{\mathbf{b}} + u_{\star, \triangleright}^{\mathbf{b}}) + \sum_{\mathbf{a}' \in [\mathbf{a}^{\triangleleft}, \mathbf{a}^{\triangleright}]} \text{sgn}(\sigma_{\mathbf{a}'} - \sigma_{\mathbf{a}}) \mathbf{m}'_{\mathbf{a}'}. \quad (8.20)$$

Let  $\sigma_{\mathbf{b}} := \sum_{\mathbf{c} \in \mathbf{b}} \mathbf{m}_{\mathbf{c}} \sigma_{\mathbf{c}} / \mathbf{m}_{\mathbf{b}}$ . Multiply both sides of (8.20) by  $\mathbf{m}_{\mathbf{a}}$ , sum both sides over all  $\mathbf{a} \in [\mathbf{a}^{\triangleleft}, \mathbf{a}^{\triangleright}]$ , and divide the result by  $\mathbf{m}_{\mathbf{b}}$ . Doing so gives  $\dot{\sigma}_{\mathbf{b}} = \frac{1}{2} (u_{\star, \triangleleft}^{\mathbf{b}} + u_{\star, \triangleright}^{\mathbf{b}})$ , which is constant. Write the left hand side as  $\dot{\sigma}_{\mathbf{b}} = \frac{1}{2} (\dots - \mathbf{m}_{\mathbf{b} \ominus 1} + \mathbf{m}_{\mathbf{b} \oplus 1} + \dots) + v'_{\mathbf{b}}$ , for some constant  $v'_{\mathbf{b}}$ . Integrate  $\dot{\sigma}_{\mathbf{b}}$  over  $s \in [0, \mathfrak{t}]$ , use  $\sigma_{\mathbf{b}}(0) = \xi_{\mathbf{b}}(0)$  and  $\sigma_{\mathbf{b}}(\mathfrak{t}) = 0 = \xi_{\mathbf{b}}(\mathfrak{t})$ , and compare the result with (2.24). Doing so shows that  $v'_{\mathbf{b}} = v_{\mathbf{b}}$ . Hence  $\frac{1}{2} (u_{\star, \triangleleft}^{\mathbf{b}} + u_{\star, \triangleright}^{\mathbf{b}}) = \frac{1}{2} (\dots - \mathbf{m}_{\mathbf{b} \ominus 1} + \mathbf{m}_{\mathbf{b} \oplus 1} + \dots) + v_{\mathbf{b}}$ . Inserting this into (8.20) gives the desired result (8.17).

#### 8.4. Legendre transform

Here we prove the first statement in Theorem 2.6(a), by showing that the Legendre transform of  $I_{\text{KPZ}}(\vec{\mathbf{r}})$  gives  $L_{\text{SHE}}(\vec{\mathbf{m}})$ . Once this is done, since  $I_{\text{KPZ}}$  is strictly convex and since  $\nabla I_{\text{KPZ}} : \mathcal{R}_{\text{conc}} \rightarrow [0, \infty)^n$  is a homeomorphism, it will follow that  $L_{\text{SHE}}(\vec{\mathbf{m}})$  is also strictly convex and is the Legendre transform of  $I_{\text{KPZ}}(\vec{\mathbf{r}})$ . Still use the notation in Theorem 2.6. Set  $G := \vec{\mathbf{m}} \cdot \vec{\mathbf{r}} - I_{\text{KPZ}}(\vec{\mathbf{r}})$ , where the dot denotes the Euclidean inner product. Our goal is to show  $G = L_{\text{SHE}}(\vec{\mathbf{m}})$ . To this end, we devise a time-dependent version of  $G$ :

$$G(t) := \sum_{\mathbf{c}=1}^n \mathbf{m}_{\mathbf{c}} h_{\star}(t, \sigma_{\mathbf{c}}(t-t)) - \sum_{\mathbf{b} \in \mathfrak{B}_{\star} t \mathbf{y}_{\triangleleft}^{\mathbf{b}}/t}^{t \mathbf{y}_{\triangleright}^{\mathbf{b}}/t} \int dx \left( \frac{1}{2} u_{\star}(t)^2 - \frac{1}{2} (\partial_x p(t))^2 \right). \quad (8.21)$$

At  $t = \mathfrak{t}$ , using  $h_{\star}(t) = f_{\star}$ ,  $\sigma_{\mathbf{c}}(t-t) = \mathbf{x}_{\mathbf{c}}$ , and Property (I) verifies that  $G(\mathfrak{t}) = G$ . Using Properties (I)–(II), it is not hard to check that  $h_{\star}(t, \sigma_{\mathbf{c}}(t-t)) \rightarrow 0$  as  $t \rightarrow 0$ , for all  $\mathbf{c}$ , and that  $u_{\star}$  and  $\partial_x p(t)$  are uniformly bounded on  $\mathcal{C}_{\mathbf{b}}$  for all  $\mathbf{b}$ . These properties together give that  $G(t) \rightarrow 0$  as  $t \rightarrow 0$ . We seek to show that, for all  $\mathfrak{t}'$  except when shocks merge,

$$\left( \frac{d}{dt} G \right)(\mathfrak{t}') = \frac{1}{24} \sum_{\mathbf{a}=1}^{n'} \mathbf{m}_{\mathbf{a}}'^3 - \sum_{\mathbf{a}=1}^{n'} \frac{1}{2} \mathbf{m}_{\mathbf{a}}' \dot{\sigma}_{\mathbf{a}}^2. \quad (8.22)$$

Once this is done, using  $\sigma_{\mathbf{c}} = \xi_{\mathbf{c}}$ , integrating both sides over  $t \in (0, \mathfrak{t}]$ , and comparing the result with  $L_{\text{SHE}}(\vec{\mathbf{m}}) = \mathbb{L}_{[0, \mathfrak{t}]}(\sum_{\mathbf{c}} \mathbf{m}_{\mathbf{c}} \xi_{\mathbf{c}})$  (see (2.21)) will yield the desired result  $G = L_{\text{SHE}}(\vec{\mathbf{m}})$ .



Let us differentiate the first sum in (8.21) at  $t = t'$  and simplify the result. Take any  $t'$  at which no shocks merge, write the first sum in (8.21) as  $\sum_{a=1}^{n'} m'_a h_*(t, \sigma_a(t-t))$ , differentiate this expression in  $t$  with the aid of Properties (IV). Doing so gives

$$\frac{d}{dt}(\text{first sum in (8.21)})|_{t=t'} = \sum_{a=1}^{n'} m'_a \cdot \left( \frac{1}{2} (u_{*,a}^-)^2 - \dot{\sigma}_a \cdot u_{*,a}^- \right). \quad (8.23)$$

Combine (8.16), (8.3), and (2.34) to get the relation  $u_{*,a}^- = \frac{1}{2} m'_a + \dot{\sigma}_a$  and insert it into the right hand side of (8.23). The result gives  $\frac{d}{dt}(\text{first sum in (8.21)})|_{t=t'} = \frac{1}{8} \sum_{a=1}^{n'} m_a'^3 - \sum_{a=1}^{n'} \frac{1}{2} m'_a \dot{\sigma}_a^2$ .

Next we treat the second sum in (8.21). The contribution of differentiating the boundary points  $ty_{\triangleleft}^b/t$  and  $ty_{\triangleright}^b/t$  is zero, because the integrand evaluates to zero at those points by Property (I). Next, recall from Property (II) that  $u_*$  is piecewise constant within the integral in (8.21). This gives

$$\int_{ty_{\triangleright}^b/t}^{ty_{\triangleleft}^b/t} dx \partial_t \frac{1}{2} (u_*)^2 = \frac{1}{2} \sum_{a: \mathcal{C}(a) \subset b} ((u_{*,a}^+)^2 - (u_{*,a}^-)^2) \dot{\sigma}_a. \quad (8.24)$$

Next, straightforward calculations verify the following relation. (One can also use the machinery of entropy-entropy flux pairs to see it, which we will not do here.)

$$\int_{ty_{\triangleright}^b/t}^{ty_{\triangleleft}^b/t} dx \partial_t \frac{1}{2} (\partial_x p(t))^2 = \frac{1}{3} (\partial_x p(t, ty_{\triangleleft}^b/t))^3 - \frac{1}{3} (\partial_x p(t, ty_{\triangleright}^b/t))^3. \quad (8.25)$$

By Property (I), the right hand side of (8.25) is equal to  $\frac{1}{3} u_*(t, ty_{\triangleleft}^b/t)^3 - \frac{1}{3} u_*(t, ty_{\triangleright}^b/t)^3$ . Telescope the last expression into  $\frac{1}{3} \sum_{a: \mathcal{C}(a) \subset b} ((u_{*,a}^+)^3 - (u_{*,a}^-)^3)$ . So far, we have

$$\frac{d}{dt}(\text{second sum in (8.21)})|_{t=t'} = \sum_{a=1}^{n'} \left( \frac{((u_{*,a}^+)^2 + (u_{*,a}^-)^2)(-\dot{\sigma}_a)}{2} - \frac{(u_{*,a}^+)^3 - (u_{*,a}^-)^3}{3} \right). \quad (8.26)$$

Inset (8.16) into the right hand side of (8.26) and simplify the result. Doing so gives  $\frac{d}{dt}(\text{second sum in (8.21)})|_{t=t'} = -\sum_{a=1}^{n'} (u_{*,a}^+ - u_{*,a}^-)^3/12$ , which, by (8.19), is equal to  $-\sum_{a=1}^{n'} m_a'^3/12$ .

Combining the results in the last two paragraphs gives (8.22) and hence completes the proof.

## Data availability

No data was used for the research described in the article.

## Appendix A. Basic properties

**Proof of the scaling identities (3.1).** First, it is not hard to check that, for any  $\lambda \in \mathfrak{m}\mathcal{P}(\mathbb{R})$ ,

$$F[\tfrac{1}{\mathfrak{m}}\mathfrak{S}_{\mathfrak{m}}\lambda](x) = \tfrac{1}{\mathfrak{m}}F[\lambda](\mathfrak{m}x), \quad \mathfrak{X}[\tfrac{1}{\mathfrak{m}}\mathfrak{S}_{\mathfrak{m}}\lambda](a) = \tfrac{1}{\mathfrak{m}}\mathfrak{X}[\lambda](\mathfrak{m}a), \quad \Phi[\tfrac{1}{\mathfrak{m}}\mathfrak{S}_{\mathfrak{m}}\lambda](x) = \tfrac{1}{\mathfrak{m}}\Phi[\lambda](\mathfrak{m}x). \quad (\text{A.1})$$

Using the preceding scaling relations for  $\mathfrak{X}$  and  $\Phi$  in (2.12) and performing a change of variables  $\mathfrak{m}a \mapsto a$  give the desired scaling relation  $\mathbb{I}(\tfrac{1}{\mathfrak{m}}\mathfrak{S}_{\mathfrak{m}}\mu) = \mathbb{I}(\mu)/\mathfrak{m}^3$  for  $\mathbb{I}$ . To prove the scaling relation for  $\mathcal{I}$ , use (A.1) in (2.10) and call  $\mathbf{g}(s, x) := \mathbf{h}(s, x/\mathfrak{m})$ . After being simplified, the result reads

$$\mathcal{I}(\tfrac{1}{\mathfrak{m}}\mathfrak{S}_{\mathfrak{m}}\mu) = \sup_{\mathbf{g}} \left\{ \tfrac{1}{\mathfrak{m}}\Lambda(\mu, \mathbf{g}) - \tfrac{\mathfrak{m}}{2} \int_0^t ds \langle \mu, (\partial_x \mathbf{g})^2 \rangle \right\}, \quad (\text{A.2})$$

where the supremum runs over  $\mathbf{g} \in \mathcal{C}_b^{1,1}([0, t], \mathbb{R})$ . A supremum of this form can be expressed as a supremum of a Rayleigh quotient: For  $u, v \in (0, \infty)$ ,

$$\sup_{\mathbf{g}} \left\{ u\Lambda(\mu, \mathbf{g}) - \tfrac{v}{2} \int_{[0, t]} ds \langle \mu, (\partial_x \mathbf{g})^2 \rangle \right\} = \frac{\sup_{\mathbf{g}} (u\Lambda(\mu, \mathbf{g}))^2}{2v \int_{[0, t]} ds \langle \mu, (\partial_x \mathbf{g})^2 \rangle}. \quad (\text{A.3})$$

with the convention that  $0/0 := 0$  and that  $b^2/0 := \infty$  for  $b \neq 0$ . To see why, rewrite the supremum on the left hand side as a supremum over  $b\mathbf{g}$ , with  $b \in \mathbb{R}$ , and optimize over  $b$  for a fixed  $\mathbf{g}$ . Using the Rayleigh quotient expression in (A.2) gives  $\mathcal{I}(\tfrac{1}{\mathfrak{m}}\mathfrak{S}_{\mathfrak{m}}\mu) = \mathcal{I}(\mu)/\mathfrak{m}^3$ .  $\square$

**Proof of Lemma 3.2(b)–(c).** To prove Part (b), assume the contrary: There exist  $s_1, s_2, \dots \in [0, t]$  and  $\varepsilon_0 > 0$  such that  $\langle \mu(s_k), \mathbf{1}_{\mathbb{R} \setminus (-k, k)} \rangle \geq \varepsilon_0 > 0$  for all  $k$ . After passing to a subsequence, we have  $s_k \rightarrow s_0 \in [0, t]$ . Since  $\mathbb{R} \setminus (-k, k)$  is closed and since  $\mu(s)$  is continuous in  $s$ ,  $\limsup_{\ell \rightarrow \infty} \langle \mu(s_\ell), \mathbf{1}_{\mathbb{R} \setminus (-k, k)} \rangle \leq \langle \mu(s_0), \mathbf{1}_{\mathbb{R} \setminus (-k, k)} \rangle$  for every  $k \in \mathbb{Z}_{>0}$ . The left hand side is at least  $\varepsilon_0$ . Sending  $k \rightarrow \infty$  gives  $0 < \varepsilon_0 \leq \liminf_{k \rightarrow \infty} \langle \mu(s_0), \mathbf{1}_{\mathbb{R} \setminus [-k, k]} \rangle$ , contradicting  $\mu(s_0) \in \mathcal{P}(\mathbb{R})$ .

To prove Part (c), write  $\mathfrak{X}[\mu(s)](a) = \mathfrak{X}(s, a)$  and let  $\omega(s, a) := \limsup_{s' \rightarrow s} \mathfrak{X}(s', a) - \liminf_{s' \rightarrow s} \mathfrak{X}(s', a)$  denote the oscillation in time of  $\mathfrak{X}$  at  $(s, a)$ . Fix  $k \in \mathbb{Z}_{>0}$ . By (3.4),

$$\begin{aligned} & \text{dist}(\mu(s_1), \mu(s_2)) \\ & \geq 2^{-k} \min \left\{ 1, \int_0^1 da |\mathfrak{X}(s_1, a) - \mathfrak{X}(s_2, a)| \mathbf{1}_{\{|\mathfrak{X}(s_1, a)| \leq k\}} \mathbf{1}_{\{|\mathfrak{X}(s_2, a)| \leq k\}} \right\}. \end{aligned} \quad (\text{A.4})$$

This inequality and the continuity of  $\mu(s)$  in  $s$  imply that the set  $\{a \in [0, 1] : \omega(s, a) > 0, \text{ for some } s \text{ with } |\mathfrak{X}(s, a)| < k\}$  has zero Lebesgue measure. Combining this with the result of Part (b) gives the desired result.  $\square$

**Proof of Properties (a)–(e) after (4.6).** To prove Property (a), observe that  $\mathcal{C}_b^{1,1}([s_1, s_3], \mathbb{R})$  can be embedded into  $\mathcal{C}_b^{1,1}([s_1, s_2], \mathbb{R})$  and  $\mathcal{C}_b^{1,1}([s_2, s_3], \mathbb{R})$ , so the inequality  $\geq$  follows. For the reverse inequality, take any  $h_{12} \in \mathcal{C}_b^{1,1}([s_1, s_2], \mathbb{R})$  and  $h_{23} \in \mathcal{C}_b^{1,1}([s_2, s_3], \mathbb{R})$  and concatenate them as follows: Take any  $\varphi \in \mathcal{C}^\infty(\mathbb{R})$  that is increasing, with  $\varphi|_{(-\infty, 0]} = 0$  and  $\varphi|_{[1, \infty)} = 1$ , set  $\varphi_\delta(s) := \varphi(s/\delta)$ , and consider

$$h_\delta(s, x) := h_{12}(s, x)\varphi_\delta(s_2 - s) + h_{23}(s, x)\varphi_\delta(s - s_2) \in \mathcal{C}_b^{1,2}([s_1, s_3], \mathbb{R}). \quad (\text{A.5})$$

Indeed,  $\int_{s_1}^{s_3} ds \langle \mu, (\partial_x h_\delta)^2 \rangle \rightarrow \int_{s_1}^{s_2} ds \langle \mu, (\partial_x h_{12})^2 \rangle + \int_{s_2}^{s_3} ds \langle \mu, (\partial_x h_{23})^2 \rangle$  as  $\delta \rightarrow 0$ . Next, write

$$\int_{s_1}^{s_3} ds \langle \mu, \partial_s h_\delta \rangle - \int_{s_1}^{s_2} ds \langle \mu, \partial_s h_{12} \rangle - \int_{s_2}^{s_3} ds \langle \mu, \partial_s h_{23} \rangle \quad (\text{A.6})$$

$$= - \int_{s_2 - \delta}^{s_2} ds \langle \mu, h_{12} \partial_s \varphi_\delta \rangle + \int_{s_2}^{s_2 + \delta} ds \langle \mu, h_{23} \partial_s \varphi_\delta \rangle + O(\delta) \quad (\text{A.7})$$

It is not hard to check that (A.7) converges to  $-\langle \mu(s_2), h_{12}(s_2) \rangle + \langle \mu(s_2), h_{23}(s_2) \rangle$  as  $\delta \rightarrow 0$ . The preceding results together verify Property (a).

Property (b) follows from Property (a) and the property that  $\mathcal{I}^\circ_{[s', s'']} \geq 0$ ; the latter property is readily seen from (4.6). Property (c) follows since the expression within the supremum in (4.6) is linear in  $\mu$ . Property (d) follows by renaming the test function  $h(s, x - y) := \tilde{h}(s, x)$  in (4.6). Property (e) follows since for a time-independent  $\mu = \lambda \in \mathcal{P}(\mathbb{R})$  the first two terms within the supremum in (4.6) cancel with each other, and the last term is nonpositive.  $\square$

**Proof of (2.22).** We consider  $\mathfrak{m} = 1$  only; the result for general  $\mathfrak{m} > 0$  follows from the result for  $\mathfrak{m} = 1$  through the scaling argument in the proof of (3.1). Without loss of generality, assume  $\partial_s \mathfrak{X}[\mu] \in \mathcal{L}^2([0, t] \times [0, 1])$ ; otherwise both sides of (2.22) are  $+\infty$  by definition. Write  $\mathfrak{X}[\mu(s)](a) = \mathfrak{X}(s, a)$  and  $\Phi[\mu(s)](x) = \Phi(s, x)$ . Expand the integrand in (2.12) and insert the result into the Left Hand Side (LHS) of (2.22) to get

$$\begin{aligned} \text{LHS of (2.22)} &= \frac{t}{24} + \langle \mu(s)^{\otimes 2}, \frac{1}{4}|x - x'| \rangle \Big|_{s=0}^{s=t} - \int_0^t ds \int_0^1 da \frac{1}{2} (\partial_s \mathfrak{X})^2 \\ &\quad + A_2 - A_3, \end{aligned} \quad (\text{A.8})$$

where  $A_2 := \int_{[0, t]} ds \int_{[0, 1]} da (\partial_s \mathfrak{X}) \Phi$ , and  $A_3 := \int_{[0, t]} ds \int_{[0, 1]} da \frac{1}{2} \Phi^2$ . For  $A_2$ , we claim that we can replace  $\Phi(s, a)$  with  $(\frac{1}{2} - a)$  to get

$$A_2 = \int_0^t ds \int_0^1 da (\partial_s \mathfrak{X}[\mu]) \cdot \left(\frac{1}{2} - a\right). \quad (\text{A.9})$$

Recall  $a_{\pm}[\mu(s)](a)$  from (3.7) and write them as  $a_{\pm}(s, a)$ . Given (3.8), this claim follows if we can show, for almost every  $(s, a)$ ,  $(\partial_s \mathfrak{X})(s, \cdot)$  is almost everywhere a constant on  $(a_-(s, a), a_+(s, a))$ . Fix any  $(s_0, a_0)$  such that  $\mathfrak{X}$  is differentiable in  $s$  at  $(s_0, a_0)$  and write  $a_{\pm}(s_0, a_0) =: a_{0, \pm}$ . For every  $a \in (a_{0, -}, a_0)$ ,  $\mathfrak{X}(s_0, a) = \mathfrak{X}(s_0, a_0)$  and  $\mathfrak{X}(s, a) \leq \mathfrak{X}(s, a_0)$  for all  $s$ . If  $\mathfrak{X}$  is differentiable in  $s$  at  $(s_0, a)$ , the preceding properties force  $(\partial_s \mathfrak{X})(s_0, a) = (\partial_s \mathfrak{X})(s_0, a_0)$ . The same conclusion holds for  $a \in (a_0, a_{0, +})$ . Under the assumption that  $\partial_s \mathfrak{X}[\mu] \in \mathcal{L}^2([0, t] \times [0, 1])$ ,  $\mathfrak{X}$  is differentiable in  $s$  Lebesgue a.e., so the claim follows. Evaluate the  $s$  integral in (A.9), use (3.9) for  $f(x) = x$  in reverse, and use Lemma 3.2(a) in reverse to get  $A_2 = \langle \mu(s), x \Phi[\mu(s)] \rangle|_{s=0}^{s=t}$ . Recalling  $\Phi[\mu(s)]$  from (2.8), we recognize the last expression as  $A_2 = \langle \mu(s)^{\otimes 2}, \frac{1}{2} x \operatorname{sgn}(x' - x) \rangle|_{s=0}^{s=t}$ . Since the measure  $\mu(s)^{\otimes 2}$  is symmetric in  $x$  and  $x'$ , we can symmetrize the function  $\frac{1}{2} x \operatorname{sgn}(x' - x)$  to get

$$A_2 = \langle \mu(s)^{\otimes 2}, \frac{1}{4}(x - x') \operatorname{sgn}(x' - x) \rangle|_{s=0}^{s=t} = -\langle \mu(s)^{\otimes 2}, \frac{1}{4}|x - x'| \rangle|_{s=0}^{s=t}. \quad (\text{A.10})$$

As for  $A_3$ , the identity (3.8) gives

$$A_3 = \frac{1}{2} \int_0^t ds \left( \int_0^1 da \left(\frac{1}{2} - a\right)^2 - \sum_{a_-} \int_{a_-}^{a_+} da \left( \left(\frac{1}{2} - a\right)^2 - \left(\frac{1-a_- - a_+}{2}\right)^2 \right) \right), \quad (\text{A.11})$$

where the sum runs over all pairs of  $(a_-, a_+)$  such that  $a_- < a_+$ . Evaluating the integrals and simplifying the result give

$$A_3 = \frac{t}{24} - \int_0^t ds \sum \frac{1}{24} (a_+ - a_-)^3 = \frac{t}{24} - \int_0^t ds \sum_x \frac{1}{24} \langle \mu, \mathbf{1}_{\{x\}} \rangle^3, \quad (\text{A.12})$$

where the sum over  $x$  runs over atoms of  $\mu(s)$ . Inserting (A.10) and (A.12) into (A.8) gives (2.22).  $\square$

## Appendix B. Proof of (6.40)

We take  $[s', s''] = [0, t]$  (whence  $\mathbb{I}_{[0, t]}(\cdot) := \mathbb{I}(\cdot)$ ) to simplify notation.

We say  $\mu$  **well approximates**  $\mu'$  **to within**  $\varepsilon$  if  $\operatorname{dist}_{[0, t]}(\mu, \mu') < \varepsilon$  and  $|\mathbb{I}(\mu) - \mathbb{I}(\mu')| < \varepsilon$ . We say **a class of deviations well approximates**  $\mu'$  if, for every  $\varepsilon > 0$ , there exists a deviation in that class that well approximates  $\mu'$  to within  $\varepsilon$ .

As the first step, we well approximate  $\mu$  by clustering deviations. For  $n \in \mathbb{Z}_{>0}$ , apply the procedure in Section 3.4 with  $\mathbf{m}_1 = \dots = \mathbf{m}_n = 1/n$  to get  $\mu_c$ ,  $\mathbf{c} = 1, \dots, n$ . As explained there, each  $\mu_c(s)$  is continuous in  $s$ . Let  $x_{\gamma, c}(s) := n \langle \mu_c(s), x \rangle$ , which belongs to  $\mathcal{C}[0, t]$  thanks to the continuity of  $\mu_c$  and thanks to Assumption 6.3. Set

$\gamma^{(n)}(s) := \frac{1}{n} \sum_{\mathbf{c}} \delta_{x_{\gamma, \mathbf{c}}}(s)$ , which is clustering. Consider (twice of) the total variation norm  $\|\lambda - \lambda'\|_{\text{tv}} := \sup_{x \in \mathbb{R}} |F[\lambda](x) - F[\lambda'](x)|$ . We claim that  $\|\gamma^{(n)}(s) - \mu(s)\|_{\text{tv}} \leq 1/n$ . To see why, consider  $S_{\mathbf{c}} := \{x : F[\mu(s)](x) \in [(\mathbf{c} - 1)/n, \mathbf{c}/n]\}$  for  $\mathbf{c} = 1, \dots, n$  separately. Within each  $S_{\mathbf{c}}$ , the function  $F[\gamma^{(n)}(s)]$  makes a single jump from the level  $(\mathbf{c} - 1)/n$  to the level  $\mathbf{c}/n$ ; this is seen from the construction of  $\gamma^{(n)}$ . Hence the difference  $|F[\gamma^{(n)}(s)](x) - F[\mu(s)](x)|$  is at most  $1/n$ , so the claim follows. We have that  $\gamma^{(n)}(s) \rightarrow \mu(s)$  in the total variation norm, uniformly over  $s \in [0, \mathfrak{t}]$ . This implies  $\text{dist}_{[0, \mathfrak{t}]}(\gamma^{(n)}, \mu) \rightarrow 0$ . Next, we show  $\mathbb{I}(\gamma^{(n)}) \rightarrow \mathbb{I}(\mu)$ . Note that by construction,  $\mathfrak{X}[\gamma^{(n)}(s)]|_{a \in I_{\mathbf{c}}} = \gamma_{\mathbf{c}}(s) := n \langle \mu(s), x \rangle = n \int_{I_{\mathbf{c}}} da \mathfrak{X}[\mu(s)](a)$ , where  $I_{\mathbf{c}} := [(\mathbf{c} - 1)/n, \mathbf{c}/n]$ . Namely,  $\mathfrak{X}[\gamma^{(n)}(s)]$  is obtained by averaging  $\mathfrak{X}[\mu(s)]$  over intervals of length  $1/n$ . By using this property and  $\partial_s \mathfrak{X}[\mu] \in \mathcal{L}^2([0, \mathfrak{t}] \times [0, 1])$ , which follows from the assumption  $\mathcal{I}(\mu) = \mathbb{I}(\mu) < \infty$ , it is not hard to check that  $\partial_s \mathfrak{X}[\gamma^{(n)}] \rightarrow \partial_s \mathfrak{X}[\mu]$  in  $\mathcal{L}^2([0, \mathfrak{t}] \times [0, 1])$ . Also,  $\|\Phi[\gamma^{(n)}(s)] - \Phi[\mu(s)]\|_{\infty} \leq \|\frac{1}{2} \text{sgn}\|_{\infty} \|\gamma^{(n)}(s) - \mu(s)\|_{\text{tv}} = \frac{1}{2} \|\gamma^{(n)}(s) - \mu(s)\|_{\text{tv}} \leq 1/(2n)$ . Combining these properties gives  $\mathbb{I}(\gamma^{(n)}) \rightarrow \mathbb{I}(\mu)$ .

As the second step, writing  $\gamma^{(n)} =: \gamma$  to simplify notation and fixing an arbitrary  $\varepsilon > 0$ , we will approximate  $\gamma$  by a finitely-changing-clustering  $\zeta$  to within  $\varepsilon$ . A deviation  $\zeta$  being **finitely-changing-clustering** means that it is clustering and that there exist  $0 = s_0 < s_1 < \dots < s_{\ell} = \mathfrak{t}$  such that, within each  $(s_{k-1}, s_k)$ , each pair of clusters either never touch or completely coincide. We will construct a sequence  $\zeta^{(1)}, \dots, \zeta^{(n)}$  of clustering deviations that satisfy the following conditions.

- (i)  $\text{dist}_{[0, \mathfrak{t}]}(\zeta^{(j-1)}, \zeta^{(j)}) \leq \varepsilon/n$  and  $|\mathbb{I}(\zeta^{(j-1)}) - \mathbb{I}(\zeta^{(j)})| \leq \varepsilon/n$ , with the convention  $\zeta^{(0)} := \gamma$ .
- (ii) Each  $\zeta^{(j)}$  has  $n$  clusters  $\zeta_1^{(j)} \leq \dots \leq \zeta_n^{(j)}$  of mass  $1/n$ , and the finitely-changing property holds up to the index  $j$ . More precisely, there exist  $0 = s_0 < s_1 < \dots < s_{\ell} = \mathfrak{t}$ , which may depend on  $j$ , such that within each  $(s_{k-1}, s_k)$ , each pair in  $\zeta_1^{(j)}, \dots, \zeta_j^{(j)}$  either never touches or completely coincides.

Once constructed,  $\zeta^{(n)} =: \zeta$  gives the finitely-change clustering deviation.

We now construct the  $\zeta^{(j)}$ s by induction on  $j$ . For  $j = 1$ , simply take  $\zeta^{(1)} = \gamma$ . Assume  $\zeta^{(1)}, \dots, \zeta^{(j-1)}$  have been constructed. We will construct  $\zeta^{(j)} =: \zeta^{\text{new}}$  out of  $\zeta^{(j-1)} =: \zeta^{\text{old}}$ . To begin the construction, keep all but the  $j$ th clusters unchanged:  $\zeta_{j'}^{\text{new}} := \zeta_{j'}^{\text{old}}$  for all  $j' \neq j$ . Next, to construct the  $j$ th cluster for  $\zeta^{\text{new}}$ , consider  $O = \{s \in (0, \mathfrak{t}) : \zeta_{j-1}^{\text{old}}(s) < \zeta_j^{\text{old}}(s)\}$ , which is open, and write  $O$  as the union of countable disjoint open intervals  $(a_1, b_1), (a_2, b_2), \dots$ . By Condition (i),  $\mathbb{I}(\zeta^{\text{old}}) < \infty$ , so  $\dot{\zeta}_{\mathbf{c}}^{\text{old}} \in \mathcal{L}^2[0, \mathfrak{t}]$  for all  $\mathbf{c}$ . Given this property, find an  $\ell_0$  large enough such that, with  $O'' := \cup_{\ell > \ell_0} (a_{\ell}, b_{\ell})$ ,

$$\int_{O''} ds \sum_{\mathbf{c}=j-1, j} \frac{1}{n} |\dot{\zeta}_{\mathbf{c}}^{\text{old}}| \leq \frac{\varepsilon}{n}, \quad \int_{O''} ds \frac{1}{2n} \sum_{\mathbf{c}=j-1, j} (|\dot{\zeta}_{\mathbf{c}}^{\text{old}}| + 1)^2 \leq \frac{\varepsilon}{n}. \quad (\text{B.1})$$

Set  $O' := (a_1, b_1) \cup \dots \cup (a_{\ell_0}, b_{\ell_0})$ . Keep the  $j$ th cluster unchanged within  $O'$  and perturb it to match the  $(j-1)$ th cluster outside  $O'$ . More explicitly,  $\zeta_j^{\text{new}}|_{O'} := \zeta_j^{\text{old}}|_{O'}$  and  $\zeta_j^{\text{new}}|_{[0, \mathfrak{t}] \setminus O'} := \zeta_{j-1}^{\text{old}}|_{[0, \mathfrak{t}] \setminus O'}$ .

We next check that the  $\zeta^{\text{new}}$  so constructed satisfies the required conditions. First, given that  $\zeta^{\text{old}}$  satisfies Condition (ii) up to the index  $j-1$ , it is not hard to check that  $\zeta^{\text{new}}$  satisfies Condition (ii) up to the index  $j$ . Move on to checking Condition (i). Recall that  $\zeta^{\text{new}}$  differs from  $\zeta^{\text{old}}$  only at the  $j$ th cluster. Further, the difference occurs only on  $O''$ . On  $O''$ , where  $\zeta_j^{\text{new}}$  and  $\zeta_j^{\text{old}}$  differ, we have  $\zeta_j^{\text{new}}|_{O''} = \zeta_{j-1}^{\text{old}}|_{O''}$ . These properties together with (6.10) and  $|\Phi[\cdot]| \leq 1/2$  give

$$\sup_{s \in [0, \mathfrak{t}]} \frac{1}{n} \sum_{\mathfrak{c}=1}^n |\zeta_{\mathfrak{c}}^{\text{new}}(s) - \zeta_{\mathfrak{c}}^{\text{old}}(s)| \leq \int_{O''} ds \frac{1}{n} |\dot{\zeta}_{j-1}^{\text{old}} - \dot{\zeta}_j^{\text{old}}|, \quad (\text{B.2})$$

$$|\mathbb{I}(\zeta^{\text{new}}) - \mathbb{I}(\zeta^{\text{old}})| \leq \int_{O''} ds \frac{1}{2n} \sum_{\mathfrak{c}=j-1, j} (|\dot{\zeta}_{\mathfrak{c}}^{\text{old}}| + \frac{1}{2})^2. \quad (\text{B.3})$$

By (B.1), the right hand sides of (B.2)–(B.3) are bounded by  $\varepsilon/n$ . By (3.6), the left hand side of (B.2) bounds  $\text{dist}_{[0, \mathfrak{t}]}(\zeta^{\text{new}}, \zeta^{\text{old}})$  from above.

Finally, we will approximate  $\zeta$  by PL-clustering deviations. Let  $0 = s_0 < s_1 < \dots < s_{\ell} = \mathfrak{t}$  prescribe the intervals on which the clusters of  $\zeta$  either never touch or completely coincide. Further partition each  $[s_{k-1}, s_k]$  into smaller intervals of equal length. Linearly interpolate the trajectories of the clusters of  $\zeta$  with respect to the smaller intervals. Use the resulting piecewise linear trajectories to build a PL-clustering  $\xi$ . Within each  $(s_{k-1}, s_k)$ , since the clusters of  $\zeta$  either never touch or completely coincide, the same property holds for  $\xi$ . By using this property, (3.6), and (6.10), it is not hard to check that, as the mesh of the sub partitions tends to zero,  $\text{dist}_{[0, \mathfrak{t}]}(\xi, \zeta) \rightarrow 0$  and  $\mathbb{I}(\xi) \rightarrow \mathbb{I}(\zeta)$ .

## Appendix C. Girsanov's transform

Here, we pack Girsanov's transform in ways convenient for our applications. Fix  $[s', s''] \subset [0, \mathfrak{t}]$ , let  $\mathbb{P}$  be the law of  $(X_i(s))_{s \in [Ts', Ts''], i=1, \dots, N}$  under (2.2) given  $(X_i(Ts'))_{i=1, \dots, N}$ , take deterministic  $v_i$  and  $J_i \subset \{1, \dots, N\}$ ,  $i = 1, \dots, N$ , and consider the law  $\mathbb{Q}$  such that

$$\text{under } \mathbb{Q}, \quad dX_i = \sum_{j \in \{1, \dots, N\} \setminus J_i} \frac{1}{2} \text{sgn}(X_j - X_i) ds + N v_i ds + dB_i, \quad s \in [Ts', Ts''], \quad (\text{C.1})$$

with the same given  $(X_i(Ts'))_{i=1, \dots, N}$ .

### Lemma C.1.

(a) Assume  $J_i = \emptyset$  for all  $i$  and take any  $r' \in (0, \infty)$ . For any measurable  $\mathcal{E} \subset \mathcal{C}([s', s''], \mathbb{R})^N$  with  $\mathcal{E} \subset \{|B_i(Ts) - B_i(Ts')| < NTr', \forall s \in [s', s''], i = 1, \dots, N\}$ ,

$$\frac{1}{N^3 T} \log \mathbb{P}[\mathcal{E}] \geq \frac{1}{N^3 T} \log \mathbb{Q}[\mathcal{E}] - \frac{(s' - s)}{2N} \sum_{i=1}^N v_i^2 - \frac{r'}{N} \sum_{i=1}^N |v_i|. \quad (\text{C.2})$$

(b) For any  $p > 1$  and any measurable  $\mathcal{E} \subset \mathcal{C}([s', s''], \mathbb{R})^N$ ,

$$\frac{1}{N^3 T} \log \mathbb{P}[\mathcal{E}] \geq \frac{p}{N^3 T} \log \mathbb{Q}[\mathcal{E}] - \frac{p(s'' - s')}{p-1} \frac{1}{2N} \sum_{i=1}^N \left( |v_i| + \frac{|J_i|^2}{2N} \right)^2. \quad (\text{C.3})$$

**Proof.** Set  $A := \int_{Ts'}^{Ts''} \sum_{i=1}^N dB_i(Nv_i - \sum_{j \in J_i} \text{sgn}(X_j - X_i)/2)$  and let  $\langle A \rangle = \int_{Ts'}^{Ts''} \sum_{i=1}^N ds (Nv_i - \sum_{j \in J_i} \text{sgn}(X_j - X_i)/2)^2$  denote its quadratic variation. Slightly abusing notation, we write the *expectation* under  $\mathbb{Q}$  also as  $\mathbb{Q}$ . Girsanov's transform gives  $\mathbb{P}[\mathcal{E}] = \mathbb{Q}[\mathbf{1}_{\mathcal{E}} \exp(A - \langle A \rangle/2)]$ . For Part (a), using  $\mathbf{1}_{\mathcal{E}} A \geq -N^2 T r' \sum_i |v_i|$  and  $\langle A \rangle = N^2 T (s'' - s') \sum_i v_i^2$  gives the desired result. For Part (b), use Hölder's inequality

$$\mathbb{Q}[\mathcal{E}] \leq \mathbb{Q}[\mathbf{1}_{\mathcal{E}} \exp(A - \frac{1}{2}\langle A \rangle)]^{1/p} \mathbb{Q}[\exp(\frac{1}{(p-1)}(-A + \frac{1}{2}\langle A \rangle))]^{(p-1)/p}, \quad (\text{C.4})$$

evaluate the last expectation

$$\mathbb{Q}\left[e^{\frac{1}{p-1}(-A + \frac{1}{2}\langle A \rangle)}\right] = \mathbb{Q}\left[\exp\left(\frac{p}{2(p-1)^2} \int_{Ts'}^{Ts''} ds \sum_{i=1}^N \left(Nv_i + \sum_{j \in J_i} \frac{1}{2} \text{sgn}(X_j - X_i)\right)^2\right)\right], \quad (\text{C.5})$$

bound the summand in  $\sum_i$  by  $(N|v_i| + |J_i|/2)^2$ , and replace  $\mathbb{Q}[\mathbf{1}_{\mathcal{E}} \exp(A - \langle A \rangle/2)]$  with  $\mathbb{P}[\mathcal{E}]$  in (C.4). Doing so gives

$$\mathbb{Q}[\mathcal{E}] \leq \mathbb{P}[\mathcal{E}]^{1/p} \exp\left(\frac{(s'' - s')T}{2(p-1)} \sum_{i=1}^N \left(N|v_i| + \frac{|J_i|}{2N}\right)^2\right). \quad (\text{C.6})$$

Applying  $\frac{p}{N^3 T} \log(\cdot)$  to both sides and simplifying the result give (C.3).  $\square$

## References

- [1] Tomer Asida, Eli Livne, Baruch Meerson, Large fluctuations of a Kardar–Parisi–Zhang interface on a half line: the height statistics at a shifted point, *Phys. Rev. E* 99 (4) (2019) 042132.
- [2] Sayan Banerjee, Amarjit Budhiraja, Domains of attraction of invariant distributions of the infinite Atlas model, *Ann. Probab.* 50 (4) (2022) 1610–1646.
- [3] Sayan Banerjee, Brendan Brown, Dimension-free local convergence and perturbations for reflected Brownian motions, *Ann. Appl. Probab.* 33 (1) (2023) 376–416.
- [4] Alexei Borodin, Ivan Corwin, Moments and Lyapunov exponents for the parabolic Anderson model, *Ann. Appl. Probab.* 24 (3) (2014) 1172–1198.
- [5] Adrian D. Banner, E. Robert Fernholz, Tomoyuki Ichiba, Ioannis Karatzas, and Vassilios Papathanakos, Hybrid Atlas models, *Ann. Appl. Probab.* 21 (2) (2011) 609–644.
- [6] Adrian D. Banner, Robert Fernholz, Ioannis Karatzas, Atlas models of equity markets, *Ann. Appl. Probab.* 15 (4) (2005) 2296–2330.

- [7] Mattia Cafasso, Tom Claeys, A Riemann–Hilbert approach to the lower tail of the Kardar–Parisi–Zhang equation, *Commun. Pure Appl. Math.* 75 (3) (2022) 493–540.
- [8] Mattia Cafasso, Tom Claeys, Giulio Ruzza, Airy kernel determinant solutions to the KdV equation and integro-differential Painlevé equations, *Commun. Math. Phys.* 386 (2) (2021) 1107–1153.
- [9] Chen Le, Robert C. Dalang, Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions, *Ann. Probab.* 43 (6) (2015) 3006–3051.
- [10] Manuel Cabezas, Amir Dembo, Andrey Sarantsev, Vladas Sidoravicius, Brownian particles with rank-dependent drifts: out-of-equilibrium behavior, *Commun. Pure Appl. Math.* 72 (7) (2019) 1424–1458.
- [11] Ivan Corwin, Promit Ghosal, KPZ equation tails for general initial data, *Electron. J. Probab.* 25 (2020).
- [12] Ivan Corwin, Promit Ghosal, Lower tail of the KPZ equation, *Duke Math. J.* 169 (7) (2020) 1329–1395.
- [13] Ivan Corwin, Promit Ghosal, Alexandre Krajenbrink, Pierre Le Doussal, Li-Cheng Tsai, Coulomb-gas electrostatics controls large fluctuations of the Kardar–Parisi–Zhang equation, *Phys. Rev. Lett.* 121 (6) (2018) 060201.
- [14] Ivan Corwin, Alan Hammond, Brownian Gibbs property for Airy line ensembles, *Invent. Math.* 195 (2) (2014) 441–508.
- [15] Ivan Corwin, Alan Hammond, KPZ line ensemble, *Probab. Theory Relat. Fields* 166 (2016) 67–185.
- [16] Xia Chen, Precise intermittency for the parabolic Anderson equation with an  $(1+1)$ -dimensional time-space white noise, in: *Annales de l’IHP Probabilités et Statistiques*, vol. 51, 2015, pp. 1486–1499.
- [17] LeChen, Yaozhong Hu, David Nualart, Regularity and Strict Positivity of Densities for the Nonlinear Stochastic Heat Equation, vol. 273, American Mathematical Society, 2021.
- [18] Daniel Conus, Mathew Joseph, Davar Khoshnevisan, On the chaotic character of the stochastic heat equation, before the onset of intermittency, *Ann. Probab.* 41 (3B) (2013) 2225–2260.
- [19] Ivan Corwin, The Kardar–Parisi–Zhang equation and universality class, *Random Matrices: Theory Appl.* 1 (01) (2012) 1130001.
- [20] Ivan Corwin, Hao Shen, Some recent progress in singular stochastic PDEs, *Bull. Am. Math. Soc.* 57 (2019) 409–454.
- [21] Ajay Chandra, Hendrik Weber, Stochastic PDEs, regularity structures, and interacting particle systems, *Ann. Fac. Sci. Toulouse Math.* 26 (2017) 847–909.
- [22] Sayan Das, Promit Ghosal, Law of iterated logarithms and fractal properties of the KPZ equation, *arXiv:2101.00730*, 2021.
- [23] Amir Dembo, Milton Jara, Stefano Olla, The infinite Atlas process: convergence to equilibrium, *Ann. Inst. Henri Poincaré Probab. Stat.* 55 (2) (2019) 607–619.
- [24] Amir Dembo, Mykhaylo Shkolnikov, SR Srinivasa Varadhan, Ofer Zeitouni, Large deviations for diffusions interacting through their ranks, *Commun. Pure Appl. Math.* 69 (7) (2016) 1259–1313.
- [25] Amir Dembo, Li-Cheng Tsai, Equilibrium fluctuation of the Atlas model, *Ann. Probab.* 45 (6B) (2017) 4529–4560.
- [26] Sayan Das, Li-Cheng Tsai, Fractional moments of the stochastic heat equation, *Ann. Inst. Henri Poincaré Probab. Stat.* 57 (2) (2021) 778–799.
- [27] Lawrence C. Evans, *Partial Differential Equations*, vol. 19, American Mathematical Society, 2022.
- [28] E. Robert Fernholz, Stochastic portfolio theory, in: *Stochastic Portfolio Theory*, Springer, 2002, pp. 1–24.
- [29] Robert Fernholz, Ioannis Karatzas, Stochastic portfolio theory: an overview, *Handb. Numer. Anal.* 15 (2009) 89–167.
- [30] Shirshendu Ganguly, Milind Hegde, Sharp upper tail estimates and limit shapes for the KPZ equation via the tangent method, *arXiv:2208.08922*, 2022.
- [31] Jürgen Gärtner, Wolfgang König, Stanislav Molchanov, Geometric characterization of intermittency in the parabolic Anderson model, *Ann. Probab.* 35 (2) (2007) 439–499.
- [32] Promit Ghosal, Yier Lin, Lyapunov exponents of the SHE under general initial data, in: *Annales de l’Institut Henri Poincaré (B) Probabilités et Statistiques*, vol. 59, Institut Henri Poincaré, 2023, pp. 476–502.
- [33] Yves Pierre, Gaudreau Lamarre, Yier Lin, Li-Cheng Tsai, KPZ equation with a small noise, deep upper tail and limit shape, *Probab. Theory Relat. Fields* (2023) 1–36.
- [34] Jürgen Gärtner, Stanislav A. Molchanov, Parabolic problems for the Anderson model, *Commun. Math. Phys.* 132 (3) (1990) 613–655.



- [35] Nicos Georgiou, Timo Seppäläinen, Large deviation rate functions for the partition function in a log-gamma distributed random potential, *Ann. Probab.* 41 (6) (2013) 4248–4286.
- [36] Alexander K. Hartmann, Alexandre Krajenbrink, Pierre Le Doussal, Probing large deviations of the Kardar–Parisi–Zhang equation at short times with an importance sampling of directed polymers in random media, *Phys. Rev. E* 101 (1) (2020) 012134.
- [37] Yaozhong Hu, Khoa Lê, Asymptotics of the density of parabolic Anderson random fields 58 (1) (2022) 105–133.
- [38] Alexander K. Hartmann, Pierre Le Doussal, Satya N. Majumdar, Alberto Rosso, Gregory Schehr, High-precision simulation of the height distribution for the KPZ equation, *Europhys. Lett.* 121 (6) (2018) 67004.
- [39] Alexander K. Hartmann, Baruch Meerson, Pavel Sasorov, Optimal paths of nonequilibrium stochastic fields: the Kardar–Parisi–Zhang interface as a test case, *Phys. Rev. Res.* 1 (3) (2019) 032043.
- [40] Alexander K. Hartmann, Baruch Meerson, Pavel Sasorov, Observing symmetry-broken optimal paths of the stationary Kardar–Parisi–Zhang interface via a large-deviation sampling of directed polymers in random media, *Phys. Rev. E* 104 (5) (2021) 054125.
- [41] Yaozhong Hu, David Nualart, Stochastic heat equation driven by fractional noise and local time, *Probab. Theory Relat. Fields* 143 (1) (2009) 285–328.
- [42] Tomoyuki Ichiba, Ioannis Karatzas, On collisions of Brownian particles, *Ann. Appl. Probab.* 20 (3) (2010) 951–977.
- [43] Tomoyuki Ichiba, Ioannis Karatzas, Mykhaylo Shkolnikov, Strong solutions of stochastic equations with rank-based coefficients, *Probab. Theory Relat. Fields* 156 (1–2) (2013) 229–248.
- [44] Tomoyuki Ichiba, Andrey Sarantsev, Yet another condition for absence of collisions for competing Brownian particles, *Electron. Commun. Probab.* 22 (8) (2017) 1–7.
- [45] Chris Janjigian, Large deviations of the free energy in the O’Connell–Yor polymer, *J. Stat. Phys.* 160 (4) (2015) 1054–1080.
- [46] Michael Janas, Alex Kamenev, Baruch Meerson, Dynamical phase transition in large-deviation statistics of the Kardar–Parisi–Zhang equation, *Phys. Rev. E* 94 (3) (2016) 032133.
- [47] Khoshnevisan Davar, *Analysis of Stochastic Partial Differential Equations*, vol. 119, American Mathematical Soc., 2014.
- [48] Yujin H. Kim, The lower tail of the half-space KPZ equation, *Stoch. Process. Appl.* 142 (2021) 365–406.
- [49] I.V. Kolokolov, S.E. Korshunov, Optimal fluctuation approach to a directed polymer in a random medium, *Phys. Rev. B* 75 (14) (2007) 140201.
- [50] I.V. Kolokolov, S.E. Korshunov, Universal and nonuniversal tails of distribution functions in the directed polymer and Kardar–Parisi–Zhang problems, *Phys. Rev. B* 78 (2) (2008) 024206.
- [51] I.V. Kolokolov, S.E. Korshunov, Explicit solution of the optimal fluctuation problem for an elastic string in a random medium, *Phys. Rev. E* 80 (3) (2009) 031107.
- [52] Davar Khoshnevisan, Kunwoo Kim, Yimin Xiao, Intermittency and multifractality: a case study via parabolic stochastic pdes, *Ann. Probab.* 45 (6A) (2017) 3697–3751.
- [53] Claude Kipnis, Claudio Landim, *Scaling Limits of Interacting Particle Systems*, vol. 320, Springer Science & Business Media, 1998.
- [54] Alexandre Krajenbrink, Pierre Le Doussal, Exact short-time height distribution in the one-dimensional Kardar–Parisi–Zhang equation with Brownian initial condition, *Phys. Rev. E* 96 (2) (2017) 020102.
- [55] Alexandre Krajenbrink, Pierre Le Doussal, Large fluctuations of the KPZ equation in a half-space, *SciPost Phys.* 5 (2018) 032.
- [56] Alexandre Krajenbrink, Pierre Le Doussal, Simple derivation of the  $(-\lambda H)^{5/2}$  tail for the 1D KPZ equation, *J. Stat. Mech. Theory Exp.* 2018 (6) (2018) 063210.
- [57] Alexandre Krajenbrink, Pierre Le Doussal, Linear statistics and pushed Coulomb gas at the edge of  $\beta$ -random matrices: four paths to large deviations, *Europhys. Lett.* 125 (2) (2019) 20009.
- [58] Alexandre Krajenbrink, Pierre Le Doussal, Inverse scattering of the Zakharov–Shabat system solves the weak noise theory of the Kardar–Parisi–Zhang equation, *Phys. Rev. Lett.* 127 (6) (2021) 064101.
- [59] Alexandre Krajenbrink, Pierre Le Doussal, Inverse scattering solution of the weak noise theory of the Kardar–Parisi–Zhang equation with flat and Brownian initial conditions, *Phys. Rev. E* 105 (2022) 054142.
- [60] Alexandre Krajenbrink, Pierre Le Doussal, Crossover from the macroscopic fluctuation theory to the Kardar–Parisi–Zhang equation controls the large deviations beyond Einstein’s diffusion, *Phys. Rev. E* 107 (1) (2023) 014137.

- [61] Alexandre Krajenbrink, Pierre Le Doussal, Sylvain Prolhac, Systematic time expansion for the Kardar–Parisi–Zhang equation, linear statistics of the GUE at the edge and trapped fermions, *Nucl. Phys. B* 936 (2018) 239–305.
- [62] Alex Kamenev, Baruch Meerson, Pavel V. Sasorov, Short-time height distribution in the one-dimensional Kardar–Parisi–Zhang equation: starting from a parabola, *Phys. Rev. E* 94 (3) (2016) 032108.
- [63] Claude Kipnis, Stefano Olla, SR Srinivasa Varadhan, Hydrodynamics and large deviation for simple exclusion processes, *Commun. Pure Appl. Math.* 42 (2) (1989) 115–137.
- [64] Mehran Kardar, Giorgio Parisi, Yi-Cheng Zhang, Dynamic scaling of growing interfaces, *Phys. Rev. Lett.* 56 (9) (1986) 889.
- [65] Alexandre Krajenbrink, Beyond the typical fluctuations: a journey to the large deviations in the Kardar–Parisi–Zhang growth model, PhD thesis, PSL Research University, 2019.
- [66] Alexandre Krajenbrink, From Painlevé to Zakharov–Shabat and beyond: Fredholm determinants and integro-differential hierarchies, *J. Phys. Math. Theor.* 54 (3) (2020) 035001.
- [67] Pierre Le Doussal, Large deviations for the Kardar–Parisi–Zhang equation from the Kadomtsev–Petviashvili equation, *J. Stat. Mech. Theory Exp.* (2020) 043201.
- [68] Pierre Le Doussal, Ranked diffusion, delta Bose gas, and Burgers equation, *Phys. Rev. E* 105 (1) (2022) L012103.
- [69] Pierre Le Doussal, Satya N. Majumdar, Alberto Rosso, Grégory Schehr, Exact short-time height distribution in the one-dimensional Kardar–Parisi–Zhang equation and edge fermions at high temperature, *Phys. Rev. Lett.* 117 (7) (2016) 070403.
- [70] Pierre Le Doussal, Satya N. Majumdar, Grégory Schehr, Large deviations for the height in 1D Kardar–Parisi–Zhang growth at late times, *Europhys. Lett.* 113 (6) (2016) 60004.
- [71] Yier Lin, Lyapunov exponents of the half-line SHE, *J. Stat. Phys.* 183 (3) (2021) 1–34.
- [72] Yier Lin, Multi-point Lyapunov exponents of the stochastic heat equation, arXiv preprint, arXiv:2305.19966, 2023.
- [73] Yier Lin, Li-Cheng Tsai, Short time large deviations of the KPZ equation, *Commun. Math. Phys.* 386 (1) (2021) 359–393.
- [74] Yier Lin, Li-Cheng Tsai, A lower-tail limit in the weak noise theory, arXiv:2210.05629, 2022.
- [75] Yier Lin, Li-Cheng Tsai, Spacetime limit shapes of the KPZ equation in the upper tails, preprint, 2023.
- [76] Baruch Meerson, Eytan Katzav, Arkady Vilenkin, Large deviations of surface height in the Kardar–Parisi–Zhang equation, *Phys. Rev. Lett.* 116 (7) (2016) 070601.
- [77] Baruch Meerson, Johannes Schmidt, Height distribution tails in the Kardar–Parisi–Zhang equation with Brownian initial conditions, *J. Stat. Mech. Theory Exp.* 2017 (10) (2017) 103207.
- [78] James Munkres, *Topology*, 2 edition, Pearson College Div, 2000.
- [79] Baruch Meerson, Arkady Vilenkin, Large fluctuations of a Kardar–Parisi–Zhang interface on a half line, *Phys. Rev. E* 98 (3) (2018) 032145.
- [80] Soumik Pal, Jim Pitman, One-dimensional Brownian particle systems with rank-dependent drifts, *Ann. Appl. Probab.* (2008).
- [81] Jeremy Quastel, Herbert Spohn, The one-dimensional KPZ equation and its universality class, *J. Stat. Phys.* 160 (4) (2015) 965–984.
- [82] Jeremy Quastel, Introduction to KPZ, *Curr. Dev. Math.* 2011 (1) (2011).
- [83] Andrey Sarantsev, Triple and simultaneous collisions of competing Brownian particles, *Electron. J. Probab.* 20 (29) (2015) 1–28.
- [84] Andrey Sarantsev, Infinite systems of competing Brownian particles, *Ann. Inst. Henri Poincaré Probab. Stat.* 53 (4) (2017) 2279–2315.
- [85] Andrey Sarantsev, Two-sided infinite systems of competing Brownian particles, *ESAIM Probab. Stat.* 21 (2017) 317–349.
- [86] Andrey Sarantsev, Comparison techniques for competing Brownian particles, *J. Theor. Probab.* 32 (2) (2019) 545–585.
- [87] Mykhaylo Shkolnikov, Competing particle systems evolving by interacting Lévy processes, *Ann. Appl. Probab.* 21 (5) (2011) 1911–1932.
- [88] Naftali R. Smith, Alex Kamenev, Baruch Meerson, Landau theory of the short-time dynamical phase transitions of the Kardar–Parisi–Zhang interface, *Phys. Rev. E* 97 (4) (2018) 042130.
- [89] R. Naftali Smith, Baruch Meerson, Exact short-time height distribution for the flat Kardar–Parisi–Zhang interface, *Phys. Rev. E* 97 (5) (2018) 052110.
- [90] Pavel Sasorov, Baruch Meerson, Sylvain Prolhac, Large deviations of surface height in the 1+1-dimensional Kardar–Parisi–Zhang equation: exact long-time results for  $\lambda h < 0$ , *J. Stat. Mech. Theory Exp.* 2017 (6) (2017) 063203.

- [91] R. Naftali Smith, Baruch Meerson, Pavel Sasorov, Finite-size effects in the short-time height distribution of the Kardar–Parisi–Zhang equation, *J. Stat. Mech. Theory Exp.* 2018 (2) (2018) 023202.
- [92] R. Naftali Smith, Baruch Meerson, Arkady Vilenkin, Time-averaged height distribution of the Kardar–Parisi–Zhang interface, *J. Stat. Mech. Theory Exp.* 2019 (5) (2019) 053207.
- [93] Andrey Sarantsev, Li-Cheng Tsai, Stationary gap distributions for infinite systems of competing Brownian particles, *Electron. J. Probab.* 22 (56) (2017) 1–20.
- [94] Li-Cheng Tsai, Stationary distributions of the Atlas model, *Electron. Commun. Probab.* 23 (10) (2018) 1–10.
- [95] Li-Cheng Tsai, Exact lower-tail large deviations of the KPZ equation, *Duke Math. J.* 1 (1) (2022) 1–44.
- [96] Li-Cheng Tsai, Integrability in the weak noise theory, [arXiv:2204.00614](https://arxiv.org/abs/2204.00614), 2022.