Exact results in stochastic processes with division, death, and diffusion: Spatial correlations, marginal entropy production, and macroscopic currents

Samuel Cameron and Elsen Tjhung
School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom

(Received 18 March 2025; accepted 21 July 2025; published 2 September 2025)

We consider a generic class of stochastic particle-based models whose state at an instant in time is described by a set of continuous degrees of freedom (e.g., positions), and the length of this set changes stochastically in time due to birth-death processes. Using a master equation formalism, we write the dynamics of the corresponding (infinite) set of probability distributions: this takes the form of coupled Fokker-Planck equations with model-dependent source and sink terms. We derive the general expression of the entropy production rate for this class of models in terms of path irreversibility. To demonstrate the practical use of this framework, we analyze a biologically motivated model incorporating division, death, and diffusion, where spatial correlations arise through the division process. By systematically integrating out excess degrees of freedom, we obtain the marginal probability distribution, enabling exact calculations of key statistical properties such as average density and correlation functions. We validate our analytical results through numerical Brownian dynamics simulations, finding excellent agreement between theory and simulation. Our method thus provides a powerful tool for tackling previously unsolved problems in stochastic birth-death dynamics.

DOI: 10.1103/b9n6-mynw

I. INTRODUCTION

Active matter research has largely focused on systems where energy is locally injected (and subsequently dissipated) to drive the self-propelled motion of individual particles [1]. Synthetic examples include self-propelled colloids with asymmetric surface chemistry [2], while biological examples range from motile cells and swimming microorganisms to flocks of birds [3]. While most studies focus on motility-driven activity, birth-death processes [4]—common in biological systems—have received less attention. These processes play a key role in tissue dynamics, leading to instabilities in otherwise stable glassy systems [5–7], bacterial surface proliferation [8,9], and biofilm formation [10–12]. In particular, cell division introduces spatial correlations along the division axis, yet there remains a lack of probabilistic frameworks capable of accurately tracking these division events in both space and time.

The simplest models of birth and death processes involve only a discrete-state space and are often used to describe population dynamics assuming there is no spatial heterogeneity [13]. For example, characterizing the growth of a bacterial colony in terms of the total number of bacteria or cells present can be sufficient if one is only interested in the global averages, and all physics can be encoded into simple rate constants. However, tracking spatial degrees of freedom is often crucial, for example, when simulating the propagating front of a bacterial colony. In such cases, the theoretical literature is largely dominated by two approaches: (1) reaction-diffusion models, which describe coarse-grained particle concentrations at each point in space, incorporating phenomenological terms for division, death, and diffusion [12,14,15]; and (2) direct particle-based simulations, such as molecular dynamics coupled with stochastic rate equations to track individual trajectories [16].

A third approach is to formulate a master equation using coupled Fokker-Planck equations with source and sink terms. To track both the number of particles and their spatial positions, the state transitions must allow changes in the number of continuous degrees of freedom and properly account for permutations of particle indices due to indistinguishability [17]. Unlike mean field reaction-diffusion models, which focus on effective single-body concentrations in the long-wavelength (hydrodynamic) limit, the master equation framework retains the full many-body probability distributions. However, analytical methods for solving such high-dimensional stochastic systems are very limited. The most established method is to map the master equation onto a quantum field theory using creation and annihilation operators in a Fock space representation [18–23]. The resulting Doi-Peliti action is generally nonlinear and must be treated perturbatively, typically via renormalization group techniques near the upper critical dimension.

In this paper, we develop a novel method for solving a general master equation in which the number of continuous-state variables varies with the discrete-state variable. Our approach is to integrate out the excess continuous degrees of freedom to obtain a hierarchy of marginal probability distributions, analogous to the BBGKY hierarchy in kinetic theory [24]. The time evolution of the one-body marginal distributions, for example, is obtained by integrating the full master equation over all

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

particle positions except that of the first particle (note that the particles are indistinguishable). This yields a set of one-body distributions that depend on the total number of particles but are functions of only a single spatial coordinate.

As a specific example, we focus on a one-dimensional model of particle division, death, and diffusion (DDD), motivated by biological processes such as cell division and apoptosis. The formalism we present is related to "sizer-timer" models of biological cell division [25], where a collection of cells divides based on their (deterministic) cell size and cell age. A master equation formalism has recently been applied to this class of models [26], where cell size dynamics were generalized to stochastic dynamics. Phenomenologically, our DDD model is distinct from work in [26] in that division and death is dependent on particle position, and so Brownian motion due to thermal fluctuations is the main stochastic variable. Additionally, we do not consider the age or size of particles in the dynamics (i.e., division and death occurs with time-independent rate constants). Therefore, our model is applicable to the dynamics of cells, tissues, and bacterial populations, but may also be relevant to chemical systems where diffusive and reactive timescales are comparable, leading to heterogeneous particle distributions.

More generally, our formalism can describe any birth- and death-type system where the discrete components are localized in space with coordinates $\{x_1, \ldots, x_m\}$. Aside from cell-level descriptions, one could also consider a system of receptor molecules in a sea of ligands, the latter of which have local concentration gradients. Then the locations of unbound receptors, $\{x_1, \ldots, x_m\}$, will change due to diffusion and (nonbinding) interactions with other receptors, as well as by binding and unbinding with ligands. A reaction-diffusion-like equation can then be recovered from the one-body marginal probability distributions. Thus our formalism provides a direct link between the phenomenological parameters of reaction-diffusion models and the underlying microscopic dynamics.

In the absence of pairwise interactions such as Lennard-Jones potentials, the steady-state one- and two-body distributions can be solved exactly in Fourier space. From the set of one-body distributions, we can derive the steady-state concentration field of the particles, which is spatially homogenous in the absence of an external field. Interestingly, in the presence of an asymmetric ratchet external potential, the system can exhibit a steady-state macroscopic current, driven by local particle division. From the set of two-body distributions, we derive the steady-state spatial correlation function, which quantifies the probability of finding a pair of particles separated by a given distance. Finally, we compare our analytical predictions with particle-based Gillespie simulations, finding excellent agreement. This confirms the accuracy and utility of our approach as a powerful analytical tool for studying many-body stochastic systems with particle birth and death.

Finally, we extend the concept of entropy production in stochastic thermodynamics [27] to our hybrid-state system by evaluating path probabilities in the full high-dimensional phase space. In addition, we introduce the notion of marginal entropy production, which arises when only a subset of the system's degrees of freedom is accessible, that is, when the dynamics are described by marginal probability distributions. To illustrate this, we compute the entropy production rate in our DDD model. While the total entropy production, evaluated from the full master equation, diverges, the marginal entropy production—obtained by integrating out the excess continuous degrees of freedom—can be finite and even vanish.

II. GENERAL FORMALISM

In this section, we introduce a general formalism to describe a stochastic process in which the number of continuous random variables is itself a random quantity. Additionally, we extend the concept of entropy production, previously defined only for purely discrete or continuous-state systems, to our hybrid-state systems. By integrating out excess degrees of freedom, we show how statistical properties, such as spatial densities and correlations, can be obtained in general. Finally, we also introduce the notion of marginal entropy production, which quantifies the time irreversibility of the system when not all degrees of freedom are accessible.

A. Master equation for hybrid-state systems

At a given time t, the state of the system is characterized by a positive integer $m \in \mathbb{Z}^+$ and a set of continuous variables $\{x\}_m \equiv \{x_1, x_2, \ldots, x_m\}$ of length m. Each continuous variable x_k is constrained to a real interval $x_k \in [-L/2, L/2]$, where $L > 0$ denotes the system size. Generalization to higher dimensional space should be straightforward. As the system evolves in time, the discrete variable m may change stochastically, and consequently, the length of the set $\{x\}_m$ may also vary in time. We refer to such a system as a hybrid-state system. Physically, m may represent the total number of particles at time t (excluding the vacuum state $m = 0$), while $\{x\}_m$ specifies the positions of these particles (assumed to be indistinguishable). Due to birth and death processes, we may add or remove particles from the system and thus both m and $\{x\}_m$ are stochastic variables.

We denote $p_m(\{x\}_m, t)d\{x\}_m$ as the probability of having m particles with positions lying within the interval $d\{x\}_m \equiv dx_1 \ldots dx_m$, centered around $\{x\}_m$, at time t. The probability density is normalized such that

$$\sum_{m=1}^{\infty} \int d\{x\}_m \, p_m(\{x\}_m, t) = 1, \quad \text{for all} \quad t. \tag{1}$$

The space integral is implicitly assumed to be taken over the interval $[-L/2, L/2]$ for each coordinate x_k. We define the transition rate (matrix) operator $\hat{W}_{mn}(\{y\}_m \to \{x\}_n; t)$ to be the rate of going from a state of m particles with coordinates $\{y\}_m$ to a state of n particles with coordinates $\{x\}_n$. Therefore, $\hat{W}_{mn}$ with $m < n$ represents birth processes, with $n - m$ new particles being generated, while $\hat{W}_{mn}$ with $m > n$ represents death processes, with $m - n$ particles being killed. $\hat{W}_{nn}$ is a transition rate in which particle number does not change, but their positions can change. We now introduce the hybrid-state master equation, which describes the time evolution of the probability distributions $\{p_m(\{x\}_m, t)\}$ in terms of these

transition rate operators:

$$\partial_t p_n(\{x\}_n, t) = \sum_{m=1}^{\infty} \int d\{y\}_m [\hat{W}_{mn}(\{y\}_m \to \{x\}_n) p_m(\{y\}_m, t) - \hat{W}_{nm}(\{x\}_n \to \{y\}_m) p_n(\{x\}_n, t)], \quad n \in \mathbb{Z}^+. \tag{2}$$

Note that x_k's and y_k's are just different labels for the particles' positions along the same coordinate axis. Equation (2) evolves from some set of initial conditions $\{p_n(\{x\}_n, 0)\}$ and (periodic) boundary conditions on x_k's to conserve probability. From now on, we assume that the transition rate operators $\hat{W}_{mn}(\{y\}_m \to \{x\}_n; t)$ do not depend explicitly on time t, and thus we have removed the t dependence from $\hat{W}_{mn}$ in Eq. (2).

A particularly interesting situation is when the initial condition is deterministic. We define the propagator $G_{mn}(\{x\}_m \to \{y\}_n; \tau, t)$ as the probability that the system is in a state of n particles with coordinates $\{y\}_n$ at time $t + \tau$, given that it was initially in a state of m particles with coordinates $\{x\}_m$ at time t. In other words, $G_{mn}(\{x\}_m \to \{y\}_n; \tau, t)$ is the solution to Eq. (2) with the initial condition $G(\{x\}_m \to \{y\}_n; 0, t) = \delta_{mn}\delta(\{y\}_n - \{x\}_n)$, where $\delta(\{x\}_n - \{y\}_n) \equiv \prod_{k=1}^{n} \delta(x_k - y_k)$. To lowest order in τ, the solution to this propagator is found to be

$$G_{mn}(\{x\}_m \to \{y\}_n; \tau) = \delta_{mn}\delta(\{y\}_n - \{x\}_n) + \tau \int d\{y'\}_m \hat{W}_{mn}(\{y'\}_m \to \{y\}_n)\delta(\{y'\}_m - \{x\}_m), \tag{3}$$

up to $O(\tau^2)$. We have omitted the dependence on absolute time t in the argument of the propagator, since $\hat{W}_{mn}$ does not depend explicitly on time t.

B. Entropy production rate for hybrid-state systems

The entropy production rate is a measure of the time irreversibility of the stochastic trajectories [27]. More specifically, we define a path or trajectory $\boldsymbol{X}_M \equiv (\{x(t_0)\}_{n_0}, \{x(t_1)\}_{n_1}, \ldots, \{x(t_M)\}_{n_M})$ as a time-ordered set of points in the phase space, starting at time t_0 and visiting state $\{x(t_j)\}_{n_j}$ at successive discrete times $t_j = t_0 + j\tau$ with $j = 0, 1, \ldots, M$. τ is the time interval of successive steps and $\Delta t = M\tau$ the total time passed. The joint probability of observing this path is

$$\mathcal{P}(\boldsymbol{X}_M; t_0, \Delta t) = p_{n_0}(\{x(t_0)\}_{n_0}, t_0) G_{n_0 n_1}(\{x(t_0)\}_{n_0} \to \{x(t_1)\}_{n_1}; \tau) \cdots \times G_{n_{M-1} n_M}(\{x(t_{M-1})\}_{n_{M-1}} \to \{x(t_M)\}_{n_M}; \tau), \tag{4}$$

where $p_{n_0}(\{x(t_0)\}_{n_0}, t_0)$ is the probability of observing the system in state $\{x(t_0)\}_{n_0}$ at time t_0. The reverse path probability is

$$\mathcal{P}^{\mathcal{R}}(\boldsymbol{X}_M; t_0, \Delta t) = \mathcal{P}\big(\boldsymbol{X}_M^{\mathcal{R}}; t_0, \Delta t\big), \tag{5}$$

where $\boldsymbol{X}_M^{\mathcal{R}}$ is the time-reversed trajectory: $\boldsymbol{X}_M^{\mathcal{R}} \equiv (\{x(t_M)\}_{n_M}, \{x(t_{M-1})\}_{n_{M-1}}, \ldots, \{x(t_0)\}_{n_0})$. We define the entropy change of the hybrid-state system over the time interval Δt as the average log ratio of the forward to backward path probabilities:

$$\Delta S_M = \sum_{\mathbf{X}_M} \mathcal{P}(\boldsymbol{X}_M; t_0, \Delta t) \ln\left(\frac{\mathcal{P}(\boldsymbol{X}_M; t_0, \Delta t)}{\mathcal{P}^{\mathcal{R}}(\boldsymbol{X}_M; t_0, \Delta t)}\right). \tag{6}$$

In the equation above, the summation is over all possible trajectories $\boldsymbol{X}_M$, visiting every possible state $\{x\}_n$. By substituting Eqs. (4) and (5) into (6), then dividing by Δt and taking the limit $M \to \infty$ and $\tau \to 0$, we can derive the rate of entropy production $\dot{S}_{\mathrm{irr}}(t)$ at time t (see Appendix A):

$$\dot{S}_{\mathrm{irr}}(t) = \lim_{\tau \to 0} \frac{1}{\tau} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \int d\{x\}_m \int d\{y\}_n \times p_m(\{x\}_m, t) G_{mn}(\{x\}_m \to \{y\}_n; \tau) \times \ln\left(\frac{p_m(\{x\}_m, t) G_{mn}(\{x\}_m \to \{y\}_n; \tau)}{p_n(\{y\}_n, t) G_{nm}(\{y\}_n \to \{x\}_m; \tau)}\right), \tag{7}$$

where the dot represent the total time derivative. It can be shown that $\dot{S}_{\mathrm{irr}}(t)$ is semipositive definite, i.e., $\dot{S}_{\mathrm{irr}}(t) \geqslant 0$ [since the summation can be rewritten as a sum of terms in the form $(x - y)\ln(x/y) > 0$ when $x > 0$ and $y > 0$]. Equation (7) is the first main result of this paper, which generalizes the expression of entropy production previously established for strictly continuous- or discrete-state systems [28,29]. It accounts for the coupling between the variable number of particles present in the system and their corresponding position states, distinguishing it from previous work in which no such coupling was present [30–32]. A similar expression for entropy production for hybrid continuous and discrete states has been derived in Ref. [33]. However, in [33], the number of continuous degrees of freedom, i.e., positions, are assumed to be fixed, whereas in our model they can vary depending on the number of particles. Alternatively, $\dot{S}_{\mathrm{irr}}(t)$ can also be defined as the semipositive definite component of the rate of change of the Shannon entropy. For our hybrid-state systems, the Shannon entropy is defined to be

$$S(t) = -\sum_{n=1}^{\infty} \int d\{x\}_n \, p_n(\{x\}_n, t) \ln[p_n(\{x\}_n, t) L^n], \tag{8}$$

which is an extension of the similar expression given in [28,29]. Here L is the system size (units of length) which is required to ensure the argument of the logarithm is dimensionless. The rate of change $\dot{S}(t)$ can be decomposed into $\dot{S}(t) = \dot{S}_{\mathrm{irr}}(t) + \dot{S}_{\mathrm{e}}(t)$, where $\dot{S}_{\mathrm{irr}}(t) \geqslant 0$. In steady state, $\dot{S}(t \to \infty)$ is zero but $\dot{S}_{\mathrm{irr}}(t \to \infty)$ is not zero in general. Thus, the presence of a finite $\dot{S}_{\mathrm{irr}}(t \to \infty) > 0$ is a hallmark of a nonequilibrium steady state.

If the operators $\hat{W}_{nm}$ decompose additively into terms associated with translations of particle positions (e.g., diffusion or drift processes) and terms that modify the particle number (i.e., death and birth), then the rate of entropy production similarly separates into two distinct contributions. In this case, $\dot{S}_{\mathrm{irr}}$ consists of two terms, one of which is the expression for entropy production found in discrete-state systems (master equation [28]), while the other is that of continuous-state systems (Fokker-Planck equation [29]). Thus, our approach represents a generalization of the entropy production formalism used in the study of, for example, ratchet dynamics in molecular motors [34].

C. Marginal probability distributions

Solving the full probability distribution functions $\{p_n(\{x\}_n, t)\}$ is neither analytically feasible nor particularly useful. One may instead calculate the l-marginal probability distribution functions by integrating out a large proportion of the position variables, thereby tracking only the first l particle positions:

$$p_n^{(l)}(\{x\}_l, t) \equiv \int \left(\prod_{k=l+1}^{n} dx_k \right) p_n(\{x\}_n, t), \tag{9}$$

where the positions of l particles out of n particles are given by $\{x_1, \ldots, x_l\}$, while the positions of the remaining $n-l$ particles are not tracked. Note that when $n \leqslant l$, $p_n^{(l)}(\{x\}_l, t) = p_n(\{x\}_n, t)$; see Fig. 1 for illustration. This is closely related to the l-body density of a system, which may be calculated as

$$\rho^{(l)}(\{x\}_l, t) \equiv \sum_{n=l}^{\infty} \frac{n!}{(n-l)!} p_n^{(l)}(\{x\}_l, t). \tag{10}$$

For example, the average density of a population of particles as a function of space x at time t is given by $\rho(x,t) = \rho^{(1)}(x_1 = x, t)$, whereas the two-body correlation function is given by $\rho^{(2)}(x_1, x_2; t)$.

We determine the marginal distribution dynamics by integrating the total dynamics, Eq. (2), over the excess $n-l$ positions as in Eq. (9). These integrals can be carried out formally for any set of transition matrices $\hat{W}_{mn}(\{y\}_m \to \{x\}_n)$. However, in general, the dynamics of the l-marginal distribution $p_n^{(l)}(\{x\}_l, t)$ might depend explicitly on the $l+1$-marginal distribution $p_n^{(l+1)}(\{x\}_{l+1}, t)$. This coupling to the $l+1$ marginal is a familiar problem in (constant particle number) statistical mechanics for interacting systems [24], where a closure relation such as the "molecular chaos hypothesis" [24] is necessary to write the $l+1$ marginal in terms of the l marginal. (Note that in this paper we do not consider explicit interactions between the particles so no closure relation is necessary.) Once this is done, the l-marginal transition rate operators $\hat{W}_{mn}^{(l)}$ can be defined formally.

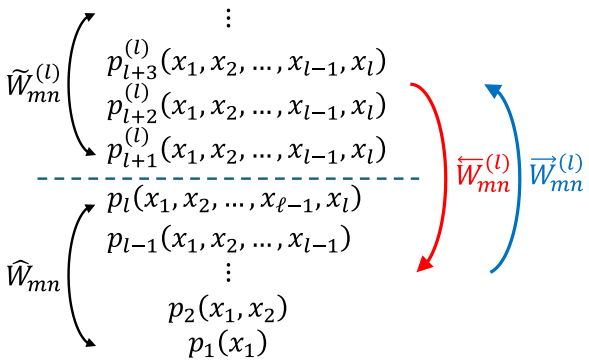


FIG. 1. The l-marginal distribution functions $p_n^{(l)}(x_1, x_2, \ldots, x_l)$, where $n > l$, are defined by integrating out the excess coordinates $x_{l+1}, x_{l+2}, \ldots, x_n$ from the full distribution functions $p_n(x_1, x_2, \ldots, x_n)$. The l-marginal transition rate operators $W_{mn}^{(l)}(\{y\}_m \to \{x\}_n)$ can also be defined formally. We distinguish $\overrightarrow{W}_{mn}^{(l)}$ for the transition rate from a state with $m \leqslant l$ particles to a state with $n > l$ particles, and $\overleftarrow{W}_{mn}^{(l)}$ for the transition rate from a state with $m > l$ particles to a state with $n \leqslant l$ particles. $\widetilde{W}_{mn}^{(l)}$ represents the transition rate from a state with $m > l$ particles to a state with $n > l$ particles.

To this end, it is also necessary to distinguish the l-marginal transition rate operators into: $\widetilde{W}_{mn}^{(l)}$, $\overrightarrow{W}_{mn}^{(l)}$, or $\overleftarrow{W}_{mn}^{(l)}$, depending on whether they correspond to transitions (1) from a state with $m > l$ to one with $n > l$ particles, (2) from a state with $m \leqslant l$ to one with $n > l$ particles, or (3) vice versa; see Fig. 1 for illustration. This is because constraints must be imposed in two cases: $m > l$ and $n > l$ [in the definition of $\overrightarrow{W}_{mn}^{(l)}$, Eq. (12)], and when $n \leqslant l < m$ [in the definition of $\overleftarrow{W}_{mn}^{(l)}$, Eq. (13)]. Then the dynamics for the l-marginal distribution, $p_n^{(l)}(\{x\}_l, t)$ can be written as

$$\partial_t p_n^{(l)}(\{x\}_l, t) = \begin{cases} \displaystyle\sum_{m=1}^{l} \int d\{y\}_m \left[\overrightarrow{W}_{mn}^{(l)}(\{y\}_m \to \{x\}_l) p_m(\{y\}_m, t) - \overleftarrow{W}_{mn}^{(l)}(\{x\}_l \to \{y\}_m) p_n^{(l)}(\{x\}_l, t) \right] \\ \quad + \displaystyle\sum_{m=l+1}^{\infty} \int d\{y\}_l \left[\widetilde{W}_{mn}^{(l)}(\{y\}_l \to \{x\}_l) p_m^{(l)}(\{y\}_l, t) - \widetilde{W}_{nm}^{(l)}(\{x\}_l \to \{y\}_l) p_n^{(l)}(\{x\}_l, t) \right], \quad n > l, \\ \displaystyle\sum_{m=1}^{l} \int d\{y\}_m \left[\hat{W}_{mn}(\{y\}_m \to \{x\}_n) p_m(\{y\}_m, t) - \hat{W}_{nm}(\{x\}_n \to \{y\}_m) p_n(\{x\}_n, t) \right] \\ \quad + \displaystyle\sum_{m=l+1}^{\infty} \int d\{y\}_l \left[\overleftarrow{W}_{mn}^{(l)}(\{y\}_l \to \{x\}_n) p_m^{(l)}(\{y\}_l, t) - \overrightarrow{W}_{nm}^{(l)}(\{x\}_n \to \{y\}_l) p_n(\{x\}_n, t) \right], \quad n \leqslant l, \end{cases} \tag{11}$$

$$\widetilde{W}_{mn}^{(l)}(\{y\}_l \to \{x\}_l) \equiv \int \left(\prod_{k=l+1}^{n} dx_k \right) \hat{W}_{mn}(\{y\}_m \to \{x\}_n), \quad m > l, \quad n > l, \tag{12}$$

$$\overleftarrow{W}_{mn}^{(l)}(\{y\}_l \to \{x\}_n) \equiv \hat{W}_{mn}(\{y\}_m \to \{x\}_n), \quad m > l \geqslant n, \tag{13}$$

$$\overrightarrow{W}_{mn}^{(l)}(\{y\}_m \to \{x\}_l) \equiv \int \left(\prod_{k=l+1}^{n} dx_k \right) \hat{W}_{mn}(\{y\}_m \to \{x\}_n), \quad n > l \geqslant m, \tag{14}$$

$$G^{(l)}_{mn}(\{y\}_{m/l} \to \{x\}_{n/l};\tau) = \begin{cases} G_{mn}(\{y\}_m \to \{x\}_n;\tau), & m \leqslant l, \quad n \leqslant l, \\ \tau \int d\{y'\}_l \overleftarrow{W}^{(l)}_{mn}(\{y'\}_l \to \{x\}_n)\delta(\{y'\}_l - \{y\}_l), & m > l \geqslant n, \\ \delta_{mn}\delta(\{y\}_l - \{x\}_l) + \tau \int d\{y'\}_l \widetilde{W}^{(l)}_{mn}(\{y'\}_l \to \{x\}_l)\delta(\{y'\}_l - \{y\}_l), & m > l, \quad n > l, \\ \tau \int d\{y'\}_m \overrightarrow{W}^{(l)}_{mn}(\{y'\}_m \to \{x\}_l)\delta(\{y'\}_m - \{y\}_m), & n > l \geqslant m. \end{cases} \tag{15}$$

The constraint for $m > l$ and $n > l$ [Eq. (12)] requires $\hat{W}_{mn}$ to be independent of $y_{l+1}, y_{l+2}, \ldots, y_m$, since the left-hand side of the equation is only a function of $y_1, y_2, \ldots, y_l$. The constraint for $m > l \geqslant n$ [Eq. (13)] requires $\hat{W}_{mn}$ to be independent of $y_{l+1}, y_{l+2}, \ldots, y_m$ for the same reason. Equation (14) simply integrates out degrees of freedom, so does not constrain the transition matrix operator when $n > l \geqslant m$, but is defined simply for convenience. Thus, Eq. (11) is *only true* when these constraints on $\hat{W}_{mn}$ are met. So, subject to Eqs. (12) and (13), we find that the marginal distributions can be solved recursively starting from $l = 1$ (one-body distributions).

Finally, Eq. (15) defines the short-time marginal propagator subject to Eqs. (12) and (13) up to $O(\tau^2)$. The subscript notations m/l and n/l in $G^{(l)}_{mn}$ are meant to identify that there are m particles transitioning to n particles, but we emphasize that if $m > l$ and/or $n > l$, the propagator depends only on $\{y\}_l$ and/or $\{x\}_l$ instead of $\{y\}_m$ and/or $\{x\}_n$, since the last $l - m$ and/or $l - n$ position variables are integrated out. Note that when $n \leqslant l$ and $m \leqslant l$, this marginal propagator just reduces to Eq. (3) as it should.

D. Marginal entropy production rates

In addition to the marginal probability distributions, one can also construct an analog to the entropy production rate Eq. (7), considering up to the first l position variables. We refer to this quantity as the l-marginal entropy production rate, and denote it by $\dot{S}^{(l)}_{\rm irr}$. Similar to the total entropy production rate derived in Sec. II B and Appendix A, $\dot{S}^{(l)}_{\rm irr}$ is defined as the expectation of the logarithm of the ratio of forward to reverse path probabilities in the reduced phase space. In our derivation, we require that Eqs. (12) and (13) are satisfied. Our l-marginal entropy production rate is related to, yet distinct from, the *coarse-grained* entropy production rate obtained by spatially averaging the particle current $j(x) \to \langle j\rangle(x)$ [35–37], similar to a renormalization group procedure [38].

Following the procedure for the full entropy production rate calculation described in Sec. II B, one can define the l-marginal path probability

$$\mathcal{P}^{(l)}(\boldsymbol{X}_M; t_0, \Delta t) = p^{(l)}_{n_0}(\{x(t_0)\}_{n_0/l}, t_0) G^{(l)}_{n_0 n_1}(\{x(t_0)\}_{n_0/l} \to \{x(t_1)\}_{n_1/l};\tau) \cdots G^{(l)}_{n_{M-1} n_M}(\{x(t_{M-1})\}_{n_{M-1}/l} \to \{x(t_M)\}_{n_M/l};\tau), \tag{16}$$

where the propagators $G^{(l)}_{nm}$ (and the initial probability $p^{(l)}_{n_0}$) depend at most on the first l positions as mentioned below Eq. (15). Additionally, the l-marginal reverse path probability is defined as in Eq. (5), but using Eq. (16) instead of Eq. (4). With a very similar calculation to the full entropy production rate, we arrive at the positive definite, l-marginal entropy production rate, which is the second main result of our paper,

$$\begin{aligned} \dot{S}^{(l)}_{\rm irr}(t) = \lim_{\tau\to 0} \Bigg\{ & \frac{1}{\tau} \sum_{n=1}^{l} \sum_{m=1}^{l} \int d\{x\}_n \int d\{y\}_m \, G_{mn}(\{y\}_m \to \{x\}_n;\tau) p_m(\{y\}_m) \ln \left[\frac{G_{mn}(\{y\}_m \to \{x\}_n;\tau) p_m(\{y\}_m)}{G_{nm}(\{x\}_n \to \{y\}_m;\tau) p_n(\{x\}_n)} \right] \\ & + \frac{1}{2\tau} \sum_{n=1}^{l} \sum_{m=l+1}^{\infty} \int d\{x\}_n \int d\{y\}_l \, \overleftarrow{G}^{(l)}_{mn}(\{y\}_l \to \{x\}_n;\tau) p^{(l)}_m(\{y\}_l) \ln \left[\frac{\overleftarrow{G}^{(l)}_{mn}(\{y\}_l \to \{x\}_n;\tau) p^{(l)}_m(\{y\}_m)}{\overrightarrow{G}^{(l)}_{nm}(\{x\}_n \to \{y\}_l;\tau) p_n(\{x\}_n)} \right] \\ & + \frac{1}{\tau} \sum_{n=l+1}^{\infty} \sum_{m=l+1}^{\infty} \int d\{x\}_l \int d\{y\}_l \, \widetilde{G}^{(l)}_{mn}(\{y\}_l \to \{x\}_l;\tau) p^{(l)}_m(\{y\}_l) \ln \left[\frac{\widetilde{G}^{(l)}_{mn}(\{y\}_l \to \{x\}_l;\tau) p^{(l)}_m(\{y\}_m)}{\widetilde{G}^{(l)}_{nm}(\{x\}_l \to \{y\}_l;\tau) p^{(l)}_n(\{x\}_n)} \right] \Bigg\}, \end{aligned} \tag{17}$$

where $\overleftarrow{G}_{mn}$, $\overrightarrow{G}_{mn}$, and $\widetilde{G}_{mn}$ are the propagators associated with $\overleftarrow{W}_{mn}$, $\overrightarrow{W}_{mn}$, and $\widetilde{W}_{mn}$ respectively. Additionally, the l-marginal Shannon entropy can be similarly defined,

$$\begin{aligned} S^{(l)}(t) = & -\sum_{n=1}^{l} \int d\{x\}_n \, p_n(\{x\}_n, t) \ln[p_n(\{x\}_n, t) L^n] \\ & - \sum_{n=l+1}^{\infty} \int d\{x\}_l \, p^{(l)}_n(\{x\}_l, t) \ln\left[p^{(l)}_n(\{x\}_l, t) L^l\right], \end{aligned} \tag{18}$$

and the marginal rate of entropy production rate $\dot{S}^{(l)}_{\rm irr}(t)$ can also be obtained from the semipositive definite part of $\dot{S}^{(l)}(t)$.

III. DIVISION, DEATH, AND DIFFUSION IN THE ABSENCE OF EXTERNAL FIELD

Having presented a formal description of stochastic hybrid-state systems and their (marginal) probability distributions and entropy production rates, we next proceed to apply these results to a specific model, which we refer to as the division, death, and diffusion (DDD) model. In this model particles can divide into two with rate r_b, die with rate $r_a(n-1)$, or undergo

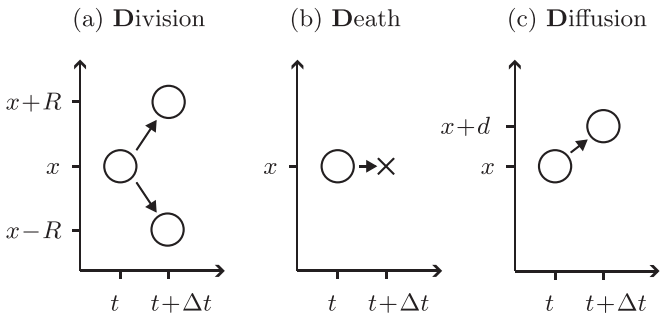


FIG. 2. Schematic of (a) division, (b) death, and (c) diffusion processes in the DDD model. Exactly one of these processes occurs in timestep Δt for each particle (in the limit $\Delta t \to 0$). We highlight three possible transitions of a single particle which starts at position x and time t. Division or birth occurs with rate r_b and death with rate $r_a(n-1)$. Unbiased diffusion occurs at every time step where division or death does not occur. For every division event, the parent particle at x is removed and two new particles are created at $x+R$ and $x-R$. We also assume periodic boundary conditions at $x=-L/2$ and $x=L/2$, and $x \in [-L/2, L/2]$.

random walks with diffusion constant D, where n denotes the total number of particles at a given time (see Fig. 2). The death rate scales with $n-1$ to prevent transitions into the vacuum state ($n=0$). In stochastic thermodynamics, the rate constants r_a and r_b are typically expressed as exponentials of activation energies divided by temperature [27,39]. Although particles do not interact directly, division events create spatial correlations between the resulting particles [see Fig. 2(a)]. We will explore how nonequilibrium processes, such as particle division and death, influence the entropy production rate. By first deriving the 1-marginal and 2-marginal probability distributions, we will also demonstrate how the average density and two-body correlations can be computed exactly.

A. Perfectly homogeneous system (division and death only)

As a warm-up to our later calculations, we first present the entropy production rate for a division and death system which neglects spatial fluctuations completely. In this case, the state of the system is characterized entirely by the number of particles n at a given time t with the corresponding probability mass function $p_n(t)$. The time evolution for $p_n(t)$ is exactly described by the autocatalytic process [40]

$$\frac{dp_n(t)}{dt} = \underbrace{r_a n(n+1)}_{W_{n+1,n}} p_{n+1}(t) + \underbrace{r_b(n-1)}_{W_{n-1,n}} p_{n-1}(t) - \underbrace{[r_a n(n-1) + r_b n]}_{W_{nn}} p_n(t), \tag{19}$$

from which we can read off the transition rates:

$$W_{mn} = r_a n(n+1)\delta_{m,n+1} + r_b(n-1)\delta_{m,n-1} - [r_a n(n-1) + r_b n]\delta_{m,n}. \tag{20}$$

Note that the transition rate only increments the total number of particles by ± 1 at most. Using the generating function formalism [13], Eq. (19) can be solved exactly in steady state:

$$\pi_n \equiv \lim_{t\to\infty} p_n(t) = \left(\frac{r_b}{r_a}\right)^n \frac{1}{n!} \frac{1}{e^{r_b/r_a} - 1}. \tag{21}$$

Consequently, the average number of particles $\langle n \rangle$ and the variance in the number of particles $\langle n^2 \rangle - \langle n \rangle^2$ can also be computed in steady state:

$$\langle n \rangle_{\rm ss} = \sum_{n=1}^{\infty} n\pi_n = \left(\frac{r_b}{r_a}\right) \frac{e^{r_b/r_a}}{e^{r_b/r_a} - 1}, \tag{22}$$

$$\langle n^2 \rangle_{\rm ss} - \langle n \rangle_{\rm ss}^2 = \frac{r_b e^{r_b/r_a}(e^{r_b/r_a} - 1 - r_b/r_a)}{r_a(e^{r_b/r_a} - 1)^2}, \tag{23}$$

where the subscript "ss" indicates averaging in the steady state. When $r_b \gg r_a$, the average number of particles in the system is given by $\langle n \rangle_{\rm ss} \approx r_b/r_a$ in steady state.

Direct substitution of this steady-state mass function, π_n, along with the transition rate, W_{mn}, into the definition of the entropy production rate for discrete-state systems [28,29] yields, in steady state,

$$\lim_{t\to\infty} \dot{S}_{\rm irr}(t) = \sum_{m,n} \pi_n W_{nm} \ln\left(\frac{\pi_n W_{nm}}{\pi_m W_{mn}}\right) = 0. \tag{24}$$

Therefore, there is no entropy production in the spatially homogeneous case. We note that the assumption of homogeneously distributed particles is often employed when using master equations to describe chemical reactions, given the relatively slow timescale at which reactions occur relative to the spatial diffusion of reacting molecules.

B. Spatial diffusion included (DDD model)

We now extend the previous analysis to a one-dimensional system of particles that can undergo division and death processes while diffusing independently under periodic boundary conditions with a period length L; see Fig. 2. Although the particles do not interact directly, each division event generates spatial correlations, as division produces two new particles separated by a fixed distance $2R$; see Fig. 2(a). Therefore, the DDD model is an extension of the simple branching Brownian motion processes [41] by including a nonzero splitting radius R and a finite death rate r_a. We also assume the absence of an external field. In Sec. IV we will explore the effect of an external field, in particular a ratchet potential. The master equation for the full dynamics is

given by

$$\partial_t p_n(\{x\}_n,t) = \sum_{i=1}^{n} -\partial_{x_i}[\hat{j}_{x_i} p_n(\{x\}_n,t)] + r_a n(n+1)\int dx_{n+1}\, p_n(x_1,\ldots,x_n,x_{n+1}) - [r_a n(n-1)+bn]p_n(\{x\}_n,t)$$
$$+ \frac{r_b}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\int dy\,\delta[x_i - f_{-R}(y)]\delta[x_j - f_{+R}(y)]p_{n-1}(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_{i-1},y,x_{i+1},\ldots,x_n;t)$$
$$+ \frac{r_b}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\int dy\,\delta[x_i - f_{-R}(y)]\delta[x_j - f_{+R}(y)]p_{n-1}(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_{j-1},x_{j+1},\ldots,x_n;t), \quad (25)$$

where we have introduced the operator $\hat{j}_{x_i}$ and the function $f_{\pm R}(y)$ to be

$$\hat{j}_{x_i} = -D\partial_{x_i} \quad \text{and} \quad f_{\pm R}(y) = y \pm R. \quad (26)$$

In Sec. IV we will consider a generalization to $\hat{j}_{x_i}$ and $f_{\pm R}(y)$. The first term in Eq. (25) represents diffusion. The second term accounts a death event occurring at position x_{n+1}. The third term acts as a sink term for the n-particle state, capturing the loss of probability due to division or death events of any particle. The fourth and the fifth terms describe a division event at position y, where a particle splits into two new particles located at $x_i = y - R$ and $x_j = y + R$. The summation ensures permutation over all position indices, maintaining the invariance of the probability distribution $p_n(x_1,\ldots,x_n)$ under the exchange of any two particle indices, as required by the indistinguishability condition.

A more realistic variation of the model incorporates stochasticity in the division separation distance $2R$. Instead of the two new particles being created at a fixed separation $\pm R$ relative to the parent particle, their separation is given by $\pm r$, where r is a random variable drawn from a Gaussian distribution with mean R and variance σ_R^2. In this case, the Dirac delta function $\delta(x)$ in Eq. (25) is replaced by a Gaussian function:

$$\delta_{\sigma_R}(x) = \frac{1}{\sqrt{2\pi\sigma_R^2}} e^{-\frac{x^2}{2\sigma_R^2}}. \quad (27)$$

The deterministic division separation can then be recovered by taking the limit $\sigma_R \to 0$. In our simulation results, we assume a deterministic division separation of $2R$ for simplicity. However, in Sec. III C we also discuss the effect of random separation distance on the rate of entropy production.

To get the 0-marginal probability distributions $p_n^{(0)}(t)$, we integrate Eq. (25) over all position variables $x_1, x_2, \ldots$. The result is identical to the perfectly homogenous master equation (19) with the identification $p_n^{(0)}(t) = p_n(t)$. Thus at this level, the 0-marginal entropy production rate $\dot{S}_{\rm irr}^{(0)}$ is zero in steady state.

C. 1-marginal distributions and 1-marginal entropy production rate

Integrating over $x_2, x_3, \ldots$ in Eq. (25) gives us the time evolution for the 1-marginal probability distributions

$$\partial_t p_n^{(1)}(x_1) = D\partial_{x_1}^2 p_n^{(1)}(x_1) + r_a n(n+1)p_{n+1}^{(1)}(x_1) - [r_a n(n-1) + r_b n]p_n^{(1)}(x_1) + \frac{r_b}{n}(n-1)(n-2)p_{n-1}^{(1)}(x_1) + \frac{r_b}{n}(n-1)\big[p_{n-1}^{(1)}(x_1 - R) + p_{n-1}^{(1)}(x_1 + R)\big], \quad (28)$$

for $n \in \mathbb{Z}^+$. Note that we have not written the time dependence explicitly inside the argument of $p_n^{(1)}$ for brevity. From Eq. (28), we can then read off the 1-marginal transition rates:

$$\widetilde{W}_{n-1,n}^{(1)}(y_1 \to x_1) = \frac{r_b(n-1)(n-2)}{n}\delta(y_1 - x_1) + \frac{r_b(n-1)}{n}[\delta(y_1 - x_1 + R) + \delta(y_1 - x_1 - R)] \quad (29)$$

$$\widetilde{W}_{n,n}^{(1)}(y_1 \to x_1) = D\partial_{y_1}^2\delta(y_1 - x_1) \quad (30)$$

$$\widetilde{W}_{n+1,n}^{(1)}(y_1 \to x_1) = r_a n(n-1)\delta(y_1 - x_1), \quad (31)$$

and $\hat{W}_{11} = \widetilde{W}_{11}^{(1)}$, $\overleftarrow{W}_{21}^{(1)} = \widetilde{W}_{21}^{(1)}$, and $\overrightarrow{W}_{12}^{(1)} = \widetilde{W}_{12}^{(1)}$. Thus for 1-marginal distributions, it turns out that we do not need to distinguish between states with $n \leqslant l$ particles and those with $n > l$ particles.

Since $x_1 \in [-L/2, L/2]$ and assuming periodic boundary conditions, we can Fourier transform $p_n^{(1)}(x_1,t)$:

$$p_n^{(1)}(x_1,t) = \sum_{k\in\mathbb{Z}} c_n^{(1)}(k,t)e^{i\frac{2\pi k}{L}x_1}. \quad (32)$$

Equation (28) can then be Fourier transformed to get

$$\partial_t \boldsymbol{c}^{(1)}(k,t) = \boldsymbol{A}^{(1)}(k)\boldsymbol{c}^{(1)}(k,t), \quad (33)$$

where $k \in \mathbb{Z}$. In the above equation, $\boldsymbol{c}^{(1)}(k,t) = [c_1^{(1)}(k,t), c_2^{(1)}(k,t), \ldots]^T$ is a semi-infinite column vector and $\boldsymbol{A}^{(1)}(k)$ is a semi-infinite tridiagonal matrix with nonzero elements given by

$$A_{n,n-1}^{(1)}(k) = \frac{r_b(n-1)}{2n}\left[\cos\left(\frac{2\pi Rk}{L}\right) + \frac{n-2}{2}\right], \quad (34)$$

$$A_{n,n}^{(1)}(k) = -\frac{D(2\pi k)^2}{L^2} - [r_a n(n-1) + r_b n], \quad (35)$$

$$A_{n,n+1}^{(1)}(k) = r_a n(n+1). \quad (36)$$

From a probabilistic argument, the real parts of all eigenvalues of $\boldsymbol{A}(k)$ must be nonpositive for all values of k. Otherwise,

the probability distribution would diverge over time instead of remaining normalized to unity. Thus, the steady-state solution corresponds to the eigenvector associated with the zero eigenvalue, which happens to occur at $k = 0$. The $k = 0$ mode represents a spatially homogenous solution, where $p_n^{(1)}(x_1)$ is independent of the position x_1 for all $n \in \mathbb{Z}^+$. Moreover, the steady-state 1-marginal distribution matches the perfectly homogenous solution (apart from the normalization factor $1/L$): $p_n^{(1)}(x_1, \infty) = \pi_n/L$, where π_n is given in Eq. (21) in Sec. III A.

Correspondingly, the average density of the particles can be computed using Eq. (10), and in steady state, this is given by

$$\rho(x, \infty) = \lim_{t\to\infty} \sum_{n=1}^{\infty} n p_n^{(1)}(x, t) = \sum_{n=1}^{\infty} n \frac{\pi_n}{L} = \frac{\langle n \rangle_{\rm ss}}{L}, \quad (37)$$

where $\langle n \rangle_{\rm ss}$ is the average total number of particles in steady state, given in Eq. (22). Thus, while the particle density $\rho(x, t)$ may initially exhibit spatial heterogeneities, over time, these spatial variations gradually diminish. In the steady state, the average density becomes spatially uniform, $\rho(x, \infty) = \text{const}$, as expected due to translational symmetry in the system.

We can now substitute the steady-state distributions π_n/L and the transition rates $\widetilde{W}_{mn}^{(1)}$ [Eqs. (29)–(31)] into Eq. (17) to get the 1-marginal entropy production rate:

$$\begin{aligned}\dot{S}_{\rm irr}^{(1)} = \sum_{n=2}^{\infty} \left(\frac{r_b}{r_a}\right)^n \frac{r_a}{n!L} \frac{(n-1)}{e^{r_b/r_a} - 1} \int dx \int dy [\delta(x + R - y) \\ + \delta(x - R - y) - 2\delta(x - y)] \\ \times \ln\left[\frac{(n-2)\delta(x-y) + \delta(x+R-y) + \delta(x-R-y)}{n\delta(x-y)}\right].\end{aligned} \quad (38)$$

To compute the integral of a logarithm of a delta function in Eq. (38), we first need to regularize the delta function as a Gaussian function $\delta_{\sigma_R}(x)$; see Eq. (27). Physically, we now allow the division separation distance to be random, rather than fixed at $\pm R$, and σ_R^2 represents the variance of the separation distance; see discussion above Eq. (27). The 1-marginal entropy production rate can then be computed exactly

$$\dot{S}_{\rm irr}^{(1)} = \begin{cases} \infty, & R \neq 0 \text{ and } \sigma_R \to 0, \\ \frac{R^2}{\sigma_R^2}\left(\frac{r_b}{e^{r_b/r_a} - 1} + r_b - r_a\right), & R \neq 0 \text{ and } \sigma_R > 0, \\ 0, & R = 0. \end{cases} \quad (39)$$

Thus, even though the steady-state average density $\rho(x, t \to \infty)$ is spatially uniform, we still have a positive and finite rate of entropy production in the steady state (for the case of $R \neq 0$ and $\sigma_R > 0$). This is because entropy production captures the dynamics between different particle number states, e.g., between $p_n^{(1)}(x_1)$ and $p_{n+1}^{(1)}(x_1)$. In the limit of deterministic separation distance $\sigma_R \to 0$, the rate of entropy production becomes infinite. This infinite entropy production stems from the improbability of the reverse process: two particles approaching each other to a separation distance $2R$, being removed, and replaced by a single particle at their center of mass.

D. 2-marginal distributions and two-body correlation function

In this section we assume a fixed division separation of $2R$ after each division event, corresponding to the limit $\sigma_R \to 0$, to make our calculations analytically tractable. Integrating over $x_3, x_4, x_5, \ldots$ in Eq. (25), we obtain the time evolution for the 2-marginal probability distributions:

$$\begin{aligned}\partial_t p_n^{(2)}(x_1, x_2) \\ = \left\{D\left(\partial_{x_1}^2 + \partial_{x_2}^2\right) - [r_a n(n-1) + r_b n]\right\} p_n^{(2)}(x_1, x_2) \\ + r_a n(n+1) p_{n+1}^{(2)}(x_1, x_2) + \frac{r_b}{n}(n-2)(n-3) p_{n-1}^{(2)}(x_1, x_2) \\ + \frac{r_b}{n}\left[p_{n-1}^{(2)}(x_1 + R, x_2) + p_{n-1}^{(2)}(x_1 - R, x_2)\right] \\ + \frac{r_b}{n}\left[p_{n-1}^{(2)}(x_1, x_2 + R) + p_{n-1}^{(2)}(x_1, x_2 - R)\right] \\ + \frac{r_b}{n}\delta(x_2 - x_1 - 2R) p_{n-1}^{(1)}(x_1 + R) \\ + \frac{r_b}{n}\delta(x_2 - x_1 + 2R) p_{n-1}^{(1)}(x_1 - R),\end{aligned} \quad (40)$$

for $n \geqslant 2$, and,

$$\partial_t p_1(x_1) = D\partial_{x_1}^2 p_1(x_1) + 2r_a p_2^{(1)}(x_1) - r_b p_1(x_1), \quad (41)$$

for $n = 1$. Notice that the time evolution of the 2-marginal distributions depends on the 1-marginal distributions. Therefore, solving Eqs. (40) and (41) also requires solving Eq. (28) simultaneously. To solve these equations, we again use the Fourier transform (assuming periodic boundary conditions):

$$p_n^{(2)}(x_1, x_2; t) = \sum_{k_1, k_2 \in \mathbb{Z}} c_n^{(2)}(k_1, k_2; t) e^{i\frac{2\pi k_1}{L} x_1} e^{i\frac{2\pi k_2}{L} x_2}, \quad (42)$$

while the Fourier transform of $p_n^{(1)}(x_1, t)$ is still given by Eq. (32). Using Eqs. (32) and (42), Eq. (40) can then be Fourier transformed to

$$\begin{aligned}\partial_t c_n^{(2)}(k_1, k_2; t) \\ = \left\{-\frac{D(2\pi)^2\left(k_1^2 + k_2^2\right)}{L^2} - [r_a n(n-1) + r_b n]\right\} c_n^{(2)} \\ + r_a n(n+1) c_{n+1}^{(2)} + \frac{r_b}{n}(n-2)(n-3) c_{n-1}^{(2)} \\ + \frac{2r_b}{n}(n-2)\left[\cos\left(\frac{2\pi k_1 R}{L}\right) + \cos\left(\frac{2\pi k_2 R}{L}\right)\right] c_{n-1}^{(2)} \\ + \frac{2r_b}{Ln}\cos\left[\frac{2\pi R}{L}(k_1 - k_2)\right] c_{n-1}^{(1)}(k_1 + k_2; t),\end{aligned} \quad (43)$$

where $n \geqslant 2$ and the argument of $c_n^{(2)}$ on the right-hand side of the equation above is implied to be $(k_1, k_2; t)$. The time evolution for $c_n^{(1)}(k, t)$ and $c_n^{(2)}(k_1, k_2; t)$ can then be written as a matrix form:

$$\partial_t \boldsymbol{c}^{(1)+(2)}(k_1, k_2; t) = \boldsymbol{A}^{(1)+(2)}(k_1, k_2) \boldsymbol{c}^{(1)+(2)}(k_1, k_2; t), \quad (44)$$

where k_1 and k_2 are integers. In the above equation, $\boldsymbol{c}^{(1)+(2)}(k_1, k_2; t)$ is the column vector

$$\boldsymbol{c}^{(1)+(2)}(k_1, k_2; t) = \begin{bmatrix} c_1^{(1)}(k_1 + k_2; t) \\ c_2^{(1)}(k_1 + k_2; t) \\ \vdots \\ c_2^{(2)}(k_1, k_2; t) \\ c_3^{(2)}(k_1, k_2; t) \\ \vdots \end{bmatrix}, \tag{45}$$

and $\boldsymbol{A}^{(1)+(2)}(k_1, k_2)$ is the block matrix

$$\boldsymbol{A}^{(1)+(2)}(k_1, k_2) = \begin{bmatrix} \boldsymbol{A}^{(1)}(k_1 + k_2) & \boldsymbol{0} \\ \boldsymbol{B}(k_1 - k_2) & \boldsymbol{A}^{(2)}(k_1, k_2) \end{bmatrix}. \tag{46}$$

In the definition above, $\boldsymbol{A}^{(1)}(k)$ is a semi-infinite tridiagonal square matrix, as defined in Eqs. (34)–(36), $\boldsymbol{B}(k)$ is a semi-infinite diagonal matrix, with its diagonal elements given by

$$B_{n-1,n-1}(k) = \frac{2r_b}{nL} \cos\left(\frac{2\pi kR}{L}\right), \tag{47}$$

and finally $\boldsymbol{A}^{(2)}(k_1, k_2)$ is another semi-infinite tridiagonal square matrix, with its nonzero elements given by

$$A^{(2)}_{n-1,n-2}(k_1, k_2) = \frac{r_b(n-2)}{n}\left[(n-3) + 2\cos\left(\frac{2\pi k_1 R}{L}\right) + 2\cos\left(\frac{2\pi k_2 R}{L}\right)\right], \tag{48}$$

$$A^{(2)}_{n-1,n-1}(k_1, k_2) = -\frac{D(2\pi)^2\left(k_1^2 + k_2^2\right)}{L^2} - [k_a n(n-1) + k_b n], \tag{49}$$

$$A^{(2)}_{n-1,n}(k_1, k_2) = r_a n(n+1). \tag{50}$$

$\boldsymbol{0}$ in Eq. (46) is a semi-infinite matrix with zero as its elements.

The matrix equation (44) cannot be solved directly. Instead, we first need to truncate the semi-infinite matrix $\boldsymbol{A}^{(1)}$ to obtain a finite $N \times N$ matrix, denoted as $\boldsymbol{A}^{(1)}_{N\times N}$, which consists of the first N rows and N columns of the original matrix $\boldsymbol{A}^{(1)}$. This approximation is valid as long as N is much larger than the average number of particles, i.e., $N \gg \langle n\rangle_{\rm ss}$. Similarly, we also truncate $\boldsymbol{A}^{(2)}$ to obtain $\boldsymbol{A}^{(2)}_{(N-1)\times(N-1)}$, and truncate $\boldsymbol{B}$ to obtain $\boldsymbol{B}_{(N-1)\times N}$. The matrix $\boldsymbol{A}^{(1)+(2)}$ in Eq. (46) can then be truncated as a $(2N-1)\times(2N-1)$ finite block matrix:

$$\boldsymbol{A}^{(1)+(2)}_{(2N-1)\times(2N-1)} = \begin{bmatrix} \boldsymbol{A}^{(1)}_{N\times N} & \boldsymbol{0}_{N\times(N-1)} \\ \boldsymbol{B}_{(N-1)\times N} & \boldsymbol{A}^{(2)}_{(N-1)\times(N-1)} \end{bmatrix}, \tag{51}$$

and we have $\boldsymbol{A}^{(1)+(2)} = \lim_{N\to\infty} \boldsymbol{A}^{(1)+(2)}_{(2N-1)\times(2N-1)}$. We are generally interested in the steady-state solution, i.e., we are looking for the zero eigenvalue of the matrix $\boldsymbol{A}^{(1)+(2)}_{(2N-1)\times(2N-1)}$. To achieve this, we shall use the following result.

Proposition (Eigenvalues of block matrices). If λ is an eigenvalue of $\boldsymbol{A}^{(1)}_{N\times N}$ or $\boldsymbol{A}^{(2)}_{(N-1)\times(N-1)}$, then λ is also an eigenvalue of $\boldsymbol{A}^{(1)+(2)}_{(2N-1)\times(2N-1)}$. The proof follows directly from the properties of the determinant of a block matrix [42].

From Sec. III C, we know that the matrix $\boldsymbol{A}^{(1)}_{N\times N}(k_1 + k_2)$ has a zero eigenvalue when $k_1 + k_2 = 0$. The condition $k_1 + k_2 = 0$ is consistent with the translational symmetry of the system, which dictates that the probability distribution $p_n^{(2)}(x_1, x_2)$ depends solely on $x_2 - x_1$ in the steady state. The corresponding eigenvector for this zero eigenvalue is $\boldsymbol{\Pi}_N = \alpha(N)[\pi_1, \pi_2, \ldots, \pi_N]^T$, where π_n is given in Eq. (21). $\alpha(N)$ is some normalization factor and in the limit $N \to \infty$, $\alpha(N) \to 1$. Thus, from the proposition above, it follows that zero is also an eigenvalue of the larger matrix $\boldsymbol{A}^{(1)+(2)}_{(2N-1)\times(2N-1)}(k_1, -k_1)$. From numerical investigations, $\boldsymbol{A}^{(2)}_{(N-1)\times(N-1)}$ does not appear have a zero eigenvalue and is assumed to be invertible. Using the inverse, we can find the zero eigenvector of the matrix $\boldsymbol{A}^{(1)+(2)}_{(2N-1)\times(2N-1)}(k_1, -k_1)$:

$$\boldsymbol{c}^{(1)+(2)}(k_1, k_2; t \to \infty) = \begin{bmatrix} \boldsymbol{\Pi}_N \\ -\left(\boldsymbol{A}^{(2)}_{(N-1)\times(N-1)}\right)^{-1} \boldsymbol{B}_{(N-1)\times N} \boldsymbol{\Pi}_N \end{bmatrix} \delta_{k_1,-k_2}. \tag{52}$$

The steady-state 2-marginal distribution $p_n^{(2)}(x_1, x_2; t \to \infty)$ can then be obtained by taking the inverse Fourier transform of the expression above. This has to be done using numerical integration, however, for the special case of $R = 0$, a closed-form expression for $p_n^{(2)}(x_1, x_2; t \to \infty)$ can be obtained:

$$p_n^{(2)}(x_1, x_2) = \frac{\pi_n}{2L^2} \frac{\omega}{\sinh(\frac{\omega}{2})} \cosh\left(\frac{\omega|x_1 - x_2|}{L} - \frac{\omega}{2}\right), \tag{53}$$

where $\omega = \sqrt{L^2 r_a/D}$. For $R = 0$, the 2-marginal entropy production rate can also be computed and is found to be divergent, $\dot{S}^{(2)}_{\rm irr} = \infty$ (see Appendix B). In contrast, the 1-marginal entropy production rate vanishes for $R = 0$, i.e., $\dot{S}^{(1)}_{\rm irr} = 0$ [see Eq. (39)]. This highlights that the rate of entropy production strongly depends on the number of degrees of freedom being tracked, as expected.

A more useful quantity is the two-body correlation function as defined in Eq. (10):

$$\rho^{(2)}(x_1, x_2) = \sum_{n=2}^{\infty} (n^2 - n) p_n^{(2)}(x_1, x_2). \tag{54}$$

Translation and inversion symmetries imply that $\rho^{(2)}(x_1, x_2)$ is solely a function of $|x_2 - x_1|$. Physically, $L^2\rho^{(2)}(|x_1 - x_2|)/\langle n\rangle^2_{\rm ss}$ is the probability that we find another particle at a distance $|x_2 - x_1|$ relative to a given particle at x_1. Figure 3 shows the two-body correlation function as a function of the pairwise distance $|x_1 - x_2|$. Interestingly, the correlation function displays a peak at $|x_1 - x_2| = 2R$, which corresponds to the separation between two daughter particles immediately following a division event. This peak indicates the most probable distance at which one is likely to find another particle relative to a given reference particle.

E. Brownian dynamics simulations

To verify our results, we perform Brownian dynamics simulations of the DDD model, where the diffusing particles at positions $\{x\}_{n(t)}$ perform simple diffusion according to the

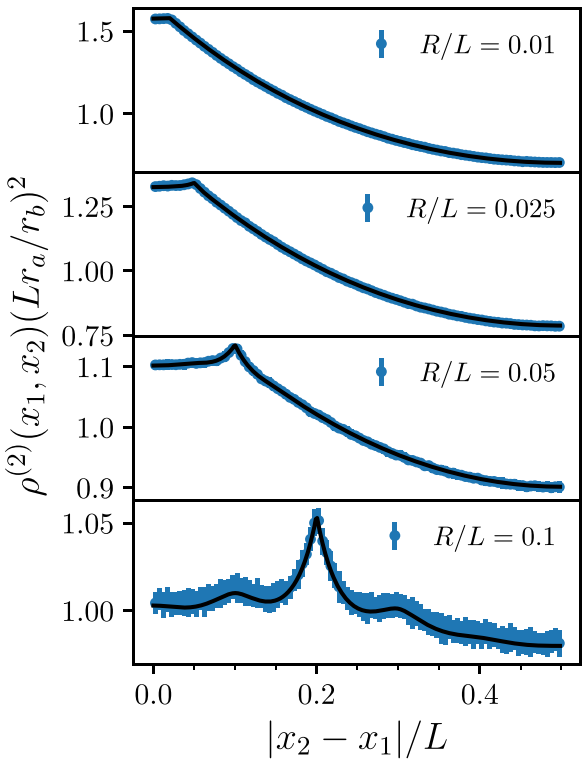


FIG. 3. Two-body correlation function of the division, death, and diffusion (DDD) model in one dimension under periodic boundary conditions (period length L) with division separation $2R$. Markers are measured directly from simulation in LAMMPS [43,44] (error bars are standard error of the mean). Solid lines are analytical results from the 2-marginal probability distributions. Other parameters are fixed at $L = 1$, $r_b/r_a = 100$, and $D/(L^2 r_a) = 0.1$.

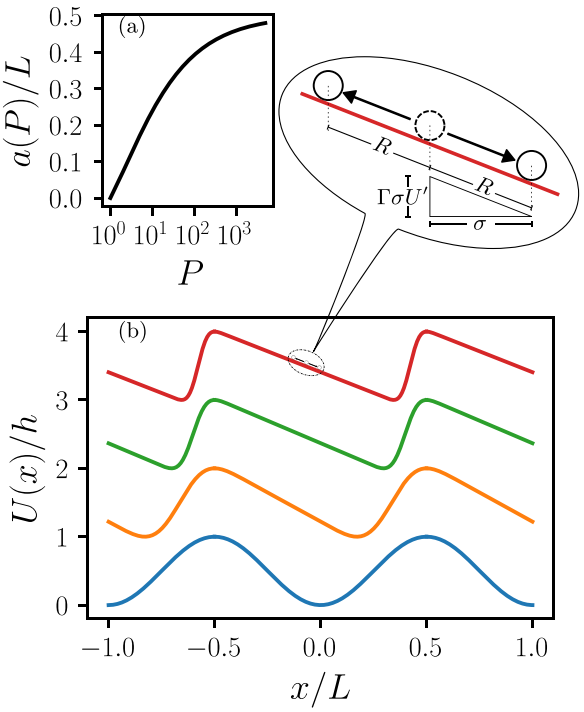


FIG. 4. Ratchet potential applied to the DDD process. In (b) the applied, L-periodic ratchet potential, $U(x)$ in Eq. (66), is shown for integer values P which dictate the location of the minima in x. From bottom to top, $P = 1$ (blue), $P = 5$ (orange), $P = 20$ (green), and $P = 40$ (red). A constant offset is added to better visualize the different curves shown. The location of the minimum in $x \in [0, L/2]$, $a(P)$ in Eq. (67), is shown in (a). The inset of (b) illustrates the geometric projection of a division event onto the local slope of the potential $U(x)$ with $\sigma = R/\sqrt{1 + [\Gamma U'(x)]^2}$ as the projected separation radius; see Eq. (57).

stochastic differential equations

$$\frac{dx_i}{dt} = \sqrt{2D}\xi_i(t), \quad i = 1, \ldots, n(t), \tag{55}$$

where $n(t)$ is the total number of particles at time t. $\xi_i(t)$ are white-noise with zero mean and variance $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{ij}\delta(t - t')$. The diffusion dynamics is coupled to division and death dynamics according to the rules described in Fig. 2. For all simulations, we initialize the system with a single particle at the origin and allow it to evolve towards a steady state. We verify that the system has reached steady state by monitoring the number of particles $n(t)$ and confirming that it fluctuates around its steady-state average value, $n(t) \sim \langle n \rangle_{\rm ss}$. To compute the correlation function, we construct a histogram of the distances between all particle pairs. This procedure is repeated over multiple simulations with independent noise realizations to obtain the ensemble average. The simulations results, shown in Fig. 3 as solid markers, demonstrate excellent agreement with the analytical results (solid lines) from the previous section.

IV. A MINIMAL MODEL OF PARTICLE CURRENT VIA DIVISION, DEATH, AND DIFFUSION

As a final application of our formalism, we examine the DDD model within an asymmetric ratchet potential $U(x)$, as illustrated in Fig. 4. This setup serves as a minimal model of nonmotile cells confined in a periodic, asymmetric channel, where they undergo division, death, and diffusion. Previous simulations of dividing particles in a two-dimensional ratcheted channel have demonstrated the emergence of a nonzero macroscopic particle current, driven by steric interactions among the particles and with the channel walls [45]. To capture the essential physics of this process, we consider a one-dimensional system of noninteracting particles subject to a periodic ratchet potential $U(x)$ [46]. This is achieved by incorporating a drift term into the current operator $\hat{j}_{x_i}$ in Eqs. (25) and (26):

$$\hat{j}_{x_i} = -D\partial_{x_i} - \frac{1}{\gamma}U(x_i), \tag{56}$$

where $D > 0$ is the diffusion constant and $\gamma > 0$ is the friction coefficient. If fluctuation-dissipation holds, we have the relationship $D = k_B T/\gamma$, where T represents the (effective) temperature of the surrounding heat bath; however, this is not always true in an active system. Additionally, we extend our previous DDD model to account for a geometric projection of particle division along the slope of the potential $U(x)$; see Fig. 4(b). This modification prevents particles from "tunneling" through the potential barrier during division. Consequently, when a particle divides at position x, the two

daughter particles are positioned at

$$f_{\pm R}(x) = x \pm \frac{R}{\sqrt{1+[\Gamma U'(x)]^2}}, \tag{57}$$

where R is now the maximum splitting radius (in contrast to the Sec. III where particles were always shifted by $\pm R$ relative to the parent particle). The phenomenological parameter Γ couples the projected splitting radius $\sigma = R/\sqrt{1+[\Gamma U'(x)]^2}$ to the ratchet potential $U(x)$. The splitting radius is given by R only where the force due to the potential is zero $U'(x) = 0$ or if $\Gamma = 0$. Otherwise, the model is given by Eq. (25) with the new definition of $f_{\pm R}(x)$ given in Eq. (57) and the additional drift term in the current operator in Eq. (56).

Our aim is to determine the average density $\rho(x,t)$, which can be obtained from the set of 1-marginal probability distributions $\rho(x,t) = \sum_{n=1}^{\infty} n p_n^{(1)}(x,t)$; see Eq. (10). The constraints of Eqs. (12) and (13) are still satisfied in spite of the addition of a nonzero external field and the new splitting rule in Eq. (57). Therefore, one may multiply Eq. (11) by n and sum over all integers $n \geqslant 1$ to determine $\rho(x,t)$. The resulting dynamics of $\rho(x,t)$ depends explicitly on the 1-marginal probability distributions $p_n^{(1)}(x,t)$. To close the equation for $\rho(x,t)$, we make the approximation $\langle n\rangle/\langle n^2\rangle \approx \sum_m m p_m^{(1)}(x)/\sum_n n^2 p_n^{(1)}(x)$, i.e., the ratio of average particle number to its square scales the same locally as it does globally (in steady state). This approximation is explained Appendix C and justified in Appendix D for a smooth ratchet potential. The steady-state average density is then given by the differential equation

$$\partial_x J[\rho(x), U(x)] = r_b S[\rho(x), U(x)], \tag{58}$$

where the divergence of the current

$$J[\rho, U] = -\frac{1}{\gamma} U'(x)\rho(x) - D\rho'(x) \tag{59}$$

balances a source or sink term

$$S[\rho, U] \equiv \left[\sum_{\sigma\in\{-R,R\}} \int dy\, \delta[f_\sigma(y) - x]\rho(y)\right] - 2\rho(x) \tag{60}$$

driven by division and death. Note that a density equation similar to Eqs. (58)–(60) is usually the starting point of dynamic density functional theory [47]. In this paper we derive the time evolution of the density (and two-body correlation function) from the full Fokker-Planck equation through a hierarchy of marginal probability distributions. As we show in Appendixes C and D, if one requires that the maximum splitting radius is much smaller than the channel period, $R/L \ll 1$, the differential equation may be recast as

$$0 = \partial_x \left\{ \left[\frac{1}{\gamma} \partial_x \tilde{U}(x) + \tilde{D}(x)\partial_x \right] \rho(x) \right\} + O\left(\frac{R^3}{L^3}\right) \tag{61}$$

with the definitions of an effective potential and diffusion coefficient,

$$\tilde{U}(x) \equiv U(x) + \frac{\gamma r_b R^2}{1+[\Gamma U'(x)]^2}, \tag{62}$$

$$\tilde{D}(x) \equiv D + \frac{r_b R^2}{1+[\Gamma U'(x)]^2}, \tag{63}$$

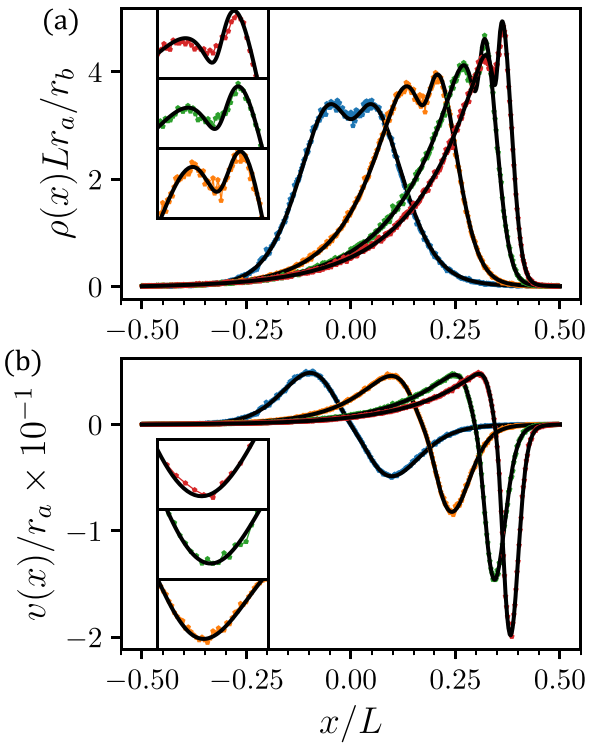


FIG. 5. (a) Average density $\rho(x)$ and (b) average local velocity (per particle per length) $v(x)$ of the DDD process subject to an external field and geometric projection of particle division in steady state. Our analytical results are represented by black solid lines, while direct simulations performed using LAMMPS [43,44] are shown as solid markers. From left to right in terms of the maxima of the curves, the externally applied ratchet potential is asymmetrically skewed such that its minimum is at $x/L = 0$ (blue), $x/L = 0.17$ (orange), $x/L = 0.30$ (green), and $x/L = 0.34$ (red). Other parameters are fixed at $R/L = 0.015$, $r_b/r_a = 400$, $\Gamma D\gamma/L = 0.15$, $D/(L^2 r_a) = 0.15$, and $D\gamma/h = 0.15$. Insets highlight differences between the analytics and simulation at their most prominent regions of disagreement.

respectively. The solution of Eq. (61) can be solved using quadrature [48] with the normalization requirement

$$\langle n\rangle = \int_{-L/2}^{L/2} dx\, \rho(x) \tag{64}$$

and periodic boundary conditions $\rho(x+L) = \rho(x)$.

As mentioned in the previous paragraph, Eq. (61) assumes that average particle number and its square scale in a spatially invariant way, and such assumption is not obvious *a priori*. In Fig. 5 we compute $\rho(x)$ using Eq. (61) and compare it to the average (steady-state) density measured in Brownian dynamics simulations, where particle $x_i(t)$ obeys the stochastic differential equation

$$\frac{dx_i}{dt} = -\frac{1}{\gamma} U'(x_i) + \sqrt{2D}\xi_i(t), \quad i = 1, \ldots, n(t), \tag{65}$$

where $n(t)$ is the number of particles, which changes due to division and death events. Each division event shifts the daughter particles by a separation radius, which is geometrically projected onto the local slope of $U(x)$, according to

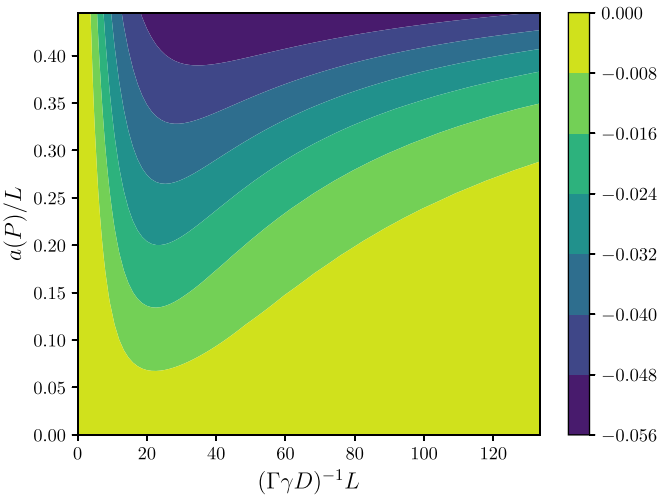


FIG. 6. Global velocity per particle (in units of Lr_a) for the geometrically projected DDD process. The ratchet potential asymmetry $a(P)/L$ is varied by changing the integer $P \in [1, 40]$, and inverse dimensionless geometrical coupling parameter $(\Gamma D\gamma)^{-1}L$ is changed while holding all other parameters fixed at $R/L = 0.015$, $r_b/r_a = 400$, $D/(L^2 r_a) = 0.15$, and $D\gamma/h = 0.15$.

Eq. (57). For the potential $U(x)$ itself, we choose a smooth ratchet potential

$$U(x) = \frac{h}{2}\left(\frac{\sum_{k=1}^{P}\binom{2P}{P-k}\sin(2\pi kx/L)/k}{\sum_{k=1}^{P}\binom{2P}{P-k}\sin[k\pi(1-a(P)/L)]/k} + 1\right), \tag{66}$$

which is differentiable everywhere and periodic with period L. h is the energy scale of the potential and $\binom{\cdots}{\cdots}$ is the binomial coefficient. The minimum of this potential is located at

$$a(P) = \frac{L}{2} - \frac{2L}{\pi}\cos^{-1}\left[\frac{1}{2}\binom{2P}{P}^{\frac{1}{2P}}\right], \tag{67}$$

which is monotonically increasing in P (see Fig. 4). The integer $P \in \mathbb{Z}^+$ in Eq. (66) controls the asymmetry of the ratchet with $a(P = 1) = 0$ (completely symmetric) and $a(P \to \infty) \to L$. In Fig. 5(a) we find good agreement between the calculated marginal density and Brownian dynamics simulations. Similarly, in Fig. 5(b) our calculation of the steady-state average local velocity per particle per length, $v(x) \equiv -\gamma^{-1}\rho(x)U'(x)/\langle n\rangle$, is in agreement with simulation.

Given the excellent agreement between the calculated $\rho(x)$ and $v(x)$ with simulation, we compute the average global velocity of this geometrically projected DDD process by integrating $v(x)$ over the period L, enabling us to scan parameter space much more efficiently than using direct simulation. In Fig. 6 we show that one does indeed get a macroscopic particle current with our minimal model of cells in a ratchet [45] due to the broken symmetry of the geometrically projected splitting radius. Of course, to quantitatively match our model to previous simulations and/or experiments, one would need to include, e.g., particle interactions, two-dimensional correlations due to division, and explicit wall interactions, which are beyond the scope of this paper.

V. DISCUSSION AND CONCLUSIONS

In this work we have presented a formalism which suitably describes systems with a state space that has both discrete and continuous degrees of freedom. In the context of existing literature, textbooks allude to the possibility of extending formalism of, e.g., the Fokker-Planck equation or the master equation to hybrid state space, but in practice this is not done explicitly [13,48,49]. Additionally, in this paper we specify that the number of continuous degrees of freedom is itself the discrete degree of freedom, i.e., having two continuous degrees of freedom x_1 and x_2 implies the state $\{2, x_1, x_2\}$.

We have studied a toy model of division, death, and diffusion (DDD) to illustrate a simple application of our formalism. In the absence of an external potential, the set of probability distribution functions $\{p_n(\{x\}_n, t)\}$ can be solved recursively using a hierarchy similar to the BBGKY hierarchy of deterministic (Hamiltonian) mechanics, which we solve analytically up to second order (two-body probability distributions). The resulting two-body correlation function demonstrates excellent agreement with numerical results obtained from direct Brownian dynamics simulations.

Additionally, we also computed the 1-marginal and the 2-marginal entropy production rate, revealing a strong dependence on the number of degrees of freedom being tracked. For example when $R = 0$, the 1-marginal entropy production rate vanishes, whereas the 2-marginal entropy production rate is infinite. The entropy production rate is also sensitive to the nature of the division separation—whether it is random or fixed at $\pm R$ relative to the original position. We demonstrate this by treating the delta function as the singular limit of a Gaussian distribution in our calculation of the one-marginal entropy production rate.

In the presence of an external potential $U(x)$, we derive the effective equation of state for the average particle density $\rho(x)$. In the minimal model, an asymmetric ratchet potential combined with geometric projection is sufficient to generate a macroscopic current. Despite its simplicity, this coupling produces a nonzero particle flow in qualitative agreement with particle-based simulations.

Future extensions of the model, or applications of the formalism presented here—particularly those motivated by experiments on confluent cell layers—will almost certainly require incorporating particle interactions. This would necessitate a closure relation analogous to the molecular chaos hypothesis in kinetic theory to enable further analytical progress. Another promising extension involves introducing an absorbing state transition into the vacuum state ($n = 0$), enabling the study of directed percolation and related models [50,51].

Overall, the framework developed in this work provides a theoretical foundation for constructing more complex models that integrate division, death, and diffusion, bridging microscopic dynamics with hydrodynamic descriptions in terms of reaction-diffusion systems.

ACKNOWLEDGMENTS

We acknowledge support from EPSRC Grant No. EP/W027194/1. E.T. would like to thank Daniel Sussman,

Rob Jack, and Mike Cates for useful discussions. E.T. would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the SPL program, funded by EPSRC Grant No. EP/R014604/1.

DATA AVAILABILITY

Simulation software is available through a custom package which is run under the LAMMPS [43] software interface. This package is available publically [44] under the GitHub branch "birthdeath_package."

APPENDIX A: ENTROPY PRODUCTION IN A HYBRID-STATE SYSTEM

In this appendix, we shall use a short-hand notation to simplify the presentation:

$$p(n_j) = p_{n_j}(\{x(t_j)\}_{n_j}, t_j), \quad G(n_j \to n_k) = G_{n_j n_k}(\{x(t_j)\}_{n_j} \to \{x(t_k)\}_{n_k}; \tau), \quad \int dn_j = \sum_{n_j=1}^{\infty} \int d\{x(t_j)\}_{n_j}. \tag{A1}$$

Substituting Eqs. (4) and (5) into (6), we get

$$\Delta \mathcal{S}_M = \underbrace{\left(\int dn_0 \int dn_1 \cdots \int dn_M\right)}_{\sum_{\boldsymbol{x}_M}} \underbrace{p(n_0) \prod_{k=0}^{M-1} G(n_k \to n_{k+1})}_{\mathcal{P}(\boldsymbol{X}_M; t_0, \Delta t)} \ln\left(\frac{p(n_0)\prod_{j=0}^{M-1} G(n_j \to n_{j+1})}{p(n_M)\prod_{j=0}^{M-1} G(n_{j+1} \to n_j)}\right) \tag{A2}$$

$$= \sum_{j=0}^{M-1} \left(\int dn_0 \int dn_1 \cdots \int dn_M\right) p(n_0) \prod_{k=0}^{M-1} G(n_k \to n_{k+1}) \ln\left(\frac{p(n_j)G(n_j \to n_{j+1})}{p(n_{j+1})G(n_{j+1} \to n_j)}\right) \tag{A3}$$

$$= \sum_{j=0}^{M-1} \int dn_0 \int dn_j \int dn_{j+1} \int dn_M \, p(n_0)G(n_0 \to n_j)G(n_j \to n_{j+1})G(n_{j+1} \to n_M) \ln\left(\frac{p(n_j)G(n_j \to n_{j+1})}{p(n_{j+1})G(n_{j+1} \to n_j)}\right), \tag{A4}$$

where we have used the Chapman-Kolmogorov relation to get the last line in the equation above:

$$G(n_0 \to n_j) = \left(\int dn_1 \cdots \int dn_{j-1}\right) \prod_{k=0}^{j-1} G(n_k \to n_{k+1}). \tag{A5}$$

Finally we use the Markov property in the last line of Eq. (A4):

$$\int dn_0 \, p(n_0)G(n_0 \to n_j) = p(n_j), \tag{A6}$$

$$\int dn_j \, G(n_0 \to n_j) = 1 \tag{A7}$$

so that Eq. (A4) simplifies to

$$\Delta \mathcal{S}_M = \sum_{j=0}^{M-1} \int dn_j \int dn_{j+1} \, p(n_j)G(n_j \to n_{j+1}) \ln\left(\frac{p(n_j)G(n_j \to n_{j+1})}{p(n_{j+1})G(n_{j+1} \to n_j)}\right). \tag{A8}$$

Finally, we take the limit $M \to \infty$ and $\tau \to 0$, the summation $\sum_{j=0}^{M-1} \tau$ then becomes a Riemann integral $\int_{t_0}^{t_0+\Delta t} dt$. We therefore identify the integrand to be the rate of entropy production [28,29] at time $t_j = t_0 + j\tau$:

$$\dot{S}_{\rm irr}(t_j) = \lim_{\tau \to \infty} \frac{1}{\tau} \sum_{n_j=1}^{\infty} \sum_{n_{j+1}=1}^{\infty} \int d\{x\}_{n_j} \int d\{y\}_{n_{j+1}} p_{n_j}(\{x\}_{n_j}, t_j) G_{n_j n_{j+1}}(\{x\}_{n_j} \to \{y\}_{n_{j+1}}; \tau)$$

$$\times \ln\left(\frac{p_{n_j}(\{x\}_{n_j}, t_j) G_{n_j n_{j+1}}(\{x\}_{n_j} \to \{y\}_{n_{j+1}}; \tau)}{p_{n_{j+1}}(\{y\}_{n_{j+1}}, t_{j+1}) G_{n_{j+1} n_j}(\{y\}_{n_{j+1}} \to \{x\}_{n_j}; \tau)}\right). \tag{A9}$$

After relabeling $t_j \to t$, $t_{j+1} \to t$, $\{x\}_{n_j} \to \{x\}_m$ and $\{y\}_{n_{j+1}} \to \{y\}_n$, we finally get Eq. (7).

APPENDIX B: 2-MARGINAL ENTROPY PRODUCTION RATE

In the case of $R = 0$, the steady-state 2-marginal probability distributions have the simple analytical form of Eq. (53). We now show that the 2-marginal entropy production rate is infinite even when $R = 0$. Just as in the 1-marginal case, only the off-diagonal terms in the summation of the 2-marginal entropy production [Eq. (17)] are nonzero. Therefore, to show that $\dot{S}_i^{(2)} = \infty$ for $R = 0$, one just needs to compute the first off-diagonal contribution to the marginal entropy production rate ($1 \leftrightarrow 2$ particles) and prove that it is infinite. Using Eq. (53), this first term is

$$\begin{aligned}
&\lim_{\tau\to 0}\frac{1}{\tau}\int dy_1 dx_1 dx_2\, G(\{y\}_1 \to \{x\}_2;\tau)p_1(\{y\}_1)G(\{x\}_2 \to \{y\}_1;\tau)p_2(\{x\}_2)\ln\left[\frac{G(\{y\}_1 \to \{x\}_2;\tau)p_1(\{y\}_1)}{G(\{x\}_2 \to \{y\}_1;\tau)p_2(\{x\}_2)}\right]\\
&= \int dy_1 dx_1 dx_2\, \ln\left[\frac{4L\sinh(\omega/2)\delta(x_1-y_1)\delta(x_2-y_1)}{\omega\cosh(\omega|x_2-x_1|/L-\omega/2)[\delta(x_2-y_1)+\delta(x_1-y_1)]}\right]\\
&\quad\times\left\{\frac{r_b\pi_1}{L}\delta(x_1-y_1)\delta(x_2-y_1) - \frac{r_a\pi_2}{L^2}\frac{\omega}{\sinh(\omega/2)}\cosh\left(\frac{\omega|x_2-x_1|}{L}-\frac{\omega}{2}\right)[\delta(x_1-y_1)+\delta(x_2-y_1)]\right\},
\end{aligned} \tag{B1}$$

where $\omega = \sqrt{L^2 r_a/D}$. As done for the 1-marginal case, if one approximates the delta functions in Eq. (B1) as infinitesimal width Gaussians, we find that the 2-marginal entropy production rate is infinite even when $R = 0$. Notably, if one takes the infinite diffusion limit $D \to \infty$, the entropy production rate remains infinite even though the 2-marginal distribution $p_n^{(2)}(x_1, x_2)$ becomes spatially homogeneous. Therefore, inclusion of additional information on spatial locations of particles in the DDD process changes the entropy production rate even when the system is spatially homogeneous, i.e., the entropy production is not simply given by Eq. (24).

APPENDIX C: DDD GEOMETRICALLY COUPLED TO $U(x)$

In this appendix we derive Eq. (61) with the effective potential and diffusion coefficients given in Eqs. (62) and (63). We start with the DDD dynamics defined in Eq. (25) but with the general current operator $\hat{j}_{x_i}$ given in Eq. (56) and general splitting rule $f_{\pm R}(x)$ given in Eq. (57). Integrating Eq. (25) over all but x_1, we find that the dynamics of 1-marginal distribution functions obey the set of equations

$$\begin{aligned}
&\partial_t p_n^{(1)}(x,t) + \partial_x\left[\hat{j}_x p_n^{(1)}(x,t)\right]\\
&\quad= r_a(\mathbb{E}-1)n(n-1)p_n^{(1)}(x,t) + (\mathbb{E}^{-1}-1)np_n^{(1)}(x,t)\\
&\qquad+\frac{r_b(n-1)}{n}\int dy[\delta[f_{+R}(y)-x]+\delta[f_{-R}(y)-x]\\
&\qquad-2\delta(y-x)]p_{n-1}^{(1)}(y,t),
\end{aligned} \tag{C1}$$

where we have defined the step operator $\mathbb{E}[f(n)] = f(n+1)$ and its inverse $\mathbb{E}^{-1}[f(n)] = f(n-1)$. The terms which shift the particle by Eq. (57) can be integrated,

$$\begin{aligned}
&\int dy\,\delta[f_{\pm R}(y)-x]p_{n-1}^{(1)}(y,t)\\
&\quad=\sum_{\alpha=1}^{S}\frac{p_{n-1}^{(1)}\left(f_{\pm R,\alpha}^{-1}(x),t\right)}{1\pm RQ'\left[f_{\pm R,\alpha}^{-1}(x)\right]}\\
&\quad=\sum_{\beta=0}^{\infty}\frac{(\mp R)^\beta}{\beta!}(\partial_x)^\beta\left[p_{n-1}^{(1)}(x,t)Q^\beta(x)\right],
\end{aligned} \tag{C2}$$

where we have defined

$$Q(x) = \sqrt{\frac{1}{1+[\Gamma U'(x)]^2}} \tag{C3}$$

for convenience; cf. Eq. (57). The second line in Eq. (C2) is obtained by a substitution of variables to $v = f_{\pm R}(y)$ and then direct integration over v. $f_{\pm R,\alpha}^{-1}(x)$ are the S inverses of Eq. (57) over the period $x \in [-L/2, L/2]$, i.e., $f_{\pm R,\alpha}^{-1}(x) \in \{y : x = f_{\pm R}(y),\ x \in [-L/2, L/2]\}$. [Note that $f_{\pm R,\alpha}^{-1}(x)$ is, in general, a multivalued function.] The last line in Eq. (C2) follows from formally Taylor expanding $\delta(f_{\pm R}(y)-x)$ about small R. Both forms are equivalent. We will use the direct substitution form to motivate what parameter regimes can be reasonably approximated by the second-order differential equation (61) in the main text. The Taylor expansion form will be used otherwise as it is simpler to use in algebraic manipulations. In particular, it highlights that when $R = 0$ and fluctuation-dissipation theorem $D = k_BT/\gamma$ holds, the steady-state solution for the 1-marginal is $p_n^{(1)}(x) = \pi_n e^{-U(x)/k_BT}/Z$ where Z is a normalization factor, i.e., when the splitting radius is zero the population dynamics decouple from the spatial dynamics in steady state.

Multiplying Eq. (C1) by n and then summing over all n, the average density obeys the differential equation

$$\begin{aligned}
&\partial_t\rho(x,t) + \partial_x[\hat{j}_x\rho(x,t)]\\
&\quad=(r_a+r_b)\rho(x,t) - r_a\sum_{n=1}^{\infty}n^2p_n^{(1)}(x,t) - 2r_b\rho(x,t)\\
&\qquad+r_b\sum_{\sigma\in\{-R,R\}}\int dy\,\delta[f_\sigma(y)-x]\rho(y,t).
\end{aligned} \tag{C4}$$

Therefore, $\rho(x,t)$ depends explicitly on the second moment of particle number. One either needs to solve for $p_n^{(1)}(x,t)$ first and then sum over n to determine $\rho(x,t)$, or introduce an approximation to close Eq. (C4). We opt for the latter approach, taking

$$\sum_{n=1}^{\infty}n^2p_n^{(1)}(x,t) \approx \frac{\langle n^2(t)\rangle}{\langle n(t)\rangle}\rho(x,t) + O(R^2) \tag{C5}$$

to close Eq. (C4), where $\langle n(t)\rangle$ and $\langle n^2(t)\rangle$ are the position independent averages of particle number and its square at time

t. Eq. (C5) is equivalent to the assumption that $p_n^{(1)}(x,t)$ is separable in n and x to first order in R^2. The error introduced by this approximation is multiplied by r_a in Eq. (C4), so we expect our results to be valid so long as $r_b > r_a$ such that the final term in Eq. (C4) still dominates the $O(R^2)$ correction. Since we choose $r_b \gg r_a$ in this paper, we expect this approximation to introduce negligible error in our calculations of $\rho(x)$. We will directly verify in Appendix D that this approximation is reasonable. Finally, by using Eq. (C5), cancellation occurs between the two terms in the second line of Eq. (C4). The closed equation for $\rho(x,t)$ is

$$\begin{aligned}&\partial_t \rho(x,t) + \partial_x[\hat{j}_x \rho(x,t)] \\ &= r_b\left(\sum_{\alpha=1}^{S} \frac{\rho\left(f_{+R,\alpha}^{-1}(x),t\right)}{1+RQ'\left[f_{+R,\alpha}^{-1}(x)\right]}\right. \\ &\quad \left. + \sum_{\alpha=1}^{S} \frac{\rho\left(f_{-R,\alpha}^{-1}(x),t\right)}{1-RQ'\left[f_{-R,\alpha}^{-1}(x)\right]} - 2\rho(x,t)\right). \end{aligned} \tag{C6}$$

The terms on the right-hand side of this equation represent a source and a sink term due to division and death, given by $S[\rho,U]$ in Eq. (60). By expanding this source or sink term to $O(R^2)$ using the second Taylor expansion approximation of the splitting events, [cf. Eq. (C2)], we recover Eq. (61) with the effective potential and diffusion coefficients given in Eqs. (62) and (63).

APPENDIX D: SMOOTH RATCHET POTENTIAL

In Appendix C we introduced two approximations which allowed us to derive Eq. (61) for general potential $U(x)$. Here we justify these approximations specifically for the ratchet potential used in the main text [Eq. (66)].

The first approximation made is to assume Eq. (C5) is valid. We check in Fig. 7 that Eq. (C5) is reasonable by directly simulating particle trajectories via Eq. (65) (coupled to the geometrically projected division and death), and measuring $\rho(x,t)$, $\sum_n n^2 p_n^{(1)}(x,t)$, $\langle n(t)\rangle$, and $\langle n^2(t)\rangle$. We find excellent overlap between the left-hand side (black lines) and right-hand side (dots) of Eq. (C5). As discussed in Appendix C, when r_a approaches r_b and both are large, Eq. (C5) should begin to fail. We find good agreement when we take $r_a = 10$ and $r_b = 40$ (green dots and overlapping black line in Fig. 7). Therefore, Eq. (C5) is reasonable for all parameter regimes studied in this paper when $U(x)$ is given by Eq. (66).

The second assumption made in deriving Eq. (61) is that $R/L \ll 1$ such that the Taylor series expansion in Eq. (C2) is sufficiently approximated by the first correction $O(R^2)$, or equivalently that there is a unique inverse of $f_{\pm R}(y)$ [so $S = 1$ in Eq. (C2)] and the summations in Eq. (C1) reduces to a single term. This depends strongly on the choice of potential $U(x)$ being used. We verify here the regions of parameter space which affect the Taylor expansion validity and unique invertibility of $f_{\pm R}(y)$ for the ratchet potential [Eq. (66)]. Since $f_{\pm R}(y)$ depends only on the dimensionless parameters R/L, $\Gamma h/L$, and the asymmetry number P [see Fig. 4(a)], these three parameters completely dictate the regions of validity for Eq. (61).

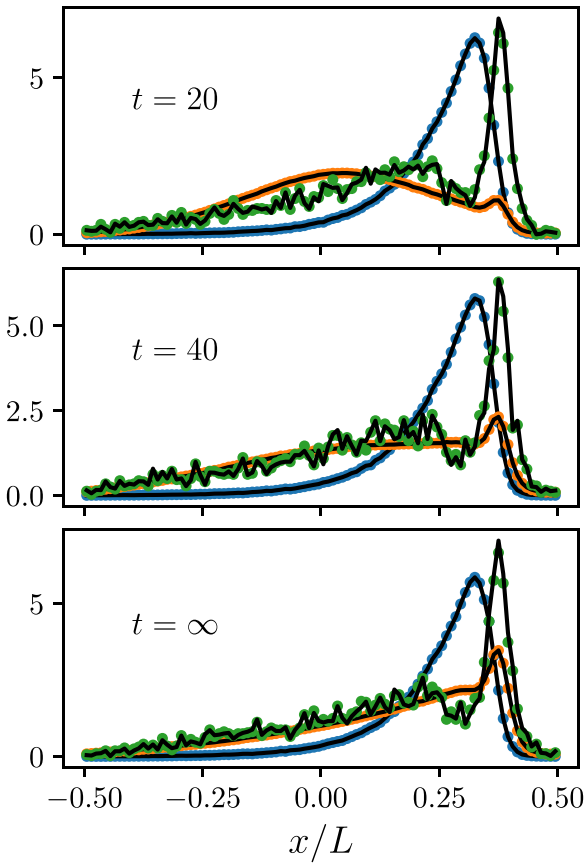


FIG. 7. Comparison of $\langle n^2\rangle \rho(x,t)/\langle n\rangle^3$ (dots) and $\sum_n n^2 p_n^{(1)}(x,t)/\langle n\rangle^2$ (black lines). Both quantities are measured directly from molecular dynamics simulations coupled to division and death events as discussed in the main text. Different colors correspond to different values of the parameter three-tuple (r_a, r_b, L): $(10^{-3}, 0.4, 10)$ (blue), $(10^{-3}, 4, 66)$ (orange), and $(10, 40, 10)$ (green). Other parameters are set at $P = 30$ (ratchet potential asymmetry corresponding to $a(P)/L \approx 0.3$), $D = 1$, $R = 1$, $\Gamma = 2$, $h = 7.9$, and $\gamma = 1$.

The inverse of $f_{\pm R}(y)$ is unique if it is one-to-one, i.e., there is no maximum. In Fig. 8 we show a map of the region where at least one maximum is present (absent) in light red (gray) when the ratchet asymmetry parameter $P = 100$. Therefore, the gray regions of Fig. 8 should permit one to use Eq. (61) to compute the average density of the DDD process with Eq. (66). Additionally, we note that the approximation to $O(R^2)$ in Eq. (61) requires Taylor expansion of $Q(f_{\pm R}^{-1}(x))$ (near x) as defined in Eq. (C3). If we define $y = f_\sigma^{-1}(x)$, where $\sigma \in \{-R, R\}$, then we have [where $Q = Q(x)$ when no argument is stated explicitly]

$$Q(y) \approx Q - \sigma QQ' + \sigma^2\left[Q(Q')^2 + \frac{1}{2!}Q^2Q''\right] + O(\sigma^3). \tag{D1}$$

When the first and second terms are of the same order, we expect the above Taylor expansion to fail. This gives the criterion $R < \frac{1}{\max(Q')}$. We show this criterion in Fig. 8 as a black solid line, which agrees with the previous invertibility condition via the presence of a maxima.

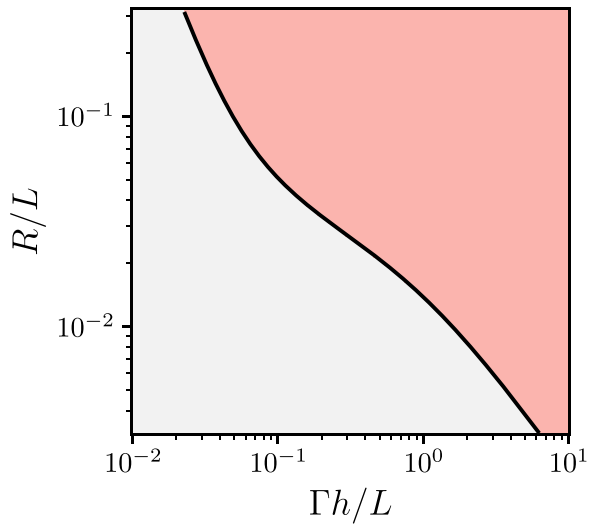


FIG. 8. Region in which Eq. (61) is valid (gray shading) and when higher order terms in R must be considered (red shading) for the asymmetry parameter $P = 100$. The black solid line is the criterion $R = 1/\max(Q')$. The three parameters R/L, $\Gamma h/L$, and P determine whether the expansion of Eq. (C6) to second order in splitting radius R is sufficient.

Above, we have confirmed that the region of parameter space in which the truncation of Eq. (C2) at $O(R^2)$ is valid when $P = 100$. We have additionally checked that the region of validity always decreases with increasing asymmetry (P) in the ratchet. Finally, we note that the requirement of single-valued inversion of Eq. (57), or equivalently the validity of Taylor expanding to lowest order in R, means that a piecewise ratchet potential cannot be used when approximating Eq. (58) by Eq. (61) since no unique inverse exists, hence our use of the smooth ratchet potential of Eq. (66).

APPENDIX E: ESTIMATORS OF DISTRIBUTIONS MEASURED IN SIMULATION

For this appendix only, a "hat" over a variable indicates that it is a random variable, not an operator as in the main text. This random variable may be sampled from a time-dependent stochastic process. In this paper, $\{\hat{y}(t)\}_{\hat{m}(t)}$ will always represent the positions of particles generated from Brownian dynamics simulations with division and death determining $\hat{m}(t)$ at each time step t. We omit the time dependence of all random variables for brevity. The expectation of any function $g_{\hat{m}}(\{\hat{y}\}_{\hat{m}})$ of this set of random variables (at a specified time t) is calculated by

$$\langle g_{\hat{m}}(\{\hat{y}\}_{\hat{m}})\rangle \equiv \sum_{\hat{m}=1}^{\infty}\int d\{\hat{y}\}_{\hat{m}} g_{\hat{m}}(\{\hat{y}\}_{\hat{m}}) p_{\hat{m}}(\{\hat{y}\}_{\hat{m}}, t), \tag{E1}$$

and an unbiased estimator of $\langle g_{\hat{m}}(\{\hat{y}\}_{\hat{m}})\rangle$ is

$$\bar{g}_{\hat{m}}(\{\hat{y}\}_{\hat{m}}) = \frac{1}{N_s}\sum_{s=1}^{N_s} g_{\hat{m}_s}(\{\hat{y}_s\}_{\hat{m}_s}), \tag{E2}$$

where the sum is over a set of length N_s containing uncorrelated samples taken from simulation at time t. Typically in this paper, we take $N_s = 10\,000$.

We now state the functions that, when averaged over, are unbiased estimators of distribution functions such as $p_n^{(l)}$ and $\rho^{(l)}$. The estimator, $\bar{p}_n$, of the full distribution for n particles, is taken to be a sum over samples of the function

$$p_{n,\hat{m}}(\{x\}_n, \{\hat{y}\}_{\hat{m}}) = \delta_{n,\hat{m}}\frac{1}{n!}\sum_{\pi\in S_n}\prod_{k=1}^{n}\delta(x_k - \hat{y}_{\pi(k)}), \tag{E3}$$

where the sum is over all $\pi \in S_n$, where S_n is the set of permutations in the symmetric group on n objects (this set is length $n!$). It follows from this equation that

$$p^{(l)}_{n,\hat{m}}(x_1, \{\hat{y}\}_{\hat{m}}) = \frac{\delta_{n,\hat{m}}(n-l)!}{n!}\sum_{i_1=1}^{n}\delta(x_1 - \hat{y}_{i_1}) \times \sum_{i_2=1}^{n}{}'\delta(x_2 - \hat{y}_{i_2})\cdots\sum_{i_l=1}^{n}{}'\delta(x_l - \hat{y}_{i_l}), \tag{E4}$$

where the prime on the summation means excluding all previous indices (so the second sum is over $i_2 \neq i_1$, the third sum is over $i_3 \notin \{i_1, i_2\}$, up to the l sum which is over $i_l \notin \{i_1, i_2, \ldots, i_{l-1}\}$. Similarly, from Eq. (10) we have

$$\rho^{(l)}(\{x\}_l, \{\hat{y}_{\hat{m}}\}) = \sum_{i_1=1}^{\hat{m}}\delta(x_1 - \hat{y}_{i_1})\sum_{i_2=1}^{\hat{m}}{}'\delta(x_2 - \hat{y}_{i_2})\cdots\sum_{i_l=1}^{\hat{m}}{}'\delta(x_l - \hat{y}_{i_l}). \tag{E5}$$

[1] G. Gompper *et al.*, J. Phys.: Condens. Matter **32**, 193001 (2020).
[2] J. R. Howse, R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian, Phys. Rev. Lett. **99**, 048102 (2007).
[3] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Rev. Mod. Phys. **85**, 1143 (2013).
[4] O. Hallatschek *et al.*, Nat. Rev. Phys. **5**, 407 (2023).
[5] M. Delarue *et al.*, Nat. Phys. **12**, 762 (2016).
[6] P. Gniewek, C. F. Schreck, and O. Hallatschek, Phys. Rev. Lett. **122**, 208102 (2019).
[7] E. Tjhung and L. Berthier, Phys. Rev. Res. **2**, 043334 (2020).
[8] H. Fujikawa and M. Matsushita, J. Phys. Soc. Jpn. **58**, 3875 (1989).
[9] A. Barabási and H. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995).
[10] I. Klapper and J. Dockery, SIAM Rev. **52**, 221 (2010).
[11] M. R. Mattei *et al.*, J. Math. Biol. **76**, 945 (2018).
[12] R. J. Allen and B. Waclaw, Rep. Prog. Phys. **82**, 016601 (2019).

[13] N. Van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland Personal Library (Elsevier, North-Holland, Amsterdam, 2011).
[14] R. A. Fisher, Ann. Eugenics **7**, 355 (1937).
[15] A. Kolmogorov, I. Petrovskii, and N. Piscunov, Byul. Moskovskogo Gos. Univ. **1**, 1 (1937).
[16] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[17] M. J. del Razo, S. Winkelmann, R. Klein, and F. Höfling, J. Math. Phys. **64**, 013304 (2023).
[18] R. Garcia-Millan, J. Pausch, B. Walter, and G. Pruessner, Phys. Rev. E **98**, 062107 (2018).
[19] R. Garcia-Millan and G. Pruessner, J. Stat. Mech. (2021) 063203.
[20] U. C. Täuber, *Critical Dynamics: A Field Theory Approach to Equilibrium and Non-Equilibrium Scaling Behavior* (Cambridge University Press, Cambridge, 2014).
[21] M. J. del Razo *et al.*, Lett. Math. Phys. **112**, 49 (2022).
[22] L. Peliti, J. Phys. France **46**, 1469 (1985).
[23] M. Doi, J. Phys. A: Math. Gen. **9**, 1479 (1976).
[24] M. Kardar, *Statistical Physics of Particles* (Cambridge University Press, Cambridge, 2007).
[25] C. A. Vargas-Garcia, M. Soltani, and A. Singh, IEEE Life Sci. Lett. **2**, 47 (2016).
[26] M. Xia and T. Chou, J. Phys. A: Math. Theor. **54**, 385601 (2021).
[27] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[28] P. Gaspard, J. Stat. Phys. **117**, 599 (2004).
[29] L. Cocconi, R. Garcia-Millan, Z. Zhen, B. Buturca, and G. Pruessner, Entropy **22**, 1252 (2020).
[30] T. Speck, Europhys. Lett. **123**, 20007 (2018).
[31] R. D. Astumian and M. Bier, Phys. Rev. Lett. **72**, 1766 (1994).
[32] P. Gaspard and E. Gerritsma, J. Theor. Biol. **247**, 672 (2007).
[33] L. Hong and H. Qian, Phys. Rev. E **104**, 044113 (2021).
[34] E. Boksenbojm and B. Wynants, J. Phys. A: Math. Theor. **42**, 445003 (2009).
[35] L. Cocconi, G. Salbreux, and G. Pruessner, Phys. Rev. E **105**, L042601 (2022).
[36] L. Defaveri, E. Barkai, and D. A. Kessler, Phys. Rev. E **107**, 024122 (2023).
[37] T. Agranov, R. L. Jack, M. E. Cates, and E. Fodor, New J. Phys. **26**, 063006 (2024).
[38] F. Caballero and M. E. Cates, Phys. Rev. Lett. **124**, 240604 (2020).
[39] P. Pietzonka and U. Seifert, J. Phys. A: Math. Theor. **51**, 01LT01 (2018).
[40] See Sec. 6.9 in [13].
[41] A. V. Skorokhod, Theory Probab. Appl. **9**, 445 (1964).
[42] J. R. Silvester, Math. Gaz. **84**, 460 (2000).
[43] A. P. Thompson *et al.*, Comput. Phys. Commun. **271**, 108171 (2022).
[44] S. Cameron (2024), https://github.com/samueljmcameron/lammps.
[45] E. Mitchell and E. Tjhung, Soft Matter **18**, 1082 (2022).
[46] P. S. Burada, P. Hänggi, F. Marchesoni, G. Schmid, and P. Talkner, ChemPhysChem **10**, 45 (2009).
[47] H. L. Michael te Vrugt and R. Wittkowski, Adv. Phys. **69**, 121 (2020).
[48] H. Risken and T. Frank, *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer Series in Synergetics (Springer, Berlin, 2012).
[49] C. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Proceedings in Life Sciences (Springer-Verlag, Berlin, 1985).
[50] K. J. Wiese, Phys. Rev. Lett. **133**, 067103 (2024).
[51] H. Hinrichsen, Adv. Phys. **49**, 815 (2000).