Theoretical and computational tools to model multistable gene regulatory networks

Federico Bocci[1,2,*], Dongya Jia[3,#], Qing Nie[1,2], Mohit Kumar Jolly[7], José Onuchic[3-6,*]

[1]The NSF-Simons Center for Multiscale Cell Fate Research; [2]Department of Mathematics, University of California, Irvine, CA 92697, USA

[3]Center for Theoretical Biological Physics, [4]Departments of Physics and Astronomy, [5]Chemistry and [6]Biosciences, Rice University, Houston, TX 77005, USA

[7]Indian Institute of Science

[#]Current address: Immunodynamics Group, Laboratory of Integrative Cancer Immunology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20814,USA

Corresponding authors:
*Federico Bocci: fbocci@uci.edu
*José Onuchic: jonuchic@rice.edu

Abstract

The last decade has witnessed a surge of theoretical and computational models to describe the dynamics of complex gene regulatory networks, and how these interactions can give rise to multistable and heterogeneous cell populations. As the use of theoretical modeling to describe genetic and biochemical circuits becomes more widespread, theoreticians with mathematics and physics backgrounds routinely apply concepts from statistical physics, non-linear dynamics, and network theory to biological systems. This review aims at providing a clear overview of the most important methodologies applied in the field while highlighting current and future challenges, and includes hands-on tutorials to solve and simulate some of the archetypical biological system models used in the field. Furthermore, we provide concrete examples from the existing literature for theoreticians that wish to explore this fast-developing field. Whenever possible, we highlight the similarities and differences between biochemical and regulatory networks and "classical" systems typically studied in non-equilibrium statistical and quantum mechanics.

Table of Contents

1. Introduction

Despite sharing similar genetic information, cells in living organisms can diversify into many different cell types, each of whom carrying a distinct and unique function [1]. For example, cells in the human body differentiate into more than 200 different cell types that can be distinguished based on the different profiles of molecular signatures [2]. The process of molecule production can be summarized at a very simplified level by stating that special sequences of the genome (genes) are used as templates to synthetize messenger RNA (mRNA) molecules, which are in turn translated into proteins by the cell's machinery. Yet, transcription (gene to mRNA) and translation (mRNA to protein) are not independent processes, but rather happen in a complex cellular environment. For example, special molecules called transcription factors can regulate the production of other molecular species by binding to the DNA in proximity to genes and either facilitating or obstructing transcription [3]. This seemingly simple mechanism can give rise to a spectacular level of complexity when considering that living cells have tens of thousands of genes, and their mutual interactions are only partially understood [4,5].

The purpose of this review is to provide an overview of the mathematical strategies used to describe this complex phenomenon and predict how cellular heterogeneity emerges from the underlying regulatory dynamics. The fundamental assumptions connecting all these modeling strategies is that it is possible to describe, with some extent of success, the intricated interactions between genes and transcription factors with mathematical models (such as differential equations or logic circuits), and that the outputs of these models (such as probability distributions or attractors) can be related to the different states that cells can attain. We acknowledge the existence of other recent reviews that discuss the application of mathematical and physical concepts to chemical-reaction systems, stochastic events, and simulation strategies in biological systems [6–9]. Compared to these reviews, which explore the theoretical aspects in deeper details, we aim at presenting a more hands-on and simplified overview of the main theoretical strategies necessary to investigate the multistability and heterogeneity in biological systems, explicitly discussing examples and offering hands-on tutorials and simulations. With this approach, we aim to create a unifying

ground not only for theoreticians interested in biological systems, but also for biologists who are interested in the theoretical tools to model complex biological systems. In this sense, this review can be seen as a set of suggestions to tackle gene regulation and biochemical interactions at different levels of complexity.

In the first section of the review, we introduce theoretical and computational methods to model the emerging dynamics of gene regulatory networks and their contribution in many biological open problems. First, we discuss discrete, stochastic frameworks to model transcription, translation, and interactions between transcription factors. Further, we introduce a continuous framework that allows to model larger circuits with several transcription factors, and discuss how to characterize multistable systems with the pseudopotential landscape, the "systems biology" analogy to the potential landscape. Finally, we discuss strategies to tackle large circuits, where missing information about the model's parameters require approximation methods such as Boolean networks and parameter randomization, and provide an introduction to how gene regulatory networks can be inferred using high-resolution single cell sequencing data. This section also includes [hands-on tutorials](#) to simulate some of the simplest and most widely used circuit structures in the literature. In the second section of the review, we discuss in detail three specific biological examples where the application of these methods led to new and significant biological insights, including the Epithelial-Mesenchymal Transition, differentiation of Stem Cells, and Notch signaling. These examples, which were chosen not least due to the expertise of the authors, cannot fully do justice to the extensive available literature, and we would like to acknowledge the many outstanding works that could not be included here due to the limitation of space. While discussing these three examples, we introduce further mathematical tools that have been especially important in their respective fields, including: the modeling of non-coding RNA in the epithelial-mesenchymal transition, the application of quantum many-body formalism in the modeling of stem cell differentiation, and spatial models with many cells in the description of cell-cell communication through Notch signaling.

2. Mathematical methods to study the multistability of regulatory networks

2.1 Discrete models of circuits: from reactions to the Chemical Master Equation

2.1.1 Construction of the CME for protein synthesis

In this first section, we start by analyzing theoretical approaches that aim at modeling the stochastic nature of individual molecular reactions. The basic steps in molecule production include transcription, where a messenger RNA (mRNA) transcript is created, and translation, when the final molecule is produced from the mRNA transcript. This working model can be formalized by assuming that a gene transcribes mRNA molecules with constant rate k_m. Moreover, each mRNA molecule translates into proteins with rate constant k_p. Further, if assuming that mRNA and protein molecules degrade with rate constants γ_m, γ_p, we can write the following set of reactions:

$$\xrightarrow{k_m} m \xrightarrow{\gamma_m m} \emptyset \quad (1a)$$

$$\xrightarrow{k_p m} p \xrightarrow{\gamma_p p} \emptyset \quad (1b)$$
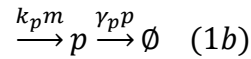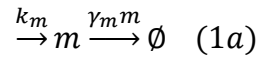
Reactions (1a-b) are summarized in figure 1A. Therefore, starting from given initial copy numbers of mRNA and protein (m, p) at time t, the next reaction will move the system to a new configuration (m', p') as shown in figure 1B. Since the rates in the reactions (1a-b) are either constant or depend linearly on (m, p), the copy number of mRNA or protein can be changed only by one molecule at a time by any of these reactions. This is not necessarily the case; in a more complex process molecules might, for instance, bind or degrade as oligomer complexes, which cannot be described by linear rates; we will discuss this case in more details at a later point when introducing Hill functions. It is also interesting to note that, in eqs. (1a-b), four rates control the copy number of two molecular species. Therefore, infinite combinations of the parameters $(k_m, k_p, \gamma_m, \gamma_p)$ can give rise to the same (m, p) state. Recent studies using high throughput data suggest, however, that only certain parameter combinations are naturally observed as a result of balancing between transcription

precision, which is achieved with large rate constants, and costs, which are minimized by low rate constants [10].

Starting from reactions (1a-b), we aim at constructing an equation to predict the probability to observe a certain combination of copy numbers (m, p) at any given time. To construct the Chemical Master equation (CME), we define a discrete, infinite set of probabilities $\{P(m, p, t)\}$, so that $P(m, p, t)$ describes the probability to observe m mRNA molecules and p protein molecules at time t. These probabilities satisfy a normalization condition (i.e., they sum up to one at any time t):

$$\sum_{n=0}^{+\infty} \sum_{p=0}^{+\infty} P(m, p, t) = 1 \quad (2)$$

Following the reaction scheme of Figure 1B, we can write an evolution equation for the set of probabilities:

$$\frac{dP(m, p)}{dt} = k_m[P(m - 1, p) - P(m, p)] + \gamma_m[(m + 1)P(m + 1, p) - mP(m, p)]$$
$$+ k_p m[P(m, p - 1) - P(m, p)] + \gamma_p[(p + 1)P(m, p + 1) - pP(m, p)] \quad (3)$$

In eq. (3), all the positive and negative terms represent influx and outflux of probability. In other words, positive terms in eq. (3) are associated with reactions that bring the system to the (m, p) state, whereas negative terms are associated with reactions that bring the system out of the (m, p) state (see again the reactions in Fig. 1B). Moreover, eq. (3) can be further simplified with reasonable assumptions when different terms on the right-hand size are associated to well-separated timescales. Specifically, one common approximation is to separate timescales for mRNA transcription and translation. Given the fast dynamics of transcription compared to translation, it can be assumed quasi steady state approximation (QSSA) for mRNA, thus $m = k_m/\gamma_m$. In other words, mRNA copy number equilibrates much faster than protein copy number and can thus be treated as constant. The resulting CME describes only the protein copy number p:

$$\frac{dP(p)}{dt} = +k[P(p-1) - P(p)] + \gamma_p[(p+1)P(p+1) - pP(p)] \quad (4)$$

Where $k = k_p m = k_p \, k_m / \gamma_m$. Eq. (4) describes a birth-death process with rate constants (k, γ_p). The steady state solution of eq. (4) can be evaluated with different analytical, numerical or simulation methods that are explored in detail in the following section.
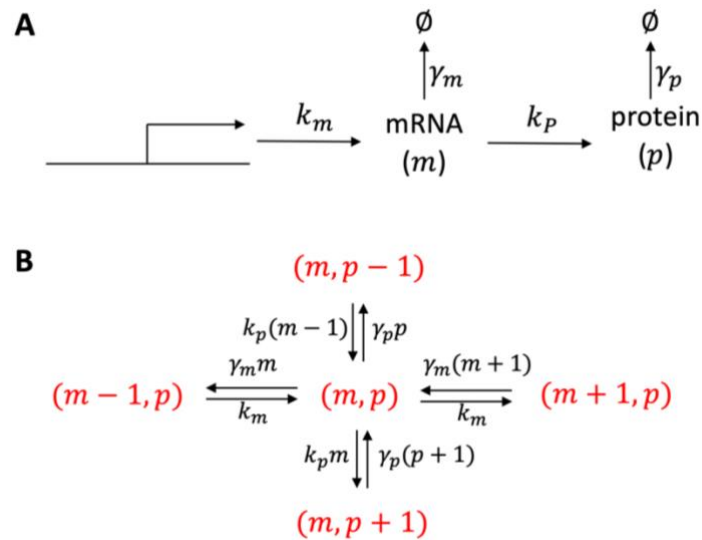
Figure 1. **Circuit and reactions for transcription-translation chain. (A)** *mRNA molecules are transcribed from the promoter with constant rate k_m. Each mRNA molecule is translated into proteins with rate constant k_p. Both mRNA and protein degrade with degradation rates γ_m, γ_p.* **(B)** *Schematic representation of all the possible reactions leading to or out of the configuration with m mRNA transcripts and p proteins.*

2.1.2 Solving and simulating the birth-death CME

The steady state solution of the CME (eq. 4) is the complete set $\{P_n{}^{SS}\}$ obtained by solving $\frac{dP_n{}^{SS}}{dt} = 0$. In this section, we discuss three methods that may be suitable to tackle different types of CMEs using the birth-death process (eq. 4) as an example. The insight box 1 offers the explicitly solution/implemented of the three methods. First, it is possible to find analytical solutions of the CME in certain special cases when the number of variables is

small, and the reactions are zeroth or first order. Typically, one first takes advantage of the fact that the equation for $P_0{}^{ss}$ has a simpler form since n is strictly semipositive. Then, all the other $P_n{}^{ss}$ can be found iteratively. This method can be successfully applied to the birth-death process, yielding $P_n = \frac{1}{n!}\left(\frac{k}{\gamma}\right)^n e^{-\frac{k}{\gamma}}$ (see insight box 1).

A more general framework to solve CMEs is obtained by converting the infinite-dimensional system of equations into the single differential equation of the generating function. Given the set of probabilities $\{P_n, n \propto [0, \infty]\}$, the generating function $\phi(z)$ is defined as:

$$\phi(z) = \sum_{n=0}^{\infty} P_n z^n \quad (5)$$

An equation for the temporal dynamics of $\phi(z)$ can be obtained by taking a time derivative and substituting the actual expression of $\frac{dP_n}{dt}$. Then, the probabilities P_n are obtained by differentiation the generating function:

$$P_n = \frac{1}{n!} \frac{\partial^n \phi}{\partial z^n} (z = 0) \quad (6)$$

Generating functions advantageously transform an infinite-dimensional problem into a one-dimensional problem. In the simple case of the birth-death process, this approach can be developed analytically (see insight box 1). Nonetheless, large number of variables and higher-order interaction terms typically prevent analytical solutions and potentially make numerical approaches expansive.

Finally, a third possibility to study the CME is to explicitly simulate the dynamics using algorithms such as non-rejective Monte Carlo, more commonly known as Gillespie algorithm [11,12]. In the Gillespie algorithm, all chemical reactions are treated as independent events (see for example [13,14]). Under this assumption, the waiting time until the next realization of any given reaction with rate w is an exponentially-distributed random variable with mean and standard deviation equal to $1/w$. Therefore, an iteration of a Gillespie simulation with $i = 1, \dots, M$ reactions includes the following steps. First, compute the individual reaction rates w_i. Second, draw an exponentially distributed random variable (τ) with rate parameter $W = \sum_{i=1}^{M} w_i$ that represents the waiting time until the next reaction occurs. Finally, draw a uniformly distributed random variable (r) to select which reaction occurs. The probability of

each reaction is given by its relative weight w_i/W (a pseudocode is presented in the insight box 1 for the birth-death process). The result of the simulation is a trajectory of the number of molecules $n(t)$ as a function of time starting from an initial condition specified by the user. In the long-time limit, the statistics of $n(t)$ will resemble closely and closely the steady-state distribution P_n obtained with iterative and generating function methods (provided that these approaches are feasible for the specific system under study). The long-time limit is identified as a timescale that is substantially larger than any typical timescale in the reaction scheme, and therefore depends on the parameters of the model (see the insight box 1 for this specific calculation in the birth-death process).

The steady state solution of the simplified birth-death process modeled by eq. (4) is a Poisson distribution with mean (μ_p) and variance (σ_p) both equal to k_P/γ_P. However, as discussed in more detail by Tsimring [7], a full model that takes into account mRNA dynamics leads to a broader protein copy number distribution with variance equal to:

$$\sigma_p = \mu_p \left(\frac{k_p}{\gamma_p + \gamma_m} + 1 \right) \quad (7)$$

Experimentally, it is well known that mRNA and protein count distributions in cells can exhibit variances much larger than the predicted Poisson distribution of the birth-death CME [15]. This effect arises due to the so-called transcriptional bursting. Namely, promoters do not transcribe mRNA molecules at a constant rate, but rather switch between 'ON' and 'OFF' periods, thus leading to bursts of mRNA molecule production on a typical timescale of a few minutes [15], thus leading to broader copy number distributions. The contribution of mRNA dynamics on protein copy number dispersion can be quantified with Gillespie simulations. The protein copy number distribution over the course of a long simulation is narrower in the simpler mRNA QSSA model, compared to the full mRNA-protein model (Fig. 2A). The relation between width of the distribution and mRNA QSSA can be quantified by testing the performance at different mRNA dynamics speed. The average, expected value of mRNA copy number is determined by the model's parameters: $m^{ss} = k_m/\gamma_m$. Therefore, increasing the both the mRNA production rate k_m and degradation rate γ_m by the same factor does not modify the expected mRNA copy number, but makes the mRNA equilibration dynamics faster, thus making mRNA QSSA more and more precise. As these parameters are increased, the standard deviation of protein (p) copy number in Gillespie

simulations decreases, and finally becomes comparable to the QSSA standard deviation when the mRNA dynamics is sufficiently fast (Fig. 2B). Further discussion of transcriptional bursting and mathematical attempts to quantitatively capture model it are reviewed in [7,15].
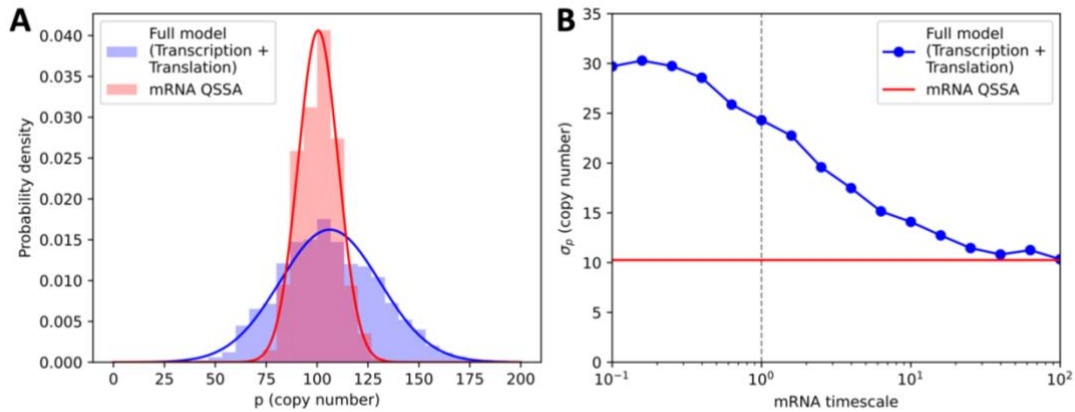
*Figure 2. **Testing the mRNA Quasi Steady State Approximation in protein synthesis. (A)** The distribution of protein copy number for the full model presented in eq. (3) (blue) and for mRNA QSSA corresponding to eq. (4) (red). Solid lines indicate Gaussian fit. **(B)** Blue: The standard deviation of protein (p) distribution for increasingly fast mRNA dynamics. The dashed black line indicates the parameters used in panel (A). Red: standard deviation for mRNA QSSA. The code to simulate protein synthesis is available in our [protein synthesis tutorial](#).*

2.1.3 CME for the toggle switch and larger networks

The CME formalism can be extended to study the coupled dynamics of an interconnected network of species. Perhaps the simplest circuit with the potential to give rise to multistability is the toggle switch composed by two genes that mutually repress each other, which can be seen as a basic mechanism to model differentiation between two distinct cell types [16]. In this section, we will set up the CMEs to describe the toggle switch. This method can serve as a general example to tackle more complex networks of interconnected biochemical species.

In the toggle switch, two genes (x and y) transcribe mRNA for the transcription factors (TF for brevity in the following) X and Y. Similar to the birth-death process of eq. (4), we will assume QSSA for mRNA dynamics to reduce the number of degrees of freedom and parameters. After being produced, TF X can inhibit the production of Y by binding to the promoter transcribing for gene y, and *vice versa* (see Figure 3A).

TFs do not necessarily bind and regulates the transcription of other species as single molecules; often, they first dimerize (or create higher-order oligomer complexes) and then bind to DNA. Here, we assume that both TFs X and Y dimerize with forward and backward rate constants k_D, k_M, as often assumed in existing mathematical modeling [17]. Considering the fast timescale of molecular binding compared to protein synthesis, we further assume that the dimerization/monomerization reactions are equilibrated (see Figure 3A). Thus, the concentrations of the dimers are $[X_2] = (k_D/k_M)X^2$ and $[Y_2] = (k_D/k_M)Y^2$. Then, the dimers bind and unbind to the DNA with rate constants k_+, k_-. Both TFs are produced at a rate k_0 when their respective promoters are not bound; the rate, however, decreases to $k_1 < k_0$ when the promoters are bound to the dimerized form of the TF. Therefore, the toggle switch can be viewed as two separate birth-death processes where the final molecular products X and Y bind to the other protein's gene to inhibit its production. All the considered reactions are depicted in Figure 3A.

The toggle switch can be characterized with four variables: two positive integer variables (n, m) accounting for the number of molecules for species X and Y, respectively, and two Boolean variables (S_X, S_Y) that encode the state of the two promoters. $S_i = 0$ implies that promoter of species i is unbound, while $S_i = 1$ implies that promoter is bound, respectively. Thus, starting from a generic configuration (n, m, S_x, S_y) at time t, a subsequent reaction can either modify the copy number of X and Y or change the state of one of the two promoters. All reactions and rates are illustrated in Figure 3B. Following the CME approach, we write an evolution equation for the probability $P(n, m, S_x, S_y)$. To write the CME in a compact form, we define the following functions. First, a generalized production rate function:

$$k(S) = k_0(1 - S) + k_1(S) \quad (8)$$

This function assumes the value k_0 when the promoter is unbound ($S = 0$) and k_1 when the promoter is bound ($S = 1$). Furthermore, a generalized rate function to describe binding/unbinding of a promoter:

$$f(S,k) = k_- S + \frac{k_+ k_D}{k_M} k^2 (1 - S) \quad (9)$$

This function returns the unbinding rate k_- when the promoter is occupied ($S = 1$) and the binding rate $\frac{k_+ k_D}{k_M} k^2$ if the promoter is free ($S = 0$). In this definition, k is the molecular copy number of the TF that inhibits the promoter. With these definitions, we can formally write the CME for the toggle switch by listing all the reactions depicted in Figure 3B:

$$\frac{dP(n, m, S_x, S_y, t)}{dt}$$

$$= -\left[\gamma n + \gamma m + k(S_x) + k(S_Y) + f(S_x, m) + f(S_y, n)\right] P(n, m, S_x, S_y, t) +$$

$$\gamma(n + 1)P(n + 1, m, S_x, S_y, t) + \gamma(m + 1)P(n, m + 1, S_x, S_y, t) +$$

$$k(S_x)P(n - 1, m, S_x, S_y, t) + k(S_y)P(n, m - 1, S_x, S_y, t)$$

$$+ f(1 - S_x, m)P(n, m, 1 - S_x, S_y, t) - f(1 - S_y, n)P(n, m, S_x, 1 - S_y, t) \quad (10)$$

In eq. (10), the first-row accounts for all the outflux terms from the configuration (n, m, S_x, S_y), including degradation, production and change in the promoters' states. The second-row accounts for influx due to degradation in configurations with higher molecule copy numbers; the third-row counts influx due to molecular production in configurations with lower copy number; and the fourth-row considers influx due to changes in the state of one of the promoters. This equation is far too complex for an exact solution with iterative methods; moreover, the relatively large number of variables makes eq. (7) stiff for generating function approaches. Simulations with the Gillespie algorithm, however, easily provide information about the relaxation dynamics and steady state (see for example [18,19]).

This treatment of the toggle switch can be considered as a footprint to tackle larger circuits of interconnected genes and transcription factors comprising transcription, translation, degradation, molecular binding and transcriptional regulation. As the circuits grow in size, understanding reasonable approximations to decrease the complexity of the model becomes crucial. Here, we applied some common approximations including QSSA for mRNA

dynamics and protein-protein binding. In general, these approximations yield good results when timescales are well-separable. Therefore, depending on the parameters of the specific system of interest, QSSA might or might not be a suitable approximation.
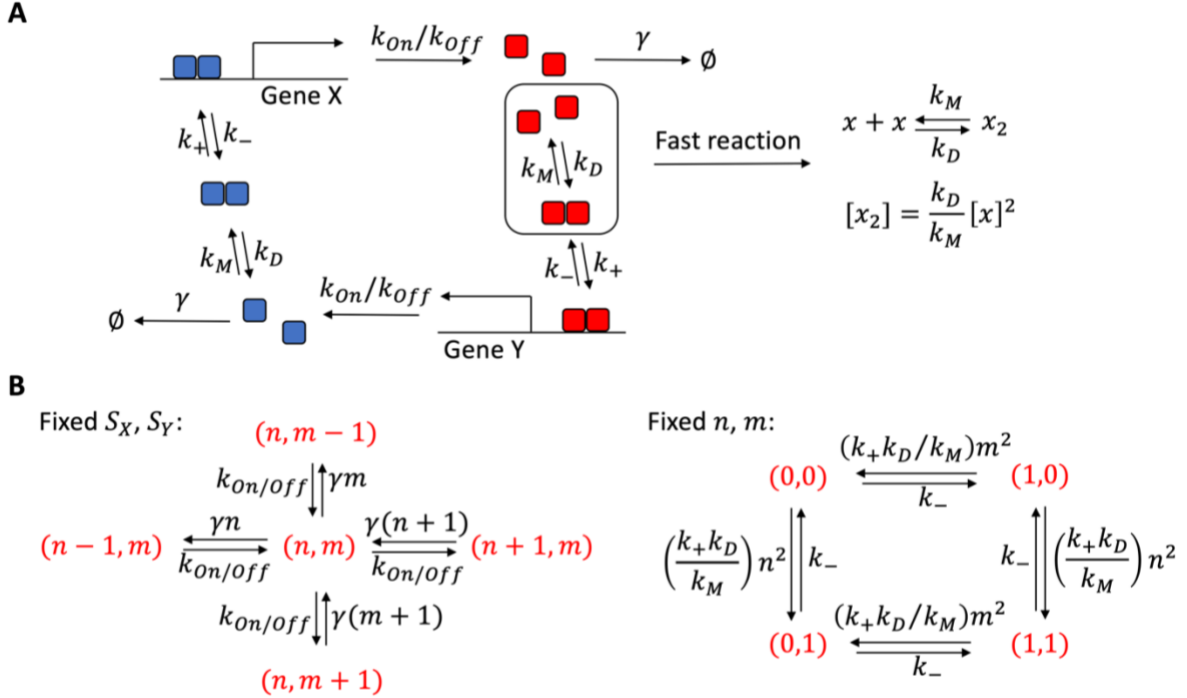
*Figure 3. **Circuit and reactions in a CME model of the toggle switch. (A)** In the toggle switch, gene x encodes for protein X (red blocks). X can dimerize and bind to DNA to inhibit the production of Y. Similarly, Y dimerizes and inhibits X. The dimerization reactions are fast compared to protein production and can thus be treated as instantaneously equilibrated. **(B)** Schematic of all the possible reactions leading to or out of the configuration (n, m, S_x, S_y). For convenience, reactions are separated into two groups: changes in molecule copy number and changes in the state of activity of one of the promoters.*

2.1.4 Insight box 1: three ways to compute the steady state of the birth-death CME

In this insight box, we discuss three approaches to tackle the birth-death CME derived in section 2.1:

$$\frac{dP_n}{dt} = -kP_n + kP_{n-1} - \gamma n P_n + \gamma(n+1)P_{n+1} \quad (11)$$

Iterative solution of the CME

This infinite system of equations can be solved by taking advantage of the simplified form of the equation for P_0 and then using induction. Since the copy number is strictly positive ($n \geq 0$):

$$\frac{dP_0}{dt} = -kP_0 + \gamma P_1 \quad (12)$$

Steady state $dP_0/dt = 0$ implies $P_1 = (k/\gamma)P_0$. Similarly, $dP_1/dt = 0$ yields:

$$P_2 = \frac{1}{2\gamma}(kP_1 - kP_0 + \gamma P_1) = \frac{1}{2}\left(\frac{k}{\gamma}\right)^2 P_0 \quad (13)$$

Where $P_1 = (k/\gamma)P_0$ was used to obtain the final expression. Continuing the process to higher-order equations, one finds the general expression for P_n:

$$P_n = \frac{1}{n!}\left(\frac{k}{\gamma}\right)^n P_0 \quad (14)$$

The constant P_0 can be calculated by imposing normalization of the total probability:

$$\sum_{n=0}^{\infty} P_n = P_0 \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{k}{\gamma}\right)^n = P_0 e^{\frac{k}{\gamma}} = 1 \quad (15)$$

Thus, the steady state set of probabilities follows a Poisson distribution with rate $\lambda = k/\gamma$:

$$P_n = \frac{1}{n!}\left(\frac{k}{\gamma}\right)^n e^{-\frac{k}{\gamma}} \quad (16)$$

Generating functions

Using the definition of generating function $\phi(z)$ introduced in section 1.2, we write the single ODE that describes the temporal variation of $\phi(z)$:

$$\frac{d\phi(z)}{dt} = \sum_{n=0}^{\infty} \frac{dP_n}{dt} z^n$$

$$= -k \sum_{n=0}^{\infty} P_n z^n + k \sum_{n=0}^{\infty} P_{n-1} z^n - \gamma \sum_{n=0}^{\infty} n P_n z^n + \gamma \sum_{n=0}^{\infty} (n+1) P_{n+1} z^n \quad (17)$$

This equation can be tackled as follows. The first term on the RHS is simply $-k\phi(z)$. In the second term, we shift indexing to $m = n - 1$; then, it becomes apparent that the term

equals $+kz\phi(z)$. In the third term, we can eliminate the factor n by exploiting $\frac{d}{dz}(z^n) = nz^{n-1}$; then, the term can be rewritten as $-\gamma z \frac{d\phi(z)}{dz}$. Finally, combining the tricks used for the second and third terms, the fourth term becomes $\gamma \frac{d\phi(z)}{dz}$. The long-time limit solution $\frac{d\phi(z)}{dt} = 0$ is then:

$$(1-z)\frac{d\phi(z)}{dz} = (1-z)\left(\frac{k}{\gamma}\right)\phi(z) \quad (18)$$

In any point $z \neq 1$, this equation has the simple general solution $\phi(z) = C_1 e^{\frac{k}{\gamma}z} + C_2$. The probabilities $\{P_n, n \propto [0, \infty]\}$ are then obtained by differentiating the generating function:

$$P_n = \frac{\phi^n(0)}{n!} = \frac{1}{n!}\left(\frac{k}{\gamma}\right)^n C_1 \quad (19)$$

This is exactly the same expression found with the iterative method, with the parameter C_1 in place of P_0. Therefore, normalization of the probability set $\{P_n\}$ yields $P_n = \frac{1}{n!}\left(\frac{k}{\gamma}\right)^n e^{-\frac{k}{\gamma}}$.

Gillespie simulation

In the simple case of the birth-death process, starting from a generic configuration at time t with n molecules, two reactions are possible: birth of a new protein molecule with rate k and death of an existing molecule with rate γn. Below we provide the pseudocode for the implementation of the Gillespie scheme for the birth-death process.

Pseudocode:

$n = n_{in}$ # set the initial number of molecules

$t = 0$ # start the simulation at t=0

while $t < T$:

      # compute total rate

      $W = k + \gamma n$

      # sample an exponentially-distributed random variable with rate parameter $1/W$

      $\tau = \exp\_rnd(1/W)$

      # sample a uniformly distributed random number in [0,1] to pick the reaction

```
# select birth reaction
```
if $uniform([0,1]) < \frac{k}{w}$:
```
    n = n + 1
# select death reaction
else:
    n = n - 1
# update time
```
$t = t + \tau$

In the birth-death process case, the birth rate is configuration-independent and always equal to k. The death rate, however, depends on protein number and is minimal at $n = 1$. Therefore, the simulation time (T) required to obtain substantial information about the steady state must at least satisfy $T \gg \max(^1/_k, ^1/_\gamma)$.

2.1.5 Explore further: RNA splicing

In this section, we mostly focused on transcription (production of mRNA) and translation (production of the finalized molecule) as the building blocks of cellular reaction systems. Newly transcribed RNA molecules, however, require a specific reaction where pre-determined portions of the RNA sequence are removed. This "splicing" reaction converts nascent to mature RNA, which can be then effectively translated into protein [20]. mRNA splicing has recently drawn interest in the systems biology field due to the development of new single cell sequencing technologies that now enable to precisely measure the amount of unspliced (nascent) and spliced (mature) RNA within individual cells [21,22]. This high-dimensional information has been used to build biochemical reaction models that infer interactions between chemical species and predict transitions between cellular states [23–25]. A detailed discussion of the mathematical approaches and consequences of splicing in the molecular dynamics can be found in [26,27].

2.2 Continuous models of regulatory networks

2.2.1 The chemical rate equation emerges as the average dynamics of the CME

In the previous section, we developed a probabilistic framework to account for the inherent stochasticity of biochemical reactions in regulatory networks. This probabilistic framework is especially important when the copy number of a protein is low within a cell, and thus stochastic fluctuations can substantially influence a biochemical circuit's response and potentially play a role in cellular function. A well-known example of the role of fluctuations is the selection of a competent state for DNA uptake in *Bacillus subtilis* [28]. In the limit of large molecular copy number, however, one can assume that stochastic fluctuations become less and less relevant. In this limit, the concentration or copy number can be described as a continuous variable. In more practical terms, when moving from the CME to a continuous formalism, we coarse-grain the model of a biochemical network by describing each chemical species with a single, continuous variable, whose temporal dynamics can be described by a single ordinary differential equation (ODE), rather than an infinite-dimensional system of CMEs.

The relation between the CME and the continuous chemical rate equation can be understood in the simple case of a birth-death process presented in eq. 4 by considering the dynamics of the average number of proteins $\langle n(t) \rangle$:

$$\frac{d\langle n(t) \rangle}{dt} = \sum_{n=0}^{+\infty} n \frac{d\langle P_n(t) \rangle}{dt}$$

$$= k \sum_{n=0}^{+\infty} [nP_{n-1}(t) - nP_n(t)] - \gamma \sum_{n=0}^{+\infty} [n^2 P_n(t) - n(n+1)P_{n+1}(t)] \quad (20)$$

After some manipulation of the summation indexes, eq. (20) can be simplified as:

$$\frac{d\langle n(t) \rangle}{dt} = k - \gamma \langle n(t) \rangle \quad (21)$$

Therefore, the evolution of the average number of molecules, which is now a continuous variable, is described by a single ODE.

This approach facilitates the treatment of large networks with many interconnected regulations that would be unfeasible with the CME approach. In general, a circuit of N interconnected species can be described as a dynamical system:

$$\dot{x} = F(x) \quad (22)$$

Where $x = (x_1, x_2, \ldots, x_N)$ is the vector of concentration/copy number of the N species in the circuit and the force field $F(x)$ describes their coupled dynamics and the possible interconnections between them. Therefore, all the prior knowledge on the circuit of interest is encoded by the choice of functions and parameters in the force field. The force field $F(x)$ captures all the relevant reactions, including transcription, translation, chemical binding and degradation. The continuous formulation of eq. (22) conveniently allows to apply all the tools of nonlinear dynamics, including linear stability analysis, phase diagrams and bifurcations that are usually applied to classical physical problems. In the following section, we will develop some of these methods in the study of the continuous toggle switch.

2.2.2 Continuous model of the toggle switch

Following the parallel with the treatment of the CME, we develop a continuous formalism for the toggle switch system composed of two genes that mutually repress each other. We also exploit this example to present some common functions used to model transcriptional activation and inhibition in the gene regulatory network literature. In the CME, the mutual inhibition between two TFs was explicitly modeled by considering TF dimerization and binding to the DNA. The main challenge of the continuous version is to find a suitable force field $F(x)$ that effectively captures the same dynamics. This is typically achieved by introducing a continuous function that modulates the transcription rate. First, we can introduce transcriptional regulation in the simpler case of the noise-free chemical rate equation of the birth-death process:

$$\frac{dp}{dt} = k \, H(R) - \gamma p \quad (23)$$

Here, R is the concentration of a transcription factor that inhibits or activates the production of p; $H(R)$ is a continuous function that modulates the production rate. In general, $H(R)$ must satisfy the condition $H(0) = 1$ so that the basal transcription rate k is recovered in the absence of external regulation. Then, $H(R)$ increases or decreases

monotonically as a function of R depending on whether R activates or inhibits p. Assuming equilibration of both TF dimerization/multimerization and TF-DNA binding leads to the so-called Hill function, which was originally introduced in 1910 to describe the binding of oxygen to hemoglobin [29], and has been since applied to a variety of biological contexts [30]:

$$H(R) = \frac{1 + \lambda \left(\frac{R}{R_0}\right)^n}{1 + \left(\frac{R}{R_0}\right)^n} \quad (24)$$

The functional form of the Hill function in Eq. (24) is derived exactly from the underlying chemical reactions in the insight box 2. In this expression, R_0 represents a threshold concentration that is related to the rate constants for TF-TF and TF-DNA binding (see insight box 2). Therefore, when $R > R_0$ the regulatory effect of R becomes important. λ is a fold-change that represents the change in transcription rate due to R. In the limit $R \gg R_0$, the production rate relaxes to $k\lambda$. Therefore, $0 < \lambda < 1$ indicates transcriptional inhibition and $\lambda > 1$ indicates transcriptional activation. Lastly, the Hill coefficient n indicates the cooperativity of the transcription factor. For instance, $n = 1$ indicates that R binds to the promoter of p as a single molecule, $n = 2$ indicates dimerization of the TF, and so forth. It is important to stress that a Hill-like relation between protein and TF does not necessarily imply TF cooperation and binding, as other biological processes upstream or downstream of transcriptional regulation that are not accounted for in the model can lead this nonlinear response of eq. (24). The insight box 2 further discusses how these parameters can be fitted or inferred from experiment observations.

In this continuous formalism, the toggle switch can be thus described by a set of two ODEs of the form:

$$\frac{dx}{dt} = k \frac{1 + \lambda \left(\frac{y}{t_0}\right)^n}{1 + \left(\frac{y}{t_0}\right)^n} - \gamma x \quad (25a)$$

$$\frac{dy}{dt} = k \frac{1 + \lambda \left(\frac{x}{t_0}\right)^n}{1 + \left(\frac{x}{t_0}\right)^n} - \gamma y \quad (25b)$$

For simplicity, we assumed that both x and y are produced and degraded with the same rate constants. Furthermore, the system is completely symmetric as the Hill thresholds, fold-changes and coefficients are the same for the two species. The ODE system of eq. (25) can exhibit bistability between a (high X, low Y) and a (low X, high Y) states (Fig. 4A). A potential pitfall of this modeling approach is the uncertainty in the estimation of the model's parameters. The dependence on the parameter choice can be investigated with parameter sensitivity approaches. Figure 4B shows the change of steady state X level in the (high X, low Y) state upon varying the model's parameter one at a time by 5%. More sophisticated approaches to parameter sensitivity exist in the current literature that can be reviewed in references[31,32].
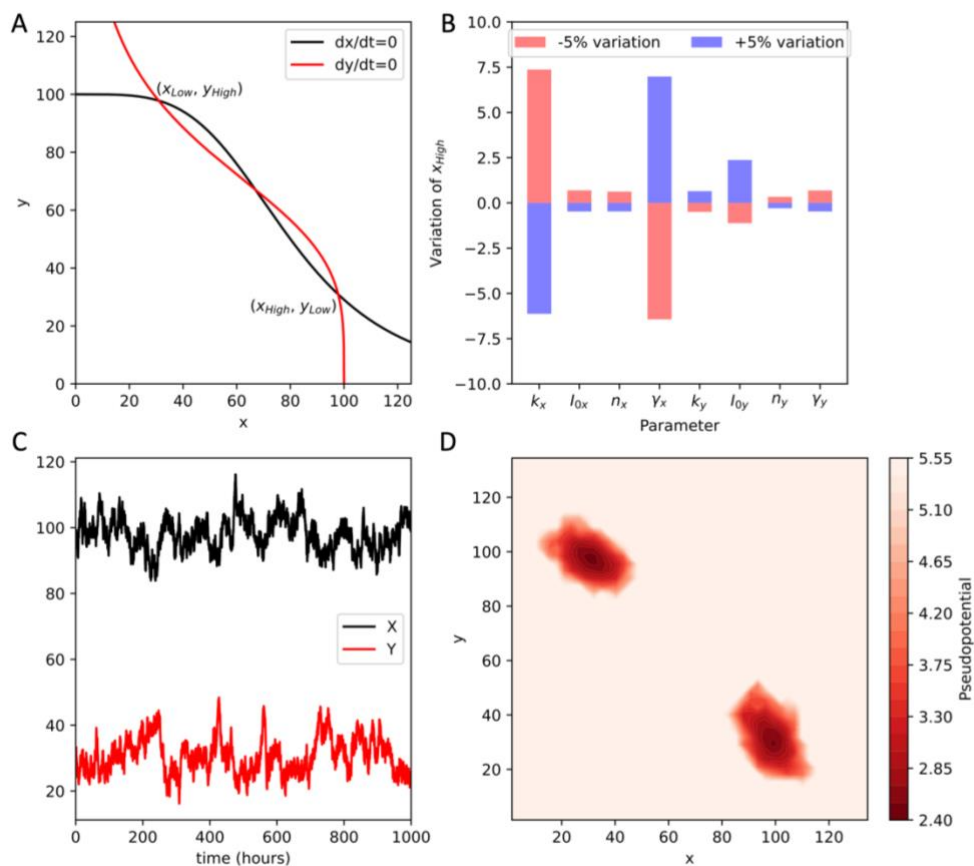

Figure 4. The deterministic and stochastic toggle switch. (A) The phase space of the toggle switch described in eq. (25) including nullclines and fixed points. (B) Local sensitivity analysis of the toggle switch parameter around the (high X, low Y) stable fixed point. (C) A temporal trajectory for the stochastic toggle switch. (D) The pseudopotential landscape of the toggle

2.2.3 Introducing noise in continuous circuits with the overdamped Langevin equation

We derived the continuous chemical rate equation as a noise-free approximation to the CME. In many situations, however, the continuous limit is utilized because the complexity and size of the circuit of interest prevents a treatment with discrete modeling. The effect of stochastic fluctuations, however, can be still incorporated in continuous models in a coarse-grained manner by considering the stochastic differential equation:

$$\dot{\boldsymbol{x}} = F(\boldsymbol{x}) + \xi(\boldsymbol{x}, t) \quad (26)$$

This equation describes the motion of a particle in the overdamped limit ($\ddot{\boldsymbol{x}} = 0$) under the presence of a force $F(\boldsymbol{x})$ and a noise term $\xi(\boldsymbol{x}, t)$ that satisfies $\langle \xi(\boldsymbol{x}, t)\xi(\boldsymbol{x}, t')\rangle = 2D\,\boldsymbol{D}\,\delta(t - t')$. In other words, the noise term is a random variable whose intensity at any time t is completely uncorrelated to the intensity at any other given time. D is a constant with the interpretation of an effective diffusion coefficient. In most of the current models of genetic circuits, the noise is assumed to be white. Thus, $\boldsymbol{D}$ is simply a unitary matrix. In other words, the noise term for each species has the form $\xi(x, y, t) = \sigma\, N(0,1)$, where σ is the noise amplitude and $N(0,1)$ is a Gaussian random variable with zero mean and unitary variance. In the case of the bistable toggle switch, white noise introduces stochastic fluctuations can induce transitions between the two stable states (Fig. 4C).

It is important to stress that stochastic fluctuations modeled in the previous section with Gillespie simulations are functions of the models' parameters. In other words, each chemical species has its own 'personal' noise amplitude, which could furthermore be a function of time as the species' reaction rates depend on other species' copy numbers. For this reason, the stochastic fluctuations arising from biochemical reactions and modeled with Gillespie-style simulations are referred to as 'intrinsic noise'. Conversely, the so-called 'extrinsic noise' broadly captures all other perturbations related to other cellular

components and/or external factors in the cell's local microenvironment [33,34]. Modeling fluctuations with white noise implies that all species in a circuit are subjected to a fixed noise intensity. This assumption could become especially problematic when certain species are near an extinction point, such as an attractor where the copy number of a given species is very low. For this reason, other ways to model noise have been explored. One notable example is shot noise where $\boldsymbol{D}$ is a diagonal matrix with $D_{ii} = x_i$. In other words, the intensity of noise for the i-th species is proportional to the square root of the variable's copy number ($\sqrt{x_i}$). Therefore, variables with higher copy number experience fluctuations with larger amplitude. On the other hand, the noise-to-signal ratio $\sqrt{x_i}/x_i = 1/x_i$ is lower for high copy number species. As a practical example, a species with average copy number $n = 10^2$ molecules will have typical fluctuations of ~10 molecules, or 10% of its average copy number. Conversely, the average fluctuations of a species with copy number of $n = 10^4$ will be larger in absolute intensity (~10^2 molecules) but will only correspond to 1% of the variable's copy number. Thus, species with low copy number can be subjected to large deviations somewhat similar to the bursting effect discussed in the context of the CME. As noted by Lu and collaborators, a combination of white and shot noise is likely a more precise way to model stochastic fluctuations in continuous, ODE model of biochemical regulatory networks as shot noise captures low copy number deviations while white noise captures external inputs [35].

2.2.4 Characterizing the landscape of multistable gene regulatory networks
In general, a multistable circuit governed by the overdamped Langevin equation will exhibit multiple attractors and, as in the case of the bistable toggle switch, stochastic fluctuations introduce the possibility of transitions from one state to another. Landscape theory is a popular way to quantify the global stability in a multistable system by computing the equivalent of a potential function.

The idea that cells follow "paths" in some underlying, high-dimensional landscape that guide the transition toward new cell states was first formalized by Waddington, albeit at a qualitative level. In Waddington's representation, cells are visualized as marbles that roll through a downhill landscape where valleys and ridges define alternative routes leading to

different final states. In turn, the location and structure of these "valleys" is controlled by the underlying interactions between genes, mRNAs and transcription factors [36].

The construction of the landscape begins by mapping the Langevin dynamics onto a Fokker-Planck (FP) equation describing the joint probability distribution $p(x_1, x_2, \ldots x_N, t)$:

$$\frac{\partial p(\boldsymbol{x},t)}{\partial t} = -\nabla \cdot [F(\boldsymbol{x})p(\boldsymbol{x},t)] + \nabla^2[D(\boldsymbol{x})p(\boldsymbol{x},t)] \quad (27)$$

In this equation, $F(\boldsymbol{x})$ is the force field driving the Langevin dynamics and $D(\boldsymbol{x})$ is a diffusion tensor related to the noise amplitude by $D(\boldsymbol{x}) = \sigma(X)\sigma^t(x)/2$. We will not discuss the derivation of the FP equation form the Langevin equation, which can be found in many textbooks on non-equilibrium statistical mechanics. The FP equation can be written more compactly in the form of a continuity equation:

$$\partial p(\boldsymbol{x},t)/_{\partial t} + \nabla \cdot J = 0 \quad (28)$$

where $J(\boldsymbol{x},t) = F(\boldsymbol{x})p(\boldsymbol{x},t) - \nabla \cdot [D(\boldsymbol{x})p(\boldsymbol{x},t)]$. From eq. (2), it is evident that the steady state solution $\partial p(\boldsymbol{x},t)/_{\partial t} = 0$ requires a divergence-free flux $\nabla \cdot J = 0$. This condition can be satisfied in different ways. Systems that respect detailed balance satisfy the stringent condition $J = 0$; in this case, there is a clear relation between force, potential and equilibrium probability distribution. Indeed, enforcing $J = 0$ implies:

$$F(\boldsymbol{x}) = \frac{\nabla \cdot [D(\boldsymbol{x})p_{eq}]}{p_{eq}} = D(\boldsymbol{x})\frac{\nabla P_{SS}}{P_{SS}} + \nabla \cdot D(\boldsymbol{x}) \quad (29)$$

Rewriting $\frac{\nabla P_{eq}}{P_{eq}} = -\nabla \log(P_{eq})$ clarifies the relation between potential and the equilibrium probability distribution $U = -\log(P_{eq})$. The dynamics of the system is completely determined by the gradient of the potential U (up to a shift due to local variation of the diffusion matrix $\nabla \cdot D(\boldsymbol{x})$). In the simple case of white noise, the diffusion is uniform ($\nabla \cdot D(\boldsymbol{x}) = 0$) and the correspondence between force and potential gradient is exact. Typically, force fields describing regulatory networks are constructed in a phenomenological manner to capture biochemical interactions such as transcription, binding or degradation with the use of functional forms such as Hill functions. For this reason, they cannot be

derived from an underlying potential and do not satisfy $J = 0$. Therefore, at steady state, the force field can be written as the sum of two components. A gradient of the log-scaled steady-state probability distribution and flux term:

$$F(\boldsymbol{x}) = D(\boldsymbol{x}) \cdot \frac{\nabla P_{SS}}{P_{SS}} + \nabla D(\boldsymbol{x}) + \frac{J_{SS}}{P_{SS}} \quad (30)$$

The flux term J_{SS} effectively quantifies the deviation from gradient-driven dynamics. Since there are no sources or sinks of probability in the domain of existence of $\boldsymbol{x}$, J_{SS} must be locally curled, and is therefore often referred to as the curl flux. As observed by Wang, the dynamics of 'gradient' dynamical systems that respect detailed balance can be compared to the motion of an electron in an electric field; conversely, the dynamics of 'non-gradient' systems such as gene regulatory networks is similar to the motion of an electron in an electromagnetic field [37].

Fig. 4D shows the pseudopotential landscape of the bistable toggle switch under the effect of white noise, which include two deep minima corresponding to the two stable fixed points of the deterministic system.

Therefore, in non-gradient systems such as coupled biochemical networks, the steady-state probability distribution $p(\boldsymbol{x}, t)$ cannot be established *a priori*, but only as a result of stochastic simulations of the Langevin equation or steady state solution of the FP equation. Nonetheless, the steady-state probability distribution and curl flux provide key information about (1) the number of accessible states, (2) their relative stability, and (3) the deviation from gradient-driven dynamics.

2.2.5 Insight box 2: Modeling transcriptional regulation with Hill functions

In the simplest model of protein production, a gene transcribes mRNA at a constant rate k_m; in turn, mRNA molecules are translated into proteins with rate constant k_p. Therefore, the copy number of mRNA (m) and protein (p) can be described by first order equations:

$$\frac{dm}{dt} = k_m D_0 - \gamma_m m \quad (31a)$$

$$\frac{dp}{dt} = k_p m - \gamma_p p \quad (31b)$$

Where D_0 is the total number of promoters transcribing for the protein p, and can be treated as a constant unless DNA is being replicated during cell division; γ_m and γ_p are degradation rate constants for mRNA and protein. Assuming fast equilibration of mRNA ($dm/dt = 0$), the mRNA copy number is $m = k_m D_0 / \gamma_m$ and the protein copy number directly depends on the number of genes D_0:

$$\frac{dp}{dt} = k D_0 - \gamma_p p \quad (32)$$

Where $k = k_m k_p / \gamma_m$. In presence of a transcriptional factor x that activates or inhibits gene expression, the mRNA production rate is modulated by a function $k_m \to k_m f(x)$ (Figure 5A). First, we assume that the transcription factor can bind and unbind to DNA as a single molecule with rate constants k_+, k_-. Assuming fast equilibration of the TF-DNA binding, the fraction of genes that are unbound (U) and bound (B) to the transcription factor are:

$$U = D_0 \frac{1}{1 + \frac{x}{x_0}} \quad (33a), \quad B = D_0 - U = D_0 \frac{\frac{x}{x_0}}{1 + \frac{x}{x_0}} \quad (33b)$$

Where $x_0 = k_+/k_-$ represents the half-concentration so that $U = B$ when $x = x_0$. The mRNA is transcribed with rate constant k_u by an unbound promoter and k_b by a bound promoter, respectively. $k_b > k_u$ implies that x is an activator while $k_b < k_u$ implies that x is an inhibitor. The corresponding protein (p) is then produced with rate $(k_u k_p / \gamma_m) U + (k_b k_p / \gamma_m) B$. Thus, the dynamics of protein p is governed by the following equation:

$$\frac{dp}{dt} = k^p{}_U \frac{1}{1 + \frac{x}{x_0}} + k^p{}_B \frac{\frac{x}{x_0}}{1 + \frac{x}{x_0}} - \gamma_p p = k^p{}_U \frac{1 + \lambda \frac{x}{x_0}}{1 + \frac{x}{x_0}} - \gamma_p p \quad (34)$$

Where $k^p{}_U = (k_u k_p / \gamma_m)$ and $k^p{}_B = (k_b k_p / \gamma_m)$. In the compact form on the right-hand side, $\lambda = k^p{}_B / k^p{}_U$ represents a fold-change in the transcription of p due to the regulatory activity of x.

In the more complex case where the transcription factor dimerizes reversibly ($x + x \leftrightarrow x_2$) with dimerization and monomerization rate constants k_2, k_1, and binds and unbind to DNA only in its dimer form x_2, the fractions of bound and unbound DNA become:

$$U = D_0 \frac{1}{1 + \left(\frac{x}{x_0}\right)^2} \quad (35a), \quad B = D_0 - U = D_0 \frac{\left(\frac{x}{x_0}\right)^2}{1 + \left(\frac{x}{x_0}\right)^2} \quad (35b)$$

Where $x_0 = \sqrt{k_1 k_+ / k_2 k_-}$. The corresponding equation for protein p becomes:

$$\frac{dp}{dt} = k^p{}_U \frac{1 + \lambda \left(\frac{x}{x_0}\right)^2}{1 + \left(\frac{x}{x_0}\right)^2} - \gamma_p p \quad (36)$$

In general, a transcription factor that binds to DNA as an n-oligomer is described by the n-th power of its monomeric concentration. If the transcription factor is an activator, it is often assumed that the transcription rate of the unbound gene is zero (i.e. $k_u = 0$), hence:

$$\frac{dp}{dt} = k \frac{\left(\frac{x}{x_0}\right)^n}{1 + \left(\frac{x}{x_0}\right)^n} - \gamma_p p \quad (37)$$

Where the concentration of DNA (D_0) has been absorbed into the production rate constant k. Similarly, if a transcription factor is a very effective inhibitor, no transcription is assumed for the bound genes (i.e. $k_b = 0$), hence:

$$\frac{dp}{dt} = k \frac{1}{1 + \left(\frac{x}{x_0}\right)^n} - \gamma_p p \quad (38)$$

The expressions in eq. (37) and eq. (38) are referred to as positive and negative Hill functions, whereas the general form presented in eq. (36) is referred to as shifted Hill function. Detailed data about TF oligomerization is not always available. The Hill coefficient, half-concentration and fold-change (n, x_0, λ), however, can still be inferred from an experiment describing protein concentration as a function of TF concentration (Figure 5B). In this case, however, there is no guarantee that the inferred Hill coefficient truly reflects TF oligomerization, as other processes can modulate transcription and change the protein-TF relation.
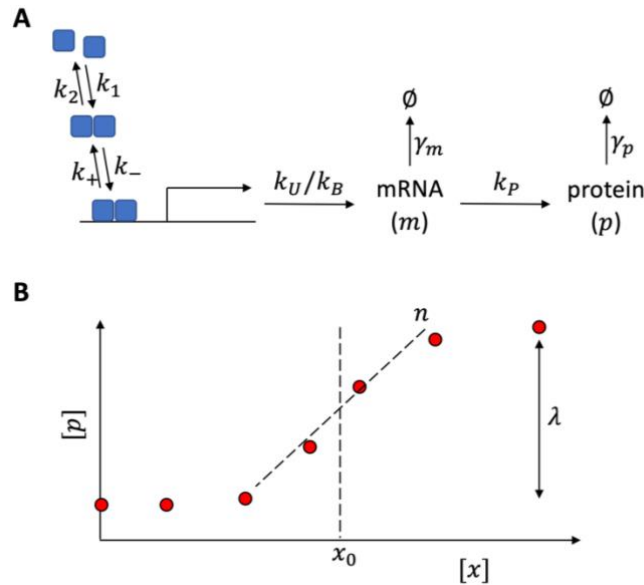
*Figure 5. **Derivation and fitting of the Hill function.** **(A)** The derivation of the Hill function considers a case where the regulator can dimerize (or create a higher-order oligomer) and then bind to DNA. **(B)** The parameters of the Hill functions can be inferred from experimental data reporting the expression of the molecule p as a function of transcription factor concentration.*

2.2.6 Explore further: Transitions in the pseudopotential landscape

One of the main quests in the modeling of gene expression and regulatory networks is understanding noise-driven transitions between stable states. For example, the attractors in a multistable pseudopotential landscape might represent the states that a cell can assume. Therefore, a transition from one attractor to the next represents a transition from one cell state to another. These transitions can be studied with long simulations of the overdamped Langevin equation (such as the trajectory in Fig. 4C). This approach, however, is time consuming and often cannot provide large statistics because escape events from attractors are rare. More complex methods based on the path integral formalism have been widely applied to biochemical systems to reconstruct transition trajectories and rates in a more computationally efficient manner. Interesting examples of these approaches include [38,39], and are summarized in [40] for interested readers.

2.3 Large networks: Boolean models, parameter randomization and interfacing with data

In the previous two sections we discussed CME and continuous models that describe genetic and biochemical interactions in a mechanism-based, detailed manner. These approaches can become impractical when applied to large networks with dozens, or even hundreds, of chemical species. These considerations are becoming more and more relevant as detailed experimental techniques provides detailed insight into the organization of large regulatory networks, thus providing an opportunity to build and benchmark larger models of gene regulation. A first challenge, from a computational standpoint, is to perform long simulations on models with hundreds of reactions and/or ODEs. A second, more fundamental problem is the difficulty to precisely estimate all the parameters of the model. Quite often, only a qualitative relation can be established between species in a network, such as inhibition or induction. In this section, we discuss two classes of models that resolve these issues and allow the modeling of large networks. First, we discuss Boolean models where nodes can only be active or inactive, and transcriptional and biochemical interactions are described via logic operators. Second, we present parameter randomization approaches that assume the functional form of the interactions and test many different combinations of parameter values. Finally, we introduce how the problem of determining the gene regulatory network architecture and connections can be tackled by interfacing quantitative modeling and high-resolution single cell sequencing data.

2.3.1 Boolean Models

Boolean models perhaps represent the most coarse-grained approach to describe regulatory networks since they require a minimal amount of information including solely the circuit's connections of positive or negative interactions among genes [41]. For this reason, Boolean approaches stand out as some of the first attempts to model gene regulation [42]. In Boolean models, each node can either be ON ($\sigma = 1$) or OFF ($\sigma = 0$), thus representing a gene or molecular species that is active or inactive, respectively. Therefore, the state of a regulatory network of N genes at any given time t is specified by an array of N Boolean

variables $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$. The value of each node i (σ_i) is a function of all the incoming signal from nodes that directly regulate node i. These functions B_i are typically constructed in an *ad hoc* manner starting from experimental evidence and using the logic commands AND, OR and NOT. Similar to their continuous counterparts, Boolean networks can be represented graphically as a set of nodes and arrows to indicate mutual activation or inhibition, such as shown in figure 6A. It is important to stress that the functional form of the functions B_i cannot be inferred by simply looking at the schematic representation. For example, in the circuit of figure 6A, x and z both activate the node y. Their input, however, can be combined in an AND or OR manner, which must be specified.

Starting from an initial configuration, the values of the nodes are updated according to an integration scheme. In particular, the synchronous scheme is the simplest update method on a Boolean network. Under this deterministic scheme, the state of each node (i) at iteration $t + 1$ is the output of the Boolean function B_i computed at time t:

$$\sigma_i(t+1) = B_i(\boldsymbol{\sigma}(t)) \quad (39)$$

This method presents some important drawbacks. Specifically, reactions in regulatory networks can be time-separated, with given reactions evolving on faster timescales than others. A popular solution is given by asynchronous updating schemes, which can be considered as a Boolean equivalent of the Gillespie algorithm. With this method, any reaction in the network is associated with a typical timescale. At any updating step, one node is selected with a probability proportional to the frequency of that reaction. Therefore, faster reactions are more probable and thus selected more often. These reference timescales are not known *a priori*, so they must be inferred from literature and given as an additional input in the construction of the model. A more in-depth discussion and comparison between different integration schemes is provided in [43,44].

Starting from a given initial condition, a Boolean network relaxes to an attractor or limit cycle by following the update scheme. In figure 6A, we present a simple example of a 3-nodes network originally proposed by Wang and collaborators [45] with the following Boolean functions:

$$B_x = x \ OR \ (NOT \ z) \quad (40a)$$
$$B_y = x \ AND \ z \quad (40b)$$

$$B_z = y \quad (40c)$$

Three different initial conditions lead to three different steady states, including two stable attractors and a limit cycle (Figure 6B). More generally, a large Boolean network with N nodes has 2^N different initial conditions that can lead to multiple attractors and/or limit cycles. Typically, their basin of attraction is quantified with a large number of simulations starting from randomized initial conditions.
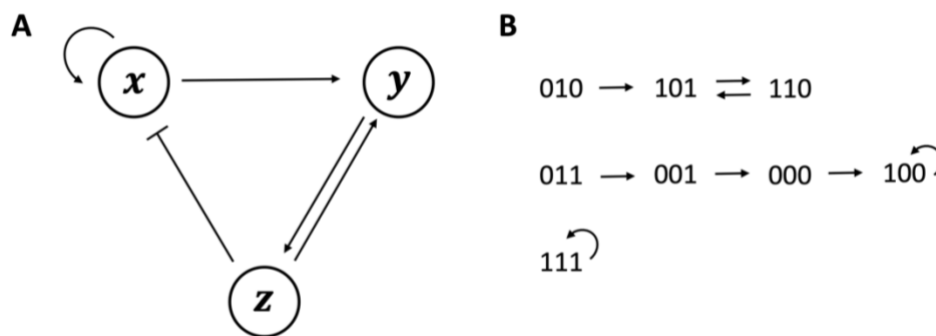
Figure 6. (A) Example of a three-node Boolean model. (B) Three different initial conditions that lead to oscillations between two configurations

2.3.2 Parameter randomization

Boolean models overcome the insufficient knowledge of the model's parameters by assuming more coarse-grained relationships between species based on logical operators, thus sacrificing the detailed description of biochemical reactions. An alternative approach is to hypothesize more detailed functional forms to describe the circuit's reactions, such as hill functions for transcriptional regulation, and explore several parameter combinations.

For example, *ra*ndom *ci*rcuit *pe*rturbation (RACIPE) takes the topological information of a network (i.e., the list of positive and negative interactions between species) as the input and generates an ensemble of mathematical models[46]. Each mathematical model, also referred to as a RACIPE model, is simulated with the same set of chemical rate equations but with distinct sets of parameters. Each set of parameters are randomly sampled from a fixed distribution (uniform, Gaussian etc.) within their biologically reasonable ranges. For

each RACIPE model, multiple (typically tens of thousands) initial conditions are used to perform the numerical simulation to solve all possible stable-state solutions. The stable-state solutions from the ensemble RACIPE models and their corresponding parameters are collected for statistical analysis, by which the pattern of solutions and the differences in parameters leading to different solutions can be identified. RACIPE originally focuses on transcriptional regulatory networks and uses the shifted Hill function[47] to simulate the transcriptional regulations. The method RACIPE has been implemented as a free open source software which can be accessed in GitHub[48]. To reach convergence of RACIPE simulation results, there are two key parameters - the number of RACIPE models and the number of initial conditions for each model should be evaluated and decided. As the average simulation time of a RACIPE model is linearly proportional to the number of parameters of that model, RACIPE can be potentially applied to large gene networks. Notably, the main purpose of RACIPE is to determine the robust stable states enabled by the gene networks by considering the large variations in parameters. For example, parameter randomization of the toggle switch topology using RACIPE showed that, even though the toggle switch topology is typically associated with bistability, more than 75% of parameter combinations lead to a single stable fixed point[46].

To detect the stability of different stable states exhibited by the gene network, a method integrating RACIPE with stochastic analysis, referred to as **sRACIPE**, has been recently developed[49]. The sRACIPE employs a simulated annealing-based scheme to estimate the stability of different stable states resulting from various initial conditions. One interesting observation by applying sRACIPE to toggle-switch-like circuits is that high expression noise induced state merging. Recently, the Boolean and randomized approaches were compared on the same set of network topologies, showing that both methods can successfully recover the important dynamical features of the circuits[50].

2.3.3 Learning regulatory networks from single cell transcriptomics

So far, we have always assumed the existence of biological knowledge that serves as input to build GRN models. In many cases, however, the interactions between genes are not well known, and alternative, data-driven strategies are necessary. Over the last decade, many

single cell Omics technologies have started to provide unprecedented resolution over gene regulation at the single cell level [51]. In this section, we focus specifically on single cell RNA sequencing (scRNA-seq), a technology capable to measure the copy number of tens of thousands of RNA species simultaneously within a cell [52]. The output of a scRNA-seq experiment can be intuitively understood as a count matrix where the rows identify individual cells, and the columns identify the genes in the cell's genome. While this data can be very useful to investigate the gene expression patterns associated to different types of cells, recently several models have been proposed to infer the regulation between genes [53,54]. Here, we briefly introduce the topic and focus on the emerging mathematical challenges. The inference of GRN from the scRNA-seq data introduces the challenge of inferring causal relationships from data that lacks temporal information. Generalizing the mathematical models of GRN developed in previous sections, the GRN inference problem can be generally set up as a nonlinear, high-dimensional system of the form:

$$\dot{\boldsymbol{X}} = f(\boldsymbol{X}) \quad (41)$$

In eq. (24), $\boldsymbol{X}$ represents the vector of mRNA counts of the N genes in the cell, and $f(\boldsymbol{X})$ is a general nonlinear function that describes how genes mutually regulate each other. In most existing applications, the problem is further simplified by assuming linear interactions between genes and degradation:

$$\dot{\boldsymbol{X}} = A\boldsymbol{X} - \gamma\boldsymbol{X} \quad (42)$$

Where A is a NxN matrix encoding linear interactions between genes, with the implicit assumption that the positive and negative coefficients A correspond to activation and inhibition, and γ is a degradation rate vector. Interestingly, the existing bioinformatics methods for GRN inference can be broadly divided into two main classes based on whether steady state is assumed in eq. (42), with both approaches exhibiting strengths and pitfalls. First, steady state methods assume that $\dot{X} = 0$ [55,56]. Without further assumptions, however, this leads to the trivial, non-interacting solution $A = \gamma$. This problem is typically circumvented by eliminating self-interactions by setting the diagonal elements of A to zero. This way, the off-diagonal elements of A can be resolved, one row at a time, by solving the regression problem:

$$x_i = \sum_{j \neq i} a_{ij} x_j - \gamma_i x_i \quad (43)$$

Bocci and collaborators showed that this method can detect the presence of GRN edges (i.e., causal connections between genes), but might lead to inaccurate sign prediction [24]. Alternatively, eq. (42) can be interpreted as an out-of-equilibrium problem [57–59]. This approach, however, requires an estimation of the first derivative $\dot{X}$ from scRNA-seq data that typically does not have temporal information. Thus, physical time is typically substituted with pseudotime. Pseudotime (pst) is a cell ordering technique based on cell-cell gene expression similarity, so that, starting from a root cell (pst=0), a differentiation trajectory is reconstructed [60]. Cells with gene expression patterns that are similar to the root cell will have smaller pseudotime while cells with different gene expression patterns will have larger pseudotime. After substituting physical time with pseudotome in the RHS, eq. (42) can be solved for A. While pseudotime is a powerful tool to reveal the lineages of biological systems, there might be pitfalls when using it as a proxy for physical time. For example, in a scenario of multistability cells at different points of a differentiation trajectory coexist at the same physical time while exhibiting different pseudotime values. Interested readers can find more in-depth description of GRN inference methods and their application to biological datasets in topical reviews [53,54].

3. Examples from various biological contexts

In this section, we examine examples where the tools to model multistable regulatory networks have been successfully applied and unraveled new biological functions. First, we consider the epithelial-mesenchymal transition (or EMT in short), a trans-differentiation process that regulates cell motility in physiological processes and diseases. In our description of EMT, we further review mathematical approaches to describe the regulation of non-coding RNAs which play an essential part in the control of EMT. Second, we review theoretical and computational models of stem cell differentiation, along with theoretical approaches that rephrase multistable gene regulatory networks as a many-body problem. Finally, we examine models of cell-cell communications through the Notch signaling

pathway. These models introduce spatial patterning and multicellular communication that arise when cells can exchange information with their neighbors.

3.1 Intermediate states in the Epithelial-Mesenchymal Transition

3.1.1. Epithelial-Mesenchymal Transition in development and cancer

The Epithelial-Mesenchymal Transition (EMT) is a complex biochemical and biophysical process where epithelial cells, typically characterized by strong adhesion to neighboring cells and apicobasal polarity, loosen their adhesion and gain motility [62]. EMT and its reverse, MET, play a fundamental role in several developmental and physiological processes, such as organogenesis and wound healing, when cells need to transiently acquire motility to travel and rearrange their spatial organization. Moreover, EMT is implicated in several aspects of cancer progression, including metastasis and resistance to therapies [63,64]. A particularly interesting topic in EMT is the existence of intermediate cellular states separating the epithelial and mesenchymal phenotypes. These hybrid epithelial/mesenchymal cells (called E/M for short in the rest of the section) retain both cell-cell adhesion typical of epithelial cells and migration potential typical of mesenchymal cells. For this reason, these hybrid E/M states play an important role in collective cell migration in both physiological and pathological processes [65–67]. In recent years, theoretical modeling helped targeting several elusive questions about hybrid epithelial/mesenchymal states: is the EMT spectrum discretized in a set of intermediate states? Are these states truly stable or just metastable intermediates? Are they redundant or do they serve different biological functions? In this section we explore mathematical models that applied a wide array of methods to seek answers to these poignant questions.

3.1.2. Multistability predicted by continuous models of a core EMT regulatory network

Experiments have revealed intricated and interconnected network of genes, transcriptional factors and protein that regulate the epigenetic and biophysical transformations associated with EMT. Nonetheless, mathematical modeling has shown that some of the main aspects of EMT including the existence of intermediate cell states can be captured by a simple model of a core gene regulatory network including two micro-RNAs (miR-34 and miR-200)

and two transcription factors (Zeb and Snail). In this core circuit, miR-34 and miR-200 inhibit, and are in turn inhibited, by Zeb and Snail, respectively. Moreover, Zeb regulates its own activity, while Snail activates Zeb and self-inhibits (Figure 7A). Zeb and Snail are typically activated when a cell undergoes EMT and are therefore associated with a mesenchymal phenotype. Conversely, miR-34 and miR-200 inhibit the action of Zeb and Snail and are therefore associated with an epithelial phenotype. A continuous framework proposed by Zhang and collaborators [68] models all these coupled interactions with Hill functions. Under this approximation, both the miR-34/Snail and miR-200/Zeb motifs behave as bistable toggle switches. Therefore, EMT is achieved by first flipping the miR-34/Snail switch from (high miR-34, low Snail) to (low miR-34, high Snail), and subsequently flipping the miR-200/Zeb switch from (high miR-200, low Zeb) to (low miR-200, high Zeb). The intermediate state with (low miR-34, high miR-200, high Snail, low Zeb) does not exhibit clear epithelial or mesenchymal signatures and is therefore interpreted as a hybrid E/M transition state (Figure 7B).
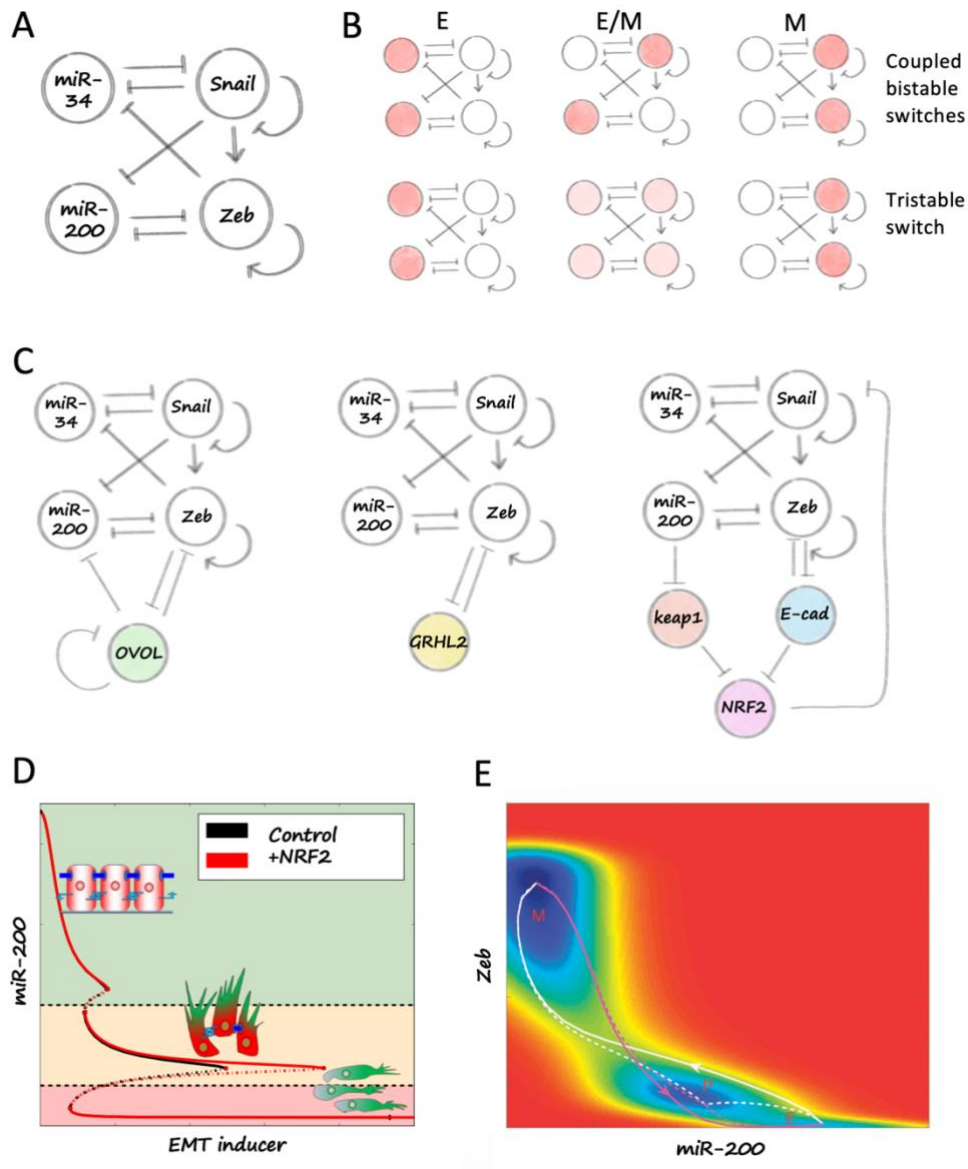
*Figure 7. **Continuous models of EMT. (A)** A core EMT gene regulatory network comprising the epithelial microRNAs miR-34, miR-200 and the mesenchymal transcription factors Zeb, Snail. **(B)** In the model by Zhang and collaborators [68], the transition state with (low miR-34, high miR-200, high Snail, low Zeb) is identified as the hybrid E/M. Conversely, in the model by Lu and collaborators [47], the hybrid E/M state has intermediate expression of all nodes. **(C)** Example of 'phenotypic stability factor' motifs that increase the stability of the hybrid E/M state by coupling to the core EMT circuit. **(D)** The stabilizing action of the PSF NRF2 is evaluated with a bifurcation diagram of miR-200 as a function of EMT inducer (adapted from Bocci and collaborators [69]); PSFs extend the hybrid E/M branch. **(E)** Pseudopotential landscape of the EMT circuit adapted from Li and collaborators [70]. E, P*

and M indicate the locations of the Epithelial, hybrid E/M (Partial) and Mesenchymal attractors. White and purple continuous lines indicate the minimum action paths (MAPs) of EMT and MET, while the dashed lines indicate the EMT and MET MAPs that pass through the hybrid E/M attractor.

3.1.3. Micro-RNA based regulation of EMT

As we previously discussed, Hill functions can effectively model transcriptional activation/inhibition where a transcription factor binds to DNA – either alone, as a dimer, or as an oligomer, to enhance or repress the transcription of its target gene (see Insight box 2). Micro-RNAs such as miR-34 and miR-200, however, inhibit their target at a post-translational stage by binding to the transcribed mRNA and facilitating its degradation. This reaction, which is always inhibitory, is therefore better described with an explicit model of microRNA-mRNA binding, rather than Hill functions [71,72]. Targeted mRNA molecules have a variable number (n) of available binding sites for the micro-RNAs. The higher the number of micro-RNA molecules bound to the mRNA, the larger the degradation rate of the complex. Therefore, micro-RNAs inhibit their target by preventing translation of the mRNA into protein. Lu and collaborators [47] developed a model of EMT regulation with similar circuit topology but different functional form to describe the action of miR-34 and miR-200 [47,73]. The micro-RNA based chimeric circuit approach models the chemical binding between micro-RNA and target mRNA molecules and the possible translation and degradation rates as a function of number of occupied binding sites. This mathematical model is discussed in detail in the insight box 3.

Similar to the model of Zhang and collaborators [68], Lu and collaborators [47] predicted the existence of a hybrid E/M state. Different modeling assumptions on micro-RNA regulation, however, are reflected into distinct features of the hybrid E/M state. In the model of Lu and collaborators, the miR-200/Zeb circuit acts as a tristable switch with an additional stable fixed point with intermediate expression of both miR-200 and Zeb (Figure 7B). The miR-34/Snail circuit, conversely, is monostable, and was proposed to act as a buffer that filter noise and confers robustness to the miR-200/Zeb switch [47].

This models has been further extended to investigate the coupling with various 'phenotypic stability factors' that modify the topology of the core EMT circuit and increase the stability of the hybrid E/M phenotype [69,74–76]. This approach shifts the focus from the underlying biology characterizing a specific EMT regulator, and successfully predict its qualitative impact on EMT based on how it is connected to the core EMT gene regulatory network (Figure 7C-D).

3.1.4. Landscape and path integral analysis highlights EMT transition routes

To gain further insights about the stability of the various EMT states and the transitions between them, Li and collaborators [70] studied the pseudopotential landscape of the EMT circuit of Lu and collaborators by explicitly solving the Fokker-Planck equation associated with the circuit. Specifically, the authors focused on the miR-200/Zeb switch that is responsible for the tristability of the circuit. This way, the pseudopotential can be easily visualized on a two-dimensional (miR-200, Zeb) space. The pseudopotential is obtained from the steady state probability $U = -\log \left(p_{SS}(miR-200, Zeb) \right)$ and is shown in Figure 7E. As expected, the landscape features three attractors corresponding to Epithelial, hybrid E/M, and Mesenchymal phenotypes. The authors further identified the minimum action paths (MAPs) connecting the three states. As discussed previously, the transitions are determined by the pseudopotential as well as the curl flux arising from broken detailed balance. Due to the curl flux, the MAP connecting the E and M attractors does not pass through the saddle points of the landscape, thus deviating from the behavior of 'classical', gradient-driven physical systems. There are several interesting biological implications stemming from this observation. First, cells undergoing E-to-M transition do not necessarily pass through the hybrid E/M phenotype, because the action of the E-E/M-M path is higher than the action of the direct E-M path. Moreover, the paths for the E-M transition and its reverse, M-E, are different, hence making them irreversible (Figure 7E). This suggests that cells undergoing EMT and MET might exhibit different molecular signatures. Finally, the action required for the E-E/M-M transition is higher than the action of the irreversible E-M transition; the action required for the first step (E-E/M transition), however, is lower than that of the E-M transition. Therefore, the E-E/M transition path could be chosen when the cell is not exposed to enough stimulus to undergo a complete E-M transition.

3.1.5. Boolean networks identify biological regulators and intermediate states in EMT

While small circuits of a core regulatory network enable a detailed analysis of EMT, several molecular players and signaling pathways participate in the regulation of EMT. Steinway and collaborators [77] reconstructed a large EMT Boolean network with 70 nodes and 135 connections by integrating experimental observations on known EMT factors and regulators from human hepatocellular carcinoma (HHC). To simulate the dynamics of EMT driven by TGF-β – a well-known EMT inducer – the authors initialized the network in an epithelial state and updated the model following a stochastic asynchronous scheme. To decrease the complexity of the network, they also applied reduction methods to eliminate 'redundant' nodes and connections. Interestingly, they identified a reduced network with 19 nodes and 70 connections with a very similar dynamical behavior (Fig. 8A). This reduced network enabled the identification of transition trajectories from the initial epithelial state to the final mesenchymal state, thus offering interesting information on the temporal ordering in the activation and deactivation of relevant genes. Furthermore, the authors explored the effect of knockout and/or constitutive activation of nodes or combinations of nodes by artificially enforcing an 'ON' or 'OFF' state for certain nodes in the circuit [78]. Strikingly, some of these perturbations prevent EMT but do not reverse the transition to the original epithelial attractor; rather, alternative attractors with both epithelial and mesenchymal active genes become stable, which are identified as candidates for hybrid epithelial/mesenchymal states.

Font-clos and collaborators [79] described the same EMT network with a pseudo Hamiltonian based on the formalism of spin systems $H = - \sum_{i,j} J_{ij} s_i s_j$. Compared to the Hamiltonian typically used to describe spin systems, here the $s_i = 0, 1$ describe an inactive or active gene, respectively, and the J_{ij} describe inhibition ($J_{ij} = -1$), activation ($J_{ij} = 1$) or lack of interaction ($J_{ij} = 0$) from gene i to gene j. Thus, differently from standard spin system Hamiltonians, interactions are not symmetric (i.e., $J_{ij} \neq J_{ji}$) and therefore it is not guaranteed that the fixed points of the system correspond to the minima of H. The authors sampled the landscape of the network with a large number of stochastic simulations using the same update scheme originally used by Steinway and collaborators [77,78]. This analysis revealed two deep minima associated with the epithelial and mesenchymal states separated

by a complex landscape with several local minima that are interpreted as intermediate, less stable states. On a two-dimensional PCA map, the distribution of steady states shows two dense regions corresponding to Epithelial (E-cadherin node is ON) and mesenchymal (E-cadherin node is OFF), separated by a sparser region where E-cadherin can be ON or OFF (Fig. 8B-C).
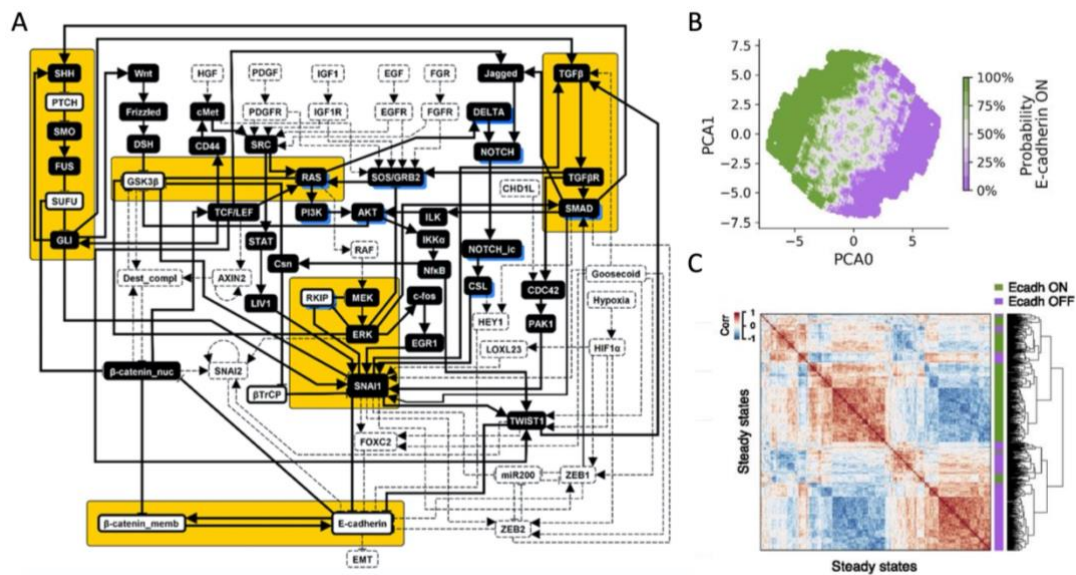
*Figure 8. **Boolean models of EMT.** (A) A large EMT network used by Steinway and collaborators [78]. Yellow areas highlight the motifs that are found to stabilize the Epithelial state. White-colored nodes are active in the Epithelial state while black-colored nodes are inactive. Nodes with blue background are predicted as sensible knockout targets that suppress EMT. (B) Density of steady states on the PCA0-PCA1 plane in the model of Font-Clos and collaborators [79]. Green and purple color shading indicate high or low probability that E-cadherin is on in the steady state. (C) Clustering and correlation of 500 steady state solutions in the model of Font-Clos and collaborators.*

3.1.6 Insight box 3: Describe micro-RNA mediated inhibition with mass action and chimeric circuits

Micro-RNAs (miRs) inhibit the expression of a protein by binding to the target mRNA molecules and degrade them before translation. Several models that describe miR-mediated inhibition employ the principles of mass action [71,72]. Suppose that mRNA of a given protein (m) is transcribed with rate k_m and the corresponding molecule (p) is translated with rate k_p. Moreover, we consider a miR species (μ) that is produced and degraded with rate constants k_μ, γ_μ, respectively. μ binds and unbinds to m with rate constants k_B and k_U, respectively, and the mRNA-miR complex $[m\mu]$ degrades with rate constant $\gamma_{m\mu}$. The coupled dynamics of mRNA, miR, mRNA-miR complex and protein ($m, \mu, [m\mu], p$) is described by the following system of ODEs:

$$\frac{dm}{dt} = k_m - k_B m\,\mu + k_U[m\mu] - \gamma_m m \quad (44a)$$

$$\frac{d\mu}{dt} = k_\mu - \gamma_\mu \mu \quad (44b)$$

$$\frac{d[m\mu]}{dt} = k_B m\,\mu - k_U[m\mu] - \gamma_{m\mu}[m\mu] \quad (44c)$$

$$\frac{dp}{dt} = k_p m - \gamma_p p \quad (44d)$$

Various approximations can be made from this basic framework, for instance by assuming fast mRNA-miR binding and writing the loss term of m as a function that only depends on m and μ. Moreover, as discussed in the insight box 1, fast mRNA dynamics can be another valid approach. More details on this are provided in other reviews [71,72].

The general scheme of eqs. (44), however, assumes a 1-to-1 stoichiometry in the mRNA-miR binding. Often, mRNA molecules have multiple binding sites, whose occupancy modulates translation and degradation rates (Figure 9). In the following, we discuss a generalization of eqs. (44) developed by Lu and collaborators that accounts for multiple miR binding sites that has been successfully applied to the gene circuit that regulates the Epithelial-Mesenchymal Transition [47,73].

Similar to eqs. (44), let us consider a system with an mRNA species (m) and an inhibiting miR (μ). Moreover, let us assume that each mRNA molecule has n binding sites for miR molecules. miR molecules bind and unbind to the mRNA with rate constants $r_{\mu+}$ and $r_{\mu-}$, respectively. In the model, binding at different sites are treated as independent events due to the small size of the miR compared to the mRNA molecule. Assuming fast equilibration of mRNA-miR binding, the fraction of mRNA molecules with i occupied binding sites satisfies:

$$r_{\mu+}\, \mu\, [m_i] = r_{\mu-}\, [m_{i+1}] \quad (45)$$

Iteratively, $[m_i]$ can be expressed as a function of $[m_0]$, i.e. the fraction of mRNAs without any occupied binding site:

$$[m_i] = \left(\frac{\mu}{\mu_0}\right)^i [m_0] \quad (46)$$

Where $\mu_0 = r_{\mu-}/r_{\mu+}$. The set of concentrations $[m_i]$, $i = 0, 1, \dots n$, must also satisfy:

$$\sum_{i=0}^{n} C_n{}^i\, [m_i] = m \quad (47)$$

Where $C_n{}^i = n!/i!\,(n-1)!$ is the number of arrangements for i micro-RNA molecules on n binding sites. Plugging eq. (46) into eq. (47) yields:

$$m = [m_0] \sum_{i=0}^{n} C_n{}^i \left(\frac{\mu}{\mu_0}\right)^i = [m_0] \left(1 + \frac{\mu}{\mu_0}\right)^n \quad (48)$$

Thus, using eq. (46) and (48), the fraction of mRNA molecules with i miR bound ($[m_i]$) can be expressed as:

$$[m_i] = m \frac{\left(\frac{\mu}{\mu_0}\right)^i}{\left(1 + \frac{\mu}{\mu_0}\right)^n} \quad (49)$$

Finally, we introduce sets of rate constants $[l_i]$, $[\gamma_{mi}]$ and $[\gamma_{\mu i}]$ that represent the translation rate of a mRNA molecule with i miR bound, the degradation rate of a mRNA molecule with i miR bound, and the degradation rate of miR molecules associated with i-bound mRNAs. Therefore, it is possible to write total translation and degradation terms:

$$\frac{dm}{dt} = k_m - m \sum_{i=0}^{n} \gamma_{mi}\, C_n{}^i\, M_n{}^i(\mu) - \gamma_m m \quad (50a)$$

$$\frac{d\mu}{dt} = k_\mu - m \sum_{i=0}^{n} \gamma_{\mu i}\, C_n{}^i\, M_n{}^i(\mu) - \gamma_\mu \mu \quad (50b)$$

$$\frac{dp}{dt} = k_p m \sum_{i=0}^{n} l_i \, C_n{}^i \, M_n{}^i(\mu) - \gamma_p p \quad (50c)$$

Where $M_n{}^i(\mu) = \left(\frac{\mu}{\mu_0}\right)^i \Big/ \left(1 + \frac{\mu}{\mu_0}\right)^n$.

Compared to modeling inhibition with Hill functions or simpler mass-action dynamics, the chimeric circuit approach requires knowledge of more parameters which might or might not be readily accessible from experiments. In the case of the EMT circuit proposed by Lu and collaborators, the expression levels of translated protein with various number of occupied mRNA binding sites were used to calibrate the model's parameters. Further information about this procedure can be found in the original publication [47].
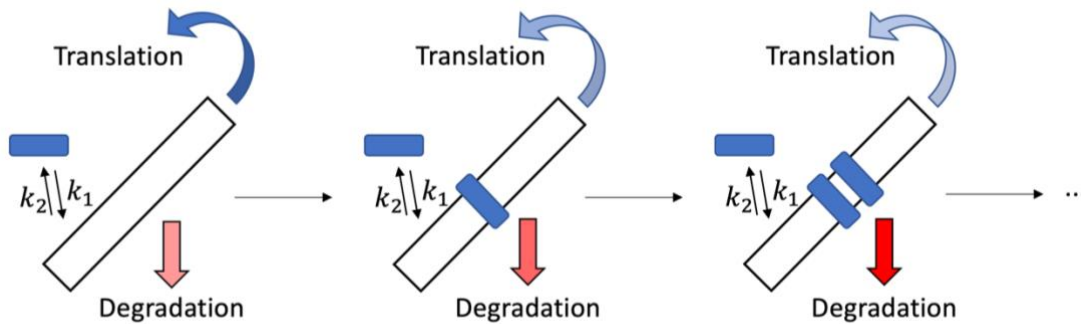
Figure 9. ***Schematic representation of microRNA-based regulation in the chimeric circuit model.*** *A mRNA molecule (large white-filled rectangle) can bind to multiple microRNA molecules (small blue rectangles). mRNA-microRNA binding does not depend on the number of already occupied binding sites (same binding and unbinding rate constants in the different pictures). The more microRNA molecules bound, the higher the complex degradation rate (indicated by brighter red arrow from leftmost to rightmost panel) and the lower the mRNA translation rate (indicated by lighter blue arrow from leftmost to rightmost panel).*

3.2 Stem cell differentiation

It has been a continuing research interest to understand how stem cells accurately specify cell fates during differentiation, referred to as stemness regulation. Mathematical modeling approaches have been applied extensively to elucidate the gene regulatory mechanisms underlying stem cell differentiation.

3.2.1 A toggle switch topology regulates the differentiation of stem cells

Mathematical models focusing on the acquisition of stable states representing cell phenotypes are often derived using chemical rate equations. Similar to the approaches discussed in EMT that focus on small circuits, the modeling of stem cell differentiation are often applied to the proposed "core" stemness regulatory networks that contain master stemness regulators.

Huang and co-workers developed a continuous model focusing on the binary decision-making during the lineage commitment of erythroid and myelomonocytic fates[80]. The model is derived to represent the temporal dynamics of a core circuit containing the two lineage specifying transcription factors (TFs) GATA1 and PU.1. GATA1 and PU.1 exhibit mutual inhibition and auto-stimulation, which are modeled by Hill functions (Fig. 10A). Modeling analysis shows that the circuit can generate stable attractors corresponding to erythroid (GATA1 high, PU.1 low) and myelomonocytic (GATA1 low, PU.1 high) phenotypes, and a metastable state representing the "multilineage priming" stage, characterized by coexpression of both GATA1 and PU.1. This study describes a binary cell fate decision-making for a bipotent progenitor cell.

Another kinetic model of stemness regulation has been developed by Jolly and co-workers focusing on the LIN28/let-7 circuit[81]. LIN28 (an RNA-binding protein) and let-7 (a microRNA) are mutually inhibitory and self-excitatory (Fig. 10B). Two ordinary differential equations have been derived to represent the birth and death process of LIN28 and let-7. The shifted Hill functions are used to represent the let-7/LIN28 mediated inhibition and self-activation. The modeling analysis suggests that the LIN28/let-7 circuit can operate as a

three-way switch enabling three stable states characterized by (high LIN28, low let-7), (high let-7, low LIN28) and (intermediate LIN28, intermediate let-7), respectively. As the experimental studies suggest that intermediate levels of OCT4 account for pluripotency and OCT4 is a downstream target of LIN28, this model proposes the stable state characterized by (intermediate LIN28, intermediate let-7) associates with pluripotency. From a theoretical perspective, this model highlights the effect of additional regulatory interactions on the "standard" toggle switch topology. While mutual inhibition typically leads to bistability, the additional self-activation of both LIN28 and let-7 introduces a third state with intermediate expression of both factors.

To elucidate the transcriptional dynamics of the master stemness regulatory TFs - OCT4, SOX2 and Nanog during stem cell differentiation, Chickarmane and co-workers developed a kinetic model to study the stemness circuit containing three master stemness regulatory TFs - OCT4, SOX2 and Nanog[82] (Fig. 10C). These three TFs play a critical role in positively regulating the stemness genes and negatively regulating the differentiation genes. The modeling analysis in this study shows that the OCT4/SOX2/Nanog circuit can function as a binary switch, and enable two stable states that are characterized by (high OCT4-SOX2, high Nanog) and (low OCT4-SOX2, low Nanog), respectively. Here OCT4-SOX2 is a heterodimer formed by OCT4 and SOX2 and can regulate OCT4, SOX2 and Nanog individually. The stable state characterized by (high OCT4-SOX2, high Nanog) exhibits high expression of stemness genes and low expression of differentiation genes thus corresponding to stem cell phenotype. Conversely, the stable state characterized by (low OCT4-SOX2, low Nanog) corresponds to cells undergoing differentiation. Through bifurcation analysis, the study shows how stem cells - (high OCT4-SOX2, high Nanog) can maintain their stemness state upon removal of the stimulus signals that upregulate OCT4 and SOX2. Later Chickarmane and Peterson [83] extended their model by including additional stemness TFs (Cdx2, Gcnf, Gata-6) to elucidate the Trophectoderm and Endoderm lineage commitment. By performing perturbation analysis on the circuit, the study suggests strategies to reprogram cells back to stemness state, such as activation of Nanog and suppression of Gata6.
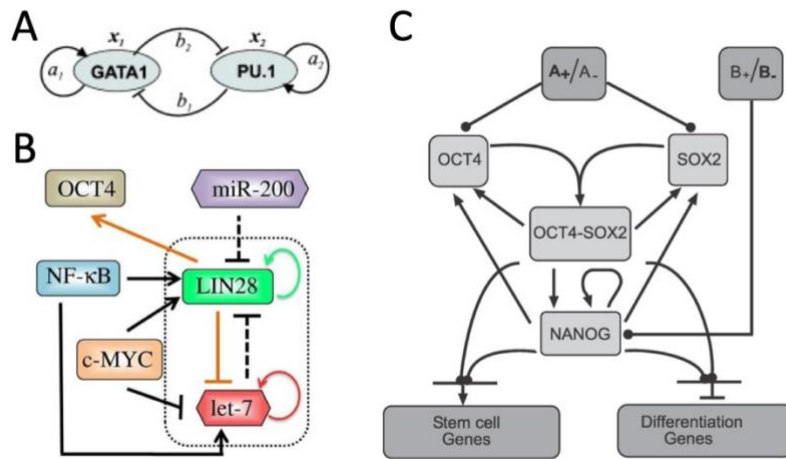
Figure 10. **Core regulatory circuits to model stem cell differentiation.** *(A) Huang and coworkers [80] modeled stem cell differentiation via mutual inhibition between GATA1 and PU.1. (B) The stemness circuit constructed by Jolly and collaborators [81]. to model the acquisition of cancer stem cell traits in cancer. (C) The circuit developed by Chickarmane and co-workers [82] focuses in the stemness regulators OCT4, SOX2 and Nanog.*

3.2.2 Role of stochastic fluctuations in stem cell differentiation

While the deterministic models discussed in the previous section successfully capture the stem cells decision-making dynamics, it remains unclear how this differentiation is regulated in the fluctuating cellular environment. To capture the stochastic effects, Kalmar and co-workers developed a continuous model based on the Oct4/Nanog circuit[84]. The dynamics of the Oct4/Nanog circuit are represented by two coupled differential equations. The transcription noise (represented by Gaussian white noise) is incorporated into the differential equation representing the temporal dynamics of Nanog. This study sheds light on how cells can vary their Nanog level without losing their pluripotency. The study suggests an interesting view of pluripotency. The pluripotency, instead of being viewed as a stable state characterized by fixed amount of gene expression levels, can be viewed as a heterogenous population driven by transcriptional noise such as the noise during Nanog transcription.

To continue studying the mechanisms underlying the heterogenous levels of Nanog in mouse embryonic stem cells, Glauche and co-works developed an ODE model to simulate the Oct2/Sox2/Nanog circuit[85]. They showed that the experimental characteristics of Nanog variation can be recapitulated by the model via either adding a Gaussian white noise term in the ODE representing Nanog (noise-induced transition between Nanog low and Nanog high states) or including a transcriptional repressor of Nanog (oscillation). This study also shows how Nanog low cells can be more prone to differentiation relative to Nanog high cells. Another modeling study to elucidate the variation of Nanog levels in embryonic cells has been performed by Herberg and co-workers by integrating a negative feedback loop FGF4/Erk signaling with the Oct4/Sox2/Nanog/Rex1 circuit. The effect of the transcriptional noise, defined by the zero-mean Gaussian process, has been incorporated into the ODEs representing the temporal dynamics of Oct4-Sox2, Nanog and Rex1. The model shows a bimodal distribution of Nanog levels and blocking the Erk signaling pathway can lead to a merged Nanoga distribution[86].

3.2.3 Chemical master equation models characterize the role of Nanog during differentiation

The heterogeneous expression of Nanog in pluripotent cells has also been studied by applying master equations to describe the stochastic dynamics of a stemness network containing eight stemness TFs (Oct4, Sox2, Nanog, GATA6, Gcnf, Cdx2, Pbx1, Klf4) and one heterodimer (Oct4-Sox2) formed by Oct4 and Sox2[87]. The master equations for the stemness network were derived by considering both the concentration of proteins and the occupancy of the DNA sites. This study shows how the stemness network can enable up to five steady states, corresponding to two stem cell phenotypes expressing (high Oct4, high Sox2 and high Klf4), one primitive endoderm phenotype, one trophectoderm phenotype and one differentiated cell type. The stem cell steady states are differentiated by distinct levels of Nanog, which is consistent with its experimental characteristics. Through analyzing the most probable transition paths among the steady states, the study suggests a sequential transition path that *Nanog high* stem cell steady state first decreases the Nanog levels to enter a novel, *Nanog low* stem cell steady state, and then transition to the primitive endoderm steady state. The modeling results support the critical role of Nanog in safeguarding stemness against differentiation.

Furthermore, to overcome the computational limitations of simulating large networks, Sasai and Wolynes described the stochastic gene expression and feedback interactions between transcription factors using the many-body approach typically employed in quantum statistical mechanics [17]. In the Insight box 4, we introduce this approach and show how the chemical master equation of an autoregulating transcription factor can be effectively mapped with this formalism. By extending this approach to an 8-gene circuit of embryonic stem cell development, they characterized the multistable landscape and the most probable transitions between attractors [87].

3.2.4 Landscape theory identifies the trajectories of differentiation and reprogramming

To quantify the kinetic flow during embryonic stem cell differentiation, epigenetic landscape approach has been applied. Li and Wang developed a global potential landscape for a stem cell development network containing fifty-two genes[88]. There are two basins of attractions identified on the landscape corresponding to a stem cell state and a differentiated state. The model shows the development process can be viewed as a transition from the stem cell attractor to the differentiated attractor. Interestingly, the reversed "reprograming" trajectory where differentiated cells acquire pluripotent traits can be different in the 52-gene expression space. Later, the authors applied the landscape and flux path approach to quantify the landscape for a reduced cancer stem cell (CSC) circuit that only contains six key genes (two microRNAs - miR200 and miR-145, four proteins - ZEB, P53, MDM2 and OCT4), to identify the basins of attractions and transition paths in between[89]. This study shows the CSC circuit enables four basins of attractions corresponding to a normal state, a normal stem cell state, a cancer state, a cancer stem cell state. Specifically, the study suggests that P53 activation promotes the transition from the cancer stem cell attractor to the normal cell attractor, which is consistent with the tumor suppressor role of P53.

To understand the large cell-to-cell variation of Nanog and other stemness TFs in embryonic stem cells, Sasai and co-workers developed an epigenetic landscape approach for a stemness regulatory network contains six genes - three stemness genes (Sox2, Oct4, Nanog), and three lineage-specific genes (Gata6, Cdx2, and Gcnf) [90]. One interesting prediction

from this study is the distribution of time scales of the regulatory processes (binding/unbinding of TFs, histone modification etc.) is critical to determine the dynamic behaviors of the network. Specifically, the slow transcriptional switching of Nanog leads to large fluctuations among multiple transient states.

3.2.5 Parameter randomization on the stemness circuit

To identify the robust gene states during stem cell differentiation, parametric randomization methods have been applied to the stemness network. For example, RACIPE has been applied to a proposed core stemness regulatory network containing eight master regulatory TFs (Oct4, Sox2, Cdx2, Gata6, Gcnf, Pbx1, Klf4 and Nanog) [91]. The stemness regulatory network contains mostly transcriptional regulation with protein-protein interaction, i.e., the binding/unbinding interactions between SoX2, Oct4 and the Oct4-Sox2 complex. Using exclusively the topological information of the network as input, RACIPE generated an ensemble of mathematical models with various numbers of steady-state solutions that can be explored with statistical analysis. Hierarchical clustering results of the RACIPE solutions exhibit distinct gene expression patterns that can be associated with different development stages and match the single-cell mRNA expression data of mouse embryonic stem cells (ESCs). One intriguing result of this study is the multi-stable behaviors of the stemness GRN and the recapitulation of the experimentally measured gene expression patterns cannot be achieved by networks with randomized topologies, even though the randomized networks can have similar amount of mutual inhibitory and self-excitatory feedback loops relative to the stemness circuit. The results suggest that the stemness GRN may indeed contain some topological properties that are beyond the expectation based on the counting of the simple motifs. By analyzing the difference of the parameters that lead to different states, the key parameters whose variation can induce the transitions among different states are identified. The physiological representation of those parameters can be the stimuli that trigger the transition. Through systematically perturbing each gene and each regulatory link in the stemness GRN, RACIPE elucidates a hierarchical decision-making structure of the stemness GRN, with the OCT4/CDX2 motif functioning as the first decision-making module followed by the GATA6/Nanog module. By simulating the effect of external signals that perturb the TFs in the stemness GRN, RACIPE analysis results suggest that the presence of external signals often restrains the gene states that can be accessed instead of creating new states. The

results suggest an alternative explanation of the Waddington landscape that the stem cell population, instead of being viewed as a fixed stable "stemness" state, can be regarded as a mixture of heterogeneous cell phenotypes. Along the differentiation, upon external signals, the cell population heterogeneity decreases partially because cells lose access to some stable states. In other words, cells are committed to specific phenotypes, therefore, differentiated.

3.2.6 Combing Landscape theory with random parametric perturbation

To calculate the landscape of stem cell networks by considering the large variation of parameters, Li developed an approach that combines landscape theory with random parameter perturbation, referred to as LRPP[92]. The LRPP approach contains the following steps: the time evolution of each gene in the stemness network is represented by ODEs; the modeling parameters are sampled from a range with either uniform or Gaussian distribution; multiple initial conditions are chosen to identify all possible steady states; the landscape is calculated using the self-consistent mean field approximation; repeat the aforementioned procedure for multiple times and Sum over the landscape of different sets of parameters to get the landscape for the ensemble models. By applying the LRPP approach to a stemness GRN that contain four stem cell marker genes OCT4, SOX2, NANOG, and KLF4 and three differentiation marker genes GCNF, CDX2, and GATA6, the author identifies three main stable basins of attractions corresponding to an embryonic stem cell state (high Nanog, low Gata6), a differentiate cell state (low Nanog, high Gata6) and an intermediate state (high Nanog, high Gata6). The intermediate state characterized by co-expression of stemness marker Nanog and differentiation marker Gata6 may account for the heterogeneity of these TFs observed during single-cell experiments, and may play a critical role in regulating stem cell plasticity.

3.2.7 Insight box 4: spin-boson formalism for an autoregulatory gene

In this insight box, we discuss the application of spin-boson formalism to describe stochastic gene switches. Here, we only aim at giving the reader a flavor of this approach. Following

Sasai and Wolynes, we examine a single autoregulatory gene (i.e., a gene that regulates its own transcription). An extensive review on the uses of quantum field theory in diffusion-reaction systems is offered by Mattis and Glasser[93].

The autoregulatory gene is described by the protein copy number n and a Boolean variable $s = 0,1$ describing whether the gene is bound or unbound.

We have previously developed the CME for a birth-death process (see insight box 1). In addition to production and degradation, produced molecules can bind to DNA and regulate their own transcription. Depending on whether molecules bind to DNA as monomers or oligomers, the protein-DNA binding rate depends on a power of protein level n (see insight box 2). Since we are not interested in solving a specific case but rather layout the mapping procedure, we assume a generic binding rate $h(n)$. The DNA-protein unbinding rate is a constant f independent from n. The unbound and bound promoter produces molecules at rates g_0 and g_1, respectively. Similar to the birth-death process, we are condensing mRNA transcription and translation into a single production rate. Finally, protein molecules degrade with rate constant k. A set of two probabilities $\left(P_1(n,t), P_0(n,t)\right)$ describes the probabilities to have n molecules and an unbound/bound promoter at time t, respectively. Combining all reactions, CMEs for P_0 and P_1 assume the form:

$$\frac{dP_0(n)}{dt} = g_0 P(n-1,0) - g_0 P(n,0) + k(n+1)P(n+1,0) - knP(n,0)$$

$$+fP(n,1) - h(n)P(n,0) \quad (51a)$$

$$\frac{dP_1(n)}{dt} = g_1 P(n-1,1) - g_0 P(n,1) + k(n+1)P(n+1,1) - knP(n,1)$$

$$-fP(n,1) + h(n)P(n,0) \quad (51b)$$

To map this CME onto a many-body problem, the first step is to rearrange eqs. (51) as a single equation for the probability vector $P(n,t) = \left(P_1(n,t), P_0(n,t)\right)$:

$$\frac{\partial}{\partial t}P(n,t) = \begin{pmatrix} g_1 & 0 \\ 0 & g_0 \end{pmatrix}[P(n-1,t) - P(n,t)] + k[(n+1)P(n+1,t) - nP(n,t)]$$

$$+ \begin{pmatrix} -h(n) & f \\ h(n) & -f \end{pmatrix}P(n,t) \quad (52)$$

Further, we define a state vector ψ:

$$\psi = \sum_{n=0}^{+\infty} P(n,t)\,|n\rangle \quad (53)$$

To make the connection to the quantum many-body problem more evident, we introduce ladder operators[93]:

$$a^*|n\rangle = |n+1\rangle \quad (54a)$$

$$a|n\rangle = n|n-1\rangle \quad (54b)$$

Notably, the ladder operators still satisfy the commutation relation $[a, a^*] = 1$, but their definition slightly differ from the standard harmonic oscillator operators ($a^*|n\rangle = \sqrt{n+1}|n\rangle$, $a|n\rangle = \sqrt{n}|n-1\rangle$). This unconventional definition arises because the coefficients in the expansion of eq. (53) are probabilities, whereas typically probabilities correspond to the squared coefficients in 'standard' expansions of the wave function.

With these definitions in hand, the CME can be rewritten in the form:

$$\frac{\partial \psi}{\partial t} = \Omega \psi \quad (55)$$

Where all the information about chemical reactions is encoded by the operator Ω:

$$\Omega = (\bar{g} + \delta g\,\sigma_Z)(a^+ - 1) + k(a - a^+ a) + \mu^+(-1 + \sigma_X) + \mu^-(-i\sigma_Y - \sigma_Z) \quad (56)$$

With the following definitions:

$$\bar{g} = \frac{(g_1 + g_0)}{2} \quad (57a)$$

$$\overline{\delta g} = \frac{(g_1 - g_0)}{2} \quad (57b)$$

$$\mu^+ = \frac{(h(aa^+) + f)}{2} \quad (57c)$$

$$\mu^- = \frac{(h(aa^+) - f)}{2} \quad (57d)$$

And the introduction of the matrices:

$$\sigma_X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (58a)$$

$$i\sigma_Y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (58b)$$

$$\sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (58c)$$

Sasai and Wolynes further apply this formalism to a more complex 8-gene circuit of stemness transcription factors to predict stable attractors and transitions corresponding to differentiation. More details about this work can be found in reference [17].

3.3 From multistability to spatial patterning: cell communication through Notch signaling

So far, we have considered intracellular signaling networks that regulate cell-fate dynamics. In all these cases, the decision on cell fate was cell-autonomous, i.e., it did not depend on the signaling state of other cells in the surrounding environment. A cell population where cells can assume one of multiple states will exhibit heterogeneity according to the states' relative stability but without any spatial organization. If cells exchange information with their neighbors, however, the cell-fate decision processes become correlated, thus giving rise to spatial organization. In this section, we consider the example of Notch signaling, one of the most well-conserved and studied signaling pathways that regulate both physiological and pathological processes [94,95]. Notch signaling operates via binding of ligands and receptors belonging to neighboring cells, thus serving as a nearest neighbor communication mechanism that couples cell-fate decisions in a spatially-dependent manner. Signaling through different classes of Notch ligands can lead to either converging or diverging cell states between neighbors, thus raising interesting parallels with the behavior of spin systems such as ferromagnets and antiferromagnets.

3.3.1 Notch signaling relays single cell multistability to spatial patterning

Notch signaling is initiated when the extracellular domain of the Notch receptor binds to the transmembrane domain of a ligand at the surface of a neighboring cell. Mammalian species typically exhibit one class of receptors (Notch) and two classes of ligands (Delta and Jagged),

which can be further divided into a variable number of subtypes with different molecular structures. Upon binding, pulling forces by endocytosis and sequential cleavage actions by assisting enzymes lead to the release of the Notch intracellular domain (NICD). The NICD is transported to the cell nucleus where it activates or inhibits the transcription of several target genes [94,96,97]. In particular, NICD transcriptionally activates Notch and Jagged while inhibiting Delta. Therefore, a cell with high expression of Delta ligands activates the Notch receptors in its neighbors, thus in turn implying the repression of Delta. Conversely, when a cell is exposed to low levels of Delta from its neighbors, the Notch receptor is not activated, thus allowing production of Delta. Hence, Notch-Delta signaling leads neighboring cells to divergent cell states: a (low Notch, high Delta) state typically referred to as Sender, and a (high Notch, low Delta) state typically referred to as Receiver. At the multicellular level, this patterning principle leads to alternation of Senders and Receivers, typically referred as "lateral inhibition", which plays a crucial role in the differentiation of cell states in several physiological processes including somitogenesis, angiogenesis and neurogenesis [98–100]. Conversely, a cell with high expression of Jagged activates Notch receptors in its neighbors, which in turn activates the production of both Notch and Jagged. Therefore, Notch-Jagged signaling promotes a convergent (high Notch, high Jagged) state among neighbors that is often referred to as hybrid Sender/Receiver. On a multicellular level, this patterning principle leads to a homogeneous population of hybrid Sender/Receiver cells, or "lateral induction", which is observed, for instance, in the spatial propagation of a pluripotent cell state during inner ear development [98,99,101]. In the following paragraphs, we will review models that investigate how intracellular signaling and ligand-receptor binding relay cell-fate decisions in individual cells and multicellular patterning, the competition between lateral inhibition and lateral induction, and the role of stochastic fluctuations in enforcing or disrupting ordered patterns.

3.3.2 Notch-Delta lateral inhibition: a two-cell toggle switch

The first mathematical model of Notch signaling proposed by Collier and collaborators [102] directly generalizes the single cell toggle switch and focuses on lateral inhibition driven by Notch-Delta signaling between neighboring cells (Figure 11A). It considers a two-

dimensional lattice where the temporal dynamics of Notch and Delta in a cell (p) is described by the set of ODEs:

$$\frac{d(N_p/N_0)}{dt} = \frac{(\overline{D_p}/D_0)^k}{a + (\overline{D_p}/D_0)^k} - \mu(N_p/N_0) \quad (59a)$$

$$\frac{d(D_p/D_0)}{dt} = \frac{1}{1 + b(N_p/N_0)^h} - \rho(D_p/D_0) \quad (59b)$$

Where $\overline{D_p} = \frac{1}{c}\sum_{p'} D_{p'}$ is the average level of Delta ligand in the nearest neighbors; in this expression c is the number of nearest neighbors and the summation goes over all nearest neighbor cells p'. N_0 and D_0 are typical levels of Notch and Delta used to scale the model while μ, ρ are dimensionless degradation rate constants. Compared to all the circuits reviewed so far, this model considers a multicellular lattice (typically two-dimensional) where each individual cell is described by the Notch-Delta circuit, and circuits of neighboring cells are connected through the nearest neighbor summations in the Hill functions of eq. (59a).

Therefore, if neighbors of cell p are Senders with high Delta, cell p represses the production of Delta while increasing the production of Notch, hence assuming a (high Notch, low Delta), Receiver state. Conversely, if neighbors of cell p are Receivers with low Delta, Notch is weakly activated in cell p, thus maintaining a Sender state with (low Notch, high Delta) [100,102]. Therefore, if two neighboring cells start with very similar but not exactly equal levels of Notch and Delta at time $t = 0$, the mutual inhibition mechanism will amplify the small initial difference and ultimately differentiate between a Sender cell and a Receiver cell.

On an extended two-dimensional lattice, this model allows patterns where Sender cells are surrounded by Receiver cells. Specifically, on a square lattice, these interactions lead to a chessboard-like pattern with alternating Senders and Receivers [103]. On a hexagonal lattice that perhaps better represents the arrangement of cells in an epithelial tissue, however, a perfect alternation of Senders and Receivers cannot be achieved due to lattice frustration. Therefore, Sender cells are typically surrounded by six Receivers that are in contact with one another, leading to a 3-to-1 Receiver/Sender ratio (figure 11B). Indeed, Receiver cells can be

in contact as they do not express Delta, and therefore do not actively regulate each other, while Sender-Sender contacts give rise to mutual inhibition until one cell is converted to the Receiver state [104]. In other words, a contact between Receivers simply results in lack of cell-cell signaling, whereas a contact between Senders leads to the mutual inhibition that ultimately breaks the symmetry and forces one cell to the Receiver state. This mechanism is not flawless and sometimes patterning mistakes can be observed; in the last section of the chapter, we will discuss in more details the nature of these "defects", their biological implications and the role of stochastic fluctuations in modulating these patterns.
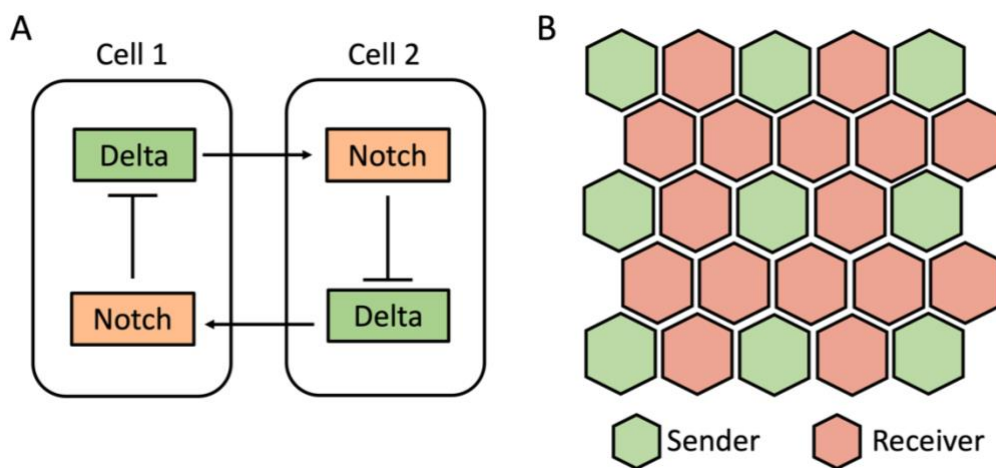

*Figure 11. **Lateral inhibition and patterning of Senders and Receivers in the Notch-Delta system. (A)** Schematic of the model proposed by Collier and collaborators in a system of two cells [102]. Pointing arrows indicate Delta-mediated activation of Notch in neighboring cells, while t-shaped arrows indicate Notch-mediated inhibition of Delta within the same cell. (B) On a hexagonal lattice, Notch-Delta lateral inhibition gives rise to a pattern where high-Delta Senders (green hexagons) are surrounded by high-Notch Receivers (red hexagons).*

3.3.3 Variable cell shape as an imprint to guide lateral inhibition

Notch receptors and ligands become available for signaling only once they reach the cell membrane. Therefore, the signaling between pairs of neighbors depends on the geometry

of cell-cell contact and their shared membrane area. From a theoretical perspective, this observation offers the possibility to integrate aspects of cell biophysics into the 'standard' Notch-Delta signaling.

Shaya and collaborators [105] investigated the relation between cell size and cell fate by integrating experimental and computational methods. By introducing reports to track the activity of Notch and Delta, they showed that signaling between pairs of nearest neighbors correlates with their shared contact area. Including the role of cell shape requires a more detailed model that explicitly describes ligand-receptor binding as well as transcriptional regulation by the NICD thereafter [100,105–107]. This can be achieved by directly generalizing the model of eqs. (1a-b) to explicitly include variable cell shapes. In this generalized model, a cell is described by a set of three ordinary differential equations for Notch, Delta and NICD:

$$\frac{dN_p}{dt} = g_N \frac{1 + \lambda_N \left(\frac{I_p}{I_0}\right)^{n_N}}{1 + \left(\frac{I_p}{I_0}\right)^{n_N}} - kN_p \sum_{j \in N(p)} I_{pj} D_j - \mu N_p \quad (60a)$$

$$\frac{dD_p}{dt} = g_D \frac{1 + \lambda_D \left(\frac{I_p}{I_0}\right)^{n_D}}{1 + \left(\frac{I_i}{I_0}\right)^{n_D}} - kD_p \sum_{j \in N(p)} I_{pj} N_j - \mu N_p \quad (60b)$$

$$\frac{dI_p}{dt} = kN_p \sum_{j \in N(p)} D_j - \mu_I I_p \quad (60c)$$

In eqs. (60a-b), the production of Notch and Delta is regulated by Hill functions that depend on the level of NICD, with $\lambda_N > 1$ and $0 < \lambda_D < 1$ to capture transcriptional activation of Notch and transcriptional inhibition of Delta. Moreover, this model considers binding of Notch and Delta molecules of a given cell (p) to ligands and receptors of neighboring cells (j); here, k represents a bimolecular binding rate constant between a ligand and a receptor at the surfaces of neighboring cells, and the summation spans over the nearest neighbors of cell (p). Binding of Notch with external ligands leads to release of NICD and degradation of the remaining molecular complex; therefore, the binding term results in a loss term in eq. (60a) and a corresponding production term in NICD's equation (60c). Furthermore, the binding term of Notch receptors with external ligands in eq. (60a) is weighted by a contact

area term (I_{pj}) that describes the shared cell contact area between cells p and j. The limit of a regular lattice is achieved by imposing that all weightage terms are equal ($I_{pj} = 1/c$).

On a disordered lattice with variable cell sizes, smaller cells with smaller contact area tend to acquire a Sender state while larger cells tend to acquire a Receiver state [105]. In the perfectly regular lattice, Senders are selected from a homogeneous initial condition simply due to spontaneous breaking of symmetry and amplification of small initial differences [102]. Instead, variation of cell size bias cell fate selection by weighting the amount of signaling between pairs of neighboring cells, and can be thus viewed as an imprint that guides the patterning by breaking the symmetry between cells. This prediction was directly validated in the context of chicken inner ear development, where smaller cells produce Delta at a high rate and eventually become hair cells, while larger cells generally committed to a non-hair, supporting phenotype [105,108].

3.3.4 Notch-Jagged signaling guides a transition between lateral inhibition and lateral induction

So far, we have reviewed models that primarily focus on Notch-Delta lateral inhibition. Ligands of the Jagged type, however, can give rise to a positive feedback between neighbors and thus promoting lateral induction of the (high Notch, high Jagged) hybrid Sender/Receiver phenotype. The conflicting effects of Delta and Jagged ligands on cell fate raise interesting questions about their competition in multicellular models where many cells collectively converge to diverse patterns. Boareto and collaborators [107,109] proposed a model of Notch-Delta-Jagged signaling based on the circuit schematic of Figure 12A that directly generalizes the Notch-Delta circuit of eqs. (60a-c) to include NICD transcriptional activation of Jagged and Notch-Jagged binding:

$$\frac{dN_i}{dt} = g_N \frac{1 + \lambda_N \left(\frac{I_i}{I_0}\right)^{n_N}}{1 + \left(\frac{I_i}{I_0}\right)^{n_N}} - kN_i \sum_{j \in N(i)} (D_j + J_j) - \mu N_i \quad (61a)$$

$$\frac{dD_i}{dt} = g_D \frac{1 + \lambda_D \left(\frac{I_i}{I_0}\right)^{n_D}}{1 + \left(\frac{I_i}{I_0}\right)^{n_D}} - kD_i \sum_{j \in N(i)} N_j - \mu N_i \quad (61b)$$

$$\frac{dJ_i}{dt} = g_J \frac{1 + \lambda_J \left(\frac{I_i}{I_0}\right)^{n_J}}{1 + \left(\frac{I_i}{I_0}\right)^{n_J}} - kJ_i \sum_{j \in N(i)} N_j - \mu J_i \quad (61c)$$

$$\frac{dI_i}{dt} = kN_i \sum_{j \in N(j)} (D_j + J_j) - \mu_I I_i \quad (61d)$$

In this model, a cell behaves as a three-way switch that can assume a Sender, Receiver, or hybrid Sender/Receiver state based on initial conditions, parameters and state of the neighbors. In the limit of a dominant Notch-Delta signaling ($g_D \gg g_J$), this model recovers a lateral inhibition pattern with alternated Senders and Receivers. In the opposite limit of a dominant Notch-Jagged signaling ($g_D \ll g_J$), however, there is a homogeneous solution where all cells are hybrid Sender/Receivers with (high Notch, high Jagged).

In this model, a single cell exposed to a fixed level of external Notch receptors and ligands can be monostable S, R or S/R, or fall in a regime of multistability based on the levels of external Delta and Jagged (Figure 12B). In a two cells scenario, this model undergoes a sharp transition from lateral inhibition to lateral induction triggered by an increasing production rate of Jagged (Figure 12C). This transition has been used to explain how TNF-α, and inflammatory signal that activates Jagged, prevents physiological angiogenesis that relies on Notch-Delta lateral inhibition [110]. On a one-dimensional chain of cells with periodic boundary conditions, intermediate conditions where both Notch-Delta and Notch-Jagged 'modes' of the signaling are relevant ($g_D \approx g_J$) give rise to disordered configurations with mixtures of Senders, Receivers and hybrid Sender/Receivers [107] reminiscent of partially disordered configurations in a spin system (Figure 12D).

Interestingly, both mathematical models and experimental observations suggest a dual role for Jagged. While a strong Notch-Jagged signaling promotes homogeneous patterns of hybrid Sender/Receiver cells, as observed during inner ear development and angiogenesis, a weaker expression of Jagged assists Notch-Delta signaling to organize a precise lateral inhibition. In the context of inner ear development, Petrovic and collaborators [111] showed experimentally that Jagged ligands help refine the pattern of Senders and Receivers by competing with Delta over Notch receptors. This leads to even more NICD, and thus an even stronger inhibition of Delta, in the Receivers. Similarly, a weak activation of Jagged improves

Notch-Delta-driven angiogenesis in an *in vitro* model developed by Kang and collaborators [110].
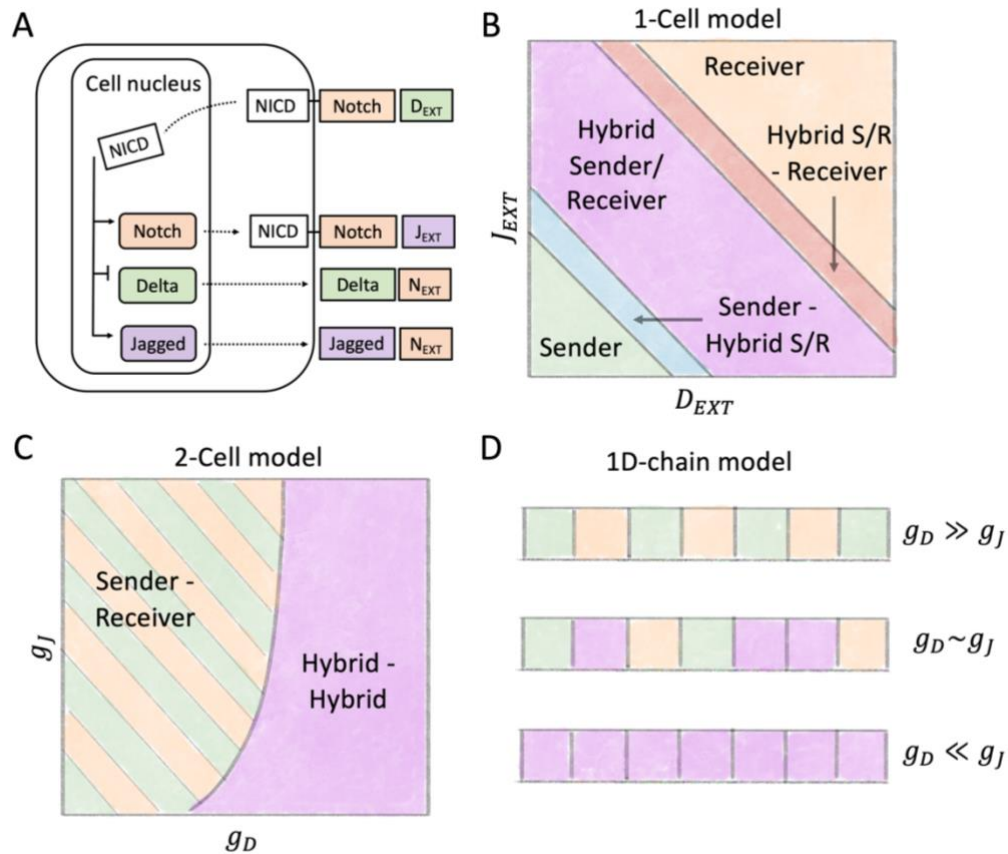

*Figure 12. **Lateral inhibition and lateral induction patterns in the Notch-Delta-Jagged model. (A)** Schematic of the Notch-Delta-Jagged circuit in the model of Boareto and collaborators [107]. Dotted arrows summarize production and transport to cell membrane of Notch, Delta and Jagged molecules. **(B)** Phase diagram of single cell exposed to constant external levels of Delta and jagged ligands (D_{EXT} and J_{EXT}). **(C)** Phase diagram of a 2-cell system as a function of cellular production rate of Jagged and Delta (g_J and g_D). **(D)** Typical patterns observed in a 1D-chain of cells with periodic boundary conditions. Green, orange and purple squares indicate Sender, Receiver and hybrid Sender/Receiver cells, respectively. Panels B, D are adapted from Boareto and collaborators [107]; panel C is adapted from Kang and collaborators [110].*

3.3.5 Stochastic fluctuations lead to optimal lateral inhibition patterning

While the standard paradigm of Notch-Delta lateral inhibition assumes a precise patterning of alternated Senders and Receivers, it is reasonable to assume that spatial constraints might lead to frustrated patterns, in analogy with the relaxation of spin systems. Evidence of patterning mistakes was recently quantified in the context of Tip-Stalk differentiation during sprouting angiogenesis, where Tip cells are occasionally separated by multiple Stalk cells [112,113]. Galbraith and collaborators studied the relaxation of a multicell Notch-Delta system under the effect of white and shot noise, respectively [103]. First, the authors showed that a deterministic Notch-Delta multicell model on a square lattice equilibrates to the "standard", lateral inhibition alternate pattern only for a small set of well-defined initial conditions, such as the "seeding" of a single Sender cell, while typically reaching disordered patterns when starting from more randomized initial conditions. By quantifying the patterning order based on the fraction of correct Sender-Receiver contacts, they demonstrated that intermediate levels of stochastic fluctuations help achieving more ordered patterns, independently of the specific type of noise. The authors suggested an analogy with the navigation of rugged energy landscapes of spin glass systems. The highly ordered salt-and-pepper configuration can be interpreted as a low energy attractor, or global minimum, in a complex, high-dimensional landscape. Conversely, more disordered configurations can be interpreted as local minima with higher energy. Therefore, the relaxation of the Notch-Delta systems is characterized by two timescales. First, on a shorter timescale, the equilibration of the chemical reaction terms leads to the closest local minimum. Second, on a longer timescale, stochastic fluctuations allow a more thorough exploration of the landscape, finally leading to the global minimum with ordered pattern. Intermediate stochastic fluctuations are key to achieve an ordered pattern because low noise levels are not sufficient to escape local minima and navigate the landscape and, conversely, strong fluctuations in the high noise regime become larger than the typical barrier height separating attractors, thus preventing relaxation toward any specific attractor [103].

4. Final remarks and future challenges

In this review, we have provided a general overview of modeling and computational strategies to study the dynamics of gene regulatory networks at different scales and levels of detail, ranging from stochastic simulations of individual chemical reactions up to coarse-grained Boolean descriptions of large networks. In particular, our main goal was to showcase tools to study complex biochemical systems that can give rise to multistability, or, in biological terms, the coexistence of multiple states that can be associated with different cell phenotypes. We urge to point out that there is not an intrinsically better approach, but rather different biological questions require choosing the most suitable tools. For example, stochastic models based on the chemical master equation might be suitable to investigate the dynamics of smaller, well-defined circuits where prior knowledge is available about the reaction parameters. More coarse-grained approaches such as continuous models, parameter randomization or Boolean circuits might be more appropriate to study the emerging dynamics of larger circuits where an informed guess of the model's parameters is unfeasible. Throughout the review, we have demonstrated how some of these strategies can be implemented using small, archetypical systems, such as a single transcribing gene and the bistable toggle switch.

Furthermore, we have provided three specific examples to showcase how the "basic" biological building blocks including transcription and translation can be complemented and generalized to include additional biological processes. In the case of the epithelial-mesenchymal transition, mathematical models that only focus on the transcriptional interactions between genes and transcription factors might not be sufficient to fully capture the biology, as post-translational regulation of non-coding RNAs follow a different dynamic and thus requires different mathematical formulations (see section 3.1 and insight box 3). Moreover, the discussion of stem cell differentiation provided an example of how methodologies originally developed for more traditional physical problems, such as quantum mechanics, can be successfully applied to biology, in this case to capture the stochastic fluctuations of an auto-regulatory gene (see section 3.2 and insight box 4). Finally, while most of the existing modeling efforts tend to focus on the dynamics of individual cells,

the discussion of Notch signaling showed how these models can be applied to multicell, spatial models, thus raising interesting connections between regulation of cell fate and spatial patterning (see section 3.3).

In conclusion, we stress that, while existing methodologies reviewed here and elsewhere provide an exhaustive framework to describe gene regulation, many open questions and challenges lie ahead. For example, the fast development of transcriptomics methods has radically improved our resolution on gene regulation at the single cell level [21]. Single cell RNA sequencing (scRNA-seq) allows to estimate the number of transcripts of each RNA species in individual cells[52]. Some computational methods have been proposed recently to infer the interactions between genes from scRNA-seq data by coupling modeling and statistical regression [24,53,114]. Furthermore, it is becoming increasingly clear that the decision-making of cell fate specification does not solely rely on regulatory interactions between genes. For example, in both the cases of EMT and Notch signaling, phenotypic transitions also imply changes in the cell's mechanical properties, which in turn regulate the transcriptional signaling, thus giving rise to mechano-chemical feedbacks[67]. Therefore, the integration of chemical and mechanical regulation in cell-fate specification represents a novel and intriguing challenge for theoretical and computational modeling. Tackling these exciting open questions will require even more integration of existing and new physical and mathematical concepts in the description of complex biological systems.

Funding statement

Author contribution

F.B. and D.J. wrote the manuscript. All authors edited the manuscript. Q.N, M.K.J., and J.N.O. supervised the research.

References

1. Coco van Boxtel, Johan H. van Heerden, Niclas Nordholt PS and FJB. 2017 Taking chances and making mistakes: non-genetic phenotypic heterogeneity and its consequences for surviving in dynamic environments. *J. R. Soc. Interface* **14**, 20170141.

2. Regev A *et al.* 2017 The human cell atlas. *Elife* **6**, e27041.

3. Bustamante C, Cheng W, Mejia YX. 2011 Revisiting the central dogma one molecule at a time. *Cell* **144**, 480–497.

4. Davidson EH. 2010 Emerging properties of animal gene regulatory networks. *Nature* **468**, 911–920.

5. Alon U. 2006 *An introduction to systems biology: design principles of biological circuits*. Chapman and Hall/CRC.

6. Bressloff PC. 2017 Stochastic switching in biology: from genotype to phenotype. *J. Phys. A Math. Theor.* **50**, 133001.

7. Tsimring LS. 2014 Noise in Biology. *Reports Prog. Phys.* **77**, 026601.

8. Schnoerr D, Sanguinetti G, Grima R. 2018 Approximation and inference methods for

stochastic biochemical kinetics—a tutorial review. *J. Phys. A Math. Theor.* **51**, 169501.

9.   Earnest TM, Cole JA, Luthey-Schulten Z. 2018 Simulating biological processes: stochastic physics from whole cells to colonies. *Reports Prog. Phys.* **81**, 052601.

10. Hausser J, Mayo A, Keren L, Alon U. 2019 Central dogma rates and the trade-off between precision and economy in gene expression. *Nat. Commun.* **10**, 68.

11. Gillespie D. 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434.

12. Gillespie DT. 1977 Exact stochastic simulation of coupled chemical reactions. *Hournal Phys. Chem.* **81**, 2340–2361.

13. Gibson MA, Bruck J. 2000 Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *J. Phys.Chem. A* **104**, 1876–1889.

14. Kierzek AM. 2002 STOCKS: STOChastic Kinetic Simulations of biochemical systems with Gillespie algorithm. *Bioinformatics* **18**, 470–481.

15. Tunnacliffe E, Chubb J. 2020 What Is a Transcriptional Burst? *Trends Genet.* **36**, 288–297.

16. Wang J, Zhang K, Xu L, Wang E. 2011 Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci.* **108**, 8257–8262.

17. Sasai M, Wolynes P. 2003 Stochastic gene expression as a many-body problem. *Proc. Natl. Acad. Sci.* **100**, 2374–2379.

18. Warren PB, Wolde PR ten. 2005 Chemical Models of Genetic Toggle Switches. *J. Phys.Chem. B* **109**, 6812–6823.

19. Hillenbrand P, Fritz G, Gerland U. 2013 Biological Signal Processing with a Genetic Toggle Switch. *PLoS One* **8**, e68345.

20. Wilkinson ME, Charenton C, Nagai K. 2020 RNA Splicing by the Spliceosome. *Annu. Rev. Biochem.* **89**, 359–388.

21. Aldridge S, Teichmann SA. 2020 Single cell transcriptomics comes of age. *Nat. Commun.* **11**, 4307.

22. La Manno G *et al.* 2018 RNA velocity of single cells. *Nature* **560**, 494–498.

23. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. 2020 Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414.

24. Bocci F, Zhou P, Nie Q. 2022 spliceJAC: Identify transition driver genes and cell state

specific regulatory interactions from single-cell transcriptome data. *Mol. Syst. Biol.* **18**, e11176.

25. Qiu X *et al.* 2022 Mapping transcriptomic vector fields of single cells. *Cell* **185**, 690–711.

26. Li T, Shi J, Wu Y, Zhou P. 2021 On the Mathematics of RNA Velocity I: Theoretical Analysis. *CSIAM Trans. Appl. Math.* **2**, 1–55. (doi:https://doi.org/10.1101/2020.09.19.304584)

27. Gorin G, Fang M, Chari T, Pachter L. 2022 RNA velocity unraveled. *PLoS Comput. Biol.* **18**, e1010492. (doi:https://doi.org/10.1101/2022.02.12.480214)

28. Hédia Maamar, Arjun Raj DD. 2007 Noise in Gene Expression Determines Cell Fate in Bacillus subtilis. *Science (80-.).* **317**, 526–529.

29. Hill A. 1910 The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *J. Physiol.*

30. Santillán M. 2008 On the Use of the Hill Functions in Mathematical Modelsof Gene Regulatory Networks. *Math. Model. Nat. Phenom.* **3**, 85–97.

31. Qian G, Mahdi A. 2020 Sensitivity analysis methods in the biomedical sciences. *Math. Biosci.* **323**, 108306.

32. Morio J. 2011 Global and local sensitivity analysis methods for a physical system. *Eur. J. Phys.* **32**, 1577.

33. Swain PS, Elowitz MB, Siggia ED. 2002 Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci.* **99**, 12795–12800. (doi:10.1073/pnas.162041399)

34. Hilfinger A, Paulsson J. 11 ADSeparating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proc. Natl. Acad. Sci.* **108**, 12167–12172.

35. Lu M, Onuchic J, Ben-Jacob E. 2014 Construction of an Effective Landscape for Multistate Genetic Switches. *Phys. Rev. Lett.* **113**, 078102.

36. Waddington C. 1957 The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology. *Georg. Allen Unwin, London*

37. Wang J. 2015 Landscape and flux theory of non-equilibrium dynamical systems with application to biology. *Adv. Phys.* **64**, 1–137.

38. Dikman M, Mori E, Ross J, Hunt P. 1994 Large fluctuations and optimal paths in chemical kinetics. *J. Chem. Phys.* **100**, 5735–5750.

39. Roma DM, O'Flanagan RA, Ruckenstein AE, Sengupta AM, Mukhopadhyay R. 2005 Optimal Path to Epigenetic Switching. *Phys. Rev. E* **71**, 011902.

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

41. Albert R. 2004 Boolean Modeling of Genetic Regulatory Networks. *Complex Networks. Lect. Notes Phys.* **650**.

42. Kauffman S. 1969 Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* **22**, 437–467.

43. Garg A, Cara A Di, Xenarios I, Mendoza L, Micheli G De. 2008 Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics* **24**, 1917–1925.

44. Barbuti R, Gori R, Milazzo P, Nasti L. 2020 A survey of gene regulatory networks modelling methods: from differential equations, to Boolean and qualitative bioinspired models. *J. Membr. Comput.* **2**, 207–226.

45. Wang R-S, Saadatpour A, Albert R. 2012 Boolean modeling in systems biology: an overview of methodology and applications. *Phys. Biol.* **9**, 055001.

46. Huang B, Lu M, Jia D, Ben-Jacob E, Levine H, Onuchic JN. 2017 Interrogating the topological robustness of gene regulatory circuits by randomization. *PLoS Comput. Biol.* **13**, 1–21.

47. Lu M, Jolly MK, Levine H, Onuchic JN, Ben-Jacob E. 2013 MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proc. Natl. Acad. Sci.* **110**, 18174–9.

48. Huang B, , Dongya Jia JF, , Herbert Levine JNO and ML. 2018 RACIPE: a computational tool for modeling gene regulatory circuits using randomization. *BMC Syst. Biol.* **12**.

49. Kohar V, Lu M. 2018 Role of noise and parametric variation in the dynamics of gene regulatory circuits. *npj Syst. Biol. Appl.* **4**. (doi:10.1038/s41540-018-0076-x)

50. Hari K, Sabuwala B, Subramani BV, Porta C La, Zapperi S, Font-Clos F, Jolly MK. 2020 Identifying inhibitors of epithelial–mesenchymal plasticity using a network topology-based approach. *npj Syst. Biol. Appl.* **6**, 15.

51. Burgess DJ. 2019 Spatial transcriptomics coming of age. *Nat. Rev. Genet.* **20**, 317.

52. Luecken MD, Theis FJ. 2019 Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **15**, e8746.

53. Pratapa A, Jalihal AP, Law JN, Bharadwaj A, Murali TM. 2020 Benchmarking

algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat. Methods* **17**, 147–154.

54. McCalla SG, Siahpirani AF, Li J, Pyne S, Stone M, Periyasamy V, Shin J, Roy S. 2023 Identifying strengths and weaknesses of methods for computational network inference from single-cell RNA-seq data. *G3 Genes Genomes Genet.* , jkad004.

55. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. 2010 Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. *PLoS One* **5**, e12776.

56. Moerman T, Santos SA, González-Blas CB, Simm J, Moreau Y, Aerts J, Aerts S. 2019 GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **35**, 2159–2161.

57. Matsumoto H, Kiryu H, Furusawa C, Ko MSH, Ko SBH, Gouda N, Hayashi T, Nikaido I. 2017 SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. *Bioinformatics* **33**, 2314–2321.

58. Woodhouse S, Piterman N, Wintersteiger CM, Göttgens B, Fisher J. 2018 SCNS: a graphical tool for reconstructing executable regulatory networks from single-cell genomic data. *BMC Syst. Biol.* **12**, 59.

59. Aubin-Frankowski P-C, Vert J-P. 2020 Gene regulation inference from single-cell RNA-seq data with linear differential equations and velocity inference. *Bioinformatics* **36**, 4774–4780.

60. Haghverdi L, Büttner M, Wolf FA, Buettner F, Theis FJ. 2016 Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848.

61. Baek S, Lee I. 2020 Single-cell ATAC sequencing analysis: From data preprocessing to hypothesis generation. *Comput. Struct. Biotechnol. J.* **18**, 1429–1439.

62. Nieto MA, Huang RY, Jackson RA, Thiery JP. 2016 EMT: 2016. *Cell* **166**, 21–45.

63. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. 2018 EMT in cancer. *Nat. Rev. Cancer* **18**, 128–134.

64. Jia D, Li X, Bocci F, Tripathi S, Deng Y, Jolly MK, Onuchic JN, Levine H. 2019 Quantifying cancer epithelial-mesenchymal plasticity and its association with stemness and immune response. *J. Clin. Med.* **8**, 725.

65. Campbell K, Casanova J. 2016 A common framework for EMT and collective cell migration. *Development* **143**, 4291–4300.

66. Jolly MK, Boareto M, Huang B, Jia D, Lu M, Ben-Jacob E, Onuchic JN, Levine H. 2015

Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. *Front. Oncol.* **5**.

67. Mercedes SAV, Bocci F, Levine H, Onuchic JN, Jolly MK, Wong PK. 2021 Decoding leader cells in collective cancer invasion. *Nat. Rev. Cancer* **21**, 592–604.

68. Zhang J, Tian X, Zhang H, Teng Y, Li R, Bai F. 2015 TGF-beta– induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci. Signal.* **7**, 1–12.

69. Bocci F *et al.* 2019 NRF2 activates a partial Epithelial-Mesenchymal Transition and is maximally present in a hybrid Epithelial/Mesenchymal phenotype. *Integr. Biol.* **11**, 251–263.

70. Li C, Hong T, Nie Q. 2016 Quantifying the landscape and kinetic paths for epithelial-mesenchymal transition from a core circuit. *Phys. Chem. Chem. Phys.* **18**, 17949–17956. (doi:10.1039/c6cp03174a)

71. Zhdanov VP. 2011 Kinetic models of gene expression including non-coding RNAs. *Phys. Rep.* **500**, 1–42.

72. Bocci F, Levine H, Onuchic JN, Jolly MK. 2019 Quantitative characteristic of ncRNA regulation in gene regulatory networks. *Lai X., Gupta S., Vera J. Comput. Biol. Non-Coding RNA. Methods Mol. Biol.* **1912**, 341–366.

73. Mingyang Lu, Mohit Kumar Jolly, Ryan Gomoto, Bin Huang, José Onuchic EB-J. 2013 Tristability in Cancer-Associated MicroRNA-TF Chimera Toggle Switch. *J. Phys. Chem. B* **117**, 13164–13174.

74. Jia D, Jolly MK, Boareto M, Parsana P, Mooney SM, Pienta KJ, Levine H, Ben-Jacob E. 2015 OVOL guides the epithelial-hybrid-mesenchymal transition. *Oncotarget* **6**, 15436–48.

75. Jolly MK *et al.* 2016 Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget* **7**, 27067–84.

76. Bocci F, Jolly MK, Tripathi SC, Aguilar M, Onuchic N, Hanash SM, Levine H, Levine H. 2017 Numb prevents a complete epithelial – mesenchymal transition by modulating Notch signalling. *J. R. Soc. Interface* **14**, 20170512.

77. Steinway SN, Zanudo JGT, Ding W, Rountree CB, Feith DJ, Loughran TP, Albert R. 2014 Network modeling of TGF-beta signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint sonic hedgehog and Wnt pathway activation.

Cancer Res. **74**, 5963–5977. (doi:10.1158/0008-5472.CAN-14-0225)

78. Steinway SN, Zañudo JGT, Michel PJ, Feith DJ, Loughran TP, Albert R. 2015 Combinatorial interventions inhibit TGFβ-driven epithelial-to-mesenchymal transition and support hybrid cellular phenotypes. *npj Syst. Biol. Appl.* **1**, 15014. (doi:10.1038/npjsba.2015.14)

79. Font-Clos F, Zapperi S, La Porta CAM. 2018 Topography of epithelial–mesenchymal plasticity. *Proc. Natl. Acad. Sci.* **115**, 5902–5907.

80. Sui Huang, Yan-Ping Guo, Gillian May TE. 2007 Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695–713.

81. Jolly MK, Huang B, Lu M, Mani SA, Levine H, Ben-Jacob E. 2014 Towards elucidating the connection between epithelial--mesenchymal transitions and stemness. *J. R. Soc. Interface* **11**, 20140962.

82. Vijay Chickarmane, Carl Troein, Ulrike A Nuber, Herbert M Sauro CP. 2006 Transcriptional Dynamics of the Embryonic Stem Cell Switch. *PLoS Comput. Biol.* **2**, e123.

83. Chickarmane V, Peterson C. 2008 A Computational Model for Understanding Stem Cell, Trophectoderm and Endoderm Lineage Determination. *PLoS One* **3**, e3478.

84. Tibor Kalmar , Chea Lim , Penelope Hayward , Silvia Muñoz-Descalzo, Jennifer Nichols, Jordi Garcia-Ojalvo AMA. 2009 Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells. *PLoS Biol.* **7**, e1000149.

85. Ingmar Glauche, Maria Herberg IR. 2010 Nanog Variability and Pluripotency Regulation of Embryonic Stem Cells - Insights from a Mathematical Model Analysis. *PLoS One* **5**, e11238.

86. Maria Herberg , Tüzer Kalkan, Ingmar Glauche, Austin Smith IR. 2014 A Model-Based Analysis of Culture-Dependent Phenotypes of mESCs. *PLoS One* **9**, e92496.

87. Zhang B, Wolynes PG. 2014 Stem cell differentiation as a many-body problem. *Proc. Natl. Acad. Sci.* **111**, 10185–10190.

88. Chunhe Li JW. 2013 Quantifying Cell Fate Decisions for Differentiation and Reprogramming of a Human Stem Cell Network: Landscape and Biological Paths. *PLoS Comput. Biol.* **9**, e1003165.

89. Li C, Wang J. 2015 Quantifying the landscape for development and cancer from a core cancer stem cell circuit. *Cancer Res.* **75**, 2607–2618. (doi:10.1158/0008-5472.CAN-15-

0079)

90. Masaki Sasai, Yudai Kawabata, Koh Makishi, Kazuhito Itoh TPT. 2013 Time scales in epigenetic dynamics and phenotypic heterogeneity of embryonic stem cells. *PLoS Comput. Biol.* **9**, e1003380.

91. Bin Huang, Mingyang Lu, Madeline Galbraith, Herbert Levine JNO and DJ. 2020 Decoding the mechanisms underlying cell-fate decision-making during stem cell differentiation by random circuit perturbation. *J. R. Soc. Interface* **17**, 20200500.

92. Li C. 2018 Landscape of gene networks for random parameter perturbation. *Integr. Biol.* **10**, 92–99.

93. Glasser DCM and ML. 1998 The uses of quantum field theory in diffusion-limited reactions. *Rev. Mod. Phys.* **70**, 979.

94. Bray SJ. 2016 Notch signaling in context. *Nat. Rev. Mol. Cell Biol.* **17**, 722–735.

95. Siebel C, Lendahl U. 2017 Notch Signaling in Development, Tissue Homeostasis, and Disease. *Physiol. Rev.* **97**, 1235–1294.

96. Kovall RA, Gebelein B, Sprinzak D, Kopan R. 2017 The Canonical Notch Signaling Pathway: Structural and Biochemical Insights into Shape, Sugar, and Force. *Dev. Cell* **41**, 228–241.

97. Andersson ER, Sandberg R, Lendahl U. 2011 Notch signaling: simplicity in design, versatility in function. *Development* **138**, 3593–3612.

98. Sjöqvist M, Andersson ER. 2019 Do as I say , Not (ch) as I do : Lateral control of cell fate. *Dev. Biol.* **447**, 58–70.

99. Bocci F, Onuchic JN, Jolly MK. 2020 Understanding the principles of pattern formation driven by Notch signaling by integrating experiments and theoretical models. *Front. Physiol.* **11**, 929.

100. Shaya O, Sprinzak D. 2011 From Notch signaling to fine-grained patterning: Modeling meets experiments. *Curr. Opin. Genet. Dev.* **21**, 732–739.

101. Joana Neves, Gina Abello JP and FG. 2013 Patterning and cell fate in the inner ear: a case for Notch in the chicken embryo. *Dev. Growth Differ.* **55**, 96–112.

102. R.Collier J, A.M.Monk N, K.Maini P, H.Lewis J. 1996 Pattern Formation by Lateral Inhibition with Feedback: a Mathematical Model of Delta–Notch Intercellular Signalling. *J. Theor. Biol.* **183**, 429–446.

103. Galbraith M, Bocci F, Onuchic JN. 2022 Stochastic fluctuations promote ordered

pattern formation of cells in the Notch-Delta signaling pathway. *PLoS Comput. Biol.* **18**, e1010306.

104. Teomy E, Levine H, Kessler DA. 2021 Ordered Hexagonal Patterns via Notch-Delta Signaling. *Phys. Biol.* **18**, 066006.

105. Shaya O *et al.* 2017 Cell-cell contact area affects Notch signaling and Notch-dependent patterning. *Dev. Cell* **40**, 505–511.

106. Pau Formosa-Jordan MĬ es. 2014 Competition in Notch Signaling with Cis Enriches Cell Fate Decisions. *PLoS One* **9**, e95744.

107. Boareto M, Jolly MK, Lu M, Onuchic JN, Clementi C, Ben-Jacob E. 2015 Jagged-Delta asymmetry in Notch signaling can give rise to a Sender/Receiver hybrid phenotype. *Proc. Natl. Acad. Sci.* **112**, 402–409.

108. Goodyear R, Richardson G. 1997 Pattern formation in the basilar papilla: evidence for cell rearrangement. *J. Neurosci.* **17**, 6289–6301.

109. Jolly MK, Boareto M, Lu M, Onuchic JN, Clementi C, Ben-Jacob E, Jose'N O, Clementi C, Ben-Jacob E. 2015 Operating principles of Notch-Delta-Jagged module of cell-cell communication. *New J. Phys.* **17**, 55021.

110. Kang T-Y, Bocci F, Jolly MK, Levine H, Onuchic JN, Levchenko A. 2019 Pericytes enable effective angiogenesis in the presence of pro-inflammatory signals. *Proc. Natl. Acad. Sci.* **116**, 23551–23561.

111. Petrovic J, Formosa-Jordan P, Luna-Escalante JC, Abello G, Ibanes M, Neves J, Giraldez F. 2014 Ligand-dependent Notch signaling strength orchestrates lateral induction and lateral inhibition in the developing inner ear. *Development* **141**, 2313–2324.

112. Yen Ling Koon, Songjing Zhang, Muhammad Bakhait Rahmat CGK& K-HC. 2018 Enhanced Delta-Notch Lateral Inhibition Model Incorporating Intracellular Notch Heterogeneity and Tension-Dependent Rate of Delta-Notch Binding that Reproduces Sprouting Angiogenesis Patterns. *Sci. Rep.* **8**, 9519.

113. Kang T-Y, Bocci F, Nie Q, Onuchic JN, Levchenko A. 2022 Spatial-temporal order-disorder transition in angiogenic NOTCH signaling controls cell fate specification. *bioRxiv* (doi:https://doi.org/10.1101/2022.12.07.519534)

114. Ramirez D, Kohar V, Lu M. 2020 Toward Modeling Context-Specific EMT Regulatory Networks Using Temporal Single Cell RNA-Seq Data. *Front. Mol. Biosci.* **7**, 54.