

Noise Reduction in Complex Biological Switches

Luca Cardelli, Microsoft Research & Oxford University

related work: Attila Csikász-Nagy, Neil Dalchau, Mirco Tribastone, Max Tschaikowski, Andrea Vandin

NETADIS, London 2015-10-23

Introduction

Noise vs. Complexity

- Cells operate in noisy molecular environments
 - Via complex regulatory networks produced by evolution
 - For each network, we can analyze the noise
 - But how does noise related to (growing) network complexity?
- For a fixed function, does complexity reduce noise?
 - Beyond the mere increase of overall molecular counts?
 - Complexity could provide and advantage counteracting its costs

Noise in Multistable Systems

- A little noise can lead to different outcomes
 - We investigate biochemical switches – bistable systems
- In previous work
 - The (classical) cell cycle switch implements an optimal(-speed) switching algorithm
 - More recent and more complex models do the same
 - All that with deterministic (ODE) semantics
- On that basis
 - We can compare networks of different complexity "fairly"
 - And investigate how they differ in terms of noise characteristics

Comparing Networks

- For chosen initial conditions
 - Certain networks of different complexity have identical output (trajectories)
 - Hence they have compatible function
- Why would then evolution choose complexity?
 - Likely many different reasons and tradeoffs
 - We investigate reduction in noise levels
 - Trying to separate it from other effects

Methods

- Bounding the problem by different techniques
 - Chemical Master Equation
 - Slow and accurate at low molecular counts, unfeasible at high counts
 - Linear Noise Approximation
 - Fast and accurate “in the thermodynamic limit”, inaccurate at low counts
 - The biological regime falls in the middle of the two
 - Computationally (and analytically) inaccessible, but bounded by consistent results
- We observe that
 - For equivalent (deterministic) function, more complex networks “tend to” exhibit a reduction in intrinsic noise. Both size and structure matter
 - Not simply attributable to the larger molecular counts of the larger networks

To carry this out we need

- A notion of “function”
 - Many different networks of different size that all “do the same thing”
- A baseline for comparison
 - Deterministic traces
- Ways of investigating noise
 - Numerical simulations of exact or approximate kinetics
 - The fundamentally non-linear aspect of chemical kinetics prevent analytical methods for most examples of interest

Biochemical Networks

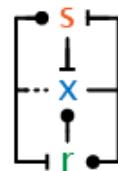
Network model

- Influence networks

- Influence species: two main molecular states (high/low or modified/unmodified)
- High-low transitions are nonlinear (e.g. sigmoidal)
- Exact transition kinetics varies (but we fix one uniformly)



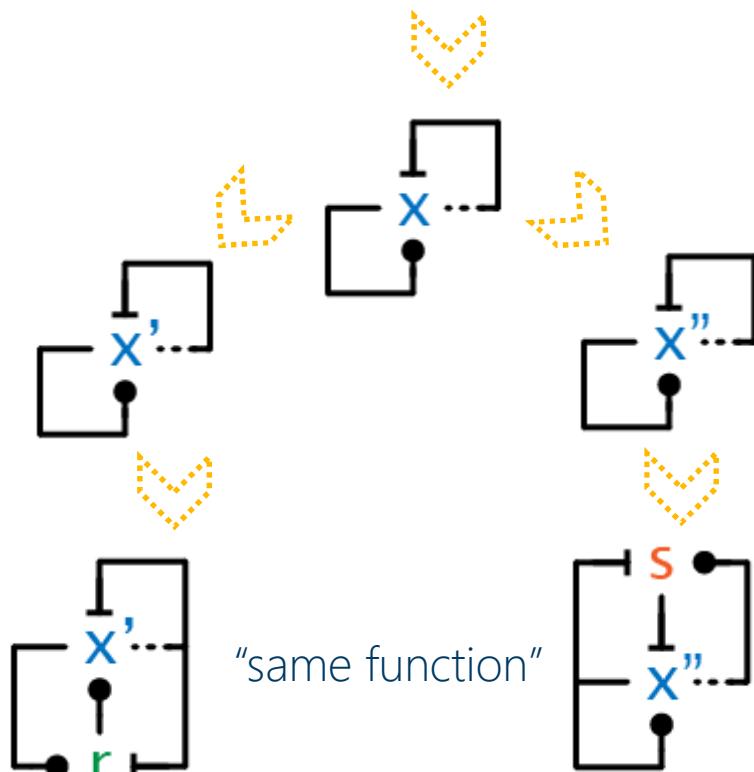
Nodes



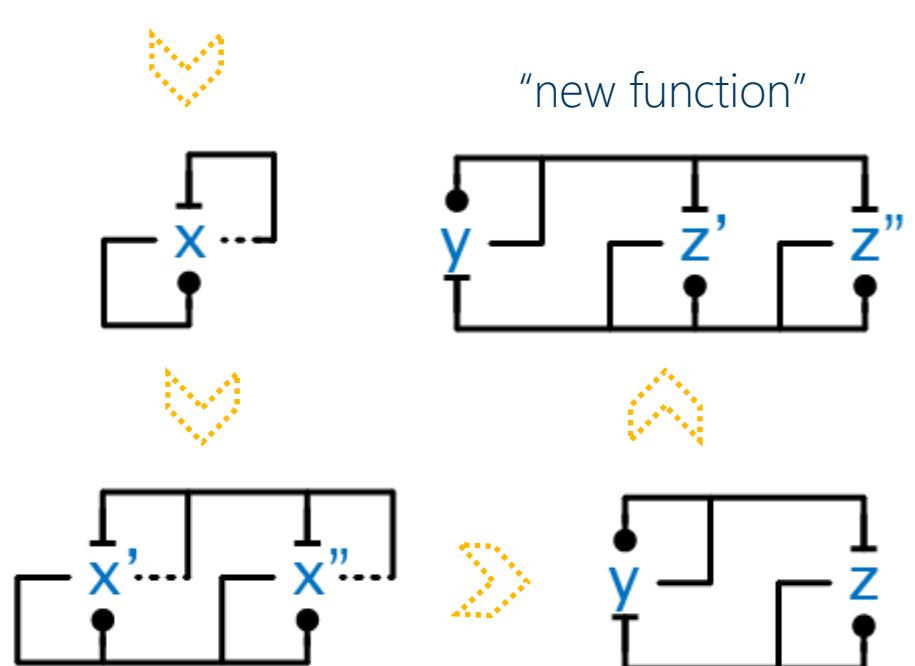
Ex.: a cell cycle switch model

Network Evolution

Across species: *Ortholog genes*



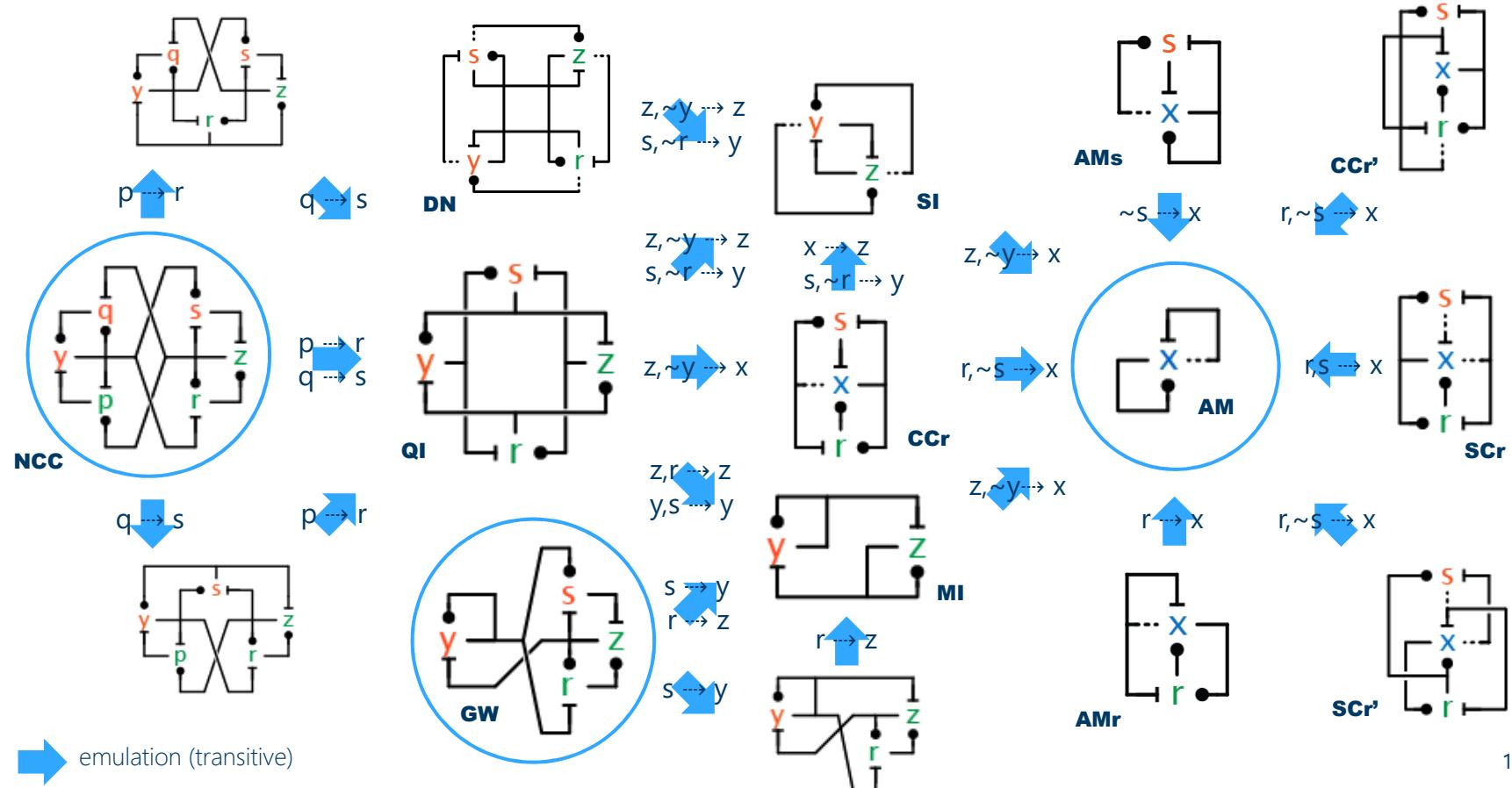
Within species: *Paralog genes*



Comparing Networks

- High-value activity:
 - 2001 Nobel prize in Physiology for the discovery of *"Key regulators of the cell cycle ... they have identified key molecules that regulate the cell cycle in all eukaryotic organisms, including yeast, plants, animals, and human."*
 - These are *not* the same molecules in all organisms, but it is still "the same network"
- Network differences expose evolution
 - Tracing back ancestral networks from current ones
- Networks are algorithms
 - Algorithms fall in different performance classes (is nature "optimal"?)
 - Different networks for the same function may or may not be in the same class
- How do we compare networks?

Network Emulation



How to model “Influence”

“True” molecular interactions.

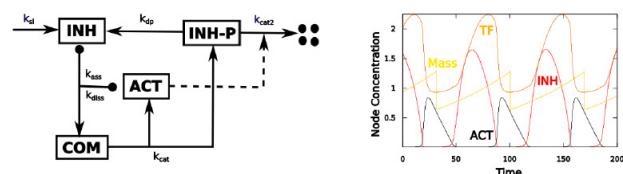


Figure 3: a) Schematic diagram of a simplified SIMM model [17]. The activation

“Equivalent” influence interactions.

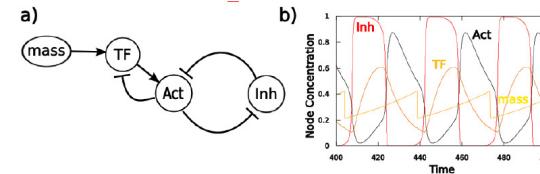


Figure 4: a) Schematic diagram of a primitive cell cycle in the reinitz framework.

Chemical Reaction Network



Influence Network

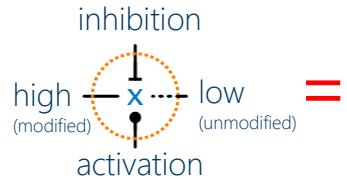
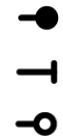
Evolving a Primitive Eukaryotic Cell Cycle Model

Malte Lücken, Jotun Hein, Bela Novak

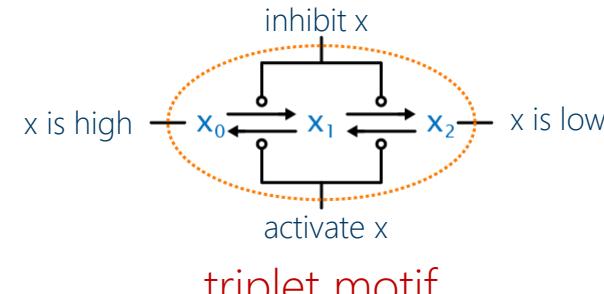
Instead of modeling basic interactions, such as binding, synthesis, and degradation of molecular components, this framework models interactions simply as activation or inhibition. This approach also reduces the number of nodes necessary in the network, as e.g. the inhibitor binding tightly to the activator to form a complex, which produces phosphorylated inhibitor to be degraded under catalysis by the activator, is now simply a double negative feedback loop shown in Figure 1. This type of interaction is the basis of both aforementioned molecular model, therefore they can both be summarized in a single Reinitz model.

The Triplet Model of Influence

activation
inhibition
catalysis



Usually modeled by sigmoid (e.g. Hill or Reinitz) functions



triplet motif

We model them by
4 mass action reactions over
3 species x_0, x_1, x_2

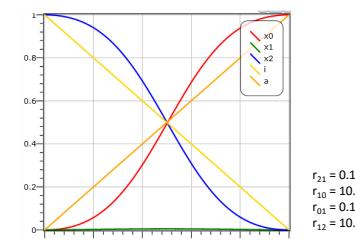
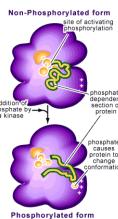
They actually implement a
Hill function of coefficient 2:

Functional Motifs in
Biochemical Reaction Networks
John J. Tyson¹ and Béla Novák²

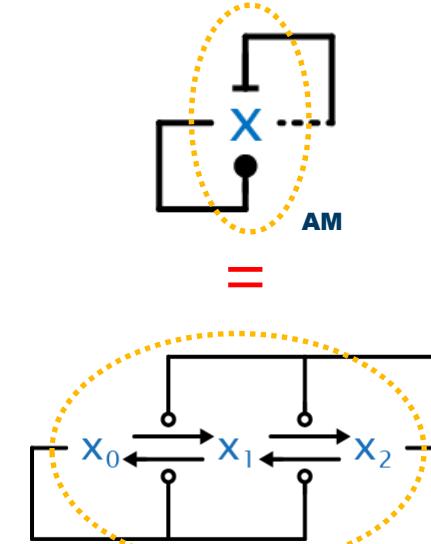
$$\frac{dX_i}{dt} = \gamma_i \frac{A_i(1 - X_i) - B_i X_i}{A_i + B_i}, \quad i = 1, \dots, N,$$

$$A_i = \exp \left\{ \sigma_i \left(\alpha_{i0} + \sum_{j=1}^N \alpha_{ij} X_j \right) \right\}, \quad B_i = \exp \left\{ \sigma_i \left(\beta_{i0} + \sum_{j=1}^N \beta_{ij} X_j \right) \right\}, \quad (4)$$

biological mechanism:
(e.g.:) multisite
phosphorylation



For example:



Approximate Majority

Consensus Networks

A Consensus Problem

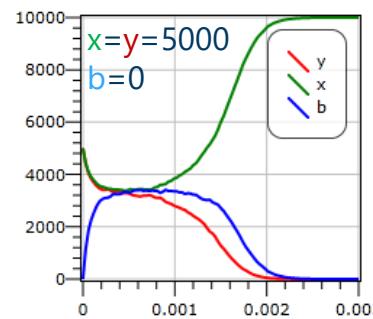
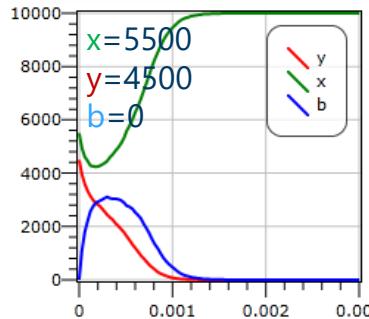
- Population Consensus
 - Given two populations of x and y "agents"
 - We want them to "reach consensus"
 - By converting *all* agents to x or to y depending on which population was in majority initially
- Population Protocols Model
 - Finite-state identity-free agents (**molecules**) interact in randomly chosen pairs (\Rightarrow stochastic symmetry breaking)
 - Each interaction (**collision**) can result in state changes
 - Complete connectivity, no centralized control (**well-mixed solution**)

specification

$$\begin{aligned} X, Y &:= X+Y, 0 \quad \text{if } X_0 \geq Y_0 \\ X, Y &:= 0, X+Y \quad \text{if } Y_0 \geq X_0 \end{aligned}$$

A Consensus Algorithm

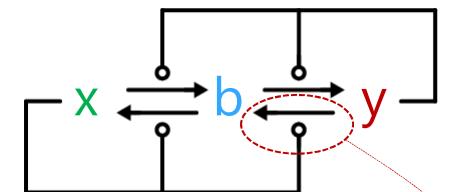
- Approximate Majority (AM) Algorithm
 - Uses a third “undecided” population b
 - Disagreements cause agents to become undecided
 - Undecided agents agree with any non-undecided agent



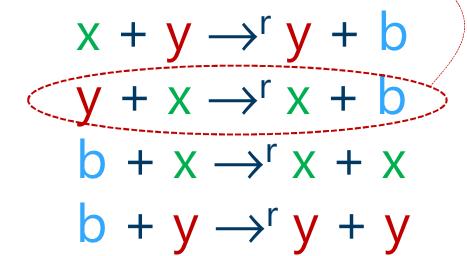
Dana Angluin · James Aspnes · David Eisenstat

A Simple Population Protocol for Fast Robust Approximate Majority

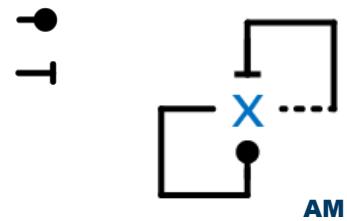
catalysis



chemical reaction network

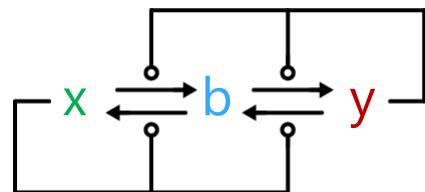


activation
inhibition



A Biological Implementation

Approximate Majority (AM)



- 1) **Bistable**
Even when initially $x=y$ (stochastically)
- 2) **Fast (asymptotically optimal)**
 $O(\log n)$ convergence time
- 3) **Robust to perturbation**
above a threshold, initial majority wins *whp*

Dana Angluin · James Aspnes · David Eisenstat

A Simple Population Protocol for Fast Robust Approximate Majority

2007

Epigenetic Switch

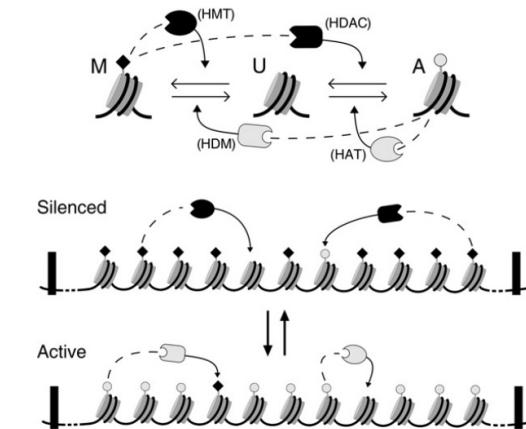


Figure 1. Basic Ingredients of the Model

Theory

Cell

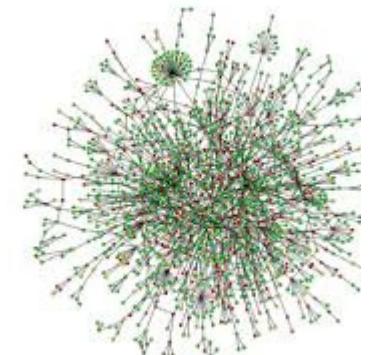
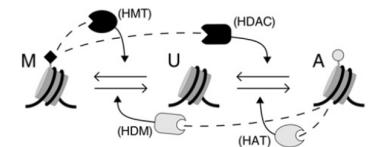
Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification

Ian B. Dodd,^{1,2} Miles A. McEvoy,³ Kim Sneppen,^{4,*} and Genevieve Thor⁵
¹Center for Models of Life, Niels Bohr Institute, Blegdamsvej 17, DK-2100, Copenhagen D, Denmark
²Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100, Copenhagen, Denmark
³Department of Molecular Biology, University of Copenhagen Biocenter, Øst Møllevæn Vej 5, DK-2200 Copenhagen N, Denmark
⁴Department of Physics, University of Copenhagen, Øst Møllevæn Vej 5, DK-2200 Copenhagen N, Denmark
⁵Department of Cell Biology, University of Copenhagen, Øst Møllevæn Vej 5, DK-2200 Copenhagen N, Denmark

2007

Not always that simple

- The epigenetic switch seems a *direct* biological implementation of an algorithm
 - Although we may have to qualify that with some notion of approximation of the (enzymatic) kinetics
- In most cases the biological implementation seems more *indirect* or *obfuscated*
 - “*Nature is subtle but not malicious - Einstein*” Ha! think again!
 - Other implementations of Approximate Majority seem more convoluted and approximate



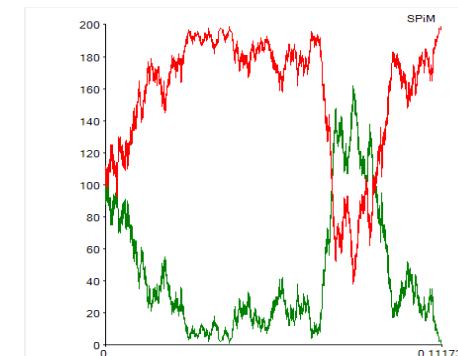
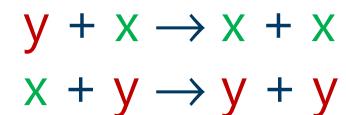
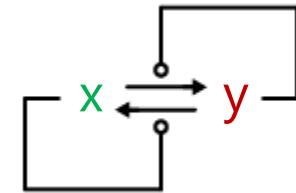
How to Build a Good Switch

- We need first a **bistable** system: one that has two *distinct* and *stable* states. I.e., given any initial state the system must settle into one of two states
- The settling must be **fast** (not get stuck in the middle for too long) and **robust** (must not spontaneously switch back)
- Finally, we need to be able to **flip** the switch by external inputs

A Bad Algorithm

- Direct Competition
 - x catalyzes the transformation of y into x
 - y catalyzes the transformation of x into y
 - when all- x or all- y , it stops
- This system has two end states, but
 - Convergence to an end state is slow (a random walk)
 - Any perturbation of an end state can start a random walk to the other end state (hence not really *bistable*)

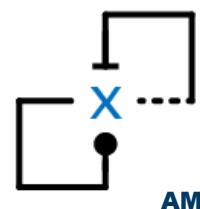
catalysis 



A Good Algorithm

- Approximate Majority (AM)
 - Third, undecided, state b
 - Disagreements cause agents to become undecided
 - Undecided agents believe any non-undecided agent
- With high probability, for n agents
 - The total number of interactions before converging is $O(n \log n)$
⇒ fast (optimal)
 - The final outcome is correct if the initial disparity is $\omega(\sqrt{n} \log n)$
⇒ solution states are robust to perturbations
- Logarithmic time bound in parallel time
 - *Parallel time* is the number of steps divided by the number of agents
 - In parallel time the algorithm converges with high probability in $O(\log n)$

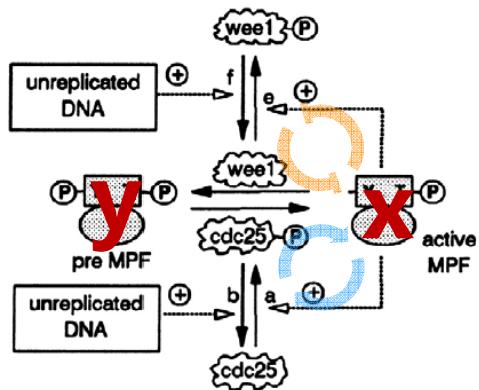
activation
inhibition



Dana Angluin · James Aspnes · David Eisenstat

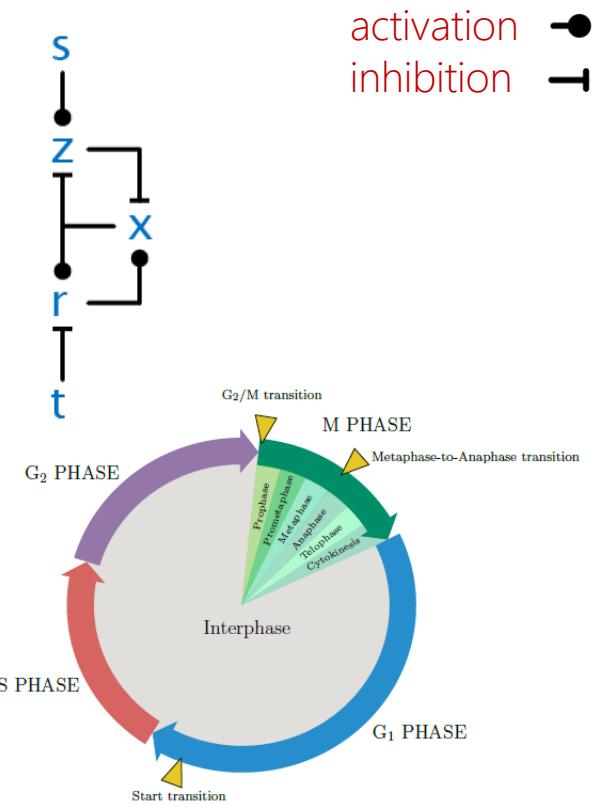
A Simple Population Protocol for Fast Robust
Approximate Majority

An “Ugly” Algorithm: Cell Cycle Switch



Nobel-prize
winning network

Obfuscation of a
distributed
algorithm?



- Is it a good algorithm? Is it bad?
- Is it optimal or suboptimal?

Journal of Cell Science 106, 1153-1168 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

**Numerical analysis of a comprehensive model of M-phase control in
Xenopus oocyte extracts and intact embryos**

Bela Novak* and John J. Tyson†

Department of Biology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24060-0406, USA

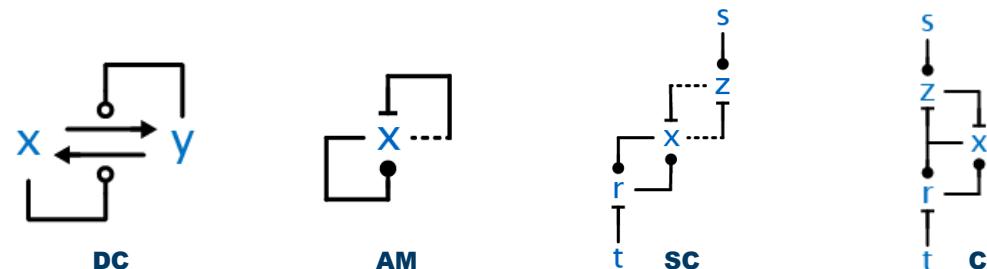
*Permanent address: Department of Agricultural Chemical Technology, Technical University of Budapest, 1521 Budapest Gellert Ter 4, Hungary

†Author for correspondence

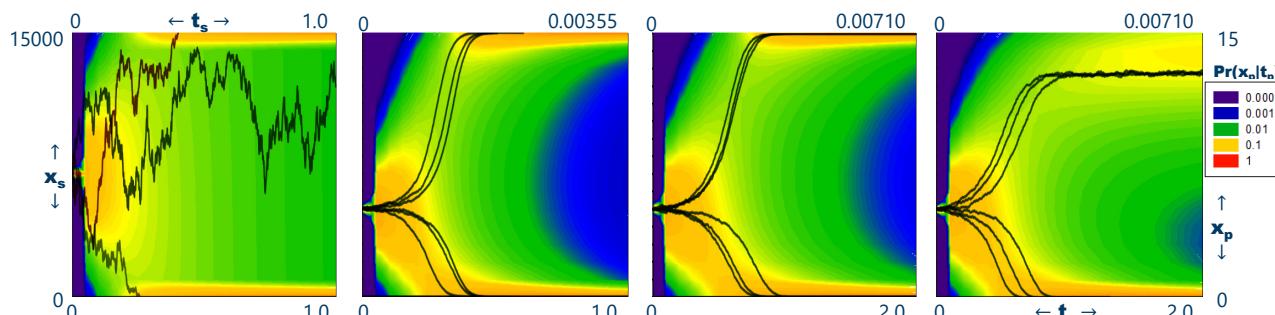
Convergence Analysis - CONSENSUS

- Switches as computational systems

CC converges in $O(\log n)$ time (like AM)
(but 2x slower than AM, and does not fully switch)



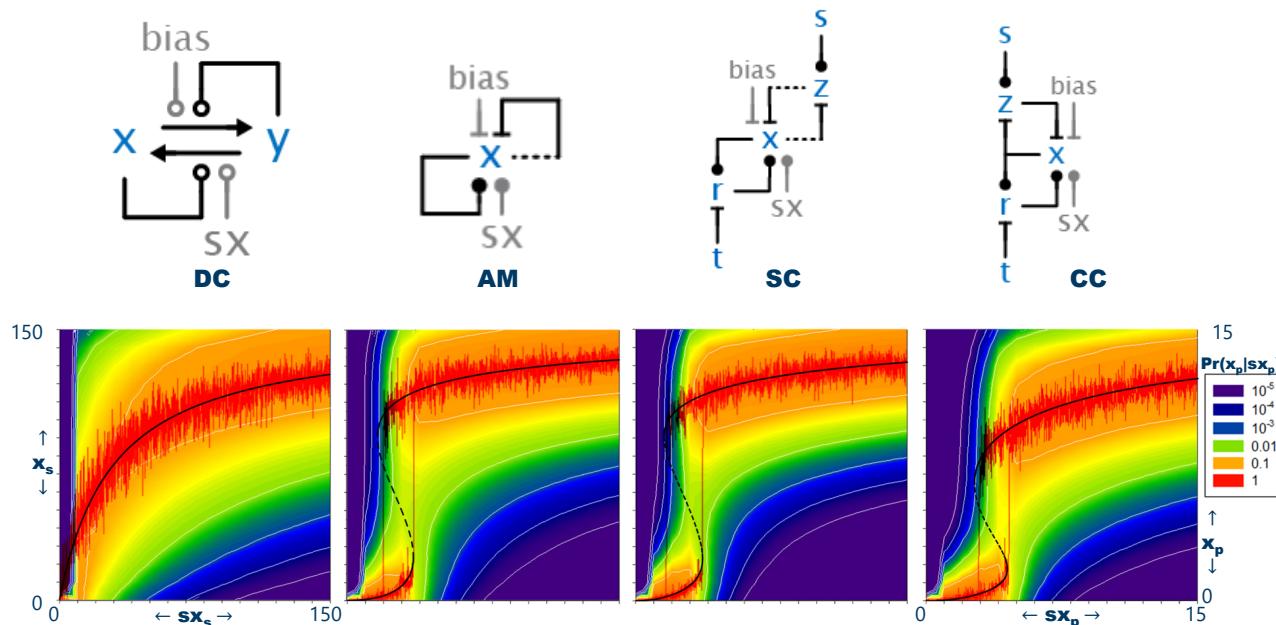
Start symmetrical
($x_0 = x_1 = x_2$ etc.)



Black lines: several stochastic simulation traces
Color: full probability distribution of small-size system

Steady State Analysis - SWITCH

- Switches as dynamical systems



Black lines: deterministic ODE bifurcation diagrams

Red lines: noisy stochastic simulations

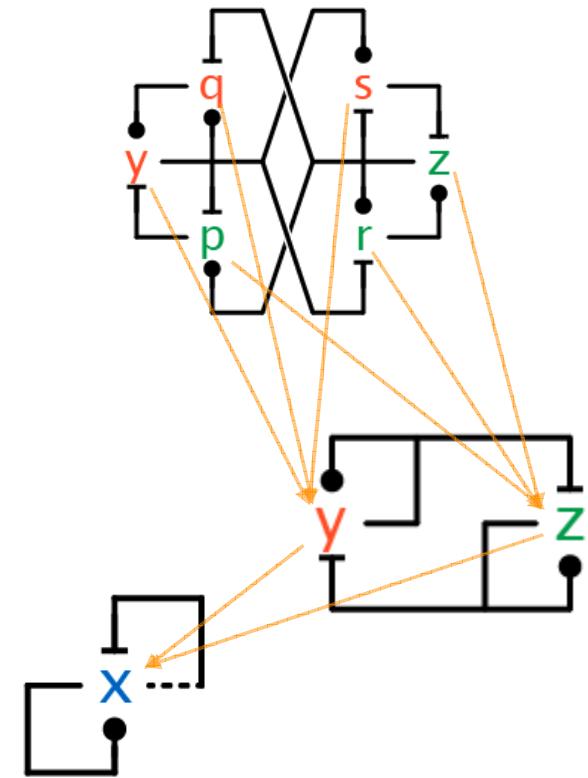
Color: full probability distribution of small-size system

Network Morphisms

When does a (complex) network
implement a (simpler) algorithm?

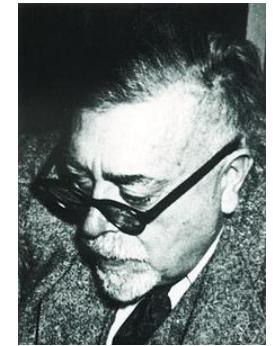
Comparing networks

- How can we compare different networks?
 - Different number of species
 - Different number of reactions
 - Apparently unrelated connectivity
- So that we can compare their function?
 - Does antagonism (in network structure) guarantee bistability (in function)?
- We do it by *mapping* networks onto one another so that they *emulate* each other
 - Deterministic semantics version of “simulation” of systems
 - (Stochastic semantics was the starting point, but too difficult/demanding for typical biological networks.)



Mapping one network into another

- Notion is strangely missing from the literature
 - Seen in Biology: single-network analysis (e.g. structure of feedback loops) and network reduction (e.g. while preserving steady states). Study of common or frequent subnetworks.
 - Seen in C.S.: comparing network *behaviors* (e.g. morphisms of event structures).
 - Nothing much resembling (bi)simulation "on the syntax" (structure) of whole biochemical networks.
- Model reduction is unavoidable and pervasive, but
 - Often criticized/ignored by biologists when it leads to quantities that are "not biologically meaningful". E.g. a fusion or change a variables in the ODEs where the new variables do not correspond to biological parts. The reduced model should "inform" the original one.
- Science's ethos
 - The "truth" is the big network, not the small one!
If you depart from the truth in any way, you have to explain how you can get back to it.
 - The point is not to reduce the size of the network (although that's neat), but to understand aspects of *the big network* by reference to a smaller one.
 - The mapping is more important than either networks.

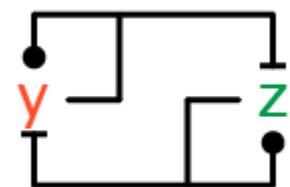


Norbert Wiener
Pioneer of stochastic processes
and inventor of Cybernetics.

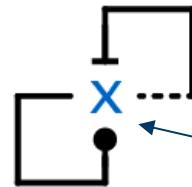
"The best material model of a cat is another, or preferably the same, cat"

Network Emulation MI emulates AM

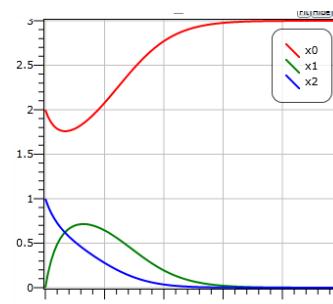
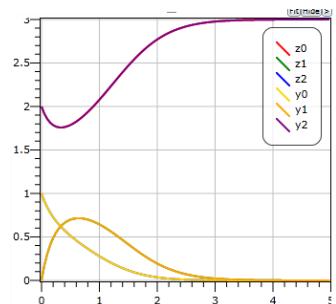
- For *any rates and initial conditions* of **AM**, we can find *some* rates and initial conditions of **MI** such that the (6) trajectories of **MI** retrace those (3) of **AM**:



$\sim y, z \rightarrow x$



(3 species)



initialize:

$$z = x$$

$$\sim y = x$$

$$(y_2 = x_0)$$

$$y_1 = x_1$$

$$y_0 = x_0$$

- How do we find these matching parameters? By a **network morphism!**

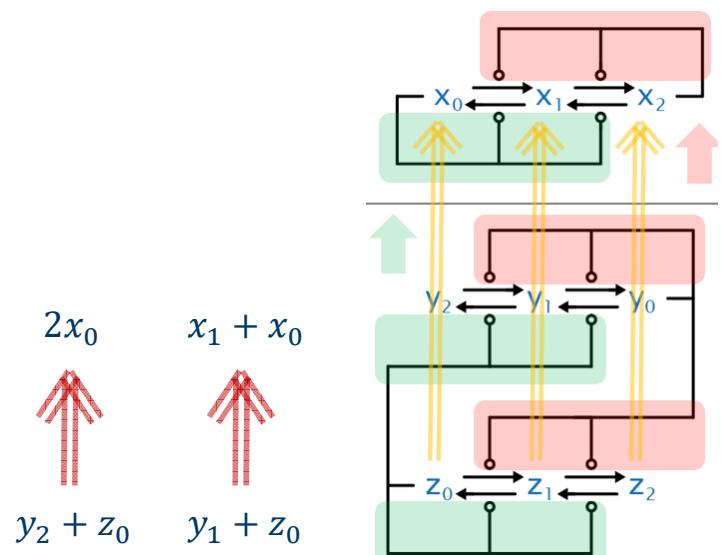
CRN Morphisms

A *CRN morphism* from (S, R) to (\hat{S}, \hat{R})
written $m \in (S, R) \rightarrow (\hat{S}, \hat{R})$

is a pair of maps $m = (m_S, m_R)$
a *species map* $m_S \in S \rightarrow \hat{S}$
a *reaction map* $m_R \in R \rightarrow \hat{R}$

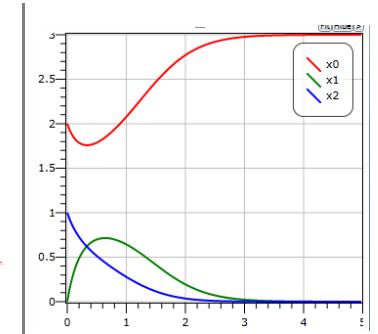
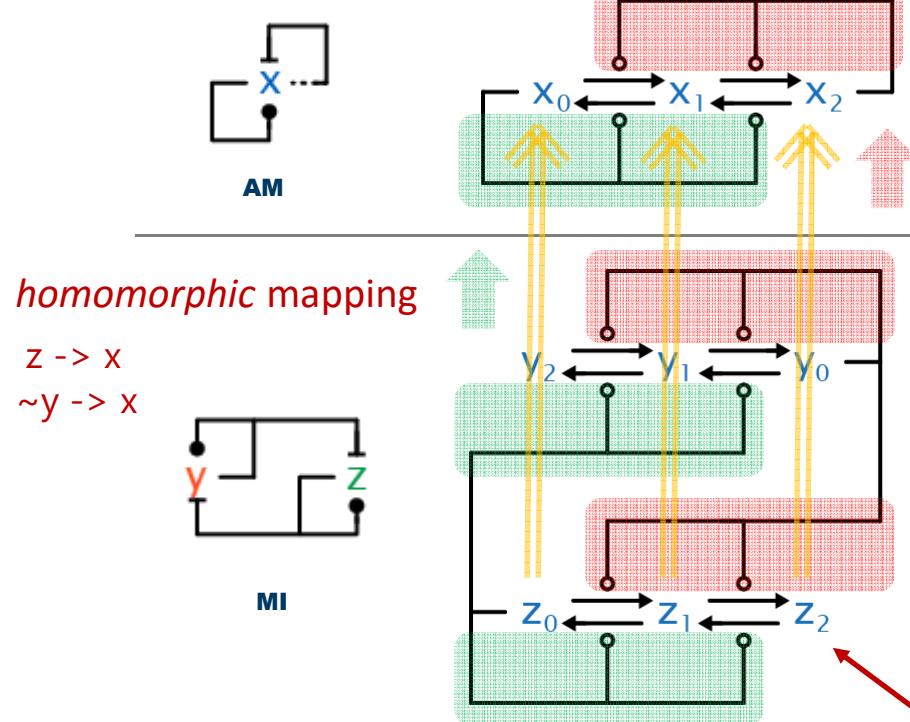
extended to a *complex map* $m_S \in \mathbb{N}^S \rightarrow \mathbb{N}^{\hat{S}}$
linearly: $m_S(\rho_{\hat{s}}) = \sum_{s \in m_S^{-1}(\hat{s})} \rho_s$

Mappings (symmetries)
between two networks

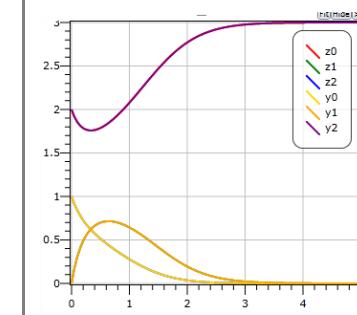


Network Emulation: MI emulates AM

A mapping of species and reactions



any initial conditions



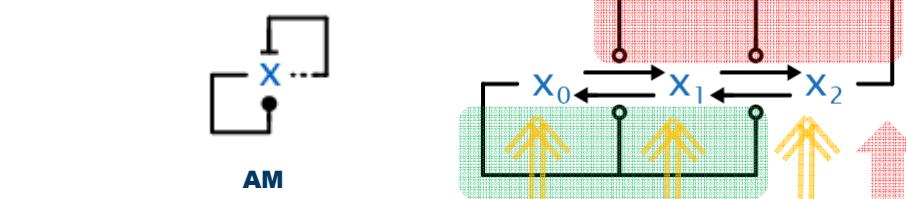
initial conditions:

$$\begin{aligned} z_0 &= y_2 = x_0 \\ z_1 &= y_1 = x_1 \\ z_2 &= y_0 = x_2 \end{aligned}$$

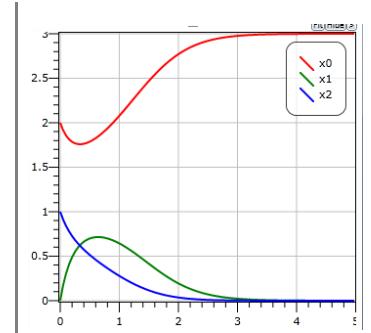
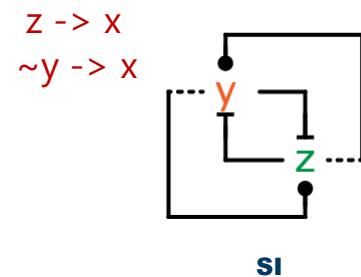
less trivial than you might think:
it need not preserve the out-degree of a node!

Network Emulation: SI emulates AM

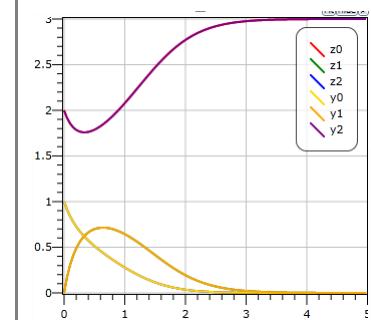
A mapping of species and reactions



homomorphic mapping



any initial conditions

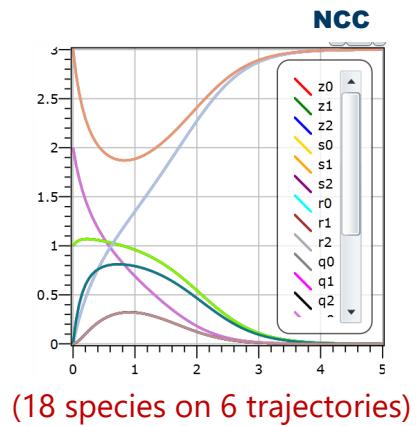
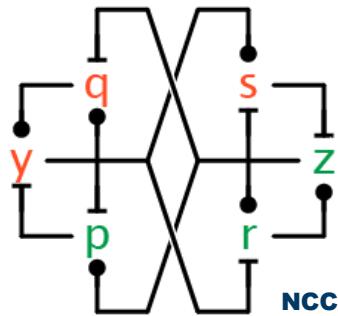


initial conditions:

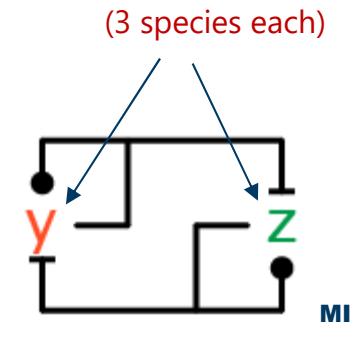
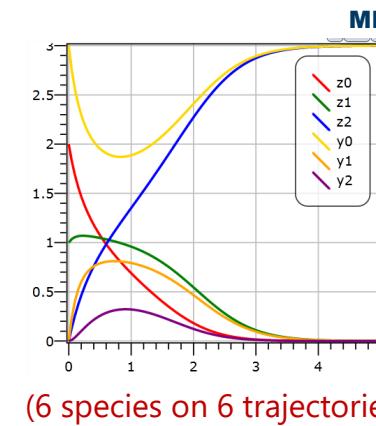
$$\begin{aligned} z_0 &= y_2 = x_0 \\ z_1 &= y_1 = x_1 \\ z_2 &= y_0 = x_2 \end{aligned}$$

Network Emulation: NCC emulates MI

- For *any* rates and initial conditions of MI we can find some rates and initial conditions of NCC such that the (18) trajectories of NCC retrace those (6) of MI



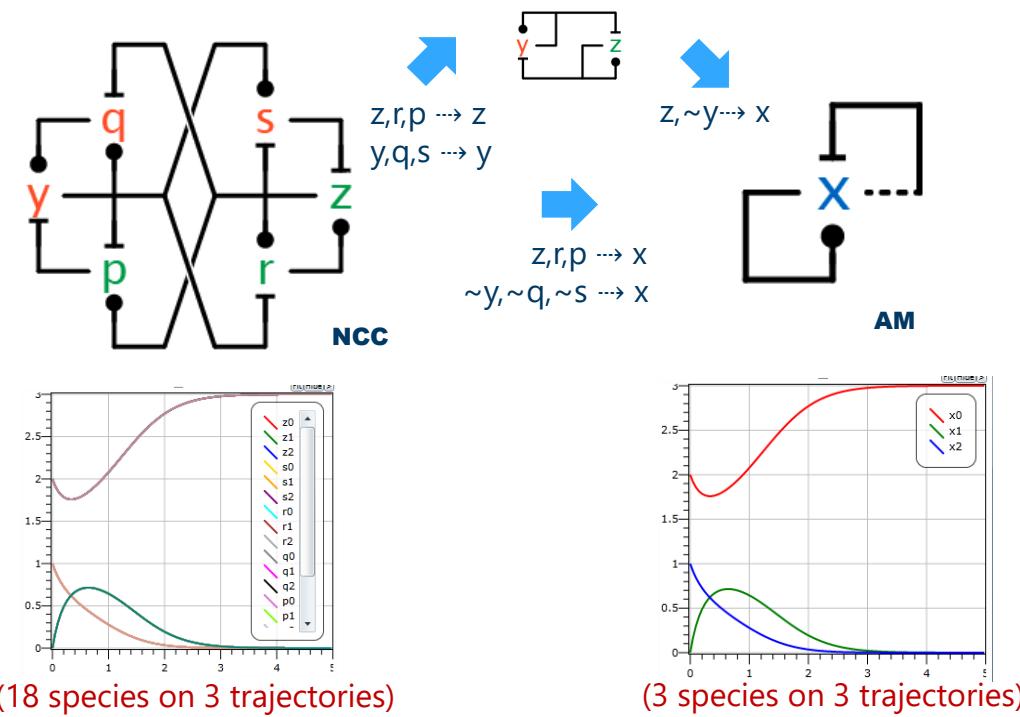
$z, r, p \rightarrow z$
 $y, q, s \rightarrow y$



initialize
 $z, r, p = z$
 $y, q, s = y$

Emulations Compose

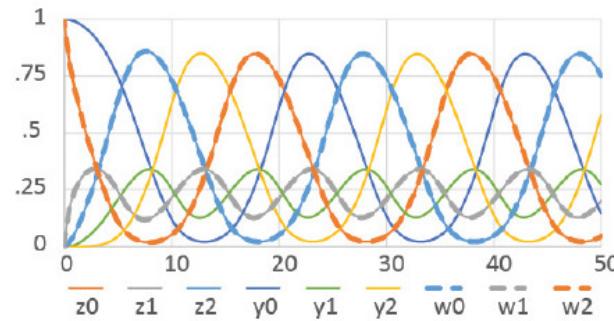
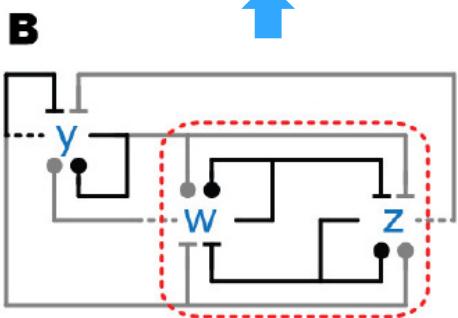
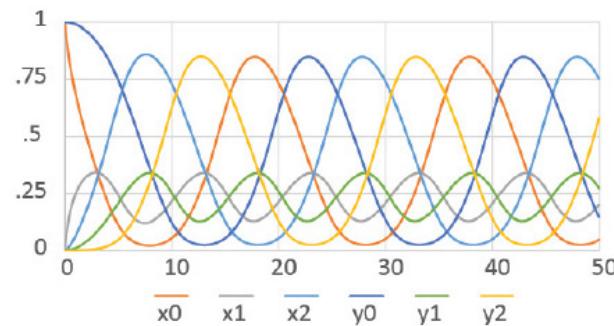
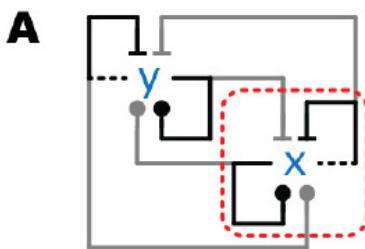
- The (18) trajectories NCC can *always* retrace those (3) of AM



The new cell cycle switch
can emulate AM exactly.
For *any* initial conditions
of AM.

And for *any* rates of AM.

Emulations are Modular



How to check for emulation

- How do we check a potential emulation morphism
for all possible initial conditions of the target?
 - Statically! Check conditions on the joint stoichiometric matrices of the two networks under the mapping.
- How do we check a potential emulation morphism
for all possible rates of the target?
 - Can't; but if one emulation is found, then the rates of the target network can be changed *arbitrarily* and a related emulation will again exist.

Static Criteria for Emulation

Emulation Theorem: If $m \in (S, R) \rightarrow (\hat{S}, \hat{R})$ is a CRN reactant morphism and stoichiomorphism then it is a CRN emulation

reactant morphism

$$\mathbf{m}_S^T \cdot \boldsymbol{\rho} = \hat{\boldsymbol{\rho}} \cdot \mathbf{m}_R^T$$

stoichiomorphism

$$\boldsymbol{\varphi} \cdot \mathbf{m}_R = \mathbf{m}_S \cdot \hat{\boldsymbol{\varphi}}$$



emulation

$$\forall \hat{\boldsymbol{v}}. \quad F(\hat{\boldsymbol{v}} \circ m_S) = \hat{F}(\hat{\boldsymbol{v}}) \circ m_S \quad \text{preserve derivatives}$$

preserve enough network structure
preserve enough chemical stoichiometry

F is the differential system of (S, R) , given by the law of mass action, $\hat{\boldsymbol{v}}$ is a state of (\hat{S}, \hat{R}) . $\boldsymbol{\varphi}$ is the stoichiometric matrix and $\boldsymbol{\rho}$ is the related reactant matrix. \mathbf{m}_S and \mathbf{m}_R are the characteristic 0-1 matrices of the morphism maps m_S (on species) and m_R (on reactions). $-^T$ is transpose.
Homomorphism implies reactant morphism.

Cardelli BMC Systems Biology 2014, 8:84
<http://www.biomedcentral.com/1752-0509/8/84>

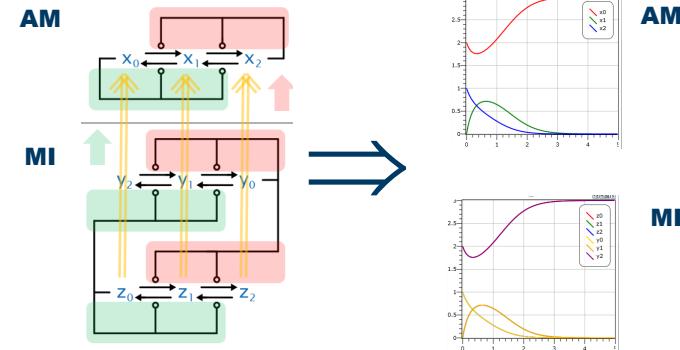


RESEARCH ARTICLE

Open Access

Morphisms of reaction networks that couple structure to function

Luca Cardelli^{1,2}



Stoichiomorphisms condition is sufficient for "networks of interest" and actually "close" to a necessary condition.

Applications of Emulation

- Model Reduction
 - Find reduced networks
 - Compute quotient CRNs
 - Find network symmetries that may be of biological interest
- Morphism Generation
 - Find morphisms between networks (e.g. all the ones for a fixed rate assignment)

Benchmarks from
Sneddon et al., Nature Methods, 2011

Model	Reactions	Species	FB	Time (s)	BB	Time (s)
e9	3538944	262146	222	4.61E+4	222	7.65E+4
e8	786432	65538	167	1.92E+3	167	3.68E+3
e7	172032	16386	122	8.15E+1	122	1.77E+2
e6	36864	4098	86	3.00E+0	86	7.29E+0
e5	7680	1026	58	1.54E-1	58	4.06E-1
e4	1536	258	37	9.00E-3	37	1.09E-1
e3	288	66	22	1.00E-3	22	3.00E-3
e2	48	18	12	1.00E-3	12	2.00E-3

Aggregation reduction Emulation reduction
₃₉

Forward and Backward Bisimulations for Chemical Reaction Networks

Luca Cardelli¹, Mirco Tribastone², Max Tschaikowski³, and Andrea Vandin⁴

¹ Microsoft Research & University of Oxford, UK
luca@ microsoft.com

²⁻⁴ University of Southampton, UK
(n.tribastone,m.tschaikowski,a.vandin)@ soton.ac.uk

Satisfiability Modulo Differential Equivalence Relations

Luca Cardelli
Microsoft Research and University of Oxford, UK
luca@microsoft.com

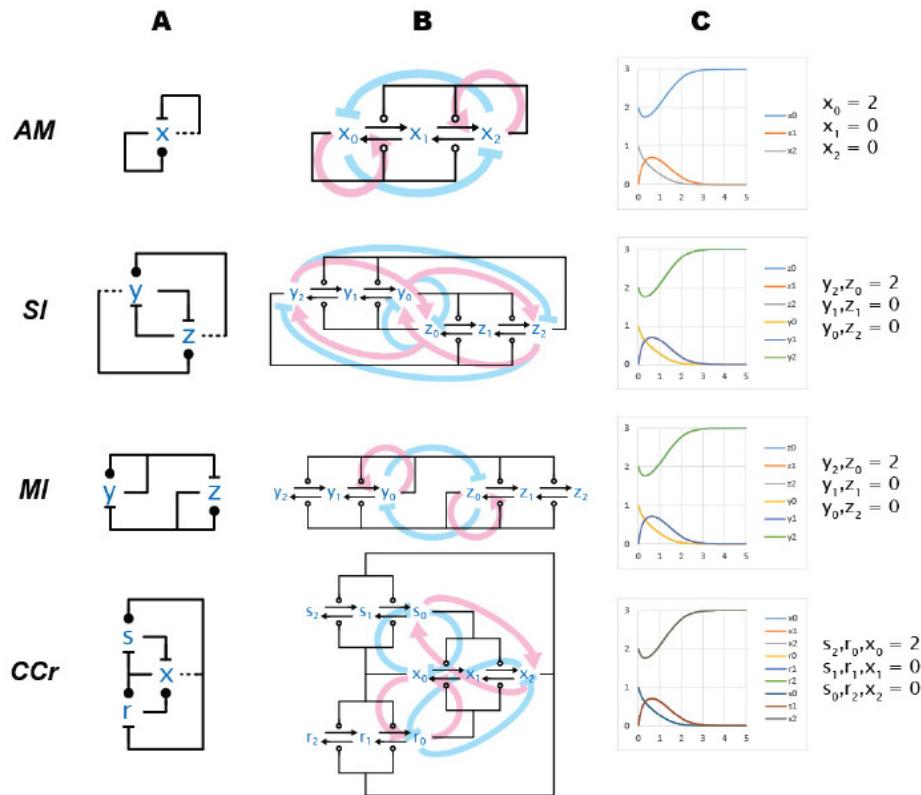
Mirco Tribastone Max Tschaikowski
Andrea Vandin
IMT - Institute for Advanced Studies Lucca, Italy
(name.surname)@imtlucca.it

Concur 2015

POPL 2016

Noise Reduction in Complex Switches

Basic Switches (deterministic)



(A) Influence network diagrams

(B) Chemical reaction network diagrams and feedback loops

(C) Numerical solutions of the deterministic kinetics of the networks:
 Horizontal axis is time
 Vertical axis is species concentration

First some arbitrary initial conditions are chosen for AM.

Then the initial conditions of the other networks are chosen in such a way that each trace of each of the other networks retraces exactly one trace of AM.

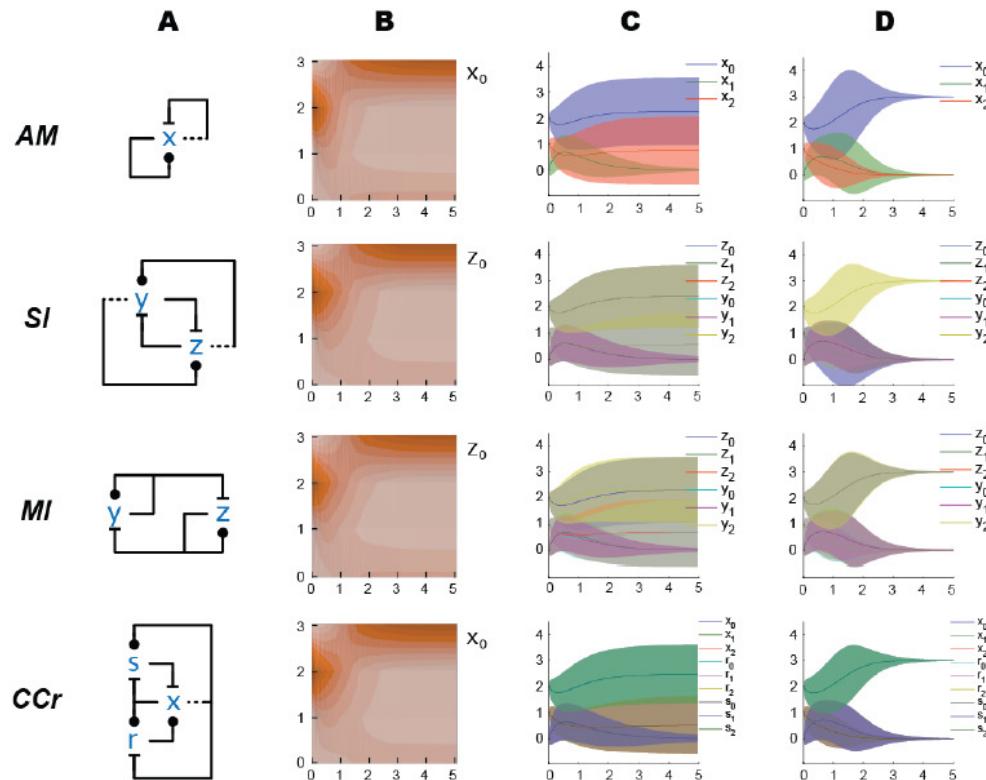
This can be done for any initial conditions chosen for AM, and indicates the potential of each of the other networks to operate as a simpler switch.

Noise Reduction in Complex Biological Switches

Luca Cardelli^{1,2,¶,*}, Attila Csikász-Nagy^{3,4,¶}, Neil Dalchau^{1,¶}, Mirco Tribastone^{5,¶}, Max Tschaikowski^{5,¶}

(To appear.)

Basic Switches (stochastic)



Horizontal axes is time
Vertical axes is number of molecules.

(A) Influence networks.

(B) Chemical Master Equation solution: probability distribution, with color (in 10 bands from light = 0 to dark = 1) indicating the probability that at time t there are y molecules of the single indicated species.

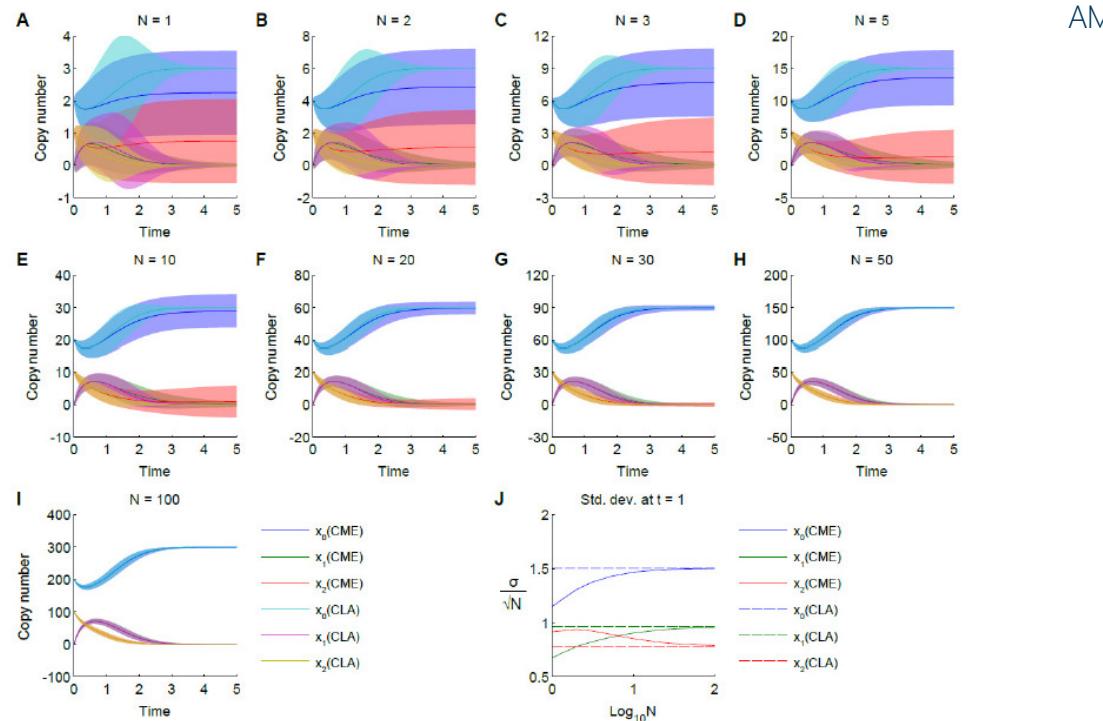
(C) Chemical Master Equation solution: mean (solid lines) and standard deviation (color bands) for the species in the network.

(D) Central Limit Approximation solution: mean (solid lines) and standard deviation (color bands) for the species in the network.

Disentangle the contribution of complexity to stochasticity

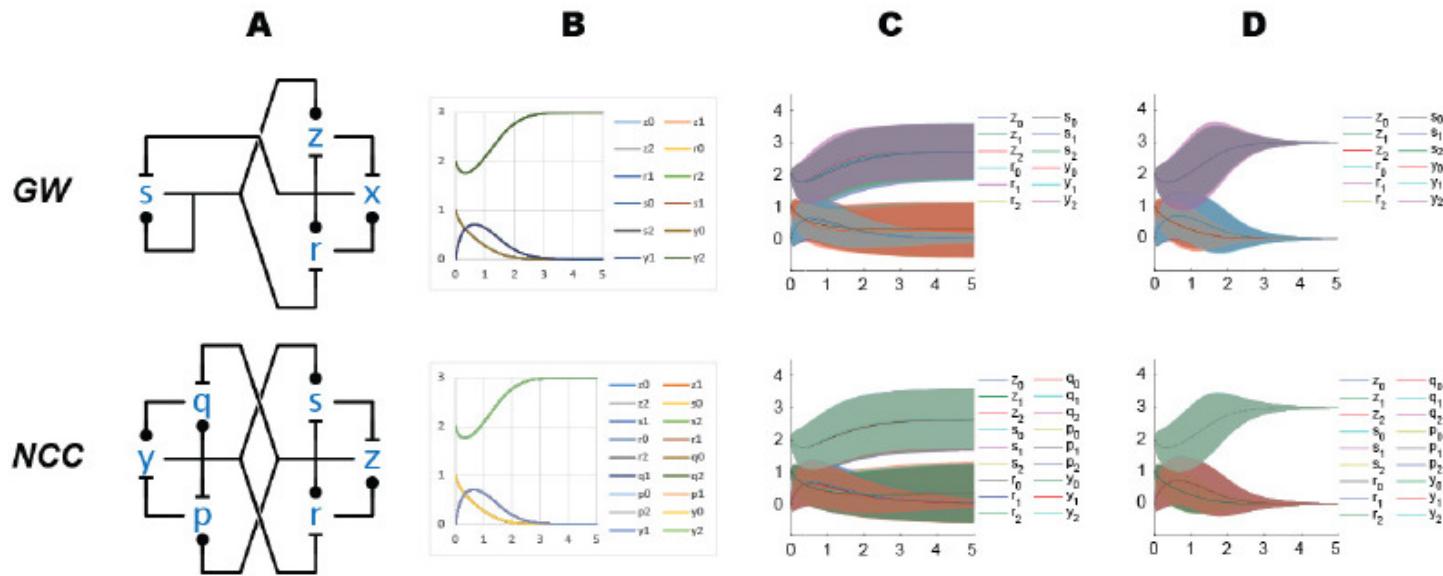
Compare network noise on the baseline of deterministic emulation, across networks of different size and structure

CME vs LNA in the limit



AM at various system sizes

More Complex Switches



Horizontal axes are time, vertical axes are number of molecules.

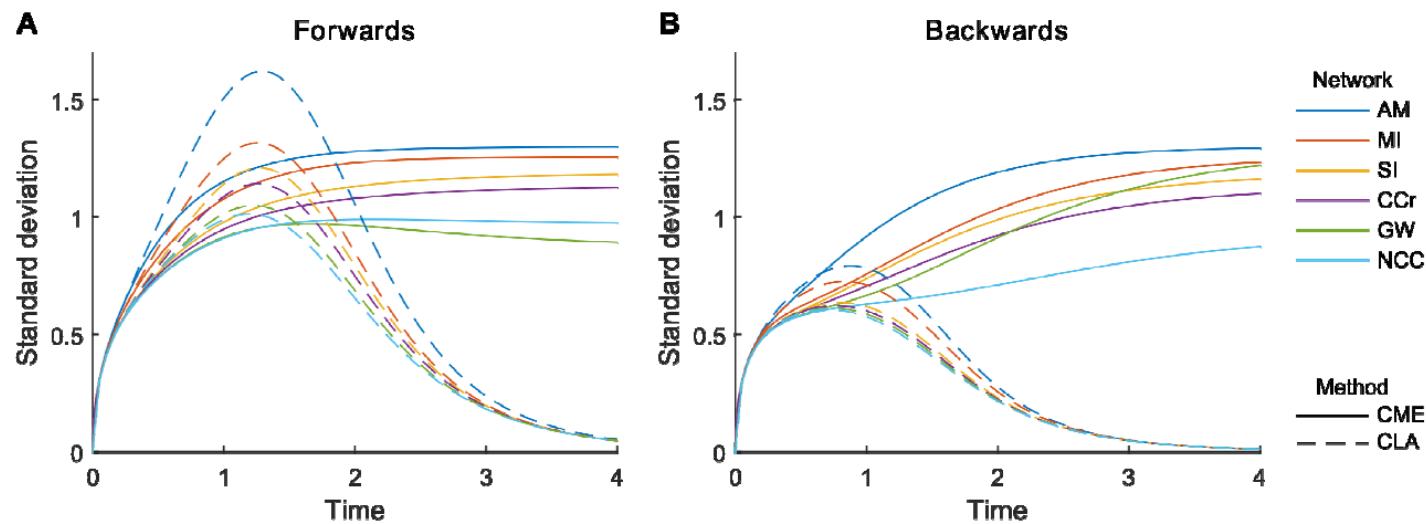
(A) Influence networks.

(B) ODE solutions for comparison

(C) Chemical Master Equation solution: mean (solid lines) and standard deviation (color bands) for the species in the network.

(D) Central Limit Approximation solution: mean (black lines) and standard deviation (color bands) for the species in the network.

Intrinsic Noise

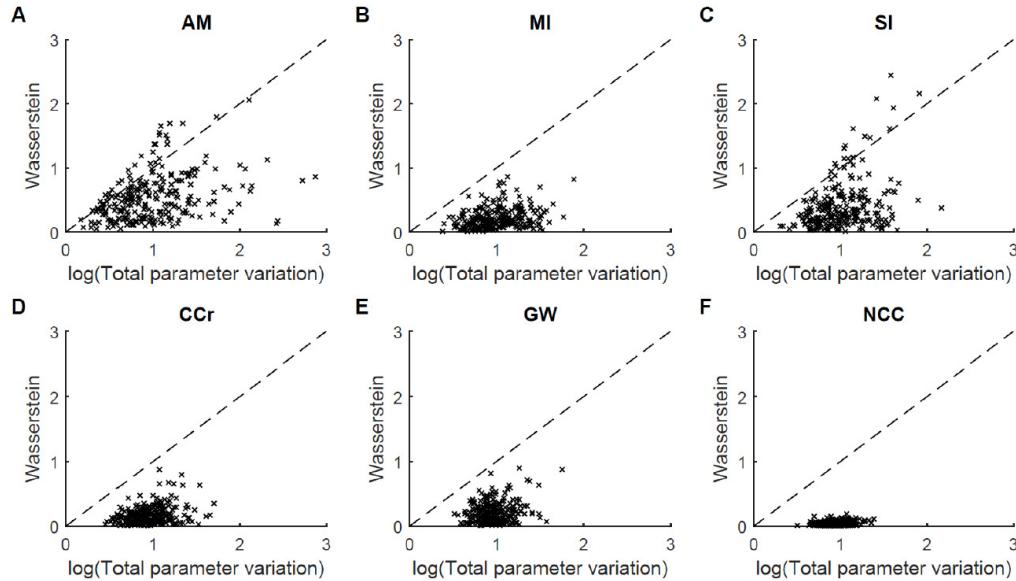


Complexity improves overall performance of the cell cycle switch. The performance of different networks was evaluated by calculating the standard deviation of the main molecular states over time.

Standard deviations are calculated via numerical integration of the chemical master equation (CME) using the Visual GEC software, and via numerical integration of the central limit approximation (CLA) in Matlab. We investigate switching in one direction or the other by providing different initial conditions that settle (more likely) in different steady states.

- (A) In the forward direction, principal molecular states were initialised at 2 copies, and complementary molecular states were initialised at 1 copy.
- (B) In the reverse direction, principal molecular states were initialised at 1 copy, and complementary molecular states were initialised at 2 copies.

Extrinsic Noise



Complexity confers switching networks robustness to extrinsic noise. Extrinsic noise was analyzed by randomly perturbing the reaction rates of each model. Variations in network behaviour were assessed in comparison to the behaviour of the default parameterisation, in which all reaction rates are set equal to 1. Network variation was quantified using the summed Wasserstein metric over the whole probability distribution over time.

Noise vs. Complexity

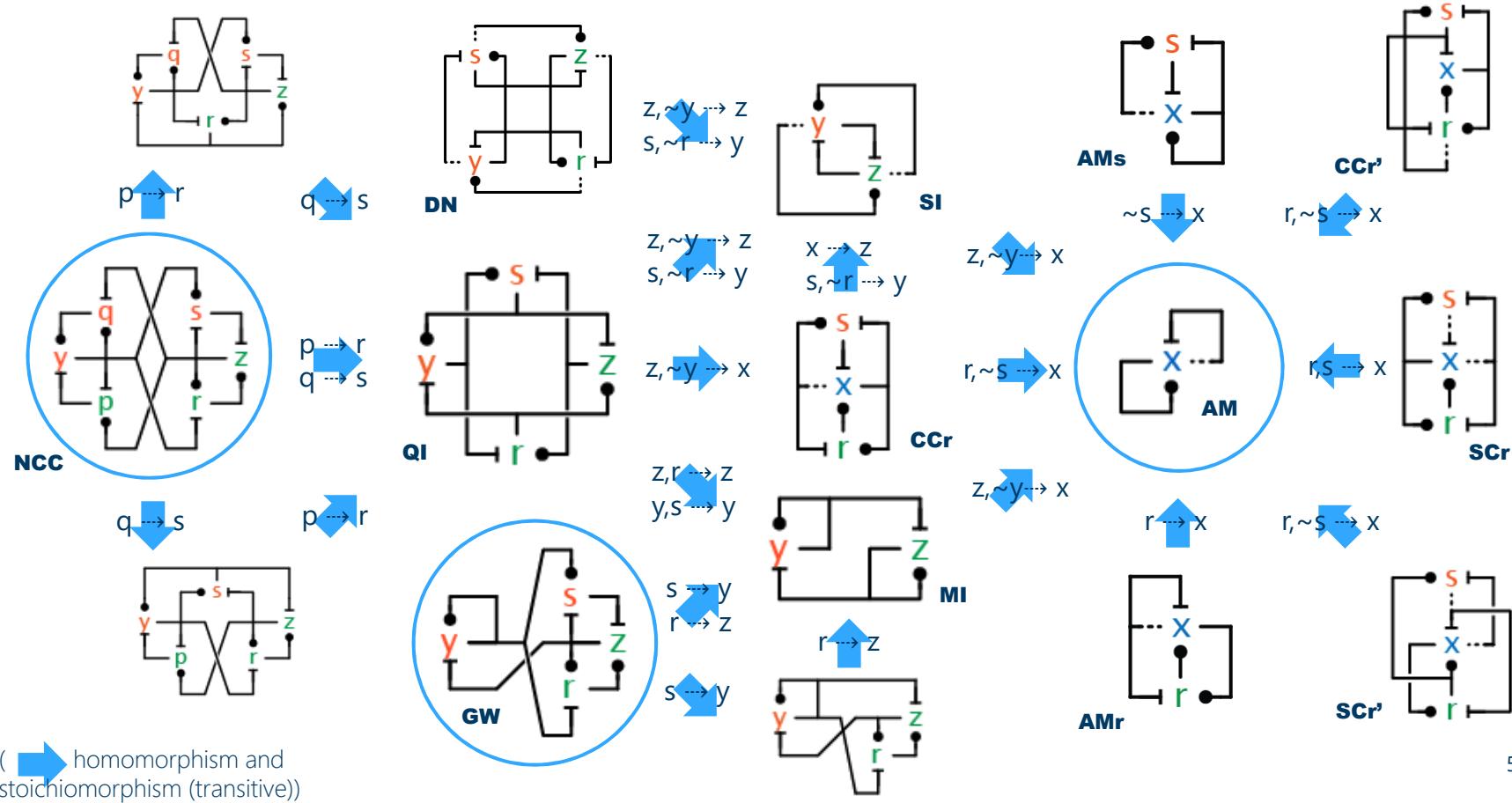
- With corresponding initial conditions, all studied networks show the same mean behavior
- CCr emulating AM is the simplest explanation of the core cell cycle switching function
- Many other biological switches can be so reduced to an algorithm with well-understood properties
- On the basis of kinetic similarity of mean behavior, we show variations in noise behavior.
- Intrinsic noise tends to decrease with complexity, but this also depends on network structure and *not* directly on total molecular counts

Complexity vs. Cost

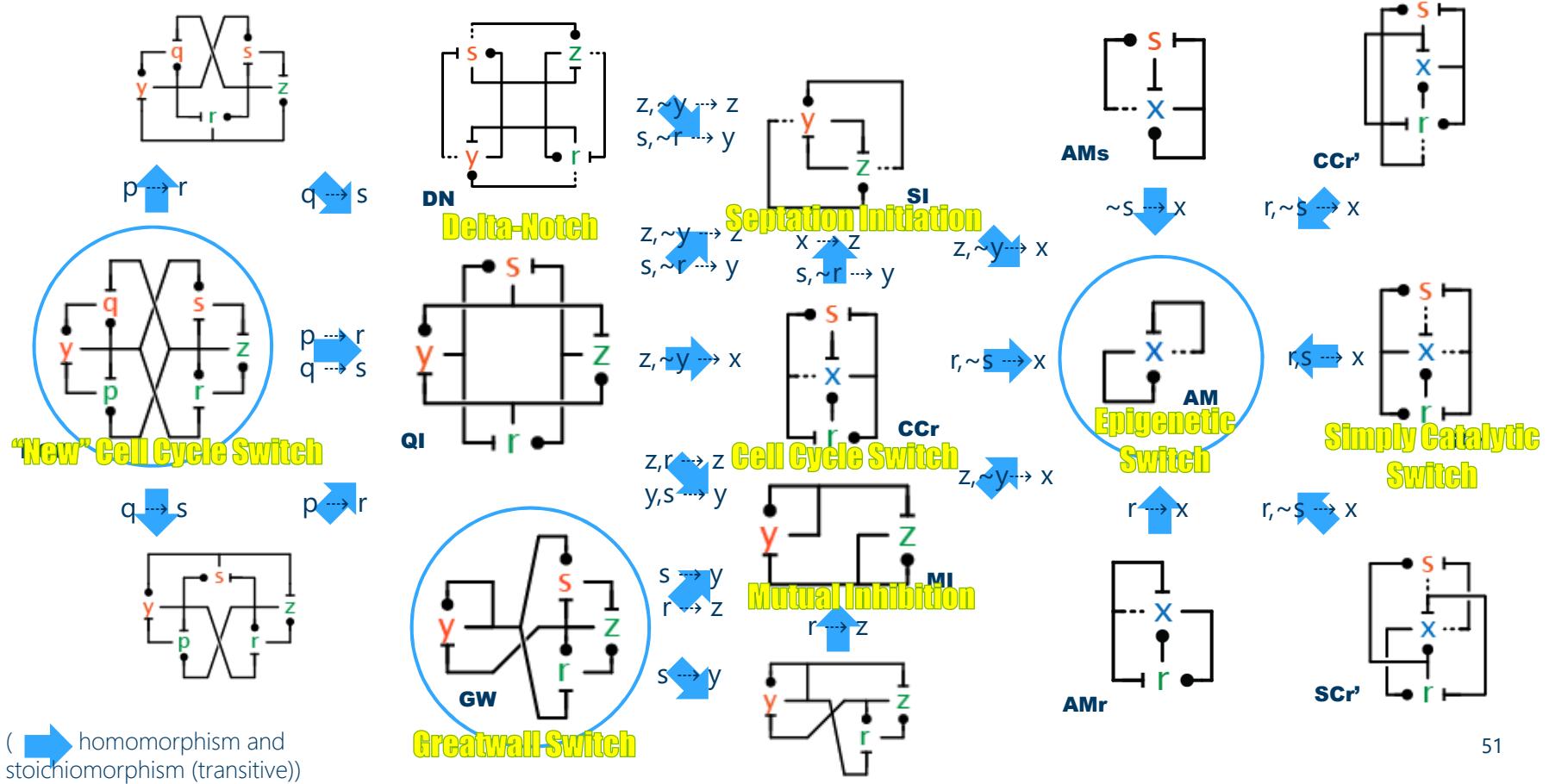
- Complex networks, while more expensive, are less of a burden in energy rich situations.
- The cell cycle operates only in such “wellness” conditions.
- Hence complex switches may have evolved to work better by using more resources
- Complex network also reduce noise levels, so for a fixed noise level that can be tolerated, they work at lower molecular level for each species.

Conclusions

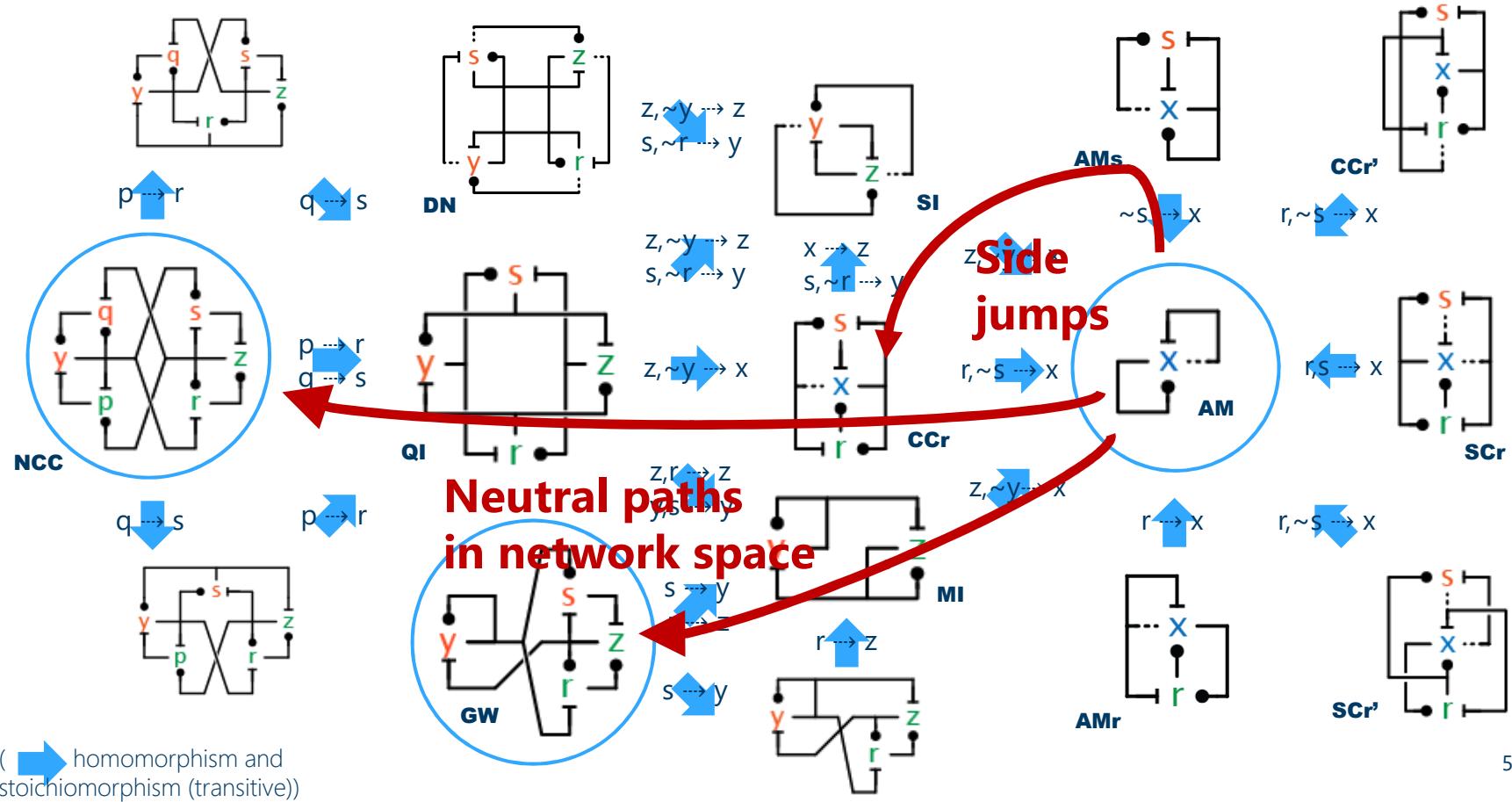
Walks in Network Space



Walks in Network Space



Walks in Network Space



Networks are Algorithms

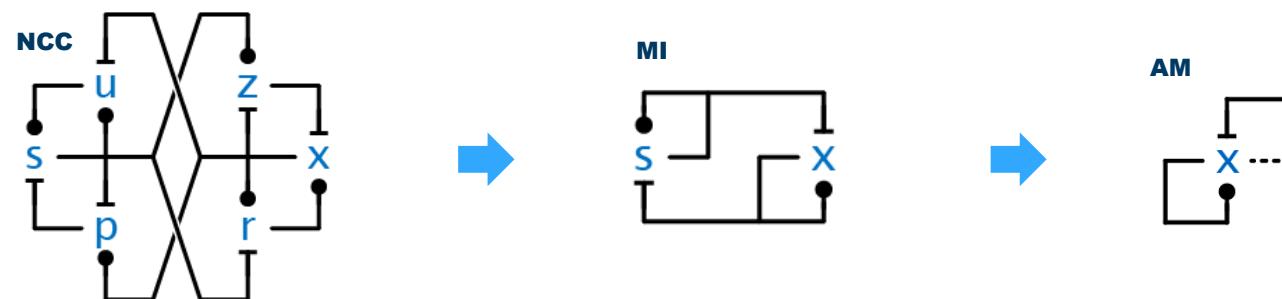
- They are *methods* for achieving a function
 - We need to understand how these methods relate to each other
 - In addition to how and how well they implement function
 - Algorithms can be obfuscated, and nature can obfuscate networks (to what end?)
- Network emulation can be checked *statically*
 - By stoichiometric/reaction-rate (*structural*) properties
 - That is, no need to compare ODE (*functional*) properties
 - For *any* initial conditions and rates of (one of) the networks
- We can efficiently discover emulations
 - Automatic model reduction of large networks

Nature likes good algorithms



These additional feedbacks *do exist* in real cell cycles (via indirections)

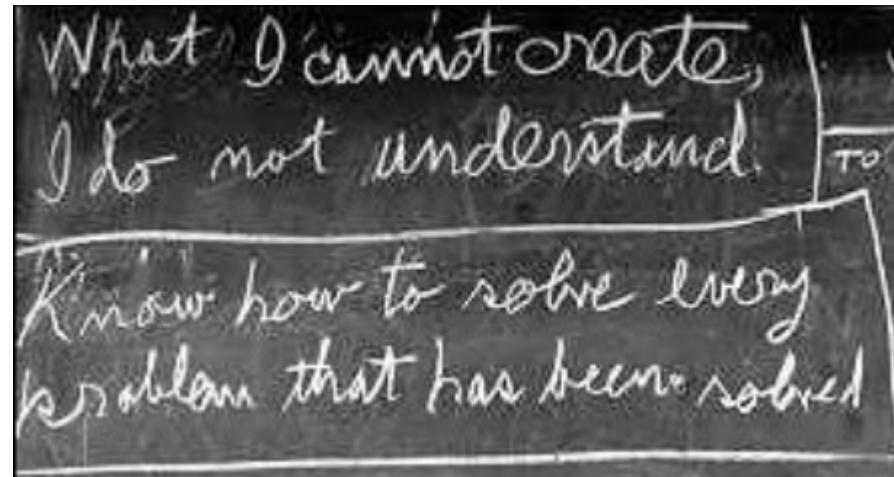
The cell cycle switch *can exactly* emulate AM



What Contributors to Complexity?

- Indifference? (does not really cost much)
- Robustness? (resist point failures)
- Adaptability? (neutral paths)
- Noise resistance? (improve signal processing)
- Temperature compensation?
- Etc.

Feynman's Blackboard



© Copyright California Institute of Technology. All rights reserved.
Commercial use or modification of this material is prohibited.