Maximum of *N* independent Brownian walkers till the first exit from the half-space

2010 J. Phys. A: Math. Theor. 43 315001

(http://iopscience.iop.org/1751-8121/43/31/315001)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **43** (2010) 315001 (22pp) doi:10.1088/1751-8113/43/31/315001

Maximum of *N* independent Brownian walkers till the first exit from the half-space

P L Krapivsky[1], Satya N Majumdar[2] and Alberto Rosso[2]

[1] Department of Physics, Boston University, Boston, MA 02215, USA
[2] Université Paris Sud, CNRS, LPTMS, UMR 8626, Orsay F-91405, France

E-mail: paulk@bu.edu, majumdar@lptms.u-psud.fr and alberto.rosso@lptms.u-psud.fr

Received 29 April 2010, in final form 3 June 2010
Published 5 July 2010
Online at stacks.iop.org/JPhysA/43/315001

Abstract

We consider the one-dimensional target search process that involves an immobile target located at the origin and N searchers performing independent Brownian motions starting at the initial positions $\vec{x} = (x_1, x_2, \ldots, x_N)$, all on the positive half-space. The process stops when the target is first found by one of the searchers. We compute the probability distribution of the maximum distance m visited by the searchers till the stopping time and show that it has a power-law tail: $P_N(m|\vec{x}) \simeq B_N(x_1 x_2 \cdots x_N)/m^{N+1}$ for large m. Thus, all moments of m up to the order $(N-1)$ are finite, while the higher moments diverge. The prefactor B_N increases with N faster than exponentially. Our solution gives the exit probability of a set of N particles from a box $[0, L]$ through the left boundary. Incidentally, it also provides an exact solution of Laplace's equation in an N-dimensional hypercube with some prescribed boundary conditions. The analytical results are in excellent agreement with Monte Carlo simulations.

PACS numbers: 02.50.Cw, 05.40.Jc, 87.10.Mn

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The probability distribution of the maximum of a single one-dimensional Brownian motion (and its variants such as a Brownian bridge or an excursion) over a fixed interval of time $[0, t]$ has a long history in the probability literature [1–7]. The statistics of the maximum has diverse applications. One example is the Kolmogorov–Smirnov test in statistics that is used to compare, in a nonparametric way, two different probability distributions [8, 9]. Similarly, the distribution of the global maximum of a discrete-time random flights (including Lévy flights) has also been studied in the probability literature [10, 11], with more recent applications in computer science [12], physics [13] and chemistry [14].

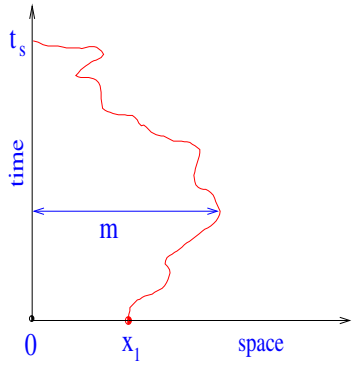

Figure 1. The trajectory (red line) of a single Brownian walker starting initially at x_1 till the stopping time t_s at which it hits the origin for the first time. The maximum distance traveled by the particle till t_s is denoted by m.

In statistical physics, there has been a recent revival of interest in related problems in the context of the distribution of the maximal height, measured with respect to a reference point, of $(1+1)$-dimensional fluctuating interfaces [15–17]. In the stationary state of a finite sample of size L, such fluctuating interfaces are often described by a Brownian bridge in space over an interval $[0, L]$, albeit with certain global constraints [16]. The statistics of the maximum has also been computed for continuous-time subdiffusive processes [18, 19] and has been used to analyze single particle trajectories [20].

The distribution of the maximum for a single Brownian motion (or its variants such as bridge, excursion, etc) has been extended to many Brownian motions, including certain strongly interacting random walkers, e.g. non-intersecting, so-called vicious random walkers [21–24]. (The latter problem has an intriguing connection to the Gaussian ensembles of the random matrix theory [21, 24, 25].) For independent walkers, the results on the distribution of the maximum have recently been used to compute the mean perimeter and the mean area of the convex hull of N independent planar Brownian motions [26, 27].

These results on Brownian motion and its variants represent rare exact analytical results for the extreme value statistics of *correlated* random variables, a subject of increasing current interest [for a brief review on 'extreme value statistics of correlated random variables' see 28]. However, all these results about the distribution of the maximum, for a single or multiple walkers, have been derived in the case when one considers the walkers over a fixed interval of time $[0, t]$. An interesting variation of this problem, with several applications, arises when the interval $[0, t]$ is not fixed, but itself varies from realization to realization, i.e. one observes the walker (or walkers) over a time interval $[0, t_s]$ where the stopping time t_s of the process itself is a random variable. For example, t_s may represent the first-passage time (through the origin) of a walker.

To be more precise, consider first a single Brownian walker that starts at time $t = 0$ at a position $x_1 > 0$. The position $x_1(t)$ of the walker evolves via the continuous-time stochastic equation, $\mathrm{d}x_1/\mathrm{d}t = \eta_1(t)$, where $\eta_1(t)$ is a Gaussian white noise with mean $\langle \eta_1(t) \rangle = 0$ and a correlator $\langle \eta_1(t)\eta_1(t') \rangle = \delta(t - t')$. The process stops at the stopping time t_s when the walker hits the origin for the first time (see figure 1).

Let m be the maximum displacement of the particle till the stopping time t_s. The statistics of the random variable m is interesting and it represents an example of the so-called first-passage Brownian functional [29]. The problem is a toy model of 'random search', where the

origin represents a fixed 'target' and the Brownian walker represents a random searcher. The search is called off when the searcher finds its target and m represents the maximum distance traveled by the searcher before it finds its target. For concreteness, we shall mostly used terminology related to random search, although there are several applications of this problem. For example, in the context of trapping [30–32] or predator–prey [33] models, the origin may represent an immobile target (prey) and the Brownian walker may represent a diffusing chemical trap (predator). The stopping time t_s is then the reaction time or the survival time of the prey and m denotes the maximum distance the predator travels before finding its prey. In the context of the directed Abelian sandpile model in $(1+1)$ dimensions [34], m represents the maximum lateral size of an avalanche [35]. The random variable m also plays an important role in characterizing the so-called staircase polygons [36]. In the context of queueing theory, where the position of the walker represents the length of a queue, m represents the maximum length of a queue during the so-called busy period [35, 36].

The probability density function (pdf) $P_1(m|x_1)$ of m (for fixed x_1) can easily be computed [35] and it turns out to be a pure power law

$$P_1(m|x_1) = \frac{x_1}{m^2}; \qquad m \geqslant x_1. \tag{1}$$

While this pdf is evidently normalized to unity, the average $\langle m\rangle$ and higher integer moments are infinite! The cumulative distribution of the maximum is given by

$$Q_1(L|x_1) = \text{Prob}[m \leqslant L|x_1] = \int_{x_1}^{L} P_1(m|x_1)\,\mathrm{d}m = 1 - \frac{x_1}{L}\,. \tag{2}$$

This distribution has a very simple interpretation: it just represents the exit probability of a Brownian particle [37], starting at $0 \leqslant x_1 \leqslant L$, from a box $[0, L]$ through its left boundary at 0.

In this paper, we study a generalization of this search problem, where there is still one fixed target at the origin, but there are N searchers who perform independent Brownian motions on the $x > 0$ axis, starting at the initial positions $\vec{x} \equiv (x_1, x_2, \ldots, x_N)$. The position $x_i(t)$ of the ith walker evolves with time t via the Brownian evolution

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \eta_i(t), \tag{3}$$

where $\eta_i(t)$ is a Gaussian white noise satisfying $\langle\eta_i(t)\rangle = 0$ and $\langle\eta_i(t)\eta_j(t')\rangle = \delta(t-t')\delta_{i,j}$. Since the walkers are independent, they can cross each other. The process stops at a stopping time t_s when the origin is hit for the first time by anyone of the walkers (e.g. the second walker (red) in figure 2). Note that t_s varies from one history of the process to another.

In the context of chemical kinetics [31], where the problem is generally referred to as the 'target annihilation' problem, various generalizations of this problem have been investigated including e.g. the situation where the target itself diffuses [38–42]. In the following, we shall limit ourselves to the case of an immobile target and focus on the statistics of the maximum distance m (from the target) traveled by any of the walkers till the stopping time t_s when the target is found. Thus, m denotes the distance of the farthest point on the x axis visited by any one of the walkers till t_s. Clearly m is a random variable fluctuating from one realization of the process to another. Our object of interest is the probability density $P_N(m|\vec{x})$ of this maximum distance m, given the number N of walkers and their initial positions $\vec{x}$. Thus, m provides an estimate (worst case) of the distance that needs to be covered by a team of N walkers to find a fixed target.

As in the single searcher case, let $Q_N(L|\vec{x}) = \text{Prob}[m \leqslant L|\vec{x}] = \int_0^L P_N(m|\vec{x})\,\mathrm{d}m$ be the cumulative probability that the maximum m till t_s is less than or equal to L. This cumulative distribution of the maximum can be interpreted as the solution of a different problem as in the

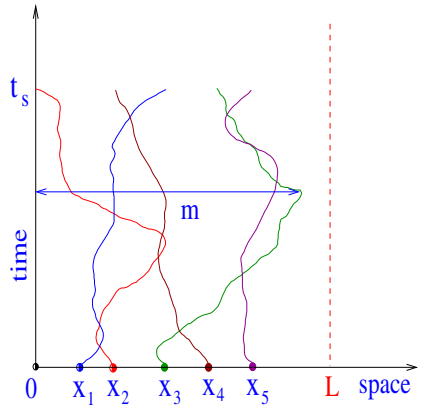

Figure 2. The trajectories of $N = 5$ independent Brownian walkers starting at the initial positions x_1, x_2, x_3, x_4 and x_5 till the stopping time t_s when one of the walkers (the second one (red) in this figure) hits the origin. The maximum displacement along the x direction till t_s (undergone, e.g., by the third particle (green)) is denoted by m. The cumulative probability $Q_N(L|\vec{x}) = \text{Prob}[m \leqslant L|\vec{x}]$ also represents the exit probability of the first particle from a box $[0, L]$ through its left boundary.

$N = 1$ case. Consider, for instance, a slightly different problem where again we have a set of N independent walkers, but now inside a box $[0, L]$, starting at the initial positions $\vec{x}$. Let us define the exit probability as the probability that the first particle that exits the box $[0, L]$ does so through 0 (and not through the upper boundary at L), see figure 2. As in the $N = 1$ case, this exit probability is precisely the cumulative distribution $Q_N(L|\vec{x})$ of the maximum m till the stopping time in the semi-infinite system, as it counts all those events where one of the trajectories hits the lower boundary 0 before hitting the upper boundary at L while all the others stay inside the box $[0, L]$ till this event of first-hitting the origin.

We will see that for this seemingly simple one-dimensional model of independent walkers, the statistics of m has a rich and nontrivial dependence on the number N of walkers. This is partly due to the fact that the same stopping time t_s for all the walkers effectively introduces a correlation between the trajectories of the walkers, even though each executes an independent Brownian motion. While for $N = 1$ the solution is simple, it becomes rather nontrivial even for $N = 2$!

Let us first summarize our main results. We compute the pdf $P_N(m|\vec{x})$ exactly for all $N \geqslant 1$ by a path counting (or path integral) method. We show that, for arbitrary $N \geqslant 1$, the pdf of the maximum has an asymptotic power-law tail

$$P_N(m|\vec{x}) \simeq B_N \frac{x_1\, x_2\, \cdots x_N}{m^{N+1}} \qquad \text{as } m \to \infty, \tag{4}$$

where the prefactor B_N has a nontrivial N dependence which we compute explicitly. For $N = 1$, we have $B_1 = 1$ and the asymptotic result in (4) is actually valid exactly for all $m \geqslant x_1$. For $N = 2$, we will see that

$$B_2 = \frac{1}{4\pi^2}\left[\Gamma\left(\frac{1}{4}\right)\right]^4 = 4.376\,88\ldots. \tag{5}$$

In particular, for large N, the prefactor B_N grows faster than exponentially

$$B_N \simeq N\left[\frac{4}{\pi}\ln(N)\right]^{N/2}. \tag{6}$$

Our asymptotic result (4) indicates that for N walkers, integer moments of m up to order $(N-1)$ are finite, while higher integer moments are infinite. Evidently, as N increases, the distribution becomes narrower and narrower as expected but it does so in a nontrivial fashion.

Recently, the cumulative distribution of the maximum m till the first-passage time t_s, or equivalently the exit probability $Q(L|x)$ from the box $[0, L]$ through the origin, was studied [43] for a generic self-affine stochastic process $x(t)$ starting at the initial position x. The process $x(t)$ typically grows with time as $x(t) \sim t^H$, where H is the Hurst exponent. This power-law growth of distance with time makes the process self-affine. An example is the ordinary Brownian motion, where $H = 1/2$. For such a generic self-affine process, it was argued [43] that the cumulative distribution of the maximum $Q(L|x) = \mathrm{Prob}(m \leqslant L|x) \sim 1 - A\,(x/L)^{\phi}$ in the limit $x/L \to 0$, where A is a constant. The exponent ϕ was found to be related to the persistence exponent θ via the scaling relation $\phi = \theta/H$ [43]. The persistence exponent θ characterizes the late time power-law decay of the survival probability, i.e. the probability that the process stays on the positive half-axis up to the time t [44]. Thus, the pdf of the maximum decays for large m as $P(m|x) \sim m^{-\phi-1}$ with $\phi = \theta/H$. The exact result (4) shows that if we think of the assembly of N independent Brownian motions as a single self-affine stochastic process in the N-dimensional space, then $\phi = N$. We will see later that the persistence exponent for this collective process is $\theta = N/2$ and the Hurst exponent $H = 1/2$. Thus, our exact result for this model supports the general scaling relation $\phi = \theta/H$ found in [43].

The paper is organized as follows. In section 2, we provide a simple heuristic argument in favor of our main result (4). This argument is not sufficient to compute the prefactor B_N exactly for all N. However, we show that this heuristic argument becomes asymptotically exact for large N and one can extract the limiting behavior of B_N for large N using an extreme value argument. In section 3, we set up the general method for computing the cumulative distribution $Q_N(L|\vec{x})$ of the maximum m. This requires solving Laplace's equation in an N-dimensional space with appropriate boundary conditions. We present explicit solutions for the cases $N = 1$ and $N = 2$. In section 4, we present an alternative path counting method that is more general, physically transparent and provides explicit results for all $N \geqslant 1$. In section 5, we present numerical results to verify our analytical predictions. Finally in section 6, we conclude with a summary and a list of interesting open problems. Some of the details of the computations are relegated to the appendices.

2. Heuristic argument

We begin with a simple heuristic argument in favor of (4). Consider the semi-infinite geometry with the independent Brownian motions, evolving via (3), starting at the initial positions $\{x_1, x_2, \ldots, x_N\}$. Let t_s be the stopping time when one of the walkers hits the origin. The probability distribution of t_s can exactly be computed as follows. As an input to solving the N-particle problem, let us first consider a single Brownian motion starting initially at $x_0 > 0$ and let $p(x_0, t)$ be the survival probability, i.e. the probability that the walker does not hit the origin up to the time t. This can easily be computed by various standard methods and is given by the well-known formula [29, 45]

$$p(x_0, t) = \mathrm{erf}\left(\frac{x_0}{\sqrt{2t}}\right); \qquad \mathrm{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y \mathrm{e}^{-u^2}\, \mathrm{d}u. \tag{7}$$

Turning to N walkers we note that since the walkers are independent, the probability that none of them hits 0 up to the time t is simply the product $\prod_{i=1}^{N} p(x_i, t)$. This is precisely the

probability that the stopping time $t_s > t$. Thus, the cumulative distribution of the stopping time t_s, given the initial positions $\vec{x}$, reads [31, 39]

$$\text{Prob}[t_s > t|\vec{x}] = \prod_{i=1}^{N} \text{erf}\left(\frac{x_i}{\sqrt{2t}}\right). \tag{8}$$

For large t, this cumulative distribution has a power-law tail

$$\text{Prob}[t_s > t|\vec{x}] \simeq \left(\frac{2}{\pi}\right)^{N/2} \frac{x_1\, x_2\, \cdots x_N}{t^{N/2}}. \tag{9}$$

This asymptotic for the survival probability states that the persistence exponent is $\theta = N/2$.

Result (8) and its asymptotic counterpart (9) are exact. Next comes the heuristic part. We note that for large t_s, the typical maximal displacement m in time t_s must scale as $m \sim \sqrt{t_s}$. Taking this relationship between the two random variables m and t_s seriously, we see that $\text{Prob}[m > L|\vec{x}] \sim \text{Prob}[t_s > L^2|\vec{x}]$ for large L. Since $Q_N(L|\vec{x}) = \text{Prob}[m \leqslant L|\vec{x}]$, we conclude that for large L

$$1 - Q_N(L|\vec{x}) = \text{Prob}[m > L|\vec{x}] \sim \text{Prob}[t_s > L^2|\vec{x}] \sim \frac{x_1\, x_2\, \cdots x_N}{L^N}, \tag{10}$$

where we used the result in equation (9). Taking derivatives with respect to L and putting $L = m$ then gives an approximate behavior of the probability density $P(m|\vec{x})$ of the maximum for large m

$$P_N(m|\vec{x}) \sim \frac{x_1\, x_2\, \cdots x_N}{m^{N+1}} \tag{11}$$

as mentioned in (4). This heuristic scaling argument thus provides, up to an overall N-dependent prefactor B_N, the leading asymptotic power-law tail of the distribution of m in (4).

To compute the prefactor B_N exactly for any N, one needs to go beyond this scaling argument. This requires a more sophisticated mathematical analysis that is carried out in the rest of the paper. However, it is possible to refine this heuristic argument, as shown below, that even provides the prefactor B_N exactly for large N.

In the above argument, the main approximation was to replace $m \sim \sqrt{t_s}$ for large m and then use the exact asymptotic distribution of t_s in equation (9) to compute the tail of the distribution of m. This approximation clearly ignores the fluctuations of m for a fixed t_s. We now use an extreme value argument to show that this approximation actually becomes exact for large N. We consider again a group of Brownian motions starting at the initial positions $\vec{x}$ and examine their trajectories over a *fixed* time interval $[0, t_s]$, with m denoting their global maximum in $[0, t_s]$. To compute the cumulative probability $Q_N(L, t_s|\vec{x}) = \text{Prob}[m \leqslant L|\vec{x}]$, we consider the trajectories that stay below the level L till t_s and also above the level 0. Now, for large L, the trajectories that contribute to $Q_N(L, t_s|\vec{x})$ typically have large excursions. So, to a first approximation, one can ignore the lower boundary at 0. For the ith walker, starting at x_i, the probability that its maximum stays below L can easily be computed: it is just the survival probability $p(x_0, t_s)$ in equation (7) with the initial position $x_0 = L - x_i$. Thus, the joint probability that all walkers stay below L till t_s (ignoring the lower boundary at 0) is just the product

$$Q_N(L, t_s|\vec{x}) \simeq \prod_{i=1}^{N} \text{erf}\left(\frac{L - x_i}{\sqrt{2t_s}}\right). \tag{12}$$

For large argument, the error function behaves as $1 - \text{erf}(x) = \text{e}^{-x^2}/(x\sqrt{\pi}) \simeq \text{e}^{-x^2}$ to leading order. Hence, for large L and large N, one can write

$$Q_N(L, t_s|\vec{x}) \simeq \exp[-N\, \text{e}^{-L^2/2t_s}] \to f[(L - a_N)/b_N], \tag{13}$$

where the scale factors $a_N = \sqrt{2t_s \ln(N)}$ and $b_N = \sqrt{t_s/(2\ln N)}$ and the scaling function $f(x) = \exp[-e^{-x}]$ is the standard Gumbel function. The pdf of m is just the derivative of the cumulative distribution. The derivative of the scaling function $f'(x) = \exp[-x - e^{-x}]$ has a peak at $x = 0$. This indicates that the random variable m has a peak at $m = a_N = \sqrt{2t_s \ln(N)}$ and the width of m around its peak is $b_N = \sqrt{t_s/(2\ln N)}$ that actually decreases with increasing N. Thus, for large N, the random variable m approaches to its mean value a_N with probability 1, i.e. $m = \sqrt{2t_s \ln(N)}$ with fluctuations around this value essentially negligible for large N. Using this relation in equation (9) provides the following tail for the cumulative distribution of m for large N:

$$1 - Q_N(L|\vec{x}) = \text{Prob}[m > L|\vec{x}] \simeq \text{Prob}\left[t_s > \frac{L^2}{2}(\ln N)^{-2}|\vec{x}\right] \simeq A_N \frac{x_1\, x_2 \cdots x_N}{L^N} \tag{14}$$

with the prefactor $A_N \simeq [4\ln(N)/\pi]^{N/2}$. Taking derivative with respect to L then gives the tail of the pdf of m in equation (4) with the prefactor, for large N,

$$B_N = N A_N \simeq N\left[\frac{4}{\pi}\ln(N)\right]^{N/2}. \tag{15}$$

We will see later that the same asymptotic result also follows from a more rigorous approach.

3. A backward Fokker–Planck method: Laplace's equation

In this section we show that the cumulative distribution $Q_N(L|\vec{x}) = \text{Prob}[m \leqslant L|\vec{x}]$ of the maximum m satisfies Laplace's equation in an N-dimensional hypercube with appropriate boundary conditions. To see this, it is first useful to consider $Q_N(L|\vec{x}) \equiv Q_N(\vec{x}|L)$ as a function of the coordinates $\vec{x}$ for a given fixed L. Note that $Q_N(\vec{x}|L)$ is the probability that starting at $\vec{x}$, the maximum of the process till t_s stays below the level L. The idea is to derive a differential equation for $Q_N(\vec{x}|L)$ using a backward approach where one focuses on the evolution of the system via (3) over a small time interval $[0, \mathrm{d}t]$ starting from the initial positions $\vec{x}$. According to (3), in this small time interval $\mathrm{d}t$, the ith particle moves from x_i to $x_i{}' = x_i + \eta_i(0)\mathrm{d}t$ where $\eta_i(0)$ is the noise at $t = 0$ that kicks the ith particle. Now, starting from this 'new' initial coordinates $x_i{}'$ the maximum of the system has to subsequently stay below L till the stopping time. Finally, one must sum over all possible values of the new coordinates $x_i{}'$. Thus, one must have

$$Q_N(x_1, x_2, \ldots, x_N|L) = \langle Q_N(x_1 + \eta_1(0)\,\mathrm{d}t, x_2 + \eta_2(0)\mathrm{d}t, \ldots, x_N + \eta_N(0)\,\mathrm{d}t|L)\rangle, \tag{16}$$

where $\langle\cdot\rangle$ denotes the averages over the initial noises $\eta_i(0)$. Expanding the right-hand side in a Taylor series and using (i) $\langle\eta_i(0)\rangle = 0$ and (ii) $\langle\eta_i(0)\eta_i(0)\rangle = 1/\mathrm{d}t$ (which follows from the delta correlator), one finds that $Q_N(\vec{x}|L)$ satisfies Laplace's equation in the N-dimensional hypercube, $0 \leqslant x_i \leqslant L$,

$$\nabla^2 Q_N(x_1, x_2, \ldots, x_N|L) = 0. \tag{17}$$

The information about the maximum is captured in the boundary conditions. For example, if $x_i = 0$, for any i, $Q_N = 1$ since if the ith particle starts at the origin, the process stops immediately ($t_s = 0$) and hence the maximum is necessarily (with probability 1) less than L. On the other hand, if $x_i = L$, for any i, one has $Q = 0$. This follows from the fact that if the ith particle starts at L, it will immediately cross the level L and the probability that the maximum will stay below L till t_s is necessarily zero.

To summarize, $Q_N(\vec{x}|L)$ satisfies Laplace's equation (17) in an N-dimensional hypercube $0 \leqslant x_i \leqslant L$ with $Q = 1$ for any $x_i = 0$ and $Q = 0$ for any $x_i = L$. Thus, it reduces to

an electrostatic problem where one needs to find the potential $Q_N(\vec{x}|L)$ inside the hypercube $[0, L]^N$, whose N faces touching the origin are held at a constant potential $Q_N = 1$ while the rest of the N faces are earthed ($Q_N = 0$). We present the solutions explicitly for $N = 1$ and $N = 2$ in the next two subsections.

Scaling. Let us remark that since the only length scales are the initial positions $\{x_i\}$ of the particles and the size of the box L, it is evident that the exit probability $Q_N(\vec{x}|L)$ satisfies the scaling property

$$Q_N(\vec{x}|L) = Q_N\left(\frac{x_1}{L}, \frac{x_2}{L}, \ldots, \frac{x_N}{L}\right) = Q_N(z_1, z_2, \ldots, z_N), \tag{18}$$

where the dimensionless scaled variables $0 \leqslant z_i = x_i/L \leqslant 1$.

Special initial condition and a duality relation. It is useful to consider a special initial condition where all the particles start from the same initial positions: $x_i = x$ for all $1 \leqslant i \leqslant N$. In this case, the exit probability is a function of a single scaled variable $0 \leqslant z = x/L \leqslant 1$: $Q_N(x_1 = x,\ x_2 = x, \ldots,\ x_N = x|L) = q_N(x/L)$ where the scaling function $q_N(z)$ satisfies

$$q_N(z) + q_N(1 - z) = 1. \tag{19}$$

This duality relation states that the exit probability through the right boundary at L starting from the initial position $L-x$ (of all the particles) is exactly identical to the exit probability through the left boundary 0 starting from the initial positions x (of all the particles). The duality relation (19) in particular states that

$$q_N(z = 1/2) = 1/2 \tag{20}$$

for all N. The general solution of Laplace's equation must satisfy equation (19) which actually provides a useful check for the validity of the solution.

3.1. $N = 1$

For $N = 1$, we have a second-order ordinary differential equation

$$\frac{\mathrm{d}^2 Q_1}{\mathrm{d}x_1^2} = 0 \qquad \text{with} \quad Q_1(0|L) = 1 \qquad \text{and} \qquad Q_1(L|L) = 0, \tag{21}$$

whose solution is

$$Q_1(x_1|L) = 1 - \frac{x_1}{L}; \qquad 0 \leqslant x_1 \leqslant L. \tag{22}$$

Since $Q_1(x_1|L) = Q_1(L|x_1) = \mathrm{Prob}[m \leqslant L|x_1]$, it follows, by differentiation, that the probability density of the maximum m has a strict power-law form for all m

$$P_1(m|x_1) = \frac{x_1}{m^2} \qquad \text{for} \quad m \geqslant x_1 \tag{23}$$

which is normalized to unity over $m \in [x_1, \infty]$ and all its integer moments diverge [35]. Thus, the prefactor $B_1 = 1$ in (4).

Clearly for $N = 1$, the exit probability $Q_1(x_1|L)$ is only a function of the scaled variable $z = x_1/L$: $Q_1(x_1|L) = q_1(x_1/L)$, where the scaling function $q_1(z)$ is simple:

$$q_1(z) = 1 - z \tag{24}$$

and evidently it satisfies the duality relation (19).

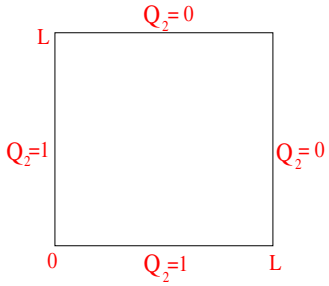

Figure 3. Laplace's equation to be solved in a square $(L \times L)$ with the boundary conditions as shown.

3.2. $N = 2$

The solution of Laplace's equation becomes more involved in higher dimensions $N > 1$. For $N = 2$, one needs to solve Laplace's equation in an $(L \times L)$ square with the boundary conditions shown in figure 3. The solution can explicitly be written down for this case, although the expression is rather cumbersome. Using separation of variables, one finds the appropriate solution [46]

$$Q_2(x_1, x_2|L) = 1 - \frac{x_1}{L} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \frac{\left[\sinh\left(n\pi\left(1 - \frac{x_2}{L}\right)\right) + (-1)^n \sinh\left(\frac{n\pi x_2}{L}\right)\right]}{\sinh(n\pi)} \times \sin\left(\frac{n\pi x_1}{L}\right). \tag{25}$$

It is straightforward to verify that (25) is a solution of Laplace's equation. Next we need to check that it satisfies the four boundary conditions (see figure 3). It is easy to check the two conditions: (i) $Q_2(x_1 = 0, x_2|L) = 1$ for all x_2 and (ii) $Q_2(x_1 = L, x_2|L) = 0$ for all x_2. The other two conditions can also be verified. For instance, putting $x_2 = L$ in (25), we get

$$Q_2(x_1, L|L) = 1 - \frac{x_1}{L} - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi x_1}{L}\right). \tag{26}$$

Using the identity

$$\sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi x_1}{L}\right) = \frac{\pi}{2}\left(1 - \frac{x_1}{L}\right), \tag{27}$$

we verify that $Q_2(x_1, L|L) = 0$. Similarly, putting $x_2 = 0$ in (25) and using the identity

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sin\left(\frac{n\pi x_1}{L}\right) = \frac{\pi\, x_1}{2L}, \tag{28}$$

one verifies the last boundary condition $Q_2(x_1, x_2 = 0|L) = 1$. Note that even though the solution $Q_2(x_1, x_2|L)$ in (25) is not manifestly symmetric under the exchange of x_1 and x_2, it is actually symmetric in x_1 and x_2 as it should be. Later in section 4, we will derive an alternative expression via the path integral method which is manifestly symmetric in x_1 and x_2.

Let us analyze the large L behavior of (25) in the situation when x_1 and x_2 are kept finite. Since $Q_2(x_1, x_2|L) = Q_2(x_1/L, x_2/L)$ is a function of only the scaled variables $z_1 = x_1/L$ and $z_2 = x_2/L$, the $L \to \infty$ limit is equivalent to taking the limits $z_1 \to 0$ and $z_2 \to 0$.

Clearly $Q_2(z_1 = 0, z_2 = 0) = 1$. It is easy to check that the first derivatives $\partial_{z_1} Q$ and $\partial_{z_2} Q$ vanish at the origin ($z_1 = 0,\ z_2 = 0$). Similarly, the second derivatives $\partial_{z_1}^2 Q_2$ and $\partial_{z_2}^2 Q_2$ also vanish at the origin. So, in a Taylor expansion of $Q_2(z_1, z_2)$ the first nonzero term is the cross derivative, indicating the following leading-order behavior as $z_1 \to 0$ and $z_2 \to 0$:

$$Q_2(z_1, z_2) = 1 - A_2 z_1 z_2 + \cdots. \tag{29}$$

The amplitude A_2 is given by

$$A_2 = \frac{\partial^2 Q_2}{\partial z_1 \partial z_2}\bigg|_{z_1=0, z_2=0} = 2\pi \left[\sum_{n=1}^{\infty} (-1)^n\, n\, \coth(n\pi) - \sum_{n=1}^{\infty} \frac{n}{\sinh(n\pi)}\right], \tag{30}$$

where the last line follows from (25). Fortunately, the sums can explicitly be carried out using some known identities [47] to yield

$$A_2 = \frac{1}{8\pi^2}[\Gamma(1/4)]^4 = 2.188\,44\ldots. \tag{31}$$

Hence for large L

$$\text{Prob}[m \leqslant L|x_1, x_2] = Q_2(x_1, x_2|L) \simeq 1 - A_2 \frac{x_1\, x_2}{L^2}, \tag{32}$$

which leads to the announced power-law tail for the probability density of the maximum m:

$$P_2(m|x_1, x_2) \simeq B_2 \frac{x_1\, x_2}{m^3}, \qquad \text{with} \quad B_2 = 2A_2 = 4.376\,88\ldots. \tag{33}$$

To compare with the $N = 1$ case, let us consider the special initial condition where both particles start from the same initial position: $x_1 = x_2 = x$. In this case, the exit probability $Q_2(x_1 = x, x_2 = x|L) = q_2(x/L)$ with the scaling function $q_2(z)$ given by

$$q_2(z) = 1 - z + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \frac{[\sinh(n\pi(1-z)) + (-1)^n \sinh(n\pi z)]}{\sinh(n\pi)} \sin(n\pi z). \tag{34}$$

One can verify that $q_2(z)$ satisfies the duality relation (19). Near $z = 0$ and $z = 1$, the scaling function has the asymptotics

$$q_2(z) \simeq \begin{cases} 1 - A_2\, z^2 & \text{as } z \to 0, \\ A_2\, (1-z)^2 & \text{as } z \to 1. \end{cases} \tag{35}$$

Comparing with the expression for the $N = 1$ case, equation (24), one finds that $q_2(z)$ has a much richer functional form. An analytic prediction for the function $q_2(z)$ is in excellent agreement with the results obtained from Monte Carlo simulations (figure 5).

4. Exact solution for all N by the path counting method

The approach based on the Laplace equation (section 3) is difficult to extend for $N > 2$. The technical problem is to find the exact solution of the Laplace equation $\nabla^2 Q_N(z_1, z_2, \ldots, z_N) = 0$ in the N-dimensional hypercube of the scaled variables $0 \leqslant z_i = x_i/L \leqslant 1$, with the prescribed boundary conditions that $Q_N = 1$ for any $z_i = 0$ and $Q_N = 0$ for any $z_i = 1$. For $N > 2$, it is not easy to find an explicit solution to this problem. In this section, we use an alternative path counting method that is physically more explicit and, in addition, allows us to write down the exact solution $Q_N(z_1, z_2, \ldots, z_N)$ for all N. This method thus provides an alternative way to solving Laplace's equation in a hypercube with the prescribed boundary conditions.

To set up the path counting method for general N, we need two basic ingredients from the single particle problem in a box $[0, L]$ with absorbing boundary conditions at the two boundaries 0 and L.

(i) The survival probability $S(x_0, t, L)$ which counts the probability that a single particle, starting at x_0 at time 0, remains inside the box $[0, L]$ during the time interval $(0, t)$.
(ii) The first passage probability density $F(x_0, t, L)$ that denotes the probability density that the particle, starting initially at x_0 ($0 \leqslant x_0 \leqslant L$), exits for the first time the box through the boundary at 0 (and not through the other boundary at L) at time t.

One can write the survival probability $S(x_0, t, L) = \int_0^L G(x, x_0, t, L)\,\mathrm{d}x$, where $G(x, x_0, t, L)$ denotes the Green's function counting the probability density that the particle reaches x at time t, starting from x_0 at time 0, while staying inside the box $[0, L]$ during time t. This Green's function can be computed (a) either by solving the diffusion equation $\partial_t G = D\partial_x^2 G$ (the diffusion constant is $D = 1/2$ for our choice of the noise term) with the absorbing boundary condition $G(x = 0, x_0, t, L) = G(x = L, x_0, t, L) = 0$ and the initial condition $G(x, x_0, 0, L) = \delta(x - x_0)$, or (b) by the path integral method. The resulting Green's function admits the following representation:

$$G(x, x_0, t, L) = \frac{2}{L}\sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{n\pi x_0}{L}\right) \mathrm{e}^{-\frac{n^2\pi^2}{2L^2}t}. \tag{36}$$

The survival probability, after integrating over the final position x, is

$$S(x_0, t, L) = \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{[1 - (-1)^n]}{n} \sin\left(\frac{n\pi x_0}{L}\right) \mathrm{e}^{-\frac{n^2\pi^2}{2L^2}t}. \tag{37}$$

To calculate the first passage probability density $F(x_0, t)$ through 0 at time t, we note that this just counts the flux of particles going out of the box through 0 at time t. The flux through a point x in the positive direction is the probability current $-D\partial_x G$ (with $D = 1/2$). Hence, the first-passage probability through the origin that counts the flux through the origin in the negative direction is simply $F(x_0, t) = D\partial_x G|_{x=0}$. Using G from (36), we get

$$F(x_0, t) = \frac{\pi}{L^2}\sum_{n=1}^{\infty} n \sin\left(\frac{n\pi x_0}{L}\right) \mathrm{e}^{-\frac{n^2\pi^2}{2L^2}t}. \tag{38}$$

Armed with these two ingredients from the single particle problem, we are now ready to compute the exit probability $Q_N(\vec{x}|L)$ for the N-particle problem. Consider first the event shown in figure 2, where one of the particles (say the ith one) exits the box for the first time between times t_s and $t_s + \mathrm{d}t_s$ while the $N - 1$ other particles stay inside $[0, L]$ till this time t_s. Clearly, the probability for this event, using independence of walkers, is given by the product $F(x_i, t_s) \prod_{j\neq i} S(x_j, t_s)\,\mathrm{d}t_s$. Now, the particle that hits 0 (whose label is i) can, in fact, be any one of the N particles. Hence we have to sum over the hitting index i from 1 to N. In addition, this event can occur at any time t_s, so we need to integrate over t_s. This path counting method then gives the following net contribution to the exit probability:

$$Q_N(\vec{x}|L) = \sum_{i=1}^{N} \int_0^{\infty} F(x_i, t_s) \prod_{j\neq i} S(x_j, t_s)\,\mathrm{d}t_s. \tag{39}$$

Substituting the results for $S(x,t)$ and $F(x,t)$ respectively from equations (37) and (38), integrating over t_s and setting $z_i = x_i/L$ gives our main exact result, valid for all N:

$$Q_N(\vec{x}|L) = Q_N(z_1, z_2, \ldots, z_N) = \left(\frac{2}{\pi}\right)^N \sum_{\{n_i\}} \frac{a(n_1, n_2, \ldots, n_N)}{n_1^2 + n_2^2 + \cdots + n_N^2} \prod_{k=1}^{N} \frac{\sin(n_k\pi z_k)}{n_k}, \tag{40}$$

where each index $n_i = 1, 2, \ldots,$ for all $1 \leqslant i \leqslant N$ and

$$a(n_1, n_2, \ldots, a_N) = \sum_{i=1}^{N} n_i^2 \prod_{j\neq i} [1 - (-1)^{n_j}]. \tag{41}$$

Equation (40) is an exact solution of Laplace's equation in the N-dimensional hybercube $0 \leqslant z_i \leqslant 1$ with the prescribed boundary conditions which has a virtue of being manifestly symmetric with respect to the interchange of the coordinates z_i's.

To extract the behavior in the limit of large L, with x_i's fixed, we need to take the limit $z_i \to 0$ in equation (40). This limiting behavior reads

$$Q_N(z_1, z_2, \ldots, z_N) = 1 - A_N z_1 z_2 \cdots z_N + O\left(z_i^2\right). \tag{42}$$

The coefficient A_N can be computed by taking the derivatives of equation (40) with respect to each z_i and then putting $z_i = 0$ for all i. This gives

$$A_N = -N\, 2^N \lim_{z_i \to 0} \sum_{\{n_i\}} \frac{n_1^2 \cos(n_1 \pi z_1)}{n_1^2 + n_2^2 + \cdots + n_N^2} \prod_{j=2}^{N} [1 - (-1)^{n_j}] \cos(n_j \pi z_j). \tag{43}$$

Noting that $Q_N(\vec{x}|L) = Q_N(z_1, z_2, \ldots, z_N)$ is precisely the cumulative distribution of the maximum m, one thus gets

$$\text{Prob}[m \leqslant L|\vec{x}] \simeq 1 - A_N \frac{x_1 x_2 \cdots x_N}{L^N} \quad \text{as } L \to \infty. \tag{44}$$

Differentiating with respect to L and putting $L = m$ gives the exact power-law tail of the pdf of the maximum m in equation (4) with the prefactor B_N which is given by the formal sum

$$B_N = N A_N = -N^2\, 2^N \lim_{z_i \to 0} \sum_{\{n_i\}} \frac{n_1^2 \cos(n_1 \pi z_1)}{n_1^2 + n_2^2 + \cdots + n_N^2} \prod_{j=2}^{N} [1 - (-1)^{n_j}] \cos(n_j \pi z_j), \tag{45}$$

where each index n_i runs over all positive integers.

The formal sum in equation (45) can explicitly be evaluated for $N = 1$ and $N = 2$. For $N = 1$,

$$B_1 = -2 \lim_{z_1 \to 0} \sum_{n_1 = 1, 2, \ldots} \cos(n_1 \pi z_1) = 1, \tag{46}$$

which is in agreement with already known results. For $N = 2$, equation (45) gives

$$\begin{aligned} B_2 &= -16 \lim_{z_i \to 0} \sum_{n_1, n_2} \frac{n_1^2 \cos(n_1 \pi z_1)}{n_1^2 + n_2^2} [1 - (-1)^{n_2}] \cos(n_2 \pi z_2) \\ &= \frac{1}{4\pi^2} \left[\Gamma\left(\frac{1}{4}\right)\right]^4 = 4.376\,88\ldots. \end{aligned} \tag{47}$$

In appendix A, we show how to compute the above sum explicitly.

For $N > 2$, we have not deduced explicit expressions for the sum in (45). However, one can reduce it to a simpler form where the sum is rapidly convergent and can then be evaluated by Mathematica. As an example, for $N = 3$,

$$B_3 = -6 + \frac{9}{8\pi^2} \Gamma^4(1/4) + 72\pi \sum_{n_1, n_2} \frac{n_1^2 + n_2^2\, (-1)^{n_2}}{\sqrt{n_1^2 + n_2^2}\, \sinh\left(\pi \sqrt{n_1^2 + n_2^2}\right)} = 15.3369\ldots. \tag{48}$$

In the limit of large N, one can evaluate the formal sum (see appendix B) to obtain the limiting behavior

$$B_N \simeq N \left[\frac{4}{\pi} \ln(N)\right]^{N/2} \tag{49}$$

in perfect agreement with the heuristic result in equation (15).

Special initial condition. Finally, let us consider the special initial condition when all the particles start from the same point: $x_1 = x_2 = \cdots = x$, where $0 \leqslant x \leqslant L$. In this case,

the exit probability in equation (40) reduces to a function of one scaled variable $z = x/L$: $Q_N(x, x, \ldots, x|L) = q_N(z)$ given by

$$q_N(z) = N \left(\frac{2}{\pi}\right)^N \sum_{\{n_i\}} \frac{n_1 \sin(n_1 \pi z)}{\left(n_1^2 + n_2^2 + \cdots + n_N^2\right)} \prod_{j\neq 1} [1 - (-1)^{n_j}] \frac{\sin(n_j \pi z)}{n_j}. \tag{50}$$

One can check that $q_N(z)$ satisfies the duality relation $q_N(z) + q_N(1 - z) = 1$ and thus $q_N(1/2) = 1/2$ for all N.

Performing this multiple sum directly by Mathematica is difficult as it converges slowly. To circumvent this problem, we first perform the sum over n_1 in equation (50) using the following identity [47]:

$$\sum_{k=1}^{\infty} \frac{k \sin(k\pi x)}{k^2 + a^2} = \frac{\pi}{2} \frac{\sinh(\pi(x - a))}{\sinh(\pi a)}. \tag{51}$$

This gives

$$q_N(z) = N \left(\frac{4}{\pi}\right)^{N-1} \sum_{n_2, n_3, \ldots, n_N \to \text{odd}} \frac{\sinh\left[\pi(1 - z)\sqrt{n_2^2 + n_3^2 + \cdots + n_N^2}\right]}{\sinh\left[\pi\sqrt{n_2^2 + n_3^2 + \cdots + n_N^2}\right]} \prod_{j=2}^{N} \frac{\sin(n_j \pi z)}{n_j}. \tag{52}$$

The multiple sum in equation (52) is now rapidly convergent and can easily be evaluated by Mathematica. In figure 2 we plot this function $q_N(z)$ for $N = 2$ and $N = 3$. For $N = 2$, it of course coincides with the earlier expression (34) obtained via Laplace's method in the previous section.

Average maximum. For a fixed identical initial position of all particles $x_i = x$, another interesting question is: how does the average maximum (till the stopping time t_s) depend on N? For $N = 1$, the average is infinite but for all $N > 1$ it is finite. However, does the average maximum for $N > 1$ increase or decrease as the number of walkers N increases? The answer to this question is not intuitively obvious. However, knowing the function $q_N(z)$, one can compute the average maximum in the following way. We have $\text{Prob}[m \leqslant L|x] = q_N(x/L)$, where $q_N(z)$ is given in equation (52). Therefore, the pdf of m reads

$$P_N(m|x) = -\frac{x}{m^2} q_N'\left(\frac{x}{m}\right); \qquad m \geqslant x, \tag{53}$$

where $q_N'(z) = \mathrm{d}q_N(z)/\mathrm{d}z$. The first moment is then given by, for all $x \geqslant 0$ and $N > 1$,

$$\langle m \rangle = -\int_x^{\infty} \frac{x}{m} q_N'\left(\frac{x}{m}\right) \mathrm{d}m = C_N\, x \tag{54}$$

with the prefactor

$$C_N = -\int_0^1 \frac{q_N'(z)}{z} \mathrm{d}z. \tag{55}$$

Thus, the average maximum, for $N > 1$, is proportional to x for all x and the proportionality constant C_N is given by equation (55). We were unable to carry out the integral in equation (55) in closed form. However, it is clear that as $N \to \infty$, using $q_N(z) = \theta(1/2 - z)$, one gets $C_N \to 2$. On the other hand C_N diverges as $N \to 1$. Thus, C_N decreases when N increases. These results are supported by Monte Carlo simulations (figure 8).

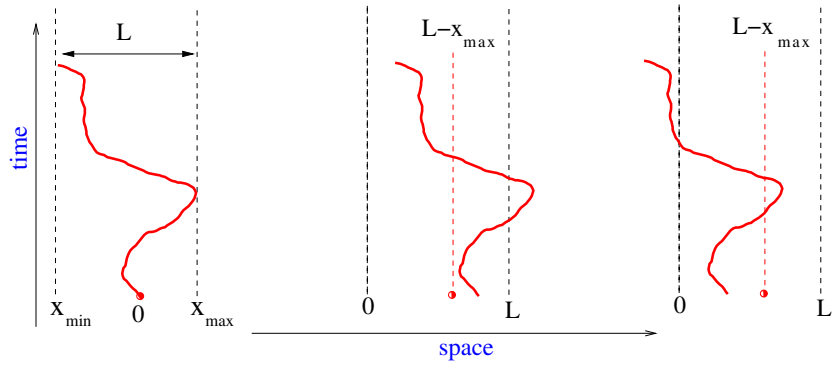

Figure 4. The illustration of the algorithm to compute the exit probability $q_N(z)$ for a single ($N = 1$) Brownian motion.

5. Numerical simulations

Our analytical result for the exit probability through the origin $q_N(z)$, when all particles start at the same scaled position $0 \leqslant z = x/L \leqslant 1$, is tested by Monte Carlo simulations. For a fixed box of size L, the exit probability for the N-particle problem can efficiently be computed using a method proposed in [43]. Naively, to compute $q_N(z)$, one would first fix the starting point $0 \leqslant z \leqslant 1$ of all the Brownian motions and then generate different realizations of the process and compute the fraction of realizations where the first exit occurs through the origin. One would then repeat the procedure for each value of $0 \leqslant z \leqslant 1$ (with an appropriate bin size) in order to compute the full function $q_N(z)$ over the range $z \in [0, 1]$. Instead of repeating the simulation for each starting point z, it turns out to be more efficient to follow a different algorithm described briefly below.

For illustration, we take the example of just one Brownian motion ($N = 1$). The method is easily generalized for all N. We start the Brownian motion at the origin; let it evolve in time and record the maximal ($x_{\max}$) and the minimal ($x_{\min}$) position reached by the walker up to the time t. The process is halted when $x_{\max} - x_{\min} \geqslant L$ for the first time (see figure 4). Keeping $x_{\min}$ and $x_{\max}$ fixed, we now horizontally slide the whole configuration thereby changing the starting point (see the second and the third panel of figure 4). Measuring all distances with respect to $x_{\min}$, it is then clear that this configuration contributes 1 to $q_1(z)$ for $z \in [0, 1 - x_{\max}/L]$ and 0 for $z \in [1 - x_{\max}/L, 1]$. So, for this configuration, we just record the number $y = 1 - x_{\max}/L$ and the fact that it contributes $\theta[y - z]$ to $q_1(z)$. We then repeat this procedure for another configuration starting at 0, do the sliding and record the value of y and the associated $\theta(y - z)$. We repeat this, say, for N_s number of samples. To sum all the contributions, we first sort the values of y associated with the configurations in increasing order. Let $\{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{N_s}\}$ denote the sorted values. Then, we assign $q_1(\tilde{y}_1) = 1$, $q_1(\tilde{y}_2) = 1 - 1/N_s$ and in general $q_1(\tilde{y}_i) = 1 - (i - 1)/N_s$. This generates the full curve $q_1(z)$.

Thus, this algorithm has two advantages: (i) it does not require to repeat the simulations for each value of z and (ii) it does not require any specific choice of bin sizes. This clearly makes the simulation much faster. The results of our simulations for $N = 2$, $N = 3$ and $N = 6$ are plotted in figure 5. The agreement with the analytical prediction is excellent. For all N, $q_N(z)$ satisfies the duality relation: $q_N(z) + q_N(1 - z) = 1$.

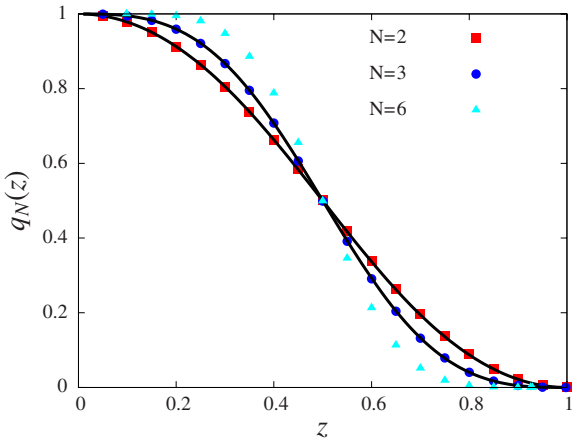

Figure 5. The exit probability $q_N(z)$ through the origin obtained from Monte Carlo simulations for $N = 2$ (red squares), $N = 3$ (deep blue circles) and $N = 8$ (blue triangles). The numerical data are compared to the analytical results (solid lines) for $N = 2$ and $N = 3$ (obtained from equation (52) respectively for $N = 2$ and $N = 3$).

In [43] we had studied the exit probability from a box $[0, L]$ through the upper boundary L (rather than the lower boundary 0 considered here) for a single particle whose motion is described by a generic self-affine stochastic process. This probability was called the 'hitting probability' in [43]. In the present paper, we are considering the complimentary event of exiting the box $[0, L]$ through the lower boundary 0. In the notation of the present paper, the hitting probability (of the boundary L) would correspond to $1 - q_N(z)$.[3] For a generic self-affine process, it was shown in [43] that the exit probability should have a power-law behavior close to the origin $1 - q(z) \sim z^{\phi}$ with $\phi = \theta/H$ as mentioned in the introduction. In addition, it was observed in [43] that for many processes (but not all), once we know the exponent ϕ, the full function $q(z)$ over the range $z \in [0, 1]$ is described by a universal one-parameter (parametrized by ϕ) form [43]

$$q_\phi(z) = 1 - I_z(\phi, \phi) = 1 - \frac{\Gamma(2\phi)}{\Gamma(\phi)^2} \int_0^z [u(1-u)]^{\phi-1}\, \mathrm{d}u. \tag{56}$$

The function $I_z(\phi, \phi)$ is the incomplete regularized beta function. In our present problem, we have seen in equation (42) that for small z, $q_N(z) \sim 1 - A_N z^N$ indicating $\phi = N$. It is then natural to investigate if our result for the full function $q_N(z)$ can be re-expressed as the universal functional form in equation (56) with $\phi = N$. Interestingly, the answer is no, as it is clearly shown in figure 6, thus providing us with a counterexample.

We have also computed the prefactor B_N numerically from the Monte Carlo simulations up to $N = 7$. The results are shown in figure 7 by squares. For $N = 1$, $N = 2$ and $N = 3$, they agree with our exact analytical predictions $B_1 = 1$, $B_2 = 4.376\,88\ldots$ and $B_3 = 15.3369\ldots$ as discussed in section 4. Had the $q_N(z)$ be described by the universal function in equation (56) with $\phi = N$, one would get from the small z expansion in equation (56), $q_N(z) \to 1 - A_N z^N$ with $A_N = \Gamma(2N)/N\Gamma^2(N)$. This would predict the prefactor $B_N = NA_N = \Gamma(2N)/\Gamma^2(N)$. In figure 7, this prediction from the universal curve is shown by the dashed line. Clearly, it does not match the simulation results, confirming once more that $q_N(z)$ is not described by

[3] Please note that in [43] the notation $q(z)$ was used for the hitting probability (for exiting through L), but here we use $q_N(z)$ to denote the complimentary event of exiting the box through 0. We apologize if this causes any confusion.

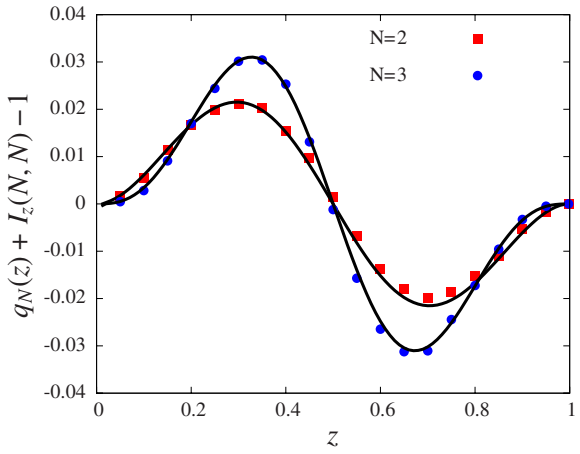

Figure 6. Difference between the exit probability for $N = 2$ and $N = 3$ and the universal function in equation (56) with $\phi = N$. The solid lines show the analytical results from equation (52) with $N = 2$ and $N = 3$. The symbols show the numerical results for 10^5 realizations of $N = 2$ (squares) and $N = 3$ (circles) Brownian motions starting from the same initial position in a box of size $L = 1000$.

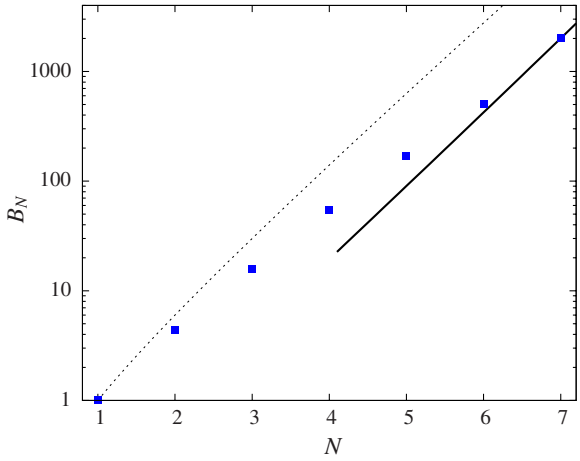

Figure 7. Dashed line: $B_N = \Gamma(2N)/\Gamma^2(N)$ from the universal function in equation (56). Solid line: asymptotic behavior for large N given in equation (49) and squares: direct simulations of N Brownian motions starting from the same position in a box of size $L = 1000$. Averages are performed over 10^6 samples.

the universal function in equation (56). We have also plotted the exact asymptotic prediction of B_N in equation (49) as a solid line for comparison. While it is difficult to extract the small z behavior of $q_N(z)$ and hence B_N for larger values of N, we note that the asymptotic large N behavior is already approached for $N = 7$.

Finally we have also computed the average maximum (till the stopping time t_s) of N Brownian motions starting from the same initial positions $x > 0$. We verified that for all $N > 1$, the average maximum exists, and it is proportional to x, as predicted analytically in

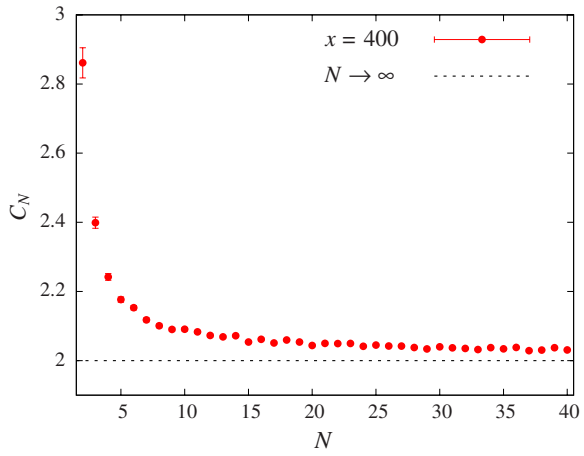

Figure 8. Monte Carlo results for the average maximum of N Brownian motions till the first passage time through the origin, all starting at the same initial position $x > 0$. The coefficient $C_N = \langle m \rangle / x$, see equation (55), is plotted versus N. Averages are performed over 10^3 realizations.

equation (54). In figure 8 we plot the amplitude C_N, given analytically in equation (55), as a function of N. As predicted, C_N approaches 2 for large N.

6. Conclusion

To summarize, we have presented an exact solution for the probability distribution of the maximum m of a set of N independent Brownian motions starting at the initial positions $\vec{x} \equiv \{x_1, x_2, \ldots, x_N\}$ on the positive half-axis and the process terminating when any one of the walkers crosses the origin. We have shown that for large m, the pdf of m decays as a power law, $P_N(m|\vec{x}) \simeq B_N(x_1 x_2 \cdots x_N)/m^{N+1}$, where the prefactor B_N has an interesting N dependence. For a fixed $N > 1$, integer moments of m up to order $(N-1)$ are finite, while all higher integer moments are infinite. The cumulative distribution of this maximum also provides an exact solution to the first-exit probability through the origin (rather than through L) of N walkers from a box $[0, L]$. Incidentally, our path counting method also provides an exact solution to the N-dimensional Laplace's equation $\nabla^2 Q_N = 0$ in a hypercube $[0, L]^N$ with the boundary conditions $Q_N = 1$ on any face of the hypercube passing through the origin and $Q_N = 0$ on the rest of the faces. Monte Carlo simulations confirm our analytical results.

This work raises some interesting open questions. We have focused only on the maximum m till the stopping time t_s of N independent walkers. Another interesting observable is not just the actual value of the maximum, but the time t_m at which this maximum occurs before the stopping time t_s. This random variable has recently been studied in a number of contexts. For a stochastic process over a fixed time interval $[0, t]$, the distribution of the time t_m has been computed for a variety of Brownian paths, such as a free Brownian motion, Brownian bridges, Brownian excursions and Brownian meanders, using the path integral method [7] and also by an alternative functional renormalization group method [18]. The distribution of t_m was also computed exactly for the random acceleration process which is a non-Markov process [50]. It has also been computed both for independent Brownian walkers [26, 27] and very recently for vicious walkers [51]. On the other hand, when the process stops at a random stopping time t_s, where for instance t_s is the first time a walker hits the origin, the distribution of t_m has been

computed for a single Brownian motion ($N = 1$), where it is already nontrivial [52]. It would be interesting to extend the results of [52] to the case of $N > 1$ independent Brownian motions.

Another interesting challenging problem would be to compute the distribution of m as well as that of t_m for a set of vicious walkers till the stopping time t_s when the walker closest to the origin crosses the origin for the first time.

Acknowledgments

PLK thanks Laboratoire de Physique Théorique et Modèles Statistiques for hospitality and the NSF grant CCF-0829541 for support. AR acknowledges support by ANR grant 09-BLAN-0097-02. The authors thank A Schweiger for useful discussions.

Appendix A. Evaluation of B_2

To evaluate the sum in equation (47), we first split the sum into two parts: $B_2 = I_1 + I_2$ with

$$I_1 = -16 \lim_{z_1,z_2\to 0} \sum_{n_1,n_2} \frac{n_1^2}{n_1^2+n_2^2} \cos(n_1\pi z_1)\cos(n_2\pi z_2) \tag{A.1}$$

$$I_2 = 16 \lim_{z_1,z_2\to 0} \sum_{n_1,n_2} \frac{(-1)^{n_2} n_1^2}{n_1^2+n_2^2} \cos(n_1\pi z_1)\cos(n_2\pi z_2), \tag{A.2}$$

where all the sums run over positive integers.

Let us first evaluate I_1. Due to the symmetry we can rewrite I_1 as

$$\begin{aligned} I_1 &= -8 \lim_{z_1,z_2\to 0} \sum_{n_1,n_2} \frac{n_1^2+n_2^2}{n_1^2+n_2^2} \cos(n_1\pi z_1)\cos(n_2\pi z_2) \\ &= -8 \lim_{z_1,z_2\to 0} \sum_{n_1,n_2} \cos(n_1\pi z_1)\cos(n_2\pi z_2) = -2. \end{aligned} \tag{A.3}$$

In the last step we have used the identity, $\sum_{n\geqslant 1}\cos(n\pi z) = -1/2$.

To evaluate I_2 we use another standard identity [47]

$$\sum_{k=1}^{\infty} \frac{(-1)^k \cos(k\pi z)}{k^2+a^2} = \frac{\pi}{2a}\frac{\cosh(a\pi z)}{\sinh(a\pi)} - \frac{1}{2a^2}. \tag{A.4}$$

We now sum over n_2 in equation (A.2) using the above identity. This gives

$$\begin{aligned} I_2 &= -8 \lim_{z_1\to 0} \sum_{n_1=1}^{\infty} \cos(n_1\pi z_1) + 8\pi \lim_{z_1,z_2\to 0} \sum_{n_1=1}^{\infty} \frac{n_1}{\sinh(n_1\pi)} \cos(n_1\pi z_1)\cos(n_1\pi z_2) \\ &= 4 + 8\pi \sum_{n_1=1}^{\infty} \frac{n_1}{\sinh(n_1\pi)}. \end{aligned} \tag{A.5}$$

The remaining sum in (A.5) can explicitly be evaluated using the identity [47]

$$\sum_{n=1}^{\infty} \frac{n}{\sinh(n\pi)} = \frac{1}{32\pi^3}\Gamma^4\left(\frac{1}{4}\right) - \frac{1}{4\pi}. \tag{A.6}$$

Adding I_1 and I_2 we arrive at the announced expression (5) for B_2.

Appendix B. Evaluation of B_N for large N

Here, we show that B_N in equation (45) behaves, to leading order for large N, as

$$B_N \simeq N\left[\frac{4}{\pi}\ln(N)\right]^{N/2}. \tag{B.1}$$

First we rewrite the denominator in equation (45) using the integral representation

$$\frac{1}{n_1^2+n_2^2+\cdots+n_N^2} = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t\,(n_1^2+n_2^2+\cdots+n_N^2)}. \tag{B.2}$$

Using this representation, one can then decouple the sums over different indices n_i in equation (45) giving

$$B_N = -N^2\,2^N \lim_{z_i\to 0}\int_0^\infty \mathrm{d}t \sum_{n_1=1}^\infty n_1^2\cos(n_1\pi z_1)\,\mathrm{e}^{-n_1^2 t}\prod_{j\neq 1}\sum_{n_j=1}^\infty [1-(-1)^{n_j}]\cos(n_j\pi z_j)\,\mathrm{e}^{-tn_j^2}. \tag{B.3}$$

If we now take the limits $z_j \to 0$ inside the sums, each of the sums is convergent. However, it is easy to check that the integrand, as a function of t, diverges as $\sim t^{1+N/2}$ as $t \to 0$. Thus, the integral is longer convergent. To circumvent this difficulty, one can use a standard regularization scheme used often in evaluating the Madelung constant in the context of lattice sums in crystals [48] whereby we introduce a parameter s and rewrite equation (B.3) as, upon taking the limits $z_j \to 0$ inside the sums,

$$B_N(s) = -\frac{N^2\,2^N}{\Gamma(s)}\int_0^\infty \mathrm{d}t\, t^{s-1}\sum_{n_1=1}^\infty n_1^2\,\mathrm{e}^{-n_1^2 t}\left[\sum_{n_j=1}^\infty [1-(-1)^{n_j}]\,\mathrm{e}^{-n_j^2 t}\right]^{N-1}. \tag{B.4}$$

Note that this integral is convergent for all $s > (1+N/2)$. The idea is to first evaluate $B_N(s)$ for large N with $s > (1+N/2)$ and then analytically continue this result to $s \to 1$ to evaluate $B_N = B_N(s\to 1)$.

Let us next define the function

$$g(t) = 4\sqrt{\frac{t}{\pi}}\sum_{k=0}^\infty \mathrm{e}^{-(2k+1)^2 t}. \tag{B.5}$$

Up to the factor $\sqrt{t/\pi}$, this function $g(t)$ can be expressed in terms of standard Jacobi theta functions [49]. Then one can rewrite equation (B.4) as

$$B_N(s) = -\frac{2N^2(\sqrt{\pi})^{N-1}}{\Gamma(s)}\int_0^\infty \mathrm{d}t\, t^{s-(N+1)/2}\,[g(t)]^{N-1}\left[\sum_{n=1}^\infty n^2\mathrm{e}^{-n^2 t}\right]. \tag{B.6}$$

To evaluate this integral for large N, we need to know how the function $[g(t)]^{N-1}$ behaves for large N.

Let us first focus on the function $g(t)$ in equation (B.5). Clearly, for large t, the dominant contribution comes from the $k=0$ term in the sum and hence $g(t) \sim \sqrt{t}\mathrm{e}^{-t}$ as $t\to\infty$. In contrast, the opposite limit $t \to 0$ is more tricky. To derive its behavior as $t\to 0$, we first use the following Jacobi identity [49]:

$$1+2\sum_{k=1}^\infty (-1)^k \mathrm{e}^{-k^2 z} = 2\sqrt{\frac{\pi}{z}}\sum_{k=0}^\infty \mathrm{e}^{-(2k+1)^2\pi^2/4z} \tag{B.7}$$

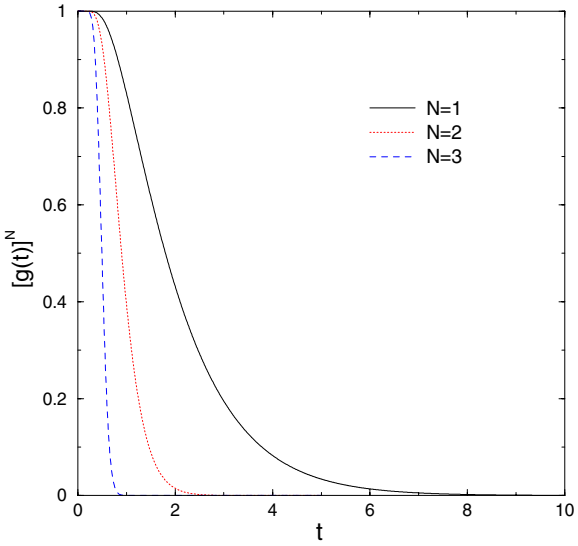

Figure B1. The function $[g(t)]^N$ plotted as a function of t for $N = 1$ (black solid), $N = 5$ (red dotted) and $N = 50$ (blue dashed). As N becomes large, $[g(t)]^N$ almost approaches to a step function.

to rewrite the function $g(t)$ (upon identifying $t = \pi^2/4z$) as

$$g(t) = 1 + 2\sum_{k=1}^{\infty}(-1)^k\, \mathrm{e}^{-\pi^2 k^2/4t}. \tag{B.8}$$

This representation of $g(t)$ is more amenable to the asymptotic analysis in the $t \to 0$ limit. We obtain from equation (B.8) the leading behavior of $g(t)$ as $t \to 0$

$$g(t) \simeq 1 - 2\,\mathrm{e}^{-\pi^2/4t} + \cdots. \tag{B.9}$$

Thus, $g(t)$ has an essential singularity at $t = 0$ and it approaches to 1 as $t \to 0$ in an extremely flat way. Thus, the function $g(t)$ starts at $g(0) = 1$ stays flat for a while and then decreases exponentially fast to 0 as t increases (see figure B1).

Next consider the function $[g(t)]^N$ that appears in the integral in equation (B.6). As N increases, the function $[g(t)]^N$ almost approaches a step function (see figure B1)

$$[g(t)]^N \to \theta(t^*(N) - t), \tag{B.10}$$

where the characteristic scale $t^*(N)$ decreases very slowly with increasing N. One can easily estimate $t^*(N)$ for large N from the asymptotic behavior in equation (B.9). For small t, one finds, to leading order for large N,

$$[g(t)]^N \sim \left[1 - 2\,\mathrm{e}^{-\pi^2/4t}\right]^N \simeq \exp\left[-2N\,\mathrm{e}^{-\pi^2/4t}\right]. \tag{B.11}$$

Thus, as N increases, it approaches to 0 rapidly for all $t > t^*$, where $2N\mathrm{e}^{-\pi^2/4t^*} \approx 1$. This provides an estimate of $t^*(N)$ which, to leading order for large N, reads

$$t^*(N) \simeq \frac{\pi^2}{4\ln(N)}. \tag{B.12}$$

Therefore, for large N, using equation (B.10), we can cut off the upper limit of the integral in equation (B.6) at $t = t^*(N)$ and replace $[g(t)]^{N-1}$ by 1 over the interval $t \in [0, t^*(N)]$. Furthermore, over this small interval $t \in [0, t^*(N)]$, one can replace the function $\sum_{n=1}^{\infty} n^2\,\mathrm{e}^{-n^2 t}$

by its small t behavior $\simeq \sqrt{\pi}/4t^{3/2}$. Substituting these results into equation (B.6) then yields, for large N,

$$B_N(s) \simeq -\frac{N^2\,\pi^{N/2}}{2\Gamma(s)} \int_0^{t^*(N)} \mathrm{d}t\; t^{s-(N+4)/2} \simeq -\frac{N^2\pi^{N/2}}{2\Gamma(s)}\,\frac{[t^*(N)]^{s-N/2-1}}{(s-N/2-1)}. \tag{B.13}$$

In deriving this result, we have assumed $s > (1 + N/2)$. After obtaining this large N formula for $B_N(s)$, we can now analytically continue it to $s \to 1$ which finally yields

$$B_N = B_N(s \to 1) = N\,\pi^{N/2}\,[t^*(N)]^{-N/2}. \tag{B.14}$$

Upon using the expression for $t^*(N)$ from equation (B.12) gives the final large N expression for B_N in equation (B.1).

References

[1] Chung K L 1975 *Bull. Am. Math. Soc.* **81** 742
[2] Kennedy D P 1976 *J. Appl. Probab.* **13** 371
[3] Durrett R T and Iglehart D L 1977 *Ann. Probab.* **5** 130
[4] Darling D A 1983 *Ann. Probab.* **11** 803
[5] Pitman J and Yor M 1999 *Electron. J. Probab.* **4** 15
[6] For a brief review see Finch S R Variants of Brownian motion, 2004 Online at http://algo.inria.fr/csolve/br.pdf
[7] Majumdar S N, Randon-Furling J, Kearney M J and Yor M 2008 *J. Phys. A: Math. Theor.* **41** 365005
[8] Kolmogorov A N 1941 *Ann. Math. Stat.* **12** 461
[9] Smirnov N V 1939 *Mat. Sb.* **6** 3 (in Russian)
[10] Pollaczek F 1952 *Comptes Rendus* **234** 2334
[11] Spitzer F 1956 *Trans. Am. Math. Soc.* **82** 323
Spitzer F 1957 *Duke Math. J.* **24** 327–43
[12] Coffman E G, Flajolet P, Flato L and Hofri M 1998 *Probab. Eng. Inform. Sci.* **12** 373–86
[13] Comtet A and Majumdar S N 2005 *J. Stat. Mech.* P06013
Majumdar S N, Comtet A and Ziff R M 2006 *J. Stat. Phys.* **122** 833
For a review see Majumdar S N 2009 arXiv:0912:2586
[14] Ziff R M, Majumdar S N and Comtet A 2007 *J. Phys. C: Condens. Matter* **19** 065102
Ziff R M, Majumdar S N and Comtet A 2009 *J. Chem. Phys.* **130** 204104
[15] Raychowdhuri S, Cranston M, Przybyla C and Shapir Y 2001 *Phys. Rev. Lett.* **87** 136101
[16] Majumdar S N and Comtet A 2004 *Phys. Rev. Lett.* **92** 225501
Majumdar S N and Comtet A 2005 *J. Stat. Phys.* **119** 777
[17] Schehr G and Majumdar S N 2006 *Phys. Rev.* E **73** 056103
Gyorgyi G *et al* 2007 *Phys. Rev.* E **75** 021123
Burkhardt T W *et al* 2007 *Phys. Rev.* E **76** 041119
Rambeau J and Schehr G 2009 *J. Stat. Mech.* P09004
[18] Schehr G and Le Doussal P 2010 *J. Stat. Mech.* P01009
[19] Carmi S, Turgeman L and Barkai E 2010 arXiv:1004.0943
[20] Tejedor V *et al* 2010 arXiv:1001.4412
[21] Schehr G, Majumdar S N, Comtet A and Randon-Furling J 2008 *Phys. Rev. Lett.* **101** 150601
[22] Katori M, Izumi M and Kobayashi N 2008 *J. Stat. Phys.* **131** 1067
Kobayashi N, Izumi M and Katori M 2008 *Phys. Rev.* E **78** 051102
[23] Feierl T 2009 *Proc. of IOWA 2009 (Lecture Notes in Computer Science* vol 5874*)* (Berlin: Springer)
[24] Nadal C and Majumdar S N 2009 *Phys. Rev.* E **79** 061117
[25] Johansson K 2002 *Probab. Theory Relat. Fields* **123** 225
Katori M and Tanemura H 2004 *J. Math. Phys.* **45** 3058
Ferrari P 2008 *Lecture Notes of Beg-Rohu Summer School* Online at http://ipht.cea.fr/Meetings/BehRohu2008/
[26] Randon-Furling J, Majumdar S N and Comtet A 2009 *Phys. Rev. Lett.* **103** 140602
[27] Majumdar S N, Comtet A and Randon-Furling J 2010 *J. Stat. Phys.* **138** 955
[28] Majumdar S N and Krapivsky P L 2003 *Physica* A **318** 161
[29] Majumdar S N 2005 Brownian functionals in physics and computer science *Curr. Sci.* **89** 2076 (arXiv:/cond-mat/0510064)
[30] Tachiya M 1983 *Radiat. Phys. Chem.* **21** 167

[31] Blumen A, Zumofen G and Klafter J 1984 *Phys. Rev.* B **30** 5379
[32] Burlatsky S F and Ovchinikov A A 1987 *Sov. Phys.—JETP* **65** 908
[33] Redner S and Krapivsky P L 1999 *Am. J. Phys.* **67** 1277
[34] Dhar D and Ramaswamy R 1989 *Phys. Rev. Lett.* **63** 1659
[35] Kearney M J and Majumdar S N 2005 *J. Phys. A: Math. Gen.* **38** 4097
[36] Kearney M J 2004 *J. Phys. A: Math. Gen.* **37** 8421
[37] Feller W 1968 *An Introduction to Probability Theory and its Applications* (New York: Wiley)
[38] Bramson M and Lebowitz J L 1988 *Phys. Rev. Lett.* **61** 2397
[39] Bray A J and Blythe R A 2002 *Phys. Rev. Lett.* **89** 150601
Blythe R A and Bray A J 2003 *Phys. Rev.* E **67** 041101
[40] Oshanin G, Benichou O, Coppey M and Moreau M 2002 *Phys. Rev.* E **66** 060101
[41] Bray A J, Majumdar S N and Blythe R A 2003 *Phys. Rev.* E **67** 060102
Moreau M *et al* 2003 *Phys. Rev.* E **67** 045104
[42] Yuste S B *et al* 2008 *Phys. Rev.* E **78** 021105
[43] Majumdar S N, Rosso A and Zoia A 2010 *Phys. Rev. Lett.* **104** 020602
[44] Majumdar S N 1999 Persistence in nonequilibrium systems *Curr. Sci.* **77** 370 (arXiv:cond-mat/9907407)
[45] Redner S 2001 *A Guide to First-Passage Processes* (Cambridge: Cambridge University Press)
[46] Lebedev N N, Skalskaya I P and Uflyand Y S 1965 *Problems of Mathematical Physics* (Englewood Cliffs, NJ: Prentice-Hall)
[47] Prudnikov A P, Brychkov Yu A and Marichev O I 1986 *Integrals and Series* vol 1 (Amsterdam: Overseas Publishers Association)
[48] Borwein D, Borwein J M and Taylor K F 1985 *J. Math. Phys.* **26** 2999
[49] Gradshteyn I S and Ryzhik I M 1980 *Tables of Integrals, Series and Products* (New York: Academic)
[50] Majumdar S N, Rosso A and Zoia A 2010 *J. Phys. A: Math. Theor.* **43** 115001
[51] Rambeau J and Schehr G 2010 arXiv:1004:0141
[52] Randon-Furling J and Majumdar S N 2007 *J. Stat. Mech.* P10008