Kinetic theories of state- and generation-dependent cell populations

Mingtao Xia [1] and Tom Chou [2]
[1] *Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, USA*
[2] *Department of Mathematics, UCLA, Los Angeles, Calfornia 90095-1555, USA*

(Received 16 February 2024; revised 20 July 2024; accepted 23 October 2024; published 24 December 2024)

We formulate a general, high-dimensional, partial integrodifferential equation (PIDE) kinetic theory describing the internal state (such as gene expression or protein levels) of cells in a stochastically evolving population. The resolution of our kinetic theory also allows one to track subpopulations associated with each generation. Both intrinsic noise of the cell's internal attribute and randomness in a cell's division times (demographic stochasticity) are fundamental to the development of our model. Using our framework, we are able to marginalize the high-dimensional PIDEs in a number of different ways to derive equations which can be PIDEs themselves) that describe the dynamics of marginalized or "macroscopic" quantities such as structured population densities, moments of generation-dependent cellular states, and moments of the total population. We also show how nonlinear "interaction" terms in lower-dimensional integrodifferential equations can arise from high-dimensional *linear* kinetic models that contain rate parameters of a cell (birth and death rates) that depend on variables associated with other cells, generating couplings in the dynamics. Our analysis provides a general, more complete mathematical framework that resolves the coevolution of cell populations and cell states. The approach may be tailored for studying, e.g., gene expression in developing tissues, or other more general particle systems which exhibit Brownian noise in individual attributes and population-level demographic noise.

DOI: 10.1103/PhysRevE.110.064146

I. INTRODUCTION

Mathematical models have been formulated to describe the evolution of populations according to a number of individual attributes such as age, size, and/or added size since birth. Such structured population models have various applications across diverse fields. For example, deterministic age-structured models that incorporate age-dependent birth and death were first developed by McKendrick [1–4] and have been applied to human populations [5]. Structured population models have also been applied to model cell size control [6,7], cellular division mechanisms [8], and structured cell population models [9,10].

In a proliferating cell population, individual cell growth is interrupted by cell division events that generate daughter cells. Kinetic theory is a natural framework to capture the link between individual cellular growth and division, within a proliferating population of cells. Kinetic theories of simple birth-death. processes that track the chronological age of each cell have been developed [11–14] that establish a rigorous mathematical framework to describe how individual cell aging, growth, and division affect population-level quantities such as population-averaged cell size. The kinetic theory PIDE can be marginalized in different ways and reduce, in different limits, to masterlike equations or structured population-like PIDEs, thus unifying deterministic "moment" equations (the structured population PIDEs) with Markovian birth-death-like models. Stochastic fluctuations in parameters such as the cellular growth rate have also been included [15], but integrating fluctuations of internal variables with random birth-death events (demographic stochasticity) is challenging due to the combinatorial complexity and unwieldiness of the relevant equations.

Besides simple individual-cell dynamical variables such as cell age or cell size, gene (mRNA) or protein expression levels are also measured cellular attributes that are important in cell biology, particularly during development. Since there are many different species of mRNA or proteins, the expression pattern is a vector of fluctuating variables.

Although modern computational and statistical techniques can be used to quantitatively infer single cellular mRNA [16], protein [17,18], or chlorophyll [19] levels from experimental data, mathematical models of how expression levels or cell states evolve is often couched in terms of transport along Waddington or fitness landscapes [20,21]. The value of the landscape may represent an "energy" function that is shaped by different genes, or a proliferation rate that is a function different gene expression rates. However, how populations of cells are represented in such high-dimensional "landscapes" is unclear. Moreover, since cellular division rates and death rates typically depend in depend on internal stochastic cell variables such as gene expression levels [22–24], it is important to model how fluctuating-gene-expression-dependent birth or death rates feature in the evolution of a population along an appropriate landscape.

Kinetic models have the capability of precisely describing both the stochastic dynamics of individual cell states and the stochastic birth-death processes associated with an evolving population. Not only is the coupling between individual cell states and the evolution of the population explicit in a kinetic equation, but potential functions governing intracell state dynamics and proliferation (defining a fitness function) arise naturally in the kinetic framework.

Previously derived kinetic models such as the timer-sizer model for cell populations distributed across size [13,25]

incorporate stochastic differential equations (SDEs) to track the dynamics individual internal cell states such as size or mRNA/protein levels. Marginalization of the kinetic equations results in equations for the correlation functions that explicitly show how individual cell states are linked to key macroscopic quantities of the overall population. However, these kinetic theories could not track lineages or generational subpopulations of cells nor did they incorporate cell death or cell division that may also depend on other stochastic variables associated with the cell.

In this paper, we formally develop a complete kinetic model that tracks continuous-valued, stochastically evolving variables (e.g., gene expression, cellular size, mRNA level, protein level, etc.) and the discrete generation number of each cell. The mathematical framework we use for delineating cell of different attribute values across different generations shares a related structure to one recently used to describe ages across different cell stages [14]. In our problem, noise in gene expression is described by a continuous-time stochastic process while noise in division events is described by a Markov jump process. Our model couples these stochastic processes through an SDE-jump-process hybrid model in which the division and death rates explicitly depend on fluctuating gene expression levels [26,27]. All of these quantities are tracked along different generations. The mathematical framework we use for delineating cell of different attribute values across different generations shares a related structure to one recently used to describe ages across different cell stages [14].

In the next section, we define the kinetic model and show how potentials that govern the intracellular dynamics and the population fitness can be motivated. Since the development of our generation-dependent kinetic equations requires intensive book-keeping and associated notation to resolve the time-dependent attributes of each member of the entire population, many of the steps are detailed in extensive mathematical Appendices. However, eventually, in Sec. III we marginalize our high-dimensional kinetic PIDE to derive a number of more meaningful "reduced" equations that describe the evolution of key quantities of biological interest. These new results are summarized and listed in the Summary and Conclusions. We also carry out a numerical experiment on a simple example to show how cellular gene expression levels evolve over generations and how the macroscopic cellular density (with respect to gene expression level), when interrupted by cellular division, can be prevented from returning to the equilibrium distribution. In the Conclusions, we discuss potential applications and extensions.

II. KINETIC EQUATION FRAMEWORK

For simplicity, we first assume the internal state of each cell is characterized by a one-dimensional scalar quantity $X \in \mathbb{R}$. This continuous stochastic variable may represent, for example, the expression level of a single mRNA transcript or protein abundance (or log-abundance). Besides this continuous variable, associated with each cell is the discrete generation $i \in \mathbb{N}^+$ to which it belongs (assuming it is part of a lineage derived from an ancestor).

We model the evolution of $X_{i,j}$ (the internal state of the jth cell in the ith generation) using an SDE of the standard form

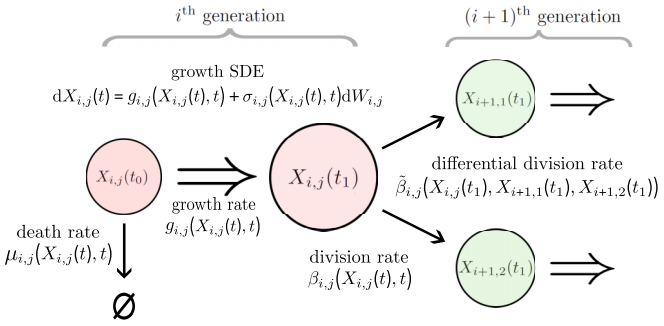

FIG. 1. A schematic of our generation-dependent cellular state evolution model Eq. (1). Here we let $X_{i,j}(t)$ refers to the cell size of the jth cell in the ith generation. Intradivision growth (shown here to occur between times t_0 and t_1) is described by the SDE 1 where $g_{i,j}$ is the cellular size growth rate and $\sigma_{i,j}$ is the amplitude growth rate fluctuations. When a cell in the ith generation divides at time t_1, it will give birth to two new cells in the $(i+1)$th generation. In the particular case of cell size measured by volume, we expect total volume to be conserved immediately after division and $X_{i+1,1}(t_1) + X_{i+1,2}(t_1) = X_{i,j}(t_1)$.

[28,29]

$$\mathrm{d}X_{i,j}(t) = g_{i,j}(X_{i,j},t)\mathrm{d}t + \sigma_{i,j}(X_{i,j},t)\mathrm{d}W_{i,j}, \tag{1}$$

where $g_{i,j}(X_{i,j},t)\mathrm{d}t$ is the deterministic convection that depends on both $X_{i,j}$ and the generation i, and $\mathrm{d}W_{i,j}$ are increments of independent Wiener processes for each i, j. Thus, the term $\sigma_{i,j}(X_{i,j},t)\mathrm{d}W_{i,j}$ represents the "intrinsic" fluctuation in the evolution of $X_{i,j}(t)$. Often, one can assume that the convection arises from gradients of a potential "energy function" Φ: $g_{i,j}(X_{i,j},t) \coloneqq -\nabla\Phi(x,t)|_{x=X_{i,j}}$ [21]. Although a gradient of $\Phi(x,t)$ may conveniently describes a time-dependent force that changes gene expression, nonconservative driving with metabolically driven fluxes, which cannot be described by a potential, is also to be expected [30].

We assume that both $g_{i,j}$ and $\sigma_{i,j}$ are Lipschitz continuous so the solution $X_{i,j}(t)$ of Eq. (1) exists and is almost surely unique given any initial condition $X_{i,j}(0)$. The evolution of $X_{i,j}$ is interrupted by the cell division; an ith generation cell with internal state $X_{i,j}$ divides in time $\mathrm{d}t$ with total probability $\beta_i(X_{i,j})\mathrm{d}t$. This Markovian birth rate can be further stratified by internal state of the two resulting daughter cells immediately after their birth. We denote the differential birth rate density of producing one daughter with internal state X_1 and the other with state X_2 as $\tilde{\beta}_{i,j}(X_{i,j},X_1,X_2)$. Integrating over all possible daughter cell states X_1, X_2 defines the total division rate:

$$\int \tilde{\beta}_{i,j}(X_{i,j},X_1,X_2)\mathrm{d}X_1\mathrm{d}X_2 = \beta_{i,j}(X_{i,j}). \tag{2}$$

A form for $\tilde{\beta}$ might be

$$\tilde{\beta}_{i,j}(X_{i,j},X_1,X_2) \propto e^{-\phi(X_1,X_2|X_{i,j})}, \tag{3}$$

which defines a "free energy" function $\phi(X_1,X_2|X_{i,j})$ for the rate of a mother cell with attribute value X to divide into daughters cell with attribute values X_1 and X_2. If the states of the daughter cells tend towards being similar in value to that of their mother cell, then $\phi(X_1,X_2|X_{i,j})$ would exhibit a

minimum at $X_1, X_2 \approx X$. Although Φ and ϕ might be loosely described in terms of Waddington and fitness landscapes, our unifying kinetic framework allows them to be unambiguously described in terms of the intracellular advection $g_{i,j}(X_{i,j}, t)$ and proliferation function $\tilde{\beta}$, respectively.

Figure 1 summarizes our generation-dependent cellular state evolution model Eq. (1), incorporating cell division and cell death. In this example, $X_{i,j}(t)$ represents the continuous cell size of the jth cell in the ith generation.

Since the derivation of our kinetic theory requires the use of a number of variables and indices, we define some simplifying notation. Specifically, each of the n_i elements of the bold vector $\boldsymbol{X}_i$ represents the expression level $X_{i,j}$ of the jth, $1 \leqslant j \leqslant n_i$ cell in the ith-generation subpopulation. These vectors $\boldsymbol{X}_i$ for the subpopulations across generations $1 \leqslant i \leqslant k$ can be collected as a matrix defined as $\boldsymbol{X}_{\boldsymbol{n}} \coloneqq (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k)$, where $\boldsymbol{n} \coloneqq (n_1, \ldots, n_k)$ is a vector representing the total number of cells in each generation $1 \leqslant i \leqslant k$. Each value n_i evolves stochastically defined by random birth and death events. Below is a table of the various definitions and overall notation used throughout this paper.

Next, define $p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ as the probability density function that the population has $\boldsymbol{n}$ cells with internal states $\boldsymbol{X}_{\boldsymbol{n}}$ given the initial condition that the system has $\boldsymbol{n}(0)$ cells with internal state values $\boldsymbol{X}(0)_{\boldsymbol{n}(0)}$ at $t = 0$. For notational simplicity, we name the cell state random variables (at time t) $X_{i,j}(t)$, $\boldsymbol{X}_i(t)$, and $\boldsymbol{X}(t)_{\boldsymbol{n}(t)}$, and denote their values by and $X_{i,j}$, $\boldsymbol{X}_i$, and $\boldsymbol{X}_{\boldsymbol{n}}$, respectively. The probability density $p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ can be defined as the expectation over trajectories from $(\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ to $(\boldsymbol{X}_{\boldsymbol{n}}, t)$:

$$p_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big) = \begin{cases} \mathbb{E}\big[\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)} - \boldsymbol{X}_{\boldsymbol{n}}\big) S\big(t; \boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big) \big| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0; \boldsymbol{n}(0 < s < t) = \boldsymbol{n}(0)\big], & \boldsymbol{n} = \boldsymbol{n}(0) \\ + \displaystyle\int_0^t \mathbb{E}\big[\tilde{J}\big(t, \tau; \boldsymbol{X}_{\boldsymbol{n}}, \boldsymbol{n}(0)\big) S\big(\tau; \boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)}\big) \big| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0; \boldsymbol{n}(0 < s < \tau) = \boldsymbol{n}(0)\big] \mathrm{d}\tau, & \\ \mathbb{E}\Big[\displaystyle\int_0^t \tilde{J}(t, \tau; \boldsymbol{X}_{\boldsymbol{n}}, \boldsymbol{n}(0)) S\big(\tau; \boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)}\big)\, \mathrm{d}\tau \Big| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\Big], & \boldsymbol{n} \neq \boldsymbol{n}(0), \end{cases} \tag{4}$$

where

$$S(t; \boldsymbol{X}(t)_{\boldsymbol{n}(t)}) \equiv \exp\Big(-\int_0^t \sum_{i=1}^{k(0)} \sum_{j=1}^{n_i(0)} \big(\beta_{i,j}(X_{i,j}(s)) + \mu_{i,j}(X_{i,j}(s))\big) \mathrm{d}s\Big)$$

$$\tilde{J}(t, \tau; \boldsymbol{X}_{\boldsymbol{n}}, \boldsymbol{n}(0)) \equiv \sum_{i=1}^{k(0)} \sum_{j=1}^{n_i(0)} \Big(\tilde{\beta}_{i,j}\big(X_{i,j}(\tau), X_1(\tau), X_2(\tau)\big) p_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}}, t - \tau \big| \boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}, 0\big) + \mu_{i,j}(X_{i,j}(\tau)) p_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}}, t - \tau \big| \boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}}, 0\big)\Big). \tag{5}$$

Here $S(t; \boldsymbol{X}(t)_{\boldsymbol{n}(t)})$ is the joint "survival" probability such that no cell death or cell division occurred up to time t; $\mathbb{E}[\tilde{J}(t, \tau; \boldsymbol{X}(t)_{\boldsymbol{n}(t)}, \boldsymbol{n}(0)) S(\tau; \boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)}) | \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0; \boldsymbol{n}(0 < s < \tau) = \boldsymbol{n}(0)]$ in Eq. (4) can be regarded as the probability such that when a cell in the initial population dies of divides at τ, the state of final structured cell population is $\boldsymbol{X}_{\boldsymbol{n}}$. Definitions of $\boldsymbol{X}^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}(s)$ and $\boldsymbol{X}^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}}$ are given in Table I. The term $S(t; \boldsymbol{X}(t)_{\boldsymbol{n}(t)})$ represents the survival probability up to time t while $\tilde{J}(t, \tau; \boldsymbol{X}(t)_{\boldsymbol{n}(t)}, \boldsymbol{n}(0))$ describes the probability flux from a given state $\boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)}$ to the current state $\boldsymbol{X}(t)_{\boldsymbol{n}(t)}$ due to division or death at time τ. The first form on the right-hand side of Eq. (4) is the probability that no division or death happens in the system during time $[0, t]$ and the final internal states of the cell population are $\boldsymbol{X}(t)_{\boldsymbol{n}(t)}$ while the second form in Eq. (4) denotes the probability that at least one division or death happened within $[0, t]$ to arrive at the final internal state $\boldsymbol{X}(t)_{\boldsymbol{n}(t)}$.

We shall show that under certain conditions, $p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ satisfies the partial differential equation

$$\begin{aligned} \frac{\partial p_{\boldsymbol{n}}}{\partial t} + \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{\partial (g_{i,j} p_{\boldsymbol{n}})}{\partial X_{i,j}} = \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{\partial^2 \big(\sigma^2_{i,j} p_{\boldsymbol{n}}\big)}{(\partial X_{i,j})^2} - \sum_{i=1}^{k} \sum_{j=1}^{n_i} \big(\beta_{i,j}(X_{i,j}) + \mu_{i,j}(X_{i,j})\big) p_{\boldsymbol{n}} \\ + \sum_{i=2}^{k} \sum_{j=1}^{n_{i-1}+1} \int \tilde{\beta}_{i,j}\big(Y, X_{i,n_i-1}, X_{i,n_i}\big) p_{\boldsymbol{n}_{\mathrm{b},i-1}}\big(\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{b},i-1}}, t \big| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big)\, \mathrm{d}Y \\ + \sum_{i=1}^{\infty} \sum_{j=1}^{n_i+1} \int \mu_{i,j}(Y) p_{\boldsymbol{n}_{\mathrm{d},i}}\big(\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{d},i}}, t \big| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big)\, \mathrm{d}Y. \end{aligned} \tag{6}$$

In Eq. (6), the predivision cell population $\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{b},i-1}}$ and the predeath cell population $\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{d},i}}$ are explicitly defined in Table I. The mathematical steps and necessary conditions needed to show that $p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ defined in Eq. (4) satisfies Eq. (6) is given in Appendix A. We impose the normalization condition $\sum_{\boldsymbol{n}} \int p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0) \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}} = 1$ for every $\boldsymbol{X}(0)_{\boldsymbol{n}(0)}$ and average over an initial distribution of $\boldsymbol{X}(0)_{\boldsymbol{n}(0)}$ (denoted by $q_{\boldsymbol{n}(0)}(\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$) to define an unconditional probability density

$$p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t) \coloneqq \sum_{\boldsymbol{n}(0)} \int_{\boldsymbol{X}_{\boldsymbol{n}(0)}} p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0) q_{\boldsymbol{n}(0)}(\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)\, \mathrm{d}\boldsymbol{X}(0)_{\boldsymbol{n}(0)} \tag{7}$$

that also satisfies Eq. (6).

TABLE I. Overview of variables. A list of the main variables and parameters used. The specific labels and definitions of state vectors given provide the proper bookkeeping of all possible initial and final states upon birth and death.

Symbol	Definition and explanation
$\boldsymbol{n}(t)$	$\boldsymbol{n}(t) \coloneqq (n_1(t), \ldots, n_{k(t)}(t))$: time-dependent vector of random numbers of cells in the ith generation, $i = 1, \ldots, k$
$\boldsymbol{n}$	$\boldsymbol{n} \coloneqq (n_1, \ldots, n_k)$: vector of integer values n_i of the number of cells in generation $i = 1, \ldots, k$
$\boldsymbol{X}(t)_{\boldsymbol{n}(t)}$	$\boldsymbol{X}(t)_{\boldsymbol{n}(t)} \coloneqq (\boldsymbol{X}_1(t), \ldots, \boldsymbol{X}_{k(t)}(t))$, $\boldsymbol{X}_i(t) \coloneqq (X_{i,1}(t), \ldots, X_{i,n_i(t)}(t))$: time-dependent random variable describing the state of each cell, e.g., gene expression level $X_{i,n_i(t)}$ of the n_i^{th} cell in the ith generation
$\boldsymbol{X}_{\boldsymbol{n}}$	$\boldsymbol{X}_{\boldsymbol{n}} \coloneqq (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k)$, $\boldsymbol{X}_i \coloneqq (X_{i,1}, \ldots, X_{i,n_i})$: values of $\boldsymbol{X}(t)_{\boldsymbol{n}(t)}$
$\vec{X}_n$	$\vec{X}_n \coloneqq (X_1, \ldots, X_n)$, the vector of state values for any collection of n cells
$g_{i,j}(X_{i,j}, t)$	Deterministic growth rate of the jth cell in the ith generation
$\sigma_{i,j}(X_{i,j}, t)$	Noise in the growth of the jth cell in the ith generation
$\beta_{i,j}(X_{i,j})$	Division rate of the jth cell in the ith generation
$\mu_{i,j}(X_{i,j})$	Death rate of the jth cell in the ith generation
$\tilde{\beta}_{i,j}(X_{i,j}, X_1, X_2)$	Differential division rate of the jth cell in the ith generation into two cells in the $(i+1)$th generation with states X_1, X_2
$\boldsymbol{X}^{-j}_{\boldsymbol{n}_{\text{b},-i}}$	States of the cell population right after the jth cell in the ith generation divides. $\boldsymbol{X}^{-j}_{\boldsymbol{n}_{\text{b},-i}}$ differs from $\boldsymbol{X}_{\boldsymbol{n}}$ in that the state variables for the cells in the $(i-1)$th generation is $(X_{i-1,1}, \ldots, X_{i-1,j-1}, X_{i-1,j+1}, \ldots, X_{i-1,n_i})$ and the state variables for the cells in the ith generation are $(X_{i,1}, \ldots, X_{i,n_i}, X_1, X_2)$
$\boldsymbol{X}^{-j}_{\boldsymbol{n}_{\text{d},-i}}$	States of the cell population right after the jth cell in the ith generation dies. $\boldsymbol{X}^{-j}_{\boldsymbol{n}_{\text{d},-i}}$ differs from $\boldsymbol{X}_{\boldsymbol{n}}$ in that the state variables for the cells in the $(i-1)$th generation are $(X_{i-1,1}, \ldots, X_{i-1,j-1}, X_{i-1,j+1}, \ldots, X_{i-1,n_i})$
$\boldsymbol{X}^{j}_{\boldsymbol{n}_{\text{b},i-1}}$	Pre-division cellular population that differs from $\boldsymbol{X}_{\boldsymbol{n}}$ as state variables for $(i-1)$th-generation cells are $(X_{i-1,1}, \ldots, X_{i-1,j-1}, Y, X_{i-1,j}, \ldots)$ while state variables for ith-generation cells are $(X_{i,1}, \ldots, X_{i,n_i-2})$ (an additional cell with Y in the $(i-1)$th generation divides and gives birth to two new daughter cells X_{i,n_i-1}, X_{i,n_i} in the ith generation)
$\boldsymbol{X}^{j}_{\boldsymbol{n}_{\text{d},i}}$	Pre-death cell population states. This differs from $\boldsymbol{X}_{\boldsymbol{n}}$ in that the state variables for the cells in the ith generation are $(X_{i,1}, \ldots, X_{i,j-1}, Y, X_{i,j}, \ldots)$ (an additional cell in the ith generation with Y dies)
$\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\text{b},i}}$	Pre-division state which differs from $\boldsymbol{X}_{\boldsymbol{n}}$ in that the state vector associated with the ith generation is $(Y, X_{i,1}, \ldots, X_{i,n_i})$ and the state of the $(i+1)$th generation does not contain components X_{i+1,j_1} and X_{i+1,j_2}

Next, we define the symmetric probability density distribution

$$\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t) \coloneqq \prod_{i=1}^{k} \frac{1}{n_i!} \sum_{\pi} p_{\boldsymbol{n}}(\pi(\boldsymbol{X}_{\boldsymbol{n}}), t), \tag{8}$$

where $p_{\boldsymbol{n}}$ is defined in Eq. (7) and $\pi(\boldsymbol{X}_{\boldsymbol{n}})$ is a permutation operator that reorders the sequence of the state variables $X_{i,j}$ of cells within each generation, for all generations. For example, for $\boldsymbol{n} = (n_1)$ with $n_1(t) = 2$, we have $\rho_{\boldsymbol{n}}(X_{1,1}, X_{1,2}, t) = \rho_{\boldsymbol{n}}(X_{1,2}, X_{1,1}, t)$ but $p_{\boldsymbol{n}}(X_{1,1}, X_{1,2}, t)$ might not be equal to $p_{\boldsymbol{n}}(X_{1,2}, X_{1,1}, t)$. Thus, the summation is taken over all such grouped permutations ($\prod_{i=1}^{k} n_i!$ permutations in total). In the special case

$$g_{i,j} = g_i, \quad \sigma_{i,j} = \sigma_i, \quad \beta_{i,j} = \beta_i, \quad \tilde{\beta}_{i,j} = \tilde{\beta}_i, \quad \mu_{i,j} = \mu_i, \tag{9}$$

i.e., when the rate parameters depend at most on the generation of a cell, $\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)$ defined in Eq. (8) obeys

$$\begin{aligned} \frac{\partial \rho_{\boldsymbol{n}}}{\partial t} + \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{\partial (g_i \rho_{\boldsymbol{n}})}{\partial X_{i,j}} &= \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{\partial^2 (\sigma_i^2 \rho_{\boldsymbol{n}})}{(\partial X_{i,j})^2} - \sum_{i=1}^{k} \sum_{j=1}^{n_i} \big(\beta_i(X_{i,j}) + \mu_i(X_{i,j})\big) \rho_{\boldsymbol{n}} \\ &\quad + \sum_{i=1}^{k-1} \frac{n_i + 1}{n_{i+1}(n_{i+1} - 1)} \sum_{1 \leqslant j_1 \neq j_2 \leqslant n_{i+1}} \int \tilde{\beta}_i\big(Y, X_{i+1,j_1}, X_{i+1,j_2}\big) \rho_{\boldsymbol{n}_{\text{b},i}}\big(\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\text{b},i}}, t\big)\, \mathrm{d}Y \\ &\quad + \sum_{i=1}^{\infty} \sum_{j=1}^{n_i+1} \int \mu_i(Y) \rho_{\boldsymbol{n}_{\text{d},i}}\big(\boldsymbol{X}^{j}_{\boldsymbol{n}_{\text{d},i}}, t\big)\, \mathrm{d}Y, \end{aligned} \tag{10}$$

where $\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\text{b},i}}$ differs from $\boldsymbol{X}_{\boldsymbol{n}}$ in that the state vector associated with cells in the ith generation are $(Y, X_{i,1}, \ldots, X_{i,n_i})$ and the state vector for cells in the $(i+1)$th generation does not have the components X_{i+1,j_1} and X_{i+1,j_2}.

We note that the complex formulas derived above can be reduced to a kinetic theory for populations stratified by a wide range of continuous attributes, intrinsically stochastic or deterministic. For example, if $X_{i,j}$ is age a (time since birth), then Eq. (10) (with $\sigma = 0$) becomes a kinetic theory for age-structured populations with stochastic birth and death times [11–13]. Further marginalization of independent particles (cells or individuals) reduces the age-structured kinetic theory for the expected density of individuals at age a to the classic McKendrick-von Foerster age-structured PDE [1–3]. Alternatively, as alluded to in the Introduction, if the kinetic parameters g, $\tilde{\beta}$, and μ are *independent* of $X_{i,j}$, then Eq. (10) can be integrated over all $\boldsymbol{X_n}$ leading to a birth-death master equation for $P(\boldsymbol{n}, t)$ [11,12,31], the probability of having a population $\boldsymbol{n}$ at time t [cf. Eqs. (31) and (32)]. Thus, our general kinetic theory framework is foundational to many types of structured (size, age, etc.) population models that arise in demography and ecology [32], while also connecting them to standard memoryless birth-death processes in certain limits. In fact, our framework extends beyond classical linear models by providing a framework that can include interactions among individuals that generate nonlinear terms in structured-population PDE models [cf. Eq. (47)]. For example, cannibalistic interactions between individuals of different ages or sizes has been recently derived from an age-structured kinetic model to explain population overcompensation behavior observed in several ecological systems [33].

In many systems, the attribute variable is a multidimensional vector instead of a scalar, i.e., $X_{i,j} \coloneqq (X_{i,j,1}, \ldots, X_{i,j,d}) \in \mathbb{R}^d$ may also represent d different gene or protein expression levels in the jth cell in the ith generation. This vector may represent, for example, d different gene or protein expression levels. We assume that the evolution of $X_{i,j}$ (each element now implicitly a vector of attributes) follows the Brownian SDE

$$\mathrm{d}X_{i,j} = \boldsymbol{g}_{i,j}(X_{i,j}, t)\mathrm{d}t + \boldsymbol{\Sigma}_{i,j}(X_{i,j}, t)\mathrm{d}\boldsymbol{W}_{i,j}, \tag{11}$$

where $\boldsymbol{W}_{i,j}$ is a d_0-dimensional vector of independent Wiener processes ($d_0 \leqslant d$) for each i, j and the coefficients $\boldsymbol{g}_{i,j}(X_{i,j}, t) \coloneqq (g_{i,j,1}(X_{i,j}, t), \ldots, g_{i,j,d}(X_{i,j}, t)) : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^d$, $\boldsymbol{\Sigma}_{i,j} : \mathbb{R}^h \times \mathbb{R}^+ \to \mathbb{R}^{d\times d_0}$ is the diffusion coefficient matrix such that its entries $(\boldsymbol{\Sigma}_{i,j}(X_{i,j}, t))_{mn}$, $m = 1, \ldots, d, n = 1, \ldots, d_0$ are all smooth, uniform Lipschitz continuous, and uniformly bounded. We can also define the symmetric probability density distribution $\rho_{\boldsymbol{n}}(\boldsymbol{X_n}, t)$ as in Eqs. (8). Suppose the rate parameters depend at most on the generation of a cell, i.e.,

$$\boldsymbol{g}_{i,j} = \boldsymbol{g}_i,\ \boldsymbol{\Sigma}_{i,j} = \boldsymbol{\Sigma}_i,\ \boldsymbol{\beta}_{i,j} = \boldsymbol{\beta}_i,\ \tilde{\beta}_{i,j} = \tilde{\beta}_i,\ \mu_{i,j} = \mu_i; \tag{12}$$

then, after applying the multidimensional forward Feynman-Kac equation case in Ref. [34] we can show that the differential equation satisfied by such $\rho_{\boldsymbol{n}}$ is

$$\begin{aligned}\frac{\partial \rho_{\boldsymbol{n}}}{\partial t} + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\sum_{\ell=1}^{d}\frac{\partial(g_{i\ell}\rho_{\boldsymbol{n}})}{\partial X_{i,j,\ell}} &= \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\sum_{\ell_1,\ell_2=1}^{d}\frac{\partial^2\left(\sum_{h=1}^{d_0}(\Sigma_i)_{\ell_1,h}(\Sigma_i)_{\ell_2,h}\,\rho_{\boldsymbol{n}}\right)}{\left(\partial X_{i,j,\ell_1}\partial X_{i,j,\ell_2}\right)} - \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\beta_i(X_{i,j}) + \mu_i(X_{i,j})\right)\rho_{\boldsymbol{n}} \\ &+ \sum_{i=1}^{k-1}\frac{n_i+1}{n_{i+1}(n_{i+1}-1)}\sum_{1\leqslant j_1\neq j_2\leqslant n_{i+1}}\int\tilde{\beta}_i\left(Y, X_{i+1,j_1}, X_{i+1,j_2}\right)\rho_{\boldsymbol{n}_{\mathrm{b},i}}\left(\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\mathrm{b},i}}, t\right)\mathrm{d}Y \\ &+ \sum_{i=1}^{\infty}\sum_{j=1}^{n_i+1}\int\mu_i(Y)\rho_{\boldsymbol{n}_{\mathrm{d},i}}\left(\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{d},i}}, t\right)\mathrm{d}Y.\end{aligned} \tag{13}$$

In Appendix C, we also derive kinetic equations for the population density associated with cells that are also labeled by their age. The derivation assumes the budding model of birth where on daughter cell's age is set to zero immediately after birth [11,12].

III. MASS-ACTION DIFFERENTIAL EQUATIONS

Henceforth, we will consider the "simpler" single-gene model which are derived from Eqs. (6), (10), or (13) and could be effectively simulated. Extending the model to include d-dimensional attributes can be implemented following the structure in Eqs. (11) and (13). It is usually very difficult to numerically determine $p_{\boldsymbol{n}}$ or $\rho_{\boldsymbol{n}}$ as defined in Eqs. (6), (10), or (13) because the variable $\boldsymbol{X_n}$ can be very high dimensional. However, by proper marginalization of the kinetic equation (10) we can derive the differential equations that describe the evolution of certain "macroscopic" and interpretable quantities such as the expected total-population levels of X. Such macroscopic quantities track a reduced number of variables, allowing functions over them to be more efficiently simulated or computed numerically. In this section, we derive governing equations for specific examples of macroscopic quantities by marginalizing Eq. (10), which are then solved numerically to show how quantities such as cellular gene expression levels can evolve over generations.

A. Evolution of the population density

First, we can track the marginal cell distributions of certain cells in specified generations by defining the macroscopic quantity

$$u_{\boldsymbol{n}}(\boldsymbol{X_n}, t) \coloneqq \sum_{\boldsymbol{m}\geqslant\boldsymbol{n}}\prod_{\ell=1}^{\infty}(\boldsymbol{m}_\ell)_{\boldsymbol{n}_\ell}\int_{\boldsymbol{X}_{\boldsymbol{m}\setminus\boldsymbol{n}}}\rho_{\boldsymbol{m}}(\boldsymbol{X_m}, t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{m}\setminus\boldsymbol{n}}, \tag{14}$$

where $\boldsymbol{m} \geqslant \boldsymbol{n}$ means that for each component in $\boldsymbol{m} \coloneqq (m_1, \ldots, m_\ell)$, $m_\ell \geqslant n_\ell$ and $(m_\ell)_{n_\ell} \coloneqq m_\ell(m_\ell - 1)\cdots(m_\ell - n_\ell + 1)$ is the falling factorial. The integration is taken over the remaining variables $\boldsymbol{X_m}$, but excludes the variables of interest $\boldsymbol{X_n}$ which are retained [i.e., $u_{\boldsymbol{n}}(\boldsymbol{X_n}, t)$ only tracks

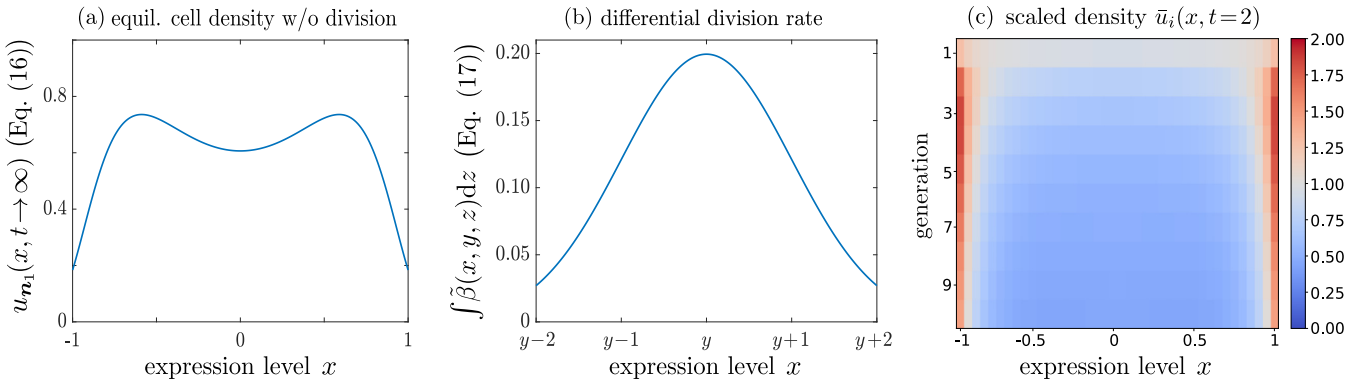

FIG. 2. (a) The equilibrium cellular density without division [Eq. (17)]. (b) A differential birth rate $\int\tilde{\beta}_{i,j}(X_{i,j}, Y, Z)\mathrm{d}Z$ using the form given in Eq. (18). (c) Using the differential birth rate in (b) and Eq. (17) for normalization, we plot the associated cellular density $\bar{u}_i(x, t=2)$ [Eq. (19)] across different generations. The differentiation process prevents the population from reaching an equilibrium ($i \geqslant 2$) even when the death rate and division rate are x independent. However, as time increases for a certain generation (such as $i=1$) in which no cell has entered, the structured population in that generation gradually returns to equilibrium.

the joint "density" of cells with states $\boldsymbol{X}_{\boldsymbol{n}}$ at generations $\boldsymbol{n}$ while ignoring all other cells]. Assuming the relationships in Eqs. (9) hold, we find that $u_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)$ satisfies the differential equation

$$\begin{aligned}\frac{\partial u_{\boldsymbol{n}}}{\partial t}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial(g_i u_{\boldsymbol{n}})}{\partial X_{i,j}}=&\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2(\sigma_i^2 u_{\boldsymbol{n}})}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_i(X_{i,j})+\mu_i(X_{i,j})\big)u_{\boldsymbol{n}}\\&+\sum_{i=1}^{k-1}\sum_{j_1\neq j_2}\int\tilde{\beta}_i\big(Y, X_{i+1,j_1}, X_{i+1,j_2}\big)u_{\boldsymbol{n}_{\mathrm{b},i}}\big(\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\mathrm{b},i}}, t\big)\,\mathrm{d}Y\\&+\sum_{i=1}^{k-1}\sum_{j=1}^{n_{i+1}}\int\big(\tilde{\beta}_i(Y, X_{i+1,j}, Z)+\tilde{\beta}_i(Y, Z, X_{i+1,j})\big)u_{\boldsymbol{n}_{\mathrm{b},i}}\big(\boldsymbol{X}^{1}_{\boldsymbol{n}_{\mathrm{b},i,j}}, t\big)\,\mathrm{d}Y\,\mathrm{d}Z.\end{aligned}\tag{15}$$

In the above equation, $\boldsymbol{X}^1_{\boldsymbol{n}_{\mathrm{b},i,j}}$ indicates that, compared to $\boldsymbol{X}_{\boldsymbol{n}}$, the i^{th} generation has an extra variable Y in the beginning but the $(i+1)^{\text{st}}$ generation is missing the variale $X_{i+1,j}$. From Eq. (15), the set of macroscopic quantities $\{u_{\boldsymbol{n}}\}$ satisfies "sequential" closed-form equations in that the PIDE satisfied by $u_{\boldsymbol{n}}$ depends only on $u_{\boldsymbol{n}_{\mathrm{b},i}}(\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\mathrm{b},i}}, t)$ and $u_{\boldsymbol{n}_{\mathrm{b},i}}(\boldsymbol{X}^{1}_{\boldsymbol{n}_{\mathrm{b},i,j}}, t)$. In the specific case in which only the population density in the structured, ith-generation one-dimensional variable $X_{i,1}$ is tracked, we define $\boldsymbol{n}_i := (0,\ldots,0,1)\in\mathbb{N}^i$ and solve for $u_{\boldsymbol{n}_i}(\boldsymbol{X}_{\boldsymbol{n}_i}, t)$. The quantity $\{u_{\boldsymbol{n}_i}(\boldsymbol{X}_{\boldsymbol{n}_i}, t)\}_{i=1}^{\infty}$ indicates how the cellular population density evolves across generations through division and differentiation. Below, we present two examples of generation-specific structured population densities $u_{\boldsymbol{n}_i}(\boldsymbol{X}_{\boldsymbol{n}_i}, t)$ as defined in Eq. (15).

Example 1. Consider the specific example studied in Ref. [29] where the coefficients in Eq. (15) take the form

$$g_i(X_{i,j}, t)=-X_{i,j},\quad \sigma_i^2(X_{i,j}, t)=\exp\big(-X_{i,j}^2\big).\tag{16}$$

Here we assume the quantity (e.g., gene expression level) to be standardized so that $X_{i,j}\in\mathbb{R}$ and a negative feedback captured by a convection rate proportional to $-X_{i,j}$ represents a negative feedback. In this case, if the cells do not divide or die (i.e., then the entire population stays in the first generation), and their attributes converge to an equilibrium distribution

$$u^*_{\boldsymbol{n}_1}(X_{1,1}=x, t\to\infty)=\frac{\exp\big[2x^2-\frac{1}{2}e^{2x^2}\big]}{Z},\tag{17}$$

where $Z=\int_{-\infty}^{\infty}\exp[2x^2-\frac{1}{2}e^{2x^2}]\mathrm{d}x$ is the normalization constant and $u^*_{\boldsymbol{n}_1}(X_{1,1}=x, t\to\infty)$ represents the population-marginalized density in the absence of birth and death. The equilibrium density is shown in Fig. 2(a).

To include birth and death, we choose rates of the form

$$\begin{aligned}\beta_{i,j}=\frac{1}{2},\quad \mu_{i,j}=\frac{i-1}{2i},\\ \tilde{\beta}_{i,j}(X_{i,j}, Y, Z)=\tilde{\beta}_{i,j}(X_{i,j}, Z, Y),\\ \int\tilde{\beta}_{i,j}(X_{i,j}, Y, Z)\mathrm{d}Z\equiv\frac{1}{2\sqrt{2\pi}}e^{-\frac{(Y-X_{i,j})^2}{2}}\end{aligned}\tag{18}$$

and set the initial condition to be $u_{\boldsymbol{n}_i}(\boldsymbol{X}_{\boldsymbol{n}_i,1}, 0)=0.2\delta_{i,1}$, where $\delta_{i,j}=1$ if $i=j$ and $\delta_{i,j}=0$ otherwise is the Kronecker δ function. The differential division rate $\tilde{\beta}$ is shown in Fig. 2(b).

Using these parameters and initial conditions, we use a finite volume method to numerically solve Eq. (15) satisfied by $u_{\boldsymbol{n}_i}$. We used spatial and temporal mesh sizes $\Delta x=0.05$, $\Delta t=0.0025$ on a spatial domain $\mathcal{D}=[-2.5, 2.5]$ and imposed a Neumann boundary condition: $\partial_x u(x,t)|_{x\in\partial\mathcal{D}}=0$. To illustrate the results, we plot the scaled [using Eq. (17)] generation-dependent cellular density,

$$\bar{u}_i(x, t)\equiv\left(\frac{1}{u^*_{\boldsymbol{n}_1}(x, t\to\infty)}\right)\frac{u_{\boldsymbol{n}_i}(\boldsymbol{X}_{\boldsymbol{n}_i}, t)}{\int_{-\infty}^{\infty}u_{\boldsymbol{n}_i}(\boldsymbol{X}_{\boldsymbol{n}_i}, t)\mathrm{d}X_{i,1}}\tag{19}$$

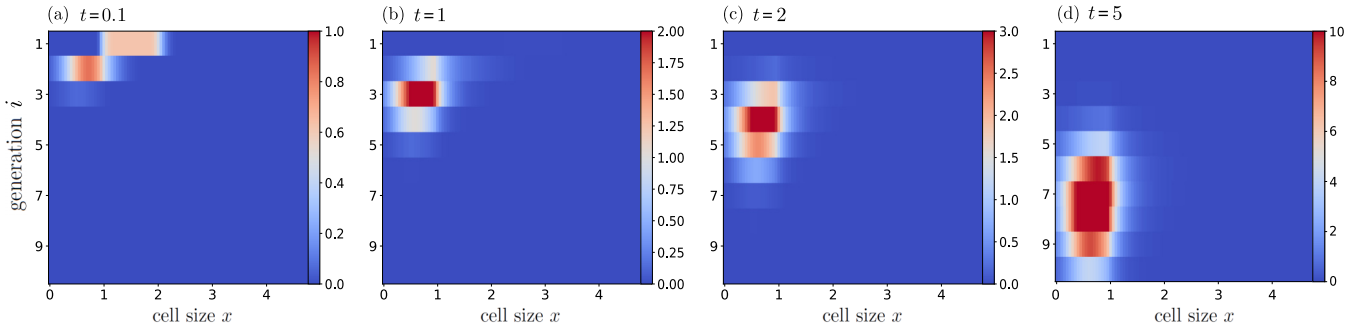

FIG. 3. Cell size densities $u_{\boldsymbol{n}_i}(\boldsymbol{X}_{i,1}=x,t)$ within each generation i at various times t: (a) $t=0.1$, (b) $t=1$, (c) $t=2$, and (d) $t=5$. The numerical results of this sizer division mechanism indicate that the size distribution (across all generations) reaches a steady state while the mean and variance of the distribution over generation number increases linearly with time.

across the first 10 generations at $t=2$ when $x\in[-1,1]$. Figure 2(c) shows that division events, which bring newborn cells into later generations $i\geqslant 2$, prevent structured cellular density in later generations from reaching the equilibrium.

In the next example, instead of gene expression or mRNA level, we define $X_{i,j}$ as cell size or volume. Our reduced kinetic theory can then be used to quantify how the size-structured cell population evolves over time.

Example 2. In this example, we consider a "sizer" cell division model which describes a mechanism whereby a cell divides upon reaching a certain size [25]. In between divisions, we assume exponential cell growth as described in Ref. [35], along with a Langevin noise:

$$g_{i,j}(X_{i,j},t)=g_0X_{i,j},\quad \sigma^2_{i,j}(X_{i,j},t)=\sigma_0^2X_{i,j}. \tag{20}$$

To determine conditions for cell division, we implement the sizer mechanism proposed in Ref. [35]: $X_d=X_0+\eta$, where X_d is the cell size at division, X_0 is a fixed cell size, and η is a random variable independent of cell size. Without loss of generality, we nondimensionalize and set $X_0=1$. The values of the uncertainty η are assumed to follow an exponential distribution $\eta\sim\mathrm{Exp}(4)$. In this case, the division rate is described by $\beta_{i,j}(x)\equiv 4\mathbb{I}_{x\geqslant 1}$, where $\mathbb{I}$ is the indicator function. We set the death rate $\mu_{i,j}=0.01$, assume symmetry in the differential birth rate $\tilde\beta_{i,j}(X_{i,j},Y,Z)=\tilde\beta_{i,j}(X_{i,j},Z,Y)$, and define $\int\tilde\beta_{i,j}(X_{i,j},Y,Z)\mathrm{d}Z=0$ for $Y\leqslant 0$ or $Y>X_{i,j}$. For $0<Y<X_{i,j}$, we define

$$\int\tilde\beta_{i,j}(X_{i,j},Y,Z)\mathrm{d}Z=\frac{\beta_{i,j}(X_{i,j})}{C(X_{i,j})\sqrt{32\pi}}e^{-\frac{(Y-X_{i,j}/2)^2}{32}},$$

$$C(X_{i,j})\coloneqq\frac{1}{\sqrt{32\pi}}\int_0^{X_{i,j}}e^{-\frac{(Y-X_{i,j}/2)^2}{32}}\mathrm{d}Y, \tag{21}$$

such that after division, the summation of the sizes of two daughter cells equals the size of their mother cell. In this example, we also track the evolution of $u_{\boldsymbol{n}_i}(\boldsymbol{X}_{\boldsymbol{n}_i},t)$, the structured population density within the ith generation with respect to size at time t. The initial condition is set to $u_{\boldsymbol{n}_i}(\boldsymbol{X}_{i,1},0)=\mathbb{I}_{1\leqslant X\leqslant 2}\cdot\delta_{i,1}$. Using the above parameters, we use a finite volume method to numerically solve Eq. (15) for $u_{\boldsymbol{n}_i}$. We employed spatial and temporal mesh sizes $\Delta x=0.05$, $\Delta t=0.001$ in a spatial domain $D=[0,5]$. We impose the Neumann boundary condition $\partial_x u(x,t)|_{x\in\partial D}=0$ and plot in Fig. 3 the structured population $u_{\boldsymbol{n}_i}(\boldsymbol{X}_{i,1}=x,t)$ at different t. We see that cell size, under the sizer division mechanism, is well regulated and very few cells reach a size $X\geqslant 4$ before division or death. As time increases, cells divide and their siblings will enter the next generation. The generation-resolved sizer model yields a stationary distribution in size x, but mean generation and variance that increases linearly in time, i.e, the implicit one-term discrete generation recursion [cf. Eq. (30)] is consistent with a convection-diffusion process in the "hydrodynamic" large generation limit.

If the coefficients $g,\sigma,\beta,\tilde\beta$ depend only on the internal state X and time t and not on the cells' generation, i.e.,

$$g_{i,j}=g,\ \sigma_{i,j}=\sigma,\ \beta_{i,j}=\beta,\ \mu_{i,j}=\mu,\ \tilde\beta_{i,j}=\tilde\beta, \tag{22}$$

then we can define

$$\hat\rho_n(\vec X_n,t)\coloneqq\sum_{\sum n_i=n}\frac{1}{n!}\sum_{\pi}p_{\boldsymbol{n}}(\pi(\boldsymbol{X}_{\boldsymbol{n}}),t), \tag{23}$$

where $p_{\boldsymbol{n}}$ is defined in Eq. (7) and the summation over π is over all possible rearrangements of $\boldsymbol{X}_n$ (defined in Table I) of a generation-resolved cell population $\boldsymbol{n}$ such that the union of states of all cells in all generations is $\vec X_n$ (i.e., if we pad $\boldsymbol{X}_{\boldsymbol{n}}$ into one vector $(X_{1,1},X_{1,2},\ldots,X_{k,n_k})$, then such a vector is a rearrangement of $\vec X_n$, the vector of attributes of all n cells as defined in Table I).

It can be shown that the differential equation satisfied by $\hat\rho_n$ is

$$\begin{aligned}&\frac{\partial\hat\rho_n}{\partial t}+\sum_{j=1}^n\frac{\partial(g\hat\rho_n)}{\partial X_j}\\&=\frac12\sum_{j=1}^n\frac{\partial^2(\sigma^2\hat\rho_n)}{(\partial X_j)^2}-\sum_{j=1}^n\big(\beta(X_j)+\mu(X_j)\big)\hat\rho_n\\&\quad+\frac1n\sum_{j_1\neq j_2}\int\tilde\beta\big(Y,X_{j_1},X_{j_2}\big)\hat\rho_{n-1}\big(\vec X_{n_\mathrm{b}}^{j_1,j_2},t\big)\mathrm{d}Y\\&\quad+(n+1)\int\mu(Y)\hat\rho_{n+1}\big(\vec X_{n_\mathrm{d}},t\big)\,\mathrm{d}Y,\end{aligned} \tag{24}$$

where $\vec X_{n_\mathrm{b}}^{j_1,j_2}$ is the predivision cell states that are different from $\vec X_n$ in that it does not have contain the X_{j_1},X_{j_2} terms but

has an extra Y at the end; $\vec{X}_{n_\mathrm{d}}$ is the predeath cell states that is different from $\vec{X}$ in that it has an extra Y component. In this case, we can define the generation-independent marginalized cell density

$$u_n(\vec{X}_n,t) \coloneqq \sum_{m\geqslant n}(m)_n \int \hat{\rho}_m(\vec{X}_m,t)\mathrm{d}\vec{X}_m\backslash\vec{X}_n \tag{25}$$

which satisfies

$$\begin{aligned}
&\frac{\partial u_n(\vec{X}_n,t)}{\partial t}+\sum_{j=1}^{n}\frac{\partial(gu_n)}{\partial X_j}\\
&=\frac{1}{2}\sum_{j=1}^{n}\frac{\partial^2(\sigma^2u_n)}{(\partial X_j)^2}-\sum_{j=1}^{n}\big(\beta(X_j)+\mu(X_j)\big)u_n\\
&\quad+\sum_{j_1\neq j_2}\int\tilde{\beta}\big(Y,X_{j_1},X_{j_2}\big)u_{n-1}\big(\vec{X}_{n_\mathrm{b}}^{j_1,j_2},t\big)\,\mathrm{d}Y\\
&\quad+\sum_{j=1}^{n}\int\big(\tilde{\beta}(Y,X_j,Z)+\tilde{\beta}(Y,Z,X_j)\big)u_n\big(\vec{X}_{n_\mathrm{b}}^{j},t\big)\,\mathrm{d}Y\,\mathrm{d}Z.
\end{aligned} \tag{26}$$

Here $\vec{X}_{n_\mathrm{b}}^{j}$ is different from $\vec{X}_n$ in that X_j is deleted, but an extra variable Y is added as the last component. If we take $n=1$, then we can obtain a closed-form PIDE for describing the cell density with respect to the scalar state variable X

$$\begin{aligned}
&\frac{\partial u_1(X,t)}{\partial t}+\frac{\partial(gu_1)}{\partial X}\\
&=\frac{1}{2}\frac{\partial^2(\sigma^2u_1)}{(\partial X)^2}-\big(\beta(X)+\mu(X)\big)u_1\\
&\quad+\int\big(\tilde{\beta}(Y,X,Z)+\tilde{\beta}(Y,Z,X)\big)u_1(Y,t)\,\mathrm{d}Y\,\mathrm{d}Z.
\end{aligned} \tag{27}$$

Equation (27) is equivalent to the cell sizer model, or a timer-sizer model of cell division [13] after marginalizing over the cells' ages. As an implementation of this model, one can numerically solve Eq. (26) or Eq. (27) using different inferred single-cell-level gene expression dynamics as candidates for g [36].

B. Evolution of cell numbers

In the simplest case where all model parameters are constants, we can marginalize over all cell state variables to obtain total cell populations. More specifically, if we define the generation vector $\boldsymbol{i}\coloneqq(i_1,\ldots,i_k)$, $0<i_1<\cdots<i_k$ and the associated orders of moments $\boldsymbol{\ell}\coloneqq(\ell_1,\ldots,\ell_k)$, $\ell_s>0$, then we can track the expectation of the product of different orders of the number of cells in different generations

$$\mathbb{E}\Big[\prod_{s=1}^{k}n_{i_s}^{\ell_s}\Big]\coloneqq\sum_{\boldsymbol{n}}\prod_{s=1}^{k}n_{i_s}^{\ell_s}\int\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\,\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}. \tag{28}$$

The differential equation satisfied by $\mathbb{E}[\prod_{i=1}^{k}n_i^{\ell_i}]$ can be shown to be

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[\prod_{s=1}^{k}n_{i_s}^{\ell_s}]}{\mathrm{d}t}&=\sum_{r=1,i_r>1}^{k}\beta_{i_r-1}\Big(\mathbb{E}\Big[\prod_{s=1}^{k}\big(n_{i_s}-\delta_{i_r-1,i_s}+2\delta_{i_r,i_s}\big)^{\ell_s}n_{i_r-1}\Big]-\mathbb{E}\Big[\prod_{s=1}^{k}n_{i_s}^{\ell_s}n_{i_r-1}\Big]\Big)\\
&\quad+\sum_{r=1}^{k}\beta_{i_r}\Big(\mathbb{E}\Big[\prod_{s=1}^{k}\big(n_{i_s}-\delta_{i_r,i_s}+2\delta_{i_r+1,i_s}\big)^{\ell_s}n_{i_r}\Big]-\mathbb{E}\Big[\prod_{s=1}^{k}n_{i_s}^{\ell_s}n_{i_r}\Big]\Big)\\
&\quad-\sum_{r=1}^{k-1}\beta_{i_r}\Big(\delta_{i_{r+1}-i_r,1}\Big(\mathbb{E}\Big[\prod_{s=1}^{k}\big(n_{i_s}-\delta_{i_r,i_s}+2\delta_{i_r+1,i_s}\big)^{\ell_s}n_{i_r}\Big]-\mathbb{E}\Big[\prod_{s=1}^{k}n_{i_s}^{\ell_s}n_{i_r}\Big]\Big)\Big)\\
&\quad-\sum_{r=1}^{\infty}\mu_{i_r}\Big(\mathbb{E}\Big[\prod_{s=1}^{k}n_{i_s}^{\ell_s}n_{i_r}\Big]-\mathbb{E}\Big[\prod_{s=1}^{k}\big(n_{i_s}-\delta_{i_s,i_r}\big)^{\ell_s}n_{i_r}\Big]\Big),
\end{aligned} \tag{29}$$

where $\delta_{i_r,i_s}=1$ if $i_r=i_s$ and $\delta_{i_r,i_s}=0$ otherwise is the Kronecker δ function. Here β_{i_r} and μ_{i_r} are constants that can depend on the generation number i_r. Note that if $\boldsymbol{i}=(i)$ is one dimensional, and $\boldsymbol{\ell}=(1)$, then Eq. (29) reduces to the evolution of the average cell number in the ith generation

$$\frac{\mathrm{d}\mathbb{E}[n_i]}{\mathrm{d}t}=2\beta_{i-1}\mathbb{E}[n_{i-1}]-\beta_i\mathbb{E}[n_i]-\mu_i\mathbb{E}[n_i]. \tag{30}$$

Finally, we can consider another special simplifying case where

$$P(\boldsymbol{n},t)\coloneqq\int\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}} \tag{31}$$

is the probability that the population contains $\{n_1,n_2,\ldots,n_k$ cells in generations $1,\ldots,k$, respectively, regardless of the individual's values of X. It turns out that $P(\boldsymbol{n})$ satisfies the series of interdependent master equations

$$\begin{aligned}
\frac{\mathrm{d}P(\boldsymbol{n},t)}{\mathrm{d}t}&=\sum_{i=2}^{k}\beta_{i-1}(n_{i-1}+1)P(\boldsymbol{n}_{\mathrm{b},i-1},t)\\
&\quad-\sum_{i=1}^{k}(\beta_i+\mu_i)n_iP(\boldsymbol{n},t)\\
&\quad+\sum_{i=1}^{\infty}\mu_i(n_i+1)P(\boldsymbol{n}_{\mathrm{d},i},t),
\end{aligned} \tag{32}$$

when the division rates and the birth rates are constants within the same generation, i.e., $\mu_{i,j} \equiv \mu_i$, $\beta_{i,j} \equiv \beta_i$. Equation (32) is a multigenerational birth-death master equation for the number of individuals in each generation i which carries the same structure as birth-death processes for cells grouped by different attributes other than generation [37]. Note that generating new members of a successive generation arises only from birth, while death only decreases the numbers within a generation.

C. Evolution of "biomass"

Another quantity of specific interest is the biomass (e.g., the total amount of protein or mRNA within a subpopulation). Assuming the relationships in Eqs. (9) hold, the expectation of the total mass $X_i \equiv \sum_{j=1}^{n_i} X_{i,j}$ within cells of the ith generation can be evaluated from

$$\mathbb{E}[X_i(t)] = \sum_{\boldsymbol{n}} \int \left(\sum_{j=1}^{n_i} X_{i,j} \right) \rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t) \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}, \tag{33}$$

where $\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)$ is defined in Eq. (10).

In general, the differential equation satisfied by $X_i(t)$ involves higher moment quantities; thus, the model is not closed. However, given certain constraints on the parameters, the dynamics for $X_i(t)$ can be closed, and a solution can be explicitly computed (analytically or numerically). For example, if $\beta_{i,j}(X_{i,j}) \coloneqq \beta_i$, $\mu_{i,j}(X_{i,j}) \coloneqq \mu_i$ are constants, then $g_{i,j}(X_{i,j}) \coloneqq g_i X_{i,j}$ is linear, and the quantity X is conserved across cell division (that is, if the mother cell carries the state variable X and the two daughter cells acquire state values Y_1 and Y_2 at birth, $Y_1 + Y_2 = X$), then

$$\frac{\mathrm{d}\mathbb{E}[X_i(t)]}{\mathrm{d}t} = (g_i - \mu_i - \beta_i)\mathbb{E}[X_i(t)] + \beta_{i-1}\mathbb{E}[X_{i-1}(t)]. \tag{34}$$

Furthermore, if the growth rate and division rate are independent of the generation number i, then we can define expectations over any moment of the total biomass summed over cells of all generations as

$$\mathbb{E}[X^q(t)] = \sum_{\boldsymbol{n}} \int \left(\sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{i,j} \right)^q \rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)\, \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}, \quad q > 1. \tag{35}$$

Specifically, if μ is a constant and $g_i(X) = g_0 X$, $\sigma_{i,j}^2(X) = \sigma_0^2 X$ (and X is conserved across cell divisions), then the differential equations satisfied by the first and second moments of the *total* biomass $X(t)$ and $X^2(t)$ are

$$\begin{aligned} \frac{\mathrm{d}\mathbb{E}[X(t)]}{\mathrm{d}t} &= (g_0 - \mu) \int x u_1(x,t)\mathrm{d}x \equiv (g_0 - \mu)\mathbb{E}[X(t)], \\ \frac{\mathrm{d}\mathbb{E}[X^2(t)]}{\mathrm{d}t} &= (g_0 - 2\mu)\mathbb{E}[X^2(t)] + \sigma^2 \mathbb{E}[X(t)] \\ &\quad + \mu \int x^2 u_1(x,t)\mathrm{d}x. \end{aligned} \tag{36}$$

Only the equation for the mean total biomass $\mathbb{E}[X(t)]$ is closed. Its second moment depends on averages over $u_1(x,t)$ requiring the solution to Eq. (27). General cases for the equations satisfied by $\mathbb{E}[X^q(t)]$ for arbitrary $q \in \mathbb{N}^+$ are discussed in Appendix B.

D. Tracking dead cells

Thus far, we have assumed that the "biomass" X originates from live cells. Once cells die, they are no longer counted in the population and the biomass X associated with them is no longer included. However, experimentally, the protein and/or mRNA extracted from a solution of cells may come from both living and dead cells (at the time of extraction). To describe these types of measurements, we keep track of cells that have died and assign them to the 0th generation $g_0 = \beta_0 = 0$. We denote their states by $\boldsymbol{X}_0 \coloneqq (X_{0,1}, \ldots, X_{0,n_0})$. We then define $\tilde{p}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t | \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ to include the zero-generation (cells that have died) population. Using arguments similar to those in Proposition 2 we can show that under certain conditions $\tilde{p}_{\boldsymbol{n}}$ satisfies the differential equation

$$\begin{aligned} \frac{\partial \tilde{p}_{\boldsymbol{n}}}{\partial t} + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial (g_{i,j}\tilde{p}_{\boldsymbol{n}})}{\partial X_{i,j}} &= \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial^2 (\sigma_{i,j}^2 \tilde{p}_{\boldsymbol{n}})}{(\partial X_{i,j})^2} - \sum_{i=1}^{k}\sum_{j=1}^{n_i} (\beta_{i,j} + \mu_{i,j})\tilde{p}_{\boldsymbol{n}} \\ &\quad + \sum_{i=1}^{k-1}\sum_{j=1}^{n_{i-1}+1} \int \tilde{\beta}_{i,j}\left(Y, X_{i+1,n_{i+1}-1}, X_{i+1,n_{i+1}}\right) \tilde{p}_{\boldsymbol{n}_{\mathrm{b},i}}\left(\boldsymbol{X}_{\boldsymbol{n}_{\mathrm{b},i}^j}, t \middle| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\right) \mathrm{d}Y \\ &\quad + \sum_{i=1}^{\infty}\sum_{j=1}^{n_i+1} \mu_{i,j}\left(X_{0,n_0}\right) \tilde{p}_{\boldsymbol{n}_{\emptyset,i}}\left(\boldsymbol{X}_{\boldsymbol{n}_{\emptyset,i}}^j t \middle| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\right), \end{aligned} \tag{37}$$

where $\boldsymbol{n}_{\emptyset,i}$ differs from in that its 0th component is $n_0 - 1$ but its ith component is $n_i + 1$, and $\boldsymbol{X}_{\boldsymbol{n}_{\emptyset,i}}^j$ differs from $\boldsymbol{X}_{\boldsymbol{n}}$ in that the internal states of the 0th generation (dead cells) are $(X_{0,1}, \ldots, X_{0,n_0-1})$ and the internal states of the ith generation are $(X_{i,1}, \ldots, X_{i,j-1}, X_{0,n_0}, X_{i,j}, \ldots, X_{i,n_i})$ (X_{0,n_0} is in the jth component). Similarly, we can define the unconditional probability density function $\tilde{p}_{\boldsymbol{n}}^*(\boldsymbol{X}_{\boldsymbol{n}}, t)$ as defined in Eq. (7) as well as the symmetrized probability density function

$$\tilde{\rho}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t) \coloneqq \prod_{i=0}^{k} \frac{1}{n_i!} \sum_{\pi} \tilde{p}_{\boldsymbol{n}}^*(\pi(\boldsymbol{X}_{\boldsymbol{n}}), t). \tag{38}$$

If the assumptions given in Eqs. (9) hold, then the PIDE satisfied by $\tilde{\rho}_{\boldsymbol{n}}$ is

$$\frac{\partial \tilde{\rho}_{\boldsymbol{n}}}{\partial t} + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial (g_i \tilde{\rho}_{\boldsymbol{n}})}{\partial X_{i,j}} = \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial^2 (\sigma_i^2 \tilde{\rho}_{\boldsymbol{n}})}{(\partial X_{i,j})^2} - \sum_{i=1}^{k}\sum_{j=1}^{n_i} (\beta_i + \mu_i)\tilde{\rho}_{\boldsymbol{n}} + \sum_{i=1}^{k-1} \frac{n_i + 1}{n_{i+1}(n_{i+1}-1)} \sum_{j_1 \neq j_2} \int \tilde{\beta}_i(Y, X_{i+1,j_1}, X_{i+1,j_2}) \tilde{\rho}_{\boldsymbol{n}_{\mathrm{b},i}}\left(\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\mathrm{b},i}}, t\right) \mathrm{d}Y + \frac{1}{n_0}\sum_{i=1}^{\infty}(n_i+1)\sum_{j=1}^{n_0} \mu_i(X_{0,j}) \rho_{\boldsymbol{n}_{\tilde{\mathrm{d}},i}}\left(\boldsymbol{X}^{1}_{\boldsymbol{n}_{\tilde{\mathrm{d}},i}}, t\right). \tag{39}$$

The expectation of the total biomass $X_0 \equiv \sum_{j=1}^{n_0} X_{0,j}$ associated with dead cells can be found from

$$\mathbb{E}[X_0(t)] \equiv \sum_{\boldsymbol{n}} \int \left(\sum_{j=1}^{n_0} X_{0,j}\right) \tilde{\rho}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)\, \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}. \tag{40}$$

If the death rates μ_i of cells are equal and constant within each generation i, then $\mathbb{E}[X_0(t)]$ satisfies

$$\frac{\mathrm{d}\mathbb{E}[X_0(t)]}{\mathrm{d}t} = \sum_{i=1}^{\infty} \mu_i \mathbb{E}[X_i(t)], \tag{41}$$

where $\mathbb{E}[X_i(t)]$ is the total expected biomass from cells in the ith generation, as defined in Eq. (33).

We can also define second moments involving the biomass from dead cells

$$\mathbb{E}\left[X_0^2(t)\right] = \sum_{\boldsymbol{n}} \int \left(\sum_{j=1}^{n_0} X_{0,j}\right)^2 \tilde{\rho}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)\, \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}} \tag{42}$$

and

$$\mathbb{E}[X_0(t)X(t)] = \sum_{\boldsymbol{n}} \int \left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{i,j}\right)\left(\sum_{\ell=1}^{n_0} X_{0,\ell}\right) \tilde{\rho}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)\, \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}. \tag{43}$$

If we assume that the death rate is a constant μ for all cells, then the growth rate $g(X) = g_0 X$, and the state variable X is conserved at division, then we can derive the differential equations

$$\frac{\mathrm{d}\mathbb{E}\left[X_0^2(t)\right]}{\mathrm{d}t} = 2\mu\mathbb{E}[X_0(t)X(t)] + \mu \sum_{\boldsymbol{n}} \int \left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{i,j}^2\right) \tilde{\rho}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)\, \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}} \tag{44}$$

$$\frac{\mathrm{d}\mathbb{E}[X_0(t)X(t)]}{\mathrm{d}t} = (g_0 - \mu)\mathbb{E}[X_0(t)X(t)] + \mu\mathbb{E}[X^2(t)] - \mu \sum_{\boldsymbol{n}} \int \left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{i,j}^2\right) \tilde{\rho}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)\, \mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}. \tag{45}$$

Higher moments of X_0, X can also be evaluated, which we do not include for brevity.

E. Correlations and interactions

Although examples so far have involved simple forms of g, σ, β, μ that depend only on the state of of the cell being tracked, these rates can depend on the states of other cells in the population. These more complex dependencies prevent closure of the PIDEs and signal more complex correlations, or "interactions." Simple interactions can be incorporated in the "mean-field" limit if we consider the parameters g, σ, β, and μ to be functions of only averaged macroscopic quantities such as $X(t)$.

As an intuitive example, if we allow the death rate of the jth cell in the ith generation to also depend on the total "biomass" from all living cells, then $\mu_{i,j} = \mu_{i,j}(X_{i,j}, \sum_i \sum_{j=1}^{n_i} X_{i,j})$. Furthermore, using the assumptions given in Eqs. (9) leads to a symmetric population density $\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)$ that satisfies

$$\frac{\partial \rho_{\boldsymbol{n}}}{\partial t} + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial (g_i \rho_{\boldsymbol{n}})}{\partial X_{i,j}} = \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial^2 (\sigma_i \rho_{\boldsymbol{n}})}{(\partial X_{i,j})^2} - \sum_{i=1}^{k}\sum_{j=1}^{n_i} \left(\beta_i(X_{i,j}) + \mu_i\left(X_{i,j}, \sum_{\ell=1}^{k}\sum_{m=1}^{n_\ell} X_{\ell,m}\right)\right)\rho_{\boldsymbol{n}} + \sum_{i=1}^{k-1} \frac{n_i + 1}{n_{i+1}(n_{i+1}-1)} \sum_{1 \leqslant j_1 \neq j_2 \leqslant n_{i+1}} \int \tilde{\beta}_i\left(Y, X_{i+1,j_1}, X_{i+1,j_2}\right) \rho_{\boldsymbol{n}_{\mathrm{b},i}}\left(\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\mathrm{b},i}}, t\right) \mathrm{d}Y + \sum_{i=1}^{\infty}\sum_{j=1}^{n_i+1} \int \mu_i\left(Y, \sum_{\ell=1}^{k}\sum_{m=1}^{n_\ell} X_{\ell,m} + Y\right) \rho_{\boldsymbol{n}_{\mathrm{d}}}\left(\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{d},i}}, t\right) \mathrm{d}Y. \tag{46}$$

Due to the dependencies on the mean-field term $\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{i,j}$, we cannot obtain a closed-form equation for macroscopic quantities such as the cellular density $u_1(X_1,t)$ defined in Eq. (27). However, assuming Eqs. (22) and the approximation $\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{i,j} \approx \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{i,j} + Y \approx \mathbb{E}[X(t)]$ hold [with $\mathbb{E}[X(t)]$ defined in Eq. (33)], an approximate PIDE for $u_1(X,t)$ defined in Eq. (25) can be motivated:

$$\frac{\partial u_1(X_1,t)}{\partial t} + \frac{\partial (g u_1)}{\partial X_1} = \frac{1}{2}\frac{\partial^2(\sigma^2 u_1)}{(\partial X_1)^2} - \left(\beta(X_1) + \mu\left(X_1, \int Y u_1(Y,t)\mathrm{d}Y\right)\right)u_1(X,t) + \int \left(\tilde{\beta}(Y,X_1,Z) + \tilde{\beta}(Y,Z,X_1)\right)u_1(Y,t)\,\mathrm{d}Y\,\mathrm{d}Z. \tag{47}$$

Equation (47) is nonlinear because the mean-field term depends on $\int x u_1(x,t)\mathrm{d}x$. Similarly, if other coefficients depend on mean-field quantities or some specific interaction terms among cells exist, then by making assumptions and marginalization, it might still be possible to find self-consistent integrodifferential equations for macroscopic quantities of interest. For example, death rates that depend on the values of X of two different cells have been shown to reduce to a nonlinear interaction term in kinetic derivations of single-species predator-prey type models [33].

IV. SUMMARY AND CONCLUSIONS

In this work, we used the forward-type Feynman-Kac formula and Markov jump process to formulate a kinetic theory for describing the cellular population density of a generation-resolved cellular population with fluctuating rates of changing internal states as well as random division times. Such a general kinetic theory not only tracks each cell's continuous-valued state attribute such as its volume, protein or mRNA abundance, but also its generation (i.e., how many times its ancestors have divided). In general, our kinetic theory framework can apply to any collection of particles that experience demographic noise from birth-death processes as well as noise in specific individual-level attributes. It is a natural framework for one to quantify age-structured, size-structured, or gene expression-structured populations and can be applied to problems in ecology [19], conservation biology, and population genetics [38,39]. The kinetic framework allows one to tailor the convection and death rates to construct systems that exhibit rich crossover behavior [19].

A number of new results were presented. The underlying kinetic theory describing the intra-generation-symmetrized cell populations is given by Eq. (10) (or Eq. (13) for a vector of attributes). We find that this fully resolved, high-dimensional probability density can be marginalized in to different directions. First, one can sum over moments of the discrete populations/subpopulations to find the dynamics of a generalized cell population density $u_{\boldsymbol{n}}(\boldsymbol{X_n},t)$ [Eq. (14)], which is found to obey Eq. (15) when generations are tracked, and Eq. (27) in the generation-independent case. Further marginalizing over all cell attributes $\boldsymbol{X_n}$ allows one to derive simpler equations for useful quantities such as the expected total number of cells in each generation [Eq. (30)] and the generation structure of the total population [Eq. (32)].

Alternatively, the full probability densities can be used to define moments of mean-field quantities such as total gene expression levels or biomass X across the entire population. These are derived in Eqs. (34) and (36), which depend on integrals over the single-particle number density $u_1(x,t)$. We also show how the biomass X_0 from dead calls can also be tracked, as is often the case in experiments. Expressions for the lowest moments are given in Eqs. (41), (44), and (45). Our results are tabulated below:

Note that the PIDEs for marginalized densities $u_{\boldsymbol{n}}(\boldsymbol{X_n},t)$ can be solved numerically using newly developed adaptive spectral methods suited for unbounded domains [40–42], providing an "Eulerian" picture of the structured population density. Our kinetic theory/PIDE framework does not directly track the structure of populations along lineages of cells (a more "Lagrangian" picture) but connecting our Eulerian representation with representations that delineate cell lineages would be useful area of future analysis.

Many of our results can be directly compared to data. At the intracellular level, statistical techniques recently developed for reconstructing general diffusion processes [43,44] make it possible to directly reconstruct the single-cell level gene expression dynamics from single-cell data, i.e., g and σ given by Eq. (1), providing parameters that can then be used in our population-level PIDE models summarized in Table II. Inference of the intracellular SDE from trajectories $X_{i,j}(t_k)$ measured at time points k is also amenable to recently developed machine-learning-based approaches [43,44].

At the population level, the PIDEs we derived for the macroscopic quantities can also be more directly compared to available population-level data, allowing for better inference of unknown coefficients such as g, σ, β, and μ. Our macroscopic PIDEs can also inform inverse-type problems by providing constraints for neural network-based machine learning approaches for inferring model parameters (such as interacting birth and death rates) from data [45,46]. For example, consider Eq. (27) with X representing cell size and $u_1(X,t)$ being the population density of a size-structured cell population. In deriving Eq. (27), the coefficients g, σ, β, $\tilde{\beta}$, and μ are implicitly equivalent across all cells in all generations. Suppose we measure the size-structured cell population density $u_1(X,t)$ at different time points. One can then apply machine-learning-based approaches to train parameterized neural networks that approximate *the functions* $g,\sigma,\beta,\tilde{\beta}$, and μ [45–47]. It will potentially be interesting to compare the g,σ learned from single-cell trajectories to those learned from population density data for further validation. Additionally, the g,σ reconstructed from single-cell dynamics may be more accurate and can be used as priors for inference of β, $\tilde{\beta}$, and μ in the PIDEs.

Finally, our approach shows how cell-cell "interactions" can arise through functions g, σ, β, $\tilde{\beta}$, and μ that depend

TABLE II. Summary of our main results. Functions describing cell numbers and overall attributes are listed, along with the equation numbers of their mathematical definitions and dynamical equations.

Quantity	Meaning	Equation
$u_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t)$	Partially marginalized cell population density of any order	Eq. (15). Closed set of PIDEs for noninteracting systems
$u_n(\vec{X}_n, t)$	Generation-independent cell population density (may include intercellular dependence)	Eqs. (25) and (26)
$\mathbb{E}[\boldsymbol{n}(t)]$	Expectation of moments of total cell number	Eqs. (28), (29), and (30)
$P(\boldsymbol{n}, t)$	Probability of $\boldsymbol{n} = \{n_i\}$ in each generation i	Eqs. (31) and (32)
$\mathbb{E}[X^q(t)]$	Expectation of moments of total biomass or expression levels	Eqs. (35), (36), and (27)
$\mathbb{E}[X_0^p(t)X^q(t)],\ p+q \leqslant 2$	Moments of biomass from dead and living cells	Eqs. (41), (42), (43), (44), (45), and (39)

on the attribute $X_{i,j}$ of multiple cells. Such functional dependencies preclude full marginalization, leading to higher-order correlation terms for which an approximation must be imposed to close the equations. We explicitly showed how a death rate that also depends on the total biomass results in the implicitly nonlinear [in $u_1(x,t)$] PIDE given in (47). Here by inferring some information about the functional form of μ from data, one may also recover information about cell-cell interactions.

Other directions for future analysis include developing tractable models of interactions that arise through complex dependencies of birth and death rates on $\boldsymbol{X}_{\boldsymbol{n}}$ and developing effective transition state models by considering attractors in the landscape dynamics [48]. Structured populations that vary spatially also arise in many applications [49–51]. For models in which convection and diffusion depend on expression levels, the dynamics of $\boldsymbol{X}_{\boldsymbol{n}}$ can be modeled as being coupled to transport. Our model focuses on the case when the evolution of cellular states X follows a pure diffusion process. However, other types of stochastic processes such as the Lévy processes are also used to describe the dynamics of bursty gene expression products, e.g., mRNA and protein copy numbers [52]; compound Poisson processes are also used to describe the dynamics of chemicals related to gene expression dynamics [53]. Generalizing our pure diffusion model [Eq. (1)] to include compound Poisson processes would expand the applicability of our kinetic framework to bursty gene expression dynamics.

The datasets and codes generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

APPENDIX A: DERIVATION OF THE DIFFERENTIAL EQUATION SATISFIED BY THE CELL POPULATION PROBABILITY DENSITY FUNCTION

To show $p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ defined in Eq. (4) satisfies Eq. (6), we require the following two propositions.

Proposition 1. (Forward-type Feynman-Kac formula for the joint survival probability) If the coefficients $g_{i,j}, \sigma_{i,j}, \beta_{i,j}, \mu_{i,j}$ are smooth, uniform Lipschitz continuous, and uniformly bounded, then, under certain conditions, the solution to

$$\frac{\partial \hat{p}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)}{\partial t} + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial (g_{i,j}(X_{i,j}, t)\hat{p}_{\boldsymbol{n}})}{\partial X_{i,j}} = \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{\partial^2 \left(\sigma_{i,j}^2(X_{i,j}, t)\hat{p}_{\boldsymbol{n}}\right)}{(\partial X_{i,j})^2} - \sum_{i=1}^{k}\sum_{j=1}^{n_i} \left(\beta_{i,j}(X_{i,j}) + \mu_{i,j}(X_{i,j})\right)\hat{p}_{\boldsymbol{n}}, \quad \text{(A1)}$$

with initial condition $\hat{p}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, 0) = \delta(\boldsymbol{X}(0)_{\boldsymbol{n}(0)} - \boldsymbol{X}_{\boldsymbol{n}})$ if $\boldsymbol{n} = \boldsymbol{n}(0)$ (and $\hat{p}_{\boldsymbol{n}}(\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0) = 0$ if $\boldsymbol{n} \neq \boldsymbol{n}(0)$), is

$$\hat{p}_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0) := \mathbb{E}\left[\delta\left(\boldsymbol{X}(t)_{\boldsymbol{n}(t)} - \boldsymbol{X}_{\boldsymbol{n}}\right) S\left(t; \boldsymbol{X}(t)_{\boldsymbol{n}(t)}\right) \middle| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0; \boldsymbol{n}(0 \leqslant s \leqslant t) = \boldsymbol{n}(0)\right], \quad \text{(A2)}$$

where each component in $\boldsymbol{X}(t)_{\boldsymbol{n}(t)}$ satisfies Eq. (1).

Proposition 1 provides the PDE satisfied by the density function for all cells with states $\boldsymbol{X}_{\boldsymbol{n}}$ in the absence of division and death. The proof of Proposition 1 and the associated specific technical assumptions are given in Sec. A 1 below.

When cell division or death occurs, the total number of cells changes according to a Markov jump process. Thus, we need the following proposition to derive the differential equation satisfied by the conditional probability density function $p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$ defined in Eq. (4).

Proposition 2. (Markov jump process for describing the transition of cellular states and change in generations resulting from cell division) Given the initial condition $\boldsymbol{n}(0)$ with states $\boldsymbol{X}(0)_{\boldsymbol{n}(0)}$ at $t = 0$ and a target state at time t with $\boldsymbol{n}$ cells and their internal states $\boldsymbol{X}_{\boldsymbol{n}}$, we start with the conditions

$$p_{\boldsymbol{n}}^0\left(\boldsymbol{X}_{\boldsymbol{n}}, t \middle| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\right) := 0, \quad p_{\boldsymbol{n}}^1\left(\boldsymbol{X}_{\boldsymbol{n}}, t \middle| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\right) := \hat{p}_{\boldsymbol{n}}\left(\boldsymbol{X}_{\boldsymbol{n}}, t \middle| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\right), \quad \text{(A3)}$$

and recursively define

$$p_{\boldsymbol{n}}^{m+1}\big(\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big):=\hat{p}_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)+\int_0^t\mathbb{E}\big[S\big(\tau;\boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)}\big)J^m\big(t,\tau;\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{n}(0)\big)\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0<s<\tau)=\boldsymbol{n}(0)\big]\mathrm{d}\tau, \tag{A4}$$

where the birth-death probability flux is defined by

$$J^m(t,\tau;\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{n}(0)):=\sum_{i=1}^{k(0)}\sum_{j=1}^{n_i(0)}\Big(\tilde{\beta}_{i,j}\big(X_{i,j}(\tau),X_1(\tau),X_2(\tau)\big)p_{\boldsymbol{n}}^m\big(\boldsymbol{X}_{\boldsymbol{n}},t-\tau\big|\boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}},0\big)+\mu_{i,j}(X_{i,j}(\tau))p_{\boldsymbol{n}}^m\big(\boldsymbol{X}_{\boldsymbol{n}},t-\tau\big|\boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}},0\big)\Big). \tag{A5}$$

Then, $p_{\boldsymbol{n}}^{m+1}$ satisfies

$$\begin{aligned}\frac{\partial p_{\boldsymbol{n}}^{m+1}}{\partial t}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial\big(g_{i,j}p_{\boldsymbol{n}}^{m+1}\big)}{\partial X_{i,j}}&=\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma_{i,j}^2p_{\boldsymbol{n}}^{m+1}\big)}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j})+\mu_{i,j}(X_{i,j})\big)p_{\boldsymbol{n}}^{m+1}\\&+\sum_{i=1}^{k-1}\sum_{j=1}^{n_{i-1}^{\mathrm{b}}}\int\tilde{\beta}_{i,j}\big(Y,X_{i+1,n_{i+1}-1},X_{i+1,n_{i+1}}\big)p_{\boldsymbol{n}_{\mathrm{b},i}}^m\big(\boldsymbol{X}^j_{\boldsymbol{n}_{\mathrm{b},i}},t\,|\,\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}Y\\&+\sum_{i=1}^{\infty}\sum_{j=1}^{n_i^{\mathrm{d}}}\int\mu_{i,j}(\mathbf{Y})p_{\boldsymbol{n}_{\mathrm{d},i}}^m\big(\boldsymbol{X}^j_{\boldsymbol{n}_{\mathrm{d},i}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}\mathbf{Y}.\end{aligned} \tag{A6}$$

Furthermore, $p_{\boldsymbol{n}}^m$ is nondecreasing in m.

The proof of Proposition 2 will be given in Sec. A 2 below. Intuitively, m in Eq. (A4) is the maximal number of birth or death events allowed within the cell population. Since $p_{\boldsymbol{n}}^m$ is increasing in m, there exists a p^* such that $p^m\to p^*$ a.s. for all $\boldsymbol{X}(0)_{\boldsymbol{n}(0)}$ and $\boldsymbol{X}_{\boldsymbol{n}}$. After integrating over $\boldsymbol{X}_{\boldsymbol{n}}$ and summing over all $\boldsymbol{n}$ on both sides of Eq. (A4) and assuming

$$\sum_{\boldsymbol{n}}\int p_{\boldsymbol{n}}^{m-1}\big(\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\,\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\leqslant 1 \tag{A7}$$

for $m\in\mathbb{N}$ and any initial condition $\boldsymbol{X}(0)_{\boldsymbol{n}(0)}$, we have $\sum_{\boldsymbol{n}}\int p_{\boldsymbol{n}}^m(\boldsymbol{X}_{\boldsymbol{n}},t\,|\,\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\leqslant F^m(t;\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0)$, where

$$F^m\big(t;\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big):=\int\hat{p}_{\boldsymbol{n}(0)}\big(\boldsymbol{X}_{\boldsymbol{n}(0)},t\,\big|\,\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}(0)}+\int_0^t\mathbb{E}\Big[S\big(\tau;\boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)}\big)\sum_{i=1}^{k(0)}\sum_{j=1}^{n_i(0)}\big(\beta_{i,j}(X_{i,j}(\tau))+\mu_{i,j}(X_{i,j}(\tau))\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0<s<\tau)=\boldsymbol{n}(0)\Big]\,\mathrm{d}\tau \tag{A8}$$

and $S(\tau;\boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)})$ is defined in Eq. (5). Taking the derivative of $F^m(t;\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0)$, we find $\mathrm{d}F^m(t;\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0)/\mathrm{d}t=0$. It is straightforward to verify that $F^m(0;\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0)=1$; therefore, we have $F^m(t;\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0)\equiv 1,\forall t\geqslant 0$, which indicates that

$$\sum_{\boldsymbol{n}}\int p_{\boldsymbol{n}}^m\big(\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\leqslant 1. \tag{A9}$$

By induction, Eq. (A7) holds true for all $m\in\mathbb{N}^+$. Finally, it is easy to show that $p_{\boldsymbol{n}}^m(\boldsymbol{X}(0)_{\boldsymbol{n}},t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0)\geqslant 0$, so by the monotone convergence theorem,

$$\sum_{\boldsymbol{n}}\int p_{\boldsymbol{n}}^*\big(\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\leqslant 1, \tag{A10}$$

which indicates $0\leqslant p^*<\infty$ exists a.e..

If (i) the convergence $p^m \to p^*$ is uniform and (ii) taking the limit with respect to m is interchangeable with taking the partial derivatives in Eq. (A6), then p^* is the solution to

$$\begin{aligned}\frac{\partial p^*_{\boldsymbol{n}}}{\partial t} + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial (g_{i,j}p^*_{\boldsymbol{n}})}{\partial X_{i,j}} &= \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2(\sigma_{i,j}p^*_{\boldsymbol{n}})}{(\partial X_{i,j})^2} - \sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j})+\mu_{i,j}(X_{i,j})\big)p^*_{\boldsymbol{n}}\\ &+ \sum_{i=1}^{k-1}\sum_{j=1}^{n_i+1}\int \tilde{\beta}_{i,j}\big(Y, X_{i+1,n_{i+1}-1}, X_{i+1,n_{i+1}}\big)p^*_{\boldsymbol{n}_{\mathrm{b},i}}\big(\boldsymbol{X}^j_{\boldsymbol{n}_{\mathrm{b},i}}, t\,\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big)\mathrm{d}Y\\ &+ \sum_{i=1}^{\infty}\sum_{j=1}^{n_i+1}\int \mu_{i,j}(\mathbf{Y})p^*_{\boldsymbol{n}_{\mathrm{d},i}}\big(\boldsymbol{X}^j_{\boldsymbol{n}_{\mathrm{d},i}}, t\,\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big)\mathrm{d}\mathbf{Y}. \end{aligned} \tag{A11}$$

Since by taking the limit $m \to \infty$ in Eq. (A4), p^* can also be written as

$$\begin{aligned}p^*_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}}, t\,\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big) &= \hat{p}_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}}, t\,\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big)\\ &+ \int_0^t \mathbb{E}\big[S\big(\tau;\boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)}\big)J^m\big(t,\tau;\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{n}(0)\big)\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}(0),0;\boldsymbol{n}(0<s<\tau)=\boldsymbol{n}(0)\big]\,\mathrm{d}\tau, \end{aligned} \tag{A12}$$

which solves the differential equation (6).

Finally, the definition of p^* in Eq. (A12) coincides with the definition of p in Eq. (4). Thus, if Eq. (A12) defined a unique p^*, then $p^*_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0) = p_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}}, t|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$. Therefore, $p_{\boldsymbol{n}}$ also solves the differential equation Eq. (6). Specifically, if

$$\sum_{\boldsymbol{n}}\int p_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}}, t\,\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\big)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}} = 1, \tag{A13}$$

then $p_{\boldsymbol{n}}$ is indeed a probability density function of the total cell population that satisfies Eq. (6).

1. Proof of Proposition 1

Here we prove Proposition 1 and provide the needed technical assumptions. We shall apply Theorem 6.2 in Ref. [34]. If $\boldsymbol{n} \neq \boldsymbol{n}(0)$, then by definition $\hat{p}_{\boldsymbol{n}} = 0$, which solves Eq. (A1). If $\boldsymbol{n}(s) \equiv \boldsymbol{n}(0)$, $s \in [0,t]$, then for any smooth function $\phi \in C^\infty(\mathbb{R}^{|\boldsymbol{n}|_1})$, $|n|_1 := \sum_{i=1}^k n_i$, we define the measure

$$\gamma^m(\phi, t) := \int_{\mathcal{C}^{|\boldsymbol{n}|_1}} \phi(\boldsymbol{X}_{\boldsymbol{n}}(t;\omega))S\big(t;\boldsymbol{X}(t,\omega)_{\boldsymbol{n}(t)}\big)\mathrm{d}m(\omega), \quad \boldsymbol{X}(0;\omega)_{\boldsymbol{n}} = \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, \tag{A14}$$

where $\mathcal{C}^d := \mathcal{C}([0,t],\mathbb{R}^d)$ (the integration is taken all realizations of $\boldsymbol{X}(t;\omega)_{\boldsymbol{n}}$). Using Theorem 6.2 in Ref. [34], $\gamma^m(\phi)$ solves the PDE

$$\frac{\partial \gamma^m}{\partial t} + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial (g_{i,j}(X_{i,j}(t),t)\gamma^m)}{\partial X_{i,j}(t)} = \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma^2_{i,j}(X_{i,j}(t),t)\gamma^m\big)}{(\partial X_{i,j}(t))^2} - \sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j}(t))+\mu_{i,j}(X_{i,j}(t))\big)\gamma^m \tag{A15}$$

in the sense of distributions. Let $K^\epsilon = \frac{1}{\epsilon^{|\boldsymbol{n}|_1}}K(\cdot)$, where $K(\cdot)$ is a smooth mollifier, and define

$$v^\epsilon(\boldsymbol{X}_{\boldsymbol{n}}, t) := \gamma^m(K^\epsilon(\cdot - \boldsymbol{X}_{\boldsymbol{n}}), t), \tag{A16}$$

or

$$v^\epsilon(\boldsymbol{X}_{\boldsymbol{n}}, t) = \mathbb{E}\big[K^\epsilon\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)} - \boldsymbol{X}_{\boldsymbol{n}}\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\big], \tag{A17}$$

where $S(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)})$ is the survival probability defined in Eq. (5). By Eq. (A15), we have

$$\begin{aligned}\frac{\partial v^\epsilon(\boldsymbol{X}_{\boldsymbol{n}}, t)}{\partial t} &= \mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\partial_{X_{i,j}(t)}K^\epsilon\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)} - \boldsymbol{X}_{\boldsymbol{n}}\big)g_{i,j}\big(X_{i,j}(t),t\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big]\\ &+ \mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{1}{2}\partial^2_{X_{i,j}(t)}K^\epsilon\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)} - \boldsymbol{X}_{\boldsymbol{n}}\big)\sigma^2_{i,j}\big(X_{i,j}(t),t\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big]\\ &- \mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j}(t))+\mu_{i,j}(X_{i,j}(t))\big)K^\epsilon\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)} - \boldsymbol{X}_{\boldsymbol{n}}\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big]. \end{aligned} \tag{A18}$$

The assumptions that we shall impose for Proposition 1 are that (i) the limit

$$v \coloneqq \lim_{\epsilon\to 0^+} v^\epsilon = \mathbb{E}\big[\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)}-\boldsymbol{X}_{\boldsymbol{n}}\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\big] \tag{A19}$$

exists and (ii) taking the limit $\epsilon\to 0^+$ commutes with taking the expectation and the derivative with respect to t and $X_{i,j}$, i.e.,

$$\begin{aligned}\frac{\partial v(\boldsymbol{X}_{\boldsymbol{n}(t)},t)}{\partial t} &= \mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\partial_{X_{i,j}(t)}\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)}-\boldsymbol{X}_{\boldsymbol{n}}\big)g_{i,j}\big(X_{i,j}(t),t\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big]\\ &+\mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{1}{2}\partial^2_{X_{i,j}(t)}\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)}-\boldsymbol{X}_{\boldsymbol{n}}\big)\sigma^2_{i,j}\big(X_{i,j}(t),t\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big]\\ &-\mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j}(t))+\mu_{i,j}(X_{i,j}(t))\big)\delta(\boldsymbol{X}(t)_{\boldsymbol{n}(t)}-\boldsymbol{X}_{\boldsymbol{n}})S(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)})\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big].\end{aligned} \tag{A20}$$

After integration by parts and noticing that

$$\begin{aligned}g_{i,j}(X_{i,j},t)v &\equiv \mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)}-\boldsymbol{X}_{\boldsymbol{n}}\big)g_{i,j}\big(X_{i,j}(t),t\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big],\\ \sigma^2_{i,j}(X_{i,j},t)v &\equiv \mathbb{E}\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)}-\boldsymbol{X}_{\boldsymbol{n}}\big)\sigma^2_{i,j}\big(X_{i,j}(t),t\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\Big],\end{aligned} \tag{A21}$$

the partial differential equation satisfied by v is

$$\frac{\partial v}{\partial t}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial(g_{i,j}(X_{i,j},t)v)}{\partial X_{i,j}}=\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma^2_{i,j}(X_{i,j},t)v\big)}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j})+\mu_{i,j}(X_{i,j})\big)v, \tag{A22}$$

which proves Proposition 1.

2. Proof of Proposition 2

We prove Proposition 2 by induction on m. Clearly, when $m=0$, 1, p^0 and p^1 solve Eq. (A6) by using Proposition 1. If the conclusion holds for $m\geqslant 1$, then when $\boldsymbol{n}\neq\boldsymbol{n}(0)$, we have

$$\begin{aligned}\frac{\partial p^{m+1}_{\boldsymbol{n}}}{\partial t} &= \mathbb{E}\big[S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)J^m(t,t;\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{n}(0))\big|\boldsymbol{X}_{\boldsymbol{n}(0)}(0),0;\boldsymbol{n}(0<s<t)=\boldsymbol{n}(0)\big]\\ &+\int_0^t\mathbb{E}\big[S(\tau;\boldsymbol{X}_{\boldsymbol{n}})\partial_t J^m(t,\tau;\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{n}(0))\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0<s<\tau)=\boldsymbol{n}(0)\big]\,\mathrm{d}\tau\\ &=-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial\big(g_{i,j}(X_{i,j},t)p^{m+1}_{\boldsymbol{n}}\big)}{\partial X_{i,j}}+\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma^2_{i,j}(X_{i,j},t)p^{m+1}_{\boldsymbol{n}}\big)}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j})+\mu_{i,j}(X_{i,j})\big)p^{m+1}_{\boldsymbol{n}}\\ &+\sum_{i=1}^{k-1}\sum_{j=1}^{n_i+1}\int\tilde{\beta}_{i,j}\big(Y,X_{i+1,n_{r+1}-1},X_{r+1,n_{i+1}}\big)\\ &\qquad\times\mathbb{E}\Big[\delta(X_{i,j}(t)-Y)\delta\big(\boldsymbol{X}(t)^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}-\boldsymbol{X}_{\boldsymbol{n}}\big)\delta_{\boldsymbol{n}(0)_{\mathrm{b},-i},\boldsymbol{n}}S(t;\boldsymbol{X}_{\boldsymbol{n}})\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0<s<t)=\boldsymbol{n}(0)\Big]\mathrm{d}Y\\ &+\sum_{i=1}^{k-1}\sum_{j=1}^{n_i+1}\int\tilde{\beta}_{i,j}\big(Y,X_{i+1,n_{i+1}-1},X_{i+1,n_{i+1}}\big)\\ &\qquad\times\Big(\int_0^t\mathbb{E}\big[S(\tau;\boldsymbol{X}_{\boldsymbol{n}})J^{m-1}\big(t,\tau;\boldsymbol{X}^j_{\boldsymbol{n}_{\mathrm{b},i}},\boldsymbol{n}(0)\big)\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0<s<t)=\boldsymbol{n}(0)\big]\mathrm{d}\tau\Big)\,\mathrm{d}Y\\ &+\sum_{i=1}^{\infty}\sum_{j=1}^{n_i+1}\int\mu_{i,j}(Y)\mathbb{E}\big[\delta(X_{i,j}-Y)\delta\big(\boldsymbol{X}^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}}(t)-\boldsymbol{X}_{\boldsymbol{n}}\big)\delta_{\boldsymbol{n}(0)_{\mathrm{d},-i},\boldsymbol{n}}S(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)})\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0<s<t)=\boldsymbol{n}(0)\big]\mathrm{d}Y\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=1}^{\infty}\sum_{j=1}^{n_i+1}\int \mu_{i,j}(Y)\int_0^t \mathbb{E}\big[S(\tau;\boldsymbol{X}_{\boldsymbol{n}})J^{m-1}\big(t,\tau;\boldsymbol{X}^{-j}_{\boldsymbol{n}_{\mathrm{d},i}},\boldsymbol{n}(0)\big)\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big]\mathrm{d}\tau\,\mathrm{d}Y\\
&=-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial\big(g_{i,j}p^{m+1}_{\boldsymbol{n}}\big)}{\partial X_{i,j}}+\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma_{i,j}p^{m+1}_{\boldsymbol{n}}\big)}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(X_{i,j})+\mu_{i,j}(X_{i,j})\big)p^{m+1}_{\boldsymbol{n}}\\
&+\sum_{i=1}^{k-1}\sum_{j=1}^{n_i+1}\int\tilde{\beta}_{i,j}\big(Y,X_{i+1,n_{i+1}-1},X_{i+1,n_{i+1}}\big)p^{m}_{\boldsymbol{n}_{\mathrm{b},i}}\big(\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{b},i}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}Y\\
&+\sum_{i=1}^{\infty}\sum_{j=1}^{n_i+1}\int\mu_{i,j}(Y)p^{m}_{\boldsymbol{n}_{\mathrm{d},i}}\big(\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{d},i}},t\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}Y.
\end{aligned}
\tag{A23}
$$

Here the function $\delta_{\boldsymbol{n}(0)_{\mathrm{b},-i},\boldsymbol{n}}=1$ if $\boldsymbol{n}(0)_{\mathrm{b},-i}=\boldsymbol{n}$ and $\delta_{\boldsymbol{n}(0)_{\mathrm{b},-i},\boldsymbol{n}}=0$ otherwise; similarly, $\delta_{\boldsymbol{n}(0)_{\mathrm{d},-i},\boldsymbol{n}}=1$ if $\boldsymbol{n}(0)_{\mathrm{d},-i}=\boldsymbol{n}$ and $\delta_{\boldsymbol{n}(0)_{\mathrm{d},-i},\boldsymbol{n}}=0$ otherwise. Proposition 1 shows that

$$
\mathbb{E}\big[\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}}-\boldsymbol{X}_{\boldsymbol{n}}\big)S\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)}\big)\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant s\leqslant t)=\boldsymbol{n}(0)\big]
\tag{A24}
$$

satisfies Eq. (A1), so we can verify that Eq. (A6) also holds for $m+1$ when $\boldsymbol{n}=\boldsymbol{n}(0)$. Thus, we have proved that Eq. (A6) holds true for $m+1$. Additionally, it is obvious that $p^{m+1}_{\boldsymbol{n}}\geqslant p^{m}_{\boldsymbol{n}}$ holds for $m=0$. If $p^{m}_{\boldsymbol{n}}\geqslant p^{m-1}_{\boldsymbol{n}}$ for any $\boldsymbol{n},\boldsymbol{X}_{\boldsymbol{n}}$, then we have for $\Delta^{m}_{\boldsymbol{n}}:=p^{m}_{\boldsymbol{n}}-p^{m-1}_{\boldsymbol{n}}$,

$$
\begin{aligned}
\Delta^{m+1}_{\boldsymbol{n}}=\int_0^t\mathbb{E}\Big[&S(\tau;\boldsymbol{X}_{\boldsymbol{n}})\sum_{i=1}^{k(0)}\sum_{j=1}^{n_i(0)}\big(\tilde{\beta}_{i,j}(X_{i,j}(\tau),X_1,X_2)\big)\Delta^{m}_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}},t-\tau\big|\boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}},0\big)\\
&+\mu_{i,j}(X_{i,j})\Delta^{m}_{\boldsymbol{n}}\big(\boldsymbol{X}_{\boldsymbol{n}},t-\tau\big|\boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}_{\mathrm{d},-i}},0\big)\big)\Big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant s\leqslant\tau)=\boldsymbol{n}(0)\Big]\mathrm{d}\tau\geqslant 0.
\end{aligned}
\tag{A25}
$$

Therefore, we have proved that $p^{m+1}_{\boldsymbol{n}}$ satisfies Eq. (A6) and that $p^{m+1}_{\boldsymbol{n}}\geqslant p^{m}_{\boldsymbol{n}}$ for all $m\in\mathbb{N}$ by induction.

APPENDIX B: DIFFERENTIAL EQUATIONS SATISFIED BY $X^q(t)$, $q\in\mathbb{N}^+$

With $X^q(t)$ according to Eq. (35), it can be shown that for $q>1$,

$$
\begin{aligned}
\frac{\mathrm{d}X^q(t)}{\mathrm{d}t}=&\,q\sum_{\boldsymbol{n}}\int\Big(\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{i,j}\Big)^{q-1}\Big(\sum_{\ell=1}^{k}\sum_{m=1}^{n_\ell}g_\ell(X_{\ell,m},t)\Big)\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\\
&+\frac{q(q-1)}{2}\sum_{\boldsymbol{n}}\int\Big(\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{i,j}\Big)^{q-2}\Big(\sum_{\ell=1}^{k}\sum_{m=1}^{n_\ell}\sigma^2_\ell(X_{\ell,m},t)\Big)\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\\
&-\sum_{\boldsymbol{n}}\int\Big[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\mu_i(X_{i,j},t)\sum_{r=1}^{q}(-1)^{r-1}\binom{q}{r}\Big(\sum_{\ell=1}^{k}{}'\sum_{m=1}^{n_\ell}{}'X_{\ell,m}\Big)^{q-r}X^r_{i,j}\Big]\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\\
&-\sum_{\boldsymbol{n}}\int\Big(\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{i,j}\Big)^{q}\sum_{\ell=1}^{k}\sum_{m=1}^{n_\ell}\beta_\ell(X_{\ell,m})\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\\
&+\sum_{\boldsymbol{n}}\int\Big(\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{i,j}\Big)^{q}\Big(\sum_{\ell=1}^{k-1}\frac{n_\ell+1}{n_{\ell+1}(n_{\ell+1}-1)}\int\sum_{j_1\neq j_2}\tilde{\beta}_\ell(Y,X_{\ell+1,j_1},X_{\ell+1,j_2})\rho_{\boldsymbol{n}_{\mathrm{b},\ell}}\big(\boldsymbol{X}^{j_1,j_2}_{\boldsymbol{n}_{\mathrm{b},\ell}},t\big)\mathrm{d}Y\Big)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}},
\end{aligned}
\tag{B1}
$$

where $\rho_{\boldsymbol{n}}$ is the symmetric probability density function defined in Eq. (8). Here $\sum_{\ell=1}^{k}{}'\sum_{m=1}^{n_\ell}{}'$ denote sums over which $\ell\neq i$ or $m\neq j$.

In particular, if X is a conserved quantity at division, then the evolution of the second-order moment can be further simplified as

$$\begin{aligned}\frac{\mathrm{d}X^q(t)}{\mathrm{d}t} &= q\sum_{\boldsymbol{n}}\int\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{i,j}\right)^{q-1}\sum_{\ell=1}^{k}\sum_{m=1}^{n_\ell}g_\ell(X_{\ell,m},t)\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\\ &\quad+\sum_{\boldsymbol{n}}\frac{q(q-1)}{2}\int\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{i,j}\right)^{q-2}\sum_{\ell=1}^{k}\sum_{m=1}^{n_\ell}\sigma_\ell^2(X_{\ell,m},t)\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\\ &\quad-\sum_{\boldsymbol{n}}\int\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\mu_i(X_{i,j})\sum_{r=1}^{q}(-1)^{r-1}\binom{q}{r}\left(\sum_{\ell=1}^{k}{}'\sum_{m=1}^{n_\ell}{}'X_{\ell,m}\right)^{q-r}X_{i,j}^r\right]\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}.\end{aligned}\tag{B2}$$

Equation (B2) can be further simplified if the coefficients g_i and σ_i satisfy certain conditions. For example, if the cells grow exponentially, i.e., $g_i(X_{i,j},t)=\lambda X_{i,j}$ and $\sigma_i^2(X_{i,j},t)=\sigma^2X_{i,j}$, then Eq. (B2) can be more simply expressed as

$$\frac{\mathrm{d}X^q(t)}{\mathrm{d}t}=\lambda qX^q(t)+\frac{q(q-1)}{2}\sigma^2X^{q-1}(t)-\sum_{\boldsymbol{n}}\int\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\mu_i(X_{i,j})\sum_{r=1}^{q}(-1)^{i-1}\binom{q}{r}\left(\sum_{\ell=1}^{k}{}'\sum_{m=1}^{n_\ell}{}'X_{\ell,m}\right)^{q-r}X_{i,j}^r\right]\rho_{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{n}},t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}.\tag{B3}$$

APPENDIX C: BIRTH-INDUCED BOUNDARY CONDITIONS

We can also consider variables that describe cellular quantities that reset upon cell division. Example of such variables include cell size and cell age [13,25]. Specifically, consider simple "timer" models where a new daughter cell acquires age 0 at its birth, while the other cell is assumed to be the "mother" that continues to age. This assignment of age across a proliferating population is described as "budding" birth [11,12]. A kinetic theory can track both cell volume and cell age through the variables $\boldsymbol{X}_{\boldsymbol{n}}$ and $\boldsymbol{A}_{\boldsymbol{n}}\coloneqq(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_k)$, respectively. Here in analogy with $X_{i,j}$ $(j\leqslant n_i)$ (Table I), $\boldsymbol{A}_i\coloneqq(A_{i,1},\ldots,A_{i,n_i})$ and $A_{i,j}$ $(j\leqslant n_i)$ is the age of the jth cell of generation i. Here the corresponding parameters $g_{i,j}$, $\sigma_{i,j}$, $\beta_{i,j}$, $\tilde{\beta}_{i,j}$, and $\mu_{i,j}$ can depend also on $A_{i,j}$.

We can show that the solution to

$$\begin{aligned}&\frac{\partial\hat{p}_{\boldsymbol{n}}(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t)}{\partial t}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial\hat{p}_{\boldsymbol{n}}}{\partial A_{i,j}}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial\big(g_{i,j}(A_{i,j},X_{i,j},t)\hat{p}_{\boldsymbol{n}}\big)}{\partial X_{i,j}}\\ &\qquad=\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma_{i,j}^2(A_{i,j},X_{i,j},t)\hat{p}_{\boldsymbol{n}}\big)}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(A_{i,j},X_{i,j})+\mu_{i,j}(A_{i,j},X_{i,j})\big)\hat{p}_{\boldsymbol{n}}\\ &\hat{p}_{\boldsymbol{n}}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},0\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)=\delta\big(\boldsymbol{X}(0)_{\boldsymbol{n}(0)}-\boldsymbol{X}_{\boldsymbol{n}}\big)\delta\big(\boldsymbol{A}(0)_{\boldsymbol{n}(0)}-\boldsymbol{A}_{\boldsymbol{n}}\big),\quad\text{if }\boldsymbol{n}=\boldsymbol{n}(0),\\ &\hat{p}_{\boldsymbol{n}}(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},0)=0\quad\text{if }\boldsymbol{n}\neq\boldsymbol{n}(0)\end{aligned}\tag{C1}$$

can be expressed as

$$\begin{aligned}&\hat{p}_{\boldsymbol{n}}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\\ &\quad\coloneqq\mathbb{E}\big[\delta\big(\boldsymbol{X}(t)_{\boldsymbol{n}(t)}-\boldsymbol{X}_{\boldsymbol{n}}\big)\delta\big(\boldsymbol{A}(t)_{\boldsymbol{n}(t)}-\boldsymbol{A}_{\boldsymbol{n}}\big)S_A\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)},\boldsymbol{A}(t)_{\boldsymbol{n}(t)}\big)\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant s\leqslant t)=\boldsymbol{n}(0)\big],\\ &\hat{p}_{\boldsymbol{n}}(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t)=0,\quad\text{if }\boldsymbol{n}\neq\boldsymbol{n}(0),\end{aligned}\tag{C2}$$

where here

$$S_A\big(t;\boldsymbol{X}(t)_{\boldsymbol{n}(t)},\boldsymbol{A}(t)_{\boldsymbol{n}(t)}\big)\coloneqq\exp\left(-\int_0^t\sum_{i=1}^{k(0)}\sum_{j=1}^{n_i(0)}\big(\beta_{i,j}(A_{i,j}(\tau),X_{i,j}(\tau))+\mu_{i,j}(A_{i,j}(\tau),X_{i,j}(\tau))\big)\mathrm{d}\tau\right).\tag{C3}$$

Furthermore, if we set

$$\begin{aligned}&p^0\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)=0,\\ &p^1\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)=\hat{p}_{\boldsymbol{n}}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big),\end{aligned}\tag{C4}$$

then we can define the recursion

$$\begin{aligned}&p_{\boldsymbol{n}}^{m+1}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)=\hat{p}_{\boldsymbol{n}}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\\ &\quad+\mathbb{E}\left[\int_0^t S_A\big(\tau;\boldsymbol{X}(\tau)_{\boldsymbol{n}(\tau)},\boldsymbol{A}(\tau)_{\boldsymbol{n}(\tau)}\big)J_A^m\big(t,\tau;\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{n}(0)\big)\mathrm{d}\tau\middle|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0;\boldsymbol{n}(0\leqslant\tau\leqslant t)=\boldsymbol{n}(0)\right],\quad\text{if }\boldsymbol{A}_{\boldsymbol{n}}>0,\end{aligned}$$

$$p_{\boldsymbol{n}}^{m+1}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\,\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)$$
$$=\mathbb{E}\Big[\sum_{j=1}^{n_i}\tilde{\beta}_{i,j}\big(A_{i,j},Y(t),X_{i,j}(t),X_{i+1,n_{i+1}}\big)p_{\boldsymbol{n}}^{m}\big(\boldsymbol{A}^{j}_{\boldsymbol{n}_{\mathrm{b},i}},\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{b},i}},t\,\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\Big],\quad \text{if } A_{i+1,n_{i+1}}=0,$$
$$p_{\boldsymbol{n}}^{m+1}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\,\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)=0,\quad \text{otherwise.} \tag{C5}$$

Here $\boldsymbol{A}_{\boldsymbol{n}}>0$ indicates that each component in $\boldsymbol{A}_i$ of $\boldsymbol{A}_{\boldsymbol{n}}$ is greater than 0. $\tilde{\beta}_{i,j}(A_{i,j},Y(t),X_{i,j},X_{i+1,n_{i+1}})$ is the rate of a cell in the ith generation giving birth to a cell in the $(i+1)$th generation with the state $X_{i+1,n_{i+1}}$ and its own state shifting to $X_{i,j}$. $\boldsymbol{A}^{j}_{\boldsymbol{n}_{\mathrm{b},i}}$ differs from $\boldsymbol{A}_{\boldsymbol{n}}$ in that its $(i+1)$th generation does not contain the (n_{i+1})th component. $\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{b},i}}$ differs from $\boldsymbol{X}_{\boldsymbol{n}}$ in that its jth component in the ith generation is $Y_{i,j}(t)$, not $X_{i,j}$ and it does not have the (n_{i+1})th component in the $(i+1)$th generation. In analogy to Eq. (A5), $J_A(t,\tau;\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}})$ in Eq. (C5) is defined as

$$J_A^m(t,\tau;\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{n}(0)):=\sum_{i=1}^{k(0)}\sum_{j=1}^{n_i(0)}\Big(\tilde{\beta}_{i,j}\big(A_{i,j}(\tau),X_{i,j}(\tau),X_1(\tau),X_2(\tau)\big)p_{\boldsymbol{n}}^{m}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t-\tau\,\big|\boldsymbol{A}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}},\boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}},0\big)$$
$$+\,\mu_{i,j}\big(A_{i,j}(\tau),X_{i,j}(\tau)\big)p_{\boldsymbol{n}}^{m}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t-\tau\,\big|\boldsymbol{A}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}},\boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}},0\big)\Big). \tag{C6}$$

In Eq. (C6), $\boldsymbol{A}^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}$ differs from $\boldsymbol{A}_{\boldsymbol{n}(0)}$ in that its $(i+1)$th generation has an extra component $A_{i+1,n_{i+1}(0)+1}=0$. $\boldsymbol{X}^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}(\tau)$ is different from $\boldsymbol{X}_{\boldsymbol{n}(0)}(\tau)$ in that compared to $\boldsymbol{X}_{\boldsymbol{n}(0)}(\tau)$, the jth component of the ith generation of $\boldsymbol{X}^{-j}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}(\tau)$ is $Y_{i,j}(\tau)$ in the jth, but the jth component of the ith generation is $X_{i,j}(\tau)$ for $\boldsymbol{X}_{\boldsymbol{n}(0)}(\tau)$; furthermore, the $(i+1)$th generation of $\boldsymbol{X}_{\boldsymbol{n}(0)}(\tau)$ does not have the $(n_{i+1}+1)$th component $X_{i+1,n_{i+1}(0)+1}(\tau)$. $\boldsymbol{A}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}}$ differs from $\boldsymbol{A}(\tau)_{\boldsymbol{n}}$ in that its ith generation is $(A_{i,1}(\tau),\ldots,A_{i,j-1}(\tau),A_{i,j+1}(\tau),\ldots,A_{i,n_i(0)}(\tau))$, and $\boldsymbol{X}(\tau)^{-j}_{\boldsymbol{n}(0)_{\mathrm{d},-i}}$ differs from $\boldsymbol{X}(\tau)_{\boldsymbol{n}}$ in that its ith generation is $(X_{i,1}(\tau),\ldots,X_{i,j-1}(\tau),X_{i,j+1}(\tau),\ldots,X_{i,n_i(0)}(\tau))$.

Then, similar to the proof of Proposition 2, $p_{\boldsymbol{n}}^{m+1}$ satisfies the following PIDE

$$\frac{\partial p_{\boldsymbol{n}}^{m+1}(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t)}{\partial t}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial p_{\boldsymbol{n}}^{m+1}}{\partial A_{i,j}}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial\big(g_{i,j}(A_{i,j},X_{i,j},t)p_{\boldsymbol{n}}^{m+1}\big)}{\partial X_{i,j}}$$
$$=\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma_{i,j}^2(A_{i,j},X_{i,j},t)p_{\boldsymbol{n}}^{m+1}\big)}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(A_{i,j},X_{i,j})+\mu_{i,j}(A_{i,j},X_{i,j})\big)p_{\boldsymbol{n}}^{m+1}$$
$$+\sum_{i=1}^{\infty}\sum_{j=1}^{n_i^{\mathrm{d}}}\int\mu_{i,j}(A,Y)p_{\boldsymbol{n}_{\mathrm{d},i}}^{m}\big(\boldsymbol{A}^{j}_{\boldsymbol{n}_{\mathrm{d},i}},\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{d},i}},t\,\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}Y\,\mathrm{d}A,\quad \text{if } \boldsymbol{A}_{\boldsymbol{n}}>0$$
$$p_{\boldsymbol{n}}^{m+1}\big(\boldsymbol{X}_{\boldsymbol{n}},\boldsymbol{A}_{\boldsymbol{n}},t\,\big|\boldsymbol{X}(0)_{\boldsymbol{n}(0)},\boldsymbol{A}(0)_{\boldsymbol{n}(0)},0\big)$$
$$=\int\sum_{i=1}^{k-1}\sum_{j=1}^{n_i}\tilde{\beta}_{i,j}\big(A_{i,j},Y_{i,j},X_{i,j},X_{i+1,n_{i+1}}\big)p_{\boldsymbol{n}}^{m}\big(\boldsymbol{A}^{j}_{\boldsymbol{n}(0)_{\mathrm{b},i}}(t),\boldsymbol{X}^{j}_{\boldsymbol{n}(0)_{\mathrm{b},i}}(t),t\,\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}Y_{i,j},\quad \text{if } A_{i+1,n_{i+1}}=0. \tag{C7}$$

Likewise, it can be shown that $p_{\boldsymbol{n}}^{m}$ is non-negative, increasing in m, and satisfies

$$\sum_{\boldsymbol{n}}\int p_{\boldsymbol{n}}^{m}\big(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t\,\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\mathrm{d}\boldsymbol{A}_{\boldsymbol{n}}\leqslant 1,\quad \forall\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)}. \tag{C8}$$

Therefore, under certain technical conditions such as commuting derivatives, there exists a limit $p_{\boldsymbol{n}}^{*}=\lim_{m\to\infty}p_{\boldsymbol{n}}^{m}$ that satisfies the PIDE

$$\frac{\partial p_{\boldsymbol{n}}^{*}(\boldsymbol{A}_{\boldsymbol{n}},\boldsymbol{X}_{\boldsymbol{n}},t)}{\partial t}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial p_{\boldsymbol{n}}^{*}}{\partial A_{i,j}}+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial\big(g_{i,j}(A_{i,j},X_{i,j},t)p_{\boldsymbol{n}}^{*}\big)}{\partial X_{i,j}}$$
$$=\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\partial^2\big(\sigma_{i,j}^2(A_{i,j},X_{i,j},t)p_{\boldsymbol{n}}^{*}\big)}{(\partial X_{i,j})^2}-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\big(\beta_{i,j}(A_{i,j},X_{i,j})+\mu_{i,j}(A_{i,j},X_{i,j})\big)p_{\boldsymbol{n}}^{*}$$
$$+\sum_{i=1}^{\infty}\sum_{j=1}^{n_i^{\mathrm{d}}}\int\mu_{i,j}(A,Y)p_{\boldsymbol{n}_{\mathrm{d},i}}^{*}\big(\boldsymbol{A}^{j}_{\boldsymbol{n}_{\mathrm{d},i}},\boldsymbol{X}^{j}_{\boldsymbol{n}_{\mathrm{d},i}},t\,\big|\boldsymbol{A}(0)_{\boldsymbol{n}(0)},\boldsymbol{X}(0)_{\boldsymbol{n}(0)},0\big)\mathrm{d}Y\,\mathrm{d}A\,,\quad \text{if}\,\boldsymbol{A}_{\boldsymbol{n}}>0, \tag{C9}$$

$$p_{\boldsymbol{n}}^{*}\left(\boldsymbol{A}_{\boldsymbol{n}}, \boldsymbol{X}_{\boldsymbol{n}}, t \middle| \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, \boldsymbol{A}(0)_{\boldsymbol{n}(0)}, 0\right)$$
$$= \int \sum_{i=1}^{k-1} \sum_{j=1}^{n_i} \tilde{\beta}_{i,j}\left(A_{i,j}, Y_{i,j}, X_{i,j}, X_{i+1,n_{i+1}}\right) p_{\boldsymbol{n}}^{*}\left(\boldsymbol{A}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}^{-j}, \boldsymbol{X}_{\boldsymbol{n}(0)_{\mathrm{b},-i}}^{-j}, t \middle| \boldsymbol{A}(0)_{\boldsymbol{n}(0)}, \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\right) \mathrm{d}Y_{i,j}, \quad \text{if } A_{i+1,n_{i+1}} = 0.$$

If p^* satisfies the normalization conditions, i.e., $\sum_{\boldsymbol{n}} \int p_{\boldsymbol{n}}^{*}(\boldsymbol{A}_{\boldsymbol{n}}, \boldsymbol{X}_{\boldsymbol{n}}, t)\mathrm{d}\boldsymbol{X}_{\boldsymbol{n}}\mathrm{d}\boldsymbol{A}_{\boldsymbol{n}} \equiv 1, \forall t \geqslant 0$, then we can also define the unconditional probability density by averaging over the initial probability density $q_{\boldsymbol{n}(0)}(\boldsymbol{A}(0)_{\boldsymbol{n}(0)}, \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0)$,

$$p_{\boldsymbol{n}}(\boldsymbol{A}_{\boldsymbol{n}}, \boldsymbol{X}_{\boldsymbol{n}}, t) := \sum_{\boldsymbol{n}(0)} \int p_{\boldsymbol{n}}^{*}\left(\boldsymbol{A}_{\boldsymbol{n}}, \boldsymbol{X}_{\boldsymbol{n}}, t \middle| \boldsymbol{A}(0)_{\boldsymbol{n}(0)}, \boldsymbol{X}(0)_{\boldsymbol{n}(0)}, 0\right) q_{\boldsymbol{n}(0)}\left(\boldsymbol{A}(0)_{\boldsymbol{n}(0)}, \boldsymbol{X}_{\boldsymbol{n}(0)}, 0\right) \mathrm{d}\boldsymbol{X}(0)_{\boldsymbol{n}(0)}\mathrm{d}\boldsymbol{A}(0)_{\boldsymbol{n}(0)}. \tag{C10}$$

From Eq. (C10), we can define the symmetric probability density function,

$$\rho_{\boldsymbol{n}}(\boldsymbol{A}_{\boldsymbol{n}}, \boldsymbol{X}_{\boldsymbol{n}}, t) := \prod_{i=1}^{k} \frac{1}{n_i!} \sum_{\pi} p_{\boldsymbol{n}}^{*}(\pi(\boldsymbol{A}_{\boldsymbol{n}}), \pi(\boldsymbol{X}_{\boldsymbol{n}}), t), \tag{C11}$$

where π is the same rearrangement for the age variables $\boldsymbol{A}_{\boldsymbol{n}}$ and state variables $\boldsymbol{X}_{\boldsymbol{n}}$. From Eq. (C11), we could derive the macroscopic quantities such as the marginalized cell density. We shall omit detailed discussions on those macroscopic quantities for brevity.

[1] A. G. M'Kendrick, Applications of mathematics to medical problems, Proc. Edinb. Math. Soc. **44**, 98 (1925).

[2] H. von Foerster, Some remarks on changing populations, in *The Kinetics of Cellular Proliferation*, edited by F. Stohlman, Jr. (Grune & Stratton, New York, 1959), pp. 382–407.

[3] B. L. Keyfitz and N. Keyfitz, The McKendrick partial differential equation and its uses in epidemiology and population study, Math. Comput. Model. **26**, 1 (1997).

[4] A. Ponosov, L. Idels, and R. Kadiev, Stochastic McKendrick–Von Foerster models with applications, Physica A **537**, 122641 (2020).

[5] Y. Wang, R. Dessalles, and T. Chou, Modelling the impact of birth control policies on China's population and age: Effects of delayed births and minimum birth age constraints, R. Soc. Open Sci. **9**, 211619 (2022).

[6] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun, Cell-size control and homeostasis in bacteria, Curr. Biol. **25**, 385 (2015).

[7] S. Burov and D. Kessler, Effective potential for cellular size-control, arXiv:1701.01725.

[8] L. Robert, M. Hoffmann, N. Krell, S. Aymerich, J. Robert, and M. Doumic, Division in *Escherichia coli* is triggered by a size-sensing rather than a timing mechanism, BMC Biol. **12**, 17 (2014).

[9] B. Perthame, *Introduction to Structured Equations in Biology*, (CNA Summer School Lecture Notes, Paris, France, 2023).

[10] J. A. J. Metz and O. Diekmann, *The Dynamics of Physiologically Structured Populations* (Springer, Berlin, 1986).

[11] C. D. Greenman and T. Chou, Kinetic theory of age-structured stochastic birth-death processes, Phys. Rev. E **93**, 012112 (2016).

[12] T. Chou and C. D. Greenman, A hierarchical kinetic theory of birth, death and fission in age-structured interacting populations, J. Stat. Phys. **164**, 49 (2016).

[13] M. Xia and T. Chou, Kinetic theory for structured populations: Application to stochastic sizer-timer models of cell proliferation, J. Phys. A: Math. Theor. **54**, 385601 (2021).

[14] J. C. Kynaston, C. Guiver, and C. A. Yates, Equivalence framework for an age-structured multistage representation of the cell cycle, Phys. Rev. E **105**, 064411 (2022).

[15] P.-Y. Ho, J. Lin, and A. Amir, Modeling cell size regulation: From single-cell-level statistics to molecular mechanisms and population-level effects, Annu. Rev. Biophys. **47**, 251 (2018).

[16] G. La Manno, R. Soldatov, A. Zeisel, E. Braun, H. Hochgerner, V. Petukhov, K. Lidschreiber, M. E. Kastriti, P. Lönnerberg, A. Furlan *et al.*, RNA velocity of single cells, Nature (London) **560**, 494 (2018).

[17] X. Qiu, Y. Zhang, J. D. Martin-Rufino, C. Weng, S. Hosseinzadeh, D. Yang, A. N. Pogson, M. Y. Hein, K. H. Joseph Min, L. Wang *et al.*, Mapping transcriptomic vector fields of single cells, Cell **185**, 690 (2022).

[18] G. Gorin, V. Svensson, and L. Pachter, Protein velocity and acceleration from single-cell multiomics experiments, Genome Biol. **21**, 39 (2020).

[19] J. Held, T. Lorimer, F. Pomati, R. Stoop, and C. Albert, Second-order phase transition in phytoplankton trait dynamics, Chaos **30**, 053109 (2020).

[20] S. Bhattacharya, Q. Zhang, and M. E. Andersen, A deterministic map of Waddington's epigenetic landscape for cell fate specification, BMC Syst. Biol. **5**, 85 (2011).

[21] J. Wang, K. Zhang, L. Xu, and E. Wang, Quantifying the Waddington landscape and biological paths for development and differentiation, Proc. Natl. Acad. Sci. USA **108**, 8257 (2011).

[22] F. Peng, M. Liao, R. Qin, S. Zhu, C. Peng, L. Fu, Y. Chen, and B. Han, Regulated cell death (RCD) in cancer: Key pathways and targeted therapies, Sign. Transduct. Target. Ther. **7**, 286 (2022).

[23] W. F. Marzluff and R. J. Duronio, Histone mRNA expression: Multiple levels of cell cycle regulation and important developmental consequences, Curr. Opin. Cell Biol. **14**, 692 (2002).

[24] N. Heintz, H. L. Sive, and R. G. Roeder, Regulation of human histone gene expression: Kinetics of accumulation and changes

in the rate of synthesis and in the half-lives of individual histone mRNAs during the HeLa cell cycle, Mol. Cell. Biol. **3**, 539 (1983).

[25] M. Xia, C. D. Greenman, and T. Chou, PDE models of adder mechanisms in cellular proliferation, SIAM J. Appl. Math. **80**, 1307 (2020).

[26] C. Buccitelli and M. Selbach, mRNAs, proteins and the emerging principles of gene expression control, Nat. Rev. Genet. **21**, 630 (2020).

[27] C. Xia, J. Fan, G. Emanuel, J. Hao, and X. Zhuang, Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression, Proc. Natl. Acad. Sci. USA **116**, 19490 (2019).

[28] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, 2009).

[29] M. Coomer, L. Ham, and M. P. H. Stumpf, Noise distorts the epigenetic landscape and shapes cell-fate decisions, Cell Syst. **13**, 83 (2022).

[30] J. Wang, Landscape and flux theory of non-equilibrium dynamical systems with application to biology, Adv. Phys. **64**, 1 (2015).

[31] C. D. Greenman, A path integral approach to age dependent branching processes, J. Stat. Mech. (2017) 033101.

[32] H. Caswell, *Matrix Population Models* (Sinauer, Sunderland, MA, 2000), Vol. 1.

[33] M. Xia, X. Li, and T. Chou, Overcompensation of transient and permanent death rate increases in age-structured models with cannibalistic interactions, Physica D **470**, 134339 (2024).

[34] A. Le Cavil, N. Oudjane, and F. Russo, Probabilistic representation of a class of non conservative nonlinear Partial Differential Equations, Latin Am. J. Probab. Math. Stat. **13**, 1189 (2016).

[35] C. Jia, A. Singh, and R. Grima, Characterizing non-exponential growth and bimodal cell size distributions in fission yeast: An analytical approach, PLoS Comput. Biol. **18**, e1009793 (2022).

[36] C. Jia, A. Singh, and R. Grima, Concentration fluctuations in growing and dividing cells: Insights into the emergence of concentration homeostasis, PLoS Comput. Biol. **18**, e1010574 (2022).

[37] R. Dessalles, M. D'Orsogna, and T. Chou, Exact steady-state distributions of multispecies birth-death-immigration processes: Effects of mutations and carrying capacity on diversity, J. Stat. Phys. **173**, 182 (2018).

[38] R. Lande, Genetics and demography in biological conservation, Science **241**, 1455 (1988).

[39] R. Lande, S. Engen, and B.-E. Saether, *Stochastic Population Dynamics in Ecology and Conservation* (Oxford University Press, New York, 2003).

[40] M. Xia, S. Shao, and T. Chou, Efficient scaling and moving techniques for spectral methods in unbounded domains, SIAM J. Sci. Comput. **43**, A3244 (2021).

[41] M. Xia, S. Shao, and T. Chou, A frequency-dependent p-adaptive technique for spectral methods, J. Comput. Phys. **446**, 110627 (2021).

[42] T. Chou, S. Shao, and M. Xia, Adaptive Hermite spectral methods in unbounded domains, Appl. Numer. Math. **183**, 201 (2023).

[43] M. Xia, X. Li, Q. Shen, and T. Chou, Squared Wasserstein-2 distance for efficient reconstruction of stochastic differential equations, arXiv:2401.11354.

[44] M. Xia, X. Li, Q. Shen, and T. Chou, An efficient Wasserstein-distance approach for reconstructing jump-diffusion processes using parameterized neural networks, Mach. Learn.: Sci. Technol. **5**, 045052 (2024).

[45] N. Thuerey, P. Holl, M. Mueller, P. Schnell, F. Trost, and K. Um, *Physics-based Deep Learning* (WWW, 2021).

[46] M. Xia, L. Böttcher, and T. Chou, Spectrally adapted physics-informed neural networks for solving unbounded domain problems, Mach. Learn.: Sci. Technol. **4**, 025024 (2023).

[47] M. Xia, X. Li, Q. Shen, and T. Chou, Learning unbounded-domain spatiotemporal differential equations using adaptive spectral methods, J. Appl. Math. Comput. **70**, 4395 (2024).

[48] R. Huang, Q. Situ, and J. Lei, Dynamics of cell-type transition mediated by epigenetic modifications, J. Theor. Biol. **577**, 111664 (2024).

[49] B. P. Ayati and I. Klapper, A multiscale model of biofilm as a senescence-structured fluid, Multiscale Model. Simul. **6**, 347 (2007).

[50] P. Auger, P. Magal, and S. Ruan, Aggregation of variables and applications to population dynamics, in *Structured Population Models in Biology and Epidemiology*, Pierre Magal and Shigui Ruan (Springer, Berlin, 2008), pp. 209–263.

[51] C. D. Nadell, K. Drescher, and K. R. Foster, Spatial structure, cooperation and competition in biofilms, Nat. Rev. Microbiol. **14**, 589 (2016).

[52] C. Jia, M. Q. Zhang, and H. Qian, Emergent Lévy behavior in single-cell stochastic gene expression, Phys. Rev. E **96**, 040402(R) (2017).

[53] C. Jia, L. Y. Wang, G. G. Yin, and M. Q. Zhang, Single-cell stochastic gene expression kinetics with coupled positive-plus-negative feedback, Phys. Rev. E **100**, 052406 (2019).