Incorporating Stochastic Gene Expression, Signaling-Mediated Cell-Cell Interactions, and Regulated Cell Proliferation in Models of Coordinated Tissue Development

Casey O. Barkan [*]
Department of Physics & Astronomy, University of California, Los Angeles, California 90095-1547, USA

Tom Chou [†]
Department of Biomathematics, University of California, Los Angeles, California 90095-1766, USA
and Department of Mathematics, University of California, Los Angeles, California 90095-1555, USA

(Received 31 December 2024; accepted 1 April 2026; published 8 May 2026)

Formulating quantitative and predictive models for tissue development requires consideration of complex, stochastic gene expression dynamics, their regulation via cell-cell interactions, and cell proliferation. Including all of these processes in a practical mathematical framework requires complex expressions that are difficult to apply. We construct a theory that incorporates intracellular stochastic gene expression dynamics, signaling chemicals that influence these dynamics and mediate cell-cell interactions, and cell proliferation and its accompanying differentiation. The dynamics of cellular states are described by a Waddington vector field, where no *landscape* assumption is made. We provide a procedure to systematically simplify the theory by applying explicit assumptions—this can reduce the theory to a master equation or even to a system of ordinary differential equations. We define an epigenetic fitness landscape that describes the proliferation of different cell types and elucidate how this fitness landscape is related to Waddington's vector field. We illustrate the applicability of our framework by analyzing two model systems: an interacting two-gene differentiation process and a spatiotemporal organism model inspired by planaria.

DOI: 10.1103/thys-1k71

I. INTRODUCTION

Cell and tissue development is a complex and multifaceted dynamical process involving gene regulatory dynamics, cell-cell interactions, and variable cell proliferation and death rates. The conceptual model of Waddington's landscape has been a guiding paradigm for understanding the mechanisms of cell development [1–3]; it proposes that the internal dynamics of a differentiating cell are driven by gradients of a rugged high-dimensional "potential" landscape. This gradient dynamics picture of Waddington has been translated into several mathematical models that describe single-cell gene expression states [3–9] and population level phenotypes [10,11]. In the typical single-cell interpretation, a cell's state is modeled as a passive particle with position $\boldsymbol{x}$ in a potential energy landscape $U(\boldsymbol{x})$ in a noisy environment. If the noise acts as an equilibrium heat bath (obeying an equilibrium fluctuation dissipation relation), then this model implies that one can reconstruct $U(\boldsymbol{x})$ from data using the simple relationship $U(\boldsymbol{x}) \propto -\log p_{\mathrm{ss}}(\boldsymbol{x})$, where $p_{\mathrm{ss}}(\boldsymbol{x})$ is the empirically measured steady-state distribution across state space. This formula has been applied to real datasets [12], but its use is limited to single cells (or to noninteracting and nonproliferating populations) and an assumption that the cell state behaves as a passive particle in an equilibrium environment that typically does not hold in biological processes. For example, most models of gene regulation, including mass-action models of autocatalytic feedback circuits [13], genetic toggle switches [14], repressilators [15], and rock-paper-scissors type of dynamics [16] do not behave like a passive particle in an equilibrium environment.

Besides the internal dynamics of cell states (e.g., gene expression profiles), proliferation and death of individual cells govern population dynamics akin to Darwinian evolution [17,18]. Additionally, intracellular gene dynamics and population-level birth-death processes are coupled through cell-cell signaling and metabolic interactions [9,18]. Thus, a complete description of development must describe an interacting, varying population of cells, not just the gene dynamics within a single cell.

When cell division and death are considered, the concept of a fitness landscape also becomes critical. In evolutionary population dynamics, high fitness genotypes grow to dominate the population. A similar phenomenon can occur during development, though the definition of the fitness landscape must be adapted to this context in which all cells have the same genotype but differ in their internal state, and where cell-cell interactions are crucial to the population dynamics. While mathematical models of developmental fate and state changes often consider only single-cells [3,5,12,19–27], approaches for reconstructing landscapes in the presence of

[*]Contact author: barkanc@ucla.edu
[†]Contact author: tomchou@ucla.edu

chemical signaling-mediated interactions have been recently developed [28,29]. However, models that explicitly incorporate cell population dynamics typically neglect cell-cell interactions [30–35].

In more fine-grained kinetic theories that combine internal cell state dynamics with state-dependent division and death rates, tractable dynamical equations describing population-level stochastic quantities can be derived if there are no direct cell-cell interactions [36–39]. However, in developing tissues, cell division and death are tightly coordinated through cell-cell interactions mediated by secreted chemical signals that influence gene expression in other cells [40,41]. Failure of this regulated tissue development can lead to pathologies such as cancer [42,43]. Thus, effective theoretical approaches that incorporate regulated intracellular gene expression with population-level birth-death processes will provide critical tools for modeling and analyzing tissue development.

Here we present a procedure for modeling developmental processes in which cell-cell interactions regulate gene expression and cell proliferation and in which spatial organization develops. First, we relax the assumption that the cell state dynamics follow the gradient of a potential landscape, replacing the Waddington *landscape* with a Waddington *vector field*. This is necessary for properly modeling the active processes of gene expression that generally violate detailed balance. Next, we describe a population of individual (discrete) cells, rather than assuming an infinite continuum of cells, as has been done in other works [9,18,32,33]. Our approach provides an important connection between intracellular state dynamics and cell division. The state of daughter cells immediately after birth can be different from that of the mother cell. State changes following cell division can be described by a state-dependent differential birth rate function which can be used to construct an effective *epigenetic fitness landscape*. Although a fitness landscape describes differential proliferation of different genotypes, our epigenetic fitness landscape describes the differential proliferation of cells of the same genotype but differing cell state, capturing the fact that different attractors of the gene regulatory dynamics induce different cell division and death rates.

Finally, we show how our general model can be reduced to simpler forms by applying common assumptions. Often, such assumptions are implicitly included in models of development without a clear delineation of regimes of validity. Our general framework provides a starting point for deriving simpler, interpretable models by applying explicit assumptions. Using this framework, we present two minimal models of the robust development of a stable tissue population as proof-of-concept examples. The first is a well-mixed model of gene-regulated stem-cell differentiation; the second is a spatiotemporal model of regeneration in organisms such as planaria.

II. MATHEMATICAL FRAMEWORK

Our framework is composed of several coupled physical processes, described as follows.

A. Intracellular state dynamics

Let $\boldsymbol{z} \in \mathbb{R}^{d_z}$ denote the internal state of a cell such as its gene expression, methylation levels of DNA, chromatin structure, spatial organization of organelles, and/or any other relevant intracellular variables. While a comprehensive description of the internal state would require a very large d_z, it can be productive to study minimal models where d_z is small and $\boldsymbol{z}$ is an abstract representation of the cell state. We label cells by a superscript α, i.e., $\boldsymbol{z}^\alpha$ is the state of cell α, where $\alpha = 1, \ldots, N(t)$, and $N(t)$ is the total number of cells which varies due to cell division and death.[1]

Interactions between spatially separated cells are largely mediated through chemical signaling molecules. Suppose there are L relevant signaling molecule species and let $c_\ell(\boldsymbol{r}, t)$ denote the concentration of species ℓ ($\ell = 1, \ldots, L$) at position $\boldsymbol{r} \in \mathbb{R}^3$ at time t. For notational convenience, let $\boldsymbol{c}(\boldsymbol{r}, t) = (c_1(\boldsymbol{r}, t), \ldots, c_L(\boldsymbol{r}, t))$ denote the list of signaling molecule concentrations. For cell α at position $\boldsymbol{r}^\alpha$, the dynamics of its state $\boldsymbol{z}^\alpha(t)$ are modeled by a stochastic differential equation (SDE) in the Itô convention [44]:

$$\mathrm{d}\boldsymbol{z}^\alpha = \boldsymbol{F}(\boldsymbol{z}^\alpha; \boldsymbol{c}^\alpha)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{z}^\alpha; \boldsymbol{c}^\alpha)\mathrm{d}\boldsymbol{B}^\alpha, \tag{1}$$

where $\boldsymbol{c}^\alpha \equiv \boldsymbol{c}(\boldsymbol{r}^\alpha, t)$ are the signaling molecule concentrations experienced by cell α at time t. $\boldsymbol{F}(\boldsymbol{z}; \boldsymbol{c})$ defines a $\boldsymbol{c}$-dependent vector field (mapping $\mathbb{R}^{d_z} \to \mathbb{R}^{d_z}$) which specifies the expected rate of change of $\boldsymbol{z}^\alpha$. We call $\boldsymbol{F}(\boldsymbol{z}; \boldsymbol{c})$*Waddington's vector field*, and it is presumed to have stable attractors corresponding to stable cell types. $\boldsymbol{F}$ could describe a specific gene regulatory network, or it could be a minimal abstract model of cell state dynamics. In Sec. IV, we compare the *Waddington's vector field* picture to the *Waddington's landscape* picture (where $\boldsymbol{F}$ is assumed to be the gradient of a potential) and explain why we consider the Waddington vector field to be the appropriate modeling choice. Noise is characterized by the d_b-dimensional Brownian motion $\mathrm{d}\boldsymbol{B}^\alpha$ and $\boldsymbol{\sigma}(\boldsymbol{z}; \boldsymbol{c})$, a $d_z \times d_b$ matrix which may depend on the cell state and signaling molecule concentrations at the cell's location.

Importantly, the signaling molecule concentration $\boldsymbol{c}^\alpha$ influences the dynamics of the state $\boldsymbol{z}^\alpha$ and vice versa, the dynamics of the concentrations $\boldsymbol{c}(\boldsymbol{r}, t)$ are governed by the states and positions of all cells (as described in the following section). Hence, the signaling molecules can mediate cell-cell interactions, allowing the tissue to develop in a coordinated way.

By describing cell state dynamics with an SDE, we assume that the cell state can vary continuously. This prevents the model from capturing discrete properties, such as methylation at specific DNA sites, though a discrete component to the state space could be included if deemed relevant for a particular system. Additionally, protein, mRNA, and signaling molecule counts are fundamentally discrete; our continuous description assumes large counts, which is often a very good approximation.[2]

[1]The labeling of cells is arbitrary. One scheme labels cells from oldest to youngest, i.e., cell $\alpha = 1$ is the oldest cell. When a cell dies, all cells younger than the dying cell are then relabeled.

[2]SDE/Fokker-Planck approximations to discrete reactions often accurately predict population means and variances even at modest concentrations; however, their predictions for extinction times can deviate from those of the discrete process [45].

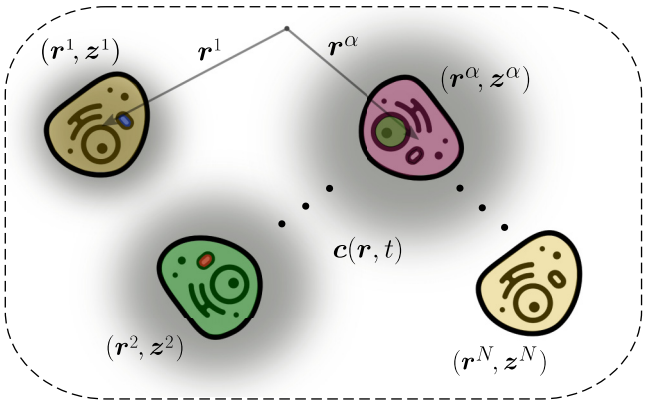

FIG. 1. Schematic of cells at various positions $\boldsymbol{r}^\alpha$ and in various gene expression states $\boldsymbol{z}^\alpha$, $\alpha = 1, 2, \ldots, N$. Cell-cell interactions mediated by position-dependent signaling molecule concentrations $\boldsymbol{c}(\boldsymbol{r}, t) \equiv (c_1(\boldsymbol{r}, t), \ldots, c_L(\boldsymbol{r}, t))$ which are represented in gray scale. Cells can produce signaling molecules at rates that depend on $\boldsymbol{z}$; gene expression is in turn influenced by the concentrations that a cell senses, leading to potentially complex cell-cell interactions.

B. Cell-cell interactions through signaling molecules

In our framework, cell-cell interactions are mediated by signaling molecules that are excreted by cells and that diffuse through the tissue environment. Interactions via direct membrane-membrane contact (e.g., juxtacrine signaling) are not modeled explicitly but can be incorporated via an approximation discussed below.

The dynamics of signaling molecule concentrations can be modeled by any variety of molecular transport models. For example, a deterministic PDE describing excretion, absorption, and diffusion takes the typical form [46]

$$\partial_t c_\ell(\boldsymbol{r}, t) = D_\ell \nabla^2 c_\ell - d_\ell c_\ell + \sum_{\text{cells } \alpha} \lambda_\ell(\boldsymbol{r} - \boldsymbol{r}^\alpha, \boldsymbol{z}^\alpha), \qquad (2)$$

where $\lambda_\ell(\boldsymbol{r} - \boldsymbol{r}^\alpha, \boldsymbol{z}^\alpha)$ is the excretion rate of signaling molecule ℓ at position $\boldsymbol{r}$ from a cell in state $\boldsymbol{z}^\alpha$ centered at $\boldsymbol{r}^\alpha$, and the sum runs over all cells. The function $\lambda_\ell(\boldsymbol{r} - \boldsymbol{r}^\alpha, \boldsymbol{z}^\alpha)$ is localized about $\boldsymbol{r}^\alpha$ and reflects the finite size of the cells. Diffusion and degradation of molecule ℓ is described by the coefficients D_ℓ and d_ℓ, respectively. A natural choice of boundary condition for Eq. (2) is $c_\ell(\boldsymbol{r}) \to 0$ as $|\boldsymbol{r}| \to \infty$. Alternatively, to model subcompartments or spatially defined microenvironments, Eq. (2) can be considered in finite regions of space each with appropriate boundary conditions. For a closed environment, one would use reflecting boundary conditions for $c_\ell(\boldsymbol{r})$. Equation (2) neglects reactions amongst the signaling molecules and models deterministic, locally averaged concentration dynamics, appropriate for molecular transport at high concentrations and/or fast molecular timescales. At very low signaling molecule densities, intrinsic molecular stochasticity may be important and any number of stochastic models [47–49] and simulation approaches [50,51] are available. Figure 1 provides a schematic of signaling-molecule-mediated interactions among cells at positions $\boldsymbol{r}^\alpha$, with potentially different internal states $\boldsymbol{z}^\alpha$.

Interactions via direct membrane-membrane contact cannot be incorporated in an exact way because our framework does not include cell shape. However, such interactions can be approximated through the assumption that the degree of membrane-membrane contact depends only on the distance between cells. More specifically, suppose cell α_1 receives a signaling molecule of type ℓ (e.g., notch intracellular domain, in the case of the notch-delta pathway [52,53]) from cell α_2 due to direct membrane-membrane contact. Now assume that the intracellular concentration of ℓ in cell α_1 is a function only of the distance $r = |\boldsymbol{r}^{\alpha_1} - \boldsymbol{r}^{\alpha_2}|$ and the state $\boldsymbol{z}^{\alpha_2}$. This function, which we denote $m_\ell(r, \boldsymbol{z}^{\alpha_2})$, would be presumed to be zero for large r when the cells are not in communicating contact. Under this assumption,

$$c_\ell(\boldsymbol{r}, t) = \sum_{\text{cells } \alpha} m_\ell(|\boldsymbol{r} - \boldsymbol{r}^\alpha|, \boldsymbol{z}^\alpha). \qquad (3)$$

Note that $m_\ell(0, \boldsymbol{z}) = 0$ to avoid self-interaction. As a matter of interpretation, Eq. (3) should be understood as specifying intracellular concentrations only, not extracellular concentrations (as juxtacrine signaling does not involve excretion of signaling molecules into the extracellular space)—however, this interpretation is immaterial to the mathematics of the model.

C. Cell movement

Coordinated movement is essential for tissue development [54,55]; we employ a simple model of cell movement in which the velocity of a given cell is a function of the positions and cell states of nearby cells. This can capture basic mechanical forces and chemical signals that influence movement, but it lacks an explicit description of membrane-membrane and membrane-substrate contact.

A general stochastic model for cell motion can be expressed using an SDE such as

$$\mathrm{d}\boldsymbol{r}^\alpha(t) = \boldsymbol{g}(\boldsymbol{r}^\alpha, S(t))\mathrm{d}t + \eta(\boldsymbol{r}^\alpha, S(t))\mathrm{d}\boldsymbol{W}^\alpha, \qquad (4)$$

where $S(t) = \{(\boldsymbol{r}^\alpha(t), \boldsymbol{z}^\alpha(t))\}_{\alpha=1}^{N(t)}$ is the full state of the tissue population, $\mathrm{d}\boldsymbol{W}^\alpha$ is a three-dimensional Brownian noise, and integration over the noise is evaluated using Itô rules. The mean velocity $\boldsymbol{g}$ and noise amplitude η (a 3×3 matrix) can depend on the cell's position $\boldsymbol{r}^\alpha$ as well as on the full state of the tissue $S(t)$. To model chemotaxis, $\boldsymbol{g}$ and η could also be chosen to depend on $\nabla \boldsymbol{c}(\boldsymbol{r}^\alpha, t)$ and/or higher derivatives of the concentration profile. Many related models for cell motion, including anomalous diffusion and longer-ranged hopping have been developed [56–58]. Note that when modeling cells within microenvironments, appropriate boundary conditions on the cell motion need to be imposed, which are distinct from the boundary conditions for the signaling molecule field $c_\ell(\boldsymbol{r}, t)$.

D. Cell proliferation (and death)

During cell division—a process involving cellular-scale disruption and reorganization—a mother cell divides into two daughter cells which may acquire states that are different from the state of the mother. Within our framework, cells have a state-dependent and concentration-dependent division rate, denoted $\beta(\boldsymbol{z}; \boldsymbol{c})$ for a cell in state $\boldsymbol{z}$ and with signaling molecule concentrations $\boldsymbol{c}$ at its location [hence, cell α has division rate $\beta(\boldsymbol{z}^\alpha; \boldsymbol{c}^\alpha)$]. Note that $\beta(\boldsymbol{z}; \boldsymbol{c})$ is a total division rate derived from a differential division rate density $\tilde{\beta}(\boldsymbol{z}; \boldsymbol{c}; \boldsymbol{z}', \boldsymbol{z}'', \Delta\boldsymbol{r}', \Delta\boldsymbol{r}'')$ for a cell in state $\boldsymbol{z}$ to divide into

daughters with states z', z'' and with positional displacements from the mother $\Delta\boldsymbol{r}', \Delta\boldsymbol{r}''$. $\beta(z;\boldsymbol{c})$ is defined as the marginal

$$\beta(z;\boldsymbol{c}) \equiv \int \tilde{\beta}(z;\boldsymbol{c};z',z'',\Delta\boldsymbol{r}',\Delta\boldsymbol{r}'')\mathrm{d}z'\mathrm{d}z''\mathrm{d}\Delta\boldsymbol{r}'\mathrm{d}\Delta\boldsymbol{r}''. \quad (5)$$

Finally, cell death occurs at rate $\mu(z;\boldsymbol{c})$, which can depend on both z and $\boldsymbol{c}$. In cases where cell movement causes cells to enter or exit the spatial domain of interest, β and μ can be augmented to include the entry and exit rates, respectively (note that this would typically require incorporating an $\boldsymbol{r}$ dependence into β and μ, as cell entry and exit occurs on the boundaries of the spatial domain).

Equations (1), (2), and (4), along with birth and death rates $\tilde{\beta}(z;\boldsymbol{c};z',z'',\Delta\boldsymbol{r}',\Delta\boldsymbol{r}'')$ and $\mu(z;\boldsymbol{c})$, form our general description of developing cell populations. For continuous cell states z, the above stochastic processes can be described by joint state and cell number probability densities that obey variable-dimensional kinetic equations [39], as described in the Introduction. By judiciously marginalizing over states and/or cell populations, reduced density functions can be shown to obey partial integrodifferential equations [39,59]. These reduced densities and integrodifferential equations may not have clear or experimentally accessible interpretations.

In the following section, we illustrate how other commonly invoked assumptions that preserve stochasticity can be incorporated to significantly simplify the framework, without relying on marginalization of cumbersome kinetic equations.

III. SIMPLIFYING ASSUMPTIONS

Here we describe a series of assumptions that can be applied to systematically simplify our mathematical framework. Our examples in Secs. V and VI show how various combinations of these assumptions can be used to obtain models that capture different degrees of detail.

A. Fast equilibration of signaling molecule concentrations

Suppose the concentrations $c_\ell(\boldsymbol{r},t)$ reach steady state on timescales shorter than those associated with gene expression dynamics, cell movement, and cell division and death. Using a quasisteady approximation, $c_\ell(\boldsymbol{r},t)$ can be treated as a function of the instantaneous state of the full population $S(t) = \{z^\alpha(t), \boldsymbol{r}^\alpha(t)\}_{\alpha=1}^{N(t)}$. We denote this quasi-steady-state concentration profile by $c_\ell^{\mathrm{ss}}(\boldsymbol{r};S(t))$. For signaling molecule dynamics described by Eq. (2), $c_\ell^{\mathrm{ss}}(\boldsymbol{r};S(t))$ takes the form

$$c_\ell^{\mathrm{ss}}(\boldsymbol{r};S(t)) = \sum_{\text{cells }\alpha} c_\ell^{(1)}(\boldsymbol{r}-\boldsymbol{r}^\alpha, z^\alpha), \quad (6)$$

where $c_\ell^{(1)}(\boldsymbol{r},z)$ is the concentration profile produced by a single isolated cell at position $\boldsymbol{r}=\boldsymbol{0}$, given by the solution to

$$0 = D_\ell \nabla_{\boldsymbol{r}}^2 c_\ell^{(1)}(\boldsymbol{r},z) - d_\ell c_\ell^{(1)}(\boldsymbol{r},z) + \lambda_\ell(\boldsymbol{r},z) \quad (7)$$

with appropriate boundary conditions.

B. Spatial uniformity

Some developing tissues may be treated as spatially uniform. For example, stem-cell microenvironments may be approximately well mixed, so that the spatial distributions of cells and signaling molecules can both be treated as spatially averaged quantities. We also note that blood is a tissue which can be well mixed within small regions. Under this assumption, Eq. (4) can be ignored and c_ℓ is independent of $\boldsymbol{r}$. If the dynamics of c_ℓ are given by Eq. (2) with reflecting boundary conditions, then Eq. (2) can be replaced by an ODE for the mean concentration $\bar{c}(t) = (1/V)\int c_\ell(\boldsymbol{r})\mathrm{d}\boldsymbol{r}$ over the closed volume V,

$$\frac{\mathrm{d}\bar{c}_\ell}{\mathrm{d}t} = \sum_{\text{cells }\alpha} \bar{\lambda}_\ell(z^\alpha) - d_\ell \bar{c}_\ell, \quad (8)$$

where $\bar{\lambda}_\ell(z) = (1/V)\int \lambda_\ell(\boldsymbol{r},z)\mathrm{d}\boldsymbol{r}$ is the spatially averaged excretion rate of cells in state z. If there is fast equilibration of signaling molecule concentrations (assumption III A above), then the signaling molecule concentrations are

$$c_\ell^{\mathrm{ss}} = \frac{1}{d_\ell}\sum_{\text{cells }\alpha} \bar{\lambda}_\ell(z^\alpha). \quad (9)$$

C. Fast z dynamics and discrete cell types

For fast cell state dynamics and well separated, distinct attractors of $\boldsymbol{F}(z^\alpha;\boldsymbol{c}^\alpha)$, the state z^α of cell α rapidly reaches an attractor, and noise-induced fluctuations can cause the cell state to jump to other attractors. In this limit, the cell state space can be approximated by a discrete network of nodes, labeled $q = 1,\ldots,Q$, with associated transition rates between them [60]. Such a discrete approximation of an otherwise continuous or highly granular set of cell states has been shown to be useful in many developmental contexts such as embryonic stem-cell differentiation [26,27,61]. In this approximation, the Waddington vector field $\boldsymbol{F}(z^\alpha;\boldsymbol{c}^\alpha)$ is replaced by a matrix of jump rates, $F_{q,q'}(\boldsymbol{c}^\alpha)$, specifying the jump rate from q to q'. These jump rates can, in principle, be estimated numerically via simulations of Eq. (1), derived from first passage time computations along the vector field, or, if the noise is small, jump rates can be estimated with a generalization of the Erying-Kramers transition rate formula [62].

The division rate at which cell α in attractor q divides into daughter cells at attractor states q', q'' with positional displacements $\Delta\boldsymbol{r}', \Delta\boldsymbol{r}''$ is denoted $\tilde{\beta}_{q,q',q''}(\boldsymbol{c};\Delta\boldsymbol{r}',\Delta\boldsymbol{r}'')$. Letting z_q denote the attractor state q, these division rates are given by $\tilde{\beta}_{q,q',q''}(\boldsymbol{c};\Delta\boldsymbol{r}',\Delta\boldsymbol{r}'') = \int_{q'} dz_{q'}\int_{q''} dz_{q''}\tilde{\beta}(z_q;\boldsymbol{c};z_{q'},z_{q''},\Delta\boldsymbol{r}',\Delta\boldsymbol{r}'')$. Here, $\int_{q'} dz_{q'}\int_{q''} dz_{q''}$ denotes integrals over the basins of attraction about $z_{q'}$ and $z_{q''}$. The diagonal elements $\tilde{\beta}_{q,q,q}$ correspond to symmetric division in which both daughters have the same state as the mother, whereas elements with $q' \neq q''$ correspond to asymmetric division. The death rate in attractor q is $\mu_q(c) = \mu(z_q;\boldsymbol{c})$.

D. Unifying the assumptions

If all three of the above simplifying approximations (fast equilibration of signaling molecules, spatially uniform tissue, and discrete cell types) hold, then the framework simplifies significantly. Let the vector $\mathbf{n} \equiv (n_1,\ldots,n_Q)$ specify the cell populations in the tissue, where n_q is the number of cells of type q. Let $\beta_{q,q',q''}(\boldsymbol{c}) \equiv \int \tilde{\beta}_{q,q',q''}(\boldsymbol{c};\Delta\boldsymbol{r}',\Delta\boldsymbol{r}'')\mathrm{d}\Delta\boldsymbol{r}'\mathrm{d}\Delta\boldsymbol{r}''$ be the division rate of a mother of type q into daughters of type q', q''. Equilibrated signaling molecule concentrations $\boldsymbol{c}(\mathbf{n})$ are now functions only of the number of cells α comprising state $\mathbf{n}$ and are determined by Eq. (9).

The probability $P(\mathbf{n}, t)$ that the system is in state $\mathbf{n}$ at time t evolves according to the following master equation:

$$\begin{aligned}\partial_t P(\mathbf{n},t) = &\sum_{q,q',q''}[(n_q+1)\beta_{q,q',q''}(\boldsymbol{c}(\mathbf{n}_{q+,q'-,q''-}))P(\mathbf{n}_{q+,q'-,q''-}) - n_q\beta_{q,q',q''}(\boldsymbol{c}(\mathbf{n}))P(\mathbf{n})]\\ &+\sum_{q,q'}[(n_q+1)F_{q,q'}(\boldsymbol{c}(\mathbf{n}_{q+,q'-}))P(\mathbf{n}_{q+,q'-}) - n_q F_{q,q'}(\boldsymbol{c}(\mathbf{n}))P(\mathbf{n})]\\ &+\sum_{q}[(n_q+1)\mu_q(\boldsymbol{c}(\mathbf{n}_{q+}))P(\mathbf{n}_{q+}) - n_q\mu_q(\boldsymbol{c}(\mathbf{n}))P(\mathbf{n})]. \end{aligned} \tag{10}$$

In analogy with the bookkeeping notation used in Refs. [38,39], $\mathbf{n}_{q+,q'-,q''-}$ denotes a system state with, compared to system state $\mathbf{n}$, one more type q cell, one fewer type q' cell, and one fewer type q'' cell. Hence, $\mathbf{n}_{q+,q'-,q''-}$ is the system state that converts into state $\mathbf{n}$ when a mother cell of type q divides into daughter cells of types q' and q''. Similarly, $\mathbf{n}_{q+,q'-}$ denotes a system state with one more type q cell and one fewer type q' cell relative to system state $\mathbf{n}$. $\mathbf{n}_{q+}$ denotes a system state with one more type q cell than in system state $\mathbf{n}$. Equation (10) represents a master equation for the probability density for populations of cells of multiple types and is similar to the master equations studied in Refs. [38,39,63]. For small populations and few accessible cell types, solving this linear master equation or simulating the process is computationally feasible.

E. Developmental fitness landscape

Proliferation of certain cell types over others due to differing division and death rates is a crucial aspect of development. Differential proliferation is also key to evolutionary dynamics, where genotypes that induce high net growth rates eventually comprise larger portions of the overall population. In evolutionary dynamics, preferential growth is described by a fitness landscape, typically defined as a function which specifies the net growth (division minus death) rate ϕ_g of a cell or organism given its genotype g. Here we adapt this concept to the context of cell development to define an epigenetic fitness landscape which specifies which cell types proliferate in the developing tissue.

A key difference between evolutionary dynamics (or speciation) and development is that, in a developing tissue, all cells have the same genotype (aside from rare mutations) and differences in cells' division rates are due to differences in their cell type, determined by the internal state $\boldsymbol{z}$. In all cases, the fitness of a multicomponent cell population describes the growth of a specific cell type across time *within the tissue environment or ecosystem*, which might be interpreted as the "diagonal" component $\beta_{q,q,q}(\boldsymbol{c}(\mathbf{n}))$ in population growth equations [60,64]. This interpretation excludes "off-diagonal" in-differentiation and out-differentiation rates. A "diagonal" representation of fitness would be appropriate if we tracked the expansion of all cells of a single clone. If one is concerned only about the total population of a particular cell type, then the off-diagonal in-differentiation and out-differentiation rates should be included.

Here we define the developmental fitness of cell type q, denoted $\phi_q(\mathbf{n})$, as the net proliferation rate of type q cells, absent production of type q cells through division of other cell types or stochastic jumps from other cell types. $\phi_q(\mathbf{n})$ depends on the tissue population state $\mathbf{n}$ due to cell-cell interactions via signaling molecules. In defining $\phi_q(\mathbf{n})$, we adopt the assumptions from the previous sections, namely fast equilibration of signaling molecules, well-mixed populations, and discrete cell types. Within our framework, we can define fitness mathematically as

$$\begin{aligned}\phi_q(\mathbf{n}) = &\beta_{q,q,q}(\boldsymbol{c}(\mathbf{n})) - \beta_{q,\mathrm{lost}}(\boldsymbol{c}(\mathbf{n}))\\ &- F_{q,\mathrm{lost}}(\boldsymbol{c}(\mathbf{n})) - \mu_q(\boldsymbol{c}(\mathbf{n})),\end{aligned} \tag{11}$$

where $\beta_{q,q,q}(\boldsymbol{c})$ is the rate of symmetric division in which a mother cell of type q divides into two daughter cells of type q. $\beta_{q,\mathrm{lost}}(\boldsymbol{c}) \equiv \sum_{q'\neq q,q''\neq q}\beta_{q,q',q''}(\boldsymbol{c})$ is the rate at which a mother cell of type q divides into daughter cells, both of whose cell types differ from that of the mother cell. Note that asymmetric division, where a cell of type q divides into daughters of type q and $q' \neq q$, leaves the population of type q cells unchanged and therefore does not factor into the developmental fitness landscape. $F_{q,\mathrm{lost}}(\boldsymbol{c}) = \sum_{q'\neq q} F_{q,q'}(\boldsymbol{c})$ is the jump rate of type q cells to a different type.

The following subsection shows how, in the limit of large tissue size, the developmental fitness landscape contributes to tissue dynamics through a simple equation.

F. Deterministic limit

In the limit of large tissue size, Eq. (10) predicts that cell densities in the tissue follow deterministic dynamics. This limit is analogous to how deterministic mass action equations (ODEs) emerge from a more microscopic chemical master equation description in the limit of large system size. Recall that Eq. (10) describes the population within a region of fixed volume, and denote this volume by V_0. Then $\boldsymbol{\rho} \equiv \mathbf{n}/V_0$ are the cell-type densities in the tissue described by Eq. (10). Now consider a system with volume $V > V_0$ and take the limit that $V \to \infty$ while $\boldsymbol{\rho}$ remains fixed. The mass-action dynamics in this limit are described by

$$\frac{\mathrm{d}\boldsymbol{\rho}}{\mathrm{d}t} = \mathrm{diag}(\boldsymbol{\phi}(\boldsymbol{\rho}V_0))\boldsymbol{\rho} + \boldsymbol{m}(\boldsymbol{\rho}V_0)\boldsymbol{\rho}, \tag{12}$$

where $\boldsymbol{\phi}(\mathbf{n}) = (\phi_1(\mathbf{n}), \ldots, \phi_Q(\mathbf{n}))$ is the developmental fitness landscape defined in the previous section and $\mathrm{diag}(\boldsymbol{\phi}(\boldsymbol{\rho}V_0))$ denotes the $Q \times Q$ matrix with the elements of $\boldsymbol{\phi}(\boldsymbol{\rho}V_0)$ on the diagonal. $\boldsymbol{m}(\boldsymbol{\rho}V_0)$ is a zero-diagonal matrix which contains the jump rates and asymmetric division rates through which cells of one type produce cells of another type.

The matrix elements are explicitly

$$m_{q,q'}(\mathbf{n}) = F_{q',q}(\boldsymbol{c}(\mathbf{n})) + \sum_{q''}[\beta_{q',q,q''}(\boldsymbol{c}(\mathbf{n})) + \beta_{q',q'',q}(\boldsymbol{c}(\mathbf{n}))]. \tag{13}$$

Deterministic dynamical models of interacting population dynamics, as could be described by Eq. (12), have been used to model aspects of development [65]. Note that the stochastic fluctuations in the dynamics of $\boldsymbol{\rho}$ are proportional to $(V/V_0)^{-1/2}$, which vanishes as $V \to \infty$. For large but finite V, the stochastic component of the dynamics can be obtained with the van Kampen system size expansion [44,66].

IV. WADDINGTON'S LANDSCAPE VERSUS WADDINGTON'S VECTOR FIELD

In the most common mathematical interpretation of Waddington's landscape, two assumptions are imposed on Eq. (1):

(i) The vector field $\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{c})$ is assumed to be the gradient of a Waddington landscape function $U(\boldsymbol{z}, \boldsymbol{c})$. Hence, under Itô calculus, the expected change in state is downhill: $\mathbb{E}[\mathrm{d}\boldsymbol{z}] = -\boldsymbol{\nabla}_z U\,\mathrm{d}t$.

(ii) $\boldsymbol{\sigma}(\boldsymbol{z}, \boldsymbol{c})$ is assumed to be independent of $\boldsymbol{z}$ (additive noise) and satisfy $\boldsymbol{\sigma}\boldsymbol{\sigma}^\top \propto \mathbb{I}$.

Under assumption (i), Eq. (1) becomes

$$\mathrm{d}\boldsymbol{z} = -\boldsymbol{\nabla}_z U(\boldsymbol{z};\boldsymbol{c})\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{c})\mathrm{d}\mathbf{W}, \tag{14}$$

where, as before, the Itô convention is used. If assumptions (i) and (ii) are both satisfied, then U can be reconstructed (up to a multiplicative constant) from the steady-state distribution of cell states under fixed signaling molecule concentrations ($p_{\mathrm{ss}}(\boldsymbol{z};\boldsymbol{c})$ for fixed $\boldsymbol{c}$) using the formula

$$U(\boldsymbol{z};\boldsymbol{c}) \propto -\log p_{\mathrm{ss}}(\boldsymbol{z};\boldsymbol{c}). \tag{15}$$

If assumptions (i) and (ii) are not *both* satisfied, then Eq. (15) is generally inapplicable except in special (and unrealistic) cases; for example, Eq. (15) can hold without assumption (ii) if $\boldsymbol{\nabla}_z U$ is parallel to an eigenvector of $\boldsymbol{\sigma}$ for all $\boldsymbol{z}$.

When are assumptions (i) and (ii) valid? For typical gene regulatory dynamics, $\boldsymbol{F}$ will not be a gradient [13–15,19,21,35,67] *and* noise is nonadditive [6]; in other words, both (i) and (ii) are violated. It was recently pointed out that the vector field $\boldsymbol{F}$ can often be converted via coordinate transformation to a gradient form satisfying assumption (i) [5]; this is possible when $\boldsymbol{F}$ has no limit cycles and satisfies a few other technical requirements ($\boldsymbol{F}$ must be Morse-Smale [68]). Even though most gene regulatory dynamic models are Morse-Smale, under Itô calculus the convection terms transform nontrivially. However, it is in general impossible to satisfy both assumption (i) and (ii) via a coordinate transformation, so this technique does not provide a Waddington landscape to which Eqs. (14) and (15) can be applied.

If Eq. (15) is applied to data from a system that does not satisfy both assumptions (i) and (ii), then the reconstructed landscape U can have spurious attractors that are not present in the true vector field $\boldsymbol{F}$; there can also be attractors of $\boldsymbol{F}$ that are missing in the reconstructed U. Previous work has shown examples of this phenomenon for systems where assumption (ii) is violated [6]. We provide an example of this for a system where assumption (i) is violated (Appendix A) and show how it can lead to spurious and missing attractors. Hence, Eq. (15) should not be used unless there is reason to believe that the system under study obeys assumptions (i) and (ii); for most gene regulatory dynamics, this will not be the case. A key implication of these properties is that it is generally not possible to reconstruct the dynamics of cell state from the steady-state distribution over cell states. Such a reconstruction would require knowledge of the state dependence of $\boldsymbol{\sigma}(\boldsymbol{z};\boldsymbol{c})$, and this is typically unknown when $F_m(\boldsymbol{z};\boldsymbol{c})$ is unknown.

The conceptual utility of Waddington's landscape is that it gives a clear intuition of the dynamics of a cell state that is drawn into an attractor as cell differentiation occurs. By using a more general Waddington vector field picture, the concept of evolving state dynamics is retained, but the unjustified assumptions (i) and (ii) are no longer needed.

V. EXAMPLE 1: STEM-CELL DIFFERENTIATION IN A WELL-MIXED MICROENVIRONMENT

Here we present a model of tissue development which is derived by augmenting a prototypical model of Waddington's landscape studied in Refs. [19,21,69]. In our model, $\boldsymbol{F}(\boldsymbol{z};\boldsymbol{c})$ describes a hypothetical gene regulatory circuit involving two proteins, A and B. $\boldsymbol{z} = (z_A, z_B)$ specifies the concentrations of these two proteins. $\boldsymbol{F}(\boldsymbol{z};\boldsymbol{c})$ has three attractors corresponding to three cell types, which, for reasons that will become clear, we call *stem cells* $\mathcal{S}$, *triggered stem cells* $\mathcal{S}^*$, and *differentiated cells* $\mathcal{D}$. Cells interact via two signaling molecules with concentrations $\boldsymbol{c} = (c_1, c_2)$. These interactions induce stem cells to differentiate in a coordinated way, so that a stable population (i.e., a tissue) of differentiated cells is produced and maintained. Hence, while the previous models in Refs. [19,21,69] describe only single-cell differentiation, our model describes the coordinated development of tissue. After analyzing the dynamics of the full model, we show how the model can be simplified into a simple system of ODEs using the approximations developed in Sec. III.

The gene regulatory circuit that governs the dynamics of $\boldsymbol{z}$ is illustrated in Fig. 2(a). Each gene (A and B) upregulates itself and downregulate the other. We make the common quasi-steady-state assumption that mRNA transcription and degradation are fast relative to protein translation and degradation, so the state and dynamics are described in terms of only the protein concentrations $\boldsymbol{z} = (z_A, z_B)$. Signaling molecule "1" influences expression of gene A but not gene B. Specifically, B tetramer (which downregulates gene A) can form a complex with signaling molecule 1, and this complex prevents any expression of A when bound to the silencer of A. This binding scheme gives rise to a Waddington vector field, shown in Fig. 2(b), whose components are

$$\begin{aligned} F_m(\boldsymbol{z};\boldsymbol{c}) = {} & r_{0,m}(1 - E_m(z_m))(1 - S_m(z_{m'})) \\ & + r_{E,m}E_m(z_m)(1 - S_m(z_{m'})) \\ & + r_{S,m}(c_1)(1 - E_m(z_m))S_m(z_{m'}) \\ & + r_{ES,m}(c_1)E_m(z_m)S_m(z_{m'}) - k_m z_m, \end{aligned} \tag{16}$$

where $m, m' = \mathrm{A}, \mathrm{B}$ or B, A. $E_m(z_m)$ and $S_m(z_{m'})$ are, respectively, the probabilities that the enhancer and silencer of gene

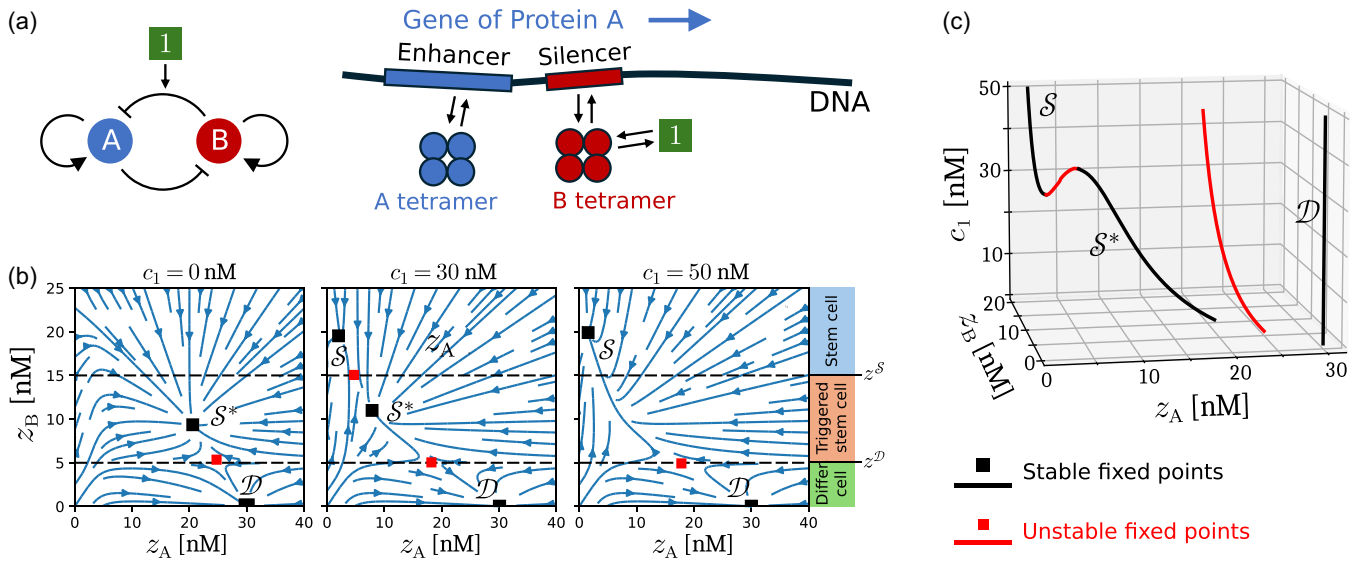

FIG. 2. Example 1: Model of regulated stem-cell differentiation. (a) Gene regulatory circuit. Left: Schematic of circuit showing upregulation (arrows), downregulation (flat arrows), and the effect of signaling molecule 1 concentration c_1 on amplification of the downregulation of A by B. Right: Diagram of molecular mechanisms. Black arrows indicate molecular binding and unbinding of the A and B tetramers to the enhancer and silencer, and of signaling molecule 1 (green square) to B tetramer. Signaling molecule 2 controls cell proliferation and is not featured in this picture of individual-cell gene regulation dynamics. (b) Vector field $\boldsymbol{F}(\boldsymbol{z};\boldsymbol{c})$ at three values of c_1. Attractors (black squares) are labeled by cell type: stem cells $\mathcal{S}$, triggered stem cells $\mathcal{S}^*$, and differentiated cells $\mathcal{D}$. Unstable fixed points indicated by red squares. Black dashed lines at $z_B = z^{\mathcal{S}}$ and $z^{\mathcal{D}}$ denote the boundary between stem cells and triggered stem and between triggered stem cells and differentiated cells, respectively. (c) Fixed points of $\boldsymbol{F}(\boldsymbol{z};\boldsymbol{c})$ as a function of c_1. Two saddle-node bifurcations result in the gain/loss of attractors $\mathcal{S}$ and $\mathcal{S}^*$.

m have a bound tetramer. These probabilities are modeled by Hill functions with Hill coefficient 4,

$$E_m(z_m) = \frac{z_m^4}{K_{E,m}^4 + z_m^4}, \quad S_m(z_{m'}) = \frac{z_{m'}^4}{K_{S,m}^4 + z_{m'}^4}. \tag{17}$$

The rates $r_{0,m}$, $r_{E,m}$, $r_{S,m}(c_1)$, and $r_{ES,m}(c_1)$ are the mean expression rates of gene m with, respectively, neither enhancer nor silencer occupied, only enhancer occupied, only silencer occupied, and both enhancer and silencer occupied. The rates $r_{S,A}(c_1)$ and $r_{ES,A}(c_1)$ are given by

$$r_{S,A}(c_1) = \frac{r_{S,A}(0)\gamma}{\gamma + c_1}, \quad r_{ES,A}(c_1) = \frac{r_{ES,A}(0)\gamma}{\gamma + c_1}. \tag{18}$$

The factor $\gamma/(\gamma + c_1)$ is the probability that a given B tetramer is *not* bound to a signaling molecule. All other rate parameters, $r_{0,m}$, $r_{E,m}$, $r_{S,B}$, $r_{ES,B}$, $r_{S,A}(0)$, and $r_{ES,A}(0)$ are constants (listed in Table I, along with all other model parameters). Figure 2(b) shows the Waddington vector field for three values of c_1: $c_1 = 0$ (left), $c_1 = 30$ nM (middle), and $c_1 = 50$ nM (right). For a particular choice of parameters, $\boldsymbol{F}(\boldsymbol{z};\boldsymbol{c})$ reduces to the model in Refs. [19,69]. The noise $\boldsymbol{\sigma}(\boldsymbol{z};\boldsymbol{c})$ used in this example is multiplicative in $\boldsymbol{z}$ and is motivated through a chemical Langevin equation [51]. Appendix B provides details on the noise implementation, including an analysis of predictions if an alternative additive noise model were used.

The key role of signaling molecule 1 is to regulate how stem cells are triggered for differentiation. Figure 2(b) shows the attractors (black squares) for the three cell types ($\mathcal{S}$, $\mathcal{S}^*$, and $\mathcal{D}$). At intermediate c_1, stable attractors exist for both $\mathcal{S}$ and $\mathcal{S}^*$, but when c_1 drops sufficiently low, the $\mathcal{S}$ attractor disappears and stem cells move toward the triggered state $\mathcal{S}^*$. Hence, low c_1 triggers stem cells for differentiation, and as will be described below, these triggered stem cells undergo asymmetric division to produce differentiated cells. On the other hand, very high c_1 reverses the triggering (the $\mathcal{S}^*$ attractor disappears and triggered stem cells revert to untriggered stem cells). Figure 2(c) shows how the locations of the attractors (black) and unstable fixed points that separate the attractors (red) vary continuously with c_1. The unstable fixed point that separates the $\mathcal{S}$ and $\mathcal{S}^*$ attractors merges with $\mathcal{S}$ at low c_1 but merges with $\mathcal{S}^*$ at high c_1. These mergers of fixed points are suggestive of saddle-node bifurcations.

Cell division rates depend on both the expression level z_B of protein B and the concentration c_2 of signaling molecule 2. The dependence on z_B gives rise to the defining behavior of the three cell types: Stem cells divide symmetrically (both daughters are type $\mathcal{S}$), triggered stem cells divide asymmetrically (one daughter is $\mathcal{S}$ and the other $\mathcal{D}$), and differentiated cells do not divide. Mathematically, we model the birth rate by

$$\tilde{\beta}(\boldsymbol{z};\boldsymbol{c};\boldsymbol{z}',\boldsymbol{z}'') = \begin{cases} 0 & z_B \leqslant z^{\mathcal{D}} \\ \frac{1}{2}\beta_{\mathcal{S}^*}\delta(\boldsymbol{z}-\boldsymbol{z}'-\boldsymbol{w}(\boldsymbol{z})) & \\ \quad \times\delta(\boldsymbol{z}-\boldsymbol{z}''+\boldsymbol{w}(\boldsymbol{z})) & z^{\mathcal{S}} \geqslant z_B > z^{\mathcal{D}} \\ \dfrac{a_1\delta(\boldsymbol{z}-\boldsymbol{z}')\delta(\boldsymbol{z}-\boldsymbol{z}'')}{1+a_2\exp(a_3c_2)} & z_B > z^{\mathcal{S}}. \end{cases} \tag{19}$$

where $z^{\mathcal{D}} < z^{\mathcal{S}}$ are threshold values for z_B. Note that we do not include the $\Delta\boldsymbol{r}'$, $\Delta\boldsymbol{r}''$ dependencies because this model

TABLE I. Parameters for Examples 1 and 2.

Parameter	Value	[units]	Description
Example 1 model parameters			
a_1	1	[day^{-1}]	Max. stem-cell division rate [Eq. (19)]
a_2	0.15	[unitless]	Parameter in Eq. (19)
a_3	0.5	[nM^{-1}]	Parameter in Eq. (19)
$\beta_{\mathcal{S}^*}$	0.6	[day^{-1}]	Triggered stem-cell division rate [Eq. (19)]
$\boldsymbol{w}(\boldsymbol{z})$	$0.95(z_A, -z_B)$	[z]	Controls asymmetric division [Eq. (19)]
μ	0.02	[day^{-1}]	Death rate
L_1, L_2	1	[nM]	Mean excretion rates of signaling molecules [Eqs. (20) and (21)]
$z^{\mathcal{S}}$	15	[nM]	Threshold value of z_B between stem cells and triggered stem cells [Eqs. (19) and (21)]
$z^{\mathcal{D}}$	5	[nM]	Threshold value of z_B between triggered stem and differentiated cells [Eqs. (19) and (20)]
$K_{E,m}$	5	[nM]	Equilibrium const. for tetramerization and binding of m=A, B to Enhancer [Eq. (17)]
$K_{S,m}$	5	[nM]	Equilibrium const. for tetramerization and binding of m=A, B to Silencer [Eq. (17)]
γ	20	[nM]	Equilibrium const. for signaling molecule 1 binding to B tetramer [Eq. (18)]
$r_{0,A}$	500	[nM/day]	Expression rate of A when A enhancer and silencer of A are unoccupied [Eq. (16)]
$r_{0,B}$	500	[nM/day]	Expression rate of B when B enhancer and silencer of B are unoccupied [Eq. (16)]
$r_{S,A}(0)$	250	[nM/day]	Expression rate of A when A silencer is occupied and $c_1 = 0$ [Eq. (18)]
$r_{ES,A}(0)$	1000	[nM/day]	Expression rate of A when A enhancer and silencer are occupied and $c_1 = 0$ [Eq. (18)]
$r_{E,A}$	1500	[nM/day]	Expression rate of A when A enhancer is occupied [Eq. (16)]
$r_{E,B}$	1000	[nM/day]	Expression rate of B when B enhancer is occupied [Eq. (16)]
$r_{S,B}$	0	[nM/day]	Expression rate of B when B silencer is occupied [Eq. (16)]
$r_{ES,B}$	500	[nM/day]	Expression rate of B when B enhancer and silencer are occupied [Eq. (16)]
k_m	50	[day^{-1}]	Degradation rate of m=A, B [Eq. (16)]
$\boldsymbol{\sigma}(\boldsymbol{z};\boldsymbol{c})$	Eq. (B2)	[day$^{-1/2}$]	Noise matrix of gene dynamics [Eq. (1)]
Example 2 model parameters			
β_0	1	[day^{-1}]	Division rate [Eq. (27)]
ξ	2.5	[μm]	Std. dev. of displacement of daughter from mother [below Eq. (27)]
c_2^*	23	[nM]	Threshold of c_2 above which cell division does not occur [Eq. (27)]
μ	0.1	[day^{-1}]	Death rate
C_ℓ	1	[nM]	Parameter controlling excretion rate of signaling molecule $\ell = 1, 2$
s_1	7	[μm]	Std. dev. of spatial profile of excretion of signaling molecule 1
s_2	30	[μm]	Std. dev. of spatial profile of excretion of signaling molecule 2
r_0	1000	[nM/day]	Expression rate of A when enhancer is unoccupied [Eq. (25)]
r_1	2100	[nM/day]	Enhancement in expression rate when enhancer is occupied [Eq. (25)]
K	35	[nM]	Equilibrium const. for dimerization and heterodimer-Enhancer binding [Eq. (25)]
b	250	[day^{-1}]	Degradation rate of protein A [Eq. (25)]
σ	$0.1z$	[day$^{-1/2}$]	Noise magnitude of gene dynamics [Eq. (1)]
v_0	10^3	[μm^2/day]	Parameter controlling velocity of cell movement [Eq. (26)]
$r_{\min}$	8	[μm]	Spatial extent of attractive forces between cells [Eq. (26)]
η	0		Noise magnitude of cell movement [Eq. (4)]

assumes a well-mixed (spatially uniform) system. The threshold concentrations $z^{\mathcal{S}}$ and $z^{\mathcal{D}}$ are indicated by dashed lines in Fig. 2(b), and are determined by the (approximate) z_B values of the unstable fixed points of $\boldsymbol{F}$.

The vector $\boldsymbol{w}(\boldsymbol{z})$ induces asymmetric division: When a triggered stem cell divides, one daughter has a state shifted by $\boldsymbol{w}(\boldsymbol{z})$ relative to the mother, and the other daughter has state shifted by $-\boldsymbol{w}(\boldsymbol{z})$. As a result, one daughter returns to the $\mathcal{S}$ attractor, while the other moves to the $\mathcal{D}$ attractor. Note that for $\boldsymbol{z}$ defined by concentrations, the concentrations $\boldsymbol{z}'$ and $\boldsymbol{z}''$ in the daughter cells are related to that of the mother through $(\boldsymbol{z}' + \boldsymbol{z}'')/2 = \boldsymbol{z}$. The death rate μ for all cells is assumed constant.

In this model, signaling molecule 1 is excreted by differentiated cells and signaling molecule 2 is excreted by stem cells. We use Eq. (9) to model signaling molecule concentrations, which assumes a well-mixed population and that concentrations rapidly reach steady state. The excretion rates of c_1 and c_2 are, respectively,

$$\bar{\lambda}_1(\boldsymbol{z}) = \begin{cases} 0 & z_B > z^{\mathcal{D}} \\ L_1 & z_B \leqslant z^{\mathcal{D}} \end{cases} \tag{20}$$

and

$$\bar{\lambda}_2(\boldsymbol{z}) = \begin{cases} L_2 & z_B > z^{\mathcal{S}} \\ 0 & z_B \leqslant z^{\mathcal{S}}. \end{cases} \tag{21}$$

Using the fast equilibration and well-mixed assumptions (Secs. III A and III B) the signaling molecule concentrations are simply proportional to the cell numbers $n_{\mathcal{S}}$ and $n_{\mathcal{D}}$,

$$\left(c_1^{\text{ss}}, c_2^{\text{ss}}\right) = (n_{\mathcal{D}} L_1 d_1, n_{\mathcal{S}} L_2 d_2). \tag{22}$$

The signaling molecules' influence on gene dynamics and birth rates gives rise to robust tissue development: From

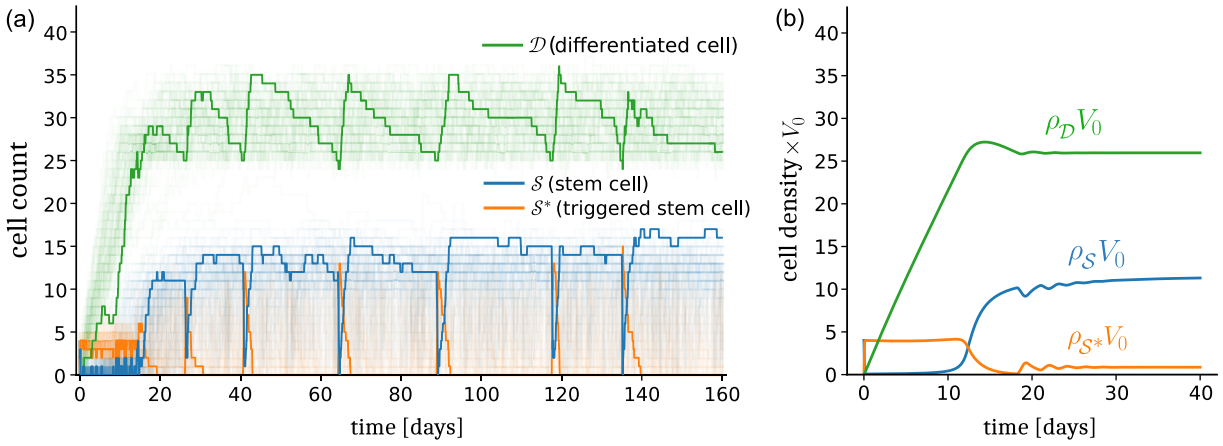

FIG. 3. Example 1: Dynamics of cell populations. (a) Dynamical simulation starting from an initial condition with $n_{\mathcal{S}} = 4$ and $n_{\mathcal{S}^*} = 0$. The population dynamics exhibit thermostat-like behavior: when $n_{\mathcal{D}}$ falls below 26, stem cells are triggered for asymmetric division, which results in $n_{\mathcal{D}}$ being replenished. One-hundred trajectories were simulated by evaluating Eq. (1), coupled with a discrete cell birth-death process, using the Euler-Maruyama method (see Appendixes B and C). One representative trajectory is highlighted, while the other ninety-nine are lightly shaded. With the parameters specified in Table I and an initial population of 4 stem cells, the population did not go extinct within the 110 days simulation for any of the 100 trials. (b) Evolution of cell populations in the deterministic limit. The densities are multiplied by the normalizing volume V_0 to express densities on the same scale as particle counts. Note that the deterministic trajectories qualitatively reflect the overall dynamics shown in (a). In the fast signaling molecule production and decay approximation used in this model, the concentrations $c_1 \approx c_1^{\rm ss}$, $c_2 \approx c_2^{\rm ss}$ are given by Eq. (22) and are proportional to the $\mathcal{D}$ and $\mathcal{S}$ cell populations $n_{\mathcal{D}}$ and $n_{\mathcal{S}}$, respectively.

a wide range of initial conditions, the system dynamically approaches a stable tissue state with regulated relative cell-type populations, as illustrated in Fig. 3(a). The two signaling molecules govern the two key regulatory mechanisms. First, the stem-cell population $n_{\mathcal{S}}$ self-regulates through the dependence of stem-cell division on c_2. When $n_{\mathcal{S}}$ is low (high), c_2 is low (high), so stem-cell division rate is high (low), replenishing (diminishing) $n_{\mathcal{S}}$ towards the stable, developed state. The dependence of $\boldsymbol{F}(\boldsymbol{z};\boldsymbol{c})$ on c_1 leads to a thermostat-like behavior of stem-cell triggering. When $n_{\mathcal{D}}$ is low (high), c_1 is low (high), so more (less) stem cells are triggered for asymmetric division. Regeneration of differentiated cells switches on when $n_{\mathcal{D}}$ drops too low (leading to low c_1), then switches off when $n_{\mathcal{D}}$ and c_1 rises sufficiently. This thermostat-like behavior is apparent from Fig. 3(a); whenever $n_{\mathcal{D}}$ drops below a threshold, stem cells are triggered, rapidly replenishing $n_{\mathcal{D}}$. Specifically, an initial population of only four stem cells develops into the stable tissue state.

A. Deterministic limit

The deterministic limit described in Sec. III F can be applied to obtain a highly simplified form of the model. Denote the cell densities for the three cell types as $\boldsymbol{\rho} = (\rho_{\mathcal{S}}, \rho_{\mathcal{S}^*}, \rho_{\mathcal{D}})$. The developmental fitness landscape [Eq. (11)] is

$$\boldsymbol{\phi}(\mathbf{n}) = (\beta_{\mathcal{S},\mathcal{S},\mathcal{S}}(n_{\mathcal{S}}) - F_{\mathcal{S},\rm lost}(n_{\mathcal{D}}) - \mu,\ -\beta_{\mathcal{S}^*,\rm lost} - \mu,\ -\mu), \tag{23}$$

where $\beta_{\mathcal{S},\mathcal{S},\mathcal{S}}(n_{\mathcal{S}}) = a_1/(1 + a_2 \exp(a_3 n_{\mathcal{S}} L_2/d_2))$, as determined by Eqs. (19) and (22). $F_{\mathcal{S},\rm lost}(n_{\mathcal{D}})$ is the rate at which stem cells transition to triggered stem cells, which can be estimated numerically by simulating the gene dynamics (see Appendix D; also note that the transition rate from stem cell directly to differentiated cell is negligible). Note that many of the terms from Eq. (11) are zero; in particular, $F_{\mathcal{S}^*,\rm lost}$ and $F_{\mathcal{D},\rm lost}$ are negligibly small for the relevant values of c_1 ($c_1 < 40$ nM). This can be seen visually from the Waddington vector fields in Fig. 2(b), and confirmed numerically (Appendix D).

The deterministic equations given by Eq. (12) are

$$\begin{aligned}
\frac{{\rm d}\rho_{\mathcal{S}}}{{\rm d}t} &= (\beta_{\mathcal{S},\mathcal{S},\mathcal{S}}(\rho_{\mathcal{S}}V_0) - F_{\mathcal{S},\rm lost}(\rho_{\mathcal{D}}V_0) - \mu)\rho_{\mathcal{S}} \\
&\quad + \beta_{\mathcal{S}^*,\rm lost}\rho_{\mathcal{S}^*} \\
\frac{{\rm d}\rho_{\mathcal{S}^*}}{{\rm d}t} &= F_{\mathcal{S},\rm lost}(\rho_{\mathcal{D}}V_0)\rho_{\mathcal{S}} - (\beta_{\mathcal{S}^*} + \mu)\rho_{\mathcal{S}^*} \\
\frac{{\rm d}\rho_{\mathcal{D}}}{{\rm d}t} &= \beta_{\mathcal{S}^*}\rho_{\mathcal{S}^*} - \mu\rho_{\mathcal{D}},
\end{aligned} \tag{24}$$

where we have used the fact that $\beta_{\mathcal{S}^*,\mathcal{S},D} + \beta_{\mathcal{S}^*,D,\mathcal{S}} = \beta_{\mathcal{S}^*}$ due to Eq. (19). Figure 3(b) plots the solutions of these equations for the same initial conditions as those used to generate Fig. 3(a). The deterministic limit captures the essential behavior of the full stochastic model in which over about 10 days, five triggered stem cells produce a growing population of differentiated cells. After this initial expansion period, the deterministic model relaxes to a steady state which is highly robust to perturbations; the population will return to this developed tissue state from any initial condition with a positive number of stem or triggered stem cells. While the steady-state populations are similar to the average populations in the full model, the persistent thermostat-like behavior in the full model is absent in the deterministic model. The deterministic model exhibits decaying fluctuations as it relaxes to steady state; the mechanism of these fluctuations is analogous to the mechanism of the thermostat-like behavior in the full model, but these fluctuations occur at a higher frequency and are not persistent.

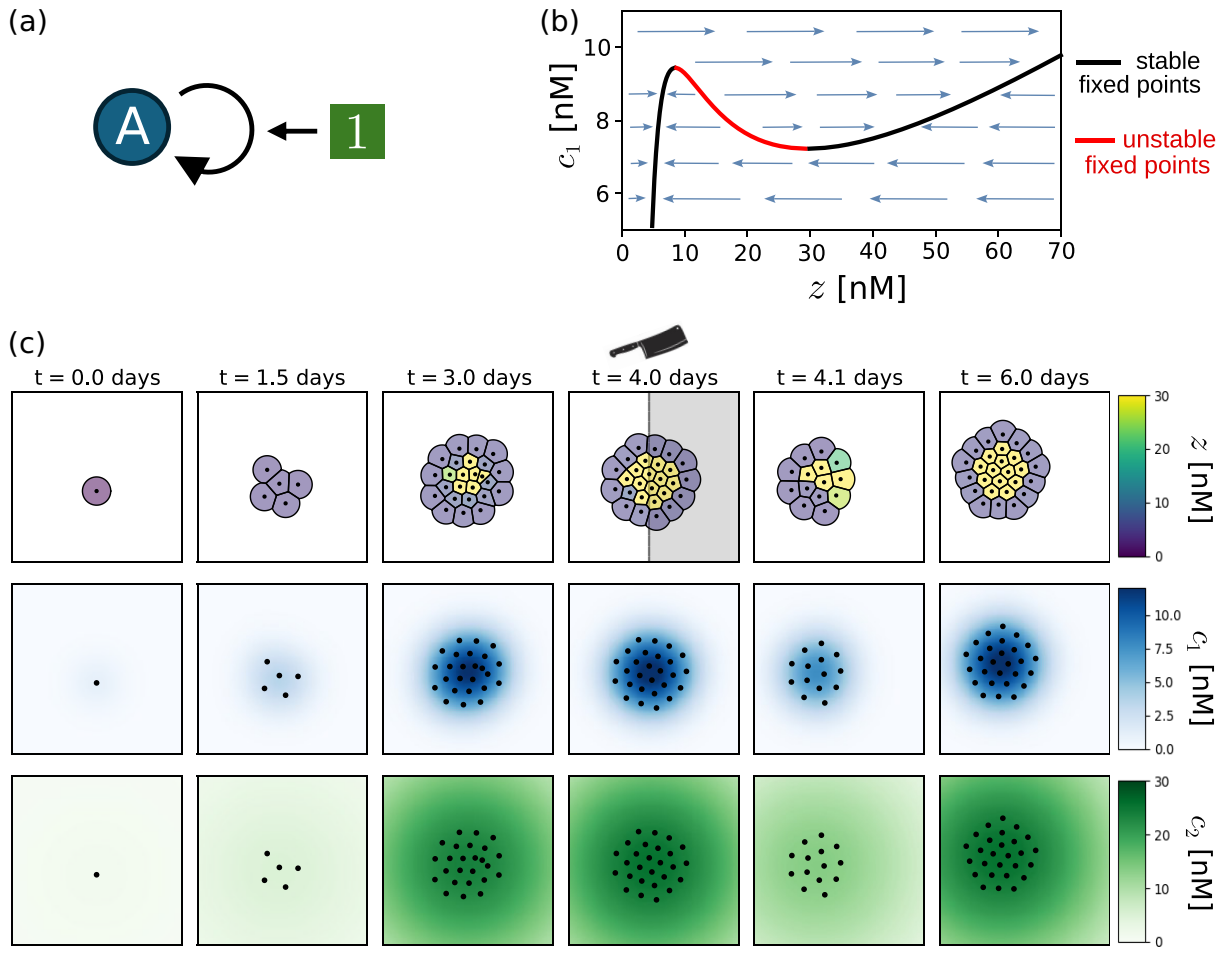

FIG. 4. Example 2: Model of a spatially structured organism with fissiparous reproduction. (a) Gene regulatory circuit, where gene A upregulates itself in the presence of signaling molecule 1. (b) Waddington vector field (blue arrows), attractors (black curves) and unstable fixed points (red curve) as a function of c_1. The left-hand attractor represents exterior cells and the right-hand attractor represents interior cells. (c) Simulation of the dynamics, showing a single cell developing into a stable developed organism (day 0 through 4). At time $t = 4.0$ days, the organism is bisected and the left half is removed. The organism then regrows (4.1 to 6.0 days). Top row: Cell position (black dots) and the state level z given by the color shading. Middle and bottom rows depict the corresponding concentrations c_1 and c_2, respectively. The cell boundaries are Voronoi constructions and are displayed for visual aid; only the cell centers (black dots) factor into our mathematical model.

VI. EXAMPLE 2: SPATIAL MODEL INSPIRED BY PLANARIA

Here we present a model of a spatially structured multicellular organism which develops from a single cell. Additionally, if the developed organism is cut into pieces, then each piece will regenerate into a new fully developed organism (a phenomenon called fissiparous reproduction). This model is inspired by planaria worms, which exhibit a remarkable capacity for fissiparous reproduction: A fragment of a planarian as small as 1/279th of the full worm can regenerate into a new worm [70,71].

In this model, cell state is described by the expression of a single gene, denoted gene A, whose expression is governed by a bistable regulatory circuit. The state z is the concentration of protein A (we assume fast mRNA dynamics, as in example 1). Two stable states corresponding to two values of z define two cell types: exterior cells (which have low z) and interior cells (which have high z). Cells excrete two signaling molecules, one which governs the z dynamics and another which governs cell division rate. The cells live and migrate on a two-dimensional substrate ($\boldsymbol{r} \in \mathbb{R}^2$), and they exert forces on one another which govern their movement.

The gene regulatory circuit is illustrated in Fig. 4(a). Gene A upregulates itself when signaling molecule 1 is present. Specifically, protein A can bind to signaling molecule 1, and this heterodimer binds to the enhancer of gene A, increasing the expression of protein A. The Waddington vector field thus takes the form

$$F(z; c_1) = r_0 + \frac{r_1 c_1 z^2}{K^2 + z^2} - bz. \tag{25}$$

The values of the parameters r_0, r_1, K, and b, as well of the parameters introduced in the following equations, are given in Table I. Because z is one dimensional, F can always be written as the negative gradient of some function, i.e., there exists a $U(z; c_1)$ such that $F(z; c_1) = -\nabla_z U(z; c_1)$.

Figure 4(b) shows the dynamics (blue arrows) as well as the locations of the attractors (black) and unstable fixed points (red) of $F(z; c)$, for values of c_1 between 4 and 12 nM. The attractor at low (high) z, which corresponds to exterior (interior)

cells, merges with the unstable fixed point at high (low) c_1 via a saddle-node bifurcation; when c_1 reaches this bifurcation, exterior (interior) cells convert to interior (exterior) cells. In this way, cell type is regulated by c_1.

Cell movement is modeled by Eq. (4) with

$$\boldsymbol{g}(\boldsymbol{r}^\alpha) = \sum_{\alpha' \neq \alpha} \boldsymbol{v}(\boldsymbol{r}^\alpha - \boldsymbol{r}^{\alpha'}), \tag{26}$$

where $\boldsymbol{v}(\boldsymbol{r}) = v_0(\frac{r_{\min}^2}{r^3} - \frac{r_{\min}}{r^2})\hat{\boldsymbol{r}}$, where $\hat{\boldsymbol{r}}$ is the unit vector in the $\boldsymbol{r}$ direction and v_0 is a parameter that sets the velocity scale (see Table I). We assume the motion is deterministic, so $\eta = 0$ in Eq. (4). The velocity $\boldsymbol{v}(\boldsymbol{r})$ moves cells separated by a distance greater than $r_{\min}$ closer together (e.g., response to an attractive force) but moves cells separated by less than $r_{\min}$ further apart, corresponding to a close-ranged repulsive force. Hence, the cells that comprise the tissue adhere together with an average spacing of approximately $r_{\min}$. We choose this simple form for illustration; it simply makes cells clump together, independent of gradients in signaling molecule concentrations. Related kinematic interaction terms have been used in Lagrangian models of self-propelled/active particles [72,73], including cells [74]. The qualitative predictions in this example do not depend on the precise form of $\boldsymbol{g}(\boldsymbol{r}^\alpha)$.

Concentration of signaling molecule 1, c_1, regulates cell type. We assume rapid equilibration of signaling molecule concentrations (Sec. III A), with c_ℓ^{ss} given by Eq. (6) with z-independent $c_\ell^{(1)}(\boldsymbol{r}) = C_\ell \exp(-r^2/(2s_\ell^2))$ for $\ell = 1, 2$ and $r = |\boldsymbol{r}|$. While the concentration field $c_\ell^{(1)}(\boldsymbol{r})$ due to a single cell can be computed from Eq. (7), it can also be approximated by a Gaussian spread function with appropriately chosen variance s_ℓ^2 (see Table I). The concentration c_1 informs cells as to whether they are surrounded by many (high c_1) or few (low c_1) neighboring cells. The gene regulatory circuit ensures that cells with many neighbors become interior cells and cells with fewer neighbors become exterior cells.

As in Example 1, the concentration of molecule 2, c_2, mediates the total cell population by influencing cell division rates, which we model by

$$\tilde{\beta}(\boldsymbol{c}; \Delta\boldsymbol{r}', \Delta\boldsymbol{r}'') = \begin{cases} \beta_0 \delta(\Delta\boldsymbol{r}')\mathcal{N}(\Delta\boldsymbol{r}'') & c_2 < c_2^* \\ 0 & c_2 \geqslant c_2^*, \end{cases} \tag{27}$$

where $\mathcal{N}(\boldsymbol{r}) \equiv e^{-r^2/2\xi^2}/\sqrt{2\pi\xi^2}$ is the two-dimensional Gaussian density function with variance ξ^2. Note that the division rate is independent of z.

Figure 4(c) shows a simulation of the model. Starting from a single cell at $t = 0$, the organism grows into a fully developed organism in roughly four days. For approximately the first two days, all cells are of the exterior type (purple). After the organism grows sufficiently large that c_1 concentration reaches a sufficient level, conversion to interior-type cells (yellow) occurs. In Fig. 4(c), the top row indicates the positions and z states of cells, the middle row shows signaling molecule 1 concentrations, and the bottom row shows signaling molecule 2 concentrations.

If the fully developed organism is bisected, then each half regrows. In the simulation shown in Fig. 4(c), the organism is bisected at time $t = 4.0$ days along the dotted line shown in the figure, and the right-hand portion is removed. The remaining left-hand portion then regrows into a new fully developed organism in approximately 2 days. These results illustrate an extremely robust developmental process.

VII. SUMMARY AND CONCLUSIONS

We have systematically constructed a general model for tissue development that incorporates intracellular gene expression dynamics, the effects of gene expression state on proliferation and death rates, heritability of cellular states during cell division, and cell-cell interactions via chemical signaling. Dissecting our interacting-cell framework into individual components, we provide a mathematically unambiguous definition of a "Waddington landscape" (described as a Waddington vector field to allow for general active dynamics) and, if cell states are discretized, then a developmental "fitness landscape."

Our general framework integrates submodels for four processes: stochastic changes in gene expression within each cell of the tissue [Eq. (1)], chemical signals produced and received by cells affecting their gene expression dynamics [Eq. (2)], spatial movement of cells [Eq. (4)] which may affect their exposure to signaling molecules, and finally, stochastic cell proliferation [Eq. (10)] that allows daughter cells to acquire different states than their mother [through the differential birth rate density function $\tilde{\beta}(\boldsymbol{z}; \boldsymbol{c}; \boldsymbol{z}', \boldsymbol{z}'', \Delta\boldsymbol{r}', \Delta\boldsymbol{r}'')$]. These distinct processes are coupled together through the chemical signaling field $\boldsymbol{c}$, the drift and diffusivity of cells (which may depend on cell state $\boldsymbol{z}$ and chemical field $\boldsymbol{c}$), and the differential birth rate which may also depend on mother-cell state and local chemical fields. Our framework represents an intermediate level of coarse-graining that allows for better interpretability than fine-grained computational models (such as CompuCell 3D [75]) or high- and variable-dimensional kinetic theories [38,39] that are difficult to marginalize to obtain equations for interpretable and measurable quantities, particularly if cells interact.

While the importance of cell-cell interactions and differential cell division and death rates for robust development is well understood [41,76], the complexity of incorporating these features into mathematical models has been an obstacle in modeling studies. Many theoretical works study cell fate transitions for single cells [3,5,12,19–25] or for populations of dividing but noninteracting cells without spatial structure [30–35]. For example, cell-cell interactions mediated by a chemical signaling concentration field, as in Eq. (2), typically cannot be incorporated into kinetic equations [38,39,77] in a straightforward way. Models that include spatial structure governed by cell-cell interactions, but which do not incorporate tissue growth via coordinated cell division or death, have also been studied [46,78]. The models developed in Refs. [9,18] incorporate cell-cell interactions and differential cell division rates to describe coordinated tissue development; these models also provide a more detailed description of the cell cycle than our framework. However, unlike our framework, they describe the cell population as a continuous distribution of infinitely many cells rather than describing individual cells as discrete entities, and they do not incorporate spatial structure.

While our description of signaling among cells invokes freely diffusing components, many signaling and tissue development processes involve direct binding of

membrane-associated ligands and receptors, such as the Delta-Notch [52,53] and Ephrin-Eph [79] signaling pathways, that regulate development/differentiation and cell motility, respectively. Cell shape and contact-mediated forces have been treated using confluent cell-type models to study tissue-level mechanical properties [80]. Within our more detailed multiscale framework, cell-cell contact-mediated signaling can be incorporated by setting the signaling molecule source profile and decay rate so that the concentration is confined to a region with size comparable to typical cell size, mimicking its localization to a cell membrane. This approximation can be used to effectively describe short-ranged communication between cells. We applied our intermediate-grained stochastic dynamical framework to two biological canonical processes: stem-cell differentiation in a well-mixed environment and a spatially structured growth and regeneration process. Even though the examples in this work are simplified models intended to provide conceptual clarity of general mechanisms of robust tissue development, one should be able to apply our framework to many other developmental processes that involve chemical signaling such as morphogenesis, embryogenesis, organogenesis, and patterning [40,81,82].

DATA AVAILABILITY

The data that support the findings of this article are openly available [83].

APPENDIX A: MISAPPLICATION OF THE WADDINGTON LANDSCAPE EQUATIONS

The misapplication of Eqs. (14) and (15) to systems that do not satisfy the assumptions (i) and (ii) in Sec. IV can lead to artifacts in the reconstructed dynamics. As in Eq. (15), let $U_R(\boldsymbol{z}) = -\log p_{\rm ss}(\boldsymbol{z})$ denote a landscape reconstructed from the steady-state distribution of cell states $p_{ss}(\boldsymbol{z})$ generated by $\mathrm{d}\boldsymbol{z} = \boldsymbol{F}(\boldsymbol{z})\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{z})\mathrm{d}\mathbf{W}$. The reconstructed vector field $\boldsymbol{F}_R(\boldsymbol{z}) \equiv -\boldsymbol{\nabla} U_R(\boldsymbol{z})$ can have "false" attractors [attractors that are absent in $\boldsymbol{F}(\boldsymbol{z})$]. An example of this is shown in Fig. 5(a), with the false attractor indicated by a red square. Additionally, $\boldsymbol{F}_R(\boldsymbol{z})$ may fail to have attractors that do exist in $\boldsymbol{F}(\boldsymbol{z})$. An example of this is shown in Fig. 5(b), with the true attractor that disappears in the reconstructed vector field indicated by an orange square.

The toy examples in Fig. 5 were selected to illustrate the phenomenon of false and missing attractors and were not derived from a specific biochemical system. The vector field used in Fig. 5(a) is $\boldsymbol{F}(\boldsymbol{z}) = -\boldsymbol{\nabla} U(\boldsymbol{z}) + \mathbf{A}(\boldsymbol{z})$, where

$$U(\boldsymbol{z}) = \frac{z_1^4}{20} + \frac{z_2^4}{50} - 2.4\, e^{-(z_1+1)^2 - z_2^2} - e^{-(z_1-1)^2 - z_2^2}$$
$$\mathbf{A}(\boldsymbol{z}) = \left(\frac{3}{e^{3(z_2+0.1)}+1} - 1.5,\ 0 \right). \tag{A1}$$

The vector field used in Fig. 5(b) is

$$\boldsymbol{F}(z_1, z_2) = \left(\frac{a_1\, z_1^4}{1+z_1^4} + \frac{1}{1+z_2^4} - a_2\, z_1, \right.$$
$$\left. \frac{a_1\, z_2^4}{1+z_2^4} + \frac{1}{1+z_1^4} - a_3\, z_2 \right) \tag{A2}$$

with $a_1 = 1.1, a_2 = 1.3, a_3 = 0.7$.

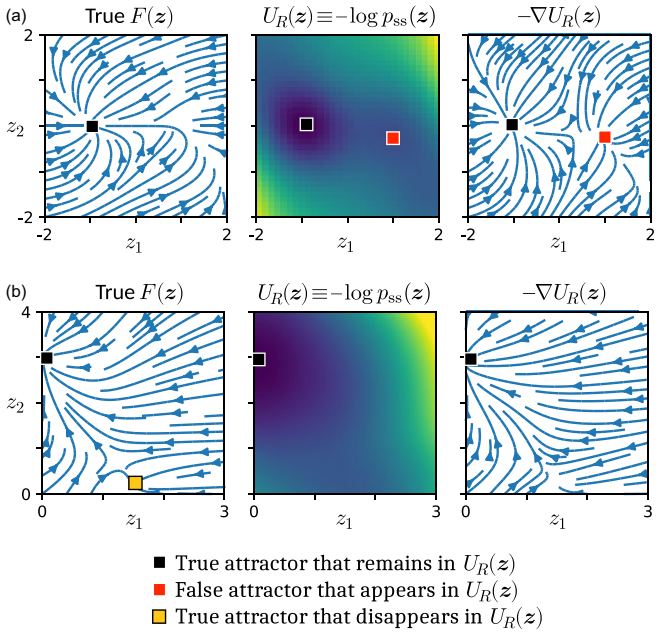

FIG. 5. Artifacts due to the misapplication of Eqs. (14) and (15). (a) An example showing a false attractor (red square) which does not exist in the true vector field but which appears in the reconstructed vector field. (b) An example showing a true attractor which disappears in the reconstructed vector field (orange square). The reconstructed landscapes $U_R(z)$ are color coded in which blue represents small values and yellow represents large values.

Additive noise was used for both examples in Fig. 5. For Fig. 5(a), $\sigma = \sqrt{2}$, and for Fig. 5(b), $\sigma = \sqrt{6}$. Previous work demonstrated a similar effect with multiplicative noise for one-dimensional systems [6].

APPENDIX B: MULTIPLICATIVE VERSUS ADDITIVE NOISE IN EXAMPLE 1

We derive the multiplicative noise (function of state value z) used in Example 1 and illustrate how the model's behavior changes if this noise is replaced with additive (constant) noise, with all else remaining unchanged. The gene regulation in Example 1 is described by dynamical production-decay processes for A and B molecules. Assume that the Waddington vector field $F_m(z, c)$ [defined in Eq. (16)] *without* the first order decay term $-k_m z_m$ represents a Markovian "production rate" $G_m(z; \boldsymbol{c}) = F_m(z; \boldsymbol{c}) + k_m z_m$ that includes molecule generation and consumption events but not the spontaneous, uncorrelated decay events. By using the Hill form for $F_m(z; \boldsymbol{c})$ as the approximate continuous production rate of A and B molecules, we are implicitly assuming that molecule production (e.g., through upregulated gene expression) is rate limiting as it responds to tetramer concentrations. The general SDE [Eq. (1)] can be regrouped with the convection term as the difference between production and decay, and the system-size-rescaled noise as the square root of the sum of the production and decay terms, to define an effective chemical Langevin equation associated with

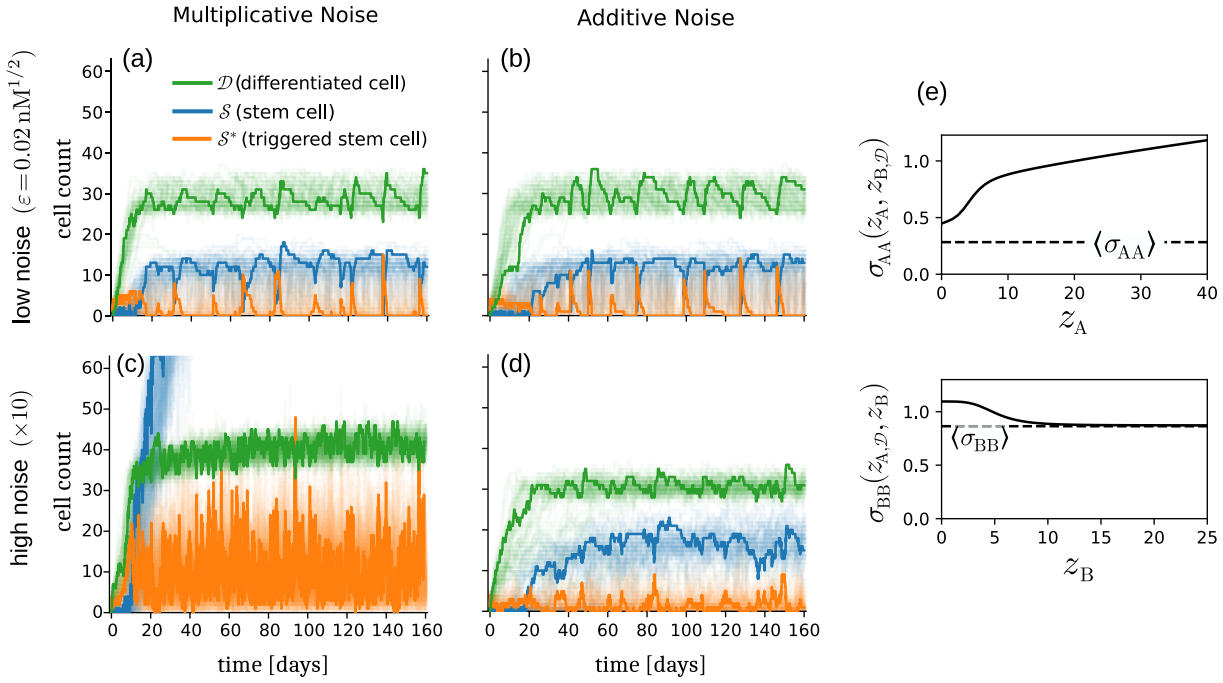

FIG. 6. Comparison of multiplicative (mult.) and additive (add.) noise in Example 1. [(a) and (b)] For low noise levels ($\varepsilon = 0.02\ \mathrm{nM}^{1/2}$), the choice of noise has very little effect on the population dynamics. [(c) and (d)] For larger noise [$\boldsymbol{\sigma}^{\mathrm{add}}(\varepsilon = 0.02) \to 10\boldsymbol{\sigma}^{\mathrm{add}}(\varepsilon = 0.02)$] the population regulation is destabilized under multiplicative noise, causing unbounded growth of $\mathcal{S}$ cells; additive noise maintains the stability. (e) Comparison of noise amplitudes as a function of cell state for multiplicative (solid) and additive (dashed) noise. Top panel shows σ_{AA} as a function of z_A with z_B set at its value in the $\mathcal{D}$ state ($z_{B,\mathcal{D}}$). Bottom panel shows σ_{BB} as a function of z_B with z_A set at its value in the $\mathcal{D}$ state ($z_{A,\mathcal{D}}$).

Example 1 [51]:

$$\begin{aligned} \mathrm{d}z_m &= (G_m(\boldsymbol{z};\boldsymbol{c}) - k_m z_m)\mathrm{d}t \\ &\quad + \varepsilon\sqrt{G_m(\boldsymbol{z};\boldsymbol{c}) + k_m z_m}\,\mathrm{d}W_m \\ &\equiv F_m(\boldsymbol{z};\boldsymbol{c})\mathrm{d}t + \varepsilon\sqrt{F_m(\boldsymbol{z};\boldsymbol{c}) + 2k_m z_m}\,\mathrm{d}W_m, \end{aligned} \tag{B1}$$

where $m \in \{\mathrm{A}, \mathrm{B}\}$, $\varepsilon \equiv 1/\sqrt{N_0 V}$, V is the cell volume, and N_0 is a reference count that defines the concentration unit (10^{-9} times Avogadro's number for units of nM). Hence,

$$\begin{aligned} \sigma_{m,m}(\boldsymbol{z};\boldsymbol{c}) &= \varepsilon\sqrt{F_m(\boldsymbol{z};\boldsymbol{c}) + 2k_m z_m} \\ \sigma_{m,m'}(\boldsymbol{z};\boldsymbol{c}) &= 0 \quad \text{for } m \neq m'. \end{aligned} \tag{B2}$$

In Example 1, we set $\varepsilon = 0.02\ \mathrm{nM}^{1/2}$, corresponding to a 4pL cell volume.

How does the model's behavior change if this multiplicative noise is replaced with additive noise? We compare the multiplicative noise model used in Example 1 to one with a modified, diagonal additive noise of the form

$$\boldsymbol{\sigma}^{\mathrm{add}} = \begin{bmatrix} \langle\sigma_{\mathrm{AA}}\rangle & 0 \\ 0 & \langle\sigma_{\mathrm{BB}}\rangle \end{bmatrix}, \tag{B3}$$

where $\langle\sigma_{\mathrm{mm}}\rangle$ are fixed to the value of $\varepsilon\sqrt{F_m(\boldsymbol{z};\boldsymbol{c}) + 2k_m z_m}$, averaged over all cells in the fully developed (steady state) tissue. Specifically, to assign a constant noise value that generates dynamics that can be compared properly with those generated with multiplicative noise, we first use the full multiplicative noise in Eq. (B2) to generate cell state trajectories and cell populations. Then we find the values of $\varepsilon\sqrt{F_m(\boldsymbol{z};\boldsymbol{c}) + 2k_m z_m}$ over all cells and average these values over a time window in the steady state (120–160 days). These specified values are then used as the value of the constant noise $\boldsymbol{\sigma}^{\mathrm{add}}$ [Eq. (B3)] to generate the associated additive noise trajectories and cell populations.

When noise is low ($\varepsilon = 0.02\ \mathrm{nM}^{1/2}$), there is very little difference between the additive and multiplicative noise models [Figs. 6(a) and 6(b)]. We explore how these dynamics change when noise is increased. Figures 6(c) and 6(d) show additive and multiplicative noise when the $\boldsymbol{\sigma}^{\mathrm{add}}$ found using ($\varepsilon = 0.02\ \mathrm{nM}^{1/2}$) is simply multiplied by a factor of 10. In this case, significant differences arise [Figs. 6(c) and 6(d)]. In the multiplicative noise case, larger noise amplitude causes $\mathcal{D}$ cells to transition to the $\mathcal{S}^*$ regime, which then divide to produce $\mathcal{S}$ and $\mathcal{D}$ cells, causing unbounded growth of $\mathcal{S}$ cells and disrupting population regulation. However, additive noise provides more robust regulation, as indicated in Fig. 6(d).

Figure 6(e) compares the noise magnitude in the case of the multiplicative (solid) and additive (dashed) noise with $\varepsilon = 0.2\ \mathrm{nM}^{1/2}$. The top panel shows σ_{AA} as a function of z_A with z_B set at its value in the $\mathcal{D}$ state (denoted $z_{B,\mathcal{D}}$). The lower panel shows σ_{BB} as a function of z_B with z_A set at its value in the $\mathcal{D}$ state. The noise experienced by $\mathcal{D}$ cells is stronger with multiplicative noise than with additive noise, explaining the greater noise-induced transition rate from $\mathcal{D}$ to $\mathcal{S}^*$.

APPENDIX C: SIMULATION METHODS

In examples in Secs. V and VI, we simulated the dynamics by discretizing time into increments $\mathrm{d}t$. Within each timestep, $\mathrm{d}\boldsymbol{z}^\alpha$ and $\mathrm{d}\boldsymbol{r}^\alpha$ are determined by Eqs. (1) and (4) associated with each cell. Each cell divides with probability $\beta\mathrm{d}t$ and/or

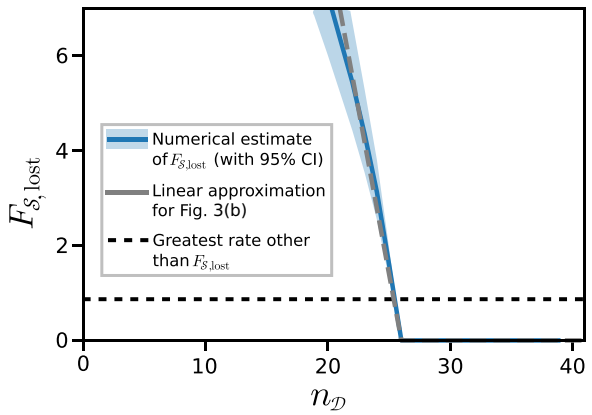

FIG. 7. The transition rate $F_{\mathcal{S},\text{lost}}(n_{\mathcal{D}})$ of stem cells to triggered stem cells in Example 1, estimated from simulations; 95% confidence intervals (CI) were estimated with the Clopper-Pearson method. A linear fit (dashed gray) provides a close approximation. Note that $F_{\mathcal{S},\text{lost}}$ rapidly transitions from very high [much higher than any other rate in Eq. (23)] to zero. The greatest rate in Eq. (12) other than $F_{\mathcal{S},\text{lost}}(n_{\mathcal{D}})$, which is $\beta_{\mathcal{S}}(0) \approx 0.9$ day^{-1}, is shown for comparison.

dies with probability μdt, where β and μ are determined as described in Sec. II D. Using this simple Monte Carlo procedure, a simulation of one trajectory in Example 1 (as shown in Fig. 3) takes approximately 3 s on a Macbook Pro M1 Pro laptop, while a realization in Example 2 [Fig. 4(c)] required approximately 12 s. Python codes to run these simulations are available. Simulations of more complex, higher-dimensional systems can be straightforwardly optimized by using, e.g., the Gillespie and related algorithms [84].

APPENDIX D: DETERMINISTIC LIMIT IN EXAMPLE 1

To numerically estimate $F_{\mathcal{S},\text{lost}}(n_{\mathcal{D}})$ in Eq. (23), simulations of the Waddington vector field dynamics were run at fixed values of $n_{\mathcal{D}}$ and with a collection of cells initially in state $\mathcal{S}$. The time at which each cell transitioned from $\mathcal{S}$ to $\mathcal{S}^*$ was recorded and the transition rate was estimated from this data. Code for this simulation and calculation is available [83].

Figure 7 shows the estimated $F_{\mathcal{S},\text{lost}}(n_{\mathcal{D}})$. A linear approximation (bounded below by zero) provides a close fit to $F_{\mathcal{S},\text{lost}}(n_{\mathcal{D}})$; we used this linear approximation to compute the deterministic dynamics depicted in Fig. 3(b).

[1] C. H. Waddington, Canalization of development and the inheritance of acquired characters, Nature (London) **150**, 563 (1942).

[2] J. E. Ferrell, Bistability, bifurcations, and Waddington's epigenetic landscape, Curr. Biol. **22**, R458 (2012).

[3] G. Schiebinger, Reconstructing developmental landscapes and trajectories from single-cell data, Curr. Opin. Syst. Biol. **27**, 100351 (2021).

[4] J. Wang, L. Xu, and E. Wang, Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations, Proc. Natl. Acad. Sci. USA **105**, 12271 (2008).

[5] D. A. Rand, A. Raju, M. Sáez, F. Corson, and E. D. Siggia, Geometry of gene regulatory dynamics, Proc. Natl. Acad. Sci. USA **118**, e2109729118 (2021).

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

[7] M. Sáez, R. Blassberg, E. Camacho-Aguilar, E. D. Siggia, D. A. Rand, and J. Briscoe, Statistically derived geometrical landscapes capture principles of decision-making dynamics during cell fate transitions, Cell Syst. **13**, 12 (2022).

[8] P. Zhou and T. Li, Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, A-type integral and beyond, J. Chem. Phys. **144**, 094109 (2016).

[9] J. Lei, Mathematical modeling of heterogeneous stem cell regeneration: From cell division to Waddington's epigenetic landscape, arXiv:2309.08064.

[10] A. Raju, B. Xue, and S. Leibler, A theoretical perspective on Waddington's genetic assimilation experiments, Proc. Natl. Acad. Sci. USA **120**, e2309760120 (2023).

[11] D. J. Cislo, M. J. Delás, J. Briscoe, and E. D. Siggia, Reconstructing Waddington's landscape from data, Proc. Natl. Acad. Sci. **122**, e2521762122 (2025).

[12] D. R. Sisan, M. Halter, J. B. Hubbard, and A. L. Plant, Predicting rates of cell state change caused by stochastic fluctuations using a data-driven landscape model, Proc. Natl. Acad. Sci. USA **109**, 19262 (2012).

[13] L. Edelstein-Keshet, *Mathematical Models in Biology* (SIAM, Philadelphia, PA, 1988).

[14] T. S. Gardner, C. R. Cantor, and J. J. Collins, Construction of a genetic toggle switch in *Escherichia coli*, Nature (London) **403**, 339 (2000).

[15] M. B. Elowitz and S. Leibler, A synthetic oscillatory network of transcriptional regulators, Nature (London) **403**, 335 (2000).

[16] B. Sinervo and C. M. Lively, The rock–paper–scissors game and the evolution of alternative male strategies, Nature (London) **380**, 240 (1996).

[17] L. Shahriyari and N. L. Komarova, Symmetric vs. asymmetric stem cell divisions: An adaptation against cancer? PLoS One **8**, e76195 (2013).

[18] J. Lei, A general mathematical framework for understanding the behavior of heterogeneous stem cell regeneration, J. Theor. Biol. **492**, 110196 (2020).

[19] 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).

[20] A. T. Fard, S. Srihari, J. C. Mar, and M. A. Ragan, Not just a colourful metaphor: Modelling the landscape of cellular development using Hopfield networks, npj Syst. Biol. Appl. **2**, 16001 (2016).

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

[22] A.-S. J. Frank, K. Larripa, H. Ryu, and S. Röblitz, Macrophage phenotype transitions in a stochastic gene-regulatory network model, J. Theor. Biol. **575**, 111634 (2023).

[23] Y. T. Nakamura, Y. Himeoka, N. Saito, and C. Furusawa, Evolution of hierarchy and irreversibility in theoretical cell differentiation model, PNAS NEXUS **3**, pgad454 (2023).

[24] G. Xue, X. Zhang, W. Li, L. Zhang, Z. Zhang, X. Zhou, D. Zhang, L. Zhang, and Z. Li, A logic-incorporated gene regulatory network deciphers principles in cell fate decisions, eLife **12**, RP88742 (2024).

[25] X. Chen, Y. Sheng, L. Chen, M. Tang, and F. Jiao, Quantifying cell fate change under different stochastic gene activation frameworks, Quant. Biol. **13**, e82 (2025).

[26] W. Wang, D. Poe, Y. Yang, T. Hyatt, and J. Xing, Epithelial-to-mesenchymal transition proceeds through directional destabilization of multidimensional attractor, eLife **11**, e74866 (2022).

[27] W. Wang, K. Ni, D. Poe, and J. Xing, Transiently increased coordination in gene regulation during cell phenotypic transitions, PRX Life **2**, 043009 (2024).

[28] F. Corson and E. D. Siggia, Gene-free methodology for cell fate dynamics during development, eLife **6**, e30743 (2017).

[29] E. Camacho-Aguilar, A. Warmflash, and D. A. Rand, Quantifying cell transitions in *C. elegans* with data-fitted landscape models, PLoS Comput. Biol. **17**, e1009034 (2021).

[30] C. Weinreb, S. Wolock, B. K. Tusi, M. Socolovsky, and A. M. Klein, Fundamental limits on dynamic inference from single-cell snapshots, Proc. Natl. Acad. Sci. USA **115**, E2467 (2018).

[31] D. S. Fischer, A. K. Fiedler, E. M. Kernfeld, R. M. J. Genga, A. Bastidas-Ponce, M. Bakhti, H. Lickert, J. Hasenauer, R. Maehr, and F. J. Theis, Inferring population dynamics from single-cell RNA-sequencing time series data, Nat. Biotechnol. **37**, 461 (2019).

[32] J. Shi, T. Li, L. Chen, and K. Aihara, Quantifying pluripotency landscape of cell differentiation from scRNA-seq data by continuous birth-death process, PLoS Comput. Biol. **15**, e1007488 (2019).

[33] S. Zhang, A. Afanassiev, L. Greenstreet, T. Matsumoto, and G. Schiebinger, Optimal transport analysis reveals trajectories in steady-state systems, PLoS Comput. Biol. **17**, e1009466 (2021).

[34] A. Forrow and G. Schiebinger, LineageOT is a unified framework for lineage tracing and trajectory inference, Nat. Commun. **12**, 4940 (2021).

[35] J. Shi, K. Aihara, T. Li, and L. Chen, Energy landscape decomposition for cell differentiation with proliferation effect, Natl. Sci. Rev. **9**, nwac116 (2022).

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

[37] 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).

[38] 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).

[39] M. Xia and T. Chou, Kinetic theories of state- and generation-dependent cell populations, Phys. Rev. E **110**, 064146 (2024).

[40] M. Barresi and S. Gilbert, *Developmental Biology*, 13th ed. (Oxford University Press, Oxford, 2023).

[41] D. E. Discher, D. J. Mooney, and P. W. Zandstra, Growth factors, matrices, and forces combine and control stem cells, Science **324**, 1673 (2009).

[42] S. Huang, Genetic and non-genetic instability in tumor progression: Link between the fitness landscape and the epigenetic landscape of cancer cells, Cancer Metastasis Rev. **32**, 423 (2013).

[43] S. Huang, Reconciling non-genetic plasticity with somatic evolution in cancer, Trend in Cancer **7**, 309 (2021).

[44] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer Series in Synergetics (Springer, Berlin, 1985).

[45] C. R. Doering, K. V. Sargsyan, and L. M. Sander, Extinction times for birth-death processes: Exact results, continuum asymptotics, and failure of the Fokker-Planck approximation, Multiscale Model. Simul. **3**, 283 (2005).

[46] H. Fooladi, P. Moradi, A. Sharifi-Zarchi, and B. H. Khalaj, Enhanced Waddington landscape model with cell–cell communication can explain molecular mechanisms of self-organization, Bioinformatics **35**, 4081 (2019).

[47] N. Kumar, T. Platini, and R. V. Kulkarni, Exact distributions for stochastic gene expression models with bursting and feedback, Phys. Rev. Lett. **113**, 268105 (2014).

[48] C. Jia and R. Grima, Small protein number effects in stochastic models of autoregulated bursty gene expression, J. Chem. Phys. **152**, 084115 (2020).

[49] X. Wang, Y. Li, and C. Jia, Poisson representation: A bridge between discrete and continuous models of stochastic gene regulatory networks, J. R. Soc. Interface. **20**, 20230467 (2023).

[50] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).

[51] D. T. Gillespie, The chemical Langevin equation, J. Chem. Phys. **113**, 297 (2000).

[52] S. Artavanis-Tsakonas, M. D. Rand, and R. J. Lake, Notch signaling: Cell fate control and signal integration in development, Science **284**, 770 (1999).

[53] G. L. Hunter, Z. Hadjivasiliou, H. Bonin, L. He, N. Perrimon, G. Charras, and B. Baum, Coordinated control of Notch/Delta signalling and cell cycle progression drives lateral inhibition-mediated tissue patterning, Development **143**, 2305 (2016).

[54] W. Xi, T. B. Saw, D. Delacour, C. T. Lim, and B. Ladoux, Material approaches to active tissue mechanics, Nat. Rev. Mater. **4**, 23 (2019).

[55] J. Ranft, M. Basan, J. Elgeti, J.-F. Joanny, J. Prost, and F. Jülicher, Fluidization of tissues by cell division and apoptosis, Proc. Natl. Acad. Sci. USA **107**, 20863 (2010).

[56] P. Dieterich, R. Klages, R. Preuss, and A. Schwab, Anomalous dynamics of cell migration, Proc. Natl. Acad. Sci. USA **105**, 459 (2008).

[57] T. H. Harris, E. J. Banigan, D. A. Christian, C. Konradt, E. D. T. Wojno, K. Norose, E. H. Wilson, B. John, W. Weninger, A. D. Luster, *et al.*, Generalized Lévy walks and the role of chemokines in migration of effector $CD8^+$ T cells, Nature (London) **486**, 545 (2012).

[58] F. Höfling and T. Franosch, Anomalous transport in the crowded world of biological cells, Rep. Prog. Phys. **76**, 046602 (2013).

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

[60] P. Greulich, B. D. MacArthur, C. Parigini, and R. J. Sánchez-García, Universal principles of lineage architecture and stem cell identity in renewing tissues, Development **148**, dev194399 (2021).

[61] S. Jang, S. Choubey, L. Furchtgott, L.-N. Zou, A. Doyle, V. Menon, E. B. Loew, A.-R. Krostag, R. A. Martinez, L. Madisen, B. P. Levi, and S. Ramanathan, Dynamics of embryonic stem cell differentiation inferred from single-cell transcriptomics show a series of transitions through discrete cell states, eLife **6**, e20487 (2017).

[62] F. Bouchet and J. Reygner, Generalisation of the Eyring–Kramers transition rate formula to irreversible diffusion processes, in *Annales Henri Poincaré,* Vol. 17 (Springer, Berlin, 2016), pp. 3499–3532.

[63] 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).

[64] C. Parigini and P. Greulich, Universality of clonal dynamics poses fundamental limits to identify stem cell self-renewal strategies, eLife **9**, e56532 (2020).

[65] P. Jain, R. Kizhuttil, M. B. Nair, S. Bhatia, E. W. Thompson, J. T. George, and M. K. Jolly, Cell-state transitions and density-dependent interactions together explain the dynamics of spontaneous epithelial-mesenchymal heterogeneity, iScience **27**, 110310 (2024).

[66] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[67] B. C. Goodwin, Oscillatory behavior in enzymatic control processes, Adv. Enzyme Regul. **3**, 425 (1965).

[68] S. Yu Pilyugin, *Morse—Smale Systems* (Birkhäuser Basel, Basel, CH, 1992), pp. 108–118.

[69] S. Huang, Y.-P. Guo, G. May, and T. Enver, Bifurcation dynamics in lineage-commitment in bipotent progenitor cells, Dev. Biol. **305**, 695 (2007).

[70] T. H. Morgan, *Experimental Studies of the Regeneration of Planaria Maculata* (W. Engelmann, Leipzig, 1898), Vol. 2.

[71] D. Lobo, W. S. Beane, and M. Levin, Modeling planarian regeneration: A primer for reverse-engineering the worm, PLoS Comput. Biol. **8**, e1002481 (2012).

[72] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, Collective memory and spatial sorting in animal groups, J. Theor. Biol. **218**, 1 (2002).

[73] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, Active Brownian particles, Eur. Phys. J.: Spec. Top. **202**, 1 (2012).

[74] B. A. Camley and W.-J. Rappel, Physical models of collective cell motility: From cell to tissue, J. Phys. D: Appl. Phys. **50**, 113002 (2017).

[75] M. Swat, G. L. Thomas, J. M. Belmonte, A. Shirinifard, D. Hmeljak, and J. A. Glazier, Multi-scale modeling of tissues using CompuCell3D, Methods Cell Biol. **110**, 325 (2012).

[76] B. G. Godard and C.-P. Heisenberg, Cell division and tissue mechanics, Curr. Opin. Cell Biol. **60**, 114 (2019).

[77] 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).

[78] M. Smart and A. Zilman, Emergent properties of collective gene-expression patterns in multicellular systems, Cell Rep. Phys. Sci. **4**, 101247 (2023).

[79] D. Arvanitis and A. Davy, Eph/ephrin signaling: networks, Genes Dev. **22**, 416 (2008).

[80] D. Bi, X. Yang, M. C. Marchetti, and M. L. Manning, Motility-driven glass and jamming transitions in biological tissues, Phys. Rev. X **6**, 021011 (2016).

[81] J. B. A. Green and J. Sharpe, Positional information and reaction-diffusion: Two big ideas in developmental biology combine, Development **142**, 1203 (2015).

[82] H. Meinhardt, Hierarchical inductions of cell states: A model for segmentation in *Drosophila*, J. Cell Sci. **1986**, 357 (1986).

[83] C. O. Barkan, Minimal models of biological tissue development (2026), https://github.com/cbarkan1/developmental-dynamics.

[84] D. T. Gillespie, Stochastic simulation of chemical kinetics, Annu. Rev. Phys. Chem. **58**, 35 (2007).