MULTISCALE MODEL. SIMUL.
Vol. 20, No. 1, pp. 404–432

A METHOD TO COARSE-GRAIN MULTIAGENT STOCHASTIC SYSTEMS WITH REGIONS OF MULTISTABILITY*

DARIA STEPANOVA†, HELEN M. BYRNE‡, PHILIP K. MAINI‡, AND TOMÁS ALARCÓN§

Abstract. Hybrid multiscale modeling has emerged as a useful framework for modeling complex biological phenomena. However, when accounting for stochasticity in the internal dynamics of agents, these models frequently become computationally expensive. Traditional techniques to reduce the computational intensity of such models can lead to a reduction in the richness of the dynamics observed, compared to the original system. Here we use large deviation theory to decrease the computational cost of a spatially extended multiagent stochastic system with a region of multistability by coarse-graining it to a continuous time Markov chain on the state space of stable steady states of the original system. Our technique preserves the original description of the stable steady states of the system and accounts for noise-induced transitions between them. We apply the method to a bistable system modeling phenotype specification of cells driven by a lateral inhibition mechanism. For this system, we demonstrate how the method may be used to explore different pattern configurations and unveil robust patterns emerging on longer timescales. We then compare the full stochastic, coarse-grained, and mean-field descriptions via pattern quantification metrics and in terms of the numerical cost of each method. Our results show that the coarse-grained system exhibits the lowest computational cost while preserving the rich dynamics of the stochastic system. The method has the potential to reduce the computational complexity of hybrid multiscale models, making them more tractable for analysis, simulation, and hypothesis testing.

Key words. large deviation theory, coarse-graining, phenotype pattern formation, multiscale modeling, hybrid modeling

AMS subject classifications. 60F10, 92C15, 92C42, 92B05, 92-08

DOI. 10.1137/21M1418575

1. Introduction.

When modeling a biological process, one has to make choices on how detailed the model should be in order to capture the characteristic features of the system. At the same time, the model should be as simple as possible in order to facilitate its analysis and numerical simulations. The evolution of systems with large numbers of agents (e.g., molecules, cells, species) can be described by the average behavior of their agents, or their mean-field limits using (ordinary or

*Received by the editors May 7, 2021; accepted for publication (in revised form) September 28, 2021; published electronically March 22, 2022.
https://doi.org/10.1137/21M1418575
Funding: The first author was supported by a grant of the Obra Social La Caixa Foundation on Collaborative Mathematics awarded to the Centre de Recerca Matemàtica through a scholarship. The first and fourth authors have been partially funded by the CERCA Programme of the Generalitat de Catalunya and by MINECO (https://www.ciencia.gob.es/) for funding under grants MTM2015-71509-C2-1-R and RTI2018-098322-B-I00. The first and fourth authors participate in project 2017SGR01735 which was awarded by AGAUR (https://agaur.gencat.cat/en/inici/index.html) but with no actual funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The second and third authors received no specific funding for this work.

†Centre de Recerca Matemàtica, Bellaterra (Barcelona) 08193, Spain, and Departament de Matemàtiques, Universitat Autònoma de Barcelona, Bellaterra (Barcelona) 08193, Spain (dstepanova@crm.cat).

‡Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK (Helen.Byrne@maths.ox.ac.uk, Philip.Maini@maths.ox.ac.uk).

§Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona 08010, Spain, Centre de Recerca Matemàtica, Bellaterra (Barcelona) 08193, Spain, and Departament de Matemàtiques, Universitat Autònoma de Barcelona, Bellaterra (Barcelona) 08193, Spain (talarcon@crm.cat).

partial) differential equations [6, 9, 35]. Dynamical systems theory provides methods and techniques for the analysis and numerical simulations of such systems. This description might become insufficient when the system comprises agents with internal variables that change in time, thus altering the agents' behavior, or when the system is not "large enough" to be described accurately by the mean-field equations. For these systems, stochastic descriptions are employed [39] (for example, continuous time Markov chains (CTMCs) or stochastic differential equations (SDEs)). In biological systems, the number of agents is finite and some level of noise is always present which can affect the system dynamics [39]. While exhibiting richer dynamics than deterministic systems, stochastic models are more computationally intensive.

Furthermore, in order to formulate a theoretical model of a biological phenomenon, it is often necessary to account for dynamics that act on different temporal and/or spatial scales [2, 17]. This has led to the development of hybrid multiscale models, in which different modeling techniques may be applied at each scale and then efficient coupling algorithms are used to integrate these models (see, e.g., [8, 16, 37] and references therein). In many of these models, individual entities (cells, species, etc.) are considered as discrete agents which are, themselves, equipped with models for their internal states determining the behavior (e.g., subcellular signaling, cell cycle, response to extracellular stimuli). Such models have great potential for generating insights into the behavior of a system (e.g., endothelial cell rearrangements [3], cell differentiation and tissue organization in intestinal crypts [8], and multiscale cancer modeling [10]). However, they frequently become numerically intractable because of their complexity (e.g., the internal dynamics of agents) [2]. This limits possible applications of these models.

In this work, we explain how to reduce the computational complexity of a hybrid model by coarse-graining the internal dynamics of its agents when these are described by a stochastic system with multiple steady states. The method involves applying large deviation theory (LDT) to reduce the dynamics of the stochastic system to a CTMC on the state space of its stable steady states. LDT provides a theoretical framework with which to quantify how small time-dependent fluctuations can lead to significant deviations from the mean-field behavior (*rare events*) such as transitions between stable steady states which cannot occur in deterministic systems [23]. This approach has previously been used to study rare, noise-induced events in individual stochastic systems [14, 15, 19, 39, 40, 42], but to our knowledge, this is its first application to a multiagent model.

In previous work, we developed a multiscale model of angiogenesis [45], the process of growth of new blood vessels from preexisting ones [28], which accounts for gene expression patterns (phenotypes) of endothelial cells at the subcellular scale. For prescribed levels of extracellular stimuli, the system is either monostable (i.e., only one cell phenotype exists) or bistable (i.e., two stable steady states, cell phenotypes, coexist). Cell phenotype is specified via contact-dependent cross-talk with neighboring endothelial cells via the VEGF-Delta-Notch signaling pathway [5, 24]. VEGF, or vascular endothelial growth factor, is the activating external stimulus; Delta and Notch are transmembrane ligands and receptors, respectively, which can trans-bind, (i.e., a ligand on one cell can bind to a receptor on another cell, thus allowing the two cells to "communicate"). Cells adjust their gene expression in order to maintain a pattern of two distinct phenotypes, Delta-high and Delta-low cells (see Figures 1(a) and 1(b)). We use the internal level of Delta as a proxy to distinguish between the phenotypes. In angiogenesis, the Delta-high (Delta-low) cells are referred to as tip (stalk) cells [5]. The number of transmembrane proteins in this signaling pathway is

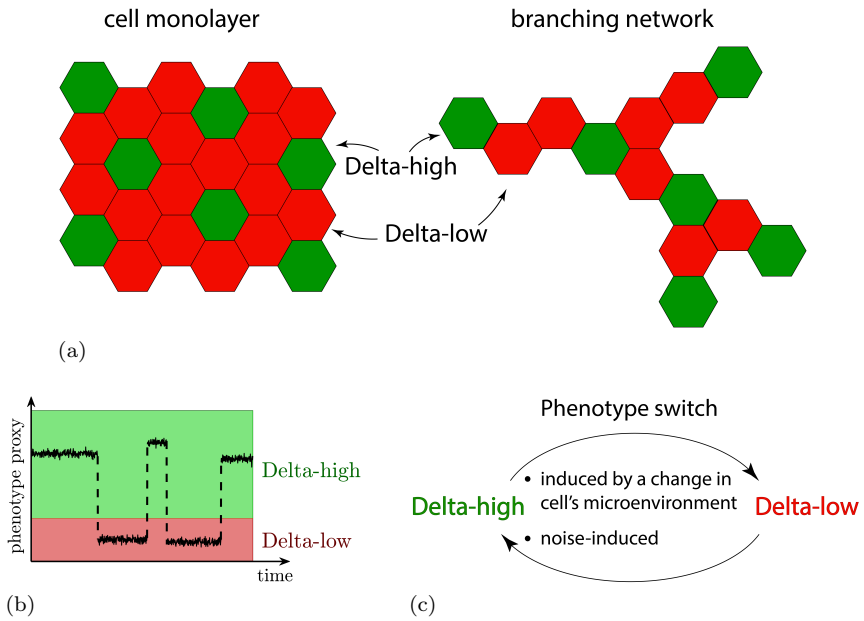


FIG. 1. *Cell phenotype specification.* (a) *Phenotype (Delta-high and Delta-low cells) patterning of cells induced by a mechanism of lateral inhibition in two different domains: a cell monolayer and a branching network.* (b) *Dynamic time evolution of phenotype adaptation of an individual cell. Using a phenotype proxy, e.g., level of Delta, allows for identification of a continuous cell phenotype.* (c) *Phenotype switches, as in* (b) *(dashed vertical lines), occur due to either a change in a cell's microenvironment or naturally present noise in intracellular signaling.*

on the order of thousands for each cell [6]. Therefore, in order to formulate a mathematical model, it is tempting to use deterministic mean-field equations to describe the kinetic reactions of this signaling pathway. However, deterministic descriptions cannot account for noise-induced transitions between stable steady states or, in the case of this signaling pathway, phenotypic switches, which can occur in regions of bistability (see Figures 1(b) and 1(c)). Since branching patterns of vascular networks are affected by the distribution of cells with different phenotypes, such phenotype transitions are potentially significant. Therefore, we modeled the subcellular signaling pathways stochastically, which increased the computational cost of the model. This example illustrates a general problem associated with computational and, in particular, hybrid models: in order to preserve emergent features of the system, such as continuous cell phenotypes and noise-induced phenotype switches, the model becomes computationally intractable for large lattice simulations. Instead of simulating the full system of stochastic kinetic reactions for the cell crosstalk, as was done in our model of angiogenesis [45], the coarse-graining (CG) technique reduces the subcellular system dynamics to a jump process involving phenotype switches (i.e., between stable steady states of the system). This allows us to preserve the continuous description of the steady states and noise-induced transitions between them, while substantially reducing the computational effort required for simulation.

We illustrate the CG method by reference to the subcellular model of the VEGF-Delta-Notch signaling pathway that defines cell phenotype. The core Delta-Notch signaling pathway plays a key role in phenotype adaptation in cell types which can form cell monolayers, such as epithelial sheets [36, 44], bristle patterning in *Drosophila* [11, 30, 13], and neural precursor cells [22]. In all of these biological processes, the

lateral inhibition mechanism of the Delta-Notch signaling pathway generates spatial patterns of cells with alternating fates (phenotypes). For the particular case of endothelial cells, we additionally take into account cell activation by a stationary distribution of an extracellular stimulus, VEGF. In other cell types, which use lateral inhibition to communicate, another external stimulus may play a similar role in strengthening the respective roles of cell phenotypes (e.g., Wnt-activity in epithelial cells in intestinal crypts [8]). In this paper we perform our simulations for two spatial geometries: a cell monolayer and a branching network (Figure 1(a)). For our model of multicellular VEGF-Delta-Notch signaling, we show typical simulation results of the coarse-grained system which allow us to explore different configurations of spatial patterns in a single realization of the model (due to phenotypic switches). We then demonstrate how this dynamic exploration of possible patterns may be used to uncover robust patterns emerging at long timescales. We finally compare the spatio-temporal dynamics and computational cost of the full stochastic CTMC, the coarse-grained and the deterministic mean-field descriptions. Our results show that the coarse-grained model, while preserving the continuous description of cell phenotype and rare events of phenotype switching, is more computationally efficient than the other two systems. Thus, it significantly reduces the computational complexity of the model without sacrificing the rich dynamics of the original stochastic system.

The remainder of the paper is organized as follows. In section 2, we review the hybrid (multiscale) modeling approach (subsection 2.1) and summarize LDT (subsection 2.2). This provides us with the information needed to formulate the coarse-grained model in section 3. In subsection 3.1, we start by coarse-graining the individual agent system and checking the accuracy of the method. We then extend the technique to a multiagent system in subsection 3.2, where we outline a general algorithm for formulating and simulating the coarse-grained model. In section 4, we present typical simulation results for the model of the VEGF-Delta-Notch signaling pathway (subsection 4.2) and compare the full stochastic, coarse-grained and mean-field systems via metrics which quantify the spatial patterns formed by the two cell phenotypes and we also compare computational cost of simulations (subsection 4.3). The paper concludes in section 5 with a summary of our findings and suggestions for future research directions.

2. Theoretical background.

2.1. Hybrid models. Biological systems are often highly complex, involving processes that may interact across multiple spatial and temporal scales (see Figure 2). From a general perspective, the subcellular scale is characterized by intracellular chemistry (e.g., gene expression, signal transduction, and receptor/ligand dynamics). Subcellular processes determine behavior at the cellular scale and may generate emergent properties at the tissue scale. In addition to this upward coupling across spatial scales, there is downward coupling whereby extracellular chemicals and biomechanical cues influence the subcellular chemistry/mechanics within a cell. In this way, dynamic interactions, encompassing all the scales, can occur (Figure 2).

From the theoretical perspective, models which consider only processes at a single spatial/temporal scale do not allow for investigation of emergent features which manifest at other scales (for example, collective migration or phenotype patterning which arises from individual cell dynamics and governs tissue scale organization). Equally, difficulties associated with the physical interpretation of parameters in phenomenological models, i.e., large-scale models which capture the overall evolution of a biological

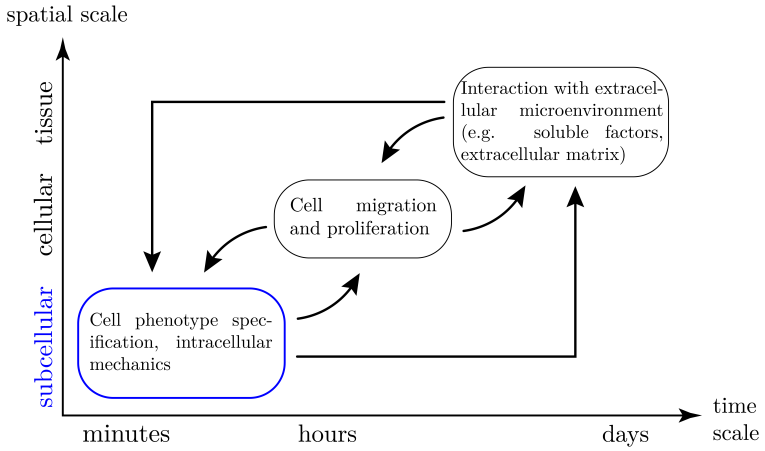


FIG. 2. *A schematic diagram illustrating characteristic spatial and temporal scales of a typical biological process and coupling between them. The VEGF-Delta-Notch signaling pathway, which serves as an illustrative example for application of the CG method, acts at the subcellular scale (highlighted in blue) on a timescale shorter than other processes (e.g., cell migration, cell-extracellular matrix interaction at the tissue scale) involved in the multiscale model of angiogenesis* [45]. *As a result, we may use LDT theory to coarse-grain its dynamics.*

process, make it challenging to fit the model to biological data. In particular, this abstract parameter construct hinders model calibration/validation and limits potential applications of the models. Multiscale models, which couple processes at different spatial and/or temporal scales, have the potential to address these issues [4].

A challenge in formulating a multiscale model relates to the number of entities (protein, cells, extracellular components, etc.) that should be included at each scale of interest. Using the same mathematical formalism to model processes involving entities which vary in number by several orders of magnitude may lead to the omission of essential features or make the model computationally intractable. Hybrid approaches are increasingly being recognized as suitable tools for trying to overcome problems of this type and have become a key part of multiscale modeling [16, 17]. The central idea is to employ the modeling framework most suitable to each subprocess and then to couple them. For example, the extracellular environment and signaling cues are usually modeled deterministically due to the large number of proteins involved. On the other hand, cells may be treated as individual entities, equipped with a subcellular model which determines their behavior (e.g., proliferation, cell polarity and migration). This framework has been used to develop multiscale models of cancer (see the reviews [16, 41] and references therein), angiogenesis [28], and collective cell migration [17], among other examples [2].

Hybrid modeling allows for efficient parameter estimation and model visualization, forging interdisciplinary collaboration between researchers in theoretical modeling and experimental biology [2, 37]. There is also the potential of using high-throughput experimental data to develop more detailed multiscale models. As an example, one of the aspects of biological systems that has received little attention in theoretical modeling is the effect of stochasticity in the response of individual entities to external stimuli [17]. Hybrid modeling allows investigation of this effect on the collective, emergent behavior. However, increasing computational complexity makes these models intractable for large-scale simulations [16].

This challenge motivated us to develop a technique which reduces the computational complexity of a model while preserving its stochasticity. The method is

applicable to systems characterized by stochastic processes which exhibit multistability and which evolve on timescales shorter than those associated with other system processes. The example that we study in this paper is of this type: the subcellular dynamics of cell fate determination via lateral inhibition (a bistable, stochastic system) act on a shorter timescale than those associated with, for example, cell migration, and tissue scale processes such as the dynamics of extracellular soluble factors (e.g., diffusion, secretion by cells, degradation) [28] (Figure 2). This observation motivates us to use LDT to coarse-grain the dynamics associated with intracellular signaling to produce a jump process (i.e., a Markov chain) on the stable state space of the steady states of the original system which describes the VEGF-Delta-Notch pathway.

2.2. Large deviation theory. In the presence of noise, small fluctuations can drive significant deviations from mean-field behavior such as, for example, transitions from one stable steady state to another. These transitions are usually referred to as *rare events* since their likelihood is small. LDT is predicated on the assumption that when rare events occur, the system follows the least unlikely paths. Deviations from these paths occur with very small probability (i.e., smaller than the probability of a rare event). Specifically, the Freidlin–Wentzell theory of large deviations predicts that the deviations are exponentially suppressed [23], making such transitions "predictable." LDT provides the means to analyze the frequency of rare events and to identify the maximum likelihood path (minimum action path (MAP)) along which these transitions can occur.

An SDE of a diffusion process, $x^\epsilon \in \mathbb{R}^n$, has the following form:

$$\mathrm{d}x^\epsilon(t) = b(x^\epsilon)\mathrm{d}t + \sqrt{\epsilon}\sigma(x^\epsilon)\mathrm{d}W, \tag{2.1}$$

where $b : \mathbb{R}^n \to \mathbb{R}^n$ is a drift vector, $a(x^\epsilon) = (\sigma\sigma^T)(x^\epsilon)$ is a diffusion tensor ($\sigma : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$, m corresponds to the number of kinetic reactions in the system), W is a Wiener process in $\mathbb{R}^m$, and $\epsilon = \Omega^{-1}$ is noise amplitude.

The mean-field limit of (2.1), $x(t) \in \mathbb{R}^n$, solves the following differential equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = b(x). \tag{2.2}$$

Assume that (2.2) has two stable steady states, $x_1,\ x_2 \in \mathbb{R}^n$, whose basins of attraction form a complete partition of $\mathbb{R}^n$. We are interested in transitions from $x_1 \to x_2$ (and $x_2 \to x_1$) which cannot be accounted for unless noise is present in the system.

A key player in LDT is the action functional

$$S_T(\psi) = \begin{cases} \displaystyle\int_0^T L(\psi, \dot{\psi})\, dt & \text{if } \psi \in C(0,T) \text{ is absolutely continuous and} \\ & \text{the integral converges,} \\ +\infty & \text{otherwise,} \end{cases}$$

which is computed for a transition path $\psi : [0, T] \to \mathbb{R}^n$ from x_1 to x_2 ($\psi(0) = x_1$ and $\psi(T) = x_2$, T is the transition time). Here, $\dot{\psi}$ denotes the gradient of the transition path, ψ; $L(x, y) = \sup_{\theta\in\mathbb{R}^n} (\langle y, \theta\rangle - H(x, \theta))$ is the large deviation Lagrangian, with $\langle\cdot,\cdot\rangle$ being the Euclidean scalar product in $\mathbb{R}^n$ and $H(x, \theta)$ being the Hamiltonian associated with $L(x, y)$. The particular form of the Hamiltonian depends on the dynamical system under consideration (in section SM2, we explain how to define the Hamiltonian for an SDE such as (2.1) and a general birth-death CTMC).

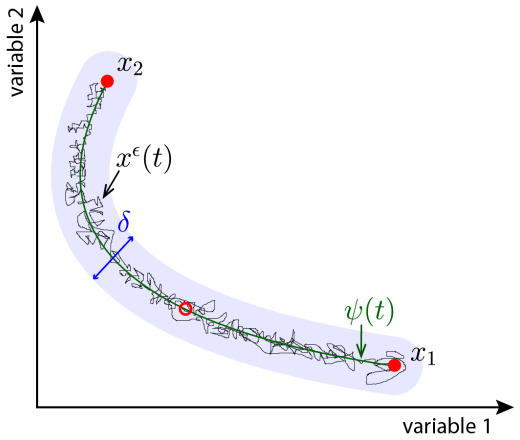


FIG. 3. *An illustration of a transition path between two stable steady states of an arbitrary bistable system. The two stable steady states, x_1 and x_2, are marked by filled red circles; an unstable saddle point is marked by an unfilled red circle. The transition path, $\psi(t)$, from x_1 to x_2 is shown by a thick green line, whereas a single stochastic trajectory, $x^\epsilon(t)$, is indicated by a thin black path. The shaded blue region indicates a δ-neighborhood around $\psi(t)$ (δ as defined in* (2.3)*).*

The action functional is used to estimate the probability that a trajectory $x^\epsilon(t)$ lies in a narrow neighborhood, of width $\delta > 0$, of a given path $\psi \in C(0,T)$ (see Figure 3 for an illustration):

$$\mathrm{P}\left\{ \sup_{0\le t\le T} \mid x^\epsilon(t) - \psi(t) \mid < \delta \;\middle|\; x^\epsilon(0) = x_1 \right\} \approx \exp\{-\epsilon^{-1} S_T(\psi)\}. \tag{2.3}$$

Since the probability function in (2.3) decreases as the action functional, $S_T(\psi)$, increases, the maximum likelihood path, ψ^*, is the minimizer of $S_T(\cdot)$. This leads naturally to the idea of the quasipotential:

$$V(x_1, x_2) = \inf_{T>0} \inf_{\psi \in \overline{C}^{x_2}_{x_1}(0,T)} S_T(\psi). \tag{2.4}$$

Here $\overline{C}^{x_2}_{x_1}(0,T)$ is the space of absolutely continuous functions $f : [0,T] \to \mathbb{R}^n$ such that $f(0) = x_1$ and $f(T) = x_2$. Roughly speaking, the quasipotential gives an estimate of how "difficult" it is to move from x_1 to x_2. Thus, the quasipotential value depends on the direction of a transition path and, in general, $V(x_1, x_2) \neq V(x_2, x_1)$.

On timescales which are much longer than those associated with relaxation to a stable steady state, the dynamics of (2.1) can be reduced, or *coarse-grained*, to that of a CTMC on the state space of the two stable steady states, $\{x_1, x_2\}$, with transition rates [23, 47]

$$k_{x_1 \to x_2} \asymp \exp\left(-\epsilon^{-1} V(x_1, x_2)\right), \qquad k_{x_2 \to x_1} \asymp \exp\left(-\epsilon^{-1} V(x_2, x_1)\right). \tag{2.5}$$

Here the symbol $\asymp$ denotes log-asymptotic equivalence so that $f(\epsilon) \asymp g(\epsilon)$ if and only if $\lim_{\epsilon\to 0} \frac{\log f(\epsilon)}{\log g(\epsilon)} = 1$.

In practice, most double minimization problems, such as (2.4), do not have a solution for finite $T > 0$. Furthermore, closed-form Lagrangians exist for SDEs of the type defined by (2.1) but not for general birth-death CTMCs. (2.4) can be reformulated in terms of a Hamiltonian system of the form

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\partial H(\phi,\theta)}{\partial \theta}, \qquad \frac{\mathrm{d}\theta}{\mathrm{d}t} = -\frac{\partial H(\phi,\theta)}{\partial \phi}.$$

This problem must be solved as a boundary-value problem, i.e., $\phi(0) = x_1$ and $\phi(T) = x_2$, on an infinite time interval, $T \to \infty$, [26] which makes it a nontrivial numerical problem. Thus the traditional LDT methods are inapplicable in most cases.

One way to resolve these problems is to reformulate the minimization problem defined by (2.4) on the space of curves (i.e., transition paths from one stable steady state to another). In [29], Heymann and Vanden-Eijnden proved that the minimization problem defined by (2.4) is equivalent to

$$V(x_1,x_2) = \inf_{\phi} \widehat{S}(\phi) \quad \text{with} \quad \widehat{S}(\phi) = \sup_{\substack{\hat{\theta}:[0,1]\to\mathbb{R}^n \\ H(\phi,\hat{\theta})=0}} \int_0^1 \langle \phi', \hat{\theta} \rangle \, d\alpha, \tag{2.6}$$

where $\phi : [0,1] \to \mathbb{R}^n$ is a curve from x_1 to x_2 parametrized by standard arc length.

The geometric reformulation, (2.6), resolves analytically the issue of the infinite time, T, in the original minimization problem. Furthermore, only the Hamiltonian is needed. In this respect, the method is more general as it can be applied to SDEs, CTMCs, and other systems for which the Hamiltonian is known (see section SM2 in Supplementary Material).

In [29], an algorithm was developed to efficiently compute $V(x_1,x_2)$ and the corresponding minimizer, ϕ^*, from the geometric reformulation. The algorithm is known as the geometric minimum action method (gMAM) and the minimizer, ϕ^*, of the action functional is referred to as the minimum action path, or MAP (for more details see section SM2).

Once the quasipotential has been computed, the coarse-grained system is given by a CTMC, with rates defined by (2.5).

3. Coarse-graining. We now illustrate how the theory described in the previous section can be used to coarse-grain a specific hybrid multiscale model, one for which the internal dynamics of the agents are described by multistable stochastic systems. This property is characteristic of, for example, systems driving cell fate (phenotype) determination. We begin by using LDT to formulate a CG model for a system comprising a single agent (here a cell). The subcellular signaling pathway, which we use to illustrate the method, is the VEGF-Delta-Notch pathway (see section SM1 in Supplementary Material and [45] for details). This pathway regulates phenotypic adaptation via lateral inhibition [12, 36]. This system meets the requirements for application of the CG technique: (a) it is bistable; its stable steady states are associated with cellular phenotypes (Delta-high and Delta-low cells); (b) we are interested in its evolution on timescales longer than the typical time for relaxation to an equilibrium since other processes (e.g., cell migration and dynamics of extracellular matrix) act on longer timescales (see Figure 2).

We then extend the method to the general case of multiagent systems. Here the dynamics of each entity is coarse-grained to a CTMC on the state space of its stable states, and coupling between the internal dynamics of individual agents is achieved via the external variables whose dynamics depend on the states of neighboring agents and/or the time evolution of these variables. We outline below how we apply this method to a monolayer of cells (motivated by phenotype patterning via the core Delta-Notch pathway in cell monolayers [36]) and a branching network (angiogenesis-motivated application [45]) that interact via VEGF-Delta-Notch signaling.

3.1. Individual agent system. Our algorithm for coarse-graining a stochastic system with a region of multistability involving a single entity is illustrated in Figure 4. For the particular case of VEGF-Delta-Notch signaling, a cell's internal state (phenotype) depends on two model parameters (inputs) corresponding to the extracellular levels of Delta and Notch, $v = (d_{ext}, n_{ext}) \in \mathbb{R}^2$ (corresponding to the levels of Delta and Notch, respectively, that the cell under consideration perceives from the cells in its external microenvironment; see section SM1). We fix the values of the model parameters and the external variables, v (see Table SM3). We then use the mean-field system defined by equation (SM1.2) to compute the steady state solutions. For this example, the values of the external variables, v, are chosen so that the system is bistable; the two stable steady states correspond to Delta-high and Delta-low cell phenotypes, $\{x_1, x_2\} = \{\text{Delta-high, Delta-low}\}$, and the unstable steady state is an unstable saddle. Our goal is to compute the transition rates of the CG system which we approximate as follows:

$$k_{x_s \to x_l} \approx C_{x_s \to x_l} \exp\left(-\Omega V(x_s, x_l)\right), \quad s,\ l \in \{1, 2\},\ s \neq l. \tag{3.1}$$

We note that the prefactor, $C_{x_s \to x_l}$, arises from the asymptotic equivalence relation defined by (2.5) [23, 39, 34]. The system size is given by $\Omega = \epsilon^{-1}$, where ϵ is the noise level.

We use the gMAM to compute the quasipotential values and corresponding paths (MAPs) for transitions between the Delta-high and Delta-low phenotypes (for more details, see section SM4 in Supplementary Material). An illustrative example is shown in Figure 5, where we compare the MAPs and sample paths of the full stochastic CTMC for an individual cell (see also Table SM2 in section SM1). Several characteristic features of the phenotype transitions are noteworthy. First, the dynamics of the MAP can be split into two parts: the transition from the steady state of origin to the saddle point (for example, from the Delta-low phenotype to the saddle point, indicated

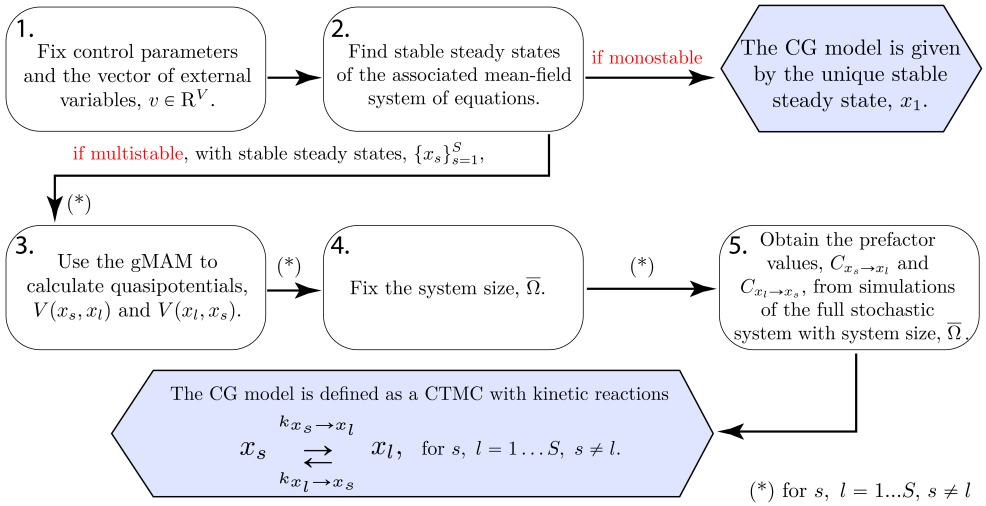


FIG. 4. *A flowchart of the procedure used to coarse-grain a multistable stochastic system for an individual entity. The steady state solutions, quasipotential, and prefactor depend on the model parameters and external variables, $v \in \mathbb{R}^V$ (V indicates the dimension of the vector of external variables). Here the transition rates, $k_{x_s \to x_l}$, are defined by* (3.1)*, the prefactor, $C_{x_s \to x_l}$, is determined from* (3.2b)*, and $\overline{\Omega}$ is given by* (3.3).

by the blue circle in Figure 5(a)) which is possible due to the presence of noise. The main contribution to the quasipotential comes from this transition. The MAP from the unstable saddle point to the stable steady state of destination (from the saddle point indicated by the blue circle to the Delta-high phenotype in Figure 5(a)) follows the fastest route given by the deterministic heteroclinic orbit connecting the steady states (i.e., the unstable saddle and the stable Delta-high cell state). For systems that possess a single unstable saddle point and no other limit sets such as periodic orbits, it is possible to show that the MAP crosses the separatrix at the unique saddle point (see, e.g., [23, 46, 26] and references therein). Thus, Figure 5 confirms the accuracy of the implemented gMAM for the system of interest. The second noteworthy feature of the phenotype transitions is that, as the level of noise, ϵ, decreases, the stochastic sample path follows the MAP more closely (compare Figures 5(a) and 5(b) for which $\Omega = \epsilon^{-1} = 70$ and $\Omega = \epsilon^{-1} = 450$, respectively). In addition, Figure SM4 illustrates the corresponding transition tubes (tubular neighborhoods around the MAPs within which transitions between steady states occur) for these phenotype transitions.

To fully determine the CG transition rates, the prefactor value, $C_{x_s \to x_l}$, must be estimated. From (3.1), for $s,\ l \in \{1, 2\},\ s \neq l$, we have

$$\log\langle T^{\Omega}_{x_s \to x_l}\rangle \approx \Omega V(x_s, x_l) - \log C_{x_s \to x_l}\ , \tag{3.2a}$$

$$\log C_{x_s \to x_l} \approx \Omega V(x_s, x_l) - \log\langle T^{\Omega}_{x_s \to x_l}\rangle\ , \tag{3.2b}$$

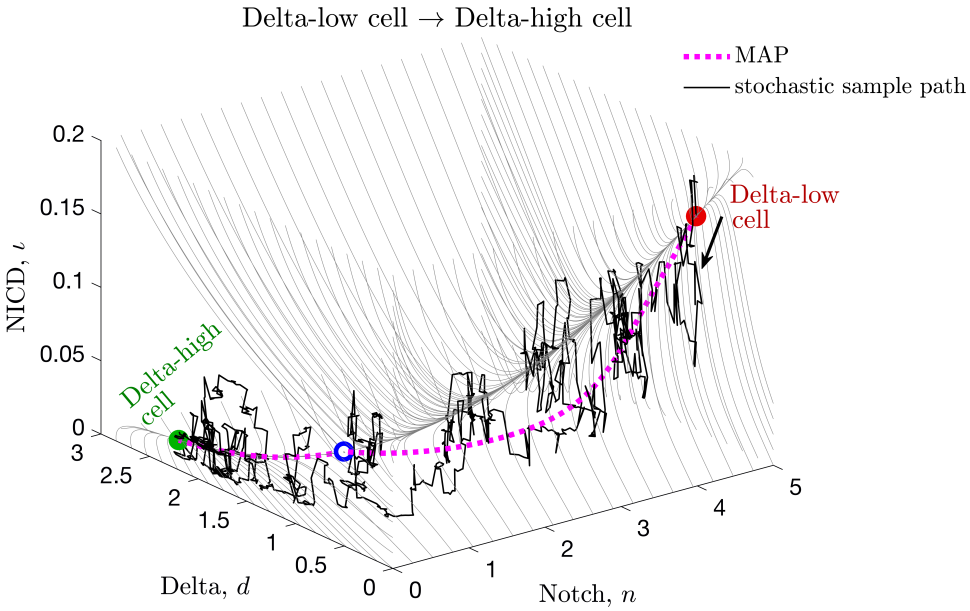
where $\langle T^{\Omega}_{x_s \to x_l}\rangle = 1/k_{x_s \to x_l}$ is the mean passage time between the stable steady states, x_s and x_l (Delta-high and Delta-low phenotypes), for a fixed value of the system size, Ω. $\langle T^{\Omega}_{x_s \to x_l}\rangle$ can be determined from direct simulation of the full stochastic model using the reaction kinetics given in Table SM2.

An accurate estimate of the quasipotential (as obtained via the gMAM) allows us to obtain the prefactor given the mean passage time, $\langle T^{\Omega}_{x_s \to x_l}\rangle$, for a single value of the system size, Ω. However, the approximate relation in (3.2) is valid in the limit $\Omega \to \infty$ (see Figure 6). Thus, Ω should be chosen sufficiently large to achieve convergence in (3.2) and, at the same time, not too large in order to ensure that transitions between the phenotypes occur in a computationally feasible time, since the waiting times for transitions between stable steady states increase exponentially as Ω grows. Specifically, since larger values of the quasipotential, $V(x_s, x_l)$, in (3.2) lead to longer mean passage times, a maximum simulation time, T_{max}, can be determined computationally by simulating the original stochastic system for the values of the external variables, v, for which the quasipotential is large (either $V(x_s, x_l)$ or $V(x_l, x_s)$). For the VEGF-Delta-Notch signaling pathway, the quasipotential is characterized by larger values close to the border of the bistability region. Thus, we performed several realizations of the full stochastic system, choosing several values of v for which at least one of the quasipotentials is large, and recorded the maximum simulation time, T_{max}, and the average prefactor estimate obtained, $\bar{C}$. The corresponding system size, $\overline{\Omega}$, is then approximated as follows:

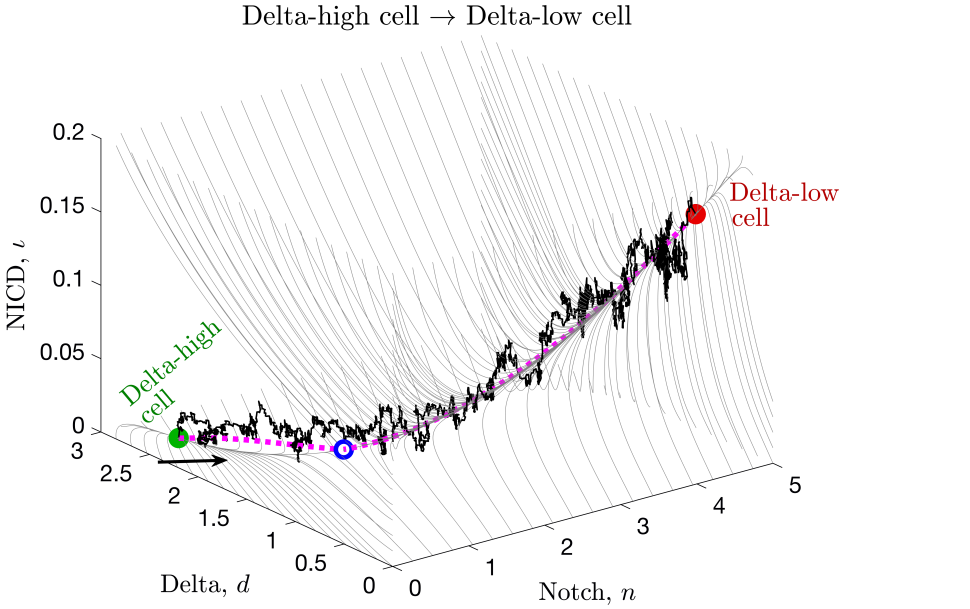
$$\overline{\Omega} \approx \frac{\log T_{max} + \log \bar{C}}{V(x_s, x_l)}. \tag{3.3}$$

Then the prefactor, $C_{x_s \to x_l}$, can be approximated using (3.2b) with $\Omega = \overline{\Omega}$.

From (3.2a), we know that $\log\langle T^{\Omega}_{x_s \to x_l}\rangle$ is a linear function of Ω whose slope and intercept are given by the quasipotential, $V(x_s, x_l)$, and $(-\log C_{x_s \to x_l})$, respectively. Thus, in order to check the accuracy of our estimate for the system size, $\overline{\Omega}$ (3.3), we



(a) system size, $\Omega = 70$ (noise level, $\epsilon = 1/\Omega \approx 0.014$)



(b) system size, $\Omega = 450$ (noise level, $\epsilon = 1/\Omega \approx 0.002$)

FIG. 5. *An illustration of the MAPs and stochastic sample paths for transitions between the Delta-high and Delta-low cell phenotypes. We computed the MAPs (indicated by the dotted magenta lines) for the subcellular VEGF-Delta-Notch system in an individual cell using the gMAM for transitions from* (a) *Delta-low to Delta-high cell and* (b) *Delta-high to Delta-low cell. The stochastic sample paths obtained by simulating the full stochastic CTMC model (Table* SM2*) with the system sizes* (a) $\Omega = 70$, (b) $\Omega = 450$, *are plotted in black. The thin gray lines indicate streamlines of the corresponding mean-field system (equation* (SM1.2)*). The Delta-high (Delta-low) cell stable steady state is indicated by a green (red) filled circle, the unstable saddle by a blue unfilled circle. The plots represent three-dimensional projections of the full five-dimensional system as defined by equation* (SM1.2)*. Parameter values are fixed as indicated in Table* SM3.

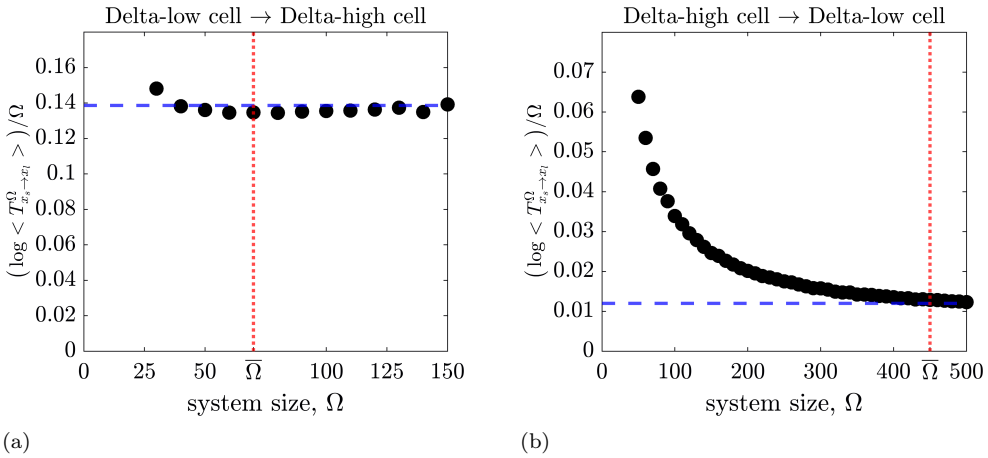


FIG. 6. *Convergence of the quasipotential, $V(x_s, x_l)$, as the system size, Ω, increases. We ran* 1000 *realizations of the stochastic VEGF-Delta-Notch model for an individual cell (see Table* SM2*) for fixed values of $d_{ext} = 0.2$, $n_{ext} = 0.5$ and increasing system size, Ω. We plotted the convergence to the quasipotential value* (a) $V(\text{Delta-low}, \text{Delta-high})$ *and* (b) $V(\text{Delta-high}, \text{Delta-low})$ *as a function of Ω (black circle markers). For these parameter values, transitions from the Delta-low to the Delta-high phenotype are less likely to occur (higher noise levels, $\epsilon = \Omega^{-1}$, and/or longer transition times are needed) than transitions from the Delta-high to Delta-low phenotype (see* (3.2a)*). Therefore, the perturbations of this random event are smaller and convergence is reached for higher values of noise. This is why lower values of Ω in* (a) *suffice to accurately determine the prefactor value from* (3.2)*. The blue dashed lines indicate the value of the corresponding quasipotential computed via the gMAM; the red dotted lines indicate $\overline{\Omega}$ from* (3.3)*. All other parameter values are fixed as indicated in Table* SM3*.*

compared linear fitting of data obtained from the full stochastic CTMC model for increasing Ω, with the estimate obtained from the gMAM quasipotential and the prefactor extracted from simulations with system size, $\overline{\Omega}$. The results presented in Figure 7 show that the estimates converge as Ω increases, confirming the accuracy of the two methods.

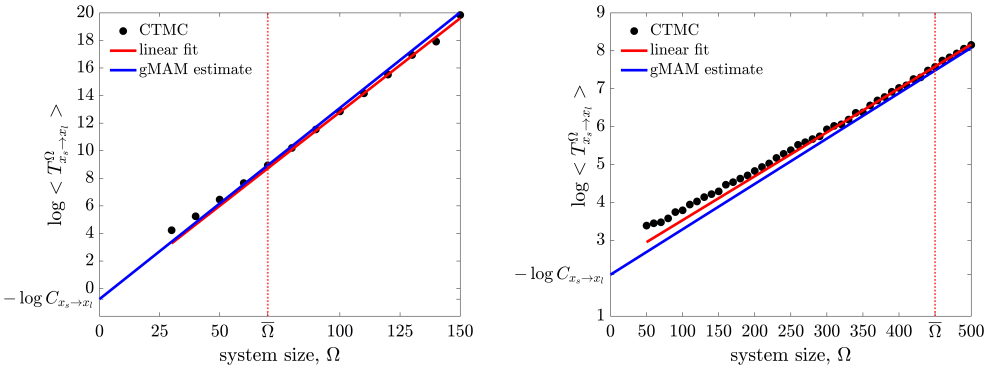
To summarize, we coarse-grain the stochastic VEGF-Delta-Notch dynamics as follows (see Figure 4):

I Fix the model parameter values and the vector of external variables, v, which, for this system, is given by the extracellular levels of Delta and Notch, $v = (d_{ext}, n_{ext})$.

II Compute the steady states of the corresponding mean-field system (equation (SM1.2)).

III Formulate the CG model:

- i If, for the given $v = (d_{ext}, n_{ext})$, the system is monostable (either Delta-high or Delta-low cell steady state exists), then the quasipotential value to arrive at this state is 0. The value of the other quasipotential can be assumed infinite (since the system is monostable, this transition is impossible). For example, if the only stable steady state is the Delta-high cell, then $V(\text{Delta-low}, \text{Delta-high}) = 0$ and $V(\text{Delta-high}, \text{Delta-low}) = \infty$. The CG model is defined by its unique stable steady state.
- ii If the system is within the bistable regime (both Delta-high and Delta-low steady states are stable), then the CG model is defined as a CTMC on the state space of $\{x_s, x_l\} = \{\text{Delta-high}, \text{Delta-low}\}$. The transition



(a) Delta-low → Delta-high (b) Delta-high → Delta-low

FIG. 7. *Prefactor estimation. Comparison of prefactor estimates obtained from simulations of the full stochastic CTMC model (black circles) and estimates obtained using the gMAM-quasipotential and mean passage times for a single value of the system size,* $\overline{\Omega}$ *(blue line); see* (3.2a)*. The linear fit of the full stochastic data (red line) was performed for values of* Ω *such that the corresponding sample* $\{T^{\Omega}_{x_s \to x_l}\}$ *is exponentially distributed (high levels of noise might affect the distribution of these transitions). Panel* (a) *corresponds to the transition from Delta-low to Delta-high phenotype; panel* (b) *corresponds to the transition from Delta-high to Delta-low phenotype. The red dotted lines indicate* $\overline{\Omega}$ *from* (3.3)*. All other parameter values are fixed as indicated in Table* SM3*.*

rates are given by (3.1). The quasipotential, $V(x_s, x_l)$, is approximated using the gMAM; the prefactor value, $C_{x_s \to x_l}$, is obtained via (3.2b) from stochastic simulations of the full VEGF-Delta-Notch model for a fixed value of the system size, $\overline{\Omega}$, defined by (3.3).

IV The CG model can be simulated using any variant of the stochastic simulation algorithm, such as, for example, the classical Gillespie algorithm [25].

The above method generalizes naturally for systems with an arbitrary number of stable steady states (see Figure 4). In this case, the quasipotential and the corresponding prefactor must be approximated for each pair of stable steady states. The method can also be applied to systems which possess other attractors, e.g., limit cycles [15, 23].

3.2. Multiagent system. In this section we show how the CG method can be applied to multiagent systems with a region of multistability. In this case, the dynamics of *each* agent is coarse-grained to that of a CTMC between its stable steady states for given values of the external variables, v, which establish the coupling between the internal dynamics of individual agents (v depends on the state of agents in the local environment of the focal agent and/or time and defines its internal state, e.g., phenotype). If the dynamics of an individual agent are independent of its neighbors and time (i.e., the values of the external variables are constant), then we use the CG method described in subsection 3.1 (see also Figure 4). A suitable range of values for the external variables, $v \in \mathcal{V}$, where $\mathcal{V} \subset \mathbb{R}^V$, can be determined by simulating the original multiscale model. Here V indicates the dimension of the vector of external variables, v. In order to reduce the computational cost in the multiagent CG system, it is convenient to calculate a priori look-up tables for the steady states, quasipotential and prefactor values for a discretization, $\{v_j\}_{j \in \mathcal{J}} \subset \mathcal{V}$ (here, j indexes entries in the generated discretization; $\mathcal{J}$ is the size of the discretization). Interpolation routines can then be used to establish an input-output relationship between an arbitrary $v \in \mathcal{V}$

and the values of the corresponding steady states and the transition rates between them. Therefore, we split the general CG method for multiagent systems into two steps (see Figure 8):

(i) **Presimulation:** calculate look-up tables for the system steady states, quasipotential, and prefactor values for each entry in a discretization, $\{v_j\}_{j\in\mathcal{J}}$, for a range of values of the external variables, $\mathcal{V}\subset\mathbb{R}^V$.

(ii) **Simulation:** the CG model is simulated (via, e.g., the Gillespie algorithm) as a CTMC on a state space defined by the steady states of all of its entities, with the coupling maintained via the external variables, v, updated at each simulation step according to entities' local environments and/or time.

We now provide more details on the presimulation and simulation steps.

3.2.1. Presimulation: Look-up tables. Precomputed look-up tables of system steady states, quasipotential, and prefactor values are used to interpolate the values of the system steady states and the CG transition rates between them for an arbitrary set of values of the external variables, $v\in\mathcal{V}$, without calculating them explicitly at each step during simulations of the CG model. By an accurate estimation of the range of the external variables, $\mathcal{V}$, we ensure that these look-up tables need to be computed only once, prior to the simulation of the CG system. In a general setting, the dimension of each table is equal to V, the dimension of the vector of external variables.

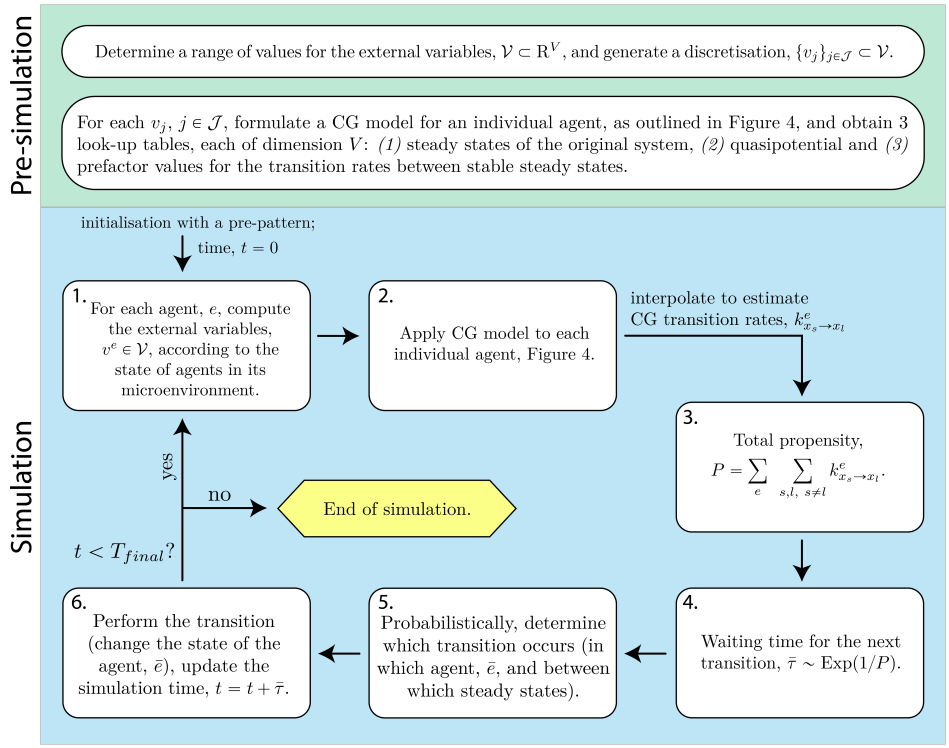


FIG. 8. *A flowchart of the procedure to coarse-grain a multiagent stochastic system with a region of multistability. A pseudocode of the simulation algorithm for the multiagent CG model is presented in section* SM5. *The simulation part of the diagram illustrates an iteration of the Gillespie algorithm for simulation of multiagent CG systems. Here* T_{final} *stands for the final simulation time;* $\mathrm{Exp}(\lambda)$ *is an exponential distribution of intensity,* λ.

The steady states must be computed numerically for each entry v_j in the discretization, $\{v_j\}_{j\in\mathcal{J}}$, using the mean-field limit for an individual entity (as described in subsection 3.1). For values of v_j that fall within the multistability region, the quasipotential is computed via the gMAM in a pairwise manner, for each pair of stable steady states, $\{x_s\}_{s=1}^{\mathcal{S}}$. The last look-up table corresponds to the prefactor, $C_{x_s\to x_l}$, $x_s,\ x_l \in 1\dots\mathcal{S}$, which must be approximated for each v_j within the multistability region. The prefactor values are obtained from (3.2b) as before, using the mean passage times, $\langle T^{\Omega}_{x_s\ \to\ x_l}\rangle$, which are determined by simulating the full stochastic model with the system size, $\Omega = \overline{\Omega}$, defined by (3.3).

3.2.2. Simulation algorithm. Once all the look-up tables have been computed, the multiagent CG system can be simulated as a standard Gillespie algorithm (or one of its variants, e.g., the next subvolume method [20]) in which the total propensity, P, at each time step is computed as a sum of transitions, $k^e_{x_s\to x_l}$, for each entity, e, to switch its (stable) state (see Figure 8). The steady states corresponding to each entity (and the transition rates between them) for the exact value of the external variables $v^e \in \mathcal{V}$ (v^e has to be computed for each entity, e, according to its microenvironment) are interpolated via appropriate numerical routines. We present pseudocode for the simulation procedure in section SM5.

Note that our CG method does not account for the initial, relatively short (compared to the LDT timescale), relaxation time during which the system relaxes onto the timescale on which the CG approximation is valid. Thus, it is necessary to obtain an initial stable steady state configuration, i.e., to *prepattern* the system, using either the full stochastic CTMC or the mean-field model (see Figure 8 and line 5 in Algorithm SM5.1). The final simulation time for the prepatterning should be large enough to ensure that the system relaxes to an equilibrium. Since this procedure is performed only once, it does not affect the computational complexity of the CG simulations. We have chosen to use the mean-field system to prepattern our simulations since it is less time-consuming and the stochasticity (i.e., transitions between phenotypes) is preserved later in the CG simulation loop.

4. Results. For illustrative purposes, we consider the specific example of spatial phenotype patterning via the Delta-Notch lateral inhibition mechanism in response to an external signaling cue (VEGF). First, we provide more details about our implementation of the CG model and present typical simulation results and the robust patterns that emerge at long times. We then discuss the relative merits of the CG method, using a variety of metrics to compare its performance with the original stochastic and mean-field systems. We used the next subvolume method [20] for simulations of the full stochastic CTMC and the Euler–Lagrange method (explicit scheme) for the numerical integration of the mean-field equations.

4.1. CG model of spatial cell phenotype patterning. The multicellular VEGF-Delta-Notch (i.e., the Delta-Notch signaling pathway coupled with external VEGF stimulation) model is bistable (see section SM1 in Supplementary Material). When simulated in a two-dimensional geometry, it produces "salt-and-pepper" patterns in which the phenotypes of neighboring cells alternate between Delta-high and Delta-low states [45]. For this model, cross-talk between individual cells is achieved via external variables, d_{ext} and n_{ext}, which represent the levels of Delta and Notch, respectively, summed over cells in a circular neighborhood with a fixed interaction radius, R_s (see Figure 9 and section SM1). Hence, for this system, $v = (d_{ext}, n_{ext})$ defines a cell's internal state (phenotype) and the dimension of the precomputed look-up tables is 2

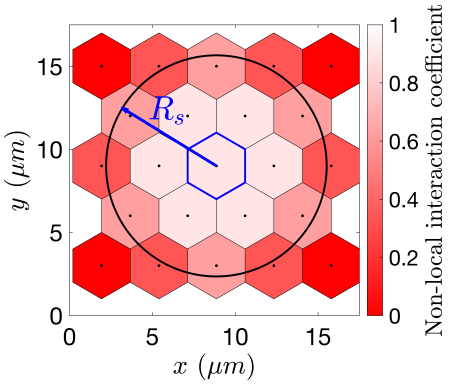


FIG. 9. *A schematic diagram showing the nonlocal interactions in the multicellular VEGF-Delta-Notch model. Cell-to-cell interactions may be nonlocal (i.e., beyond immediate neighbors on a given lattice) provided they lie within an interaction radius, R_s. The diagram illustrates the weights of interactions between the focal cell (highlighted in blue) and cells in its neighborhood for a regular hexagonal lattice (the weights are defined as a normalized area of the overlap between a neighboring voxel and the circular neighborhood of the focal cell; see equation* (SM1.3) *in section* SM1*).*

(see subsection 3.2.1). We determined a suitable range, $\mathcal{V} = [0, d_{ext}^{max}] \times [0, n_{ext}^{max}] \subset \mathbb{R}^2$, for these variables by running 100 realizations of the multiscale model of angiogenesis (the number of realizations depends on the model of interest).

We then generated a regular discretization of $\mathcal{V}$, $\{v_j\}_{j \in \mathcal{J}}$, with a grid 100×100. For each v_j in this grid, we computed the steady states for the mean-field limit defined by equation (SM1.2) using nonlinear solvers from the C++ GNU Scientific Library. We note that, once the steady states of the full system have been computed, the subcellular variables ι, r_2, and r_2^*, corresponding to the Notch intracellular domain, VEGF receptor 2 (VEGFR2), and VEGF-VEGFR2 complexes, respectively (see definitions in section SM1 in Supplementary Material), are redundant; it is not necessary to track these variables because the input-output relationship between $v = (d_{ext}, n_{ext})$ and the steady states completely defines the configuration of the system.

For values of v_j that fall within the bistability region, we computed the quasipotential values of the transitions between phenotypes (see Figure 10) using the gMAM (see section SM2 in Supplementary Material). We also used the full stochastic system to check those values of the quasipotential for which a phenotype switch is more likey to occur. As expected, most phenotype transitions occur close to the boundary of the bistability region, where values of the quasipotential are lower. For example, Figures 10(a) and 10(b) show a sample path of the full stochastic system for an individual cell during a simulation of the multiagent model [45]. The cell undergoes a noise-induced switch from a Delta-high to a Delta-low phenotype. Figures 10(c) and 10(d) show the same sample path projected onto the quasipotential surfaces. These plots show that phenotypic switches are more likely to occur when the values of external Delta and Notch, (d_{ext}, n_{ext}), are such that the quasipotential, $V(x_1, x_2) = V(\text{Delta-high}, \text{Delta-low})$, is small.

We constructed a look-up table of prefactor values, $C_{x_s \to x_l}$, x_s, $x_l \in \{\text{Delta-high}, \text{Delta-low}\}$, by approximating the mean passage times, $\langle T^{\overline{\Omega}}_{x_s \to x_l} \rangle$ (sample size of 1000 realizations), for an individual cell to switch its phenotype from simulations of the full stochastic CTMC (Table SM2) with the system size, $\overline{\Omega}$, given by (3.3).

We then implemented the CG model in C++ using Algorithm SM5.1. In order to establish an input-output relationship between an arbitrary $v = (d_{ext}, n_{ext})$ and

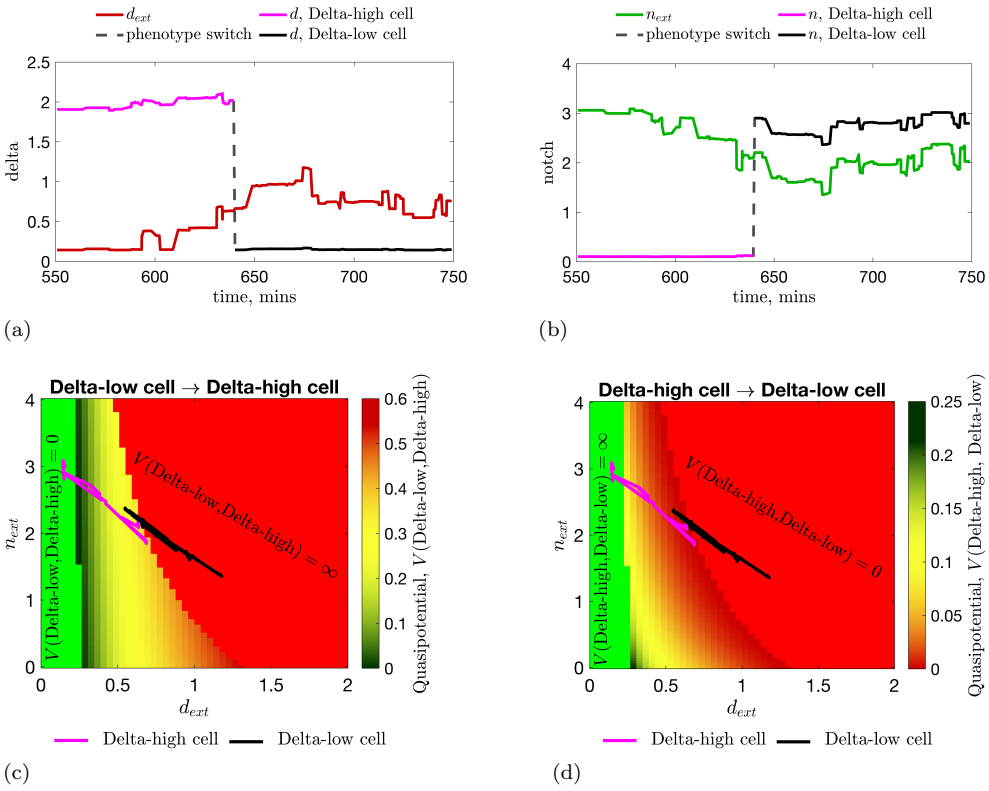


FIG. 10. *An illustration of the quasipotential surfaces. Upper panels: a noise-induced transition from Delta-high (in magenta) to Delta-low (in black) phenotype of a single cell during a simulation of the angiogenesis model* [45] *plotted as a function of the focal cell's* (a) *Delta and* (b) *Notch levels. The external Delta, d_{ext}, (Notch, n_{ext}) for the focal cell is computed as a weighted sum of the Delta (Notch) levels of its neighbors as defined by equation* (SM1.4). *Lower panels: two-dimensional projections of the quasipotential surfaces* (c) V(*Delta-low, Delta-high*) *and* (d) V(*Delta-high, Delta-low*) *as functions of d_{ext} and n_{ext}. The monostability region in which the unique stable steady state corresponds to a Delta-high (Delta-low) cell is colored green (red). The color bar indicates the value of the corresponding quasipotential. The trajectory (as in panels* (a) *and* (b)*) plotted on the quasipotential surfaces (in* (c) *and* (d)*) illustrates that phenotype switches are more likely to occur for lower values of the quasipotential. Parameter values are fixed as indicated in Table* SM3.

the corresponding cell phenotypes and transition rates, we used bilinear interpolation routines from the C++ GNU Scientific Library (gsl_interp2d routines). The model was then simulated using the standard Gillespie algorithm. We used no-flux boundary conditions to compute for each cell the extracellular levels of Delta and Notch in all our simulations.

4.2. Spatial patterning in the CG model. In order to illustrate the CG model, we first ran numerical simulations on a small cell monolayer (10×12 voxels). The results presented in Figures 11(a) to 11(d) show how the distribution of Delta-high and Delta-low cells changes over time during a typical CG realization (see also Movie S1 [local/web 6.5MB]). Starting from an initial prepattern (Figure 11(a)), noise-induced phenotype transitions enable the system to explore different pattern configurations for the given geometry, while the proportion of Delta-high cells remains on average constant (see Figure 11(e)).

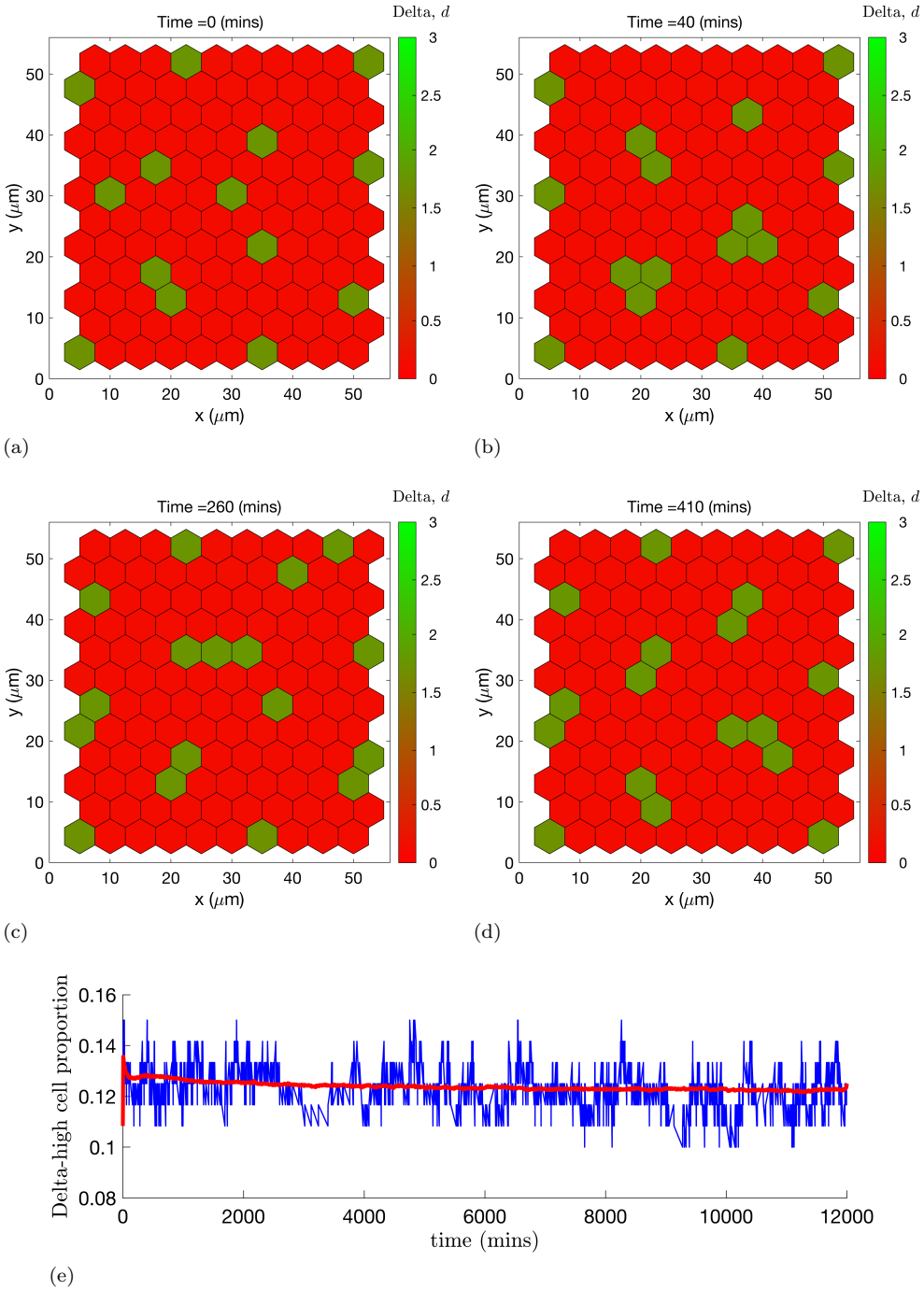


FIG. 11. *Different pattern configurations explored by the CG model.* (a)–(d) *Series of plots showing how the distribution of cell phenotypes changes over time during a single simulation of the CG model. The color bar indicates the level of Delta.* (a) $t = 0$; (b) $t = 40$; (c) $t = 260$; (d) $t = 410$ *minutes.* (e) *Time evolution of the Delta-high cell proportion (defined as a ratio of cells with the Delta-high phenotype to the total cell number) for a single simulation of the CG model (blue line) and averaged over* 1000 *realizations (red line). For these simulations, the interaction radius and system size were fixed at* $R_s = 15\mu m$ *and* $\Omega = 100$*, respectively; the values of the remaining parameters were fixed as indicated in Table* SM3.

The mean proportion of Delta-high cells (and, thus, the spatial pattern) during simulations of the CG system depends on the interaction radius, R_s. For values of R_s corresponding to nearest-neighbors interaction ($R_s \leq 1.5h$, where h is the voxel width), we observe classical patterns of alternating Delta-high and Delta-low cells (i.e., the so-called salt-and-pepper pattern [12]; see Figure SM6a). As R_s increases, the number of Delta-low cells that may be inhibited by a focal Delta-high cell increases, causing the proportion of Delta-high cells in the spatial patterns to decrease [45]. Thus, for larger values of R_s ($R_s > 1.5h$), Delta-high cells are separated by larger distances (see Figures SM6b to SM6d). These results for CG simulations are consistent with those obtained for the full multicellular stochastic model of the VEGF-Delta-Notch signaling pathway [45]. The ability of the CG system to explore different spatial patterns increases as the size of the interaction radius, R_s, grows, and the corresponding emerging patterns are more diverse (see Figures SM6b to SM6d and Figures 11(a) to 11(d)).

It is noteworthy that spatial patterns explored in simulations of the CG model differ in their robustness to noise. In particular, the mean passage time for a phenotype switch, and, thus, a change in the pattern, to occur, which is equal to the inverse of the total propensity, P, depends on the values of the quasipotential, $V(x_s, x_l)$, for all entities in the system. Here, the total propensity, P, for a phenotype switch event is defined as a sum of, transition rates, $k^e_{x_s \to x_l}$, for each cell with index, e, to change its state from x_s to x_l; see Figure 8. When, via random exploration, the system finds a configuration for which the values of $V(x_s, x_l)$ are larger, the waiting time for a phenotype switch increases and the configuration is more resilient to further changes.

This feature of the CG method facilitates exploration of new robust spatial patterns which cannot practically be achieved using other numerical frameworks: (i) simulations of the full stochastic model are too computationally intensive, which makes the exploration of these patterns infeasible because of the longer timescales needed; (ii) the deterministic framework does not allow for transitions between stable steady states, which makes this exploration impossible; (iii) the complexity of analytic methods needed to verify the stability of a pattern of a system with nonlocal interactions does not permit exploration of complex pattern configurations [38].

We now present simulation results which illustrate the ability of the CG method to uncover new spatial patterns for the VEGF-Delta-Notch system at long times. We fixed the interaction radius at $R_s = 3.0h = 15\mu m$ ($h = 5\mu m$ is the voxel width), so that interactions occur between cells that are first and second order neighbors in the lattice; the noise amplitude was fixed at $\epsilon = \Omega^{-1} = 0.001$. We ran a CG simulation on a medium-size monolayer of cells (see Figure 12(a) and Movie S2 [local/web 5.9MB]). Starting from the initial prepattern, the CG model explores various patterns until it eventually settles on a more robust configuration (shown in Figure 12(a)). In order to confirm our prediction regarding pattern robustness, we plotted the temporal evolution of the total propensity of the lattice, P, in Figure 12(b). As its value decreases, $P \to 0$, the mean waiting time for a change in the spatial pattern becomes infinite, which accounts for the robustness of the emerging pattern. We also considered the dynamics of an individual cell (its position in the monolayer is highlighted by a cyan line in Figure 12(a)). Figure 12(c) shows how the phenotype of this cell changes over time: at early times, the cell switches between Delta-high and Delta-low phenotypes (low (high) values of subcellular Delta, d, correspond to Delta-low (Delta-high) phenotype). As the spatial pattern settles to a robust configuration, the cell's environment, i.e., the levels of Delta of its neighbors, d_{ext}, stops changing and the cell acquires a Delta-high phenotype that remains unchanged for the rest of the

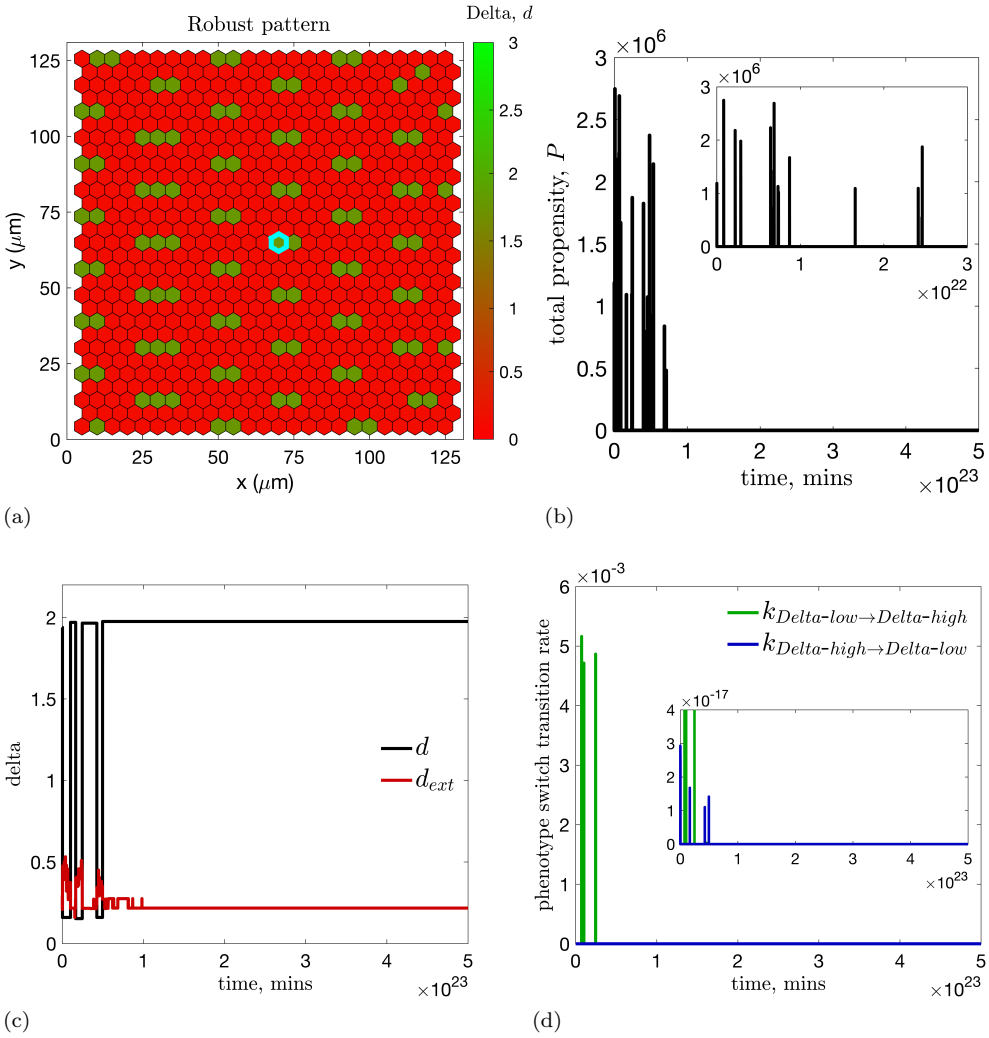


FIG. 12. *Emergence of robust pattern configurations in simulations of the CG model. At long times, via exploration of different pattern configurations, the dynamics of the CG system evolve to a robust pattern in which any further phenotype switches are unlikely.* (a) *A typical emergent pattern for a single realization of the CG model (the color bar indicates the level of Delta, d, for each cell).* (b) *The time evolution of the total propensity, P, for a phenotype switch to occur. Cells in the border rim (three-cell width) are excluded from P since, due to the model geometry, they do not possess a "robust" configuration of neighbors. As P decreases to 0, the waiting time for a phenotype switch to occur approaches infinity, and the pattern becomes more robust to change.* (c)–(d) *The dynamics of an individual cell (outlined in cyan in* (a)*) during this simulation.* (c) *Temporal evolution of the internal level of Delta, d (defining cell phenotype: high (low) values of d correspond to Delta-high (Delta-low) phenotype), and that in its microenvironment, d_{ext}.* (d) *Temporal evolution of transition rates for a phenotype switch for this cell. We note that the large difference in the order of values for transition rates for the total propensity, P, of the lattice ($O(10^6)$), plot* (b)*, and for an individual cell ($O(10^{-17}) - O(10^{-3})$), plot* (d)*, comes from the contribution to P of transition rates for cells which are, for the given values of the external variables, on the border of the bistability region (see Figure* 10*). For these simulations, the interaction radius and system size were fixed at $R_s = 15\mu m$ and $\Omega = 1000$, respectively; the values of all remaining parameters were fixed as indicated in Table* SM3.

simulation. The transition rates for phenotype switches for this cell (Figure 12(d)) exhibit similar dynamics to the total propensity, P, of the whole lattice (Figure 12(b)).

Our CG simulation results show that this robust pattern configuration is not unique. However, we note that the spatial patterns tend to have a regular structure; for example, Delta-high cells may be organized in similar clusters comprising two or three cells as in the pattern shown in Figure 12(a). These configurations have lower values of the total propensity, P. Cells on the border of the lattice undergo phenotype switches (see Movie S2 [local/web 5.9MB]), since they cannot attain this "more robust" combination of neighbors for the given geometry (since we use no-flux boundary conditions in our simulations).

4.3. Comparison of the full stochastic, coarse-grained, and mean-field frameworks. We compared the dynamics of the multicellular VEGF-Delta-Notch model using three frameworks:

(i) full stochastic CTMC in which each cell's dynamics is given by the set of kinetic reaction rates listed in Table SM2 of Supplementary Material;
(ii) CG description formulated as in subsection 4.1;
(iii) deterministic mean-field description in which each cell is equipped with a set of deterministic ODEs (see equation (SM1.4) in Supplementary Material).

Simulated (using any of these frameworks) on a two-dimensional domain, the model produces a characteristic pattern of ECs with two cell phenotypes (see, for example, Figure SM6 and Figure 11). Since the CG approximation describes the long-term behavior of the system, when its evolution is dominated by the timescale associated with phenotypic switches, it does not account for the initial relaxation onto a quasi–steady state pattern. Thus, the three frameworks cannot be compared with respect to their behavior at early evolution times. Instead, we quantified the final pattern and the computational cost of simulations. The final simulation time, $t = T_{final}$, was chosen sufficiently large to ensure that a steady state pattern had been established for the mean-field simulations (since stochastic systems do not have a steady state pattern in a classical sense). In order to systematically compare the three frameworks, we used the same final simulation time, $t = T_{final}$, for the other two systems.

We used the following set of metrics to compare the dynamics of the three mathematical descriptions (Supplementary Material (Supplement.pdf [local/web 16.2MB])):

- Delta-high cell proportion, which is defined as the ratio of the number of cells with Delta-high phenotype to the total number of cells in the system;
- distribution of Delta-high cell clusters, which provides a breakdown of sizes of Delta-high cell clusters (adjacent cells with Delta-high phenotype, e.g., a single Delta-high cell, two adjacent Delta-high cells) in a steady pattern configuration;
- computational cost, which is defined as the average CPU time (in seconds) to perform a single realization of model simulation.

Since the precalculated look-up tables for the CG simulations (subsection 3.2.1) were computed for a fixed set of model parameters (see Table SM3), we held them fixed for all simulations. However, the cell-to-cell interaction radius, R_s, which is used in the multicellular simulations to determine for each cell, e, the vector of extracellular variables, $v^e = (d^e_{ext}, n^e_{ext})$, may vary. In our simulations, we used $R_s \in \{5,\ 7.5,\ 10,\ 12.5,\ 15\}\ \mu m$ which correspond to experimental observations of the distance over which cell-to-cell interaction can occur in endothelial cells [18] (which corresponds to up to three cells in the interaction circle). Nonetheless, from a theoretical point of view, this quantity can take any value greater than the half-width

of a voxel, $R_s > 0.5h$, where h is the voxel width (we fix $h = 5\mu m$ in our simulations). In addition, for the full stochastic CTMC and CG descriptions, we vary the noise amplitude, $\epsilon = 1/\Omega$, by changing the system size parameter, Ω. We used $\Omega \in \{50,\ 100,\ 200,\ 500,\ 1000\}$. The larger the value of Ω, the closer will be the dynamics of a stochastic system to its mean-field description. For each numerical setup (R_s and Ω), we ran 100 realizations.

We considered two simulation geometries: a 2D cell monolayer and a branching network.

Setup 1: A cell monolayer. We first ran numerical simulations on a cell monolayer (see Figure SM7). This spatial geometry was motivated by the biological process of cell fate specification induced by lateral inhibition via Delta-Notch signaling in flat domains. Examples of such cell fate specification include bristle patterning in *Drosophila* notum [11, 30, 13] and differentiation of neural precursors in neurogenesis [22] (see [7, 36] and references therein for other examples). The fixed stationary distribution of the VEGF serves as an external stimulus which enhances lateral inhibition via Delta-Notch signaling. We chose VEGF as an illustrative example, although, depending on the specific system, other extracellular signals will provide cell stimulus.

We began by considering the dynamics of the Delta-high cell proportion for this spatial geometry (see Figure 13(a)). Consistent with the previous results [45], for all simulation frameworks (i.e., the full stochastic (CTMC), CG, and mean-field descriptions), the Delta-high cell proportion decreases as the cell interaction radius, R_s, increases. Figure 13(a) confirms that, as expected, differences in this metric between the three systems decrease as the level of noise is reduced (i.e., as Ω increases). In particular, for high noise levels (i.e., lower values of Ω), the patterns generated by the stochastic systems (full CTMC and CG frameworks) are more diverse, and the Delta-high cell proportions differ from those for the associated mean-field description. We note that each cell is not an isolated system; its dynamics are affected by the noisy behavior of its neighbors and the model geometry. This explains why we observe variations in the Delta-high cell proportion for lower values of Ω for the full CTMC and CG frameworks. We also note that the dynamics of the Delta-high cell proportion for the mean-field system (red lines) are identical in all subplots in Figure 13(a) since noise is absent in deterministic systems (i.e., the system size parameter, Ω, is irrelevant).

We also quantified the size distribution of the Delta-high cell clusters associated with the final patterns established on the cell monolayers. Since the dynamics of the three systems converge for larger values of the system size, Ω (as shown in Figure 13(a)), Figure 13(b) shows results for this metric computed for simulations with $\Omega = 1000$. The distributions are in good quantitative agreement for the three systems. The discrepancy for simulations with larger cell interaction radius (e.g., $R_s = 15\mu m$) arises because (for this value of Ω) the CG system is more likely to explore long timescale patterns which have a more "regular" structure and are more robust to noise (cells with Delta-high phenotype organized in similar clusters; see subsection 4.2).

Setup 2: A branching network. We next considered a more complex spatial geometry of a small branching network (see Figure SM8) extracted from a simulation of a hybrid model of angiogenesis [45]. Figure SM9 shows a series of patterns explored by the CG system at different time points during a typical simulation for this configuration (for the full simulation, see Movie S3 [local/web 3.4MB]).

For this spatial configuration, we compared the three simulation frameworks using the same metrics as for the cell monolayer. The results for the Delta-high cell proportion are presented in Figure SM10a. We find that the number of possible patterns

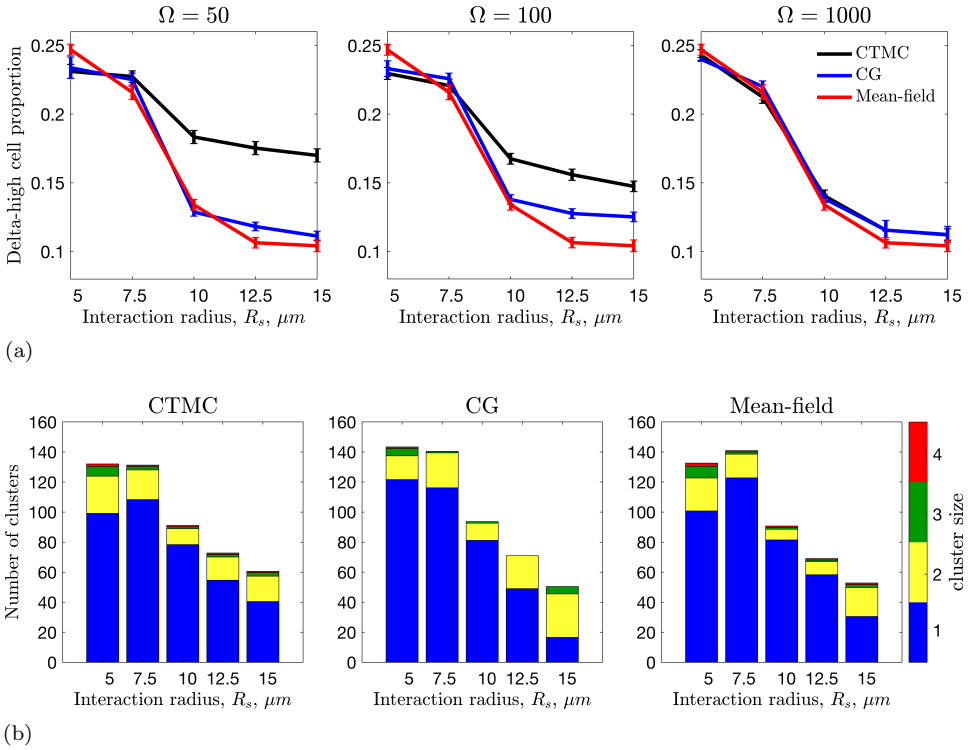


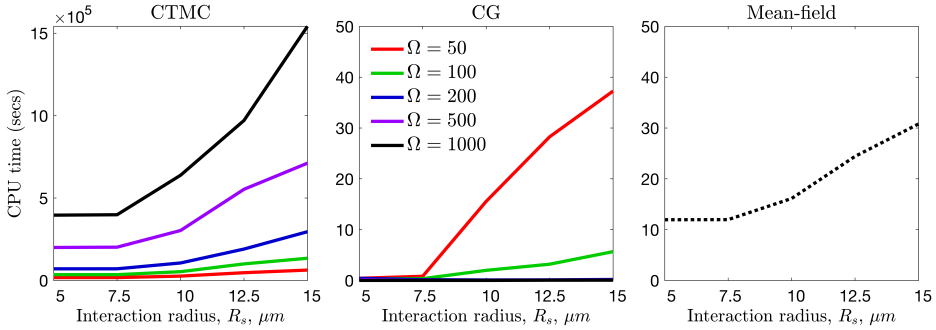
FIG. 13. *Comparison of the dynamics of the multicellular VEGF-Delta-Notch model simulated on a cell monolayer using the full stochastic (CTMC), CG, and mean-field descriptions.* (a) *The Delta-high cell proportion as a function of the cell-to-cell interaction radius, R_s, for varying noise amplitude, $\epsilon = 1/\Omega$ (the value of Ω is indicated in the title of each plot), for the full stochastic CTMC (black), CG (blue), and mean-field (red) descriptions. To explore different possible patterns in the deterministic mean-field system, we created a small initial perturbation to the initial configuration (Figure* SM7*).* (b) *A series of barplots showing how the long-time distribution of Delta-high cell clusters changes as the interacton radius, R_s, varies for the full stochastic CTMC (left panel), CG (middle panel), and mean-field (right panel) systems. The number of single Delta-high cells in the final pattern (i.e., at a fixed final simulation time) is shown in blue; the number of clusters with* 2, 3, *and* 4 *adjacent Delta-high cells is shown in yellow, green, and red, respectively. For these simulations, we fixed* $\Omega = 1000$ *($\epsilon = 0.001$). The results are averaged over* 100 *realizations. The remaining parameter values were fixed as indicated in Table* SM3.

generated by lateral inhibition is lower for the branching network geometry than for the cell monolayer (see Figure SM9). Consequently, the Delta-high cell proportions converge for smaller values of Ω (compare Figure SM10a and Figure 13(a)). We also note that since, in the network configuration, cells have fewer neighbors, the values of this metric are higher than those computed for a cell monolayer.

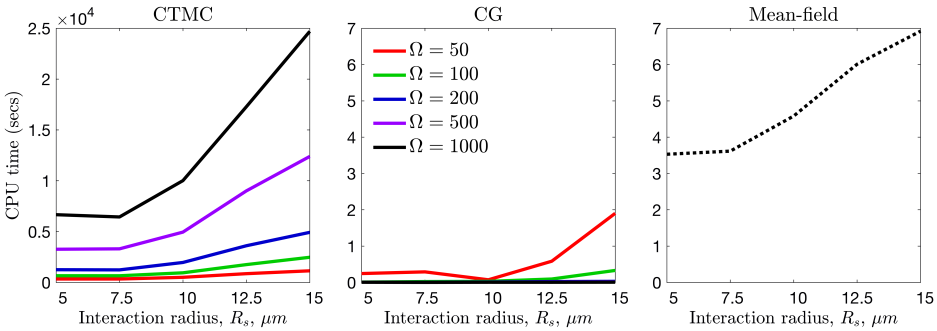
Figure SM10b shows the size distribution of Delta-high cell clusters for simulations on the branching network. We note that, for this configuration, isolated Delta-high cells (i.e., cells not adjacent to another Delta-high cell) are predominant in the final spatial patterns and the patterns generated by the three frameworks are comparable.

Regarding the computational cost (see technical specifications of computers used in File S1 [local/web 10KB]), the CG method showed a great reduction in the average CPU time compared to the original stochastic system when performing a single realization (see Figure 14). Whereas the numerical cost of simulations of the full stochastic system (Figure 14, left panels) increases exponentially as the system size,

Ω, grows, simulations of the CG system decrease in average computational time as Ω increases (Figure 14, middle panels). This is because, as the noise level decreases (i.e., Ω increases), fewer transitions occur in a CG simulation for a fixed final simulation time. Interestingly, the CG simulations are also faster (see Figure 14, right panels) than the numerical integration scheme used for the mean-field system (we used the explicit scheme for the Euler–Lagrange method, although other schemes for numerical integration may show better performance). This scheme required evaluation of the nonlinear right-hand side of the mean-field equations (see equation (SM1.4) in Supplementary Material) at each time step for every voxel in the lattice, whereas for the CG simulations only one voxel undergoes a change (i.e., a phenotype switch) at each iteration and an update is required only for a local neighborhood of this voxel. Therefore, fewer updates are required in the CG system (which are further decreased as Ω grows). In addition, the transition rates for phenotype switches needed for CG updates are interpolated directly using the precalculated look-up tables (see subsection 3.2.1) which reduces the amount of computations required (as compared to the evaluation of the right-hand side of the deterministic system). This explains why the



(a) cell monolayer



(b) branching network

FIG. 14. *Comparison of the mean CPU times to simulate the multicellular VEGF-Delta-Notch model. The plots show how the average (*100 *realizations) CPU times (in seconds) to perform a single realization using the full stochastic CTMC (left panels), CG (middle panels), and mean-field (right panels) descriptions change as the cell-to-cell interaction radius, R_s, and the system size, Ω, vary. The color code indicates the system size, Ω (shown as insets in the middle panels); the mean-field description (dotted black line) corresponds to the limit $\Omega \to \infty$. The simulation setup was* (a) *a medium size cell monolayer and* (b) *a small branching network. For both spatial geometries, the average CPU time for simulating the full stochastic CTMC is several orders of magnitude larger that those for CG and mean-field descriptions. For these simulations, the parameter values were fixed as indicated in Table* SM3.

mean CPU time for CG simulations is smaller (except for high noise levels, $\Omega = 50$) than for the numerical scheme we used for the associated deterministic system (see Figure 14, middle and right panels).

To summarize, the CG method, while preserving stochasticity of transitions between cell phenotypes and producing spatial patterns comparable to those generated using the original stochastic and mean-field descriptions, significantly reduces computational time of simulations.

5. Discussion and conclusions. Hybrid (multiscale) models of complex biological phenomena are often computationally inefficient, which hinders their potential utility. To address this issue, we have developed a coarse-graining, or CG, method that reduces the numerical cost of simulations of multiagent stochastic systems with multiple stable steady states. The CG technique is based on large deviation theory that allows the dynamics of a stochastic system to be reduced to a jump process (i.e., a continuous time Markov chain) on a discrete state space which comprises the stable steady states of all agents in the system. The CG system operates on a timescale on which transitions between these steady states take place. This allows the method to be applied to models whose dynamics act on timescales longer than the typical timescale for relaxation to an equilibrium (e.g., molecular or subcellular processes act on longer timescales when compared to higher spatial scales such as cell migration, dynamics of extracellular cues, etc.). Our results show good qualitative and quantitative agreement between CG simulations and other simulation methods (Figure SM10 and Figure 13). Furthermore, the CG algorithm is numerically more efficient in terms of CPU time even when compared with the corresponding mean-field simulations (see Figure 14). Likewise, the CG framework allows exploration of new emergent properties of the system, such as long timescale patterns in multicellular systems (Figure 12).

The implementation of the CG method requires precalculation of several look-up tables (for stable steady state solutions of the system that is being coarse-grained, quasipotential values for transitions between them, and the corresponding prefactor of these transitions) which are used later in simulations. To do this, the values of model parameters must be fixed (except for the external variables). However, in order to perform sensitivity analysis with respect to any specific parameter, this parameter may be added to the set of external variables (thus, adding a new dimension to the look-up tables). Since the procedure of precalculating the look-up tables is done once, prior to model simulation, it does not increase the numerical cost of the algorithm. Likewise, the computational cost of computing the quasipotential via the geometric minimum action method is independent of the system size, Ω, and an estimate for the required prefactor can be obtained from simulations of the full stochastic model for a single value of the system size parameter, Ω, for which we provided an accurate estimate (see (3.3) and Figure 7). Then the CG model can be efficiently simulated using the standard Gillespie algorithm for any value of Ω (or, equivalently, noise level, $\epsilon = 1/\Omega$).

After introducing the CG method (section 3), we applied it to a multiagent model of phenotypic specification of cells via the VEGF-Delta-Notch signaling pathway. For this system, we demonstrated how the spatial patterning of cells with different phenotypes changes as CG transitions between these steady states (phenotypes) occur (Figure 11). We then confirmed that the patterns generated by the CG system are quantitatively similar to steady state configurations of the original stochastic system and the associated mean-field limit for this model (see Figure SM10 and Figure 13). We conclude that the CG method preserves the continuous cell phenotypes

and stochasticity of the original system, while reducing the computational cost of simulations by several orders of magnitude (as compared to the numerical cost of simulations of the full stochastic system; see Figure 14).

In this paper, we used the VEGF-Delta-Notch model to illustrate the benefits of the CG method. We note, however, that the CG method can be applied to a wider class of multiagent models in which the behavior of the agents is regulated by stochastic models with multiple stable attractors (e.g., steady states, limit cycles) and whose dynamics are controlled by external cues (e.g., morphogens, growth factors, levels of specific ligands/receptors in neighboring cells). Examples of systems with subcellular dynamics which satisfy the requirements for application of the CG method include fate specification of cells in intestinal crypts [32, 8], epithelial to mesenchymal phenotypic transition (and its reverse) in cancer invasion [31] and development [43], cell differentiation in neurogenesis [22], and a general class of models describing cell decision switches [27]. These models are multistable and the timescale of simulations is longer than the timescale of the relevant subcellular signaling pathway. Nonetheless, the spectrum of models which are suitable for coarse-graining via the CG algorithm is not restricted to intracellular signaling pathways in animal cells; other examples include vegetation patterning in arid ecosystems [33] or plant morphogenesis mediated via the auxin hormone [1, 21]. The exact implementation of the CG system for the aforementioned models is beyond the scope of this paper.

To conclude, the CG method developed in this paper paves the way for a systematic reduction of the dynamics of a wide class of multistable stochastic models. It allows for investigation of their behavior on longer timescales than is possible with other frameworks (e.g., full stochastic simulations or deterministic equations). To our knowledge, this is the first example in which large deviation theory has been used to coarse-grain the dynamics of a multiagent system. In future work we intend to further investigate the performance of the CG method by incorporating the CG system for the VEGF-Delta-Notch signaling into a multiscale model of angiogenesis [45].

Data management. All of the computational data output is included in the manuscript and/or in the supplementary material (Supplement.pdf [local/web 16.2MB]). The code of the numerical procedures used in this work is available upon request.

Supplementary materials.

Supplementary Material (Supplement.pdf [local/web 16.2MB]). The file contains a more detailed description of the VEGF-Delta-Notch model, implementation of the CG method, and additional figures and tables.

File S1 [local/web 10KB]. Technical specifications of the computers used to perform simulations in this work.

Movie S1 [local/web 6.5MB]. **A simulation movie showing different pattern configurations explored by the CG system in a small two-dimensional cell monolayer.** The movie is the complete animation of the simulation snapshots shown in Figure 11 of the main text. This movie demonstrates how the spatial distribution of cells with two phenotypes changes over time in the CG system due to phenotype switches in individual cells. The color bar indicates the levels of Delta. For this simulation, the interaction radius and system size were fixed at $R_s = 15\mu m$ and $\Omega = 100$, respectively; the values of the remaining parameters were fixed at the values given in Table SM3.

Movie S2 [local/web 5.9MB]. **A simulation movie showing the emergence of robust pattern configurations in simulations of the CG system.** The movie is the complete animation of the simulation snapshot shown in Figure 12 of

the main text. This movie illustrates the emergence of a robust spatial pattern in the dynamics of the CG system at long times. Due to the exploration of different pattern configurations, the CG system settles on a configuration whose total propensity, P, is small. Thus, the mean waiting time for a phenotype switch for this pattern (given by $1/P$) tends to infinity and it becomes robust to any further phenotype switches. The lattice site highlighted in cyan indicates the position of a cell whose dynamics are shown in Figures 12(c) and 12(d). The color bar indicates the levels of Delta. For this simulation, the interaction radius and system size were fixed at $R_s = 15\mu m$ and $\Omega = 1000$, respectively; the values of all remaining parameters were fixed at the values given in Table SM3.

Movie S3 [local/web 3.4MB]. A simulation movie showing different pattern configurations explored by the CG system in a branching network. The movie is the complete animation of the simulation snapshots shown in Figure SM9 of Supplementary Material. It illustrates the evolution of the CG system simulated on a branching network. The color bar indicates the levels of Delta. For this simulation, the interaction radius and system size were fixed at $R_s = 15\mu m$ and $\Omega = 100$, respectively; the values of the remaining parameters were fixed at the values given in Table SM3.

REFERENCES

[1] V. Baldazzi, N. Bertin, H. de Jong, and M. Génard, *Towards multiscale plant models: integrating cellular networks*, Trends in Plant Science, 17 (2012), pp. 728–736.

[2] R. Bardini, G. Politano, A. Benso, and S. Di Carlo, *Multi-level and hybrid modeling approaches for systems biology*, Computational Structural Biotechnology J., 15 (2017), pp. 396–402.

[3] K. Bentley, C. A. Franco, A. Philippides, R. Blanco, M. Dierkes, V. Gebala, F. Stanchi, M. Jones, I. M. Aspalter, G. Cagna, et al., *The role of differential Ve-cadherin dynamics in ell rearrangement during angiogenesis*, Nature Cell Biology, 16 (2014), pp. 309–321.

[4] S. Bernard, *How to build a multiscale model in biology*, Acta Biotheoretica, 61 (2013), pp. 291–303.

[5] R. Blanco and H. Gerhardt, *Vegf and notch in tip and stalk cell selection*, Cold Spring Harbor Perspectives Medicine, 3 (2013), a006569.

[6] M. Boareto, M. K. Jolly, M. Lu, J. N. Onuchic, C. Clementi, and E. Ben-Jacob, *Jagged–delta asymmetry in notch signaling can give rise to a sender/receiver hybrid phenotype*, Proc. Natl. Acad. Sci. USA, 112 (2015), pp. E402–E409.

[7] F. Bocci, J. N. Onuchic, and M. K. Jolly, *Understanding the principles of pattern formation driven by notch signaling by integrating experiments and theoretical models*, Frontiers Physiology, 11 (2020).

[8] P. Buske, J. Galle, N. Barker, G. Aust, H. Clevers, and M. Loeffler, *A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt*, PLoS Comput. Biol., 7 (2011), e1001045.

[9] H. Byrne and M. Chaplain, *Mathematical models for tumour angiogenesis: numerical simulations and nonlinear wave solutions*, Bull, Math, Biol., 57 (1995), pp. 461–486.

[10] M. A. Chaplain, *Multiscale modeling of cancer: Micro-, meso-and macro-scales of growth and spread*, in Approaching Complex Diseases, Springer, New York, 2020, pp. 149–168.

[11] M. Cohen, M. Georgiou, N. L. Stevenson, M. Miodownik, and B. Baum, *Dynamic filopodia transmit intermittent delta-notch signaling to drive pattern refinement during lateral inhibition*, Developmental Cell, 19 (2010), pp. 78–89.

[12] J. R. Collier, N. A. Monk, P. K. Maini, and J. H. Lewis, *Pattern formation by lateral inhibition with feedback: A mathematical model of delta-notch intercellular signaling*, J. Theoret. Biol., 183 (1996), pp. 429–446.

[13] F. Corson, L. Couturier, H. Rouault, K. Mazouni, and F. Schweisguth, *Self-organized notch dynamics generate stereotyped sensory organ patterns in drosophila*, Science, 356 (2017).

[14] R. de la Cruz, P. Guerrero, J. Calvo, and T. Alarcón, *Coarse-graining and hybrid methods for efficient simulation of stochastic multi-scale models of tumour growth*, J. Comput. Phys., 350 (2017), pp. 974–991.

[15] R. De La Cruz, R. Perez-Carrasco, P. Guerrero, T. Alarcon, and K. M. Page, *Minimum action path theory reveals the details of stochastic transitions out of oscillatory states*, Physical Rev. Lett., 120 (2018), 128102.
[16] T. S. Deisboeck, Z. Wang, P. Macklin, and V. Cristini, *Multiscale cancer modeling*, Ann. Rev. Biomed. Engrg., 13 (2011), pp. 127–155.
[17] A. Deutsch, P. Friedl, L. Preziosi, and G. Theraulaz, *Multi-scale analysis and modeling of collective migration in biological systems*, Philos. Trans. B, 375(2020).
[18] Y. Du, S. C. Herath, Q.-g. Wang, D.-a. Wang, H. H. Asada, and P. C. Chen, *Three-dimensional characterization of mechanical interactions between endothelial cells and extracellular matrix during angiogenic sprouting*, Scientific Reports, 6 (2016), 21362.
[19] M. I. Dykman, E. Mori, J. Ross, and P. Hunt, *Large fluctuations and optimal paths in chemical kinetics*, J. Chemi. Phys., 100 (1994), pp. 5735–5750.
[20] J. Elf and M. Ehrenberg, *Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases*, Systems Biol., 1 (2004), pp. 230–236.
[21] E. Farcot, C. Lavedrine, and T. Vernoux, *A modular analysis of the auxin signaling network*, PLoS One, 10 (2015), e0122231.
[22] P. Formosa-Jordan, M. Ibañes, S. Ares, and J.-M. Frade, *Lateral inhibition and neurogenesis: Novel aspects in motion*, Internat. J. Developmental Biology, 57 (2013), pp. 341–350.
[23] M. I. Freidlin and A. D. Wentzell, *Random perturbations*, in Random Perturbations of Dynamical Systems, Springer, New York, 1998, pp. 15–43.
[24] H. Gerhardt, M. Golding, M. Fruttiger, C. Ruhrberg, A. Lundkvist, A. Abramsson, M. Jeltsch, C. Mitchell, K. Alitalo, D. Shima, et al., *VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia*, J. Cell Biol., 161 (2003), pp. 1163–1177.
[25] D. T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Comput. Phys., 22 (1976), pp. 403–434.
[26] T. Grafke, T. Schäfer, and E. Vanden-Eijnden, *Long term effects of small random perturbations on dynamical systems: Theoretical and computational tools*, in Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science, Springer, New York, 2017, pp. 17–55.
[27] R. Guantes and J. F. Poyatos, *Multistable decision switches for flexible control of epigenetic differentiation*, PLoS Comput. Biol., 4 (2008), e1000235.
[28] T. Heck, M.-M. Vaeyens, and H. Van Oosterwyck, *Computational models of sprouting angiogenesis and cell migration: Towards multiscale mechanochemical models of angiogenesis*, Math. Model. Nat. Phenom., 10 (2015), pp. 108–141.
[29] M. Heymann and E. Vanden-Eijnden, *The geometric minimum action method: A least action principle on the space of curves*, Commun. Pure Appl. Math., 61 (2008), pp. 1052–1117.
[30] G. L. Hunter, Z. Hadjivasiliou, H. Bonin, L. He, N. Perrimon, G. Charras, and B. Baum, *Coordinated control of notch/delta signaling and cell cycle progression drives lateral inhibition-mediated tissue patterning*, Development, 143 (2016), pp. 2305–2310.
[31] M. K. Jolly, M. Boareto, B. Huang, D. Jia, M. Lu, E. Ben-Jacob, J. N. Onuchic, and H. Levine, *Implications of the hybrid epithelial/mesenchymal phenotype in metastasis*, Frontiers Oncology, 5 (2015), 155.
[32] S. K. Kay, H. A. Harrington, S. Shepherd, K. Brennan, T. Dale, J. M. Osborne, D. J. Gavaghan, and H. M. Byrne, *The role of the hes1 crosstalk hub in notch-wnt interactions of the intestinal crypt*, PLoS Comput. Biol., 13 (2017), e1005400.
[33] S. Kéfi, M. B. Eppinga, P. C. de Ruiter, and M. Rietkerk, *Bistability and regular spatial patterns in arid ecosystems*, Theoret. Ecol., 3 (2010), pp. 257–269.
[34] C. Lv, X. Li, F. Li, and T. Li, *Constructing the energy landscape for genetic switching system driven by intrinsic noise*, PLoS One, 9 (2014), e88167.
[35] P. Macklin, S. McDougall, A. R. Anderson, M. A. Chaplain, V. Cristini, and J. Lowengrub, *Multiscale modeling and nonlinear simulation of vascular tumour growth*, J. Math. Biol., 58 (2009), pp. 765–798.
[36] N. A. Monk, J. A. Sherratt, and M. R. Owen, *Spatiotemporal patterning in models of juxtacrine intercellular signaling with feedback*, in Mathematical Models for Biological Pattern Formation, Springer, New York, 2001, pp. 165–192.
[37] J. M. Osborne, A. Walter, S. Kershaw, G. Mirams, A. Fletcher, P. Pathmanathan, D. Gavaghan, O. Jensen, P. Maini, and H. Byrne, *A hybrid approach to multi-scale modeling of cancer*, Philos. Trans. A, 368 (2010), pp. 5013–5028.
[38] R. D. O'Dea and J. R. King, *Continuum limits of pattern formation in hexagonal-cell monolayers*, J. Math. Biol., 64 (2012), pp. 579–610.
[39] R. Perez-Carrasco, P. Guerrero, J. Briscoe, and K. M. Page, *Intrinsic noise profoundly alters the dynamics and steady state of morphogen-controlled bistable genetic switches*, PLoS Comput. Biol., 12 (2016), e1005154.

[40] G. L. Poppe, Jr., *Physical Applications of the Geometric Minimum Action Method*, Ph. D. thesis, City University of New York, 2018.
[41] K. A. Rejniak and A. R. Anderson, *State of the art in computational modeling of cancer*, Math. Medicine Biology, 29 (2012), pp. 1–2.
[42] D. M. Roma, R. A. O'Flanagan, A. E. Ruckenstein, A. M. Sengupta, and R. Mukhopadhyay, *Optimal path to epigenetic switching*, Phys. Rev. E, 71 (2005), 011902.
[43] Y. Sha, D. Haensel, G. Gutierrez, H. Du, X. Dai, and Q. Nie, *Intermediate cell states in epithelial-to-mesenchymal transition*, Physical Biology, 16 (2019), 021001.
[44] D. Sprinzak, A. Lakhanpal, L. LeBon, L. A. Santat, M. E. Fontes, G. A. Anderson, J. Garcia-Ojalvo, and M. B. Elowitz, *Cis-interactions between notch and delta generate mutually exclusive signaling states*, Nature, 465 (2010), pp. 86–90.
[45] D. Stepanova, H. M. Byrne, P. K. Maini, and T. Alarcón, *A multiscale model of complex endothelial cell dynamics in early angiogenesis*, PLoS Comput. Biol., 17 (2021), e1008055.
[46] M. Tao, *Hyperbolic periodic orbits in nongradient systems and small-noise-induced metastable transitions*, Phys. D, 363 (2018), pp. 1–17.
[47] H. Touchette, *The large deviation approach to statistical mechanics*, Phys. Rep., 478 (2009), pp. 1–69.