

## RESEARCH ARTICLE

# A single-cell resolved cell-cell communication model explains lineage commitment in hematopoiesis

Megan K. Rommelfanger and Adam L. MacLean\*

## ABSTRACT

Cells do not make fate decisions independently. Arguably, every cell-fate decision occurs in response to environmental signals. In many cases, cell-cell communication alters the dynamics of the internal gene regulatory network of a cell to initiate cell-fate transitions, yet models rarely take this into account. Here, we have developed a multiscale perspective to study the granulocyte-monocyte versus megakaryocyte-erythrocyte fate decisions. This transition is dictated by the *GATA1-PU.1* network: a classical example of a bistable cell-fate system. We show that, for a wide range of cell communication topologies, even subtle changes in signaling can have pronounced effects on cell-fate decisions. We go on to show how cell-cell coupling through signaling can spontaneously break the symmetry of a homogenous cell population. Noise, both intrinsic and extrinsic, shapes the decision landscape profoundly, and affects the transcriptional dynamics underlying this important hematopoietic cell-fate decision-making system.

This article has an associated 'The people behind the papers' interview.

**KEY WORDS:** Cell fate, Hematopoiesis, Mathematical modeling, Gene regulatory network, Bistability

## INTRODUCTION

The production of mature cell types from stem cells and progenitors is essential for development and organ homeostasis. Nevertheless, in few cases, are we able to fully specify the conditions necessary to drive stem cell differentiation towards a particular cell lineage. Stem cell differentiation is controlled by cell-internal and -external signals that, in turn, control the transcriptional state of the cell and specify its eventual fate (Roeder, 2006; Laurenti and Göttgens, 2018; Graf and Enver, 2009). These changes in transcriptional state are often irreversible and involve binary choices. Thus, multipotent cells, through successive binary lineage specification choices, eventually become committed to a specific lineage and cell-fate. Significant unanswered questions remain regarding the cell-fate decisions that dictate lineage specification of stem/progenitor cells: How large is the role of extracellular signaling in

regulating cell differentiation? What mechanisms allow cells from initially homogeneous or clonal populations to converge to different lineages? How does one population of progenitor cells maintain stable heterogeneous subpopulations of committed cell types?

During lineage specification, changes in gene expression are controlled by gene regulatory networks (GRNs), consisting of genes and their protein products (transcription factors), which are able to regulate the expression of various genes, including other transcription factors, creating feedback loops (Tyson et al., 2003). Codified through mathematical models, GRNs can be studied in light of their steady states, where each steady state can represent a committed cell fate. Certain GRN topologies permit bistability, i.e. more than one steady state can be reached for a single set of biological conditions (Craciun et al., 2006): such topologies are frequently observed in networks that control cell-fate decisions (Alon, 2007; Kamenz et al., 2021). The GRN topology that dictates a particular lineage decision is generally conserved across cells, thereby providing insight into the intracellular dynamics that occur during cell differentiation. Crucially, the GRNs that instigate or reinforce lineage decisions are not only controlled by cell-intrinsic networks, but also by cell-extrinsic signals.

The GRN topology considered in this work consists of two mutually repressive genes, a topology frequently observed among gene networks mediating lineage decisions (Pal et al., 2014; Ferrell, 2002). One such lineage decision occurs during hematopoiesis: myeloid progenitor cells make a binary choice between commitment to the megakaryocyte-erythroid (ME) lineage or the granulocyte-monocyte (GM) lineage. *GATA1* and *PU.1* (*SPI1*) mutually inhibit one another; *GATA1* is expressed in the ME lineage and *PU.1* is expressed in the GM lineage. This mutually repressive GRN has been extensively studied and characterized in models, mostly consisting of ordinary differential equations that permit bistability, thus enabling investigation into the dynamics of this myeloid lineage decision (Chickarmane et al., 2009; Duff et al., 2012; Huang et al., 2007; Chang et al., 2008; Roeder and Glauche, 2006).

Given the *GATA1-PU.1* mutual inhibitory loop that leads to bistability, changing the initial conditions (gene expression levels) within a bistable region is sufficient to change the cell fate. It has thus been proposed that random fluctuations of *GATA1* and *PU.1* levels are primarily responsible for determining cell fate in the bipotent progenitor population that has dual ME and GM lineage potential (Chang et al., 2008). More recently, this notion was challenged by a study that used a double reporter mouse (*PU.1<sup>eGFP</sup>GATA1<sup>mCherry</sup>*) to show that that random fluctuations of *PU.1* and *GATA1* (*Spi1* and *Gata1* in mouse) are insufficient to initiate the cell-fate decision between ME and GM lineages (Hoppe et al., 2016). Hoppe et al. also provide strong evidence that ME versus GM lineage specification cannot be determined solely from initial ratios of *PU.1* to *GATA1* expression.

As suggested by Hoppe et al. (Hoppe et al., 2016), extracellular signaling could resolve this controversy. Cell signaling pathways, including Jak/Stat, Wnt/ $\beta$ -catenin, MAPK/ERK/p38 and PI3K/Akt

Department of Quantitative and Computational Biology, University of Southern California, 1050 Childs Way, Los Angeles, CA 90089, USA.

\*Author for correspondence (macleana@usc.edu)

 M.K.R., 0000-0003-3071-7419; A.L.M., 0000-0003-0689-7907

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Handling Editor: Paul François

Received 3 May 2021; Accepted 6 November 2021

have been recently identified as being crucially important during GM and ME cell-fate specification (Wang et al., 2021). Previous studies of hematopoiesis both *in vivo* and *in vitro* have also hinted that cell-cell signaling influences cell-fate decision-making. *In vivo*, it has been shown that the number of hematopoietic cells transplanted affects the phenotype of the regenerated blood system in the recipient (Brewer et al., 2016). *In vitro*, hematopoietic stem and progenitor cells placed into wells of different cavity sizes, and therefore with different numbers of neighboring cells, gave rise to different cell-fate patterns (Müller et al., 2015). These two lines of evidence suggest that cell-cell signaling plays a role in hematopoietic cell-fate specification. Both the theoretical gap and the experimental data thus point to a pressing need to consider the impact of extracellular signaling dynamics when modeling hematopoietic differentiation.

The role of extracellular signaling, however, has been chronically understudied in gene regulatory network models. Extracellular signaling through paracrine factors enables populations of cells to share information and coordinate behaviors over short distances, and is implicated in a range of cell-fate behaviors from developmental branching (Lambert et al., 2018; Iber and Menshkyau, 2013) to patterning (Scholes et al., 2019) and migration (McLennan et al., 2015). A seemingly simple question motivated this work: how do intercellular signaling networks impact myeloid lineage specification, as controlled by the *GATA1-PU.1* gene regulatory network?

Several studies have characterized the impact on cell fate of cell-cell communication through extracellular signals. One approach taken is to model the internal gene dynamics with differential equations and model cell-cell communication by allowing for molecular diffusion of proteins between neighboring cells (Ellison et al., 2016; Mugler et al., 2016; Smith and Grima, 2018). These methods must neglect anisotropic effects, and assume that signaling between cells occurs on the same timescale as intracellular dynamics; yet the processes can occur on vastly different timescales. Other approaches have relaxed this assumption, permitting different timescales by modeling the response time of a cell to a signal received as a random variable (Thurley et al., 2018). This approach, however, omits any detail of the intracellular dynamics, which are often crucial to the response; GRNs are the enforcing mechanisms of the phenotypic switches. In order to model the individual behaviors of a population of communicating cells, the gap between intracellular dynamics and extracellular signaling must be bridged. A model of cell-fate decision-making is needed that permits cell-cell communication while allowing for both description of the cell-internal dynamics and of the extracellular signaling dynamics without making diffusion-like approximations.

We present a multiscale model that bridges this gap. We assume deterministic dynamics within each cell and thus model the cell-internal GRN with ordinary differential equations (ODEs). We assume that signals sent between cells can be described by a Poisson process. Signals received by cells alter the internal GRN dynamics through their effects on parameters of the ODE model. We test this model in a large range of different intercellular signaling topologies. We find that the addition of cell-cell communication to GRN dynamics leads to model outcomes (cell-fate distributions) becoming probabilistic: cell-fate choice probabilities now depend on the position of the cell in a particular signaling topology. We discover that the model can intuitively characterize cell-cell coupling, changes to which impact the cell-fate decisions made, which can lead to mixtures of heterogeneous cell types within a

population. We also study how noise impacts fate outcomes: we find that although both intrinsic and extrinsic noise alter the cell-fate decision-making boundaries, extrinsic noise is the dominant driver of cell-fate variability. Finally, we study how cell fates are influenced by spatial architecture, and we demonstrate that the extent to which cell fates are coupled is controlled by proximity in spatially structured cell-cell signaling systems.

## RESULTS

### Cell-cell communication over a wide range of signaling topologies leads to divergent cell fates

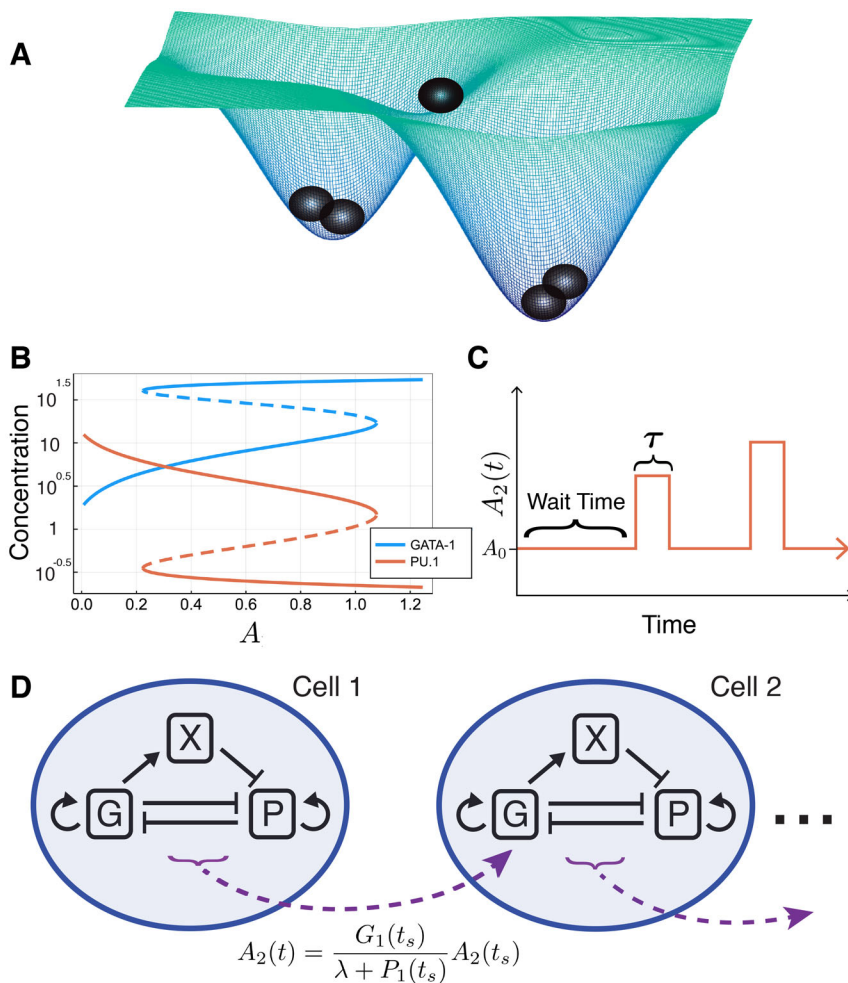
To determine how cell-cell communication instructs hematopoietic lineage specification, we construct a multiscale model of the bifurcation point in hematopoiesis that separates erythroid/megakaryocyte from granulocyte/monocyte lineages (Fig. 1A). This lineage decision is controlled in part by the mutually antagonistic genes *GATA1* (*G*) and *PU.1* (*P*), giving rise to a bistable model (Fig. 1B). We model the cell-internal GRN coupled to a cell-cell communication motif that can alter the cell-internal expression of *G* in neighboring cells, according to a Poisson process (Fig. 1C,D). Full details of the model specification can be found in the Materials and Methods.

We test the model for a wide range of signaling topologies. We sought to ensure that core assumptions were upheld, namely that all cells eventually reach a steady state, even in the presence of noisy signaling, and that cells exhibit bistability with respect to  $A_0$  (the initial value of *A*). These behaviors were preserved in all signaling topologies tested. Moreover, we found that the addition of cell-cell signaling to a previously deterministic model results in probabilistic cell-fate choices.

In Fig. 2, simulated trajectories for a 20-cell loop topology (see also Fig. 3B) are shown. We plot only the concentration of *G* as a proxy for fate choice, as its state is sufficient to determine the steady state of the full system (if *G* is high at steady state, *P* is low, and vice versa); as illustrated by the trajectories of both *G* and *P* (Fig. S1). These trajectories illustrate the variety of cell-fate distributions that are observed in the presence of cell-cell communication (fate convergence to either the high or low state, and fate divergence). In Fig. 2C,D, the values of  $A_0$  are equal, demonstrating that  $A_0$  does not determine the proportion of cells in each state. Rather, for each value of  $A_0$ , the probability of each cell converging to a certain state can be computed.

In analyzing how signals propagate down chains of cells, two striking observations are made. First, each additional signal shifts the  $P(G \text{ high})$  curve (the probability of reaching the *G* high state) to the left, although these shifts are successively smaller down the chain such that eventually the distributions converge. Second, a ‘sharpening’ of the  $P(G \text{ high})$  distributions occurs whereby the region of uncertain fate decreases for cells further down a chain (Fig. 3A). This is a hallmark of cooperativity: accumulated signals by cell-cell communication act to reinforce cell-fate decision-making, leading to regions of uncertainty that are decreasing with the total number of signals that are acting on a cell.

In loops, all cells converge to the same fate for each simulated trajectory, therefore every cell in the model has the same probability distribution (below we will see that the coordination of fates in all cells within a loop signaling topology can be broken by changing the cell coupling strength). The probability distribution for a cell loop depends on the number of cells in the loop, and we see that for loops of size  $n$ , the behaviors observed above for cell chains are recapitulated (Fig. 3B). For both cell loops and chains, we observe that, as the number of cells in the signaling topology increases, so



**Fig. 1. Overview of multiscale cell-fate commitment model.** (A) Illustration of cells in a two-attractor state model (such as the GATA1-PU.1 inhibitory loop). (B) Bifurcation diagram for the system of ODEs described by Eqns 1-3, with respect to the external signaling parameter  $A$ . (C) Cell-cell communication via signals sent between cells (e.g. from cell 1 to cell 2) is modeled by a Poisson process, with wait times sampled from an exponential distribution and fixed signaling ‘pulses’ of length  $\tau$ , where the value of  $A_2$  is set by Eqns 4 or 5. (D) Full model schematic, depicting the cell-internal gene regulatory network of  $\{G, P, X\}$  modeled by ODEs and the external signal  $A_2$  modeled by the Poisson process in C.

does cell response fidelity, i.e. sharpness of the curve (Fig. 3A,B), a prediction that matches experimental data on paracrine signaling in wound healing (Handly et al., 2015). This holds true for large topologies up to at least 100 cells in size (Fig. S2). Moreover, the limiting distributions for the cell chain and the cell loop are the same, highlighting an important underlying convergence.

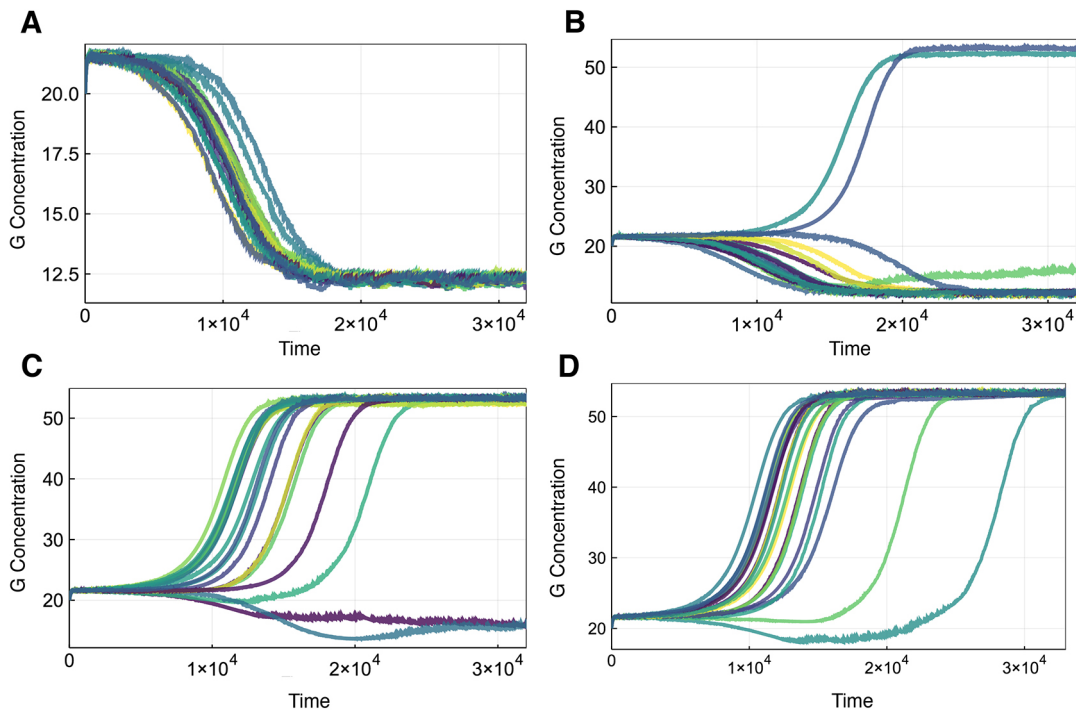
If we consider a dissensus signal (the opposite of a consensus signal), the fate probability distribution of cell 2 shifts in the opposite direction than previously (Fig. 3C), relative to the value of  $A_0$ , at which cell 1 switches lineages. When an inhibition signal is incorporated into a loop (Fig. 3D), the two cells in the loop no longer coordinate lineages for every trajectory as they did when all signals were regular. Notably, in Fig. 3D there exists a large range of  $A_0$  for which cell 1 converges to the  $G$  high state and cell 2 converges to the  $G$  low state, i.e. dissensus signals significantly change cell-fate behaviors in both chains and loops of cells.

In a feed-forward signaling motif (Fig. 3E), multiple promoting signals reinforce the lineage choice of cell 3, thus increasing cooperativity and as a result the cell commits to the  $G$  high state at smaller values of  $A_0$  than it would in a cell chain. Other three-cell signaling topologies also display interesting behaviors (Fig. 3F,G and Fig. S3). In an incoherent feed-forward motif, we observe a multiplicative effect between the dissensus signal from cell 1 and the consensus signal from cell 1 via cell 2 (Fig. 3F). The addition of the dissensus signal results in the distribution of cell 3 being shifted to the right relative to a corresponding cell in a chain. For a doubly

dissensus topology (Fig. 3G), we find that for a region of  $A_0$  the cell fates can diverge (two cells in the high state; the third in the low state) with non-zero probability.

Consensus or dissensus signaling can occur when the signaling pathway mediating the signal contains a positive- or a negative-feedback loop, respectively. There are many such recorded examples of signaling in hematopoiesis and other contexts. One example of a consensus signal directly implicated in the GM versus ME cell-fate choice is mediated by the cytokine  $TNF\alpha$ , which activates the p38 MAPK pathway, in turn activating  $TNF\alpha$  as a target gene (Zarubin and Han, 2005; Wang et al., 2021). A principal example of dissensus signaling, both in the context of hematopoiesis and other cellular processes, is the Delta-Notch signaling pathway (Ohishi et al., 2003), which leads to consecutively divergent patterns of activity.

We expanded our analysis of cell-cell communication to consider other extracellular signals, and their effects alone or in competition with each other. We studied a model of consensus signaling that impacts the activation of  $PU.1$  rather than  $GATA1$ , i.e. changes the parameter  $B$ . Cell-cell communication in this model impacts cell fate similarly to in the original, i.e. for a bistable region with respect to  $B_0$ , lineage specification becomes probabilistic for simple cell signaling topologies (Fig. S4). We also studied a model with two competing extracellular signals, i.e. we changed both parameters  $A$  and  $B$  through independent Poisson processes. In the context of a two cell chain, we found that the joint influence of the signals



**Fig. 2. Sample trajectories of GATA1 (G) for cells signaling in a loop.** Representative simulations of a 20-cell loop signaling topology. Colors used only to distinguish the trajectories. Across all trajectories, the signaling parameter is set to  $\lambda=28$ . (A)  $A_0=0.995$ . (B)  $A_0=0.9955$ . (C,D)  $A_0=1.0$ .

sharpens the cell-fate distribution. Full characterization of the two-signals model has yet to be conducted as the size of the parameter space doubles, leading to challenges in fully elucidating the region of bistability.

The signaling topologies studied here pertain to small and idealized cell signaling networks. Nonetheless, their behaviors point to general and important effects of cell-cell communication. In all of the above experiments, the addition of cell-cell signaling transformed a population of homogeneous, independent and deterministic cells into one of heterogeneous cells that choose fates non-deterministically. Thus, given cells with identical initial conditions (i.e. transcriptional states), external signaling changes cell-fate outcomes at the population level. This corroborates a major finding of Hoppe et al. (Hoppe et al., 2016): that GMP/MEP cell-fate decisions are not predictable from initial transcription factor ratios alone. Our model not only supports this result, but offers an explanation. The missing determinant of cell-fate commitment – which acts to break the symmetry of an initially homogenous population of cells – is cell-cell communication.

### Varying the strength of cell-cell coupling results in different ratios of stable subpopulations of heterogeneous cell types

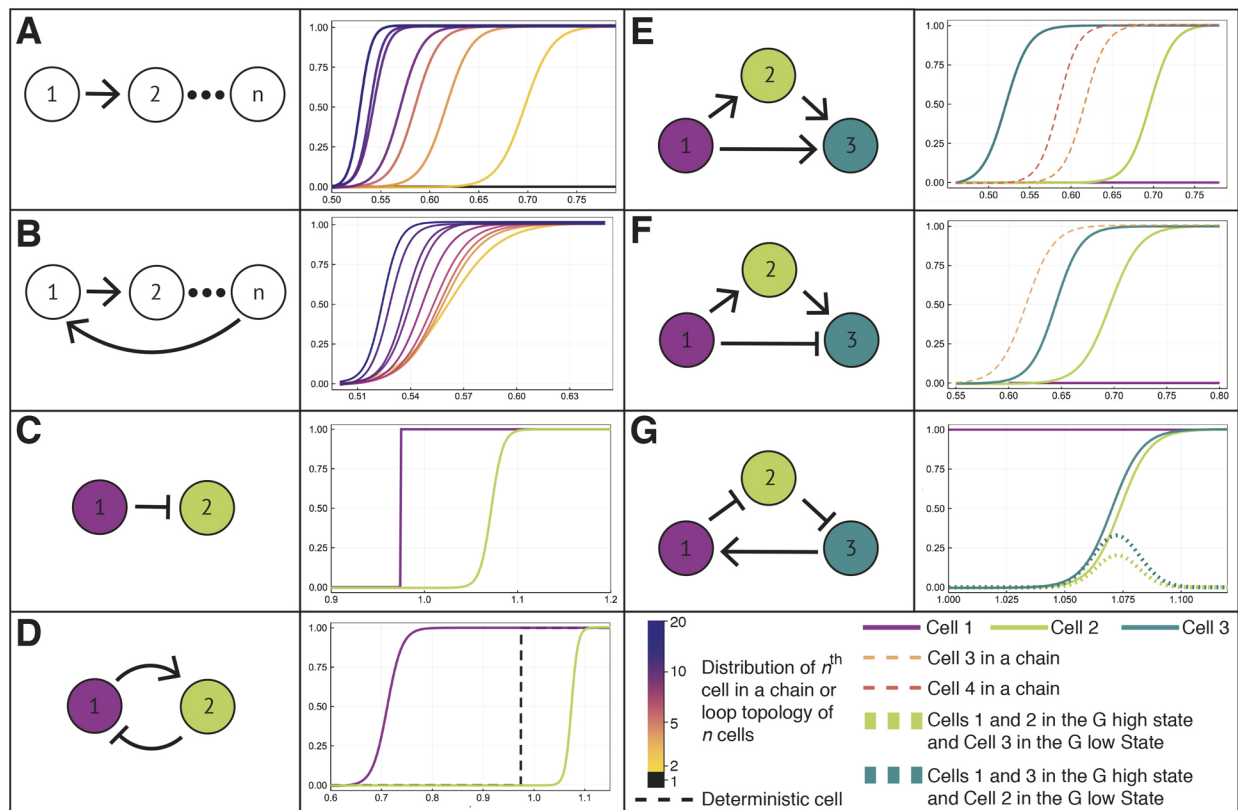
Through investigation of cell-cell communication in the multiscale model introduced above, we discovered that cell-fate divergence can occur under the control of the model parameter  $\lambda$ . We thus characterize  $\lambda$  as a ‘cell-cell decoupling’ mechanism. Analysis of cell-cell coupling and decoupling led to two significant findings: cell-cell communication can explain how a population of cells maintains stable subpopulations of heterogeneous cell types; and the distribution of cells that converge to each fate depends both on cell-cell coupling and on the external environment.

Analysis of the effects of  $\lambda$  on consensus signaling topologies shows that – for a chain of two cells – as  $\lambda$  increases, the probability distribution of the second cell,  $P(G \text{ high})$ , shifts to the right

(Fig. 4A). Opposite results are observed when we analyzed the effects of  $\kappa$  on dissensus signaling topologies, where the probability distribution of the second cell shifts to the left as  $\kappa$  increases (Fig. S5). We identified  $\lambda=22$  as close to a critical value, where cell-fate decision-making becomes switch-like. The value of  $A_0$  where the curve for  $\lambda=22$  switches from one lineage to another is extremely close to the value at which a deterministic cell with the same initial conditions switches. Overall,  $\lambda$  determines the range over which the probability curve for the second cell is non-zero.

We next assessed how different values of  $\lambda$  changed coordination of cell-fate decisions between lineages. Previously, we observed that for any trajectory of a loop topology of any size, where  $\lambda=1$ , all cells converged to the same state, regardless of the value of  $A_0$ . Furthermore, although all cells in the topology converge to the same lineage, the probability that they all converge to a given lineage depended on  $A_0$ . To see whether or not this behavior was conserved for other values of  $\lambda$ , we tested a two cell loop topology with a range of  $\lambda$  values. We observed that there exists a value,  $\lambda^*\approx 22$ , such that, for  $\lambda<\lambda^*$ , all the cells coordinate their lineage just as we had seen before. For  $\lambda>\lambda^*$ , the two cells do not always coordinate their lineage decisions. For example, Fig. S6 gives sample trajectories showing two cells in a loop converging to each combination of states. Fig. 4B shows the two cell loop results for  $\lambda=38$ , giving the probabilities that each cell converges to the  $G$  high state as well as the probabilities that the cells are decoupled, converging to opposite lineages. For each  $\lambda$ , we recorded the maximum value of the probability that the cells converge to different lineages. Fig. 4C shows how the maximum probability of cells converging to different lineages increases with  $\lambda$ .

Next, we looked at larger loops of 10 cells. For each value of  $A_0$ , we simulated 500 trajectories and recorded how many cells were in the  $G$  high state versus the  $G$  low state. Fig. 4D shows results for  $\lambda=30$ , and  $A_0=1.0, 1.002$  and  $1.004$ . We see that the distribution of cells converging to each state changes with the value of  $A_0$ .



**Fig. 3. Cell-fate commitment across a range of signaling topologies.** Schematic of different cell signaling topologies (left) and the corresponding probability distributions: the probability that each cell in the network will commit to the G high steady-state,  $P(G \text{ high})$ , for different values of  $A_0$ . (A) Chains of cells of length  $n$ . Each probability distribution corresponds to one cell in the chain. (B) Loops of cells of size  $n$ . In a loop of fixed size, each cell from 2 to  $n$  has the same probability distribution; thus, one curve is plotted for each loop of  $n$  cells. (C) A chain of two cells with a dissensus signal: the distribution of fates of cell 2 is shifted to the right, i.e. it commits at higher values, compared with consensus signal in A, where the distribution of cell 2 is shifted to the left. (D) A loop of two cells with one consensus and one dissensus signal; dashed line indicates the distribution of a single cell that receives no signals (deterministic). We cannot observe the fate where cell 1 is in the G low state and cell 2 is in the G high state. (E) Three-cell feed-forward motif; cells 3 (dashed orange) and 4 (dashed red) from a chain of cells marked for comparison. We see that receiving multiple signals results in a non-additive synergistic effect. (F) Feed-forward motif with a dissensus signal; cell 3 (dashed orange) from a chain of cells marked for comparison. (G) Loop of three cells with two dissensus signals. Dashed lines indicate the probabilities that: (1) cells 1 and 2 are in G high state, while cell 3 is in G low state (dark green); or (2) cells 1 and 3 are in G high state, whereas cell 2 is in the G low state (light green).

Further analysis of distributions with different values of  $\lambda$  can be found in Fig. S7.

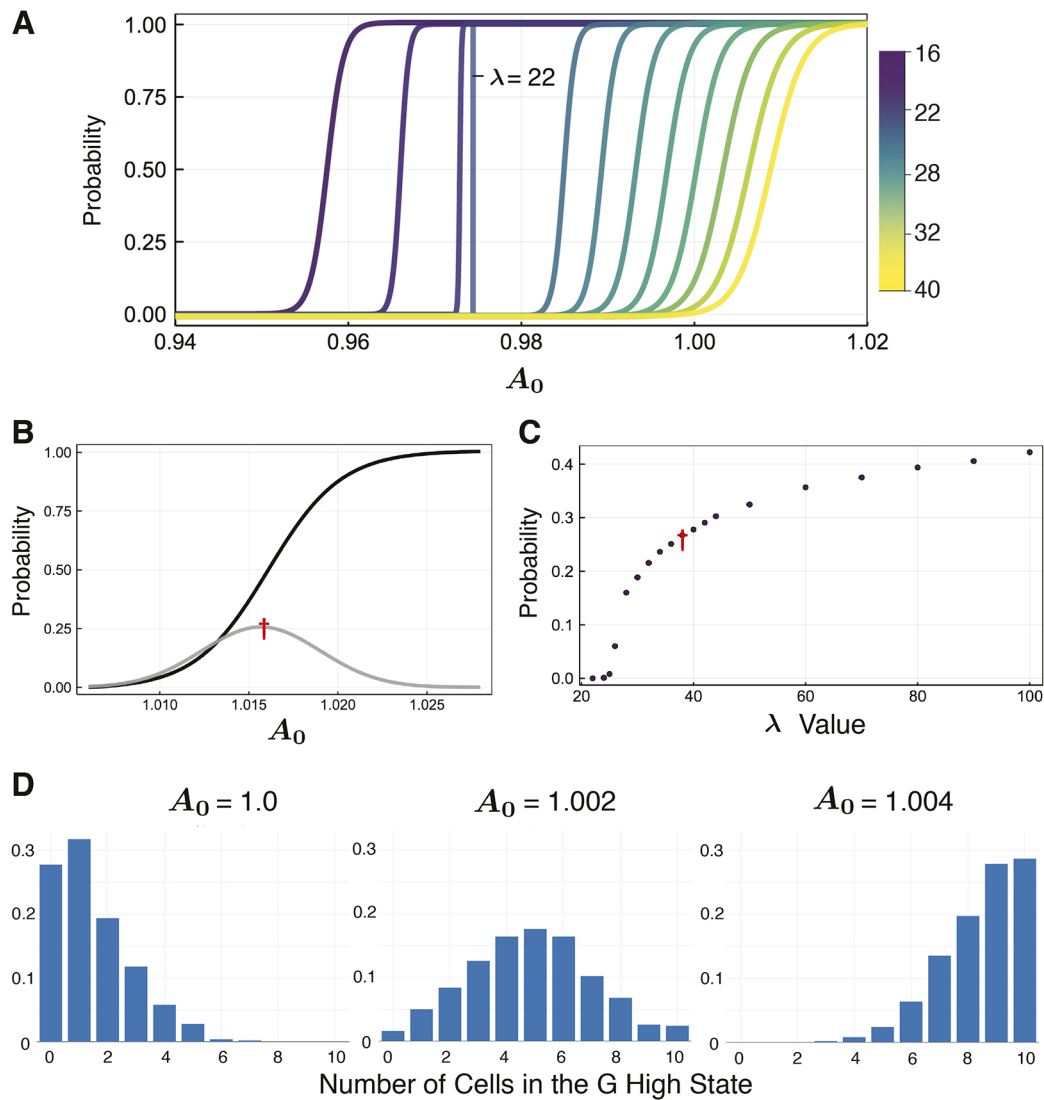
We have identified an explicit mechanism by which cell-cell communication can break the symmetry of a homogeneous population of progenitor cells, and give rise to stable, heterogeneous populations of lineage-committed cells. The proportion of cells committed to each lineage depends on the external environment,  $A_0$ , and the strength of cell-cell coupling due to signaling,  $\lambda$ . Moreover, these results show that fluctuations in the external environment can lead to shifts in the relative abundances of committed cell types. These results corroborate previous work that studied the generation of heterogeneous cell populations through stem cell differentiation (Mojtahedi et al., 2016), and showed that external signals (e.g. through cell-cell communication) are required both to maintain heterogeneous cell populations and to shift relative cell types abundances in response to environmental perturbations.

### Intrinsic and extrinsic noise alter cell-fate decision-making boundaries

We have until now assumed that signals are passed between cells with perfect fidelity. In fact, multiple sources of noise contribute to imperfect communication between cells, and the modeling framework here lends itself well to the investigation of the effects of intrinsic versus extrinsic noise (Hilfinger and Paulsson, 2011;

Swain et al., 2002). We investigated the impact of two different sources of noise in cell-cell communication: due to cell-extrinsic factors, i.e. noise with respect to the extracellular environment (Eqn 7); or due to cell-intrinsic factors, i.e. noise due to signal transduction downstream of a paracrine signaling factor (Eqn 8). These sources of noise are represented in the model by varying either the baseline level of cell-cell communication (extrinsic noise) or the cell signaling pulse level, i.e. the intrinsic signal transduction noise (Fig. 5A).

For simple topologies, as we would expect, as the variance of either the extrinsic or the intrinsic noise increases, the observed variability in cell fate outcomes also increases (Fig. S8). This result also holds for larger cell-cell communication topologies, e.g. a ten-cell loop: intrinsic ( $\eta_p$ ) and extrinsic ( $\eta_e$ ) noise (Fig. 5B-D) both reduce the sensitivity of the cell-fate decision-making boundary. We also observe unexpected and striking results. First, not only are the probability distributions flattened by either noise source, but they are shifted to the left, i.e. intrinsic and extrinsic noise directly affect the decision-making boundary (by shifting its mean), as well as the sensitivity of cell-fate decision-making. Second, even under the presence of noise, we still observe the effects of cooperativity at play through the sharpening of distributions down a chain of cells (Fig. 5D). That is, individual noisy signals increase the variability (reduce the sensitivity) of cell-fate decision-making, but the



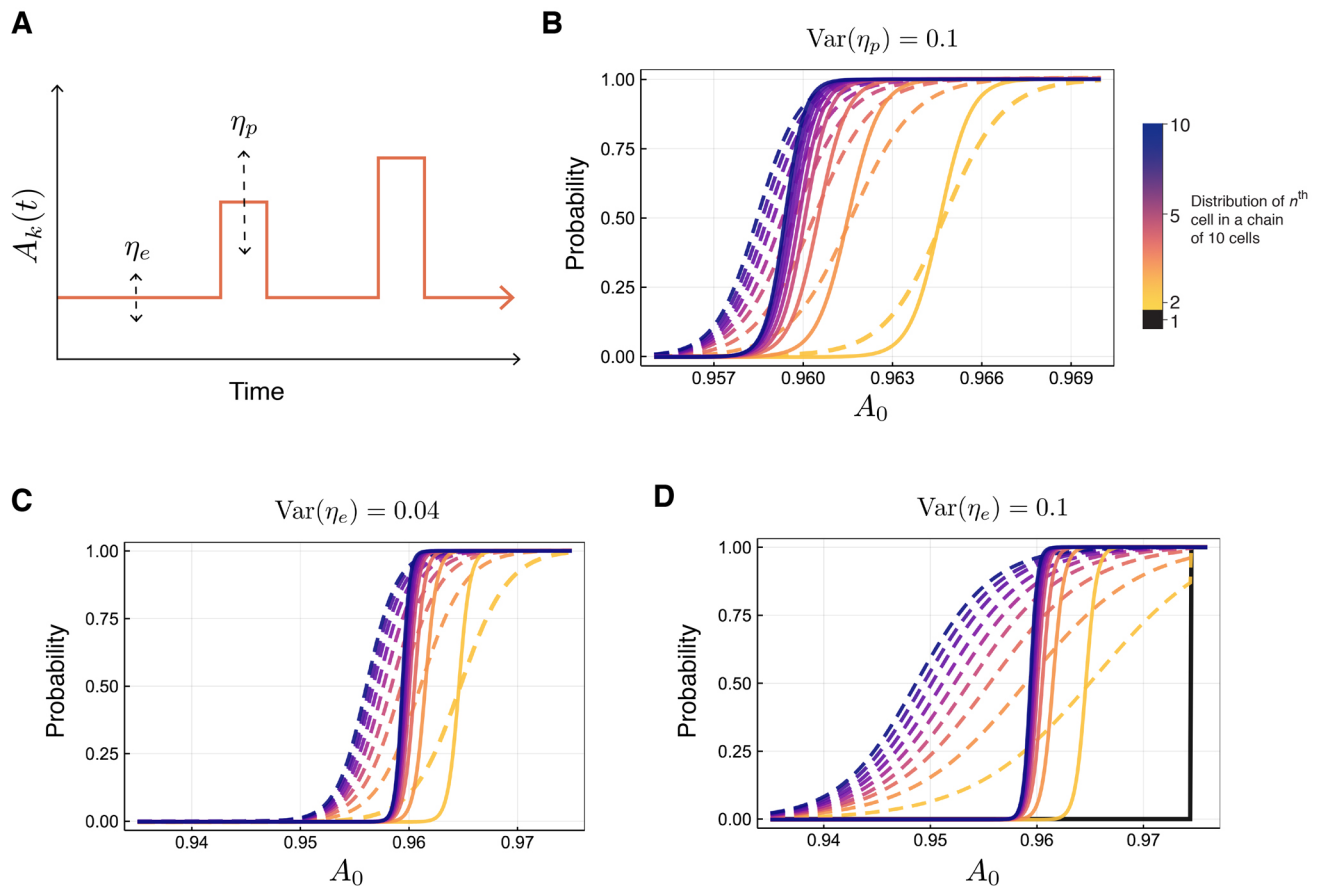
**Fig. 4. Cell-fate coupling is controlled by the external signaling parameter.** (A) Probability distributions of cell-fate commitment to G high state for cell 2 in a chain of two cells for  $\lambda \in [16, 40]$ . (B) Probability distributions for a loop of two cells as a function of  $A_0$ , with  $\lambda = 38$ . Probability of cell-fate commitment to G high state (black); probability that the two cells will not commit to the same fate (gray), with the maximum of the distribution marked (red dagger). (C) For a loop of two cells, the maximum probability that the cells do not converge to the same lineage as  $\lambda$  varies; red dagger equivalent to that in B is marked for comparison. (D) Comparison of distributions for the number of cells that commit to the G high state in a loop of 10 cells with  $\lambda = 30$ .

cumulative effects of noisy signaling can at least partially compensate for this, and decrease the variability in cell-fate decision-making.

Through comparison of the relative effects of the intrinsic signal transduction noise (Fig. 5B) and the extrinsic extracellular noise (Fig. 5C,D), we see that the impact on the cell-fate decision-making boundary is much larger for extrinsic rather than intrinsic noise contributions. From inspection of Eqns 6 and 7, this is in part due to smaller duration of the pulse  $\tau$  relative to the mean wait time  $\mu$ . The result of which is that when  $\eta_e \sim \mathcal{N}(0, 0.1)$  (Fig. 5D), the probability curves for cells 2-10 in the chain flatten to the extent that the cell fates are indeterminate regardless of position along the chain. Moreover, the probability curves for cells  $\geq 2$  along the chain intersect with the point at which cell one (which is deterministic) switches fates from the low to the high state (black line in Fig. 5D), thus forcing all other cells also into the high state with probability = 1. This coordination of cell fates is influenced by cell-cell coupling [here, we use  $\lambda = 18 < \lambda^*$ , i.e. a value

of  $\lambda$  for which we observe complete coupling (fate coordination) between cells]. The dominant impact of extrinsic over intrinsic noise is in agreement with previous works, including a study that quantified the contributions of extrinsic and intrinsic noise in the MAPK signaling pathway, and showed that extrinsic noise is the dominant driver of cell-to-cell variability (Filippi et al., 2016). It has also been shown that explicit extrinsic noise contributions are necessary to explain mRNA abundance distributions (Ham et al., 2020).

In summary, the effects of intrinsic and extrinsic noise on cell-cell communication topologies are to increase the variance of the resulting cell-fate distributions and (surprisingly) to alter the mean values of these probability distributions. In other words, the presence of noise alone can force cells to change lineages. The observed increases in the variability of cell-fate decision-making are maintained for large cell-cell communication topologies. A similar result was described in a study of cell-fate decision-making during early mouse gastrulation (Mohammed et al., 2017):



**Fig. 5. Extrinsic and intrinsic sources of noise reduce fidelity of cell-fate choice.** (A) Schematic depicting how extrinsic noise ( $\eta_e$ ) and intrinsic noise ( $\eta_p$ ) are modeled through their effects on the signaling strength  $A_k$ . (B) Probability distributions of cell-fate commitment (to G high state) for each cell in a ten-cell chain, with  $\lambda=18$ : no added noise (solid lines); and with added noise (dashed lines). The colors darken as the position of the cell along the chain increases. The type of noise modeled is intrinsic, with  $\text{Var}(\eta_p) = \sigma_p^2 = 0.1$ . (C) As for B, but for extrinsic noise modeled, with  $\text{Var}(\eta_e) = \sigma_e^2 = 0.04$ . (D) As for B, but for extrinsic noise modeled, with  $\text{Var}(\eta_e) = \sigma_e^2 = 0.1$ .

transcriptional noise is greatest at the point of cell-fate decision-making (when epiblasts begin to differentiate). Our results reiterate the same point made by Mohammed et al.: that gene expression noise is beneficial during windows of cell-fate decision-making as it leads to an increased possible repertoire of cell fates. Our findings go further in that they suggest a rationale: that the increase in transcriptional noise results from noisy extracellular factors influencing cell-cell signaling during differentiation.

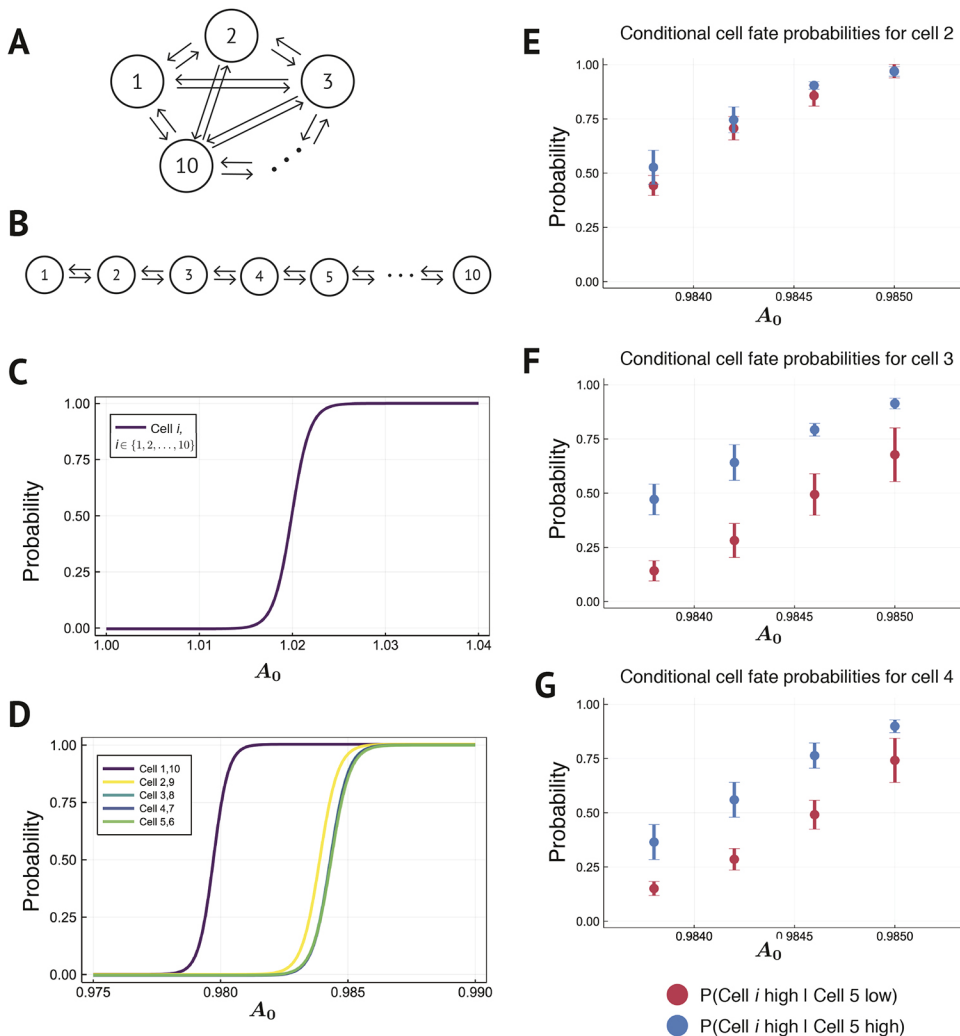
### Spatial structure in communication networks impacts cell-fate outcomes

Years of thorough investigation into the spatial organization of the bone marrow have revealed complex environmental interactions and stem cell niches (Lo Celso and Scadden, 2011; Morrison and Scadden, 2014; Baryawno et al., 2019; Upadhaya et al., 2020; Christodoulou et al., 2020); however, much less focus has been given to the locations of progenitor cells, including multipotent cells of GM and ME lineage potential. Moreover, the freedom of cell movement relative to, for example, epithelial tissues limit our ability to characterize spatial cell locations. Coupled with the current limitations on detecting cell-cell communication directly from data, it is not yet possible to infer spatial cell-cell communication networks in hematopoietic cells from data. Thus, to investigate the impact of spatial organization on cell fate, we investigated two topologies that characterize spatial extrema: a fully connected

topology and a bidirectional chain of cells (Fig. 6A,B). These models use the ‘connectedness’ of a cell to characterize environments either lacking in spatial structure (fully connected topology) or highly structured (bidirectional cell chain).

We simulate ten-cell models for both the well-mixed and spatially restricted topologies, with the same model parameters and initial conditions. We used  $\lambda=28$  and reduced the signaling pulse period to 1.0, to better distinguish the signals due to the large total number of signals being sent and received in the fully connected topology. In the fully connected model, each cell is topologically equivalent and has the same probability distribution of converging to the GM lineage (G high) state (Fig. 6C). However, in the spatially restricted bidirectional chain, the fate choice probability distribution is position dependent: the farther from the center of the chain, the more the distributions diverge (Fig. 6D). The spatial organization encapsulated in the bidirectional chain impacts hematopoietic lineage specification, even in a population of homogeneous cells.

To explore the spatial patterns of cell fate that emerge, we computed the probability of the commitment of an individual cell to the GM lineage conditioned on the fate of cell 5, at the center of the bidirectional chain (Fig. 6E-G). We found that the influence of the fate of cell 5 decreased with increasing distance from cell 5 in the chain. The probability that cell 2 commits to the GM lineage is only marginally affected by the fate of cell 5 (Fig. 6E).



**Fig. 6. Spatial organization of cells regulates patterns of cell-fate commitment.**

Schematic representation of models for (A) a fully connected signaling topology and (B) a bidirectional chain signaling topology. (C,D) Probability distributions of the probability that each cell in the bidirectional chain (C) and fully connected (D) signaling topologies will commit to the G high steady state for different values of  $A_0$ . (E-G) Probabilities that: (E) cell 2, (F) cell 3 or (G) cell 4 will commit to the G high steady state, conditional on the fate of cell 5. Cell 5 committing to the G high state (blue); cell 5 committing to the G low state (red). The model was simulated in  $n=10$  independent experiments each with 100 trajectories. The mean of all 1000 simulations is given, and the error bars indicate the s.d. observed across the 10 samples.

As a function of  $A_0$ , over range of uncertainty in the fate of cell 5, cell 1 commits to the GM lineage with certainty. Cells near to cell 5, in contrast, are much more likely to the GM lineage if cell 5 itself commits to the GM lineage (Fig. 6F,G). That is, local spatial architecture is captured in this model and manifests through the fate coupling of cells in close proximity. This is in stark contrast with cells signaling between each other in a fully connected topology, whereby cells all share the same distributions of cell-fate choice.

We expect hematopoietic organization in the bone marrow to reflect a niche intermediate between the two spatially extreme models studied here. As we have shown through simulation, spatial cell arrangements in the bone marrow will lead to a location-dependent coupling of cell fates in local neighborhoods of hematopoietic progenitor cells. This prediction of the reliance of cell fate on the fate of nearby cells is supported by experimental evidence. In recent work studying hematopoietic progenitor cells *in vivo*. Hérault et al. (2017) show that, unlike at steady state, during regeneration (when the requirement for new myeloid cell production is high), progenitor cells cluster together in groups. In agreement with our model, the spatial structure introduced by spatially proximal progenitor cells influences lineage commitment: the clustered GMPs activate cell cycle pathways and are driven towards differentiation into granulocytes, relative to the steady-state GMPs. Our results could be further tested in culture using similar methods to Hoppe et al. (2016) and analyzing the impact of spatial proximity

on lineage commitment. Although it is currently beyond the capacity of experimental technologies to infer single-cell communication networks directly from data, this is changing (Giladi et al., 2020), leading to future opportunities to fit single-cell signaling models to data.

## DISCUSSION

Despite many theoretical and experimental advances in our understanding of gene regulatory network (GRN) dynamics, our ability to use GRN models to explain cell-fate decision-making during differentiation of multipotent progenitor cells remains limited. Here, with application to a well-studied cell fate GRN, the *GATA1-PU.1* mutual inhibition loop (Chickarmane et al., 2009; Chang et al., 2008; Roeder and Glauche, 2006), we introduced a new model that can simultaneously describe GRN dynamics and single cell-resolved cell-cell communication. Notably, although cell-cell communication is often assumed to be a crucial component of cell differentiation, it is rarely incorporated into models. The previous studies that have characterized cell-cell communication in models did not capture the detailed complexity of these dynamics, by making simplifying assumptions regarding either the GRN dynamics (Thurley et al., 2018) or the mechanisms by which cells signal (Smith and Grima, 2018; Mugler et al., 2016). We found that by combining these dynamic processes that describe noisy cell-fate decision-making at single-cell resolution,



we were able to reconcile several outstanding controversies in the field.

Over a large domain of possible cell-cell communication topologies, we found that cell-cell communication alters cell-fate decision-making boundaries, which become probabilistic in response to the levels of external signaling factors. This helped to reconcile a controversy: that of whether or not transcriptional stochasticity is sufficient to initiate the granulocyte-monocyte versus megakaryocyte-erythrocyte cell-fate decision. Previous models supported the hypothesis; however, Hoppe et al. presented compelling evidence to contradict it (Hoppe et al., 2016). Our results agree with Hoppe et al., in that we show that eventual cell fate cannot be inferred from the initial gene expression state alone – and go further in emphasizing the stochastic nature of these cell-fate decisions (Moris et al., 2016). Analysis of cell-cell coupling led to the discovery that stable distributions of heterogeneous cell types can be robustly maintained by a set of external signals. This result offered insight into another open question: that of how population-level cell-fate behaviors emerge during cell differentiation (Mojtahedi et al., 2016). We also showed how (primarily extrinsic) noise increases the variability of cell-fate decision-making, in line with previous analyses of transcriptional noise during development (Mohammed et al., 2017). Finally, simulation of two topologies representing spatial extrema (well-mixed versus highly structured) suggest that increased coupling of cell fates occurs in proximal cells, a behavior observed experimentally during hematopoietic regeneration (Hérault et al., 2017).

We sought to tightly constrain model complexity here, for reasons of parsimony and interpretability. Relaxing some of the constraints imposed will likely yield further interesting results. We assumed throughout that cells were initially homogeneous, i.e. they shared the same initial conditions and internal GRN networks. Heterogeneous initial conditions and heterogeneous cell-fate decision-making (i.e. different GRNs in different cells) ought to be explored, e.g. by considering interactions between two progenitor cell types, each controlled by their own GRN. An alternative approach naturally lending itself to the analysis of spatially organized tissues and spatial heterogeneity is agent-based modeling. However, the methods that our framework relies on (tractability of the steady states and bifurcations) remain out of reach for most agent-based modeling approaches.

There is also much room for exploration of larger and more varied cell signaling network topologies. Here, future work should be guided by data, as it becomes harder to justify large signaling networks chosen *a priori*. Spatial transcriptomics (Moor and Itzkovitz, 2017) and new technologies (Giladi et al., 2020) will assist in the inference of cellular networks from data; inferred topologies could then be input to our model framework. These may include dissensus as well as consensus signals.

A central challenge for the model introduced here is that of fitting parameters to data. Ideally, this would require both spatially and temporally resolved single-cell transcriptomic data – at the limits of current technologies, although this is changing (Foreman and Wollman, 2020). Thus, in the current work we rely on comparison of qualitative features arising from the model with previous experimental studies. Moreover, due to the hybrid deterministic-stochastic formulation of the model, resulting in time-dependent signaling parameters, we doubt that it will be possible to derive a likelihood function for Bayesian parameter inference. Thus, approximate Bayesian computation will likely be required. Yet, even here, simulation times may be prohibitive or require further approximations to be made.

Future work should consider more thoroughly the stochastic nature of gene expression, e.g. by replacing the deterministic GRN dynamics with discrete stochastic simulation or stochastic differential equations (SDEs). This would be straightforward to accomplish numerically, although it will complicate the analysis of model bifurcations. Although the complete tractability of the deterministic model made it an ideal first candidate with which to study the effects of cell-cell communication, noise undoubtedly plays important roles in regulating single-cell phenotypes (Coomer et al., 2021).

In conclusion, the introduction of a single-cell resolved cell-cell communication model of GRN dynamics has helped to explain various cell-fate decision-making phenomena. Even in a tightly constrained model space, we have shown that changes in the distribution of cell fates due to cell-cell communication can be broad and varied. More generally, we have highlighted the need to consider multiscale effects in models of cell-fate dynamics. We anticipate that the application of these methods to other GRNs will lead to greater understanding of specific cell-fate decision-making control points, as well as general principles of control of stem cell differentiation.

## MATERIALS AND METHODS

### A multiscale model of cell-cell communication between single cells

To investigate how external signaling impacts intracellular dynamics during cell-fate decisions, we first must select an internal GRN topology to model. We choose to model a mutual inhibitory GRN: such network topologies arise frequently in cell-fate decision-making and developmental biology (Yu et al., 2018; Hong et al., 2015). Within certain parameter regions, such models give rise to two stable steady states and bistability (Fig. 1A). For example, it has been observed that the GRN determining the bifurcation of erythroid/megakaryocyte lineages from granulocyte/monocyte lineages contains a core mutual inhibitory loop topology consisting of mutual antagonism of *GATA1* (*G*) and *PU.1* (*P*). For this genetic switch, high expression of *G* corresponds to commitment to the erythroid/megakaryocyte lineage and high expression of *P* corresponds to commitment to the granulocyte/monocyte lineage.

This myeloid progenitor cell-lineage decision has been rigorously studied, and robust models exist of the intracellular GRN dynamics (Chickarmane et al., 2009; Duff et al., 2012; Chang et al., 2008; Huang et al., 2007; Roeder and Glauche, 2006; Strasser et al., 2012). Here, we implement the ODE model defined by Chickarmane et al., which follows the Shea-Ackers formalism for transcription factor dynamics (Shea and Ackers, 1985; Chickarmane et al., 2009). The ODE model is given by Eqns 1-3, where *X* is a transcription factor recruited by *GATA1* to bind to and inhibit *PU.1* expression. These ODEs give rise to two stable steady states for a region of parameter space with respect to the parameters *A*, *B* and *C*, where *A*, *B* and *C* are parameters summarizing environmental signals (Fig. 1B). The parameters through which the extracellular environment is summarized will be modified below to implement a multiscale model of cell-cell communication. The parameter values used throughout this work can be found in Table S1.

$$\frac{dG}{dt} = -\beta_1 G + \frac{\alpha_1 A + \alpha_2 G}{1 + \gamma_1 A + \gamma_2 G + \gamma_3 GP} \quad (1)$$

$$\frac{dP}{dt} = -\beta_2 P + \frac{\alpha_3 B + \alpha_4 P}{1 + \gamma_4 B + \gamma_5 P + \gamma_6 GP + \gamma_7 GX} \quad (2)$$

$$\frac{dX}{dt} = -\beta_3 X + \frac{\alpha_5 G}{1 + \gamma_8 G + \gamma_9 C} \quad (3)$$

To incorporate cell-cell signaling into the model, we must modify the cell-internal ODEs in a way that reflects the signal received by the cell without

unrealistically changing the internal dynamics (e.g. a loss of bistability). For simplicity, we choose a single parameter  $A$  to summarize the effects of external signaling. We study both signals that recruit nearby cells to commit to the same lineage and the opposite lineage of the sender. For the first case, we consider two cells, cell 1 and cell 2, where cell 1 signals to cell 2. Let  $G_1(t)$  and  $P_1(t)$  denote the concentrations of  $G$  and  $P$  in cell 1 at time  $t$ , respectively, and let  $A_2(t)$  be the value of the parameter  $A$  in cell 2 at time  $t$ . As cell 1 signals to cell 2 to coordinate fate decisions ('be like me' signal), then  $A_2(t)$  should increase if the ratio  $G_1(t):P_1(t)$  indicates that cell 1 has committed to – or is increasingly likely to commit to – the ME state (where  $G$  is highly expressed). Conversely,  $A_2(t)$  should decrease if the ratio  $G_1(t):P_1(t)$  indicates that cell 1 has committed to – or is increasingly likely to commit to – the GM state where  $P$  is highly expressed. To model the arrival process of the signals received by cells, we use a Poisson process. This choice is based upon previous results, which have demonstrated that the arrival of independent Brownian particles to an absorbing boundary is a Poisson process (Nadler et al., 2001). This result does assume that ligands are continuously produced and secreted from the 'sender' cell, which does not fully reflect the underlying biology, but we reason that it is an appropriate limiting assumption to allow for a simple, tractable signaling model. Based on these assumptions, we constructed a communication model: we treat signaling as a Poisson process, where the wait time between signals being sent by a cell follows an exponential random variable with mean  $\mu$ . After this wait time, at time  $t_s$ , a signal is sent and we change the value of  $A_2$  by:

$$A_2(t) = \begin{cases} \frac{G_1(t_s)}{\lambda + P_1(t_s)} A_2(t_s), & \text{for } t \in [t_s, t_s + \tau), \lambda > 0, \\ A_0, & \text{otherwise} \end{cases} \quad (4)$$

After a delay,  $\tau$ ,  $A_2$  returns to its original value,  $A_0$  (Fig. 1C). Returning to the same value  $A_0$  between signals ensures that attractor states (Fig. 1B) do not change. Sampled wait time values are rounded to integer values to avoid integration errors. In Fig. 1D is a schematic of this signaling model. Fig. S9 gives a sample simulation of a chain of four cells and the values of  $A_2, A_3$  and  $A_4$  plotted over time. Eqn 4 defines a signal between cells 1 and 2, but holds equally in the case of any signal between cells  $l$  and  $k$ .

In the case of signals that promote nearby cells to commit to the opposite lineage, we can follow the same principles to define a signal:

$$A_2(t) = \begin{cases} \frac{P_1(t_s)}{\kappa + G_1(t_s)} A_2(t_s), & \text{for } t \in [t_s, t_s + \tau), \kappa > 0. \\ A_0, & \text{otherwise} \end{cases} \quad (5)$$

These definitions of signaling also allow for multiple cells to signal a single cell at similar times, while still having  $A_2$  well defined. We implemented this model in Julia (Bezanson et al., 2017), where we used the DifferentialEquations.jl (Rackauckas and Nie, 2017) and Distributions.jl (Besançon et al., 2021) packages for numerical simulation.

The model developed here shares similarities with models specified by piecewise-deterministic Markov processes (PDMPs), first introduced by Davis (Davis, 1984). Such models comprise continuous dynamics and a discrete Markov process that can alter the continuous dynamics at discrete time points. PDMPs have previously been applied to biological systems for the study of gene expression and genetic networks (Rudnicki and Tyranskińska, 2017; Zeiser et al., 2008). In the case of two cells sharing one directed signal between them, the model described here closely resembles a PDMP, with the exception that here, due to information transfer between cells, the discrete process is not Markovian.

### Model specification and choice of parameters

To elucidate our choice of signaling model, we briefly discuss some alternatives. Examples of alternative signaling model choices include: signals that change more than one ODE parameter, additive rather than multiplicative signaling, or signals that permanently change ODE parameters. We choose to modify a single parameter for interpretability and to constrain the model space: the ODE system exhibits bifurcations with respect to many of its parameters, and we do not seek to explore bifurcations of co-dimension 2 or greater here. Modeling with additive signals brings

complications, such as the possibility of obtaining negative values of  $A$ . In addition, if a large number of committed cells are communicating with a single cell, the value of  $|A|$  may be unrealistically large and dominate the ODE. This problem is dealt with in the signaling model we present here through the parameter  $\lambda$ . Last, permitting signals to cause permanent changes to ODE parameters (rather than over time intervals) can result in the divergence of  $A$ . Thus, we have selected a signaling model that can capture how changes in the extracellular environment can alter cell internal GRN dynamics while still preserving the overall behaviors (i.e. the attractor states) of the dynamical system.

In this work, we will assume that cells are homogeneous, i.e. all cells share the initial conditions  $(G, P, X) = (20, 1, 20)$ . The internal GRNs are identical with the same parameter values, including the same initial value of  $A$ , denoted  $A_0$ . We set the signaling period to be  $\tau = 5$ , and the mean of the signaling wait time distribution to be  $\mu = 50$ . Changing the values of  $\mu$  and  $\tau$  does cause qualitative changes as long as  $\tau$  is sufficiently large relative to  $\mu$ . For more insight on the relationship between the wait time distribution and  $\tau$ , see Fig. S10. Unless otherwise specified, we set  $\lambda = \kappa = 1$ . We initially chose  $\lambda = \kappa = 1$  to avoid unrealistic behaviors when  $G$  or  $P \approx 0$ . We will provide an in depth discussion on the roles of these parameters in the next section. For all topologies depicted schematically in Fig. 3, regular arrowhead arrows correspond to consensus signaling (Eqn 4) and flat (inhibition) arrowhead arrows correspond to dissensus signaling (Eqn 5).

As a result of signaling, cell-fate decisions became probabilistic rather than deterministic. Then, for each cell in a signaling topology, we can examine the probability distribution of the cell converging to a certain lineage. To do so, for each signaling topology tested, we selected a range of values of  $A_0$  to simulate. The step size between the  $A_0$  values was scaled between each topology, depending on the range of the probability distribution. For each value of  $A_0$ , we simulated  $N$  trajectories, counting the number of times each cell converged to both fates. From these counts, we estimated the probability that each cell converged to the  $G$  high state and plotted:

$$P(G \text{ High}) \approx \frac{\text{Number of trajectories converging to } G \text{ high state}}{N}$$

A sample of the simulated data points can be found in Fig. S11. Unless stated otherwise, probabilities are estimated by running  $N = 1000$  simulations.

While  $\lambda = 1$  was an intuitive first choice to avoid divergence in signaling parameters, it relies on the assumption that a cell converging to the  $P$  high state implies that  $P$  is more highly expressed than  $G + 1$ , and vice versa. This turns out to not necessarily be true. Rather, for a given value of  $A_0$ , being in the  $P$  high state means that  $P$  is more highly expressed relative to the other stable steady state value of  $P$ . Recalling Fig. 1B, we see that for some values of  $A_0$ , e.g.  $A_0 = 0.8$ ,  $G$  is always more highly expressed than  $P$  regardless of which stable steady state a cell converges to. Looking at the bifurcation diagram, we see that in the  $P$  high state, the maximum steady state value of  $G$  is  $G^* \approx 16.97$  and the minimum value of  $P$  is  $P^* \approx 1.475$ . Similarly, in the  $G$  high state, the maximum steady state value of  $P$  is  $P^* \approx 0.3529$  and the minimum value of  $G$  is  $G^* \approx 40.35$ . From here, we wanted to select values of  $\lambda$  that satisfy the following:

$$\begin{aligned} \frac{G_{\{\max P \text{ High State}\}}}{\lambda + P_{\{\min P \text{ High State}\}}} &< 1, \\ \frac{G_{\{\min G \text{ High State}\}}}{\lambda + P_{\{\max G \text{ High State}\}}} &> 1, \\ \lambda &> 0. \end{aligned}$$

These inequalities are satisfied for  $\lambda \in [15.495, 39.997]$ . Similarly, for  $\kappa$ , in order to satisfy the analogous inequalities for the dissensus signal, we redefine the dissensus signal as

$$A_2(t) = \begin{cases} \frac{P_1(t_s) + \kappa}{G_1(t_s)} A_2(t_s), & \text{for } t \in [t_s, t_s + \tau), \kappa > 0, \\ A_0, & \text{otherwise} \end{cases} \quad (6)$$

where  $\kappa \in [15.495, 39.997]$ . We will explore how selecting values of  $\lambda$  in or near this range changes the behaviors of both individual cells and populations of cells.

### Intrinsic and extrinsic cell-cell communication noise

In this work, we focus on how intercellular signaling alters cell decision making. Therefore, we only consider noise with respect to signaling (not the internal GRN dynamics). Noise in the signal can arise due to the noisy extracellular environment (extrinsic noise) or can arise during the signal transduction within a cell (intrinsic noise) (Smith and Grima, 2018; Thurley et al., 2018). Here, we define models for both sources of noise.

For extrinsic noise, i.e. noise with respect to the extracellular environment, let  $\eta_e$  be a random variable such that  $\eta_e \sim \mathcal{N}(0, \sigma_e^2)$ , where  $\sigma_e^2 > 0$ . For cell  $k$  in a given topology, at the start of each new wait period, we sample a value of  $\eta_e$  and update  $A_k(t)$  by

$$A_k(t) = \max(0, A_0 + \eta_e). \quad (7)$$

We truncate  $A_k(t)$  at zero to avoid negative values. However, we select values of  $\sigma_e^2$  small enough that the truncation is rarely necessary in simulation. Note that noise in  $A_k(t)$  also results in noise during the signaling period according to Eqns 4 and 5.

For intrinsic noise, i.e. noise with respect to signal transduction, let  $\eta_p$  be a random variable such that  $\eta_p \sim \mathcal{N}(0, \sigma_p^2)$ , where  $\sigma_p^2 > 0$ . For cell  $k$  in a given topology, at the start of each new pulse period, we sample a new value of  $\eta_p$  and update  $A_k(t)$  as follows:

$$A_k(t_s) = \max\left(0, \frac{G_1(t_s)}{\lambda + P_1(t_s)} A_k(t_s) + \eta_p\right). \quad (8)$$

In the analysis of noise effects, we decrease the mean wait time to  $\mu=40$ , resulting in a larger signal-to-wait time ratio, allowing signals to have a greater influence on target cells. Doing so accentuates the effects of intrinsic and extrinsic noise because cells are spending more time in a pulse state.

### Acknowledgements

We thank Rong Lu and members of the Lu lab for many helpful discussions. We are grateful to Michael P. H. Stumpf for insightful comments and feedback on the manuscript.

### Competing interests

The authors declare no competing or financial interests.

### Author contributions

Conceptualization: M.K.R., A.L.M.; Methodology: M.K.R., A.L.M.; Software: M.K.R.; Investigation: M.K.R., A.L.M.; Writing - original draft: M.K.R., A.L.M.; Supervision: A.L.M.; Funding acquisition: A.L.M.

### Funding

This work was partially supported by a University of Southern California WiSE Top-up Fellowship (to M.K.R.) and by the National Science Foundation (DMS 2045327 to A.L.M.). Deposited in PMC for immediate release.

### Data availability

The model was developed in Julia, and all the code associated with simulation and analysis is available under an MIT license at: <https://github.com/maclean-lab/Cell-Cell-Communication>

### Peer review history

The peer review history is available online at <https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.199779>

### References

Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450-461. doi:10.1038/nrg2102

Baryawno, N., Przybylski, D., Kowalczyk, M. S., Kfoury, Y., Severe, N., Gustafsson, K., Kokkalis, K. D., Mercier, F., Tabaka, M., Hofree, M. et al. (2019). A cellular taxonomy of the bone marrow stroma in homeostasis and leukemia. *Cell* **177**, 1915-1932.e16. doi:10.1016/j.cell.2019.04.040

Besançon, M., Anthoff, D., Arslan, A., Byrne, S., Lin, D., Papamarkou, T. and Pearson, J. (2021). Distributions.jl: definition and modeling of probability distributions in the julia ecosystem. *Environ. Sci. Technol.* **55**, 1699-1709. doi:10.18637/jss.v098.i16

Bezanson, J., Edelman, A., Karpinski, S. and Shah, V. B. (2017). Julia: a fresh approach to numerical computing. *SIAM Rev.* **59**, 65-98. doi:10.1137/141000671

Brewer, C., Chu, E., Chin, M. and Lu, R. (2016). Transplantation dose alters the differentiation program of hematopoietic stem cells. *Cell Reports* **15**, 1848-1857. doi:10.1016/j.celrep.2016.04.061

Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E. and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**, 544-547. doi:10.1038/nature06965

Chickarmane, V., Enver, T. and Peterson, C. (2009). Computational modeling of the hematopoietic erythroid-myeloid switch reveals insights into cooperativity, priming, and irreversibility. *PLoS Comput. Biol.* **5**, e1000268. doi:10.1371/journal.pcbi.1000268

Christodoulou, C., Spencer, J. A., Yeh, S.-C. A., Turcotte, R., Kokkalis, K. D., Panero, R., Ramos, A., Guo, G., Seydhasantehrani, N., Esipova, T. V. et al. (2020). Live-animal imaging of native haematopoietic stem and progenitor cells. *Nature* **578**, 278-283. doi:10.1038/s41586-020-1971-z

Coomer, M. A., Ham, L. and Stumpf, M. P. H. (2021). Noise distorts the epigenetic landscape and shapes cell-fate decisions. *Cell Systems* **13**, 1-20. doi:10.1016/j.cels.2021.09.002

Craciun, G., Tang, Y. and Feinberg, M. (2006). Understanding bistability in complex enzyme-driven reaction networks. *Proc. Natl. Acad. Sci. USA* **103**, 8697-8702. doi:10.1073/pnas.0602767103

Davis, M. H. (1984). Piecewise-deterministic markov processes: a general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society: Series B (Methodological)* **46**, 353-388. doi:10.1111/j.2517-6161.1984.tb01308.x

Duff, C., Smith-Miles, K., Lopes, L. and Tian, T. (2012). Mathematical modelling of stem cell differentiation: the pu. 1-gata-1 interaction. *J. Math. Biol.* **64**, 449-468. doi:10.1007/s00285-011-0419-3

Ellison, D., Mugler, A., Brennan, M. D., Lee, S. H., Huebner, R. J., Shamir, E. R., Woo, L. A., Kim, J., Amar, P., Nemenman, I. et al. (2016). Cell-cell communication enhances the capacity of cell ensembles to sense shallow gradients during morphogenesis. *Proc. Natl. Acad. Sci. USA* **113**, E679-E688. doi:10.1073/pnas.1516503113

Ferrell, J. E. (2002). Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr. Opin. Cell Biol.* **14**, 140-148. doi:10.1016/S0955-0674(02)00314-9

Filippi, S., Barnes, C. P., Kirk, P. D., Kudo, T., Kunida, K., McMahon, S. S., Tsuchiya, T., Wada, T., Kuroda, S. and Stumpf, M. P. (2016). Robustness of MEK-ERK dynamics and origins of cell-to-cell variability in MAPK signaling. *Cell Reports* **15**, 2524-2535. doi:10.1016/j.celrep.2016.05.024

Foreman, R. and Wollman, R. (2020). Mammalian gene expression variability is explained by underlying cell state. *Mol. Syst. Biol.* **16**, e9146. doi:10.15252/msb.20199146

Giladi, A., Cohen, M., Medaglia, C., Baran, Y., Li, B., Zada, M., Bost, P., Blecher-Gonen, R., Salame, T.-M., Mayer, J. U. et al. (2020). Dissecting cellular crosstalk by sequencing physically interacting cells. *Nat. Biotechnol.* **38**, 629-637. doi:10.1038/s41587-020-0442-2

Graf, T. and Enver, T. (2009). Forcing cells to change lineages. *Nature* **462**, 587-594. doi:10.1038/nature08533

Ham, L., Brackston, R. D. and Stumpf, M. P. (2020). Extrinsic noise and heavy-tailed laws in gene expression. *Phys. Rev. Lett.* **124**, 108101. doi:10.1103/PhysRevLett.124.108101

Handly, L., Pilko, A. and Wollman, R. (2015). Paracrine communication maximizes cellular response fidelity in wound signaling. *eLife* **4**, e09652. doi:10.7554/eLife.09652

Hilfinger, A. and Paulsson, J. (2011). Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12167-12172. doi:10.1073/pnas.1018832108

Hong, T., Watanabe, K., Ta, C. H., Villarreal-Ponce, A., Nie, Q. and Dai, X. (2015). An Ovol2-Zeb1 mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states. *PLoS Comput. Biol.* **11**, e1004569. doi:10.1371/journal.pcbi.1004569

Hoppe, P. S., Schwarzfischer, M., Loeffler, D., Kokkalis, K. D., Hilsenbeck, O., Moritz, N., Ende, M., Filipczyk, A., Gambardella, A., Ahmed, N. et al. (2016). Early myeloid lineage choice is not initiated by random pu. 1 to gata1 protein ratios. *Nature* **535**, 299-302. doi:10.1038/nature18320

Hérault, A., Binnewies, M., Leong, S., Calero-Nieto, F. J., Zhang, S. Y., Kang, Y.-A., Wang, X., Pietras, E. M., Chu, S. H., Barry-Holson, K. et al. (2017). Myeloid progenitor cluster formation drives emergency and leukaemic myelopoiesis. *Nature* **544**, 53-58. doi:10.1038/nature21693

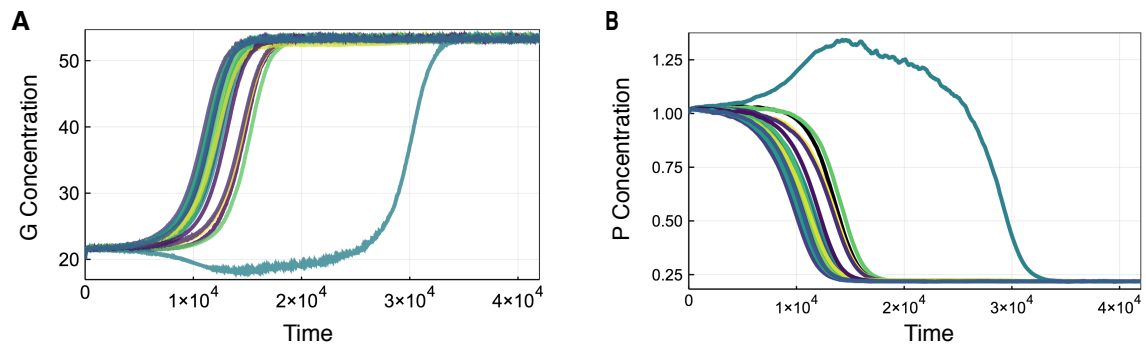
Huang, S., Guo, Y. P., May, G. and Enver, T. (2007). Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695-713. doi:10.1016/j.ydbio.2007.02.036

Iber, D. and Menshykau, D. (2013). The control of branching morphogenesis. *Open Biol.* **3**, 130088. doi:10.1098/rsob.130088

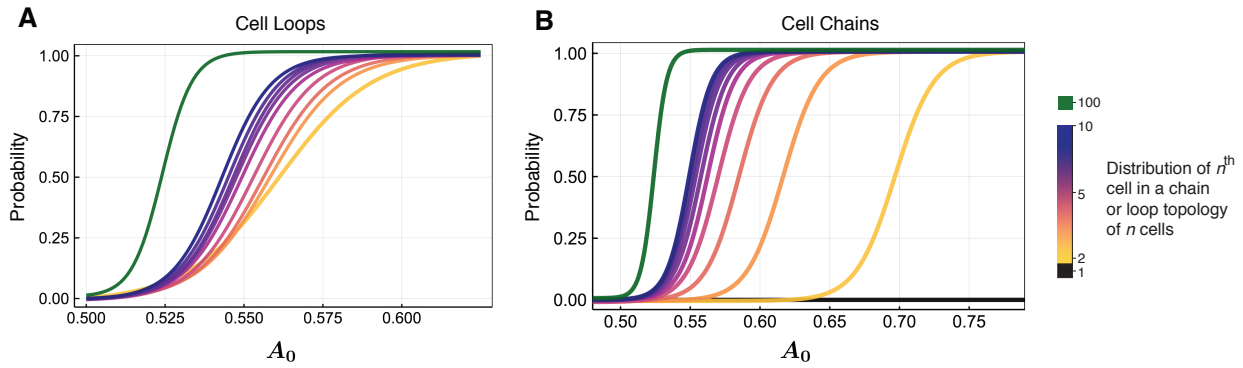
Kamenz, J., Gelens, L. and Ferrell, J. E. (2021). Bistable, biphasic regulation of PP2A-B55 accounts for the dynamics of mitotic substrate phosphorylation. *Curr. Biol.* **31**, 794-808. doi:10.1016/j.cub.2020.11.058

Lambert, B., MacLean, A. L., Fletcher, A. G., Combes, A. N., Little, M. H. and Byrne, H. M. (2018). Bayesian inference of agent-based models: a tool for

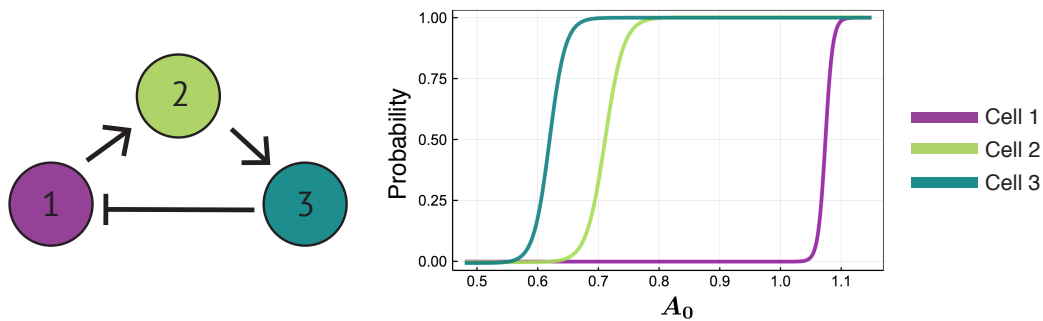
- studying kidney branching morphogenesis. *J. Math. Biol.* **76**, 1673-1697. doi:10.1007/s00285-018-1208-z
- Laurenti, E. and Göttgens, B.** (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**, 418-426. doi:10.1038/nature25022
- Lo Celso, C. and Scadden, D. T.** (2011). The haematopoietic stem cell niche at a glance. *J. Cell Sci.* **124**, 3529-3535. doi:10.1242/jcs.074112
- McLennan, R., Schumacher, L. J., Morrison, J. A., Teddy, J. M., Ridenour, D. A., Box, A. C., Semerad, C. L., Li, H., McDowell, W., Kay, D. et al.** (2015). Vegf signals induce trailblazer cell identity that drives neural crest migration. *Dev. Biol.* **407**, 12-25. doi:10.1016/j.ydbio.2015.08.011
- Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J. et al.** (2017). Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. *Cell Reports* **20**, 1215-1228. doi:10.1016/j.celrep.2017.07.009
- Mojtahedi, M., Skupin, A., Zhou, J., Castaño, I. G., Leong-Quong, R. Y., Chang, H., Trachana, K., Giuliani, A. and Huang, S.** (2016). Cell fate decision as high-dimensional critical state transition. *PLoS Biol.* **14**, e2000640. doi:10.1371/journal.pbio.2000640
- Moor, A. E. and Itzkovitz, S.** (2017). Spatial transcriptomics: paving the way for tissue-level systems biology. *Curr. Opin. Biotechnol.* **46**, 126-133. doi:10.1016/j.copbio.2017.02.004
- Moris, N., Pina, C. and Arias, A. M.** (2016). Transition states and cell fate decisions in epigenetic landscapes. *Nat. Rev. Genet.* **17**, 693-703. doi:10.1038/nrg.2016.98
- Morrison, S. J. and Scadden, D. T.** (2014). The bone marrow niche for haematopoietic stem cells. *Nature* **505**, 327-334. doi:10.1038/nature12984
- Mugler, A., Levchenko, A. and Nemenman, I.** (2016). Limits to the precision of gradient sensing with spatial communication and temporal integration. *Proc. Natl Acad. Sci. USA* **113**, E689-E695. doi:10.1073/pnas.1509597112
- Müller, E., Grinenko, T., Pompe, T., Waskow, C. and Werner, C.** (2015). Space constraints govern fate of hematopoietic stem and progenitor cells in vitro. *Biomaterials* **53**, 709-715. doi:10.1016/j.biomaterials.2015.02.095
- Nadler, B., Naeh, T. and Schuss, Z.** (2001). The stationary arrival process of independent diffusers from a continuum to an absorbing boundary is poissonian. *SIAM J. Appl. Math.* **62**, 433-447. doi:10.1137/S0036139900372363
- Ohishi, K., Katayama, N., Shiku, H., Varnum-Finney, B. and Bernstein, I.** (2003). Notch signalling in hematopoiesis. *Semin. Cell Dev. Biol.* **14**, 143-150. doi:10.1016/S1084-9521(02)00183-0
- Pal, M., Ghosh, S. and Bose, I.** (2014). Non-genetic heterogeneity, criticality and cell differentiation. *Phys. Biol.* **12**, 016001. doi:10.1088/1478-3975/12/1/016001
- Rackauckas, C. and Nie, Q.** (2017). Differentialequations.jl—a performant and feature-rich ecosystem for solving differential equations in julia. *J. Open Res. Soft.* **5**, 15. doi:10.5334/jors.151
- Roeder, I.** (2006). Quantitative stem cell biology: computational studies in the hematopoietic system. *Curr. Opin Hematol.* **13**, 222-228. doi:10.1097/O1.moh.0000231418.08031.48
- Roeder, I. and Glauche, I.** (2006). Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors GATA-1 and PU.1. *J. Theor. Biol.* **241**, 852-865. doi:10.1016/j.jtbi.2006.01.021
- Rudnicki, R. and Tyran-Kamińska, M.** (2017). *Piecewise Deterministic Processes in Biological Models*, vol. 1. Springer.
- Scholes, N. S., Schnoerr, D., Isalan, M. and Stumpf, M. P.** (2019). A comprehensive network atlas reveals that Turing patterns are common but not robust. *Cell systems* **9**, 243-257. doi:10.1016/j.cels.2019.07.007
- Shea, M. and Ackers, G.** (1985). The or control system of bacteriophage lambda. a physical-chemical model for gene regulation. *J. Mol. Biol.* **181**, 211-230. doi:10.1016/0022-2836(85)90086-5
- Smith, S. and Grima, R.** (2018). Single-cell variability in multicellular organisms. *Nat. Commun.* **9**, 345. doi:10.1038/s41467-017-02710-x
- Strasser, M., Theis, F. J. and Marr, C.** (2012). Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression. *Biophys. J.* **102**, 19-29. doi:10.1016/j.bpj.2011.11.4000
- Swain, P. S., Elowitz, M. B. and Siggia, E. D.** (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12795-12800. doi:10.1073/pnas.162041399
- Thurley, K., Wu, L. F. and Altschuler, S. J.** (2018). Modeling cell-to-cell communication networks using response-time distributions. *Cell systems* **6**, 355-367. doi:10.1016/j.cels.2018.01.016
- Tyson, J. J., Chen, K. C. and Novak, B.** (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.* **15**, 221-231. doi:10.1016/S0955-0674(03)00017-6
- Upadhaya, S., Krichevsky, O., Akhmetzyanova, I., Sawai, C. M., Fooksman, D. R. and Reizis, B.** (2020). Intravital imaging reveals motility of adult hematopoietic stem cells in the bone marrow niche. *Cell Stem Cell* **27**, 336-345.e4. doi:10.1016/j.stem.2020.06.003
- Wang, J., Tu, C., Zhang, H., Huo, Y., Menu, E. and Liu, J.** (2021). Single-cell analysis at the protein level delineates intracellular signaling dynamic during hematopoiesis. *BMC Biol.* **19**, 201. doi:10.1186/s12915-021-01138-6
- Yu, P., Nie, Q., Tang, C. and Zhang, L.** (2018). Nanog induced intermediate state in regulating stem cell differentiation and reprogramming. *BMC Syst. Biol.* **12**, 22. doi:10.1186/s12918-018-0552-3
- Zarubin, T. and Han, J.** (2005). Activation and signaling of the p38 map kinase pathway. *Cell Res.* **15**, 11-18. doi:10.1038/sj.cr.7290257
- Zeiser, S., Franz, U., Wittich, O. and Liebscher, V.** (2008). Simulation of genetic networks modelled by piecewise deterministic markov processes. *IET Syst. Biol.* **2**, 113-135. doi:10.1049/iet-syb:20070045



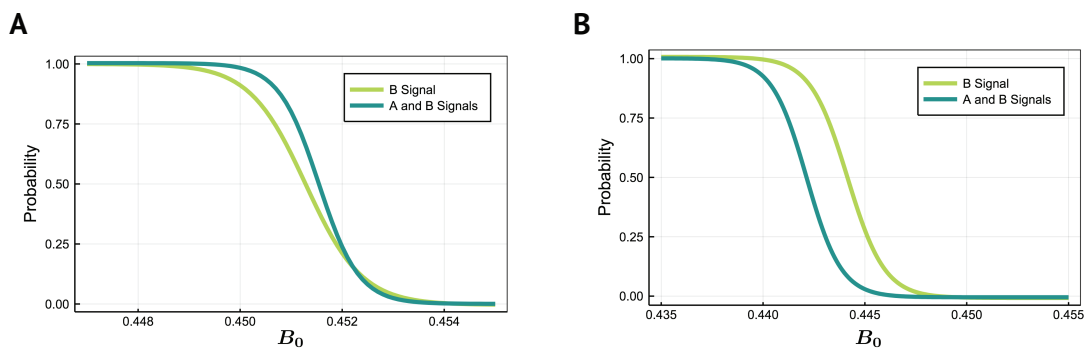
**Fig. S1.** Sample simulation of a loop of 20 cells with concentrations of both *G* and *P* over time. Parameters are set to  $\lambda = 28.0$  and  $A_0 = 1.0$ .



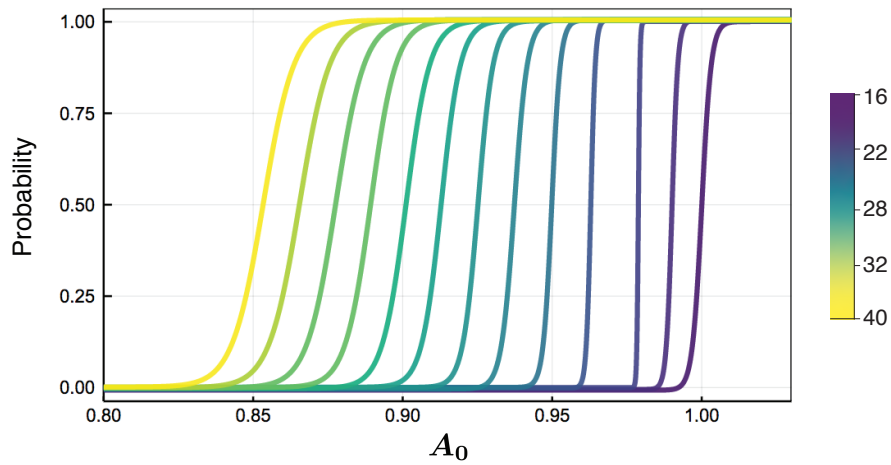
**Fig. S2.** Chains and loops of cells from Fig. 3A-3B along with curves where the number of cells  $n = 100$ . The number of simulation iterations was reduced to 100 per  $A_0$  value in simulations of 100 cell topologies.



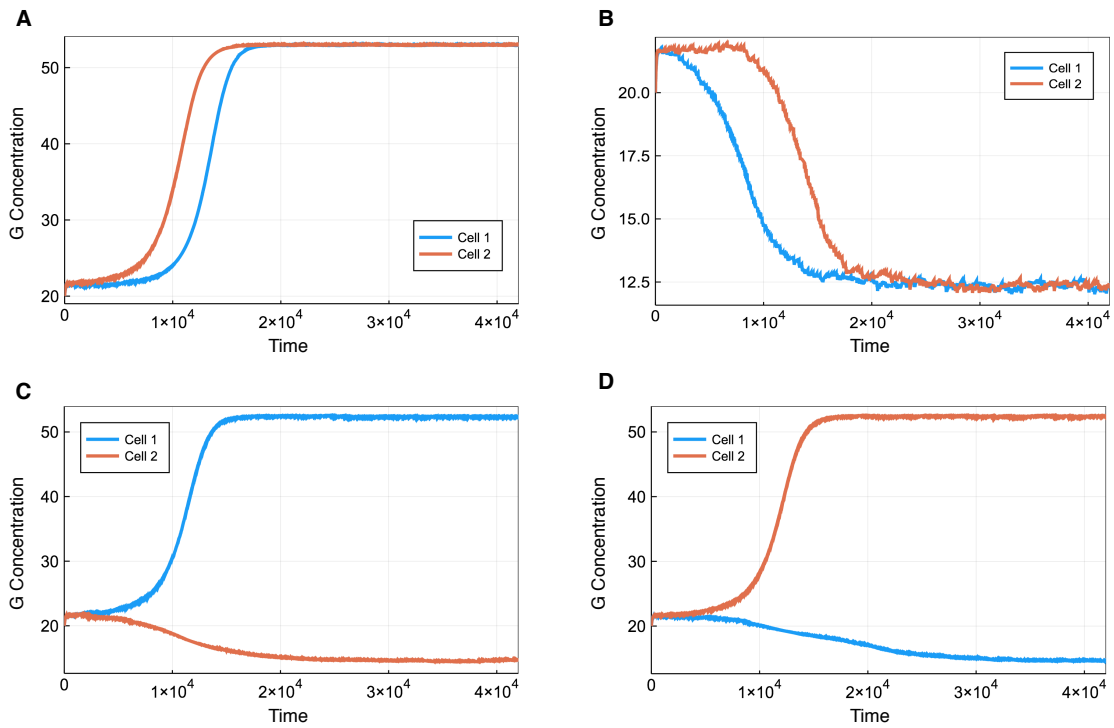
**Fig. S3.** Cell fate probability distributions for a loop of three cells with one inverse signal.



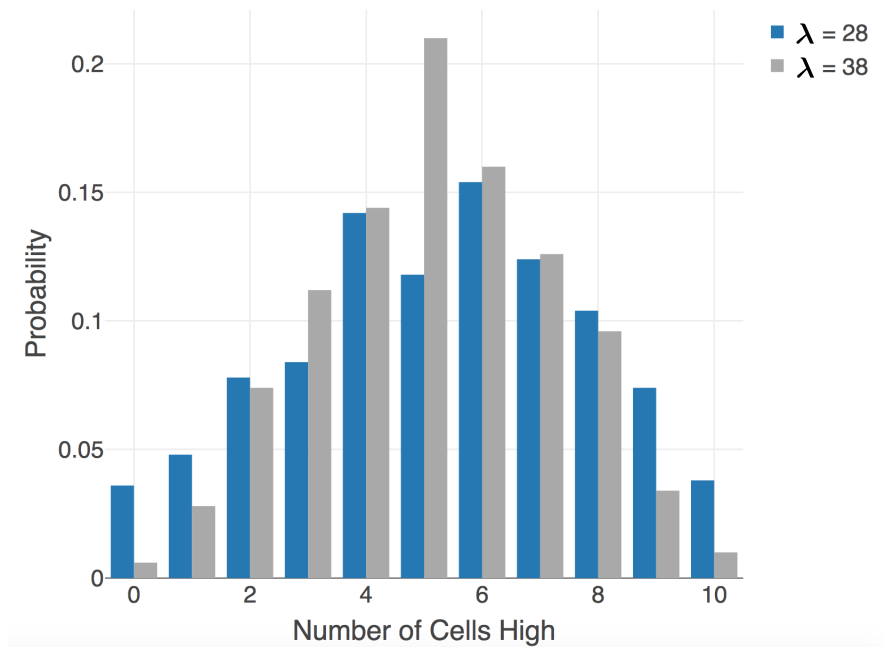
**Fig. S4.** Probability distributions of (A) cell two in a chain and (B) two cell loop converging to the G high state resulting from a consensus signal to the parameter  $B$  and dual consensus signals to both parameters  $A$  and  $B$ .  $A_0$  is fixed as 0.65. The signaling parameters for the signals changing  $A$  and  $B$  are  $\lambda_A = 28.0$  and  $\lambda_B = 28.0$  respectively.



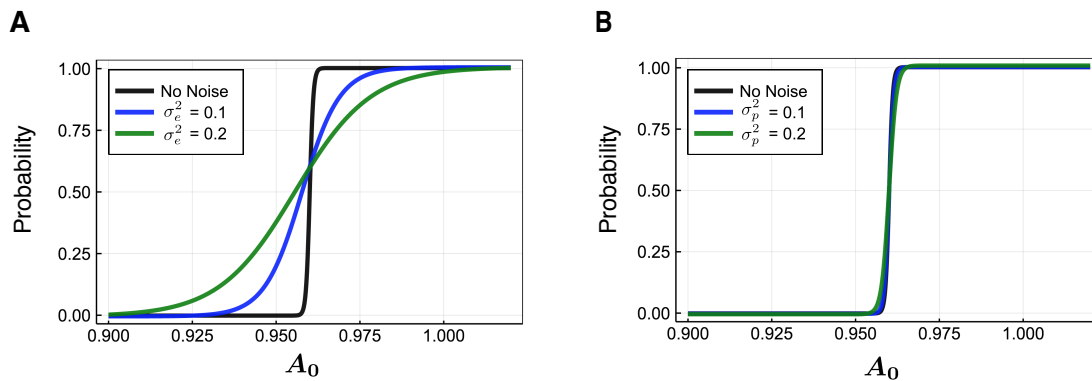
**Fig. S5.** Probability distributions of cell fate commitment to  $G$  high steady state for cell 2 in a chain of two cells with a dissensus signal where  $\kappa \in [16, 40]$ .



**Fig. S6.** (A)-(D) Sample trajectories of a loop of two cells with  $A_0 = 1.0175$  and  $\lambda = 40$ .

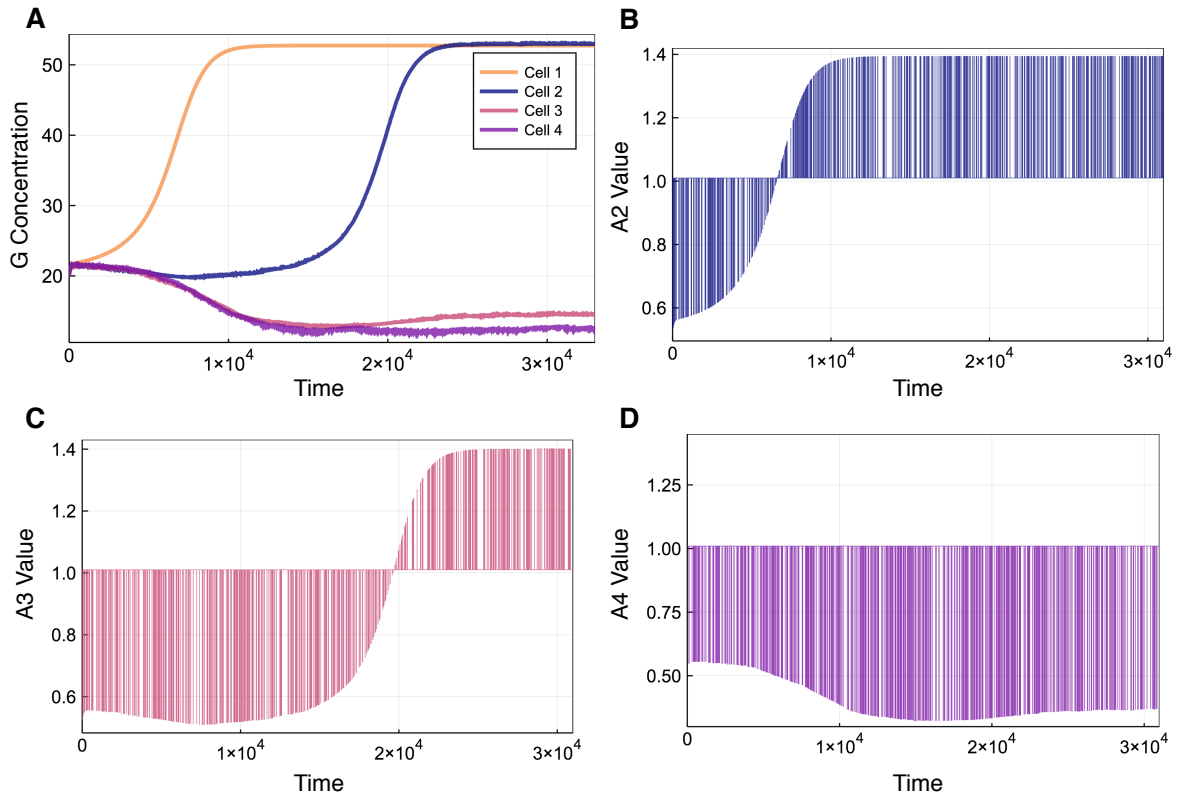


**Fig. S7.** Distributions of loops of 10 cells with different values of  $\lambda$ . The values of  $A_0$  were selected so that the distributions had similar expected values. For the blue distribution,  $A_0 = 0.9973$  and the expected value is 5.284. For the grey distribution,  $A_0 = 1.016$  and the expected value is 5.154

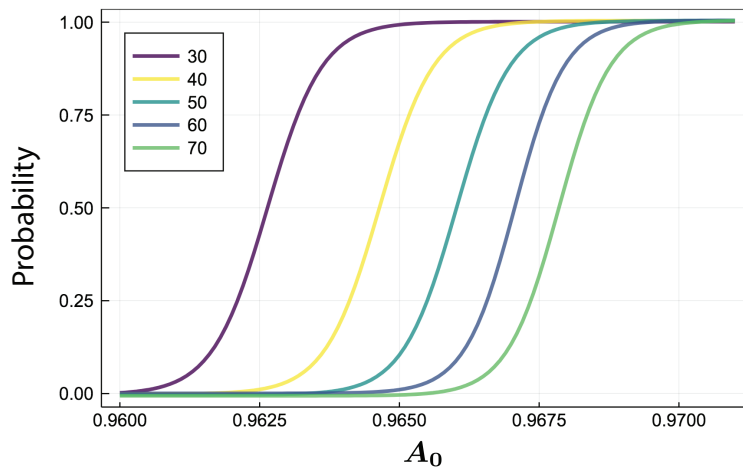


**Fig. S8.** Probability distributions that a given cell in a two cell loop will converge to the erythroid state with varying amounts of (A) extrinsic noise or (B) intrinsic noise.

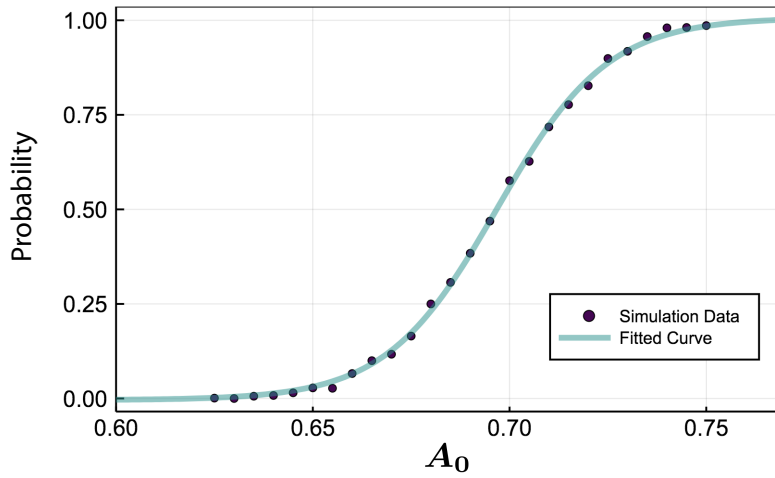




**Fig. S9.** (A) Sample trajectory of a chain of four cells where  $A_0 = 1.01$  and  $\lambda = 38.0$ . (B)-(D) Plots of  $A_2(t)$ ,  $A_3(t)$ , and  $A_4(t)$  over the same timescale as the trajectory.



**Fig. S10.** Probability distributions of cell 2 in a chain of cells converging to the  $G$  high state where  $\lambda = 18$  with different mean wait times,  $\mu$ , given in the legend.



**Fig. S11.** Sample of simulated data points along with the fitted curve for cell 2 in a chain of cells where  $\lambda = 1$ .

**Table S1.** Parameter values used in the ODE system.

$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\beta_1$	$\beta_2$	$\beta_3$	$\gamma_1$	$\gamma_2$
1.0	0.25	1.0	0.25	0.01	0.01	0.01	0.01	1.0	0.25
$\gamma_3$	$\gamma_4$	$\gamma_5$	$\gamma_6$	$\gamma_7$	$\gamma_8$	$\gamma_9$	$B$	$C$	
1.0	1.0	0.25	1.0	0.13	0.01	10.0	0.5	0	