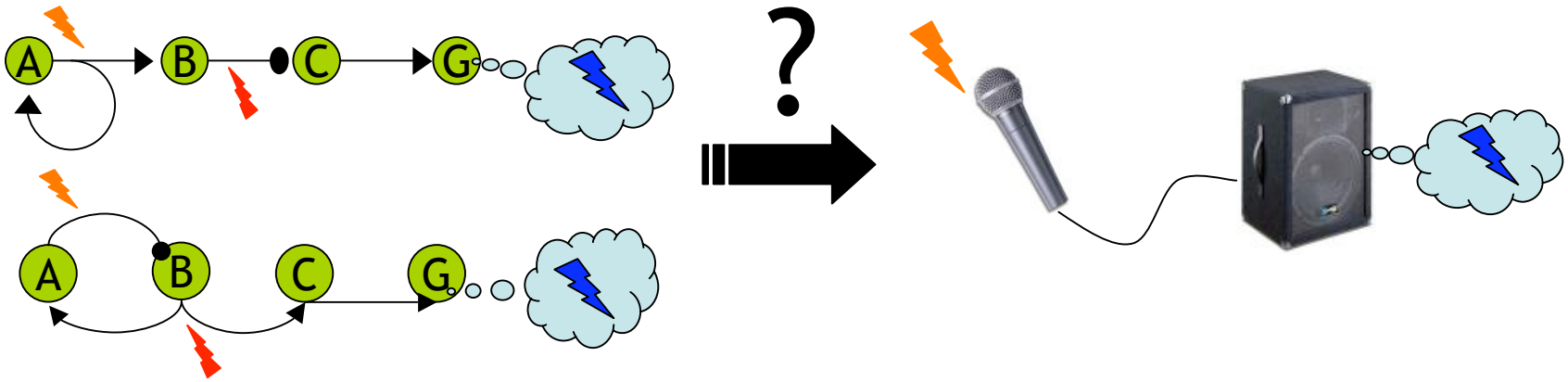
Optimal Information Processing in Small Stochastic Biochemical Networks



1. Columbia University, New York, NY 10013
2. Los Alamos National Laboratory, Los Alamos, NM 87545

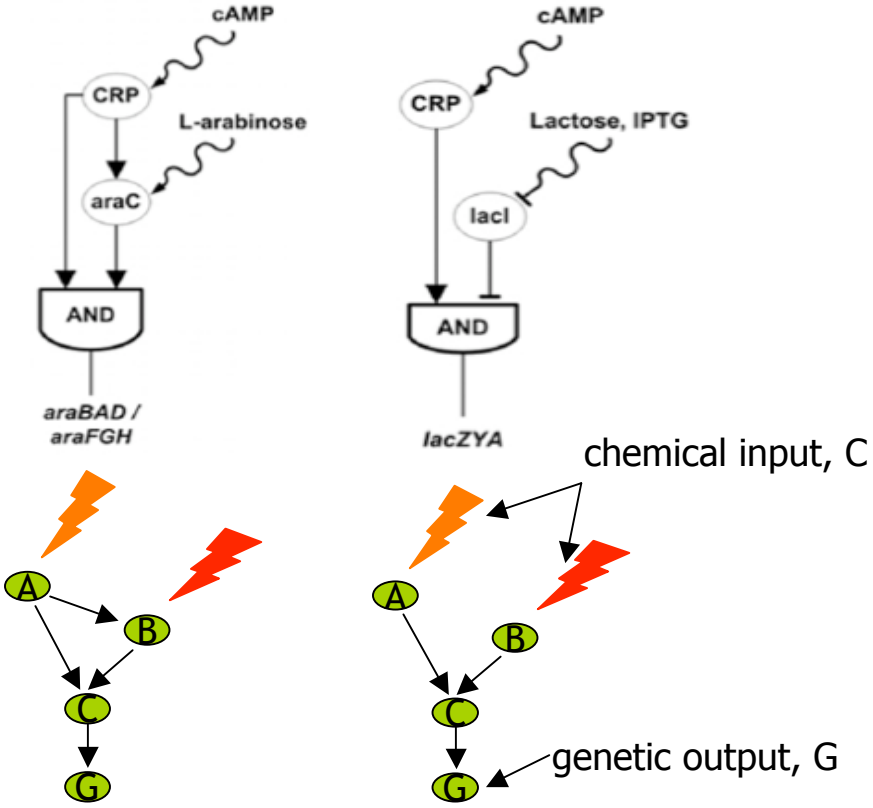
Etay Ziv[1]
Manuel Middendorf[1]
Ilya Nemenman[1,2]
Chris Wiggins[1]

contact: ez87@columbia.edu, nemenman@lanl.gov

Abstract

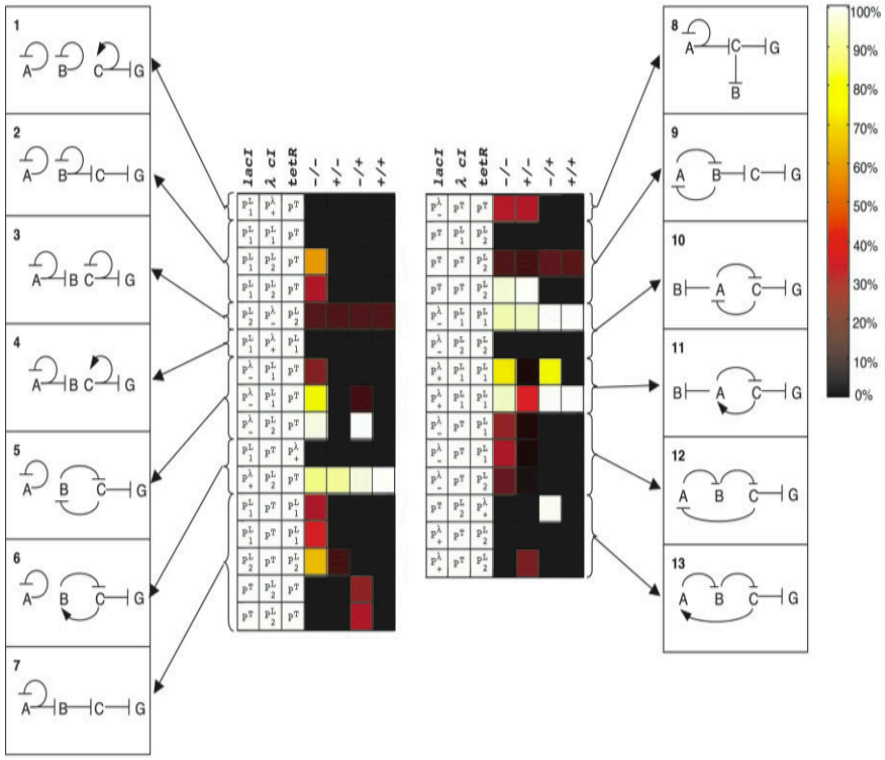
We argue that the functional quality of a biochemical signaling pathway or a regulatory circuit should be measured in terms of the amount of information (in bits) between the copy numbers of the input and the output signaling molecules that is attainable by the circuit. Treating stochastic effects by the linear noise (semiclassical) expansion around a deterministic solution of a biochemical dynamical system (which we verify by direct Gillespie simulations), we systematically analyze this mutual information in many small biochemical circuits, including various feedback loops, that can be built out of 4 chemical species coupled by Hill-type interactions as a function of ~20 chemical kinetics parameters. We study this information for a certain distribution of the input signals and maximize it over biologically realistic ranges of the parameters. Surprisingly, all the circuits manage to attain almost the maximum information possible (which we calculate analytically) for the given mean molecular copy numbers and the times it takes for the circuits to relax to the steady states. Additionally, these high information solutions are robust to rather large fluctuations in the parameters. These findings suggest potential explanations for the "cross-talk" paradox and other molecular information processing phenomena. Furthermore, they lead us to question an assumption behind many recent publications that naturally occurring biochemical networks are special in their information processing properties.

Experiments



A Sign-Sensitive Delay

From Mangan et al., 2003

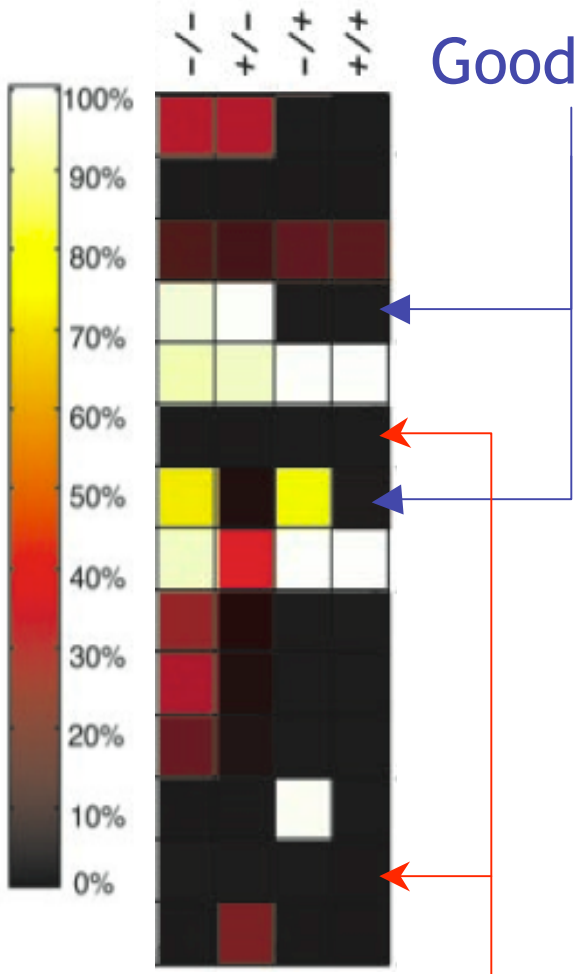


Logic Gates

From Guet et al., 2002

How to characterize the function of these systems?

Function = Information Processing



Good Circuits

Bad Circuits

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

Circuit Quality:

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

Simplify: Steady State Inf. Processing

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

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

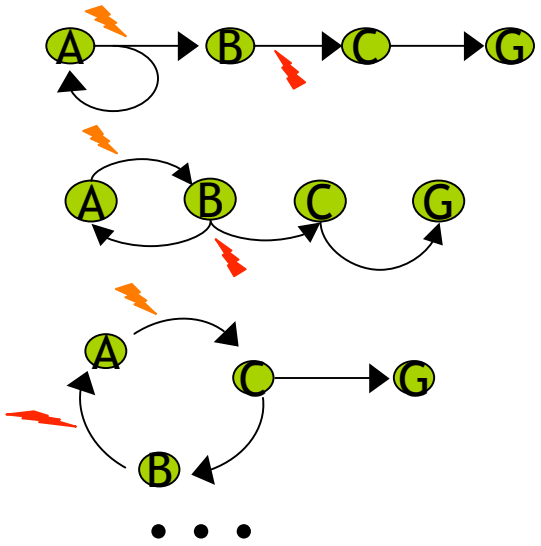
$$0 \le I(C,G) \le \min\{S(C), S(G)\}$$

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

How good are the circuits?

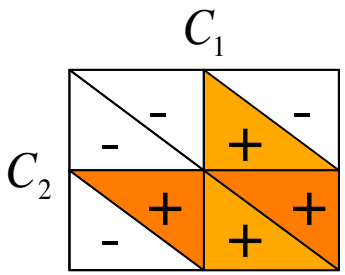
1. For a given topology,

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



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

each input is binary



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

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

θ ...

4. And maximize information.

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

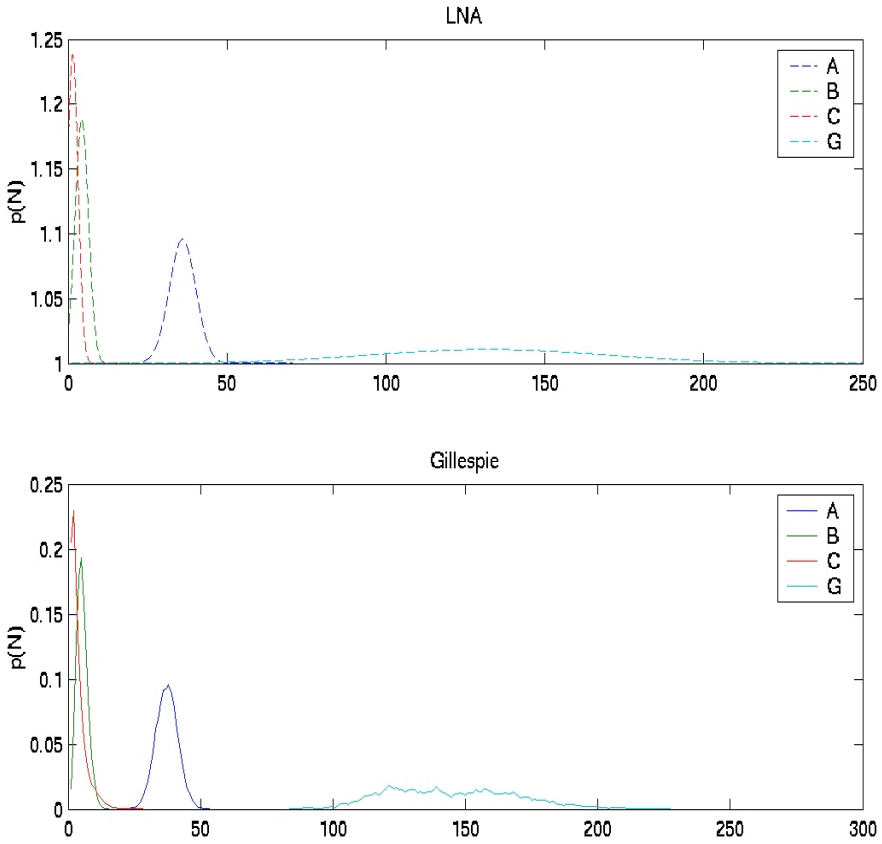
$$\begin{aligned}
\hat{\theta} &= \text{argmax}_\theta I(G,C) - \lambda_1 N \\
\hat{\theta} &= \text{argmax}_\theta I(G,C) - \lambda_2 T \\
\hat{\theta} &= \text{argmax}_\theta I(G,C) - \lambda_1 N - \lambda_2 T
\end{aligned}$$

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?

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:

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

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

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

At steady state:

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

1. For copy # as low as 10, LNA agrees with Gillespie (by KL measure).
2. We can go to higher order in $1/\Omega$.
3. Contrary to Baras et al, 1996, LNA is sound if $A = A(g),\ B = B(g)$.
4. For eig(Σ) of very different sizes, need to adiabatically integrate out the fast modes.

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

Model + parameters: details

$$0 = \frac{dg_i}{dt} = -Rg_i + a_0 + \alpha\left(g_j, s_j\right)$$

$$\text{Inhibition: } \alpha(\phi_j, s_j) = a\frac{K^n}{K^n + (\phi_j / s_j)^n}$$

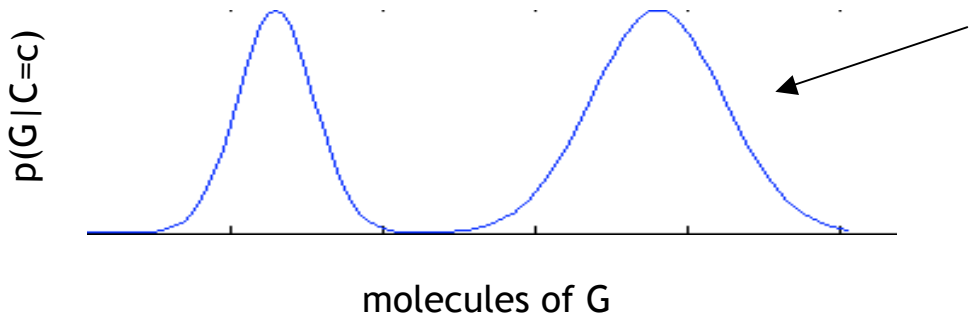
$$\text{Excitation: } \alpha(\phi_j, s_j) = a\frac{(\phi_j / s_j)^n}{K^n + (\phi_j / s_j)^n}$$

$$s_j = \begin{cases} 1, & \text{signal +} \\ \text{optimized}, & \text{signal -} \end{cases}$$

equivalent to rescaling K

g_i	determ. conc. of ith TF
R	protein decay rate
K	dissociation constant
n	Hill coefficient (set to 2)
a	range of promoter
a_0	leak of promoter
s	effect of signal molecule

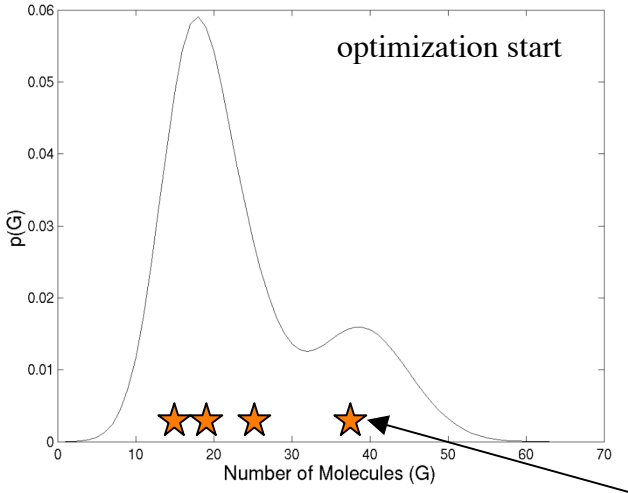
Up to 22 parameters



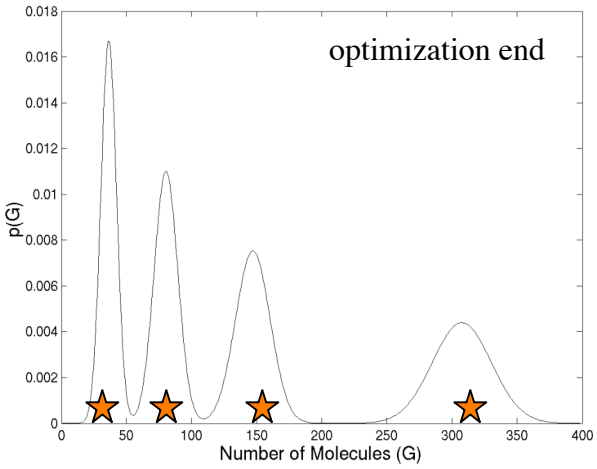
Example: two distinct steady states with Gaussian noise; P(each state | C=c)=const; no stochastic stability analysis.

However: we can consider cycles ($g \to \infty$ is never a solution, so at $t \to \infty$, we either have cycles or fixed points, and we have not observed chaos).

Numerics: Increasing MI

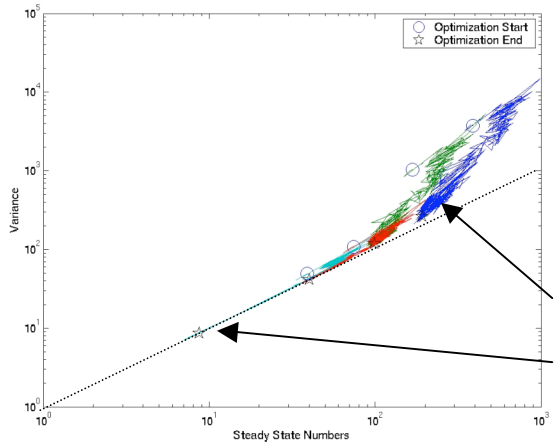


separating peaks



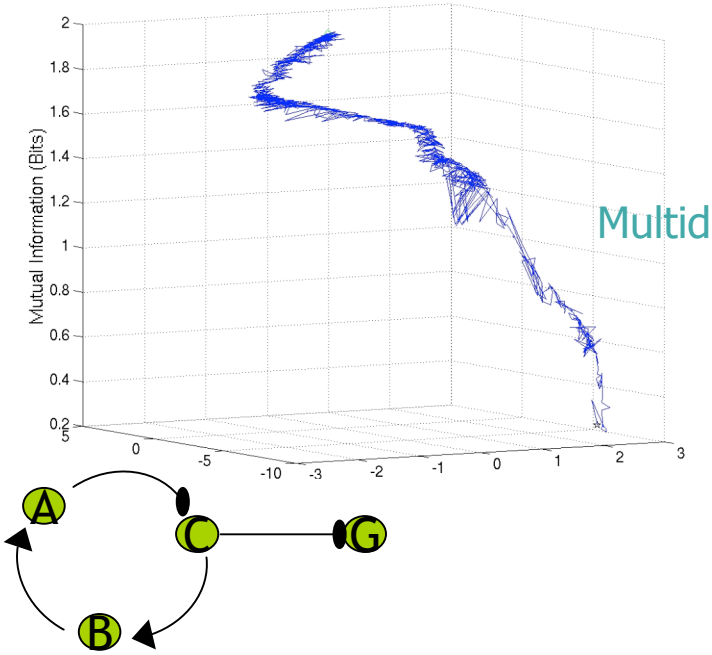
Mean outputs in response to different inputs

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)

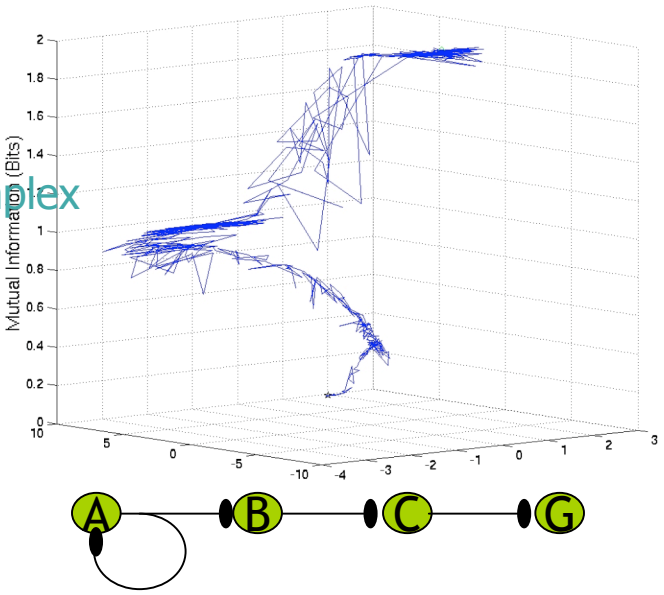


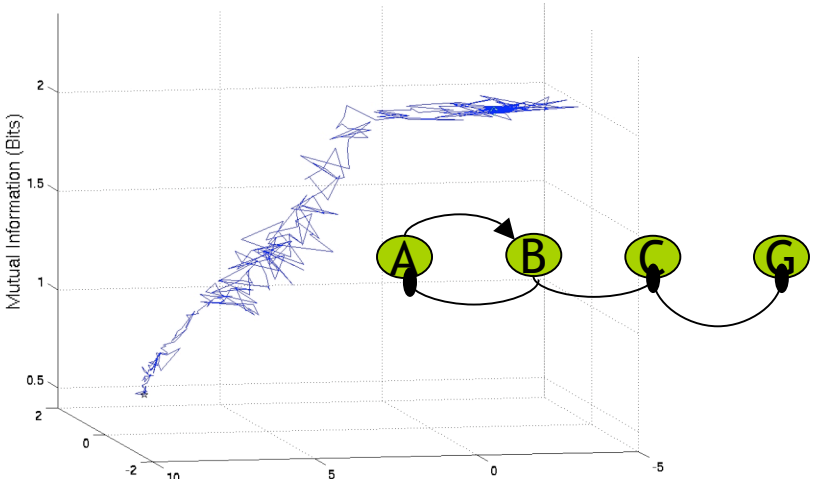
other species reporter

Achieving 2 bits (unconstrained)



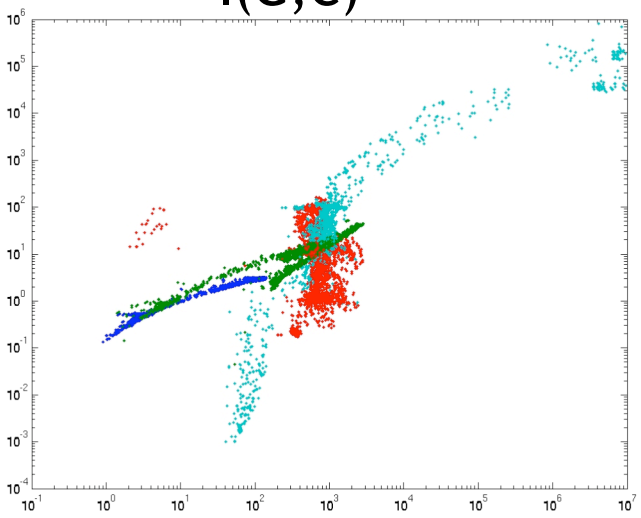
Multidimensional simplex minimization



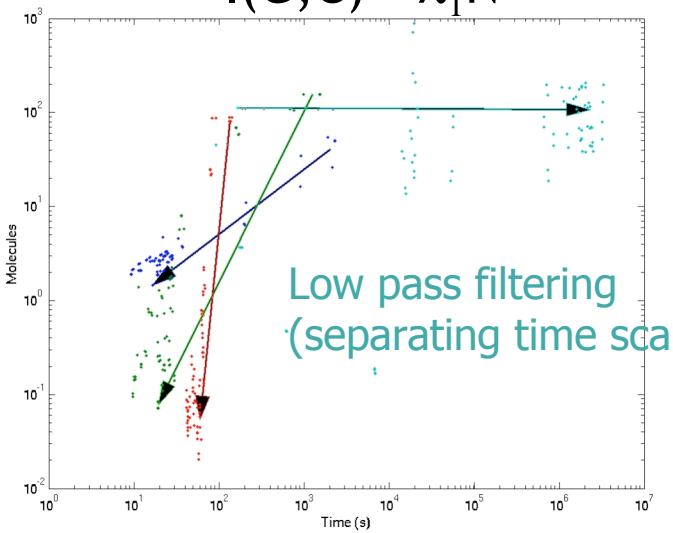


Achieving 2 bits (t,#)

I(G,C)



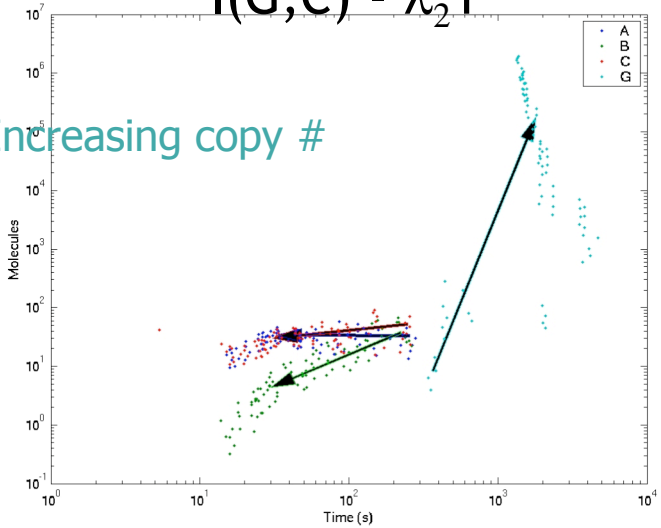
I(G,C) - λ_1N



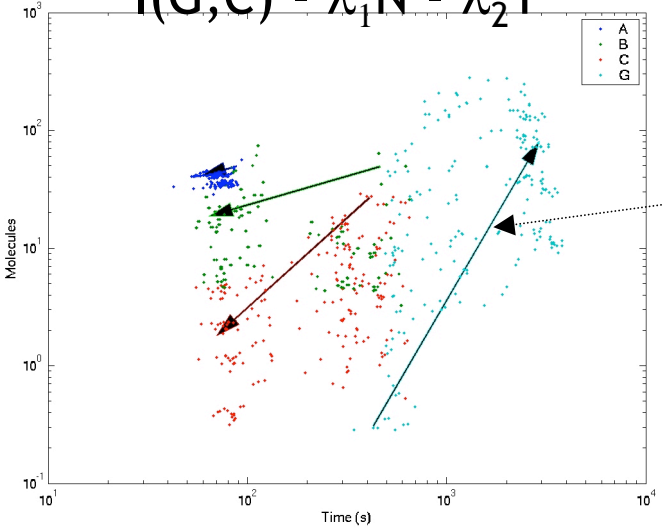
Low pass filtering
(separating time scales)

I(G,C) - λ_2T

Increasing copy #

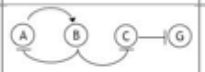
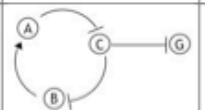
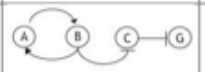
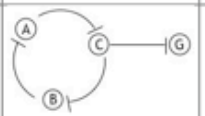
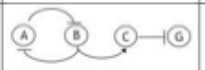
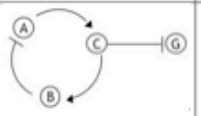
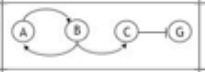
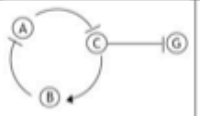
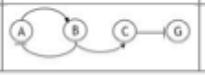


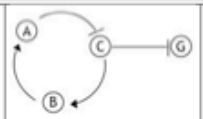
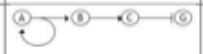
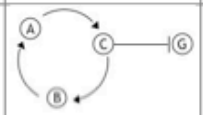
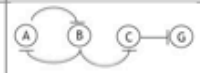
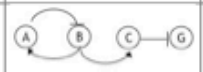
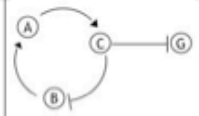
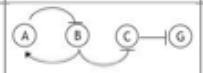
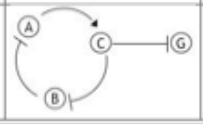
I(G,C) - λ_1N - λ_2T



optimization
flow

Topology is unimportant

Number	Topology	max $I(X,Y)$	max $I(X,Y)$-$\lambda_N N$-$\lambda_T T$
1		1.9913	1.5698
2		1.9992	1.9016
3		1.9915	0.9745
4		1.9950	0.9842
5		1.9998	0.9740
6		1.9934	0.9894
7		1.9985	0.9809
8		2.000	1.8811
9		1.9972	1.8228
10		1.9913	0.9649
11		2.0000	0.9890
12		2.0000	0.9902

Number	Topology	max $I(X,Y)$	max $I(X,Y)$-$\lambda_N N$-$\lambda_T T$
13		1.9983	0.9860
14		1.9999	1.7752
15		1.9994	1.4309
16		1.9995	1.8979
17		1.9952	0.9704
18		1.9996	1.4079
19		1.9978	1.8581
20		2.0000	0.9563
21		1.9955	0.9900
22		2.0000	0.8232
23		2.0000	0.9871
24		1.9912	1.1193

$$\langle N \rangle \leq 100,\ T \leq \text{hours}$$

More bits?

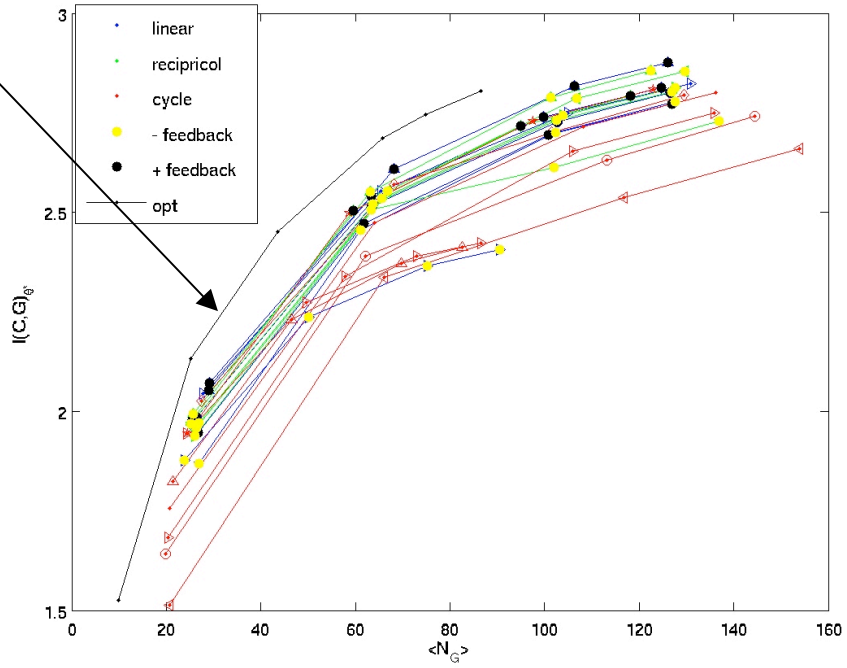
Inputs with >2 states

Since G has no feedback, minimum noise for fixed N_G is Poisson.

$$I(C,G) \le S(G) \le \begin{cases} 0.5 \log N_G + O(N_G^{-1/2}), & \text{for } S(C) = \infty \\ \text{numerical curve}, & \text{for } S(C) = 3 \end{cases}$$

Results:

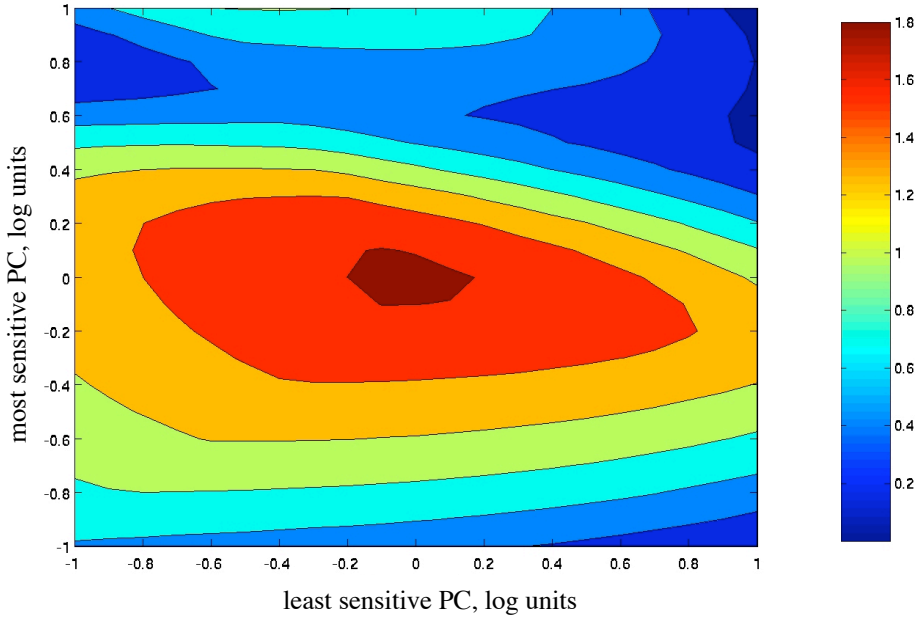
1. All topologies close to the maximum (the worst at N=80 and $S(C)$=3, is about 2.4 bits with $\max_3$=2.7 and $\max_\infty$=3.2).
2. 3-cycles seem to be consistently worse.



$T \le$ hours

Insensitivity to parameters

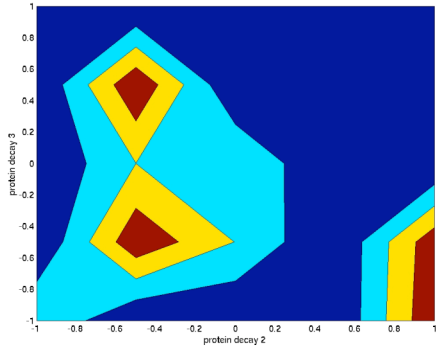
I for topology 1:



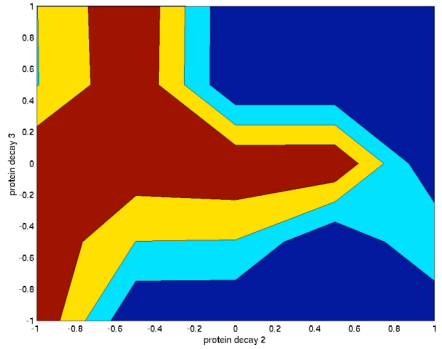
1. Almost 10-fold parameter changes may still lead to I>1.4 bits (holds for some other topologies).
2. High I is generic! No fine-tuning.

Upcoming results (same authors, "Biological Networks: Does Function Follow Form?"):

Multiple maxima (functions)



Plateaus and ridges



• • •

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.