

# Universal Kardar-Parisi-Zhang fluctuations for moderate deviations of random walks in random environments

Jacob Hass <sup>1</sup>, Hindy Drillick <sup>2</sup>, Ivan Corwin <sup>2</sup>, and Eric I. Corwin <sup>1</sup>

<sup>1</sup>*Department of Physics and Materials Science Institute, University of Oregon, Eugene, Oregon 97403, USA*

<sup>2</sup>*Department of Mathematics, Columbia University, New York, New York 10027, USA*



(Received 23 April 2025; revised 26 July 2025; accepted 23 September 2025; published 26 January 2026)

The theory of diffusion seeks to describe the motion of particles in a chaotic environment. Classical theory models individual particles as independent random walkers, effectively forgetting that particles evolve together in the same environment. *Random walks in a random environment* models treat the environment as a random space-time field that biases the motion of particles based on where they are in the environment. We provide a universality result for the moderate deviations of the transition probability of this model over a wide class of choices of random environments. In particular, we show the convergence of moments to those of the multiplicative noise stochastic heat equation, whose logarithm is the Kardar-Parisi-Zhang equation. The environment only filters into the scaling limit through one parameter, which depends explicitly on the statistical description of the environment. This forms the basis for our introduction, in Hass *et al.* [*Phys. Rev. Lett.* **133**, 267102 (2024)], of the extreme diffusion coefficient.

DOI: [10.1103/4mhw-4xkj](https://doi.org/10.1103/4mhw-4xkj)

## I. INTRODUCTION

Classical diffusion theory is used to describe the statistical behavior of agents in a wide variety of systems with random or chaotic fluctuations such as stock prices [1,2], the movement of photons in a scattering medium [3], and the spread of viruses [4]. The theory is built upon the assumption that each particle can be modeled as effectively independent random walkers [5–12], thus reducing everything to one parameter—the Einstein diffusion coefficient—that characterizes the variance per unit time of those walkers. Despite the simplicity of this model, it is remarkably effective in describing the statistics of the bulk or typical diffusing particle in a system of many particles [13–15]. However, a recent series of works [16–25] has provided evidence that the independent random walkers model for many-particle diffusion fails to accurately predict the behavior of extreme particles, i.e., those that travel the farthest or fastest and are often of considerable interest [26–36]. Those works have focused on models of *random walks in a random environment* (RWRE) with environments that are quickly mixing in space and time. Although typical particles in such models for many-particle diffusion behave as if the environment was statistically averaged, i.e., reduced to the independent random walker model, the extreme particles manifest novel behavior in the presence of a common random environment. The purpose of this paper is to unwind how the statistical description of the random environment translates into the statistical behavior of the extreme particles.

In the model considered here, an environment is a collection of transition probabilities, indexed by discrete one-dimensional space and time. At each time, particles choose their next spatial location independently according to the transition probabilities at that space-time point. The random environment comes from choosing those transition probability

distributions randomly so as to be independent and identically distributed over all of space and time. We consider the many-step transition probability—which is random in light of the random environment. We show that as the time span grows in a scale  $N$ , the *moderate deviations* of this transition probability in a spatial scale  $N^{3/4}$  (i.e., the probability of a single particle moving  $N^{3/4}$  to the right of its mean velocity) converges to the solution of the *stochastic heat equation* (SHE), whose logarithm solves the *Kardar-Parisi-Zhang* (KPZ) equation, and which is given by

$$\partial_T Z = \frac{1}{2} \partial_X^2 Z + \sqrt{2D_0} Z \eta, \quad (1)$$

where  $X$  is distance and  $T$  is time. Above we have  $Z(0, X) = \delta(X)$  (Dirac delta function) initial data and take  $\eta(X, T)$  to be space-time Gaussian white noise [i.e.,  $\mathbb{E}[\eta(X, T)] = 0$  and  $\mathbb{E}[\eta(X, T)\eta(X', T')] = \delta(X - X')\delta(T - T')]$ .

The only parameter in this limit is the *noise strength*  $D_0 \in \mathbb{R}_{>0}$ , which we determine in Eq. (4) explicitly in terms of the statistical description of the random environment. We show convergence at the level of moments, and, though our methods can be applied to general moments, we restrict ourselves to the first and second moments which suffice to pin down the value of  $D_0$ . Similar results have been demonstrated recently in the mathematics literature in work of Parekh [23]. That work, as well as ours, can be seen as a generalization of the nearest-neighbor or sticky-Brownian motion models studied in Refs. [19,22,37] to arbitrary random environments. Since moderate deviations translate into the behavior of the maximum of many draws from a probability distribution, our results translate into results about the statistical behavior of the extreme particles under certain scalings of time and the number of particles [18,20,25], as well as the location of first

passage barriers [18,21,25] and the correlations between the positions of particles [38].

The remainder of the paper proceeds as follows. In Sec. II, we clearly describe the RWRE model and some relevant notation. Section III contains our main results, namely the convergence of the first two moments and the determination of  $D_0$ . Our approach is explained in Sec. IV—in particular we use a variant of the replica method to relate moments to discrete local times and then employ two probabilistic tools—the Tanaka formula for Brownian local times and the stationary measures for certain Markov chains—to relate this to the replica formulas for the SHE moments. Section VI fleshes out this local time convergence approach, and Sec. VII relates those calculations back to the moments of the moderate deviations.

## II. THE RWRE MODEL

The random environment in which we will consider random walks is defined as  $\xi := \{\xi_{t,x} : x \in \mathbb{Z}, t \in \mathbb{Z}_{\geq 0}\}$ , where each  $\xi_{t,x}$  is a probability distribution on  $\mathbb{Z}$ . The  $\xi_{t,x}$  are themselves random, chosen to be independent and identically distributed over all choices of  $t$  and  $x$ , with a distribution  $\nu$  that completely determines the RWRE model. Since  $\xi_{t,x}$  are not correlated in time, the environment is memoryless. Given an environment  $\xi$ , we will consider  $N$  independent random walkers; when a walker is at position  $x$  and  $t$  time, they choose how far to jump in the next time interval according to the distribution  $\xi_{t,x}$ . Different walkers at the same site use the same distribution  $\xi_{t,x}$  but sample from it independently. Thus, there are two levels of randomness in the model, that of the environment in which all particles evolve together and that of the independently sampled walker trajectories. We introduce some notation to keep track of these two levels of randomness and illustrate it with an example.

We let  $\mathcal{M}_1(\mathbb{Z})$  denote the space of probability distributions on  $\mathbb{Z}$  and  $\nu$  a probability distribution on  $\mathcal{M}_1(\mathbb{Z})$ . Each choice of  $\nu$  corresponds to a different law on the random environment  $\xi$ , where we sample  $\xi_{t,x} \in \mathcal{M}_1(\mathbb{Z})$  according to  $\nu$ , independently for each  $(t, x)$ . We will restrict ourselves to only consider  $\nu$  which are supported on finite range probability distributions, i.e., if  $\xi$  is distributed according to  $\nu$ , then for some  $M$  sufficiently large,  $\xi(j) = 0$  for all  $j > M$ . We let  $\mathbb{E}_\nu[\bullet]$  denote the expectation of a function  $\bullet$  of  $\xi$  with respect to the product measure where each  $\xi_{t,x}$  is independent with distribution  $\nu$ . To illustrate this notation, let us note one example. For integer  $k > 0$ , and  $\alpha_{-k}, \dots, \alpha_k > 0$ , let  $(\xi(-k), \dots, \xi(k))$  be Dirichlet distributed with the specified  $\alpha$  parameters [and let all other  $\xi(j) = 0$  for  $j \notin \{-k, k\}$ ]. Precisely, restricted to the simplex where  $\xi(-k), \dots, \xi(k) \in [0, 1]$  and  $\xi(-k) + \dots + \xi(k) = 1$ , the probability density function is proportional to  $\prod_{j=-k}^k \xi(j)^{\alpha_j-1}$ . This defines a random probability distribution on  $\mathbb{Z}$  (or, rather,  $\{-k, \dots, k\}$ ) and hence defines one choice of  $\nu$ .

Given an instance of the environment  $\xi$ , we will consider  $N \in \mathbb{N}$  random walkers evolving independently with jump distributions determined by  $\xi$  and denote the corresponding probability measure by  $\mathbb{P}^\xi$ . Such a model naturally lends itself to the study of extreme particles and allows one to extract extreme value statistics such as those of the

particle which is at the maximum distance at a given time,  $\max(\{R^1(t), \dots, R^N(t)\})$ , or that which is first to cross a given boundary [25].

More precisely,  $\mathbb{P}^\xi$  is a probability measure on the sample space  $(\mathbb{Z}^N)^{\mathbb{Z}_{\geq 0}}$ , where we think of an element  $\mathbf{R} = (R^1, \dots, R^N)$  of this space as the time-indexed spatial trajectory of  $N$  walkers  $R^1(t), \dots, R^N(t)$  for  $t \in \mathbb{Z}_{\geq 0}$ . For a given realization  $\xi$  of the environment, we define  $\mathbb{P}^\xi$  to be the measure on  $(\mathbb{Z}^N)^{\mathbb{Z}_{\geq 0}}$  such that  $(R^1, \dots, R^N)$  are distributed as  $N$  independent random walks all started at 0, with independent increments given, for each  $k \in \{1, \dots, N\}$  by

$$\mathbb{P}^\xi(R^k(t+1) = x+i \mid R^k(t) = x) = \xi_{x,t}(i).$$

In other words, each walker uses the jump distribution at their current time and position to determine the size of their next jump. When the environment  $\xi$  is random, the measure  $\mathbb{P}^\xi$  is random as well. We will be interested below in understanding the random probability distribution of a single walker, i.e.,  $\mathbb{P}^\xi(R^k(t) = x)$ . In that case, we will adopt the notation  $R(t) = R^1(t)$ , dropping the superscript.

We define the *annealed* probability measure  $\mathbf{P}_\nu$  on the same  $N$ -path sample space  $(\mathbb{Z}^N)^{\mathbb{Z}_{\geq 0}}$  by averaging  $\mathbb{P}^\xi$  over  $\xi$ , according to the earlier described  $\nu$ -dependent product measure  $\mathbb{E}_\nu$ . In other words, we define the annealed probability measure  $\mathbf{P}_\nu(\cdot) := \mathbb{E}_\nu[\mathbb{P}^\xi(\cdot)]$  and the corresponding expectation  $\mathbf{E}_\nu[\cdot] = \mathbb{E}_\nu[\mathbb{E}^\xi[\cdot]]$ . For any  $k \leq N$ , we call the process  $(R^1, \dots, R^k)$  with law given by  $\mathbf{P}_\nu$  the *k-point motion*.

We denote the *average* environment as  $\bar{\xi} = \mathbb{E}_\nu[\xi_{t,x}]$ , which can also be thought of as the ensemble-averaged environment, meaning what you see if you average over a large swath of space and time. This average is the same at each site, hence, no  $t, x$  subscript is needed. For simplicity, we will only consider models with  $\nu$  that have a drift-free average environment, i.e., such that  $\sum_{i \in \mathbb{Z}} \bar{\xi}(i)i = 0$ , though we expect similar results to hold for general  $\nu$  after changing to a suitable moving reference frame.

Since we will make extensive use of them, we describe here the one- and two-point motions  $R^1$  and  $(R^1, R^2)$  under the annealed measure  $\mathbf{P}_\nu$ . The one-point motion  $R^1$  is an independent and identically distributed increment random walk that jumps from position  $x$  to position  $x+j$  with probability  $\bar{\xi}(j)$  for all  $j \in \mathbb{Z}$ . The same is true marginally of  $R^2$ ; however, it is not quite true that  $R^1$  and  $R^2$  are independent. When  $R^1(t) \neq R^2(t)$ ,  $(R^1, R^2)$  evolves according to independent  $\bar{\xi}$  distributed increments, i.e., if  $R^1(t) = x$  and  $R^2(t) = y$ , then  $R^1(t+1) = x+j$  and  $R^2(t+1) = y+k$  with probability given by the product  $\bar{\xi}(j)\bar{\xi}(k)$ . However, when  $R^1(t) = R^2(t) = x$ , the probability that  $R^1(t+1) = x+j$  and  $R^2(t+1) = x+k$  is equal to  $\mathbb{E}_\nu[\xi_{t,x}(j)\xi_{t,x}(k)]$ .

It follows from the above two-point motion transition formulas that we can think of them as a pair of *sticky* random walks since when the two walkers are at the same site at time  $t$ , they are more likely to also be at the same site at time  $t+1$  than two independent random walks. An important object in our analysis will be the difference  $V(t) := R^1(t) - R^2(t)$  between the two-point motion  $(R^1, R^2)$ . In particular, we will study the gap  $\Delta(t) := |V(t)|$ . The walk  $V(t)$  is a Markov chain

on  $\mathbb{Z}$  where the transition probability from state  $i$  to  $j$  is

$$p(i, j) = \begin{cases} \sum_{k \in \mathbb{Z}} \mathbb{E}_v[\xi_{t,x}(k)\xi_{t,x}(k-j)] & \text{if } i = 0 \\ \sum_{k \in \mathbb{Z}} \bar{\xi}(k)\bar{\xi}(k+j-i) & \text{if } i \neq 0 \end{cases} \quad (2)$$

which only depends on  $j-i$  and whether or not  $i=0$ . Furthermore, for  $i, j \neq 0$ , we have  $p(i, j) = p(j, i)$ .

For clarity, we now summarize the class of models we study. We consider RWRE models with the following assumptions:

- (1) Particles take finite-range jumps.
- (2) The environment is given by an independent and identically distributed family of random probability distributions  $\xi_{t,x}$ , indexed by times  $t$  and positions  $x$ .
- (3) The average environment is drift free such that  $\sum_{i \in \mathbb{Z}} \bar{\xi}(i)i = 0$ .

### III. MAIN RESULT: RWRE LIMIT TO SHE

Based on previous results for nearest neighbor RWRE jump models [17–19,22], we expect that in the moderate deviation regime, the tail probability will have environmental dependent fluctuations that have a SHE scaling limit. We will precisely state the scaling of this tail probability and demonstrate convergence to the SHE at the level of the first and second moments. Our methods can be readily generalized to show convergence of higher moments, but the first and second moments suffice in identifying the *a priori* unknown noise strength coefficient  $D_0$ .

We study the (random) tail probability [recalling our shorthand convention that  $R(t) = R^1(t)$ ]

$$\mathbb{P}^\xi \left( R(NT) \geq N^{3/4}T + \frac{\sum_{i \in \mathbb{Z}} \bar{\xi}(i)i^3}{2(2D)^2} N^{1/2}T + \sqrt{2DN}X \right) \quad (3)$$

in the limit that  $N \rightarrow \infty$ , where  $N \in \mathbb{Z}_{>0}$ ,  $T \in N^{-1}\mathbb{Z}_{\geq 0}$ ,  $X \in \mathbb{R}$ , and  $D := \frac{1}{2} \sum_{i \in \mathbb{Z}} \bar{\xi}(i)i^2$  is the diffusion coefficient of a random walk in the average environment. Note that here and below we use  $x$  and  $t$  to represent discrete positions in space and time while  $X$  and  $T$  represent continuous variables corresponding to the KPZ equation. We have dropped the superscript on the random walk in Eq. (3) since all  $N$  random walks are independent and identically distributed.

Our main result, given by Eq. (40), states that the fluctuations of Eq. (3) converge to the SHE in Eq. (1) with time  $T$ , position  $X$ , and noise strength

$$D_0 = \frac{\lambda_{\text{ext}}}{(2D)^{3/2}}, \quad (4)$$

where  $\lambda_{\text{ext}}$  is a characteristic length defined as follows. Let  $Y$  be a random variable distributed according to  $\xi_{t,x}$ , which can be thought of as a single step of a random walk  $R^1$ . Further let  $\mu(l)$  be the unique invariant measure of  $V(t)$  with the normalization  $\mu(0) = 1$  [we address the necessary modifications for the case where the invariant measure of  $V(t)$  is not unique in Sec. VIA 1]. We then have

$$\lambda_{\text{ext}} = \frac{\text{Var}_v(\mathbb{E}^\xi[Y])}{\sum_{l=0}^{\infty} \tilde{\mu}(l) \mathbf{E}_v[\Delta(t+1) - \Delta(t) \mid \Delta(t) = l]}, \quad (5)$$

where

$$\text{Var}_v(\mathbb{E}^\xi[Y]) = \mathbb{E}_v \left[ \left( \sum_{i \in \mathbb{Z}} \xi_{t,x}(i)i \right)^2 \right],$$

$$\tilde{\mu}(l) = \begin{cases} 1 & \text{if } l = 0 \\ 2\mu(l) & \text{if } l > 0 \end{cases}.$$

Although the form of  $\lambda_{\text{ext}}$  is rather complex, it simplifies significantly for a number of distributions, which we discuss in [25].

The term  $\text{Var}_v(\mathbb{E}^\xi[Y])$  satisfies  $\text{Var}_v(\mathbb{E}^\xi[Y]) \in [0, 2D]$  and represents the variance, over all environments, of the drift of a single jump of a random walk. When  $\text{Var}_v(\mathbb{E}^\xi[Y]) = 2D$ , the jump distribution is only supported on a single site, so the walks are perfectly sticky. In the limit that  $\text{Var}_v(\mathbb{E}^\xi[Y]) \rightarrow 2D$ , sticky Brownian motion can be recovered as studied in Ref. [37]. When  $\text{Var}_v(\mathbb{E}^\xi[Y]) = 0$ , the drift of the system is deterministic. Since we only study environments that are net drift free, this means  $\mathbb{E}^\xi[Y] = 0$  with probability 1 (see Refs. [23,24] for a discussion of this case).

The invariant measure,  $\tilde{\mu}(l)$ , can be interpreted as how much time two particles spend a distance  $l$  apart as compared to being at the same site in the long-time limit. One could potentially extract the invariant measure using microrheology as follows [39,40]. Start two particles in the same environment. Let them diffuse for a long time and then measure their distance apart. Repeat this many times in different environments to build a histogram of the distance apart. After normalizing the histogram to 1 when the particles are at the same location, the histogram represents the invariant measure,  $\tilde{\mu}(l)$ . Alternatively, the invariant measure could be measured by observing a large system of diffusing particles. Letting the system run for a long time and then building a histogram of the distances between pairs of particles yields the invariant measure,  $\tilde{\mu}(l)$ .

The term  $\mathbf{E}_v[\Delta(t+1) - \Delta(t) \mid \Delta(t) = l]$  does not depend on  $t$  as we are conditioning on  $\Delta(t) = l$ . This term quantifies the change in the distance between two random walks. The expectation value could be extracted using microrheology as follows. Start two particles a distance  $l$  apart and record how much farther apart they are after a small time step. Repeating this for many trials yields the expectation value. Since we are assuming a finite-size jump distribution, the expected value will be 0 for large-enough  $l$ , and thus, the sum in the denominator will be finite.

Our main result makes a remarkable connection between the microrheological quantity,  $\lambda_{\text{ext}}$ , and extreme values. Our results agree with those for nearest neighbor RWRE models [19,22] and a specific case of a more general class of RWRE models studied in Ref. [23].

### IV. OVERVIEW OF OUR DERIVATION

To show convergence of the tail probability to the SHE, we work with a more general setup of which the tail probability in Eq. (3) is a special case. We study the probability mass function smoothed by a spatial test function. We include this smoothing because the probability mass distribution at a single lattice site is too noisy to work with (i.e., it depends

on the order-one behavior of the noise). The tail probability can be recovered by choosing the spatial test function to sum over all lattice sites in the tail, as discussed below. Of course, any physical measurement of diffusion involves some smoothing, so this is a natural lens through which to study the model.

Since the RWRE probability mass function is given in terms of a discrete path integral through the random environment, its moments admit representations in terms of interacting random walks, where the interaction relates to the law of the environment and is active when the walks coincide. This description is a discrete analog of the replica method formulas for moments of the SHE in terms of Brownian motions interacting through their local times. The argument presented below identifies a scaling under which the discrete moment formulas converge to their continuum SHE counterparts.

As shown below, the fluctuations of the RWRE tail probability converge to that of the KPZ equation when at lattice positions which scale like  $N^{3/4}$ . This scaling regime is called the “moderate deviation regime” [19]. The central limit regime is found for lattice positions which scale like  $N^{1/2}$ . In this regime, we expect fluctuations in the tail probability to be determined by the Edward-Wilkinson universality class. The large deviation regime is found for lattice positions of order  $N$ . In the large deviation regime, fluctuations in the tail probability have been conjectured to be Tracy-Widom distributed [16–18]. However, this has only been proven rigorously for a specific choice of environment [16]. If one assumes that there are no abrupt changes in the distribution of fluctuations, then one would expect the distribution in the moderate deviation regime to smoothly interpolate between the Gaussian and Tracy-Widom distributions. As the SHE is known to do precisely this, it is natural that we find the fluctuations of the tail probability in the moderate deviation regime converge to the SHE.

We now introduce a smoothed out version of the probability mass function under this moderate deviations scaling. To start, we re-express the location about which we are studying the tail probability as a window of size  $\sqrt{2DN}$  about  $x = c(N)T$ , where

$$c(N) = N^{3/4} + \frac{\sum_{i \in \mathbb{Z}} \tilde{\xi}(i) i^3}{2(2D)^2} N^{1/2}.$$

This specific choice places us in the moderate deviation regime and is justified in detail below. The probability mass distribution at this location in the tail decays to first order like a Gaussian as  $N \rightarrow \infty$ , mimicking the behavior of the tail in the central limit regime. About this Gaussian behavior there are random fluctuations in the probability mass function. To study these fluctuations, we rescale the probability mass function by the first-order Gaussian behavior so the fluctuations stay of order one. This asymptotic Gaussian behavior is encoded in the prefactor  $C(N, T, X)$ , whose particular functional form is chosen for future mathematical convenience. Since the fluctuations at a single lattice point are too noisy to work with directly, we smooth the probability mass function in space by a spatial test function,  $\phi$ .

Putting all of these elements together, we study

$$\mathcal{U}_N(T, \phi) := \mathbb{E}^\xi \left[ C \left( N, T, \frac{R(NT) - c(N)T}{\sqrt{2DN}} \right) \times \phi \left( \frac{R(NT) - c(N)T}{\sqrt{2DN}} \right) \right], \quad (6)$$

where  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is a spatial test function (i.e., a smooth and compactly supported function),  $R = R^1$  is a single random walk in the random environment  $\xi$ , and

$$C(N, T, X) := \frac{\exp \left\{ \frac{c(N)}{2DN^{1/4}} T + \frac{1}{\sqrt{2D}} N^{1/4} X \right\}}{\left( \sum_{i \in \mathbb{Z}} \tilde{\xi}(i) \exp \left\{ \frac{i}{2DN^{1/4}} \right\} \right)^{NT}}$$

encodes the first-order Gaussian behavior of the probability mass distribution in the moderate deviation regime while its higher-order terms conspire to yield convergence to the moments of the KPZ equation. Thus,  $\mathcal{U}_N(T, \phi)$  is a weighted (by the  $C(N, T, X)$  factor) probe of the distribution of  $R(NT)$ , smoothed by a test function, in the vicinity of  $c(N)T$  and in the scale  $\sqrt{2DN}$ .

We now give our general strategy to show the moments (for different  $k$ )  $\mathbb{E}_v[\mathcal{U}_N(T, \phi)^k]$  of Eq. (6) converge to those of the SHE. We fully realize this for the first and second moments, though the approach works similarly for general moments. Consider the first moment. It is immediate, see (9), that  $\mathbb{E}_v[\mathcal{U}_N(T, \phi)]$  can be expressed in terms of an expectation over the one-point motion. Then, due to its multiplicative nature, the  $C(N, T, X)$  prefactor in the definition of  $\mathcal{U}_N(T, \phi)$  can be absorbed as a tilt of the jump distribution of the one-point motion, and hence in the limit yields a Brownian expectation.

Before going to the  $k = 2$  case, let us note that the  $k$ th moment of the SHE integrated against a spatial test function can be written (via the replica method) in terms of the expectation of the local time of  $k$  Brownian motions

$$\begin{aligned} & \mathbb{E} \left[ \left( \int_{\mathbb{R}} \phi(X) Z(X, T) dX \right)^k \right] \\ &= \mathbb{E} \left[ \Phi(\vec{B}) \exp \left\{ D_0 \sum_{i < j} L^{B^i - B^j}(T) \right\} \right] \end{aligned} \quad (7)$$

where  $\mathbb{E}$  on the left is the expectation over the noise  $\eta$  of the SHE, and on the right  $\Phi(\vec{x}) := \phi(x_1) \cdots \phi(x_k)$ , and the expectation  $\mathbb{E}$  is over independent Brownian motions  $B^1, \dots, B^k$  with  $L^{B^i - B^j}(t)$  their pair local time at zero, defined as follows. The local time of a space-time continuous Brownian motion  $B(t)$  with variance  $\sigma^2 t$  is

$$L^B(t) := \lim_{\epsilon \rightarrow 0^+} \frac{\sigma^2}{2\epsilon} \int_0^t \mathbb{1}_{\{-\epsilon < B(s) < \epsilon\}} ds. \quad (8)$$

For  $k = 2$  (and higher), the moment is rewritten in terms of an expectation with respect to the two-point motion. The same tilting argument works provided the two-point motions occupy different spatial locations. When they are at the same site, the two-point motion jump distribution [i.e.,  $p(i, j)$  from (2) with  $i = 0$ ] has some residual effect after tilting which can be written, as in Eq. (21), in terms of a local time contribution. The tilted two-point motion clearly converges to independent



Brownian motions, so the whole challenge is to understand how the discrete local time converges to a limiting Brownian local time.

To illustrate this challenge, note that a random walk that lives entirely on the odd sublattice of  $\mathbb{Z}$  will have zero discrete local time at 0, yet will still converge to a Brownian motion. More relevant to the current situation, the introduction of some stickiness at zero for a random walk will not impact its Brownian limit (provided the stickiness is not tuned in the scaling limit) but will cause the discrete local time at 0 to converge to a constant multiple of the Brownian local time at 0. That constant dilation factor of the local time depends on the degree of stickiness.

Thus, to address the discrete to continuous local time convergence—in particular computing the dilation factor (which translates into the noise strength coefficient  $D_0$ )—we develop an argument based on a discrete version of the *Tanaka formula* and the *Doob-Meyer decomposition*, both tools based on studying random walks and Brownian motions as *martingales*. This will identify a discrete local time, not entirely concentrated at 0, which converges to the Brownian local time at 0. Then we will use the invariant measure of the gap between the two-point motion to identify what portion of that discrete local time comes from the discrete local time at 0 (which is what arises in our moment formulas). This will yield the desired dilation factor and explains the form of  $D_0$  given earlier. The details of this argument are presented below.

## V. CONVERGENCE OF THE FIRST MOMENT

The convergence of the first moment will show that our choice of  $C(N, T, X)$  is the correct prefactor to ensure  $\mathcal{U}_N(T, \phi)$  converges to the first moment of the SHE. We begin by averaging  $\mathcal{U}_N(T, \phi)$  in Eq. (6) over the random environment to obtain

$$\mathbb{E}_v[\mathcal{U}_N(T, \phi)] = \mathbb{E}_v \left[ C \left( N, T, \frac{R(NT) - c(N)T}{\sqrt{2DN}} \right) \times \phi \left( \frac{R(NT) - c(N)T}{\sqrt{2DN}} \right) \right]. \quad (9)$$

The notation used above should be recalled from Sec. II. In particular,  $R = R^1$  is the one-point motion under the annealed measure  $\mathbf{P}_v$  and is the expectation with respect to that measure  $\mathbf{E}_v$ .

We now absorb the prefactor into the expectation by interpreting it as an *exponential tilting* of the independent and identically distributed increments of the one-point motion. We start by breaking up the one-point into its increments,  $R(NT) = \sum_{i=1}^{NT} Y_i$  where  $Y_i$  are independent and identically distributed according to  $\tilde{\xi}$ . Thus, we can rewrite

$$\begin{aligned} C \left( N, T, \frac{R(NT) - c(N)T}{\sqrt{2DN}} \right) &= \frac{\exp \left\{ \frac{R(NT)}{2DN^{1/4}} \right\}}{\left( \sum_{i \in \mathbb{Z}} \tilde{\xi}(i) \exp \left\{ \frac{i}{2DN^{1/4}} \right\} \right)^{NT}} \\ &= \prod_{i=1}^{NT} \frac{\exp \left\{ \frac{Y_i}{2DN^{1/4}} \right\}}{\sum_{i \in \mathbb{Z}} \tilde{\xi}(i) \exp \left\{ \frac{i}{2DN^{1/4}} \right\}}. \end{aligned}$$

Since the  $Y_i$  are independent and identically distributed according to  $\tilde{\xi}$ , this factor can be absorbed as a tilt of the jump distribution for  $Y_i$ . Define the tilted measure  $\tilde{\mathbf{E}}_v$  under which

the  $Y_i$  are independent and identically distributed but now with the probability that  $Y_i = j$  for  $j \in \mathbb{Z}$  given by

$$\frac{\tilde{\xi}(j) \exp \left\{ \frac{j}{2DN^{1/4}} \right\}}{\sum_{i \in \mathbb{Z}} \tilde{\xi}(i) \exp \left\{ \frac{i}{2DN^{1/4}} \right\}}. \quad (10)$$

Notice how the prefactor was chosen to make sure this tilting results in a probability measure. Recalling that  $R(NT) = \sum_{i=1}^{NT} Y_i$ , we thus have shown that Eq. (9) can be rewritten in terms of the tilted measure as

$$\mathbb{E}_v[\mathcal{U}_N(T, \phi)] = \tilde{\mathbf{E}}_v \left[ \phi \left( \frac{R(NT) - c(N)T}{\sqrt{2DN}} \right) \right]. \quad (11)$$

Under the tilted measure, the centered and scaled one-point motion converges to a Brownian motion, i.e.,

$$\lim_{N \rightarrow \infty} \frac{R(NT) - c(N)T}{\sqrt{2DN}} = B(T), \quad (12)$$

where  $B$  is a standard unit variance Brownian motion. To see this, since the increments of  $R$  are independent and identically distributed, it suffices to check that that  $\frac{R(NT) - c(N)T}{\sqrt{2DN}}$  has mean 0 and variance  $T$ , at least up to terms that vanish as  $N \rightarrow \infty$ . Under the tilted distribution Eq. (10), the increments of  $R(NT)$  have mean

$$\begin{aligned} &\frac{\sum_{i \in \mathbb{Z}} \tilde{\xi}(i) i \exp \left\{ \frac{i}{2DN^{1/4}} \right\}}{\sum_{j \in \mathbb{Z}} \tilde{\xi}(j) \exp \left\{ \frac{j}{2DN^{1/4}} \right\}} \\ &= \frac{\sum_{i \in \mathbb{Z}} \tilde{\xi}(i) i \left( 1 + \frac{i}{2DN^{1/4}} + \frac{i^2}{4DN^{1/2}} + \mathcal{O}(N^{-3/4}) \right)}{\sum_{j \in \mathbb{Z}} \tilde{\xi}(j) \left( 1 + \mathcal{O}(N^{-1/2}) \right)} \\ &= \sum_{i \in \mathbb{Z}} \tilde{\xi}(i) \frac{i^2}{2DN^{1/4}} + \sum_{i \in \mathbb{Z}} \tilde{\xi}(i) \frac{i^3}{4D^2N^{1/2}} + \mathcal{O}(N^{-3/4}) \\ &= \frac{c(N)}{N} + \mathcal{O}(N^{-3/4}) \end{aligned}$$

where  $\mathcal{O}(x)$  denotes terms of order  $x$  and lower. The second moment of the increments of  $R(NT)$  are

$$\begin{aligned} &\frac{\sum_{i \in \mathbb{Z}} \tilde{\xi}(i) i^2 \exp \left\{ \frac{i}{2DN^{1/4}} \right\}}{\sum_{j \in \mathbb{Z}} \tilde{\xi}(j) \exp \left\{ \frac{j}{2DN^{1/4}} \right\}} = \frac{\sum_{i \in \mathbb{Z}} \tilde{\xi}(i) i^2 (1 + \mathcal{O}(N^{-1/4}))}{\sum_{j \in \mathbb{Z}} \tilde{\xi}(j) (1 + \mathcal{O}(N^{-1/2}))} \\ &= 2D + \mathcal{O}(N^{-1/4}). \end{aligned}$$

Thus, under the tilted measure  $R(NT)$  has mean  $c(N)T + \mathcal{O}(N^{1/4})$  and variance  $2DNT + \mathcal{O}(N^{3/4})$ , or equivalently  $\frac{R(NT) - c(N)T}{\sqrt{2DN}}$  has a mean of order  $\mathcal{O}(N^{-1/4})$  and variance  $T + \mathcal{O}(N^{-1/4})$ .

Combining Eq. (12) with Eq. (11), we see the first equality,

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}_v[\mathcal{U}_N(T, \phi)] &= \mathbf{E}[\phi(B(T))] \\ &= \mathbf{E} \left[ \int_{\mathbb{R}} \phi(X) Z(X, T) dX \right]. \end{aligned} \quad (13)$$

Here  $\mathbf{E}$  is the expectation with respect to a standard Brownian motion starting at 0, and the second equality (to the first moment of the SHE) follows from in Eq. (7).

## VI. CONVERGENCE OF THE SECOND MOMENT

We now show that the second moment of  $\mathcal{U}(T, \phi)$  converges to the second moment of the SHE with the strength of the noise given by  $D_0 = \frac{\lambda_{\text{ext}}}{(2D)^{3/2}}$ . In doing so, we identify the characteristic length scale  $\lambda_{\text{ext}}$ .

Recall that given an instance of the environment  $\xi$ , the random walks  $R^1, R^2, \dots$  are independent and identically distributed. Thus, in Eq. (6), we could have just as well replaced  $R$  by  $R^1$  or  $R^2$  without changing anything. Doing that, and owing to the independent (given  $\xi$ ) of  $R^1$  and  $R^2$ , it follows that

$$\mathcal{U}_N(T, \phi)^2 = \mathbb{E}^\xi \left[ C\left(N, T, \frac{R^1(NT) - c(N)T}{\sqrt{2DN}}\right) C\left(N, T, \frac{R^2(NT) - c(N)T}{\sqrt{2DN}}\right) \phi\left(\frac{R^1(NT) - c(N)T}{\sqrt{2DN}}\right) \phi\left(\frac{R^2(NT) - c(N)T}{\sqrt{2DN}}\right) \right].$$

This is the first step of the *replica method*.

The next step is to take the expectation over the random environment to get a formula for the second moment. Using the explicit form of the prefactor  $C$ , this yields

$$\mathbb{E}_v[\mathcal{U}_N(T, \phi)^2] = \mathbb{E}_v \left[ \frac{\exp\left\{\frac{R^1(NT) + R^2(NT)}{2DN^{1/4}}\right\}}{\left(\sum_{i \in \mathbb{Z}} \bar{\xi}(i) \exp\left\{\frac{i}{2DN^{1/4}}\right\}\right)^{2NT}} \phi\left(\frac{R^1(NT) - c(N)T}{\sqrt{2DN}}\right) \phi\left(\frac{R^2(NT) - c(N)T}{\sqrt{2DN}}\right) \right], \quad (14)$$

where the pair  $(R^1, R^2)$  is the two-point motion defined earlier. As we did for the first moment, we want to absorb the prefactor as a tilt of the transition probabilities for the two-point motion. When  $R^1 \neq R^2$ , this works exactly as before since the two random walks take independent jumps. When  $R^1 = R^2$ , the two jumps are no longer independent, and an additional factor is needed to ensure that the tilting results in a probability measure. This produces a discrete local time at 0 term in our moment formula, see Eq. (20) below.

To derive Eq. (20), first observe that the denominator inside the expectation in Eq. (14) can be broken into  $NT$  equal factors with each term given by

$$\left( \sum_{i \in \mathbb{Z}} \bar{\xi}(i) \exp\left\{\frac{i}{2DN^{1/4}}\right\} \right)^2 = \sum_{i, j \in \mathbb{Z}} \bar{\xi}(i) \bar{\xi}(j) \exp\left\{\frac{i+j}{2DN^{1/4}}\right\} \quad (15)$$

after expanding into a double sum. When  $R^1(t) \neq R^2(t)$ , this is the right normalization for the tilting factor needed to produce a probability measure for the increments of  $R^1$  and  $R^2$ . When  $R^1(t) = R^2(t)$ , we require a different normalization to get a tilted probability measure since  $\bar{\xi}(i) \bar{\xi}(j)$  should be replaced by  $\mathbb{E}_v[\xi_{t,x}(i) \xi_{t,x}(j)]$  (as the walks are at the same site).

To account for this, we really should have started with a different prefactor,

$$C\left(N, T, \frac{R^1(NT) - c(N)T}{\sqrt{2DN}}\right) C\left(N, T, \frac{R^2(NT) - c(N)T}{\sqrt{2DN}}\right) \exp \left\{ -g((2D)^{-1}N^{-1/4}) \sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}} \right\} \quad (16)$$

where

$$g(\lambda) = \ln \left( \sum_{i, j \in \mathbb{Z}} \mathbb{E}_v[\xi_{t,x}(i) \xi_{t,x}(j)] e^{\lambda(i+j)} \right) - 2 \ln \left( \sum_{i \in \mathbb{Z}} \bar{\xi}(i) e^{\lambda i} \right). \quad (17)$$

If we use this tilting factor, then we obtain a tilted version of the two-point motion which is a Markov chain with the following transition probabilities: When  $R^1(t) \neq R^2(t)$ , then  $R^1(t+1) = R^1(t) + i$  and  $R^2(t+1) = R^2(t) + j$  with probability

$$\frac{\bar{\xi}(i) \bar{\xi}(j) \exp\left\{\frac{i+j}{2DN^{1/4}}\right\}}{\sum_{k, l \in \mathbb{Z}} \bar{\xi}(k) \bar{\xi}(l) \exp\left\{\frac{k+l}{2DN^{1/4}}\right\}}. \quad (18)$$

When  $R^1(t) = R^2(t) = x$ , then  $R^1(t+1) = x + i$  and  $R^2(t+1) = x + j$  with probability

$$\frac{\mathbb{E}_v[\xi_{t,x}(i) \xi_{t,x}(j)] \exp\left\{\frac{i+j}{2DN^{1/4}}\right\}}{\sum_{k, l \in \mathbb{Z}} \mathbb{E}_v[\xi_{t,x}(k) \xi_{t,x}(l)] \exp\left\{\frac{k+l}{2DN^{1/4}}\right\}}. \quad (19)$$

Noting the overload with the notation from the first moment calculation, we let  $\tilde{\mathbb{E}}_v[\bullet]$  denote the expectation value with respect to the tilted probability distribution on  $(R^1, R^2)$  given by the above transition probability.

Taking into account the  $g$  factor we included in Eq. (16), we find that

$$\mathbb{E}_v[\mathcal{U}_N(T, \phi)^2] = \tilde{\mathbb{E}}_v \left[ \exp \left\{ g((2D)^{-1}N^{-1/4}) \sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}} \right\} \phi\left(\frac{R^1(NT) - c(N)T}{\sqrt{2DN}}\right) \phi\left(\frac{R^2(NT) - c(N)T}{\sqrt{2DN}}\right) \right]. \quad (20)$$

By Taylor expansion we find that  $g((2D)^{-1}N^{-1/4}) = \frac{\mathbb{E}_v[(\sum_{i \in \mathbb{Z}} i \xi_{t,x}(i))^2]}{(2D)^2 N^{1/2}} + \mathcal{O}(N^{-3/4}) = \frac{\text{Var}_v(\mathbb{E}^\xi[Y])}{(2D)^2 N^{1/2}} + \mathcal{O}(N^{-3/4})$ , where the second equality comes from observing that  $\text{Var}_v(\mathbb{E}^\xi[Y]) = \mathbb{E}_v[(\sum_{i \in \mathbb{Z}} i \xi_{t,x}(i))^2]$ . The fact that this term behaves like  $N^{-1/2}$  is key since the discrete local time [i.e.,  $\sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}}$ ] needs to be scaled by that factor to have a limit. This consideration is what forces the scaling regime for SHE convergence to be in the  $N^{3/4}$ -depth moderate deviations of the tail probability.

Substituting this expansion in Eq. (20) and using  $\approx$  to denote that we have dropped the lower-order terms, we have

$$\mathbb{E}_v[\mathcal{U}_N(T, \phi)^2] \approx \tilde{\mathbb{E}}_v \left[ \exp \left\{ \frac{\text{Var}_v(\mathbb{E}^\xi[Y])}{(2D)^2 N^{1/2}} \sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}} \right\} \phi \left( \frac{R^1(NT) - c(N)T}{\sqrt{2DN}} \right) \phi \left( \frac{R^2(NT) - c(N)T}{\sqrt{2DN}} \right) \right]. \quad (21)$$

Notice that Eq. (21) looks like a discrete analog of the second moment of the SHE in Eq. (7) since  $\sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}}$  is the discrete local time of  $R^1(i) - R^2(i)$  at 0. Furthermore, besides their sticky interaction when  $R^1 = R^2$ , the tilted two-point motion behaves like independent random walks just as in the first moment case. The only way that the sticky interaction could impact the scaling limit is if it is scaled to become stronger as  $N \rightarrow \infty$  (i.e., so that the time they stay together increases in the scaling limit), or if the size of the jump from  $R^1 = R^2$  were scaled to become longer as  $N \rightarrow \infty$  (i.e., so as to result in a discontinuous jump in the limit process). This is not the case, and hence under the measure  $\tilde{\mathbb{E}}_v$ , the pair  $(\frac{R^1(NT) - c(N)T}{\sqrt{2DN}}, \frac{R^2(NT) - c(N)T}{\sqrt{2DN}})$  converge to independent variance one Brownian motions, as in the first moment case.

Our final step is to identify the scaling limit of the discrete local time  $\sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}}$ , namely that it converges to a constant (which we identify) times the local time of the difference of two independent standard Brownian motions. This is the most subtle and interesting part of the argument since it more broadly explains in what sense discrete local times converge to their continuum Brownian analogs. In particular, although the discrete random walks converge to independent Brownian motions, the discrete local time need not converge to the local time of their Brownian motion limits. For example, consider a simple symmetric random walk on  $\mathbb{Z}$  modified so as to stay at 0 with probability 1/2 and go to  $\pm 1$  each with probability 1/4 (i.e., it is sticky at the origin). The discrete random walks will converge to Brownian motion, but the discrete local time will converge to a constant (explicitly calculable and exceeding one) times the local time of standard Brownian motion. Identifying this constant which rescales the local time limit is the key to identifying the noise strength of the SHE. Specifically, in what follows, we show that

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{\sqrt{2DN}} \sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}} &= \frac{1}{\sum_{l=0}^{\infty} \tilde{\mu}(l) \mathbb{E}_v[\Delta(t+1) - \Delta(t) \mid \Delta(t) = l]} L_0^{B^1-B^2}(T), \end{aligned} \quad (22)$$

where  $B^1$  and  $B^2$  are independent standard Brownian motions starting at 0. Notice that the prefactor of the local time on the right-hand side is the denominator of  $\lambda_{\text{ext}}$  in Eq. (5). We show this in Sec. VIA.

Putting the above together, we conclude that

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}_v[\mathcal{U}_N(T, \phi)^2] &= \mathbf{E} \left[ e^{\frac{\lambda_{\text{ext}}}{(2D)^{3/2}} L_0^{B^1-B^2}(T)} \phi(B^1(T)) \phi(B^2(T)) \right], \end{aligned}$$

where  $\mathbf{E}$  is the expectation with respect to two independent Brownian motions  $B^1$  and  $B^2$  start at 0 and with variance 1; and  $\lambda_{\text{ext}}$  is given by Eq. (5). This is indeed the second moment of the multiplicative stochastic heat equation defined in Eq. (7) with noise strength given by Eq. (4).

A similar argument applies for the case of general moments. *A priori*, when dealing with higher moments, there are terms that come from multiparticle interactions, e.g., when  $R^1 = R^2 = R^3$ . Each of these contributions needs to be accounted for when writing down the tilted measure and contribute differently to the local time factor. However, in our scaling, these different factors end up factorizing as  $N \rightarrow \infty$  and thus only contribute in the form of two-body local times, as needed to recover the general moment formula Eq. (7).

### A. Convergence to local time

The purpose of this section is to demonstrate the claimed convergence of Eq. (22). Recall  $V(t) = R^2(t) - R^1(t)$ , so  $\sum_{i=0}^{NT-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}} = \sum_{i=0}^{NT-1} \mathbb{1}_{\{|V(i)|=0\}}$  is the sum of occupation times at 0 for the random walk  $V$ . To derive Eq. (22), we identify a discrete analog of Tanaka's formula to identify a smoothed, discrete local time that converges to the local time of Brownian motion. We then use the invariant measure of the two-point gap process,  $\Delta(t) = |V(t)|$ , to relate the smoothed, discrete local time to the discrete local time at 0.

Before giving Tanaka's formula, we recall the definition of a martingale. A continuous time martingale is a time-parameterized collection of random variables,  $(Y_t)_{t \geq 0}$ , which obeys the martingale property whereby  $\mathbb{E}[Y_t \mid \mathcal{F}_s] = Y_s$  for all  $t > s \geq 0$ , where  $\mathcal{F}_s$  is the  $\sigma$ -algebra generated by  $Y_r$  for all  $s \geq r \geq 0$ . In words, this property says that the expected value at time  $t$ , given the history of the process up to time  $s$ , is the value at time  $s$ . Brownian motion  $B_t$  with drift zero is an example of a martingale, as is  $B_t^2 - t$  (the quadratic martingale) or  $e^{\lambda B_t - \lambda^2 t/2}$  (the exponential martingale) for any  $\lambda$ .

Tanaka's formula gives a decomposition of the absolute value of a Brownian motion into the sum of a martingale and Brownian local time. More precisely, given a Brownian motion  $B(t)$ , the Tanaka formula states that

$$|B(t)| = \int_0^t \text{sgn}(B_s) dB_s + L^B(t), \quad (23)$$

where  $L^B(t)$  is the local time at zero, defined in Eq. (8), and

$$\text{sgn}(x) = \begin{cases} +1 & x > 0 \\ 0 & x = 0. \\ -1 & x < 0 \end{cases}$$

The stochastic integral  $\int_0^t \text{sgn}(B_s) dB_s$  is a martingale, and, in fact, is another Brownian motion.

Tanaka's formula gives a way to decompose  $|B(t)|$  into the sum of a martingale (the stochastic integral) and an increasing predictable process (the local time). Such a decomposition is called the *Doob-Meyer decomposition*. A key fact is that the Doob-Meyer decomposition of a *submartingale* [a process such as  $Y_t = |B(t)|$  that satisfies  $\mathbb{E}[Y_t | \mathcal{F}_s] \geq Y_s$  for all  $t > s \geq 0$ ] is unique under fairly general assumptions. Therefore, if we can obtain *any* decomposition of  $|B(t)|$  into the sum of a martingale term and an increasing predictable process, then we can identify that increasing predictable process as Brownian local time.

Motivated by this fact, we obtain a Doob decomposition for  $\Delta(t) = |V(t)|$  by identifying a discrete analog of Tanaka's formula. This decomposition involves the sum of a martingale and a discrete local time term. We then take the term-by-term limit of this decomposition. We know that  $\Delta(t)$  will converge to  $|B(t)|$ , where  $B(t)$  will be a Brownian motion with variance 2. The martingale term will converge to a limiting martingale, and the limit of the discrete local time term will still be an increasing predictable process. Hence, by the uniqueness of the Doob-Meyer decomposition for  $|B(t)|$ , we will conclude that the limit of the discrete local time term is Brownian local time.

There are several mathematical subtleties that we brush over but quickly note here. From Eq. (21), we are concerned with convergence of an expectation with respect to the tilted measure  $\tilde{\mathbf{E}}_v$  of an expression involving the expectation of the local time at zero for  $\Delta(t)$ . The argument presented below shows that this local time under the untilted measure  $\mathbf{E}_v$  converges in distribution to a constant multiple of Brownian local time. The replacement of the tilted measure by its untilted counterpart is mostly out of convenience (the argument could be done under the tilted measure, too) and is justified since as  $N \rightarrow \infty$ , the tilted ( $N$ -dependent) measure converges to the untilted one. This can be seen by expanding  $e^x$  around  $x = 0$  to simplify the exponential in the tilted measure. Then the sums in the denominators of (18) and (19) evaluate to 1, and we recover the transition probabilities of the untilted two-point motion. The issue around convergence of expectations of exponentials would require some further careful justification that we do not pursue here.

Turning to the details of this argument, we decompose  $\Delta(t)$  as follows:

$$\Delta(t) = M(t) + \sum_{i=0}^{t-1} \sum_{l=0}^{\infty} \kappa(l) \mathbb{1}_{\{\Delta(i)=l\}}, \quad (24)$$

where we define

$$\kappa(l) = \mathbf{E}_v[\Delta(i+1) - \Delta(i) | \Delta(i) = l] \quad \text{and}$$

$$M(t) := \Delta(t) - \sum_{i=0}^{t-1} \sum_{l=0}^{\infty} \kappa(l) \mathbb{1}_{\{\Delta(i)=l\}}.$$

Note that  $\Delta(i+1) - \Delta(i)$  does not actually depend on  $i$  since we are conditioning on  $\Delta(i) = l$ .

To see that  $M(t)$  is a martingale with respect to the untilted measure  $\mathbf{E}_v$ , we rewrite  $M(t)$  as

$$\begin{aligned} M(t) &= \Delta(t) - \sum_{i=0}^{t-1} \sum_{l=0}^{\infty} \mathbf{E}_v[\Delta(i+1) - \Delta(i) | \Delta(i) = l] \mathbb{1}_{\{\Delta(i)=l\}} \\ &= \Delta(t) - \sum_{i=0}^{t-1} \mathbf{E}_v[\Delta(i+1) - \Delta(i) | \Delta(i)], \end{aligned}$$

where we have taken the sum over  $l$  so that we now condition on the random value  $\Delta(i)$ . Notice  $\Delta(i+1) - \Delta(i)$  is independent of the values of  $\Delta(1), \dots, \Delta(i-1)$  and only depends on  $\Delta(i)$ . Thus, we can rewrite

$$M(t) = \Delta(t) - \sum_{i=0}^{t-1} \mathbf{E}_v[\Delta(i+1) - \Delta(i) | \Delta(1), \dots, \Delta(i)].$$

It follows from this and telescoping that  $M(t)$  is a martingale.

We now show that the term  $\sum_{i=0}^{t-1} \sum_{l=0}^{\infty} \kappa(l) \mathbb{1}_{\{\Delta(i)=l\}}$  is increasing in  $t$  and predictable. We first show that  $\kappa(l) \geq 0$ . To see this, we use our results from Ref. [25] which show  $\kappa(l)$  simplifies to

$$\kappa(l) = \begin{cases} \sum_{i,j \in \mathbb{Z}} |i-j| \mathbb{E}_v[\xi(i)\xi(j)] & l = 0 \\ \sum_{|i-j| > l} (|i-j| - l) \bar{\xi}(i) \bar{\xi}(j) & l > 0 \end{cases} \quad (25)$$

Thus,  $\kappa(l) \geq 0$  and hence  $\sum_{i=0}^{t-1} \sum_{l=0}^{\infty} \kappa(l) \mathbb{1}_{\{\Delta(i)=l\}}$  is a sum of non-negative terms which increases in time. Furthermore, this process is predictable as it only depends on the values of  $\Delta(0) \dots \Delta(t-1)$ . Therefore, (24) is a decomposition of  $\Delta(t)$  into the sum of a martingale and an increasing predictable process.

We look at the limiting behavior of this decomposition under a diffusive scaling and match it term by term with Tanaka's formula in Eq. (23). As argued in the paragraph after Eq. (21), the term  $\Delta(t) = |R^1(t) - R^2(t)|$  converges under diffusive scaling to  $|B^1(t) - B^2(t)|$ . Since the term  $M(t)$  is a martingale, it should likewise converge to a limiting martingale. Therefore, the second term on the right-hand side of Eq. (24) should converge to the local time  $L^{B^1-B^2}(t)$  in Eq. (23) (where we take  $B = B^1 - B^2$ ). Thus, taking  $t = NT$ , under the diffusive scaling

$$\lim_{N \rightarrow \infty} \frac{1}{\sqrt{2DN}} \sum_{l=0}^{\infty} \kappa(l) \sum_{i=0}^{NT-1} \mathbb{1}_{\{\Delta(i)=l\}} = L^{B^1-B^2}(T). \quad (26)$$

This completes the first step towards establishing Eq. (22). Of course, that result calls for taking the limit of the discrete local time at 0,

$$\lim_{N \rightarrow \infty} \frac{1}{\sqrt{2DN}} \sum_{i=0}^{NT-1} \mathbb{1}_{\{\Delta(i)=0\}}, \quad (27)$$

whereas Eq. (26) is in terms of a combination of local time at every position  $l \in \mathbb{Z}_{\geq 0}$ .

To compare the limits in Eqs. (26) and (27), we will show in Sec. VIA 1 that for  $l \geq 0$ ,

$$\lim_{t \rightarrow \infty} \frac{\sum_{i=0}^t \mathbb{1}_{\{\Delta(i)=l\}}}{\sum_{i=0}^t \mathbb{1}_{\{\Delta(i)=0\}}} = \tilde{\mu}(l), \quad (28)$$



where  $\tilde{\mu}$  is the normalized invariant measure of  $V(t)$ , see the discussion after Eq. (5). For large  $t$ , we can therefore approximately relate the local time at  $l$  and at 0 so that

$$\sum_{i=0}^t \mathbb{1}_{\{\Delta(i)=l\}} \approx \tilde{\mu}(l) \sum_{i=0}^t \mathbb{1}_{\{\Delta(i)=0\}}. \quad (29)$$

Using this approximation, we obtain

$$\sum_{l=0}^{\infty} \kappa(l) \sum_{i=0}^{t-1} \mathbb{1}_{\{\Delta(i)=l\}} \approx \left( \sum_{l=0}^{\infty} \kappa(l) \tilde{\mu}(l) \right) \sum_{i=0}^{t-1} \mathbb{1}_{\{\Delta(i)=0\}}.$$

Substituting this into Eq. (26) and rearranging, we obtain

$$\lim_{N \rightarrow \infty} \frac{1}{\sqrt{2DN}} \sum_{i=0}^{Nt-1} \mathbb{1}_{\{\Delta(i)=0\}} = \frac{1}{\sum_{l=0}^{\infty} \kappa(l) \tilde{\mu}(l)} L^{B^1-B^2}(t). \quad (30)$$

After substituting in our definitions of  $\kappa(l)$  and  $\Delta(t)$ , we conclude that

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{\sqrt{2DN}} \sum_{i=0}^{Nt-1} \mathbb{1}_{\{R^1(i)=R^2(i)\}} \\ = \frac{1}{\sum_{l=0}^{\infty} \tilde{\mu}(l) \mathbf{E}_v[\Delta(t+1) - \Delta(t) \mid \Delta(t)=l]} L^{B^1-B^2}(t), \end{aligned}$$

which indeed matches Eq. (22).

### 1. Invariant measure of $V$

In this section, we demonstrate the claim in Eq. (28). If a Markov chain is irreducible (starting from any state, the chain can eventually reach any other state) and recurrent (it returns to every state infinitely many times), then it has a unique invariant measure up to a constant multiple. As explained earlier, we consider  $V$  under the untilted measure  $\mathbf{E}_v$  on the two-point motion rather than the tilted probability distribution. We can see that  $V(t)$  is recurrent as follows: When  $V(t)$  is away from 0 its increments are independent and identically distributed and have mean 0 and thus, by the Chung-Fuchs theorem [41], the walk will almost surely eventually return to 0. Once at 0, it will eventually leave, and the above argument can be repeated to show that  $V(t)$  will return to 0 infinitely many times, which suffices to show recurrence. However,  $V(t)$  may not always be irreducible. For example, consider the case where  $R^1(t)$  and  $R^2(t)$  only take nearest-neighbor jumps. Then  $V(t)$  can only take jumps that are multiples of 2, so if it starts from 0, it can only visit sites in the sublattice  $2\mathbb{Z}$ .

Let us first deal with the case where  $V(t)$  is irreducible so that it has a unique (up to a constant multiple) invariant measure. Let

$$T_0 = 0, \quad T_k = \inf\{n > T_{k-1} : V(n) = 0\}$$

so that  $T_k$  is the  $k$ th return time to 0. Define  $Y_k^l = \sum_{i=T_k}^{T_{k+1}-1} \mathbb{1}_{\{V(i)=l\}}$ . For an irreducible and recurrent Markov chain, we know that the measure  $\mu$  on  $\mathbb{Z}$  defined by  $\mu(l) := \mathbb{E}[Y_0^l]$  is an invariant measure for  $V$  where  $\mathbb{E}$  is the expectation of the walker starting at 0. Furthermore, since  $V$  is a Markov chain, the random variables  $(Y_k^l)_{k \geq 0}$  are independent and identically distributed for a fixed  $l$ .

We have that

$$\frac{Y_0^l + \cdots + Y_{n-1}^l}{n} = \frac{\sum_{i=0}^{T_n-1} \mathbb{1}_{\{V(i)=l\}}}{\sum_{i=0}^{T_n-1} \mathbb{1}_{\{V(i)=0\}}}. \quad (31)$$

The equality in the denominator is due to the fact that we visit 0 exactly  $n$  times before time  $T_n - 1$ . By the law of large numbers for independent and identically distributed random variables,

$$\lim_{n \rightarrow \infty} \frac{Y_0^l + \cdots + Y_{n-1}^l}{n} = \mathbb{E}[Y_0^l] = \mu(l), \quad (32)$$

almost surely. On the other hand, we have that

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=0}^{T_n-1} \mathbb{1}_{\{V(i)=l\}}}{\sum_{i=0}^{T_n-1} \mathbb{1}_{\{V(i)=0\}}} = \lim_{t \rightarrow \infty} \frac{\sum_{i=0}^t \mathbb{1}_{\{V(i)=l\}}}{\sum_{i=0}^t \mathbb{1}_{\{V(i)=0\}}}. \quad (33)$$

Thus, by combining Eqs. (31)–(33), we obtain

$$\lim_{t \rightarrow \infty} \frac{\sum_{i=0}^t \mathbb{1}_{\{V(i)=l\}}}{\sum_{i=0}^t \mathbb{1}_{\{V(i)=0\}}} = \mu(l). \quad (34)$$

Note that by symmetry  $\mu(l) = \mu(-l)$ . Furthermore, for  $l = 0$ , we have

$$\mathbb{1}_{\{\Delta(i)=l\}} = \mathbb{1}_{\{V(i)=l\}}$$

and for  $l > 0$ ,

$$\mathbb{1}_{\{\Delta(i)=l\}} = \mathbb{1}_{\{V(i)=l\}} + \mathbb{1}_{\{V(i)=-l\}}.$$

Putting this all together with Eq. (34), we conclude that

$$\lim_{t \rightarrow \infty} \frac{\sum_{i=0}^t \mathbb{1}_{\{\Delta(i)=l\}}}{\sum_{i=0}^t \mathbb{1}_{\{\Delta(i)=0\}}} = \tilde{\mu}(l),$$

where

$$\tilde{\mu}(l) = \begin{cases} 1 & \text{if } l = 0 \\ 2\mu(l) & \text{if } l > 0. \end{cases}$$

Finally, we consider the case where  $V(t)$  is not irreducible. We can decompose  $\mathbb{Z}$  into the union of finitely many closed and irreducible sets (sublattices)  $E_i$ . Let  $E_0$  be the set containing 0. Since we have  $V(0) = 0$ , we know that  $V(t) \in E_0$  for all  $t$ . Let  $\mu$  be the unique invariant measure of the Markov chain  $V(t)$  when restricted to the state space  $E_0$ . All of the above analysis goes through, but the sum  $\sum_{l=0}^{\infty} \kappa(l) \tilde{\mu}(l)$  gets replaced by  $\sum_{l \in E_0} \kappa(l) \tilde{\mu}(l)$ . An example of this situation is when  $R^1$  and  $R^2$  are nearest-neighbor random walks. In this case,  $V(t)$  is restricted to the even integer sublattice, i.e.,  $E_0 = 2\mathbb{Z}$ .

## VII. CONVERGENCE OF THE TAIL PROBABILITY TO THE SHE

The above discussion shows that the first two moments of the rescaled probability mass distribution converge to the moments of the SHE. In this section, we discuss the convergence of the tail probability, which can be written as

$$\begin{aligned} \mathbb{P}^\xi(R(NT) \geq c(N)T + \sqrt{2DN}X) \\ = \mathbb{E}^\xi \left[ \mathbb{1}_{\{R(NT) \geq c(N)T + \sqrt{2DN}X\}} \right]. \end{aligned} \quad (35)$$

Defining (and then simplifying)

$$\begin{aligned}\phi_N^{\text{tail}}(X') &:= \frac{C(N, T, X)}{C(N, T, X')} \mathbb{1}_{\{X' \geq X\}} \\ &= \exp \left\{ -\frac{N^{1/4}}{\sqrt{2D}}(X' - X) \right\} \mathbb{1}_{\{X' \geq X\}},\end{aligned}\quad (36)$$

we can rewrite Eq. (35) as

$$\begin{aligned}\mathbb{P}^\xi(R(NT) \geq c(N)T + \sqrt{2DN}X) \\ &= \frac{1}{C(N, T, X)} \mathbb{E}^\xi \left[ C \left( N, T, \frac{R^1(NT) - c(N)T}{\sqrt{2DN}} \right) \right. \\ &\quad \left. \phi_N^{\text{tail}} \left( \frac{R^1(NT) - c(N)T}{\sqrt{2DN}} \right) \right] \\ &= \frac{1}{C(N, T, X)} \mathcal{U}_N(T, \phi),\end{aligned}\quad (37)$$

where we recall  $\mathcal{U}_N(T, \phi)$  from Eq. (6). Observe that as  $N \rightarrow \infty$ ,

$$\exp \left\{ -\frac{N^{1/4}}{\sqrt{2D}}(X' - X) \right\} \mathbb{1}_{\{X' \geq X\}} \approx \delta \left( \frac{N^{1/4}}{\sqrt{2D}}(X' - X) \right) \mathbb{1}_{\{X' \geq X\}}.$$

Note here that the convergence shown earlier was for fixed  $\phi$  while we are now permitting  $\phi$  to vary in  $N$  and tend to a Dirac delta function. Thus, in light of the convergence of  $\mathcal{U}_N(T, \phi)$  to the SHE, this implies that

$$\begin{aligned}\mathbb{P}^\xi(R(NT) \geq c(N)T + \sqrt{2DN}X) \\ \approx \frac{1}{C(N, T, X)} \int_X^\infty \delta \left( -\frac{N^{1/4}}{\sqrt{2D}}(X' - X) \right) \\ \times Z(T', X') dX'\end{aligned}\quad (38)$$

$$= \frac{\sqrt{2D}}{N^{1/4}C(N, T, X)} Z(T, X).\quad (39)$$

Moving the prefactor to the left-hand side,

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{N^{1/4}C(N, T, X)}{\sqrt{2D}} \mathbb{P}^\xi(R(NT) \geq c(N)T + \sqrt{2DN}X) \\ = Z(T, X).\end{aligned}\quad (40)$$

Thus, the tail probability converges to the SHE with the same noise strength but a modified prefactor.

### VIII. THE DISTRIBUTION OF THE EXTREME LOCATION

In a system of many diffusing particles, the location of the particle that has moved the furthest is often more important than that of a typical particle, as in oocyte fertilization [26,33,34], neuronal activation [32], and the information flow in networks [35,36]. Thus, understanding the distribution of the extreme particle is of critical importance.

We summarize our main result in Ref. [25], where we used the above results for the tail probability to characterize the statistics of the *extreme location*, which is the particle that has traveled the furthest, as a function of time. In that work, we argued that a measurement of the extreme location can be used to reveal microscopic information about the underlying environment through the coefficient  $\lambda_{\text{ext}}$ . To do this, we studied

the extreme location of  $N$  random walks at time  $t$ ,

$$\text{Max}_t^N := \max(\{R^1(t), \dots, R^N(t)\}),$$

recalling that  $R^1(t), \dots, R^N(t)$  for  $t \in \mathbb{Z}_{\geq 0}$  are random walks distributed according to the measure  $\mathbb{P}^\xi$ . Using our results for the distribution of the tail probability in Eq. (40), we found that as  $t \rightarrow \infty$ ,

$$\begin{aligned}\text{Max}_t^N \approx \sqrt{4Dt \ln(N)} + \text{Gumbel} \left( 0, \sqrt{\frac{Dt}{\ln(N)}} \right) \\ + \sqrt{\frac{Dt}{\ln(N)}} h \left( 0, \frac{8\lambda_{\text{ext}}^2 \ln(N)^2}{Dt} \right),\end{aligned}\quad (41)$$

where  $\text{Gumbel}(\mu, \beta)$  is a Gumbel random variable with location  $\mu$  and shape  $\beta$ ,  $h(X, T)$  solves the KPZ equation with noise strength  $D_0 = 1/2$  and these two random variables are independent.

The first and second terms in Eq. (41) agree with results from classical diffusion on the distribution of the extreme location [28,29]. The derivation of the third term in Eq. (41) is the main result of Ref. [25]. Since only the second and third terms are random variables, the variance of  $\text{Max}_t^N$  has two contributions. We find that as  $t \rightarrow \infty$

$$\text{Var}(\text{Max}_t^N) \approx \frac{\pi^2 Dt}{6 \ln(N)} + \lambda_{\text{ext}} \sqrt{2\pi Dt},$$

where  $\text{Var}(\cdot)$  is the variance under the annealed measure  $\mathbf{P}_v$ , the first term is the classical diffusion result which comes from taking the variance of the Gumbel random variable in Eq. (41), and the second term comes from taking the variance of the KPZ term and then using the small time expansion of  $h(0, T)$  [42,43].

By measuring the variance of the extreme location and subtracting off the contribution due to classical diffusion, the KPZ term, and subsequently  $\lambda_{\text{ext}}$ , can be directly measured. Thus, a macroscopic measurement of the fluctuations of the extreme location over time yields microscopic information about the underlying environment through the coefficient  $\lambda_{\text{ext}}$ .

### IX. CONCLUSION

We have demonstrated that under a moderate deviations scaling, there are universal KPZ fluctuations for a large class of RWRE models. This class of model only includes environments which are delta correlated in space and time making the environment memoryless. However, as our model is a discretized or coarse-grained model, it can be equally well applied to environments with memories that decay faster than a power law by choosing the time steps in the model to be larger than the correlation time of the memory. It should be noted that this excludes a large number of interesting systems with power-law memories [39] which would be expected to deviate even farther from the Einstein model in their extremes.

We show that the strength of the noise of the KPZ equation is characterized by the variable  $\lambda_{\text{ext}}$ , which depends on the statistics of the underlying environment. Our results can be extended to characterize the distribution of the extreme

value statistics of a system of  $N$  diffusing particles as in Refs. [20,21,25] (e.g., the first time a particle reaches a position  $L$  or the position of the farthest particle at time  $t$ ). Thus, we have established a link between the extreme value statistics of a system of diffusing particles and the microscopic rheological properties of the underlying environment.

### ACKNOWLEDGMENTS

We thank Guillaume Barraquand, Pierre Le Doussal, and Shalin Parekh for helpful discussions. This work was funded by the W. M. Keck Foundation Science and Engineering grant on “Extreme Diffusion.” I.C. also wishes to acknowledge ongoing support from NSF DMS:1811143, 1937254,

and 2246576 and the Simons Foundation Investigator Grant No. 929852. H.D. wishes to acknowledge support from the NSF GRFP DGE:2036197. This material is partly based upon work supported by the National Science Foundation under Grant No. DMS-2424139, while H.D. was in residence at the Simons Laufer Mathematical Sciences Institute in Berkeley, California. This work benefited from access to the University of Oregon high performance computing cluster, Talapas.

### DATA AVAILABILITY

No data were created or analyzed in this study.

- 
- [1] H. Liu, C.-Y. Liao, J.-Y. Ko, and J.-S. Lih, Anchoring effect on first passage process in taiwan financial market, *Physica A* **477**, 114 (2017).
  - [2] F. Black and M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.* **81**, 637 (1973).
  - [3] A. Ishimaru, Diffusion of light in turbid material, *Appl. Opt.* **28**, 2210 (1989).
  - [4] P. H. Kao and R. J. Yang, Virus diffusion in isolation rooms, *J. Hosp. Infect.* **62**, 338 (2006).
  - [5] A. Einstein, Zur theorie der brownischen bewegung, *Ann. Phys. (NY)* **324**, 371 (1906).
  - [6] A. Einstein, Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen, *Ann. Phys. (NY)* **322**, 549 (1905).
  - [7] A. Einstein, Theoretische bemerkungen über die brownische bewegung, *Z. Elektrochem. Angew. Phys. Chem.* **13**, 41 (1907).
  - [8] P. Langevin, Sur la théorie du mouvement brownien, *Compt. Rendus* **146**, 530 (1908).
  - [9] W. Sutherland, The viscosity of gases and molecular force, *Philos. Mag. Ser. V* **36**, 507 (1893).
  - [10] W. Sutherland, LXXV, A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin, *Lond. Edinb. Dubl. Philos. Mag. J. Sci.* **9**, 781 (1905).
  - [11] M. von Smoluchowski, Zur kinetischen theorie der brownischen molekularbewegung und der suspensionen, *Ann. Phys. (NY)* **326**, 756 (1906).
  - [12] M. Von Smoluchowski, Notiz uiber die Berechnung der Brownischen Molekularbewegung bei der Ehrenhaft-Millikanschen Versuchsanordnung, *Phys. Z.* **16**, 318 (1915).
  - [13] J. Perrin, Mouvement brownien et molécules, *J. Chim. Phys.* **8**, 57 (1910).
  - [14] Jean B. Perrin, Le Mouvement Brownien et la Realite Moleculaire, *Ann. Chim. Phys.* **18**, 5 (1909).
  - [15] R. Brown, XXVII, A brief account of microscopical observations made in the months of June, July and 1827, on the particles contained in the pollen of plants; And on the general existence of active molecules in organic and inorganic bodies, *Philos. Mag.* **4**, 161 (1828).
  - [16] G. Barraquand and I. Corwin, Random-Walk in beta-distributed random environment, *Probab. Theory Relat. Fields* **167**, 1057 (2017).
  - [17] T. Thiery and P. Le Doussal, Exact solution for a random walk in a time-dependent 1D random environment: The point-to-point beta polymer, *J. Phys. A: Math. Theor.* **50**, 045001 (2017).
  - [18] P. Le Doussal and T. Thiery, Diffusion in time-dependent random media and the Kardar-Parisi-Zhang equation, *Phys. Rev. E* **96**, 010102(R) (2017).
  - [19] G. Barraquand and P. Le Doussal, Moderate deviations for diffusion in time dependent random media, *J. Phys. A: Math. Theor.* **53**, 215002 (2020).
  - [20] J. B. Hass, A. N. Carroll-Godfrey, I. Corwin, and E. I. Corwin, Anomalous fluctuations of extremes in many-particle diffusion, *Phys. Rev. E* **107**, L022101 (2023).
  - [21] J. B. Hass, I. Corwin, and E. I. Corwin, First-passage time for many-particle diffusion in space-time random environments, *Phys. Rev. E* **109**, 054101 (2024).
  - [22] S. Das, H. Drillick, and S. Parekh, Multiplicative SHE limit of random walks in space-time random environments, *Probab. Theory Relat. Fields* (2024).
  - [23] S. Parekh, Hierarchy of KPZ limits arising from directed random walk models in random media, *arXiv:2401.06073*.
  - [24] J. Hass, Super-universal behavior of outliers diffusing in a space-time random environment, *arXiv:2505.01533*.
  - [25] J. B. Hass, H. Drillick, I. Corwin, and E. I. Corwin, Extreme diffusion measures statistical fluctuations of the environment, *Phys. Rev. Lett.* **133**, 267102 (2024).
  - [26] Z. Schuss, K. Basnayake, and D. Holcman, Redundancy principle and the role of extreme statistics in molecular and cellular biology, *Phys. Life Rev.* **28**, 52 (2019).
  - [27] S. Linn and S. D. Lawley, Extreme hitting probabilities for diffusion, *J. Phys. A: Math. Theor.* **55**, 345002 (2022).
  - [28] S. D. Lawley, Distribution of extreme first passage times of diffusion, *J. Math. Biol.* **80**, 2301 (2020).
  - [29] S. D. Lawley, Universal formula for extreme first passage statistics of diffusion, *Phys. Rev. E* **101**, 012413 (2020).
  - [30] J. B. Madrid and S. D. Lawley, Competition between slow and fast regimes for extreme first passage times of diffusion, *J. Phys. A: Math. Theor.* **53**, 335002 (2020).
  - [31] K. Basnayake, Z. Schuss, and D. Holcman, Asymptotic formulas for extreme statistics of escape times in 1, 2 and 3-dimensions, *J. Nonlin. Sci.* **29**, 461 (2019).
  - [32] K. Basnayake, D. Mazaud, A. Bemelmans, N. Rouach, E. Korkotian, and D. Holcman, Fast calcium transients in dendritic spines driven by extreme statistics, *PLOS Biol.* **17**, e2006202 (2019).

- [33] J. Yang, I. Kupka, Z. Schuss, and D. Holcman, Search for a small egg by spermatozoa in restricted geometries, *J. Math. Biol.* **73**, 423 (2016).
- [34] B. Meerson and S. Redner, Mortality, redundancy, and diversity in stochastic search, *Phys. Rev. Lett.* **114**, 198101 (2015).
- [35] R. Germano and A. P. S. de Moura, Traffic of particles in complex networks, *Phys. Rev. E* **74**, 036117 (2006).
- [36] S.-P. Wang and W.-J. Pei, First passage time of multiple Brownian particles on networks with applications, *Physica A* **387**, 4699 (2008).
- [37] S. Das, H. Drillick, and S. Parekh, KPZ equation limit of sticky Brownian motion, *J. Funct. Anal.* **287**, 110609 (2024).
- [38] P. Le Doussal, Dynamics at the edge for independent diffusing particles, *Phys. Rev. E* **109**, 024101 (2024).
- [39] T. G. Mason and D. A. Weitz, Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids, *Phys. Rev. Lett.* **74**, 1250 (1995).
- [40] T. G. Mason, H. Gang, and D. A. Weitz, Diffusing-wave-spectroscopy measurements of viscoelasticity of complex fluids, *J. Opt. Soc. Am. A* **14**, 139 (1997).
- [41] K. L. Chung and W. H. J. Fuchs, On the distribution of values of sums of random variables, *Mem. Am. Math. Soc.* **6**, 12 (1951).
- [42] G. Amir, I. Corwin, and J. Quastel, Probability distribution of the free energy of the continuum directed random polymer in  $1 + 1$  dimensions, *Commun. Pure Appl. Math.* **64**, 466 (2011).
- [43] P. Calabrese, P. Le Doussal, and A. Rosso, Free-energy distribution of the directed polymer at high temperature, *Europhys. Lett.* **90**, 20002 (2010).