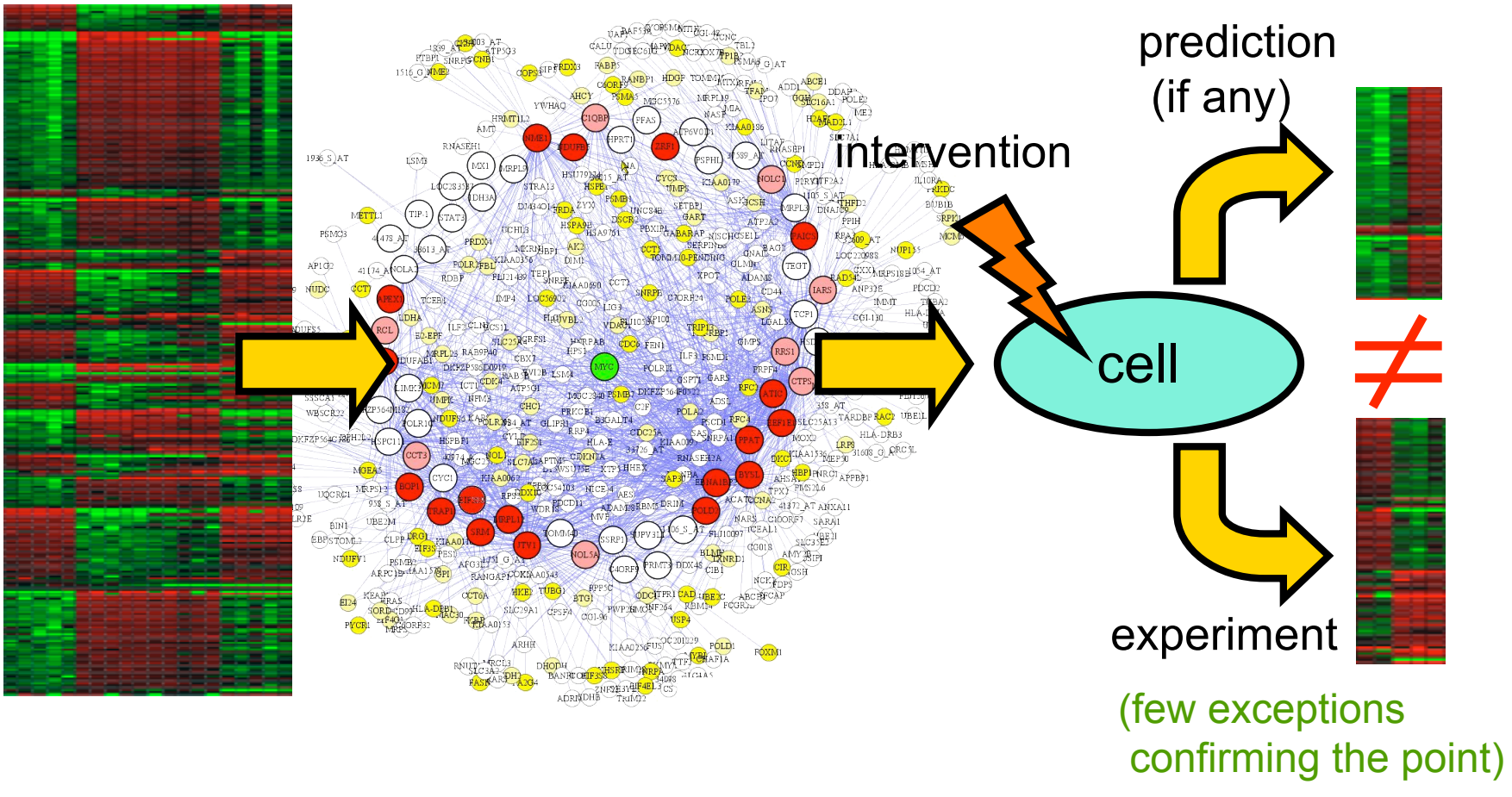
Two notes on transcriptional regulatory networks

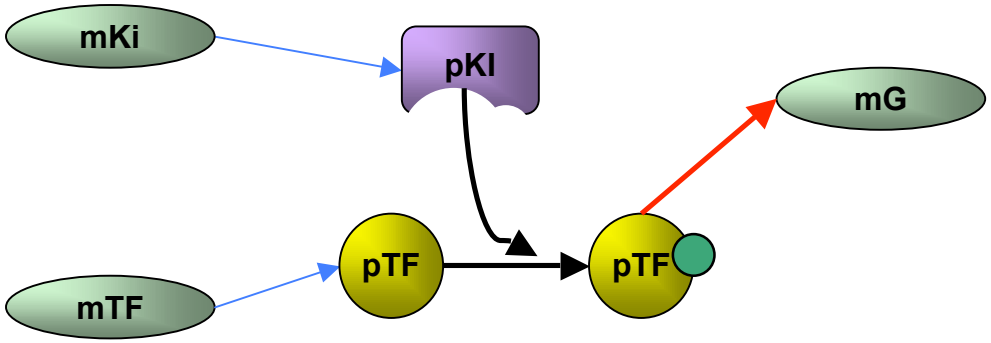
Ilya Nemenman

(CCS-3/LANL)

Reconstructing transcriptional networks



Why?
Maybe incomplete networks…

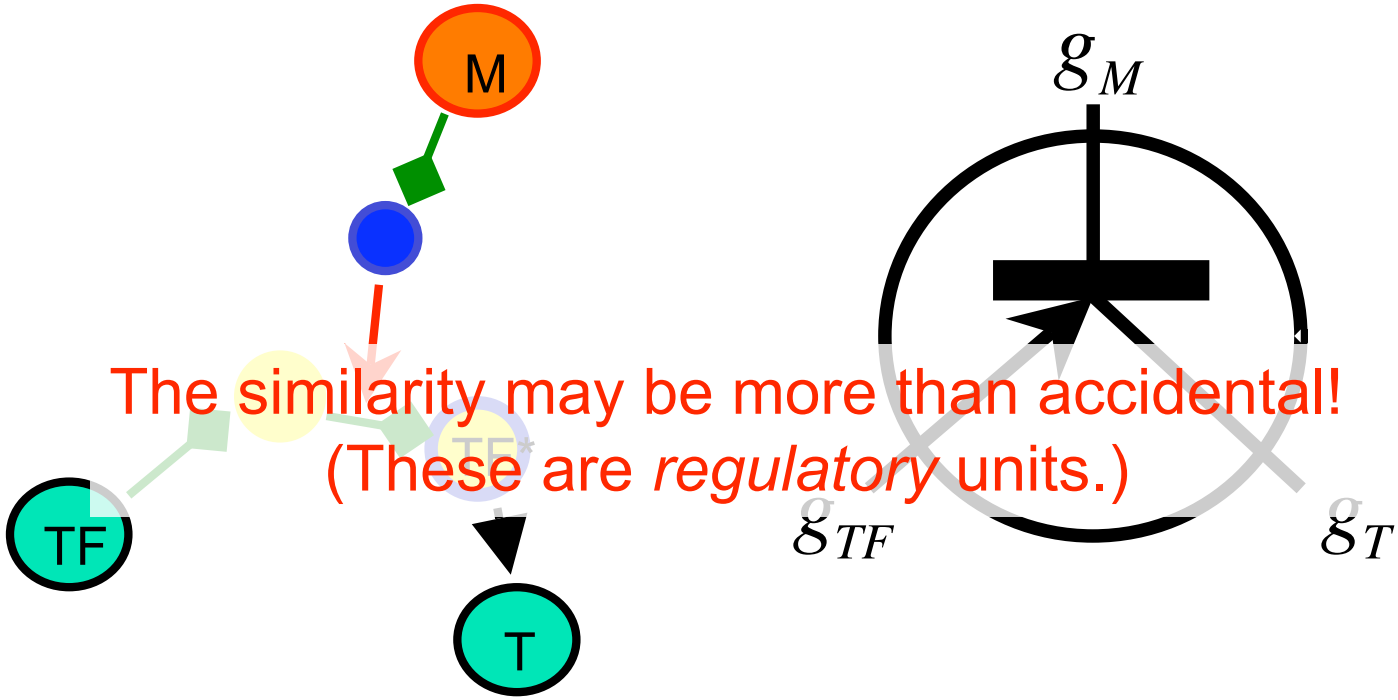


- Context specificity
- Post-translational/post-transcriptional modifications
- Many mRNA constitutively expressed (p53)
- mRNA data carries no information about these modulation events

Or does it?

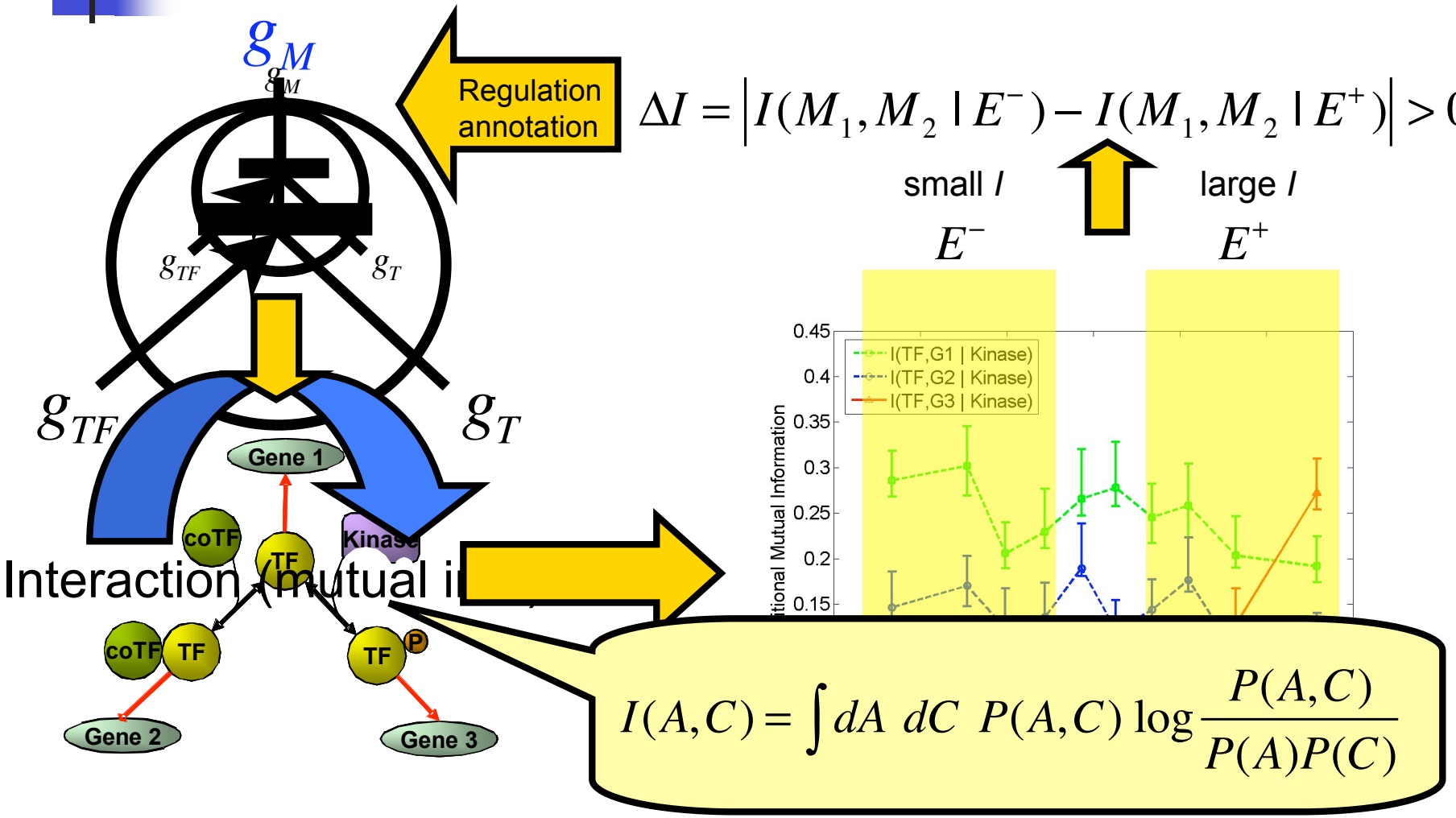
Posttranslational modulation in mRNA data

Solution: Phenotypic and population variability (even in constitutively expressed genes) induces higher order dependencies between TFs, targets, and modulators.

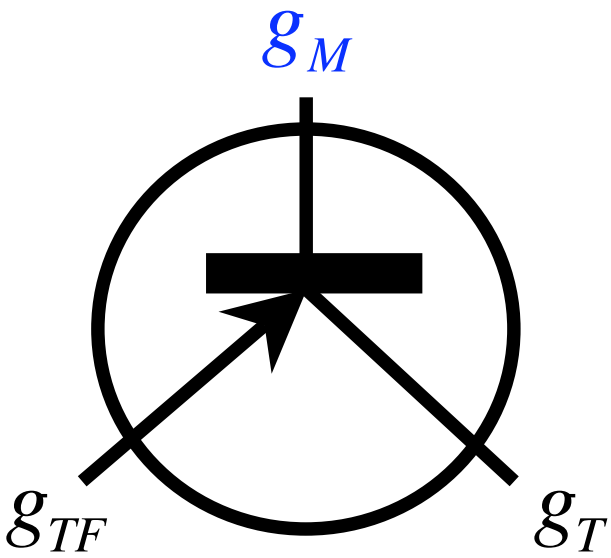


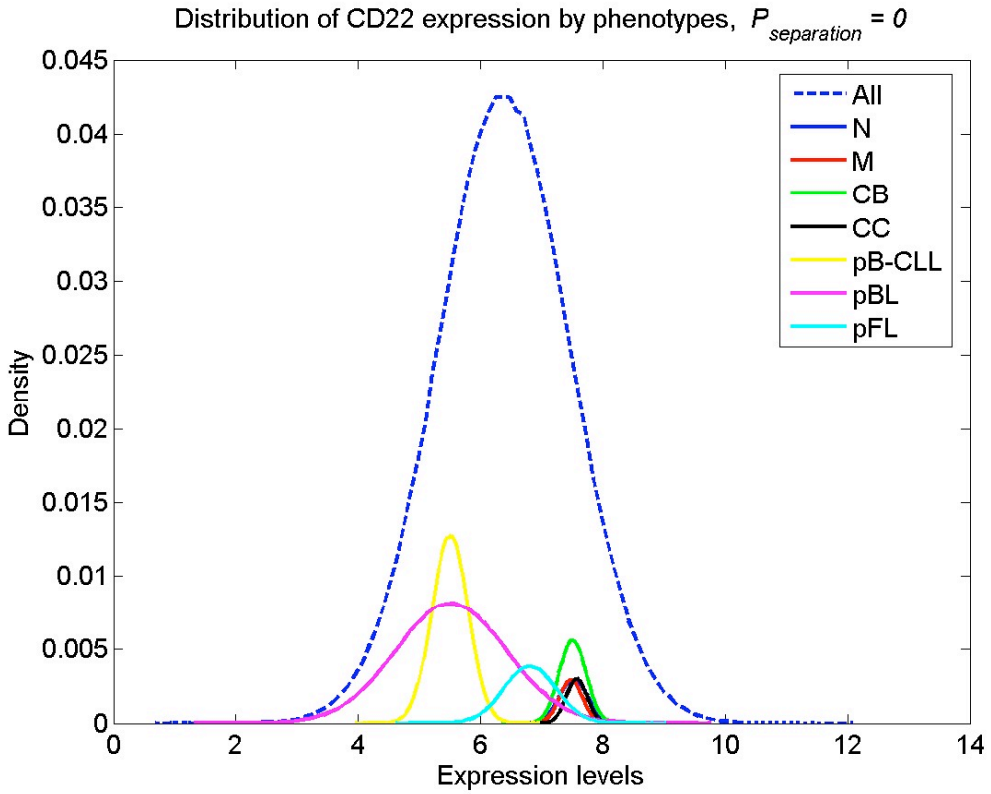
The similarity may be more than accidental!
(These are *regulatory* units.)

Posttranslational modulation: MI signature

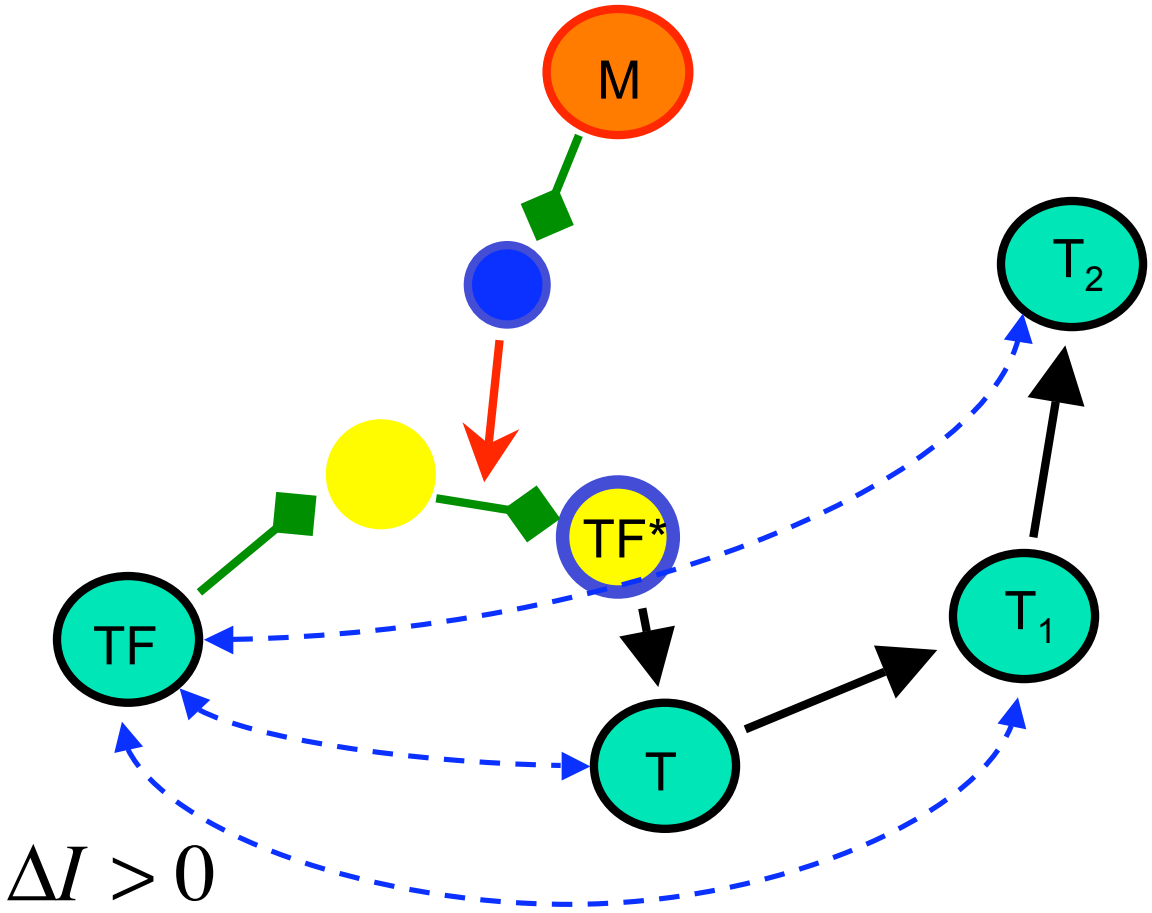


Phenotypic variability of constitutive modulators

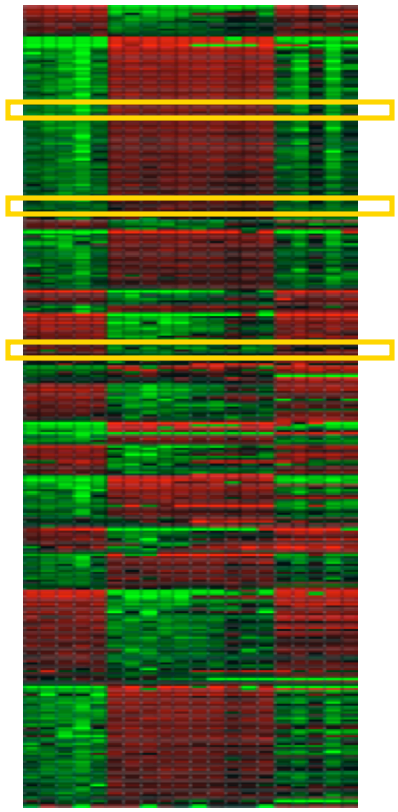




Genuine modulation?



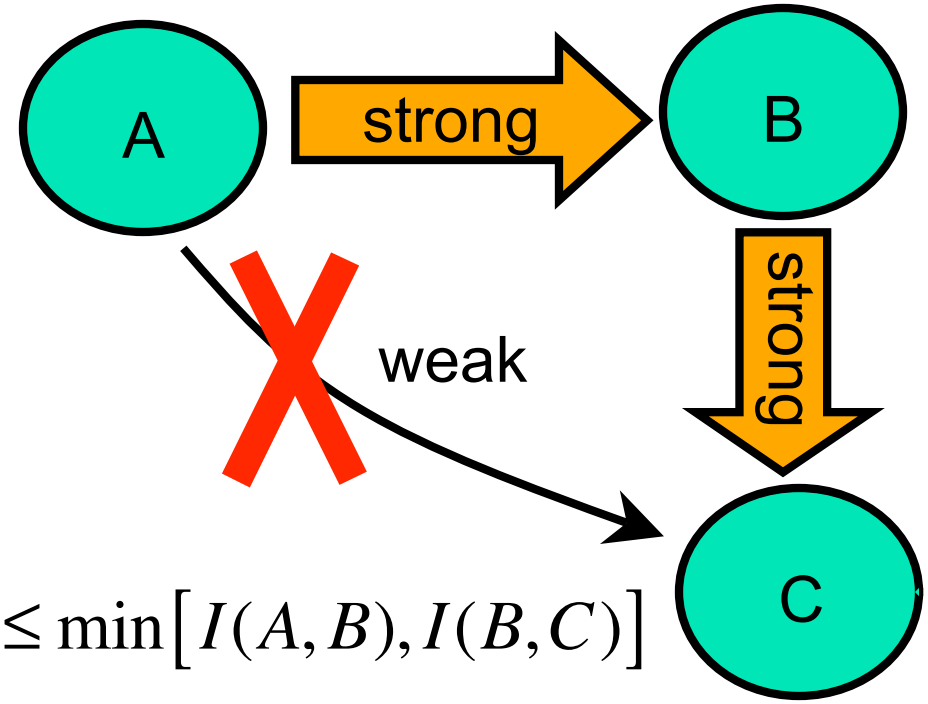
Distinguishing genuine modulation (ARACNE, DPI)



A

B

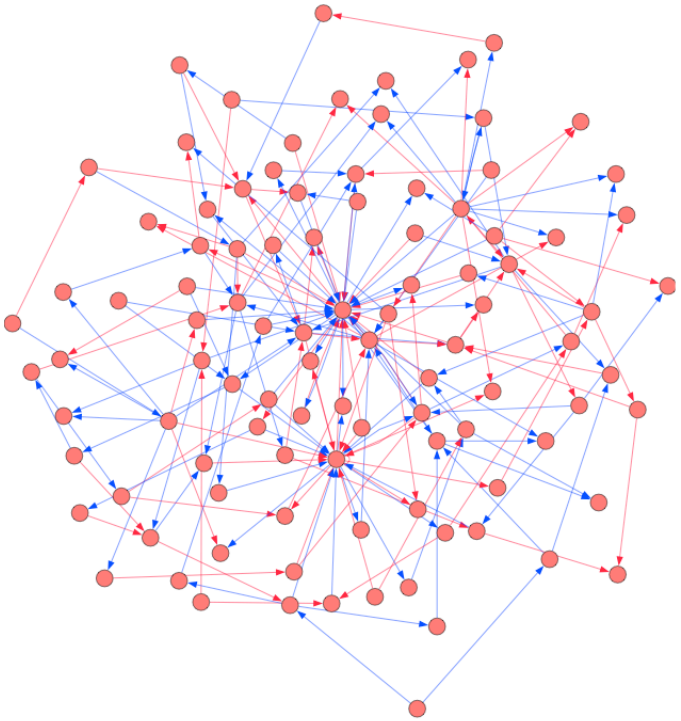
C

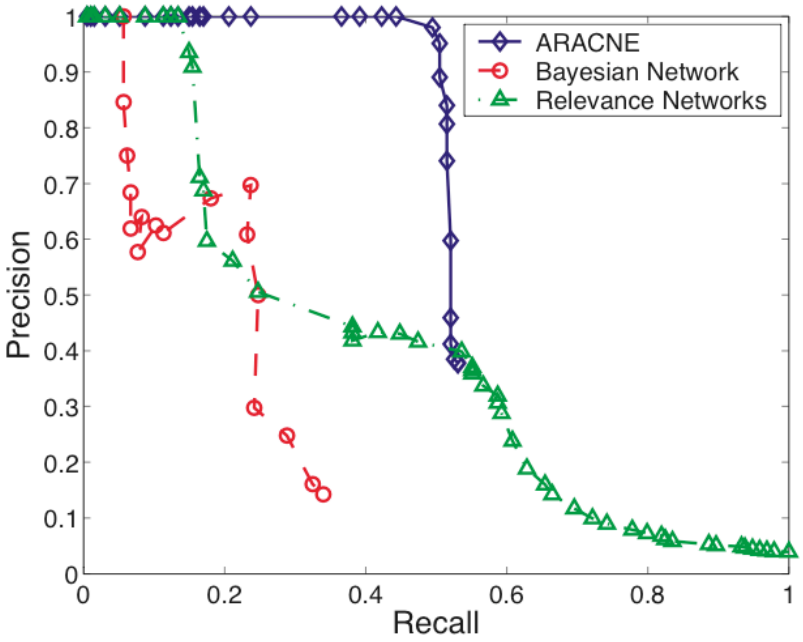


$$I(A,C) \le \min\left[I(A,B), I(B,C)\right]$$

Reparm. invariance; small sample; low complexity; good performance; low false positives.

Synthetic networks

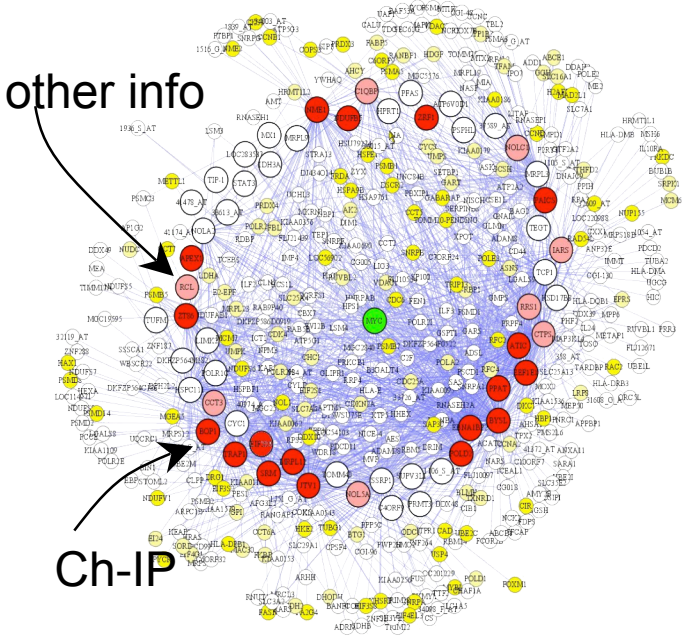




$$\frac{dx_i}{dt} = a_i \prod_j \frac{I_{0,j}^{\nu_j}}{I_j^{\nu_j} + I_{0,j}^{\nu_j}} \prod_j \left(1 + \frac{A_{0,j}^{\nu_j}}{A_j^{\nu_j} + A_{0,j}^{\nu_j}} \right) - b_i x_i$$

B-cell dataset: cMYC network

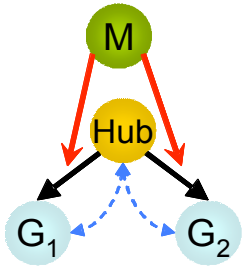
- ~400 arrays (Dalla-Favera et al.)
- No dynamics
- ~250 naturally occurring, ~150 perturbed
- ~25 phenotypes (normal, tumors, experimental perturbations)

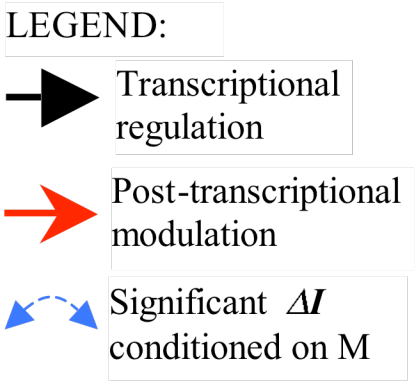


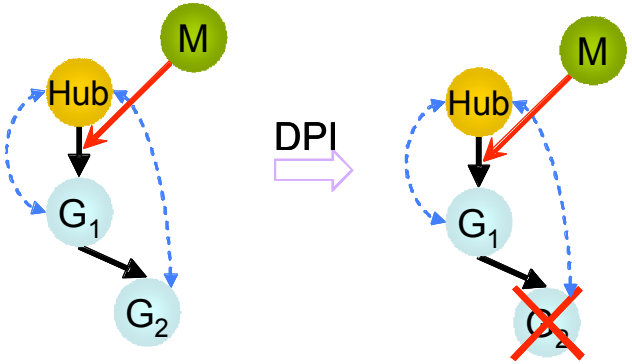
- Protooncogene,
- 12% background binding,
- one of top 5% hubs
- significant MI with 2000 genes

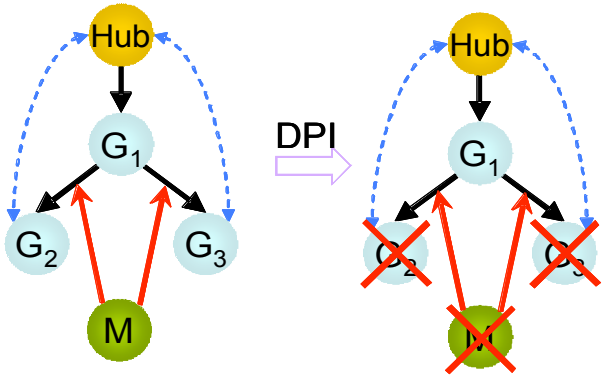
Total interactions: 56
Pre-known: 22
New Ch-IP validated: 11/12

Enforcing irreducibility: ARACNE on a TF-hub





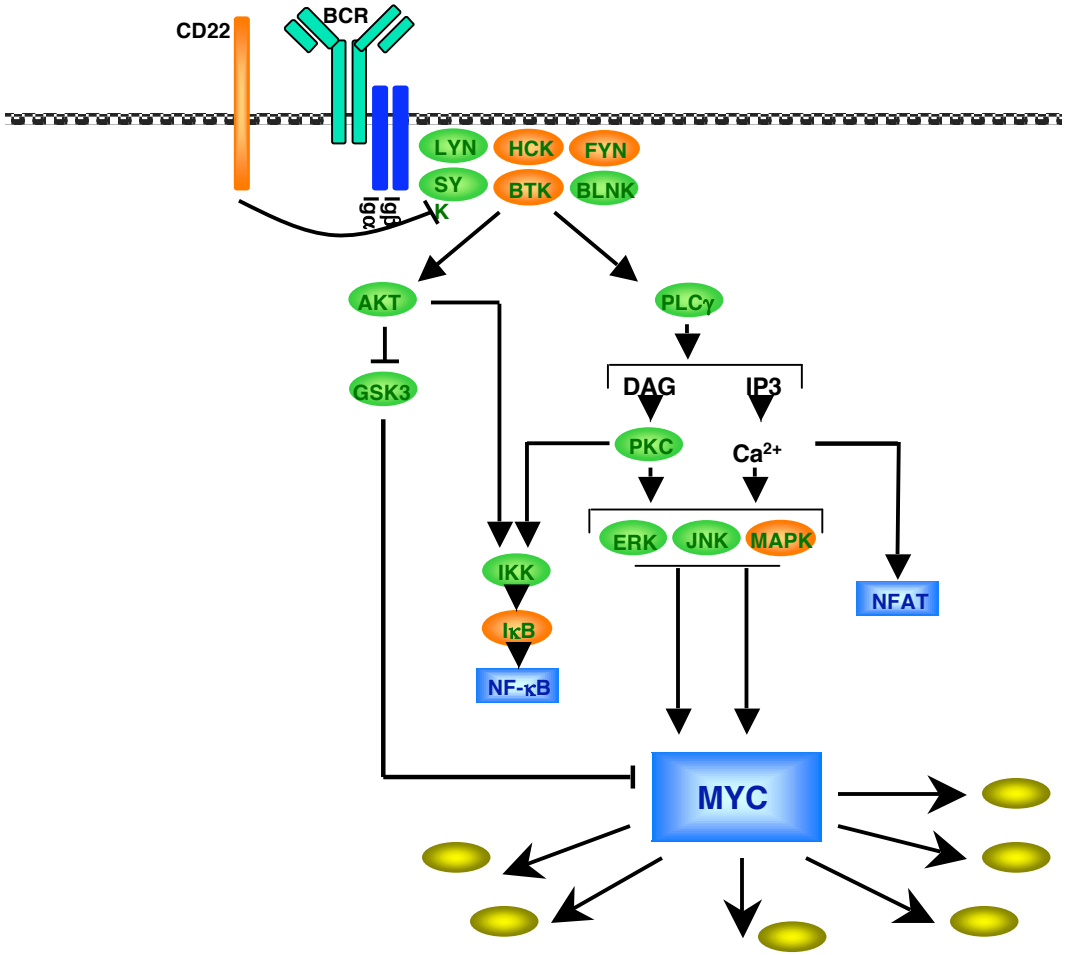




c-MYC modulators

- 1117 candidate modulators
- 100 modulators, 130 targets, 205 interactions
- GO enrichment of the modulator set: kinases, acyltransferases, TFs (all p<5%)
- Modulators in known MYC regulation pathways (e.g., BCR)
- TFs: 15/100, p=1e-6.
- 4/5 TF modulators (e.g., E2F5) with TRANSFAC signatures have binding sites in modulated targets promoter regions.
- Modulators with many (>=4) targets are not-specific (proteolisis, upstream signaling components, receptor signaling molecules).
- Modulators with few (1-2) effected targets are mostly co-TFs, interaction-specific.
- ~1/3 modulators are literature-validated.
- Biochemical validation of some of the predictions in progress.

BCR pathway: Reducibility



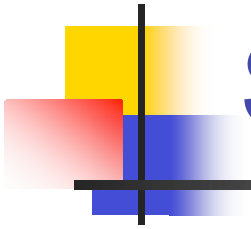
predicted modulators

not in the candidate list

TF's not predicted

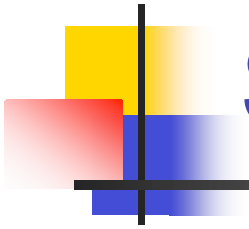
Protein complex

Targets



Summary of part 1

- Post-translational regulatory mechanisms visible from transcriptional data
- Sparseness of species sampling probably not the reason for bad predictability

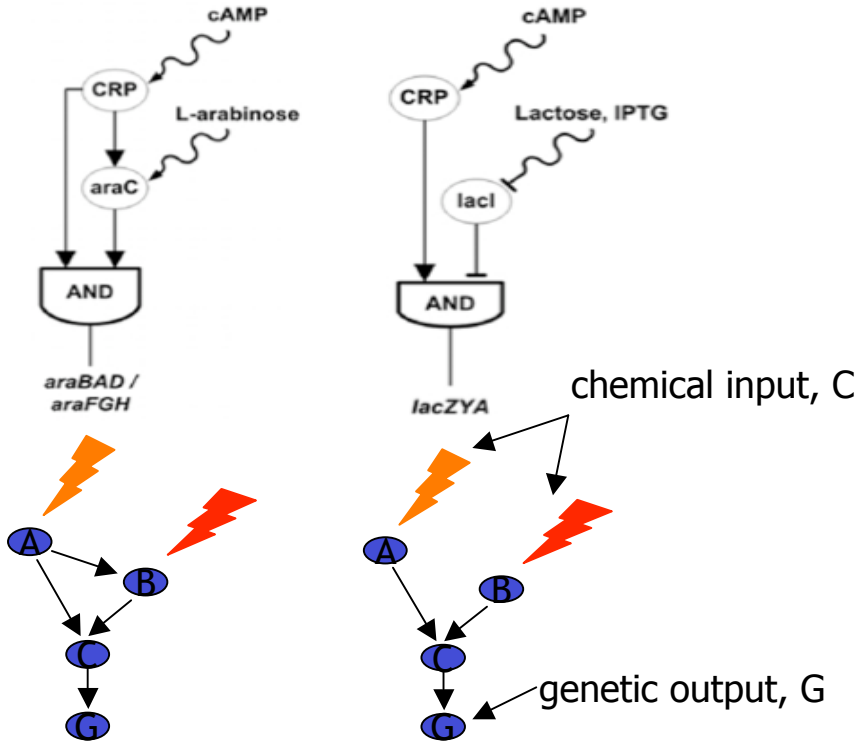


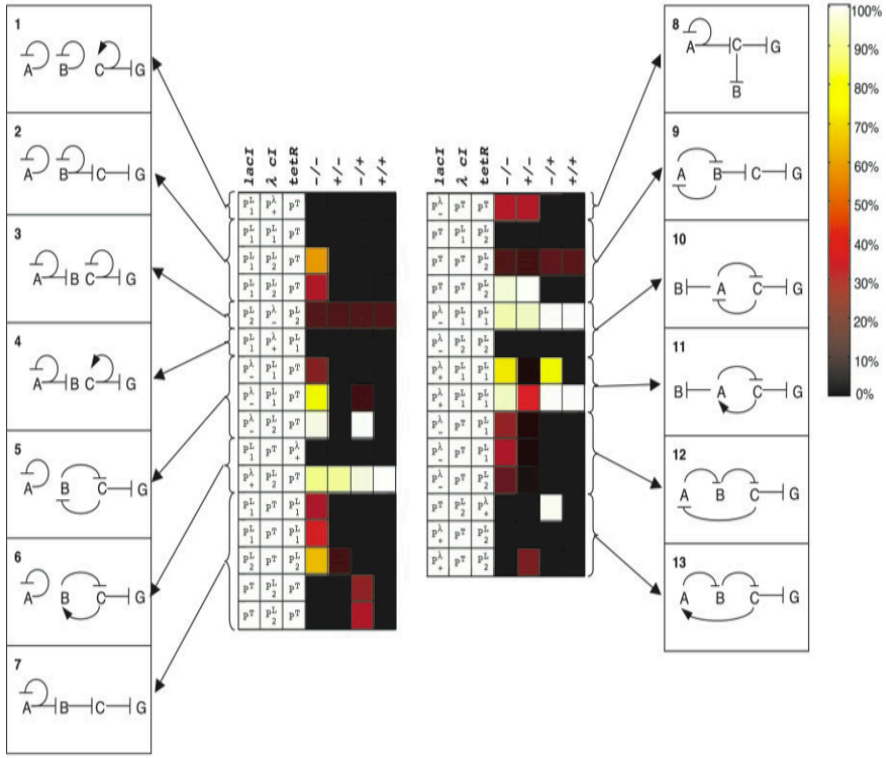
So why low predictive power?

- Maybe: noisiness due to small TF concentration?
- NB: reconstruction models that keep strength of the interaction besides topolgy do better (Leslie et al.)
- Maybe: adjacency matrix description just not enough (soft parameters needed)?
- Maybe: networks can adapt soft parameters to perform the tasks they want to?

Let’s check this for simple topologies!

Experiments





A Sign-Sensitive Delay Element

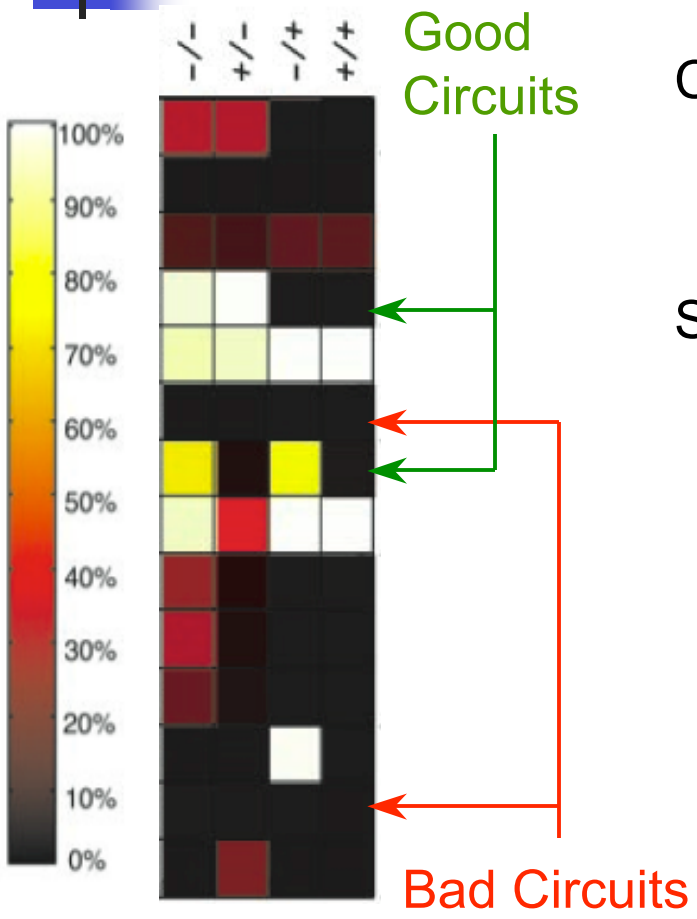
Logic Gates

How to characterize the function of these systems?

From Mangan et al., 2003

From Guet et al., 2002

Function = Information processing



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

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

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

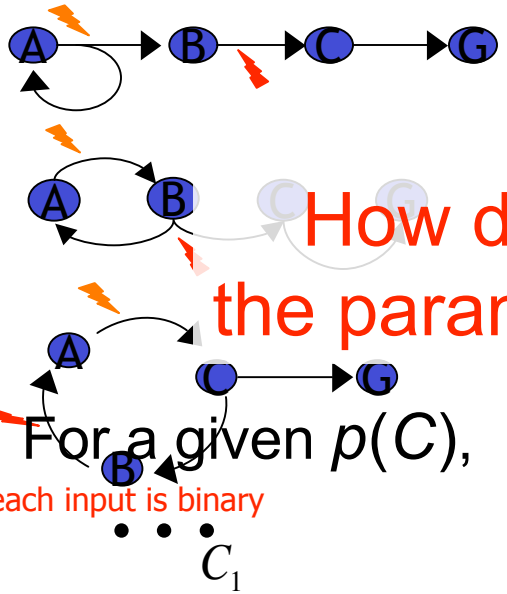
$$0 \le I(C,G) \le \min\{S(C),S(G)\}$$

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?

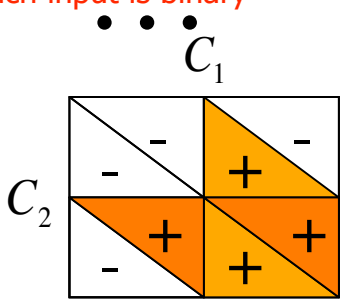
1. For a given topology,

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⊂C

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

...

θ

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

4. And maximize information.

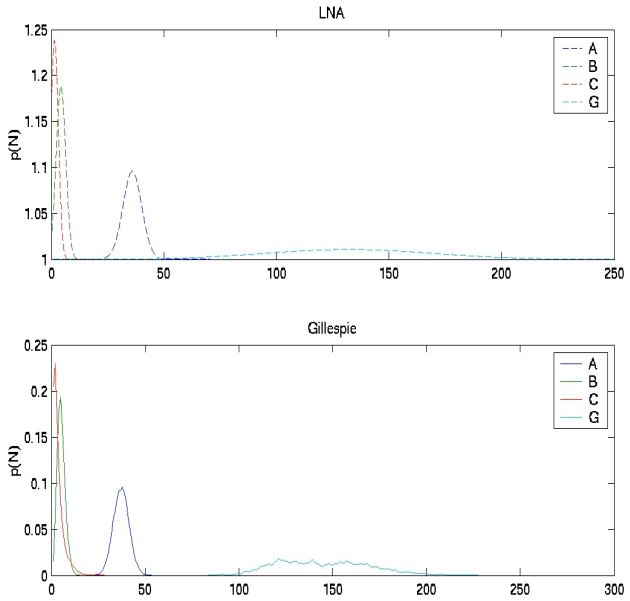
$$\theta = \arg\max_{\theta \subset \left(\begin{smallmatrix}\text{biologically}\\ \text{realistic}\end{smallmatrix}\right)} I(C,G)$$

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

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

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

Calculating P(g|c): linear noise approximation (LNA)



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/Ω.
3. Contrary to Baras et al, 1996, LNA is sound if $A = A(g)$, $B = B(g)$
4. For eig(S) of very different sizes, need to adiabatically integrate out the fast modes.

Model + parameters: details

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

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

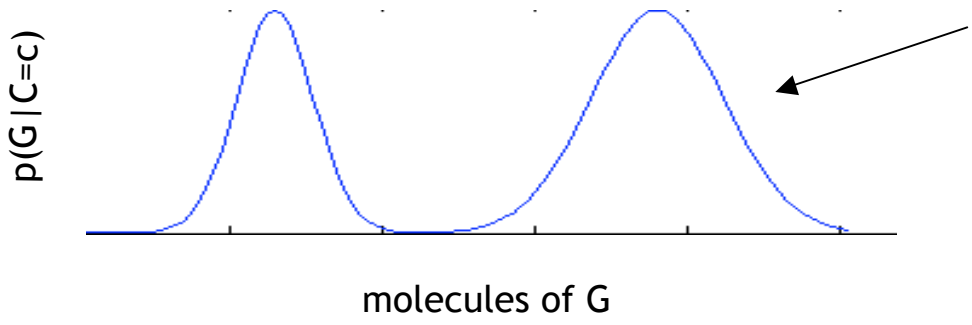
$$\text{Excitation: } \alpha(\phi_j, s_j) = a\frac{(\phi_j / s_j)^n}{K^n + (\phi_j / s_j)^n}$$

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

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

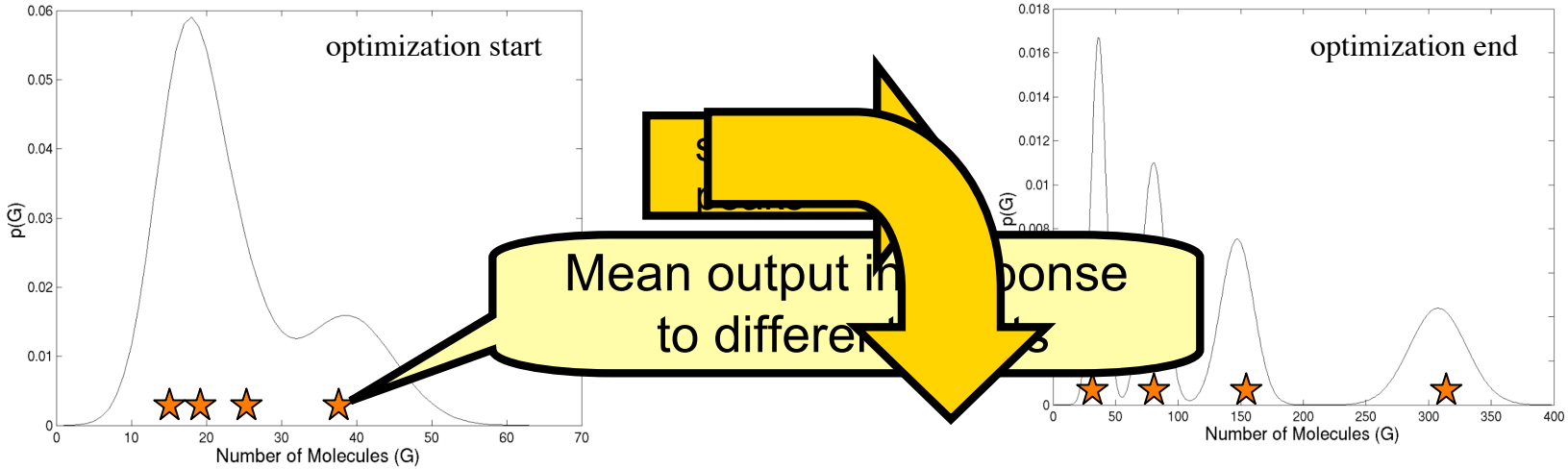
Up to 22 parameters



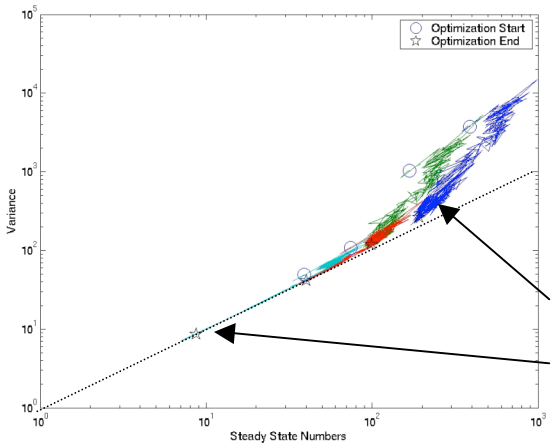
Example: two distinct steady states with Gaussian noise; P(each state | C=c)=const; no stochastic stability analysis.

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



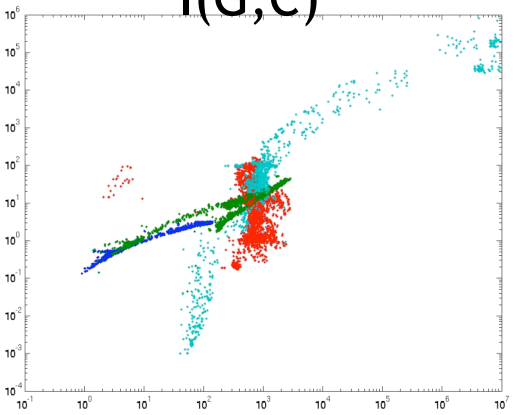
decreasing the reporter variance to the Poisson limit (low pass filtering upstream noise by slow reporter); variance of the other species may be sub-Poisson (negative feedback)



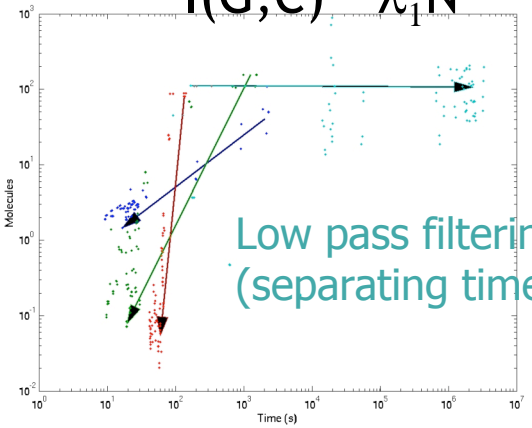
other species reporter

Achieving 2 bits (T, #)

I(G,C)



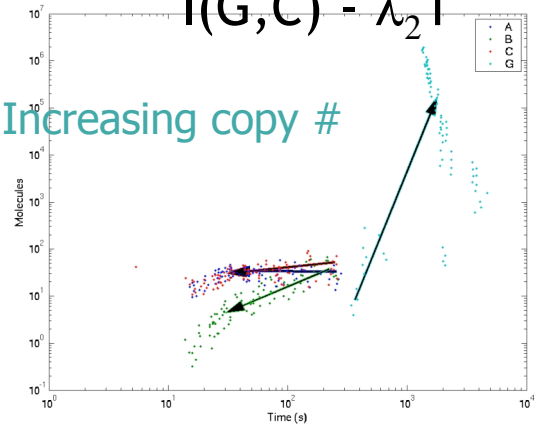
I(G,C) - λ_1N



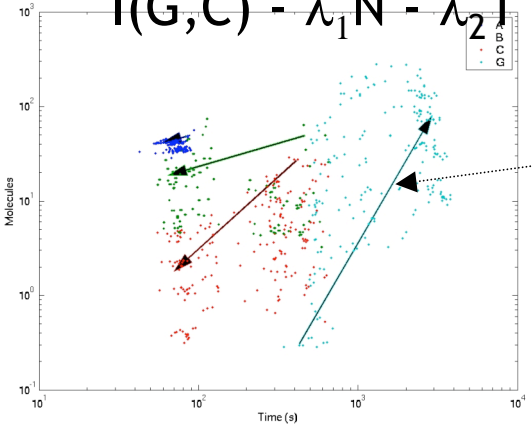
Low pass filtering
(separating time scales)

I(G,C) - λ_2T

Increasing copy #

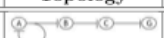
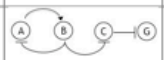
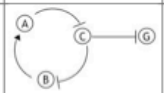
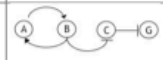
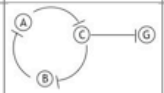
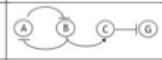
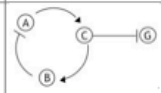
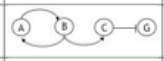
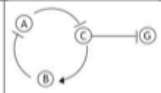
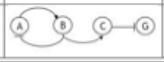


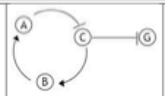
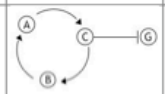
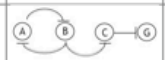
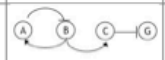
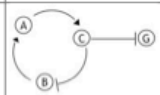
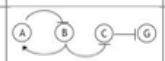
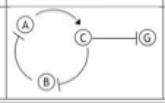
I(G,C) - λ_1N - λ_2T



optimization
flow

Adaptation makes any topology functional

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

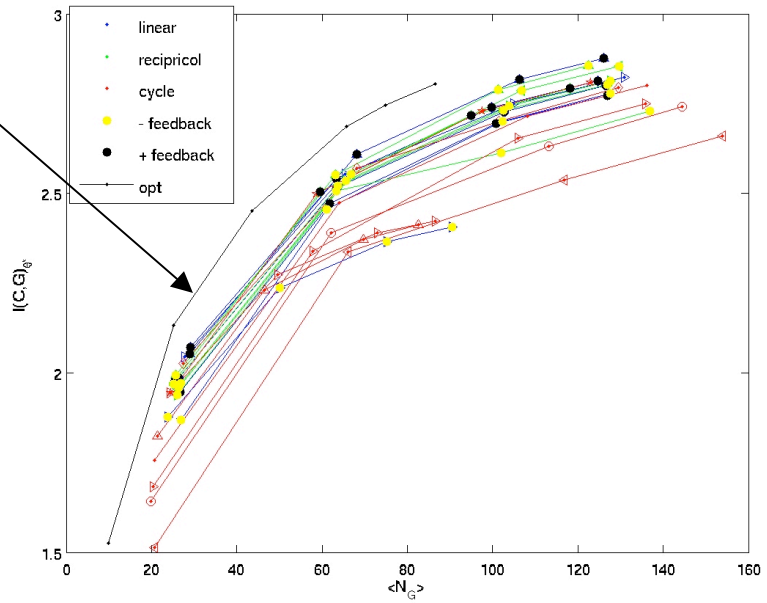
Number	Topology	max $I(X,Y)$	max $I(X,Y)$-$\lambda_N N$-$\lambda_T T$
13		1.9983	0.9860
14		1.9999	1.7752
15		1.9994	1.4309
16		1.9995	1.8979
17		1.9952	0.9704
18		1.9996	1.4079
19		1.9978	1.8581
20		2.0000	0.9563
21		1.9955	0.9900
22		2.0000	0.8232
23		2.0000	0.9871
24		1.9912	1.1193

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

More bits? Other $P(c)$? More states of c?

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

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

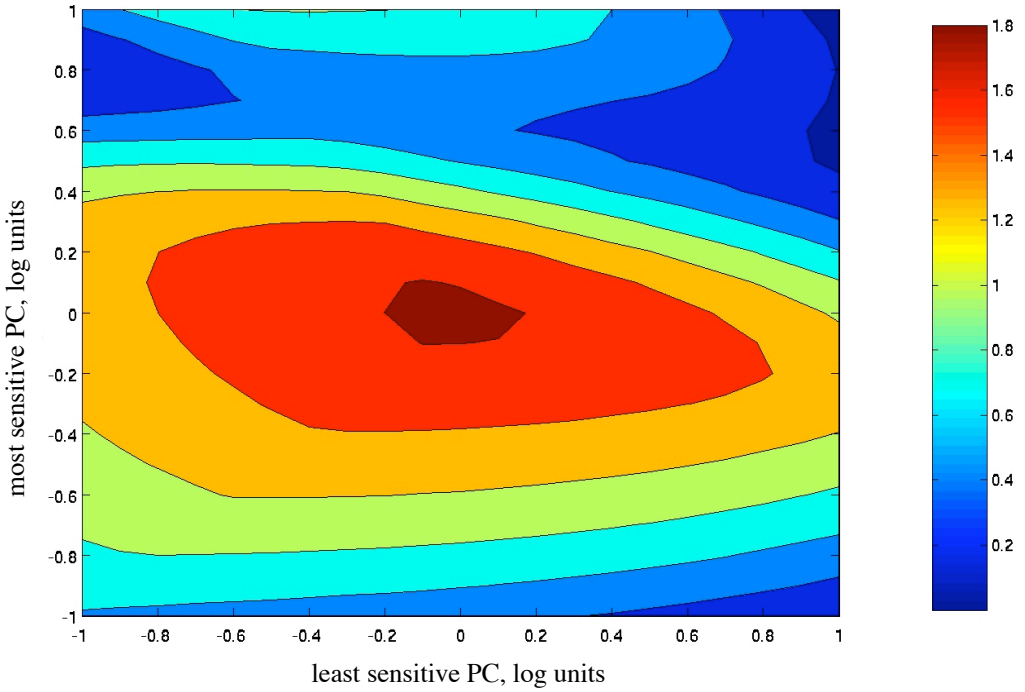


$T \le$ hours

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.

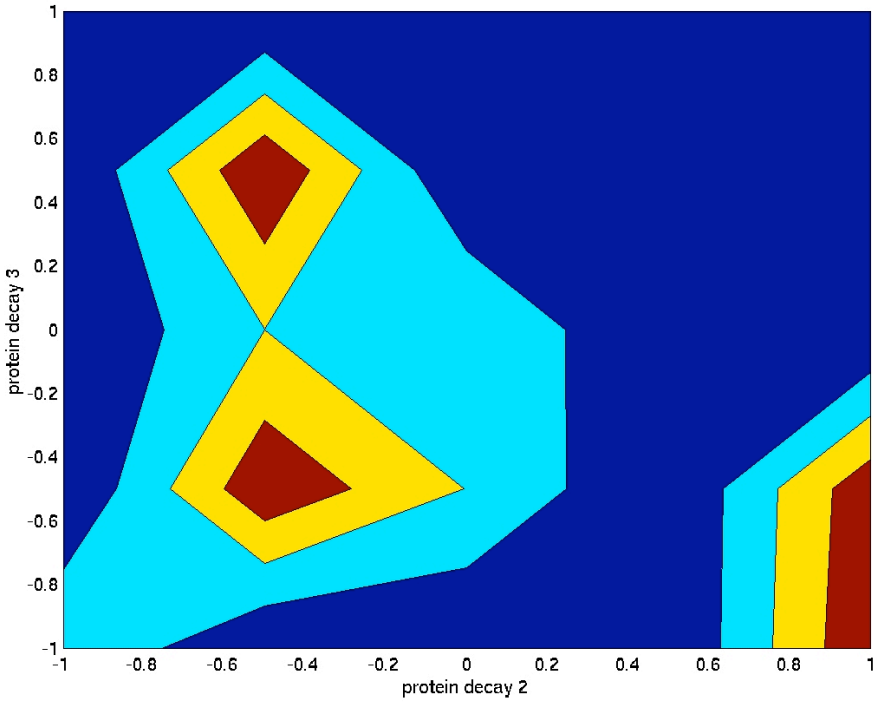
Insensitivity to parameters



(e.g., *I* for topology 1)

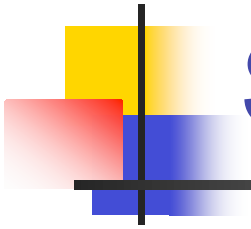
1. Almost 10-fold parameter changes may still lead to *I*>1.4 bits (holds for some other topologies).
2. High *I* is generic! No fine-tuning.

Multiple functions



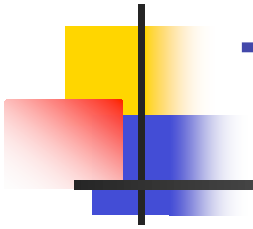
Function:

input	reporter rank peak 1	reporter rank peak 2
00	2	2
01	1	1
10	4	3
11	3	4



Summary of part 2

- Generically, can send many bits through simple networks (cross-talk?)
- Noise in adapted circuits is not the reason for low predictability
- Circuits can adapt (e.g., after a knock-out) to perform a function reliably
- Circuits can perform many functions (without “hard” changes)
- “Soft” parameters must be known to specify networks, or must focus on functions



The work was done by…

- Post-translational regulation: Kai Wang, Adam Margolin, Andrea Califano et al., Katia Basso, Riccardo Dala-Favera et al., LANL comp-bio team
- Information processing by circuits: Etay Ziv, Chris Wiggins