Does topology of a biochemical network influence its function?

Ilya Nemenman (LANL)

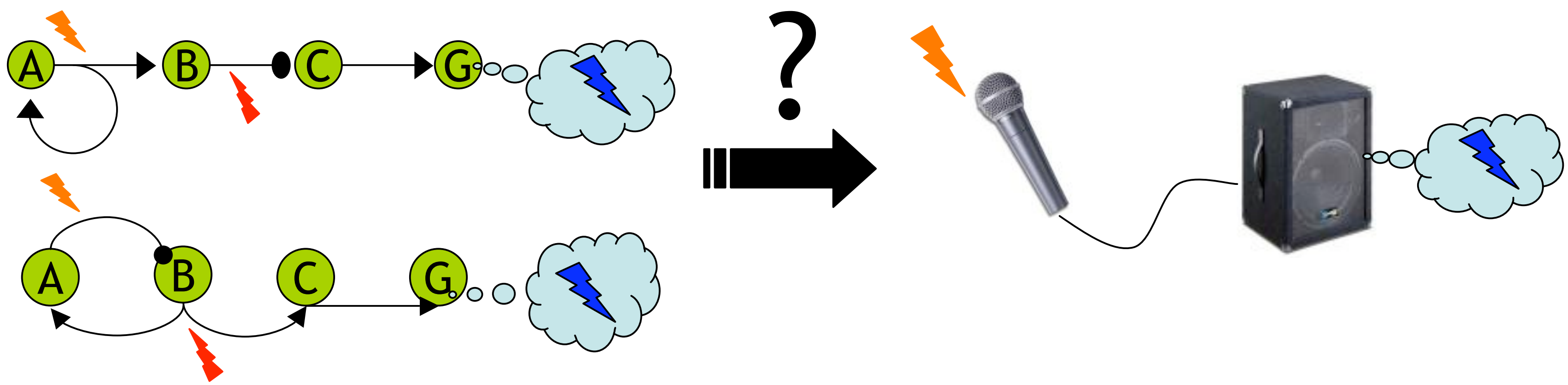
with

Etay Ziv, Chris Wiggins (Columbia)

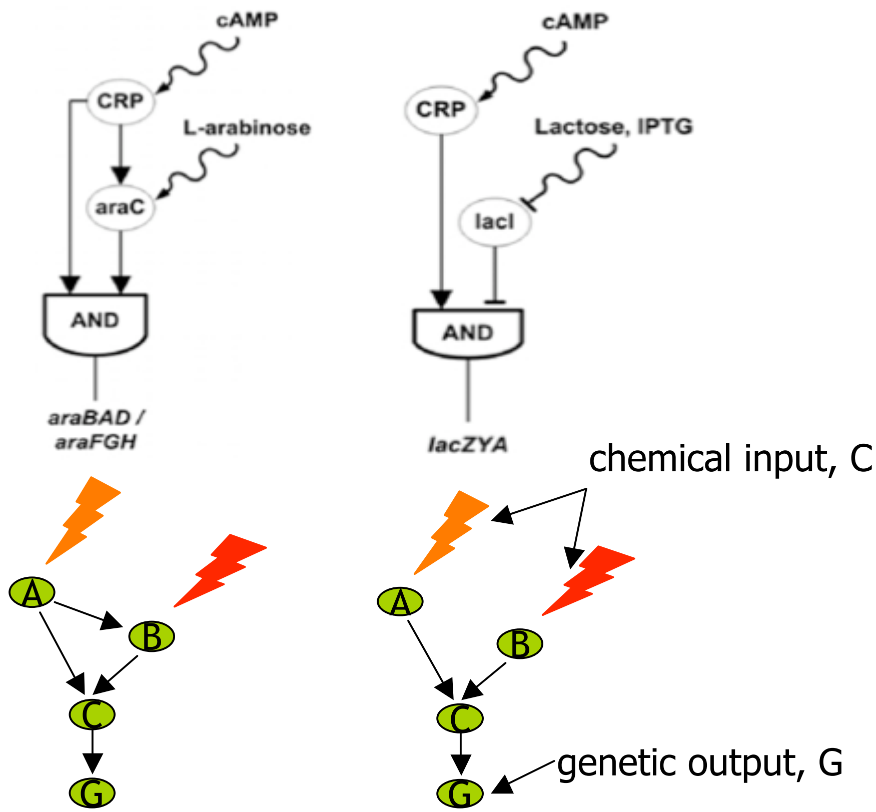
Los Alamos NATIONAL LABORATORY EST. 1943

Operated by the Los Alamos National Security, LLC for the DOE/NNSA

How can a function of a bionet be characterized?



Topology has a function
But… does it?

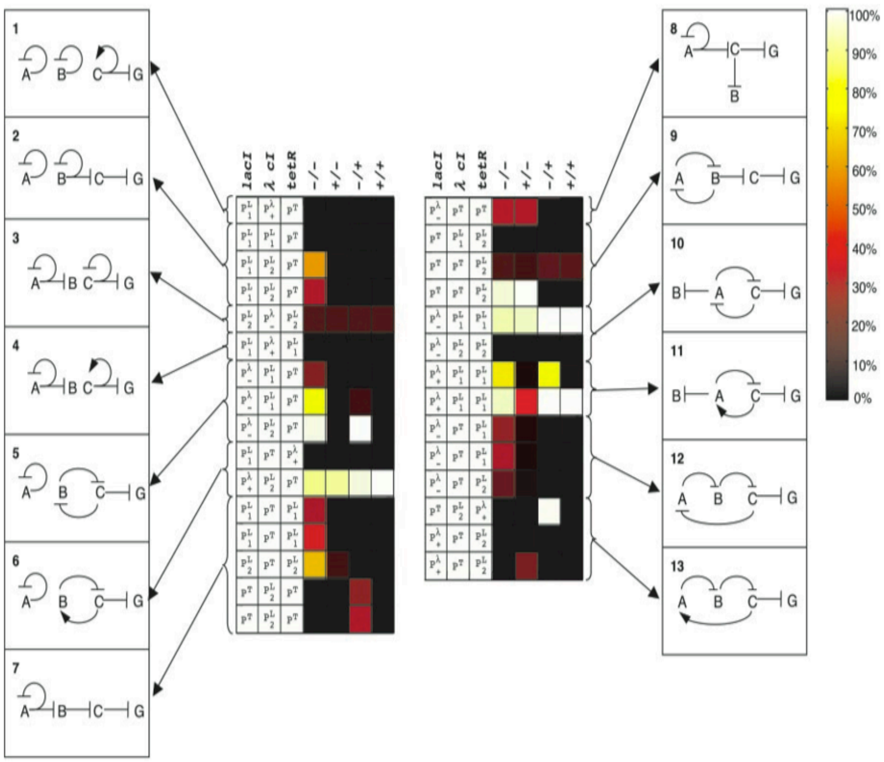


See Wall et al.: multiple functions

A Sign-Sensitive Delay

From Mangan et al., 2003

Some topologies are better than the others
But… are they?

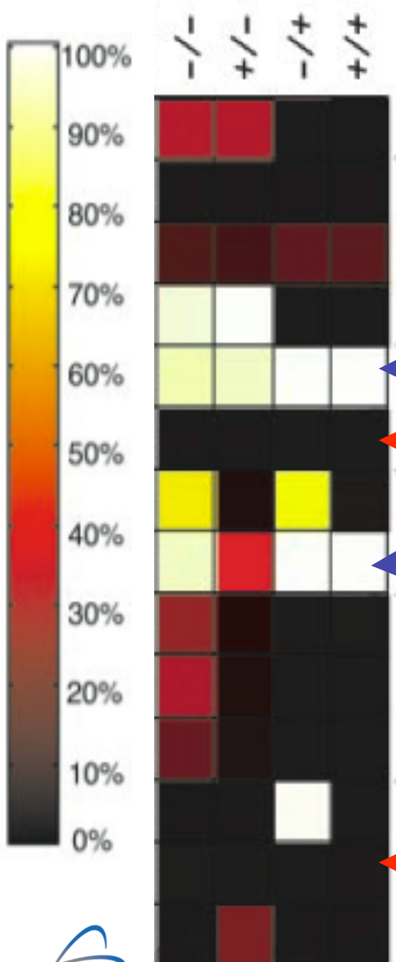


What if wrong parameters were explored?

Logic Gates

From Guet et al., 2002

How to measure circuit quality without knowing its function?



Good Circuits

Bad Circuits

Circuit Quality:

Functional integral

Need to know $P[C(t)]$

$$I[c(t),g(t)] = \int dP[c(t),g(t)] \log \frac{dP[c(t),g(t)]}{dP[c(t)]\, dP[g(t)]}$$

Simplify: Steady State Inf. Processing

$$g = g(c,t)\big|_{t\to\infty} + \text{noise}$$

$$I(C,G) = \int dc\, dg\;\; p(c,g) \log \frac{p(c,g)}{p(c)p(g)}$$

$$0 \le I(C,G) \le \min\{S(C),S(G)\}$$

Guet example:
C={(0,0),(1,0),(0,1),(1,1)}
G={+1, -1}
Broken circuit: $I(C,G)=S(G)=0$

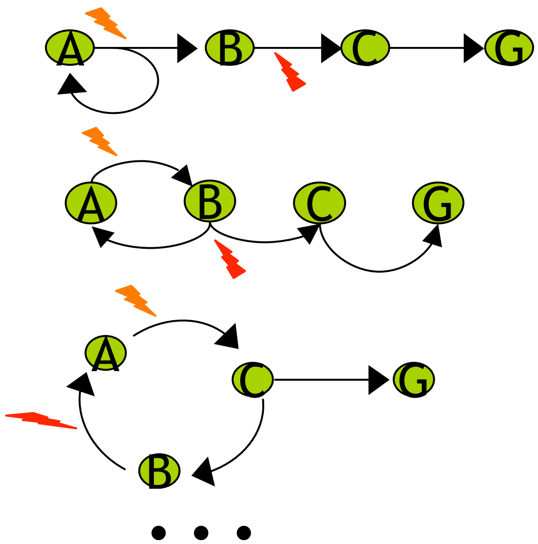
What hides beneath?

- Circuits may not have oscillations
  - Neglect solutions (low information capacity)
- Circuits may have multiple fixed points
  - Enumerate as many as possible
- Fixed points may have different basins of attraction
  - Assume equal weighting, though this is not a big problem

How good are circuits?

1. For a given topology,

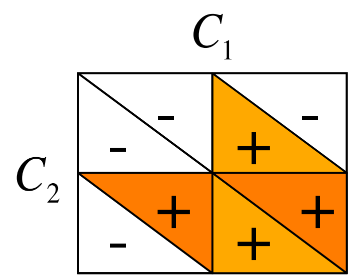
exactly one promoter per gene,
each TF binds to one promoter type



(actually do for 3 inputs)

2. For a given $p(C)$,

each input is binary



3. Calculate $g=g(c)$ for all $c\subset C$

$$\frac{dg}{dt} = -R_g g + a_0 + \alpha\left(\{g,c\}\right) \ldots$$

θ

Los Alamos NATIONAL LABORATORY EST. 1943

Operated by the Los Alamos National Security, LLC for the DOE/NNSA

How good are circuits?

4. And maximize information.

$$\hat{\theta} = \arg\max_{\theta \subset \left(\text{biologically realistic}\right)} I(C,G)$$

$$\begin{aligned}\hat{\theta} &= \text{argmax}_{\theta} I(G,C) - \lambda_1 N \\ \hat{\theta} &= \text{argmax}_{\theta} I(G,C) - \lambda_2 T \\ \hat{\theta} &= \text{argmax}_{\theta} I(G,C) - \lambda_1 N - \lambda_2 T\end{aligned}$$

Constraints on time to and the copy # at the steady state.

max= high fidelity differentiation in development
high capacity signal transduction (*lac*, photoreceptor)

5. How does max(I) depend on the parameters and the topology?

Operated by the Los Alamos National Security, LLC for the DOE/NNSA

Calculating *P(g|c):* Linear noise

copy #

deterministic
\# density

$$X_i = \Omega g_i + \Omega^{1/2} \xi_i$$

volume

white noise
fluctuations

Evolution of probability density:

$$0 = \frac{\partial \Pi(\xi, t)}{\partial t} = -\sum_{ik} A_{ik} \frac{\partial (\xi_k \Pi)}{\partial \xi_i} + \frac{1}{2} \sum_{ik} B_{ik} \frac{\partial^2 \Pi}{\partial \xi_i \partial \xi_k}$$

Van Kampen, 1997
Elf and Ehrenberg, 2003
Paulsson et al., 2004

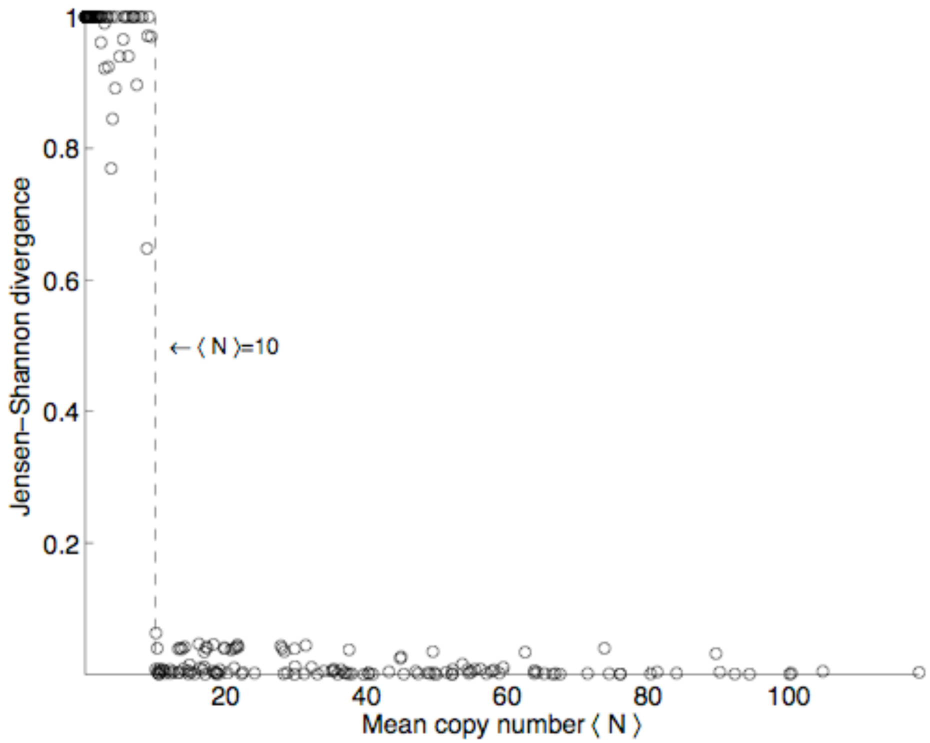
Noise covariance $\Sigma = \left\langle \xi \xi^T \right\rangle$

$$A\Sigma + \Sigma A^T + \Omega B = 0$$

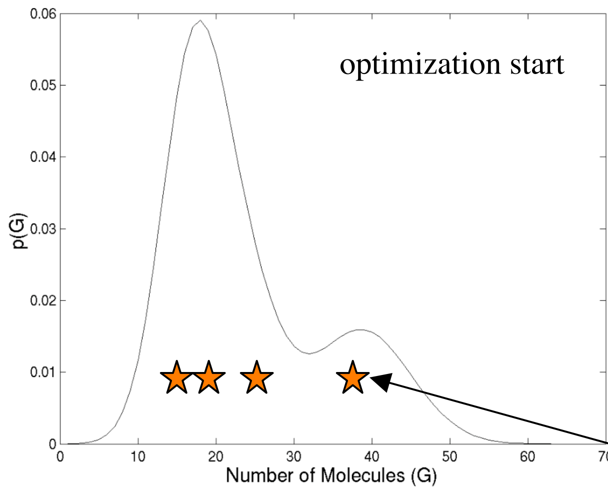
At steady state:

$$P(g \mid c) = N\left[g(c), \Sigma \right]$$

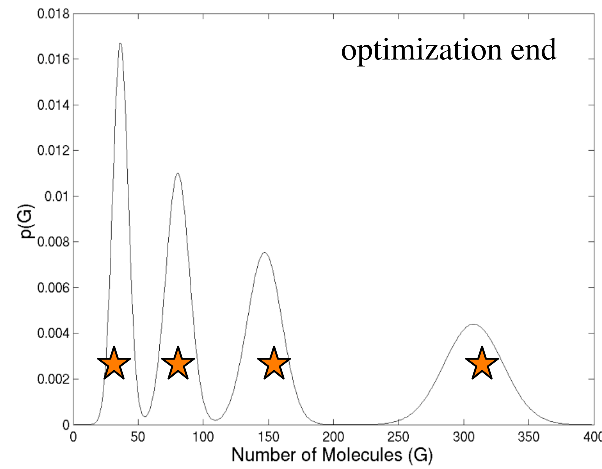
How good is LNA (JS measure)?



Numerics: increasing MI

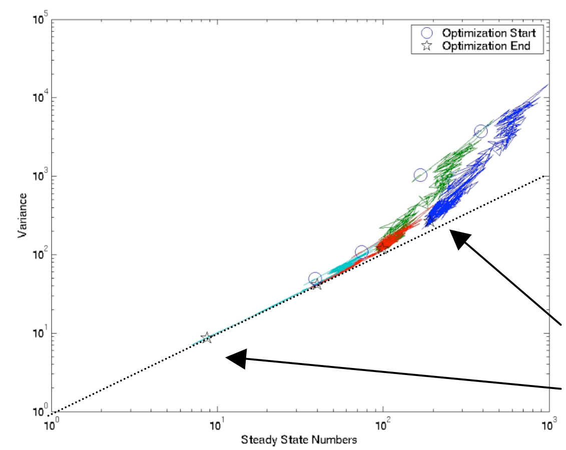


separating peaks



Mean outputs in response to different inputs

decreasing the reporter variance to the Poisson limit (low pass filtering upstream noise by slow reporter); variance of the other species may be sub-Poisson (negative feedback)



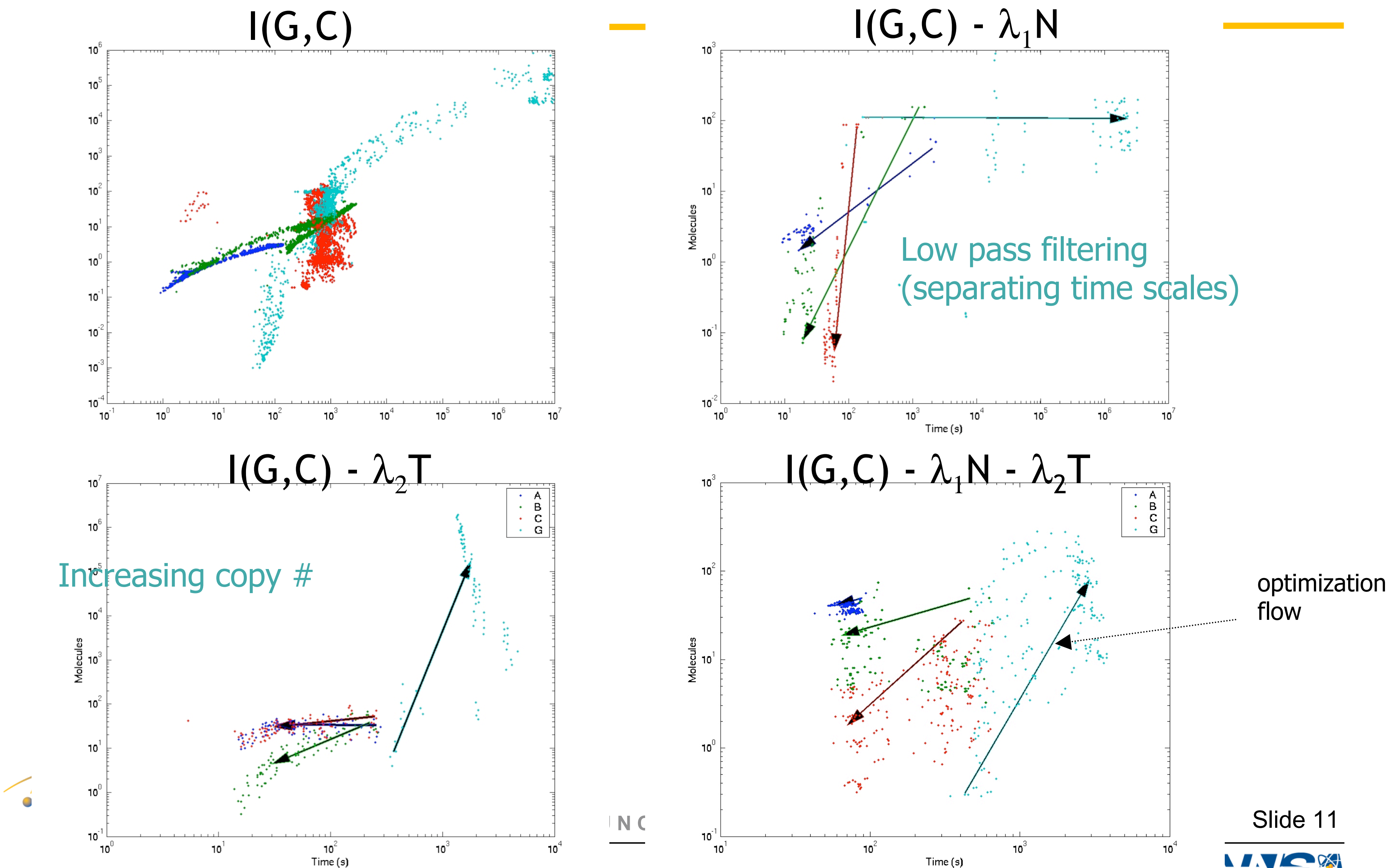
other species

reporter

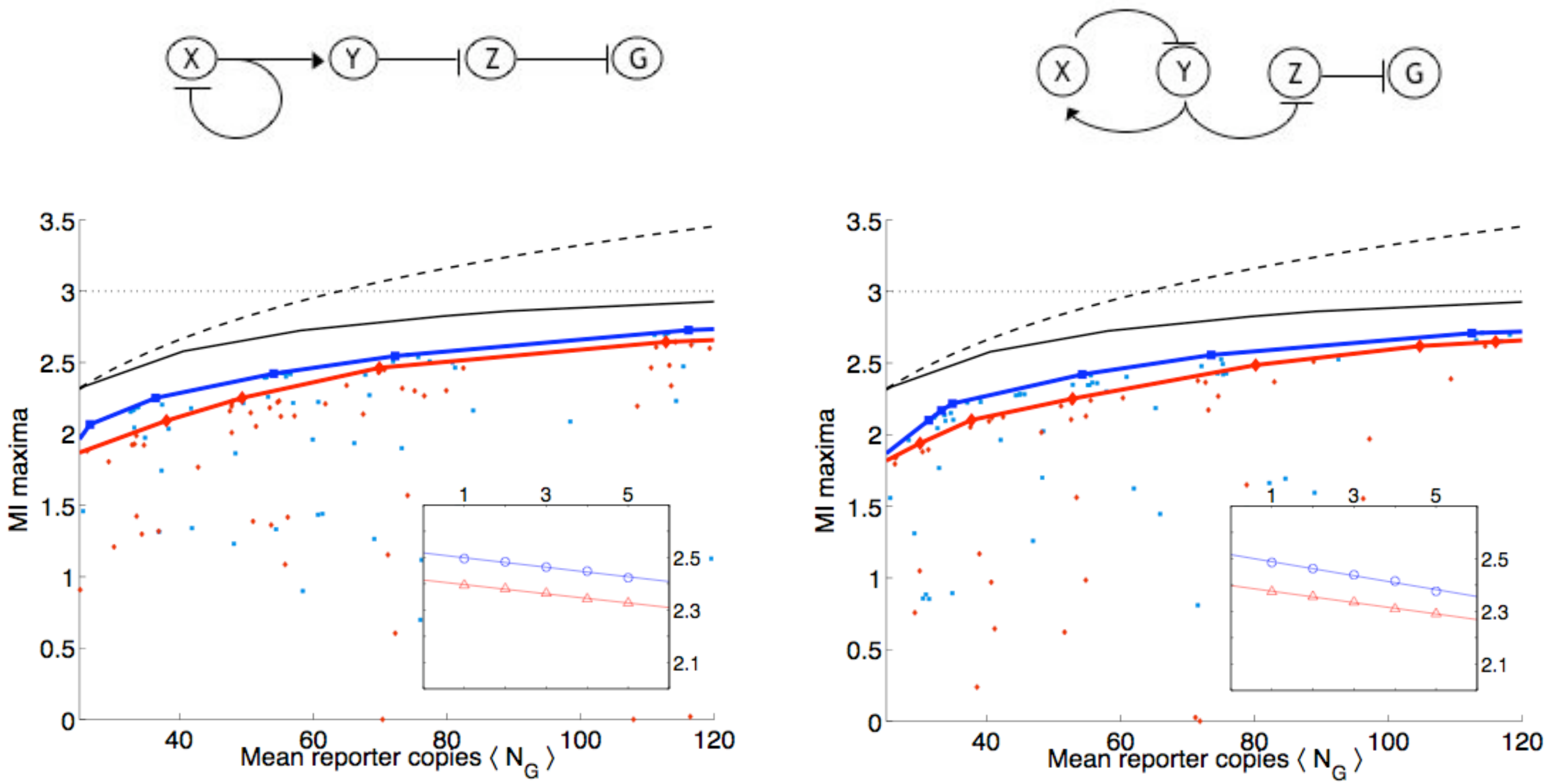
Los Alamos NATIONAL LABORATORY EST. 1943

Operated by the Los Alamos National Security, LLC for the DOE/NNSA

Achieving 2 bits

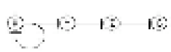
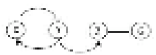
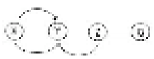
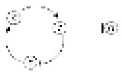
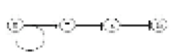
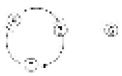
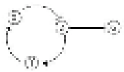
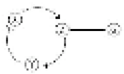
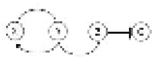
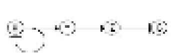


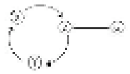
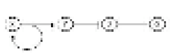
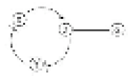
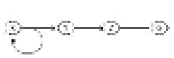
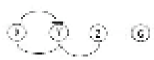
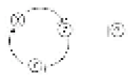
Specific circuits: more than 1 bit, almost optimal



Maximum: analytics and numerics

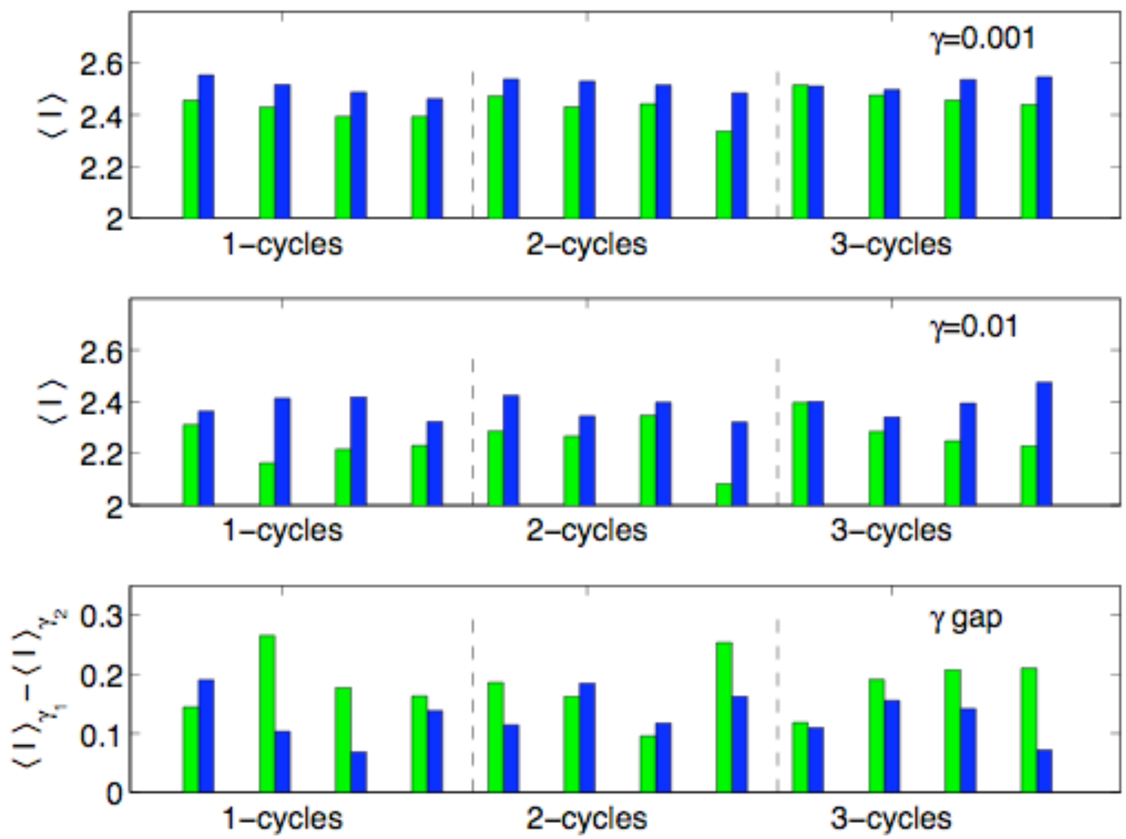
Is topology important?

Number	Topology	$\gamma = 0.001$	$\gamma = 0.01$
1		2.5570	2.3638
20		2.5524	2.3970
6		2.5451	2.4818
2		2.5357	2.3549
22		2.5354	2.3909
19		2.5218	2.3718
10		2.5172	2.3925
13		2.5055	2.4058
8		2.5002	2.3463
23		2.4976	2.3831
14		2.4874	2.4251
12		2.4809	2.3219

Number	Circuit	$\gamma = 0.001$	$\gamma = 0.01$
17		2.4695	2.2876
5		2.4659	2.2806
4		2.4624	2.2930
21		2.4605	2.23121
9		2.4497	2.3491
7		2.4420	2.2773
15		2.4244	2.1587
24		2.4234	2.2123
11		2.3958	2.2143
16		2.3943	2.2281
18		2.3603	2.0751
3		2.3099	2.2471

All are great!
Some are better than others

Positive vs. negative feedback



negative

positive

$p = 0.0002$

$p = 0.0003$

$p = 0.01$

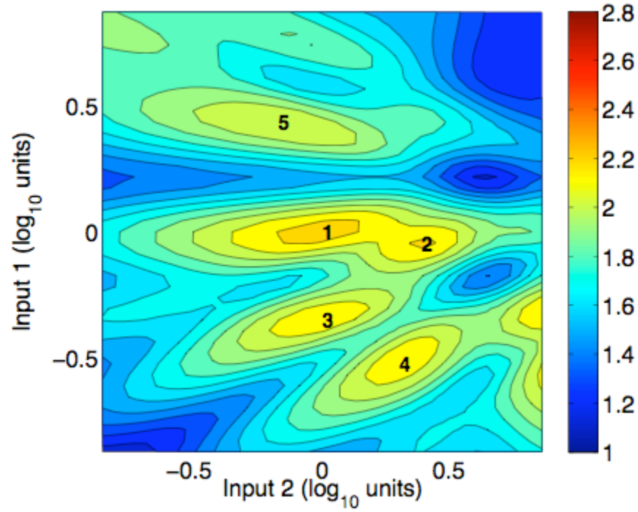
NF circuits have higher capacity and reach it easier

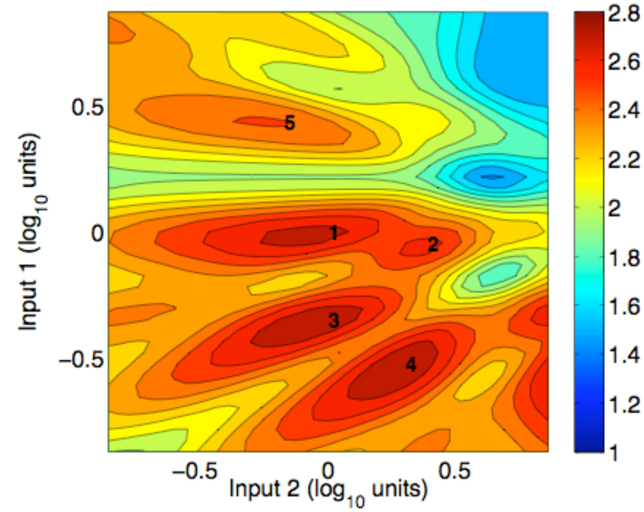
Explanation available in terms of decreased state variance

Los Alamos NATIONAL LABORATORY EST. 1943

Operated by the Los Alamos National Security, LLC for the DOE/NNSA

Multiple functions?

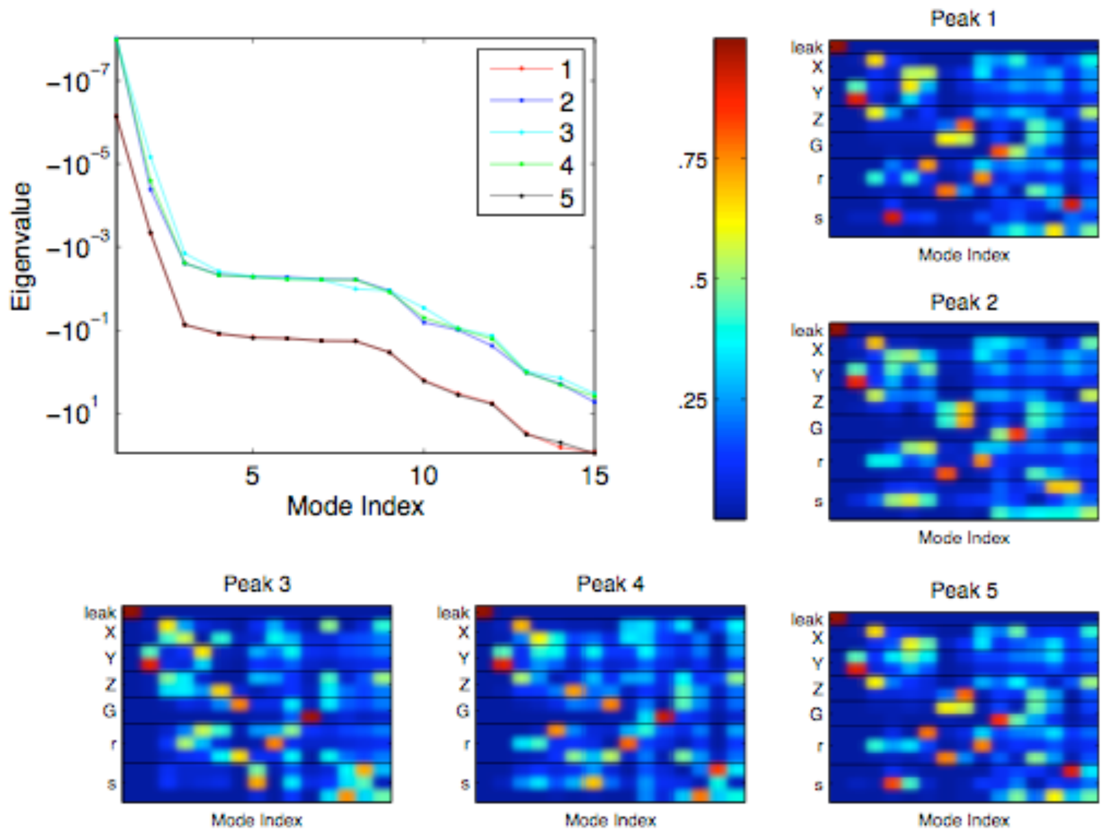




Topology 2
First and
second K_m

Chemical State	000	001	010	011	100	101	110	111
Peak 1	2	6	1	5	4	8	3	7
Peak 2	2	6	4	1	5	8	3	7
Peak 3	2	1	4	6	3	5	8	7
Peak 4	2	1	6	4	5	3	8	7
Peak 5	6	2	5	1	8	4	7	3

Robust maxima?



Los Alamos NATIONAL LABORATORY EST. 1943
Operated by the Los Alamos National Security, LLC for the DOE/NNSA

Predictions

Fast response and autorepression - correlated
Rosenfeld et al. (2002) - autorepression causes fast response
Alternative: Fast response *requires* negative feedback (cannot average)

	Negative Feedback	No Negative Feedback
Proteolysis	9	4
No Proteolysis	44	88

p=0.013

Conclusions

- Small, noisy, generic biochemical networks easily achieve >1 bit of information throughput over short times (with all biochemical parameters within realistic ranges) with only a handful of molecules. Thus the same pathway can transmit >1 binary signal, and cross-talk is not a problem even for stochastic systems.
- In a steady state, the circuits come very close to transmitting the maximum information possible given a fixed number of involved molecules. This generic optimality is intriguing. It may suggest that some regulatory topologies cannot be evolutionary selected over the others based on their signal processing properties alone.
- Transmitted information is only weakly sensitive to the biochemical parameters within large ranges: no fine tuning is required.
- It is plausible that distinctions between different topologies emerge for complex, high entropy signals.
- Negative feedback circuits perform marginally better
- Multiple functions per circuit (more exploration is needed).