

PAPER: CLASSICAL STATISTICAL MECHANICS, EQUILIBRIUM AND NON-EQUILIBRIUM

Mean perimeter and mean area of the convex hull over planar random walks

To cite this article: Denis S Grebenkov *et al* *J. Stat. Mech.* (2017) 103203

View the [article online](#) for updates and enhancements.

Related content

- [Record statistics of a strongly correlated time series: random walks and Lévy flights](#)
Claude Godrèche, Satya N Majumdar and Grégory Schehr
- [The convex hull for a random acceleration process in two dimensions](#)
Alexis Reymbaut, Satya N Majumdar and Alberto Rosso
- [Survival probability of random walks and Lévy flights on a semi-infinite line](#)
Satya N Majumdar, Philippe Mounaix and Grégory Schehr

PAPER: Classical statistical mechanics, equilibrium and non-equilibrium

Mean perimeter and mean area of the convex hull over planar random walks

Denis S Grebenkov,^{1,2} Yann Lanoiselée¹
and Satya N Majumdar³

¹ Laboratoire de Physique de la Matière Condensée (UMR 7643),
CNRS—Ecole Polytechnique, 91128 Palaiseau, France

² Interdisciplinary Scientific Center Poncelet (ISCP), Bolshoy Vlasyevskiy
Pereulok 11, 119002 Moscow, Russia⁴

³ Laboratoire de Physique Théorique et Modèles Statistiques (UMR 8626
du CNRS), Université de Paris-Sud, Bât. 100, 91405 Orsay Cedex, France
E-mail: denis.grebenkov@polytechnique.edu

Received 25 June 2017

Accepted for publication 22 August 2017

Published 24 October 2017



Online at stacks.iop.org/JSTAT/2017/103203
<https://doi.org/10.1088/1742-5468/aa8c11>

Abstract. We investigate the geometric properties of the convex hull over n successive positions of a planar random walk, with a symmetric continuous jump distribution. We derive the large n asymptotic behavior of the mean perimeter. In addition, we compute the mean area for the particular case of isotropic Gaussian jumps. While the leading terms of these asymptotics are universal, the subleading (correction) terms depend on the finer details of the jump distribution and describe a ‘finite size effect’ of discrete-time jump processes, allowing one to accurately compute the mean perimeter and the mean area even for small n , as verified by Monte Carlo simulations. This is particularly valuable for applications dealing with discrete-time jumps processes and ranging from the statistical analysis of single-particle tracking experiments in microbiology to home range estimations in ecology.

Keywords: diffusion, exact results, extreme value, classical Monte Carlo simulations

⁴ International Joint Research Unit—UMI 2615, CNRS/IUM/IITP RAS/Steklov MI RAS/Skoltech/HSE, Moscow, Russia

Contents

1. Introduction	2
2. The model and the main results	5
3. Main steps leading to the derivation of results	9
3.1. Reduction to a one-dimensional problem	9
3.2. Formal solution of the one-dimensional problem	11
3.3. Mean perimeter and mean area of the convex hull	13
3.3.1. Mean perimeter for jumps with a finite variance.	13
3.3.2. Mean perimeter for Lévy flights	13
3.3.3. Mean area for isotropic Gaussian jumps.	14
4. Examples and simulations	15
4.1. Gaussian jumps	15
4.2. Exponentially distributed radial jumps	16
4.3. Independent exponentially distributed jumps	17
4.4. Radial Lévy jumps	18
4.5. Independent Lévy α -stable symmetric jumps.	19
5. Discussion and conclusion	19
Acknowledgments	21
Appendix A. Asymptotic analysis	21
A.1. Derivation of generating functions.	21
A.2. Jumps with a finite variance	22
A.3. Lévy flights	24
Appendix B. Two exactly solvable examples	25
B.1. Symmetric exponential distribution	25
B.2. Modified exponential distribution	27
References	28

1. Introduction

Consider a set of n points with position vectors $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}$ in a d -dimensional space. The most natural and perhaps the simplest way to characterize the *shape* of this set of points is by drawing the convex hull around this set: a convex hull is the unique minimal convex polytope that encloses all the points. This unique polytope is convex since the line segment joining any two points on the surface of the polytope is fully contained within the polytope. Properties of such convex polytopes have been widely studied in mathematics, computer science (image processing and pattenr recognition)

and in the physics of crystallography (the Wulff construction). In two dimension, where the convex hull is a polygon, there are many other applications, most notably in ecology where the home range of animals or the spread of an epidemics are typically estimated by convex hulls. For a review on the history and applications of convex hulls, see [1].

When the points $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}$ are drawn randomly from a joint distribution $P(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n)$, the associated convex hull also becomes random and characterizing its statistical properties is a challenging problem, since the convex hull is a highly non-trivial functional of the random variables $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}$. For instance, what can one say about the statistics of the surface area S_d or the volume V_d of the convex hull, for a given joint distribution $P(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n)$? Even finding the mean surface area $\langle S_d \rangle$ or the mean volume $\langle V_d \rangle$, for arbitrary joint distribution, is a formidably difficult problem. In the special case when the points are independent and identically distributed, i.e. when the joint distribution factorizes as $P(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) = \prod_{k=1}^n P(\vec{r}_k)$ (with $P(\vec{r}_k)$ representing the marginal distribution), several results on the statistics of the surface and volume of the convex hull are known (see [1] for a historical review). However, for *correlated* points where the joint distribution does not factorize, very few results are available.

The simplest example of a set of correlated points corresponds to the case of a random walk in d -dimensional continuous space, where \vec{r}_k represents the position of the walker at step k , starting at the origin at step 0. The position evolves via the Markov rule, $\vec{r}_k = \vec{r}_{k-1} + \vec{\eta}_k$, where $\vec{\eta}_k$ represents the jump at step k , and one assumes that $\vec{\eta}_k$ are independent and identically distributed random variables, each drawn from some prescribed distribution $p(\vec{\eta}_k)$. The walk evolves up to n steps generating the vertices $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}$ of its trajectory. There is a unique convex hull for each sample of this trajectory. What can one say about the mean surface area or the mean volume of this convex hull, given the jump distribution $p(\vec{\eta}_k)$? This is the basic problem that we address in this paper. We show that at least for $d = 2$ (planar random walks), it is possible to obtain precise *explicit* results for all n for the mean perimeter and the mean area of the convex hull of the walk, for a large class of jump distributions $p(\vec{\eta})$, including in particular Lévy flights where the jump distribution has a fat tail. We also obtain similar results for the mean area of the convex hull but under additional assumptions on the jump distribution.

This problem concerning the convex hull of a random walk becomes somewhat simpler in the special case of the Brownian limit, where several results are known. Consider, for example a jump distribution $p(\vec{\eta}_k)$ with zero mean and a finite variance σ^2 . In this case, the walk converges in the large n limit to the Brownian motion. In other words, one can consider the continuous-time limit, as $\sigma^2 \rightarrow 0$ and $n \rightarrow \infty$ with $n\sigma^2 = 2Dt$ being fixed (here D is called the diffusion constant and t is the duration of the walk). In this Brownian limit and for $d = 2$, the mean perimeter and the mean area have been known exactly for a while. Takács [2] computed the mean perimeter

$$\langle S_2 \rangle = \sqrt{16\pi Dt}, \tag{1}$$

while El Bachir [3] and Letac [4] computed the mean area

$$\langle V_2 \rangle = \pi Dt. \tag{2}$$

For a planar Brownian bridge of duration t (where the walker returns to the origin after time t), the mean perimeter $\langle S_2 \rangle_{\text{bridge}} = \sqrt{\pi^3 D t}$ was computed by Goldman [5], while the mean area $\langle V_2 \rangle_{\text{bridge}} = (2\pi/3) D t$ was computed relatively recently by Randon-Furling *et al* [6]. An interesting extension of this problem in $d = 2$ is to the case of N independent planar Brownian motions (or Brownian bridges) [1, 6]. This is relevant in the context of the home range of animals, where N represents the size of an animal population and the trajectory of each animal is approximated by a Brownian motion during their foraging period. For a fixed population size N , the mean perimeter and the mean area of the convex hull were computed exactly: $\langle S_2 \rangle = \alpha_N \sqrt{D t}$ and $\langle V_2 \rangle = \beta_N D t$, where the prefactors α_N and β_N were found to have nontrivial N dependence [1, 6]. For $d > 2$, very few exact results are known for this problem. For a single ($N = 1$) Brownian motion, the mean surface area and the mean volume of the convex hull were recently computed by Eldan [7]: $\langle S_d \rangle = \frac{2(4\pi D t)^{(d-1)/2}}{\Gamma(d)}$ and $\langle V_d \rangle = \frac{(\pi D t)^{d/2}}{\Gamma(d/2+1)^2}$ (see also [8] for another derivation and extension to Brownian bridges). However, for $N > 1$ and $d > 2$, no exact result is available. Finally, going beyond the mean surface and the mean volume, very few results are known for higher moments (see the review [1] for results on variance) or even the full distribution of the surface or the volume of the convex hull of Brownian motion (see [9, 10] for a recent discussion on the distribution of the perimeter in $d = 2$ and $N = 1$). Very recently, the full distribution (including the large deviation tails) of the perimeter and the area of $N \geq 1$ planar Brownian motions were calculated numerically [11, 12]. Some rigorous results on the convex hulls of Lévy processes were recently derived [13, 14].

If one is interested only in the mean area or the mean volume of the convex hull of a generic stochastic process (not necessarily just a random walk), a particular simplification occurs in $d = 2$ (planar case) where several analytical results can be derived by adapting Cauchy's formula [15, 16] for arbitrary closed convex curves in $d = 2$. Indeed by employing Cauchy's formula for every realization of a random planar convex hull, it was shown in [1, 6] that the problem of computing the mean perimeter and the mean area of an *arbitrary* two dimensional stochastic process (can in general be non-Markovian and in discrete-time) can be mapped to computing the extremal statistics associated with the one dimensional component of the process (see section 3 for the precise mapping). This mapping was introduced originally in [6] to compute $\langle S_2 \rangle$ and $\langle V_2 \rangle$ exactly for $N \geq 1$ planar Brownian motions. Since then, it has been used for a number of continuous-time planar processes: random acceleration process [17], branching Brownian motion with applications to animal epidemic outbreak [18], anomalous diffusion processes [19] and also to a single Brownian motion confined to a half-space [20].

The objective of this paper is to go beyond the continuous-time limit and obtain results for the convex hull of a discrete-time planar random walk of n steps (with n large but finite) with arbitrary jump distribution, including for instance Lévy flights. Indeed, in any realistic experiment or simulation, the points of the trajectory are always discrete. For example, recently proposed local convex hull estimators [21] are based on a relatively small number of points, where we cannot apply the Brownian limiting results reviewed above. The first rigorous result for a two-dimensional discrete random walk, modeled as a sum of independent random variables in the complex plane, was derived for the mean perimeter of the convex hull by Spitzer and Widom [22],

$$\langle L_n \rangle = 2 \sum_{k=1}^n \frac{\langle |x_k + iy_k| \rangle}{k} \quad (3)$$

(here $x_k + iy_k$ is the complex-valued position of the walker after k steps). Although the formula (3) looks deceptively simple, an *explicit* computation of the mean $\langle L_n \rangle$ is difficult using equation (3), in particular its behavior for large but finite n . For the case of zero mean and finite variance jump distributions, the leading $n^{\frac{1}{2}}$ term in $\langle L_n \rangle$ was identified in [22], but the relevant subleading terms were not known, to the best of our knowledge. For other results on the statistics of L_n , see [23]. The Spitzer–Widom formula (3) was extended to generic d -dimensional random walks in [24], in which exact combinatorial expressions for the expected surface area and the expected volume of the convex hull were derived. However, these expressions are not suitable for the asymptotic analysis at large n . Several other geometrical properties of the convex hull of random walks are known: for example, the exact formula for the mean number of facets of the convex polytope of a d -dimensional random walk, for $d = 2$ [25] and $d > 2$ [26, 27]. But in this paper, we will restrict ourselves only to the mean perimeter $\langle L_n \rangle = \langle S_2 \rangle$ and the mean area $\langle A_n \rangle = \langle V_2 \rangle$ of a planar random walk of n steps and our main goal is to derive explicitly not only the leading term in $\langle L_n \rangle$ and $\langle A_n \rangle$ for large n , but also the subleading terms.

Our strategy is to adapt the mapping between the convex hull of a 2-d process and the extreme statistics of the 1-d component process, mentioned above, to the case of a single discrete-time planar random walk with generic jump distributions. Using this strategy, we are able to compute explicitly the leading and subleading terms of the mean perimeter of the convex hull of a planar random walk of n steps with arbitrary symmetric continuous jump distributions for large but finite n . The mean area is also computed but only for the particular case of isotropic Gaussian jumps. The rest of the paper is organized as follows. Section 2 outlines the class of considered planar random walks and the main results. In section 3, we explain the derivation steps. In section 4, several explicit examples are presented and used to illustrate the accuracy of the derived asymptotic relations by comparison with Monte Carlo simulations. In section 5, we discuss the main results, their applications, and conclusions. Appendices A and B contain some technical details of the derivation and exactly solvable examples, respectively.

2. The model and the main results

We consider a discrete-time random walker in the plane whose jumps are random, independent, and identically distributed. Starting from the origin, the walker produces a sequence of $(n + 1)$ points $\{(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$ after n jumps such that

$$(x_0, y_0) = (0, 0), \quad (x_k, y_k) = (x_{k-1}, y_{k-1}) + (\eta_k^x, \eta_k^y) \quad (k = 1, 2, \dots, n), \quad (4)$$

where the jumps $\vec{\eta}_k = (\eta_k^x, \eta_k^y)$ are independent from step to step, and at each step they are drawn from a prescribed joint probability density function (PDF) $p(x, y)$, i.e.

$$\mathbb{P}\{\eta_k^x \in (x, x + dx), \eta_k^y \in (y, y + dy)\} = p(x, y) dx dy. \quad (5)$$

We emphasize that the starting point (x_0, y_0) is not random and for convenience, we choose $(x_0 = y_0 = 0)$ to be the origin. The convex hull constructed over these $(n + 1)$

points is the minimal convex polygon that encloses all these points (see figure 1 for an illustration). We are interested in the perimeter L_n and the area A_n of the convex hull which are random variables given that the points are generated as successive positions of a planar random walk. We aim at computing *exactly* the leading and subleading terms of the mean perimeter, $\langle L_n \rangle$, and the mean area, $\langle A_n \rangle$, of the convex hull for large n .

As explained in section 3, our computation relies on two key results: (i) Cauchy’s formula for the perimeter and the area of a closed convex curve, that allows one to reduce the original planar problem to the analysis of one-dimensional projections, and (ii) the Pollaczek–Spitzer formula describing the distribution of the maximum of partial sums of independent symmetric continuously distributed random variables [28, 29]. To use the Pollaczek–Spitzer formula, we need thus to assume that the joint probability density $p(x, y)$ is continuous and centrally symmetric:

$$p(-x, -y) = p(x, y). \tag{6}$$

In particular, our results will not be applicable to a classical random walk on the square lattice because its distribution is not continuous. In the following, we outline the main results that will be derived in section 3.

The mean perimeter $\langle L_n \rangle$ is computed for a very general class of symmetric continuous jump distributions. Writing the Fourier transform of $p(x, y)$ as

$$\hat{\rho}_\theta(k) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy p(x, y) e^{ik(x \cos \theta + y \sin \theta)}, \tag{7}$$

one can characterize the behavior of the mean perimeter according to the asymptotic properties of $\hat{\rho}_\theta(k)$ as $k \rightarrow 0$. We assume a general expansion

$$\hat{\rho}_\theta(k) \simeq 1 - |a_\theta k|^\mu + o(|k|^\mu) \quad (k \rightarrow 0), \tag{8}$$

with the scaling exponent $0 < \mu \leq 2$, and a scale $a_\theta > 0$. When $0 < \mu \leq 1$, the mean perimeter of the convex hull is infinite. We therefore focus on the case $1 < \mu \leq 2$.

First, we derive an exact formula for the generating function of $\langle L_n \rangle$ which is valid for any $1 < \mu \leq 2$. Extracting the asymptotic large n behavior of $\langle L_n \rangle$ from this general formula is, however, nontrivial. We distinguish the following cases.

- (i) When the jump variance is finite ($\mu = 2$), the mean perimeter is shown to behave as

$$\langle L_n \rangle \simeq C_0 n^{\frac{1}{2}} + C_1 + o(1) \quad (n \gg 1), \tag{9}$$

with

$$C_0 = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^{2\pi} d\theta \sigma_\theta, \quad C_1 = \int_0^{2\pi} d\theta \sigma_\theta \gamma_\theta, \tag{10}$$

where σ_θ and γ_θ are given by

$$\sigma_\theta^2 = - \lim_{k \rightarrow 0} \frac{\partial^2 \hat{\rho}_\theta(k)}{\partial k^2} = \langle (\eta^x \cos \theta + \eta^y \sin \theta)^2 \rangle = \frac{a_\theta^2}{2}, \tag{11}$$

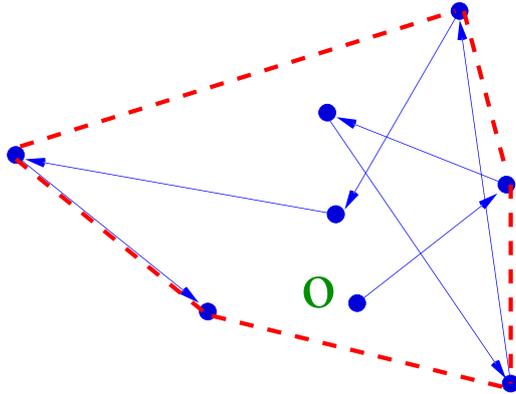


Figure 1. Illustration of the convex hull of a 7-stepped planar random walk. The walk starts at the origin O and makes independent jumps at each step (shown by arrows). After 7 steps, the convex hull (shown by dashed red lines) is constructed around the points of the trajectory.

$$\gamma_\theta = \frac{1}{\pi\sqrt{2}} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - \hat{\rho}_\theta(\sqrt{2}k/\sigma_\theta)}{k^2} \right). \quad (12)$$

If in addition the fourth-order moment of the jump distribution is finite, one gains the second subleading term,

$$\langle L_n \rangle \simeq C_0 n^{\frac{1}{2}} + C_1 + C_2 n^{-\frac{1}{2}} + o(n^{-\frac{1}{2}}) \quad (n \gg 1), \quad (13)$$

with

$$C_2 = \frac{C_0}{8} + \frac{\sqrt{2}}{24\sqrt{\pi}} \int_0^{2\pi} d\theta \sigma_\theta \mathcal{K}_\theta \quad (14)$$

and

$$\mathcal{K}_\theta = \frac{1}{\sigma_\theta^4} \lim_{k \rightarrow 0} \frac{\partial^4 \hat{\rho}_\theta(k)}{\partial k^4} = \frac{\langle (\eta^x \cos \theta + \eta^y \sin \theta)^4 \rangle}{\langle (\eta^x \cos \theta + \eta^y \sin \theta)^2 \rangle^2}. \quad (15)$$

Higher-order corrections can also be derived under further moments assumptions. Note that the integral expression for the coefficient C_0 in front of the leading term $n^{1/2}$ first appeared in [22]. In section 4, we will show that the asymptotic formula (13) is very accurate even for small n .

- (ii) When the jump variance is infinite (i.e. $1 < \mu < 2$), one needs to consider the subleading term in the small k asymptotics of $\hat{\rho}_\theta(k)$:

$$\hat{\rho}_\theta(k) \simeq 1 - |a_\theta k|^\mu + b_\theta |k|^\nu + o(|k|^\nu) \quad (k \rightarrow 0), \quad (16)$$

with the subleading exponent $\nu > \mu$ and a coefficient b_θ . Depending on the subleading exponent ν , we distinguish two cases:

(1) if $\mu < \nu < \mu + 1$, one has

$$\langle L_n \rangle \simeq C_0 n^{1/\mu} + C_1 n^{1-(\nu-1)/\mu} + o(n^{1-(\nu-1)/\mu}) \quad (n \gg 1), \quad (17)$$

with

$$C_0 = \frac{\mu \Gamma(1 - 1/\mu)}{\pi} \int_0^{2\pi} d\theta a_\theta, \quad (18)$$

$$C_1 = -\frac{\Gamma((\nu - 1)/\mu)}{\pi(\mu + 1 - \nu)} \int_0^{2\pi} d\theta a_\theta^{1-\nu} b_\theta. \quad (19)$$

Note that the coefficient C_0 also appears in the mean perimeter of the convex hull of continuous-time symmetric stable processes [13].

(2) if $\nu > \mu + 1$, one has

$$\langle L_n \rangle \simeq C_0 n^{1/\mu} + C_1 + o(1) \quad (n \gg 1), \quad (20)$$

with C_0 from equation (18) and

$$C_1 = \int_0^{2\pi} d\theta \gamma_\theta, \quad (21)$$

where

$$\gamma_\theta = \frac{1}{\pi} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - \hat{\rho}_\theta(k)}{(ak)^\mu} \right). \quad (22)$$

For instance, for a Lévy symmetric alpha-stable distribution with $\hat{\rho}(k) = \exp(-|ak|^\mu)$, one gets [33]

$$\gamma = a \frac{\zeta(1/\mu)}{(2\pi)^{1/\mu} \sin(\pi/(2\mu))}, \quad (23)$$

where $\zeta(z)$ is the Riemann zeta function.

The obtained results are indeed very general.

In turn, our method of computation of the mean area requires two additional strong assumptions: (a) the independence of the jumps along x and y coordinates, i.e. $p(x, y) = p(x) p(y)$ and (b) the isotropy of the jump PDF, i.e. $p(x, y)$ should depend only on the distance $r = \sqrt{x^2 + y^2}$ but not on the direction of the jump. According to Porter–Rosenzweig theorem [30], only the Gaussian jump distribution with identical variance σ^2 along x and y directions, i.e. $p(x, y) = \frac{1}{2\pi\sigma^2} \exp[-(x^2 + y^2)/2\sigma^2]$, satisfies both properties (a) and (b). Our result for the mean area is thus only valid for this Gaussian distribution:

$$\sigma^{-2} \langle A_n \rangle = \frac{\pi}{2} n + \gamma \sqrt{8\pi} n^{\frac{1}{2}} + \pi(\mathcal{K}/12 + \gamma^2) + o(1), \quad (24)$$

with $\mathcal{K} = 3$ and

$$\gamma = \frac{\zeta(1/2)}{\sqrt{2\pi}} = -0.582\,59\dots \quad (25)$$

Recently, an exact formula for the mean area of the convex hull of a Gaussian random walk was derived [8]. In the isotropic case, the formula reads

$$\sigma^{-2}\langle A_n \rangle = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{n-i} \frac{1}{\sqrt{ij}}. \quad (26)$$

While the result in equation (26) is very useful for finite n , deriving the large n asymptotics of this double sum (including up to two subleading terms as in equation (24)) seems somewhat complicated. Our method, in contrast, gives a more direct access to the asymptotics. Moreover, one can check numerically that our asymptotic formula (24) agrees accurately with the exact expression (26) even for moderately large n .

The leading term of equation (24) was shown to be valid for a generic random walk with increments of a finite variance (see proposition 3.3 in [9]). Moreover, our numerical simulations (see section 4.2 and figure 5) suggest that the obtained formula (24) (including the subleading terms) may be applicable for some other isotropic processes. In other words, the technical assumption about the independence of the jumps along x and y might be relaxed in future. This statement, which is uniquely based on numerical simulations for some jump distributions, is conjectural. In turn, the isotropy assumption is important, as illustrated by numerical simulations.

The large n asymptotic relations (13), (17), (20), (24) are the main results of the paper. Setting $t = n\tau$ and $D = 2\sigma^2/(4\tau)$ with a time step τ , one recovers from equations (13) and (24) the same leading terms as in equations (1) and (2) for Brownian motion (note that we write $2\sigma^2$ in D because σ^2 is the variance of jumps along one direction). It is thus not surprising that the leading term in equation (13) is universal because its derivation is valid for any planar random walk with a symmetric and continuous jump distribution and having a finite variance σ^2 .

Thinking of Brownian motion as a limit of random walks, the subleading terms in equation (13) can be understood as ‘finite size’ corrections. The first subleading term is valid under the same assumptions as the leading term, although the coefficient γ_θ depends on the jump distribution (see examples in section 4). In turn, the second subleading term depends on the kurtosis \mathcal{K}_θ and thus requires an additional assumption that \mathcal{K}_θ is finite.

3. Main steps leading to the derivation of results

3.1. Reduction to a one-dimensional problem

We start with Cauchy’s formula for the perimeter L and the area A of an arbitrary convex domain \mathcal{C} with a reasonably smooth boundary $\gamma_{\mathcal{C}}$ [15, 16]. Let the boundary $\gamma_{\mathcal{C}}$ be parameterized as $(X(s), Y(s))$ with a curvilinear coordinate s ranging from 0 to 1. Setting the origin of coordinates inside the domain, one defines the support function

$M(\theta)$ as the distance from the origin to the closest straight line that does not cross the domain and is perpendicular to the vector from the origin in direction θ (figure 2). In other words,

$$M(\theta) = \max_{0 \leq s \leq 1} \{X(s) \cos \theta + Y(s) \sin \theta\}. \tag{27}$$

Cauchy showed that [15]

$$L = \int_0^{2\pi} d\theta M(\theta), \tag{28}$$

$$A = \frac{1}{2} \int_0^{2\pi} d\theta (M^2(\theta) - [M'(\theta)]^2). \tag{29}$$

For a simple derivation of this formula see [1]. A straightforward calculation of $M(\theta)$ for a convex hull over a set of points may seem to be hopeless, as one would need first to construct the convex hull by identifying and ordering its vertices among the given set of points and then to compute $M(\theta)$. The key idea is that $M(\theta)$ can be found directly from the vertices of the trajectory as [6, 22]

$$M(\theta) = \max_{0 \leq k \leq n} \{x_k \cos \theta + y_k \sin \theta\}. \tag{30}$$

Moreover, given that the maximum for a fixed θ is realized by a certain vertex (with index k^* which discretely changes with θ), one also obtains the derivative:

$$M'(\theta) = -x_{k^*} \sin \theta + y_{k^*} \cos \theta. \tag{31}$$

When the points (x_k, y_k) are random, the perimeter and the area of the convex hull are random variables. We focus on the mean values $\langle L_n \rangle$ and $\langle A_n \rangle$:

$$\langle L_n \rangle = \int_0^{2\pi} d\theta \langle M(\theta) \rangle, \tag{32}$$

$$\langle A_n \rangle = \frac{1}{2} \int_0^{2\pi} d\theta (\langle M^2(\theta) \rangle - \langle [M'(\theta)]^2 \rangle), \tag{33}$$

i.e. the computation is reduced to the first two moments of $M(\theta)$ and to the mean $\langle [M'(\theta)]^2 \rangle$. The important observation is that, for a fixed direction θ , one needs to characterize the maximum of the projection of points (x_k, y_k) onto that direction

$$M(\theta) = \max_{0 \leq k \leq n} \{z_k^\theta\}, \quad z_k^\theta = x_k \cos \theta + y_k \sin \theta. \tag{34}$$

The projection of a random walk is also a random walk. In fact, we can write according to equation (4)

$$z_0^\theta = 0, \quad z_k^\theta = z_{k-1}^\theta + \xi_k^\theta \quad (k = 1, 2, \dots, n), \tag{35}$$

with

$$\xi_k^\theta = \eta_k^x \cos \theta + \eta_k^y \sin \theta. \tag{36}$$

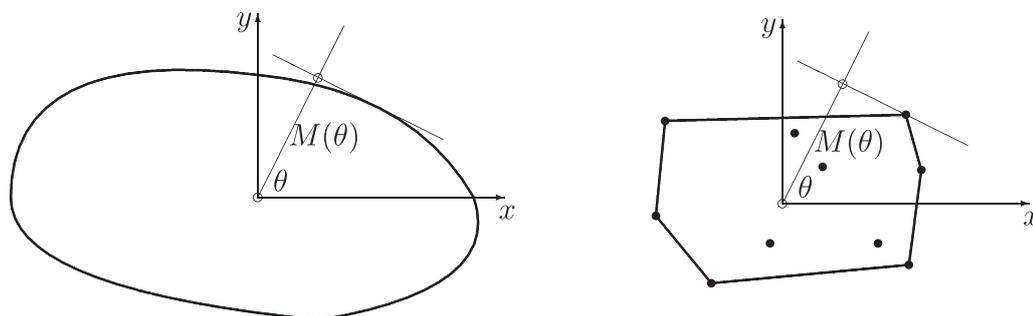


Figure 2. The support function $M(\theta)$ for a closed convex curve (left) and for a set of points $\{(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)\}$ (right). $M(\theta)$ is the distance between two open circles.

The probability density of ξ_k^θ , $\rho_\theta(z)$, is fully determined by that of the jump (η_k^x, η_k^y) . In particular, its Fourier transform $\hat{\rho}_\theta(k)$ is related to $p(x, y)$ by equation (7). The symmetry (6) implies that $\hat{\rho}_\theta(-k) = \hat{\rho}_\theta(k)$ and thus the density $\rho_\theta(z)$ is symmetric.

Having discussed the general jump distributions, let us mention two particular cases that will be important later.

- (a) in the case of independent jumps along x and y coordinates, one has $p(x, y) = p_x(x) p_y(y)$, and thus

$$\hat{\rho}_\theta(k) = \hat{p}_x(k \cos \theta) \hat{p}_y(k \sin \theta), \tag{37}$$

where \hat{p}_x and \hat{p}_y are the Fourier transforms of $p_x(x)$ and $p_y(y)$, respectively.

- (b) For isotropic jumps, $p(x, y)$ depends only on the radial coordinate, $p(x, y) dx dy = p_r(r) dr d\phi / (2\pi)$, where $p_r(r)$ is the radial density (that includes the factor r from the Jacobian). From equation (7), one gets

$$\hat{\rho}(k) = \int_0^\infty dr p_r(r) J_0(|k|r), \tag{38}$$

in which the integration over the angular coordinate ϕ eliminated the dependence on θ and resulted in the Bessel function of the first kind, $J_0(|k|r)$.

3.2. Formal solution of the one-dimensional problem

The formal exact solution of the one-dimensional problem can be obtained via the Pollaczek–Spitzer formula [28, 29]. This formula characterizes the maximum of partial sums of independent identically distributed random variables ξ_k with a symmetric and continuous density $\rho(z)$:

$$M_n = \max\{0, \xi_1, \xi_1 + \xi_2, \dots, \xi_1 + \xi_2 + \dots + \xi_n\} \tag{39}$$

(in this subsection, we temporarily drop the subscript and superscript θ from all the variables; it will be restored at the end). Considering ξ_k as jumps of a random walker, $z_k = z_{k-1} + \xi_k$ (with $z_0 = 0$), one can also write

$$M_n = \max\{z_0, z_1, z_2, \dots, z_n\}. \tag{40}$$

Pollaczek and later Spitzer showed that the cumulative distribution $Q_n(z) = \mathbb{P}\{M_n \leq z\}$ of M_n satisfies the following identity for $0 \leq s \leq 1$ and $\lambda \geq 0$

$$\sum_{n=0}^{\infty} s^n \langle e^{-\lambda M_n} \rangle = \sum_{n=0}^{\infty} s^n \int_0^{\infty} dz e^{-\lambda z} Q'_n(z) = \frac{\phi(s, \lambda)}{\sqrt{1-s}}, \tag{41}$$

with

$$\phi(s, \lambda) = \exp\left(-\frac{\lambda}{\pi} \int_0^{\infty} dk \frac{\ln(1 - s\hat{\rho}(k))}{\lambda^2 + k^2}\right), \tag{42}$$

where $Q'_n(z) = dQ_n(z)/dz$ is the probability density of the maximum [28, 29]. In principle, all the moments of M_n can be obtained from equation (41). However, in practice, deriving explicitly the moments of M_n by inverting this formula is highly nontrivial [31]. The expected maximum of a discrete-time random walk $\langle M_n \rangle$ appears in a number of different problems, from packing algorithms in computer science [32], all the way to the survival probability of a single or multiple walkers in presence of a trap [33–38]. It has also emerged in the context of the order, gap and record statistics of random walks [39–43]. Note that $\langle M_n \rangle$ has already been analyzed for large n using the Pollaczek–Spitzer formula in equation (41). Here, in addition to calculating the first *three* terms in the asymptotic expansion of $\langle M_n \rangle$ for $n \gg 1$, we also calculate the large n behavior of the second moment $\langle M_n^2 \rangle$, that we need for the computation of the mean area of the convex hull.

In fact, the Pollaczek–Spitzer formula determines the generating functions for all moments of M_n :

$$h_m(s) = \sum_{n=0}^{\infty} s^n \langle M_n^m \rangle = (-1)^m \lim_{\lambda \rightarrow 0} \frac{\partial^m}{\partial \lambda^m} \frac{\phi(s, \lambda)}{\sqrt{1-s}}. \tag{43}$$

By considering a general asymptotic expansion

$$\hat{\rho}(k) \simeq 1 - |ak|^\mu + o(|k|^\mu) \quad (k \rightarrow 0) \tag{44}$$

with an exponent $0 < \mu \leq 2$ and a scale $a > 0$, we derive in appendix A.1 the exact expressions

$$h_1(s) = \frac{1}{\pi(1-s)} \int_0^{\infty} \frac{dk}{k^2} \ln\left(\frac{1 - s\hat{\rho}(k)}{1-s}\right) \quad (0 \leq s < 1), \tag{45}$$

which is valid for any $1 < \mu \leq 2$ (note that $\langle M_n \rangle = \infty$ for $0 < \mu \leq 1$), and

$$h_2(s) = (1-s)[h_1(s)]^2 + \frac{a^2 s}{(1-s)^2} \quad (0 \leq s < 1), \tag{46}$$

which is valid for $\mu = 2$ (note that $\langle M_n^2 \rangle = \infty$ for $0 < \mu < 2$). The exact relations (45) and (46) are new results which allow one to study the first two moments of the maximum M_n . In the next subsection, we will analyze the expansion of equations (45) and (46) as $s \rightarrow 1$ in order to determine the asymptotic behavior of the moments $\langle M_n \rangle$ and $\langle M_n^2 \rangle$ as $n \rightarrow \infty$. We consider separately jumps with a finite variance, and Lévy flights.

3.3. Mean perimeter and mean area of the convex hull

3.3.1. *Mean perimeter for jumps with a finite variance.* Given a generic continuous jump distribution $p(x, y)$ satisfying the property in equation (6), we determine $\hat{\rho}_\theta(k)$ using equation (7). Furthermore, by examining the small k behavior of $\hat{\rho}_\theta(k)$, we determine the θ -dependent variance σ_θ^2 and the θ -dependent kurtosis \mathcal{K}_θ , using respectively equations (11) and (15). In addition, knowing $\hat{\rho}_\theta(k)$ from equation (7), we also determine γ_θ in equation (12). Equipped with these three quantities σ_θ , \mathcal{K}_θ and γ_θ , we show in appendix A that the leading large n terms of the first two moments of M_n are given by

$$\frac{\langle M_n \rangle}{\sigma_\theta} \simeq \frac{\sqrt{2}}{\sqrt{\pi}} n^{\frac{1}{2}} + \gamma_\theta + \frac{\mathcal{K}_\theta + 3}{12\sqrt{2\pi}} n^{-\frac{1}{2}} + o(n^{-\frac{1}{2}}), \tag{47}$$

$$\frac{\langle M_n^2 \rangle}{\sigma_\theta^2} \simeq n + \frac{\sqrt{8} \gamma_\theta}{\sqrt{\pi}} n^{\frac{1}{2}} + (\mathcal{K}_\theta/12 + \gamma_\theta^2) + o(1). \tag{48}$$

For the mean perimeter of the convex hull, we will only need the first moment in equation (47). Indeed, using equation (32), the integration of the expansion (47) over θ from 0 to 2π yields the announced result (13) for the mean perimeter of the convex hull. The result for the second moment in equation (48) will be needed later to determine the mean area $\langle A_n \rangle$.

3.3.2. *Mean perimeter for Lévy flights.* When the variance of jumps is infinite, one gets the Taylor expansion equation (8), with the scaling exponent $0 < \mu < 2$. When $0 < \mu \leq 1$, the mean perimeter of the convex hull is infinite. Throughout this section, we focus on the case $1 < \mu < 2$, in which the first moment of jumps is finite (and zero due to the assumption of a symmetric distribution), whereas the variance is infinite. In this case, the leading behavior of the mean maximum of partial sums is universal [33]

$$\langle M_n \rangle \simeq a_\theta \frac{\mu\Gamma(1 - 1/\mu)}{\pi} n^{1/\mu} + o(n^{1/\mu}) \quad (n \gg 1). \tag{49}$$

However, the subleading term depends on finer details of the jump distribution. In order to determine the subleading term, we consider the expansion (16) with the subleading term $b_\theta |k|^\nu$ such that $\nu > \mu$. We distinguish two cases: $\mu < \nu < \mu + 1$ and $\nu > \mu + 1$. In appendix A.3, we derive the following asymptotics results:

$$\langle M_n \rangle \simeq a_\theta \frac{\mu\Gamma(1 - 1/\mu)}{\pi} n^{1/\mu} - a_\theta^{1-\nu} b_\theta \frac{\Gamma((\nu - 1)/\mu)}{\pi(\mu + 1 - \nu)} n^{1-(\nu-1)/\mu} + o(n^{1-(\nu-1)/\mu}) \quad (n \gg 1) \tag{50}$$

for $\mu < \nu < \mu + 1$, and

$$\langle M_n \rangle \simeq a_\theta \frac{\mu\Gamma(1 - 1/\mu)}{\pi} n^{1/\mu} + \gamma_\theta + o(1) \quad (n \gg 1) \tag{51}$$

for $\nu > \mu + 1$, with γ_θ given by equation (22). The asymptotic relation (51) was first derived in [33] for the particular case $\nu = 2\mu$. One can see that for $\mu < \nu < \mu + 1$, the subleading term of $\langle M_n \rangle$ grows with n , whereas for when $\nu > \mu + 1$, the subleading term is constant. Higher-order corrections can be derived as well.

Finally, using the Cauchy formula (32), the integration of equations (50) and (51) over θ from 0 to 2π yields equations (17) and (20) for the mean perimeter of the convex hull, announced in section 2.

3.3.3. *Mean area for isotropic Gaussian jumps.* According to equation (33), the expansion (48) determines the first contribution to the mean area. This contribution was calculated for an arbitrary symmetric continuous jump distribution with a finite variance. The second contribution to the mean area comes from $\langle [M'(\theta)]^2 \rangle$ that has to be computed separately. We recall that $M'(\theta)$ is given by equation (31). Our computation of this contribution relies on two additional simplifying assumptions: (a) the jumps along x and y coordinates are independent and (b) the jump process is isotropic, i.e. the distribution of jumps does not depend on their direction. In this case, using the isotropy condition (b), we get

$$\langle [M(\theta)]^2 \rangle = \langle [M(0)]^2 \rangle = \langle x_{k^*}^2 \rangle, \tag{52}$$

$$\langle [M'(\theta)]^2 \rangle = \langle [M'(0)]^2 \rangle = \langle y_{k^*}^2 \rangle. \tag{53}$$

The disentanglement of $\langle [M(\theta)]^2 \rangle$ and $\langle [M'(\theta)]^2 \rangle$ allows one to compute the latter one by using the following argument. We recall that k^* is the index of the maximal position among x_k , i.e. its statistics is fully determined by the jumps along x coordinate. Once this statistics is known, the mean $\langle [M'(0)]^2 \rangle$ can be found by taking the conditional expectation of $y_{k^*}^2$ at any fixed value k^* and then the expectation with respect to the distribution of k^* . Now, once we condition k^* , i.e. the time step at which the x_k 's achieve their maximum, the y_k process will, in general, be affected by this conditioning. However, if x_k and y_k are independent (which happens when the jump process satisfies property (a) above), we get

$$\langle [M'(0)]^2 \rangle = \sigma^2 \langle k^* \rangle, \tag{54}$$

where $\langle [\eta_k^y]^2 \rangle = \sigma^2$. It remains to find $\langle k^* \rangle$. For symmetric and continuous jump distributions, it follows from symmetry that $\langle k^* \rangle = n/2$ independent of the details of the jump PDF. This can be deduced formally also, by noting that the time step k^* , at which the maximum of x_k is achieved, has a universal distribution, independent of the jump distribution (given that the latter is continuous and symmetric) [31]:

$$P_n(k^* = k) = \binom{2k}{k} \binom{2(n-k)}{n-k} 2^{-2n}. \tag{55}$$

This is the direct consequence of the Sparre Andersen theorem. From this distribution, one easily computes the mean value as

$$\langle k^* \rangle = \frac{n}{2}, \tag{56}$$

and thus

$$\langle [M'(\theta)]^2 \rangle = \sigma^2 \frac{n}{2}. \tag{57}$$

This yields equation (24) for the mean area of the convex hull for an isotropic jump process with independent jumps along x and y coordinates. As discussed in section 2,

the only process satisfying two requirements (a) and (b) is the isotropic Gaussian process. Our formula (24) is thus provided only in this case. For a more general case, one would need to compute the joint distribution of the maximum position k^* and both values x_{k^*} and y_{k^*} which is more complicated and remains an open problem.

4. Examples and simulations

In this section, we illustrate the above general results on several examples of symmetric planar random walks. We derive the explicit values of the relevant parameters that determine the mean perimeter and the mean area. We also investigate the effect of anisotropy of the jump distribution on the convex hull properties. Finally, we compare our theoretical predictions to the results of Monte Carlo simulations. For this purpose, we generate 10^5 planar random walks, compute the convex hull for each generated trajectory with $n + 1$ points by using the Matlab function ‘convhull’, and determine its perimeter and area. These simulations yield the representative statistics of perimeters and areas from which the mean values are computed.

4.1. Gaussian jumps

We first consider the basic example of Gaussian jumps which are independent along x and y coordinates and characterized by variances σ_x^2 and σ_y^2 . Substituting the jump probability density,

$$p(x, y) = \frac{\exp\left(-\frac{x^2}{2\sigma_x^2}\right)}{\sqrt{2\pi}\sigma_x} \frac{\exp\left(-\frac{y^2}{2\sigma_y^2}\right)}{\sqrt{2\pi}\sigma_y}, \tag{58}$$

into equation (7) yields $\hat{\rho}_\theta(k) = e^{-k^2\sigma_\theta^2/2}$, with $\sigma_\theta^2 = \sigma_x^2 \cos^2 \theta + \sigma_y^2 \sin^2 \theta$. One can see that the anisotropy only affects the variance σ_θ^2 , whereas the two other relevant parameters, γ and \mathcal{K} , which are rescaled by variance, do not depend on θ . One finds $\mathcal{K} = 3$, whereas the integral in equation (12) was computed exactly in [33] and provided in equation (25). Assuming (without loss of generality) that $\sigma_x \geq \sigma_y$, we set

$$\sigma \equiv \frac{1}{2\pi} \int_0^{2\pi} d\theta \sigma_\theta = \frac{2}{\pi} E\left(\sqrt{1 - (\sigma_y/\sigma_x)^2}\right) \sigma_x, \tag{59}$$

where $E(k)$ is the complete elliptic integral of the second kind (for the isotropic case, $\sigma_x = \sigma_y = \sigma$). With this notation, we get the expansion coefficients

$$\sigma^{-1}C_0 = \sqrt{8\pi}, \quad \sigma^{-1}C_1 = 2\pi\gamma = -3.6605\dots, \quad \sigma^{-1}C_2 = \frac{\sqrt{\pi}}{\sqrt{2}}. \tag{60}$$

Figure 3 shows the rescaled mean perimeter, $\langle L_n \rangle/n^{1/2}$, and the rescaled mean area, $\langle A_n \rangle/n$, for isotropic planar random walk with independent Gaussian jumps ($\sigma_x = \sigma_y = 1$). The results of Monte Carlo simulations are in perfect agreement with our theoretical predictions (13) and (24). One can see that the subleading terms play an important role. In fact, if one kept only the leading term and omitted the subleading

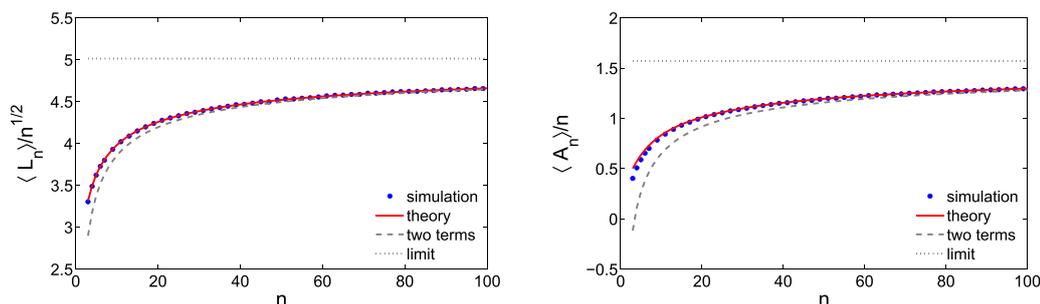


Figure 3. The rescaled mean perimeter, $\langle L_n \rangle / n^{1/2}$, (left), and the rescaled mean area, $\langle A_n \rangle / n$, (right), for isotropic planar random walks with independent Gaussian jumps, with $\sigma_x = \sigma_y = 1$. The results of Monte Carlo simulations (shown by circles) are in perfect agreement with our theoretical predictions (13) and (24) (shown by solid line). The dotted horizontal line presents the coefficients $\sqrt{8\pi}$ and $\pi/2$ of the leading term, whereas the dashed line illustrates the theoretical predictions with only two principal terms.

terms, one would get the horizontal dotted line. This corresponds to the case of Brownian motion, in which only the leading term is present, see equations (1) and (2). The subleading terms account for the discrete-time character of random walks which is particularly important for moderate values of n . In order to highlight the role of the third term in the asymptotic expansions, we also draw by dashed line the asymptotics without this term. As expected, the third term improves the quality of the theoretical prediction at small n . Note also that the asymptotic relation is slightly less accurate for the mean area.

In figure 4, we consider the convex hull for anisotropic random walk with independent Gaussian jumps, with $\sigma_x = 5$ and $\sigma_y = 1$. In this case, one can use the asymptotic formula (13), in which the expansion coefficients C_j are given by equation (60), with an effective variance σ^2 computed in equation (59). In this example, $\sigma = 5 \frac{2}{\pi} E(\sqrt{1 - 1/25}) = 3.3439 \dots$. For the mean perimeter, one observes a perfect agreement between the theoretical predictions and Monte Carlo simulations. In turn, our asymptotic formula (24) for the mean area is not applicable for anisotropic case, as also confirmed by simulations (not shown).

4.2. Exponentially distributed radial jumps

The next common model has exponentially distributed radial jumps with uniform angular distribution. This is a particular realization of a ‘run-and-tumble’ model of bacterial motion [44–46]. Substituting the radial density $p_r(r) = \sigma^{-1} e^{-r/\sigma}$ into equation (38) yields $\hat{\rho}(k) = (1 + (k\sigma)^2)^{-1/2}$. One gets $\mathcal{K} = 9$ and

$$\gamma = \frac{1}{\pi\sqrt{2}} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - (1 + 2k^2)^{-1/2}}{k^2} \right) = -0.8183 \dots \quad (61)$$

from which the expansion coefficients are

$$\sigma^{-1} C_0 = \sqrt{8\pi}, \quad \sigma^{-1} C_1 = -5.1416 \dots, \quad \sigma^{-1} C_2 = \sqrt{2\pi}. \quad (62)$$

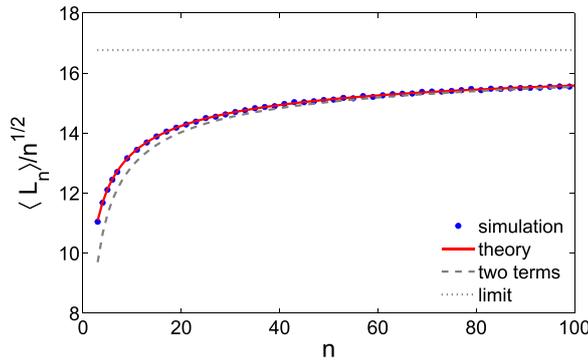


Figure 4. The rescaled mean perimeter, $\langle L_n \rangle / n^{1/2}$, for anisotropic planar random walk with independent Gaussian jumps, with $\sigma_x = 5$ and $\sigma_y = 1$. The results of Monte Carlo simulations (shown by circles) are in perfect agreement with our theoretical prediction (13) (shown by solid line). The dotted horizontal line presents the coefficient $\sigma\sqrt{8\pi}$ of the leading term (with $\sigma = 5\frac{2}{\pi}E(\sqrt{1-1/25}) = 3.3439\dots$), whereas the dashed line illustrates the theoretical predictions with only two principal terms.

Figure 5 illustrates the obtained results. As for isotropic Gaussian jumps, there is a perfect agreement between theory and simulations for the mean perimeter. We also present the results for the mean area. We recall that the theoretical formula (24) was derived under the assumption of independent jumps along x and y coordinates, which evidently fails for the exponential radial jumps. In spite of this failure, the theoretical formula (24) is in perfect agreement with Monte Carlo simulations, except for very small n . This empirical observation suggests a possibility to relax this technical assumption in future, at least for large n .

4.3. Independent exponentially distributed jumps

We also consider the case, when the jumps along x and y coordinates are independent and exponentially distributed, with two densities $p_x(x) = \frac{1}{2}\sigma_x^{-1}e^{-|x|/\sigma_x}$ and $p_y(y) = \frac{1}{2}\sigma_y^{-1}e^{-|y|/\sigma_y}$. The Fourier transforms are $\hat{p}_x(k) = (1 + (k\sigma_x)^2)^{-1}$ and $\hat{p}_y(k) = (1 + (k\sigma_y)^2)^{-1}$ so that

$$\hat{\rho}_\theta(k) = \frac{1}{1 + (k\sigma_x)^2 \cos^2 \theta} \frac{1}{1 + (k\sigma_y)^2 \sin^2 \theta}. \tag{63}$$

We get $\sigma_\theta^2 = 2(\sigma_x^2 \cos^2 \theta + \sigma_y^2 \sin^2 \theta)$ and

$$\mathcal{K}_\theta = 24 \frac{\sigma_x^4 \cos^4 \theta + \sigma_x^2 \sigma_y^2 \sin^2 \theta \cos^2 \theta + \sigma_y^4 \sin^4 \theta}{\sigma_\theta^4}. \tag{64}$$

Using the identity (B.2), we compute explicitly

$$\gamma_\theta = \frac{\sqrt{2}\sigma_x\sigma_y|\sin \theta \cos \theta| - \sigma_\theta(\sigma_x|\cos \theta| + \sigma_y|\sin \theta|)}{\sigma_\theta^2}. \tag{65}$$

In the particular case $\sigma_x = \sigma_y = \sigma$, one gets $\sigma_\theta^2 = 2\sigma^2$, $\mathcal{K}_\theta = 6(1 - \sin^2 \theta \cos^2 \theta)$, and

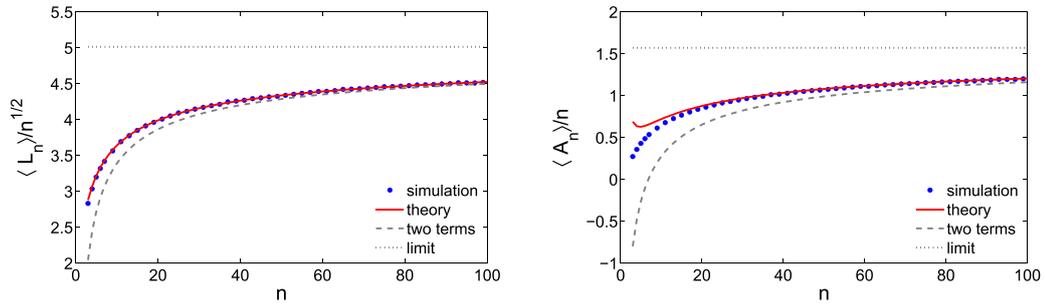


Figure 5. The rescaled mean perimeter, $\langle L_n \rangle / n^{1/2}$, (left), and the rescaled mean area, $\langle A_n \rangle / n$, (right), for isotropic planar random walk with exponentially distributed radial jumps, with $\sigma = 1$. The results of Monte Carlo simulations (shown by circles) are in perfect agreement with our theoretical predictions (13) and (24) (shown by solid line). The dotted horizontal line presents the coefficient $\sqrt{8\pi}$ and $\pi/2$ of the leading term, whereas the dashed line illustrates the theoretical predictions with only two principal terms.

$$\gamma_\theta = \frac{|\sin \theta \cos \theta| - |\cos \theta| - |\sin \theta|}{\sqrt{2}}. \tag{66}$$

For this case, we obtain from equation (10)

$$\sigma^{-1}C_0 = 4\sqrt{\pi}, \quad \sigma^{-1}C_1 = -6, \quad \sigma^{-1}C_2 = \frac{11\sqrt{\pi}}{8}. \tag{67}$$

Figure 6 illustrates the excellent agreement between theory and simulations.

4.4. Radial Lévy jumps

Now, we investigate the example of radial Lévy flights with infinite variance (but finite mean) and uniform angle distribution. Among various heavy-tailed jump distributions (e.g. Pareto distributions), we choose for our illustrative purposes the distribution

$$\mathbb{P}\{\xi > r\} = (1 + (r/R)^2)^{-\alpha}, \tag{68}$$

with a scale $R > 0$ and the scaling exponent $\mu = 2\alpha$, with $\frac{1}{2} < \alpha < 1$. For this distribution, equation (38) yields a simple closed formula [47]

$$\hat{\rho}(k) = \frac{2^{1-\alpha}}{\Gamma(\alpha)} (|k|R)^\alpha K_\alpha(|k|R), \tag{69}$$

where $K_\alpha(z)$ is the modified Bessel function of the second kind. The asymptotic behavior of $K_\alpha(z)$ near z implies as $k \rightarrow 0$

$$\hat{\rho}(k) \simeq 1 - \frac{\pi (|k|R)^{2\alpha}}{2^{2\alpha} \sin(\pi\alpha) \Gamma(\alpha) \Gamma(\alpha + 1)} + \frac{(kR)^2}{4(1 - \alpha)} + O(|k|^{2+2\alpha}). \tag{70}$$

Comparing this expansion to equation (16), one can identify

$$a = R \left(\frac{\pi}{2^{2\alpha} \sin(\pi\alpha) \Gamma(\alpha) \Gamma(\alpha + 1)} \right)^{\frac{1}{2\alpha}}, \quad b = \frac{R^2}{4(1 - \alpha)}, \quad \nu = 2. \tag{71}$$

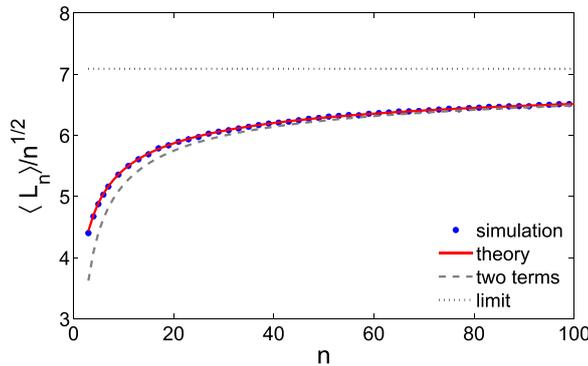


Figure 6. The rescaled mean perimeter, $\langle L_n \rangle / n^{1/2}$, for anisotropic planar random walk with independent exponentially distributed jumps, with $\sigma_x = \sigma_y = 1$. The results of Monte Carlo simulations (shown by circles) are in perfect agreement with our theoretical prediction (13) (shown by solid line). The dotted horizontal line presents the coefficient $4\sqrt{\pi}$ of the leading term, whereas the dashed line illustrates the theoretical predictions with only two principal terms.

In figure 7, the mean perimeter computed by Monte Carlo simulations for $\mu = 1.5$ and $R = 1$ is compared to the theoretical prediction (17). One can see that the agreement is good but worse than for the earlier examples with a finite variance. One obvious reason is that here we have determined only two terms, whereas equation (13) contains three terms.

4.5. Independent Lévy α -stable symmetric jumps

Finally, we investigate Lévy α -stable symmetric jumps, with independent displacements along x and y coordinates given by $\hat{p}_x(k) = \hat{p}_y(k) = \exp(-|ak|^\mu)$, with $1 < \mu < 2$. Using equation (37), one gets thus $\hat{\rho}_\theta(k) = \exp(-|a_\theta k|^\mu)$, with

$$a_\theta = a(|\cos \theta|^\mu + |\sin \theta|^\mu)^{1/\mu}. \tag{72}$$

so that $\nu = 2\mu$, and $b_\theta = a_\theta^2/2$. The mean perimeter is determined by equation (20), with

$$C_0 = a \frac{\mu\Gamma(1 - 1/\mu)}{\pi} \int_0^{2\pi} d\theta (|\cos \theta|^\mu + |\sin \theta|^\mu)^{1/\mu}, \tag{73}$$

and γ given by equation (23) which is independent of θ .

For $\mu = 3/2$, we obtain numerically $a^{-1}C_0 = 8.6275\dots$ and $a^{-1}C_1 = -5.2151\dots$. Figure 8 shows the good agreement between the theoretical prediction (20) and Monte Carlo simulations.

5. Discussion and conclusion

To summarize, we have presented exact asymptotic results for the mean perimeter of the convex hull of an n -step discrete-time random walk in a plane, with a generic continuous jump distribution satisfying the central symmetry assumption in equation (6).

Mean perimeter and mean area of the convex hull over planar random walks

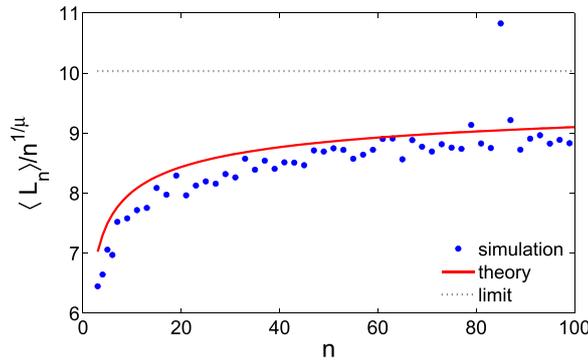


Figure 7. The rescaled mean perimeter, $\langle L_n \rangle / n^{1/\mu}$, for isotropic planar random walk with radial Lévy flights whose lengths are distributed according to equation (68) with $\mu = 1.5$ and $R = 1$. The results of Monte Carlo simulations (shown by circles) agree well with the theoretical prediction (17) (shown by the solid line).

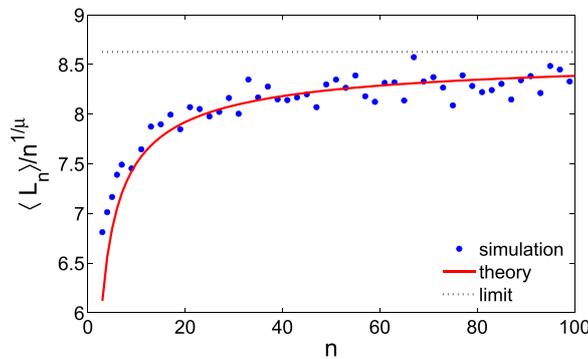


Figure 8. The rescaled mean perimeter, $\langle L_n \rangle / n^{1/\mu}$, for anisotropic planar random walk with independent Lévy α -stable jump distribution with $\mu = 1.5$ and $a = 1$. The results of Monte Carlo simulations (shown by circles) agree well with the theoretical prediction (20) (shown by the solid line).

Explicit results, along with simulations confirming them have been presented for several examples of such jump distributions. For the mean area of the convex hull, we have derived exact results for isotropic Gaussian jump distributions. For jumps with a finite variance, our results provide precise estimates of the deviations from the Brownian limit and explain the discrepancies between the asymptotic Brownian limit results and observed simulations for finite but large n .

The obtained results are particularly valuable for applications dealing with discrete-time random processes, e.g. home range estimation in ecology. Given that the tracks of animal displacements are typically recorded at discrete time steps (e.g. daily observations) and relatively short, the subleading terms play an important role. The asymptotic formulas can also be used for calibrating new estimators, based on the local convex hull, that were proposed for the analysis of intermittent processes in microbiology [21]. Finally, the knowledge of the mean perimeter of the convex hull can be used to estimate the scaling exponent and the scale of symmetric Lévy flights, for which the conventional mean and variance estimators are useless.

There are many interesting open problems that may possibly be addressed using the methods presented here. For example, the numerical evidence suggests a possible

extension of the derived asymptotic formula for the mean area to other isotropic processes, beyond the Gaussian case. Also, it would be interesting to extend our results to the case of the convex hull of planar discrete-time random bridges (where the walker is constrained to come back to the starting point after n discrete jumps). For such discrete-time bridges, there are recent exact results on the statistics of the first two maximum and the gap between them [48] which may be useful for the convex hull problem. One can also consider the problem with many independent discrete-time walkers. Finally, it would be interesting to study the statistics of the perimeter and the area for random walks with jump distributions that violate the reflection property in equation (6), for example, for walks in presence of a drift or a potential.

Acknowledgments

DG acknowledges the support under Grant No. ANR-13-JSV5-0006-01 of the French National Research Agency.

Appendix A. Asymptotic analysis

In this appendix, we derive the main results (47) and (48). The derivation is based on the asymptotic analysis of the Pollaczek–Spitzer identity (41) and extends the earlier results from [33]. Since the function $\phi(s, \lambda)$ from equation (42) is related to the Laplace transform of the probability density $Q'_n(z)$, it determines the generating functions $h_m(s)$ of the moments $\langle M_n^m \rangle$ via equation (43). In turn, the asymptotic behavior of $h_m(s)$ as $s \rightarrow 1$ determines the asymptotic behavior of $\langle M_n^m \rangle$ as $n \rightarrow \infty$.

A.1. Derivation of generating functions

We recall that M_n denotes the maximum of partial sums of independent identically distributed random variables $\{\xi_k\}$:

$$M_n = \max\{0, \xi_1, \xi_1 + \xi_2, \dots, \xi_1 + \dots + \xi_n\}. \tag{A.1}$$

We assume that the jump distribution is symmetric and continuous, whereas its characteristic function $\hat{\rho}(k)$ admits the expansion

$$\hat{\rho}(k) = 1 - |ak|^\mu + \dots \quad (k \rightarrow 0), \tag{A.2}$$

with an exponent $0 < \mu \leq 2$ and a scale $a > 0$.

First, we derive the generating function $h_1(s)$ from the Pollaczek–Spitzer formula (41). For this purpose, we need to determine the expansion of this formula in powers of λ for small λ . Let us first write

$$\phi(s, \lambda) = \exp[-I(s, \lambda)], \quad I(s, \lambda) = \frac{\lambda}{\pi} \int_0^\infty \frac{dk}{k^2 + \lambda^2} \ln(1 - s\hat{\rho}(k)). \tag{A.3}$$

It is easy to obtain the $\lambda \rightarrow 0$ limit of $I(s, \lambda)$. We rescale $k = \lambda u$ in equation (A.3) and take the $\lambda \rightarrow 0$ limit. This gives, using $\hat{\rho}(k=0) = 1$, a very simple expression

$$I(s, 0) = \frac{1}{\pi} \int_0^\infty \frac{du}{1+u^2} \ln(1-s) = \frac{1}{2} \ln(1-s). \tag{A.4}$$

Next we re-write

$$\begin{aligned} \phi(s, \lambda) &= \exp[-I(s, \lambda)] = \exp[-I(s, 0)] \exp[-(I(s, \lambda) - I(s, 0))] \\ &= \frac{1}{\sqrt{1-s}} \exp\left[-\frac{\lambda}{\pi} \int_0^\infty \frac{dk}{\lambda^2 + k^2} \ln\left(\frac{1-s\hat{\rho}(k)}{1-s}\right)\right]. \end{aligned} \tag{A.5}$$

Expanding the right-hand side of equation (A.5) up to the first order λ for small λ (with s fixed) gives

$$\phi(s, \lambda) \simeq \frac{1}{\sqrt{1-s}} \left[1 - \frac{\lambda}{\pi} \int_0^\infty \frac{dk}{k^2} \ln\left(\frac{1-s\hat{\rho}(k)}{1-s}\right) + o(\lambda) \right]. \tag{A.6}$$

Note that the integral in the second term is convergent for any jump pdf with $\hat{\rho}(k)$ satisfying equation (A.2) provided $1 < \mu \leq 2$. Taking the derivative with respect to λ , evaluated at $\lambda = 0$, yields

$$h_1(s) = \sum_{n=0}^\infty s^n \langle M_n \rangle = \frac{1}{\pi(1-s)} \int_0^\infty \frac{dk}{k^2} \ln\left(\frac{1-s\hat{\rho}(k)}{1-s}\right). \tag{A.7}$$

Note that this is an exact result for any jump PDF $\hat{\rho}(k)$ satisfying equation (A.2) with $1 < \mu \leq 2$ and it also holds for arbitrary s such that the generating function on the left-hand side of equation (A.7) is convergent. In turn, when $0 < \mu \leq 1$, the mean value $\langle M_n \rangle$ is infinite, and $h_1(s)$ is undefined.

The second moment $\langle M_n^2 \rangle$ is finite only for $\mu = 2$. Expanding equation (A.5) up to the order λ^2 for small λ (with s fixed) and taking the second derivative with respect to λ at $\lambda = 0$, one gets for any $0 \leq s < 1$

$$h_2(s) = \sum_{n=0}^\infty s^n \langle M_n^2 \rangle = (1-s)[h_1(s)]^2 + \frac{a^2 s}{(1-s)^2}. \tag{A.8}$$

A.2. Jumps with a finite variance

Here we investigate the asymptotic behavior of $h_1(s)$ as $s \rightarrow 1$ for the jump distribution with a finite variance σ^2 so that $\hat{\rho}(k) \simeq 1 - k^2\sigma^2/2 + o(k^2)$. First, we represent $h_1(s)$ from equation (45) as

$$h_1(s) = \frac{1}{\pi(1-s)} \int_0^\infty \frac{dk}{k^2} \left\{ \ln\left(\frac{1-s(1-k^2\sigma^2/2)}{1-s}\right) + \ln\left(\frac{1-s\hat{\rho}(k)}{1-s(1-k^2\sigma^2/2)}\right) \right\}. \tag{A.9}$$

The first integral can be computed explicitly, so that

$$h_1(s) = \frac{\sigma\sqrt{s/2}}{(1-s)^{3/2}} + \frac{I_\varepsilon}{1-s}, \tag{A.10}$$

where I_ε denotes the second integral in equation (A.9) that we rewrite as

$$I_\varepsilon = \frac{1}{\pi} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - (1 - \varepsilon^2)\hat{\rho}(k)}{\sigma^2 k^2/2 + \varepsilon^2(1 - \sigma^2 k^2/2)} \right), \tag{A.11}$$

with $\varepsilon = \sqrt{1 - s}$. Setting $\varepsilon \rightarrow 0$, one gets

$$I_0 = \frac{1}{\pi} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - \hat{\rho}(k)}{\sigma^2 k^2/2} \right) = \sigma \gamma, \tag{A.12}$$

with the constant γ from equation (12), with $\sigma = \sigma_\theta$.

Next, we consider

$$I_\varepsilon - I_0 = \frac{1}{\pi} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - (1 - \varepsilon^2)\hat{\rho}(k)}{\frac{1}{2}\sigma^2 k^2 + \varepsilon^2(1 - \frac{1}{2}\sigma^2 k^2)} \frac{\frac{1}{2}\sigma^2 k^2}{1 - \hat{\rho}(k)} \right). \tag{A.13}$$

Changing again the integration variable, one has

$$I_\varepsilon - I_0 = \frac{1}{\pi\varepsilon} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - (1 - \varepsilon^2)\hat{\rho}(k\varepsilon)}{\frac{1}{2}\sigma^2 k^2 + (1 - \frac{1}{2}\sigma^2 \varepsilon^2 k^2)} \frac{\frac{1}{2}\sigma^2 k^2}{1 - \hat{\rho}(k\varepsilon)} \right). \tag{A.14}$$

To get the next-order term, we assume the existence of the fourth-order moment of the jumps so that

$$\hat{\rho}(k\varepsilon) \simeq 1 - \frac{1}{2}\sigma^2 k^2 \varepsilon^2 + c\sigma^4 k^4 \varepsilon^4 + o(\varepsilon^4), \tag{A.15}$$

where $c = \mathcal{K}/24$ is related to the kurtosis \mathcal{K} given by equation (15). Substituting this expansion into equation (A.14), one gets

$$I_\varepsilon - I_0 \simeq \frac{1}{\pi\varepsilon} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - \frac{\varepsilon^2 k^4 c \sigma^4}{1 + \sigma^2 k^2/2}}{1 - 2k^2 \varepsilon^2 c \sigma^2} \right) \simeq \frac{1}{\pi\varepsilon} \int_0^\infty dk \frac{2c\varepsilon^2 \sigma^2}{1 + \sigma^2 k^2/2} = \varepsilon \sqrt{2} c \sigma. \tag{A.16}$$

Combining these results, we conclude that

$$h_1(s) \simeq \frac{\sigma}{\sqrt{2}(1-s)^{3/2}} + \frac{\sigma\gamma}{1-s} + \frac{\sigma(\mathcal{K}/6 - 1)}{\sqrt{8}(1-s)^{1/2}} + o((1-s)^{-1/2}) \quad (s \rightarrow 1). \tag{A.17}$$

Substituting this expansion into equation (46), one gets

$$h_2(s) = \frac{\sigma^2}{(1-s)^2} + \frac{\sqrt{2}\sigma^2\gamma}{(1-s)^{3/2}} + \frac{\sigma^2(\mathcal{K}/12 + \gamma^2 - 1)}{(1-s)} + o((1-s)^{-1}) \quad (s \rightarrow 1). \tag{A.18}$$

In order to invert the relations (A.17) and (A.18), we use the following identities:

$$\sum_{n=0}^\infty s^n a_n = \frac{1}{(1-s)^{1/2}}, \quad \sum_{n=0}^\infty s^n = \frac{1}{1-s}, \tag{A.19}$$

$$\sum_{n=0}^\infty s^n b_n = \frac{1}{(1-s)^{3/2}}, \quad \sum_{n=0}^\infty (n+1)s^n = \frac{1}{(1-s)^2}, \tag{A.20}$$

with

$$a_n = \binom{2n}{n} 2^{-2n} \simeq \frac{1}{\sqrt{\pi n}} \left(1 - \frac{1}{8n} + O(n^{-2}) \right), \tag{A.21}$$

$$b_n = (2n + 1) \binom{2n}{n} 2^{-2n} \simeq \frac{2\sqrt{n}}{\sqrt{\pi}} \left(1 + \frac{3}{8n} + O(n^{-2}) \right). \tag{A.22}$$

Inverting term by term, we obtain the main relations (47) and (48). Higher-order corrections can be obtained in the same way.

A.3. Lévy flights

For Lévy flights with $\hat{\rho}(k) \simeq 1 - |ak|^\mu + o(|k|^\mu)$, one can derive the asymptotic behavior of $h_1(s)$ as $s \rightarrow 1$ for the case $1 < \mu < 2$. In this case, we replace the representation (A.9) by

$$h_1(s) = \frac{1}{\pi(1-s)} \int_0^\infty \frac{dk}{k^2} \left\{ \ln \left(\frac{1 - s(1 - (ak)^\mu)}{1 - s} \right) + \ln \left(\frac{1 - s\hat{\rho}(k)}{1 - s(1 - (ak)^\mu)} \right) \right\}. \tag{A.23}$$

As previously, the first integral can be evaluated explicitly so that

$$h_1(s) = \frac{a s^{1/\mu}}{(1-s)^{1+1/\mu}} \frac{1}{\sin(\pi/\mu)} + \frac{I_\varepsilon}{1-s}, \tag{A.24}$$

where

$$I_\varepsilon = \frac{1}{\pi} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - (1 - \varepsilon^2)\hat{\rho}(k)}{\varepsilon^2 + (1 - \varepsilon^2)(ak)^\mu} \right), \tag{A.25}$$

with $\varepsilon = \sqrt{1-s}$. The first term in equation (A.24) provides the leading contribution. Using the discrete form of the Tauberian theorem,

$$\sum_{n=0}^\infty a_n s^n = \frac{1}{(1-s)^\alpha} \Rightarrow a_n \simeq \frac{n^{\alpha-1}}{\Gamma(\alpha)} \quad (n \gg 1), \tag{A.26}$$

one gets the leading term to $\langle M_n \rangle$ to be $(a\mu\Gamma(1 - 1/\mu)/\pi) n^{1/\mu}$.

In order to compute the subleading term, one needs to analyze the integral I_ε whose behavior depends on the next-order term in the small k expansion of $\hat{\rho}(k)$ in equation (16). Formally setting $\varepsilon = 0$ in equation (A.25), one would get

$$I_0 = \frac{1}{\pi} \int_0^\infty \frac{dk}{k^2} \ln \left(\frac{1 - \hat{\rho}(k)}{(ak)^\mu} \right). \tag{A.27}$$

Substituting the expansion (16) into this integral, one gets the factor $\ln(1 - bk^{\nu-\mu}/a^\mu)$. As a consequence, the integral converges when $\nu - \mu > 1$ and diverges otherwise. This condition naturally distinguishes two asymptotic regimes: $\nu > \mu + 1$ and $\nu \leq \mu + 1$. In the former case, the integral converges, and the two principal terms in $h_1(s)$ are

$$h_1(s) \simeq \frac{a s^{1/\mu}}{(1-s)^{1+1/\mu}} \frac{1}{\sin(\pi/\mu)} + \frac{I_0}{1-s} + o((1-s)^{-1}) \quad (s \rightarrow 1), \tag{A.28}$$

where I_0 is identical to γ from equation (22). The inversion of this relation yields equation (51).

In the case $\nu \leq \mu + 1$, the subtraction of the term with $1 - |ak|^\mu$ in equation (A.23) was not sufficient to get the convergent correction term. We remedy this problem by using another representation:

$$h_1(s) = \frac{1}{\pi(1-s)} \int_0^\infty \frac{dk}{k^2} \left\{ \ln \left(\frac{1-s(1-(ak)^\mu + bk^\nu)}{1-s} \right) + \ln \left(\frac{1-s\hat{\rho}(k)}{1-s(1-(ak)^\mu + bk^\nu)} \right) \right\}. \tag{A.29}$$

Changing the variable $u = ak(s/(1-s))^{1/\mu}$, we rewrite the first term as

$$\frac{as^{1/\mu}}{\pi(1-s)^{1+1/\mu}} \int_0^\infty \frac{du}{u^2} \ln(1 + u^\mu - Cu^\nu), \tag{A.30}$$

with $C = ba^{-\nu}s^{-\nu/\mu}(1-s)^{\nu/\mu}$. When $s \rightarrow 1$, C is small, and one can expand the logarithm to get

$$\frac{as^{1/\mu}}{\pi(1-s)^{1+1/\mu}} \int_0^\infty \frac{du}{u^2} \left\{ \ln(1 + u^\mu) - \frac{Cu^\nu}{1 + u^\mu} \right\}. \tag{A.31}$$

Now both integrals are convergent. The first integral yields the same leading term as earlier, whereas the second term provides the subleading correction. Finally, the second term in equation (A.29) results in higher-order corrections that we ignore. Keeping only the leading term and the first subleading term, we get as $s \rightarrow 1$

$$h_1(s) \simeq \frac{a s^{1/\mu}}{(1-s)^{1/\mu}} \frac{1}{\sin(\pi/\mu)} - \frac{b a^{1-\nu}}{(1-s)^{1+(1-\nu)/\mu}} \frac{1}{\mu \sin(\pi(\nu-1)/\mu)} (1 + o(1)). \tag{A.32}$$

The inversion of this relation yields the large n asymptotic formula (50) for $\langle M_n \rangle$.

Appendix B. Two exactly solvable examples

In this appendix, we briefly discuss two distributions, for which the function $\phi(s, \lambda)$ can be found exactly: the symmetric exponential distribution $\rho(z) = \frac{1}{2}e^{-|z|}$ (see [33]) and its modification $\rho(z) = \frac{1}{2}|z|e^{-|z|}$. These examples served us as benchmarks for checking general results.

B.1. Symmetric exponential distribution

When $\rho(z) = \frac{1}{2}e^{-|z|}$, one gets $\hat{\rho}(k) = (1 + k^2)^{-1}$ and thus equation (42) becomes

$$\phi(s, \lambda) = \frac{1 + \lambda}{\sqrt{1 - s + \lambda}}, \tag{B.1}$$

where we used the identity

$$\int_0^\infty dk \frac{\ln(a^2 + b^2 k^2)}{\lambda^2 + k^2} = \frac{\pi}{\lambda} \ln(a + \lambda b) \tag{B.2}$$

to compute the integral:

$$\int_0^\infty dk \frac{\ln(1 - s/(1 + k^2))}{\lambda^2 + k^2} = \frac{\pi}{\lambda} \ln\left(\frac{\sqrt{1 - s} + \lambda}{1 + \lambda}\right). \tag{B.3}$$

According to equations (11), (15) and (12), we also obtain

$$\sigma^2 = 2, \quad \mathcal{K} = 6, \quad \gamma = -\frac{1}{\sqrt{2}}. \tag{B.4}$$

The first derivative of $\phi(s, \lambda)$ from equation (B.1) yields

$$h_1(s) = \sum_{n=0}^\infty s^n \langle M_n \rangle = \frac{1}{(1 - s)^{3/2}} - \frac{1}{1 - s}, \tag{B.5}$$

from which, using equations (A.19) and (A.20), one retrieves the exact form of the first moment, which is valid for any n and was first derived in [33]:

$$\langle M_n \rangle = (2n + 1) \binom{2n}{n} 2^{-2n} - 1. \tag{B.6}$$

Using the Stirling formula, one can get the large n expansion to any order:

$$\sigma^{-1} \langle M_n \rangle \simeq \frac{\sqrt{2}}{\sqrt{\pi}} n^{\frac{1}{2}} - \frac{1}{\sqrt{2}} + \frac{3}{4\sqrt{2\pi}} n^{-\frac{1}{2}} - \frac{7}{64\sqrt{2\pi}} n^{-\frac{3}{2}} + O(n^{-\frac{5}{2}}). \tag{B.7}$$

The first three terms agree with the general expansion (47).

Similarly, we get from equation (B.1)

$$h_2(s) = \sum_{n=0}^\infty s^n \langle M_n^2 \rangle = \frac{2}{(1 - s)^2} - \frac{2}{(1 - s)^{3/2}}, \tag{B.8}$$

from which the exact formula follows using equations (A.19) and (A.20)

$$\sigma^{-2} \langle M_n^2 \rangle = (n + 1) - (2n + 1) \binom{2n}{n} 2^{-2n}. \tag{B.9}$$

In the large n limit, we deduce

$$\sigma^{-2} \langle M_n^2 \rangle \simeq n - \frac{2}{\sqrt{\pi}} n^{\frac{1}{2}} + 1 - \frac{3}{4\sqrt{\pi}} n^{-\frac{1}{2}} + \frac{7}{64\sqrt{\pi}} n^{-\frac{3}{2}} + O(n^{-\frac{5}{2}}). \tag{B.10}$$

The first three terms agree with the general expansion (48).

Note that one can also obtain the generating function for the cumulative distribution $Q_n(z)$. In fact, one has for the Laplace transform of $Q_n(z)$

$$\sum_{n=0}^\infty s^n \mathcal{L}\{Q_n(z)\}(\lambda) = \frac{1 + \lambda}{\lambda\sqrt{1 - s}(\sqrt{1 - s} + \lambda)}, \tag{B.11}$$

from which the inverse Laplace transform (with respect to λ) yields

$$\sum_{n=0}^{\infty} s^n Q_n(z) = \frac{1}{1-s} - \left(\frac{1}{1-s} - \frac{1}{\sqrt{1-s}} \right) e^{-z\sqrt{1-s}}. \tag{B.12}$$

From this relation, one can easily get any generating function $h_m(s)$.

B.2. Modified exponential distribution

For the case $\rho(z) = \frac{1}{2}|z|e^{-|z|}$, one gets

$$\hat{\rho}(k) = \frac{1-k^2}{(1+k^2)^2}. \tag{B.13}$$

Using again equation (B.2), one computes the integral in equation (42) as

$$\int_0^{\infty} dk \frac{\ln(1-s\hat{\rho}(k))}{\lambda^2+k^2} = \frac{\pi}{\lambda} \ln \left(\frac{(\mu_+(s)+\lambda)(\mu_-(s)+\lambda)}{(1+\lambda)^2} \right), \tag{B.14}$$

where

$$\mu_{\pm}(s) = \sqrt{\frac{2+s \pm \sqrt{s^2+8s}}{2}}. \tag{B.15}$$

We get thus

$$\phi(s, \lambda) = \frac{(1+\lambda)^2}{(\mu_+(s)+\lambda)(\mu_-(s)+\lambda)}. \tag{B.16}$$

According to equations (11), (15) and (12), we also obtain

$$\sigma^2 = 6, \quad \mathcal{K} = \frac{10}{3}, \quad \gamma = \frac{1}{3\sqrt{2}} - \frac{2}{\sqrt{6}}. \tag{B.17}$$

Taking the first derivative of $\phi(s, \lambda)$ from equation (B.16) with respect to λ yields

$$h_1(s) = \frac{-1}{\sqrt{1-s}} \frac{2\mu_-\mu_+ - (\mu_+ + \mu_-)}{\mu_+^2 \mu_-^2}. \tag{B.18}$$

Since $\mu_-\mu_+ = \sqrt{1-s}$, $\mu_+^2 + \mu_-^2 = 2+s$, and $\mu_+ + \mu_- = \sqrt{2+s+2\sqrt{1-s}}$, we get

$$h_1(s) = \frac{\sqrt{2+s+2\sqrt{1-s}}}{(1-s)^{3/2}} - \frac{2}{1-s}. \tag{B.19}$$

Expanding near $s \rightarrow 1$, one obtains

$$h_1(s) \simeq \frac{\sqrt{3}}{(1-s)^{3/2}} + \frac{\frac{1}{\sqrt{3}} - 2}{1-s} - \frac{2}{3\sqrt{3}(1-s)^{1/2}} + O(1), \tag{B.20}$$

from which

$$\sigma^{-1} \langle M_n \rangle \simeq \frac{\sqrt{2}}{\sqrt{\pi}} n^{\frac{1}{2}} + \gamma - \frac{19}{36\sqrt{2\pi}} n^{-\frac{1}{2}} + O(n^{-\frac{3}{2}}), \tag{B.21}$$

in agreement with the general expansion (47).

Similarly, we compute $h_2(s)$ as the second derivative of $\phi(s, \lambda)$:

$$\begin{aligned} h_2(s) &= 2 \frac{\mu_- \mu_+ + \mu_-^2 \mu_+^2 - 2\mu_- \mu_+ (\mu_- + \mu_+) + \mu_+^2 + \mu_-^2}{\sqrt{1-s} \mu_+^3 \mu_-^3} \\ &= \frac{6}{(1-s)^2} + \frac{2 - 4\sqrt{2+s} + 2\sqrt{1-s}}{(1-s)^{3/2}}. \end{aligned} \quad (\text{B.22})$$

As $s \rightarrow 1$, one gets

$$h_2(s) \simeq \frac{6}{(1-s)^2} + \frac{2 - 4\sqrt{3}}{(1-s)^{3/2}} - \frac{4/\sqrt{3}}{(1-s)} + O((1-s)^{-1/2}), \quad (\text{B.23})$$

from which

$$\sigma^{-2} \langle M_n^2 \rangle \simeq n + \gamma \frac{2}{\sqrt{\pi}} n^{\frac{1}{2}} + 1 - \frac{2}{3\sqrt{3}} + O(n^{-\frac{1}{2}}), \quad (\text{B.24})$$

in agreement with the general expansion (48).

References

- [1] Majumdar S N, Comtet A and Randon-Furling J 2010 *J. Stat. Phys.* **138** 1–55
- [2] Takács L 1980 *Am. Math. Mon.* **87** 5–6
- [3] El Bachir M 1983 *PhD Thesis* Université Paul Sabatier, Toulouse, France
- [4] Letac G 1993 *J. Theor. Probab.* **6** 385
- [5] Goldman A 1996 *Probab. Theory Relat. Fields* **105** 57
- [6] Randon-Furling J, Majumdar S N and Comtet A 2009 *Phys. Rev. Lett.* **103** 140602
- [7] Eldan R 2014 *Electron. J. Problems* **19** 1–34
- [8] Kabluchko Z and Zaporozhets D 2016 *Trans. Am. Math. Soc.* **368** 8873–99
- [9] Wade A and Xu C 2015 *Stoch. Process. Appl.* **125** 4300
- [10] Wade A and Xu C 2015 *Proc. Am. Math. Soc.* 143 433–45
- [11] Claussen G, Hartmann A K and Majumdar S N 2015 *Phys. Rev. E* **91** 052104
- [12] Dwenter T, Claussen G, Hartmann A K and Majumdar S N 2016 *Phys. Rev. E* **94** 052120
- [13] Kampf J, Last G and Molchanov I 2012 *Proc. Am. Math. Soc.* 140 2527–35
- [14] Molchanov I and Wespi F 2016 *Electron. Commun. Probab.* **21** 1–11
- [15] Cauchy A 1832 *La Rectification des Courbes* (Paris: Memoire de l'Académie des Sciences)
- [16] Santalo L A 1976 *Integral Geometry and Geometric Probability* (Reading, MA: Addison-Wesley)
- [17] Reymbaut A, Majumdar S N and Rosso A 2011 *J. Phys. A: Math. Theor.* **44** 415001
- [18] Dumonteil E, Majumdar S N, Rosso A and Zoia A 2013 *Proc. Natl Acad. Sci. USA* **110** 4239
- [19] Lukovic M, Geisel T and Eule S 2013 *New J. Phys.* **15** 063034
- [20] Chupeau M, Bénichou O and Majumdar S N 2015 *Phys. Rev. E* **91** 050104
- [21] Lanoiselée Y and Grebenkov D S 2017 *Phys. Rev. E* **96** 022144
- [22] Spitzer F and Widom H 1961 *Proc. Am. Math. Soc.* 12 506–9
- [23] Snyder T L and Steele J M 1993 *Proc. Am. Math. Soc.* 117 1165–73
- [24] Vysotsky V and Zaporozhets D 2015 arXiv:1506.07827
- [25] Baxter G 1961 *Ann. Math. Stat.* **32** 901
- [26] Kabluchko Z, Vysotsky V and Zaporozhets D 2016 arXiv:1612.00249
- [27] Randon-Furling J and Wespi F 2017 *Phys. Rev. E* **95** 032129
- [28] Pollaczek F 1952 *C. R.* **234** 2334
- [29] Spitzer F 1956 *Trans. Am. Math. Soc.* **82** 323
- [30] Porter C E and Rosenzweig N 1960 *Ann. Acad. Sci. Fennicae. Ser. A: VI Phys.* **44** 1
- [31] Majumdar S N 2010 *Physica A* **389** 4299
- [32] Coffman E G, Flajolet Ph, Flato L and Hofri M 1998 *Probab. Eng. Inf. Sci.* **12** 373
- [33] Comtet A and Majumdar S N 2005 *J. Stat. Mech.* P06013

- [34] Majumdar S N, Comtet A and Ziff R M 2006 *J. Stat. Phys.* **122** 833
- [35] Ziff R M, Majumdar S N and Comtet A 2007 *J. Phys.: Condens. Matter* **19** 065102
- [36] Ziff R M, Majumdar S N and Comtet A 2009 *J. Chem. Phys.* **130** 204104
- [37] Franke J and Majumdar S N 2012 *J. Stat. Mech.* P05024
- [38] Majumdar S N, Mounaix Ph and Schehr G 2017 arXiv:1704.05940
- [39] Schehr G and Majumdar S N 2012 *Phys. Rev. Lett.* **108** 040601
- [40] Schehr G and Majumdar S N 2014 Exact record and order statistics of random walks via first-passage ideas
First-Passage Phenomena and Their Applications ed R Metzler *et al* (Singapore: World Scientific)
- [41] Majumdar S N, Mounaix Ph and Schehr G 2013 *Phys. Rev. Lett.* **111** 070601
- [42] Wergen G, Majumdar S N and Schehr G 2012 *Phys. Rev. E* **86** 011119
- [43] Godreche C, Majumdar S N and Schehr G 2017 arXiv:1702.00586
- [44] Berg H C and Brown D A 1972 *Nature* **239** 500–4
- [45] Berg H C 2004 *E. Coli in Motion* (New York: Springer)
- [46] Lauga E and Powers T R 2009 *Rep. Prog. Phys.* **72** 096601
- [47] Gradshteyn I S and Ryzhik I M 1980 *Table of Integrals, Series and Products* (New York: Academic)
- [48] Majumdar S N, Mounaix P and Schehr G 2014 *J. Stat. Mech.* P09013