Persistent fluctuations of the swarm size of Brownian bees

Baruch Meerson\textsuperscript{1,}\textsuperscript{*} and Pavel Sasorov\textsuperscript{2,3,}\textsuperscript{†}

\textsuperscript{1}Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem 91904, Israel
\textsuperscript{2}Institute of Physics CAS, ELI Beamlines, 182 21 Prague, Czech Republic
\textsuperscript{3}Keldysh Institute of Applied Mathematics, Moscow 125047, Russia

The “Brownian bees” model describes a system of $N$ independent branching Brownian particles. At each branching event the particle farthest from the origin is removed, so that the number of particles remains constant at all times. Berestycki et al. (2020) proved that, at $N \to \infty$, the coarse-grained spatial density of this particle system is described by the solution of a free boundary problem for a deterministic reaction-diffusion equation. Further, they showed that, at long times, this solution approaches a unique steady state: a spherically symmetric distribution with compact support whose radius $\ell_0$ depends on the spatial dimension $d$. Here we study fluctuations in this system in the limit of large $N$ due to the stochastic character of the branching Brownian motion, and we focus on persistent fluctuations of the swarm size. We evaluate the probability density $\mathcal{P}(t, N, T)$ that the swarm size remains smaller than a specified value $\ell < \ell_0$, or larger than a specified value $\ell > \ell_0$, on a time interval $0 < t < T$, where $T$ is very large. We find that $\mathcal{P}(t, N, T)$ exhibits the large-deviation form $-\ln \mathcal{P} \simeq NTR_d(t)$. For all $d$ we obtain asymptotics of the rate function $R_d(t)$ in the regimes $\ell \ll \ell_0$, $\ell \gg \ell_0$ and $|\ell - \ell_0| \ll \ell_0$. For $d = 1$ the whole rate function can be calculated analytically. We obtain these results by determining the optimal (most probable) density profile of the swarm, conditioned on the specified $\ell$.

I. INTRODUCTION

Nonequilibrium steady states of macroscopic systems composed of reacting and diffusing particles, continue to attract attention from physicists [1–5]. The nonequilibrium steady states shed light on the physics of a plethora of important dissipative systems, both non-living and living. The search for, and the analysis of, simple models which can teach us about general properties of nonequilibrium fluctuations, continues. One of the simplest models of this type is a particle-conserving variant of branching Brownian motion that we will describe shortly.

Branching Brownian motion (BBM for short) unites two fundamental continuous-time Markov processes: the random branching and the Brownian motion, or the Wiener process. In the past BBM was extensively studied by mathematicians [6, 7], and it continues to attract interest from physicists [8–11]. Here we will consider the recently formulated Brownian bees model [12, 13].

“Brownian bees” is a variant of BBM with an imposed exact conservation law. The microscopic model is defined as follows. The system consists of $N$ independent particles (bees) located in a $d$-dimensional space. Each particle can branch, in a small time interval $\Delta t$, into two particles with probability $\Delta t$. It can also perform, with the complementary probability $1 - \Delta t$, continuous-time Brownian motion with diffusion constant 1 [14]. Whenever a branching event occurs, the particle which is farthest from the origin is instantaneously removed, so that the number of particles remains constant at all times. The name Brownian bees, coined by Jeremy Quastel [12], comes from the superficial analogy with a swarm of bees around a hive.

The Brownian bees is a particular example of the so-called Brunet-Derrida $N$-particle systems. All these systems involve the BBM with exact conservation law. They differ from each other only by the rule of elimination of “least-fit” particles, thus providing insight into different aspects of biological selection. See Ref. [13] for a brief review of these models.

Recently Berestycki et al. [12] showed that, in the limit of $N \to \infty$, the coarse-grained spatial density $u(x, t) \geq 0$ of the Brownian bees is described by the solution of the following free boundary problem in $d$ dimensions:

$$
\begin{align*}
\partial_t u(x, t) &= \nabla^2 u(x, t) + u(x, t), \quad |x| \leq L(t), \\
u(x, t) &= 0, \quad |x| > L(t), \\
\int_{|x| \leq L(t)} u(x, t) \, dx &= 1.
\end{align*}
$$

(1)

$u(x, t)$ is continuous at $|x| = L(t)$, and an initial condition must be specified. According to Eq. (1), $u(x, t)$ has a compact support which is, at all $t > 0$, a $d$-dimensional sphere. There is an effective absorbing wall at $|x| = L(t)$ (or two walls for $d = 1$), which moves so as to impose the constant number of particles at all times.

In a companion paper [13] Berestycki et al. showed that, at long times, the solution of the deterministic problem (1) approaches a unique steady state $u = u(x)$ which is described by the fundamental mode of the Helmholtz equation with a unit eigenvalue,

$$
\nabla^2 U + U = 0,
$$

(2)

inside a $d$-dimensional sphere with radius $\ell_0$ which depends on $d$. The steady state solution is spherically sym-
metric and has the following form:

\[ U(r) = \begin{cases} 
  \frac{F(r)}{4\pi\ell_0^{d/2}J_{d/2}(\ell_0)}, & r < \ell_0, \\
  0, & r > \ell_0, 
\end{cases} \tag{3} \]

where

\[ F(r) = \frac{J_{d/2-1}(r)}{r^{d/2-1}}, \tag{5} \]

and \( \ell_0 = \ell_0(d) \) is the first positive root of the Bessel function of the first kind \( J_{d/2-1}(r) \). For \( d = 1 \) one obtains

\[ U(x) = \begin{cases} 
  \frac{1}{2}\cos x, & |x| < \ell_0, \\
  0, & |x| > \ell_0. 
\end{cases} \tag{6} \]

and \( \ell_0 = \pi/2 \).

In the present work we study fluctuations of the stationary swarm of Brownian bees due to the stochastic character of the branching Brownian motion. We consider the limit of \( N \gg 1 \). In this limit the fluctuations are typically small. But large fluctuations (often called large deviations) also occur, and it is interesting to evaluate their probability as well. In this paper we focus on persistent fluctuations of the swarm size. Our objective is to evaluate the probability density \( P(\ell, N, T) \) that, on a long time interval \( 0 < t < T \), the swarm radius \( L(t) \) remains smaller than a specified value \( \ell < \ell_0 \), or larger than a specified value \( \ell > \ell_0 \). We argue that, at \( N \gg 1 \) and \( T \to \infty \), \( P(\ell, N, T) \) exhibits the large-deviation form

\[ -\ln P(\ell, N, T) \approx NTR_d(\ell). \tag{8} \]

This result follows naturally from the optimal fluctuation method (OFM) which we employ for solving this problem. The OFM (also known in other fields as the instanton method, the weak noise theory and the macroscopic fluctuation theory) is briefly described in Sec. II. The method boils down to finding the optimal (most probable) density profile of the swarm, conditioned on the specified \( \ell \) during a long time \( T \), and dominating the probability density of \( \ell \).

For all \( d \) we obtain, in Sec. III, the asymptotics of the rate function in the regimes \( |\ell - \ell_0| \ll \ell_0, \ell \ll \ell_0 \) and \( \ell \gg \ell_0 \). These asymptotics have the following form:

\[ R_d(\ell) \approx \begin{cases} 
  \alpha_d(\ell - \ell_0)^2, & |\ell - \ell_0| \ll \ell_0, \\
  \frac{\ell_0^2}{\ell^2} - \frac{2}{\sqrt{\alpha_d}} \ell \ll \ell_0, \\
  \frac{1}{3}, & \ell \gg \ell_0, 
\end{cases} \tag{9} \]

where the \( d \)-dependent factor \( \alpha_d \) is presented in Eq. (57) below. For \( d = 1 \) the whole rate function \( R_d(\ell) \) can be calculated exactly, as we explain in Sec. IV. We summarize and briefly discuss our results in Sec. V.

II. OPTIMAL FLUCTUATION METHOD: GOVERNING EQUATIONS

A convenient departure point for the derivation of the OFM equation is a lattice gas formulation for a gas of non-interacting branching random walkers. One starts from a multivariate master equation which describes the evolution with time of the probability of observing a certain number particles on each lattice site at time \( t \). Being interested in large deviations, and making a WKB-type ansatz in the master equation, one arrives at an effective multi-particle Hamilton-Jacobi equation, which can be recast in a Hamiltonian form [15–17]. Assuming in addition that the hopping rate of the random walkers is much larger than the branching rate, one obtains a continuous coarse-grained Hamiltonian field-theoretic description of large fluctuations in this reacting lattice gas, valid at distances large in comparison with the lattice constant [15–17]. In this limit the lattice constant only enters (alongside with the hopping rate) the diffusion constant, bringing back the continuous-space BBM model. In addition to the gas density field \( q(x, t) \), which formally plays the role of the “coordinate” of the Hamiltonian description, there is a canonically conjugate “momentum” density field \( p(x, t) \) which describes the most likely configuration of the noise which dominates the large deviation in question.

The OFM equations for reacting lattice gases were also derived and used by other workers [3, 18–22]. The problem of Brownian bees, however, brings an important new element: exact conservation of the total number of bees which takes the form of a non-local, integral constraint on the optimal (most likely) gas density history \( q(x, t) \geq 0 \):

\[ \int_{|x| < L(t)} q(x, t) \, dx = 1 \quad \text{for all} \quad 0 \leq t \leq T. \tag{12} \]

To accommodate this constraint in the OFM formalism we introduce a Lagrangian multiplier \( \lambda(t) \) (which in general depends on time) and add the term

\[ \int_0^T dt \lambda(t) \int_{|x| < L(t)} q(x, t) \, dx \tag{13} \]

to the Hamiltonian of the classical field-theory [15–17]. Note that, in this form, the additional term (13) holds even for a more general constraint, when the total number of particles is a specified function of time, cf. Ref. [28].

With the account of this additional term, the OFM equations [15–17] become

\[ \partial_t q = \frac{\delta H}{\delta p} = q e^p + \nabla \cdot (\nabla q - 2q \nabla p), \tag{14} \]

\[ \partial_t p = -\frac{\delta H}{\delta q} = - (e^p - 1) - \nabla^2 p - (\nabla p)^2 - \lambda(t). \tag{15} \]
Here
\[ H = H[q(x, t), p(x, t, \lambda(t))] = \int dx \mathcal{H}(q, p, \lambda) \] (16)
is the constrained Hamiltonian,
\[ \mathcal{H}(q, p, \lambda) = \mathcal{H}_0(q, p) + \lambda(t)q \] (17)
is the density of the constrained Hamiltonian, and
\[ \mathcal{H}_0(q, p) = (e^p - 1)q - \nabla q \cdot \nabla p + q (\nabla p)^2 \] (18)
is the density of the unconstrained Hamiltonian [15–17]. The boundary conditions on the absorbing wall (or on the two absorbing walls for \( d = 1 \)) are [18, 23–27].

\[ q(|x| = L(t)) = p(|x| = L(t)) = 0. \] (19)

We also have to specify some boundary conditions in time: at \( t = 0 \) and \( t = T \). These depend on whether we deal with a deterministic or fluctuating initial condition at \( t = 0 \), and on whether we specify the whole density function \( q(x, t = T) \) or only its finite support \( L(t = T) \).

The probability distribution \( \mathcal{P}(\ell, N, T) \), can be found, up to a preexponential factor, from the relation
\[-\ln \mathcal{P}(\ell, N, T) \approx NS(\ell, T), \] (20)
where
\[ S(\ell, T) = \int_0^T dt \int_{|x|<L(t)} dx (p\partial_t q - \mathcal{H}_0) \] (21)
is the action per particle. Plugging Eqs. (14) and (18) into Eq. (21), we obtain \( S \) in terms of an integral along the optimal trajectory:
\[ S(\ell, T) = \int_0^T dt \int_{|x|<L(t)} dx \left[q (pe^p - e^p + 1) + q (\nabla p)^2 \right]. \] (22)

Let us briefly discuss the physical meaning of the OFM equations. The momentum density field \( p(x, t) \) describes the history of the optimal realization of the noise. This is a deterministic time-dependent field which gives a dominant contribution to the specified large deviation of the system. In the absence of fluctuations we have \( p = 0 \) and \( \lambda(t) = 0 \). In this case Eq. (15) is obeyed trivially, and Eq. (14) coincides with the deterministic equation (1). The field \( p(x, t) \) enters Eq. (14) in two places. The factor \( e^p \), coming from the random character of branching, modifies the effective branching rate: it enhances the branching for \( p > 0 \), and suppresses it for \( p < 0 \). In its turn, the term including \( \nabla p \) describes a fluctuational contribution to the particle flux, coming from the stochastic character of Brownian motion.

For arbitrary \( T \) the OFM problem, described above, is both complicated and non-universal: it is intrinsically time-dependent, and the solution strongly depends on the initial and final conditions. Fortunately, the problem becomes much simpler in the limit of very large \( T \). Here it is natural to assume that the optimal gas density \( q(x, t) \) and the momentum density \( p(x, t) \) are stationary, \( L(t) \) is equal to the specified \( \ell \), and the Lagrange multiplier \( \lambda(t) \) is constant. The stationarity holds for most of the time interval \( 0 < t < T \) except for short non-universal transients close to \( t = 0 \) and \( t = T \) which would contribute to the action only at a subleading order, and which will be ignored in the following. The stationarity assumption leads to the steady-state OFM equations
\[ q e^p + \nabla \cdot (\nabla q - 2q \nabla p) = 0, \] (23)
\[ -(e^p - 1) - \nabla^2 p - (\nabla p)^2 = \lambda, \] (24)
whereas Eqs. (20) and (22) yield
\[ S(\ell, T) = TR_d(\ell) \] (25)
with the rate function
\[ R_d(\ell) = \int \left. dx \left[q (pe^p - e^p + 1) + q (\nabla p)^2 \right] \right|_{|x|<\ell} \] (26)
The initial and final conditions for \( q \) become irrelevant in the steady-state solution. Finally, if there are multiple stationary solutions, the one with the minimal action – hence, the minimal \( R_d(\ell) \) – must be selected. Equations (20) and (25) lead to the announced large-deviation behavior (8) of the probability density \( \mathcal{P} \).

A further simplification arises when we go over from \( q \) and \( p \) to the Cole-Hopf canonical variables \( Q = q e^{-p} \) and \( P = e^p - 1 \). Equations (23) and (24) become
\[ \nabla^2 Q + (2P + 1 + \lambda)Q = 0, \] (27)
\[ \nabla^2 P + (P + 1)(P + \lambda) = 0. \] (28)
Note that \( Q(x) \geq 0 \), whereas \( P(x) \) can vary from \(-1\) to \(+\infty\). Importantly, Eq. (28) is decoupled from Eq. (27). Also, for a given \( P(x) \), Eq. (27) is a linear and homogeneous equation for \( Q(x) \).

In the Hopf-Cole variables the boundary conditions for Eqs. (27) and (28) are
\[ Q(|x| = \ell) = P(|x| = \ell) = 0, \] (29)
the mass conservation (12) reads
\[ \int_{|x|<\ell} Q(x) [P(x) + 1] dx = 1. \] (30)
and the rate function \( R_d(\ell) \) becomes [16]
\[ R_d(\ell) = -\int_{|x|<\ell} dx \left[ Q(P + P^2) - \nabla P \cdot \nabla Q \right]. \] (31)
Applying the first Green’s identity and using the boundary condition \( Q(|x| = \ell) = 0 \) and Eq. (28), we can transform Eq. (31) to
\[ R_d(\ell) = \lambda \int_{|x|<\ell} dx Q(P + 1), \] (32)
which, by virtue of the integral constraint (30), brings us to the remarkably simple result

$$R_d(\ell) = \lambda.$$  

(33)

That is, the calculation of the rate function $R_d(\lambda)$ only requires to express the Lagrange multiplier $\lambda$ through $\ell$. Equation (33) implies that the solution exists only for $\lambda \geq 0$, as we indeed find here.

Although Eqs. (27) and (28) are written in the general $d$-dimensional form, the optimal solution for $Q$ and $P$ is spherically symmetric: $Q = Q(r)$ and $P = P(r)$, where $r$ is the radial coordinate. Therefore, only the radial part $\nabla_r^2$ of the Laplace operators actually appears in Eqs. (27) and (28). Correspondingly, the spatial integration in Eq. (30) is the following:

$$\Omega_d \int_0^\ell dr r^{d-1} Q(r)[P'(r) + 1] = 1,$$

(34)

where $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of the $d$-dimensional unit sphere, and $\Gamma(z)$ is the gamma function. For $d = 1$ we have

$$\int_{-\ell}^{\ell} dx Q(x)[P(x) + 1] = 1,$$

(35)

and $Q(x)$ and $P(x)$ are even functions:

$$Q'(0) = P'(0) = 0.$$  

(36)

Here and in the following the primes stand for the $x$- (or $r$-) derivatives.

Equations (27) and (28) possess an important conservation law which follows from their Hamiltonian (or, alternatively, Lagrangian) character:

$$[Q'P' + Q(P + 1)(P + \lambda)]' + \frac{2(d - 1)}{r} Q'P' = 0.$$  

(37)

One way of deriving this conservation law is by using a direct analogy with the energy-momentum tensor of the classical field theory [29]. In this interpretation Eq. (37) expresses the vanishing divergence of the energy-momentum tensor in the presence of spherical symmetry.

The case of $d = 1$ is special, because in this case the conservation law (37) does not depend explicitly on the coordinate, and we obtain [17, 21]:

$$Q'P' + Q(P + 1)(P + \lambda) = W = \text{const}.$$  

(38)

Furthermore, in the case of $d = 1$, and only in this case, Eq. (28) has its own conservation law:

$$\frac{1}{2} (P')^2 + V(P, \lambda) = E = \text{const}.$$  

(39)

The cubic potential

$$V(P, \lambda) = \frac{1}{3} P^3 + \frac{1 + \lambda}{2} P^2 + \lambda P$$

(40)

has extrema at $P = -1$ and $P = -\lambda$. The two conservation laws, Eqs. (38) and (39), make the one-dimensional case exactly integrable, see Sec. IV.

For all $d$, the stationary problem in the Hopf-Cole variables can be conveniently solved numerically. Since Eq. (27) is linear and homogeneous, and Eq. (28) is not coupled to Eq. (27), we can solve Eqs. (27) and (28) by setting an arbitrary nonzero boundary condition for $Q(r = 0)$, for example $Q(r = 0) = 1$, and ultimately normalize the solution by using Eq. (34). Specifying $\lambda$ and using $P(r = 0)$ as a shooting parameter, we demand that the first zeros of $Q(r)$ and $P(r)$ coincide with Eq. (2). In the subleading order we obtain

$$\nabla_r^2 P_0 + P_0 = 0,$$

(43)

$$\nabla_r^2 Q_0 + Q_0 = 0,$$

(44)

which coincide with Eq. (2). In the subleading order we obtain

$$\nabla_r^2 P_1 + P_1 = -1 - P_0^2,$$

(45)

$$\nabla_r^2 Q_1 + Q_1 = -2P_0Q_0,$$

(46)

etc. The solutions of Eqs. (43) and (44) can be written as

$$\frac{P_0(r)}{p_0} = \frac{Q_0(r)}{q_0} = 2^{d-1} \Gamma(d/2) F(r),$$

(47)

where $p_0$ and $q_0 > 0$ are (yet unknown) constants, and $F(r)$ was defined in Eq. (5). For $d = 1$, 2 and 3 we obtain

$$\frac{P_0(r)}{p_0} = \frac{Q_0(r)}{q_0} = \begin{cases} \cos r, & d = 1, \\ J_0(r), & d = 2, \\ \sin r/r, & d = 3, \end{cases}$$

(48)

III. ASYMPTOTICS OF THE RATE FUNCTION

There are three asymptotic regimes, where the rate function $R_d(\lambda)$ can be calculated analytically in any dimension.

A. $|\ell - \ell_0| \ll \ell_0$

Here $\lambda \ll 1$ is a small parameter, and the solutions of Eqs. (27) and (28) can be found via the perturbative ansatz

$$P(r) = \sqrt{\lambda} \left[ P_0(r) + \sqrt{\lambda} P_1(r) + \lambda P_2(r) + \ldots \right],$$

(41)

$$Q(r) = Q_0(r) + \sqrt{\lambda} Q_1(r) + \lambda Q_2(r) + \ldots,$$

(42)

where the functions $P_0$, $P_1$, $Q_0$, $Q_1$, etc. are of order 1. In the leading order the ansatz (41) and (42) leads to two identical Helmholtz equations

$$\nabla_r^2 P_0 + P_0 = 0,$$

(43)

$$\nabla_r^2 Q_0 + Q_0 = 0,$$

(44)

which coincide with Eq. (2). In the subleading order we obtain

$$\nabla_r^2 P_1 + P_1 = -1 - P_0^2,$$

(45)

$$\nabla_r^2 Q_1 + Q_1 = -2P_0Q_0,$$

(46)

etc. The solutions of Eqs. (43) and (44) can be written as

$$\frac{P_0(r)}{p_0} = \frac{Q_0(r)}{q_0} = 2^{d-1} \Gamma(d/2) F(r),$$

(47)

where $p_0$ and $q_0 > 0$ are (yet unknown) constants, and $F(r)$ was defined in Eq. (5). For $d = 1$, 2 and 3 we obtain

$$\frac{P_0(r)}{p_0} = \frac{Q_0(r)}{q_0} = \begin{cases} \cos r, & d = 1, \\ J_0(r), & d = 2, \\ \sin r/r, & d = 3, \end{cases}$$

(48)
respectively. For any $p_0$ and $q_0$, the two functions $P_0(r)$ and $Q_0(r)$ have their first positive zeros at the same point $r = \ell_0$. The subleading-order equations (45) and (46) lift this degeneracy. The condition that the first positive zeros of $P(r)$ and $Q(r)$ coincide also in the subleading order yield a unique value of the constant $p_0$. Having found $P(r)$ in the subleading order, we can determine the rate function $R_d(\ell) = \lambda$ in terms of the small difference $\ell - \ell_0$. The remaining constant $q_0$ can be determined from Eq. (34), but it does not affect $\lambda$.

The calculation proceeds as follows. First, we solve the linear equations (45) and (46) with the boundary conditions $P_1(0) = P_1'(0) = 0$ and $Q_1(0) = Q_1'(0) = 0$, respectively. The solutions can be written as

$$P_1(r) = \pi \int_0^r dy y^{d-1} \frac{F(r) G(y)}{G(r) G(y)} \left(1 + P_0(y)^2\right),$$

$$Q_1(r) = \pi \int_0^r dy y^{d-1} \frac{F(r) G(y)}{G(r) G(y)} P_0(y) Q_0(y).$$

where $F(r)$ is defined in Eq. (5),

$$G(r) = \frac{Y_{d-1}(r)}{r^{\frac{d}{2}-1}},$$

and $Y$ is the Bessel function of the second kind. As $P_1$ and $Q_1$ give subleading contributions, it suffices to evaluate the integrals in Eqs. (49) and (50) at $r = \ell_0$. This yields the special value of $p_0$ for which the zeros of $P$ and $Q$ coincide in the subleading order:

$$p_0^2 = \frac{\int_0^{\ell_0} dr r^{d-1} F(r)}{2^{d-2}\Gamma(d/2)^2 \int_0^{\ell_0} dr r^{d-1} [F(r)]^3}.$$  

As a result,

$$p_0^2 \approx \begin{cases} 
3/2, & d = 1, \\
2.21501, & d = 2, \\
\frac{3 \text{ Si}(\pi) - \text{Si}(3\pi)}{4\pi} = 3.23787, & d = 3.
\end{cases}$$

where Si($z$) = $\int_0^z dt \sin(t)/t$ is the sine integral function. For $\ell < \ell_0$ we have $p_0 = 0$. Here the fluctuations enhance branching ($p > 0$) and cause an inward particle flux ($\nabla_x p < 0$). The case of $\ell > \ell_0$ corresponds to $p_0 < 0$. Here the branching is suppressed, and an outward particle flux is enhanced, by fluctuations.

With $p_0^2$ at hand, we Taylor expand $P_0(r)$ in a small vicinity of $r = \ell_0$, keeping only the term linear in $\ell - \ell_0$,

$$P_0(r) \approx P_0'(\ell_0) (\ell - \ell_0),$$

and express $\lambda$ via the small shift $\Delta \ell = \ell - \ell_0$ of the first zero of $P(r) \approx P_0(r) + \sqrt{\lambda} P_1(r)$. In this way we obtain the asymptotic of the rate function $R_d(\ell) = \lambda$, announced in Eq. (9): $R_d(\ell) \approx \alpha_d(\Delta \ell)^2$ with

$$\alpha_d = \frac{\left(P_0'(\ell_0)\right)^2}{\pi G(\ell_0) \int_0^{\ell_0} dr r^{d-1} F(r)^2}.$$  

where we used Eqs. (47) and (52). Equation (57) gives $\alpha_1 = 3/8, \alpha_2 = 0.14924 \ldots$, and

$$\alpha_3 = \frac{1}{\pi [3 \text{ Si}(\pi) - \text{Si}(3\pi)]} = 0.08201 \ldots.$$  

The quadratic dependence of $R_d$ on $\Delta \ell$ describes a Gaussian asymptotic of the distribution $P(\ell, N, T)$ at $\ell$ close to the expected value $\ell_0$.

B. $\ell \ll \ell_0$

Here the fluctuations should work against a strong outward diffusion and also enhance the branching process considerably so as to make up for the particle losses at $r = \ell \ll \ell_0$. The optimal balance between these two effects is determined by Eqs. (27) and (28) at $\lambda \gg 1$. As the first zeros of $Q$ and $P$ must coincide (at $r = \ell$), we see that $P$ (which is positive and large here) must scale as $\sqrt{\lambda}$, whereas $\ell$ must scale as $1/\sqrt{\lambda}$. This leads us to the large-$\lambda$ perturbative ansatz

$$P(r) = \sqrt{\lambda} \left[ P_0(\bar{r}) + \frac{1}{\sqrt{\lambda}} P_1(\bar{r}) + \frac{1}{\lambda} P_2(\bar{r}) + \ldots \right],$$

$$Q(r) = q_0 \left[ Q_0(\bar{r}) + \frac{1}{\sqrt{\lambda}} Q_1(\bar{r}) + \frac{1}{\lambda} Q_2(\bar{r}) + \ldots \right],$$

where $\bar{r} = r/\sqrt{\lambda}$. The functions $P_0, P_1, Q_0, Q_1, \ldots$, are again of order 1, while $q_0$ must scale as $\lambda^{(d-1)/2}$ to comply with the normalization condition. This ansatz leads to the same pair of Helmholtz equations

$$\nabla^2 P_0 + P_0 = 0,$$

$$\nabla^2 Q_0 + Q_0 = 0$$

as in Sec. III A. Remarkably, it also yields the same subleading-order equations,

$$\nabla^2 P_1 + P_1 = -1 - P_2^2,$$

$$\nabla^2 Q_1 + Q_1 = -2P_0Q_0,$$

as Eqs. (45) and (46). Comparing Eqs. (41) and (42) with Eqs. (59) and (60), we see that, for the purpose of coincidence of the zeros of $P$ and $Q$, there is an exact mapping between the two cases if we replace $\sqrt{\lambda}$ by $1/\sqrt{\lambda}$ and, in the case of $\ell \ll \ell_0$, measure the distance in the units of $1/\sqrt{\lambda}$. This unexpected mapping between the two (physically very different) regimes allows us to immediately obtain the rate function asymptotic $R_d(\ell) = \lambda(\ell)$ for $\ell \ll \ell_0$ from the already found asymptotic $\lambda = \alpha_d(\ell - \ell_0)^2$ for $|\ell - \ell_0| \ll \ell_0$. After a simple algebra we obtain, in the leading and subleading orders in $1/\ell \ll 1$, the expression announced in Eq. (10). In particular,
The leading-order asymptotic $R_d(\ell \ll \ell_0) \approx \ell_0^d/\ell^2$ coincides with the rate function, corresponding to the long-time survival probability of pure Brownian motion (no branching) inside a $d$-dimensional sphere of radius $\ell$, see e.g. Ref. [30]. This result is not unexpected: at very small $\ell$ the fluctuations mostly “work” against diffusion which otherwise would rapidly spread out the swarm. The subleading term, proportional to $1/\ell$, is much larger than 1 and is therefore important. It is already affected by the branching process.

We can also determine the optimal stationary profile of the gas density $q(r) = Q(r)[1 + P(r)]$, conditioned on $\ell \ll \ell_0$. In most of the region $r < \ell$, $q(r)$ can be approximately described by the leading-order solution

$$q(r) \approx q_0\sqrt{\lambda} Q_0(\sqrt{\lambda}r) P_0(\sqrt{\lambda}r)$$

$$\simeq C_d \frac{q_0^{d-1}}{\ell^d} \left[F\left(\frac{\ell_0 r}{\ell}\right)\right]^2.$$  

(68)

where the function $F(r)$ is defined in Eq. (5), $\lambda \equiv R_d(\ell)$ is given by Eqs. (10) and (65)-(67), and the numerical factor $C_d$ is determined by normalization to unity. Equation (68) coincides with the optimal density profile, corresponding to the survival probability of pure Brownian motion [30]. Importantly, however, Eq. (68) breaks down close to $r = \ell$, where the solution is dominated by the term $q \simeq Q \approx q_0 Q_0(\sqrt{\lambda}r)$. It is this term which determines (large) particle losses at $r = \ell$ which are compensated by the strongly enhanced branching in the bulk [31].

As an example, Fig. 1 shows the optimal density profile $q(x)$ for $d = 1$ and $\ell = 0.048 \ll \ell_0$. Here $\lambda = 10^3$.

### C. $\ell \gg \ell_0$

In this regime the fluctuations must keep the swarm very large which requires a strong suppression of the branching process. A lower bound for $P(\ell, N, T)$ [that is an upper bound for $R_d(\ell \gg \ell_0)$] can be obtained by assuming that there are no branching events altogether during the whole time $T$. In this extreme scenario there are no particle losses, while the diffusion spread proceeds unhindered, that is deterministically. In this scenario the probability density $P^{\text{lower bound}}(\ell, N, T) = \exp(-NT)$ is independent of $\ell$ (and, in the physical units, independent of the diffusion constant), leading to the upper bound for $R_d^\text{upper bound}(\ell \gg \ell_0) = 1$.

As we will now show, the true optimal configuration at $\ell \gg \ell_0$ outperforms this simple upper bound by suppressing the branching and diffusion in most of the swarm, but allowing it in a close vicinity of $r = \ell$. Almost all of the bees in this regime are concentrated near the swarm boundary.

A complete suppression of the branching by fluctuations would require $\rho = -\infty$, that is $P = -1$. The true optimal trajectory is such that $P(r)$ stays very close to $-1$ on most of the interval $0 < r < \ell$, and increases and reaches $P = 0$ in the narrow boundary layer near $r = \ell$ with width $O(1)$. Let us first consider $d = 1$.

#### 1. $d = 1$

Here there are two conservation laws (38) and (39). In the next section we explain how one can use them to solve the case of $d = 1$ exactly for any $\ell$. Here we use them to find a simple limiting solution $Q(x)$ and $P(x)$ which gives the leading-order asymptotic of the exact solution at $\ell \to \infty$. Because of the asymmetry of the one-dimensional swarm with respect to $x = 0$, we can limit ourselves to positive $x$, that is consider the right half of the swarm.

Plugging the exact boundary condition $P'(x = 0) = 0$ and the asymptotically exact boundary condition $P(x = 0) = -1$ into Eq. (38), we see that $W = 0$. Then, because of the boundary condition $Q(x = \ell) = 0$, the same equation yields $P'(x = \ell) = 0$ [32].

Now we turn to the second conservation law (39). Using the equality $P'(x = \ell) = 0$, that we have just established, and the boundary condition $P(x = \ell) = 0$, we obtain $E = 0$. Then, using $P'(x = 0) = 0$ and
\( P(x = 0) = -1 \), we obtain \( \lambda = 1/3 \). This immediately
leads us to the large-\( \ell \) asymptotic of the rate function:

\[
R_1(\ell \gg \ell_0) \simeq \frac{1}{3} .
\]  

(70)

This leading-order asymptotic is independent of \( \ell \), and
is three times smaller than the simple upper bound (69). Note that the limiting solution for \( P(x) \), with \( \lambda = 1/3 \)
and \( E = 0 \), corresponds to a homoclinic trajectory on
the phase plane \( (P, P') \), see Fig. 2.

In view of Eq. (28), we can rewrite Eq. (38) for the
limiting solution as

\[
Q'P' - QP'' = 0 .
\]  

(71)

This homogeneous linear ODE gives a simple relation
between \( Q(x) \) and \( P(x) \) for the limiting solution: \( Q(x) = k P'(x) \),
where \( k > 0 \) is constant. In its turn,

\[
q(x) = [1 + P(x)]Q(x) = k [1 + P(x)]P'(x) ,
\]  

(72)

and \( k \) can be determined from the normalization condition:

\[
k \int_{-\infty}^{0} (1 + P)P' \, d\xi = \frac{k}{2} (1 + P)^2 \bigg|_{\xi = 0}^{\xi = -\infty} = \frac{1}{2} ,
\]  

(73)

so \( k = 1 \). In fact, \( P \) and \( Q \), and then \( p \) and \( q \), for the limiting
solution can be found in an explicit and elementary form:

\[
P(\xi) = -1 + \text{sech}^2 \left( \frac{\xi}{\sqrt{6}} \right) ,
\]  

(74)

\[
Q(\xi) = -\sqrt{\frac{2}{3}} \tanh \left( \frac{\xi}{\sqrt{6}} \right) \text{sech}^2 \left( \frac{\xi}{\sqrt{6}} \right) ,
\]  

(75)

\[
p(\xi) = 2 \ln \text{sech} \left( \frac{\xi}{\sqrt{6}} \right) ,
\]  

(76)

\[
q(\xi) = -\sqrt{\frac{2}{3}} \tanh \left( \frac{\xi}{\sqrt{6}} \right) \text{sech}^4 \left( \frac{\xi}{\sqrt{6}} \right) ,
\]  

(77)

where \( \xi = x - \ell \leq 0 \). Figure 3 shows the resulting plots
of \( P(x) \) and \( Q(x) \) (the top panel) and \( q(x) \) (the bottom
panel) for \( \ell = 20 \). As one can see, the optimal gas
density, conditioned on \( \ell \gg \ell_0 \), is close to zero in most of the
swarm. The bees are constantly produced in the boundary
layer with width \( O(1) \) near \( x = \ell \), diffuse to the
absorbing wall at \( x = \ell \) and get absorbed. Furthermore,
the optimal fluctuation suppresses the inward diffusion
of the bees from the periphery. Indeed, using Eq. (76),
we can see that, well outside of the boundary layer at
\( \xi = 0 \), \( p(\xi) \) behaves as \( (2/\sqrt{6}) \xi + \text{const} \). As a result, the
\( \nabla p \) term in Eq. (23) produces a constant drift velocity
directed outward. It is this outward drift which suppresses
the inward particle diffusion.

2. \( d > 1 \)

Crucially, the leading-order asymptotic \( R_d(\ell \gg \ell_0) = 1/3 \), obtained for \( d = 1 \), is valid in all dimensions, as

announced in Eq. (11). This is because, as \( \ell \) goes to in-
finity, the first-derivative terms \( \frac{d}{dr} P'(r) \) and \( \frac{d}{dr} Q'(r) \)
in the Laplace operators of Eqs. (45) and (46) become
negligible compared with the second derivative terms in
the region \( \ell - r = O(1) \), where \( P(r) + 1 \) and \( Q(r) \) are
localized.

As an example, Fig. 4 shows a numerical solution for
\( d = 3 \) and \( \lambda = 0.3 \), which corresponds to \( \ell \simeq 27.4 \simeq
8.72\ell_0 \). As one can see, this optimal solution is quali-
tatively similar to that for \( d = 1 \). The branching is al-
most completely suppressed in the bulk of the swarm:
\( P(r) \approx -1 \). This, however, does not lead to an ac-
tion proportional to \( \ell \), because the gas density \( q(r) \) is
almost zero in the bulk. The branching and diffusion act
only close to the boundary of the swarm \( r = \ell \), and the
fluctuation-induced outward drift suppresses the inward
diffusion of the bees.

An attentive reader could have noticed that an \( \ell \)-
independent asymptotic of \( R_d(\ell) \), and therefore of
\( P(\ell, N, T) \), formally leads to a divergent integral
\( \int_0^\infty P(\ell, N, T) \, d\ell \) and, therefore, to a non-normalizable
distribution. The resolution of this paradox is the fol-
lowing. The asymptotic (11) assumes stationarity of the
optimal configuration. At fixed \( \ell \), this assumption is valid
only when the observation time \( T \) is sufficiently large. If
we instead fix \( T \) and keep increasing \( \ell \), we will ultimately
enter a non-stationary regime, where \( P(\ell, N, T) \) does not
have the large-deviation form (8), and where it rapidly
falls off with a further increase of \( \ell \), thus resolving the
non-normalizability paradox.

IV. \( d = 1 \) IS INTEGRABLE

In one dimension, Eq. (39) describes a Newtonian par-
ticle of unit mass with “coordinate” \( P \) moving in “time”
We implemented this scheme in full. The explicit expressions [especially the ones involving \( Q(x, E, W) \)] are too bulky to be presented here. Therefore we only show, in Fig. 5, the final results in the form of a plot of \( R_1 \sim \ell (\ell) = \lambda(\ell) \). Also shown are the three asymptotics (9)-(11) for \( d = 1 \). The lower panel of Fig. 5 shows the product \( \ell^2 R_1(\ell) \) versus \( \ell \). Evident is a good agreement between the asymptotics and the exact results in the proper regions.

V. SUMMARY AND DISCUSSION

Here we studied persistent large deviations of the swarm size of Brownian bees. We determined the rate function \( R_d(\ell) \), which characterizes these large deviations, analytically and numerically in different limits. The optimal fluctuation method (OFM) was instrumental in obtaining these results. In addition to the rate function itself, the OFM provides an illuminating insight into the most probable configurations of the swarm that dominate the probability density of the specified unusual swarm size. As we observed, these configurations are quite fascinating in the limit of unusually large swarms, \( \ell \gg \ell_0 \).

As it is evident from Fig. 5, there exists a value of \( \ell = \ell_1 > \ell_0 \) such that, at \( \ell > \ell_1 \), the rate function \( R_d(\ell) \) is nonconvex and, therefore, cannot be obtained from the G"artner-Ellis theorem, see e.g. Ref. [34].

It would be very interesting to also evaluate the probability density of instantaneous fluctuations of the swarm size in the steady state. This probability density \( \mathcal{P}(N, \ell) \) does not depend on time and is expected to scale as \(- \ln \mathcal{P} \approx N f(\ell) \). Furthermore, the rate function \( f(\ell) \)
should come from a non-stationary optimal path of the swarm conditioned on reaching a specified \( \ell \). Solving this challenging non-stationary OFM problem would be a considerable achievement.

ACKNOWLEDGMENTS

We thank A. Vilenkin for useful discussions. The research of B.M. is supported by the Israel Science Foundation (grant No. 1499/20). The research of P.S. is supported by the project “High Field Initiative” (CZ.02.1.01/0.0/0.0/15_003/0000449) of the European Regional Development Fund.

[1] G. Jona-Lasinio, C. Landim and M.E. Vares, Probab. Theory Relat. Fields 97, 339 (1993).
[2] G. Basile and G. Jona-Lasinio, Int. J. Mod. Phys. B 18, 479 (2004).
[3] T. Bodineau and M. Lagouge, J. Stat. Phys. 139, 201 (2010).
[4] P.I. Hurtado, A. Lasanta and A. Prados Phys. Rev. E 88, 022110 (2013).
[5] B. Meerson, J. Stat. Mech. 2015, 05004.
[6] H.P. McKean, Comm. Pure Appl. Math. 28, 323 (1975).
[7] M.D. Bramson, Mem. Am. Math. Soc. vol. 44 (285) (1983).
[8] É. Brunet and B. Derrida, EPL 87, 60010 (2009); J. Stat. Phys. 143, 420 (2011).
[9] A. H. Mueller and S. Munier, Phys. Rev. E 90, 042143 (2014).
[10] K. Ramola, S. N. Majumdar, and G. Schehr, Phys. Rev. Lett. 112, 210602 (2014); Chaos, Solitons and Fractals 74, 79 (2015); Phys. Rev. E 91, 042131 (2015).
[11] B. Derrida, B. Meerson and P. V. Sasorov, Phys. Rev. E 93, 042139 (2016).
[12] J. Berestycki, É. Brunet, J. Nolen and S. Penington, arXiv:2005.09384.
[13] J. Berestycki, É. Brunet, J. Nolen and S. Penington, arXiv:2006.06486.
[14] By rescaling time and the coordinate one can set the branching rate and the diffusion constant to unity.
[15] V. Elgart and A. Kamenev, Phys. Rev. E 70, 041106 (2004).
[16] B. Meerson and P.V. Sasorov, Phys. Rev. E 83, 011129 (2011).
[17] B. Meerson, P.V. Sasorov and Y. Kaplan, Phys. Rev. E 84, 011147 (2011).
[18] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Rev. Mod. Phys. 87, 593 (2015).
[19] G. Jona-Lasinio, C. Landim and M.E. Vares, Probab. Theory Relat. Fields 97, 339 (1993).
[20] G. Basile and G. Jona-Lasinio, Int. J. Mod. Phys. B 18, 479 (2004).
[21] B. Meerson and P.V. Sasorov, Phys. Rev. E 84, 030101(R) (2011).
[22] B. Meerson, A. Vilenkin and P.V. Sasorov, Phys. Rev. E 87, 012117 (2013).
[23] T. Bodineau and B. Derrida, Phys. Rev. Lett. 92, 180601 (2004).
[24] E. Akkermans, T. Bodineau, B. Derrida and O. Shpielberg, EPL 103, 20001 (2013).
[25] O. Shpielberg and E. Akkermans, Phys. Rev. Lett. 116, 240603 (2016).
[26] B. Meerson, A. Vilenkin, and P. L. Krapivsky, Phys. Rev. E 90, 022120 (2014).
[27] B. Meerson, J. Stat. Mech. (2015) P04009.
[28] N. R. Smith, B. Meerson and A. Vilenkin, J. Stat. Mech. (2019) 053207.
[29] L.D. Landau and E.M. Lifshitz, The Classical Theory of Fields, 4th edition (Butterworth-Heinemann, Amsterdam, 1980), Sec. 32.
[30] T. Agranov, B. Meerson and A. Vilenkin, Phys. Rev. E 93, 012136 (2016).
[31] The particle flux \( j = -\nabla q + 2q\nabla p \) at \( r = \ell \) is equal to \( -\nabla q \), or to \( -\nabla Q \) in the Hopf-Cole variables. Hence, it is determined, in the leading order, by the expression \( -q_0 \sqrt{\lambda} Q'_0|_{\bar{r} = \ell_0} \). As a result, the particle flux scales as \( (\ell_0/\ell)^d \gg 1 \).
[32] \( Q(x = \ell) \) must be nonzero. If it were zero alongside with \( Q(x = \ell) \), \( Q(x) \) would vanish identically.
[33] W. A. Schwalm, Lectures on Selected Topics in Mathematical Physics: Elliptic Functions and Elliptic Integrals (Morgan and Claypool Publishers, San Rafael, CA, USA, 2015).
[34] H. Touchette, Phys. Rep. 478, 1 (2009).