Probing the large deviations for the Beta random walk in random medium

Alexander K. Hartmann,[1] Alexandre Krajenbrink,[2,3] and Pierre Le Doussal[4]

[1] *Institut für Physik, Universität Oldenburg, 26111 Oldenburg, Germany*
[2] *Quantinuum, Terrington House, 13–15 Hills Road, Cambridge CB2 1NL, United Kingdom*
[3] *Le Lab Quantique, 58 rue d'Hauteville, 75010, Paris, France*
[4] *Laboratoire de Physique de l'École Normale Supérieure, PSL University, CNRS, Sorbonne Universités, 24 rue Lhomond, 75231 Paris, France*

(Dated: October 11, 2023)

We consider a discrete-time random walk on a one-dimensional lattice with space and time-dependent random jump probabilities, known as the Beta random walk. We are interested in the probability that, for a given realization of the jump probabilities (a sample), a walker starting at the origin at time $t = 0$ is at position beyond $\xi\sqrt{T/2}$ at time T. This probability fluctuates from sample to sample and we study the large-deviation rate function which characterizes the tails of its distribution at large time $T \gg 1$. It is argued that, up to a simple rescaling, this rate function is identical to the one recently obtained exactly by two of the authors for the continuum version of the model. That continuum model also appears in the macroscopic fluctuation theory of a class of lattice gases, e.g. in the so-called KMP model of heat transfer. An extensive numerical simulation of the Beta random walk, based on an importance sampling algorithm, is found in good agreement with the detailed analytical predictions. A first-order transition in the tilted measure, predicted to occur in the continuum model, is also observed in the numerics.

PACS numbers: 05.40.-a, 02.10.Yn, 02.50.-r

I. INTRODUCTION

The macroscopic fluctuation theory (MFT) [5] provides a coarse grained continuum description of the fluctuations of the density and current [6] for a broad class of discrete stochastic systems in one dimension with a diffusive scaling at large time. One important example is the symmetric exclusion process, where particles perform symmetric jumps onto neighboring unoccupied sites on a lattice. Another example is the Kipnis-Marchioro-Presutti (KMP) model [27], a lattice model where each site has an energy and whose the dynamics is described by a random exchange of energy between neighbors. Upon introduction of an asymmetry or a driving, such as in the asymmetric exclusion process [7], the diffusive scaling breaks down above some scale, and the large scale behavior of the model is usually described by the Kardar-Parisi-Zhang (KPZ) universality class [8]. It was shown that there is a natural crossover from the MFT to the so-called weak noise theory (WNT) of the KPZ equation as the asymmetry is increased [1].

The MFT and the WNT allow to reduce the calculation of the large deviations of density and current to solving a system of two coupled non-linear differential dynamical equations, with prescribed boundary conditions both at initial and final time. Recently, starting with the WNT for the KPZ equation [14, 15, 56], exact solutions to these systems were obtained [1–4, 39, 44]. This was achieved by using the close connection of these systems to the non-linear Schrodinger equation (NLS), or to the derivative NLS equation (DNLS), and extending the inverse scattering methods of [16–18] to mixed-time boundary conditions. Another largely equivalent method used exact closure schemes [3, 39]. This allows one to compute large deviations for observables such as the integrated current or of the position of a tracer.

Here we will focus on the case where the MFT takes the form of a linear stochastic equation for a space-time coarse-grained density field $q_\eta(y,\tau)$

$$\partial_\tau q_\eta(y,\tau) = \partial_y^2 q_\eta(y,\tau) - \partial_y(\sqrt{2}\eta(y,\tau)q_\eta(y,\tau))\,, \quad (1)$$

where $\eta(y,\tau)$ is a standard space-time Gaussian white noise. It was proved in [29] that at large time the large deviations for the discrete KMP model are identical to those of the continuum stochastic model (1). At large time the dynamical action associated to model (1) is controled by a saddle point, and the corresponding saddle point equations define the MFT for this model. These MFT equations were studied in a number of works [5, 28, 30–40]. We noted in [1] an interesting connection to a continuum model of diffusion in a time dependent random environment, previously considered in [20, 23, 64–66, 72]. Indeed, Eq. (1) can also be seen as the Fokker-Planck equation for the probability distribution function (PDF) $q_\eta(y,\tau)$ of the position $y(\tau)$ at time τ of a particle convected by the random field $\eta(y,\tau)$, described by the Langevin equation

$$\frac{\mathrm{d}y(\tau)}{\mathrm{d}\tau} = \sqrt{2}\eta(y(\tau),\tau) + \chi(\tau)\,, \quad (2)$$

where χ is a standard white noise in time. The subscript in q_η emphasizes that it depends on the realization of the random field η, i.e. the sample. In [1] (see also [2, 44]) we solved the MFT equations and derived the large-time large-deviation function associated to q_η for the continuum model (1), with applications to diffusion of extremes in time-dependent continuum random media. This solution was obtained by inverse scattering methods on a non-linear system interpolating between

the DNLS and NLS equations. We also obtained the same result by performing the large-time expansion of an exact Fredholm determinant formula obtained in [23] using the Bethe ansatz.

It is thus natural to investigate whether the MFT associated to (1) can describe a discrete model of a random walk in a random environmnent (RWRE). The natural example which we will consider here is the so-called Beta random walk, introduced and studied in [19]. This model was later studied in relation to the KPZ equation in Refs. [20–24, 65, 66]. In the present paper we will first argue that the large-time large-deviations tails for the Beta polymer are indeed identical, up to some simple rescaling that we can predict, to those of the continuum model. Next, we will perform an extensive numerical study of these large deviations for the Beta random walk, using an importance sampling algorithm, to test our analytical predictions.

The outline is as follows. In Section II we define the model of the Beta random walk, introduce the observables of interest and define the associated large-deviation rate functions. In Section III we sketch the argument which allows to relate the large deviation of the discrete model to those of the continuum model. In Section IV we explain the numerical method used here, notably the importance sampling method which allows to explore the deep tails of the large-deviation regime. In Section V we give the main numerical results and discuss how they compare to the analytical predictions.

II. MODEL AND OBSERVABLES

The model of the Beta random walk is defined as follows [19]. One defines first the "environment" or sample, by choosing for each $x \in \mathbb{Z}$ and $t \in \mathbb{N}$ a variable $w_{x,t} \in [0,1]$. The $w_{x,t}$ are chosen as i.i.d random variables taken from the beta distribution with parameters $\alpha, \beta > 0$ and density

$$\mathcal{P}(w) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} w^{\alpha-1}(1-w)^{\beta-1} \tag{3}$$

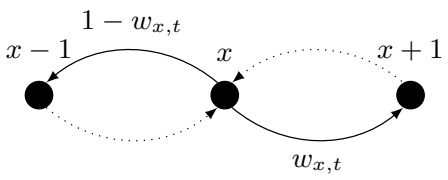

Figure 1. Lattice representation of the Beta random walk.

One now considers a particle at position $X(t) \in \mathbb{Z}$, which starts at the origin at time zero, $X(0) = 0$, and which performs a random walk defined by the following transition probabilities (see Fig. 1)

$$\begin{aligned} &\mathbb{P}(X(t+1) = x+1 | X(t) = x) = w_{x,t} \\ &\mathbb{P}(X(t+1) = x-1 | X(t) = x) = 1 - w_{x,t} \end{aligned} \tag{4}$$

We will denote by $\langle X(t) \rangle_w$, or more simply $\langle X(t) \rangle$ the mean position in a given sample, and by $\overline{\cdots}$ the averages over samples. It is easy to see that the sample averaged bias and diffusion coefficient, defined as $\overline{\langle X(t) \rangle} = \bar{v}t$ and $\overline{\langle X(t)^2 \rangle^c} = \bar{D}t$ are equal to

$$\bar{v} = 2\langle w \rangle - 1 = \frac{\alpha - \beta}{\alpha + \beta} \quad , \quad \bar{D} = \frac{4\alpha\beta}{(\alpha+\beta)^2} \tag{5}$$

One can show that at large time the typical walk in a typical sample is also characterized by the same bias and diffusion coefficient. We will choose from now on $\alpha = \beta$, i.e $\bar{v} = 0$ and $\bar{D} = 1$.

Since at large time the typical motion is diffusive ref ? we will be interested in the following probability, which will be our observable

$$Z = Z_\xi(T) = \mathbb{P}\left(\frac{X(T)}{\sqrt{T/2}} > \xi \right), \tag{6}$$

where ξ is an asymmetry parameter (here chosen positive $\xi \geq 0$) which describes how the position of the random walker deviates from its mean in a given sample. Keep in mind that Z is a random variable with respect to the sample. We will thus be interested in the PDF's of Z w.r.t. the sample, denoted $P(Z)$, equivalently of $H = \log Z$ denoted (abusively) as $P(H)$. At large time T they are expected to take the large-deviation forms

$$P(Z) \sim e^{-\sqrt{T}\hat{\Phi}_\xi^{\mathrm{RW}}(Z)} \tag{7}$$

$$P(H) \sim e^{-\sqrt{T}\Phi_\xi^{\mathrm{RW}}(H)} \tag{8}$$

and we will determine the rate functions $\hat{\Phi}_\xi^{\mathrm{RW}}(Z)$ and $\Phi_\xi^{\mathrm{RW}}(H)$, both through an analytical argument, and through extensive numerics.

III. ANALYTICAL PREDICTIONS

A. Continuum model observables

One can define a similar observable for the continuum model (1)-(2), namely

$$\tilde{Z} = \tilde{Z}_\xi(\tilde{T}) = \int_{\xi\sqrt{\tilde{T}}}^{+\infty} \mathrm{d}y \, q_\eta(y, \tilde{T}) = \mathbb{P}\left(\frac{y(\tilde{T})}{\sqrt{\tilde{T}}} > \xi \right), \tag{9}$$

where the particle is at the origin at time zero, $y(0) = 0$. Note that here and below the tilde variables are associated to the continuum model. We have shown in Ref. [1] that at large time the PDF's of $\tilde{Z}$ and $\tilde{H} = \log \tilde{Z}$ take the large-deviation forms

$$P(\tilde{Z}) \sim e^{-\sqrt{\tilde{T}}\hat{\Phi}_\xi(\tilde{Z})} \tag{10}$$

$$P(\tilde{H}) \sim e^{-\sqrt{\tilde{T}}\Phi_\xi(\tilde{H})} \tag{11}$$

and we have obtained the analytical expressions of the rate functions $\hat{\Phi}_\xi$ and Φ_ξ, which will be recalled below. These were obtained by considering the following generating function, which takes the large-deviation form at

large time

$$\overline{e^{-\tilde{z}\sqrt{\tilde{T}}\tilde{Z}}} \sim e^{-\sqrt{\tilde{T}}\Psi_\xi(\tilde{z})} \tag{12}$$

The two rate functions are related by a Legendre transform. Indeed one has

$$\Psi_\xi(\tilde{z}) = \min_{\tilde{Z}\in[0,1]} (z\tilde{Z} + \hat{\Phi}(\tilde{Z})) \tag{13}$$

In Ref. [1] we obtained the expression of $\Psi_\xi(\tilde{z})$ by two different methods, one of them will be recalled in the next section. From it we obtained the rate functions $\hat{\Phi}_\xi$ and Φ_ξ through Legendre inversion. The explicit formula for these rate functions will be given in Section III E.

B. Main prediction

Our main prediction is that the discrete model is described by the same rate functions as the continuum one, up to some scale factors, hinting to a form of universality. More precisely we claim that

$$\hat{\Phi}^{\rm RW}_\xi(Z) = \frac{\alpha}{\sqrt{2}}\hat{\Phi}_\xi(Z) \quad , \quad \Phi^{\rm RW}_\xi(H) = \frac{\alpha}{\sqrt{2}}\Phi_\xi(H) \tag{14}$$

and

$$\Psi^{\rm RW}_\xi(z) = \frac{\alpha}{\sqrt{2}}\Psi_\xi\left(\tilde{z} = \frac{z\sqrt{2}}{\alpha}\right) \tag{15}$$

We will now explain the origin of this prediction. To this aim we first need to recall one method to obtain the rate function Ψ_ξ in the continuum. Next we show how an extension of the same method for the discrete case leads to the predictions above.

C. Fredholm determinant method for the continuum model

In [23] a mathematically well posed version of the continuum model, called the sticky Brownian motion, was defined. An exact formula was derived for the Laplace transform of the PDF of $\tilde{Z}$ for any $\tilde{T}, \xi$ in terms of a (complicated) Fredholm determinant. For $u \geq 0$ one has [23, Thm 1.11]

$$\overline{e^{-u\tilde{Z}_\xi(\tilde{T})}} = \mathrm{Det}(I - K_u)|_{\mathbb{L}^2(C)} \tag{16}$$

where the kernel $K_u(v, v')$ acts on functions defined on a contour C in the v complex plane, where C is a positively-oriented circle centered at R with radius R. The kernel reads

$$K_u(v, v') = \frac{1}{2\mathbf{i}\pi}\int_{1/2+\mathbf{i}\mathbb{R}} \frac{\pi u^s}{\sin \pi s}\frac{g(v)}{g(v+s)}\frac{\mathrm{d}s}{s+v-v'} \tag{17}$$

where the function $g(v)$ is

$$g(v) = g_c(v) := e^{\xi\sqrt{\tilde{T}}\psi_0(v) + \tilde{T}\psi_1(v)}\Gamma(v) \tag{18}$$

where $\psi_{0,1}$ denote polygamma functions (see [1, Supp Mat Sec. X] for more details and correspondence of conventions).

In Ref. [1, Supp Mat Sec. X] we have studied in detail the large time limit $\tilde{T} \gg 1$ of the kernel K_u (17) and of the Fredholm determinant (16) when u is scaled as $u = \tilde{z}\sqrt{\tilde{T}}$ for the continuum model (recalling that here the observation time in the continuum model is denoted by $\tilde{T}$). From (12) and (16) this provided an independent method to obtain the rate function $\Psi_\xi(\tilde{z})$. The important point is that we showed there that the only relevant quantity is the asymptotic form at large $\tilde{T}$ of the function $g(v)$ under the rescaling

$$v = w\sqrt{\tilde{T}} \tag{19}$$

More specifically one finds that this asymptotic form reads [1, Supp Mat Eqs. (S206)-(S210)]

$$\begin{aligned} \log g_c(v) &= \sqrt{\tilde{T}}\left(\phi(w) + (w+\xi)\log\sqrt{\tilde{T}}\right) \\ &+\chi(w) - \frac{1}{2}\log(\sqrt{\tilde{T}}) + o(\tilde{T}) \end{aligned} \tag{20}$$

where we defined

$$\begin{aligned} \phi(w) &= \frac{1}{w} - w + (w+\xi)\log(w) \\ \chi(w) &= \frac{1}{2w^2} - \frac{\xi}{2w} + \frac{1}{2}\log(2\pi/w) \end{aligned} \tag{21}$$

where we corrected a misprint in the last term in [1]. The knowledge of $\phi(w)$ in this asymptotic form then allows to obtain the explicit form of $\Psi_\xi(\tilde{z})$ in Eq. (A2), see [1, Supp Mat Sec. X] for details. In particular, subdominant terms such as $\chi(w)$ are irrelevant.

D. Discrete to continuum universality

To obtain the rate functions for the discrete Beta random walk model, we can consider, similarly to (12), the generating function associated to the observable $Z = Z_\xi(T)$ which takes the form at large time

$$\overline{e^{-z\sqrt{T}Z}} \sim e^{-\sqrt{T}\Psi^{\rm RW}_\xi(z)} \tag{22}$$

One method to obtain $\Psi^{\rm RW}_\xi(z)$ is to use the result from [19] which we now recall. Reference [19, Theorem 1.13] gives an exact formula for the Laplace transform (22) of the PDF of Z for any T, ξ in terms of a (complicated) Fredholm determinant. For $u \geq 0$ one has

$$\overline{e^{-uZ_\xi(T)}} = \mathrm{Det}(I - K_u)|_{\mathbb{L}^2(C)} \tag{23}$$

where the kernel $K_u(v, v')$ acts on functions defined on a contour C in the v complex plane, where C is a positively-oriented circle centered at $\mathsf{R} \geq 0$ with radius $\mathsf{R}+\varepsilon$ so that $0 < \varepsilon \leq \min(1, \alpha+\beta)$. The kernel has the same form as for the continuum case, namely

$$K_u(v, v') = \frac{1}{2\mathbf{i}\pi}\int_{1/2+\mathbf{i}\mathbb{R}} \frac{\pi u^s}{\sin \pi s}\frac{g(v)}{g(v+s)}\frac{\mathrm{d}s}{s+v-v'} \tag{24}$$

except that now for the Beta random walk one has

$$g(v) = g_{\rm RW}(v) = \left(\frac{\Gamma(v)}{\Gamma(\alpha+v)}\right)^{\frac{T-\xi\sqrt{\frac{T}{2}}}{2}} \left(\frac{\Gamma(\alpha+\beta+v)}{\Gamma(\alpha+v)}\right)^{\frac{T+\xi\sqrt{\frac{T}{2}}}{2}} \Gamma(v) \tag{25}$$

where we recall that from now on we restrict to the case $\beta = \alpha$.

It is interesting to note that there is a way to obtain the continuum model from the discrete one, by taking the limit $\alpha \to 0$. Indeed if one sets

$$T = 2\tilde{T}/\alpha^2 \tag{26}$$

one has, with the same value of ξ

$$\lim_{\alpha\to 0} g_{\rm RW}(v) = g_c(v) \tag{27}$$

This corresponds to the convergence of the discrete random walk to the continuum one, which can be expressed as the convergence [23]

$$\alpha X(2\alpha^{-2}\tau) \underset{\alpha\to 0}{\to} y(\tau) \tag{28}$$

recalling that $X(t)$ corresponds to position in the Beta random walk with index α, and $y(\tau)$ to the position of the particle in the continuum model (2).

However this is not what we are interested in here. Instead we want to keep α fixed and take the time T of the Beta random walk to be large. We now argue that it leads to the same large-deviation rate functions as for the continuum model, up to the rescaling (26).

Since the form of the kernel is quite similar in both cases, to obtain the asymptotics of K_u in (24) and of the Fredholm determinant (23) for the Beta random walk, in the limit $T \to +\infty$ with $u = z\sqrt{T}$, we also only need to study the large time limit of the function $g_{RW}(v)$ under the same rescaling (19). Although we are working here for an arbitrary fixed α, we will choose the correspondence between the discrete and continuous time as in (26). Let us use the expansion at large v

$$v^{b-a}\frac{\Gamma(a+v)}{\Gamma(b+v)} = 1 + \frac{(a-b)(a+b-1)}{2v} + \frac{(a-b-1)(a-b)\left(3b^2+6ab-5b+a(3a-7)+2\right)}{24v^2} + \mathcal{O}\left(\frac{1}{v^3}\right) \tag{29}$$

for any $a, b = \mathcal{O}(1)$. Let us consider (25) with $\beta = \alpha$, express it as a function of $\tilde{T}$ using (26), and insert the rescaling (19). In the large $\tilde{T}$ limit one finds

$$\log g_{RW}(v) = \sqrt{\tilde{T}}\left(\phi(w) + (w+\xi)\log\sqrt{\tilde{T}}\right) + \chi_{RW}(w) - \frac{1}{2}\log(\sqrt{\tilde{T}}) + o(\tilde{T}) \tag{30}$$

with

$$\chi_{RW}(w) = (1-2\alpha)\left(\frac{1}{2w^2} - \frac{\xi}{2w}\right) + \frac{1}{2}\log(2\pi/w) \tag{31}$$

and the function $\phi(w)$ being identical to the one for the continuum model in (21).

Thus, in the large time limit we can identify $Z = \tilde{Z}$, i.e. the two random variables

$$Z_\xi(T) \equiv \tilde{Z}_\xi(\tilde{T}) \tag{32}$$

and identify separately each sides of (22) and (12) respectively which leads to

$$z\sqrt{T} \equiv \tilde{z}\sqrt{\tilde{T}} \tag{33}$$

$$\sqrt{T}\Psi_\xi^{\rm RW}(z) = \sqrt{\tilde{T}}\Psi_\xi(\tilde{z}) \tag{34}$$

which using the correspondence between continuum and discrete time in (26), finally leads to the prediction (14) and (15) for the rate function of the Beta random walk.

E. Explicit formula for the rate functions

We now recall the analytical prediction from [1] for the rate functions of the continuum model. Since $Z = \tilde{Z}$ and $H = \tilde{H}$, see section above, we use below only the notations Z and H in place of $\tilde{Z}$ and $\tilde{H}$. The rate function $\hat{\Phi}_\xi(Z)$ is obtained from the parametric representation

$$\begin{cases} \hat{\Phi}_\xi(Z) = \Psi_\xi(z) - \tilde{z}Z, \\ Z = \Psi'_\xi(\tilde{z})\,. \end{cases} \tag{35}$$

where $\Psi_\xi(\tilde{z})$ for $\xi \geq 0$ is given by

$$\Psi_\xi(\tilde{z}) = -\int_{\mathbb{R}} \frac{{\rm d}q}{2\pi} \frac{{\rm Li}_2(\tilde{z}(\mathbf{i}q - \frac{\xi}{2})e^{-q^2-\frac{\xi^2}{4}})}{(\mathbf{i}q - \frac{\xi}{2})^2} \tag{36}$$

where the principal value is required only for $\xi = 0$.

We now consider here only the case $\xi = 0$ where for any real value of $\tilde{z}$, the dilogarithm in the integrand of (36) does not have any branch cut on the real axis for q.

This expression for the rate function $\Psi_{\xi=0}(\tilde{z})$ then allows to obtain $\hat{\Phi}_{\xi=0}(Z)$ for any $Z \in [0,1]$. From this one obtains the rate function $\Phi_{\xi=0}(H) = \hat{\Phi}_{\xi=0}(Z)$ by the simple change of variable $H = \log Z$ for any $H \leq 0$. This is summarized in the Table I.

The case $\xi > 0$ is more involved and is given in the Appendix A 2. However one can give for any ξ the typical

interval of H	interval of $\tilde{z}$	$H =$	$\Phi_\xi(H) =$
$H \in \mathbb{R}^-$	$\tilde{z} \in \mathbb{R}$	$\log \Psi_0'(\tilde{z})$	$\Psi_0(\tilde{z}) - \tilde{z}\Psi_0'(\tilde{z})$

Table I. Case $\xi = 0$

value

$$Z_{\rm typ} = \overline{Z} = \Psi_\xi'(0) = \frac{1}{2}{\rm Erfc}\left(\frac{\xi}{2}\right)$$

and the variances of the PDF's $P(Z)$ and $P(H)$ for the continuum model [1]

$$\overline{Z^2}^c = \frac{1}{4\sqrt{2\pi\tilde{T}}} e^{-\xi^2/2} \tag{37}$$

$$\overline{H^2}^c = \frac{1}{\sqrt{2\pi\tilde{T}}} e^{-\xi^2/2} ({\rm Erfc}(\frac{\xi}{2}))^{-2} \tag{38}$$

The corresponding variances for the Beta random walk at large time are obtained by the correspondence $\tilde{T} = \frac{\alpha^2}{2} T$.

IV. METHODS

Next, we describe our numerical approaches. In Subsection IV A, we first state how we obtain, for each given sample $\omega = \{w_{x,t}\}$, as drawn from the beta distribution (3), the quantities Z according to (6) and therefore $H = \log Z$. We are interested in the distributions $P(Z)$ and $P(H)$. In Subsection IV B we explain how we achieve this over a large range of the support down to very small probability densities such as 10^{-50} or even smaller.

A. Random walk on a lattice

For each of the samples $\omega = \{w_{x,t}\}$, corresponding to the probabilities (4) to move left and right, we calculate the probability $Q(X|t)$ of reaching site X at step t. For this purpose we apply a dynamic programming, i.e., transfer matrix, approach by calculating

$$\begin{aligned} Q(0|0) &= 1 \\ Q(X|0) &= 0 \text{ for } X \neq 0 \\ Q(X|t+1) &= w_{X-1,t} Q(X-1|t) + \\ &\quad (1 - w_{X+1,t}) Q(X+1|t) \end{aligned} \tag{39}$$

for $t = 0, 1, \ldots, T-1$. For a walk of T steps, these probabilities can be calculated in $\mathcal{O}(T^2)$ time. This allows one to obtain the cumulative probability R of being right of some point X by simply summing

$$R(X|T) = \sum_{X'>X} Q(X'|T)\,, \tag{40}$$

which is achieved in $\mathcal{O}(T)$ steps, which is negligible compared to the $\mathcal{O}(T^2)$ steps to compute the (half) transfer matrix $Q(X|t)$. The value Z of (6) we are interest in is obtained by

$$Z = R(\sqrt{T/2}\,\xi\,|\,T) \tag{41}$$

where we round $\sqrt{T/2}\,\xi$ to the next lowest integer. Note that for small values of ξ, not all values of the matrix $Q(X|t)$ contribute. But even for $\xi = 0$, where walks contribute which reach $X = T/2$ and return to $X = 0$, one needs half of Q. Thus, the total computation time is always $\mathcal{O}(T^2)$.

The corresponding value of H is obtained simply by $H = \log(Z)$. Note that H is completely determined by the sample ω, so we can write $H = H(\omega)$.

B. Introduction to importance sampling

For the purpose of the introduction of the idea of importance sampling, we retain some elements of the presentation made in Ref. [95]. In principle one could obtain an estimate of the probability distribution $P(H)$ numerically from *direct sampling*. For this, one generates many disorder samples and calculates $H = \log Z$ for each one according to Eq. (41). Then the distribution is estimated by the suitably normalized histogram of the values of H. Nevertheless, this limits the smallest probabilities which can be resolved to the inverse of the number of samples, hence reaching probabilities as small as 10^{-50} is strictly impossible. Therefore, a different approach is required.

To estimate $P(H)$ for a much larger range, where probability densities as small as 10^{-50} may appear, we use a more powerful approach, called importance sampling as discussed in Refs. [75, 76]. This approach has been successfully applied to many problems in statistical physics and mathematics to obtain the tails of distributions arising in equilibrium and non-equilibrium situations [77–86]. The idea behind importance sampling is to sample the different disorder samples with an additional bias $\exp(-\theta H(\omega))$ where θ is an adjustable parameter interpreted as a fictive temperature. If $\theta > 0$ the samples with a negative H become more likely, conversely if $\theta < 0$ the samples with a positive H are favored. Now, it is not possible to sample the disorder samples ω directly when the bias is included. For this reason, a standard Markov-chain Monte Carlo simulation is used to sample according to the biased distribution [88, 89]. Here, one has a disorder sample ω as current configuration of the Markov chain, and the configurations change only slightly from step to step. In detail, at each step of the Markov chain, a new disorder sample ω^* is proposed by replacing on the current sample ω a certain fraction r of the random numbers $\omega = \{w_{x,t}\}$ by new random numbers which are drawn according to Eq. (3). The new disorder sample is then accepted with the usual Metropolis-Hastings probability

$$p_{\rm Met} = \min\{1, e^{-\theta[H(\omega^*) - H(\omega)]}\}, \tag{42}$$

otherwise the old configuration is kept [90]. By construction, the algorithm fulfils detailed balance and is ergodic, since within a sufficient number of steps, each possible sample may be constructed. Thus, in the limit of infinitely long Markov chains, the distribution of biased disorder samples will follow the probability

$$q_\theta(\omega) = \frac{1}{W(\theta)} P_{\rm dis}(\omega) e^{-\theta H(\omega)}\,, \tag{43}$$

where $P_{\rm dis}(\omega)$ is the original disorder distribution, i.e., the product of the Beta distributions for all disorder values, and $W(\theta) = \sum_\omega P_{\rm dis}(\omega) e^{-\theta H(\omega)}$ is the normalisation factor. Note that $W(\theta)$ also depends on the walk length T because of finite-size effects. $W(\theta)$ is generally unknown but can be determined, see below. The output of this Markov chain allows one to construct a biased histogram $P_\theta(H)$. In order to get the correct empirical probability density $P(H)$ one should unbias the result such that

$$P(H) = e^{\theta H} W(\theta) P_\theta(H). \tag{44}$$

Hence, the target distribution $P(H)$ can be estimated, up to a normalisation constant $W(\theta)$. For each value of the parameter θ, a specific range of the distribution $P(H)$ will be sampled and using a positive (respectively, negative) parameter allows one to sample the region of a distribution at the left (respectively, at the right) of its center.

For suitably chosen sets of temperature values θ, the ranges of support for neighboring densities $P_{\theta_i}(H)$ and $P_{\theta_{i+1}}(H)$ will overlap. Since after rescaling with $W(\theta_i)$ and $W(\theta_{i+1})$, respectively, they must be equal to $P(H)$. Thus, in particular they have to be in equal to each other, up to statistical fluctuations, for those values of H where they overlap. This allows one to determine rations $W(\theta_i)/W(\theta_{i+1})$ for all neighboring pairs of temperatures, and finally all absolute values $W(\theta_i)$ through the overall normalisation of $P(H)$, for details and examples see Appendix B and Ref. [75, 76]. Most accurately, the determination of the normalisation factors can be achieved using the Multi Histogram approach [96], see also the convenient tool of Peter Werner [97].

V. COMPARISON OF THE THEORETICAL PREDICTIONS WITH THE SIMULATIONS

We now compare the theoretical predictions of Section III with the numerical simulations of the finite-time random walks on a lattice, for various values of ξ. We insist on the fact that the comparison will be done without any fitting parameter.

A. Presentation of the simulations

The numerical simulations were run for walks of length $T \in \{64, 128, 256, 512, 1024\}$, the largest lengths only for some cases. Most of the walks are for distribution parameter $\alpha = 1$, which corresponds to a uniform $U(0,1)$ distribution, but in the beginning we also show some simple sampling results for other values of α, which indicate the universality with respect to α subject to simple scaling of the number of steps. We have evaluated the cumulative distribution of positions for several values of the asymmetry parameter $\xi = \{0, 1, 2, 3, 4, 5\}$.

For the large-deviation simulations, we have to make sure that the Markov chain is equilibrated. This can be confirmed by running the Markov chain for very different initial configurations of the sample ω. Extreme samples where all entries $\omega_{x,t}$ are close to 0, or all values are close to 1, correspond to extreme values of H. An impression of the convergence of the Markov chain is obtained monitoring $H(t_{\rm MC})$ as a function of the number $t_{\rm MC}$ of Monte Carlo steps and observing where these values agree within fluctuations for different initial configurations of ω, see Fig. 2. Evidently, the equilibration is obtained rather quickly, within few thousand MC steps.

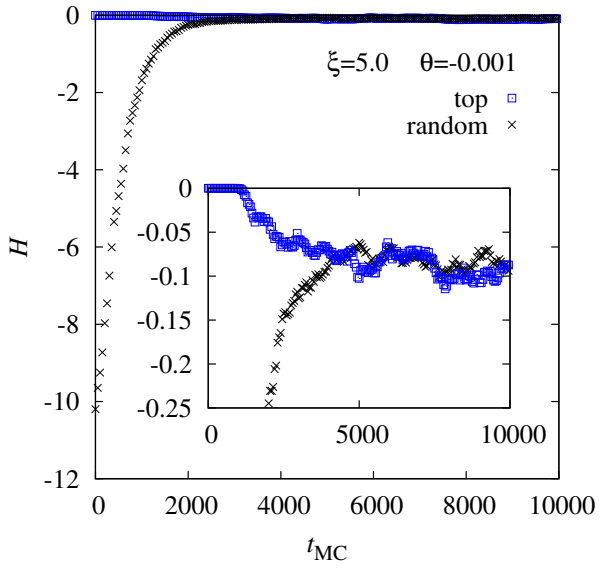

Figure 2. Equilibration of the Markov chain: Measured value of H as function of the number $t_{\rm MC}$ of Markov steps for the case $\xi = 5$, $T = 128$ and sampling temperature $\Theta = -0.001$ which corresponds to the very tail of the distribution which is hardest to reach. Two initial starting configurations ω resulting in very different initial values of H were chosen: one just typical random one, and one where $w_{x,t} = 0.999$ for all values ("top"). The inset enlarges the top part.

B. Variance for the case $\xi = 0$

First, we consider the probability distribution $P(H)$ as obtained by simple sampling for $\xi = 0$ and several values of the distribution parameter α. We determine its variance $\sigma^2 = \overline{H^2}^c$ as a function of the total length of the random walk T, up to $T = 2048$. Our analytical

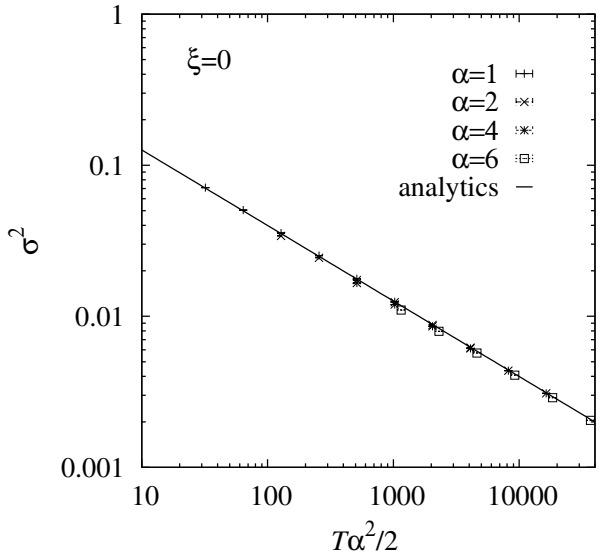

Figure 3. Variance σ^2 of the distribution $P(H)$ for different values of α, as a function of the scaled parameter $\tilde{T} = \frac{\alpha^2 T}{2}$ where T is the number of steps of the Beta random walk. It is compared to the analytical prediction (45) (solid line).

prediction at large T is, see (37)

$$\sigma^2 \simeq \frac{1}{\sqrt{2\pi\tilde{T}}} \simeq \frac{1}{\sqrt{2\pi}}\sqrt{\frac{2}{\alpha^2 T}} \tag{45}$$

In Fig. 3 we show the variance σ^2 as a function of the scaled time parameter $\tilde{T} = \frac{\alpha^2 T}{2}$. As visible, the data points fall nicely on one line, proving the universality with respect to α.

To see how well the expected limiting behavior (45) is reached, we plot in Fig. 4 the combination $\sigma^2 \tilde{T}^{1/2} = \sigma^2(\alpha^2 T/2)^{1/2}$. For all considered values of α, a convergence to the expected value $1/(2\pi)^{1/2}$ is visible. The convergence seems to be faster for smaller values of α, i.e., for more flat step distributions of the samples ω.

C. Distribution $P(H)$

The distribution of H is shown in Fig. 5 for $T = 128$ and three values of the asymmetry parameter ξ. As visible, with the large-deviation approach, here small probabilities such as 10^{-50} are reached. For increasing values of ξ, the probability of a walk ending beyond $\xi\sqrt{T/2}$ will decrease, which is reflected by a shift of $P(H)$ to more negative values. For a more detailed analysis and comparison with the analytic results, we consider from now on the rate functions.

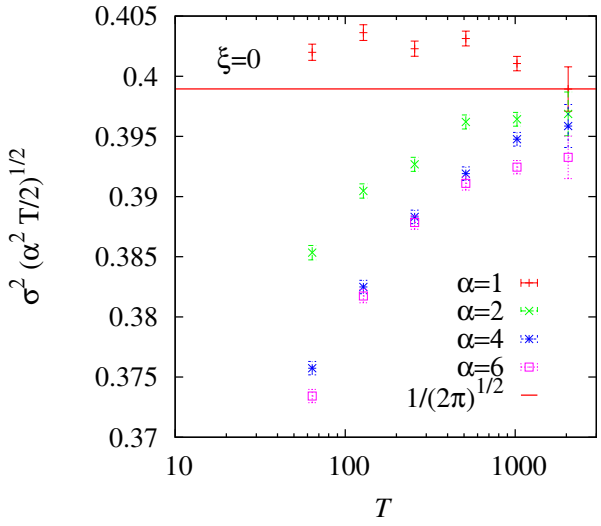

Figure 4. The variance σ^2 scaled by the expected limiting behavior, as function of step size T. The prediction $1/\sqrt{2\pi}$ from (45) is also shown.

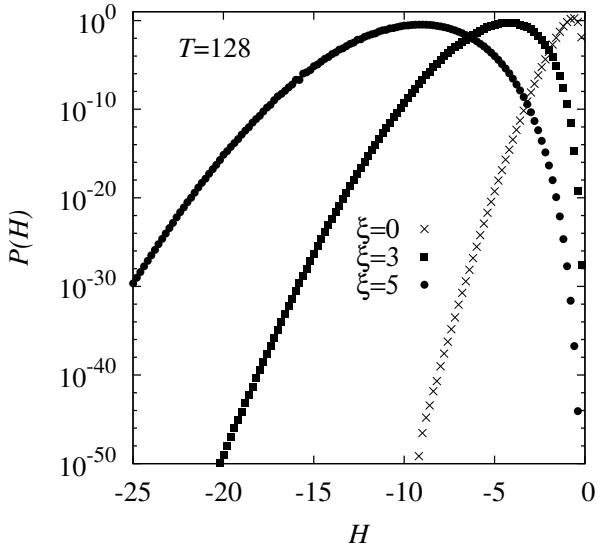

Figure 5. Distribution $P(H)$ for walk length $T = 128$ and three values of the asymmetry parameter ξ.

D. Rate functions

We will now test the analytical prediction (14) for the rate functions $\Phi_\xi^{\rm RW}(H)$ and $\hat{\Phi}_\xi^{\rm RW}(Z)$ defined in Eqs. (10) and (8) (where the rate functions on the r.h.s. of (14) are given in Sections III E and Appendix A). Note that in this numerical section we simplify notations and denote $\Phi(H) \equiv \Phi_\xi^{\rm RW}(H)$ and $\Phi(Z) \equiv \hat{\Phi}_\xi^{\rm RW}(Z)$. These rate functions are shown in Fig. 6 for walk length $T = 128$ and all considered values of the asymmetry parameter $\xi = \{0, 1, 2, 3, 4, 5\}$. Note that a value of Φ close to 20 correspond for $T = 128$ to a probability $e^{-\sqrt{128}\times 20} \approx 5 \times 10^{-99}$.

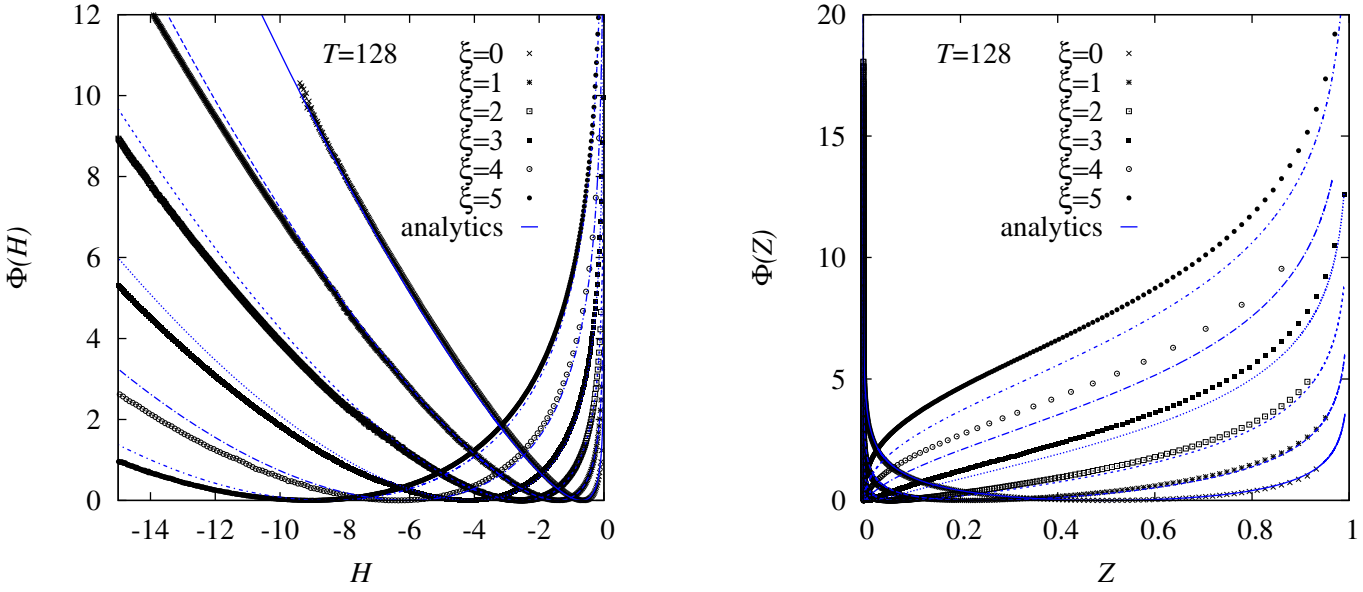

Figure 6. Rate functions $\Phi(H)$ and $\Phi(Z)$ for walk length $T = 128$ and various values of the asymmetry parameter $\xi = \{0, 1, 2, 3, 4, 5\}$. The lines show our analytical predictions for $T \to \infty$.

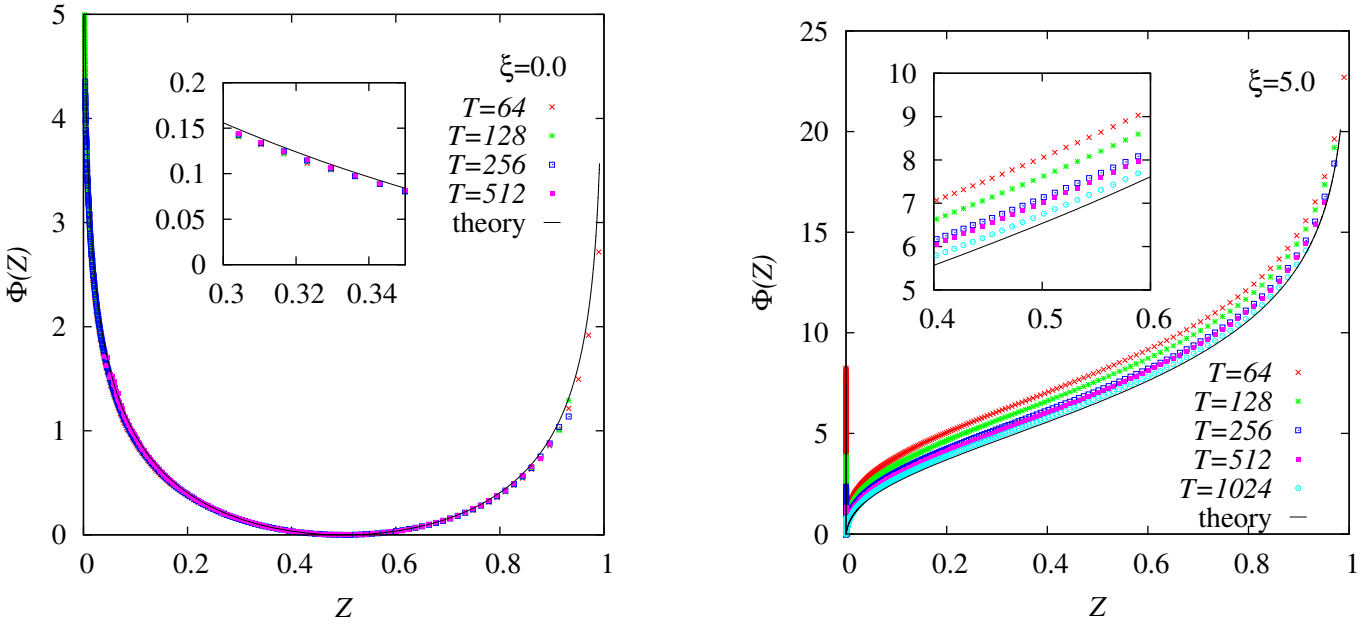

Figure 7. Rate function $\Phi(Z)$ for various walk lengths $T = \{64, 128, 256, 512\}$ and 1024, for the cases $\xi = 0$ (left, only up to $T = 512$) and $\xi = 5$ (right). The lines show the analytical results, respectively. The insets enlarge the regions $I \in [0.3, 0.35]$ (left) and $Z \in [0.4, 0.6]$ (right).

For small values of ξ already a good agreement between finite-T numerical data and analytical results is visible. Nevertheless, for values such as $\xi = 4$ and $\xi = 5$ substantial deviations are visible. For this reason, we have performed numerical simulations for the two extreme cases of the asymmetry $\xi = 0$ and $\xi = 5$ for various lengths of the walk $T = \{64, 128, 256, 512\}$, and even $T = 1024$ for $\xi = 5$. The results for $\Phi(Z)$ are shown in Fig. 7. For the case $\xi = 0$ basically all results agree, the limiting behavior is already visible for short walk length T. For $\xi = 5$ a clear convergence to the analytical result is visible. The fact that the finite-T corrections are stronger for larger values of ξ reminds one of the different of convergence speeds within the central limit theorem: The properly rescaled sum of random numbers attains a Gaussian shape near the typical values, corresponding to small values of ξ here, much faster than in the tails, corresponding to large values of ξ.

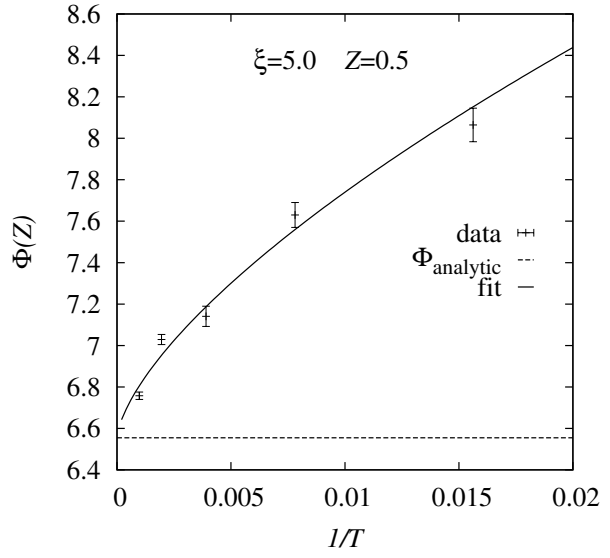

Figure 8. Extrapolation of the rate function for $\xi = 5$ to infinite walks lengths $T \to \infty$ by showing the numerical result of $\Phi(Z)$ as function of $1/T$ for a fixed value of $Z = 0.5$ and $T = 64, 128, 256, 512$ and 1024. The upper line shows the result of a fit according to Eq. (46). The horizontal line indicates the analytic asymptotic value for $\Phi(Z = 0.5)$.

We have also performed a heuristic extrapolation by fitting the behavior as function of T, for a fixed value of Z, to a power law according to

$$\Phi(Z,T) = \Phi_Z^{\infty} + a_Z T^{-bz} , \qquad (46)$$

i.e., with fitting parameters Φ^{∞}, a and b which may depend on Z. An example for such a fit is shown in Fig. 8. As visible, the extrapolated value is compatible with the analytical result.

E. Non convexity of $\Phi(Z)$ and first-order transition

A remarkable prediction of Ref. [1] is that for the continuum model the rate function $\hat{\Phi}_\xi(Z)$ in (14) is non-convex for $\xi > \sqrt{8}$. This results in a first-order phase transition in its Legendre transform (i.e. $\Psi_{\rm opt}(z)$ defined in (A11)) associated to a tilted version of $P(Z)$ (see definition below and in Eq. (A13)). For a detailed discussion see Appendix A 3, and for an illustration of the first-order transition see Fig. 10). We thus predict that the same property holds for the Beta random walk, as we will now confirm.

We find that for large enough values of ξ, the numerical rate function $\Phi(Z)$ for the random walk exhibits a non-monotonic curvature, which is already visible in Fig. 6. This leads to the appearance and disappearance of maxima in the tilted distribution $P(Z)\exp(-z(T/2)^{1/2}Z)$, depending on the choice of z. We recall that $\alpha = 1$ here, hence this factor is also $\exp(-z\tilde{T}^{1/2}Z)$, corresponding to (A13) for the continuum model. Our prediction is thus that at large T and for $\xi > \sqrt{8}$, $P(Z)\exp(-z(T/2)^{1/2}Z)$ should exhibit two peaks for $z \in [z_{c1}, z_{c2}]$ given in (A7). For $\xi = 5$ this corresponds to $z \in [-216.5, -13]$. The two peaks should become of same height for $z = z^*$ given in (A15), which for $\xi = 5$ evaluates to $z^* \simeq -17.84$.

In the numerical results, we indeed observe that, for not too negative values of z, the rescaled rate functions exhibits a peak close to $Z = 0$, see left of Fig. 9 the case $z = -12$. For intermediate values of z, a second peak appears, see the case $z = -14$. This is consistent with our analytical prediction recalled above that a second peak should appear for $z \simeq -13$. This peak becomes slightly more pronounced when increasing the number T of steps (not shown), and much more pronounced when decreasing the value of z.

For each value of T, there is a value $z^*(T)$ where both peaks exhibit the same height. The inset of left of Fig. 9 shows $z^*(T)$ together with a fit to a power law $z^*(T) = z^*_\infty + a^* T^{-b^*}$, which results in $z^*_\infty = 17(1)$ which is well compatible with the predicted value $z^* \simeq -17.84$.

For very negative values of z, the first peak disappears, see right of Fig. 9. The value of z above which this happens, which we predict to be $z_{c1} = -216.5$, is observed to be indeed very negative for small values of T, about $z = -1150$ for $T = 128$ and increases when increasing T, to about $z = -320$ for $T = 1024$. When fitting $z_{c1}(T)$ to a heuristic power law of the form $z_{c1}(T) = z_{c1}^{\infty} + a_z T^{-b_z}$ we obtain a limiting value $z_{c1}^{\infty} = -312(50)$ which is in rough agreement, i.e. within two sigma, with the limiting value $z_{c1} = -216.5$ (see inset).

The behavior of the numerically determined tilted PDF of Z is thus in agreement with the prediction, and the first-order transition in $\Psi_{\rm opt}(z)$ results when the second peak becomes higher than the first one (see Appendix A 3).

VI. CONCLUSION AND OUTLOOK

To summarize we have studied analytically and numerically the Beta random walk, a discrete time random walk on the square lattice with Beta distributed time dependent i.i.d. jump probabilities with parameter α. We have focused on the probability Z that a walk starting from the origin is at large time T at position to the right of $X = \xi\sqrt{T/2}$ for a given $\xi > 0$. We have determined analytically and numerically the law of large deviations of the observable Z. We have first predicted that the large-deviation rate function of the Beta random walk is identical, up to a scale factor involving α that we determined, to the one of the continuum model for diffusion in random media, which we recently obtained analytically (and is ξ-dependent). This prediction was based on the large time asymptotic analysis of an exact Fredholm determinant formula which exists for both the discrete and the continuum problem. The prediction holds for any value of the parameter α, hence it hints at some universality in the large-deviation rate functions. Proving this prediction rigorously remains an open question for

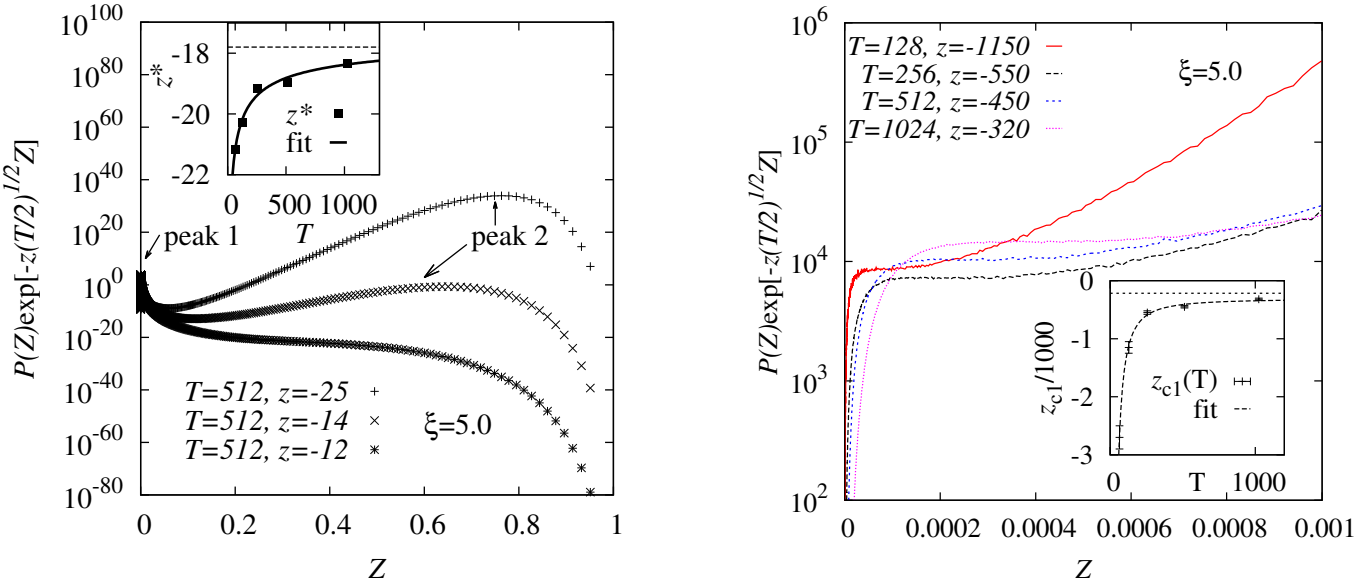

Figure 9. (left panel) Distributions $P(Z)$ rescaled with the factor $\exp(-z(T/2)^{1/2}Z)$. For intermediate values of z, the distribution is observed to exhibit two peaks, as predicted. The inset shows the dependence of the scaling value z^* where the two peaks attain the same height as function T and a fit to a shifted power law, see text. The horizontal line shows the predicted value $z^* \simeq -17.84$. (right panel) For very negative values of $z \leq z_{c1}(T)$, one observes that the first peak close to $Z = 0$ is suppressed. The inset illustrates the convergence to the predicted value $z_{c1} = -216.5$ at large T, which is included as horizontal line.

mathematicians.

In the absence of a rigorous proof, we performed a numerical test of this prediction. We used a large-deviation sampling approach to measure the PDF $P(Z)$ for various values of α and ξ. We were able to meaure the PDFs over many decades down to values as small as 10^{-100} and below. We obtained an accurate determination of the rate function and observed convergence at large T to the predicted analytical value. In addition, for a deeper investigation of the system properties beyond the rough shape of the distributions, we observed a first-order transition in the rate function $\Psi_{\rm opt}(z)$, which manifests itself as multiple peaks in the tilted PDF of Z, as predicted in Ref. [1]. Our numerical results are thus also a confirmation of the predictions obtained in that work. The numerical methods used here should be useful to study the large-deviation regime for various models of diffusion of the extremal particle in a cloud of many random walkers [67].

ACKNOWLEDGMENTS

We thank G. Barraquand for helpful discussions.

The simulations were performed at the HPC Cluster CARL, located at the University of Oldenburg (Germany) and funded by the DFG through its Major Research Instrumentation Program (INST 184/157-1 FUGG) and the Ministry of Science and Culture (MWK) of the Lower Saxony State.

Appendix A: Analytical results

We recall in this Appendix the analytical result of [1] for the continuum model for arbitrary $\xi \geq 0$. To simplify notations in this Appendix all subscripts ξ are implicit and $\tilde{z}$ is denoted z. We also recall that $Z = \tilde{Z}$ and $H = \tilde{H}$, hence we use only the notations Z and H in place of $\tilde{Z}$ and $\tilde{H}$.

The rate functions $\hat{\Phi}(Z) = \Phi(H)$ (with $H = \log Z$) is obtained quite generally from the parametric representation

$$\begin{cases} \hat{\Phi}(Z) = \Psi(z) - zZ, \\ Z = \Psi'(z)\,. \end{cases} \tag{A1}$$

While the rate function $\hat{\Phi}(Z)$ is well defined and single valued, for general $\xi > 0$, $\Psi(z)$ may have several branches. This can be seen in Fig. 10 where in some cases one value of z corresponds to one or three values of $Z = \Psi'(z)$.

ξ	$0 \le \xi \le \xi_1$	$\xi_1 \le \xi \le \xi_2$ $z_{c1} < z_{c2} < z_c$	$\xi_2 \le \xi$ $z_{c1} < z_c < z_{c2}$
$\Delta(z) =$	$\begin{cases} 0, & z_c < z \\ \Delta_1(z), & z < z_c \end{cases}$	$\begin{cases} 0, & z_c < z \\ \Delta_1(z), & z_{c1} < z < z_c \\ \Delta_2(z), & z_{c1} < z < z_{c2} \\ \Delta_3(z), & z < z_{c2} \end{cases}$	$\begin{cases} 0, & z_c < z \\ \Delta_1(z), & z_{c1} < z < z_c \\ \Delta_2(z), & z_{c1} < z < z_c \\ \Delta_2(z) - \Delta_1(z), & z_c < z < z_{c2} \\ \Delta_3(z) - \Delta_1(z), & z_c < z < z_{c2} \\ \Delta_3(z), & z < z_c \end{cases}$

Table II. Determination of the jump function $\Delta(z)$ in the different phases in the case $\xi \geq 0$. One has $z_c = -\frac{2}{\xi} e^{\xi^2/4} \leq 0$ and the points $z = z_{c1}$ and $z = z_{c2}$ are turning points which depend on ξ. In the interval $z \in [z_{c1}, z_{c2}]$, the function $\Delta(z)$ is multi-valued (i.e. it has several branches) due to these turning points. The definition of Δ_ℓ is given in (A4).

Below, we first explain how to compute $\Psi(z)$ and its various branches, and then we explain how to perform the Legendre inversion. Finally we discuss the multivaluation and the first-order transition of the "optimal" $\Psi_{\rm opt}(z)$, see Fig. 10.

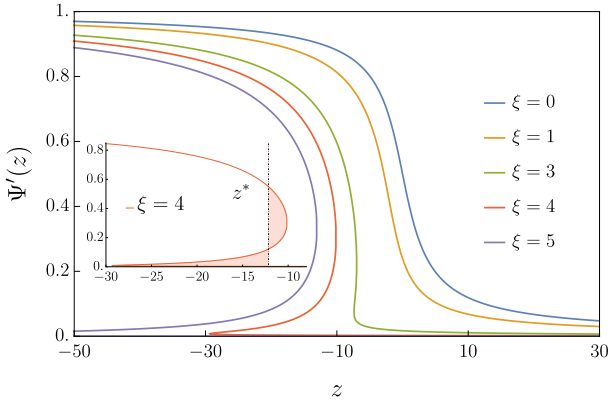

Figure 10. For $\xi = (0, 1, 2, 3, 4, 5)$ we plot the derivative rate function $\Psi'(z)$ from Table II as a function of z, with $\Psi'(+\infty) = 0$ and $\Psi'(-\infty) = 1$ (all the branches are shown). For $\xi > \xi_1$ and $z \in [z_{c1}, z_{c2}]$ the function is multi-valued (see text). (**Inset**) First-order transition: at $z = z^*$ such that the areas of the two shaded regions become equal the value of (the optimal) $\Psi'_{\rm opt}(z)$ (see definition in (A14)) from one branch to the other, shown here for $\xi = 4$.

1. How to compute $\Psi(z)$

One first defines

$$\Psi^0(z) = -\int_{\mathbb{R}} \frac{\mathrm{d}q}{2\pi} \frac{\mathrm{Li}_2(z(\mathbf{i}q - \frac{\xi}{2}) e^{-q^2 - \frac{\xi^2}{4}})}{(\mathbf{i}q - \frac{\xi}{2})^2} \tag{A2}$$

The general formula for $\Psi(z)$ takes the form

$$\Psi(z) = \Psi^0(z) + \Delta(z) \tag{A3}$$

where $\Psi^0(z)$ is the same integral as in (A2). Note that we compute $\Psi^0(z)$ numerically using the default `PolyLog` function in Mathematica together with the `NIntegrate` routine for integration (with in some cases a branch cut on the integration contour dealt automatically by Mathematica). The convention $\Delta(z) = 0$ defines the main branch of $\Psi(z)$. The other branches and the form of $\Delta(z)$ as a function of ξ and z are shown in Table II. Note that for $z \geq 0$ one has $\Delta(z) = 0$.

The jump functions $\Delta_\ell(z)$ for $\ell = \{1, 2, 3\}$ which appear in this Table are defined as follows. First one has

$$\Delta_\ell(z) = \hat{\Delta}(p_\ell(z, \xi)) \tag{A4}$$

where

$$\hat{\Delta}(p) = \frac{1}{\xi}\Big[- (\xi^2 + 2)(\log(\xi) - \log(\xi + 2p)) + 2p(p - \xi) - \frac{4p}{\xi + 2p}\Big] \tag{A5}$$

The $p_\ell(z, \xi)$ are the real roots of the equation for p

$$e^{-p^2 + \frac{\xi^2}{4}} + z(p + \frac{\xi}{2}) = 0\,. \tag{A6}$$

The behavior of these roots is as follows.

Let us define $z_c = -\frac{2}{\xi} e^{\frac{\xi^2}{4}}$. For $z_c \le z \le 0$, there is always one positive zero to (A6) denoted $p_1 = p_1(z, \xi)$. For $z < z_c$, the zeroes of (A6) are all negative and their number is:

1. for $0 < \xi < \xi_1 = \sqrt{8}$, there is one zero $p_1(z, \xi)$;

2. for $\xi_1 < \xi$ and $z \in]z_{c1}, z_{c2}[$ there are three zeroes $p_1(z, \xi) > p_2(z, \xi) > p_3(z, \xi)$. The zeroes degenerate, i.e. $p_1 = p_2$ for $z = z_{c1}$ and $p_2 = p_3$ for $z = z_{c2}$ which define z_{c1}, z_{c2}. For $z > z_{c2}$, there is only one zero $p_1(z, \xi)$. For $z < z_{c1}$, there is only one zero $p_3(z, \xi)$.

Note that $z_{c1} < z_{c2} < 0$, with $z_{c1} = z_{c2}$ at $\xi = \xi_1$, and their explicit expression and dependence on ξ is given for $\xi > \xi_1 = \sqrt{8}$ by

$$\begin{aligned} z_{c1} &= -\frac{1}{2} e^{\frac{1}{8}\left(\xi\left(\xi + \sqrt{\xi^2 - 8}\right) + 4\right)} \left(\xi - \sqrt{\xi^2 - 8}\right) \\ z_{c2} &= -\frac{1}{2} e^{\frac{1}{8}\left(\xi\left(\xi - \sqrt{\xi^2 - 8}\right) + 4\right)} \left(\xi + \sqrt{\xi^2 - 8}\right) \end{aligned} \tag{A7}$$

Note that z_c and z_{c2} become equal at the value $\xi = \xi_2$ with

$$\begin{aligned} \xi_2 &= -2\sqrt{\frac{2}{-2W_{-1}\left(-\frac{1}{2\sqrt{e}}\right) - 1}}\, W_{-1}\left(-\frac{1}{2\sqrt{e}}\right) \\ &\simeq 3.13395 \end{aligned} \tag{A8}$$

where W_{-1} is the Lambert function [92].

2. Inversion of Legendre transform

Defining the critical height $H_c = \log \Psi_0'(z_c)$, the rate function $\Phi(H)$ is given by the parametric representation displayed in Table III for $\xi \le \xi_1 = \sqrt{8}$.

interval of H	interval of z	$H =$	$\Phi(H) =$
$H \le H_c$	$z_c \le z$	$\log \Psi_0'(z)$	$\Psi_0(z) - z\Psi_0'(z)$
$0 > H > H_c$	$z_c > z$	$\log(\Psi_0'(z) + \Delta_1'(z))$	$\Psi_0(z) + \Delta_1(z) - z(\Psi_0'(z) + \Delta_1'(z))$

Table III. Case $0 < \xi \le \xi_1$

For $\xi_1 < \xi \le \xi_2$ it is given by the parametric representation displayed in Table IV where we have defined

$$\begin{aligned} H_{c1} &= \log(\Psi_0'(z_{c1}) + \Delta_1'(z_{c1})) = \log(\Psi_0'(z_{c1}) + \Delta_2'(z_{c1})), \\ H_{c2} &= \log(\Psi_0'(z_{c2}) + \Delta_2'(z_{c2})) = \log(\Psi_0'(z_{c2}) + \Delta_3'(z_{c2})), \end{aligned} \tag{A9}$$

interval of H	interval of z	$H =$	$\Phi(H) =$
$H \leq H_c$	$z_c \leq z$	$\log \Psi_0'(z)$	$\Psi_0(z) - z\Psi_0'(z)$
$H_c < H \leq H_{c1}$	$z_{c1} \leq z < z_c$	$\log(\Psi_0'(z) + \Delta_1'(z))$	$\Psi_0(z) + \Delta_1(z) - z(\Psi_0'(z) + \Delta_1'(z))$
$H_{c1} < H \leq H_{c2}$	$z_{c1} < z \leq z_{c2}$	$\log(\Psi_0'(z) + \Delta_2'(z))$	$\Psi_0(z) + \Delta_2(z) - z(\Psi_0'(z) + \Delta_2'(z))$
$H_{c2} < H < 0$	$z_{c2} > z$	$\log(\Psi_0'(z) + \Delta_3'(z))$	$\Psi_0(z) + \Delta_3(z) - z(\Psi_0'(z) + \Delta_3'(z))$

Table IV. Case $\xi_1 < \xi \leq \xi_2$

For $\xi_2 < \xi$ it is given by the parametric representation displayed in Table V where we have defined

$$\begin{aligned} H_{c10} &= \log(\Psi_0'(z_{c1}) + \Delta_1'(z_{c1})), \\ H_{c11} &= \log(\Psi_0'(z_c) + \Delta_2'(z_c)), \\ H_{c20} &= \log(\Psi_0'(z_{c2}) + \Delta_2'(z_{c2}) - \Delta_1'(z_{c2})), \\ H_{c21} &= \log(\Psi_0'(z_c) + \Delta_3'(z_c)), \end{aligned} \tag{A10}$$

interval of H	interval of z	$H =$	$\Phi(H) =$
$H \leq H_c$	$z_c \leq z$	$\log \Psi_0'(z)$	$\Psi_0(z) - z\Psi_0'(z)$
$H_c < H \leq H_{c10}$	$z_{c1} \leq z < z_c$	$\log(\Psi_0'(z) + \Delta_1'(z))$	$\Psi_0(z) + \Delta_1(z) - z(\Psi_0'(z) + \Delta_1'(z))$
$H_{c10} < H \leq H_{c11}$	$z_{c1} < z \leq z_c$	$\log(\Psi_0'(z) + \Delta_2'(z))$	$\Psi_0(z) + \Delta_2(z) - z(\Psi_0'(z) + \Delta_2'(z))$
$H_{c11} < H \leq H_{c20}$	$z_c < z \leq z_{c2}$	$\log(\Psi_0'(z)+\Delta_2'(z)-\Delta_1'(z))$	$\Psi_0(z)+\Delta_2(z)-\Delta_1(z)-z(\Psi_0'(z)+\Delta_2'(z)-\Delta_1'(z))$
$H_{c20} < H \leq H_{c21}$	$z_c \leq z < z_{c2}$	$\log(\Psi_0'(z)+\Delta_3'(z)-\Delta_1'(z))$	$\Psi_0(z)+\Delta_3(z)-\Delta_1(z)-z(\Psi_0'(z)+\Delta_3'(z)-\Delta_1'(z))$
$H_{c21} < H < 0$	$z_c > z$	$\log(\Psi_0'(z) + \Delta_3'(z))$	$\Psi_0(z) + \Delta_3(z) - z(\Psi_0'(z) + \Delta_3'(z))$

Table V. Case $\xi_2 < \xi$

3. Multi-valuation and first-order transition

To interpret the S-shape form of $\Psi'(z)$ shown with all its branches in Figure 10, we recall the definition of the "optimal" $\Psi(z)$ defined as

$$\Psi_{\rm opt}(z) = \min_{Z\in[0,1]} [\hat{\Phi}(Z) + zZ] \tag{A11}$$

It has the property that its derivative obeys

$$\Psi'_{\rm opt}(z) = \langle Z \rangle_z \tag{A12}$$

where $\langle Z \rangle_z$ is the expectation value for large $\tilde{T}$ of the random variable Z under the z-dependent tilted measure

$$P(Z)e^{-\sqrt{\tilde{T}}zZ} \sim e^{-\sqrt{\tilde{T}}(\hat{\Phi}(Z)+zZ)} \tag{A13}$$

The key point is that for $\xi > \xi_1$ the function $\hat{\Phi}(Z)$ has a concave part [1]. As a consequence, for $z \in [z_{c1}, z_{c2}]$ the tilted measure (A13) develops three extrema at $Z_j(z) = e^{H_j(z)}$, solutions of $\hat{\Phi}'(Z) = -z$. They lead to the three branches

of $\Psi'(z) = Z_j(z)$. The "optimal" $\Psi_{\rm opt}(z)$ is determined by the absolute minimum in (A11) (which corresponds to the absolute maximum in the tilted PDF of Z) hence it is given by

$$\Psi_{\rm opt}(z) = \min_{j=1,2,3} [\hat{\Phi}(Z_j) + zZ_j] \tag{A14}$$

and the optimal j switches from $j = 1$ to $j = 3$ at $z = z^*(\xi)$ where z^* is the solution of

$$\Delta_1(z^*) = \Delta_3(z^*)\,. \tag{A15}$$

It is also the point given by an equal area law on the curve $\Psi'(z)$, as in standard magnetization versus field curve for a first-order phase transition, see Fig. 10 (inset). The points $Z = \{Z_1, Z_3\}$ are "stable" whereas $Z = Z_2$ is "unstable". The optimal rate function $\Psi_{\rm opt}(z)$ thus exhibits a first-order transition. This type of transition occurs in other large-deviation problems [47].

Appendix B: Technical details of the importance sampling algorithm

To sample a wide range of values of H, one chooses a suitable set of parameters $\{\theta_{-N_n}, \theta_{-N_n+1}, \ldots, \theta_{N_p-1}, \theta_{N_p}\}$, N_n and N_p being the number of negative and positive parameters, to access the large-deviation regimes (left and right). The normalisation constants $W(\theta_i)$ are obtained by first computing the histogram using direct sampling, corresponds to $\theta = 0$. Then for θ_{+1}, one matches the right part of the biased histogram with the left tail of the unbiased one and for θ_{-1}, one matches the left part of the biased histogram with the right tail of the unbiased one. Similarly one iterates for the other values of θ and the corresponding *relative* normalisation constants can be obtained. In the end the full distribution is normalized to result in a total probability of one.

[1] A. Krajenbrink and P. Le Doussal. *Crossover from the macroscopic fluctuation theory to the Kardar-Parisi-Zhang equation controls the large deviations beyond Einstein's diffusion.* Physical Review E, 107(1):014137, (2023). See also arXiv:2204.04720.

[2] E. Bettelheim, N. R. Smith, B. Meerson, *Full Statistics of Nonstationary Heat Transfer in the Kipnis-Marchioro-Presutti Model*, J. Stat. Mech. 093103, (2022).

[3] A. Grabsch, P. Rizkallah, A. Poncet, P. Illien, and O. Bénichou. *Exact spatial correlations in single-file diffusion.* Physical Review E, 107(4):044131, 2023.

[4] K. Mallick, H. Moriya, and T. Sasamoto. *Exact solution of the macroscopic fluctuation theory for the symmetric exclusion process.* Phys. Rev. Lett. 129, 040601, (2022).

[5] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, *Macroscopic fluctuation theory.* Reviews of Modern Physics, 87(2):593, (2015).

[6] B. Derrida, *Non equilibrium steady states: fluctuations and large deviations of the density and of the current* Journal of Statistical Mechanics: Theory and Experiment, 2007(07):P07023, (2007).

[7] Derrida B, *An exactly soluble non-equilibrium system: the asymmetric simple exclusion process.* Physics Reports, 301(1-3):65–83, (1998).

[8] C. A. Tracy and H. Widom. *Asymptotics in ASEP with step initial condition.* Comm. Math. Phys., 290(1):129–154, (2009).

[9] M. Kardar, G. Parisi and Y-C. Zhang, *Dynamic Scaling of Growing Interfaces*, Phys. Rev. Lett. **56**, 889, (1986).

[10] L. Bertini and G. Giacomin. Stochastic burgers and KPZ equations from particle systems. Communications in mathematical physics, 183(3):571–607, (1997).

[11] I. V. Kolokolov, S. E. Korshunov, *Explicit solution of the optimal fluctuation problem for an elastic string in random potential*. Phys. Rev. E **80**, 031107, (2009); *Universal and non-universal tails of distribution functions in the directed polymer and KPZ problems*. Phys. Rev. B **78**, 024206, (2008); *Optimal fluctuation approach to a directed polymer in a random medium*. Phys. Rev. B **75**, 140201, (2007).

[12] B. Meerson, E. Katzav, A. Vilenkin, *Large Deviations of Surface Height in the Kardar-Parisi-Zhang Equation*, Physical Review Letters **116**, 070601, (2016).

[13] N. R. Smith, B. Meerson, and A. Vilenkin. *Time-averaged height distribution of the Kardar-Parisi-Zhang interface.* Journal of Statistical Mechanics: Theory and Experiment, 2019(5):053207, (2019).

[14] A. Krajenbrink and P. 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):064101, (2021).

[15] A. Krajenbrink and P. L. 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, 054142, (2021).

[16] Shabat, A., and V. Zakharov. *Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media.* Soviet physics JETP 34.1 (1972).

[17] Ablowitz, M. J., Kaup, D. J., Newell, A. C., and Segur, H. *The inverse scattering transform-Fourier analysis for nonlinear problems.* Studies in Applied Mathematics, 53(4), 249-315, (1974).

[18] D. J. Kaup and A. C. Newell. *An exact solution for a derivative nonlinear Schrödinger equation. Journal of Mathematical Physics, 19(4):798–801*, (1978).

[19] G. Barraquand and I. Corwin. *Random-walk in beta-distributed random environment.* Probab. Theory Rel. Fields, 167(3):1057–1116, (2017).

[20] P. Le Doussal, T. Thiery, *Diffusion in time-dependent random media and the Kardar-Parisi-Zhang equation*, Phys. Rev. E 96, 010102 (2017).

[21] I. Corwin and Y. Gu. *Kardar–Parisi–Zhang equation and large deviations for random walks in weak random environments.* J. Stat. Phys., 166(1):150–168, (2017).

[22] T. Thiery, P. Le Doussal, *Exact solution for a random walk in a time-dependent 1D random environment: the point-to-point Beta polymer*, Journal of Physics A: Mathematical and Theoretical 50 4, (2016).

[23] G. Barraquand and M. Rychnovsky. *Large deviations for sticky Brownian motions.* arXiv:1905.10280, (2019).

[24] G. Barraquand and P. Le Doussal. *Moderate deviations for diffusion in time dependent random media.* Journal of Physics A: Mathematical and Theoretical 53.21: 215002, (2020).

[25] D. Bernard and P. Le Doussal. *Entanglement entropy growth in stochastic conformal field theory and the KPZ class.* Europhysics Letters, 131(1):10007, (2020).

[26] G. Barraquand, *Some integrable models in the KPZ universality class*, Probability [math.PR]. Université Paris Diderot – Paris 7, 2015. English. tel-01167855 HAL Id: tel-01167855 https://tel.archives-ouvertes.fr/tel-01167855.

[27] C. Kipnis, C. Marchioro and E. Presutti, *Heat flow in an exactly solvable model* J. Stat. Phys. 27, 65, (1982)

[28] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Current Fluctuations in Stochastic Lattice Gases* Phys. Rev. Lett. 94, 030601 (2005).

[29] L. Bertini, D. Gabrielli, and J. L. Lebowitz. *Large deviations for a stochastic model of heat flow.* Journal of statistical physics, 121(5):843–885, (2005).

[30] B. Derrida and A. Gerschenfeld, *Current Fluctuations in One Dimensional Diffusive Systems with a Step Initial Density Profile* J. Stat. Phys. 137, 978 (2009).

[31] V. Lecomte, A. Imparato, and F. van Wijland, *Current Fluctuations in Systems with Diffusive Dynamics, in and out of Equilibrium* Prog. Theor. Phys. Suppl. 184, 276 (2010).

[32] P. L. Krapivsky and B. Meerson, *Fluctuations of current in nonstationary diffusive lattice gases* Phys. Rev. E 86, 031106 (2012).

[33] L. Zarfaty and B. Meerson, *Statistics of large currents in the Kipnis–Marchioro–Presutti model in a ring geometry* J. Stat. Mech. 033304 (2016).

[34] T. Bodineau and B. Derrida, *Distribution of current in nonequilibrium diffusive systems and phase transitions* Phys. Rev. E 72, 066110 (2005)

[35] J. Tailleur, J. Kurchan, and V. Lecomte, *Mapping Nonequilibrium onto Equilibrium: The Macroscopic Fluctuations of Simple Transport Models* Phys. Rev. Lett. 99, 150602 (2007)

[36] P. I. Hurtado and P. L. Garrido, *Spontaneous Symmetry Breaking at the Fluctuating Level* Phys. Rev. Lett. 107, 180601 (2011). A. Prados, A. Lasanta, and P. I. Hurtado, *Nonlinear driven diffusive systems with dissipation: Fluctuating hydrodynamics* Phys. Rev. E 86, 031134 (2012). C. Gutierez-Ariza and P. I. Hurtado, *The kinetic exclusion process: a tale of two fields* J. Stat. Mech. 103203 (2019) .

[37] M. A. Peletier, F. H. J. Redig, and K. Vafayi, *Large deviations in stochastic heat-conduction processes provide a gradient-flow structure for heat conduction* J. Math. Phys. 55, 093301 (2014).

[38] O. Shpielberg, Y. Don, and E. Akkermans, *Numerical study of continuous and discontinuous dynamical phase transitions for boundary-driven systems* Phys Rev E 95, 032137 (2017).

[39] A. Grabsch, A. Poncet, P. Rizkallah, P. Illien, and O. Bénichou. *Closing and solving the hierarchy for large deviations and spatial correlations in single-file diffusion.* arXiv:2110.09269, (2021).

[40] A. Poncet, A. Grabsch, P. Illien, and O. Bénichou. *Generalized correlation profiles in single-file systems.* Physical review letters, 127(22):220601, (2021).

[41] Since $\hat{\Phi}(Z)$ may not be convex it should be called a Legendre-Fenchel transform, which is not involutive [47].

[42] We will keep β as a parameter but for the application to obtain $\Psi(z)$ it is understood that it is set to $\beta = -1$.

[43] Comparing with [14] the "true" coupling constant is in fact $\hat{g} = \Lambda g = -z\frac{\xi}{2}e^{-\frac{\xi^2}{4}}$.

[44] Bettelheim, Eldad, Naftali R. Smith, and Baruch Meerson. *Inverse Scattering Method Solves the Problem of Full Statistics of Nonstationary Heat Transfer in the Kipnis-Marchioro-Presutti Model.* Phys. Rev. Lett. 128, 130602, (2021).

[45] Wadati, Miki, Kimiaki Konno, and Yoshi-Hiko Ichikawa. *A generalization of inverse scattering method.* No. IPPJ–381. Nagoya Univ.(Japan). Inst. of Plasma Physics, (1979).

[46] When the integrand has a jump at $q = 0$, one integrates respecting the symmetry $\mathbf{i}q \to -\mathbf{i}q$, equivalently one integrates on $\mathbb{R}^-$ and take twice the real part.

[47] Hugo Touchette. *The large deviation approach to statistical mechanics.* Physics Reports, 478(1-3):1–69, (2009).

[48] A. Krajenbrink, P. Le Doussal, *In preparation*

[49] P. Le Doussal, S. N. Majumdar, A. Rosso, G. Schehr, *Exact short-time height distribution in 1D KPZ equation and edge fermions at high temperature*, Phys. Rev. Lett. **117**, 070403, (2016).

[50] For the KPZ equation the partition sum $Z_{\rm KPZ} = e^{H_{\rm KPZ}}$ is an unbounded random variable. As a result the rate function $\Psi(\tilde{z})$ is undefined for $\tilde{z} < -1$. The present convergence results show precisely how, in the large ξ limit, one goes from a

bounded random variable Z to an unbounded one $Z_{\rm KPZ}$.
[51] A. Krajenbrink, P. Le Doussal, *Simple derivation of the* $(-\lambda H)^{5/2}$ *large deviation tail for the 1D KPZ equation*, J. Stat. Mech. 063210, (2018).
[52] A. Krajenbrink, P. Le Doussal, S. Prolhac, *Systematic time expansion for the Kardar-Parisi-Zhang equation, linear statistics of the GUE at the edge and trapped fermions.* Nuclear Physics B, **936** 239–305, (2018).
[53] 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.
[54] Ivan Corwin, *Private communication and in preparation.*
[55] Although it maps to the $\{P, Q\}$ system with $\delta - \delta$ initial conditions (as was the case in the study of the WNT of the KPZ equation with droplet initial condition) because of the highly non-local nature of the mapping it does not map SSEP observables to KPZ ones in any obvious way.
[56] L.-C. Tsai. *Integrability in the weak noise theory.* Transactions of the American Mathematical Society, (2023).
[57] see e.g. Section 3.3.1. in M. Dunajski. *Solitons, Instantons and Twistors.* Oxford University Press, Oxford, 2009.
[58] Kisil, Anastasia V. *The relationship between a strip Wiener–Hopf problem and a line Riemann–Hilbert problem.* IMA Journal of Applied Mathematics 80.5: 1569-1581, (2015).
[59] Chapter 1.3 in Noble, B. (1958) *Methods Based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations.* International Series of Monographs on Pure and Applied Mathematics, vol. 7. New York: Pergamon Press.
[60] A. Krajenbrink, P. Le Doussal, *Exact short-time height distribution in the one-dimensional Kardar-Parisi-Zhang equation with Brownian initial condition*, Phys. Rev. E **96**, 020102, (2017).
[61] Yu, Jinjiong. *Edwards–Wilkinson fluctuations in the Howitt–Warren flows.* Stochastic Processes and their Applications 126, no. 3: 948-982, (2016).
[62] Balázs, M., Rassoul-Agha, F., Seppalainen, T.. *The random average process and random walk in a space-time random environment in one dimension.* Communications in mathematical physics, 266(2), 499-545, (2006).
[63] Howitt, Chris, and Jon Warren. *Consistent families of Brownian motions and stochastic flows of kernels.* The Annals of Probability 37.4: 1237-1272, (2009).
[64] D. Brockington, J. Warren *The Bethe Ansatz for Sticky Brownian Motions*, arXiv:2104.06482
[65] D. Brockington and J. Warren. *At the edge of a cloud of Brownian particles.* arXiv:2208.11952, 2022.
[66] S. Das, H. Drillick, and S. Parekh. *KPZ equation limit of sticky Brownian motion.* arXiv:2304.14279, (2023).
[67] J. B. Hass, A. N. Carroll-Godfrey, I. Corwin, and E. I. Corwin. *Anomalous fluctuations of extremes in many-particle diffusion.* Physical Review E, 107(2):L022101, 2023.
[68] K. Gawedzki and P. Horvai, *Sticky behavior of fluid particles in the compressible Kraichnan model.* J. Stat. Phys., 116(5-6):1247–1300, (2004).
[69] H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer, New York, 1991)
[70] M. Wadati and K. Sogo. *Gauge transformations in soliton theory.* Journal of the Physical Society of Japan, 52(2):394–398, (1983).
[71] Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C., *Towards a nonequilibrium thermodynamics: a self-contained macroscopic description of driven diffusive systems.* J. Stat. Phys. 135, 857–872, (2009).
[72] G. Barraquand, M. Rychnovsky, *Random walk on nonnegative integers in beta distributed random environment*, arXiv:2201.07270, (2022).
[73] Tsirelson, Boris. *Scaling Limit, Noise, Stability.* In Lectures on probability theory and statistics, pp. 1-106. Springer, Berlin, Heidelberg, (2004).
[74] Y.L. Le Jan and O. Raimond. *Sticky flows on the circle and their noises.* Probability Theory and Related Fields, 129(1), 63-82 (2004) *The noise of a Brownian sticky flow is black.* arXiv math.PR/0212269 (2008).
[75] A. K. Hartmann, *Sampling rare events: Statistics of local sequence alignments Phys. Rev. E* **65** 056102, (2001).
[76] A. K. Hartmann, *Large-deviation properties of largest component for random graphs Eur. Phys. J. B* **84**, 627, (2011).
[77] A. Engel A., R. Monasson, A. K. Hartmann, *On Large Deviation Properties of Erdos-Rényi Random Graphs J. Stat. Phys.* **117**, 387, (2004).
[78] A. K. Hartmann, *Calculation of partition functions by measuring component distributions*, Physical Review Letters, 94(5), (2005).
[79] C. Monthus T. Garel, *Probing the tails of the ground-state energy distribution for the directed polymer in a random medium of dimension* $d = 1, 2, 3$ *via a Monte Carlo procedure in the disorder*, Phys. Rev. E **74**,051109, (2006).
[80] S. Wolfsheimer A. K. Hartmann, *Minimum (Free-) Energy Distribution of RNA Secondary Structures: Entropic and Thermodynamic Properties of Large Deviations*, Phys. Rev. E **82**, 021902, (2010).
[81] T. A. Driscoll, K. L. Maki *Searching for Rare Growth Factors Using Multicanonical Monte Carlo Methods* SIAM Review **49**, 673, (2007).
[82] N. Saito, Y. Iba, K. Hukushima *Multicanonical sampling of rare events in random matrices* Phys. Rev. E **82**, 031142, (2010).
[83] A. K Hartmann, S. N. Majumdar, A. Rosso *Sampling fractional Brownian motion in presence of absorption: A Markov chain method* Phys. Rev. E **88**, 022119, (2013).
[84] A. K. Hartmann, *High-precision work distributions for extreme nonequilibrium processes in large systems*, Phys. Rev. E **89**, 052103, (2014).
[85] G. Claussen, A. K. Hartmann, S. N. Majumdar, *Convex hulls of random walks: Large-deviation properties*, Phys. Rev. E **91**, 052104, (2015).

[86] T. Dewenter, G. Claussen, A. K. Hartmann, S. N. Majumdar, *Convex hulls of multiple random walks: A large-deviation study*, Phys. Rev. E **94**, 052120, (2016).
[87] J, M. Hammersley K. W. Morton, *A new Monte Carlo technique: antithetic variates*, Math. Proc. Cambr. Phil. Soc. **52**, 449–475, (1956).
[88] M. E. J. Newman , G. T. Barkema *Monte Carlo Methods in Statistical Physics*, (Clarendon Press, Oxford) (1999).
[89] D. P. Landau, K. Binder *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge), (2000).
[90] N. Metropolis, A. W. Rosenbluth , M. N. Rosenbluth, A. Teller, E. Teller, *Equation of State Calculations by Fast Computing Machines* J. Chem. Phys. **21**, 1087, (1953).
[91] A. K. Hartmann, B. Meerson, and P. Sasorov. *Optimal paths of non-equilibrium stochastic fields: the Kardar-Parisi-Zhang interface as a test case.* Physical Review Research, 1(3), (2019).
[92] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, D. E. Knuth *On the Lambert W function*, Advances in Computational Mathematics, **5** 329–359, (1996).
[93] K. Hukushima K. Nemoto, *Exchange Monte Carlo Method and Application to Spin Glass Simulations*, J. Phys. Soc. Jpn.**65**, 1604, (1996).
[94] A. K. Hartmann, Sequence alignments in *New Optimization Algorithms in Physics*, edited by Hartmann A. K. Rieger H. (Whiley-VCH, Weinheim) p. 253, (2004) .
[95] A. K. Hartmann, A. Krajenbrink, and P. Le Doussal. *Probing large deviations of the Kardar-Parisi-Zhang equation at short times with an importance sampling of directed polymers in random media.* Physical Review E, 101(1):012134, (2020).
[96] Alan M. Ferrenberg and Robert H. Swendsen *Optimized Monte Carlo data analysis*, Phys. Rev. Lett. **63**, 1195, (1989).
[97] P. Werner. *A Software Tool for "Gluing" Distributions* . arXiv:2207.08429 , (2022).