

Probing the large deviations of the Kardar-Parisi-Zhang equation at short time with an importance sampling of directed polymers in random media

Alexander K. Hartmann,¹ Alexandre Krajenbrink,² and Pierre Le Doussal²

¹*Institut für Physik, Universität Oldenburg, 26111 Oldenburg, Germany*

²*Laboratoire de Physique de l'École Normale Supérieure, PSL University, CNRS, Sorbonne Universités, 24 rue Lhomond, 75231 Paris, France*

(Dated: September 9, 2019)

The one-point distribution of the height for the continuum Kardar-Parisi-Zhang (KPZ) equation is determined numerically using the mapping to the directed polymer in a random potential at high temperature. Using an importance sampling approach, the distribution is obtained over a large range of values, down to a probability density as small as 10^{-1000} in the tails. The short time behavior is investigated and compared with recent analytical predictions for the large-deviation forms of the probability of rare fluctuations, showing a spectacular agreement with the analytical expressions. The flat and stationary initial conditions are studied in the full space, together with the droplet initial condition in the half-space.

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

I. INTRODUCTION

We study the continuum Kardar-Parisi-Zhang (KPZ) equation [1] in 1+1 dimensions, which describes the stochastic growth of an interface parameterized by a height field $h(x, t)$

$$\partial_t h(x, t) = \nu \partial_x^2 h(x, t) + \frac{\lambda_0}{2} (\partial_x h(x, t))^2 + \sqrt{D} \xi(x, t), \quad (1)$$

starting from an initial condition $h(x, 0)$. Here $\xi(x, t)$ is a centered Gaussian white noise with $\mathbb{E}[\xi(x, t)\xi(x', t')] = \delta(x - x')\delta(t - t')$, and we use from now on units of space, time and heights such that $\lambda_0 = D = 2$ and $\nu = 1$ [2]. We consider here the Cole-Hopf solution, such that $Z(x, t) = e^{h(x, t)}$ satisfies the stochastic heat equation (SHE),

$$\partial_t Z(x, t) = \partial_x^2 Z(x, t) + \sqrt{2}\xi(x, t)Z(x, t) \quad (2)$$

Then $Z(x, t)$ equals the partition sum of a continuum directed polymer in the $d = 1 + 1$ random potential ξ , with one endpoint at (x, t) . In some cases we will consider the KPZ equation in a half-space $x \geq 0$, with the Neumann boundary condition $\partial_x h(x, t)|_{x=0} = A$ for all time t , that is $\partial_x Z(x, t)|_{x=0} = AZ(0, t)$, where A is the boundary parameter. The choice $A = 0$ corresponds to a symmetric (reflective) wall at $x = 0$, while $A = +\infty$ corresponds to an infinitely repulsive wall (hard wall, or absorbing wall) imposing $Z(0, t) = 0$ for all t .

Exact solutions for the probability distribution function (PDF) of the height $h(0, t)$ of the KPZ equation, valid at all time t , have been obtained for several initial conditions [3–12], notably *flat*, *droplet* and *Brownian* (also called *stationary*). They showed convergence to Tracy-Widom distributions at large time for the typical fluctuations, depending on the class of initial condition. Exact solutions valid at all time have also been obtained for the half-space geometry [13–16].

Recently, the large deviations away from the typical behavior have been studied. At short time, a number of

results have been obtained for a variety of initial conditions. While the typical height fluctuations are Gaussian and of scale $\delta h \sim t^{1/4}$, it was shown that the PDF of the shifted random variable $H = H(t) = h(0, t) - \langle h(0, t) \rangle$ enjoys the large-deviation principle in the regime $H = \mathcal{O}(1)$ and $t \ll 1$

$$P(H, t) \sim \exp\left(-\frac{\Phi(H)}{\sqrt{t}}\right) \quad (3)$$

with $\Phi(0) = 0$. The rate function $\Phi(H)$ exhibits some universal properties : it behaves as $\sim H^2$ for small $|H| \ll 1$, in agreement with the Gaussian form of the typical fluctuations, and has asymmetric tails, $|H|^{5/2}$ on the negative side and $H^{3/2}$ on the positive side, with amplitudes depending on the initial condition, see Table I.

Initial condition	Left tail	Right tail
Full space		
Droplet	$\frac{4}{15\pi} H ^{5/2}$	$\frac{4}{3} H^{3/2}$
Brownian - Stationary	$\frac{4}{15\pi} H ^{5/2}$	$\frac{2}{3} H^{3/2}$
Flat	$\frac{8}{15\pi} H ^{5/2}$	$\frac{4}{3} H^{3/2}$
Half space		
Droplet with reflective wall	$\frac{2}{15\pi} H ^{5/2}$	$\frac{2}{3} H^{3/2}$
Droplet with repulsive hard wall	$\frac{2}{15\pi} H ^{5/2}$	$\frac{4}{3} H^{3/2}$
Brownian with repulsive hard wall	$\frac{2}{15\pi} H ^{5/2}$	$\frac{2}{3} H^{3/2}$

Table I: Tails of the large-deviation function $\Phi(H)$ in (3) for various initial conditions. In half-space, we use the shorthand notations: reflective wall for $A = 0$, repulsive hard wall for $A = +\infty$. These results have been obtained in Refs. [17–21] and Refs. [23–32].

There are two complementary methods to study the short-time large deviations. The first one uses the known

exact solutions to the KPZ equation and allows to obtain the rate function $\Phi(H)$ completely analytically, see Refs. [17–21]. It is thus, at this stage, restricted to the initial conditions where exact solutions are available. The second method, the weak noise theory (also called optimal fluctuation theory), see Refs. [22–32], is more versatile but leads to differential equations which, at this stage, can only be solved numerically, although exact results have been obtained concerning the tails $|H| \rightarrow +\infty$, in good agreement with the first method. The first method was later extended to obtain higher order corrections in the small time expansion in Ref. [33].

Having determined analytically the large-deviation function $\Phi(H)$, one can ask whether it is possible to match the theoretical predictions with numerical simulations [34]: this is an outstanding problem as large deviations are by definition extremely hard to probe. In the context of the KPZ equation, a previous work [35] has investigated this question for the droplet initial condition in full space, showing a very good agreement over hundreds of decades in probability.

The aim of this paper is to extend the previous numerical effort to obtain a similar quality of agreement with analytic predictions for other initial conditions of the KPZ equation (*i*) the flat and stationary, i.e. Brownian, initial conditions in full-space and (*ii*) the droplet initial condition in half space with an infinite hard wall. In addition, we aim to confirm the existence of a phase transition for the stationary initial condition. This transition, unveiled in [25, 27, 28] was further analyzed and confirmed in [18] where analytic formula were obtained for the critical height H_{c2} and the two branches for $\Phi(H)$ which coexist beyond the transition. Here we aim to confirm *both branches*, and in particular that the "analytic branch" obtained in Ref. [18] (which does not show any phase transition) corresponds indeed to the flat initial condition, as surmised in Ref. [28].

The outline is as follows. In Section II we explain the methods (i) the use of a lattice directed polymer in the high temperature limit to simulate the continuum KPZ equation (ii) the importance sampling method which allows to explore the deep tails of the large-deviation regime. In Section III we study first the flat and stationary initial conditions (since their large deviations are related) and, secondly, the half-space KPZ with an infinitely repulsive wall. We take the opportunity to summarize in each case the analytical predictions for the rate function $\Phi(H)$. Next we specify how the directed polymer mapping is implemented. Finally we show and comment the numerical results.

II. METHODS

A. Directed polymer on a lattice

Define a directed polymer on the rotated square lattice (y, τ) , see Fig. 1, which is allowed to grow according to the following rule

$$(y, \tau) \rightarrow (y \pm \frac{1}{2}, \tau + 1) \quad (4)$$

For each site of the square lattice, define a quenched random variable $V_{y,\tau}$, a temperature T and an associated Boltzmann weight $\exp(-\frac{V_{y,\tau}}{T})$. For a path $\gamma : (0, 0) \rightarrow (y_f, L)$, we define its weight by

$$w_\gamma = \prod_{(y,\tau) \in \gamma} e^{-\frac{V_{y,\tau}}{T}}. \quad (5)$$

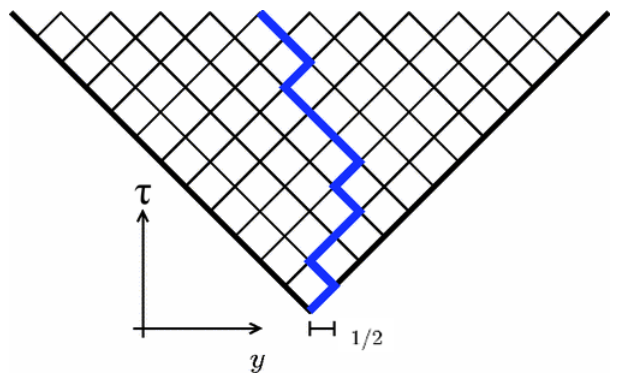


Figure 1: Representation of a directed polymer on a square lattice. The square is rotated by 45° and the path of the polymer starts from one of the corner of coordinate $(0, 0)$.

For paths of length L with a starting point $(0, 0)$ and an end-point (y_f, L) the partition sum over all possible polymers is given by the sum of the weight of all paths joining these points

$$Z_{y_f, L} = \sum_{\gamma} w_\gamma. \quad (6)$$

With these definitions, the partition sum verifies a natural recursion formula which constitutes a discretized version of the Stochastic Heat Equation [36, 37]

$$Z_{y, \tau+1} = \left(Z_{y-\frac{1}{2}, \tau} + Z_{y+\frac{1}{2}, \tau} \right) e^{-\beta V_{y, \tau+1}}. \quad (7)$$

Numerically, this problem is solved by the *transfer matrix method* and the complexity to compute the partition sum up to some discrete time τ is of order $\mathcal{O}(\tau^2)$. It is interesting to investigate the extreme regimes of zero temperature and high temperature.

At zero temperature, $T \rightarrow 0$, the free energy defined as $F_{y, \tau} = -T \log Z_{y, \tau}$ verifies the optimization/recursion equation

$$F_{y, \tau+1} = \min(F_{y-\frac{1}{2}, \tau}, F_{y+\frac{1}{2}, \tau}) + V_{y, \tau+1}. \quad (8)$$

In the high temperature regime, $T \rightarrow \infty$, the lattice polymer converges to the continuum one. The continuum polymer, whose partition function is the solution of the SHE (2), is obtained from the lattice one using the following parametrization [3, 36]

$$x = \frac{4y_f}{T^2}, \quad t = \frac{2L}{T^4}. \quad (9)$$

The proper convergence of the partition function is expressed as

$$\lim_{\substack{L \rightarrow +\infty \\ T \rightarrow +\infty \\ t \text{ fixed}}} 2^{-L} Z_{y_f, L} = Z(x, t) \quad (10)$$

If the random variables $V_{y, \tau}$ are chosen as independent unit Gaussian (which is the choice made everywhere in this paper) then $Z(x, t)$ is solution of the SHE (2). The coordinates (x, t) are associated to the continuum model. From this parameterization, we see that taking the limits of infinite temperature T and infinite lattice length L , while keeping the time t fixed and small, allows to access the short-time dynamics of the SHE, equivalently of the KPZ equation, which is our present aim.

Having defined the temperature regime, we now need to choose the geometry of the polymer problem. We consider two type of configurations.

- The point to point polymer whose starting and ending points are fixed.
- The point to line polymer whose starting point is fixed, but the ending point is arbitrary on some line. In that case the partition sum is obtained by summing over all ending points on this line, with possibly an additional weight. In the limit of large temperature T and the large lattice size L , this sum will be viewed as a Riemann sum approaching an integral.

We recall that if we want to evaluate the final KPZ height at $x = 0$, then its formal expression at time t is given by

$$e^{h(0, t)} = \int_{\mathbb{R}} dx' Z(0, t | x', 0) e^{h(x', t=0)} \quad (11)$$

where $Z(x, t | x', 0) = Z(x, t)$ denotes the solution of the SHE with initial condition $Z(x, t = 0 | x', 0) = \delta(x - x')$, which corresponds to the so-called droplet initial condition (at position x') for the KPZ equation. From the above, we thus need to encode the initial condition of the KPZ equation in the configuration of the polymer using the correspondence

$$Z(0, t | x, 0) \longleftrightarrow 2^{-L} Z_{y_f, L} \quad (12)$$

The usual initial conditions will be encoded as follows:

- The droplet initial condition is $e^{h(x, t=0)} = \delta(x)$ so $e^{h(0, t)} = Z(0, t | 0, 0)$, this is precisely the point to point polymer;

- The flat initial condition is $e^{h(x, t=0)} = 1$ so $e^{h(0, t)} = \int_{\mathbb{R}} dx' Z(0, t | x', 0)$. This integral would be obtained on the discrete lattice by summing over all points over the initial line. It is however fully equivalent (and more convenient for our study) to flip the figure upside down, and sum over all points over the final line: this is precisely the point to line polymer as represented in Fig 1, where the notion of final line is made clear (we will use the same notion for the Brownian initial condition below).
- The Brownian initial condition is $e^{h(x, t=0)} = e^{B(x)}$ so $e^{h(0, t)} = \int_{\mathbb{R}} dx' Z(0, t | x', 0) e^{B(x')}$ where $B(x')$ is the unit two-sided Brownian motion with $B(0) = 0$. This integral will be obtained on the discrete lattice by summing over all points over the final line (see discussion above) with an additional weight representing the Brownian contribution, hence this is another type of point to line polymer. The exact method and scaling to introduce the Brownian contribution will be discussed later.

The analytical predictions being for the centered one-point KPZ height, we also need to center the distribution of the height obtained in the simulations. Calling Z_τ the discrete partition function up to time τ (whatever configuration), we will compare our theoretical predictions to the distribution of the quantity $H(t)$ defined by

$$H(t) = \log \frac{Z_L}{\overline{Z_L}} \quad (13)$$

where $\overline{Z_L}$ is the average value of the partition function over many realizations¹.

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. [35]. In principle one could obtain an estimate of the probability distribution $P(H, t)$ numerically from *simple sampling*. For this, one generates many disorder realisations and calculates the partition function for each. Then $\overline{Z_L}$ is estimated by averaging over all samples, and the distribution is the histogram of the values of H according to Eq. (13). 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^{-1000} is strictly impossible. Hence, a different approach is required.

To estimate $P(H, t)$ for a much larger range, where probability densities as small as 10^{-1000} may appear, we

¹ An additional centering to impose $\Phi(0) = 0$ will be applied on the numerical data (see below). It corresponds to corrections which are subdominant at small t in the continuum model (i.e. the difference between $\log Z(0, t)$ and $\log \overline{Z(0, t)}$ [3, 38]).

use a more powerful approach, called importance sampling as discussed in Refs. [39, 40]. 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 [41–50]. The idea behind importance sampling is to sample the different disorder realisations with a suitable additional bias [51]. Here we use the bias $\exp(-\theta H(V))$ where θ is an adjustable parameter interpreted as a fictive inverse temperature. Varying the value of θ allows one to sample the distribution in different regions. In particular, if $\theta > 0$ the configurations with a negative H become more likely, conversely if $\theta < 0$ the configurations with a positive H are favored. A standard Markov-chain Monte Carlo simulation is used to sample the biased configurations [52, 53]. At each time step a new disorder realisation V^* is proposed by replacing on the current realisation V a certain fraction r of the random numbers $V_{y,\tau}$ by new random numbers. The new disorder realisation is then accepted with the usual Metropolis-Hastings probability

$$p_{\text{Met}} = \min\{1, e^{-\theta[H(V^*)-H(V)]}\}, \quad (14)$$

otherwise the old configuration is kept [54]. Note that the average partition function $\overline{Z_L}$ appearing in the definition of H (13) drops out of the Metropolis probability. By construction, the algorithm fulfils detailed balance and is ergodic, since within a sufficient number of steps, each possible realisation may be constructed. Thus, in the limit of infinitely long Markov chains, the distribution of biased disorder realisations will follow the probability

$$q_\theta(V) = \frac{1}{Q(\theta)} P_{\text{dis}}(V) e^{-\theta H(V)}, \quad (15)$$

where $P_{\text{dis}}(V)$ is the original disorder distribution and $Q(\theta) = \sum_V P_{\text{dis}}(V) e^{-\theta H(V)}$ is the normalization factor. Note that $Q(\theta)$ also depends on L and T because of finite size and temperature effects. $Q(\theta)$ is generally unknown but can be determined from the numerical results, see Appendix B. The output of this Markov chain allows to construct a biased histogram $P_\theta(H, t)$. In order to get the correct empirical probability density $P(H, t)$ one should debias the result so that

$$P(H, t) = e^{\theta H} Q(\theta) P_\theta(H, t). \quad (16)$$

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

III. COMPARISON OF THE THEORETICAL PREDICTIONS WITH THE SIMULATIONS

We now compare the theoretical predictions for the probability distribution of the solutions to the Kardar-

Parisi-Zhang equation with the numerical simulations of the directed polymer on a lattice for each initial condition. We start with the stationary (i.e. Brownian) and flat initial conditions, since their analytical expressions are related. Then we address the half-space geometry. We proceed as follows. In each case we first recall and give explicitly the analytical predictions for the large-deviation rate function $\Phi(H)$. Then we provide details on how to encode the random weights on the lattice model to account for the particular initial conditions. Finally we present the numerical data and compare with the predictions. The large-deviation function $\Phi(H)$ has been constructed so that it is centered around 0 and hence for a fair comparison, the probability densities obtained in the simulations will be shifted so that their maximum is also reached at 0. Finally, we insist on the fact that the comparison will be done without any fitting parameter.

A. The flat and stationary initial conditions in full-space

1. Recall of the analytical result for $\Phi(H)$

We first recall the results of Ref. [18] (see also Ref. [19]) for the two branches of $\Phi(H)$. As discussed below, they provide the short-time probability distribution for the flat and the stationary (i.e. Brownian) initial condition in full-space. Let us first define the critical height $H_{c2} = 2 \log(2e - \mathcal{I}) - 1 \simeq 1.85316$ with

$$\mathcal{I} = \int_0^{+\infty} \frac{dy}{\pi} \left[1 + \frac{1}{y} \right] \frac{\sqrt{y}}{e^{-1} + ye^y}. \quad (17)$$

together with the function $\Psi(z)$ expressed as

$$\Psi(z) = - \int_{\mathbb{R}} \frac{dk}{2\pi} \text{Li}_2\left(-\frac{ze^{-k^2}}{k^2}\right), \quad (18)$$

and its reduced version $\psi(z) = \Psi(z) - 2z\Psi'(z)$. We further denote the intervals:

$$\begin{aligned} I_1 &= [0, +\infty[, \quad I_2 = [0, e^{-1}], \quad I_3 =]0, e^{-1}], \\ J_1 &=]-\infty, 0], \quad J_2 = [0, H_{c2}], \quad J_3 = [H_{c2}, +\infty[. \end{aligned} \quad (19)$$

We define a large-deviation function $\Phi(H)$ given in a parametric form by

$$e^H = \begin{cases} z\Psi'(z)^2, & z \in I_1, H \in J_1 \\ z \left[\Psi'(z) - 2z^{-1}[-W_0(-z)]^{\frac{1}{2}} \right]^2, & z \in I_2, H \in J_2 \end{cases}$$

$$\Phi(H) = \begin{cases} \psi(z), & z \in I_1 \\ \psi(z) + \frac{4}{3}[-W_0(-z)]^{\frac{3}{2}}, & z \in I_2 \end{cases} \quad (20)$$

For $z \in I_3, H \in J_3$, we define two continuations of $\Phi(H)$: a symmetric analytic one (resp. asymmetric non-analytic) denoted $\Phi_{\mathbf{A}}$ (resp. $\Phi_{\mathbf{NA}}$) such that their para-

metric representations read

$$\begin{cases} e^H = z \left[\Psi'(z) - 2z^{-1}[-W_{-1}(-z)]^{\frac{1}{2}} \right]^2 \\ \Phi_{\mathbf{A}}(H) = \psi(z) + \frac{4}{3}[-W_{-1}(-z)]^{\frac{3}{2}} \end{cases}$$

and

$$\begin{cases} e^H = z \left[\Psi'(z) - z^{-1} \left([-W_{-1}(-z)]^{\frac{1}{2}} + [-W_0(-z)]^{\frac{1}{2}} \right) \right]^2 \\ \Phi_{\mathbf{NA}}(H) = \psi(z) + \frac{2}{3}[-W_{-1}(-z)]^{\frac{3}{2}} + \frac{2}{3}[-W_0(-z)]^{\frac{3}{2}} \end{cases} \quad (21)$$

The large-deviation solutions for the flat and the Brownian initial conditions are by far the most difficult to study, as two successive continuations are required (while only one is required for the droplet initial condition [17]). Indeed, the large-deviation functions for the stationary and flat initial conditions in full-space are given for all H in \mathbb{R} by

$$\Phi_{\text{Brownian}}(H) = \Phi_{\mathbf{NA}}(H), \quad \Phi_{\text{flat}}(H) = 2^{-\frac{3}{2}} \Phi_{\mathbf{A}}(2H). \quad (22)$$

The indices \mathbf{A} and \mathbf{NA} are explicitly indicated for the region $H \in [H_{c2}, +\infty[$ where the two continuations are distinct and can be omitted for the region where $H \in]-\infty, H_{c2}]$ where the two functions are the same, i.e. $\Phi_{\mathbf{NA}}(H) = \Phi_{\mathbf{A}}(H)$ for $H \in]-\infty, H_{c2}]$. We further recall that W_0 and W_{-1} are the two real branches of the Lambert function, see Appendix A.

The result (22) for the Brownian initial condition which follows the \mathbf{NA} branch was obtained in Ref. [18]. The result (22) for the flat initial condition arises from (i) the identification made in the context of the WNT in Ref. [28] of the analytic branch and the large-deviation function of the flat initial condition (the rationale being that no phase transition is expected for the flat initial condition) and (ii) the analytic result for the \mathbf{A} branch in Eq. (21) obtained in Ref. [18]. The verification of (22) will be an important point to seek in numerical simulations. Before we introduce the results of the simulations obtained through the importance sampling method, we will explain in some details the lattice construction of these initial conditions.

2. Point to line directed polymer mapping

We first introduce the mapping of the flat initial condition of the KPZ equation to the directed polymer model on the lattice and then extend the discussion to the Brownian initial condition. As mentioned earlier, contrary to the droplet initial condition, the flat initial condition requires to perform a summation of the partition sum over the final line of the lattice, *i.e.*

$$Z_L = \sum_{k=-L}^L Z_{y_k, L}. \quad (23)$$

where the points $\{y_k\}$ are all the points over the final line of the lattice ordered from left to right. Note that

this amounts to set a uniform weight over the final line. Comparatively, for the droplet initial condition, all the weight was concentrated at $k = 0$ (the above sum with a Kronecker delta at $k = 0$ would yield the droplet initial condition).

Recalling the continuum parametrization of Eq. (9), the infinite temperature limit (up to a proportionality constant)

$$Z_{y_k, L} \longleftrightarrow Z(x = \frac{4y_k}{T^2}, t = \frac{2L}{T^4}), \quad (24)$$

by a Riemann summation using a step of size $\Delta x = \frac{4}{T^2} \ll 1$ we obtain the partition function of the flat condition (up to a proportionality constant)

$$\frac{4}{T^2} \sum_{k=-L}^L Z_{y_k, L} \xrightarrow{T \gg 1} \int_{-4L/T^2}^{4L/T^2} dx Z(x, t). \quad (25)$$

We emphasize that because of our construction of the height (13), the proportionality constants are irrelevant. As we renormalize the partition sum by its average according to Eq. (13), the prefactor $4/T^2$ can be discarded and will play strictly no role in the numerics. To approximate the integral as being over the full real line, the factor $4L/T^2$ should be taken as large as possible: its finiteness might induce additional finite size and discretization effects as compared to the point to point continuum polymer. To provide orders of magnitude, some of the simulations will have a time $t = 1/16$ and the factor $4L/T^2$ is then in the range [4, 8] for the different lattice sizes used.

The extension of this mapping to the stationary initial condition is obtained by adding Brownian weights $e^{B(x)}$ on the final line on the lattice

$$Z_{y_k, L} e^{B(\frac{4y_k}{T^2})} \rightarrow Z(x = \frac{4y_k}{T^2}, t) e^{B(x)}. \quad (26)$$

By self-similarity, we have $B(\frac{4y_k}{T^2}) \equiv \frac{2}{T} B(y_k)$ and hence $B(y_k)$ can be easily sampled by a random walk with Gaussian increments with unit variance. Note that the final partition function assumes an additional average over the Brownian motion compared to a deterministic initial condition, hence this provides an extra challenge on the numerical side.

3. Presentation of the simulations

The numerical simulations for the flat and Brownian initial conditions in full-space were run for polymers of length $L = \{32, 64, 128\}$ and temperature T chosen so that the corresponding time for the Kardar-Parisi-Zhang equation is fixed at $t = 1/16$. Convergence to the analytic predictions is expected for $L \rightarrow +\infty$. We present the simulations in Figs. 2 and 3. We observe for the flat and Brownian initial conditions that the agreement between the numerics and the theory is fairly good and improves as L increases. We should note that there are additional sources of finite size effects and statistical errors since for

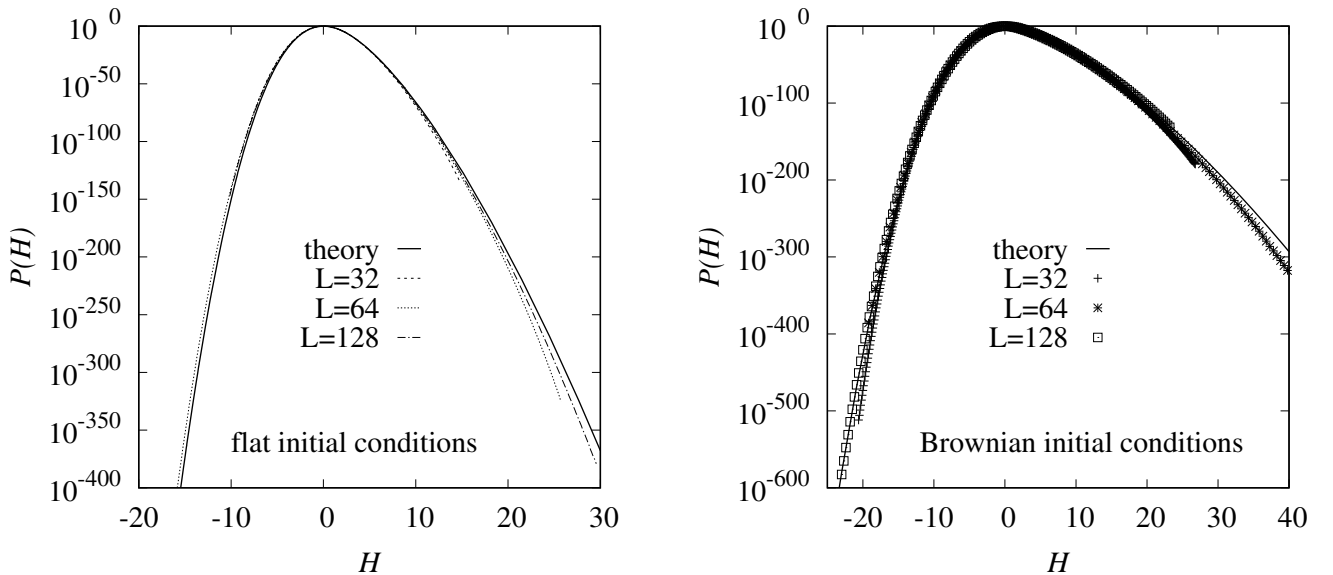


Figure 2: Probability distribution $P(H, t)$ for a short time $t = 1/16$ for three different polymer lengths $L = \{32, 64, 128\}$ and **left:** flat initial conditions, **right:** Brownian (stationary) initial conditions. In both plots the solid lines indicate the analytical prediction displayed in Eq. (22), together with Eqs. (20) and (21).

the flat and the Brownian initial conditions we need to perform a summation of the partition function over the final line in both cases and an additional average over the Brownian weights in the second case. Nonetheless, the simulations are able to probe both tails of the distribution of $P(H, t)$ for quite a range of values for both initial conditions as can be seen in Fig. 3. A very good agreement of analytical and numerical results is visible here and on this scale only very small finite-size corrections are visible. The numerical results even extend to the regime below $P(H) \sim e^{-1000}$ where the leading tails behavior starts to be dominant.

4. Discussion around the choice of branch Φ_A/Φ_{AN} for flat and Brownian initial conditions

We now turn to the crucial discussion of the choice of branch Φ_A/Φ_{AN} for both flat and Brownian initial conditions. We recall that considerations coming from Weak Noise Theory [28] together with the precise study of the Fredholm determinant associated to the solution for the Brownian initial condition [18], concluded that for all H in \mathbb{R} the prediction (22) holds, where the two branches coincide $\Phi_{NA}(H) = \Phi_A(H)$ for $H < H_{c2}$, but differ for $H > H_{c2}$. This implies that if we rescale the simulation data for the flat initial condition, they should coincide pointwise up to the critical height H_{c2} and then separate to the symmetric and asymmetric branches. We emphasize that the choice of branch is critical as it determines the tails of the distribution and the existence of a singularity in the large-deviation function.

We present in Fig. 4 the comparison in a region

around H_{c2} between the exact expression of the different branches of $\Phi(H)$ given in Eqs. (20) and (21) and the directed polymer simulations of length $L = 128$ at a time $t = 1/16$ for the Brownian and the flat initial conditions. Note that we rescaled the distribution for the flat initial condition according to the scaling predicted in Eq. (22) for a fair comparison. Figure 4 confirms in a remarkable way that each initial condition corresponds to the one of the two branches that we have obtained analytically. The coincidence of the two distributions before the point $H = H_{c2}$ is also quite spectacular.

B. The droplet initial condition in half-space with a hard-wall $A = +\infty$

We now consider the KPZ equation in the half-space $x \geq 0$ with the boundary condition $\partial_x h(x, 0)|_{x=0} = A$, equivalently $\partial_x Z(x, t)|_{x=0} = AZ(0, t)$ for the continuum polymer, with droplet initial conditions i.e. $Z(x, t = 0) = \delta(x - \epsilon)$ with $\epsilon \rightarrow 0^+$. The large-deviation function $\Phi(H)$ was obtained in [21] for several values of A (respectively $-1/2, 0$ and $+\infty$). Here we will test only the analytical prediction of Ref. [21] for $A = +\infty$, which corresponds to the infinitely repulsive hard wall (i.e. $Z(0, t) = 0$ for all t).

1. Recall of the analytical result for $\Phi(H)$

To provide the short-time probability distribution for the droplet initial condition in a half-space in presence of a hard wall, let us first define the branching height

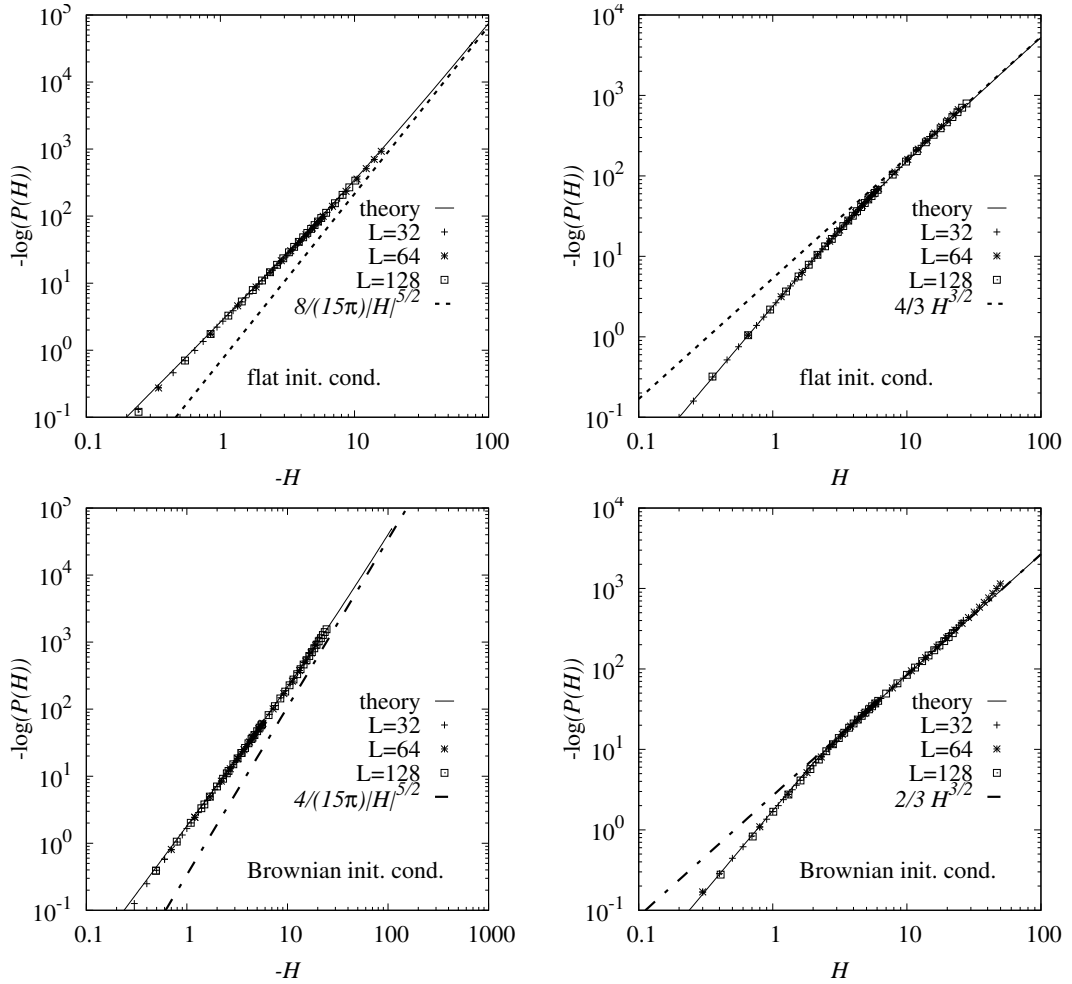


Figure 3: **Top**: blow up of the left and right tails of the data shown in the left plot of Fig. 2 for the flat initial condition. **Bottom**: blow up of the left and right tails of the data shown in the right plot of Fig. 2 for the Brownian initial condition. In all four plots the solid line is the analytical prediction given in Eq. (22), together with Eqs. (20) and (21). The leading tails at large $|H|$ given in the Table I are also plotted in each case.

$H_c = \log(8\sqrt{\pi}\mathcal{I}) \simeq 0.9795$ with

$$\mathcal{I} = \int_0^{+\infty} \frac{dy}{\pi} \left[1 - \frac{1}{y} \right] \frac{\sqrt{y}}{e^y/y - e}. \quad (27)$$

We also define the functions $\Psi(z)$, $\Delta(z)$ and its derivative $\Delta'(z)$ as

$$\begin{aligned} \Psi(z) &= - \int_{\mathbb{R}} \frac{dk}{4\pi} \text{Li}_2(-zk^2 e^{-k^2}) \\ \Delta(z) &= \frac{2}{3} \left[W_0\left(\frac{1}{z}\right) \right]^{\frac{3}{2}} + \frac{2}{3} \left[W_{-1}\left(\frac{1}{z}\right) \right]^{\frac{3}{2}} \\ &\quad + 2 \left[W_0\left(\frac{1}{z}\right) \right]^{\frac{1}{2}} + 2 \left[W_{-1}\left(\frac{1}{z}\right) \right]^{\frac{1}{2}} \\ \Delta'(z) &= -\frac{1}{z} \left[W_0\left(\frac{1}{z}\right) \right]^{\frac{1}{2}} - \frac{1}{z} \left[W_{-1}\left(\frac{1}{z}\right) \right]^{\frac{1}{2}} \end{aligned} \quad (28)$$

and we define their reduced versions ψ and δ

$$\psi(z) = \Psi(z) - z\Psi'(z), \quad \delta(z) = \Delta(z) - z\Delta'(z). \quad (29)$$

We further denote the intervals:

$$\begin{aligned} I_1 &= [-e, +\infty[, & I_2 &= [-e, 0[, \\ J_1 &=]-\infty, H_c], & J_2 &= [H_c, +\infty[. \end{aligned} \quad (30)$$

The associated large-deviation function $\Phi(H)$ is expressed by the parametric system

$$\begin{aligned} e^H &= \begin{cases} 8\sqrt{\pi}\Psi'(z), & z \in I_1, H \in J_1 \\ 8\sqrt{\pi}[\Psi'(z) + \Delta'(z)], & z \in I_2, H \in J_2 \end{cases} \\ \Phi(H) &= \begin{cases} \psi(z), & z \in I_1 \\ \psi(z) + \delta(z), & z \in I_2 \end{cases} \end{aligned} \quad (31)$$

Before we introduce the results of the simulations obtained through the importance sampling method, we will explain in some details the lattice construction of the

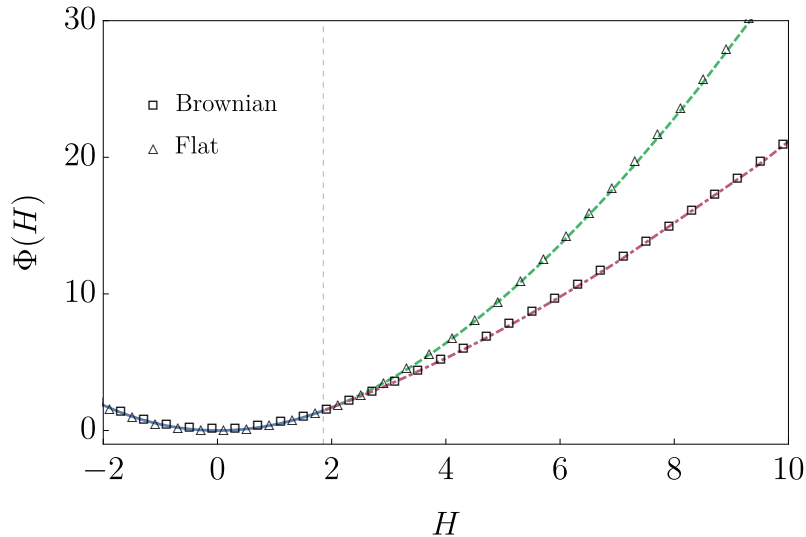


Figure 4: (color online) The analytic expression of $\Phi(H)$ obtained in Eqs. (20) and (21) is compared with the simulation data obtained for the flat and for Brownian initial conditions. In the case of the flat initial condition, the data are rescaled according to Eq. (22) for a fair comparison. The blue line corresponds to the function $\Phi(H)$ before the critical point $H_{c2} \simeq 1.85316$ which position is indicated by a vertical grey dotted line. The green dashed line corresponds to the symmetric analytic branch Φ_A and the red dot-dashed line corresponds to the asymmetric non-analytic branch Φ_{NA} . The square markers represent the Brownian initial condition simulation data and the triangle markers represent the rescaled flat initial condition data.

half-space problem in the presence of a wall.

2. Point to point directed polymer mapping in a half-space with a hard-wall $A = +\infty$

To mimic a hard-wall type problem, one can forbid the polymer to visit some edges and introduce a tabu list for edges bringing the polymer towards the diagonal of the lattice. Indeed, the partition function should be strictly zero on the diagonal so that the probability of crossing the diagonal is strictly zero. We introduce this idea of forbidden edges in Fig. 5. Finally, let us mention that an extension of this construction to a generic value of A would be interesting.

3. Presentation of the simulations

The numerical simulations for the droplet initial condition in half-space in the presence of a hard wall were run for polymers of length $L = \{64, 128, 256\}$ and temperature T chosen so that the corresponding time for the Kardar-Parisi-Zhang equation is fixed at $t = 1/16$. Convergence to the analytic predictions is expected for $L \rightarrow +\infty$. We present the simulations in Figs. 7 and 6 and we observe that the agreement between numerical and analytic results is remarkable in the tails (for $H \leq -25$ and all $H \geq 0$), but nonetheless there is an intermediate region for negative H ($-20 \leq H \leq -10$) where the matching is not entirely perfect: this could

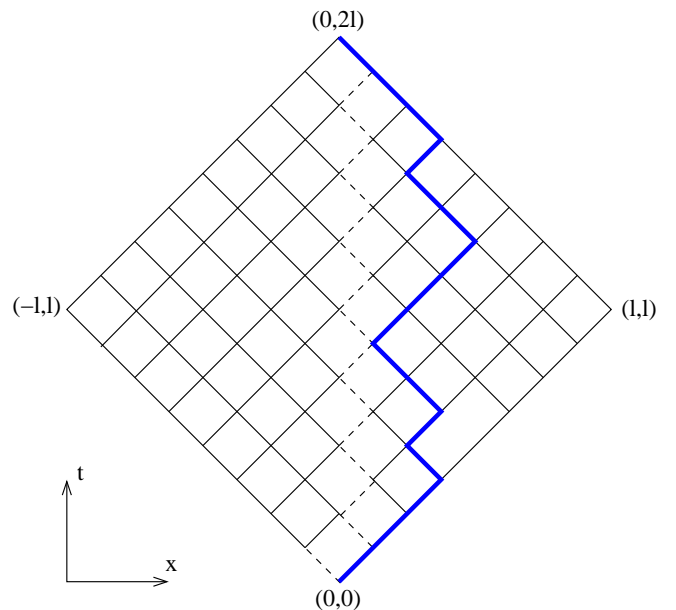


Figure 5: Square lattice designed for the half-space problem with a hard-wall. The dashed lines are forbidden edges for the polymer constraining it to stay on the right of the lattice. An example of a polymer realization is drawn on the blue line. In this representation, we have $L = 2\ell$.

possibly come from finite size effects but it still need to be further investigated.

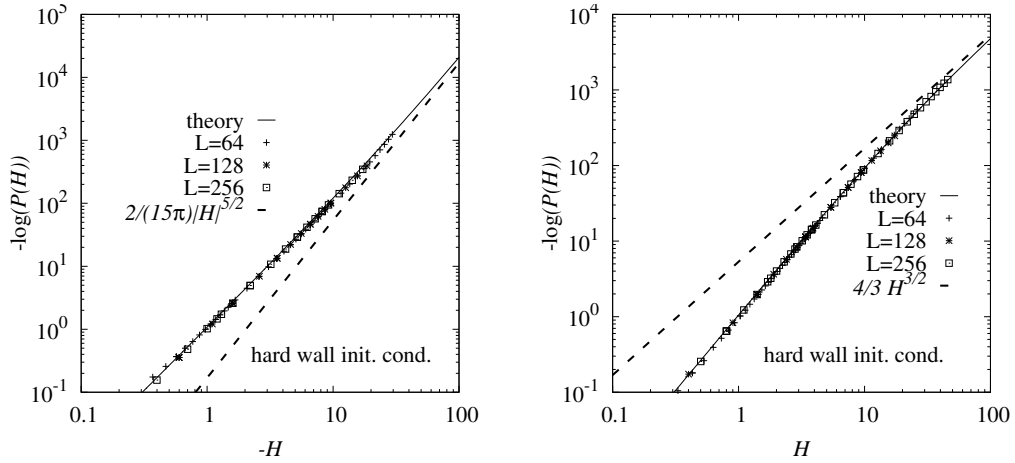


Figure 6: Blow up of the left and right tails of the data shown of Fig. 7. The solid lines indicate the analytical prediction displayed in Eq. (31) obtained in [21]. The data is also compared with the leading behavior of each tail as displayed in the Table I.

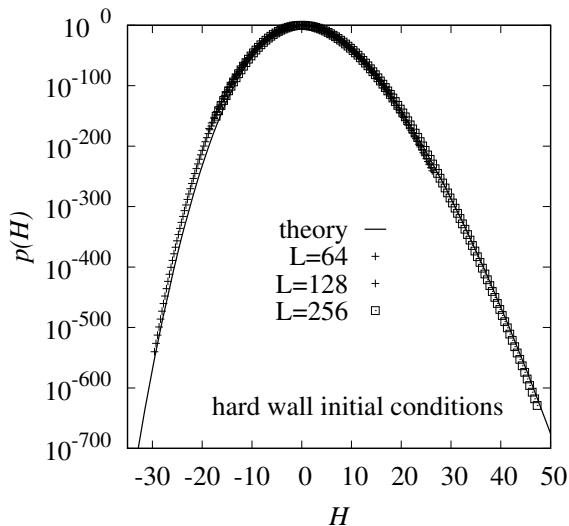


Figure 7: Probability distribution $P(H, t)$ for a short time $t = 1/16$ for three different polymer lengths $L = \{64, 128, 256\}$ for the droplet initial condition in a half-space with an infinite hard-wall at the origin. The solid line indicates the analytical result displayed in Eq. (31) obtained in [21].

IV. CONCLUSION AND OUTLOOK

To summarize, a large-deviation sampling approach has been used to measure the probability distribution $P(H, t)$ of the centered height H solution the KPZ equation, for various initial conditions: flat and stationary in full-space and droplet in half-space in presence of an infinitely repulsive wall. This was achieved using a lattice directed polymer model, whose free energy converges, in the high-temperature limit, to the height of the continuum KPZ equation. This allowed to determine numer-

ically the probability distribution of the height over a large range of values, leading to a precise comparison with the analytical predictions. We find that the agreement with the short-time large-deviation function $\Phi(H)$ predicted by the theory, which differs according to the geometry and the initial condition, is spectacular, even very far in the tails. The existence and location of a phase transition for the stationary initial condition are fully confirmed, as well as the close connection of the rate function for the flat and stationary initial conditions.

It would be interesting to extend this study in several directions. First, multi-point distributions of the height field could be investigated. Recently the optimal profile at time $t/2$, i.e. $h(x, t/2)$, conditioned to a large value of $|H|$ (at time t), predicted by the weak noise theory was successfully tested in a similar numerical simulation [55]. Although this provides some information, one could ask more detailed questions about the configurations in real space of the polymer corresponding to the tails. One expects that the polymer configurations contributing to the left tail of the distribution of H will be quite different from the ones which contribute to the right tail. It would be interesting to determine the roughness of these atypical polymers, both at short time and at large time. The knowledge of the shape of such polymers could be of great benefits to get some insight on how to sample from atypical polymer distributions.

Acknowledgments

We thank A. Rosso, G. Schehr and S. N. Majumdar for very helpful discussions. AK and PLD acknowledge support from ANR grant ANR-17-CE30-0027-01 RaMaTraF.

- [1] M. Kardar, G. Parisi and Y-C. Zhang, *Dynamic Scaling of Growing Interfaces*, *Phys. Rev. Lett.* **56**, 889, (1986).
- [2] This is equivalent to use everywhere the following units of space, time and heights
 $x^* = (2\nu)^3/(D\lambda_0^2)$, $t^* = 2(2\nu)^5/(D^2\lambda_0^4)$, $h^* = \frac{2\nu}{\lambda_0}$
- [3] P. Calabrese, P. Le Doussal, A. Rosso, *Free-energy distribution of the directed polymer at high temperature*, *Europhys. Lett.* **90**, 20002, (2010).
- [4] V. Dotsenko, *Bethe ansatz derivation of the Tracy-Widom distribution for one-dimensional directed polymers* *EPL* **90**, 20003 (2010); *Replica Bethe ansatz derivation of the Tracy-Widom distribution of the free energy fluctuations in one-dimensional directed polymers*, *J. Stat. Mech.* P07010, (2010); V. Dotsenko and B. Klumov, *Bethe ansatz solution for one-dimensional directed polymers in random media* *J. Stat. Mech.* P03022, (2010).
- [5] T. Sasamoto, H. Spohn, *One-dimensional Kardar-Parisi-Zhang equation: an exact solution and its universality*, *Phys. Rev. Lett.* **104**, 230602, (2010).
- [6] G. Amir, I. Corwin, J. Quastel, *Probability distribution of the free energy of the continuum directed random polymer in $1 + 1$ dimensions*, *Comm. Pure and Appl. Math.* **64**, 466, (2011).
- [7] P. Calabrese, M. Kormos and P. Le Doussal, *From the sine-Gordon field theory to the Kardar-Parisi-Zhang growth equation*, [arXiv:1405.2582](https://arxiv.org/abs/1405.2582), *EPL* **107** 10011, (2014).
- [8] J. Ortmann, J. Quastel, D. Remenik, *Exact formulas for random growth with half-flat initial data*, [arXiv:1407.8484](https://arxiv.org/abs/1407.8484), *Annals of Applied Probability*, Vol. **26**, No. **1**, 507-548, (2016).
- [9] P. Calabrese, P. Le Doussal, *Exact solution for the Kardar-Parisi-Zhang equation with flat initial conditions*, *Phys. Rev. Lett.* **106**, 250603, (2011); *The KPZ equation with flat initial condition and the directed polymer with one free end*, *J. Stat. Mech.* P06001, (2012).
- [10] T. Imamura, T. Sasamoto, *Exact solution for the stationary Kardar-Parisi-Zhang equation*, *Phys. Rev. Lett.* **108**, 190603, (2012).
- [11] T. Imamura, T. Sasamoto, *Stationary correlations for the 1D KPZ equation*, *J. Stat. Phys.* **150**, 908-939, (2013).
- [12] A. Borodin, I. Corwin, P. L. Ferrari, B. Veto, *Height fluctuations for the stationary KPZ equation*, *Math. Phys. Anal. Geom.* **18**, 20, (2015).
- [13] A. Krajenbrink, P. Le Doussal, *Replica Bethe Ansatz solution to the Kardar-Parisi-Zhang equation on the half-line*. [arXiv:1905.05718](https://arxiv.org/abs/1905.05718), (2019).
- [14] T. Gueudré, P. Le Doussal, *Directed polymer near a hard wall and KPZ equation in the half-space*, *Europhys. Lett.* **100**, 26006, (2012).
- [15] A. Borodin, A. Bufetov, I. Corwin, *Directed random polymers via nested contour integrals*. *Annals of Physics*, **368** 191-247, (2016).
- [16] G. Barraquand, A. Borodin, I. Corwin, M. Wheeler, *Stochastic six-vertex model in a half-quadrant and half-line open ASEP*, [arXiv:1704.04309](https://arxiv.org/abs/1704.04309), (2017).
- [17] 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).
- [18] 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).
- [19] A. Krajenbrink, *Beyond the typical fluctuations: a journey to the large deviations in the Kardar-Parisi-Zhang growth model*. In preparation, (2019).
- [20] 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).
- [21] A. Krajenbrink, P. Le. Doussal, *Large fluctuations of the KPZ equation in a half-space*. *SciPost Phys.* **5**, 032, (2018).
- [22] 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).
- [23] B. Meerson, E. Katzav, A. Vilenkin, *Large Deviations of Surface Height in the Kardar-Parisi-Zhang Equation*, *Phys. Rev. Lett.* **116**, 070601, (2016).
- [24] A. Kamenev, B. Meerson, P. V. Sasorov, *Short-time height distribution in 1D KPZ equation: starting from a parabola*, *Phys. Rev. E* **94**, 032108, (2016).
- [25] M. Janas, A. Kamenev, B. Meerson. *Dynamical phase transition in large-deviation statistics of the Kardar-Parisi-Zhang equation*, *Phys. Rev. E* **94**, 032133, (2016).
- [26] B. Meerson, J. Schmidt, *Height distribution tails in the Kardar-Parisi-Zhang equation with Brownian initial conditions*, *J. Stat. Mech.* 103207, (2017).
- [27] N. R. Smith, A. Kamenev, B. Meerson, *Landau theory of the short-time dynamical phase transition of the Kardar-Parisi-Zhang interface*, [arXiv:1802.07497](https://arxiv.org/abs/1802.07497), (2018).
- [28] N. R. Smith and B. Meerson, *Exact short-time height distribution for the flat Kardar-Parisi-Zhang interface*, [arXiv:1803.04863](https://arxiv.org/abs/1803.04863), (2018).
- [29] T. Asida, E. Livne, and B. Meerson. *Large fluctuations of a Kardar-Parisi-Zhang interface on a half-line: the height statistics at a shifted point*. *Phys. Rev. E* **99**, 042132, [arXiv:1901.07608](https://arxiv.org/abs/1901.07608), (2019).
- [30] B. Meerson and A. Vilenkin. *Large fluctuations of a Kardar-Parisi-Zhang interface on a half line*. *Physical Review E*, **98**(3):032145, (2018).
- [31] N. R. Smith, B. Meerson, and P. Sasorov. *Finite-size effects in the short-time height distribution of the Kardar-Parisi-Zhang equation*. *Journal of Statistical Mechanics: Theory and Experiment*, **2018**(2):023202, (2018).
- [32] N. R. Smith, B. Meerson, and A. Vilenkin. *Time-averaged height distribution of the Kardar-Parisi-Zhang interface*. [arXiv:1902.08110](https://arxiv.org/abs/1902.08110), (2019).
- [33] 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).
- [34] A. K. Hartmann, *Big Practical Guide to Computer Simulations*, (World Scientific, Singapore), (2015)
- [35] A. K. Hartmann, P. Le Doussal, S. N. Majumdar, A. Rosso, G. Schehr, *High-precision simulation of the height distribution for the KPZ equation*, *Europhys. Lett.* **121**,

- 67004 (2018)
- [36] E. Brunet *Influence des effets de taille finie sur la propagation d'un front - Distribution de l'énergie libre d'un polymère dirigé en milieu aléatoire*, Available on E. Brunet's webpage, (2000).
- [37] I. Corwin, *The Kardar-Parisi-Zhang equation and universality class*, [Random Matrices: Theory Appl.](#) **01 1130001**, (2012).
- [38] T. Gueudré, P. Le Doussal, A. Rosso, A. Henry, P. Calabrese, *Short time growth of a KPZ interface with flat initial conditions*, [Phys. Rev. E](#) **86**, 041151, (2012).
- [39] A. K. Hartmann, *Sampling rare events: Statistics of local sequence alignments* [Phys. Rev. E](#) **65** 056102, (2001).
- [40] A. K. Hartmann, *Large-deviation properties of largest component for random graphs* [Eur. Phys. J. B](#) **84**, 627, (2011).
- [41] A. Engel A., R. Monasson, A. K. Hartmann, *On Large Deviation Properties of Erdos-Rényi Random Graphs* [J. Stat. Phys.](#) **117**, 387, (2004).
- [42] A. K. Hartmann, *Calculation of partition functions by measuring component distributions*, [Phys. Rev. Lett.](#) **94**, 050601, (2005).
- [43] 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).
- [44] 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).
- [45] T. A. Driscoll, K. L. Maki *Searching for Rare Growth Factors Using Multicanonical Monte Carlo Methods* [SIAM Review](#) **49**, 673, (2007).
- [46] N. Saito, Y. Iba, K. Hukushima *Multicanonical sampling of rare events in random matrices* [Phys. Rev. E](#) **82**, 031142, (2010).
- [47] 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).
- [48] A. K. Hartmann, *High-precision work distributions for extreme nonequilibrium processes in large systems*, [Phys. Rev. E](#) **89**, 052103, (2014).
- [49] G. Claussen, A. K. Hartmann, S. N. Majumdar, *Convex hulls of random walks: Large-deviation properties*, [Phys. Rev. E](#) **91**, 052104, (2015).
- [50] 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).
- [51] J. M. Hammersley K. W. Morton, *A new Monte Carlo technique: antithetic variates*, [Math. Proc. Cambr. Phil. Soc.](#)**52**, 449–475, (1956).
- [52] M. E. J. Newman , G. T. Barkema *Monte Carlo Methods in Statistical Physics*, (Clarendon Press, Oxford) (1999).
- [53] D. P. Landau, K. Binder *A Guide to Monte Carlo Simulations in Statistical Physics* ([Cambridge University Press, Cambridge](#)), (2000).
- [54] 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).
- [55] A. K. Hartmann, B. Meerson, and P. Sasorov. *Optimal paths of non-equilibrium stochastic fields: the Kardar-Parisi-Zhang interface as a test case*. [arXiv:1907.05677](#), (2019).
- [56] 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).
- [57] K. Hukushima K. Nemoto, *Exchange Monte Carlo Method and Application to Spin Glass Simulations*, [J. Phys. Soc. Jpn.](#)**65**, 1604, (1996).
- [58] A. K. Hartmann, Sequence alignments in *New Optimization Algorithms in Physics*, edited by Hartmann A. K. Rieger H. (Wiley-VCH, Weinheim) p. 253, (2004) .

SUPPLEMENTARY MATERIAL

We give some additional technical details of the results of this manuscript.

Appendix A: The Lambert function W

We introduce the Lambert W function [56] which we use extensively throughout the Letter. Consider the function defined on \mathbb{C} by $f(z) = ze^z$, the W function is composed of all inverse branches of f so that $W(ze^z) = z$. It does have two real branches, W_0 and W_{-1} defined respectively on $[-e^{-1}, +\infty[$ and $[-e^{-1}, 0[$. On their respective domains, W_0 is strictly increasing and W_{-1} is strictly decreasing. By differentiation of $W(z)e^{W(z)} = z$, one obtains a differential equation valid for all branches of $W(z)$

$$\frac{dW}{dz}(z) = \frac{W(z)}{z(1+W(z))} \quad (\text{A1})$$

Concerning their asymptotics, W_0 behaves logarithmically for large argument $W_0(z) \simeq_{z \rightarrow +\infty} \ln(z) - \ln \ln(z)$ and is linear for small argument $W_0(z) \simeq_{z \rightarrow 0} z - z^2 + \mathcal{O}(z^3)$. W_{-1} behaves logarithmically for small argument $W_{-1}(z) \simeq_{z \rightarrow 0^-} \ln(-z) - \ln(-\ln(-z))$. Both branches join smoothly at the point $z = -e^{-1}$ and have the value $W(-e^{-1}) = -1$. These remarks are summarized on Fig. 8. More details on the other branches, W_k for integer k , can be found in [56].

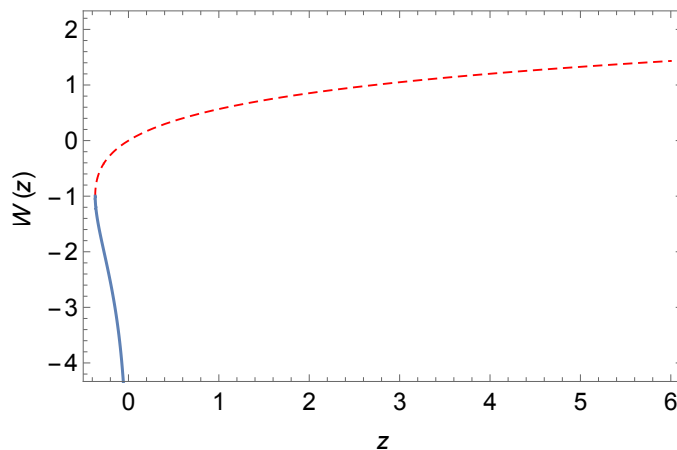


Figure 8: The Lambert function W . The dashed red line corresponds to the branch W_0 whereas the blue line corresponds to the branch W_{-1} .

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 $\{\theta_{-N_n}, \theta_{-N_n+1}, \dots, \theta_{N_p-1}, \theta_{N_p}\}$, 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 $Q(\theta)$ are obtained by first computing the histogram using simple sampling, i.e., the bulk of the distribution, which is well normalised and 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. For a perfect matching, with a suitable chosen partition function $Q(\theta_{+1})$, the two distributions would agree in the overlapping region, i.e., $e^{\theta H} Q(\theta) P_\theta(H, t) = e^{\theta_{+1} H} Q(\theta_{+1}) P_{\theta_{+1}}(H, t) \forall H$ overlapping (with $e^{\theta H} Q(\theta) = 1$ for $\theta = 0$). Due to statistical fluctuations, this is never exactly fulfilled, thus we determine $Q(\theta_{+1})$ such that the squared difference between the two rescaled distribution in the overlapping regime is minimized. In the same way 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 normalisation constants can be obtained. Note that in the end the resulting distribution is slightly non-normalized, because the tails were added to the bulk of the distribution, but this is easily repaired by a final global normalization.

The main drawback of our method is that as for any Markov-chain Monte Carlo simulation, it has to be equilibrated and this may take a large number of steps. To speed the simulation up, *parallel tempering* was used [57]. Here, a parallel implementation using the *Message Passing Interface (MPI)* was applied, such that each computing core was responsible in parallel for an independent realisation $V_i(s)$ at a given θ_i . After 1000 Monte Carlo steps, one parallel-tempering sweep was performed and the parameters θ_i and θ_{i+1} were exchanged between two computing cores. The number and values of the inverse temperatures are determined by some preliminary test simulations with criterion that the empirical acceptance rate of the parallel-tempering exchange steps is about 0.5 for all pairs of $\theta_i, \theta_i + 1$ of neighboring inverse temperatures. This resulted in few hundred different values of the inverse temperature where the systems were simulated in parallel. For each case considered here a corresponding cluster was used for few weeks. A pedagogical explanation and examples of this sampling procedure can be found in Ref. [58].
