

Control of Stochastic and Induced Switching in Biophysical Networks

Daniel K. Wells,^{1,2} William L. Kath,^{1,2,3} and Adilson E. Motter^{2,3,4,*}

¹*Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, Illinois 60208, USA*

²*Northwestern Physical Sciences-Oncology Center, Northwestern University, Evanston, Illinois 60208, USA*

³*Northwestern Institute on Complex Systems, Northwestern University, Evanston, Illinois 60208, USA*

⁴*Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60208, USA*

(Received 23 September 2014; revised manuscript received 9 July 2015; published 16 September 2015)

Noise caused by fluctuations at the molecular level is a fundamental part of intracellular processes. While the response of biological systems to noise has been studied extensively, there has been limited understanding of how to exploit it to induce a desired cell state. Here we present a scalable, quantitative method based on the Freidlin-Wentzell action to predict and control noise-induced switching between different states in genetic networks that, conveniently, can also control transitions between stable states in the absence of noise. We apply this methodology to models of cell differentiation and show how predicted manipulations of tunable factors can induce lineage changes, and further utilize it to identify new candidate strategies for cancer therapy in a cell death pathway model. This framework offers a systems approach to identifying the key factors for rationally manipulating biophysical dynamics, and should also find use in controlling other classes of noisy complex networks.

DOI: [10.1103/PhysRevX.5.031036](https://doi.org/10.1103/PhysRevX.5.031036)

Subject Areas: Biological Physics, Complex Systems, Nonlinear Dynamics

I. INTRODUCTION

Cellular systems are not entirely deterministic, but are instead impacted by small, random fluctuations in the number and activity of molecules of intracellular species [1,2]. Such fluctuations lead to macroscopic effects in a diverse array of processes. In differentiation, the resulting noise plays a central role in cell fate determination and can allow clonal populations of differentiating cells to achieve distinct final states [3,4]. Noise can also produce spontaneous transitions, whereby it causes a system to switch from one stable state to another, often producing a significant change of phenotype or function. Such stochastic state switching occurs, for example, in the *lac* system, where rare, brief transcription events in the “off” state cause large bursts in LacY expression, which in turn can be amplified and stabilized by a positive feedback loop [5]. Stochastically induced transitions also underlie recent observations of spontaneous dedifferentiation in cancer cells [6,7], in which cancer stem cells arose *de novo* from nonstem cell populations.

The response to noise and the overall behavior of many biophysical systems are determined by an underlying epigenetic landscape [8]. In this landscape, the valleys

represent the distinct achievable states of the system and the heights of the separating barriers determine their robustness to noise. A benefit of this representation is that bifurcation points—locations in the parameter space at which one or more stable states suddenly cease to exist—correspond precisely to the points where one or more of such barriers first reach zero height as parameters change. This landscape thus incorporates two distinct features of a state, namely, its robustness to noise and its deterministic stability, into one: the less robust a state is to noise, the closer it is to being eliminated through a bifurcation, and vice versa.

The landscape representation has been given a quantitative foundation as the quasipotential of the deterministic component of the system dynamics [9] and has been explored in experiments, e.g., to show how two parameters in the yeast galactose signaling network, the concentrations of galactose and intracellular Gal80p, can alter the rates of stochastic switching in this bistable circuit [10]. Despite these advances, researchers’ ability to control this landscape in order to induce prespecified biological outcomes has been generally limited to at most two parameters [11,12], and no general method currently exists to systematically tune transitions between stable states and/or eliminate undesired states altogether. The possibility of such control would offer clear opportunities. For example, under the widely supported stochastic model for induced pluripotent stem cell generation [13], a majority of cells have the possibility of being reprogrammed, even though existing technologies have achieved substantially smaller yields [14].

*motter@northwestern.edu

Published by the American Physical Society under the terms of the *Creative Commons Attribution 3.0 License*. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

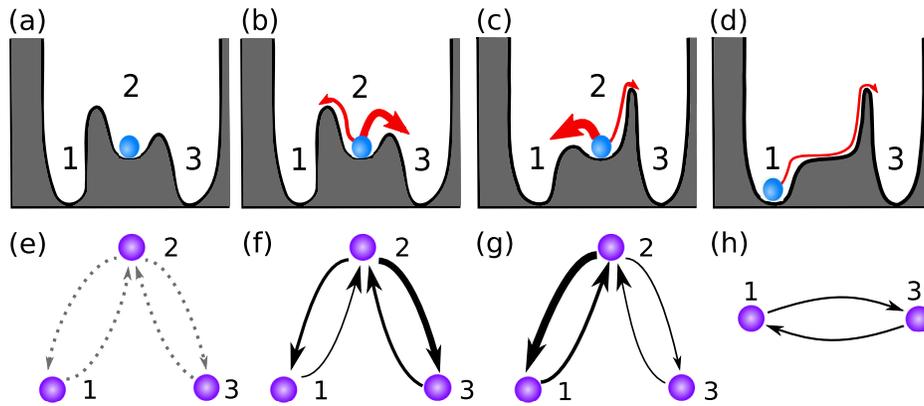


FIG. 1. Control of the response to noise illustrated for a multiwell quasipotential. (a) Without noise the state of the system is fixed and will not change over time. (b) In the presence of noise the state of the system wanders within the attraction basin and will be eventually ejected into the attraction basin of another stable state. For small noise such transitions are exponentially more likely to occur through the lowest barrier, even though other transitions are also possible in principle. (c) Using OLAC on a set of tunable parameters, we can alter the quasipotential to produce a desired response to noise, in this case tailored to increase transitions from state 2 to state 1 and to reduce transitions to state 3. (d) If desired, OLAC can also find combinations of tunable parameters that can alter the quasipotential in order to eliminate a stable state through a bifurcation; in this illustration state 2 is eliminated, in favor of state 1. (e)–(h) NEST for each of the quasipotentials in (a)–(d), respectively, where the nodes represent stable states and the continuous edges represent transition rates. A wider continuous edge indicates a higher transition rate. The dotted edges in (e) indicate transitions that could occur in the presence of noise.

The ability to control the response to noise of differentiated and stem cell states (e.g., inhibiting transitions to the first and promoting transitions to the second) could lead to enhanced procedures to create induced pluripotent stem cells. Similarly, in the context of the “cancer attractor” hypothesis [15,16], in which normal and cancer cells correspond to distinct coexisting stable states, identifying interventions that destabilize, or eliminate, the cancerous state could lead to new therapeutic strategies.

In this paper we propose a broadly applicable method, here termed optimal least action control (OLAC), that can predict and control the dynamical behavior and response to noise in a wide class of biophysical networks. As schematically illustrated in Fig. 1, the essence of our approach is that to control a biophysical system it is sufficient to identify interventions, e.g., changes to gene expression, protein levels, or interaction rates, that can reshape the topography of the underlying quasipotential in a desired way. This approach ultimately leads to a network of state transitions (NEST) describing the transitions between stable states and that can be controlled by changing the heights of the separating barriers without changing any quality of the noise. For a given system, this is achieved by determining the minimum action paths—those followed by the most likely noise-induced transition trajectories—and the corresponding transition rates between all pairs of stable states, and then optimizing these transition rates for a desired outcome. Furthermore, this general foundation in a physical least action principle allows OLAC to be applied broadly to many other complex networks as well. In particular, while we focus our application of OLAC to biophysical networks, applications to other networks where

noise and multistability play important roles, including power-grid networks [17], polymer networks [18], and food-web networks [19], among others, are immediate within the formulation we establish here.

We apply OLAC to several gene network models and illustrate how this method can be used to make biologically realizable reprogramming predictions. In the limit of zero noise intensity, OLAC automatically identifies bifurcations that eliminate undesirable states and induce purely deterministic transitions to the desired ones. The significance of the latter is demonstrated by considering a third application, to eliminate cancerous states in a cell death network model, which concerns a time scale for which stochastic switches can be neglected. As illustrated in these examples, the NEST is a powerful yet simple representation that captures the essence of the state switching dynamics and can inform counterintuitive results, e.g., the possibility of transitions through intermediate stable states when direct transitions are essentially impossible (a behavior observed even in high dimensions, where indirect transitions generally require longer paths). The method proposed here is easily implementable and the computational effort scales linearly in the number of control parameters and the dimension of the state space, allowing our approach to be applied to large networks and high-dimensional systems in general.

II. THEORY

A. Transition rates for small noise

We consider biophysical networks whose deterministic components are described by nonlinear differential equations of the form $d\vec{X}/dt = \vec{F}(\vec{X}; \Omega)$, where \vec{X} is a vector

representing the activity of the relevant biological factors, \vec{F} is the function representing the rates of change of these factors, and Ω is a set of tunable parameters, which we show can be manipulated to drive cellular processes in advantageous directions. We focus on the most prevalent case of systems with two or more stable states and, although our approach is general, for concreteness we first assume that these states are time independent. Time-independent stable states correspond to fixed points $\vec{X}^* = \vec{X}^*(\Omega)$, defined by $\vec{F}(\vec{X}^*; \Omega) = \vec{0}$, towards which neighboring trajectories converge over time. The set of all stable states represents the possible long-term behaviors of the deterministic system.

Stochasticity is modeled here as additive Gaussian white noise,

$$d\vec{X} = \vec{F}(\vec{X}; \Omega)dt + \sqrt{\varepsilon}d\vec{W}, \quad (1)$$

where ε is the variance of the distribution (other cases are discussed in the Supplemental Material [20]). With the addition of this small noise term, trajectories no longer approach the stable states asymptotically as in the deterministic case. Instead, a trajectory close to a stable state will oscillate stochastically within its basin of attraction, typically staying close to the fixed point for long periods of time. The trajectory will also make rare but large excursions from the stable state. After sufficiently long time an excursion large enough to eject the trajectory from the original basin of attraction will necessarily occur, at which point it will transition to the neighborhood of another stable state of the system. The time scales for the occurrence of such transitions may be shorter or longer than the biologically relevant ones. The manipulation of these time scales underlies much of the control approach introduced below.

For a given noise intensity ε , the transitions between two stable states i and j occur as a Poisson process with a certain rate $R_{i,j}^\varepsilon(\Omega)$. These rates can be computed by evolving Eq. (1), but, in general, at an unreasonably high computational cost. As a way to reduce this effort, we employ an asymptotic formula [21]:

$$\tilde{R}_{i,j}^\varepsilon(\Omega) \propto \exp\left(-\frac{1}{\varepsilon}S_{i,j}^*(\Omega)\right), \quad (2)$$

where $\tilde{R}_{i,j}^\varepsilon$ serves as an excellent approximation to $R_{i,j}^\varepsilon$ for ε small compared to $S_{i,j}^*$, which is typically the case for noise associated with biophysical systems. Here, $S_{i,j}^*$ is the minimum of the Freidlin-Wentzell action $S[\cdot]$, a functional over all possible transition paths $\vec{\phi}_{i,j}$ connecting the two stable states in the state space. The path $\vec{\phi}_{i,j}^*$ minimizing the action for the given pair of stable states is calculated numerically employing an implementation of the adaptive minimum action method [22], in which we determine this minimum action path through the optimization of

a discretized version of the action functional using a quasi-Newton method (see Appendix A). Conceptually, $S_{i,j}^*$ represents the cumulative height of all saddle points traversed by the minimum action path between the states i and j . It should be noted that a proportionality constant is omitted in the expression for $\tilde{R}_{i,j}^\varepsilon$. Although there is no known formula for computing this constant in general [23], the key condition for neglecting it is clear: as long as the actions $S_{i,j}$ associated with two paths differ by more than $\mathcal{O}(\varepsilon)$, the exponential term will dominate and the omission of the proportionality constant will not affect the result qualitatively.

The rate $\tilde{R}_{i,j}^\varepsilon$ represents the transition probability per unit of time along the most likely *direct* path(s) between the stable states i and j . The transition rates through intermediate stable states can be determined by composing these elementary transition rates. It is often expected that the most likely transition path between two stable states would be a direct path, but as we show below, in many cases it passes through intermediate stable states. Moreover, the transition paths and rates are generally asymmetric, with $\tilde{R}_{j,i}^\varepsilon \neq \tilde{R}_{i,j}^\varepsilon$. In the limit of long time, these transitions lead to an equilibrium probability distribution of occupied stable states, which we refer to as the limiting occupancy of the system and denote by \vec{v}^ε . Given that we generally study large populations of cells, in our applications it is convenient to interpret the limiting occupancy as an ensemble average over different realizations of noise.

To illustrate the minimum action paths and their use in controlling cell behavior, we first consider a two-dimensional model for the *Caenorhabditis elegans* vulval precursor cell (VPC) differentiation [24]. The VPC differentiation is representative of many other differentiation processes and enjoys significant experimental characterization. The model describes the differentiation of VPCs into one of three competent lineages, 1°, 2°, and 3°, corresponding to stable states in the system and marked as nodes in Fig. 2(a). These lineages depend on two dimensionless parameters, ℓ_1 and ℓ_2 , that determine the levels of epidermal growth factor (EGF) and Notch signaling, respectively (for the model equations, see Appendix B). It is known experimentally that increasing EGF and decreasing Notch signaling (relative to the base value of 0.1) will bias cells towards lineage 1°, that decreasing EGF and increasing Notch will bias cells towards lineage 2°, and that decreasing both EGF and Notch will bias cells towards lineage 3° [25]. Figure 2(a) shows the state space for the VPC system with equal, intermediate levels of EGF and Notch signaling ($\ell_1 = \ell_2 = 0.1$) and a realistic level of noise ($\varepsilon = 0.007$). The calculated transition paths, transition rates, and limiting occupancies are indicated by the edges, their width, and the size of the nodes, respectively. For these parameters, our calculations indicate that the limiting occupancy is comparable for all three stable states, with lineage 2° having the highest occupancy.

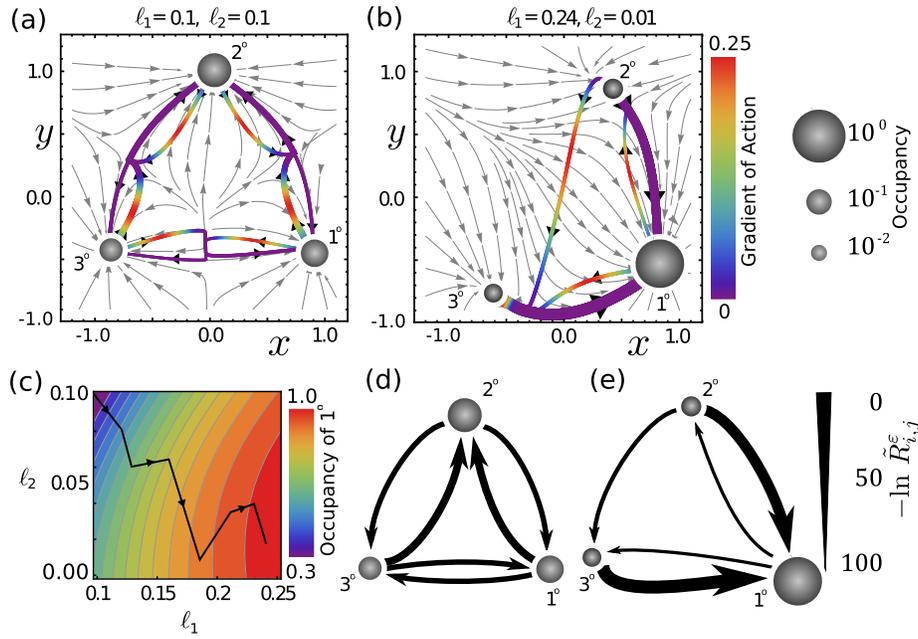


FIG. 2. OLAC applied to a two-dimensional model of VPC differentiation. (a), (b) State space representation of stable states (nodes) and optimal transition paths (edges) before intervention (a) and after OLAC is applied to maximize the limiting occupancy of lineage 1° (b). Node size indicates occupancy and edge width indicates the (negative of the) log transition rate along the corresponding minimum action path; the color code on the transition paths indicates the derivative of the quasipotential. The background shows the velocity field for the given parameters. For equal EGF and Notch signaling (a), lineage 2° is the one with highest occupancy. The OLAC solution (b) indicates that high EGF signaling and low Notch signaling will lead to the maximum occupancy of lineage 1° . (c) Trajectory in the parameter space for one realization of the optimization procedure, where the contour plot indicates the limiting occupancy of lineage 1° . Note that each step of the optimization routine increases the occupancy of this state. (d), (e) NESTs for the initial (d) and optimized (e) systems. In all panels, the noise intensity is assumed to be $\varepsilon = 0.007$ (as used previously [24]).

B. Optimal least action control

We now turn to the manipulation of the tunable parameters Ω , which may include, for example, gene expression levels, protein activity, and interaction rates. Specifically, we seek to modify Ω to optimize either specific transition rates $\tilde{R}_{i,j}^\varepsilon$ directly or the limiting occupancy \bar{v}^ε through the alteration of the transition rates, while recognizing that the parameters can be modified only within specific ranges determined by biophysical and experimental constraints. The latter include *sparsity* constraints, which we can use to effectively limit the number of targets in the control set while avoiding a combinatorial explosion (see Appendix C). The problem is formalized as the maximization of an objective function of interest G (unique to each problem) over the parameters Ω subject to the given constraints $\{g_k(\Omega) = \zeta_k\}_k$ and $\{h_k(\Omega) \leq \eta_k\}_k$:

$$\max_{\substack{\Omega \\ \{g_k(\Omega) = \zeta_k\}_k \\ \{h_k(\Omega) \leq \eta_k\}_k}} G(\{S_{i,j}^*(\Omega)\}; \varepsilon). \quad (3)$$

This procedure constitutes the central step of our implementation of OLAC. Note that this formulation is independent of the dimension and complexity of the system. Thus, given a well-defined model we can identify control interventions able to alter the switching behavior and/or

the stability of the states in desired ways without the need to know explicitly *a priori* how variations of the tunable factors affect the system (beyond the implicit dependence defined by the dynamical equations).

The objective function and the associated constraints do not need to be linear, allowing a wide range of possible dynamical behaviors to be optimized. For any set of such constraints, Eq. (3) can be solved numerically through a second quasi-Newton method step that nests the adaptive minimum action method used to determine the minimum action paths in connection with Eq. (2). Importantly, OLAC is highly scalable, with the following computational cost:

$$\text{cost} \sim \mathcal{O}[|\Omega|D(K + \gamma P)], \quad (4)$$

where D is the number of variables defining the state space, $|\Omega|$ is the number of tunable parameters under consideration, P is the number of stable states, and K is the number of transitions upon which G depends. This cost estimation follows from noting that every optimization step of OLAC relies on $|\Omega|$ evaluations of G , each of which requiring K runs of the nested optimization step, where each such run has a cost that is linear in D when using the L-BFGS optimization method [26]. Furthermore, for each top-level optimization step [Eq. (3)], every one of the P stable states

has to be continued $|\Omega|$ times, each time at an integration cost that is linear in D if the average network's degree is approximately constant, as in most network models (it would be at most quadratic in D in the most general case) (see Appendix A). Parameter γ accounts for other constants describing the relative cost between optimization and integration. Because of this high scalability, our method can be applied to complex high-dimensional multiparameter systems without excessive computational cost. In this way, OLAC expands on previous foundational work that demonstrated how barriers between stable states can be altered to control the response to fluctuations and multistability [27–32] to now address large networks with many variables and many potential control parameters.

Before turning to high-dimensional systems, we consider an illustrative example application of OLAC in which we maximize the final occupancy of lineage 1° in the VPC system. As an example constraint we stipulate that neither of the other two stable states be lost to bifurcation. This nonbifurcation constraint is intended to depict how the presence of dosing and/or experimental limitations may make complete elimination of undesired states infeasible. This condition is imposed as the constraint that the states $\vec{X}_k^* = \vec{X}_k^*(\Omega)$ representing each lineage k remain stable fixed points: $\vec{g}_k(\Omega) \equiv \vec{F}(\vec{X}_k^*, \Omega) = \vec{0}$ and $h_k(\Omega) \equiv \text{Re}(\lambda_k) \leq -\tau_0$. In this notation, $\lambda_k = \lambda_k(\Omega)$ is the eigenvalue of the Jacobian matrix $D\vec{F}(\vec{X}, \Omega)|_{\vec{X}_k^*}$ with largest real part, and τ_0 is a tolerance (set to be 0.1 in our simulations). The objective function is $G(\Omega; \varepsilon) = v_{1^\circ}^{\varepsilon}(\Omega)$ and the tunable parameters are $\Omega = \{\ell_1, \ell_2\}$. The result, shown in Fig. 2(b), indicates a substantial (threefold) increase in the limiting occupancy of the lineage 1° state for the optimal control intervention identified by OLAC. The parameter-space path for a representative realization of the optimization is shown in Fig. 2(c). The optimal intervention, defined as an average over multiple realizations, is a combination of increased EGF signaling ($\ell_1: 0.1 \rightarrow 0.24$) and reduced Notch signaling ($\ell_2: 0.1 \rightarrow 0.01$), which has the net effect of lowering the barrier for transitions from lineages 2° and 3° to lineage 1° while maintaining a high barrier for exiting this lineage. This controlled state corresponds to signaling strengths that have been observed experimentally to indeed bias cell lineages towards lineage 1° [25]. As mentioned above, for these results we utilize Eq. (2) with the proportionality constants omitted to calculate transition rates between states. The inclusion of these prefactors could potentially alter the occupancy calculations of the stable states and in turn invalidate the results of OLAC. In the Supplemental Material [20], Sec. S1, we apply sensitivity analysis to quantify the uncertainty associated with omitting prefactors and show that doing so does not significantly alter the results in this case. Therein we also discuss more generally under which conditions prefactors can be omitted.

C. Network of state transitions

Our formulation leads to a succinct and intuitive NEST representation for the transition dynamics, in which the nodes are unique stable states and the weighted, directed edges between nodes represent the rates of transition between the stable states. The basis for this representation is the observation that, for small ε , the trajectories of the system in Eq. (1) are most often close to a stable state or transitioning between stable states along a minimum action path. The trajectories will rarely venture into other regions of the state space, allowing us to focus on only this “waiting-transition” dynamics without sacrificing information about the system's behavior. For a system with P stable states we can write the $P \times P$ transition matrix

$$\tilde{\mathbf{R}}_{i,j}^{\varepsilon} = \begin{cases} \tilde{R}_{i,j}^{\varepsilon} & i \neq j \\ -\sum_j \tilde{R}_{i,j}^{\varepsilon} & i = j, \end{cases} \quad (5)$$

which defines a (continuous-time) Markov process on the stable states of the system. In our numerical calculations, we use the fact that the limiting occupancy is described by the equilibrium solution of this process. In fact, the Markov process specifies all attributes of the associated NEST. The NEST is conceptually similar to the state transition network used in physical chemistry and biochemistry [33–36], as well as to transition networks studied in mathematics [21]. There are, however, key differences between our approach and those considered previously, in addition to the fact that we focus on network systems. In particular, the NEST does not assume the existence of a potential energy landscape and is defined for nonvanishing levels of noise, which requires us to develop a new formulation that accounts, in particular, for long mixing times and for values of ε larger than $S_{i,j}^*$ (see Supplemental Material [20], Sec. S4). Furthermore, unlike static state transition networks, the transition rates in the NEST are malleable and can be rationally manipulated with OLAC.

For the VPC model, the NESTs corresponding to the unmodified signaling strengths [Fig. 2(a)] and to the signaling strengths that optimize lineage 1° occupancy [Fig. 2(b)] are shown in Figs. 2(d) and 2(e), respectively. These networks represent a substantial distillation of the dynamics of the underlying biophysical system and can be used to simplify and explain the transition dynamics in a high-dimensional system without the need to consider its entire state space. In particular, for the range of edge widths shown, the optimized NEST in Fig. 2(e) has edges between all nodes except from lineage 3° to lineage 2°, indicating that a direct transition between these two states is highly unlikely, whereas an indirect transition is possible; indeed, the two-step transition $3^\circ \rightarrow 1^\circ \rightarrow 2^\circ$ has an overwhelmingly higher rate (10^2 times higher). By comparing with the NEST of the original system, shown in Fig. 2(d), this also

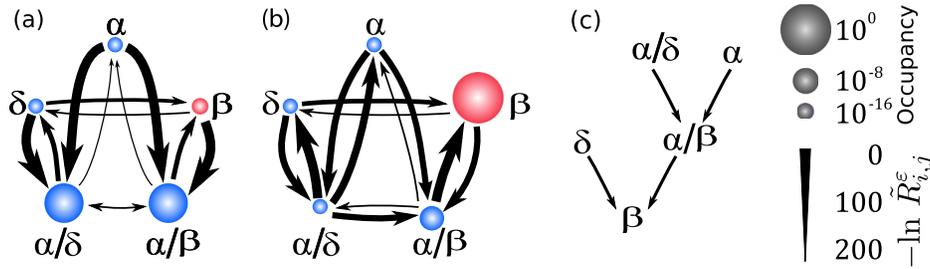


FIG. 3. Controlled stochastic lineage switching in a multidimensional model of HPC differentiation. (a), (b) NESTs of the model for the unmodified parameters (a) and for OLAC applied to optimize the limiting occupancy of the β cell state (red) (b), for $\epsilon = 0.01$. In both panels, node size indicates the limiting occupancy of the state and edge width indicates the (negative of the) log transition rate. OLAC identifies that only three (out of ten) transcription factors need to be tuned in the model to optimize β -cell occupancy: MafA (increased), Brn4 (decreased), and δ -gene (decreased). (c) Transition hierarchy into the β cell state for the optimized system. Two states (δ and α/β) transition directly; two others (α and α/δ) require passage through an intermediate state—in this case α/β .

demonstrates that direct transitions for one parameter regime can become indirect for another.

The OLAC method can be implemented directly on the NEST representation. Indeed, the objective function $G(\{S_{i,j}^*\}; \epsilon)$ is naturally defined in terms of the transition matrix of the system, $\tilde{\mathbf{R}}^\epsilon$. As our application to the VPC system shows, the combined effect of optimizing this objective function is to vary the height of the transition barrier along the minimum action path between the stable states. This shows that OLAC itself does not require determining the full quasipotential of the system, which would be computationally prohibitive in high dimensions.

III. APPLICATIONS

A. Controlling pancreas cell transdifferentiation

An important research problem in cellular reprogramming concerns the induction of insulin-producing pancreatic β cells from noninsulin-producing cell lineages; interventions capable of achieving this goal could lead to new treatments for type I diabetes. In order to computationally identify an optimal intervention to induce the desired reprogramming, we consider a ten-dimensional model of the hierarchical pancreas cell (HPC) differentiation [37]. The model has five stable fixed points—three representing differentiated endocrine pancreas cell types (α , β , and δ) and two representing intermediate states (α/β and α/δ). The expression level of the ten regulatory genes $\{x_i\}$ are each assumed to be tunable independently through a factor σ_i (details of the model are given in Appendix B).

We first consider the uncontrolled model, in which $\sigma_i = 1$ for $i = 1, \dots, 10$. As shown in Fig. 3(a), in this case the two intermediate states attract the majority ($>99\%$) of the occupancy. Furthermore, transitions to the β state from the α and δ states occur at negligibly small rates ($<10^{-100}$), indicating that such lineage respecification effectively never occurs spontaneously. We apply OLAC to this model in order to identify the optimal combination of control actions that maximize the occupancy of the β

state v_β^ϵ ; the admissible interventions are limited to those for which no bifurcations occur, which is imposed using the same constraints as in the previous example. The optimal intervention that maximizes v_β^ϵ is a three-gene one, consisting of the downregulation of Brn4 and δ -gene combined with the upregulation of MafA. The resulting NEST for $\epsilon = 0.01$ is shown in Fig. 3(b). Under this optimal intervention, the limiting occupancy of the β lineage goes from less than 0.01 to more than 0.99.

The analysis specifically shows the reliance of some, but not all, lineages on two-step transitions to reach the desired β lineage [Fig. 3(c)]. Previous research has suggested that indirect lineage respecifications might be suboptimal reprogramming strategies [38]. This is clearly the case for the $\delta \rightarrow \beta$ reprogramming, which is optimized through a direct transdifferentiation event. For the α and α/δ lineages, however, with overwhelming likelihood the transformation to the β state will pass through the intermediate α/β state. Thus, which of the two cellular reprogramming strategies (direct or indirect) is optimal is context dependent and cannot be determined without specification of the system, the initial state, and the final state. Systematically accounting for such context dependence could lead to new advances in the development of cellular reprogramming technologies.

B. Predicting anticancer therapeutic targets

Evasion of apoptosis is one of the hallmarks of cancer cells [39]. As such, identifying tunable factors in the cell death pathway that effectively eliminate proliferative (or abnormal survival) cell states without harming healthy cells could lead to new therapeutic targets. To computationally identify candidate targets, we employ OLAC to analyze a reformulation of the Boolean model of the cell death pathway proposed in Ref. [40]. This reformulation is a continuous-variable model generated using the HillCube methodology [41], which is more amenable to analysis and preserves all relevant properties of the original model, including the stable states. The model is comprised of

22 genes central to the programmed cell death and 42 parameters representing kinetic constants of the different reactions, which we denote by c_i , $i = 1, \dots, 42$ (see Appendix B). It follows that two stable states form for all values of the parameters: *apoptosis* and *necrosis*. For a specific range of parameters representing healthy cells, a third state is also stable: the so-called *naive* state. For a different parameter range, however, a different stable state arises, corresponding to a *survival* cell type that is resistant to the apoptotic signal. This survival state represents cancer and is the focus of our discussion.

Our goal is to predict therapeutic targets that can induce transition from the survival state to the apoptotic state, without increasing the rate of apoptotic death in normal cell types or causing them to become survival cells. Although noise can, in principle, induce switches from the survival to the apoptotic state, only a fraction of the cells would transform as desired when both stable states exist. We therefore neglect the effect of noise temporarily and show that even in this case OLAC can be used to identify *successful* optimal interventions, which under these conditions lead to a bifurcation that completely eliminates the undesired (survival) state. To avoid inflammatory response, it is also desirable not to induce transitions to the necrotic state. These conditions are assured by taking $G(\Omega) = -S^*(\text{survival} \rightarrow \text{apoptosis})$ as the objective function to be maximized, and by imposing the constraints $\Delta S^*(\text{naive} \rightarrow \text{apoptotic}) \geq -\vartheta_0$, $\Delta S^*(\text{naive} \rightarrow \text{necrotic}) \geq -\vartheta_0$, and $S^*(\text{survival} \rightarrow \text{necrotic}) \geq \vartheta_0$, where Δ indicates change under the control intervention and ϑ_0 is taken to be 0.05 in our simulations. To encompass the largest possible set of candidate targets, we assume that any of the 42 parameters of the model can be tuned in experiments, and hence, we take $\Omega = \{c_i\}_{i=1}^{42}$. However, since existing experimental techniques cannot be easily used to manipulate a large number of targets, we further constrain each control intervention to involve only a relatively small number of all tunable parameters. This can be achieved by stipulating that the sum of the absolute values of all changes to the parameters must be equal to a prespecified *intervention strength* χ_0 , as detailed in Appendix C.

We thus apply OLAC to the cell death model for the objective function and constraints above. To account for genetic heterogeneity between cancer cells, we analyze six different survival cell types, represented here as six sets of unique values for the parameters Ω . These parameters represent nondimensional kinetic constants for each gene-gene interaction. In our simulations the individual uncontrolled parameters for these cell types are on average 0.50 ± 0.03 and the intervention strength χ_0 is taken to be 0.1, 0.2, ..., 0.9 (Fig. S3 in the Supplemental Material [20] shows the breakdown of all cases). The number of parameters modified by optimal interventions tends to increase as χ_0 increases, ranging from an average of

2.7 to 5.7 for χ_0 varied from 0.1 to 0.9. For each cell type, a successful intervention (eliminating the survival state) is always achieved for large enough χ_0 within this interval. On average, approximately only 5 out of the 42 parameters need to be modified in the optimal successful interventions. To put this result in perspective, we note that if the sparsity constraint in Eq. (C1) is disabled, OLAC leads to an average of no less than 40 modified parameters for a successful intervention. This constraint, which generalizes immediately to any system, is therefore effective to restrain the number of control parameters.

Figure 4 summarizes the results, showing that a unique subset of only 10 parameters is needed to form the (on average) five-parameter successful target sets for any cell type. The biological functions and prevalence of these targets are explicitly indicated in Table S1 of the Supplemental Material [20]. Notably, only two targets are included in interventions found for all six cell types, namely, the parameters whose predicted increase will decrease the activation of NF κ B by IKK and the activation of IKK by RIP1ub (Fig. 4, parameters 26 and 36, respectively). The identification of these two targets is not entirely surprising since NF κ B is a central regulator of the cell death pathway whose overactivation has been implicated in the cellular transition to cancer [42]; the consistent identification of both targets across all six cell types is an indication of the robustness of our approach. Aside from these global targets, OLAC also predicts unique combinations of targets for each cell type, in many cases indicating genes and interactions that have only recently been identified as possible cancer targets, e.g., the potential of suppressing the activation of cFLIP by NF κ B [43] (Fig. 4, parameter 9). The identification of optimal target combinations that are unique to different cell types illustrates the potential of OLAC to assist the development of personalized therapeutic strategies as well as of interventions to address various forms of cancer [44] and to manipulate heterogeneous multistable cells in general [45–47].

OLAC finds an optimal control action whether or not a bifurcation has been reached, allowing its efficacy and possible adverse effects to be monitored in experiments as the strength of the intervention is increased. Theoretically, the efficacy of an intervention can be defined as the relative reduction of the action associated with the transition from the survival state to the apoptotic state [Supplemental Material [20], Fig. S3(c)]. Experimentally, the efficacy can be more easily estimated by monitoring the predicted gene expression changes induced by the interventions [Supplemental Material [20], Fig. S3(b)]. As shown in Fig. 4, for interventions that eliminate the survival state, the sensitivity to the control interventions can vary widely across different genes. For example, in every cell type considered, the expressions of CASP3, SMAC, and CYT-c (among others) are predicted not to change at all until the

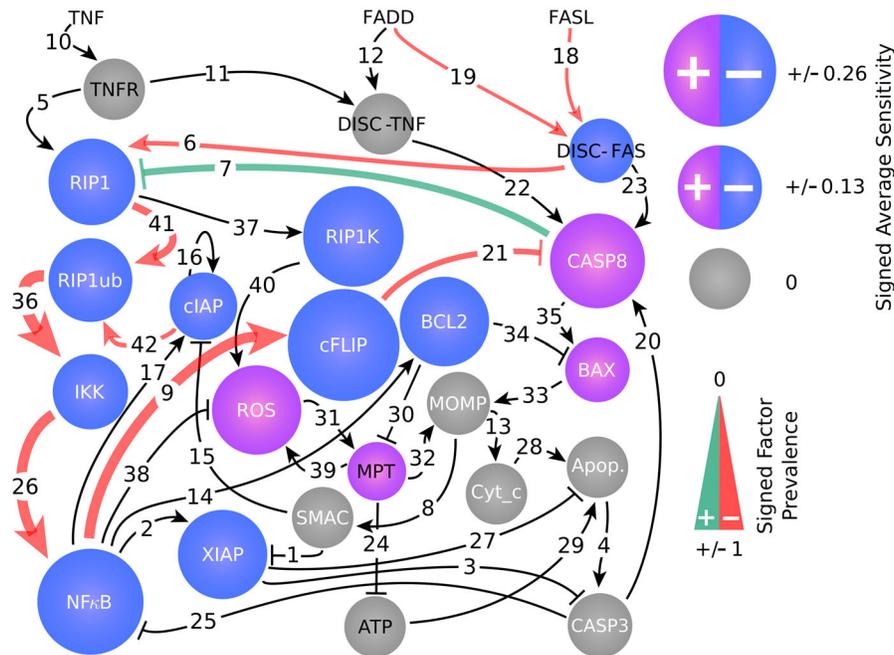


FIG. 4. Optimal therapeutic interventions for the cell death regulatory network model. The network has 22 genes (circles), 3 input parameters (top), and 42 target tunable parameters (edges), where the edge heads distinguish between excitatory (arrow) and inhibitory (bar) regulatory interactions. OLAC is applied to induce a bifurcation transition from survival to apoptotic states. The parameters involved in one or more optimal intervention are consistently upregulated (green) or downregulated (red) by the interventions; those not recruited by any optimal intervention are shown in black. The edge width indicates the prevalence of that parameter in optimal interventions for the various cell types and intervention strengths considered (up to the elimination of the survival state). The size of each circle represents the sensitivity of the gene to proapoptotic interventions, defined as the change in gene expression between the uncontrolled scenario and the smallest intervention scenario at which the survival state is eliminated; the color distinguishes up and down expression. Genes in white are those that change expression substantially between the survival and apoptotic states. These results are based on 6 simulated cell types and 9 different intervention strengths (see Supplemental Material [20], Fig. S3).

elimination of the survival state; in contrast, cFLIP, IKK, and NF κ B change expression in all cell types.

C. Biophysically relevant transition times

In cases where OLAC identifies an intervention to induce a bifurcation to eliminate an undesired state, it is expected from experimental work in deterministically reprogramming somatic cells to a pluripotent state (a prototypical example of an induced cell state transition) that such a transition will occur on a relatively short time scale, within approximately one week [48]. However, due to constraints on parameter changes, it may not always be possible to induce a bifurcation. To maintain biological relevance, the time scale over which these nonbifurcative interventions act should not be exceedingly long. Given a noise strength ϵ and a barrier height $S_{i,j}^*$ between two states, the approximate mean first exit time $\tilde{T}_{i,j}^\epsilon(\Omega)$ from that state can be estimated as [21]

$$\tilde{T}_{i,j}^\epsilon(\Omega) \propto \exp\left(\frac{1}{\epsilon} S_{i,j}^*(\Omega)\right), \quad (6)$$

where the transition time increases for lower noise levels and higher action barriers, as expected. Since $\tilde{T}_{i,j}^\epsilon$ in Eq. (6) is dimensionless, this quantity is best interpreted as a relative increase in transition time over the case of $S_{i,j}^* = 0$, which from above we take to be one week.

Figure 5 shows the mean first transition time using as a model application the cell death model considered in the previous section. The figure shows the mean transition time as a function of both noise level and barrier heights for a single cell type [as defined by each intervention strength in Fig. S3(c) of the Supplemental Material [20]]. It follows that for cases where $S_{i,j}^*(\Omega) \leq 3\epsilon$, the average transition time will be less than 20 weeks. This time scale is a reasonable upper limit for the biological relevance of any intervention: because transition times are exponentially distributed, interventions at this strength will cause a measurable fraction of the population to transition in just a couple of weeks. This benchmark thus provides a straightforward criterion to determine if an identified intervention will be relevant in practice. Figure 5 also demonstrates the important fact that it is not the size of the dynamical system that determines the switching time between states, but rather the ratio of the barrier height to the strength of the noise.

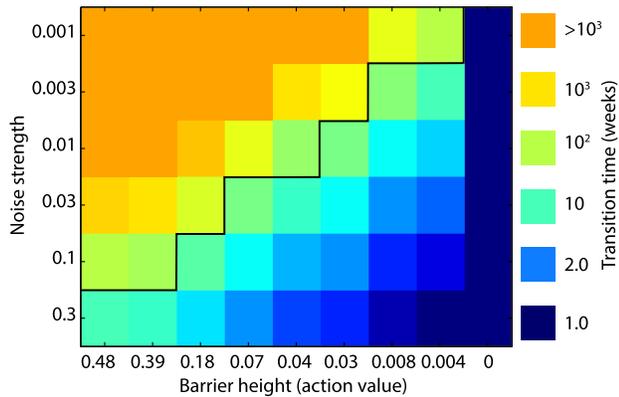


FIG. 5. Transition time for the cell death model as a function of barrier height and noise strength. Transition time is calculated using Eq. (6). The black line separates those barrier height–noise strength combinations that occur over a therapeutically relevant time scale (bottom right) and those that do not (top left).

We expand on this approach in the Supplemental Material [20], Sec. S3, where we demonstrate that OLAC can be used to identify constrained interventions that achieve a desired limiting occupancy within a prespecified time frame. Therein we also generalize OLAC to systems with (multiplicative) noise that depends on the system state and to systems that are best modeled using a chemical kinetic formulation [49–51].

IV. DISCUSSION

The method proposed here, optimal least action control, represents a new direction in the control of biophysical systems. Traditional control approaches for biophysical systems are based on manipulating the trajectories of the system, while our method is based instead on manipulating the epigenetic landscape through which these trajectories travel. This is achieved by effectively altering the barriers between different stable states in the quasipotential of the system, which could be achieved biologically through, for example, a genome editing approach [52]. In this way OLAC stands in contrast to other orthogonal methods that seek to control cellular states by directly modifying gene expression levels through, for example, siRNA strategies [53–55]. As such, our method naturally accounts for cellular noise and the incorporation of constraints on the possible control actions. But in contrast to previous work that has sought to construct the entire quasipotential for a dynamical system [56], we utilize the associated network of state transitions (NEST) to describe and control the system at a substantially reduced computational cost, which renders the approach applicable to a wide range of biophysical systems—including high-dimensional networks.

The cell death model example illustrates one of the key strengths of OLAC: its effectiveness when the problem

involves an explosion in the number of possible reprogramming combinations. In practical applications to multi-parameter systems, it is of interest to identify interventions that are not only optimal but also sparse, i.e., that target only a few out of many possible tunable factors. Such sparsity is desired since most biologically realizable interventions are able to control only a few parameters and, for example, would not be able to directly control all parameters in the cell death model. Because OLAC can benefit from the framework of convex regularization, however, this combinatorial increase in the number of possible interventions can be dealt with easily by incorporating a sparsity constraint that has only marginal impact on the computational cost.

Another key property of the approach is its flexibility. For example, previous modeling approaches to explain reprogramming experiments have focused mainly on bifurcations that destabilize or eliminate states [57], which are comparatively larger changes in the dynamics of the system. These approaches do not benefit from the presence of noise and tacitly assume that, to reach the desired state, the initial state must become deterministically unstable or disappear, which, as demonstrated for the cell death model, may not be possible given a stringent enough set of constraints on the possible set of interventions. On the other hand, for being based on the Freidlin-Wentzell action, our method is effective as a unified approach both (i) to exploit the presence of noise for control in the absence of any bifurcation (as shown for the pancreas model and the cell death model with low intervention strengths) and (ii) to identify control interventions mediated by bifurcations in the absence of any noise (as shown for the cell death model with large intervention strengths). Therefore, instead of facing reduced performance in the presence of noise, which is a common drawback of other control approaches [53], OLAC benefits from the presence of noise, utilizing noise as an additional control tool.

Through the use of the approach introduced here, we have shown that, counterintuitively, the optimal lineage respecification trajectory is often indirect; that is, they correspond to cases in which the most likely trajectory for an optimized transition between two states will pass through one or more intermediate stable states. Such cases cannot be anticipated by common sense, since for a therapeutic intervention, for instance, an indirect path may cause the cells to worsen before they improve. This result suggests a new possible method for identifying enhanced reprogramming strategies, namely, by systematically exploring combinations of intermediate transitions.

Our approach can also be applied to a much broader range of biophysical problems than those discussed in detail here. In particular, OLAC could be used in the context of synthetic biology, as researchers seek to build ever more complex synthetic systems and computer-aided design methods play an increasingly important role [58]. In that context, OLAC

can identify optimal parameter tuning to reshape the quasipotential for the rational design of systems with prespecified dynamical behavior and response to noise. The same approach can also be used to generate insights into epigenetic diseases and the mechanisms that give rise to them, including the possible dependence of their incidence rate on external versus genetic factors, as well as insights into potential preventive measures to reduce disease risks by identifying conditions that increase the barriers for transitions to disease states.

Furthermore, the foundation of OLAC in the Freidlin-Wentzell action means that the applications of this method need not be biophysical. In particular, our method can easily be used to predict control interventions in other noisy multistable networks and, along with NEST, to characterize the basins of attraction of these systems [59]. For example, in the Supplemental Material [20], Sec. S4, we construct the NEST for a dynamical system with more than 100 attractors [60] and demonstrate that, even in a system with such substantial multistability, noise can effectively eliminate the occupancy of the majority of attractors, leaving only a small fraction of them occupied. Moreover, along with the already mentioned applications to power grids, polymer networks, and food-web networks, OLAC could also find use in controlling spreading processes on social networks [61], in inducing synchronization patterns in oscillator networks [62], in manipulating associative memory networks [63], and potentially in creating new attractors [64]. Further development of this method could also expand its applications to models of disease epidemics and population dynamics. In particular, substantial foundational work has been done on the modeling of extinction events in such systems, which typically requires model-specific mathematical analysis [65–67]. Because a white-noise approximation, like in Eq. (1), cannot accurately approximate the dynamics of these systems [68], expanding OLAC to apply to extinction events in larger networks will require using situation-specific calculations of the transition rates.

Ultimately, we believe that OLAC—together with NEST—forms a flexible, scalable method that can be used to understand and control the dynamical stability and response to noise of a wide range of complex networks, including those with large number of dynamical variables, tunable parameters, and attractors. The method is easily implementable, with a ready-to-use numerical implementation included as supplemental files [20]. The method requires no *a priori* information (beyond those implicitly defined by the dynamical equations) about how variations in the control parameters affect the system, and it can be used in concert with rather general constraints on the control actions. While OLAC can be applied to many systems, as formulated here, application of the method requires a quantitative dynamical model. Extending the

method to systems for which no model is available is an important direction for future research. Future work is also expected to expand on the applications of the approach and to further demonstrate its experimental efficacy.

ACKNOWLEDGMENTS

The authors thank Jorge Nocedal for insightful discussions on optimization methods, and members of the Leonard and Thomas-Elliott labs for continued support. This work was funded by the National Cancer Institute Grants No. 1U54CA143869 (NU PS-OC) and No. 1U54CA193419 (Chicago Region PS-OC), the National Institute of General Medical Sciences Grant No. R01GM113238, the National Science Foundation Grant No. DMS-1057128, the Chicago Biomedical Consortium, and a National Science Foundation Graduate Research Fellowship.

APPENDIX A: COMPUTING THE MINIMUM ACTION VALUE

The minimum action $S_{i,j}^*$ is determined using

$$S_{i,j}^* = S[\vec{\phi}_{i,j}^*] = \min_{\substack{\vec{\phi}(t) \\ \vec{\phi}(T_1)=\vec{a}_i \\ \vec{\phi}(T_2)=\vec{a}_j}} (S[\vec{\phi}]),$$

$$S[\vec{\phi}] = \frac{1}{2} \int_{T_1}^{T_2} \left\| \frac{d\vec{\phi}}{dt}(t) - \vec{F}(\vec{\phi}(t); \Omega) \right\|^2 dt, \quad (\text{A1})$$

where \vec{a}_i and \vec{a}_j denote the coordinates of the initial and final stable states, respectively. To solve this optimization problem, we minimize the discretized version of the functional [22], given by

$$S_{t_0, \dots, t_{m_s}}[\vec{\Phi}_0, \dots, \vec{\Phi}_{m_s}]$$

$$= \frac{1}{2} \sum_{n=1}^{m_s} \left\| \frac{\vec{\Phi}_n - \vec{\Phi}_{n-1}}{\Delta t_n} - \vec{F}(\vec{\Phi}_{n-1/2}; \Omega) \right\|^2 \Delta t_n, \quad (\text{A2})$$

where we use $T_1 = t_0 < t_1 < \dots < t_{m_s} = T_2$, $\Delta t_n = t_n - t_{n-1}$, $\vec{\Phi}_n = \vec{\phi}(t_n)$, and $\vec{\Phi}_{n-1/2} = (\vec{\Phi}_n + \vec{\Phi}_{n-1})/2$. In our simulations, we set $T_2 = -T_1 = 20$ and $m_s = 100$, and verified that larger values of $T_2 - T_1$ and m_s would not improve accuracy noticeably.

To maximize efficiency, we regularly remesh the path from the time domain to the space domain and adaptively redefine t_n according to

$$\Delta\alpha_n \int_{T_1}^{T_2} w(t) dt = w(t_{n-1/2}) \Delta t_n, \quad (\text{A3})$$

where $\Delta\alpha_n = 1/m_s$, $t_{n-1/2} = (t_n + t_{n-1})/2$, and $w(t) = \sqrt{1 + C\|(d\vec{\phi}/dt)(t)\|^2}$ is the monitor function measuring the speed along the path. We use the initial t_n evenly spaced and $C = 10^{10}$. These calculations require an initial path between the stable states. The numerical results we report are obtained using a straight line as the initial path. We have checked that by using different initial paths, such as those generated through the Brownian bridge approach [69], the simulation always converges to the same optimal paths.

We note that, since the parameters of the system are modified at each step of OLAC, robust numerical continuation of the equilibria is necessary. We use a simple homotopy method to continue the stable states from initial parameters Ω' to terminal parameters Ω'' . Specifically, we use linear interpolation between these parameter values combined with a Newton step [70]. The latter includes checking at each step that the desired states are not lost to unintended bifurcations.

All optimization procedures are done employing the interior-point algorithm in the `fmincon` function of the optimization toolbox of MATLAB [71]. A MATLAB implementation of OLAC is included as supplemental files (see Supplemental Material [20], Sec. S2, for details).

APPENDIX B: EQUATIONS OF THE MODELS CONSIDERED

1. VPC differentiation model

The deterministic component of the VPC model [24] is given by

$$\frac{d\vec{r}}{dt} = \frac{1}{\tau} (\vec{\sigma}_1 - \vec{r}), \quad (\text{B1})$$

where $\vec{r} = (x, y)^T$, $\vec{\sigma}_1 = \tanh(\|\vec{f} + \vec{m}\|)(\vec{f} + \vec{m}/\|\vec{f} + \vec{m}\|)$, $\vec{f} = [2x - 2c_2xy, 2y + c_2(y^2 - x^2)]^T$, and $\vec{m} = \vec{m}_0 + \ell_1\vec{m}_1 + \ell_2\vec{m}_2$. This is an effective representation of the combined effects of EGF and Notch signaling, whose signaling strengths are determined by ℓ_1 and ℓ_2 , respectively. The other parameters are $\tau = 2.18$, $\vec{m}_0 = (-0.86, -0.50)^T$, $\vec{m}_1 = (0.86, -0.50)^T$, and $\vec{m}_2 = (0.0, 1.0)^T$.

2. HPC differentiation model.

The deterministic component of the HPC differentiation model is [37]

$$\begin{aligned} \text{Pdx1: } \frac{dx_1}{dt} &= \sigma_1 a_s \frac{x_4^n + x_7^h + x_8^h + x_{10}^h}{1 + x_4^h + x_7^h + x_8^h + x_{10}^h} - k_{\text{deg}} x_1, \\ \text{ptf1a: } \frac{dx_2}{dt} &= \sigma_2 a \frac{x_{10}^h}{1 + x_{10}^h + x_3^h} - k_{\text{deg}} x_2, \\ \text{Ngn3: } \frac{dx_3}{dt} &= \sigma_3 a \frac{x_{10}^h}{1 + x_{10}^h + x_2^h} - k_{\text{deg}} x_3, \\ \text{Pax6: } \frac{dx_4}{dt} &= \sigma_4 a \frac{\mu^h x_3^h + x_4^h}{1 + \mu^h x_3^h + x_4^h} - k_{\text{deg}} x_4, \\ \text{Pax4: } \frac{dx_5}{dt} &= \sigma_5 a_e \frac{\mu_m^h x_1^h \mu^h x_3^h}{1 + \mu_m^h x_1^h \mu^h x_3^h + x_6^h} - k_{\text{deg}} x_5, \\ \text{Arx: } \frac{dx_6}{dt} &= \sigma_6 a_e \frac{\mu_m^h x_1^h \mu^h x_3^h}{1 + \mu_m^h x_1^h \mu^h x_3^h + x_5^h} - k_{\text{deg}} x_6, \\ \text{MafA: } \frac{dx_7}{dt} &= \sigma_7 a \frac{\mu_m^h x_1^h \mu^h x_3^h x_5^h + x_7^h}{1 + \mu_m^h x_1^h \mu^h x_3^h x_5^h + x_7^h + x_8^h} - k_{\text{deg}} x_7, \\ \delta\text{-gene: } \frac{dx_8}{dt} &= \sigma_8 a \frac{\mu_m^h x_1^h \mu^h x_3^h x_5^h + x_8^h}{1 + \mu_m^h x_1^h \mu^h x_3^h x_5^h + x_7^h + x_8^h} - k_{\text{deg}} x_8, \\ \text{Brn4: } \frac{dx_9}{dt} &= \sigma_9 a \frac{\mu^h x_6^h + x_9^h}{1 + \mu^h x_6^h + x_9^h} - k_{\text{deg}} x_9, \\ \text{Hnf6: } \frac{dx_{10}}{dt} &= -\sigma_{10} k_{\text{deg}} x_{10}, \end{aligned} \quad (\text{B2})$$

where $(a_s, a_e, a, k_{\text{deg}}, \mu, \mu_m, h) = (2.2, 6.0, 4.0, 1.0, 0.25, 0.125, 4.0)$ are the fixed parameters in these equations. Each parameter σ_i represents a tunable factor to alter the expression of the gene represented by x_i .

3. Cell death model.

The cell death model was converted from the Boolean model [40] (see Ref. [72]) into a continuous version using the ODEFY package [41]. The system has 22 variables, representing gene products, and 42 tunable parameters. In addition, the model has 3 input parameters that do not change in time and 3 output variables that indicate the state of the cell (distinct combinations of which indicate whether the cell is in the survival, apoptotic, necrotic, or naive state).

APPENDIX C: IMPLEMENTING THE SPARSITY CONSTRAINT

In many biological systems there are dozens or hundreds of parameters that could potentially be changed, but in general only a few of them can be changed in any one intervention. To identify the few most promising targets from a large field of possible ones we employ a sparsity constraint. The constraint is implemented as

$$\sum_{i=1}^{|\Omega|} |\Delta\Omega_i| \leq \chi_0, \quad (\text{C1})$$

where, as above, Δ denotes the change due to the control action. While the condition in Eq. (C1) is by itself consistent with all parameters being altered, optimization under this constraint (as the one invoked by OLAC) is expected to lead to a reduced number of modified parameters. The basis for this conclusion is that this constraint works similarly to well-established methods of convex regularization, which are known to lead to sparsity under general conditions [73]. The specific number of modified parameters as well as success rate will generally depend on χ_0 , and this dependence can be explored as an additional control factor. This formulation has the remarkable advantage of involving only one optimization step, and hence avoids the combinatorial explosion that would be involved in testing all $|\Omega|!/k!(|\Omega| - k)!$ combinations of possible “ $|\Omega|$ choose k ” tunable factors. Indeed, an exhaustive strategy would be computationally prohibitive since testing for all k would require $2^{|\Omega|} - 1$ optimization steps, which is $>10^{12}$ for $|\Omega| = 42$. Naturally, if a particular parameter is not targetable under the given conditions, such information can be directly incorporated into our analysis.

-
- [1] A. Raj and A. van Oudenaarden, *Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences*, *Cell* **135**, 216 (2008).
- [2] H. H. McAdams and A. Arkin, *Stochastic Mechanisms in Gene Expression*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 814 (1997).
- [3] G. Balazsi, A. van Oudenaarden, and J. Collins, *Cellular Decision Making and Biological Noise: From Microbes to Mammals*, *Cell* **144**, 910 (2011).
- [4] H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, and S. Huang, *Transcriptome-Wide Noise Controls Lineage Choice in Mammalian Progenitor Cells*, *Nature (London)* **453**, 544 (2008).
- [5] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, *A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell*, *Science* **322**, 442 (2008).
- [6] P. B. Gupta, C. M. Fillmore, G. Jiang, S. D. Shapira, K. Tao, C. Kuperwasser, and E. S. Lander, *Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells*, *Cell* **146**, 633 (2011).
- [7] C. L. Chaffer *et al.*, *Normal and Neoplastic Nonstem Cells Can Spontaneously Convert to a Stem-like State*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7950 (2011).
- [8] C. H. Waddington, *The Strategy of the Genes* (Allen and Unwin, London, 1957).
- [9] S. Huang, G. Eichler, Y. Bar-Yam, and D. E. Ingber, *Cell Fates as High-Dimensional Attractor States of a Complex Gene Regulatory Network*, *Phys. Rev. Lett.* **94**, 128701 (2005).
- [10] M. Acar, A. Becskei, and A. van Oudenaarden, *Enhancement of Cellular Memory by Reducing Stochastic Transitions*, *Nature (London)* **435**, 228 (2005).
- [11] K. Ishimatsu, T. Hata, A. Mochizuki, R. Sekine, M. Yamamura, and D. Kiga, *General Applicability of Synthetic Gene-Overexpression for Cell-Type Ratio Control via Reprogramming*, *ACS Synth. Biol.* **3**, 638 (2014).
- [12] L. Dai, D. Vorselen, K. Korolev, and J. Gore, *Generic Indicators for Loss of Resilience before a Tipping Point Leading to Population Collapse*, *Science* **336**, 1175 (2012).
- [13] J. Hanna, K. Saha, B. Pando, J. van Zon, C. J. Lengner, M. P. Creighton, A. van Oudenaarden, and R. Jaenisch, *Direct Cell Reprogramming is a Stochastic Process Amenable to Acceleration*, *Nature (London)* **462**, 595 (2009).
- [14] D. A. Robinton and G. Q. Daley, *The Promise of Induced Pluripotent Stem Cells in Research and Therapy*, *Nature (London)* **481**, 295 (2012).
- [15] S. Huang and S. Kauffman, *How to Escape the Cancer Attractor: Rationale and Limitations of Multi-Target Drugs*, *Seminars in cancer biology* **23**, 270 (2013).
- [16] P. Creixell, E. M. Schoof, J. T. Erler, and R. Lindig, *Navigating Cancer Network Attractors for Tumor-Specific Therapy*, *Nat. Biotechnol.* **30**, 842 (2012).
- [17] A. E. Motter, S. A. Myers, M. Anghel, and T. Nishikawa, *Spontaneous Synchrony in Power-Grid Networks*, *Nat. Phys.* **9**, 191 (2013).
- [18] B. Schnurr, F. Gittes, and F. C. MacKintosh, *Metastable Intermediates in the Condensation of Semiflexible Polymers*, *Phys. Rev. E* **65**, 061904 (2002).
- [19] S. Sahasrabudhe and A. E. Motter, *Rescuing Ecosystems from Extinction Cascades through Compensatory Perturbations*, *Nat. Commun.* **2**, 170 (2011).
- [20] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevX.5.031036> for generalizations of OLAC, a MATLAB implementation of OLAC, additional analyses, supplemental table, supplemental figures, and supplemental references.
- [21] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, Berlin, 2013).
- [22] X. Zhou, W. Ren, and W. E. Barlow, *Adaptive Minimum Action Method for the Study of Rare Events*, *J. Chem. Phys.* **128**, 104111 (2008).
- [23] R. Maier and D. Stein, *Limiting Exit Location Distributions in the Stochastic Exit Problem*, *SIAM J. Appl. Math.* **57**, 752 (1997).
- [24] F. Corson and E. D. Siggia, *Geometry, Epistasis, and Developmental Patterning*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5568 (2012).
- [25] P. W. Sternberg, *Vulval Development*, in *WormBook*, edited by The *C. elegans* Research Community, doi: (2005).
- [26] J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, Berlin, 2006).
- [27] V. N. Smelyanskiy and M. I. Dykman, *Optimal Control of Large Fluctuations*, *Phys. Rev. E* **55**, 2516 (1997).
- [28] M. I. Dykman, E. Mori, J. Ross, and P. M. Hunt, *Large Fluctuations and Optimal Paths in Chemical Kinetics*, *J. Chem. Phys.* **100**, 5735 (1994).
- [29] A. N. Pisarchik and U. Feudel, *Control of Multistability*, *Phys. Rep.* **540**, 167 (2014).
- [30] B. E. Yugmeister and H. Rabitz, *Cooperating with Nonequilibrium Fluctuations through Their Optimal Control*, *Phys. Rev. E* **55**, 2522 (1997).
- [31] B. S. Lindley and I. B. Schwartz, *An Iterative Action Minimizing Method for Computing Optimal Paths in*

- Stochastic Dynamical Systems*, *Physica (Amsterdam)* **255D**, 22 (2013).
- [32] L. Billings, I. B. Schwartz, M. McCrary, A. N. Korotkov, and M. I. Dykman, *Switching Exponent Scaling Near Bifurcation Points for Non-Gaussian Noise*, *Phys. Rev. Lett.* **104**, 140601 (2010).
- [33] O. M. Becker and M. Karplus, *The Topology of Multidimensional Potential Energy Surfaces: Theory and Application to Peptide Structure and Kinetics*, *J. Chem. Phys.* **106**, 1495 (1997).
- [34] F. Rao and A. Caflisch, *The Protein Folding Network*, *J. Mol. Biol.* **342**, 299 (2004).
- [35] R. S. Wang, A. Saadatpour, and R. Albert, *Boolean Modeling in Systems Biology: An Overview of Methodology and Applications*, *Phys. Biol.* **9**, 055001 (2012).
- [36] F. Noe and S. Fischer, *Transition Networks for Modeling the Kinetics of Conformational Change in Macromolecules*, *Curr. Opin. Struct. Biol.* **18**, 154 (2008).
- [37] J. X. Zhou, L. Brusch, and S. Huang, *Predicting Pancreas Cell Fate Decisions and Reprogramming with a Hierarchical Multi-Attractor Model*, *PLoS One* **6**, e14752 (2011).
- [38] R. Passier and C. Mummery, *Getting to the Heart of the Matter: Direct Reprogramming to Cardiomyocytes*, *Cell Stem Cell* **7**, 139 (2010).
- [39] D. Hanahan and R. A. Weinberg, *The Hallmarks of Cancer*, *Cell* **100**, 57 (2000).
- [40] L. Calzone, L. Tournier, S. Fourquet, D. Thieffrey, B. Zhivotovsky, E. Barillot, and A. Zinovyev, *Mathematical Modeling of Cell-Fate Decision in Response to Death Receptor Engagement*, *PLoS Comput. Biol.* **6**, e1000702 (2010).
- [41] D. M. Wittmann, J. Krumsiek, J. Saez-Rodriguez, D. A. Lauffenburger, S. Klamt, and F. J. Theis, *Transforming Boolean Models to Continuous Models: Methodology and Application to T-Cell Receptor Signaling*, *BMC Syst. Biol.* **3**, 98 (2009).
- [42] J. L. Luo, H. Kamata, and M. Karin, *IKK/NF- κ B Signaling: Balancing Life and Death—A New Approach to Cancer Therapy*, *J. Clin. Invest.* **115**, 2625 (2005).
- [43] A. R. Safa, *c-Flip, a Master Anti-Apoptotic Regulator*, *Exp. Oncol.* **34**, 176 (2012).
- [44] A. Saadatpour, R. Wang, A. Liao, X. Liu, T. P. Loughran, I. Albert, and R. Albert, *Dynamical and Structural Analysis of a T Cell Survival Network Identifies Novel Candidate Therapeutic Targets for Large Granular Lymphocyte Leukemia*, *PLoS Comput. Biol.* **7**, e1002267 (2011).
- [45] E. R. Regan and W. C. Aird, *Dynamical Systems Approach to Endothelial Heterogeneity*, *Circ. Res.* **111**, 110 (2012).
- [46] A. H. Lang, H. Li, J. J. Collins, and P. Mehta, *Epigenetic Landscapes Explain Partially Reprogrammed Cells and Identify Key Reprogramming Genes*, *PLoS Comput. Biol.* **10**, e1003734 (2014).
- [47] S. N. Steinway, J. G. T. Zañudo, W. Ding, C. B. Rountree, D. J. Feith, T. P. Loughran, Jr., and R. Albert, *Network Modeling of TGF β Signaling in Hepatocellular Carcinoma Epithelial-to-Mesenchymal Transition Reveals Joint Sonic Hedgehog and Wnt Pathway Activation*, *Cancer Res.* **74**, 5963 (2014).
- [48] Y. Rais, A. Zviran, S. Geula, O. Gafni, E. Chomsky, S. Viukov, A. AlFatah Mansour, I. Caspi, V. Krupalnik, M. Zerbib *et al.*, *Deterministic Direct Reprogramming of Somatic Cells to Pluripotency*, *Nature (London)* **502**, 65 (2013).
- [49] T. E. Turner, S. Schnell, and K. Burrage, *Stochastic Approaches for Modelling In Vivo Reactions*, *Comp. Biol. Chem.* **28**, 165 (2004).
- [50] D. Liu, *Optimal Transition Paths of Stochastic Chemical Kinetic Systems*, *J. Chem. Phys.* **124**, 164104 (2006).
- [51] D. Liu, *A Numerical Scheme for Optimal Transition Paths of Stochastic Chemical Kinetic Systems*, *J. Comput. Phys.* **227**, 8672 (2008).
- [52] T. Gaj, C. A. Gersbach, and C. F. Barbas III, *ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering*, *Trends Biotechnol.* **31**, 397 (2013).
- [53] S. P. Cornelius, W. L. Kath, and A. E. Motter, *Realistic Control of Network Dynamics*, *Nat. Commun.* **4**, 1942 (2013).
- [54] H. Xia, Q. Mao, H. Paulson, and B. L. Davidson, *siRNA-Mediated Gene Silencing In Vitro and In Vivo*, *Nat. Biotechnol.* **20**, 1006 (2002).
- [55] J. G. T. Zañudo and R. Albert, *Cell Fate Reprogramming by Control of Intracellular Network Dynamics*, *PLoS Comput. Biol.* **11**, e1004193 (2015).
- [56] J. Wang, K. Zhang, L. Xu, and E. Wang, *Quantifying the Waddington Landscape and Biological Paths for Development and Differentiation*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8257 (2011).
- [57] V. Chickarmane and C. Peterson, *A Computational Model for Understanding Stem Cell, Trophectoderm and Endoderm Lineage Determination*, *PLoS One* **3**, e3478 (2008).
- [58] M. Marchisio and J. Stelling, *Computational Design Tools for Synthetic Biology*, *Curr. Opin. Biotechnol.* **20**, 479 (2009).
- [59] P. J. Menck, J. Heitzig, N. Marwan, and J. Kurths, *How Basin Stability Complements the Linear-Stability Paradigm*, *Nat. Phys.* **9**, 89 (2013).
- [60] U. Feudel, *Complex Dynamics in Multistable Systems*, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **18**, 1607 (2008).
- [61] M. Granovetter, *Threshold Models of Collective Behavior*, *Am. J. Sociology* **83**, 1420 (1978).
- [62] D. M. Abrams and S. H. Strogatz, *Chimera States for Coupled Oscillators*, *Phys. Rev. Lett.* **93**, 174102 (2004).
- [63] J. J. Hopfield, *Neural Networks and Physical Systems with Emergent Computational Abilities*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554 (1982).
- [64] C. Campbell and R. Albert, *Stabilization of Perturbed Boolean Network Attractors through Compensatory Interactions*, *BMC Syst. Biol.* **8**, 53 (2014).
- [65] M. Assaf, A. Kamenev, and B. Meerson, *Population Extinction in a Time-Modulated Environment*, *Phys. Rev. E* **78**, 041123 (2008).
- [66] M. Dykman, I. B. Schwartz, and A. S. Landsman, *Disease Extinction in the Presence of Random Vaccination*, *Phys. Rev. Lett.* **101**, 078101 (2008).
- [67] M. Khasin and M. I. Dykman, *Extinction Rate Fragility in Population Dynamics*, *Phys. Rev. Lett.* **103**, 068101 (2009).

- [68] K. V. Sargsyan, C. R. Doering, and L. M. Sander, *Extinction Times for Birth-Death Processes: Exact Results, Continuum Asymptotics, and the Failure of the Fokker-Planck Approximation*, *Multiscale Model. Simul.* **3**, 283 (2005).
- [69] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus* (Springer, Berlin, 1991).
- [70] R. Seydel, *From Equilibrium to Chaos: Practical Bifurcation and Stability Analysis* (Elsevier, Amsterdam, 1988).
- [71] MATLAB version 8.1.0.604, The MathWorks, Inc., 2013.
- [72] Available at http://www.ebi.ac.uk/biomodels-main/MODEL_0912180000.
- [73] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004).