Genetic network complexity: weights matter more than topology

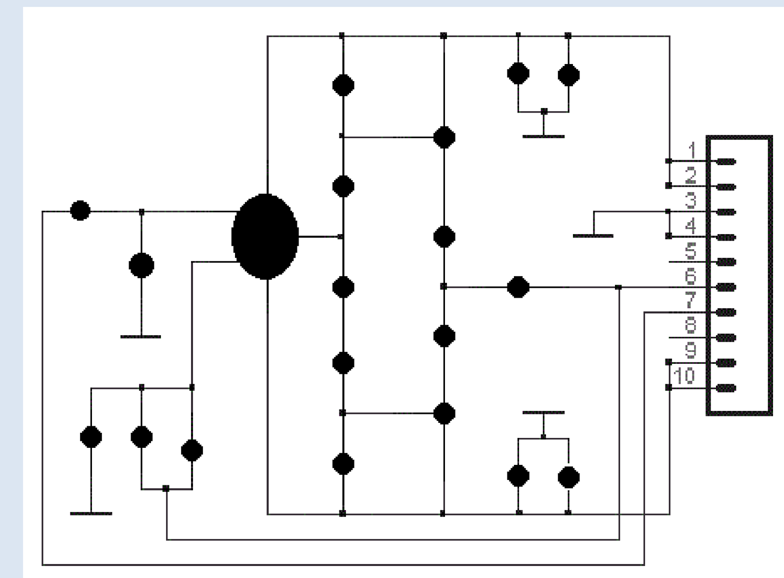
Mikhail Tikhonov and William Bialek

Joseph Henry Laboratories of Physics, Lewis-Sigler Institute for Integrative Genomics, Princeton University

1. Appropriate level of description?

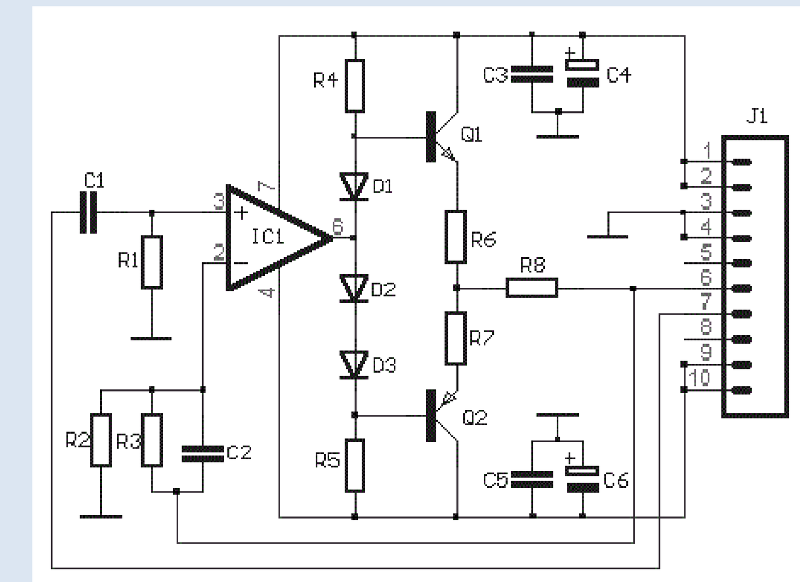
Can we understand genetic networks like we understand electronic circuits? Not every detail matters, but knowing only topology is insufficient.

Goal: build a toy model where the appropriate level of description can be constructed explicitly. How much do microscopic quantitative details matter?



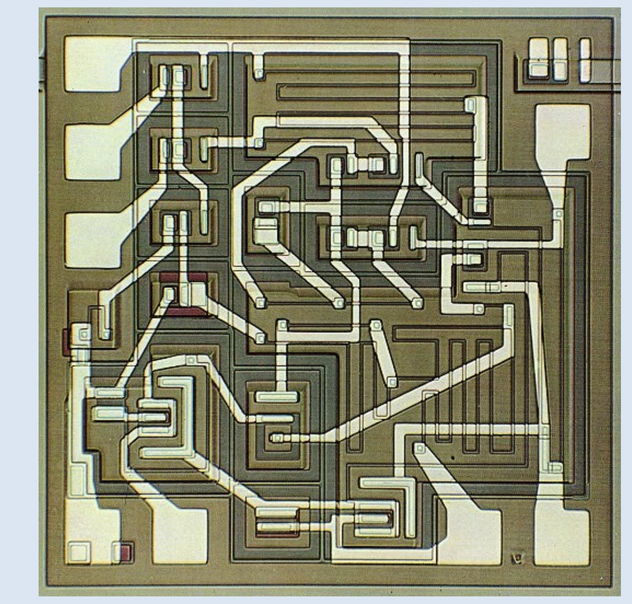
Just topology: too coarse

⊂



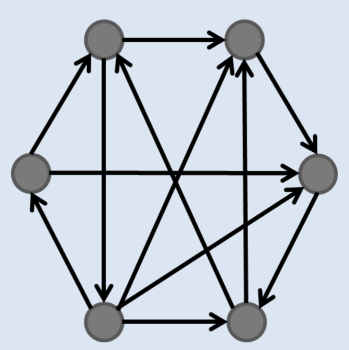
Only relevant detail

⊂



Too microscopic

2. Model: weighted graphs and complexity

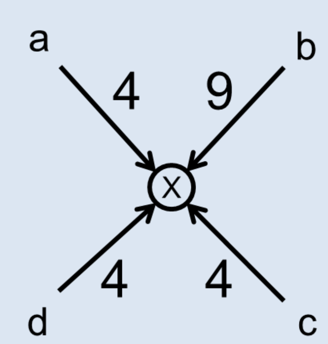


Genes: binary variables; interactions have variable strength.
A gene is activated if its inputs exceed a threshold:

$$s_i = \mathrm{sgn}\left(\sum_{j\to i} J_{ij} s_j + H_i\right) \qquad (1)$$

Capacity of a network: number of solutions (= number of cell types it can encode)

Each node implements a Boolean function from a *finite* set with a *non-arbitrary measure*.



⟺ $x = f(a,b,c,d)$

For a given topology, we can enumerate all of its *non-equivalent* and *equiprobable* (!) realizations as a weighted graph.

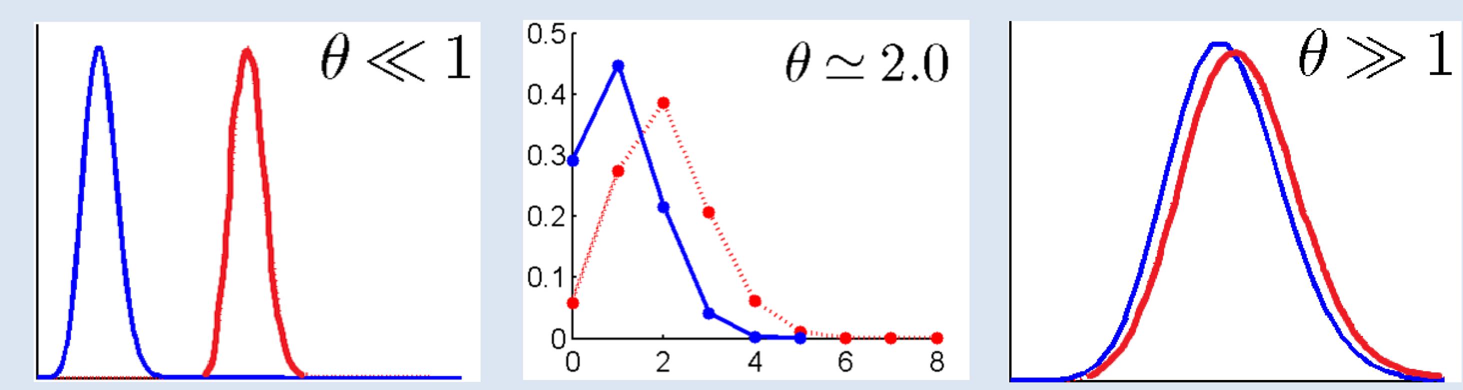
Define *complexity* as the diversity of possible causal relations in the graph.

Links are satisfied or frustrated. Define an *active* link as a link whose satisfied state is essential for equation (1) to hold. Each solution defines a binary sequence: the pattern of active links. Define diversity of a set of sequences as the length of the shortest path connecting all of them on a hypercube (traveling salesman).

How is complexity affected by the choice of topology vs. the choice of weights?

3. Conclusions: weights matter more

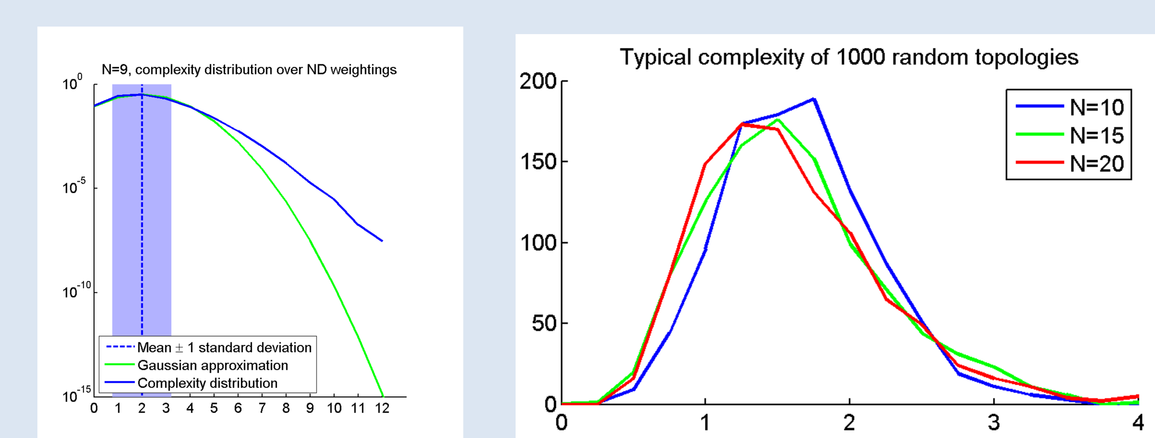
1. Relative importance of weights vs. topology is of order 1.



Capacity distributions over all choices of weights for 2 fixed topologies.
Left/right: capacity determined by topology/weights. Actual data is in the middle.
θ defined as $[\mathrm{mean}_{topology}(\sigma_{weights})]/[\sigma_{topology}(\mathrm{mean}_{weights})]$, σ is standard deviation

2. Optimal weights outperform optimal topology
(for 85% of topologies with N≤10)

3. Larger networks are not automatically more complex



Capacity distributions, both over weights and over topologies, are heavy-tailed. Typical capacity (random choice of weights) does not scale with network size.

For an information-processing network:

1. Topology and weights have effects of the same order.

2. High-complexity graphs operate in a non-generic parameter regime.

3. Evolving weights is a better strategy than changing topology / adding nodes

This work was supported by the Center for Science of Information (CSoI), NSF Science and Technology Center. CCF-0939370