

RESEARCH ARTICLE

Noncommutative Biology: Sequential Regulation of Complex Networks

William Letsou, Long Cai*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California, United States of America

* lcai@caltech.edu

Abstract

Single-cell variability in gene expression is important for generating distinct cell types, but it is unclear how cells use the same set of regulatory molecules to specifically control similarly regulated genes. While combinatorial binding of transcription factors at promoters has been proposed as a solution for cell-type specific gene expression, we found that such models resulted in substantial information bottlenecks. We sought to understand the consequences of adopting sequential logic wherein the time-ordering of factors informs the final outcome. We showed that with noncommutative control, it is possible to independently control targets that would otherwise be activated simultaneously using combinatorial logic. Consequently, sequential logic overcomes the information bottleneck inherent in complex networks. We derived scaling laws for two noncommutative models of regulation, motivated by phosphorylation/neural networks and chromosome folding, respectively, and showed that they scale super-exponentially in the number of regulators. We also showed that specificity in control is robust to the loss of a regulator. Lastly, we connected these theoretical results to real biological networks that demonstrate specificity in the context of promiscuity. These results show that achieving a desired outcome often necessitates roundabout steps.



OPEN ACCESS

Citation: Letsou W, Cai L (2016) Noncommutative Biology: Sequential Regulation of Complex Networks. *PLoS Comput Biol* 12(8): e1005089. doi:10.1371/journal.pcbi.1005089

Editor: Gabor Balazsi, Stony Brook University, UNITED STATES

Received: February 23, 2016

Accepted: July 28, 2016

Published: August 25, 2016

Copyright: © 2016 Letsou, Cai. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research is supported in part by the US National Institutes of Health single-cell analysis program award R01HD075605. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Author Summary

DNA is the blueprint of life. Yet the order in which a cell follows these instructions makes it capable of generating thousands of different fates. How this information is extracted from underlying gene regulatory networks is unclear, especially given that biological networks are highly interconnected, and that the number of signaling pathways is relatively small (approximately 5–10). The conventional approach for increasing the information capacity of a limited set of regulators is to use them in combination. Surprisingly, combinatorial logic does not increase the diversity of target configurations or cell fates, but instead causes information bottlenecks. A different approach, called sequential logic, uses noncommutative sequences of a small set of regulators to drive networks to a large number of novel configurations. If certain targets are first protected, then even promiscuous regulators can activate specific subsets of lineage-specific targets. In this paper we show how sequential logic outperforms combinatorial logic, and argue that noncommutative

sequences underlie a number of cases of biological regulation, e.g. how a small number of signaling pathways generates a large diversity of cell types in development. In addition to explaining biological networks, sequential logic may be a general experimental design strategy in synthetic and single-cell biology.

Introduction

A fundamental question in systems biology is how a small number of signaling inputs specifies a large number of cell fates through the coordinated expression of thousands of genes. This problem is especially challenging given that gene regulatory and other types of networks in biology tend to be highly interconnected and their regulators promiscuous, with regulators affecting multiple targets and targets being affected by multiple regulators. Examples of this architecture include: transcription factor binding networks in bacteria [1], yeast [2, 3], plants [4], and animals [5, 6]; cellular signalling pathways involved in growth and differentiation [7–9]; the interactome of protein kinases and phosphatases [10, 11]; and synaptic connections between different layers of the brain [12]. Furthermore, because the targets and regulators are often well-mixed and mutually accessible in the cell, most actions are likely to have nonspecific and undesired effects.

At the same time, regulatory molecules drive networks to a large number of highly specific outcomes or cell fates. Although there are approximately four hundred canonical cell types in the adult human [13], recent single-cell RNA expression profiling experiments in the developing embryo [14, 15], brain [16], hematopoietic system [17, 18], and other organs [19, 20], have indicated that there may be thousands more.

Given there are only a few signaling pathways used in metazoan development [21, 22], understanding how cells reach their final outcomes when there are fewer regulators than fates and/or targets is an unsolved problem. One extensively studied solution for the control of promiscuous gene networks is combinatorial binding of DNA-binding transcription factors (TFs) at the promoter [23–31]. At the level of individual promoters, combinatorial binding ensures that individual genes are ON only when specific combinations of TFs are present (Fig 1A). However, on the genome level, combinatorial regulation restricts which sets of genes may be ON at the same time. For example, using AND logic, gene H in Fig 1A is only ON in the case that the three TFs K_1 , K_2 , and K_3 are concurrent at the H promoter; but these stringent requirements mean that H can never be transcribed independently of the less highly-regulated genes A–G. (A similar conclusion holds for OR logic.)

In fact, using combinatorial control, there is a one-to-one correspondence between configurations of the targets and configurations of the regulators. As shown in Fig 1A, the ON/OFF states of 3 TFs uniquely define the binding combinations at $2^3 = 8$ promoters. A similar conclusion holds when the regulators are expressed in a graded fashion.

This one-to-one correspondence is the fundamental limitation of combinatorial regulation: it requires an equal number of regulators and independently controlled targets and/or cell fates. Applied to embryonic development, combinatorial control requires that hundreds or thousands of cell-type specific TF combinations be generated in a spatially precise manner at the start. However, the combinatorial scheme does not explain how the TF states are regulated in the first place, and thus it offers no new insight into how cell fate is specified.

The limitations of combinatorial logic can also be understood from an information theoretic point of view. In particular, it is impossible to specify arbitrary cell fates if the regulatory layer bottlenecks the capacity of the targets to receive messages from extracellular signals. It is

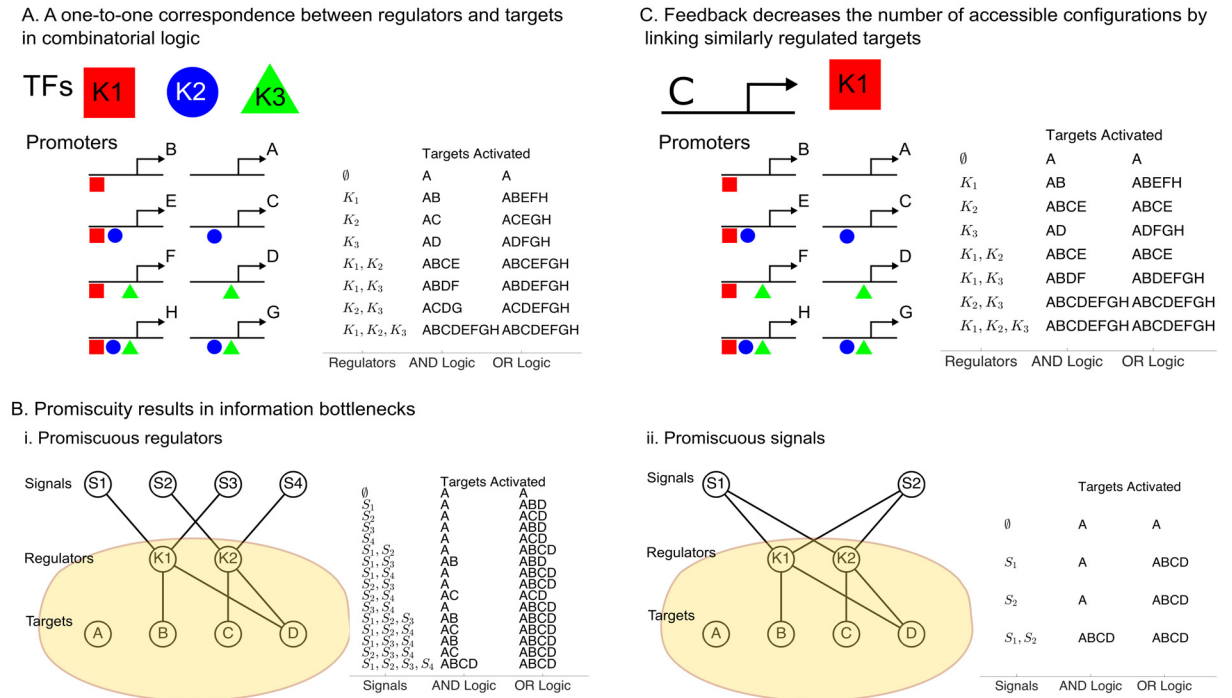


Fig 1. Combinatorial logic bottlenecks information flow in networks. (A) The number of ways that three TFs (K_1, K_2, K_3) can be ON or OFF (tabulated at right) is the same as the number of ways they can bind at promoters (left). An equal number of gene expression states are observed whether the TFs use AND logic (requiring all factors be present) or OR logic (requiring at least one of the factors). (B) Signal-to-target information flow is bottlenecked by regulators if (i) the regulators respond to multiple targets, or (ii) the signals activate multiple regulators. The allowed target states are tabulated for signals using AND logic and regulators using AND/OR logic. (C) A feedback loop causes constitutive activation of a regulator (K_1) and leads to fewer accessible configurations (tabulated at right).

doi:10.1371/journal.pcbi.1005089.g001

known that some ten to twenty types of signals [21, 22] converge onto membrane-bound regulators in many different combinations, permitting messages to be passed to the downstream targets. Much of this information stands to be lost, however, if the network relies on combinatorial logic alone: the regulatory layer simply cannot transmit messages in their entirety if there are more signals than regulators. Thus, combinatorial logic strongly circumscribes what fates are ultimately reachable. Cell fate information is lost not only if the signals are more numerous than the regulators, but also if the connections between signals and regulators are promiscuous (Fig 1B). When different signals activate the same regulators (Fig 1Bi), certain signaling inputs become redundant. On the other hand, when same signal activates different regulators (Fig 1Bii), some of the regulators become redundant. One may determine by direct enumeration exactly how redundancy decreases the number of configurations available to the targets (Materials and Methods Sections 1 and 2). These preliminary conclusions are at odds with the observation that signaling molecules are deployed over time in a complex code [32]. How do these messages in the signal space reach the targets if the regulatory layer imposes a bottleneck on information flow?

In addition, feedback regulation—a common feature of regulatory networks—exacerbates information bottlenecks when coupled with combinatorial logic. Stated another way, feedback merely widens the basin of attraction of certain promoter configurations at the expense of the number of distinct configurations. In Fig 1C, constitutive expression of K_1 by C means that C is never ON independently of the targets regulated by K_1 . Thus, the number of accessible configurations decreases from 8 to 6 without allowing new target configurations to be explored.

We need an alternative to combinatorial logic in cell fate specification that overcomes information bottlenecks. Here, we considered time-ordered control schemes, which we refer to as sequential logic. In this scheme, regulators can be applied in a stepwise manner; the entire sequence matters, so the final configurations can differ if the same regulators are permuted in time. In order for different temporal sequences to carry distinct information, the actions of the regulators must be *noncommutative*. This is the case, for example, when a regulator protects its targets from the action of another regulator, as when loci recruited to repressive chromatin compartments are protected from further modification [33, 34].

While it is not surprising that noncommutative sequences like this result in different outcomes at the single promoter level, these simple mechanisms may have nontrivial implications for regulation at the genome level. In particular, noncommutativity permits the same regulators to be used at different times with distinct effects. This is seen in development when ubiquitous signaling molecules like FGF family members exert different effects depending on the time and context of their expression [35–38]. Reuse of factors could greatly expand the information capacity of the major signaling pathways.

A number of examples show that noncommutativity may be a general strategy in other areas of biology. In hematopoietic stem cells, activation of GATA2 and C/EBP α in different orders results in different cell fates [39]. In neurobiology, different temporal orderings of the same inputs lead to distinct firing patterns [40–42]. In the field of synthetic biology, a DNA switch was developed that could detect the order in which invertase enzymes were applied [43]. And in evolutionary biology, the order in which mutations arise was recently implicated in determining a genotype's fitness [44–47]. There is also accumulating evidence for sequential logic in transcriptional control: signaling molecules and TFs in mammalian cells, including ERK [48], NF- κ B [49, 50], p53 [51], as well as in yeast [52–54] have been observed to pulse, suggesting that TF timing may be used to control the transcriptional state of the cell.

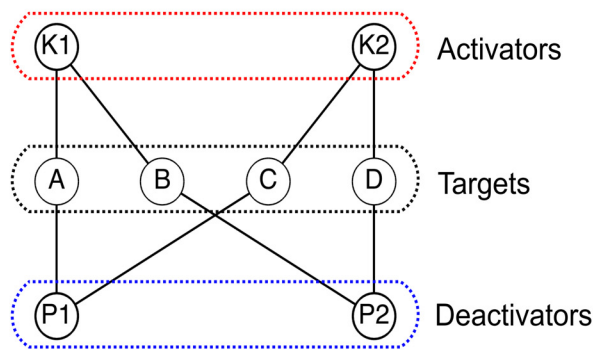
By applying sequential logic, we show that, even in complex and promiscuously regulated networks, specific target configurations can be reached using a temporal sequence of regulators. In particular, we consider two models inspired by (i) kinase/neural networks and (ii) chromosome folding and show analytically that both scale super-exponentially. We further show that noncommutative networks are robust to the loss of regulators, suggesting a mechanism for regulator evolution. We also show that regulators induce different orbits in expression space, which is related to the number of networks that can be controlled in parallel. We conclude by discussing how these models apply to interconnected networks in and outside biology and by providing possible experimental tests of the theoretical concepts. Theorems and proofs are given in the Materials and Methods.

Results

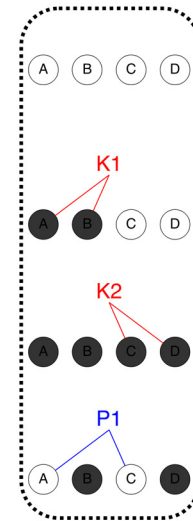
A time-sequence ratchet model generates more diversity than combinatorial logic in multiply-connected networks

To consider how time-ordered sequences of regulators can specifically control groups of targets, we begin by analyzing a generic two-layer network that is an extension of combinatorial logic (Fig 2, Materials and Methods Sections 1 and 2). In this model, each regulator controls multiple targets, and each target is accessible to any of its regulators. The model is meant to be analogous to the cellular environment wherein regulators and targets are well-mixed. For example, targets could be substrate proteins capable of multi-site phosphorylation [55, 56], and regulators the kinases and phosphatases. Targets could also be neurons and regulators their upstream excitatory and inhibitory inputs [12]. We denote by K the set of activators (i.e. kinases) and P the set of deactivators (i.e. phosphatases). Each target has a ladder of (integer-

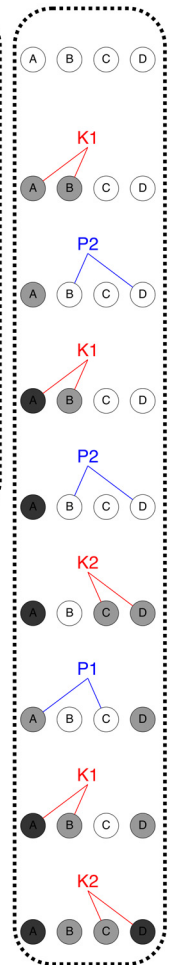
A. An example ratchet network



B. T = 1



C. T = 2



D. Scaling in the ratchet network

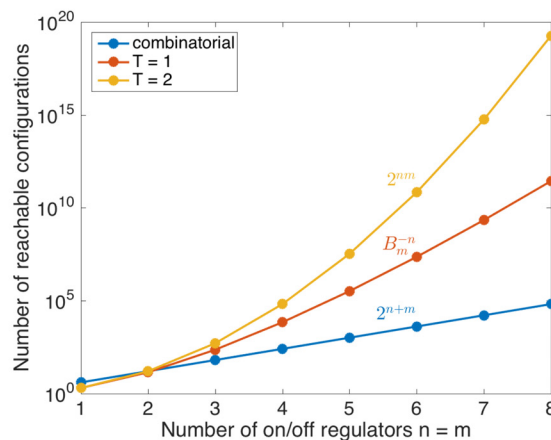


Fig 2. The ratchet model attains configurations not reachable by combinatorial logic. (A) The ratchet model for $n = m = 2$ and $l_n = l_m = 1$. Activators (K 's) and deactivators (P 's) turn targets ON and OFF, respectively. (B) An example temporal sequence for the network with a threshold equal to 1. Black targets are in the 1 state. (C) An example sequence for the same network with threshold equal to 2. Gray targets are in the 1 state, and black targets are in the 2 state. (D) Scaling laws for the threshold $T = 1$ (red) and $T = 2$ (yellow) are shown for symmetric networks ($n = m$). A comparison to combinatorial logic with an equivalent number of regulators ($n + m$) is shown in blue.

doi:10.1371/journal.pcbi.1005089.g002

valued) states, and together the states of the targets are a configuration of the network. (This distinction is in contrast to the common usage of “state” as a gene expression vector.) An additional parameter, the threshold T , determines the number of rungs on the ladder. Regulators ratchet the targets through their states, and only targets that have reached threshold will be ON at the end of a sequence of regulators. If each target in the group can be controlled by a unique combination of K 's and P 's, what ON/OFF configurations are possible?

In this model, termed the ratchet network (Fig 2A), each of n K 's and m P 's control

$N = \binom{n}{l_n} \binom{m}{l_m}$ unique targets, with the connectivity parameters l_n and l_m specifying the number of regulators to which each target connects. Consider the sequence $K_1 K_2 P_1$ acting on the targets $A, B, C,$ and D (Fig 2B). In the final configuration, B and D are ON together even though

no single K connects to both, and A and C are OFF, even though both share an activator with B and D . Therefore, this simple model illustrates the important point that similarly regulated targets can be independently controlled using sequential logic.

With threshold $T = 1$, not all configurations are reachable. Observe that there is no way to specifically activate A and D while leaving B and C OFF. This result is surprising given that A and D share no regulators: specificity depends on the network as a whole, not just individual targets. By going to $T = 2$, the forbidden configuration becomes accessible (Fig 2C), along with all ON/OFF states (below).

A combinatorial formulation of the model as a connectivity matrix

The model described above can be formalized as a combinatorial object that we refer to as the *connectivity matrix* \mathbf{A} . This formulation is useful because it is amenable to studying scaling, and it permits a direct comparison between noncommutative ratchet networks and standard combinatorial logic. For the interested reader, the models considered in this paper have a universal formulation as noncommutative matrix operators on the vector space of configurations (Materials and Methods Section 9).

Typically, the state of N targets is represented as an N -dimensional vector. If each target is controlled by a unique (K_i, P_j) pair (i.e. $l_n = l_m = 1$), the $N = nm$ -dimensional vector can be reformulated as an $n \times m$ matrix

$$\mathbf{A} = \begin{matrix} & P_1 & \cdots & P_m \\ \begin{matrix} K_1 \\ \vdots \\ K_n \end{matrix} & \begin{pmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{pmatrix} \end{matrix} \tag{1}$$

where each entry $\mathbf{A}_{i,j} \in \{0, 1, \dots, T\}$ is the state of the target regulated by K_i and P_j . For example, the connectivity matrix for the network in Fig 2 is

$$\mathbf{A} = \begin{matrix} & P_1 & P_2 \\ \begin{matrix} K_1 \\ K_2 \end{matrix} & \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{matrix} \tag{2}$$

In general, a regulator may connect to multiple targets (i.e. $l_n, l_m > 1$, see below), so that each entry of \mathbf{A} may be thought of as an M -dimensional vector (M determined in Materials and Methods Section 1). It turns out that this is an unnecessary complication; we instead let each $\mathbf{A}_{i,j} = 1$ if at least one of the M targets regulated by K_i and P_j is ON, and $\mathbf{A}_{i,j} = 0$ only if all M targets are OFF.

In this formulation K_i and P_j are raising and lowering operators that map $n \times m$ matrices to $n \times m$ matrices via the rules

$$\begin{aligned} K_i(\mathbf{A}_{i,j}) &= \begin{cases} \mathbf{A}_{i,j} + 1 & \text{if } \mathbf{A}_{i,j} < T \\ \mathbf{A}_{i,j} & \text{if } \mathbf{A}_{i,j} = T \end{cases} \\ P_i(\mathbf{A}_{i,j}) &= \begin{cases} \mathbf{A}_{i,j} - 1 & \text{if } \mathbf{A}_{i,j} > 0 \\ 0 & \text{if } \mathbf{A}_{i,j} = 0. \end{cases} \end{aligned} \tag{3}$$

From Eq (3), any sequence $K_{i_1} K_{i_2} \cdots K_{i_k}$ of all K 's is commutative, because any target controlled

by $t \leq k$ of the K 's will be in state $t \leq T$ at the end of the sequence, regardless of the order. A similar argument holds for the P 's. However, sequences consisting of both K 's and P 's are in general noncommutative. This is due to edge effects when $A_{i,j} = 0$ or T . If $A_{i,j} = T$, for example, then $K_i P_j$ results in $A_{i,j} = T - 1$, whereas $P_j K_i$ gives $A_{i,j} = T$. Therefore, \mathbf{A} gives insight into both the configuration of the targets and the noncommutativity of the regulators.

The problem of determining the number of accessible configurations in a network is reduced to finding the number of matrices satisfying certain patterns. For example, combinatorial logic with $T = 1$ corresponds to the special case in which the only sequences are the 2^n combinations of the n K 's. In an $n \times 1$ connectivity matrix, activating K_i corresponds to turning all 0's in row i into 1's. There are 2^n matrices generated by this procedure. More complicated cases of combinatorial logic can be studied this way (Materials and Methods Section 2), but it turns out that the total number of network configurations is always less than 2^{n+m} , with $n+m$ the total number of regulators. This is important because noncommutative models can bypass the exponential limit.

The ratchet model scales as the poly-Bernoulli numbers

We used the connectivity matrix representation of the ratchet network to determine the scaling as function of the number of regulators n and m , with each target connected to a unique (K, P) pair (i.e. $l_n = l_m = 1$) and the threshold $T = 1$. K_i turns 0's to 1's in row i and P_j turns 1's to 0's in column j . The rules are consistent with the one-pot reaction model in which all substrates receptive to K_i are promoted when K_i is active. For example, the sequence $K_1 K_2 P_1$ in Fig 2B can be recast as

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \xrightarrow{K_1} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \xrightarrow{K_2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \xrightarrow{P_1} \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}. \tag{4}$$

The main result is that \mathbf{A} must avoid the patterns $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ in any 2×2 sub-block (Materials and Methods Section 3). Brewbaker [57] enumerated the $n \times m$ binary matrices avoiding these patterns and showed that they scale as the poly-Bernoulli numbers [58]

$$B_m^{-n} = B_n^{-m} = \sum_{j=0}^m (-1)^{(n+j)} j!(j+1)^n \left\{ \begin{matrix} n \\ j \end{matrix} \right\} = \sum_{j=0}^{\min(n,m)} (j!)^2 \left\{ \begin{matrix} m+1 \\ j+1 \end{matrix} \right\} \left\{ \begin{matrix} n+1 \\ j+1 \end{matrix} \right\}, \tag{5}$$

where $\left\{ \begin{matrix} n \\ j \end{matrix} \right\}$ is a Stirling number of the second kind, defined combinatorially as the number of ways to put j labelled balls into n unlabelled boxes such that no box is empty [59]. These numbers scale not quite as fast as $2^N = 2^{nm}$, but much faster than 2^{n+m} , the maximum number of states in the equivalent combinatorial network (Fig 2D). Thus, a simple time-sequence model is able to generate super-exponential scaling.

All binary ON/OFF states are reachable for an increased threshold

Are more configurations accessible if multiple activation events are needed before reaching threshold? For example, neurons require the summation of multiple excitatory inputs to reach action potential, and proteins need to be phosphorylated at multiple sites before they are activated [55, 56]. We found that by increasing the threshold to $T = 2$, all 2^N ON/OFF

configurations of the N targets become reachable. In the connectivity matrix formulation, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ are no longer forbidden, which we show with an inductive proof

(Materials and Methods Section 4). This scaling law (Fig 2D), achieves the maximum of reachability and specificity; it far exceeds the scaling 2^{n+m} of the combinatorial model.

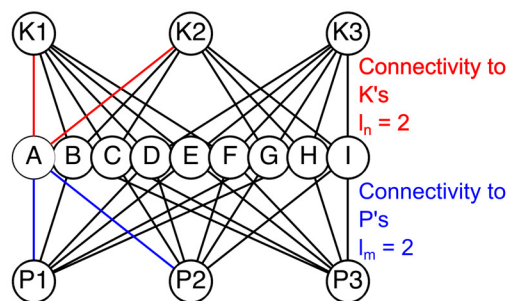
Being able to reach the entire ON/OFF space of N targets is overkill for most biological networks, which only display a relatively small number of stable configurations. The major implication of this result is that multiple levels of activity permit more targets to be controlled independently.

Increased regulatory connectivity generates robustness

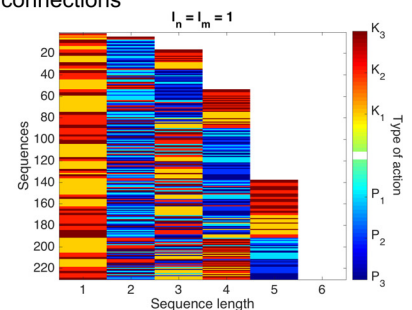
As sequential logic allows a large number of configurations to be reached in a complex network, we asked whether increasing the connectivity of the network (l_n and l_m) can maintain the specificity of the network while making it robust to the loss of a regulator. This is potentially relevant to evolution of biological networks, because redundant connections allow the network to repurpose regulators for new functions without severely impairing existing ones [60].

In the ratchet model, an increase in the connectivity parameters to $l_n = 2$ K 's and $l_m = 2$ P 's permits multiple targets to share a common (K, P) pair (Fig 3A). The connectivity matrix

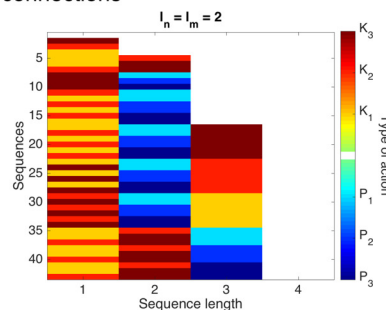
A. A network with extra connections



B. Sequences in a network with minimal connections



C. Sequences in a network with extra connections



D. Number of unique sequences

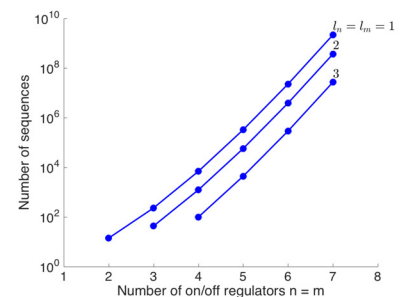


Fig 3. Multiple connections in the ratchet network decreases the number of configurations. (A) An example network where each target has $l_n = 2$ connections to the K 's (red) and $l_m = 2$ connections to the P 's (blue). (B) A list of the minimal length sequences generating unique configurations in the network in when $l_n = l_m = 1$. Red bars are K actions and blue bars are P actions. (C) The list of minimal length sequences when $l_n = l_m = 2$. Some sequences now map to the same configuration. (D) Analytical solution for the number of sequences as a function of $n = m$ for different $l_n = l_m$ families.

doi:10.1371/journal.pcbi.1005089.g003

incorporating the extra links in the network in Fig 3A is

$$\mathbf{A} = \begin{matrix} & P_1 & P_2 & P_3 \\ K_1 & \left(\begin{matrix} ABDE & ACDF & CDEF \\ ABGH & ACGI & BCHI \\ DEFG & DFGI & EFHI \end{matrix} \right) \\ K_2 & \\ K_3 & \end{matrix} \quad (6)$$

Now that each entry of \mathbf{A} is a group of $M > 1$ targets, it makes sense to track the state of the group as a whole with a single number $\mathbf{A}_{i,j}$. Even though a target appears in multiple entries of \mathbf{A} , the rules prevent a regulator from altering the state of groups at remote locations (e.g. K_1 cannot change the state of the group at $\mathbf{A}_{2,2}$).

We prove in the Materials and Methods that all sequences using at least $n - l_n + 1$ K 's and $m - l_m$ P 's are redundant with shorter sequences (Fig 3B and 3C, Materials and Methods Section 5). For example, the sequences $K_1 K_2 K_3$ is required to turn ON all targets in the case $l_n = l_m = 1$, but if $l_n = l_m = 2$, the shorter sequences $K_1 K_2$, $K_1 K_3$, and $K_2 K_3$ have the same effect. We derived a recursive formula that eliminates the redundant sequences in each (n, m, l_n, l_m) instance to derive the number of sequences in $(n, m, l_n + 1, l_m)$ and $(n, m, l_n, l_m + 1)$ (Fig 3D and S2 Fig). The formula agreed exactly with an algorithm designed to find all minimal length sequences (Materials and Methods 5). Notably, increasing l_n, l_m reduced the number of configurations. We observed a similar effect in combinatorial logic (S1 Fig).

To investigate the robustness of sequential logic networks, we studied the effect of deleting regulators in increasingly connected networks on the number of reachable configurations (Fig 4A). We hypothesized that sequences that activate similar subsets of targets should be able to recoup permanently lost configurations. To test this, we computed the normalized correlation coefficient between configurations in the network using all K 's (the full network) and configurations in the network without K_1 (the impaired network), subject to the constraint that those configurations were reached using longer sequences (Fig 4B). To focus on the recoverable fraction, we deleted all configurations that had an exact match. Highly similar configurations (yellow) clustered to the right of the plot, indicating that longer sequences can be used to recover lost configurations.

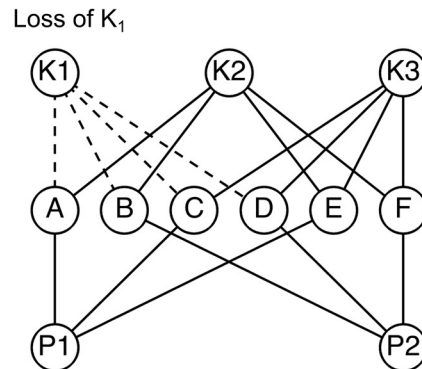
How similar are the recouped configurations? As connectivity increased, the maximum similarity became increasingly concentrated above 0.8 (Fig 4C). There is generally a tradeoff between reachability and the size of the fraction above 0.8 (Fig 4D). The tradeoff is nonlinear, however: using $l_n = 2$ gave the greatest increase recoverability for the smallest loss of configurations, showing that an intermediate level of redundancy can buffer the network to loss of regulators. The above analyses demonstrate that specificity of control is not compromised when regulators are lost or repurposed in heavily interconnected networks.

Sequestration networks generate diversity through protected states

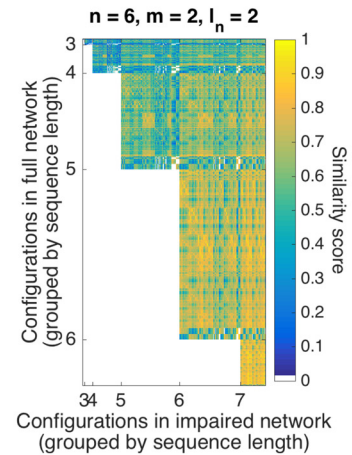
In the ratchet model, all targets are accessible to their regulators at all times. However, in some cases targets may be shielded from regulators: for example, genes can be silenced by sequestration in various nuclear compartments [61, 62]. This was seen in a landmark study by Filion *et al* [63], who used a DNase accessibility assay to show that genes associate with different regulators depending on their chromatin “color” or accessibility status.

To study the effect of accessibility and silencing on activating specific subsets of genes, we constructed the following sequestration model. In addition to the OFF state 0 and the ON state 1, each target/gene is endowed with additional orthogonal states 2 to n (allowing for a total of

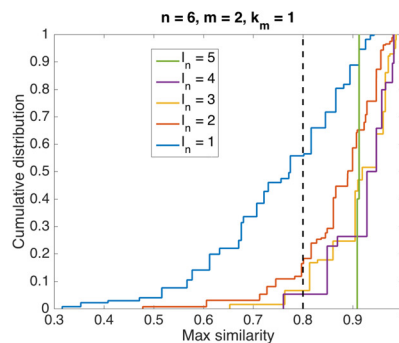
A. A network with redundant connectivity



B. Similarity of configurations in the full and impaired networks



C. Fraction of configurations that can be recovered



D. Trade-off between reachability and robustness

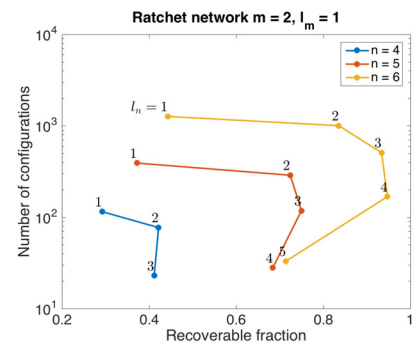
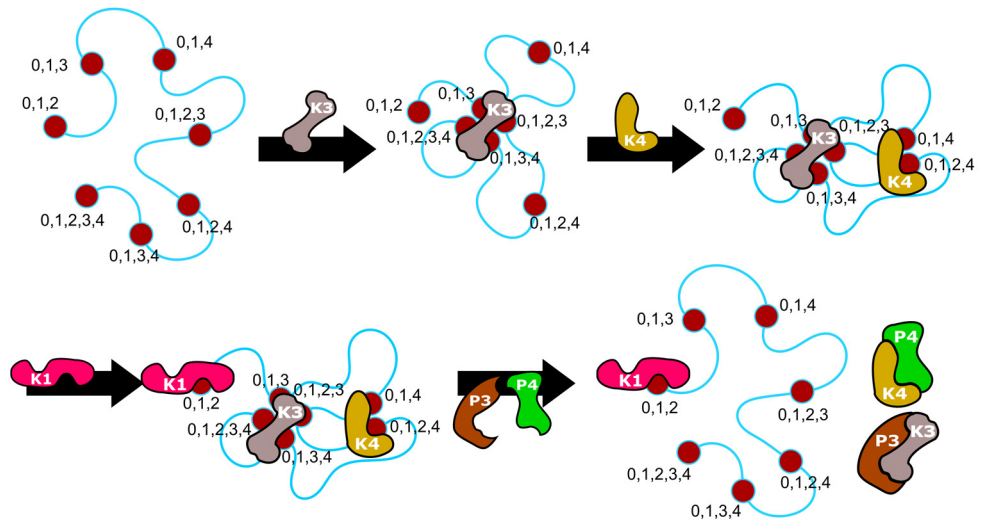


Fig 4. The ratchet network is robust to loss of a regulator. (A) A schematic illustration of the experiment. The regulator K_1 was deleted from networks with $m = 2^P$'s and variable n for different values of the connectivity l_n . The resulting number of configurations was computed by simulation. (B) Correlation coefficient between configurations in the full network (all K 's; rows) and the impaired network (without K_1 ; columns). All rows with exact matches were deleted. (C) Cumulative distribution $F(x)$ of the maximum correlation coefficient x for each row in C for different values of l_n . The dashed line is the similarity cutoff 0.8. (D) Tradeoff between reachability and robustness. The number of reachable configurations as a function of (n, l_n) is plotted vs. the fraction of states above the similarity cutoff 0.8 (i.e. $1 - F(0.8)$) for different values of n .

doi:10.1371/journal.pcbi.1005089.g004

$2^{n-1} - 1$ genes). If RNA polymerase (RNAP) is associated with K_1 , what genes can be independently activated? In this model (Fig 5) a regulator K_i promotes targets in the 0 state to state i , and P_i returns targets in state i to 0. Any target in state i is protected from regulators other than P_i . As an example of gene regulation on a three-dimensional chromosome (Fig 5A), the sequence $K_3 K_4 K_1 P_3 P_4$ first clusters all genes having a 3 in a repressive compartment, and then the remaining genes having a 4 in another repressive compartment. The net effect is that RNAP can only act on the gene represented by $\{1, 2\}$. We represent this abstractly as a configuration vector of k -armed targets (Fig 5B), where each entry corresponds to the state $\{0, 1, \dots, n\}$ of a gene able to access $k \leq n$ of the states (see below for a mathematical description of the model). Therefore, protected states in the sequestration model allow genes to be transcribed specifically in a well-mixed environment.

A. Pictorial view of noncommutative sequestration



B. The same sequence using multi-arm targets

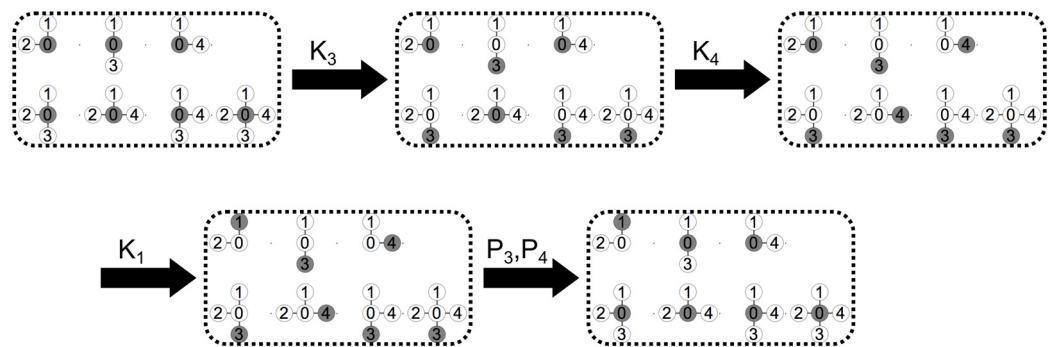


Fig 5. The sequestration network is a noncommutative model of gene regulation by chromosome folding. (A) A sequence of moves $K_3 K_4 K_1 P_3 P_4$ on a hypothetical chromosome with K and P actions represented as DNA-binding factors and K_1 playing the role of RNAP. Red circles correspond to genes and numbers correspond to allowed binding partners. (B) The same sequence in A represented as a collection of targets with up to $n = 4$ arms. For example, the target $\{0, 1, 2\}$ corresponds to the gene locus with states 1 and 2 in A. The filled circle represents the current state.

doi:10.1371/journal.pcbi.1005089.g005

We derived (see below) that the number of reachable configurations scales with the number of regulator pairs n as

$$f(n) = 2^{2^{n-1}-1} - \sum_{m=2}^{n-1} \binom{n-1}{m} \left(2^{\sum_{k=3}^m \binom{m-1}{k-1}} - 1 \right) \left(2^{\sum_{k=3}^m \binom{n-1}{k-1} - \binom{m-1}{k-1}} \right). \quad (7)$$

For $n = 1, 2, 3, 4, 5, 6$, this formula gives $f(n) = 1, 2, 7, 89, 16897, 780304385$ (Fig 6). We also relaxed the constraint that all genes have a 1 state (allowing for a total of $2^n - 1$ genes) and found that the number of configurations scales as $c_n = 2, 7, 94, 37701$ with $n = 1, 2, 3, 4$. The full model does not have an analytical solution, but it does have upper and lower bounds related to Eq (7) (Materials and Methods Section 7, S3 Fig).

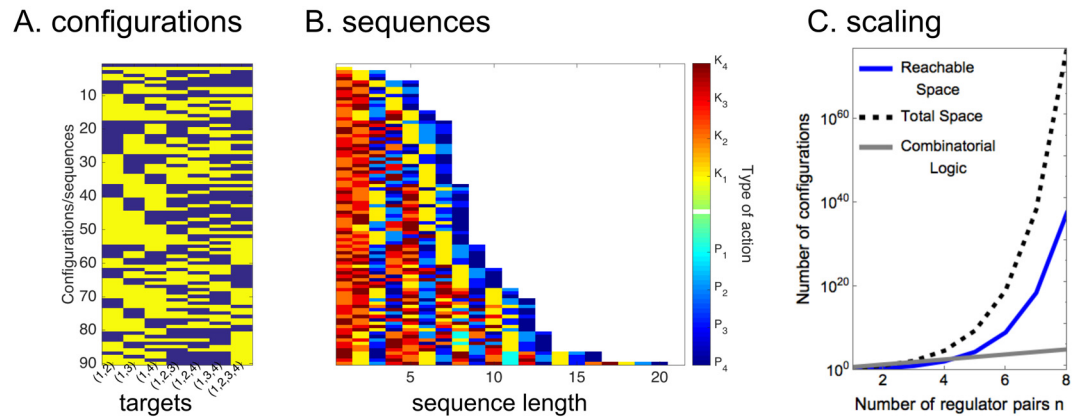


Fig 6. Scaling in the sequestration model is super-exponential. (A) A plot of all the allowed configurations of a set of targets of $n = 4$ regulator pairs in the sequestration model. Yellow represents targets that are ON, and blue those that are OFF. (B) A list of the sequences generating the corresponding states in A. K actions are shown in the red spectrum, and P in the blue. (C) A logarithmic plot of the scaling in the sequestration model. The total space is the $2^{2^n - 1} - 1$, the reachable space is calculated from Eq (7), and the combinatorial model is 2^{2^n} .

doi:10.1371/journal.pcbi.1005089.g006

Combinatorial scaling laws of this sort are not uncommon [44, 64, 65]. Edwards and Glass [64] saw an explosion in the number of states when studying trajectories on n -cubes, and Green and Rees [65] saw a super-exponential jump when enumerating certain types of nonrepeating sequences on n letters. Furthermore, a similar small number (four) of factors are necessary and sufficient to reprogram fibroblasts to stem cells [66]. Together, these results indicate that sequences can far exceed the 2^n limit set by combinatorial regulation, and that only a few regulators are necessary to make large changes in the configuration of a cell.

Regulators act on the configuration vector in the sequestration model

The sequestration network with n regulator pairs (referred to as the n -network) is described using the $1 \times 2^n - 1$ configuration vector \mathbf{x} . This is a simpler description than the connectivity matrix because a target affected by K_i is necessarily affected by P_i . The entries of \mathbf{x} are the states of each target g able to be controlled by $k \leq n$ of the regulator pairs. Each target g is a list $\{0, i_1, \dots, i_k\}$ of the k regulators to which it responds. Because of their radial appearance, such targets are said to have k arms (see Fig 5B).

The regulators act on \mathbf{x} according to the rules

$$\begin{aligned}
 K_i(\mathbf{x}_g) &= \begin{cases} \mathbf{x}_g + i & \text{if } i \in g \text{ and } \mathbf{x}_g = 0 \\ \mathbf{x}_g & \text{else} \end{cases} \\
 P_i(\mathbf{x}_g) &= \begin{cases} 0 & \text{if } i \in g \text{ and } \mathbf{x}_g = i \\ \mathbf{x}_g & \text{else.} \end{cases}
 \end{aligned}
 \tag{8}$$

Eq (8) guarantees that the regulators are *orthogonal* in the sense that a target in state j is protected from K_i and P_i if $i \neq j$; and also *idempotent* in that $K_i^2 = K_i$. Furthermore, sequences of regulators are noncommutative unless the only actions are P 's. This is a consequence of the fact that P 's put all affected targets into the 0 state. Although these rules are different from the ratchet model, a formulation exists that generalizes the K 's and P 's to matrix operators consistent with both models (Materials and Methods Section 9).

If \mathbf{x} is restricted to the $2^{n-1} - 1$ targets all able to be regulated by K_1 and at least one other K , the network is said to be *reduced*; otherwise we say \mathbf{x} is *full*. This distinction was used in Fig 5.

A *one-coloring* is a configuration of \mathbf{x} that uses only one of the states and 0. For example, the configuration $\mathbf{x} = (1, 0, 0, 1, 1, 0, 0)$ in the full $n = 3$ -network is a one-coloring of 1; so is the reduced network formed by $(\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_7) = (1, 1, 0)$. This concept is easily extended to $k > 1$ -colorings. One-colorings are particularly important because they resemble the ON/OFF configurations of genes in an RNA-seq experiment, and we would like to know how many such configurations can be reached.

A simple counting argument for the connected one-colorings illustrates super-exponential scaling in the sequestration model

As in the ratchet model, finding the accessible states of the sequestration network amounts finding restricted patterns in \mathbf{x} . We determined that the restricted one-colorings are those that violate a property referred to as *connectivity* (Materials and Methods Section 7). A configuration of \mathbf{x} is said to be connected if all $k > 3$ -arm targets $g_i^{(k)} = \{0, i_1, \dots, i_k\}$ match the state of at least one of k of the 2-arm targets $\{0, i_1, i_2\}, \dots, \{0, i_{k-1}, i_k\}$ sharing the indices i . If the network is reduced, no k -arm target may be in the 1 state when all of 2-arm targets with which it overlaps (i.e. shares an index other than 1) are in the 0 state. This restricts the one-colorings and suggests a method to determine the scaling law for the model in Fig 5.

As an example, in the $n = 4$ network on the reduced set of $2^3 - 1$ targets illustrated in Fig 5, $\{0, 1, 3\}$ and $\{0, 1, 4\}$ both being 0 constrains $\{0, 1, 3, 4\}$ to be 0 as well. Furthermore, even though $\{0, 1, 2\}$ is in the 1 state, $\{0, 1, 2, 4\}$ and $\{0, 1, 2, 3, 4\}$ may be 0. It is only the two-arm targets that constrain the possible configurations: for example, the longer sequence $K_2 K_4 P_2 K_3 K_2 P_4 K_1 P_3 K_4 P_1 K_3 P_4 K_1 P_2 P_3$ obtains the state $\mathbf{x} = (0, 0, 1, 0, 0, 0, 1)$ in which only the targets $\{1, 4\}$ and $\{0, 1, 2, 3, 4\}$ are ON, showing that $\{0, 1, 2, 3, 4\}$ need not be in the same state as $\{0, 1, 2, 3\}$, $\{0, 1, 2, 4\}$, or $\{0, 1, 3, 4\}$. In Fig 6A and 6B we show the allowed states and the sequences that generate them for $n = 4$; there are 90 out of a possible $2^{2^4 - 1 - 1} = 128$ configurations.

There are $2^{2^{n-1} - 1}$ one-colorings on $2^{n-1} - 1$ targets. How many of these violate the connectivity rule? Suppose there are m 0's among the 2-arm targets. If $m = 1$, then $\binom{m}{k-1} = \binom{1}{k-1} = 0$ of the $k \geq 3$ -arm targets are constrained to be 0, as there is always another 2-arm target (in the 1 state) that each k -arm target can match. If $m > 1$ and $m - 1 < k$, however, then $\binom{m}{k-1} > 0$, so $\binom{m}{k-1}$ k -arm targets whose states $\{i_1, \dots, i_{k-1}\}$ are completely contained within the set of 2-arm targets $\{0, 1, j_1\}, \dots, \{0, 1, j_m\}$ must be 0. Hence in any violation of the connectivity rules at least one of $\sum_{k=3}^m \binom{m}{k-1}$ k -arm targets will be in the 1 state and the remaining $\sum_{k=3}^m \binom{n-1}{k-1} - \binom{m}{k-1}$ k -arm targets will be 0 or 1. Furthermore, there are $\binom{n-1}{m}$ ways of specifying m 0's, so the total number of violations is

$$\sum_{m=2}^{n-1} \binom{n-1}{m} \left(2^{\sum_{k=3}^m \binom{m}{k-1}} - 1 \right) \left(2^{\sum_{k=3}^m \binom{n-1}{k-1} - \binom{m}{k-1}} \right). \tag{9}$$

Subtraction from $2^{2^{n-1} - 1} - 1$ gives Eq (7).

The ratchet and sequestration networks divide the configuration space into orbits

Until now we have considered the reachable space of a single group of targets each starting in 0. An ensemble of networks could each start with their targets in some arbitrary state, and when a sequence is applied to the ensemble the different networks will in general span different

configurations. Determining the number of *orbits* (defined precisely in [Materials and Methods](#) Section 8) within the set of possible configurations tells us how many networks can be controlled in parallel.

Enumerating the reachable space for both the ratchet and sequestration networks involved finding configurations that violated at least one rule. If two configurations have distinct violations, then there is no way they can communicate using the regulators. Therefore, the different orbits are the groups of configurations having the same forbidden patterns. It is possible that a violation could be alleviated by an action that changes the state of an offending target, so we require that each orbit be immune to a subset of the regulators. This could be achieved in biological networks by locking targets in protective chromatin states or by shutting down certain cellular receptors.

We determined a recursive formula for the number of orbits in the ratchet network for an arbitrary n, m ([Materials and Methods](#) Section 8). In [Fig 7A](#) we plot the orbits for the $n = 4, m = 2$ case. There is one large component of size B_n^{-m} and several smaller orbits of size B_i^{-j} with $i \leq n$ and $j \leq m$. There are only a handful of singleton orbits in [Fig 7A](#), but the number of isolated states dominates the space as n, m increase.

We were unable to find a similar solution for the sequestration network because we lack a general solution for the number of states in the main orbit. However, [Fig 7B](#) shows the computationally discovered orbits for the full network on $2^n - 1$ targets. A nontrivial feature is that there are orbits which use all pairs of regulators, but which do not communicate with the main orbit. For example, the sequence $K_2 K_3$ from $\mathbf{x} = (1, 0, 0, 0, 0, 0, 1)$ reaches the same configuration as the sequence K_1 starting from $\mathbf{x} = (0, 2, 3, 2, 2, 3, 0)$; these configurations are part of the same orbit because both violate the connectivity rule between $\mathbf{x}_7 = \{0, 1, 2, 3\}$ and the 2-arm targets $\mathbf{x}_4, \mathbf{x}_5,$ and \mathbf{x}_6 .

Another observation is that some pathways cannot be reversed by a legal action in the ratchet network orbits (indicated by a directed arrow in [Fig 7](#)), whereas there always exists a

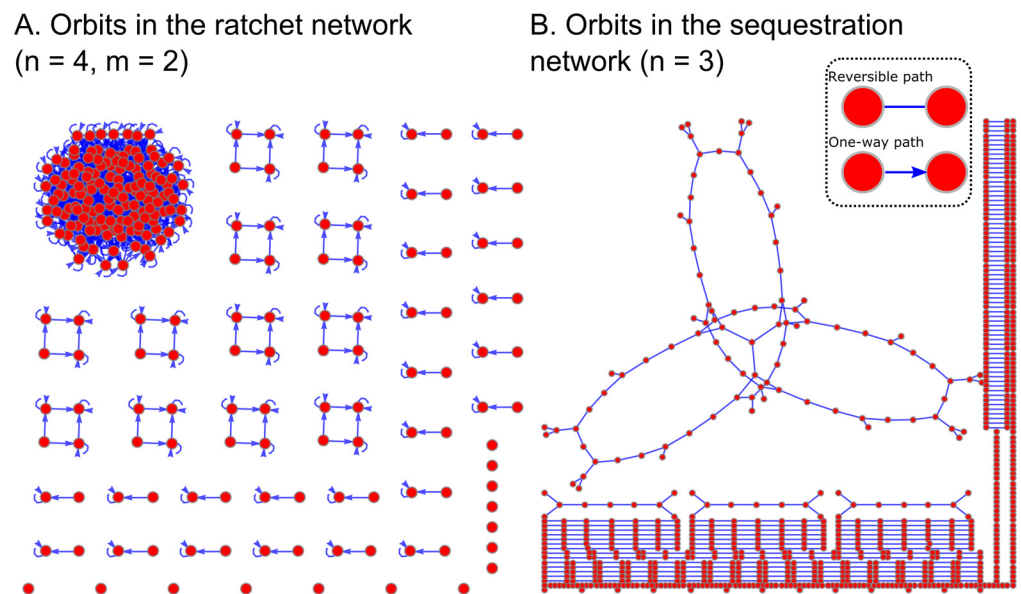


Fig 7. Noncommutative models induce orbits in the configuration space. Graphical representation of the orbits in (A) the $n = 4, m = 2$ ratchet network and (B) the full $n = 3$ sequestration network. Configurations are indicated by red circles, and those accessible to each other are connected with blue lines. Arrows in (A) indicate whether a path is irreversible.

doi:10.1371/journal.pcbi.1005089.g007

reversible path between configurations in the sequestration network orbits (no arrowheads). It can be proved that this is true in general for the sequestration network ([Materials and Methods Section 8](#)). This feature permits orbits to be found computationally by looking for reversible one-step paths in the entire configuration space.

The orbits are one explanation for the phenomenon the same signal can cause cells to behave differently [38]. More generally, the orbits demonstrate an intriguing symmetry between the targets responding to a restricted subset of the regulators on one hand, and the orbits restricted to the same subset on the other.

Discussion

In this paper we first show how noncommutative, sequential logic can relieve information bottlenecks in multilayer networks. Bottlenecks in combinatorial logic may occur whenever a downstream layer has fewer elements than the layer upstream, which poses the problem of how networks process complex signals without loss of information. Noncommutative solutions such as the ratchet and sequestration models, in which the number of configurations scales super-exponentially in the number of regulators (Eqs (5) and (7)), permit longer, more complex messages to reach the targets via information “pulses.” These pulses encode a large diversity of signals into configurations of the targets that would otherwise be lost using combinatorial logic.

Noncommutativity has long been recognized as a central concept in control theory, because it allows systems with few controllers to explore a broader configuration space. For example, one generates z rotations in 3D by $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$, so control over z is generated by a pulse sequence of rotations in x and y , as in airplane control where roll and pitch generate yaw [67]. Infinitesimal motions in the form of generating matrices are translated into flows in a vector space by exponentiation. Because matrix multiplication is noncommutative, composition of flows is not simply the addition of generators, but rather a higher order polynomial of commutators of the generators given by the Baker-Campbell-Hausdorff formula [68]. Noncommutativity also appears in experimental physical chemistry where pulse sequences can prepare spin systems in nontrivial population configurations [69]. A formal description of these phenomena is based on the Heisenberg picture of quantum mechanics, wherein evolution of a system of many variables is given by a differential equation involving the commutator of a Hamiltonian operator.

The significance of noncommutative control for systems biology is that it becomes possible to independently control targets that would otherwise be activated by the same promiscuous regulator. In this paper, we argue that noncommutative sequences permit control over new directions in gene expression space, allowing more specific sets of targets to be controlled. Several studies have shown that TFs that can bind genes in one tissue type are in fact precluded from binding the same genes in another [70, 71]. The *C. elegans* TF LIN35 fails to bind targets in the germline that it binds in the intestine [71], and the SMARCA4 complex in mouse binds enhancer elements in heart, limb, and brain tissue in a tissue-specific manner [70]. One hypothetical explanation for these observations, based on the sequestration model, is that cell-type specific gene expression is the result of noncommutative sequences like $K_1 K_2$ and $K_2 K_1$ that silence certain promoters. The three-dimensional structure of the genome is a likely setting for this type of regulation.

Gene regulation is known to take place in three-dimensions, as observations of DNA looping [72], nonrandom chromosome packing [73], and clustered transcription factories [74] have shown. However, the factors that affect chromosome structure are non-specific. One such factor is the ubiquitous zinc finger protein CCCTC binding factor (CTCF) [75], which

functions as both an activator of transcription by bringing enhancers and promoters together [76, 77] and as a repressor by insulating genes [78, 79]. Epigenetic modifications, such as histone methylation and acetylation [80–82], also affect three-dimensional structure. In addition, DNA looping was observed in the context of allelic exclusion during B- and T-cell lineage specification where individual alleles were recruited to heterochromatic regions while the other underwent recombination [33, 34]. Consequently, the sequestration model predicts that temporal permutations of a small set of chromatin modifying factors could specify a large number of potential chromosomal conformations and lead to different expression states and corresponding cell fate decisions.

New technologies such RNA-seq and ChIP-seq can be used to test the predictions of the noncommutativity hypothesis at the genome level. Epigenetic drugs such as azacytidine and trichostatin A inhibit DNA methylation [83] and histone deacetylation [84], respectively, and have been shown to cause global changes in gene expression alone and in combination [83, 85]. The sequestration hypothesis predicts that perturbations to the three-dimensional structure of the chromosome are noncommutative, so distinct gene expression states may be reached by permuting the order in which epigenetic drugs are applied. While the sequestration model may underlie chromosome folding, the ratchet model could form the basis of phosphorylation networks. For example, mass spectrometry studies have revealed complex phosphorylation patterns [86, 87], though the number of kinases and phosphatases is comparatively small and the networks are highly interconnected [10, 11]. As phosphoproteins are the mediator of extracellular signals, ordered disruption of signaling pathways could also lead to distinct gene expression configurations.

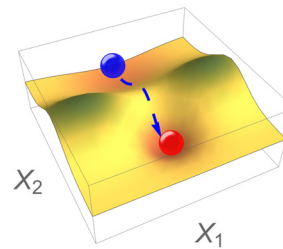
Analogously, the ratchet model may aid in the specification of distinct neural activity patterns, owing to the fact that connections between the different hippocampal layers overlap [12, 88]. While superficial neurons can be activated in response to spatial cues, deeper layers can be selectively activated by time sequences of inputs [40, 41, 89]. These results suggest the hypothesis that neural networks may be noncommutative. In particular, experimental support exists for the role of the dentate gyrus in pattern separation and orthogonalization by way of ensuring that even quite similar memory representations use distinct subsets of neurons [90, 91]. The ratchet model, by ordering inputs in time, is one way of reaching these specific subsets if the number of input neurons is smaller than the number of target neurons. Memories share many common elements, including shape, color, smell, and sound, which poses problems for recall. We hypothesize that older, “fuzzier” memories could be those relegated to very long ratchet sequences. According to this hypothesis, memories are not forgotten, but are instead increasingly difficult to access, and memories that are not consolidated are those that never formed a unique ratchet sequence.

Beyond resolving bottlenecks and generating specificity, noncommutative actions offer a new interpretation of how cell fate decisions and other stepwise processes occur on abstract regulatory landscapes. The classical Waddington landscape view of development holds that cells decay to attractor configurations representing terminal outcomes [92]; this is consistent with a boolean network with many variables X converging to a fixed point [93]. In a static landscape, the final outcome is determined *a priori* by the nearest energy minimum. What then determines the initial configuration? In organisms such as *Drosophila*, maternal patterning of the embryo may account for this initial bias [94]; but in other organisms that employ mechanisms like multilineage priming [82, 95], it is not clear that every cell fate decision is made at the beginning.

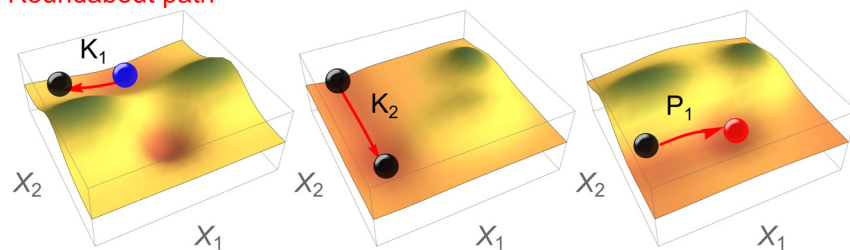
Sequential logic allows cells to reach their final fate on a dynamic landscape. In the system of Fig 8A (top), for example, it is not possible for cells in the blue configuration to transition to the red fate by increasing X_2 , because this involves an uphill climb. However, the regulators of

A. The regulatory landscape can be dynamic in sequential logic

Direct path



Roundabout path



B. Making a small change in one dimension may require a large change in a hidden dimension

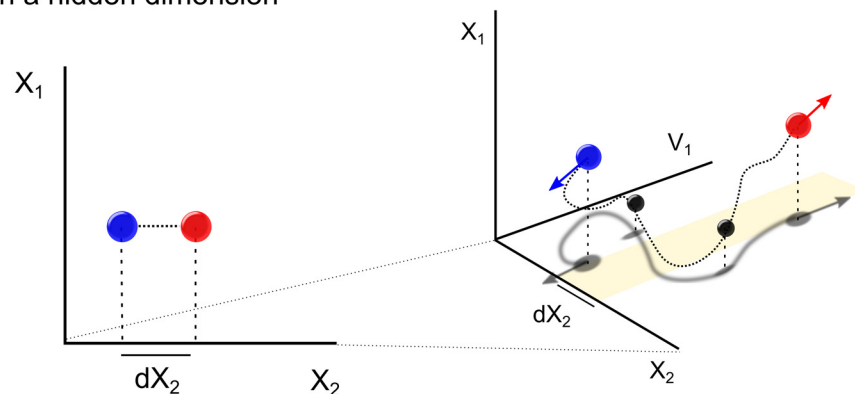


Fig 8. Sequential logic on regulatory landscapes. (A) The regulatory landscape for the 2-mRNA system X_1, X_2 for two hypothetical paths with configurations represented by balls. It is difficult to directly increase X_2 because of a potential barrier (top). In the roundabout path (bottom), visiting two intermediate configurations via $K_1 K_2 P_1$ results in an altered regulatory landscape. (B) The initial and final configurations in (A) projected onto (X_1, X_2) space (left) and (X_1, X_2, V_1) space (right). The regulators affect not only X_1 and X_2 , but also an additional variable, denoted V_1 , that alters the landscape of X_1 and X_2 . The arrows indicate the instantaneous direction of the trajectory.

doi:10.1371/journal.pcbi.1005089.g008

genetic networks may also affect the landscape directly. This is seen in Fig 8A (bottom) where the sequence $K_1 K_2 P_1$ changes the landscape in such a way that the overall cost of reaching the same endpoint is much lower than the direct path (Fig 8A, top). This can be understood as the effect of regulators acting on additional variables V , which modulates the landscape in X space. For example, TFs can recruit chromatin regulators that modify global three-dimensional chromosome structure and future TF accessibility [74, 76, 96, 97], or kinases can sequester substrates in the nucleus to prevent their subsequent activation [53, 54]. Because sequential logic acts on the V 's as well as the X 's, changes that appear to be small in one dimension (Fig 8B, left) actually involve large excursions in the full space (Fig 8B, right). As a consequence, in

noncommutative regulation, the landscape changes and cells can take on fates that were not accessible at the beginning.

Previous theoretical models have explored dynamic regulatory landscapes in the form of bifurcations [98, 99]. In these models, a set of kinetic parameters determines the positions of minima and maxima in the landscape. However, the noncommutative model advanced here is fundamentally different, in that using the regulators to move through X changes the landscape directly. This could happen, for example, if acting on X_1 with K_1 hides it from the effect of K_2 . Uncoupling of targets in this way may underlie the distinct effects of signals like FGF at different stages of development [35–38]. It will be interesting to explore time series data for hints that some genes pulse ON and OFF in order to protect their promoters from the actions of promiscuous regulators.

Multistep processes other than development can benefit from the type of noncommutative regulation highlighted in Fig 8. What seems like an intractable problem at the start becomes much more feasible if one realizes that the effects of actions change with time and context. This intuition is why thinking in terms of commutators $[A, B] = AB - BA$ can make complex problems more soluble: the desired effect is often what is leftover after performing and undoing a sequence of actions. Several examples illustrate this concept.

With its increased capacity for generating diversity, sequential logic is likely to be used in evolution. A recent theoretical example in social bacteria demonstrated that in evolving a new quorum sensing receptor-ligand pair, adding new receptors prior to ligands is preferred over the opposite path [45]. An analysis of the stability and catalytic activity of a family of bacterial β -lactamase mutants showed that the ability to evolve new substrate specificity is contingent on mutations that first stabilize the protein active site [46, 100]. Finally, biological networks evolve the same functions in different orders, but the order in which these functions arise dictates which other genotypes can be reached by neutral mutations [44]. These results suggest that permuted sequences of mutation events may have different fitness costs. With extensive artificial evolution experiments underway in protein engineering [100] and bacterial mutation accumulation [47], coupled with progress in sequencing technologies, it will be possible to test this hypothesis by permuting the conditions that promote mutation.

Sequential logic can also be applied in synthetic biology to build circuits with memory [43, 101–103]. In general, the toolkit that permits up- and downregulation of genes is small, with a few staples like Lac, Tet, and Ara [104]. Significant effort has been put into generating logic gate (AND/OR) promoters [30]. To further expand the toolkit, it has been proposed that more orthogonal regulators be developed [105]. We suggest that sequential logic may be a more promising strategy to scale up the number of targets that can be independently controlled by permuting in time a small number of controllers.

More broadly, sequential logic can be used to accomplish experimental goals not possible in single-step approaches. For example, in multiplexing mRNA detection in single cells, we previously used a sequential hybridization scheme that permits the number of barcodes to exponentially [106], whereas combinatorial schemes can only specify approximately 30 barcodes. We expect many single-cell experiments to benefit from a sequential strategy in which detours facilitate achievement of the main goal with high efficiency.

Finally, our results connect outside of biology to strategic planning in social, political, and economic arenas. Anyone familiar with negotiating knows about the limitations inherent in trying to make interconnected groups of people move in specific directions, especially when the actions affect all participants at once. Multiparty negotiations and tournaments may benefit from time-ordered strategies in which enemies temporarily team up, or fringe interest groups are transiently pacified. Indeed, a conclusion from the sequestration model is

that the most highly regulated targets need to be protected prior to satisfying the ones with fewer connections. Determining whether this prediction is borne out in congressional and international negotiations, for example, is an interesting question for political science. Evidence for noncommutative effects in games exists in that the initial seeding in a tournament can bias its outcome [107], and that long-term goals change players' strategies in the repeated prisoner's dilemma [108]. In conclusion, the direct path to an outcome in a networks with many interacting parts may have many unintended and prohibitively expensive consequences. A multi-step strategy may achieve the same outcome with minimal cost and side effects.

Materials and Methods

1. The connectivity matrix with multiple targets

In this section we determine how many targets are controlled by the same regulators in the connectivity matrix \mathbf{A} . Then we extend \mathbf{A} to more than 2 dimensions.

If $l_n = l_m = 1$ it is clear that each $\mathbf{A}_{i,j}$ corresponds to a single target and that each target appears only once. In general, however, a target can appear in multiple entries of \mathbf{A} (cf. Eq (6)). To see this, consider the bipartite graph formed by all the targets and all the K 's, but none of the P 's. The handshaking lemma from graph theory [59] says that the total number of edges is one half the sum of the degrees of each vertex, which is either l_n for a target or some number p_n for a K regulator. There are Nl_n total edges, so we find $\frac{1}{2}(Nl_n + np_n) = Nl_n$ or $p_n = \frac{2}{n}l_n$ for the number of links coming from each K . Similarly, the number of links emanating from each P is $p_m = \frac{2}{m}l_m$. In terms of the connectivity matrix, p_n and p_m correspond to the number of unique targets in each row and column, respectively.

Because K_1 connects to a fraction $\frac{p_n}{N}$ of the targets, it follows that K_1 and P_1 together connect to a fraction $\frac{p_n p_m}{N^2}$ of the targets. Therefore, the total number of targets connecting to K_1 and P_1 is $M = N \left(\frac{p_n p_m}{N^2} \right) = \frac{p_n p_m}{N}$. Another way to see this is to consider one target in the intersection of K_1 and P_1 . This one target uses up one of each of the regulators and one unit of connectivity, leaving a total of $M = \binom{n-1}{l_n-1} \binom{m-1}{l_m-1}$ ways to connect other targets to the same pair of regulators. It is easily verified that these two formulations for the number of targets per matrix entry M are equivalent. This illustrates that there is not simply a one-to-one correspondence between the entries of \mathbf{A} and the targets.

There was nothing special about the labels K and P in the above paragraphs. Thus, the connectivity matrix can easily be extended to a u -dimensional *connectivity tensor* where u is the number of pools of regulators. Each pool has n_i regulators connecting to l_{ni} targets, and each target connects to $p_{ni} = \frac{2}{n_i}l_{ni}$ regulators of pool $i, \forall i \in \{1, \dots, u\}$. The total number of targets and the total number of targets per entry are extensions of the $u = 2$ case, giving

$$N = \prod_{i=1}^u \binom{n_i}{l_{ni}} \tag{10}$$

distinct targets and

$$M = \frac{\prod_{i=1}^u p_{ni}}{N^{u-1}} = \prod_{i=1}^u \binom{n_i - 1}{l_{ni} - 1} \tag{11}$$

targets controlled by one factor from each of the u pools. [S1A Fig](#) shows an example network with $u = 3$ pools.

2. Counting configurations in combinatorial networks using the connectivity matrix

The number of configurations in combinatorial logic is the number of ways that N targets can each be bound by exactly u regulators, where each regulator comes from a different pool. In the main text we analyzed the case $u = 1$ and $l_n = 1$. Here we extend those results to arbitrary u and l_n .

First consider the case $u = 2$, corresponding to a pool of K 's and a pool of P 's. Whereas in the ratchet model, K_i and P_j acted separately on the entries of \mathbf{A} , in combinatorial logic the pair (K_i, P_j) is needed to switch $\mathbf{A}_{i,j}$ from 0 to 1. Many such pairs may be active at any one time. We write this formally as

$$(\{K\}, \{P\})[\mathbf{A}_{i,j}] = \begin{cases} 1 & \text{if } K_i \in \{K\} \text{ and } P_j \in \{P\} \\ 0 & \text{else,} \end{cases} \quad (12)$$

where $\{K\}$ denotes a subset of the K 's. The notation (\cdot, \cdot) means that a combination of factors acts on the target, instead of just a single factor.

If $l_n = l_m = 1$ there are $(2^n - 1)(2^m - 1) + 1$ ways to pick at least one of n K 's and one of m P 's, plus one way to pick nothing. If $l_m = 1$ and $l_n > 1$, then for a certain number $\alpha \leq n$ of the K 's, any subset containing α or more K 's has the same effect as activating all n K 's at once. For example, in Eq (6), the action of $(\{K_1, K_2\}, \{P_1, P_2\})$ is sufficient to activate all targets in the $n = m = 3, l_n = l_m = 2$ network. To determine α , recall that there are M targets in each entry of the connectivity matrix \mathbf{A} . Choosing i K 's means that the total number of targets is $M \times i$, but a single column of \mathbf{A} only contains p_m unique targets. Each target is connected to l_n K 's, so for a target in the intersection of i K 's and a single P , there are $l_n - i$ spots left over to choose $n - i$ K 's and $l_m - 1$ spots left over to choose $m - 1$ P 's, or $\binom{n-i}{l_n-i} \binom{m-1}{l_m-1}$ ways total. Using the principle of inclusion-exclusion [59] this means that α is the smallest i such that

$$M \times i - \sum_{i'=2}^{\min(i, l_n)} (-1)^{i'} \binom{i}{i'} \binom{n-i'}{l_n-i'} \binom{m-1}{l_m-1} \geq p_m. \quad (13)$$

By choosing α K 's, the number of unique targets in a column of \mathbf{A} that can be turned ON is exactly the number represented in that column. Because all subsets with $\alpha, \alpha + 1, \dots, n - 1$ K 's are redundant, here are only $(2^n - 1) - \sum_{i=\alpha}^{n-1} \binom{n}{i}$ subsets of K 's that contribute to unique configurations, leaving a total of $[(2^n - 1) - \sum_{i=\alpha}^{n-1} \binom{n}{i}](2^m - 1) + 1$ unique configurations.

If the P 's also have redundant connections, the result generalizes to

Theorem 1 *The number of configurations in combinatorial logic with parameters n, m, l_n, l_m and $u = 2$ is*

$$(2^n - 1)(2^m - 1) + 1 - \left(\sum_{i=\alpha}^{n-1} \binom{n}{i} \right) (2^m - 1) - (2^n - 1) \left(\sum_{i=\beta}^{m-1} \binom{m}{i} \right) + \left(\sum_{i=\alpha}^{n-1} \binom{n}{i} \right) \left(\sum_{i=\beta}^{m-1} \binom{m}{i} \right), \quad (14)$$

where α (resp. β) is the smallest number of K 's (resp. P 's) having the same effect as all K 's (resp. P 's) at once.

This result is obtained by counting all pairings of K 's and P 's, then subtracting those pairings that have a redundant effect. For example, any combination using K_3 is redundant in the connectivity matrix of Eq (6). Finally, those pairings that were excluded twice are added back in.

This result generalizes to all u with slight modifications. Because one factor from each of u pools is now required, the combinatorial equation determining state of a target is

$$(\{K\}_1, \{K\}_2, \dots, \{K\}_u) [A_{ij\dots k}] = \begin{cases} 1 & \text{if } K_{i1} \in \{K\}_1, K_{i2} \in \{K\}_2, \dots, K_{ik} \in \{K\}_u \\ 0 & \text{else.} \end{cases} \quad (15)$$

Here the double subscript K_{ik} indicates the k^{th} factor in the i^{th} pool. Determining α_i for each pool i of regulators requires finding the pool $j \neq i$ which maximizes the number N_i of targets controlled in two dimensions. If we choose α_i or more regulators in the i^{th} pool, then there is a redundancy in the j^{th} dimension, whereas any choice of fewer than α_i regulators activates fewer than N_i targets. Write $N_i = \max_{j \neq i} \left\{ \binom{n_i}{l_{ni}} \binom{n_j}{l_{nj}} \right\}$ the total number of targets and $p_{nj} = \frac{N_i}{n_j} l_{nj}$ the number of targets in any column of the equivalent $n_i \times n_j$ connectivity matrix regulated by pools i and j . It is easy to see that these parameters reduce to their previous definitions for $u = 2$. Now define $M_i = \binom{n_i-1}{l_{ni}-1} \binom{n_j-1}{l_{nj}-1}$ as the number of targets in each entry of the equivalent $n_i \times n_j$ connectivity matrix. As above, α_i is now the smallest r such that

$$M_i \times r - \sum_{r'=2}^{\min(r, l_{ni})} (-1)^{r'} \binom{n_i - r'}{l_{ni} - r'} \binom{n_j - 1}{l_{nj} - 1} \geq p_{nj}. \quad (16)$$

Once α_i is determined for each pool i , the inclusion-exclusion sum can be extended using standard arguments [59]. Define by

$$S^{(k)} = \sum_{\sigma \in \binom{u}{k}} \left\{ \prod_{i \in \sigma} \left(\sum_{j=\alpha_i}^{n_i-1} \binom{n_i}{j} \right) \prod_{i \notin \sigma} (2^{n_i} - 1) \right\}, \quad (17)$$

where σ denotes all k -subsets of $\{1, \dots, u\}$. Then we have the final result

Theorem 2 *The total number of configurations in combinatorial logic with u pools and parameters $n_i, l_{ni}, i \in \{1, \dots, u\}$ is*

$$S = 1 + \sum_{k=0}^u (-1)^k S^{(k)}. \quad (18)$$

This result reduces to Theorem 1 when there are only $u = 2$ pools. At most there are $\prod_{i=1}^u (2^{n_i} - 1)$ ways to specify at least one target, corresponding to the 0^{th} -order term in Eq (18). Increasing the connectivity through the l_{ni} can only reduce the number of configurations. This behavior is shown in S1B Fig for the symmetric case that all the n_i and l_{ni} are equal. As u is increased the number of configurations increases dramatically, but the scaling is actually sub-exponential, i.e. less than 2^N . Increasing connectivity through l_{ni} shifts the curves to the right.

3. Using the connectivity matrix to establish a one-to-one correspondence between the ratchet network and the lonesum matrices

To establish the correspondence between the reachable configurations of ratchet network ($l_n = l_m = 1, T = 1$) and the lonesum matrices, we must show (i) that A avoids the patterns $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ in any 2×2 sub-block, and (ii) that any lonesum matrix can be constructed

from K and P actions. First observe that the value 1 in $\mathbf{A}_{i,j}$ indicates the last K affecting that index must have followed a P , whereas 0 indicates the last P must have followed a K . For the first restriction we have $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ implies $\begin{pmatrix} P_1 \dots K_1 & K_1 \dots P_2 \\ K_2 \dots P_1 & P_2 \dots K_2 \end{pmatrix}$. This means P_2 follows K_1 follows P_1 follows K_2 follows P_2 , which is a contradiction, showing that this 2×2 block is unreachable. The other five unique 2×2 blocks are all reachable with elementary sequences. This establishes point (i) that the reachable configurations are a subset of the lonesum matrices.

To establish point (ii) that the lonesum matrices are a subset of the reachable configurations, we use an equivalent formulation of the lonesum matrices as *staircase matrices* composed of the rows $a_j = (1, \dots, 1, 0, \dots, 0)$ with the last 1 appearing at position i_j subject to the constraint that $i_j \leq i_{j-1}$ for all $\forall j \in \{2, \dots, n\}$ [109]. It is easy to see that the pattern of ones resembles an inverted staircase. We show via induction that any staircase matrix can be constructed from K and P actions. The n^{th} row is obtained by the sequence $K_n P_{i_n+1} \dots P_m$ which leaves 1's at the first i_n indices and 0's at the remainder. Now assume that the k^{th} row is obtained by the sequence $K_k P_{i_k+1} \dots P_m$ without affecting any of the rows $n, n-1, \dots, k+1$. Then the sequence $K_{k-1} P_{i_{k-1}+1} \dots P_m$ puts 1's at the first i_{k-1} indices of row $k-1$. Because $i_{k-1} \geq i_k \geq \dots \geq i_n$, none of the $P_{i_{k-1}+1}, \dots, P_m$ turn a 1 to a 0 in rows $n, n-1, \dots, k+1, k$. This proves the induction hypothesis and shows that the staircases matrices are a subset of the reachable configurations.

Together with the fact that the reachable configurations are a subset of the staircase matrices, this implies that the reachable configurations and the lonesum matrices are in fact the same set, and we have

Theorem 3 *The number of reachable configurations in the (n, m) ratchet network with $l_n = l_m = 1$ and threshold 1 scales as the poly-Bernoulli numbers $B_m^{-n} = B_n^{-m}$.*

4. Inductive proof that all binary ON/OFF configurations are reachable in the ratchet network with threshold greater than 1

With $T = 2$, only targets in state 2 are ON. Once a 0-1 configuration of \mathbf{A} is obtained, however, it is a simple matter to convert it into an ON/OFF configuration by applying all the K 's. Here we use the fact that 1's can be reached from above and below to prove the

Theorem 4 *In the ratchet network represented by the matrix \mathbf{A} with $l_n = l_m = 1$ and threshold $T = 2$, all binary 0-1 configurations are reachable.*

Proof. We use an induction argument analogous to the proof of Theorem 3. Suppose that in row n a set of $r \leq m$ indices $\{n_j\} = \{n_{j_1}, \dots, n_{j_r}\}$ should be ON. First prepare every target in row n in the 1 state using K_n , then use the sequence $K_n P_{j_r+1} \dots P_{j_1}$ to obtain 2's at $\{n_{j_1}, \dots, n_{j_r}\}$ and 1's at $\{n_{j_r+1}, \dots, n_{j_m}\}$. Now assume that we can prepare rows $n, n-1, \dots, k+1$ in a similar 1-2 configuration with the rest of the matrix 0. We want to show that we can add row k to this set without affecting any of the previous rows. Assuming that a set of $s \leq m$ indices $\{k_j\}, \dots, \{k_s\}$ should be ON, apply the sequence $P_{j_1} \dots P_{j_s} K_k^2 P_{j_{s+1}} \dots P_{j_m}$ to obtain 2's at $\{k_{j_1}, \dots, k_{j_s}\}$ and 1's at $\{k_{j_s+1}, \dots, k_{j_m}\}$. Now, because $\{P_{j_1}, \dots, P_{j_s}\} \cup \{P_{j_{s+1}}, \dots, P_{j_m}\} = \{P_1, \dots, P_m\}$, all 2's and 1's in rows $n, n-1, \dots, k+1$ are now 1's and 0's, respectively. Applying the sequence $K_n K_{n-1} \dots K_{k+1}$ reestablishes the 1-2 configuration we had prior to fixing row k and leaves 0's at rows $1, \dots, k-1$. Now that row k is also in the proper 1-2 configuration, we have proved the induction hypothesis. Once all rows in the proper 1-2 configuration, the sequence $P_1 \dots P_m$ obtains the matrix in the 0-1 configuration. Since this procedure can be repeated for any collection of indices $\{\{1_j\}, \dots, \{n_j\}\}$, it follows that all binary 0-1 matrices are reachable.

5. A recursive formula for the number of non-redundant sequences in the ratchet network

When the connectivity parameters l_n and l_m exceed 1, certain sequences in the threshold 1 ratchet network become redundant. Our goals in this section are to (i) to characterize the redundant sequences by the number of K 's and P 's, and (ii) count the non-redundant sequences. This will obtain an upper bound on the number of configurations.

We want the shortest sequences that can activate or (deactivate) all targets; any sequences longer than this are redundant. To see why this is so, we need the concept of a *cycle*. We say that a target has gone through a cycle if has traversed the states 0, 1, 0 at some subsequent time points. We have the following lemma.

Lemma 5 Any sequence that takes all targets through a cycle is redundant.

Proof. The final configuration of any sequence is represented by the positions of the 1's and 0's of the connectivity matrix. Recall that $A_{i,j} = 0$ if and only if all targets represented by $A_{i,j}$ are OFF in the final configuration. Permute the rows and columns of A until it is in staircase form with $r \leq \min(n, m)$ steps, where a step is a group of adjacent rows or columns having the same number of 1's and 0's. The steps partition the rows and columns of A into subsets of indices $\{i_1, i_2, \dots, i_r\}$ and $\{j_1, j_2, \dots, j_r\}$ where the k^{th} step is defined by 1's at rows i_k to $i_{k+1} - 1$ and 0's at columns j_k to $j_{k+1} - 1$. Then the sequence $\prod_{k=1}^r K_{i_k} \dots K_{i_{k-1}} P_{j_k} \dots P_{j_{k-1}}$ obtains the desired configuration of 1's and 0's. Being able to write a staircase matrix for the final configuration means that every target ON in the final configuration occurs only where there are 1's in the matrix. These targets are never affected by a P in this procedure; they do not go through a cycle. Because any allowed configuration can be reached from this procedure, it follows that any sequence that uses a cycle is redundant.

Knowing that the non-redundant sequences must avoid cycles, it suffices to find the longest sequences that can be written before cycles appear.

Lemma 6 For each value of l_n (l_m), the maximum number of K 's (P 's) that can be used before all targets are activated (deactivated) is $n - l_n + 1$ ($m - l_m$).

Proof. A sequence that activates all targets has no intervening P 's. Recall that a single K activates at most $\frac{n}{n} l_n$ targets. Then, prior to the last K being used, the number of activated targets is $N - \frac{n}{n} l_n = \frac{n}{n} (n - l_n) \leq \frac{n}{n} l_n (n - l_n)$. This means there are at most $n - l_n$ groups of targets controlled by different K 's. Thus, at most $n - l_n$ K 's are used before the last K is used, and $n - l_n + 1$ K 's must be sufficient to activate the complete set. The maximum number of P 's that can be used is only $m - l_m$ because we can think of every sequence starting in the zero configuration as having been preceded by a single P ; this modification puts the P 's on equal footing with the K 's.

With this characterization of the non-redundant sequences our goal is to recursively eliminate sequences that use $n - l_n + 1$ K 's and $m - l_m$ P 's. We first find the number of sequences that use up to $m - l_m$ P 's, which forms the top row in each (n, m) block in [S2 Fig](#). Then we use these values to recursively find the number of sequences using up to $n - l_n + 1$ K 's. The strategy is to subtract from the total number of sequences at a given (l_n, l_m) all those sequences using the forbidden number of regulators in order to get the new total.

Denote by a_n^m the number of sequences using m P 's when the total number of K 's is n . If $m = 1$, then all $B_1^{-n} = 2^n$ sequences (except for the empty sequence) use a K and none use a P . If $m = 2$, the maximum number of P 's that can be used is $m - l_m = 1$. Discarding the 2^n sequences with no P , the number of sequences using a single P is

$$a_n^1 = \frac{B_2^{-n} - 2^n}{2}. \tag{19}$$

Division by $m = 2$ is required to account for the fact that there are $\binom{m}{1} = m$ different ways of

starting each sequence with a P , and we consider both of these equivalent. Having determined a_n^m , it is straightforward to determine a_n^{m+1} . Because there are $m + 1$ P 's to choose from, there are $\binom{m+1}{m} a_n^m$ ways to write sequences with m P 's, $\binom{m+1}{m-1} a_n^{m-1}$ ways to write sequences with $m - 1$ P 's, . . . , $\binom{m+1}{0} 1$ ways to write sequences with 0 P 's, the only remaining sequences are those with $m + 1$ P 's. Knowing that the total number of sequences is B_m^{-n} , this leaves

$$a_n^{m+1} = \frac{B_m^{-n} - 2^n - \sum_{j=0}^m \binom{m+1}{j} a_n^j}{m+1} \tag{20}$$

total sequences using $m + 1$ P 's when the total number of K 's is n . Having determined this number, we can sum up all the sequences using $m - l_m$ P 's to get the first row of the (n, m) block in [S2 Fig](#). Denote by $c_n^m(l_n, l_m)$ the l_m^{th} column and l_n^{th} row of the (n, m) block. The column headers $c_n^m(1, l_m)$ are given by

$$c_n^m(1, l_m) = 2^n + \sum_{j=0}^{m-l_m} \binom{m}{j} a_n^j. \tag{21}$$

We can determine the row entries for $l_n > 1$ in the same way that we determined the column headers, the only difference being that the total number of sequences is $c_n^m(1, l_m)$, not B_m^{-n} unless $l_m = 1$. Denote by $b_m^l(l_m)$ the number of sequences using n K 's when the total number of P 's is m and the P connectivity is l_m . For fixed m, l_m and $n = 1$, there are

$$b_m^1(l_m) = 2^m - \sum_{j=0}^{l_m-1} \binom{m}{j}, \tag{22}$$

sequences, as all but the empty sequence use a single K . In complete analogy to [Eq \(20\)](#) we find there are

$$b_m^{n+1}(l_m) = c_n^m(1, l_m) - \sum_{j=0}^{n-l_n+1} \binom{n+1}{j} b_m^j(l_m) \tag{23}$$

sequences using $n + 1$ K 's when the total number of P 's is m . Unlike in the equation for a_n^m , there is no division by $n + 1$ because all sequences starting with a different K are different. Finally, we can sum up all the sequences using $n - l_n + 1$ K 's to get the

Theorem 7 *The number of minimal length sequences in the (n, m, l_n, l_m) ratchet network with threshold $T = 1$ using no more than $n - l_n + 1$ K 's and $m - l_m$ P 's is*

$$c_n^m(l_n, l_m) = \sum_{j=0}^{n-l_n+1} \binom{n}{j} b_m^j(l_m). \tag{24}$$

We used this formula to compute each entry in [S2 Fig](#). Because of the complexity of this procedure, we checked it against a computer algorithm operating with the following steps. In step 1 find all B_m^{-n} sequences in the $l_n = l_m = 1$ case. In step 2 increase the connectivity (l_n or l_m) and find all sequences of a given length; group them by the configuration they generate. Some of these sequences will not appear in the list generated by step 1: for example, both $K_1 K_2$ and $K_2 K_1$ will be found in step 2. We are interested in index permutation e.g. $1 \rightarrow 3$, not letter permutation, so in step 3 delete all sequences in each length group not appearing in step 1. Repeat steps 1–3 with this new list of sequences until $l_n = n - 1$. This code, implemented in Matlab Version 2015b, gave exact agreement with Theorem 7.

6. Proof that the reachable configurations are equivalent to the connected one-colorings

We now show that rules restrict the reachable configurations of the sequestration model in the main text to the connected one-colorings of the reduced n -network.

Theorem 8 *There is a one-to-one correspondence between the reachable configurations of the reduced n -network and the connected one-colorings.*

Proof. The converse direction, reachable implies connected, is easier to prove and will be discussed first. Assume that all configurations in the reduced n -network so far reached are connected. The next configuration will be reached by turning all 0's to i 's or all j 's to 0's by application of K_i or P_j , respectively. The k -arm targets sharing state i with the 2-arm target $\{0, 1, i\}$ are either in the same state as some other 2-arm target $\{0, 1, i'\}$ or are in the 0 state. So application of K_i cannot change the connectivity of the configuration. Furthermore, a k -arm target can be in the j state only if the target $\{0, 1, j\}$ is in the j state, so these targets will still be matched after application of P_j . Thus, any configurations reachable from a reachable configuration must be connected.

The forward direction, connected implies reachable, is less trivial. In order to prove that all connected one-colorings in the n -network are reachable, we will use the strong form of mathematical induction. Assume the theorem holds for all networks up to $n - 1$. Embedded within the full n -network of $2^n - 1$ targets is the reduced n -network on 2^{n-1} targets. Within the reduced n -network is a set of 2^{n-2} targets able to access $\{0, 1, 2\}$ and all subsets (including \emptyset) of the integers $\{3, \dots, n\}$. Thus, we can substitute $2 \rightarrow 1$ as the ON state in this embedded network and all connected one-colorings (of 2) will be reachable. The same holds in general for all 2^{n-k} targets able to access $\{0, 1, k\}$ and all subsets of the integers $\{k + 1, \dots, n\}$. In each of these embedded networks the substitution $k \rightarrow 1$ as the ON state will enable us create any connected one-coloring.

Pick any connected one-coloring (of 1) in the n -network. Its *opposite* configuration is formed by the transformation at each target g of $1 \rightarrow 0$ and $0 \rightarrow k_{min}$, where $k_{min} = \min\{k \in g \mid \mathbf{x}_{\text{pos}(\{0, j, k\})} = 0\}$ is the smallest index that g shares with a corresponding 2-arm target at position $\text{pos}(\{0, j, k\})$ of \mathbf{x} (possibly in the full network) currently in the 0 state. The opposite of a connected one-coloring is clearly connected, because all the connected 1's are now 0, and all the 0's are in the same state as the 2-arm target $\{0, j, k_{min}\}$. If it is possible to reach the opposite configuration, then application of the sequence $K_1 P_2 \dots P_n$ yields the desired one-coloring of the n -network.

To show that the opposite configuration of the chosen one-coloring is indeed reachable, isolate the embedded networks one-by-one by application of the sequence $K_k K_1 P_k$ for $k = 2, \dots, n$, so that the targets in the $n - k + 1$ -network are the only targets in the 0 state. By hypothesis, the connected one-colorings are reachable in all embedded networks which have at most $n - k$ states besides 0, 1, and k . The opposite configuration in the n -network is composed of connected one-colorings (of k) in each embedded network; these are are reachable. Therefore, the one-coloring of the n -network is reachable via $K_1 P_2 \dots P_n$. This procedure holds for any one-coloring.

7. Lower and upper bounds for the full n -network

How many configurations are reachable in the full n -network? Let this number be c_n . The following theorems derive lower and upper bounds for c_n in terms of the number of one-colorings.

Theorem 9 *The formula $f(n + 1)$ for the number of connected one-colorings in the reduced $n + 1$ -network is a lower bound for c_n .*

Proof. The full $n + 1$ -network can be partitioned into a set of 2^n targets having a 1 and all subsets of $\{2, \dots, n + 1\}$, and $2^n - 1$ targets that lack 1 but have all nonempty subsets of $\{2, \dots, n + 1\}$. The latter set of targets is an embedded full n -network, while the former is the reduced $n + 1$ -network. All $2(n + 1)$ letters are needed to form the one-colorings in the reduced $n + 1$ -network. Every one-coloring is finally obtained by applying some permutation of K_1, P_2, \dots, P_{n+1} to a configuration that uses (at most) the states $2, \dots, n + 1$ and 0, i.e. the full n -network. Because K_1 and P_1 do not affect the targets of the the embedded full n -network, there must be (at least) one sequence using only $\{K_2, \dots, K_{n+1}\}$ and $\{P_2, \dots, P_{n+1}\}$ that prepares the embedded full n -network in the aforementioned configuration, which means we may associate a one-coloring with (at least) one of the c_n sequences in the embedded full n -network. Therefore, multiple configurations in the full n -network may map to the same one-coloring in the reduced $n + 1$ -network. Conversely, if two one-colorings are different, they are distinguishable by their configurations immediately preceding the final K_1, P_2, \dots, P_{n+1} sequence, and must therefore map to different configurations in the full n -network. Together, these statements imply that the map from configurations in the full n -network to one-colorings in the reduced $n + 1$ -network is many-to-one, but the map from one-colorings to configurations in the full n -network is one-to-one. Therefore, $f(n + 1) \leq c_n$.

Theorem 10 An upper bound on c_n is

$$\begin{aligned}
 &nf(n) + n(n - 1)(f(n) - 1)f(n - 1) + \dots + n!(f(n) - 1) \dots (f(2) - 1)f(1) + 1 \\
 &= \sum_{k=1}^n (n)_k \left\{ \prod_{j=n-k+2}^n (f(j) - 1) \right\} f(n - k + 1) + 1.
 \end{aligned} \tag{25}$$

where $(n)_k = n(n - 1) \dots (n - k + 1)$ is the falling factorial.

Proof. There are $nf(n)$ one-colorings in the full n -network, plus one origin. Each one of the one-colorings can be thought of as the origin of an $n - 1$ -network, which in turn generate $(n - 1)f(n - 1)$ one-colorings in an embedded $n - 1$ -network, for a total of

$$nf(n)(n - 1)f(n - 1)$$

configurations using 1, 2, and perhaps 0, hence termed two-colorings. However, one of the $f(n)$ one-colorings is the 0 state of the n -network, so it does not generate any two-colorings. Thus, there are at most $1 + nf(n) + n(n - 1)(f(n) - 1)f(n - 1)$ zero-, one-, and two-colorings. Now assume that the number of k -colorings is

$$n(n - 1) \dots (n - k + 1)(f(n) - 1)(f(n - 1) - 1) \dots (f(n - k + 2) - 1)f(n - k + 1).$$

Of these,

$$n(n - 1) \dots (n - k + 1)(f(n) - 1)(f(n - 1) - 1) \dots (f(n - k + 2) - 1)$$

are origins of an $n - k$ -network, meaning they are actually $k - 1$ -colorings; they cannot generate any $k + 1$ -colorings. The remaining

$$\begin{aligned}
 &n(n - 1) \dots (n - k + 1)(f(n) - 1)(f(n - 1) - 1) \\
 &\dots (f(n - k + 2) - 1)(f(n - k + 1) - 1)
 \end{aligned}$$

are genuine k -colorings which can generate $f(n - k)$ one-colorings in the $n - k$ -network, or equivalently, $k + 1$ -colorings. Thus, the total number of zero-, one-, two-, \dots , $k + 1$ -colorings is

no more than

$$n(n-1) \cdots (n-k+1)(f(n)-1)(f(n-1)-1) \cdots (f(n-k+2)-1)(f(n-k+1)-1).$$

This induction argument proves the statement.

8. Properties of the orbits in the ratchet and sequestration network

First we define what it means to be an *origin* and an *orbit* in the threshold-1 ratchet network and determine the number of orbits as a function of n and m . Then we prove that the configurations in the sequestration network are defined by reversible paths.

A forbidden configuration in the ratchet network contains some row or column permutation of the pattern $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ on any 2×2 sub-block of the connectivity matrix \mathbf{A} . This is the

minimum violation, but larger blocks may violate this pattern as well, for example $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$

has 2 violations. Furthermore, application of any of the K 's or P 's in this sub-block will relieve at least one of these violations. Therefore, we define an i, j -orbit in the ratchet network as the locus of configurations having a forbidden configuration on an $i \times j$ sub-block that does not use the corresponding set of $i K$'s and $j P$'s. The *origin* of any i, j -orbit is the configuration having all remaining $nm - ij$ entries of \mathbf{A} equal to 0 (or all 1 to make the case of having only P actions symmetric with having only K 's). A matrix \mathbf{X} having the same forbidden $i \times j$ sub-block as an origin \mathbf{Y} is not considered to be in the orbit of \mathbf{Y} if (i) there is no sequence of actions that transforms \mathbf{Y} to \mathbf{X} , or (ii) if the sequence involves one of the forbidden K 's or P 's. With these restrictions, the number of origins is equal to the number of orbits.

Denote by c_i^j the number of orbits in a ratchet network of size $n \times m$ with violations involving $i \leq n K$'s and $j \leq m P$'s. If $i = j = 2$ there are $2^{ij} - B_2^{-2} = 2$ forbidden configurations that turn into origins for the remaining $n - i K$'s and $n - j P$'s. There are more orbits in these smaller networks. For every $i', j' \geq 2$ there are $\binom{i}{i'} \binom{j}{j'} c_{i-i'}^{j-j'}$ configurations reached by orbits using i' K 's and j' P 's. Only configurations *not* reached by these orbits are available as new origins when the number of K 's and P 's not to be used is i and j , respectively. Finally, there are $\binom{n}{i} \binom{m}{j}$ ways to specify $i \leq n K$'s and $j \leq m P$'s. Then we have the

Theorem 11 For a given set of $i \leq n K$'s and $j \leq m P$'s, the number of i, j -orbits is

$$c_i^j = (2^{ij} - B_i^{-j}) - \sum_{\substack{i', j' \geq 2 \\ i'+j' \leq i+j-1}} \binom{i}{i'} \binom{j}{j'} c_{i-i'}^{j-j'}, \tag{26}$$

and the total number of i, j -orbits in the $n \times m$ ratchet network is

$$C_i^j(n, m) = \binom{n}{i} \binom{m}{j} c_i^j, \tag{27}$$

where

$$B'_{i-i'}^{-(j-j')} = \begin{cases} B_{i-i'}^{-(j-j')} & \text{if } i - i' > 0 \text{ and } j - j' > 0 \\ 2^{i-i'} & \text{if } j - j' = 0 \\ 2^{j-j'} & \text{if } i - i' = 0 \end{cases} \quad (28)$$

The modification B' ensures that an orbit lacking allowable P 's (K 's) can still use K 's (P 's). A table of values of Eq (27) is given in S4 Fig.

We noted in the main text that configuration in the sequestration network can be joined by reversible paths. A path $K_i P_j$ or $P_j K_i$ is *reversible* if a configuration reached by the sequence of actions w is also reached by either the sequence $wK_i P_j$ or $wP_j K_i$, but not wK_i or wP_j , respectively. Thus we can also prove the

Theorem 12 *There always exists a reversible path between any two configurations in an orbit of the sequestration network.*

Proof. Let \mathbf{x} be a configuration in an orbit using $m \leq n$ of the actions, and let P denote the locus of configurations reached from \mathbf{x} . We now need to show that P must be reversibly reached from the origin. Denote by \bar{P} the complement of P , so that any $\mathbf{y} \in \bar{P}$ is reversibly reached from the origin. In order for there to be no reversible path between $\mathbf{x} \in P$ and $\mathbf{y} \in \bar{P}$, there must always be a state i such that K_i increases the number of targets $\{\cdot, i\}$ in the i state and P_i increases the number of targets $\{\cdot, i\}$ in the zero state. Now assume there is a configuration $\mathbf{z} \in P$ using all m allowed states. \mathbf{z} must have at least one target in the 0 state, but this is unallowed, because then \mathbf{z} would violate the connection rule. Therefore, there is a maximum number $m' < m$ of states used by any $\mathbf{x} \in P$. Now assume there is a configuration $\mathbf{z}' \in P$ using all m' allowed states. But this implies that there is a single-arm target $\{0, j\}$ that must be in the zero state. Then the action K_j takes \mathbf{z}' to a configuration $\mathbf{y} \in \bar{P}$ and P_j takes \mathbf{y} to \mathbf{z} . This path must be reversible, and \mathbf{z}' is reached reversibly from the origin. By induction we conclude that $m' = 0$ and that $P = \emptyset$. Finally, because any two configurations are reached reversibly from the origin, there is a reversible path between them.

Theorem 12 defines the orbits of the sequestration network as those configurations connected by reversible paths.

9. A universal formulation of the actions as matrix operators

In this section we show how to write the K and P regulators as matrix operators in a manner consistent with both models considered in the paper. First we define the vector space \mathcal{V} of configurations of the N targets, then we derive the operators that transform \mathcal{V} .

Let $\mathbf{x} \in \mathcal{V}$. For a network with N targets we require that $\sum_i x_i = N$. This means that \mathbf{x} has at least N entries, and in general $\dim \mathbf{x} \geq N$. Therefore, we cannot use the standard state space of N -dimensional vectors, because the operators will not conserve the number of targets. Each target has a 0 state. The number D of independent directions accessible from 0 is called the *dimension* of the network, and the number T of steps one can move along each dimension is called the *threshold*. In the ratchet model, each target has a single ladder of states with variable threshold, so $D = 1$ and T is allowed to vary; in the sequestration model $D = n$ but the threshold is $T = 1$.

Denote by A_{di} the fraction of the targets of type A in state $i \in \{0, 1, \dots, T\}$ along dimension d . For a subset of the targets a K -type action causes population transfer between states (d, j) and (d, i) with $i = j + 1$, and a P -type action the reverse. If a K regulator acts for a short time we

can write the “reaction rate” equation as

$$\begin{aligned}\dot{x}_{A_j} &= -g_A x_{A_j} \\ \dot{x}_{A_i} &= +g_A x_{A_j}\end{aligned}\tag{29}$$

where $g_A > 0$ is a proportionality constant. This defines a matrix differential equation

$$\dot{\mathbf{x}} = \mathbf{G}_{dj} \cdot \mathbf{x}\tag{30}$$

with $\mathbf{x} \in \mathbb{R}^{N(DT+1) \times 1}$ the vector of populations of the $DT + 1$ states of the N targets and $\mathbf{G}_{dj} \in \mathbb{R}^{N(DT+1) \times N(DT+1)}$ the block diagonal matrix of rate constants between the j and $j + 1$ population states along dimension d . Eq (29) can be rewritten

$$\begin{pmatrix} \dot{x}_{A_j} \\ \dot{x}_{A_i} \end{pmatrix} = \begin{pmatrix} -g_A & 0 \\ g_A & 0 \end{pmatrix} \cdot \begin{pmatrix} x_{A_j} \\ x_{A_i} \end{pmatrix}.\tag{31}$$

Because \mathbf{G}_{dj} is block diagonal, Eq (30) can be solved by exponentiation on each block:

$$\begin{pmatrix} x_{A_j}(t) \\ x_{A_i}(t) \end{pmatrix} = \exp \left\{ \begin{pmatrix} -g_A & 0 \\ g_A & 0 \end{pmatrix} t \right\} \cdot \begin{pmatrix} x_{A_j}(0) \\ x_{A_i}(0) \end{pmatrix} = \begin{pmatrix} e^{-g_A t} & 0 \\ 1 - e^{-g_A t} & 1 \end{pmatrix} \cdot \begin{pmatrix} x_{A_j}(0) \\ x_{A_i}(0) \end{pmatrix}.\tag{32}$$

The restriction of the model from a continuous range of population states $x_{A_i} \in [0, 1]$ to the boolean values $\{0, 1\}$ formally emerges by considering the “reaction” K catalyzes on its targets to have gone to completion. We do this by taking the limit $t \rightarrow \infty$ in Eq (32) to get

$$\begin{pmatrix} x_{A_j}(t) \\ x_{A_i}(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_{A_j}(0) \\ x_{A_i}(0) \end{pmatrix},\tag{33}$$

so that the matrix \mathbf{K}_{dj} defined by

$$\mathbf{K}_{dj} = \lim_{t \rightarrow \infty} \exp \left\{ \mathbf{G}_{dj} t \right\}\tag{34}$$

is the block diagonal matrix having 1’s at (row, column) positions $(1 + (d - 1)T + i, 1 + (d - 1)T + j)$ of each block that responds to K in dimension d and admits population transfer between from j to i .

Because K acts on all targets at once, it is insensitive to the initial state j . Thus the matrix corresponding to the action of K is

$$\mathbf{K}_d = \prod_j \mathbf{K}_{dj},\tag{35}$$

which is the block diagonal matrix having 1’s at (row, column) positions

$$\begin{aligned}(1 + (d - 1)T + 1, 1 + (d - 1)T + 0), \dots, (1 + (d - 1)T + T, 1 + (d - 1)T + T - 1) \\ \text{and } (1 + (d - 1)T + T, 1 + (d - 1)T + T)\end{aligned}$$

of each block that responds to K in dimension d .

This derivation can be repeated in the case that population goes in the opposite direction from at state j to a state $i < j$ using a different set of rate matrices \mathbf{H}_{dj} corresponding to the reverse of Eq (31). We obtain the block diagonal matrix \mathbf{P}_d corresponding to the action of P in

dimension d having 1's at (row, column) positions

$$(1 + (d - 1)T + 0, 1 + (d - 1)T + 1), \dots, (1 + (d - 1)T + T - 1, 1 + (d - 1)T + T)$$

and $(1 + (d - 1)T + 0, 1 + (d - 1)T + 0)$

of each block that responds to P in dimension d . Whereas \mathbf{K}_d is sub-diagonal, \mathbf{P}_d is super-diagonal.

The Baker-Campbell-Hausdorff expansion shows that \mathbf{K}_d in Eq (35) and in general any product of matrices \mathbf{K}_d and \mathbf{P}_d are generated by matrix exponentiation of commutators of the generators $\mathbf{G}_{dj}, \mathbf{H}_{dj}$. This is the origin of noncommutativity in both the ratchet and sequestration models.

An example in the sequestration network illustrates population transfer between states. In the $n = 2$ network on the targets A, B , and C the initial configuration of the network is represented by

$(A_0 \ A_{11} \ A_{21} \ B_0 \ B_{11} \ B_{21} \ C_0 \ C_{11} \ C_{21})^T = (1 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1 \ 0 \ 0)^T$. Only targets A and C can access dimension 1, and only targets B and C can access dimension 2. Therefore the $t \rightarrow \infty$ action of K_1 on the network is given by

$$\begin{pmatrix} e^{-g_A t} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 - e^{-g_A t} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & e^{-g_C t} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 - e^{-g_C t} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \xrightarrow{t \rightarrow \infty} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \tag{36}$$

Only A and C advance to state 1 and the number of targets (3) is conserved.

Supporting Information

S1 Fig. Scaling in combinatorial networks is sub-exponential. (A) An example network with $u = 3$ pools of $n = 2$ regulators each. A target is only ON if all u of its regulators bind. (B) Plots of Eq (18) vs. n for an increasing number of pools u and increasing redundancy l_n . (TIF)

S2 Fig. Number of unique words in the threshold 1 ratchet network as a function of n, m, l_n , and l_m found using Eq (24). n and m increase across the rows and up the columns. l_n and l_m increase down the columns and across the rows of the sub-blocks. (TIF)

S3 Fig. The full n -network model has upper and lower bounds. (A) A plot of all the allowed configurations of a set of targets controlled by $n = 3$ regulators pairs in the full n -network. Blue, cyan, yellow, and red correspond to states 0, 1, 2, and 3, respectively. (B) A list of the words generating the corresponding states in A. K actions are shown in the red spectrum, and P in the

blue. (C) A logarithmic plot of the bounds on the full model. The total space is $\prod_{i=0}^n (i+1)^{\binom{n}{i}}$, the upper and lower bounds are calculated from Eqs (24) and (7), respectively, and the combinatorial model is 2^{2^n} .

(TIF)

S4 Fig. Number of orbits restricted from using i of the K 's and j of the P 's in the threshold 1 ratchet network as a function of n and m calculated using Eq (27). n and m increase the across the rows and up the columns. i and j increase down the columns and across the rows of the sub-blocks.

(TIF)

Author Contributions

Conceived and designed the experiments: LC WL.

Performed the experiments: WL.

Analyzed the data: WL LC.

Contributed reagents/materials/analysis tools: WL.

Wrote the paper: WL LC.

References

1. Ma HW, Kumar B, Ditges U, Gunzer F, Buer J, Zeng AP. An extended transcriptional regulatory network of *Escherichia coli* and analysis of its hierarchical structure and network motifs. *Nucleic Acids Res.* 2004; 32:6643–6649. doi: [10.1093/nar/gkh1009](https://doi.org/10.1093/nar/gkh1009) PMID: [15604458](https://pubmed.ncbi.nlm.nih.gov/15604458/)
2. Balaji S, Babu MM, Iyer LM, Luscombe NM, Aravind L. Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. *J Mol Biol.* 2006; 360:213–227. doi: [10.1016/j.jmb.2006.04.029](https://doi.org/10.1016/j.jmb.2006.04.029) PMID: [16762362](https://pubmed.ncbi.nlm.nih.gov/16762362/)
3. Guelzim N, Bottani S, Bourgine P, Képès F. Topological and causal structure of the yeast transcriptional regulatory network. *Nat Genet.* 2002; 31:60–63. doi: [10.1038/ng873](https://doi.org/10.1038/ng873) PMID: [11967534](https://pubmed.ncbi.nlm.nih.gov/11967534/)
4. Pfeiffer A, Shi H, Tepperman JM, Zhang Y, Quail PH. Combinatorial complexity in a transcriptionally centered signaling hub in Arabidopsis. *Mol Plant.* 2014; 7:1598–1618. doi: [10.1093/mp/ssu087](https://doi.org/10.1093/mp/ssu087) PMID: [25122696](https://pubmed.ncbi.nlm.nih.gov/25122696/)
5. Rhee DY, Cho DY, Zhai B, Slattery M, Ma L, Mintseris J, et al. Transcription factor networks in *Drosophila melanogaster*. *Cell Rep.* 2014; 8:2031–2043. doi: [10.1016/j.celrep.2014.08.038](https://doi.org/10.1016/j.celrep.2014.08.038) PMID: [25242320](https://pubmed.ncbi.nlm.nih.gov/25242320/)
6. Neph S, Stergachis AB, Reynolds A, Sandstrom R, Borenstein E, Stamatoyanopoulos J. Circuitry and dynamics of human transcription factor regulatory networks. *Cell.* 2012; 150:1274–1286. doi: [10.1016/j.cell.2012.04.040](https://doi.org/10.1016/j.cell.2012.04.040) PMID: [22959076](https://pubmed.ncbi.nlm.nih.gov/22959076/)
7. Hua S, Kittler R, White KP. Genomic antagonism between retinoic acid and estrogen signaling in breast cancer. *Cell.* 2009; 137:1259–1271. doi: [10.1016/j.cell.2009.04.043](https://doi.org/10.1016/j.cell.2009.04.043) PMID: [19563758](https://pubmed.ncbi.nlm.nih.gov/19563758/)
8. Lu H, Ward MG, Adeola O, Ajuwon KM. Regulation of adipocyte differentiation and gene expression-crosstalk between TGF β and wnt signaling pathways. *Mol Bio Rep.* 2013; 40:5237–5245. doi: [10.1007/s11033-013-2623-2](https://doi.org/10.1007/s11033-013-2623-2)
9. Iwanaszko M, Kimmel M. NF- κ B and IRF pathways: cross-regulation on target genes promoter level. *BMC Genomics.* 2015; 16:307. doi: [10.1186/s12864-015-1511-7](https://doi.org/10.1186/s12864-015-1511-7) PMID: [25888367](https://pubmed.ncbi.nlm.nih.gov/25888367/)
10. Varjosalo M, Keskitalo S, Drogen AV, Nurkkala H, Vichalkovski A, Aebersold R, et al. The protein interaction landscape of the human CMGC kinase group. *Cell Rep.* 2013; 3:1306–1320. doi: [10.1016/j.celrep.2013.03.027](https://doi.org/10.1016/j.celrep.2013.03.027) PMID: [23602568](https://pubmed.ncbi.nlm.nih.gov/23602568/)
11. Sacco F, Peretto L, Castagnoli L, Cesareni G. The human phosphatase interactome: An intricate family portrait. *FEBS Lett.* 2012; 586:2732–2739. doi: [10.1016/j.febslet.2012.05.008](https://doi.org/10.1016/j.febslet.2012.05.008) PMID: [22626554](https://pubmed.ncbi.nlm.nih.gov/22626554/)
12. Sasaki T, Minamisawa G, Takahashi N, Matsuki N, Ikegaya Y. Reverse optical trawling for synaptic connections in situ. *J Neurophysiol.* 2009; 102:636–643. doi: [10.1152/jn.00012.2009](https://doi.org/10.1152/jn.00012.2009)

13. Vickaryous MK, Hall BK. Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biol Rev Camb Philos Soc.* 2006; 81:425–455. doi: [10.1017/S1464793106007068](https://doi.org/10.1017/S1464793106007068) PMID: [16790079](https://pubmed.ncbi.nlm.nih.gov/16790079/)
14. Ohnishi Y, Huber W, Tsumura A, Kang M, Xenopoulos P, Kurimoto K, et al. Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nat Cell Biol.* 2014; 16:27–37. doi: [10.1038/ncb2881](https://doi.org/10.1038/ncb2881) PMID: [24292013](https://pubmed.ncbi.nlm.nih.gov/24292013/)
15. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mo Biol.* 2013; 20:1131–1139. doi: [10.1038/nsmb.2660](https://doi.org/10.1038/nsmb.2660)
16. Zeisel A, Muñoz-Manchado AB, Codeluppi S, Lönnerberg P, La Manno G, Juréus A, et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* 2015; 347:1138–1142. doi: [10.1126/science.aaa1934](https://doi.org/10.1126/science.aaa1934) PMID: [25700174](https://pubmed.ncbi.nlm.nih.gov/25700174/)
17. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, et al. Massively parallel Single-Cell RNA-seq for marker-free decomposition of tissues into cell types. *Science.* 2014; 343:776–779. doi: [10.1126/science.1247651](https://doi.org/10.1126/science.1247651) PMID: [24531970](https://pubmed.ncbi.nlm.nih.gov/24531970/)
18. Moignard V, Woodhouse S, Haghverdi L, Lilly AJ, Tanaka Y, Wilkinson AC, et al. Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat Biotechnol.* 2015; 33:269–276. doi: [10.1038/nbt.3154](https://doi.org/10.1038/nbt.3154) PMID: [25664528](https://pubmed.ncbi.nlm.nih.gov/25664528/)
19. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature.* 2014; 509:371–375. doi: [10.1038/nature13173](https://doi.org/10.1038/nature13173) PMID: [24739965](https://pubmed.ncbi.nlm.nih.gov/24739965/)
20. Brunskill EW, Park JS, Chung E, Chen F, Magella B, Potter SS. Single cell dissection of early kidney development: multilineage priming. *Development.* 2014; 141:3093–3101. doi: [10.1242/dev.110601](https://doi.org/10.1242/dev.110601) PMID: [25053437](https://pubmed.ncbi.nlm.nih.gov/25053437/)
21. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med.* 2004; 10:789–799. doi: [10.1038/nm1087](https://doi.org/10.1038/nm1087) PMID: [15286780](https://pubmed.ncbi.nlm.nih.gov/15286780/)
22. Gerhart J, Kirschner M. *Cells, Embryos, and Evolution.* Blackwell Science; 1997.
23. Wilson NK, Foster SD, Wang X, Knezevic K, Schütte J, Kaimakis P, et al. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell.* 2010; 7:532–542. doi: [10.1016/j.stem.2010.07.016](https://doi.org/10.1016/j.stem.2010.07.016) PMID: [20887958](https://pubmed.ncbi.nlm.nih.gov/20887958/)
24. Kaplan T, Li XY, Sabo PJ, Thomas S, Stamatoyannopoulos JA, Biggin MD, et al. Quantitative models of the mechanisms that control genome-wide patterns of transcription factor binding during early *Drosophila* development. *PLoS Genet.* 2011; 7:e1001290. doi: [10.1371/journal.pgen.1001290](https://doi.org/10.1371/journal.pgen.1001290) PMID: [21304941](https://pubmed.ncbi.nlm.nih.gov/21304941/)
25. He X, Samee MA, Blatti C, Sinha S. Thermodynamics-based models of transcriptional regulation by enhancers: the roles of synergistic activation, cooperative binding and short-range repression. *PLoS Comput Biol.* 2010; 6:e1000935. doi: [10.1371/journal.pcbi.1000935](https://doi.org/10.1371/journal.pcbi.1000935) PMID: [20862354](https://pubmed.ncbi.nlm.nih.gov/20862354/)
26. Thomas R. Boolean formalization of genetic control circuits. *J Theor Biol.* 1973; 42:563–585. doi: [10.1016/0022-5193\(73\)90247-6](https://doi.org/10.1016/0022-5193(73)90247-6) PMID: [4588055](https://pubmed.ncbi.nlm.nih.gov/4588055/)
27. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol.* 1969; 22:437–467. doi: [10.1016/0022-5193\(69\)90015-0](https://doi.org/10.1016/0022-5193(69)90015-0) PMID: [5803332](https://pubmed.ncbi.nlm.nih.gov/5803332/)
28. Davidich MI, Bornholdt S. Boolean network model predicts cell cycle sequence of fission yeast. *PLoS ONE.* 2008; 3:e1672. doi: [10.1371/journal.pone.0001672](https://doi.org/10.1371/journal.pone.0001672) PMID: [18301750](https://pubmed.ncbi.nlm.nih.gov/18301750/)
29. Fauré A, Naldi A, Chaouiya C, Thieffry D. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics.* 2006; 22:e124–e131. doi: [10.1093/bioinformatics/btl210](https://doi.org/10.1093/bioinformatics/btl210) PMID: [16873462](https://pubmed.ncbi.nlm.nih.gov/16873462/)
30. Buchler NE, Gerland U, Hwa T. On schemes of combinatorial transcription logic. *Proc Natl Acad Sci U S A.* 2003; 100:5136–5141. doi: [10.1073/pnas.0930314100](https://doi.org/10.1073/pnas.0930314100) PMID: [12702751](https://pubmed.ncbi.nlm.nih.gov/12702751/)
31. Peter IS, Faure E, Davidson EH. Predictive computation of genomic logic processing functions in embryonic development. *Proc Natl Acad Sci U S A.* 2012; 109:16434–16442. doi: [10.1073/pnas.1207852109](https://doi.org/10.1073/pnas.1207852109) PMID: [22927416](https://pubmed.ncbi.nlm.nih.gov/22927416/)
32. Isshiki T, Pearson B, Holbrook S, Doe CQ. *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell.* 2001; 106:511–521. doi: [10.1016/S0092-8674\(01\)00465-2](https://doi.org/10.1016/S0092-8674(01)00465-2) PMID: [11525736](https://pubmed.ncbi.nlm.nih.gov/11525736/)
33. Chan EAW, Teng G, Corbett E, Choudhury KR, Bassing CH, Krangel DGMS. Peripheral subnuclear positioning suppresses *Tcrb* recombination and segregates *Tcrb* alleles from RAG2. *Proc Natl Acad Sci U S A.* 2013; 110:E4628–E4637. doi: [10.1073/pnas.1310846110](https://doi.org/10.1073/pnas.1310846110) PMID: [24218622](https://pubmed.ncbi.nlm.nih.gov/24218622/)

34. Sayegh CE, Jhunjhunwala S, Riblet R, Murre C. Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes Dev.* 2005; 19:322–327. doi: [10.1101/gad.1254305](https://doi.org/10.1101/gad.1254305) PMID: [15687256](https://pubmed.ncbi.nlm.nih.gov/15687256/)
35. Min H, Danilenko DM, Scully SA, Bolon B, Ring BD, Tarpley JE, et al. Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to Drosophila branchless. *Genes Dev.* 1998; 12:3156–3161. doi: [10.1101/gad.12.20.3156](https://doi.org/10.1101/gad.12.20.3156) PMID: [9784490](https://pubmed.ncbi.nlm.nih.gov/9784490/)
36. Shifley ET, Kenny AP, Rankin SA, Zorn AM. Prolonged FGF signaling is necessary for lung and liver induction in *Xenopus*. *BMC Dev Biol.* 2012; 6:27. doi: [10.1186/1471-213X-12-27](https://doi.org/10.1186/1471-213X-12-27)
37. Iyengar L, Wang Q, Rasko JE, McAvoy JW, Lovicu FJ. Duration of ERK1/2 phosphorylation induced by FGF or ocular media determines lens cell fate. *Differentiation.* 2007; 75:662–668. doi: [10.1111/j.1432-0436.2007.00167.x](https://doi.org/10.1111/j.1432-0436.2007.00167.x) PMID: [17381542](https://pubmed.ncbi.nlm.nih.gov/17381542/)
38. Wada N, Nohno T. Differential Response of Shh Expression Between Chick Forelimb and Hindlimb Buds by FGF-4. *Dev Dyn.* 2001; 221:402–411. doi: [10.1002/dvdy.1150](https://doi.org/10.1002/dvdy.1150) PMID: [11500977](https://pubmed.ncbi.nlm.nih.gov/11500977/)
39. Iwasaki H, Mizuno S, Arinobu Y, Ozawa H, Mori Y, Shigematsu H, et al. The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev.* 2006; 20:3010–3021. doi: [10.1101/gad.1493506](https://doi.org/10.1101/gad.1493506) PMID: [17079688](https://pubmed.ncbi.nlm.nih.gov/17079688/)
40. Ginther MR, Ramus DFWSJ. Hippocampal neurons encode different episodes in an overlapping sequence of odors task. *J Neurosci.* 2011; 31:2706–2711. doi: [10.1523/JNEUROSCI.3413-10.2011](https://doi.org/10.1523/JNEUROSCI.3413-10.2011) PMID: [21325539](https://pubmed.ncbi.nlm.nih.gov/21325539/)
41. Agster KL, Fortin NJ, Eichenbaum H. The hippocampus and disambiguation of overlapping sequences. *J Neurosci.* 2002; 22:5760–5768. PMID: [12097529](https://pubmed.ncbi.nlm.nih.gov/12097529/)
42. Moro SI, Tolboom M, Khayat PS, Roelfsema PR. Neuronal activity in the visual cortex reveals the temporal order of cognitive operations. *J Neurosci.* 2010; 30:16293–16303. doi: [10.1523/JNEUROSCI.1256-10.2010](https://doi.org/10.1523/JNEUROSCI.1256-10.2010) PMID: [21123575](https://pubmed.ncbi.nlm.nih.gov/21123575/)
43. Ham TS, Lee SK, Keasling JD, Arkin AP. Design and construction of a double inversion recombination switch for heritable sequential genetic memory. *PLoS One.* 2008; 3:e2815. doi: [10.1371/journal.pone.0002815](https://doi.org/10.1371/journal.pone.0002815) PMID: [18665232](https://pubmed.ncbi.nlm.nih.gov/18665232/)
44. Payne JL, Wagner A. Constraint and contingency in multifunctional gene regulatory circuits. *PLoS Comput Biol.* 2013; 9:e1003071. doi: [10.1371/journal.pcbi.1003071](https://doi.org/10.1371/journal.pcbi.1003071) PMID: [23762020](https://pubmed.ncbi.nlm.nih.gov/23762020/)
45. Eldar A. Social conflict drives the evolutionary divergence of quorum sensing. *Proc Natl Acad Sci U S A.* 2011; 108:13635–13640. doi: [10.1073/pnas.1102923108](https://doi.org/10.1073/pnas.1102923108) PMID: [21807995](https://pubmed.ncbi.nlm.nih.gov/21807995/)
46. Wang X, Minasov G, Shoichet BK. Evolution of an Antibiotic Resistance Enzyme Constrained by Stability and Activity Trade-offs. *J Mol Bio.* 2002; 320:85–95. doi: [10.1016/S0022-2836\(02\)00400-X](https://doi.org/10.1016/S0022-2836(02)00400-X)
47. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci U S A.* 2008; 105:7899–7906. doi: [10.1073/pnas.0803151105](https://doi.org/10.1073/pnas.0803151105) PMID: [18524956](https://pubmed.ncbi.nlm.nih.gov/18524956/)
48. Albeck JG, Mills GB, Brugge JS. Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. *Mol Cell.* 2013; 49:249–261. doi: [10.1016/j.molcel.2012.11.002](https://doi.org/10.1016/j.molcel.2012.11.002) PMID: [23219535](https://pubmed.ncbi.nlm.nih.gov/23219535/)
49. Nelson DE, Ihekweba AE, Elliott M, Johnson JR, Gibney CA, Foreman BE, et al. Oscillations in NF- κ B signaling control the dynamics of gene expression. *Science.* 2004; 306:704–708. doi: [10.1126/science.1099962](https://doi.org/10.1126/science.1099962) PMID: [15499023](https://pubmed.ncbi.nlm.nih.gov/15499023/)
50. Kellogg RA, Tian C, Lipniacki T, Quake SR, Tay S. Digital signaling decouples activation probability and population heterogeneity. *elife.* 2015; 4:e08931. doi: [10.7554/eLife.08931](https://doi.org/10.7554/eLife.08931) PMID: [26488364](https://pubmed.ncbi.nlm.nih.gov/26488364/)
51. Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine AJ, Elowitz MB, et al. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat Genet.* 2004; 36:147–150. doi: [10.1038/ng1293](https://doi.org/10.1038/ng1293) PMID: [14730303](https://pubmed.ncbi.nlm.nih.gov/14730303/)
52. Lin Y, Sohn CH, Dalal CK, Cai L, Elowitz MB. Combinatorial gene regulation by modulation of relative pulse timing. *Nature.* 2015; 527:54–58. doi: [10.1038/nature15710](https://doi.org/10.1038/nature15710) PMID: [26466562](https://pubmed.ncbi.nlm.nih.gov/26466562/)
53. Cai L, Dalal CK, Elowitz MB. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature.* 2008; 455:485–490. doi: [10.1038/nature07292](https://doi.org/10.1038/nature07292) PMID: [18818649](https://pubmed.ncbi.nlm.nih.gov/18818649/)
54. Dalal CK, Cai L, Lin Y, Rahbar K, Elowitz MB. Pulsatile dynamics in the yeast proteome. *Curr Biol.* 2014; 24:2189–2194. doi: [10.1016/j.cub.2014.07.076](https://doi.org/10.1016/j.cub.2014.07.076) PMID: [25220054](https://pubmed.ncbi.nlm.nih.gov/25220054/)
55. Roach PJ. Multisite and hierarchal protein phosphorylation. *J Biol Chem.* 1991; 266:14139–14141. PMID: [1650349](https://pubmed.ncbi.nlm.nih.gov/1650349/)
56. Thomson M, Gunawardena J. Unlimited multistability in multisite phosphorylation systems. *Nature.* 2009; 460:274–277. doi: [10.1038/nature08102](https://doi.org/10.1038/nature08102) PMID: [19536158](https://pubmed.ncbi.nlm.nih.gov/19536158/)
57. Brewbaker C. A combinatorial interpretation of the poly-Bernoulli numbers and two Fermat analogues. *Integers.* 2008; 8:A02. Available from: <http://www.eudml.org/doc/130297>.

58. Kaneko M. Poly-Bernoulli numbers. *Journal de théorie des nombres de Bordeaux*. 1997; 9:221–228.
59. Martin GE. *Counting: The Art of Enumerative Combinatorics*. Springer; 2001.
60. Tanay A, Regev A, Shamir R. Conservation and evolvability in regulatory networks: the evolution of ribosomal regulation in yeast. *Proc Natl Acad Sci U S A*. 2005; 102:7203–7206. doi: [10.1073/pnas.0502521102](https://doi.org/10.1073/pnas.0502521102) PMID: [15883364](https://pubmed.ncbi.nlm.nih.gov/15883364/)
61. Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell*. 2010; 38:603–613. doi: [10.1016/j.molcel.2010.03.016](https://doi.org/10.1016/j.molcel.2010.03.016) PMID: [20513434](https://pubmed.ncbi.nlm.nih.gov/20513434/)
62. Zullo JM, Demarco IA, Piqué-Regi R, Gaffney DJ, Epstein CB, Spooner CJ, et al. DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. *Cell*. 2012; 149:1474–1487. doi: [10.1016/j.cell.2012.04.035](https://doi.org/10.1016/j.cell.2012.04.035) PMID: [22726435](https://pubmed.ncbi.nlm.nih.gov/22726435/)
63. Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, et al. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell*. 2010; 143:212–224. doi: [10.1016/j.cell.2010.09.009](https://doi.org/10.1016/j.cell.2010.09.009) PMID: [20888037](https://pubmed.ncbi.nlm.nih.gov/20888037/)
64. Edwards R, Glass L. Combinatorial explosion in model gene networks. *Chaos*. 2000; 10:691–704. doi: [10.1063/1.1286997](https://doi.org/10.1063/1.1286997) PMID: [12779419](https://pubmed.ncbi.nlm.nih.gov/12779419/)
65. Green JA, Rees D. On semi-groups in which $x^r = x$. *Mathematical Proceedings of the Cambridge Philosophical Society*. 1952; 48:35–40. doi: [10.1017/S0305004100027341](https://doi.org/10.1017/S0305004100027341)
66. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126:663–676. doi: [10.1016/j.cell.2006.07.024](https://doi.org/10.1016/j.cell.2006.07.024) PMID: [16904174](https://pubmed.ncbi.nlm.nih.gov/16904174/)
67. Murray RM, Sastry SS. Nonholonomic motion planning: steering using sinusoids. *IEEE Transactions on Automatic Control*. 1993; 38:700–716. doi: [10.1109/9.277235](https://doi.org/10.1109/9.277235)
68. Sternberg S. 1. In: *Lie Algebras*. Book online; 2004. p. 7–32. Available from: <http://www.math.harvard.edu/~shlomo>.
69. Carravetta M, Levitt MH. Long-lived nuclear spin states in high-field solution NMR. *J Am Chem Soc*. 2004; 126:6228–6229. doi: [10.1021/ja0490931](https://doi.org/10.1021/ja0490931) PMID: [15149209](https://pubmed.ncbi.nlm.nih.gov/15149209/)
70. Attanasio C, Nord AS, Zhu Y, Blow MJ, Biddie SC, Mendenhall EM, et al. Tissue-specific SMARCA4 binding at active and repressed regulatory elements during embryogenesis. *Genome Res*. 2014; 24:920–929. doi: [10.1101/gr.168930.113](https://doi.org/10.1101/gr.168930.113) PMID: [24752179](https://pubmed.ncbi.nlm.nih.gov/24752179/)
71. Kudron M, Niu W, Lu Z, Wang G, Gerstein M, Snyder M, et al. Tissue-specific direct targets of *Caenorhabditis elegans* Rb/E2F dictate distinct somatic and germline programs. *Genome Biol*. 2013; 14:R5. doi: [10.1186/gb-2013-14-1-r5](https://doi.org/10.1186/gb-2013-14-1-r5) PMID: [23347407](https://pubmed.ncbi.nlm.nih.gov/23347407/)
72. Noordermeer D, de Wit E, Klous P, van de Werken H, Simonis M, Lopez-Jones M, et al. Variegated gene expression caused by cell-specific long-range DNA interactions. *Nat Cell Biol*. 2011; 13:944–951. doi: [10.1038/ncb2278](https://doi.org/10.1038/ncb2278) PMID: [21706023](https://pubmed.ncbi.nlm.nih.gov/21706023/)
73. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*. 2013; 502:59–64. doi: [10.1038/nature12593](https://doi.org/10.1038/nature12593) PMID: [24067610](https://pubmed.ncbi.nlm.nih.gov/24067610/)
74. Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet*. 2010; 42:53–61. doi: [10.1038/ng.496](https://doi.org/10.1038/ng.496) PMID: [20010836](https://pubmed.ncbi.nlm.nih.gov/20010836/)
75. Marsman J, Horsfield JA. Long distance relationships: Enhancer-promoter communication and dynamic gene transcription. *Biochim Biophys Acta*. 2012; 18–19:1217–1227. doi: [10.1016/j.bbaggm.2012.10.008](https://doi.org/10.1016/j.bbaggm.2012.10.008)
76. Handoko L, Xu H, Li G, Ngan CY, Chew E, Schnapp M, et al. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet*. 2011; 43:630–638. doi: [10.1038/ng.857](https://doi.org/10.1038/ng.857) PMID: [21685913](https://pubmed.ncbi.nlm.nih.gov/21685913/)
77. Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, et al. CTCF mediates long-range chromatin looping and local histone modification in the β -globin locus. *Genes Dev*. 2006; 20:2349–2354. doi: [10.1101/gad.399506](https://doi.org/10.1101/gad.399506) PMID: [16951251](https://pubmed.ncbi.nlm.nih.gov/16951251/)
78. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*. 1999; 98:387–396. doi: [10.1016/S0092-8674\(00\)81967-4](https://doi.org/10.1016/S0092-8674(00)81967-4) PMID: [10458613](https://pubmed.ncbi.nlm.nih.gov/10458613/)
79. Hou C, Zhao H, Tanimoto K, Dean A. CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc Natl Acad Sci U S A*. 2008; 105:20398–20403. doi: [10.1073/pnas.0808506106](https://doi.org/10.1073/pnas.0808506106) PMID: [19074263](https://pubmed.ncbi.nlm.nih.gov/19074263/)

80. Guertin MJ, Lis JT. Mechanisms by which transcription factors gain access to target sequence elements in chromatin. *Curr Opin Genet Dev.* 2013; 23:116–123. doi: [10.1016/j.gde.2012.11.008](https://doi.org/10.1016/j.gde.2012.11.008) PMID: [23266217](https://pubmed.ncbi.nlm.nih.gov/23266217/)
81. Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* 2011; 21:1273–1283. doi: [10.1101/gr.122382.111](https://doi.org/10.1101/gr.122382.111) PMID: [21632746](https://pubmed.ncbi.nlm.nih.gov/21632746/)
82. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature.* 2007; 448:553–560. doi: [10.1038/nature06008](https://doi.org/10.1038/nature06008) PMID: [17603471](https://pubmed.ncbi.nlm.nih.gov/17603471/)
83. Mossman D, Kim KT, Scott RJ. Demethylation by 5-aza-2'-deoxycytidine in colorectal cancer cells targets genomic DNA whilst promoter CpG island methylation persists. *BMC Cancer.* 2010; 10:366. doi: [10.1186/1471-2407-10-366](https://doi.org/10.1186/1471-2407-10-366) PMID: [20618997](https://pubmed.ncbi.nlm.nih.gov/20618997/)
84. Yoshida M, Kijima M, Akita M, Beppu T. Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *J Biol Chem.* 1990; 265:17174–17179. PMID: [2211619](https://pubmed.ncbi.nlm.nih.gov/2211619/)
85. Dannenberg LO, Edenberg HJ. Epigenetics of gene expression in human hepatoma cells: expression profiling the response to inhibition of DNA methylation and histone deacetylation. *BMC Genomics.* 2006; 7:181. doi: [10.1186/1471-2164-7-181](https://doi.org/10.1186/1471-2164-7-181)
86. Qi Y, Fan P, Hao Y, Han B, Fang Y, Feng M, et al. Phosphoproteomic analysis of protein phosphorylation networks in the hypopharyngeal gland of honeybee workers (*Apis mellifera ligustica*). *J Proteome Res.* 2015; 14:4647–4661. doi: [10.1021/acs.jproteome.5b00530](https://doi.org/10.1021/acs.jproteome.5b00530) PMID: [26384081](https://pubmed.ncbi.nlm.nih.gov/26384081/)
87. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, et al. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell.* 2010; 143:1174–1189. doi: [10.1016/j.cell.2010.12.001](https://doi.org/10.1016/j.cell.2010.12.001) PMID: [21183079](https://pubmed.ncbi.nlm.nih.gov/21183079/)
88. Schmidt B, Marrone DF, Markus EJ. Genomic antagonism between retinoic acid and estrogen signaling in breast cancer. *Behav Brain Res.* 2012; 226:56–65.
89. Brown TI, Ross RS, Keller JB, Hasselmo ME, Stern CE. Which way was I going? Contextual retrieval supports the disambiguation of well learned overlapping navigational routes. *J Neurosci.* 2010; 30:7414–7422. doi: [10.1523/JNEUROSCI.6021-09.2010](https://doi.org/10.1523/JNEUROSCI.6021-09.2010) PMID: [20505108](https://pubmed.ncbi.nlm.nih.gov/20505108/)
90. Leutgeb JK, Leutgeb S, Moser MB, Moser MI. Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science.* 2007; 315:961–966. doi: [10.1126/science.1135801](https://doi.org/10.1126/science.1135801) PMID: [17303747](https://pubmed.ncbi.nlm.nih.gov/17303747/)
91. Chadwick MJ, Maguire DHMA. Decoding overlapping memories in the medial temporal lobes using high-resolution fMRI. *Learn Mem.* 2011; 18:742–746. doi: [10.1101/lm.023671.111](https://doi.org/10.1101/lm.023671.111) PMID: [22086391](https://pubmed.ncbi.nlm.nih.gov/22086391/)
92. Waddington CH. The cybernetics of development. In: *The Strategy of the Genes.* George Allen & Unwin Ltd.; 1957. p. 11–58.
93. Huang S. Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. *Philos Trans R Soc Lond B Biol Sci.* 2011; 366:2247–2259. doi: [10.1098/rstb.2011.0008](https://doi.org/10.1098/rstb.2011.0008) PMID: [21727130](https://pubmed.ncbi.nlm.nih.gov/21727130/)
94. St Johnston D, Nüsslein-Volhard C. The origin of pattern and polarity in the *Drosophila* embryo. *Cell.* 1992; 68:201–219. doi: [10.1016/0092-8674\(92\)90466-P](https://doi.org/10.1016/0092-8674(92)90466-P) PMID: [1733499](https://pubmed.ncbi.nlm.nih.gov/1733499/)
95. Akashi K, He X, Chen J, Iwasaki H, Niu C, Steenhard B, et al. Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis. *Blood.* 2003; 101:383–389. doi: [10.1182/blood-2002-06-1780](https://doi.org/10.1182/blood-2002-06-1780) PMID: [12393558](https://pubmed.ncbi.nlm.nih.gov/12393558/)
96. Biddie SC, John S, Sabo PJ, Thurman RE, Johnson TA, Schiltz RL, et al. Transcription Factor AP1 Potentiates Chromatin Accessibility and Glucocorticoid Receptor Binding. *Mol Cell.* 2011; 43:145–155. doi: [10.1016/j.molcel.2011.06.016](https://doi.org/10.1016/j.molcel.2011.06.016) PMID: [21726817](https://pubmed.ncbi.nlm.nih.gov/21726817/)
97. Voss TC, Schiltz RL, Sung MH, Yen PM, Stamatoyannopoulos JA, Biddie SC, et al. Dynamic Exchange at Regulatory Elements during Chromatin Remodeling Underlies Assisted Loading Mechanism. *Cell.* 2011; 146:544–554. doi: [10.1016/j.cell.2011.07.006](https://doi.org/10.1016/j.cell.2011.07.006) PMID: [21835447](https://pubmed.ncbi.nlm.nih.gov/21835447/)
98. Huang S, Guo YP, May G, Enver T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev Biol.* 2007; 305:695–713. doi: [10.1016/j.ydbio.2007.02.036](https://doi.org/10.1016/j.ydbio.2007.02.036) PMID: [17412320](https://pubmed.ncbi.nlm.nih.gov/17412320/)
99. Ferrell J. Bistability, bifurcations, and Waddington's epigenetic landscape. *Curr Biol.* 2012; 22:R458–456. doi: [10.1016/j.cub.2012.03.045](https://doi.org/10.1016/j.cub.2012.03.045) PMID: [22677291](https://pubmed.ncbi.nlm.nih.gov/22677291/)
100. Romero PA, Arnold FH. Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Biol.* 2008; 10:866–876. doi: [10.1038/nrm2805](https://doi.org/10.1038/nrm2805)
101. Fritz G, Buchler NE, Hwa T, Gerland U. Designing sequential transcription logic: a simple genetic circuit for conditional memory. *Syst Synth Biol.* 2007; 1:89–98. doi: [10.1007/s11693-007-9006-8](https://doi.org/10.1007/s11693-007-9006-8) PMID: [19003438](https://pubmed.ncbi.nlm.nih.gov/19003438/)

102. Margulies D, Felder CE, Melman G, Shanzer A. A molecular keypad lock: a photochemical device capable of authorizing password entries. *J Am Chem Soc.* 2007; 129:347–354. doi: [10.1021/ja065317z](https://doi.org/10.1021/ja065317z) PMID: [17212414](https://pubmed.ncbi.nlm.nih.gov/17212414/)
103. Lou C, Liu X, Ni M, Huang Y, Huang Q, Huang L, et al. Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. *Mol Syst Biol.* 2010; 6:350. doi: [10.1038/msb.2010.2](https://doi.org/10.1038/msb.2010.2) PMID: [20212522](https://pubmed.ncbi.nlm.nih.gov/20212522/)
104. Voight CA. Genetic parts to program bacteria. *Curr Opin Biotechnol.* 2006; 23:548–557. doi: [10.1016/j.copbio.2006.09.001](https://doi.org/10.1016/j.copbio.2006.09.001)
105. Rao CV. Expanding the synthetic biology toolbox: engineering orthogonal regulators of gene expression. *Curr Opin Biotechnol.* 2012; 23:689–694. doi: [10.1016/j.copbio.2011.12.015](https://doi.org/10.1016/j.copbio.2011.12.015) PMID: [22237017](https://pubmed.ncbi.nlm.nih.gov/22237017/)
106. Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L. Single-cell *in situ* RNA profiling by sequential hybridization. *Nat Methods.* 2014; 11:360–361. doi: [10.1038/nmeth.2892](https://doi.org/10.1038/nmeth.2892) PMID: [24681720](https://pubmed.ncbi.nlm.nih.gov/24681720/)
107. Groh C, Moldovanu B, Sela A, Sunde U. Optimal seedings in elimination tournaments. *Economic Theory.* 2008; 49:59–80. doi: [10.1007/s00199-008-0356-6](https://doi.org/10.1007/s00199-008-0356-6)
108. Fudenberg D, Maskin E. Evolution and cooperation in noisy repeated games. In: *The American Economic Review.* vol. 80. American Economic Association; 1990. p. 274–279. Available from: <http://www.jstor.org/stable/2006583>.
109. Kim HK, Krotov DS, Lee JY. Poly-Bernoulli numbers and lonesum matrices. *arXiv.* 2011;1103.4884.