PAPER

One-dimensional Monte Carlo dynamics at zero temperature

IOP Publishing Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **54** (2021) 485001 (20pp) https://doi.org/10.1088/1751-8121/ac2dc2

One-dimensional Monte Carlo dynamics at zero temperature

Alexei D Chepelianskii[1,*], Satya N Majumdar[2], Hendrik Schawe[3] and Emmanuel Trizac[2]

[1] Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405, Orsay, France
[2] Université Paris-Saclay, CNRS, LPTMS, 91405, Orsay, France
[3] LPTM, UMR 8089, CY Cergy Paris Université, CNRS, 95000 Cergy, France

E-mail: alexei.chepelianskii@universite-paris-saclay.fr

Received 22 July 2021, revised 4 October 2021
Accepted for publication 7 October 2021
Published 4 November 2021

Abstract

We investigate, both analytically and with numerical simulations, a Monte Carlo dynamics at zero temperature, where a random walker evolving in continuous space and discrete time seeks to minimize its potential energy, by decreasing this quantity at each jump. The resulting dynamics is universal in the sense that it does not depend on the underlying potential energy landscape, as long as it admits a unique minimum; furthermore, the long time regime does not depend on the details of the jump distribution, but only on its behavior for small jumps. We work out the scaling properties of this dynamics, as embodied by the walker probability density. Our analytical predictions are in excellent agreement with direct Monte Carlo simulations.

Keywords: Monte Carlo, Metropolis algorithm, Markov chains

(Some figures may appear in colour only in the online journal)

1. Introduction

The steepest descent method is one of the oldest optimization schemes. Cauchy, with a rather minimal mention in one of his papers, is credited for its formulation [1]. It amounts to searching for the minimum of a well behaved function by following the steepest gradient 'downhill'. A stochastic reformulation has been proposed [2], to alleviate the computational cost of working in spaces with large dimensions: the gradient is then estimated over a restricted set of directions, drawn randomly [3]. This technique is widely used in machine learning [4].

*Author to whom any correspondence should be addressed.

We are interested here in a minimal version of stochastic gradient descent, in a one-dimensional setting. A random walker on the line, with position denoted by x, moves by random jumps; the move is accepted only if it leads to the decrease of some objective function $U(x)$, referred to as the potential. In this respect, the walker is greedy, performing moves that always decrease U. We assume that U admits a single minimum: we are not interested in finding this point (taken as $x = 0$ below), but rather in the dynamics of the walker upon approaching it. The specific form of the potential is immaterial; it does not need to be symmetric, as long as it has no local minima, beyond the global maximum at $x = 0$. Such an algorithm can be viewed as the vanishing temperature limit of a standard Metropolis sampling, a Monte-Carlo method where a Markov chain is constructed to sample the phase space according to a predefined target distribution [5–7]. As simple as it is—the walker position distribution function is increasingly peaked at $x = 0$—such a dynamics exhibits non trivial features, depending on the sampling chosen. A key role is indeed played by the probability distribution of attempted jumps, and in particular, its behavior for small jumps. The present $T = 0$ problem differs from widespread approaches such as simulated annealing, where the temperature ruling the evolution of a random walker is gradually decreased, in order to find minima in a given potential landscape. Here, the motivation is different: the location of the potential minimum is known, and we are interested in the dynamics towards this target.

Starting from the Metropolis rule, we define in section 2 the dynamics, in discrete time $n = 0, 1, \ldots$. The walker's density $P_n(x)$ evolves at long time toward $P_\infty(x) = \delta(x)$, where δ denotes the Dirac distribution. Our goal is to resolve the approach to this limiting form, that takes place in a self-similar way. We show indeed in section 3 that $P_n(x)$ admits a scaling form at long times, where the whole position and temporal information is encoded in a single universal scaling function, independent of $U(x)$ and the initial condition. Its generic properties are studied analytically. A number of exact solutions, presented in section 4 corroborate the general findings, and provide additional insights into the long time dynamics. All predictions fare well compared to Monte Carlo simulations, where the original dynamics is directly implemented.

2. The model

We consider a single particle moving on a line in the presence of an external confining potential $U(x)$, having a single minimum, and no local minima. In other words, $U(x)$ should be a monotonically increasing function of $|x|$. At time n, an attempted jump η_n is drawn from a distribution $w(\eta)$ and the particle moves according to the Metropolis rules

$$x_n = \begin{cases} x_{n-1} + \eta_n & \text{with prob. } p = \min\left(1, e^{-\beta\,\Delta U}\right) \quad \text{where} \quad \Delta U = U(x_{n-1} + \eta_n) - U(x_{n-1}) \\ x_{n-1} & \text{with prob. } 1 - p, \end{cases} \tag{1}$$

where β is for inverse temperature $(k_B T)^{-1}$. The position distribution $P_n(x)$ evolves via the Master equation

$$P_n(x) = \int_{-\infty}^{\infty} \mathrm{d}x'\, P_{n-1}(x')\, w(x - x')\, \min\left(1, e^{-\beta\,(U(x) - U(x'))}\right) + \left[1 - \int_{-\infty}^{\infty} \mathrm{d}x'\, w(x' - x)\, \min\left(1, e^{-\beta\,(U(x') - U(x))}\right)\right] P_{n-1}(x), \tag{2}$$

where the jump distribution $w(\eta)$ is assumed to be symmetric: $w(\eta) = w(-\eta)$. It is convenient to replace the 'min' function above by the following identity

$$\min\left(1, e^{-\beta\left(U(x)-U(x')\right)}\right) = \theta\left(U(x') - U(x)\right) + \mathrm{e}^{-\beta\left(U(x)-U(x')\right)}\,\theta\left(U(x) - U(x')\right), \tag{3}$$

where $\theta(z)$ is the Heaviside theta function: $\theta(z) = 1$ if $z > 0$ and $\theta(z) = 0$ if $z < 0$.

We now focus on the limit $T = 0$, i.e. $\beta = \infty$ [8]. In this limit, (3) becomes

$$\min\left(1, e^{-\beta\left(U(x)-U(x')\right)}\right) = \theta\left(U(x') - U(x)\right) = \theta\left(|x'| - |x|\right), \tag{4}$$

where in arriving at the last equality we used the fact that the potential increases monotonically with $|x|$ so that $U(x') > U(x)$ implies $|x'| > |x|$. Thus, in this limit, the particle can jump only downhill, and all uphill moves are forbidden. More precisely, if the particle is at x at step n, then in the next step it can jump only to the region $x' \in [-|x|, |x|]$. Any jump that takes it outside this region is forbidden at $T = 0$ (see figure 1). Thus the explicit dependence of the position distribution $P_n(x)$ on the form of the potential $U(x)$ drops out in this $T = 0$ limit, that can naturally be simulated by a Monte Carlo method, as described in appendix A. However, even in this relatively simple limit, the evolution of the position distribution remains rather nontrivial. With the simplification in (4), the Master equation (2) reduces to a simpler form

$$P_n(x) = \int_{-\infty}^{\infty} \mathrm{d}x'\, P_{n-1}(x')\, w(x - x')\, \theta\left(|x'| - |x|\right) + \left[1 - \int_{-\infty}^{\infty} \mathrm{d}x'\, w(x' - x)\, \theta\left(|x| - |x'|\right)\right] P_{n-1}(x). \tag{5}$$

One can check that equation (5) satisfies the probability conservation

$$\int_{-\infty}^{\infty} P_n(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} P_{n-1}(x)\,\mathrm{d}x. \tag{6}$$

Since the particle moves only downhill, we expect that at long times, i.e. in the limit $n \to \infty$, the position distribution should approach a delta function at the origin (irrespective of the initial condition)

$$\lim_{n\to\infty} P_n(x) = \delta(x). \tag{7}$$

We are interested in computing how the position distribution relaxes to this steady state.

For the analytical work, it is convenient to start from a symmetric initial condition. This ensures that at all times $P_n(x)$ is also symmetric, $P_n(x) = P_n(-x)$. In that case, we can just focus on $x \geqslant 0$. This symmetry assumption will be tested against numerical simulations, run in conditions of asymmetric initial conditions. It will be seen that the dynamics gradually suppresses non-symmetrical features of the position distribution. An explicit solution in a specific case will confirm this computational observation. Our Master equation (5) then reads (restricted to $x \geqslant 0$)

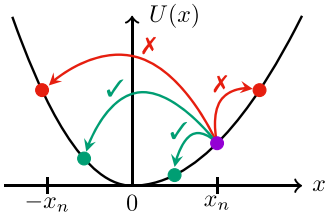

Figure 1. Possible moves of a particle at $T = 0$ in the potential $U(x)$, that monotonically increases with $|x|$. Only downhill stochastic moves are allowed (greedy motion). If at step n, the particle is at x_n, then at the next step it can jump only into the region $[-x_n, x_n]$. Any jump that takes the particle outside this interval is forbidden at $T = 0$.

$$P_n(x) = \int_{-\infty}^{-x} \mathrm{d}x'\, P_{n-1}(x')\, w(x - x') + \int_{x}^{\infty} \mathrm{d}x'\, P_{n-1}(x')\, w(x - x') + \left[1 - \int_{-x}^{x} \mathrm{d}x'\, w(x' - x)\right] P_{n-1}(x). \tag{8}$$

In the first integral on the right-hand side (rhs) we make the change of variable $x' \to -x'$ and use the symmetry $P_n(x) = P_n(-x)$ to get for all $x \geqslant 0$

$$P_n(x) - P_{n-1}(x) = \int_{x}^{\infty} \mathrm{d}x'\, P_{n-1}(x')\, \left[w(x + x') + w(x - x')\right] - \left[\int_{-x}^{x} \mathrm{d}x'\, w(x' - x)\right] P_{n-1}(x). \tag{9}$$

3. Scaling analysis of the master equation: asymptotic results

3.1. The scaling ansatz

The integral equation (9) is still hard to solve exactly for all n and for arbitrary symmetric jump distribution $w(\eta)$. However, in the large n limit, further simplifications occur. Indeed, for large n we make a scaling ansatz for $P_n(x)$ that should be verified *a posteriori* and also confirmed numerically. Our ansatz reads

$$P_n(x) = B n^{\alpha} F\left(B\, n^{\alpha}\, x\right), \tag{10}$$

where the exponent $\alpha > 0$ and the scale factor B will be selected subsequently. In doing so, we seek to 'resolve' the structure of the asymptotic Dirac delta distribution, toward which the position distribution evolves. Since $P_n(x)$ is taken here symmetric around $x = 0$, the scaling function $F(z)$ is symmetric: $F(z) = F(-z)$. In addition, from the normalization of $P_n(x)$, the scaling function $F(z)$ must satisfy the constraint

$$\int_{-\infty}^{\infty} F(z)\, \mathrm{d}z = 1. \tag{11}$$

The scaling form in (10) makes sense physically. It stipulates that the width of the distribution decreases for large time n as $\sim n^{-\alpha}$ with $\alpha > 0$, while the value at the peak increases as n^{α}.

Thus, as n increases, the position distribution function $P_n(x)$ gets more and more peaked near $x = 0$, and eventually approaches the delta function $P_n(x) \to \delta(x)$ as $n \to \infty$.

We substitute this scaling ansatz (10) in the master equation (9) and evaluate the left-hand side (lhs) and the rhs For large n, the lhs simplifies to

$$P_n(x) - P_{n-1}(x) \approx \frac{\partial P_n(x)}{\partial n} \approx B\alpha n^{\alpha-1} \left[F(z) + z\,F'(z)\right], \quad \text{where} \quad z = Bn^{\alpha}x. \tag{12}$$

We now substitute the scaling ansatz (10) on the rhs of (9). After making the change of variable $z' = Bn^{\alpha}x'$ inside the integrals on the rhs of (9), it reads

$$\text{r.h.s} \approx \int_z^{\infty} \mathrm{d}z'\, F(z') \left[w\left(\frac{z+z'}{B\,n^{\alpha}}\right) + w\left(\frac{z-z'}{Bn^{\alpha}}\right)\right] - \left(\int_{-z}^{z} \mathrm{d}z'\, w\left(\frac{|z'-z|}{Bn^{\alpha}}\right)\right) F(z). \tag{13}$$

Since we assumed $\alpha > 0$ (to be verified *a posteriori*), it follows that as $n \to \infty$, the arguments of the function w in (13) approach 0. Hence, the scaling behavior depends crucially on how the jump distribution function $w(\eta)$ behaves for small $|\eta|$. We consider the following natural class of power law behaviors near $\eta = 0$

$$w(\eta) \approx C_p\,|\eta|^p \quad \text{as} \quad \eta \to 0, \tag{14}$$

where $C_p > 0$ is a positive constant and $p > -1$ to ensure the normalization. Substituting this behavior on the rhs in (13) and equating it to the lhs in (12) we get, to leading order for large n,

$$B\,\alpha\,n^{\alpha-1}\left[F(z) + z\,F'(z)\right] = \frac{C_p}{B^p}\,n^{-\alpha p}\left[\int_z^{\infty} \mathrm{d}z'\,\left[(z+z')^p - (z'-z)^p\right]\,F(z') - \left(\int_{-z}^{z} \mathrm{d}z'\,|z'-z|^p\right)\,F(z)\right]. \tag{15}$$

Note that $F'(z) = \mathrm{d}F(z)/\mathrm{d}z$. In order that both sides scale as the same power of n, for large n, we must have the exponent

$$\alpha = \frac{1}{1+p}. \tag{16}$$

Furthermore, let us choose the scale factor B as

$$B = \left[C_p(1+p)\right]^{1/1+p}. \tag{17}$$

With this choice of B, the scaling function $F(z)$ now depends only on the single parameter p, hence we will denote it by $F_p(z)$. It then satisfies the integro-differential equation in $z \geqslant 0$, obtained from (15)

$$\begin{aligned} F_p(z) + z\,F_p'(z) &= \int_z^{\infty} \mathrm{d}z'\,\left[(z'+z)^p + (z'-z)^p\right]\,F_p(z') \\ &\quad - \left(\int_0^{z} \mathrm{d}z'\,\left[(z+z')^p + (z-z')^p\right]\right)\,F_p(z) \\ &= \int_z^{\infty} \mathrm{d}z'\,\left[(z'+z)^p + (z'-z)^p\right]\,F_p(z') - \frac{(2z)^{1+p}}{1+p}\,F_p(z). \end{aligned} \tag{18}$$

To summarise, we justified *a posteriori* that the position distribution $P_n(x)$ satisfies the scaling form

$$P_n(x) \approx B\, n^{1/(1+p)}\, F_p\left(B\, n^{1/(1+p)}\, x\right), \quad \text{where} \quad B = \left[C_p(1+p)\right]^{1/(1+p)}, \tag{19}$$

and the scaling function $F_p(z)$, indexed by p, is determined from the solution of the integro-differential equation (18). Note that $F_p(z)$ is symmetric in z and hence satisfies the normalization condition

$$\int_0^\infty F_p(z)\, \mathrm{d}z = \frac{1}{2}. \tag{20}$$

Thus the scaling function $F_p(z)$ depends only on the index p, but otherwise is universal, i.e. independent of the details of the jump distribution $w(\eta)$. For general $p > -1$, it is hard to solve the integro-differential (18) exactly. However, as we show below, one can derive the asymptotic behaviors of $F_p(z)$ as $z \to \infty$ and as $z \to 0$ for general $p > -1$. Later, we show that for the three special cases, $p = 0$, $p = 1$ and $p = 2$, it is possible to obtain exact solutions for the full scaling function $F_p(z)$.

3.2. Large z behavior of $F_p(z)$

For large z, the lhs of (18) is dominated by the second term $zF_p'(z)$. In contrast, for large z, the first term on the rhs of (18) goes to zero, while the second term dominates, as can be checked *a posteriori*. Hence, for large z, we get $zF_p'(z) \approx -[(2z)^{p+1}/(p+1)]F_p(z)$. Integrating, we obtain for large z and any $p > -1$

$$F_p(z) \sim \exp\left[-\frac{2^{p+1}}{(1+p)^2}\, z^{1+p}\right]. \tag{21}$$

This result has a neat physical interpretation. In fact, expressing the scaled variable $z = Bn^{1/(p+1)}x$ in terms of the original distance variable x, one finds (up to power law prefactors)

$$P_n(x) \sim F_p(z) \sim \exp\left[-\frac{C_p}{1+p}\, (2\,x)^{p+1}\, n\right]. \tag{22}$$

Hence $P_n(x)$ decays exponentially with increasing n for fixed x. This can be understood from a very simple Poisson type of argument. Consider the particle to be at x at large time n such that $B^{-1}n^{-1/(1+p)} \ll |x| \ll 1$. In this regime $|z| = Bn^{1/(1+p)}|x| \gg 1$. From this position the particle attempts a jump. It succeeds if the jump takes it into $x' \in [-x, x]$, otherwise it is unsuccessful and the particle stays at x. Also, the influx of probability to x from higher values of $|x|$ is negligible in this regime. Hence, the probability to stay at x after step n is approximately

$$P_n(x) \sim \left[1 - p_{\text{accept}}(x)\right]^n, \tag{23}$$

where the acceptance probability at x is given by

$$p_{\text{accept}}(x) = \int_{-x}^{x} \mathrm{d}x'\, w(x'-x) = \int_0^x \left[w\left(x+x'\right) + w\left(|x-x'|\right)\right] \mathrm{d}x'. \tag{24}$$

Since $|x| \ll 1$, the argument of w is small and we can expand in Taylor series. Keeping only the leading order term for $|x| \ll 1$ and using (14), we obtain

$$p_{\text{accept}}(x) \approx \frac{C_p}{(1+p)} (2\,x)^{1+p}. \tag{25}$$

Substituting this result in (23), we recover, for large n, the result in (22).

3.3. Small z behavior of $F_p(z)$

Extraction of the small z behavior of $F_p(z)$ from (18) is non trivial. We consider the rhs of (18) and start with the small z behavior of the first term $\int_z^\infty \mathrm{d}z' \left[(z'+z)^p + (z'-z)^p\right] F_p(z')$. Consider the first integral and rewrite it as

$$R_1 = \int_z^\infty \mathrm{d}z' (z'+z)^p F_p(z') = \int_{2z}^\infty \mathrm{d}u\, u^p F_p(u-z). \tag{26}$$

We now expand $F_p(u-z)$ in Taylor series for small z and keep terms up to $O(z^2)$. Furthermore, we rewrite the integral $\int_{2z}^\infty = \int_0^\infty - \int_0^{2z}$ and then expand the second part also for small z. After some algebra, this gives for small z

$$R_1 = b_0 + b_1\, z + b_2\, z^2 - \frac{F_p(0)}{1+p} (2z)^{p+1} + O(z^\gamma), \tag{27}$$

with

$$b_k = \frac{(-1)^k}{k!} \int_0^\infty \mathrm{d}u\, u^p\, F_p^{(k)}(u) \quad \text{where} \quad F_p^{(k)}(u) = \frac{\mathrm{d}F_p^k(u)}{\mathrm{d}u^k}. \tag{28}$$

The exponent γ is given by

$$\gamma = \min(3, p+2). \tag{29}$$

Repeating the same exercise with the second integral we get

$$\begin{aligned} R_2 &= \int_z^\infty \mathrm{d}z' (z'-z)^p F_p(z') = \int_0^\infty \mathrm{d}u\, u^p F_p(u+z) \\ &= b_0 + b_1\, z + b_2\, z^2 + O(z^3). \end{aligned} \tag{30}$$

Adding (27) and (30), and substituting on the rhs of equation (18), we get

$$F_p(z) + z F_p'(z) = \frac{\mathrm{d}}{\mathrm{d}z}[z F_p(z)] = 2b_0 + 2b_2 z^2 - 2\frac{F_p(0)}{1+p}(2z)^{p+1} + O(z^\gamma) \tag{31}$$

Integrating we get for small z

$$F_p(z) = 2b_0 + \frac{2b_2}{3} z^2 - \frac{2^{p+2} F_p(0)}{(1+p)(2+p)} z^{1+p} + O(z^\gamma). \tag{32}$$

Thus, while the leading term is always $2b_0$, the subleading term depends on the value of p. Let us distinguish between three cases.

- $\boldsymbol{p > 1}$**:** in this case, the first two terms in the small z expansion of $F_p(z)$ are

$$F_p(z) \to a_0 + a_2\, z^2, \tag{33}$$

where the two coefficients are given by

$$a_0 = 2\, b_0 = 2 \int_0^\infty \mathrm{d}u\, u^p\, F_p(u) \tag{34}$$

$$a_2 = \frac{2b_2}{3} = \frac{1}{3}\int_0^\infty \mathrm{d}u\, u^p\, F_p''(u) = \frac{p(p-1)}{3} \int_0^\infty \mathrm{d}u\, u^{p-2}\, F_p(u). \tag{35}$$

Note that since $a_2 > 0$, the function $F_p(z)$ for small z actually increases as z increases. Finally as $z \to \infty$, $F_p(z)$ has to decay as in (21). Hence, the function $F_p(z)$ is non-monotonic as a function of z for $p > 1$. If one considers the symmetrized version of $F_p(z)$, there is a local minimum (a hole) at $z = 0$ (see figure 4 for the case $p = 2$).

- $\boldsymbol{-1 < p < 1}$**:** in this case, the first two terms are given by

$$F_p(z) \to a_0 + a_{p+1} z^{p+1}, \tag{36}$$

where the coefficients read

$$a_0 = 2b_0 = 2 \int_0^\infty \mathrm{d}u\, u^p F_p(u) \tag{37}$$

$$a_{p+1} = -\frac{2^{p+2} F_p(0)}{(1+p)(2+p)}. \tag{38}$$

Thus in this case, since $a_{p+1} < 0$, the function $F_p(z)$ decreases as z increases from 0, indicating that for $-1 < p < 1$, the scaling function $F_p(z)$ is likely to be a monotonically decreasing function of z, with a single peak at $z = 0$, see figure 2 for $p = 0$.

- $\boldsymbol{p = 1}$**:** finally in the marginal case $p = 1$, where one has to merge the second and the third term together on the rhs of (32), one gets

$$F_p(z) \to a_0 + a_2\, z^2, \tag{39}$$

where the coefficients are

$$a_0 = 2\, b_0 = 2 \int_0^\infty \mathrm{d}u\, u F_1(u) \tag{40}$$

$$a_2 = \frac{1}{3}\int_0^\infty \mathrm{d}u\, u F_1''(u) - \frac{4}{3} F_1(0) = -F_1(0). \tag{41}$$

As we will see later, in this case we can derive the full solution exactly, $F_1(z) = (1/\sqrt{\pi})\, e^{-z^2}$ whose small z expansion agrees with (40) and (41).

Let us then summarize the asymptotic behavior of $F_p(z)$ for general $p > -1$. Using the symmetry of $F_p(z)$ we get

$$F_p(z) \sim \exp\left[-\frac{2^{p+1}}{(1+p)^2}\, |z|^{1+p}\right] \quad \text{as} \quad |z| \to \infty. \tag{42}$$

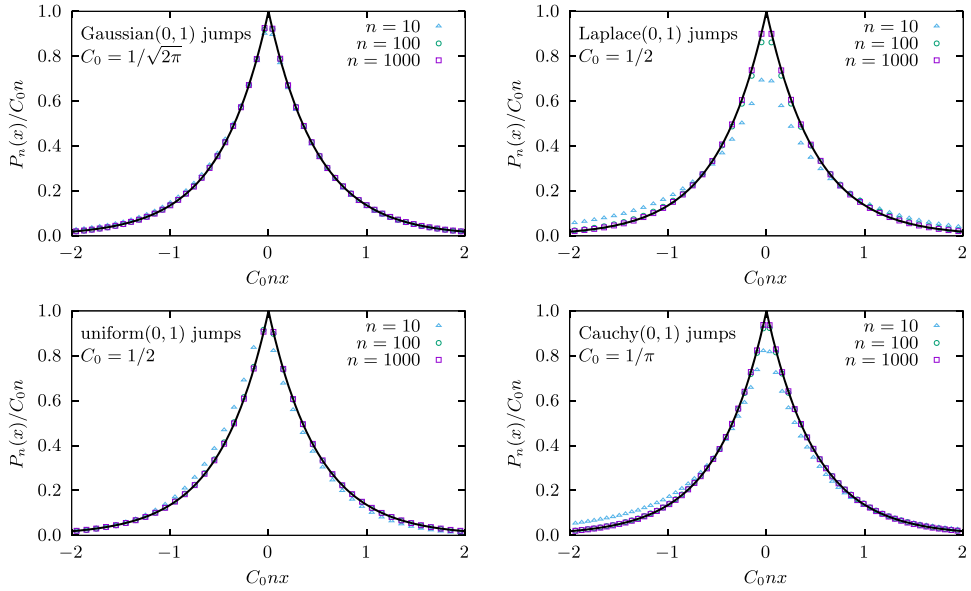

Figure 2. Case $p = 0$. Scaling function $F_0(z)$ obtained from the simulation data and the scaled $P_n(x)$ as in (46), compared with the theoretical prediction of the universal form $F_0(z) = e^{-2|z|}$. Four different jump distributions w have been used: the Gaussian, the Laplace, the uniform and the Cauchy distributions, all having $p = 0$. While the measurements for small times ($n = 10$) show deviations and an asymmetry due to the initial position chosen in the simulations ($x_0 = -1$), the convergence toward the form (46) is already very good for $n = 1000$ in all tested distributions. Measurements averaged over 10^7 independent runs.

The small z behavior depends on the index p and the first two terms are given by

$$F_p(z) \to \begin{cases} 2\displaystyle\int_0^\infty \mathrm{d}u\, u^p F_p(u) - \left[\frac{2^{p+2}F_p(0)}{(1+p)(2+p)}\right] |z|^{p+1} & \text{for} \quad -1 < p < 1 \\ 2\displaystyle\int_0^\infty \mathrm{d}u\, u^p F_1(u) - [F_1(0)]\, z^2 & \text{for} \quad p = 1 \\ 2\displaystyle\int_0^\infty \mathrm{d}u\, u^p F_p(u) + \left[\frac{p(p-1)}{3}\int_0^\infty \mathrm{d}u\, u^{p-2} F_p(u)\right] z^2 & \text{for} \quad p > 1. \end{cases} \tag{43}$$

There is thus a change of the small z behavior at $p = 1$, from a decreasing function of $|z|$ for $|p| < 1$ to an increasing one for $p > 1$. Indeed, with a probability density of jumps $w(\eta) \propto |\eta|^p$, $p > 1$ does penalize the occurrence of the small jumps, required to move, as compared to the case $p < 1$. When $p > 1$, the reduced likelihood of small jumps leads to a depletion of $F_p(z)$ at $z = 0$. Beyond the asymptotic results derived here, we present in the next section exact solutions for the three special cases $p = 0$, $p = 1$ and $p = 2$.

4. Exact scaling solutions

4.1. Exact solution for p = 0

The case $p = 0$ includes most natural symmetric jump distributions such as

- Gaussian: $w(\eta) = \frac{1}{\sqrt{2\,\pi}}\, e^{-\eta^2/2}$
- Double-exponential: $w(\eta) = \frac{1}{2}\, e^{-|\eta|}$, see also appendix B.
- Uniform distribution: $w(\eta) = \frac{1}{2}\ [\theta(\eta + 1) - \theta(\eta - 1)]$.
- Long-ranged distributions such as $w(\eta) = 1/[\pi(1 + \eta^2)]$, the Cauchy jump distribution.

In all these cases, $w(\eta) \to$ const. as $\eta \to 0$, implying $p = 0$. In this case, the integro-differential equation (18) reads

$$F_0(z) + z\, F_0'(z) = 2 \int_z^\infty \mathrm{d}z'\, F_0(z') - 2zF_0(z), \tag{44}$$

and it must satisfy the normalization condition (20) together with the large z asymptotic behavior (42). The solution of (44) turns out to be simple, as can be checked by direct substitution:

$$F_0(z) = e^{-2z}, \quad \text{for} \quad z \geqslant 0. \tag{45}$$

The large and small z behaviors in (42) and (43) are consistent with the exact solution (45), as can be checked easily. Thus for $p = 0$, our prediction is that $P_n(x)$ is given by the scaling form

$$P_n(x) \approx C_0 n F_0\,(C_0 n x)\,, \quad \text{with} \quad F_0(z) = e^{-2|z|}, \tag{46}$$

where we used the symmetry $F_0(z) = F_0(-z)$ and $C_0 = w(0)$. In figure 2, we compare our theoretical prediction for $F_0(z)$ with the numerical simulations (see appendix A for details on the numerical aspects), obtained for three different jump distributions all with $p = 0$, namely the Gaussian, the double-exponential and the uniform jump distributions. We find excellent agreement between the theoretical prediction and the simulations. The slight asymmetry observed between positive and negative values of x is a fingerprint of the asymmetric initial conditions. Its progressive disappearance gives an idea on the speed of convergence toward the scaling form. We come back to this question in appendix B.

4.2. Exact solution for p = 1

An example of a jump distribution that belongs to the $p = 1$ class is the symmetric Weibull distribution (normalized to unity)

$$w(\eta) = |\eta| e^{-\eta^2}. \tag{47}$$

In this case, setting $p = 1$ in (18) we get

$$F_1(z) + z\, F_1'(z) = 2 \int_z^\infty \mathrm{d}z'\, z' F_1(z') - 2z^2 F_1(z), \tag{48}$$

where the solution $F_1(z)$ should satisfy the normalization condition $\int_0^\infty F_1(z)\,\mathrm{d}z = 1/2$ and also the large z asymptotic behavior in (42) with $p = 1$. Remarkably, (48) also admits a simple

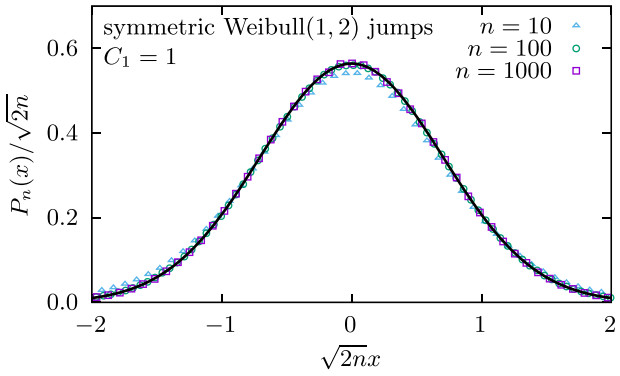

Figure 3. Case $p = 1$. The scaling function $F_1(z)$ obtained from the simulation data for the scaled $P_n(x)$ as in (50), compared with the theoretical prediction of the universal form $F_1(z) = (1/\sqrt{\pi})\, e^{-z^2}$. In the simulation, we used the symmetric Weibull jump distribution with shape factor 2 shown in (47), thus $C_1 = 1$. The agreement for $n \geqslant 100$ is excellent. Measurements obtained from 10^7 independent runs with $n = 1000$ jumps each. Here again, the slight $x \to -x$ asymmetry observed at early times is a consequence of having chosen an asymmetric initial condition.

solution (as can be checked by direct verification)

$$F_1(z) = \frac{1}{\sqrt{\pi}}\, e^{-z^2}, \tag{49}$$

where the prefactor is chosen such that $\int_0^\infty F_1(z)\, \mathrm{d}z = 1/2$. It is immediate to check that this exact solution is compatible with the large and small z behaviors in (42) and (43). Thus, for $p = 1$, our scaling prediction for the position distribution $P_n(x)$ reads

$$P_n(x) \approx \sqrt{2C_1}n^{1/2}\, F_1\left(\sqrt{2C_1}n^{1/2}x\right) \quad \text{with} \quad F_1(z) = \frac{1}{\sqrt{\pi}}\, e^{-z^2}, \tag{50}$$

where we again used the symmetry $F_1(z) = F_1(-z)$ and $B = \sqrt{2\,C_1}$ for $p = 1$ from (19). In figure 3, we compare the theoretical prediction for $F_1(z)$ in (50) with the numerically obtained $F_1(z)$ using the symmetric Weibull jump distribution in (47), finding excellent agreement.

4.3. Exact solution for p = 2

We now consider the case $p = 2$. This corresponds to symmetric jump distribution $w(\eta)$ with the small η behavior

$$w(\eta) \to C_2\, \eta^2 \quad \text{as} \quad \eta \to 0. \tag{51}$$

Then, the scaling theory for the position distribution (19) predicts that for large n

$$P_n(x) \approx (3C_2 n)^{1/3} F_2\left((3C_2 n)^{1/3} x\right), \tag{52}$$

where the scaling function $F_2(z) = F_2(-z)$ is symmetric and is normalized $\int_0^\infty F_2(z)\, \mathrm{d}z = 1/2$. The scaling function $F_2(z)$, setting $p = 2$ in (18), satisfies the integro-differential equation

(restricting only to $z \geqslant 0$)

$$F_2(z) + z\,F_2'(z) = 2z^2 \int_z^\infty \mathrm{d}z' F_2(z') + 2\int_z^\infty \mathrm{d}z'(z')^2 F_2(z') - \frac{8z^2}{3} F_2(z), \tag{53}$$

where $F_2'(z) = \mathrm{d}F_2(z)/\mathrm{d}z$. The function $F_2(z)$ must approach a constant as $z \to 0$ and should decay as $F_2(z) \sim \exp[-8z^3/9]$ as $z \to \infty$: this follows from the general asymptotics in (42) and (43). In addition, it must satisfy the normalization condition (20).

To solve this integro-differential equation (53), the strategy would be first to reduce it to a differential equation. To do this, let us define the cumulative scaling function

$$G_2(z) = \int_z^\infty F_2(z')\,\mathrm{d}z'. \tag{54}$$

We have $G_2'(z) = -F_2(z)$ and $G_2(z) \to 0$ as $z \to \infty$. Now consider the second term on the rhs of (53). With an integration by parts, we can rewrite it as

$$-\int_z^\infty \mathrm{d}z'(z')^2 G_2'(z) = z^2 G_2(z) + 2\int_z^\infty z' G_2(z')\,\mathrm{d}z'. \tag{55}$$

Thus, using $F_2(z) = -G_2'(z)$, equation (53) reads

$$-\left[1 + \frac{8z^3}{3}\right] G_2'(z) - zG_2''(z) - 4z^2 G_2(z) = 4\int_z^\infty z' G_2(z')\mathrm{d}z'. \tag{56}$$

Differentiating once more with respect to z gives a differential equation for $G_2(z)$

$$z\,G_2''(z) + \left[2 + \frac{8z^3}{3}\right] G_2''(z) + 12z^2 G_2'(z) + 4zG_2(z) = 0. \tag{57}$$

We divide by z, differentiate once more with respect to z and use $G_2'(z) = -F_2(z)$ to finally obtain a third order ordinary differential equation for $F_2(z)$

$$F_2'''(z) + \left[\frac{2}{z} + \frac{8z^2}{3}\right] F_2''(z) + \left[-\frac{2}{z^2} + \frac{52z}{3}\right] F_2'(z) + 16F_2(z) = 0. \tag{58}$$

A symbolic calculation software [9] allows to solve this equation, and it gives the general solution as a linear combination of three independent functions

$$F_2(z) = (A_1)\,{}_1F_1\left(\frac{1}{2}, \frac{1}{3}, -\frac{8z^3}{9}\right) + (A_2)\,z^2\,{}_1F_1\left(\frac{7}{6}, \frac{5}{3}, -\frac{8z^3}{9}\right) + (A_3)\,G^{2,2}_{2,2}\left(\begin{matrix} -\frac{1}{3}, \frac{1}{2} \\ -\frac{1}{3}, -\frac{2}{3} \end{matrix}\middle| -\frac{8z^3}{9}\right), \tag{59}$$

where ${}_1F_1(a, b, z)$ is the Kummer's confluent hypergeometric function and $G^{p,q}_{m,n}$ denotes the Meijer's G function. The constants A_1, A_2 and A_2 have to be determined from boundary conditions. First we note that the third function $G^{p,q}_{m,n}$ diverges as $1/z$ as $z \to 0$. Since this is not allowed (the scaling function $F_2(z)$ is normalizable and approaches a constant as $z \to 0$ from (43)). Hence we must have $A_3 = 0$. Thus

$$F_2(z) = A_1 \left[{}_1F_1\left(\frac{1}{2}, \frac{1}{3}, -\frac{8z^3}{9}\right) + B_2 z^2 {}_1F_1\left(\frac{7}{6}, \frac{5}{3}, -\frac{8z^3}{9}\right)\right], \tag{60}$$

where we define $B_2 = A_2/A_1$. The overall global constant A_1 can be fixed by the normalization condition in (20). The only remaining unknown constant B_2 has to be found from the boundary condition as $z \to \infty$, where we expect from (42) with $p = 2$ that $F_2(z) \sim e^{-8z^3/9}$. So, the constant B_2 has to be chosen such that $F_2(z)$ decays in this fashion as $z \to \infty$.

To fix B_2, we need to find the asymptotic behavior of ${}_1F_1(a, b, -x)$ when x approaches ∞ along the real line and $\theta = a - b$ is a non-integer. It turns out that this asymptotic behavior is rather subtle and has been obtained only rather recently [10]. For large x one gets

$$\begin{aligned} {}_1F_1(a, b, -x) = {} & \frac{x^{-a}\,\Gamma(b)}{\Gamma(b-a)} \sum_{k=0}^{\infty} \frac{(a)_k\,(1+a-b)_k}{k!\,x^k} \\ & + x^{a-b}\,e^{-x}\left[\cos(\pi(a-b)) \; + O(1/x)\right], \end{aligned} \tag{61}$$

where $(a)_k = \Gamma(a+k)/\Gamma(a)$ is the Pochamer symbol. Thus, generically, to leading order, it decays as a power law $\sim x^{-a}$ as $x \to \infty$. Substituting this asymptotic behavior in (60) using $x = 8z^3/9$, we get as $z \to \infty$

$$\begin{aligned} F_2(z) \approx A_1 & \left[\frac{\Gamma(5/3)}{\Gamma(1/2)}\left(-\frac{8}{9}\right)^{-7/6}\left(B_2 - \frac{\pi}{3\Gamma^2(5/6)}\left(\frac{2}{3}\right)^{1/3}\right)\right. \\ & \left. \times z^{-3/2} \sum_{k=0}^{\infty} \frac{9^k\,\Gamma(1/2+k)\Gamma(7/6+k)}{k!\,8^k\,z^{3k}}\right] \\ & + A_1\,\frac{\Gamma(1/3)\cos(\pi/6)}{\sqrt{\pi}}\left(\frac{8}{9}\right)^{1/6} z^{1/2}\,e^{-8z^3/9}. \end{aligned} \tag{62}$$

Since the boundary condition for large z predicts that $F_2(z) \sim \exp[-8z^3/9]$ (see equation (42)), we must eliminate the slow power law decay in (62). This can be done by choosing the constant B_2 as

$$B_2 = \frac{\pi}{3\Gamma^2(5/6)}\left(\frac{2}{3}\right)^{1/3}. \tag{63}$$

Finally, the global constant A_1 is obtained from the normalization condition $\int_0^\infty F_2(z)\,\mathrm{d}z = 1/2$. This gives (upon using Mathematica to do the integral)

$$A_1 = \left(\frac{3}{2}\right)^{1/3} \frac{\Gamma^2(5/6)}{\pi}. \tag{64}$$

Thus, with the two unknown constants A_1 and B_2 fixed, we then have our exact scaling function valid for all $z \geqslant 0$ (and symmetrically for $z \leqslant 0$)

$$F_2(z) = \left(\frac{3}{2}\right)^{1/3} \frac{\Gamma^2(5/6)}{\pi}\,{}_1F_1\left(\frac{1}{2}, \frac{1}{3}, -\frac{8z^3}{9}\right) + \frac{z^2}{3}\,{}_1F_1\left(\frac{7}{6}, \frac{5}{3}, -\frac{8z^3}{9}\right). \tag{65}$$

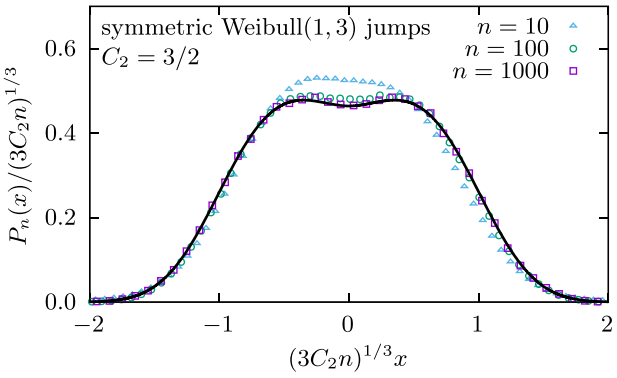

Figure 4. Case $p = 2$. Comparison between the numerically obtained scaling function $F_2(z)$ and the prediction of equations (52) and (65). The scaling function is symmetric ($F_2(-z) = F_2(z)$), non-monotonic, with a pair of maxima at $z = \pm 0.350\,322\ldots$ and a local minimum at $z = 0$. The simulation data for $P_n(x)$, scaled as in (10), are shown with the symbols; they have been obtained with a normalized jump distribution of the symmetric Weibull type and shape factor 3, i.e. $w(\eta) = \frac{3}{2}\,\eta^2\,e^{-|\eta|^3}$, thus $C_2 = 3/2$. The agreement for $n = 1000$ is excellent. Measurements obtained from 10^7 independent runs with $n = 1000$ jumps each. Here again, the asymmetry that is visible at $n = 10$ stems from our initial condition. Asymmetry is gradually washed out as time proceeds.

The function $F_2(z)$ has the following asymptotic behaviors

$$F_2(z) \approx \begin{cases} \left(\frac{3}{2}\right)^{1/3} \frac{\Gamma^2(5/6)}{\pi} + \frac{1}{3}\,z^2 + O(z^3) & \text{as } z \to 0 \\ b\,\sqrt{z}\,e^{-8z^3/9} & \text{as } z \to \infty \end{cases}, \tag{66}$$

where the constant prefactor b is given by

$$b = \frac{\sqrt{3}}{2^{5/6}\,\pi^{3/2}}\,\Gamma^2(5/6)\,\Gamma(1/3) = 0.595\,887\ldots \tag{67}$$

The large z asymptotic follows from the remaining nonzero term in (62) once B_2 is fixed. One can check that the small z behavior above is also completely in agreement with (43). A plot of this function $F_2(z)$ vs z is given in figure 4. In this figure, we also compare our theoretical prediction with numerical simulations for $p = 2$, finding excellent agreement. Interestingly, the function $F_2(z)$ is non-monotonic, and has a pair of maxima at $z \approx \pm 0.350\,322\ldots$. The reason behind this non monotonicity has been put forward below equation (43): $p > 1$ penalizes small jumps. Here, we consequently expect the dynamics to proceed in back and forth motion, from the left of the minimum to the right, and vice-versa for the subsequent successful jump. Such an oscillatory motion is fully compatible with the 'two hump' structure of the scaling function for $p > 1$. Yet, we stress that this oscillatory motion toward $x = 0$ is not specific to cases with $p > 1$: the late-time probability that the position changes sign after an accepted move is smoothly increasing when p increases. This probability is given by

$$\frac{\int_x^{2x} \eta^p\,\mathrm{d}\eta}{\int_0^{2x} \eta^p\,\mathrm{d}\eta} = 1 - \frac{1}{2^{p+1}}. \tag{68}$$

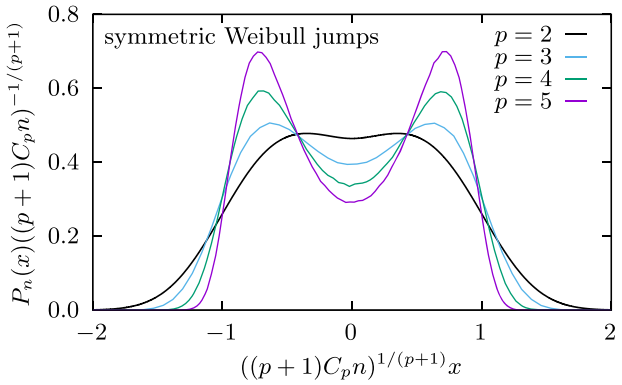

Figure 5. Numerically obtained asymptotic scaling functions $F_p(z)$ vs the scaling variable z (see Eq. (19)) for $p = 3, 4, 5$. The Monte Carlo simulations were run with a normalized jump distribution of the symmetric Weibull type and shape factor $(p+1)$, i.e. $w(\eta) = \frac{p+1}{2}\,\eta^p\, e^{-|\eta|^{p+1}}$. Measurements are collected over 10^6 independent runs with $n = 10^5$ jumps for $p = 3, 4$ and $n = 10^6$ jumps for $p = 5$. The curve for $p = 2$ is equation (65) shown for comparison.

Starting from 0 when $p \to -1^+$, it thus crosses 50% for $p = 0$, and exceeds 90% for for $p > 2.33$.

Finally, we show in figure 5 the evolution of the scaling function F_p with p, as obtained in Monte Carlo simulations. The two hump structure resulting from increasing p is visible. It is a consequence of the reduced likelihood of performing small jumps, when p increases, as discussed at the end of section 3.

5. Concluding remarks

We have studied a stochastic steepest-gradient descent on a line, in a potential landscape $U(x)$. A random walker proceeds in discrete time, with a succession of jumps (η_n at time n). The walker is greedy, in the sense that it only performs moves that decrease the energy $U(x)$. The dynamics does not depend on $U(x)$, and is driven to its minimum at $x = 0$, irrespective of the value (or even the existence) of the gradient, as long as the minimum is non-degenerate (no local minima). The long-time regime has been shown to be self-similar, with a scaling function $F_p(z)$ that only depends on the likelihood of small jump displacements, through the parameter p: Denoting the jump probability distribution by $w(\eta)$, we have $w(\eta) \propto \eta^p$ for $\eta \to 0$ (and thus $p = 0$ when $w(0)$ is finite). A Poisson-type of argument reveals that the large z tail of the scaling function is of the form $\log F_p(z) \propto -z^{1+p}$, and it also provides the scaling variable as $z = xn^{1/(1+p)}$. In other words, the root mean squared (r.m.s.) spread of the particle's position shrinks at large times n as $n^{-1/(1+p)}$. It is quite natural that large p values lead to a dynamical slow down, since they are associated to a smaller likelihood of small jumps, which is detrimental to evolution. A byproduct of the analysis is that redefining the clock from n to $\mathcal{N}$, where $\mathcal{N}$ would count the number of accepted moves, the scaling variable would become $z = \mathcal{N}x$, while the form of the scaling function F_p would of course be unaffected.

We restricted the analysis to the search of the large-time symmetric scaling solution, which is universal in the sense that it is independent of initial conditions. The excellent agreement

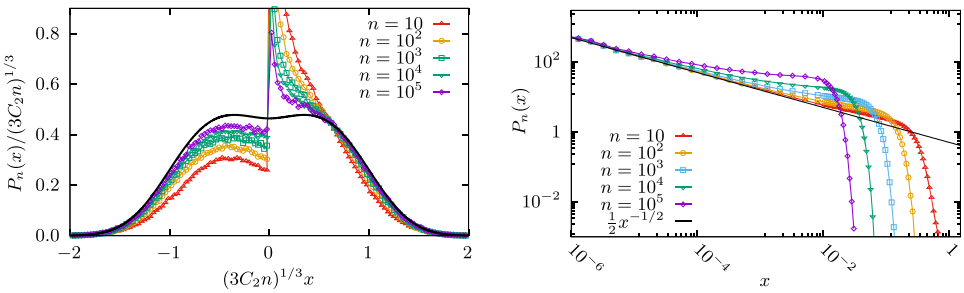

Figure 6. Monte Carlo results for initial conditions with a singularity at $x = 0$, here $P_0(x) = \frac{1}{2}x^{-1/2}$ for $0 < x < 1$, and $P_0(x) = 0$ otherwise; the jumps are of the symmetric Weibull type ($p = 2$). The left panel shows the measured $P_n(x)$ after n jumps averaged over 10^7 realizations and scaled to demonstrate the slow convergence toward the scaling form $F_2(z)$ shown as a black line. The right panel shows $P_n(x)$ on a log-log scale, as a function of the unscaled variable x. Comparison of the two panels shows that close to $x = 0$, $P_n(x)$ is of the form $x^{-1/2}$, before it crosses over to the asymptotic shape, further away from $x = 0$.

with the Monte Carlo simulations, that start from a non symmetric initial configuration, proves that asymmetric modes must decay faster than the symmetric ones. In some situations, the decay is slow. Such is the case when the initial distribution $P_0(x)$ is of the type $x^{-\mu}$ for small x with $0 < \mu < 1$, thus singular near the origin. By construction, the smaller an x-value, the least probable it will be affected by the dynamics. We then expect that in such a case, $P_n(x)$ features the same singularity as $P_0(x)$ close enough to $x = 0$. This is confirmed by figure 6, where $\mu = 1/2$. The initial asymmetry of $P_0(x)$ (note that $P_0(x) = 0$ for $x < 0$), impinges on later times, while the singularity at $x = 0$ has a rather spectacular effect on $P_n(x)$. Yet, figure 6 gives credit to the statement that for $n \to \infty$, the scaled P_n tends toward F_2, as given in equation (60). In addition, explicit exact calculations for the double exponential jump distribution, beyond scaling, show (1) the convergence toward the symmetric scaling form and (2) the faster decay of asymmetric contributions coming from the initial conditions. Details are provided in appendix B.

We did not discuss the cases where $w(\eta)$ is depleted near $\eta = 0$, with a vanishing probability in some interval $[-\eta^*, \eta^*]$: such a case would be non self-averaging, and lead to a dynamical arrest, whenever the walker arrives within the depletion segment $[-\eta^*, \eta^*]$. One would need to average over may realizations sharing the same initial conditions in order to obtain an interesting a smooth late-time dynamics, where $P_n(x) \to \delta(x)$, thereby resolving the structure of the δ peak. Conversely, in the case we have studied, and although we have averaged our Monte Carlo data over a large number of samples to garner statistics, the dynamics is self-averaging in the sense that a single trajectory leads to a well defined scaling function; statistics is increased by following the evolution on longer time scales.

In this paper, we focused on $T = 0$ where the steady state is trivial, but the late time relaxational dynamics is self-similar and typical observables (such as the r.m.s displacement) decay as a power law in time n for large n. This power law behavior is due to the vanishing acceptance probability of new moves at the minimum. In a recent paper [11], we studied the finite temperature version of this model where the dynamics satisfies detailed balance. In this case the steady state is of Gibbs–Boltzmann form and the relaxation of observables toward their steady state value becomes exponential with the error due to incomplete convergence decaying as Λ^n (with $\Lambda < 1$). The relaxation time $\tau = -1/\ln \Lambda$ has a rich behavior as a function of

the jump amplitude a, achieving a minimum value at an optimal a^*. Surprisingly for $a > a^*$, the relaxation at finite temperature is governed by self-similar scaling solutions very similar to the scaling ansatz established here for $T = 0$. The framework developed here may thus find applications beyond the $T = 0$ limit.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Monte Carlo simulations

The zero temperature dynamics investigated here defines a Markov chain, that is naturally simulated with the Monte Carlo technique [5–7]. In order to observe the relaxation of the particle distribution $P_n(x)$ we directly simulate m independent particles (here $m = 10^6$ or $m = 10^7$), all starting at $x_0 = -1$ by iterating

$$x_n = \begin{cases} x_{n-1} + \eta_n & \text{if } |x_{n-1} + \eta_n| < |x_{n-1}| \\ x_{n-1} & \text{otherwise,} \end{cases}, \tag{A1}$$

which is a simplified but equivalent version of equation (1) using equation (4). The η_n are independent random numbers drawn from the respective distributions $w(\eta)$, which (except for the Gaussian for which we used an implementation of the Ziggurat method [12]) can be generated using the inversion method [13]. For each value of n in which we are interested, e.g. $n = 10, 100, 1000$, we initialize a histogram and update it with the position of the m particles, after iteration n. We use the shape of the distribution we measured this way at different values of n to determine whether our Markov chain is long enough to reach the predicted scaling form, which typically happens very fast for jump distributions with low values of p and takes a longer time for jump distributions with larger values of p.

Appendix B. An exact solution for the double-exponential jump distribution

To find a case where the long time asymptotic properties of the probability distribution $P_n(x)$ can be obtained exactly, we use the $p = 0$ type jump distribution:

$$w(\eta) = \frac{1}{2} e^{-|\eta|}. \tag{B1}$$

It is convenient to separate $P_n(x)$ into symmetric and anti-symmetric components:

$$P_n(x) = S_n(x) + Z_n(x) \tag{B2}$$

with $S_n(x) = S_n(-x)$ and $Z_n(x) = -Z_n(-x)$. Substituting this decomposition in the master equation (5) and using the symmetry of $w(\eta)$, one finds that the associated Master equations for S_n and Z_n separate. We first focus on the symmetric component $S_n(x)$. For $x > 0$, the corresponding master equation reads:

$$S_{n+1}(x) = \cosh(x) \int_x^\infty e^{-y} S_n(y) \mathrm{d}y + R(x) S_n(x), \tag{B3}$$

where we introduced a compact notation for the probability to reject an attempted move from x:

$$R(x) = \int_{-\infty}^{\infty} \mathrm{d}y\, w(y-x)\theta\left(|y| - |x|\right) = e^{-x}\cosh x. \tag{B4}$$

This quantity ranges from 1 at small x (where most moves are rejected), to 1/2 at large x, where downhill moves only are accepted.

Introducing the generating function:

$$Q_s(x) = \sum_{n\geqslant 0} s^n S_n(x), \tag{B5}$$

we find the equation:

$$\cosh(x)\int_x^{\infty} e^{-y}Q_s(y)\mathrm{d}y + (e^{-x}\cosh x)Q_s(x) = \frac{Q_s(x) - S_0(x)}{s}. \tag{B6}$$

This integral equation can be reduced to a first order differential equation, that can be integrated explicitly:

$$Q_s(x) = \frac{S_0(x)}{1 - se^{-x}\cosh x} + 2s(e^{2x} + e^{4x})h_s(x)^{-1-\frac{1}{2-s}}\int_x^{\infty} \mathrm{d}y\, h_s(y)^{-1+\frac{1}{2-s}}S_0(y), \tag{B7}$$

where we introduced the notation:

$$h_s(x) = e^{2x}(2-s) - s. \tag{B8}$$

With this expression, we can use the general technique of singularity analysis [14] to obtain exact results on the large n behavior of $S_n(x)$, scrutinizing the complex plane position of the singularities of the generating function $Q_s(x)$ as function of the parameter s.

To make further progress, we assume $S_0(x) = \delta(|x| - y)/2$ (with $y > 0$). This choice is the symmetric part of $P_0(x) = \delta(x - y)$ and we find:

$$2Q_s(x) = \frac{\delta(x-y)}{1 - se^{-y}\cosh y} + 2s(e^{2x} + e^{4x})h_s(x)^{-1-\frac{1}{2-s}}\, h_s(y)^{-1+\frac{1}{2-s}}\theta(y-x). \tag{B9}$$

Singularities appear when $h_s(x) = 0$ or $h_s(y) = 0$. For $x < y$, the singularity in the complex s-plane that is closest to the origin $|s| = 0$ is the solution of $h_s(x) = 0$, and reads

$$s = \frac{2}{1 + e^{-2x}} = R(x)^{-1}. \tag{B10}$$

This implies that up to sub-exponential factors:

$$S_n(x) \sim R(x)^n. \tag{B11}$$

Particles cannot climb uphill in this model and $R(x') < R(x)$ for $x' > x > 0$. It is thus not surprising that the decay rate of $S_n(x)$ is determined by the rejection probability $R(x)$.

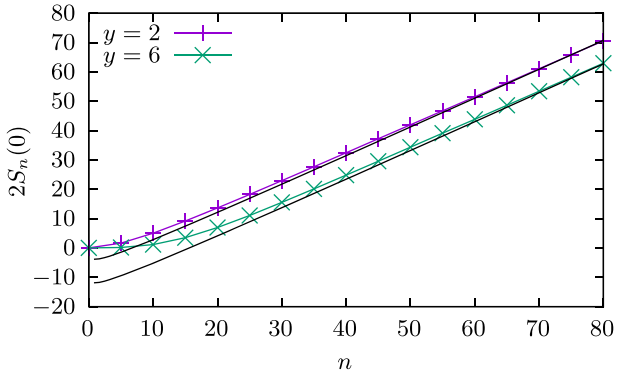

Figure 7. Comparison between the asymptotic behavior $2S_n(0)$ and numerical simulations for $S_0(x) = \delta(|x| - y)/2$, with $y = 2$ and $y = 6$. Symbols show numerical results and continuous black lines show equation (B13).

It is possible to obtain sub-exponential corrections for particular values of x. For example for $x = 0$, the singularity with the smallest modulus is at $s = 1$. Expanding the generating function around this position, we find:

$$2Q_s(0) \simeq \frac{1}{(1-s)^2} - \frac{1}{1-s}\log\frac{1}{1-s} - \frac{1+\log(e^{2y}-1)/2}{1-s} + \cdots \tag{B12}$$

This allows us to obtain the sub-exponential corrections up to the first term that depends on the initial position y:

$$2S_n(0) \simeq n - \log n - \left(1 + \gamma - \log 2 + \log[e^{2y} - 1]\right) + \cdots, \tag{B13}$$

where γ is the Euler–Mascheroni constant. This formula is in good agreement with numerical simulations which are shown shown in figure 7 for two different initial positions y of the random walker.

A similar asymptotic analysis for $x \ll 1$ gives:

$$2S_n(x) \simeq n^{1+x} R(x)^n + \cdots, \tag{B14}$$

where we derived only the leading sub-exponential term. Setting $\widetilde{x} = nx$, we can get:

$$S_n(n^{-1}\widetilde{x}) \simeq \frac{n}{2}\left(1 - \frac{|\widetilde{x}|}{n}\right)^n \simeq \frac{n}{2} e^{-|\widetilde{x}|}. \tag{B15}$$

Using the result $F_0(z) = e^{-2z}$ derived in section 4.1 for the present $p = 0$ situation, together with the fact that here, $C_0 = B = 1/2$, we recover the limiting scaling behavior derived in section 4.1: $P_n(x) = nF_0(nx/2)/2$, see equation (46).

For the anti-symmetric component of the probability distribution $Z_n(x)$, the master equation for $x > 0$ is:

$$Z_{n+1}(x) = \sinh(x) \int_x^\infty e^{-y} Z_n(y) \mathrm{d}y + R(x) Z_n(x). \tag{B16}$$

As previously, we introduce the generating function: $K_s(x) = \sum_{n\geqslant 0} Z_n(x)s^n$, and obtain

$$K_s(x) = \frac{Z_0(x)}{1 - se^{-x}\cosh x} + 4se^x \sinh x\, h_s(x)^{\frac{1}{-2+s}} \int_x^\infty e^{\frac{2(1-s)y}{2-s}} h_s(y)^{-2+\frac{1}{2-s}} Z_0(y)\mathrm{d}y. \quad \text{(B17)}$$

Focusing on the choice $P_0(x) = \delta(x-y)$, we set $Z_0(x) = [\delta(x-y) - \delta(x+y)]/2$. For $0 < x < y$, the position of the singularity with the smallest modulus lies at the same position as for the symmetric part: $s = R(x)^{-1}$. It is thus the sub-exponential factors that discriminate the decay rates of the symmetric and anti-symmetric components. Expanding the generating function to first order around $x = 0$ near the pole at $s = 1$, we find:

$$K_s(x) \simeq \frac{x(\coth y - 1)}{1-s}. \quad \text{(B18)}$$

This gives the expected small x asymptotic behavior for $Z_n(x)$:

$$Z_n(x) \simeq x(\coth y - 1). \quad \text{(B19)}$$

This confirms that the antisymmetric component is indeed negligible, in the large n limit, against its symmetric counterpart: $|Z_n(x)| \ll S_n(x)$, for small enough $|x|$. This stems from the additional $h_s(x)^{-1}$ factor present in equation (B7), as compared to equation (B17).

ORCID iDs

Alexei D Chepelianskii https://orcid.org/0000-0003-2404-9856
Satya N Majumdar https://orcid.org/0000-0001-9844-6980
Hendrik Schawe https://orcid.org/0000-0002-8197-1372
Emmanuel Trizac https://orcid.org/0000-0002-0166-1076

References

[1] Cauchy A 1847 Methode generale pour la resolution des systemes d'equations simultanees *Appl. Math. Sci.* **25** 536–8
[2] Robbins H and Monro S 1951 *Ann. Math. Stat.* **22** 400
[3] Spall J C 2012 Stochastic optimization *Handbook of Computational Statistics: Concepts and Methods* ed ed J E Gentle, W K Härdle and Y Mori (Berlin: Springer) p 173
[4] Sra S, Nowozin S and Wright S J 2011 *Optimization for Machine Learning* (Cambridge, MA: MIT Press)
[5] Frenkel D and Smith B 2002 *Understanding Molecular Simulations* 2nd edn (New York: Adademic)
[6] Newman M E J and Barkema G T 1999 *Monte Carlo Methods in Statistical Physics* (Oxford: Oxford University Press)
[7] Krauth W 2006 *Statistical Mechanics: Algorithms and Computations* (*Oxford Master Series in Physics*) (Oxford: Oxford University Press)
[8] Monthus C 2021 Convergence of time-averaged observables towards their steady values for Markov trajectories with application to skew-detailed-balance lifted-Markov processes (arXiv:2106.09429)
[9] Wolfram Research, Inc. 2021 Mathematica *(Version 12.3)* (Champaign, IL) https://support.wolfram.com/472
[10] Paris R B 2013 *Appl. Math. Sci.* **7** 6601
[11] Alexei D Chepelianskii, Satya N Majumdar, Hendrik Schawe and Emmanuel Trizac 2021 (in preparation)
[12] Marsaglia G and Tsang W W 2000 *J. Stat. Softw.* **5** 1
[13] Press W H, William H, Teukolsky S A, Vetterling W T, Saul A and Flannery B P 2007 *The Art of Scientific Computing* (*Numerical Recipes*) 3rd edn (Cambridge: Cambridge University Press)
[14] Flajolet P and Sedgewick R 2009 *Analytic Combinatorics* (Cambridge: Cambridge University Press)