IOP Publishing Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **55** (2022) 225001 (25pp) https://doi.org/10.1088/1751-8121/ac62bb

Mean area of the convex hull of a run and tumble particle in two dimensions

Prashant Singh[1,*], Anupam Kundu[1], Satya N Majumdar[2] and Hendrik Schawe[3]

[1] International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bengaluru 560089, India
[2] LPTMS, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91405, Orsay, France
[3] Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS, CY Cergy Paris Université, France

E-mail: prashant.singh@icts.res.in and anupam.kundu@icts.res.in

Received 22 December 2021, revised 8 March 2022
Accepted for publication 30 March 2022
Published 11 May 2022

Abstract

We investigate the statistics of the convex hull for a single run-and-tumble particle (RTP) in two dimensions. RTP, also known as the persistent random walker, has gained significant interest in the recent years due to its biological application in modelling the motion of bacteria. We consider two different statistical ensembles depending on whether (i) the total number of tumbles n or (ii) the total observation time t is kept fixed. Benchmarking the results on the perimeter, we study the statistical properties of the area of the convex hull for a RTP. Exploiting the connections to extreme value statistics, we obtain exact analytical expressions for the mean area for both ensembles. For fixed-t ensemble, we show that the mean area possesses a scaling form in γt (with γ being the tumbling rate) and the corresponding scaling function is exactly computed. Interestingly, we find that it exhibits a crossover from $\sim t^3$ scaling at small times $\left(t \ll \gamma^{-1}\right)$ to $\sim t$ scaling at large times $\left(t \gg \gamma^{-1}\right)$. On the other hand, for fixed-n ensemble, the mean expectedly grows linearly with n for $n \gg 1$. All our analytical findings are supported with the numerical simulations.

Keywords: run and tumble particle, convex hull, extreme value statistics, stochastic processes

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

*Author to whom any correspondence should be addressed.

 © 2022 IOP Publishing Ltd Printed in the UK

1. Introduction

Active matter refers to a class of driven non-equilibrium systems that transduces systematic movement out of the supplied energy. Contrary to the boundary-driven systems, the energy is exchanged at the local scale which endows the constituent particles with self-propulsion [1–3]. As a result, the dynamics of these systems break time-reversal symmetry and thus, violate the detailed balance. Recently, the self-propulsion (or 'active' nature) has been harnessed to produce useful work for potential therapeutic applications in various diseases like cancer and heart disease [4, 5]. Furthermore, they display a plethora of complex features like clustering [6, 7], flocking [8, 9], motility induced phase separation [10–13], non-existence of the equation of states for pressure [14] and so on. Going beyond the theoretical studies, the dynamics of active particles has been realised in many experiments based on different phoretic effects [15, 16].

Run and tumble particle (RTP) has emerged as a quintessential model in mimicking the dynamics of the active particles. Previously known in the random walk literature as persistent Brownian motion [17, 18], the RTP motion has recently been quite extensively studied due to its biological application in modelling the motion of bacteria like *Escherichia coli* [10, 19, 20]. In this model, the particle moves in a series of ballistic runs interspersed by instantaneous tumbles that occur after random time durations with constant rate γ. During tumble, the particle does not move but chooses a new direction for the next run. Exact model is introduced in section 2. Over the recent few years, this model has been substantially studied and a variety of results are known. Examples include—position distribution in free space as well as in confining potential [21–27], condensation transition [28–30], persistent properties [31–33], extremal properties [34, 35], path functionals [34, 36], current fluctuations [37], interacting multiple RTPs [27, 38–41], etc.

In this paper, we are interested in the statistics of the convex hull for a RTP in two dimensions. Consider a set of points $(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ in two dimensions which represent the positions of a RTP at various instances of time. Then, convex hull refers to the unique smallest convex polygon that encloses all these points [42, 43] (see figure 1). Since the motion is random, the set $(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ also varies over realisations. This implies that the convex hull is also different for different realisations. We are interested in the statistical properties of this random convex hull. As stated before, the run and tumble dynamics is used to model the motion of bacteria like *E. coli* [19]. One natural question is—what is the spatial extent over which bacteria move during their search activities? One way to estimate this is by constructing the convex hull for the trajectories of the particles. In fact, in ecology, convex hull has been used in estimating the home-range size of animals from their locational data [44]. Hence, a knowledge of the convex hull may also be useful for bacteria in designing and demarcating their mobility territory. This motivates us to study the convex hull problem for a RTP in our paper.

The properties of the convex hull have been of prime interest in the mathematics literature also [45–49]. In physics, the mean area and the mean perimeter of the convex hull are found to be related to the subject of extreme value statistics [50]. Exploiting this connection, the mean area and the mean perimeter have been studied for a variety of processes like Brownian motion [51], random acceleration [52], diffusion with resetting [53], random walk and its generalisations [54–61]. Extensions of these studies to higher dimensions and multi-particle case have also been considered [50, 62–64]. Going beyond the mean values, the entire distributions of the area and perimeter have also been studied using sophisticated numerical techniques [55–57]. We refer to [50] for a review on the convex hull problem.

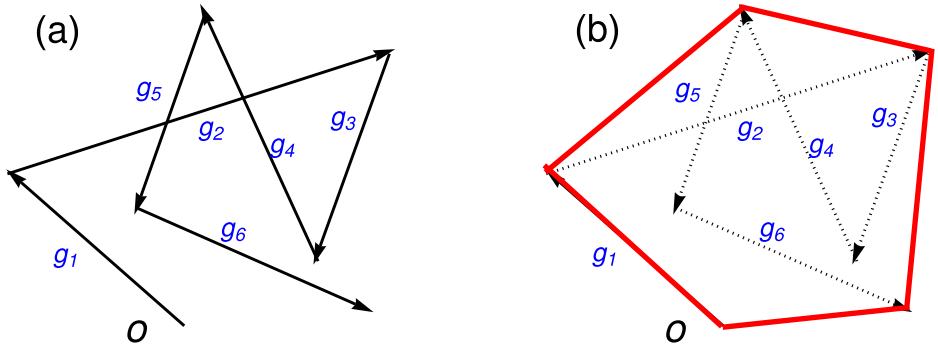

Figure 1. (a) Schematic representation of a typical trajectory of a RTP in two dimensions with total number of runs $n = 6$. The RTP moves in a series of runs interspersed by instantaneous tumbles that occur after random times with rate γ. (b) This figure shows the convex hull (red polygon) for the trajectory on the left.

Recently in [65], the mean perimeter of the convex hull for a RTP in a plane was exactly computed for the two different ensembles—(i) fixed number of tumbles n and (ii) fixed observation time t (discussed later). Here, we go beyond this work to investigate the statistics of the area of the convex hull for a two dimensional RTP. For both ensembles, we compute the mean area exactly. We verify our analytical results numerically and, also study the variance and the distribution of the area numerically.

The paper is organised as follows: in section 2, we introduce the model and summarize the main results of our paper. Section 3 contains a brief discussion on the convex hull problems for the general 2D stochastic processes. Analytic calculations for the mean area are presented in section 4 for the fixed-n ensemble and in section 5 for the fixed-t ensemble. We devote section 6 for the numerical study of the probability distribution of the area which is followed by the conclusion in section 7.

2. Model and summary of the results

We consider a RTP moving on a plane. Starting from the origin, the particle chooses an angle ϕ_1 (measured with respect to the x-axis) uniformly from $[0, 2\pi]$ and moves ballistically in that direction with a speed v_0. The ballistic motion, referred to as a 'run', persists along ϕ_1 for a random time τ_1 drawn from the exponential distribution $\rho(\tau) = \gamma\, \mathrm{e}^{-\gamma\tau}$ with constant rate γ. After this, the particle 'tumbles' instantaneously in which it chooses a new direction ϕ_2 uniformly from $[0, 2\pi]$. Then, it performs another run for a random time τ_2 again drawn independently from $\rho(\tau) = \gamma\, \mathrm{e}^{-\gamma\tau}$. The motion continues in the form of the ballistic *runs* interspersed by the instantaneous *tumbles* that occur after random times drawn independently from the exponential distribution. Let us focus on the ith run along the direction ϕ_i. Denoting the displacement (position increment) during ith interval by (x_i, y_i), we have

$$x_i = v_0 \tau_i \cos(\phi_i), \tag{1}$$

$$y_i = v_0 \tau_i \sin(\phi_i), \tag{2}$$

where τ_i is the total duration for ith run. The position of the particle (X_i, Y_i) after ith run can be written in terms of (x_i, y_i) as

$$X_i = X_{i-1} + x_i, \tag{3}$$

$$Y_i = Y_{i-1} + y_i, \tag{4}$$

where $i = 1, 2, \ldots$ and we assume $(X_0, Y_0) = (0, 0)$. As mentioned earlier, we consider the motion of the particle in two different ensembles with—(i) fixed number of tumbles (n) and (ii) fixed observation time (t). In case (i), the particle undergoes n number of runs and we stop the process after n runs have taken place. The total observation time t will fluctuate for different realisations. Moreover, we consider the starting point as a tumble which makes the number of runs equal to the number of tumblings and $n \geqslant 1$. On the other hand, for ensemble (ii), we fix the total observation time t and therefore, the number of tumblings n fluctuates for different realisations. We will refer to (i) as fixed-n ensemble and (ii) as fixed-t ensemble.

For these two ensembles, we look at the statistical properties of the convex hull. Although, in actual bacterial motion, the dynamics takes place in continuous time, it turns out that our calculation is much simpler in the fixed-n ensemble, than in the fixed-t ensemble. This is like the canonical versus grand-canonical ensembles. Usually, in the grand-canonical ensemble, where the number of particles fluctuates, it is often easier to calculate the observables than in the canonical ensemble where the number of particles is fixed. However, in the RTP problem, it is the opposite: it is easier in the fixed-n ensemble than in the fixed-t ensemble where n fluctuates. The two ensembles are equivalent at late times (as the canonical and the grand-canonical ensembles are equivalent in the thermodynamic limit). But they are different for finite n/finite t. In many experimental situations, the experimental timescale is limited (finite) which means it is important to know the effects of finite t. Besides, the 'active' nature of the process also shows up only at short times ($t \ll \gamma^{-1}$) or intermediate times ($t \sim \gamma^{-1}$). At large times $t \gg \gamma^{-1}$, the RTP process is essentially indistinguishable from the Brownian motion.

Recently, the mean perimeter of the convex hull for this model was computed exactly for the two ensembles and the distribution for the perimeter was numerically studied [65]. Here, we investigate the statistics of the area of the convex hull both analytically and numerically. Using connection to the extreme value statistics developed in [50, 51], we compute the mean area in the two ensembles exactly. Next, we also investigate the variance and the distribution of the typical fluctuations of the area. Our main results are summarised below:

(a) For fixed n ensemble, we find that the mean area $\langle A_n \rangle$ is given by

$$\langle A_n \rangle = \frac{v_0^2}{2\gamma^2} \mathcal{S}_n, \quad n > 1, \tag{5}$$

where the term $\mathcal{S}_n$ is given by

$$\mathcal{S}_n = \frac{2+\pi}{\sqrt{\pi}} \left[\frac{\Gamma\left(\frac{n-1}{2} - \left\lfloor \frac{n-3}{2} \right\rfloor\right)}{\Gamma\left(\frac{n}{2} - 1 + \left\lceil \frac{3-n}{2} \right\rceil\right)} + \frac{\Gamma\left(\frac{n}{2} + 1 - \left\lfloor \frac{n}{2} \right\rfloor\right)}{\Gamma\left(\frac{n+1}{2} - \left\lfloor \frac{n}{2} \right\rfloor\right)} - \frac{\Gamma\left(\frac{n+2}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)} - \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \right] \Theta(n-1)$$
$$+ \sum_{m=1}^{n-1} \frac{\Gamma\left(\frac{n-m+1}{2}\right)}{\Gamma\left(\frac{n-m+2}{2}\right)} \left[\frac{\Gamma\left(2 + \left\lfloor \frac{m-1}{2} \right\rfloor\right)}{\Gamma\left(\frac{3}{2} + \left\lfloor \frac{m-1}{2} \right\rfloor\right)} + \frac{\Gamma\left(\frac{3}{2} + \left\lfloor \frac{m}{2} \right\rfloor\right)}{\Gamma\left(1 + \left\lfloor \frac{m}{2} \right\rfloor\right)} \right]. \tag{6}$$

Here $\lfloor z \rfloor$ (or $\lceil z \rceil$) denotes the greatest (or least) integer lesser (or greater) than or equal to z. For large n, we find that $\mathcal{S}_n \simeq \pi n$ and inserting this in equation (5) yields

$$\langle A_n \rangle \simeq \frac{n\pi v_0^2}{2\gamma^2}, \quad \text{as } n \to \infty. \tag{7}$$

(b) On the other hand, for fixed t ensemble, we find that the mean area $\langle A(t)\rangle$ obeys the scaling relation

$$\langle A(t)\rangle = \frac{v_0^2}{2\gamma^2}\mathcal{J}(\gamma t), \tag{8}$$

where the scaling function $\mathcal{J}(w)$ is given by

$$\mathcal{J}(w) = \mathrm{e}^{-w}\sum_{n=2}^{\infty}\frac{\mathcal{S}_n}{\Gamma(n+2)}w^{n+1}. \tag{9}$$

The scaling function $\mathcal{J}(w)$ displays the following asymptotic behaviours:

$$\mathcal{J}(w) \simeq \frac{w^3}{3\pi} + O(w^4), \quad \text{as } w \to 0, \tag{10}$$

$$\simeq \pi w + O(\sqrt{w}), \quad \text{as } w \to \infty. \tag{11}$$

Inserting these forms in equation (8) yields that $\langle A(t)\rangle$ exhibits a crossover from $\sim t^3$ scaling for $t \ll \gamma^{-1}$ to $\sim t$ scaling for $t \gg \gamma^{-1}$:

$$\langle A(t)\rangle \simeq \frac{\gamma v_0^2 t^3}{6\pi} + O(t^4), \quad \text{for } t \ll \gamma^{-1} \tag{12}$$

$$\simeq \frac{\pi v_0^2}{2\gamma}t + O(\sqrt{t}), \quad \text{for } t \gg \gamma^{-1}. \tag{13}$$

The comparisons of the analytic expressions in equations (5) and (8) with the numerical simulation are illustrated in figure 2.

(c) We have also studied the variance and the distribution of the area in both ensembles numerically. We found that for large time (n in the fixed-n ensemble and t in the fixed-t ensemble) the variance of the area grows quadratically with time. In addition we found that the central part of the distribution (describing the typical fluctuations around the mean) for large time possesses a scaling form when the area is scaled by its mean and the scaling form matches with that of the Brownian motion, as expected.

In what follows, we derive the results for the mean area explicitly and study its typical fluctuations numerically.

3. Mean area of the convex hull

Let us begin by briefly summarising the central idea to compute the mean area of the convex hull for two dimensional stochastic processes. A more detailed account of this idea is given in [50, 51]. Based on the knowledge of the Cauchy's formulae for closed curve [66], it was shown that the mean area and the mean perimeter for random convex hulls are related to the subject of the extreme value statistics. To see this connection, consider a closed curve $\mathcal{C}$ parametrised by the points $\{(\mathcal{X}(s), \mathcal{Y}(s))\}$ on its boundary where s is the arc length. For this curve $\mathcal{C}$, we now define the support function $\mathcal{M}(\theta)$ along the direction θ (with respect to the x-axis) as

$$\mathcal{M}(\theta) = \max_{s\in\mathcal{C}}\left[\mathcal{X}(s)\cos\,\theta + \mathcal{Y}(s)\sin\,\theta\right]. \tag{14}$$

Geometrically, the support function $\mathcal{M}(\theta)$ represents the maximum extension of the curve $\mathcal{C}$ along the direction θ. Interestingly, the perimeter and the area of the domain enclosed by $\mathcal{C}$ are given in terms of $\mathcal{M}(\theta)$ by the Cauchy's formula as

$$L = \int_0^{2\pi} \mathrm{d}\theta\, \mathcal{M}(\theta), \tag{15}$$

$$A = \frac{1}{2}\int_0^{2\pi} \mathrm{d}\theta \left[\mathcal{M}^2(\theta) - \left(\mathcal{M}'(\theta)\right)^2\right], \tag{16}$$

where $\mathcal{M}'(\theta) = \frac{\mathrm{d}\mathcal{M}(\theta)}{\mathrm{d}\theta}$. To elaborate further, let us, for simplicity, consider a discrete time stochastic process of n steps. Let the positions of the particle at successive (discrete) times of a realisation be denoted by $\{(X_i, Y_i)\}$ where $i = 1, 2, \ldots, n$. We further consider that the curve $\mathcal{C}$ now represents the convex hull corresponding to the points $\{(X_i, Y_i)\}$. To construct the support function $\mathcal{M}(\theta)$ for $\mathcal{C}$, one clearly needs $\{(\mathcal{X}(s), \mathcal{Y}(s))\}$ which is a difficult task. However, it was shown in [50, 51] that this problem can be circumvented by noting the fact that $\mathcal{M}(\theta)$ is also the maximum of the projections of all points $\{(X_i, Y_i)\}$ along the direction θ. One can now write the support function $\mathcal{M}(\theta)$ as

$$\mathcal{M}(\theta) = \max_{1 \leqslant i \leqslant n} \left[X_i \cos\theta + Y_i \sin\theta\right]. \tag{17}$$

Using this form for $\mathcal{M}(\theta)$ in equations (15) and (16) and then taking average over different realisations one gets the mean perimeter and the mean area of the convex hull $\mathcal{C}$. Since, we are interested in area only, we provide below the expression of the mean area which follows directly from equation (16):

$$\langle A_n \rangle = \frac{1}{2}\int_0^{2\pi} \mathrm{d}\theta \left[\langle \mathcal{M}^2(\theta)\rangle - \langle \left(\mathcal{M}'(\theta)\right)^2\rangle\right]. \tag{18}$$

To proceed further, we assume that the maximum in equation (17) is attained in the k^*th step which enables us to write $\mathcal{M}(\theta)$ and $\mathcal{M}'(\theta)$ as

$$\mathcal{M}(\theta) = X_{k^*} \cos\theta + Y_{k^*} \sin\theta, \tag{19}$$

$$\mathcal{M}'(\theta) = -X_{k^*} \sin\theta + Y_{k^*} \cos\theta. \tag{20}$$

For isotropic processes, the support function $\langle \mathcal{M}^2(\theta)\rangle$ and $\langle \mathcal{M}'^2(\theta)\rangle$ are independent of θ and we can consider just the direction $\theta = 0$. Then, the mean area in equation (18) becomes

$$\langle A_n \rangle = \pi \left[\langle M_n^2\rangle - \langle Y_{k^*}^2\rangle(n)\right], \tag{21}$$

where $M_n = \max[X_1, X_2, \ldots, X_n]$ is the maximum displacement along the x-axis and Y_{k^*} is the abscissa at k^*th time-step at which the maximum M_n along the x-direction is reached. Later, equation (21) will be useful in calculating the mean area of the convex hull for a RTP in the fixed n ensemble.

Although equation (21) is derived for the discrete time isotropic stochastic process, one can derive an analogous formula for the mean area of $\mathcal{C}$ in the continuous time case [50]. For this case, the mean area of the convex hull reads

$$\langle A(t)\rangle = \pi \left[\langle M^2(t)\rangle - \langle Y(t_m)^2\rangle(t)\right], \tag{22}$$

where $M(t)$ is the maximum of the x-coordinate until observation time t i.e. $M(t) = \max[\{X(\tau)\}, \forall\, 0 \leqslant \tau \leqslant t]$ and t_m is the time at which this maximum $M(t)$ is reached. Also, $Y(t_m)$ represents the y-coordinate of the RTP at time t_m. Once again, equation (22) will be useful in computing the mean area for a RTP in the fixed t ensemble.

Before closing this section, we remark that the formulae of the mean area in equations (21) and (22) apply to general isotropic 2D stochastic process. In the following, we use these formulae to compute the mean area of $\mathcal{C}$ for the RTP model in two dimension. We first compute the mean area for the fixed-n ensemble and then focus on the fixed-t ensemble.

4. Mean area for fixed-*n* ensemble

Let us first look at the RTP in the fixed-n ensemble where the total number of runs n is fixed but the total time t varies for different realisations. As indicated by equation (21), we need the maximum M_n of the x-coordinate and the corresponding abscissa $Y_{k^*}(n)$ to compute the mean area $\langle A_n \rangle$. Recall from equations (3) and (4), the position coordinates of the RTP perform random walks with correlated increments (jumps) (x_i, y_i) which are given in equations (1) and (2). Also recall that we have chosen the initial position of the RTP to be $(X_0, Y_0) = (0, 0)$. Let us first compute the joint probability distribution $p(x_i, y_i, \tau_i)$ of the increments x_i and y_i and the time duration τ_i for the ith run. Since the RTP moves ballistically during time τ_i, we have $v_0\tau_i = \sqrt{x_i^2 + y_i^2}$. Also, the time τ_i is exponentially distributed $\rho(\tau_i) = \gamma\, \mathrm{e}^{-\gamma\tau_i}$. This enables us to write the joint distribution $p(x, y, \tau)$ as

$$p(x, y, \tau) = \frac{\gamma\, \mathrm{e}^{-\gamma\tau}}{\pi} \delta(v_0^2\tau^2 - x^2 - y^2), \tag{23}$$

where the factor $1/\pi$ comes from the normalisation condition. Finally, integrating $p(x, y, \tau)$ over τ, we get the joint distribution of the increments x and y as

$$p(x, y) = \frac{\gamma}{2\pi v_0 \sqrt{x^2 + y^2}} \exp\left(-\frac{\gamma}{v}\sqrt{x^2 + y^2}\right). \tag{24}$$

Notice that, the problem of the run and tumble motion now got mapped to a model of a random walk in two dimensions with the jump distribution $p(x, y)$ given in equation (24). Such mappings have been considered in [32] to study the persistent properties of the RTP. Since the jump distribution $p(x, y)$ is isotropic, we use the formula for the mean area in equation (21), true for the discrete time processes, for RTP also. In what follows, we use this equation to compute the mean area of the convex hull for the fixed-n ensemble. From equation (21), we see that this reduces to the problem of computing $\langle M_n^2 \rangle$ and $\langle Y_{k^*}^2 \rangle(n)$ which we calculate below.

4.1. Computation of $\langle M_n^2 \rangle$

In order to compute the second moment of the maximum $M_n = \max\{X_0, X_1, X_2, \ldots, X_n\}$ of the x-component of a given trajectory of n steps, we first recall that $\{X_0, X_1, X_2, \ldots, X_n\}$ denote just a one dimensional random walk trajectory such that $X_i = X_{i-1} + x_i$. The increment x_i is distributed according to the probability distribution $p_1(x_i)$ which is obtained by integrating the joint distribution $p(x_i, y_i)$ in equation (24) over all y_i. The resulting expression reads

$$p_1(x) = \int_{-\infty}^{\infty} \mathrm{d}y\, p(x, y) = \frac{\gamma}{\pi v_0} K_0\left(\frac{\gamma|x|}{v_0}\right), \tag{25}$$

where $K_\nu(z)$ is the modified Bessel function of second kind. Note that $p_1(x)$ is both symmetric and continuous. Hence the random walker is characterised by the identical and independent increments $\{x_i\}$ drawn from the symmetric and continuous distribution $p_1(x_i)$.

To calculate $\langle M_n^2 \rangle$, we use the Pollaczek–Spitzer formula [46, 67] which characterises the maximum M_n for a random walker with identical and independent increments drawn from

the symmetric and continuous distribution. If $Q_n(M) = \text{Prob}[M_n \leqslant M]$ denotes the cumulative probability of M_n, then according to the Pollaczek–Spitzer formula, $Q_n(M)$ satisfies [54, 68]:

$$\sum_{n=0}^{\infty} z^n \langle \mathrm{e}^{-\lambda M_n} \rangle = \sum_{n=0}^{\infty} z^n \int_0^{\infty} \mathrm{d}M\, \mathrm{e}^{-\lambda M} Q_n'(M) = \frac{\phi(z,\lambda)}{\sqrt{1-z}}, \tag{26}$$

where $0 \leqslant z \leqslant 1$ and $\lambda \geqslant 0$ and the function $\phi(z,\lambda)$ is defined as

$$\phi(z,\lambda) = \exp\left(-\frac{\lambda}{\pi}\int_0^{\infty} \mathrm{d}\xi \frac{\ln(1 - z\hat{p}_1(\xi))}{\lambda^2 + k^2}\right), \tag{27}$$

with $\hat{p}_1(\xi)$ being the Fourier transform of $p_1(x)$ and is given by

$$\hat{p}_1(\xi) = \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{\mathrm{i}\xi x} p_1(x) = \frac{1}{\sqrt{1+\xi^2\sigma^2}}. \tag{28}$$

We have inserted $p_1(x)$ from equation (25) in writing $\hat{p}_1(\xi)$ and defined $\sigma = \frac{v_0}{\gamma}$. One can suitably use equation (26) to compute all moments of M_n. By expanding equation (26) in λ, one finds that the generating functions of the first two moments are given by [54, 69]

$$h^{(1)}(z) = \sum_{n=0}^{\infty} z^n \langle M_n \rangle = \frac{1}{\pi(1-z)} \int_0^{\infty} \frac{\mathrm{d}\xi}{\xi^2}\, \ln\left(\frac{1 - z\hat{p}_1(\xi)}{1-z}\right), \tag{29}$$

$$h^{(2)}(z) = \sum_{n=0}^{\infty} z^n \langle M_n^2 \rangle = (1-z)\left[h^{(1)}(z)\right]^2 + \frac{\sigma^2 z}{2(1-z)^2}. \tag{30}$$

By appropriately differentiating $h^{(2)}(z)$ with respect to z, it is straightforward to show that the second moment $\langle M_n^2 \rangle$ can be expressed completely in terms of the first moment $\langle M_n \rangle$ as

$$\langle M_n^2 \rangle = \sum_{m=1}^{n-1} \langle M_m \rangle \left[\langle M_{n-m} \rangle - \langle M_{n-m-1} \rangle\right] + \frac{n\sigma^2}{2}. \tag{31}$$

Expanding the right-hand side of equation (29) one can in principle compute $\langle M_n \rangle$. Using Kac's formula [45] for the mean maximum displacement, it was recently computed explicitly in [65] to be

$$\langle M_n \rangle = \frac{\sigma}{2\sqrt{\pi}} \sum_{j=1}^{n} \frac{\Gamma\left(\frac{j+1}{2}\right)}{\Gamma\left(\frac{j+2}{2}\right)}. \tag{32}$$

Using this expression, we get

$$\langle M_{n-m} \rangle - \langle M_{n-m-1} \rangle = \frac{\sigma}{2\sqrt{\pi}} \frac{\Gamma\left(\frac{n-m+1}{2}\right)}{\Gamma\left(\frac{n-m+2}{2}\right)}, \quad n > 1. \tag{33}$$

Finally inserting equations (32) and (33) in the expression of $\langle M_n^2 \rangle$ in equation (31) we get

$$\langle M_n^2 \rangle = \frac{v_0^2}{2\gamma^2}\left(\frac{\mathcal{S}_n}{\pi} + n\right), \quad n > 1, \tag{34}$$

where the quantity $\mathcal{S}_n$ is defined as

$$\mathcal{S}_n = \frac{\sqrt{\pi}}{\sigma}\sum_{m=1}^{n-1}\frac{\Gamma\left(\frac{n-m+1}{2}\right)}{\Gamma\left(\frac{n-m+2}{2}\right)}\langle M_m\rangle. \tag{35}$$

Inserting $\langle M_m\rangle$ from equation (32) in the above equation and simplifying, one gets the explicit expression of $\mathcal{S}_n$ given in equation (6).

4.2. Computation of $\langle Y_{k^*}^2\rangle(n)$

We now compute the other term $\langle Y_{k^*}^2\rangle(n)$ in the expression of the mean area in equation (21). To calculate this, we first compute the joint distribution $\mathcal{P}(Y,k^*|n) = \text{Prob.}[Y_{k^*} = Y, k^*|n]$ and then compute the second moment of the displacement Y_{k^*} of the particle along y direction at step k^* in which the particle reaches its maximum along the x direction in a walk of n-steps. It is possible to show that one can compute this joint distribution for a general 2D discrete time random walk where the position coordinates (X_i, Y_i) at ith step evolve, starting from $X_0 = 0, Y_0 = 0$, as $X_i = X_{i-1} + x_i$ and $Y_i = Y_{i-1} + y_i$ with the jump increments (x_i, y_i) at different steps drawn independently from the common distribution $p(x,y)$. Note that at a given step, the increments x and y can be correlated.

To proceed let us define marginal distribution of the y-increment

$$p_2(y) = \int_{-\infty}^{\infty} \mathrm{d}x\, p(x,y), \tag{36}$$

and the marginal distribution of the y-coordinate Y_k of the walker at step k

$$P(Y,k) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\mathrm{d}y_1\,\mathrm{d}y_2\ldots\mathrm{d}y_k\,\delta\left(Y-\sum_{i=1}^{k}y_i\right)\prod_{i=1}^{k}p_2(y_i). \tag{37}$$

Now consider any trajectory in two dimensions up to n steps. Let k^* denote the time at which the x-coordinate achieves its maximum M_n and Y_{k^*} denote the y-coordinate exactly at step k^*. Recall, we want to compute the joint distribution of Y_{k^*} and k^*, given the total number of steps n, i.e., $\mathcal{P}(Y,k^*|n)$. We show that this joint distribution is given by

$$\mathcal{P}(Y,k^*|n) = q_{k^*}\, q_{n-k^*} P(Y,k^*), \tag{38}$$

where $P(Y,k^*)$ is defined in equation (37) and $q_n = \binom{2n}{n}\,2^{-2n}$ is the Sparre Andersen survival probability of a 1D random walk with arbitrary symmetric and continuous jump distribution [69, 70].

To prove the claim in equation (38), we start with the joint probability distribution for $M_n = M,\ Y_{k^*} = Y$ and k^* denoted by $\mathscr{P}(M,Y,k^*|n)$. This joint probability distribution of the n-step walk can be expressed as a multi-dimensional integral

$$\mathscr{P}(M,Y,k^*|n) = \int \vec{\mathrm{d}x}\,\vec{\mathrm{d}y}\,\mathcal{Z}_{k^*}\left(M,Y,\{x_i\},\{y_i\}\right)\prod_{j=1}^{n}p(x_j,y_j), \tag{39}$$

with $\mathcal{Z}_{k^*}\left(M,Y,\{x_i\},\{y_i\}\right)$ defined as

$$\mathcal{Z}_{k^*}(M,Y,\{x_i\},\{y_i\}) = \left[\prod_{j\neq k^*,j=0}^{n}\Theta(M-X_j)\right]\delta(M-X_{k^*})\delta(Y-Y_{k^*}), \tag{40}$$

where $X_i = \sum_{j=1}^{i} x_j$ and $Y_i = \sum_{j=1}^{i} y_j$ (recall we have chosen $X_0 = Y_0 = 0$). Here $\Theta(n)$ is the Heaviside theta function. The function $\mathcal{Z}_{k^*}\left(M, Y, \{x_i\}, \{y_i\}\right)$ ensures that $X_{k^*} = M$ and $Y_{k^*} = Y$ while all other $\{X_i\}$ are smaller than M. Finally, we integrate over all $\{x_i, y_i\}$ with appropriate joint distribution $\prod_{i=1}^{n} p(x_i, y_i)$. For simplicity, we have used the short-hand notation $\vec{\mathrm{d}x} = \mathrm{d}x_1\, \mathrm{d}x_2 \ldots \mathrm{d}x_n$ and $\vec{\mathrm{d}y} = \mathrm{d}y_1\, \mathrm{d}y_2 \ldots \mathrm{d}y_n$.

Since we are interested in the joint distribution $\mathcal{P}(Y, k^*|n)$ of $Y_{k^*} = Y$ and k^*, we integrate $\mathscr{P}(M, Y, k^*|n)$ in equation (39) over M i.e.

$$\mathcal{P}(Y, k^*|n) = \int_0^\infty \mathrm{d}M\ \mathscr{P}(M, Y, k^*|n). \tag{41}$$

To proceed further we take the Fourier transformation with respect to Y

$$\bar{\mathcal{P}}\left(\xi, k^*|n\right) = \int_{-\infty}^{\infty} \mathrm{d}Y\, \mathrm{e}^{\mathrm{i}\xi Y} \int_0^\infty \mathrm{d}M\ \mathscr{P}(M, Y, k^*|n), \tag{42}$$

and perform some algebraic simplications in equation (39). We relegate the details of the calculations to appendix A and present only the final result here. The final expression reads

$$\bar{\mathcal{P}}\left(\xi, k^*|n\right) = q_{k^*} q_{n-k^*}\ [\hat{p}_2(\xi)]^{k^*}, \tag{43}$$

where $\hat{p}_2(\xi)$ represents the Fourier transformation of the marginal distribution $p_2(y)$ of y-increment (see equation (36)) and is defined by

$$\hat{p}_2(\xi) = \int_{-\infty}^{\infty} \mathrm{d}y\, \mathrm{e}^{\mathrm{i}\xi y} p_2(y). \tag{44}$$

The term $q_n = \binom{2n}{n}\ 2^{-2n}$ in equation (43), as mentioned earlier, is the survival probability of a random walker in one dimension starting from the origin and with jumps drawn independently from a symmetric and continuous distribution. Note that the term $[\hat{p}_2(\xi)]^{k^*}$ in equation (43) is actually the Fourier transform of the marginal distribution $P(Y, k^*)$ of the y-coordinate at step k^* as can be easily seen from equation (37). Hence, performing inverse Fourier transform on both sides of equation (43) one arrives at the result in equation (38). Notice that the expression of $\mathcal{P}(Y, k^*|n)$ in equation (38) appears naturally in the following form:

$$\mathcal{P}(Y, k^*|n) = \text{Prob.}[Y_{k^*} = Y] \times\ \text{Prob.}[M_n \text{ occursatstep } k^* \quad \text{in } n - \text{stepwalk}]. \tag{45}$$

This result is quite universal and holds true for any joint distribution $p(x, y)$ as long as it is symmetric and continuous in x. This universality is a consequence of Sparre Andersen theorem [69, 70]. In addition if the joint distribution $p(x, y)$ is such that $\langle y^2 \rangle$ is finite, one finds

$$\langle Y_{k^*}^2 \rangle(n) = \sum_{k^*=1}^{n} \langle Y_{k^*}^2 \rangle\ \times\ q_{k^*} q_{n-k^*} \tag{46}$$

$$= . \sum_{k^*=1}^{n} \langle y^2 \rangle\ k^*\ q_{k^*} q_{n-k^*} \tag{47}$$

$$= \langle y^2 \rangle\ \frac{n}{2}, \tag{48}$$

where we have used $q_k = 2^{-2k} \binom{2k}{k}$. Note that this result is also universal.

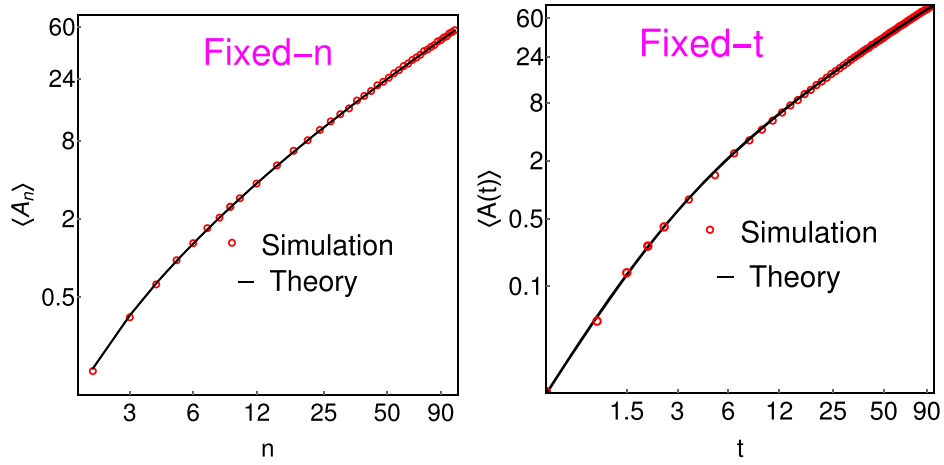

Figure 2. Comparison of the mean area of the convex hull for a RTP in fixed-n ensemble (left) and fixed-t ensemble (right) with the numerical simulations in the log–log scale. The corresponding analytic expressions are given in equations (5) and (8) respectively. For both panels, we have chosen $v_0 = 1$, $\gamma = 1.5$.

Recall, in this paper, we are interested to compute $\langle Y_{k^*}^2 \rangle(n)$ for the RTP in which case the joint distribution $p(x, y)$ is given in equation (24). For this distribution one has

$$p_2(y) = \frac{\gamma}{\pi v_0} K_0\left(\frac{\gamma|y|}{v_0}\right), \tag{49}$$

with $\langle y^2 \rangle = \frac{v_0^2}{\gamma^2}$ which gives

$$\langle Y_{k^*}^2 \rangle(n) = \frac{v_0^2}{2\gamma^2} n. \tag{50}$$

4.3. Mean area $\langle A_n \rangle$

Substituting $\langle Y_{k^*}^2 \rangle(n)$ from equation (50) along with $\langle M_n^2 \rangle$ from equation (34) in the expression of $\langle A_n \rangle$ in equation (21), we obtain the mean area of the convex hull for a 2D isotropic run and tumble motion in the fixed-n ensemble as quoted in equation (5). For large n, we find $\mathcal{S}_n \simeq \pi n$ (see appendix B for proof) which yields the asymptotic form of $\langle A_n \rangle$ in equation (5) as

$$\langle A_n \rangle \simeq \frac{n\pi}{2}\sigma^2, \quad \text{as } n \to \infty. \tag{51}$$

This matches with the mean area of the convex hull of a discrete two dimensional random walk of n steps for any jump distribution with a finite variance σ^2 [54]. In figure 2 (left panel), we have compared our analytic result of $\langle A_n \rangle$ in equation (5) with the simulation results. We observe an excellent agreement between them. A summary of the numerical scheme adopted to construct the convex hull is provided in appendix C.

To compare the mean area for different parameters, we rescale $\langle A_n \rangle$ in equation (5) with σ^2 where $\sigma = v_0/\gamma$. Moreover, from equation (51), we see that $\langle A_n \rangle$ scales linearly with n for

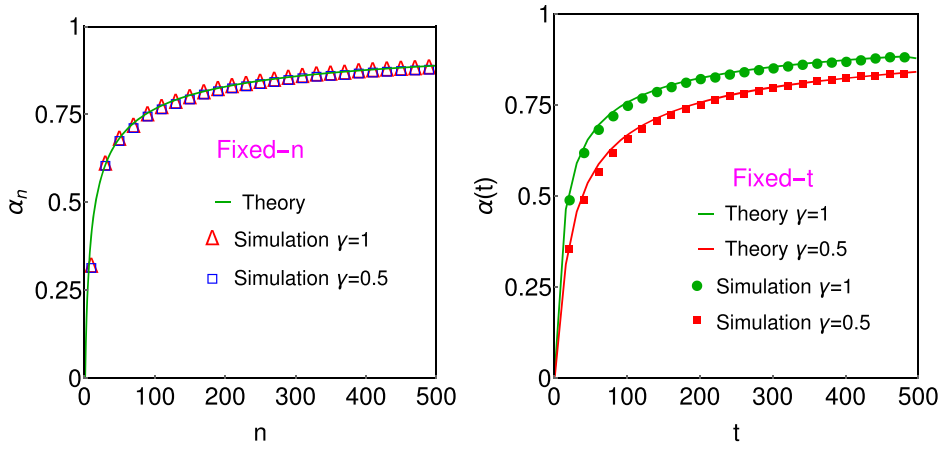

Figure 3. Comparison of the mean area for fixed-n (left) and fixed-t (right) ensembles with the numerical simulations. For fixed-n ensemble (left), we have denoted $\alpha_n = 2\langle A_n\rangle/n\pi\sigma^2$ with $\sigma = v_0/\gamma$ and $\langle A_n\rangle$ given in equation (5). On the other hand, for fixed-t ensemble (right), we have introduced the notation $\alpha(t) = 2\langle A(t)\rangle/\pi\gamma t\sigma^2$ with $\langle A(t)\rangle$ given in equation (8). For both panels, we have used $v_0 = 1$.

$n \gg 1$. Therefore, we also rescale $\langle A_n\rangle$ with n to remove the asymptotic growth with respect to n for proper visualisation. Defining

$$\alpha_n = \frac{2\langle A_n\rangle}{n\pi\sigma^2}, \tag{52}$$

one expects α_n to be independent of v_0 and γ via equation (5). Also, α_n should saturate to the value 1 for $n \to \infty$. Indeed, in figure 3 (left panel), we observe that α_n is identical for two different values of γ, namely $\gamma = 1$ and $\gamma = 0.5$. Moreover, it approaches the value $\alpha_n \to 1$ as we go to higher values of n. This comparison of α_n for two different values of γ provides another verification of $\langle A_n\rangle$ in equation (5).

5. Fixed-*t* ensemble

The previous section dealt with the mean area of the convex hull for a RTP in the fixed-n ensemble. We now consider the mean area in the fixed-t ensemble where the observation time t is fixed but the number of runs n varies from sample to sample. For this case also, we show that the run and tumble model can be suitably mapped to a random walker in two dimensions which is then used to calculate exactly the mean area $\langle A(t)\rangle$ via equation (22). To begin with, let us consider a realisation of the RTP with n runs where the ith run lasts for time τ_i with position increments x_i and y_i. Since at the end of each run except the nth one, the RTP encounters a tumbling, the times $\{\tau_i\}$for $1 \leqslant i \leqslant (n-1)$ are all drawn independently from the exponential distribution $\rho(\tau_i) = \gamma\, \mathrm{e}^{-\gamma\tau_i}$. Therefore the joint distribution $p(x_i, y_i, \tau_i)$ with $1 \leqslant i \leqslant (n-1)$ is given by equation (23). On the other hand, during the last interval τ_n, the RTP does not encounter any tumble: the probability of which is $\mathrm{e}^{-\gamma\tau_n}$. Hence, the corresponding joint distribution is

$$p_{\rm last}(x_n, y_n, \tau_n) = \frac{\mathrm{e}^{-\gamma\tau_n}}{\pi}\delta(v_0^2\tau_n^2 - x_n^2 - y_n^2) = \frac{1}{\gamma}\, p(x_n, y_n, \tau_n), \tag{53}$$

where $p(x_n, y_n, \tau_n)$ is given by equation (23). We emphasise that unlike in the fixed-n case, here the runs are correlated due to the constraint of fixed t. To see this more clearly, we write the grand joint distribution of $\{x_i\}, \{y_i\}$ and n below:

$$P\left(\{x_i\}, \{y_i\}, n|t\right) = \int_0^t \mathrm{d}\tau_1 \int_0^t \mathrm{d}\tau_2 \ldots \int_0^t \mathrm{d}\tau_n \frac{1}{\gamma} \left[\prod_{i=1}^n p(x_i, y_i, \tau_i)\right] \delta\left(\sum_{i=1}^n \tau_i - t\right). \quad (54)$$

To get rid of the δ-function, we take the Laplace transformation with respect to t ($\to s$)

$$\int_0^\infty \mathrm{d}t\, \mathrm{e}^{-st} P\left(\{x_i\}, \{y_i\}, n|t\right) = \frac{1}{\gamma} \left[\prod_{i=1}^n \frac{\gamma \exp\left(-\frac{(\gamma+s)}{v}\sqrt{x_i^2+y_i^2}\right)}{2\pi v_0 \sqrt{x_i^2+y_i^2}}\right], \quad (55)$$

which we rewrite as

$$\int_0^\infty \mathrm{d}t\, \mathrm{e}^{-st} P\left(\{x_i\}, \{y_i\}, n|t\right) = \frac{1}{\gamma}\left(\frac{\gamma}{\gamma+s}\right)^n \left[\prod_{i=1}^n g_s(x_i, y_i)\right], \quad (56)$$

$$\text{with} \quad g_s(x,y) = \frac{(\gamma+s)\exp\left(-\frac{(\gamma+s)}{v}\sqrt{x_i^2+y_i^2}\right)}{2\pi v_0 \sqrt{x^2+y^2}}. \quad (57)$$

Finally, inverting the Laplace transform in equation (56), the grand joint distribution $P\left(\{x_i\}, \{y_i\}, n|t\right)$ can be formally written as

$$P\left(\{x_i\}, \{y_i\}, n|t\right) = \int_\Gamma \frac{\mathrm{d}s}{2\pi \mathrm{i}} \mathrm{e}^{st} \frac{1}{\gamma}\left(\frac{\gamma}{\gamma+s}\right)^n \left[\prod_{i=1}^n g_s(x_i, y_i)\right], \quad (58)$$

where Γ is the Bromwich contour in the complex s plane. Like in the fixed-n ensemble, here also, we find that the jump distributions in equation (58), although correlated, are isotropic which enables us to use the Cauchy's formula for the mean area in equation (22). For this, we first note that the function $g_s(x, y)$ given in equation (57) can be interpreted as a probability distribution as it is positive over full (x, y) plane and normalised to unity. As a result the term inside the square bracket in the integrand of the equation (58) can be interpreted as the joint distribution of the increments x_i and y_i of a random walker in two dimension in steps $i = 1, 2, \ldots, n$. In the context of RTP such mapping to random walk problem was observed earlier [35, 65] and exploited to study the survival probability in higher dimension [35]. In this paper we follow a similar calculation using this mapping and compute the mean area $\langle A(t)\rangle$ of the convex hull by employing the formula in equation (22). As seen in this formula, we then need to calculate $\langle M^2(t)\rangle$ and $\langle Y(t_m)^2\rangle(t)$. In the following, we use the joint distribution $P\left(\{x_i\}, \{y_i\}, n|t\right)$ in equation (58) to calculate these two quantities explicitly.

5.1. Computation of $\langle M^2(t)\rangle$

Let us begin with the computation of $\langle M^2(t)\rangle$ where $M(t)$ is the maximum of the x-coordinate of the RTP up to observation time t. For this, we need to compute the statistics of the maximum of a 1D random walker for fixed t with n jumps in the x-coordinate: $\{x_i\}$ for $i = 1, 2, \ldots, n$. The joint distribution of the increments $\{x_i\}$ can be obtained by integrating $P\left(\{x_i\}, \{y_i\}, n|t\right)$ in equation (58) over all $\{y_i\}$ as

$$P_x\left(\{x_i\},n|t\right) = \int_{-\infty}^{\infty} \mathrm{d}y_1\,\mathrm{d}y_2\ldots\mathrm{d}y_n\,P\left(\{x_i\},\{y_i\},n|t\right), \tag{59}$$

$$= \int_{\Gamma}\frac{\mathrm{d}s}{2\pi\mathrm{i}}\,\mathrm{e}^{st}\,\frac{1}{\gamma}\left(\frac{\gamma}{\gamma+s}\right)^n\left[\prod_{i=1}^{n} g_s(x_i)\right], \tag{60}$$

$$\text{with}\quad g_s(x) = \int_{-\infty}^{\infty}\mathrm{d}y\,g_s(x,y) = \frac{(\gamma+s)}{\pi v_0}K_0\left(\frac{(\gamma+s)|x|}{v_0}\right). \tag{61}$$

Using this expression of $P_x\left(\{x_i\},n|t\right)$, we now proceed to calculate the statistics of the maximum $M(t)$. To this end, we define $Q(M,n|t)$ as the probability that $X_i < M$ for $1 \leqslant i \leqslant n$, where $X_i = \sum_{j=1}^{i} x_j$. It is easy to realise that $Q(M,n|t)$ is actually the survival probability that the walker with n steps up to time t has not crossed $X = M$. Formally, this is given as

$$Q(M,n|t) = \int_{-\infty}^{\infty}\mathrm{d}x_1\ldots\int_{-\infty}^{\infty}\mathrm{d}x_n\,\text{Prob.}\left[X_1 < M, X_2 < M,\ldots,X_n < M,n|t\right], \tag{62}$$

$$= \int_{-\infty}^{\infty}\mathrm{d}x_1\ldots\int_{-\infty}^{\infty}\mathrm{d}x_n\,\Theta(M-X_1)\ldots\Theta(M-X_n)P_x\left(\{x_i\},n|t\right). \tag{63}$$

Note that $Q(M,n|t)$ is also the probability that the maximum displacement of the 1D random walk with n steps up to time t is less than or equal to M. Differentiating $Q(M,n|t)$ with M gives the joint probability distribution for M and n which can then be used to calculate $\langle M^2(t)\rangle$. The formal expression of $\langle M^2(t)\rangle$ reads

$$\langle M^2(t)\rangle = \sum_{n=1}^{\infty}\int_0^{\infty}\mathrm{d}M\,M^2\partial_M Q(M,n|t), \tag{64}$$

$$= \int_{\Gamma}\frac{\mathrm{d}s}{2\pi\mathrm{i}}\,\mathrm{e}^{st}\,\frac{1}{\gamma}\left(\frac{\gamma}{\gamma+s}\right)^n\langle M_s^2(n)\rangle, \tag{65}$$

where $\langle M_s^2(n)\rangle$ is

$$\langle M_s^2(n)\rangle = \int_0^{\infty}\mathrm{d}M\,M^2\partial_M Q_s(M,n),\quad \text{with} \tag{66}$$

$$Q_s(M,n) = \int_{-\infty}^{\infty}\mathrm{d}x_1\ldots\int_{-\infty}^{\infty}\mathrm{d}x_n\,\Theta(M-X_1)\ldots\Theta(M-X_n)\left[\prod_{i=1}^{n} g_s(x_i)\right]. \tag{67}$$

Here $\int_{-\infty}^{\infty}\mathrm{d}x\,g_s(x) = 1$ which can be verified easily from equation (61). Hence $Q_s(M,n)$ can be deciphered as the cumulative distribution that the maximum is less than M up to n steps for an auxilliary 1D random walk with identical and independent jumps which follow the symmetric and continuous distribution $g_s(x)$ given in equation (61). Consequently, $\langle M_s^2(n)\rangle$ is the second moment of the maximum $M_s(n)$ of the auxillary random walk which can be calculated using the Pollaczek Spitzer formulae in equations (29) and (30) as done for the fixed-n ensemble. To avoid repetition, we present the details of this calculation in appendix D and write only the final expression of $\langle M_s^2(n)\rangle$ here which reads

$$\langle M_s^2(n)\rangle = \frac{v_0^2}{2(\gamma+s)^2}\left(\frac{\mathcal{S}_n}{\pi}+n\right), \tag{68}$$

where $\mathcal{S}_n$ is given in equation (6). Substituting $\langle M_s^2(n)\rangle$ in the expression of $\langle M^2(t)\rangle$ in equation (65) and performing the inverse Laplace transformation gives

$$\langle M^2(t)\rangle = \frac{v_0^2}{2\gamma^2}\left[\mathrm{e}^{-\gamma t} - 1 + \gamma t + \frac{\mathrm{e}^{-\gamma t}}{\pi}\sum_{n=1}^{\infty}\frac{\mathcal{S}_n}{\Gamma(n+2)}(\gamma t)^{n+1}\right]. \tag{69}$$

5.2. Computation of $\langle Y(t_m)^2\rangle(t)$

We next calculate $\langle Y(t_m)^2\rangle(t)$ for the mean area $\langle A(t)\rangle$ in equation (22). Recall that $Y(t_m|t)$ is the y-coordinate of the RTP at time t_m when the maximum $M(t)$ of the x-coordinate is attained in a trajectory of duration t. To calculate $\langle Y(t_m)^2\rangle(t)$, we first notice that the maximum in the x-direction occurs at the end of some complete jump step, say k^* which is a function of the total number of jumps n occurring in time t. Of course the number of jumps n is a random quantity and consequently so is k^* as they change from realisation to realisation and also they are functions of t. Hence denoting the time at the end of step k^* by t_m, we can write $Y(t_m) = \sum_{i=1}^{k^*} y_i$.

We start with the grand joint distribution $P(\{x_i\},\{y_i\},n|t)$ given in equation (58). As we have mentioned earlier, the term $\prod_{i=1}^{n} g_s(x_i,y_i)$ inside the square bracket on the right-hand side of this equation can be interpreted as the joint probability distribution of the jumps x_i and y_i for $i = 1,2,\ldots,n$ of a random walk in two dimension of n steps. Once again we emphasise that $g_s(x,y)$, given explicitly in equation (57), can be interpreted as an effective joint distribution of elementary jumps along x and y directions, similar to $p(x,y)$ as considered earlier in section 4 except now it is parametrised by s. As a result we see that for a given trajectory of duration t containing n jump steps there is a trajectory of n jumps generated by the joint distribution $\prod_{i=1}^{n} g_s(x_i,y_i)$. Hence, if the maximum displacement in the x-direction occurs at step k^* of a trajectory of duration t containing n jump steps, then in the auxiliary random walk problem generated by $g_s(x,y)$ the maximum displacement along x-direction occurs at the same step k^*. Moreover the displacements X_i and Y_i (starting from the origin) along x and y-directions at ith step are exactly same for $i = 1,2,\ldots,n$. Hence, we have

$$\langle Y(t_m)^2\rangle(t) = \int_{\Gamma}\frac{\mathrm{d}s}{2\pi\mathrm{i}}\mathrm{e}^{st}\,\frac{1}{\gamma}\sum_{n=1}^{\infty}\left(\frac{\gamma}{\gamma+s}\right)^n \langle Y_{k^*}^2\rangle_s(n), \tag{70}$$

where $\langle Y_{k^*}^2\rangle_s(n)$ should be computed following the procedure given in section 4.2 with only difference being the joint distribution $p(x,y)$ is replaced by $g_s(x,y)$ which is given in equation (57). That is why we now have a subscript s in the notation of $\langle Y_{k^*}^2\rangle_s(n)$. Executing the computation steps from equations (46)–(48) with $g_s(x,y)$ we get

$$\langle Y_{k^*}^2\rangle_s(n) = \frac{v_0^2}{(\gamma+s)^2}\frac{n}{2}, \tag{71}$$

where we have used $\langle y^2\rangle_{g_s} = \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\, y^2\, g_s(x,y) = \frac{v_0^2}{(\gamma+s)^2}$. Inserting the above expression from equation (71) in equation (70) and carrying out the sum over n we get

$$\langle Y(t_m)^2\rangle(t) = \frac{v_0^2}{2}\int_{\Gamma}\frac{\mathrm{d}s}{2\pi\mathrm{i}}\frac{\mathrm{e}^{st}}{s^2(\gamma+s)}, \tag{72}$$

which upon performing inverse Laplace transformation with respect to s gives the final expression

$$\langle Y(t_m)^2\rangle(t) = \frac{v_0^2}{2\gamma^2}\left(\gamma t - 1 + \mathrm{e}^{-\gamma t}\right). \tag{73}$$

5.3. Mean area for fixed-t ensemble

The expressions of $\langle M^2(t)\rangle$ and $\langle Y(t_m)^2\rangle(t)$ in equations (69) and (73) respectively guide us to write the mean area $\langle A(t)\rangle$ via equation (22). Inserting these forms explicitly, it is straightforward to show that $\langle A(t)\rangle$ indeed possesses the scaling form of equation (8) with the scaling function $\mathcal{J}(w)$ given in equation (9). In figure 2 (right panel), we have plotted $\langle A(t)\rangle$ and compared it against the numerical simulations. We observe an excellent agreement.

To get the Brownian limit of the expression of $\langle A(t)\rangle$ in equation (8), we look at the asymptotic behaviours of the scaling function $\mathcal{J}(w)$ which read

$$\mathcal{J}(w) \simeq \frac{w^3}{3\pi} + O(w^4), \quad \text{as } w \to 0, \tag{74}$$

$$\simeq \pi w + O(\sqrt{w}), \quad \text{as } w \to \infty. \tag{75}$$

Inserting these forms in equation (8), we find that $\langle A(t)\rangle$ exhibits a crossover from $\sim t^3$ scaling for $t \ll \gamma^{-1}$ to $\sim t$ scaling for $t \gg \gamma^{-1}$:

$$\langle A(t)\rangle \simeq \frac{\gamma v_0^2 t^3}{6\pi} + O(t^4), \quad \text{for } t \ll \gamma^{-1} \tag{76}$$

$$\simeq \frac{\pi v_0^2}{2\gamma} t + O(\sqrt{t}), \quad \text{for } t \gg \gamma^{-1}. \tag{77}$$

For $t \gg \gamma^{-1}$, we recover the result for Brownian motion with effective diffusion constant $D = v_0^2/2\gamma$. The large-t behaviour of $\langle A(t)\rangle$ in equation (77) can also be understood from the large-n behaviour of $\langle A_n\rangle$ in equation (5). For large t, the number of tumbling events experienced by the particle typically scales with time as $n \simeq \gamma t$. Plugging this in equation (5) directly yields the asymptotic form in equation (13). However, our study goes beyond this asymptotic behaviour and also gives the mean area for small and intermediate values of t and n where the effect of activity is strong. At small times $t \ll \gamma^{-1}$, the behaviour is remarkably different than that of the Brownian motion as illustrated by the $\sim t^3$ growth in equation (76). To understand this cubic growth, we proceed as follows. Recall that the particle typically takes time $\tau_{\rm tum}$ of order $\sim \gamma^{-1}$ to experience a tumble. However, the value of $\tau_{\rm tum}$ fluctuates from realisation to realisation. In fact, for some realisation, it can be quite smaller than γ^{-1} (the distribution of $\tau_{\rm tum}$ is $p(\tau_{\rm tum}) = \gamma\, \mathrm{e}^{-\gamma\tau_{\rm tum}}$ which is peaked at $\tau_{\rm tum} = 0$). The cubic growth of the mean area at smaller timescales arises essentially from those trajectories for which one also has $\tau_{\rm tum} \ll \gamma^{-1}$. The minimum number of tumbling required to construct a convex hull is two up to time t (counting the starting point as a tumble). Then, the convex hull is essentially a triangle with two sides of length $v_0\tau$ and $v_0(t-\tau)$ and some angle ζ between them. The area is given by $A(t) = |\frac{v_0^2\tau(t-\tau)\sin\,\zeta}{2}|$. To calculate mean, we recall that τ is drawn from the exponential distribution $p(\tau) = \gamma\, \mathrm{e}^{-\gamma\tau}$ and the angle ζ is chosen uniformly from $[0, 2\pi]$. Although at large times, $\langle A(t)\rangle$ behaves identical to that of the Brownian motion, the short time behaviour is rather different. Another way to demonstrate this difference is to define

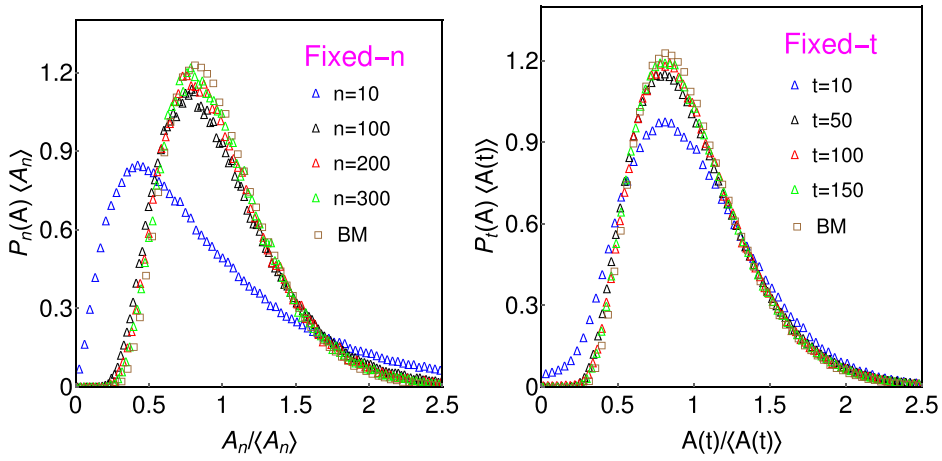

Figure 4. (Left) Simulation results for the distribution of the area A_n for fixed-n ensemble for different values of n. We have rescaled the distribution with mean area $\langle A_n\rangle$ and compared it with that of the Brownian motion. We have chosen $v_0 = 1$ and $\gamma = 1$ for all values of n. (Right) The same analysis is conducted for the fixed-t ensemble with same choice of parameters.

$$\alpha(t) = \frac{2\langle A(t)\rangle}{\gamma t \pi \sigma^2}. \tag{78}$$

For $t \to \infty$, $\alpha(t)$ saturates to the value 1. In figure 3 (right panel), we have plotted $\alpha(t)$ for two different values of γ and also compared them against the numerical simulations. We see agreement of the numerical data to the analytic expressions for both cases. Also, we obtain that $\alpha(t)$ approaches the value 1 in both cases.

6. Numerical study of the probability distribution

In the previous sections, we explicitly derived the exact analytical expressions of the mean area of the convex hull in the fixed-n and fixed-t ensembles and compared them against the numerical simulations. These expressions are given respectively in equations (5) and (8). We now investigate the probability distribution of the area of the convex hull for a single RTP. Deriving analytic forms of the distribution seems a difficult problem. In view of this, we perform a rigorous numerical study for the distribution in the two statistical ensembles. Here, we only look at the distribution corresponding to the typical fluctuations in area. By this, we mean the parts of distribution that lie within few standard deviations around the mean. To compare the distribution for different values of n or t, it turns out useful to rescale it with the mean area. In figure 4, we have illustrated the simulation data for the rescaled distribution for different values of n and t. For both ensembles, we find that the distribution converges to that of the Brownian motion in the asymptotic regime, i.e. $n \gg 1$ for fixed-n ensemble and $t \gg \gamma^{-1}$ for fixed-t ensemble. However, for other (small and intermediate) values of n and t, we expectedly see clear departure from the Brownian motion as elucidated by blue symbols in both panels of figure 4. We refer to appendix C for a discussion on the numerical scheme for constructing the histograms in figure 4.

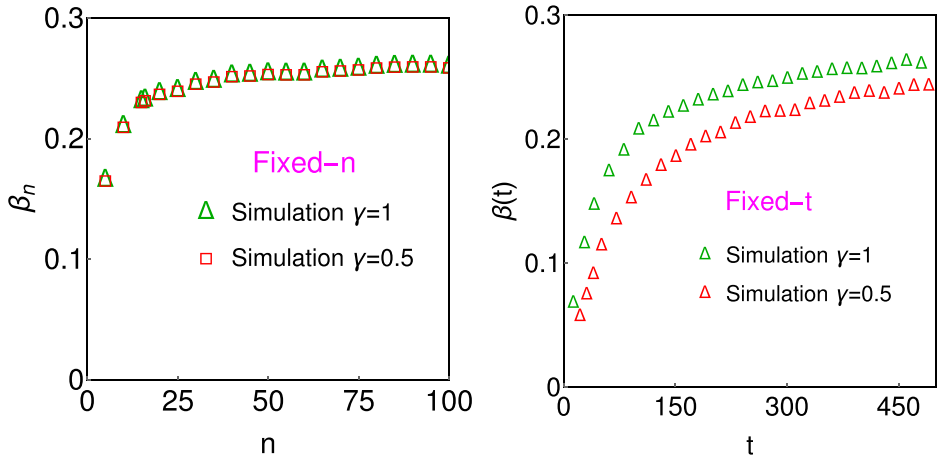

Figure 5. Simulation data for the variance of area for fixed-n (left) and fixed-t (right) ensembles. For fixed-n ensemble (left), we have plotted $\beta_n = \mathrm{Var}(A_n)/n^2\sigma^4$ vs n and for fixed-t ensemble (right), we have plotted $\beta(t) = \mathrm{Var}(A(t))/(\gamma t)^2\sigma^4$ vs t. Parameters used in these plots are $v_0 = 1$, $\gamma = 1$ (green) and $\gamma = 0.5$ (red) for both panels.

Similarly, we have also studied the variance of the area in figure 5 for two ensembles. As done for the mean area in equations (52) and (78), we define the following two quantities:

$$\beta_n = \frac{\mathrm{Var}(A_n)}{n^2\sigma^4}, \quad \text{for fixed-}n \tag{79}$$

$$\beta(t) = \frac{\mathrm{Var}(A(t))}{(\gamma t)^2\sigma^4}. \quad \text{for fixed-}t. \tag{80}$$

As seen before, this rescaling of the variance helps in better visualisation of the data since all of them converge to the same value in the asymptotic regime for both ensembles and for different values of the parameters. For both ensembles, we see in figure 5 that β_n and $\beta(t)$ tend towards the same value for different values of γ.

7. Conclusion

We have investigated the area of the convex hull of a RTP in two dimensions. We have considered this problem in two different ensembles: (i) fixed-n ensemble and (ii) fixed-t ensemble. We have obtained explicit expressions of the mean area $\langle A_n\rangle$ and $\langle A(t)\rangle$ in these two ensembles and verified them numerically. To study mean area analytically, we have used a mapping of the run-and-tumble motion to a random walk model in two dimensions similar to what was used previously in [35]. After exploiting the connection between the extreme value statistics and the computation of the mean area through Cauchy's formulae (equations (21) and (22)), we use this mapping to employ the Sparre Andersen theorem which finally leads us to arrive at the explicit expressions of the mean area in equations (5) to (9). We observed that at large times the mean area grows linearly whereas at small times it grows as $\sim t^3$ with t. We have obtained a scaling function that describes this crossover from the cubic growth to the linear growth around the natural time scale γ^{-1} provided by the tumbling rate.

Obtaining analytical results for the higher order moments and the distribution seems a challenging task. We have numerically studied the variance of the mean area as a function of time

(number of tumbles n in the fixed run ensemble and t in the fixed time ensemble) and found that it grows quadratically with time at long times. We have also studied the distribution of the area numerically. While in the asymptotic regime, i.e. $n \gg 1$ for the fixed-n ensemble and $t \gg \gamma^{-1}$ for the fixed-t ensemble, the distribution converges to that of the Brownian motion when area is scaled with its mean, we find a clear difference at small or intermediate regimes.

As mentioned before, computing the higher order moments and the full distribution of the area is a challenging problem and still remains an open problem even for a Brownian particle. We believe that our work may have potential biological applications in estimating the spatial extent over which bacteria like *E. Coli* move, since they exhibit run and tumble dynamics [19]. However, here, we have looked at the simple version of this model where tumbles are instantaneous. On the other hand, it has been experimentally found that active particles in reality spend small but non-zero time while tumbling [19, 71]. For fixed-n case, we expect the results in these realistic systems with non-instantaneous tumblings to be same as the simple model (with instantaneous tumblings) considered in our work. However, for fixed-t case, the results will be different atleast when the tumbling timescale is comparable to or larger than the run timescale. Extending our results for these realistic systems remains a promising future direction. Finally, in this work, we have focused on one class of the active particles called RTPs. It would be interesting to explore how our results get generalised for other models of the active particles like active Brownian particle and active Ornstein–Uhlenbeck particle [3, 72].

Acknowledgments

AK and PS acknowledge support of the Department of Atomic Energy, Government of India, under Project No. 12 R & D-TFR-5.10-1100. AK acknowledges support from DST, Government of India grant under Project No. ECR/2017/000634. We thank B De Bruyne and F Mori for useful discussions.

Data availability statement

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

Appendix A. Derivation of $\bar{\mathcal{P}}(\xi, k^*|n)$ in equation (43)

In this appendix we derive the expression of $\bar{\mathcal{P}}\left(\xi, k^*|n\right)$ given in equation (43). To begin with, we start with the joint distribution $\mathscr{P}(M, Y, k^*|n)$ in equation (39). Performing Fourier transform with respect to Y, one can write $\bar{\mathcal{P}}\left(\xi, k^*|n\right)$ defined in equation (42) as

$$\bar{\mathcal{P}}\left(\xi, k^*|n\right) = \int_0^\infty \mathrm{d}M\, I_{\rm left}(M, \xi, k^*) I_{\rm right}(M, k^*, n), \tag{A.1}$$

where we have defined

$$I_{\rm right}(M, k^*, n) = \int \prod_{i=k^*+1}^{N} \mathrm{d}y_i\, \mathrm{d}x_i\, p(x_i, y_i) \Theta\left(M - \sum_{j=1}^{i} x_j\right), \tag{A.2}$$

and

$$I_{\rm left}(M,\xi,k^*) = \int \prod_{i=1}^{k^*} {\rm d}x_i\, \tilde{p}(x_i,\xi) \prod_{i=1}^{k^*-1} \Theta\left(M - \sum_{j=1}^{i} x_j\right) \delta\left(M - \sum_{j=1}^{k^*} x_j\right), \tag{A.3}$$

with the definition

$$\tilde{p}(x,\xi) = \int_{-\infty}^{\infty} {\rm d}y\, {\rm e}^{{\rm i}\xi y} p(x,y). \tag{A.4}$$

Let us first consider the integral $I_{\rm right}(M,k^*,n)$ defined in equation (A.2). Since,

$$X_{k^*+j} = x_1 + x_2 + \cdots + x_{k^*} + x_{k^*+1} + \cdots + x_{k^*+j} = M + x_{k^*+1} + \cdots + x_{k^*+j} \tag{A.5}$$

upon using $M = \sum_{i=1}^{k^*} x_i$, we can re-write equation (A.2) as

$$I_{\rm right}(M,k^*,n) = \int \left[\prod_{i=k^*+1}^{n} {\rm d}x_i\, p_1(x_i)\right] \prod_{i=k^*+1}^{n} \Theta\left(-\sum_{j=k^*+1}^{i} x_j\right), \tag{A.6}$$

where $p_1(x) = \int_{-\infty}^{\infty} p(x,y)\,{\rm d}y$ is a normalized (to unity) probability density function for the increment in the x-direction. However, the integral in equation (A.6) is simply the probability that a random walk in one dimension (in the x-direction) starting at the origin, with independent and identically distributed increment x_i's drawn from $p_1(x_i)$, stays below the origin up to step $n-k^*$. This is precisely given by q_{n-k^*} via the Sparre Andersen theorem (independently of the jump distribution $p_i(x)$, where $q_n = \binom{2n}{n} 2^{-2n}$. Hence, we have

$$I_{\rm right}(M,k^*,n) = q_{n-k^*}. \tag{A.7}$$

Note that the integral $I_{\rm right}(M,k^*,n)$ does not depend on M, but only on $(n-k^*)$.

We now turn to the left integral $I_{\rm left}(M,\xi,k^*)$ in equation (A.3). Let us first re-write $\tilde{p}(x,\xi)$ in equation (A.4) in a different way. Let us first consider the integral

$$\int_{-\infty}^{\infty} {\rm d}x\, \tilde{p}(x,\xi) = \int_{-\infty}^{\infty} {\rm d}x \int_{-\infty}^{\infty} {\rm d}y\, p(x,y)\, {\rm e}^{{\rm i}\xi y} = \int_{-\infty}^{\infty} {\rm d}y\, p_2(y) {\rm e}^{{\rm i}\xi y} = \tilde{p}_2(\xi), \tag{A.8}$$

where $p_2(y) = \int_{-\infty}^{\infty} {\rm d}x\, p(x,y)$ is the marginal distribution for the y-increment. Now, let us re-write

$$\tilde{p}(x,\xi) = \frac{\tilde{p}(x,\xi)}{\int_{-\infty}^{\infty} {\rm d}x\, \tilde{p}(x,\xi)} \times \tilde{p}_2(\xi) = f(x,\xi)\, \tilde{p}_2(\xi), \tag{A.9}$$

where we used the identity in equation (A.8) and

$$f(x,\xi) = \frac{\tilde{p}(x,\xi)}{\int_{-\infty}^{\infty} {\rm d}x\, \tilde{p}(x,\xi)}. \tag{A.10}$$

Note that $f(x,\xi)$ is normalized to unity (when integrated over x) and can be thought of as an effective jump distribution in the x direction that is just parametrized by ξ assuming it is positive

for all x. We use this expression of $\tilde{p}(x,\xi)$ from equation (A.9) into the integral expression for $I_{\rm left}(M,\xi,k^*)$ in equation (A.3), to get

$$I_{\rm left}(M,\xi,k^*) = [\tilde{p}_2(\xi)]^{k^*} \int \prod_{i=1}^{k^*} \mathrm{d}x_i\, f(x_i,\xi) \prod_{i=1}^{k^*-1} \Theta\left(M - \sum_{j=1}^{i} x_j\right) \delta \times \left(M - \sum_{j=1}^{k^*} x_j\right). \tag{A.11}$$

Now, substituting this expression and the result in equation (A.7) on the right-hand side of equation (A.1) and carrying out the integral over M gives

$$\tilde{P}(\xi,k^*|n) = q_{n-k^*}[\tilde{p}_2(\xi)]^{k^*} \int \left[\prod_{i=1}^{k^*} \mathrm{d}x_i\, f(x_i,\xi)\right] \prod_{i=1}^{k^*} \Theta\left(\sum_{j=1}^{i} x_{k^*+1-j}\right) \tag{A.12}$$

However, we immediately identify the k^*-fold integral in equation (A.12) as the probability that a one dimensional random walker, starting at the origin and with jump distribution drawn from $f(x,\xi)$ (which is normalised to unity), stays above the origin up to k^* steps. By Sparre Andersen theorem, this is universal and is simply q_{k^*} and is independent of $f(x,\xi)$, and in particular then does not depend on ξ. Hence we finally have

$$\tilde{P}(\xi,k^*|n) = \int_{\infty}^{\infty} P(Y,k^*|n)\, \mathrm{e}^{\mathrm{i}\xi Y}\, \mathrm{d}Y = q_{k^*}\, q_{n-k^*}\, [\tilde{p}_2(\xi)]^{k^*}, \tag{A.13}$$

which upon Fourier inversion, yields the result in equation (38). This result is true for arbitrary joint distribution $p(x,y)$ as long as it is symmetric and continuous in x.

Appendix B. Proof $\mathcal{S}_n \simeq \pi n$ as $n \to \infty$

In this appendix, we derive the asymptotic form of $\mathcal{S}_n$ for large n which was used to obtain the large n behaviour of $\langle A_n \rangle$ in equation (51). To this end, we consider the expression of $\mathcal{S}_n$ in equation (35) and change the variable $m = zn$ to yield

$$\mathcal{S}_n = \frac{\sqrt{\pi}}{\sigma} \sum_{z=1/n}^{(n-1)/n} \frac{\Gamma\left(\frac{n(1-z)+1}{2}\right)}{\Gamma\left(\frac{n(1-z)+2}{2}\right)} \langle M_{nz} \rangle. \tag{B.1}$$

Note that $z \in \{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}\}$. For large n, we change the summation in equation (B.1) to integration as $\sum_{z=1/n}^{(n-1)/n} \to n \int_0^1 \mathrm{d}z$ and rewrite it as

$$\mathcal{S}_n \simeq \frac{\sqrt{\pi}}{\sigma} n \int_0^1 \mathrm{d}z \frac{\Gamma\left(\frac{n(1-z)+1}{2}\right)}{\Gamma\left(\frac{n(1-z)+2}{2}\right)} \langle M_{nz} \rangle, \quad \text{as } n \to \infty. \tag{B.2}$$

We next use the result of [65] to write $\langle M_{nz} \rangle$ for large n as $\langle M_{nz} \rangle \simeq \sigma \sqrt{\frac{2zn}{\pi}}$. In addition, we approximate $\frac{\Gamma\left(\frac{n(1-z)+1}{2}\right)}{\Gamma\left(\frac{n(1-z)+2}{2}\right)} \simeq \sqrt{\frac{2}{n(1-z)}}$ as $n \to \infty$. Inserting these forms in equation (B.2) and performing the integration over z, we get

$$\mathcal{S}_n \simeq \pi n, \quad \text{as } n \to \infty. \tag{B.3}$$

Appendix C. Details of the numerical simulation for constructing convex hull

To construct convex hull numerically, we first generate a trajectory of the RTP depending on the choice of ensemble. For fixed-n ensemble, the particle undergoes a fixed number of runs (say n) and we stop the process after n runs have taken place. On the other hand, for fixed-t ensemble, we stop the system when the observation time t is reached. Given this trajectory, we construct the convex hull using the *Andrew's monotone chain algorithm* [73] which is further expedited with Akl's heuristic [74]. Then, to calculate the area, we denote the m vertices of the convex hull as $\{\bar{X}_i, \bar{Y}_i\},\ 1 \leqslant i \leqslant m$ in order of their Cartesian coordinates and use

$$A = \frac{1}{2}\sum_{i=0}^{m-1} \left(\bar{Y}_i + \bar{Y}_{i+1}\right)\left(\bar{X}_i - \bar{X}_{i+1}\right), \tag{C.1}$$

with $(\bar{X}_0, \bar{Y}_0) = (\bar{X}_m, \bar{Y}_m)$. This procedure is then repeated for R number of realisations to finally construct the histogram or average of A. We have taken $R = 10^5$ for constructing the histogram and $R = 10^5$ to compute the mean.

Appendix D. Derivation of $\langle M_s^2(n)\rangle$ in equation (68)

Here, we show that the expression of $\langle M_s^2(n)\rangle$ in equation (68) can be derived using the Pollaczek–Spitzer formula in equation (26). We first recall that $Q_s(M, n)$ in equation (66) represents the cumulative distribution that the maximum is less than M up to n steps for a random walker with independent and identically distributed increments $\{x_i\}$ chosen from the symmetric and continous distribution $g_s(x_i)$ in equation (61). For this, the Pollaczek–Spitzer formula gives [46, 67]

$$\sum_{n=0}^{\infty} z^n \langle \mathrm{e}^{-\lambda M_s(n)}\rangle = \sum_{n=0}^{\infty} z^n \int_0^{\infty} \mathrm{d}M\, \mathrm{e}^{-\lambda M} Q_s'(M, n) = \frac{\phi_s(z, \lambda)}{\sqrt{1-z}}, \tag{D.1}$$

where $0 \leqslant z \leqslant 1$ and $\lambda \geqslant 0$ and the function $\phi_s(z, \lambda)$ is defined as

$$\phi_s(z, \lambda) = \exp\left(-\frac{\lambda}{\pi}\int_0^{\infty} \mathrm{d}\xi \frac{\ln(1 - z\hat{p}_s(\xi))}{\lambda^2 + \xi^2}\right), \quad \text{with} \tag{D.2}$$

$$\hat{p}_s(\xi) = \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{\mathrm{i}\xi x} g_s(x) = \frac{1}{\sqrt{1 + \xi^2\sigma_s^2}}. \tag{D.3}$$

Here $\sigma_s = v_0/(\gamma + s)$. As seen for fixed-$n$ ensemble in equations (29) and (30), one can extend this formula to determine the generating function for the moments [54, 69]. For this case, one gets

$$h_s^{(1)}(z) = \sum_{n=0}^{\infty} z^n \langle M_s(n)\rangle = \frac{1}{\pi(1-z)}\int_0^{\infty} \frac{\mathrm{d}\xi}{\xi^2} \ln\left(\frac{1 - z\hat{p}_s(\xi)}{1-z}\right), \tag{D.4}$$

$$h_s^{(2)}(z) = \sum_{n=0}^{\infty} z^n \langle M_s^2(n)\rangle = (1-z)\left[h_s^{(1)}(z)\right]^2 + \frac{\sigma_s^2 z}{2(1-z)^2}. \tag{D.5}$$

Taking derivative of $h_s^{(2)}(z)$ n-times, we get

$$\langle M_s^2(n)\rangle = \sum_{m=1}^{n-1} \langle M_s(n)\rangle \left[\langle M_s(n-m)\rangle - \langle M_s(n-m-1)\rangle\right] + \frac{n\sigma_s^2}{2}. \tag{D.6}$$

We next use the results of [65] to write $\langle M_s(n)\rangle$ as

$$\langle M_s(n)\rangle = \frac{\sigma_s}{2\sqrt{\pi}} \sum_{j=1}^{n} \frac{\Gamma\left(\frac{j+1}{2}\right)}{\Gamma\left(\frac{j+2}{2}\right)}, \tag{D.7}$$

and using this, we get

$$\langle M_s(n-m)\rangle - \langle M_s(n-m-1)\rangle = \frac{\sigma_s}{2\sqrt{\pi}} \frac{\Gamma\left(\frac{n-m+1}{2}\right)}{\Gamma\left(\frac{n-m+2}{2}\right)}. \tag{D.8}$$

Finally, we insert equations (D.7) and (D.8) in the expression of $\langle M_s^2(n)\rangle$ in equation (D.6) and perform the sum over m explicitly to yield

$$\langle M_s^2(n)\rangle = \frac{\sigma_s^2}{2}\left(\frac{\mathcal{S}_n}{\pi} + n\right), \tag{D.9}$$

where $\mathcal{S}_n$ is given in equation (6). Identifying $\sigma_s = v_0/(\gamma + s)$, we recover the result in equation (68).

ORCID iDs

Prashant Singh https://orcid.org/0000-0003-4296-8274
Hendrik Schawe https://orcid.org/0000-0002-8197-1372

References

[1] Ramaswamy S 2010 *Annu. Rev. Condens. Matter Phys.* **1** 323–45
[2] Ramaswamy S 2017 *J. Stat. Mech.* 054002
[3] Bechinger C, Di Leonardo R, Löwen H, Reichhardt C, Volpe G and Volpe G 2016 *Rev. Mod. Phys.* **88** 045006
[4] Santiago I 2018 *Nano Today* **19** 11–5
[5] Ghosh A, Xu W, Gupta N and Gracias D H 2020 *Nano Today* **31** 100836
[6] Redner G S, Hagan M F and Baskaran A 2013 *Phys. Rev.Lett.* **110** 055701
[7] Bricard A, Caussin J-B, Desreumaux N, Dauchot O and Bartolo D 2013 *Nature* **503** 95–8
[8] Toner J, Tu Y and Ramaswamy S 2005 *Ann. Phys., NY* **318** 170–244
[9] Kumar N, Soni H, Ramaswamy S and Sood A K 2014 *Nat. Commun.* **5** 4688
[10] Cates M E and Tailleur J 2015 *Annu. Rev. Condens. Matter Phys.* **6** 219–44
[11] Gonnella G, Marenduzzo D, Suma A and Tiribocchi A 2015 *C. R. Phys.* **16** 316–31
[12] Partridge B and Lee C F 2019 *Phys. Rev. Lett.* **123** 068002
[13] Caprini L, Marconi U and Puglisi A 2020 *Phys. Rev. Lett.* **124** 078001
[14] Solon A P, Fily Y, Baskaran A, Cates M E, Kafri Y, Kardar M and Tailleur J 2015 *Nat. Phys.* **11** 673–8
[15] Howse J R, Joners R A L, Ryan A J, Gough T, Vafabakhsh R and Golestanian R 2007 *Phys. Rev. Lett.* **99** 048102
[16] Jiang H-R, Yoshinaga N and Sano M 2010 *Phys. Rev. Lett.* **105** 268302

[17] Masoliver J, Porrà J M and Weiss G H 1993 *Physica* A **193** 469–82
[18] Weiss G H 2002 *Physica* A **311** 381–410
[19] Berg H C 2004 *E. coli in Motion* (Heidelberg: Springer)
[20] Tailleur J and Cates M E 2008 *Phys. Rev. Lett.* **100** 218103
[21] Malakar K, Jemseena V, Kundu A, Vijay Kumar K, Sabhapandit S, Majumdar S N, Redner S and Dhar A 2018 *J. Stat. Mech.* 043215
[22] Dhar A, Kundu A, Majumdar S N, Sabhapandit S and Schehr G 2019 *Phys. Rev.* E **99** 032132
[23] Demaerel T and Maes C 2018 *Phys. Rev.* E **97** 032604
[24] Majumdar S N and Evans M R 2018 *J. Phys. A: Math. Theor.* **51** 475003
[25] Singh P, Sabhapandit S and Kundu A 2020 *J. Stat. Mech.* 083207
[26] Santra S, Basu U and Sabhapandit S 2020 *Phys. Rev.* E **101** 062120
[27] Le Doussal P, Majumdar S N and Schehr G 2020 *Europhys. Lett.* **130** 40002
[28] Mori F, Gradenigo G and Majumdar S N 2021 *J. Stat. Mech.* 103208
[29] Gradenigo G and Majumdar S N 2019 *J. Stat. Mech.* 053206
[30] Mori F, Le Doussal P, Majumdar S N and Schehr G 2021 *Phys. Rev.* E **103** 062134
[31] Angelani L, Di Leonardo R and Paoluzzi M 2014 *Eur. Phys. J.* E **37** 59
[32] Mori F, Le Doussal P, Majumdar S N and Schehr G 2020 *Phys. Rev. Lett.* **124** 090603
[33] De Bruyne B, Majumdar S N and Schehr G 2021 *J. Stat. Mech.* 043211
[34] Singh P and Kundu A 2019 *J. Stat. Mech.* 083205
[35] Mori F, Le Doussal P, Majumdar S N and Schehr G 2020 *Phys. Rev.* E **102** 042133
[36] Singh P and Kundu A 2021 *Phys. Rev.* E **103** 042119
[37] Banerjee T, Majumdar S N, Rosso A and Schehr G 2020 *Phys. Rev.* E **101** 052101
[38] Slowman A B, Evans M R and Blythe R A 2016 *Phys. Rev. Lett.* **116** 218101
[39] Singh P and Kundu A 2021 *J. Phys. A: Math. Theor.* **54** 305001
[40] Le Doussal P, Majumdar S N and Schehr G 2021 *Phys. Rev.* E **104** 044103
[41] Le Doussal P, Majumdar S N and Schehr G 2019 *Phys. Rev.* E **100** 012113
[42] Letac G and Takács L 1980 *Am. Math. Mon.* **87** 142
[43] Hug D 2013 Random polytopes *Stochastic Geometry, Spatial Statistics and Random Fields* (Heidelberg: Springer)
[44] Worton B J 1995 *Biometrics* **51** 1206–15
[45] Kac M 1954 *Duke Math. J.* **21** 501–9
[46] Spitzer F 1956 *Trans. Am. Math. Soc.* **82** 323–39
[47] Snyder T L and Steele J M 1993 *Proc. Am. Math. Soc.* **117** 1165
[48] Kabluchko Z, Vysotsky V and Zaporozhets D 2017 *Adv. Math.* **320** 595
[49] Kabluchko Z, Vysotsky V and Zaporozhets D 2017 *Geom. Funct. Anal.* **27** 880
[50] Majumdar S N, Comtet A and Randon-Furling J 2010 *J. Stat. Phys.* **138** 955–1009
[51] Randon-Furling J, Majumdar S N and Comtet A 2009 *Phys. Rev. Lett.* **103** 140602
[52] Reymbaut A, Majumdar S N and Rosso A 2011 *J. Phys. A: Math. Theor.* **44** 415001
[53] Majumdar S N, Mori F, Schawe H and Schehr G 2021 *Phys. Rev.* E **103** 022135
[54] Grebenkov D S, Lanoiselée Y and Majumdar S N 2017 *J. Stat. Mech.* 103203
[55] Claussen G, Hartmann A K and Majumdar S N 2015 *Phys. Rev.* E **91** 052104
[56] Schawe H, Hartmann A K and Majumdar S N 2018 *Phys. Rev.* E **97** 062159
[57] Schawe H and Hartmann A K 2019 *J. Phys.: Conf. Ser.* **1290** 012029
[58] Chupeau M, Bénichou O and Majumdar S N 2015 *Phys. Rev.* E **92** 022145
[59] Kampf J, Last G and Molchanov I 2012 *Proc. Am. Math. Soc.* **140** 2527–35
[60] Luković M, Geisel T and Eule S 2013 *New J. Phys.* **15** 063034
[61] Dumonteil E, Majumdar S N, Rosso A and Zoia A 2013 *Proc. Natl Acad. Sci. USA* **110** 4239–44
[62] Eldan R 2014 *Electron. J. Probab.* **19** 1–34
[63] Kabluchko Z and Zaporozhets D 2016 *Trans. Am. Math. Soc.* **368** 8873–99
[64] Schawe H, Hartmann A K and Majumdar S N 2017 *Phys. Rev.* E **96** 062101
[65] Hartmann A K, Majumdar S N, Schawe H and Schehr G 2020 *J. Stat. Mech.* 053401
[66] Cauchy A 1832 *Mémoire sur la rectification des courbes et la quadrature des surfaces courbées* (Paris: Chez de Bure fréres)
[67] Pollaczek F 1952 *C. R.* **234** 2334
[68] Comtet A and Majumdar S N 2005 *J. Stat. Mech.* P06013
[69] Majumdar S N 2010 *Physica* A **389** 4299
[70] Sparre Andersen E 1954 *Math. Scand.* **2** 195–223

[71] Basu U, Majumdar S N, Rosso A, Sabhapandit S and Schehr G 2020 *J. Phys. A: Math. Theor.* **53** 09LT01
[72] Romanczuk P, Bär M, Ebeling W, Lindner B and Schimansky-Geier L 2012 *Eur. Phys. J. Spec. Top.* **202** 1–162
[73] Andrew A M 1979 *Inf. Process. Lett.* **9** 216
[74] Akl S G and Toussaint G T 1978 *Inf. Process. Lett.* **7** 219–22