Network topology and interaction logic determine states it supports

Check for updates

Tomáš Gedeon ✉

In this review paper we summarize a recent progress on the problem of describing range of dynamics supported by a network. We show that there is natural connection between network models consisting of collections of multivalued monotone boolean functions and ordinary differential equations models. We show how to construct such collections and use them to answer questions about prevalence of cellular phenotypes that correspond to equilibria of network models.

The goal of this review paper is to describe recent progress on describing capacity of regulatory networks to exhibit different phenotypes in different conditions. In several types of models of regulatory networks we associate different phenotypes to equilibria admitted by the model. Other dynamical phenotypes like cell cycle progression and circadian rhythm can be also investigated by the approach we describe[1,2], but will be not discussed here.

This work relies on several papers in specialized mathematical[3,4] and theoretical computer science[5] literature that were firmly rooted in problem of parameterization of ordinary differential equations (ODE) network models. The recent realization of the connection between monotone boolean functions and parameterization of switching ODEs[5,6] facilitated successful applications to study of steady states in developmental networks[7,8]. The goal of this review is to provide a concise and accessible entry point to the DSGRN approach for the systems biology audience with emphasis on description of steady states in monotone boolean models of developmental networks.

One of our motivations is analysis of developmental networks that determine cell's fate. Here equilibria of the model represent differentiated cell types and presence of multiple stable equilibria (multistability) suggests that different developmental pathways may lead to different cell types. In such networks it is important to understand what types of multistability are possible, which include the number of coexisting stable equilibria, their prevalence under changing conditions and the types of equilibria that are able to co-exist. When we will use boolean description, the *type* of the equilibrium is determined by which genes are expressed high, and which ones are expressed at low levels.

Networks are qualitative models of pairwise interactions/influences between nodes which can represent mRNA, protein concentrations, or even different conformations of proteins if they have differential impact on other network nodes. The pairwise interactions are directed from one node to another and may model transcriptional regulation, post-translational modifications like phosphorylation, ubiquitination as well as conformational changes.

Behavior of any network that include feedback loops where a sequence of nodes influence each other in a circular fashion is very difficult to understand without a mathematical model. This is especially true for larger networks. We will concentrate here on two seemingly very different types of models, *boolean models* and *ordinary differential equations models (ODE)*. Boolean models describe state of each node as "active" and assign to this state value 1, or inactive, and assign to this state value 0. In closely related *multivalued boolean models* a state of each node i is described by a set of integers $X_i = \{0, 1, \ldots, t\}$ expressing level to which the node is able to activate some, but not all, downstream nodes. To each node i we associate a boolean update function g_i, which describes the state of the node i as a function of its inputs. Boolean functions that respect the type of network interactions (activating vs. repressing) are called *monotone boolean functions (MBF)*. Dynamics of these models consists of regular updates of the state of each node i. *Synchronous update* updates all nodes at the same time, while asynchronous update updates nodes one at the time. While synchronous update leads to a deterministic dynamics, the implied presence of a clock that synchronizes the update schedule makes it biologically unrealistic. Since in the asynchronous update the future state depends on the order of nodes that are being updated, this update is represented by a multivalued map where states can evolve differently based on which node is updated first. Although boolean networks are often presented as "parameter free" models, different choices of the boolean update functions that are compatible with the same network may lead to different types of dynamics and different types of equilibria supported by the network.

In contrast to boolean models, ODE models describe evolution of state in continuous time. Regulatory network models often use monotone sigmoid Hill functions to describe network interactions. Specification of each Hill nonlinearity typically requires four parameters and these parameters are difficult to obtain experimentally. In addition, these biological parameters are fundamentally different than the parameters of physics models that are the gold standard of scientific modeling. Mass of an object is a parameter that is independent on the model used; any model attempting to describe motion will need to have this parameter present. Since network ODE models are not derived from first principles, changing a nonlinearity from a Hill function to, say, a polynomial, necessitates re-fitting of all the parameters. Thus, the values of the parameters are model dependent. It is

Department of Mathematical Sciences, Montana State University, Bozeman, MT, USA. ✉e-mail: tgedeon@montana.edu

therefore difficult to justify spending experimental effort and resources on measuring precise parameter values at which the network operates. Perhaps it is more realistic to try to establish a range for each parameter. However, even if he ranges are successfully established describing all dynamics of an ODE system across ranges of parameter values is a very difficult problem.

One approach to approximate such description is sampling the parameter space, simulating each resulting ODE system and collect statistical data about the behavior across the samples. However, since the number of parameters for even a small network is very high, such sampling is always sparse. In addition, there is no theory that would guarantee that certain sample size is sufficient to sample all possible behaviors, or even a high proportion of all behaviors. This is partially a consequence of the fact that the set of all possible behaviors for ODE is uncountable, preventing its probabilistic description. Along these lines, Randomized Circuit Perturbation (RACIPE)[9] is a sampling approach that judiciously tries to sample predominantly biologically relevant parameters.

In this review, we describe an alternative approach, DSGRN (Dynamic Signatures Generated by Regulatory Networks)[3,5,10]. DSGRN associates to a network an ODE model with piece-wise constant monotone nonlinearities consistent with the network structure. Since the nonlinearities only assume finite number of values, there are two important simplifications compared to a general ODE model. First, the ODE solutions in the phase space can be described by a finite state transition graph (STG) and second, the continuous parameter space can be decomposed to finite number of domains such that for all parameters in a domain the STG, and hence the dynamics defined by STG, is the same. This turns analysis of an ODE system with its continuum phase space and continuum parameter space into a finite combinatorial problem. In addition, the piecewise constant nonlinearities can be perturbed to Hill function models, ramp function models or any other sigmoid nonlinearities and theoretical results guarantee that the analysis of the combinatorial dynamics is valid for nearby continuum models[11].

Numerical investigation comparing the repertoire of equilibria, prevalence of bistability and multistability described DSGRN and the same repertoire described using RACIPE[9] was done in ref. 12 for two networks: toggle switch[13] and toggle triad (Fig. 1a). Since the RACIPE simulates Hill models with finite value of Hill coefficient n, the paper[12] examined how large the value of n should be for RACIPE and DSGRN results to agree. Surprisingly, DSGRN predicts RACIPE results even for relatively small values of n. Since the DSGRN analysis is computationally many orders of magnitude faster than sampling and simulation of RACIPE, this suggests that DSGRN may be a valuable tool for the first pass analysis of the range of behaviors that the network is able to support.

Importantly, DSGRN approach bridges the divide between boolean models and ODE models. It can be shown[5], that each parameter domain of a switching ODE is described by a collection of monotone boolean maps (MBF). Coarse STG dynamics of any ODE parameterized by a parameter from such a domain agrees with the dynamics of the asynchronous update of a particular multivalued monotone boolean map (mMBF). This bridge between boolean models and ODE model suggests description of potential network dynamics by enumerating all multivalued monotone Boolean functions compatible with the network and, for each such choice, describing its set of equilibria. This approach is limited by the exponential growth of number of mMBFs compatible with a network as a function of the number of its nodes and edges. We describe potential ways to address this curse of dimensionality by focusing of particular small subsets of MBFs that seem to represent the behavior of the entire set.

Example: (Toggle triad)

Before we describe our approach in detail, we illustrate it on a simple example. Consider toggle triad network in Fig. 1a. with three nodes a, b, c and pair of repressive edges between any two nodes, This network was anaylzed in refs. 7,14 as a network responsible for Th1/Th2/Th17 immune cell differentiation.

We assume each node can be either active of inactive and these are represented as boolean states in $\mathbb{B} = \{0, 1\}$. Each node receives two inputs and the state of each node is updated by a monotone boolean function $f : \mathbb{B}^2 \to \mathbb{B}$. In Fig. 1b, we list all such functions. In the first column we list all potential values of boolean inputs X and Y. Assume for the moment that these represent states of nodes a, b respectively. Since all edges of the network are repressive, second column lists the values that are transmitted by the edges to their target c. The update function f_c takes this pair of boolean values and produces the new state of node c. Therefore f_c is a composition of the map B which reverses the boolean inputs and which depends on the fact that both edges are repressive, and the second map g which takes this reversed input to its final value. There are six choices for a monotone boolean function g listed in the last six columns. These choices only depend on the fact that the node c has two inputs, but not on whether these are activating or repressing—that information is encoded in the map B. The potential functions g are the constant functions **0** and **1**; function X (Y) that repeats the values of the input X (Y) and functions $\wedge$ ($\vee$), which are logical AND (OR), respectively. This set of functions can be organized as a partially ordered set (poset) in Fig. 1c where two functions are connected by an edge when they differ in exactly one output.

All possible MBF that are consistent with the toggle triad are triples $f = (f_a, f_b, f_c)$. Therefore there are $6^3 = 216$ boolean network models.


Fig. 1 | Toggle triad network analysis. a Toggle triad network (**b**) Set of possible update boolean functions $f : \mathbb{B}^2 \to \{0, 1\}$ where $f = g \circ B$. Left column: input values u are all possible values of pair of boolean variables X and Y. Second column; values of function $B(u)$. Columns 3–8: Collection of all possible monotone boolean functions $g : \mathbb{B}^2 \to \{0, 1\}$: function **0** and the **1** function are constant; function X (Y) repeats the values of input X (Y) and functions $\wedge$ ($\vee$) are logical AND (OR) respectively. **c** The partially ordered set of all MBFs $g : \mathbb{B}^2 \to \{0, 1\}$; neighboring functions differ by exactly one values.
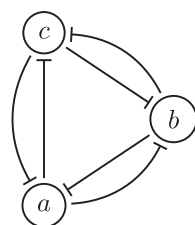
XY	$\neg X \neg Y$	**0**	$\wedge$	X	Y	$\vee$	**1**
11	00	0	0	0	0	0	1
10	01	0	0	0	1	1	1
01	10	0	0	1	0	1	1
00	11	0	1	1	1	1	1
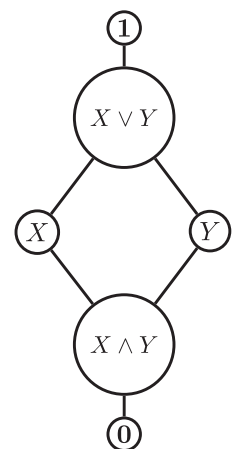
(a)　　　　　　(b)　　　　　　(c)

We investigate how many of these models support a constant equilibria (000) and (111), how many support equilibrium (001) where only one gene is active and two are suppressed and how many support equilibrium (110) where two genes are active and one is suppressed. Because of the symmetry, the number of functions f supporting (001) and the number supporting (010) (as well as (100)), is the same. The key observation is that an equilibrium Q = (uvw) is supported by $f = (f_a, f_b, f_c)$ if, and only if,

$$f_a(Q) = u, \quad f_b(Q) = v, \quad f_c(Q) = w. \tag{1}$$

With this in mind, a direct inspection of table in Fig. 1 shows that (000) is an equilibrium when

$$f_a(00) = 0, \quad f_b(00) = 0, \text{ and } \quad f_c(00) = 0.$$

There is only a single combination of functions, $f = (\mathbf{0}, \mathbf{0}, \mathbf{0})$, that satisfies these conditions. Therefore the phenotype (000) has prevalence 1/216 in toggle triad. Similarly, the equilibrium (111) is only supported by the parameter $(\mathbf{1}, \mathbf{1}, \mathbf{1})$ and has prevalence 1/216. We conclude that both constant phenotypes are rare in toggle triad.

Now we count number of boolean networks that support equilibrium (001). This requires

$$f_a(01) = 0, \quad f_b(01) = 0, \quad f_c(00) = 1.$$

The first and the second conditions are satisfied by any functions $f_a, f_b \in \{\mathbf{0}, \wedge, Y\}$ and the third condition by by $f_c \in \{\wedge, X, Y, \vee, \mathbf{1}\}$. Therefore there are $3 \times 3 \times 5 = 45$ parameters supporting this equilibrium. Similar calculation shows that also (110) is supported by 45 boolean networks, as would be expected from the symmetry considerations. We conclude that toggle triad supports mixed phenotypes, where one gene, or two genes are active with prevalence 45/216. This is higher than the prevalence of constant phenotypes. These results agree with ref. [14] which found using RACIPE sampling that the constant equilibria are negligible phenotype and most of the monostable dynamics shows convergence to either singly activated (100) or doubly activated (110) types of equilibria.

We close the introduction with a comment on *non-degenerate* boolean functions which we return to later in the text. Among the six MBFs in Fig. 1b the constant functions $\mathbf{0}, \mathbf{1}$ are considered degenerate, as their output does not depend on the input values. Further, if we assume that the network edges have been experimentally determined, there must be conditions at which these edges influence the state of the target node. Under this assumption, function $g(XY) = X$ is also degenerate, since the input Y does not influence the result; the same is true for the function $g(XY) = Y$. Therefore the only non-degenerate functions $g(XY) = X \wedge Y$ and $g(XY) = X \vee Y$. The concept of non-degenerate monotone boolean functions allows us, for larger networks, to restrict our attention to only *essential* boolean networks f where each component function is non-degenerate. Note that there are only 8 essential boolean networks out of 216 boolean networks; none of these support constant equilibria (000), (111), but 2/8 support mixed equilibria (001) and (110).

Astute reader certainly notices that there are six mixed equilibria (100), (010), (001), (110), (101), (011) each of which is supported by two out of eight essential networks. This is only possible when some of the essential networks support bistability or multistability. This is indeed the case, and we postpone the computation of prevalence of bistability and multistability to the "Multistability in toggle triad".

The paper is organized as follows. "Ensemble of multivalued monotone boolean functions compatible with the network" is devoted to theoretical description of DSGRN methodology that builds a collection of all multivalued monotone boolean functions compatible with network structure and organizes it into parameter graph PG. Each node of the parameter graph gives rise to potentially different dynamics captured by a state transition graph (STG); the long term behavior of STG dynamics is captured by a *Morse graph*. Theoretical developments are illustrated along the way on a E2F-Rb network responsible for commitment to the S-phase during the mammalian cell cycle. In "Essential boolean parameters" we use our methodology to analyze three networks: E2F-Rb network and two networks implied in immune commitment networks: toggle triad and toggle tetrahedron. We conclude by the "Discussion" section and leave the description of connection between the parameter graph PG and the ODE network models to "Connecting parameter graph PG to ODE models".

Ensemble of multivalued monotone boolean functions compatible with the network

In addition to the toggle triad example in the introduction we will illustrate our methods on a network that plays central role in transition from G1 to S phase in cell cycle in eukaryots. A mammalian network (Fig. 2a) was studied by ref. [15] and then further analyzed by refs. [1,16]. The essential elements of this network is a family of E2F transcription factors which are sequestered in a heterodimer by Rb in non-proliferating cells in G1 phase. Release of E2F by phosphorylation of Rb results in initiation of S phase of the cell cycle. The principal controls of Rb are cyclin/kinase complexes CycD/Cdk4,6 and CycE/Cdk2. The initial phosphorylation of Rb releases E2F, which up-regulates the kinase CycE/Cdk2, which then completes the phosphorylation of Rb and finishes the release of E2F[1,15–19]. In Fig. 2a the node Rb represents complex E2F-Rb and node E2F a free E2F that is able to act as a transcription factor. Interestingly, the mammalian network and the yeast *S. cerevisiae* (Fig. 2b) network have identical structure in spite of the fact that individual genes having limited homology[20,21]. The central dynamical feature of both of these networks is ability to exhibit bistability between an **On** state where E2F is high, Rb low and CycE high, and an **Off** state where E2F is low, Rb high, and CycE is low.

Fig. 2 | G1/S restriction point network. a G1/S restriction point network in mammalian cells (**b**) START network in yeast (**c**) Reduced network example.
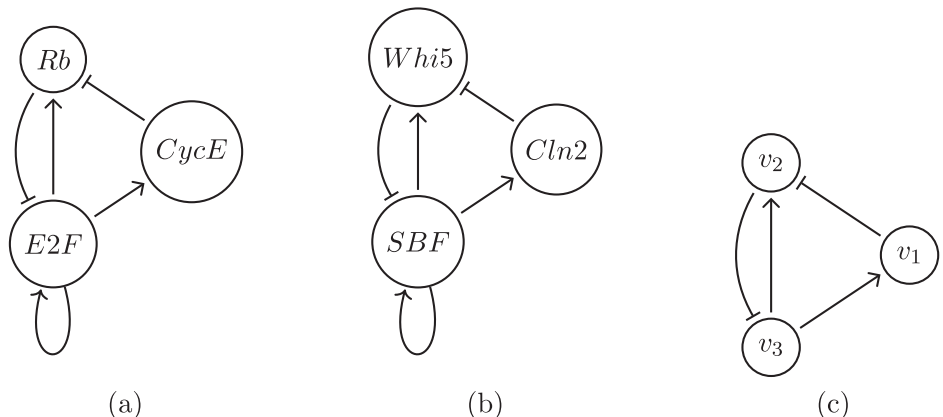
In Fig. 2c, we depict a simplified network where we removed the activating self-edge on E2F. We ask whether this simplified network still supports required bistability, and, if so, what is the prevalence of this phenotype.

A *regulatory network* $RN = (V, E, \delta)$ is a directed graph with nodes V, directed edges E, and an edge sign function $\delta: E \rightarrow \{-1, 1\}$. We denote an edge from node v_i to node v_j without indicating its sign by $v_i{\multimap}v_j$. The edge $v_i{\multimap}v_j$ is *activating* if $\delta_i^j = 1$ and *repressing* if $\delta_i^j = -1$. Graphically, an activating edge is denoted by $v_i \rightarrow v_j$ and a repressing edge by $v_i \dashv v_j$. The *sources* and *targets* of a node u are given by

$$\mathbf{S(v_i)} := \{v_k \in V | v_k v_i \in E\} \quad \mathbf{T(v_i)} := \{v_j \in V | v_i v_j \in E\},$$

respectively.

In our simplified example in Fig. 2c we have $\mathbf{T}(v_3) = \{v_1, v_2\}$ and $\mathbf{S}(v_3) = \{v_2\}$.

Parameters

In this section, we want to define the set of boolean functions that are compatible with the network RN. Consider the toggle triad example where we associated a boolean update function f_a to node a. Node a is a source of two edges to b and c. In real regulatory network, the chemical concentration x_a at node a will likely affect b and c at different levels, which we think of as different thresholds. Therefore the state of node a should have more than two states 0 (inactive) and 1 (active); it should at least have states 0 (does not activate neither b nor c), 1 (activates one but not the other), and 2 (activates both b and c). This leads naturally to considering *multivalued monotone boolean networks* where the state of a node v_i is one of the integers $X_i = \{0, 1\ldots, t_i\}$, where $t_i := |\mathbf{T}(v_i)|$ denote the number of target nodes of v_i. Since each edge $v_i \rightarrow v_j$ is associated with a threshold, this integer represents number of thresholds that get activated by the state of v_i.

This extension to multilevel boolean networks enlarges the set of functions that are compatible with RN. Instead of calling all such collections "multilevel monotone boolean networks", we will simply call them *DSGRN parameters* of RN, or just *parameters* of RN. Because of different activation thresholds, the dynamics of the network may change when the order of thresholds changes. Even the same collection of multilevel monotone boolean functions, different order of activation of downstream edges may lead to different dynamics. Therefore such orders must be included in the description of the parameters of the network.

An *order parameter* for node v_i is a bijection $\theta_i: \mathbf{T}(v_i) \rightarrow \{1, \ldots, t_i\}$ which defines an ordering of the out-edges of v_i. The set of order parameters for v_i is denoted by $\Theta(v_i)$. The set of all order parameters is given by $:= \prod_{v_i \in V}(v_i)$.

For the network in Fig. 2c, since $|\mathbf{T}(v_1)| = |\mathbf{T}(v_2)| = 1$ and we have $X_1 = X_2 = \{0, 1\}$ there is a single order parameter in both $\Theta(v_1) = \{\theta_1\}$ and $\Theta(v_2) = \{\theta_2\}$. Since $|\mathbf{T}(v_3)| = 2$, we have $X_3 = \{0, 1, 2\}$ and $(v_3) = \{\theta_3^1, \theta_3^2\}$ where

$$\theta_3^1(v_1) = 1, \quad \theta_3^1(v_2) = 2 \quad \text{and} \quad \theta_3^2(v_1) = 2, \quad \theta_3^2(v_2) = 1. \quad (2)$$

The collection of all order parameters Θ has two elements $(\theta_1, \theta_2, \theta_3^1)$ and $(\theta_1, \theta_2, \theta_3^2)$. We note that if $|\mathbf{T}(v_3)| = k$ the collection $\Theta(v_3)$ will have $k!$ permutations $\theta_3^1, \ldots, \theta_3^{k!}$.

Let $\mathbb{B} := (\{0, 1\}; 0 < 1)$ be a two element set with natural order $0 \prec 1$ and let $\mathbb{B}^n$ be a partially ordered set (poset) of n vectors with elements in $\mathbb{B}$ with order induced component-wise by $<$. $\mathbb{B}^n$ is in fact a Boolean lattice.

Similarly, the

$$\mathcal{C} := \prod_{v_i \in V} X_i.$$

is a Boolean lattice of integer vectors with partial order $\prec$ induced component-wise by $<$. That is, $a \prec b$ if for every i the i-th component satisfies $a_i \leq b_i$.

While the assumption that each v_i activates downstream nodes $v_j \in \mathbf{T}(v_i)$ at different thresholds leads to considering the set of states $\mathcal{C}$, the activation of a particular node $v_j \in \mathbf{T}(v_i)$ by v_i only happens at a single threshold. This leads to definition of the following function that associates to each state $c \in \mathcal{C}$ the information on which targets it actually activates and which ones it does not. The state $c = (c_i)_{v_i \in V}$ produces an input to a node v_j via the *input map*

$$B^j : \mathcal{C} \rightarrow \prod_{i \in V} \mathbb{B}^{|\mathbf{S}(v_j)|}, \qquad B^j := (B_i^j)_{v_i \in \mathbf{S}(v_j)}$$

where

$$B_i^j(c) := \begin{cases} 0, & \text{if } c_i < \theta_i(v_j) \text{ and } \delta_i^j = 1 \text{ or } c_i \geq \theta_i(v_j) \text{ and } \delta_i^j = -1 \\ 1, & \text{if } c_i \geq \theta_i(v_j) \text{ and } \delta_i^j = 1 \text{ or } c_i < \theta_i(v_j) \text{ and } \delta_i^j = -1 \end{cases}.$$

$$(3)$$

Note that for activating edge $v_i \rightarrow v_j$, if c_i is below (above) the activating threshold $\theta_i(v_j)$ then the input from v_i to v_j is 0 (1). This assignment is reversed if the edge is repressing. Function $B = (B^j)_{j \in V}$ depends on the structure of the network through functions δ_i^j, the number of inputs to each node and logic parameter θ.

The function B associates to each $c \in \mathcal{C}$ and each node $v_j \in V$ its boolean input vector which is an element of $\mathbb{B}^{|\mathbf{S}(v_j)|}$. Whether a particular boolean input activates an output edge $v_j{\multimap}v_i$ connecting v_j to $v_i \in \mathbf{T}(v_j)$ or not, is determined by a *logic parameter* at node v_j that we define next. Logic parameters capture all potential patterns of combinatorial activation, where only some combinations of input nodes activate an output edge. In addition, these patterns may vary from one output edge of v_j to the next.

Definition 1.1. A function $g : \mathbb{B}^n \rightarrow [0, 1, \ldots, k]$ is a *positive multivalued monotone Boolean function (mMBF)* if $b^1 \prec b^2$ implies $g(b^1) \leq g(b^2)$. When $k = 1$ the function g is *positive monotone Boolean function (MBF)*.

A *logic parameter* for node v_i is a positive boolean mMBF

$$g_i : \mathbb{B}^{|\mathbf{S}(v_i)|} \rightarrow X_i.$$

A collection $g := (g_i)_{v_i \in V}$ is a *logic parameter*. The set of all logic parameters for node v_i is denoted $\mathcal{L}(v_i)$, while the set of all logic parameters is $\mathcal{L} := \prod_{v_i \in V} \mathcal{L}(v_i)$.

Definition 1.2. Consider two logical parameters $g, h \in \mathcal{L}(v_i)$, where $g, h : \mathbb{B}^n \rightarrow X_i$. We say $g \prec h$ if $g(b) \leq h(b)$ for all $b \in \mathbb{B}^n$. With this ordering the set of logic parameters $(\mathcal{L}(v_i), \prec)$ is a *partially ordered set*.

We describe the set of logic parameters for the network in Fig. 2c. At node v_1 since $\mathbf{S}(v_1) = v_3$ and $X_1 = \{0, 1\}$, the set of logic parameters is the set of all MBFs $g_1 : \mathbb{B} \rightarrow X_1$. Two of these functions are constant: we denote by $\mathbf{0}$ the zero function and by $\mathbf{1}$ the one function. The third function, which we denote by Id maps 0 to 0 and 1 to 1 see Fig. 3a. These functions form a poset shown in Fig. 4a.

Fig. 3 | Logic parameters for the restriction point network. a Three logic parameters $g_1 \in \mathcal{L}(v_1)$; **b** Six logic parameters $g_3 \in \mathcal{L}(v_3)$.
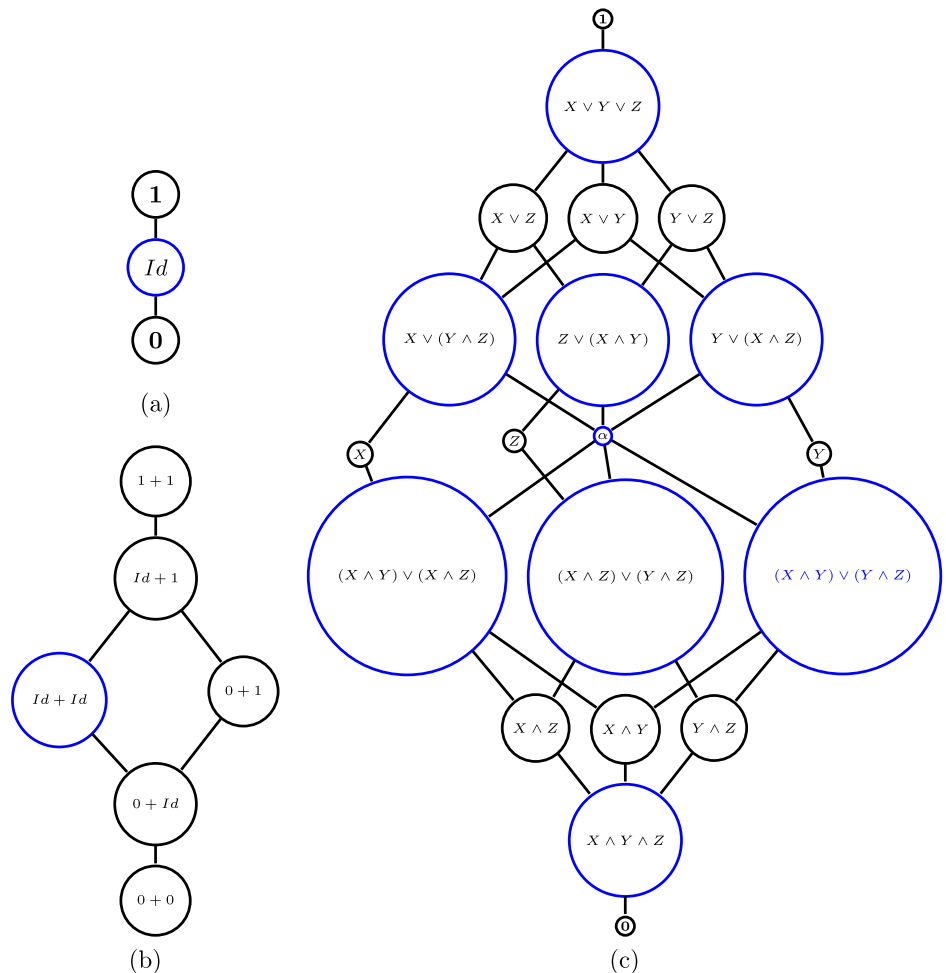
	0	Id	**1**
0	0	0	1
1	0	1	1

(a)

	0 + 0	**0+Id**	**0 + 1**	Id+Id	Id+1	**1 + 1**
0	0	0	1	0	1	2
1	0	1	1	2	2	2

(b)

Fig. 4 | Logic parameter graphs. a Logic parameter graph $\mathcal{L}(v_1)$ consists of MBFs $g_1 : \mathbb{B} \to \{0, 1\}$; **b** logic parameter graph $\mathcal{L}(v_3)$ consists of mMBFs $g_3 : \mathbb{B} \to \{0, 1, 2\}$ written as sums of two ordered functions (2-chains) from (**a**); **c** poset of MBFs with three inputs X, Y, Z, where $\alpha = (X \wedge Y) \vee (X \wedge Z) \vee (Y \wedge Z)$. Blue circles indicate non-degenerate MBFs.

Consider v_2. Since $\mathbf{S}(v_2) = \{v_1, v_2\}$ and $X_2 = \{0, 1\}$, the set $\mathcal{L}(v_2)$ is the set of MBFs that map $g_2 : \mathbb{B}^2 \to X_2$. There are 6 such functions which we denote by $\mathbf{0}, \wedge, X, Y, \vee, \mathbf{1}$. These were depicted in Fig. 1b, c.

Finally, consider the node v_3. Since $\mathbf{S}(v_3) = \{v_2\}$ and the set $X_3 = \{0, 1, 2\}$, the logical parameters are multivalued monotone boolean functions $g_3 : \mathbb{B} \to X_3$. It can be shown[5,6], that there are 6 such mMBFs that correspond to ordered pairs (2-chains) in the poset of MBFs from $\mathbb{B}$ to $\mathbb{B}$. Therefore the functions g_3 are sums $f + g$ of pairs of functions $f \prec g$ in the set $f, g \in \{\mathbf{0}, Id, \mathbf{1}\}$ in Fig. 3a. We list these function using the sum notation in Fig. 3b; the poset structure of these functions is in Fig. 4b.

The set of all parameters is the product of logic and order parameters $\mathcal{P} := \mathcal{L} \times$. We call $\mathcal{P}(v_i) := \mathcal{L}(v_i) \times (v_i)$ the set of parameters for node v_i; the set of all parameters is the product $\mathcal{P} = \prod_{v_i \in V} \mathcal{P}(v_i)$.

We have seen that the set of logical parameters has an additional structure of a partially ordered set, or a graph. In our example, the poset $\mathcal{L}(v_1)$ is in Fig. 4a, poset for $\mathcal{L}(v_2)$ is in Fig. 1c, and $\mathcal{L}(v_3)$ is in Fig. 1c. We will call this organization of the set of DSGRN parameters $\mathcal{P}$ a *parameter graph* PG of network RN. The following section shows the construction of PG by defining adjacency between elements of $\mathcal{P}$.

Parameter graph

The parameter graph PG has nodes and edges. The set of nodes is the set of DSGRN parameters $\mathcal{P}$. The undirected edges will connect adjacent nodes. Two parameter nodes for a node v_i, $(g_i, \theta_i), (\hat{g}_i, \hat{\theta}_i) \in \mathcal{P}(v_i)$ are *adjacent* if exactly one of the following conditions is satisfied.

- *Order adjacency*: $g_i = \hat{g}_i$ and the values of the order parameters θ_i and $\hat{\theta}_i$ are exchanged on a single pair of neighboring entries on which the logic parameters agree.

- *Logical adjacency*: $\theta_i = \hat{\theta}_i$ and the logic parameters g_i and $\hat{g}_i$ differ by 1 in a single input i.e., there exists exactly one $d \in \mathbb{B}^{|\mathbf{S}(v_i)|}$ such that $g_i(d) = \hat{g}_i(d) \pm 1$.

The *factor graph* for node v_i is the undirected graph $\mathsf{PG}(v_i) := (\mathcal{P}(v_i), \mathcal{E}(v_i))$ whose nodes are parameter nodes for v_i and whose edges are given by adjacency. The *parameter graph* $\mathsf{PG} := (\mathcal{P}, \mathcal{E})$ is the Cartesian product $\mathsf{PG} := \prod_{v_i \in V} \mathsf{PG}(v_i)$. That is, there is an edge $(p^1, p^2) \in \mathcal{E}$ if and only if there is a unique $v_i \in V$ such that $(p_i^1, p_i^2) \in \mathcal{E}(v_i)$ and $p_i^1 = p_i^2$ otherwise.

We return to our example in Fig. 2c. Since the set of order parameters $\Theta(v_1)$ and $\Theta(v_2)$ have both only one element, the factor graph $\mathsf{PG}(v_1) \cong \mathcal{L}(v_1)$ is isomorphic to the poset of logic parameters $\mathcal{L}(v_1)$ in Fig. 4a and factor graph $\mathsf{PG}(v_2) \cong \mathcal{L}(v_2)$ is isomorphic to the poset of logic parameters $\mathcal{L}(v_2)$ in Fig. 1c. The factor graph $\mathsf{PG}(v_3)$ consists of two copies of the poset of logic parameters $\mathcal{L}(v_3)$ in Fig. 4b, where one copy has order parameter θ_3^1 and the copy has order θ_3^2, see (2). The two copies are connected between pairs of nodes $\mathbf{0} + \mathbf{0}$, $Id + Id$ and $\mathbf{1} + \mathbf{1}$ where the change from θ_3^1 to θ_3^2 satisfies oder adjacency condition.

In Fig. 4c we show a logic factor graph for a network node with three inputs and one output, which consists of 20 monotone boolean functions. The number of MBFs, with n inputs, called *n-th Dedekind number*, increases rapidly with the number of inputs. The 9th Dedekind number has been only recently computed[22].

Essential parameters

There are two types of subsets of the logical parameters that are of special interest. First, we may only be interested in those logical parameters at which

the associated multivalued monotone boolean function is *non-degenerate*. The definition below directly generalizes the concept of non-degenerate MBF[23,24].

Definition 1.3. A variable b_k is an *essential variable* of a multivalued monotone Boolean function f if there is at least one $\mathbf{b} \in B^n$ such that $f|_{b_k=0} \neq f|_{b_k=1}$. An MBF is said to be *non-degenerate* if all variables are essential.

Definition 1.4. A parameter $p \in PG$ of network RN is *essential* if the corresponding logic parameter g is a non-degenerate mMBF.

This agrees with the definition of essential parameter nodes given in refs. 2,5,25. All non-degenerate mMBFs in $\mathcal{L}(\nu_1)$ and $\mathcal{L}(\nu_2)$ are in blue circles in Fig. 4.

Boolean parameters

Another special subset of logical parameters are those that represent MBF rather than mMBF. Consider those parameters $p = (g, \theta) \in PG$ where the range of every mMBF g_i consists of only two values 0 and t_i, i.e., the lowest possible value and the highest possible value. This can be interpreted as each input either does not activate any target nodes or it activates all target nodes. Since such mMBF has only two values, it can be represented as a MBF $\tilde{g}_i$. We call such parameters *boolean parameters* since the function $\tilde{g} = (\tilde{g}_1, \tilde{g}_2, \ldots, \tilde{g}_n)$, as a map from $\mathbb{B}^n \to \mathbb{B}^n$, represents a traditional boolean system where states 0 and 1 are assigned to each variable.

In Fig. 4a, c, all parameters are boolean since the range of all functions is $\{0, 1\}$. In Fig. 4b there are three Boolean parameters $\mathbf{1} + \mathbf{1}$, $Id + Id$ and $\mathbf{0} + \mathbf{0}$. Notice that not all boolean parameters are essential.

Essential boolean parameters

We now restrict our attention even further to just *essential boolean* parameters. These represent parameters represented by non-degenerate monotone boolean function. Figure 4c shows 9 non-degenerate MBFs with three inputs in blue circles. Paper[8] studied repressive tetrahedron network (see "Toggle tetrahedron network" below) where all four nodes have three repressive inputs from the other three nodes, and in turn repress all other nodes. Each node has three inputs and three outputs and the parameter graph PG has about 27 trillion nodes. However, since there are only nine non-degenerate MBFs, the total number of essential boolean parameters is only $9^4 = 6561$. Surprisingly, the frequency of observed equilibria that exists within this small set approximates well the overall frequency of dynamics across entire PG, as documented by random DSGRN parameter sampling as well as ODE parameter sampling by RACIPE[8]. Clearly, analyzing dynamics of 6561 parameters is computationally feasible while examining 27 trillion parameters is not. The reason why this small sample seems to match the overall dynamics remains an open problem.

Dynamics

The multi-valued boolean dynamics associated to a network RN = (V, E, δ) depends on a choice of parameter $p \in \mathcal{P}$, and on the input function B defined in (3) that reflects the position of the positive and negative in the network through function δ.

The dynamics occur on the state space $\mathcal{C} = \prod_{\nu_i \in V} X_i$, introduced earlier. We call $c \in \mathcal{C}$ a *state* of the network RN.

We return to the E2F-Rb example and define the functions $B^1 : X_3 \to \mathbb{B}$, $B^2 : X_1 \times X_3 \to \mathbb{B}^2$ and $B^3 : X_2 \to \mathbb{B}$, see Fig. 5 for the order parameter $(\theta_1, \theta_2, \theta_3^1)$, where $\theta_3^1(\nu_1) = 1$ and $\theta_3^2(\nu_2) = 2$. We only list relevant inputs $X_i \subset \mathcal{C}$ for each i. Since the edge $\nu_3 \to \nu_1$ is activating and since $\theta_3^1(\nu_1) = 1$, ν_3 activates ν_1 at the first threshold the output value changes from 0 to 1 at the first threshold (see Fig. 5a.) Because the edge from ν_1 to ν_2 is repressive, the first component B_1^2 of the function B^2 reverses the boolean input from node ν_1. The second component B_3^2 describes input from node ν_3 to node ν_2 which is activating at the second threshold, see Fig. 5b. Finally, the function B^3 in Fig. 5c reflects the fact that the edge $x_2 \dashv x_3$ is repressive.

Fig. 5 | Functions B^1, B^2, B^3, where we only list the relevant inputs. Note that the presence of negative edge between ν_1 and ν_2 is manifested in the reversal in the first coordinate B_1^2 of B^2 and negative edge between ν_2 and ν_3 is reflected in reversal in function B^3.

Definition 1.5. The dynamics for network RN with the sign function δ at parameter $(g, \theta) \in \mathcal{P}$ is defined as an *asynchronous update* of function $f := g \circ B$. More precisely,

1. The *multi-valued boolean map* $f : \mathcal{C} \to \mathcal{C}$ is defined by

$$f_i(c) := g_i(B^i(c)) \quad (4)$$

2. The *multi-level boolean dynamics* $\mathcal{F} : \mathcal{C}\mathcal{C}$ is a multi-valued map generated by f and defined by
   - If $f(c) = c$ then $\mathcal{F}(c) = \{c\}$.
   - For any ν_i and $\eta \in \{-1, 1\}$ satisfying $\eta f_i(c) > \eta c_i$ the state

$$\bar{c}_i = c_i + \eta, \quad \bar{c}_j = c_j \text{ for } j \neq i$$

   satisfies $\bar{c} \in \mathcal{F}(c)$.

The maps f and $\mathcal{F}$ depend on the choice of network RN and the choice of parameter $(g, \theta) \in \mathcal{P}$. We will explicitly include these dependencies as arguments as needed.

This definition provides a connection between each parameter $(g, \theta) \in \mathcal{P}$ and a discrete dynamics on $\mathcal{C}$ given by map $\mathcal{F}$. As we show in "Connecting parameter graph PG to ODE models" each such map $\mathcal{F}$ also represents behavior of continuous solutions of switching ODE system that models continuous network dynamics. Therefore the parameter graph PG connects continuous dynamics of ODEs and asynchronous update dynamics induced by discrete map $\mathcal{F}$.

Remark 1.1. Consider all boolean parameters $p = (g, \theta) \in PG$ where the logical parameter g gives rise to the same boolean function $\tilde{g}$, but where they differ in the order parameter. Then it is easy to see that the dynamics at all these parameters will be the same since the order of thresholds is irrelevant for the resulting update function $f(c) = \tilde{g}(B(c))$. Therefore boolean parameters are fully described by their logic parameters and such logic parameters correspond to collections of mMBFs.

Morse graph

The recurrent dynamics of $\mathcal{F}(\cdot; p)$ are encoded by a *Morse graph* MG(p). The Morse graph MG = (SCC, A) is a directed graph with nodes consisting of strongly connected components of STG$(\mathcal{C}, p)$. The Morse graph is the Haase diagram on SCC of the reachability relation A on the strongly connected components within STG$(\mathcal{C}, p)$. We label each strongly connected component $s \in SCC$ according to the following.

- If $s \in SCC$ consists of a single recurrent state, $s = \{x\}$, then x is a fixed point of $\mathcal{F}$ and we label s by FP(x).
- If $s \in SCC$ is not an FP then we label s as a *partial cycle* PC or a *full cycle* FC. The strongly connected component s is a PC if s is constant in at least one coordinate i.e., there is a node $u \in V$ and an integer k such

(X_1, X_3)	**0**	∧	X_1	X_3	∨	**1**
(0,0)	0	0	1	0	1	1
(0,1)	0	1	1	1	1	1
(0,2)	0	1	1	1	1	1
(1,0)	0	0	0	0	0	1
(1,1)	0	0	0	1	1	1
(1,2)	0	0	0	1	1	1

(a)

	0 + 0	**0+Id**	**0 + 1**	Id+Id	Id+1	**1 + 1**
0	0	1	1	2	2	2
1	0	0	1	0	1	2

(b)

Fig. 6 | Asynchronous update functions. a Compositions $f_2 = g_2 \circ B^2$ for all 6 choices of function g^2. **b** Compositions $f_3 = g_3 \circ B^3$ for all six choices of function g^3.

that $x \in s$ implies $x_u = k$. If s is not an FP or an PC then s is an FC. If $s \in$ SCC has no out-edges in MG(p), we call s *stable Morse node*. Otherwise, s is *unstable*.

Applications

We now illustrate our approach on three examples. First, we investigate the ability of E2F-Rb network without self-loop on E2F (Fig. 2c) to support the bistable phenotype between **On** and **Off** state that characterizes the switch-like entry into S phase in the cell cycle. Then we look at two networks that have been studied in the context of differentiation of immune cells subtypes: toggle triad and toggle tetrahedron. For toggle triad we look at prevalence of different types of bistability and for toggle tetrahedron we summarize the results on prevalence of different types of equilibria, that are described in more details in ref. 8.

In all of these examples we focus our attention to essential boolean parameters as this small set is amenable to theoretical analysis. The analysis across entire parameter graph PG is possible using DSGRN software[26,27].

E2F-Rb network

In this section, we find fixed points for the E2F-Rb network at different parameters. Fixed points fulfill the satisfiability condition

$$s \in \mathcal{C} \text{ is a fixed point, iff } f_i(s) = g_i(B^i(s)) = s_i, \quad i = 1, 2, 3, \quad (5)$$

where $g_i \in \mathcal{L}(v_i)$, the set of logical parameters. Recall that $\mathcal{L}(v_1)$ is in Fig. 4a, $\mathcal{L}(v_2)$ in Fig. 1c, and the $\mathcal{L}(v_3)$ in Fig. 4b. The functions B^i are listed in Fig. 5. The set of composite functions $f_1 := g_1 \circ B^1 \colon X_3 \to X_1$, where $g_1 \in \mathcal{L}_1$ is identical to that in Fig. 3a. All possibilities for the second composite function $f_2 := g_2 \circ B^2 \colon X_1 \times X_3 \to X_2$, for all choices of g_2, are listed as columns in Fig. 6a. We list all compositions $f_3 := g_3 \circ B^3 \colon X_2 \to X_3$ in Fig. 6b.

These tables make direct verification of Eq. (5) possible, albeit for larger network this poses a combinatorial problem as the satisfiability of (5) is equivalent to logical satisfiability problem[28] which is NP complete[29,30].

We consider two potential fixed points that are biologically important. First the state $s_{on} = (1, 0, 2)$ represents the state **On**; that is, a commitment to transition from G1 to S phase of the cell cycle, since both v_1 (CycE) and v_3 (free E2F) are at their highest states, and v_2 (E2F-Rb dimer) is at the lowest state. Consider two parameters $p_1 := (L_1, \theta)$ and $p_2 := (L_2, \theta)$ with the order parameter $\theta := (\theta_1, \theta_2, \theta_3^1)$ used above and the only two essential logic parameters

$$L_1 = (g_1 = Id, g_2 = \wedge, g_3 = Id + Id), \quad L_2 = (g_1 = Id, g_2 = \vee, g_3 = Id + Id).$$

Then for parameter p_1 we get

$$f(s_{on}) = f(1, 0, 2) = (f_1(2), f_2(12), f_3(0)) = (1, 0, 2),$$

so s_{on} is a fixed point, but for parameter p_2 we get

$$f(s_{on}) = f(1, 0, 2) = (f_1(2), f_2(12), f_3(0)) = (1, 1, 2).$$

Therefore s_{on} is a fixed point for p_1, but not for p_2. Similar calculation shows that $s_{off} = (0, 1, 0)$ that represents the **Off** state where cell is pausing in G1 phase, is a fixed point under p_2, but not under p_1.

We remark that there are two other essential parameters $p_3 := (L_1, \bar{\theta})$ and $p_4 := (L_2, \bar{\theta})$ with the same logical parameters L_1, L_2, but order parameter $\bar{\theta} = (\theta_1, \theta_2, \theta_3^2)$. The results are similar: one of the parameters supports only s_{on} and one of them supports only s_{off}.

Since E2F-Rb network is assumed to act as a bistable switch[1,15,16] we would like to investigate if there are other, non-essential parameters where both s_{on} and s_{off} are fixed points. Examining Fig. 6 for columns where $f_3(12) = 0$ and $f_3(00) = 1$ we find that the only such function arise from logical parameter $g_2(X_1, X_3) = X_1$. This represents function where the input to v_2 from v_3 does not affect the outcome since the function g_2 only depends on the input from v_1. Erasing the edge $v_3 \to v_2$ from the network we get a network that consists of single positive loop. Such a reduced network is known to support bistablity.

Our analysis can be interpreted in two ways:
- the smaller network consisting of a positive loop $v_1 \dashv v_2$, $v_2 \dashv v_3$ and $v_3 \to v_1$ supports the bistability;
- the self-edge E2F to E2F is needed for the original network to support the bistability at the set of essential parameters.

Both interpretations provide valuable insight into interplay between the structure of the network and bistability.

Developmental networks
Multistability in toggle triad

We now investigate parameters that support bistability and multistability in toggle triad. We only consider essential boolean parameters which, by Remark 1.1, can be represented as $g = (g_a, g_b, g_c)$ where each g_i is a non-degenerate monotone boolean function.

Using the update functions $f_i = g_i \circ B^i$ for $i = a, b, c$, where the $B^i(XY) = (\neg X \neg Y)$ reverses both inputs (see Fig. 1b), the essential boolean parameters that support so called *mirror* bistability between (001) and (110) must satisfy

$$f_a(01) = 0, \qquad f_a(10) = 1$$
$$f_b(01) = 0, \qquad f_b(10) = 1$$
$$f_c(00) = 1, \qquad f_c(11) = 0$$

Direct inspection of table in Fig. 1b shows that there is unique choice of functions for both $g_1(X, Y) = Y$ and $g_2(X, Y) = Y$. On the other hand any $g_3(X, Y) \in \{\vee, Y, X, \wedge\}$ works. Therefore there are 4 boolean parameters which support this type of bistability. However, none of these parameters are essential.

On the other hand, non-mirror bistability between (001) and (100) is supported by all parameters for which

$$f_a(01) = 0, \qquad f_a(00) = 1$$
$$f_b(01) = 0, \qquad f_b(10) = 0$$
$$f_c(00) = 1, \qquad f_c(10) = 0.$$

Then any combination of choices $g_a(X, Y) \in \{\wedge, Y\}$, $g_b(X, Y) \in \{0, \wedge,\}$ and $g_c(X, Y) \in \{\wedge, X\}$ satisfy these equations. These combinations form eight parameters supporting this bistability, and one of them, $g = (\wedge, \wedge, \wedge)$ is an essential boolean parameter.

Can the network support tristability? The natural candidates for three fixed points are (001), (100), (010). This adds additional conditions (third column below) to the conditions above

$$f_a(01) = 0, \qquad f_a(00) = 1 \quad f_a(10) = 0$$
$$f_b(01) = 0, \qquad f_b(10) = 0 \quad f_b(00) = 1$$
$$f_c(00) = 1, \qquad f_c(10) = 0. \quad f_c(01) = 0$$

Direct calculation shows that $g = (\wedge, \wedge, \wedge)$ supports the tristability (001), (100), (010). Similar calculations shows that the only parameter that supports tristability between (110), (011), (101) is $g = (\vee, \vee, \vee)$.

We close this part by discussion of all 8 essential boolean parameters. From the discussion in the introduction, the equilibrium (001) is supported by two essential parameters $g = (\wedge, \wedge, \vee)$ and $g = (\wedge, \wedge, \wedge)$. By symmetry, the equilibrium (010) is supported by $g = (\wedge, \vee, \wedge)$ and $g = (\wedge, \wedge, \wedge)$ and, finally, the equilibrium (100) by $g = (\vee, \wedge, \wedge)$ and $g = (\wedge, \wedge, \wedge)$. The second one of these, logical parameters $g(\wedge, \wedge, \wedge)$, supports tristability. Similar argument shows that the parameters with two copies of $g_i = \vee$ and one copy of $g_i = \wedge$ support a single equilibrium in the set (110), (011), (101). We conclude that no essential boolean parameter supports bistability in the toggle triad. While we examined a very small subset of essential boolean parameters, our results are consistent with[14] which found that the prevalence of tristability among (100), (100), (010) is greater than any other type of tristability and that bistability is comparatively rare.

The methodology in this paper also allows us to ask us what happens if we perturb parameters away from essential boolean parameters. Perturbing essentiality leads to consideration of 216 choices of boolean functions discussed in the introduction. Perturbing boolean functions to the class of multivalued boolean functions is possible within the structure of the parameter graph and will result in potentially different dynamics.

Toggle tetrahedron network

In this section, we briefly review extensions of the results from toggle triad to toggle tetrahedron[8]. Motivation for studying this network is differentiation of naive CD4+ T cells into four different types denoted Th1, Th2, Th17, and Treg. Each of these four types of cells is characterized by a lineage specific transcription factors and these factors repress each other.

Therefore toggle tetrahedron has four nodes a, b, c, d that are fully connected without self-edges and each node receives three repressive inputs from other three nodes. We again focus on computing the number of essential boolean parameters $g = (g_a, g_b, g_c, g_d)$, where each g_i is a non-degenerate monotone boolean function, that support a particular type of a steady state. The types we are interested in include all-high (1111) and all low (0000) states, as well as states with one (1000), two (1100) or three (1110) active components. Poset of all MBFs with three inputs in Fig. 4c has 20 functions with 9 non-degenerate MBF marked in blue[6,24].

We now summarize results from[8].

We first note that, similarly to toggle triad, the only parameters that support constant equilibria (0000) and (1111) are $(0, 0, 0, 0)$ and $(1, 1, 1, 1)$, respectively. Further, by symmetry for every parameter supporting equilibrium (1000) there is a parameter that supports equilibrium (0111), since the logic parameter functions are simply negated.

Therefore we only need to consider equilibria of type 3–1 where three genes have different expression levels than the fourth gene and the equilibria of type 2–2 where two genes are active and two are inactive. Similar analysis to the one for toggle triad gives the following.

Theorem 0.1. ([8]) Out of total of $9^4 = 6561$ essential boolean parameters, there are

- $2 * 2 * 2 * 9 = 72$ essential boolean parameters that support any 3–1 equilibrium.

- $7^4 = 2401$ essential boolean parameters that support any 2–2 equilibrium.

As discussed in detail in ref. 8 significantly higher prevalence of 2–2 equilibria than prevalence of 3–1 equilibria may indicate that the direct differentiation from precursor cell into individual cell types, represented by a 3–1 state, is less likely that a two step differentiation, where in the first step cells attain a mixed state represented by a 2–2 equilibrium, followed by a subsequent differentiation to individual cell types.

Connecting parameter graph PG to ODE models

The *switching system* dynamics[3,10,31–39] associated to a regulatory network RN is a system of ordinary differential equations

$$\dot{x}_i = \Lambda_u(x) - \gamma_i x_i, \quad i \in V \qquad (6)$$

where $x = (x_i)_{v_i \in V} \in \mathbb{R}^N_+$, $\gamma_i \in \mathbb{R}$ is the decay rate of x_i, and Λ_i is a piecewise constant function which captures the effect of the sources $\mathbf{S}(i)$ on the node i. The function $\Lambda = (\Lambda_1, \ldots, \Lambda_n)$ is defined for parameter $p = (g, \theta) \in \mathcal{P}$ as follows.

1. We associate a continuous variable x_i to each node $v_i \in V$.
2. We associate a threshold values $\theta_{ji}, j \in \mathbf{T}(i)$ to each edge $i \to j$ and assume that these thresholds are distinct $\theta_{ji} \neq \theta_{ki}$ for any $j, k \in \mathbf{T}(i)$.
3. The thresholds θ_{ji} form a rectangular grid $G := \{x \in \mathbb{R}^N_+ \mid x = \theta_{ji}, u \in V, j \in \mathbf{T}(i)\}$. The set $\mathbb{R}^n_+ \setminus G$ is a collection of a finite number of open *domains* $\mathcal{D}$ where $x \in \mathcal{D}$ if all components of vector x lie between the thresholds. Observe that a collection of all domains $\mathcal{D}$ is in one-to-one correspondence with space $\mathcal{C}$. This is expressed via a map

$$\varphi(x) : \mathbb{R}^n_+ \setminus G \to \mathcal{C} \qquad x \mapsto (k_1, \ldots, k_n),$$

where k_i is an integer $k_i \in X_i$ such that $k_i < x_i < k_i + 1$ for all i. This map associates to each $x \in \mathbb{R}^n_+ \setminus G$ a signature $c \in \mathcal{C}$ of its domain $d \in \mathcal{D}$.

4. Then we set

$$\Lambda_i(x) := \gamma_i f_i(\varphi_i(x)) = \gamma_i g_i(B(\varphi_i(x))) \qquad (7)$$

In an open domain $d \in \mathcal{D}$, the function Λ is constant and the flow of (6) is directed toward the *target point* $\Lambda(x)$. All trajectories in d are straight lines towards the target point. If the target point is contained in d then the target point is an asymptotically stable fixed point of (6). If the target point is not in d, then the trajectories continue in a straight line until they hit the boundary of d. For a generic set of initial condition in d, trajectory hits a co-dimension one boundary of d where $x = \theta_{ji}$ for single threshold of θ_{ji}. If $j \neq i$, then the sign of $\dot{x}_i$ does not change at $x_i = \theta_{ji}$ and the trajectory can be extended by continuation into a new domain d'. If $i = j$ and the edge $i \dashv j$ is repressing, then it is possible the sign of $\dot{x}_i$ may change on $x_i = \theta_{ii}$. However, since only one component of Λ changes at θ_{ii}, all the components of the vector field $\Lambda_k(x)$, $k \neq i$ remain the same between d and d'. Therefore a sliding motion along the hyperplane $x_i = \theta_{ii}$ between d and d' is well defined. As a consequence, if the target point of d does not lie in d, for generic set of initial conditions in d, the solutions can be continued to some neighboring domain d'. This observation has been used in ref. 2 to define state transition graphs even for systems with negative self-edges.

This description shows that the dynamics of (6) are well defined for every parameter $p = (g, \theta) \in \mathcal{P}$ and determined by the target point function $f = g \circ B$, see eq. (4). Furthermore it is easy to see that the trajectories of (6) that exit domain d may enter any domaing $\mathcal{F}(d)$. Therefore transition of the state transition graph defined by $\mathcal{F}$ capture all possible transitions by solutions of (6). It follows, that the Morse nodes denoted by FP of MG contain fixed points of ODE system (6) and any Morse node PC or FC has a potential to contain periodic solutions of (6).

The precise correspondence between *invariant sets* of (6) which are central objects in study of dynamical systems[40] and the Morse nodes is complex and beyond the scope of this paper[11]. The ongoing current work

aims to show that the Morse graph recovers Morse decomposition of a wide class of smooth ordinary differential equations that are approximated by the switching system (6). We describe briefly the main ideas for a restricted class of functions, where Λ has a form of *product-of-sums*[41–44] and which have been used extensively in DSGRN[3,5,26,27]. In product-of-sums systems, the functional form for Λ_u is restricted to be a product of sums of switching functions

$$\Lambda_i(x) = \prod_{j \in \mathbf{S}(i)} \sum \sigma_{ij}(x_j) \qquad \sigma_{ij}(x_j) := \begin{cases} L_{ij}, & \text{if } \delta_{ij}(x_j - \theta_{ij}) < 0 \\ U_{ij}, & \text{if } \delta_{ij}(x_j - \theta_{ij}) > 0 \end{cases} \tag{8}$$

where $0 < L_{ij} < U_{ij}$ are the lower (L) and upper (U) values for the effect of v_j on v_i. The advantage of the product-of-sums description for Λ is that the parameters L and U are easy to interpret in applications.

In particular, for every function σ_{ij} there is a sequence of Hill functions, parameterized by the Hill parameter n, of the form

$$h_{ij}^n(x_j) := L_{ij} + (U_{ij} - L_{ij})\frac{x_i^n}{\theta_{ij}^n + x_j^n} \tag{9}$$

such that

$$\lim_{n \to \infty} h_{ij}^n(x) = \sigma_{ij}(x) \qquad \text{pointwise.}$$

This allows comparison between ODE system with Hill functions and switching systems. Since the repertoire of long-term dynamics of switching system associated to a network RN is determined by collection of Morse graphs, parameterized by all multi-valued MBFs in parameter graph PG, DGSRN provides a bridge between continuous dynamics of networks and combinatorial, finite collection of Morse graphs in PG.

There is numerical evidence that DSGRN successfully predicts dynamics of ODE network dynamics[8,12]. In ref. 12, the results from DSGRN computation of equilibria were compared to results from RACIPE[9] approach that samples parameters of Hill function network models and then runs the ODE simulations. In particular, ref. 12 examines at what value of Hill coefficient n the RACIPE and DSGRN results start to agree. Surprisingly, DSGRN predicts RACIPE results even for relatively small values of n. The paper[8] considers toggle tetrahedron network which has 27 trillion DSGRN parameters which is too large for exhaustive computation. We hasten to add that computations involving several billion of parameters can be computed on a laptop in matter of hours. Two alternative approaches have been used. In one, four random samples of 10,000 DSGRN parameters from the set of all DSGRN parameters were selected and examined for different types of equilibria and different types of multistability. In the second approach the collection of all 6561 essential boolean parameters have been examined. We compared results from both of these approaches to results from RACIPE samples and again, we found good agreement between all three measurements. This is surprising as the essential Boolean parameters represents a tiny slice of the parameter space, yet it seems to predict well behavior of the network over a entire parameter space.

Since the DSGRN analysis is computationally many orders of magnitude faster that RACIPE this suggests that DSGRN is a valuable tool for the first pass analysis of the range of behaviors that the network is able to support.

Discussion

Cellular regulatory networks describe directed pairwise interactions between genes and proteins. Some small networks seems to occur statistically more frequently that others[45], which suggests that they are subject to evolutionary selection. The role of cell regulation is to dynamically respond to changes in the environment and thus dynamics supported by the regulatory networks is related to cell's fitness. It is therefore important to understand dynamics that these networks can support. Accordingly, theory of motifs[46,47] suggested that a particular dynamics of the motifs is responsible

for their overrepresentation within the set of cellular networks. However, any model of network dynamics depends on choice of parameters which represent mathematically different environmental resources, external signals as well as internal resources like number of ribosomes. Since these are difficult to measure in individual cells, it is natural to try to examine the entire range of dynamical behaviors that the network can support.

We have reviewed recent progress on the problem of describing range of dynamics supported by a network. We concentrate here on description of equilibria, or steady states, rather than more dynamic behaviors like periodic attractors. We show that there is natural connection between network models consisting of collections of multivalued monotone boolean functions and models using ordinary differential equations. These mMBFs are organized in a parameter graph PG. This structure allows us to start from a small subset of essential boolean parameters, examine dynamics at these parameters, and then explore the neighborhood of these parameters.

We examine three example networks where we discuss prevalence of different equilibria within the set of essential boolean parameters.

Our approach provides a new tool to answer the questions about range of dynamics a network may exhibit across different conditions. If this range does not include experimentally observed dynamics, the network is likely incomplete. When network does exhibit observed dynamics, its prevalence within PG may be used to rank the networks and focus experimental efforts[1,2,48], and reduce the set of potential hypotheses.

Data availability
No experimental data were used in this article. DSGRN software is available in GitHub repositories[26,27].

References
1. Gedeon, T., Cummins, B., Harker, S. & Mischaikow, K. Identifying robust hysteresis in networks. *PLoS Comput. Biol.* **14**, e1006121 (2018).
2. Gameiro, M., Gedeon, T., Kepley, S. & Mischaikow, K. Rational design of complex phenotype via network models. *PLoS Comput. Biol.* **17**, e1009189 (2021).
3. Cummins, B., Gedeon, T., Harker, S., Mischaikow, K. & Mok, K. Combinatorial representation of parameter space for switching systems. *SIAM J. Appl Dyn. Syst.* **15**, 2176–2212 (2016).
4. Cummins, B., Gedeon, T., Harker, S. & Mischaikow, K. Database of dynamic signatures generated by regulatory networks (DSGRN). In Koeppl, J. F. H. (ed.) *Computational Methods in Systems Biology*, Chap. 19, 300–308 (Springer, 2017).
5. Crawford-Kahrl, P., Cummins, B. & Gedeon, T. Joint realizability of monotone Boolean functions. *J. Theor. Comp. Sci.* **922**, 447–474 (2022).
6. Gedeon, T. Lattice structures that parameterize regulatory network dynamics. *Math. Biosci.* https://authors.elsevier.com/sd/article/S0025-5564(24)00085-3 (2024).
7. Duddu, A., Majumdar, S., Sahoo, S., Jhunjhunwala, S. & Jolly, M. Emergent dynamics of a three-node regulatory network explain phenotypic switching and heterogeneity: a case study of th1/th2/th17 cell differentiation. *Mol. Biol. Cell* **33**, 46 (2022).
8. Duddu, A. et al. Multistability and predominant double-positive states in a four node mutually repressive network: a case study of Th1/Th2/Th17/T-reg differentiation. *npj. Syst. Biol. bioRxiv.* https://doi.org/10.1101/2024.01.30.575880v1 (2024).
9. Huang, B. et al. Interrogating the topological robustness of gene regulatory circuits. *PLoS Comput. Biol.* **13**, e1005456 (2017).
10. Gedeon, T. Multi-parameter exploration of dynamics of regulatory networks. *BioSystems* **190**, 104113 (2020).
11. Gedeon, T., Harker, S., Kokubu, H., Mischaikow, K. & Oka, H. Global dynamics for steep sigmoidal nonlinearities in two dimensions. *Physica D* **339**, 18–38 (2017).

12. Hari, K. et al. Assessing biological network dynamics: comparing numerical simulations with analytical decomposition of parameter space. *NPJ Syst. Biol. Appl.* **9**, 29 (2023).

13. Gardner, T., Cantor, C. & Collins, J. Construction of a genetic toggle switch in escherichia coli. *Nature* **403**, 339–342 (2000).

14. Duddu, A., Sahoo, S., Hati, S., Jhunjhunwala, S. & Jolly, M. Multistability in cellular differentiation enabled by a network of three mutually repressing master regulators. *J. R. Soc. Interface* **17**, 20200631 (2020).

15. Yao, G., Lee, T., Mori, S., Nevins, J. & You, L. A bistable Rb-E2F switch underlies the restriction point. *Nat. Cell Biol.* **10**, 476–482 (2008).

16. Yao, G., Tan, C., West, M., Nevins, J. & You, L. Origin of bistability underlying mammalian cell cycle entry. *Mol. Syst. Biol.* **7**, 485 (2011).

17. Pardee, A. A restriction point for control of normal animal cell proliferation. *Proc. Natl Acad. Sci. USA* **71**, 1286–90 (1974).

18. Blagosklonny, M. V. & Pardee, A. B. The restriction point of the cell cycle. *Cell Cycle* **2**, 102–109 (2002).

19. Sears, R. & Nevins, J. Signaling networks that link cell proliferation and cell fate. *J. Biol. Chem.* **277**, 11617–11620 (2002).

20. Wang, H., Carey, L., Cai, Y., Wijnen, H. & Futcher, B. Recruitment of cln3 cyclin to promoters controls cell cycle entry via histone deacetylase and other targets. *PLoS Biol.* **7**, e1000189 (2009).

21. Cross, F., Buchler, N. & Skotheim, J. M. Evolution of networks and sequences in eukaryotic cell cycle control. *Philos. Trans. R. Soc. B* **366**, 3532–3544 (2011).

22. Jäkel, C. A computation of the ninth Dedekind number. *J. Comput. Algebra* **6-7**, 100006 (2023).

23. Shmulevich, I., Dougherty, E., Kim, S. & Zhang, W. Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics* **18**, 261–74 (2002).

24. Cury, J. E. R., Roxo, P. T., Manquinho, V., Chaouiya, C. & Monteiro, P. T. Immediate Neighbours of Monotone Boolean Functions. *arXiv preprint arXiv*:2407.01337 (2024).

25. Xin, Y., Cummins, B. & Gedeon, T. Multistability in the epithelial-mesenchymal transition network. *BMC Bioinformatics* **21**, 1–17 (2020).

26. Harker, S. Dsgrn software. https://github.com/shaunharker/DSGRN (2017).

27. Harker, S. & Cummins, B. Code supplemental for "identifying robust hysteresis in networks". https://github.com/shaunharker/2017-DSGRN-IdentifyingRobustHysteresisInNetworks (2017).

28. Milano, M. & Roli, A. Solving the satisfiability problem through boolean networks. In Lamma, E. & Mello, P. (eds.) *AI*IA 99: Advances in Artificial Intelligence*, 72–83 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2000).

29. Cook, S. A. The complexity of theorem-proving procedures. In *Proc. Third Annual ACM Symposium on Theory of Computing*, STOC '71, 151–158 (Association for Computing Machinery, New York, NY, USA, 1971) https://doi.org/10.1145/800157.805047

30. Trakhtenbrot, B. A survey of russian approaches to perebor (brute-force searches) algorithms. *Ann. Hist. Comput.* **6**, 384–400 (1984).

31. Glass, L. & Kauffman, S. A. Co-operative components, spatial localization and oscillatory cellular dynamics. *J. Theor. Biol.* **34**, 219–37 (1972).

32. Glass, L. & Kauffman, S. A. The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.* **39**, 103–29 (1973).

33. Glass, L. & Pasternack, J. Prediction of limit cycles in mathematical models of biological oscillations. *Bull. Math. Biol.* **40**, 27=44 (1978).

34. Snoussi, E. H. Qualitative dynamics of piecewise-linear differential equations: a discrete mapping approach. *Dyn. Stab. Syst.* **4**, 565–583 (1989).

35. Snoussi, H. & Thomas, R. Qualitative dynamics of piecewise-linear differential equations: a discrete mapping approach. *Bull. Math. Biol.* **55**, 973–991 (1993).

36. Thomas, R. Regulatory networks seen as asynchronous automata: a logical description. *J. Theor. Biol.* **153**, 1–23 (1991).

37. Thomas, R. Boolean formalization of genetic control circuits. *J. Theor. Biol.* **42**, 563–585 (1973).

38. Thomas, R., Thieffry, D. & Kaufman, M. Dynamical behaviour of biological regulatory networks-I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull. Math. Biol.* **57**, 247–76 (1995).

39. Thieffry, D. & Romero, D. The modularity of biological regulatory networks. *BioSystems* **50**, 49–59 (1999).

40. Katok, A. & Hasselblatt, B. *Introduction to Modern Theory of Dynamical Systems* (Cambridge University Press, 1995).

41. de Jong, H. et al. Qualitative simulation of genetic regulatory networks using piecewise-linear models. *Bull. Math Biol.* **66**, 301–40 (2004).

42. Ironi, L., Panzeri, L., Plahte, E. & Simoncini, V. Dynamics of actively regulated gene networks. *Phys. D Nonlinear Phenom.* **240**, 779–794 (2011).

43. Edwards, R., Machina, a, McGregor, G. & van den Driessche, P. A modelling framework for gene regulatory networks including transcription and translation. *Bull. Math. Biol.* **77**, 953–983 (2015).

44. Tournier, L. & Chaves, M. Uncovering operational interactions in genetic networks using asynchronous Boolean dynamics. *J. Theor. Biol.* **260**, 196–209 (2009).

45. Milo, R. et al. Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).

46. Alon, U. *An Introduction to Systems Biology* (Chapman & Hall/CRC, 2007).

47. Alon, U. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461 (2007).

48. Cummins, B., Gedeon, T., Harker, S. & Mischaikow, K. Model rejection and parameter reduction via time series. *SIAM J. Appl. Dyn. Syst.* **17**, 1589–1616 (2018).

Author contributions

The author conceptualized and wrote the paper.

Competing interests

The author declares no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Tomáš. Gedeon.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.