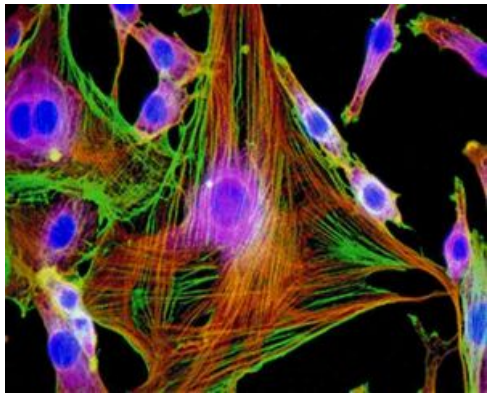
thermodynamic limits in cellular information processing

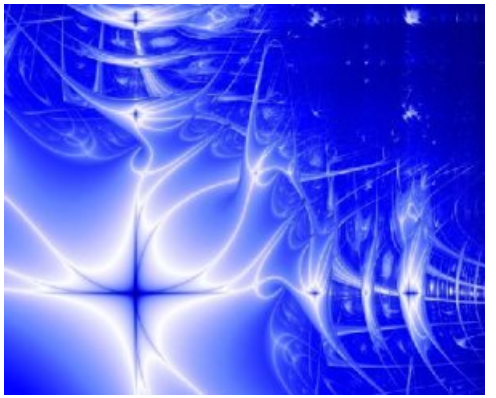
jeremy gunawardena



department of systems biology
harvard medical school
200 longwood avenue
boston, ma 02115

jeremy@hms.harvard.edu

http://vcp.med.harvard.edu/



ICTS, bangalore

10 december 2020

- hopfield and his barrier
- the linear framework
- gene regulation
- the hill function as hopfield barrier
- non-equilibrium challenges

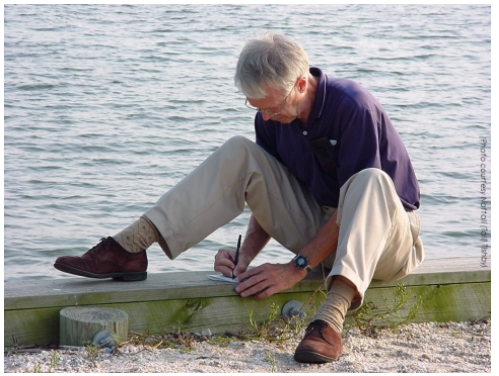
- hopfield and his barrier
- the linear framework
- gene regulation
- the hill function as hopfield barrier
- non-equilibrium challenges

energy expenditure in error reduction

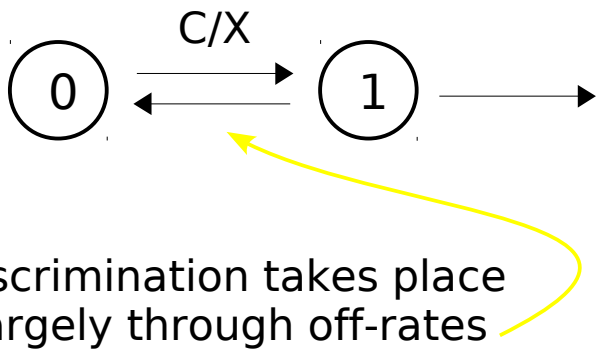
Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity

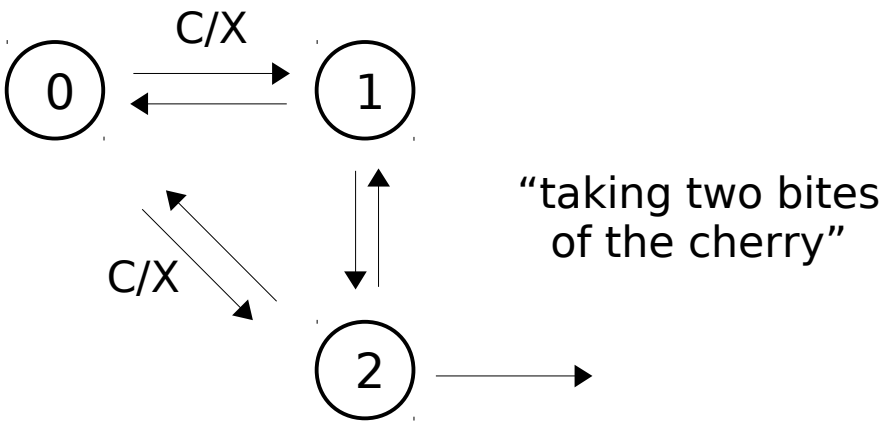
(protein synthesis/DNA replication/amino-acid recognition)

J. J. HOPFIELD *Proc. Nat. Acad. Sci. USA*
Vol. 71, No. 10, pp. 4135–4139, October 1974



discrimination between correct (C) and wrong (X) substrate





Wong, Amir, Gunawardena, “ "*Energy-speed-accuracy relation in complex networks for biological discrimination*"”, **Phys Rev E** 98:012420 2018

the hopfield barrier

given any information processing task,

if the biochemical mechanism implementing that task is operating in steady state at thermodynamic equilibrium, then there is a **fundamental upper bound** to how well that task can be undertaken;

the only way to exceed this **hopfield barrier** is to dissipate energy and maintain the mechanism in a non-equilibrium steady state.

beyond this barrier, **tradeoffs** occur in how energy is used to achieve different kinds of functionality, such as accuracy, speed, efficiency, etc.

Estrada, Wong, DePace, Gunawardena, "*Information integration and energy expenditure in gene regulation*", **Cell** 166:234-44 2016.

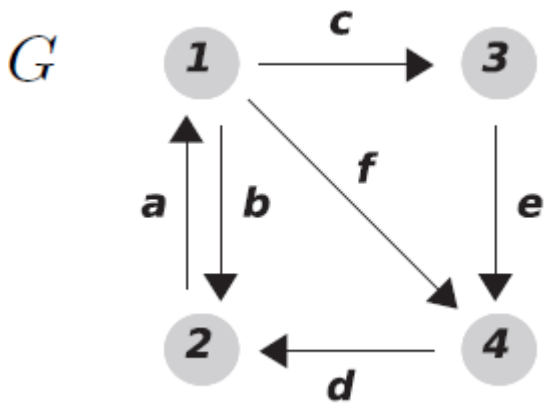
- hopfield and his barrier
- the linear framework
- gene regulation
- the hill function as hopfield barrier
- non-equilibrium challenges

linear framework for timescale separation

finite graphs with labelled, directed edges – representing “fast” components

labels encode interaction with external “slow” components in the environment

1-dimensional chemistry



$$\begin{pmatrix} -(b+c+f) & a & 0 & 0 \\ b & -a & 0 & d \\ c & 0 & -e & 0 \\ f & 0 & e & -d \end{pmatrix}$$

Laplacian matrix

$$\frac{d\mathbf{x}(t)}{dt} = \mathcal{L}(G).\mathbf{x}(t)$$

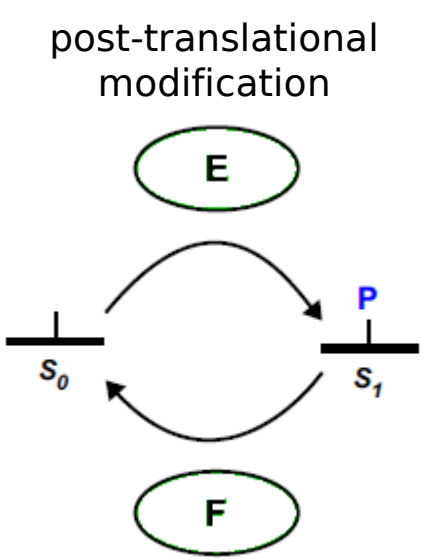
$$\mathbf{x}_1(t) + \cdots + \mathbf{x}_n(t) = \mathbf{x}_{tot}$$

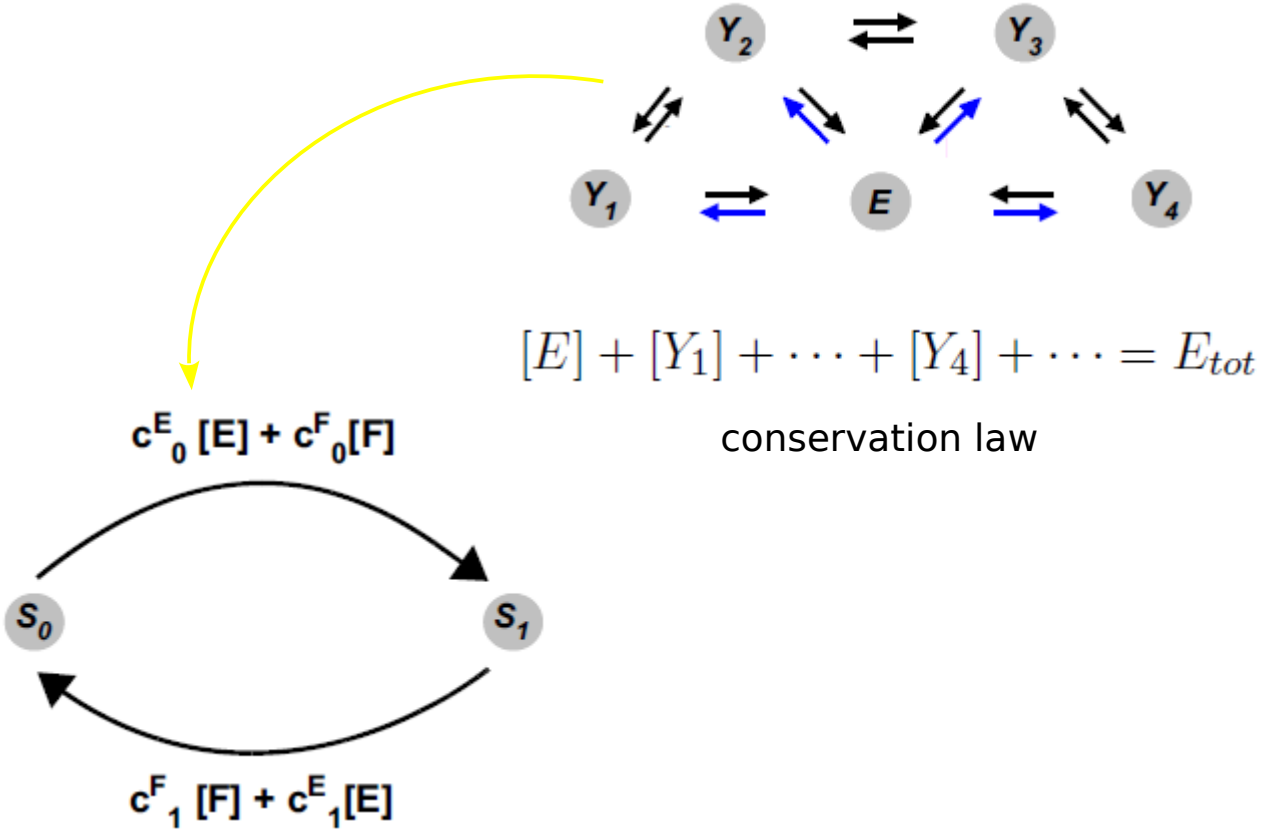
$$\mathcal{L}(G).\mathbf{1} = 0$$

Gunawardena, “*A linear framework for time-scale separation in nonlinear biochemical systems*", **PLoS ONE** 7:e36321 2012

enzyme-catalysed biochemistry

steady-state analysis of nonlinear biochemical networks



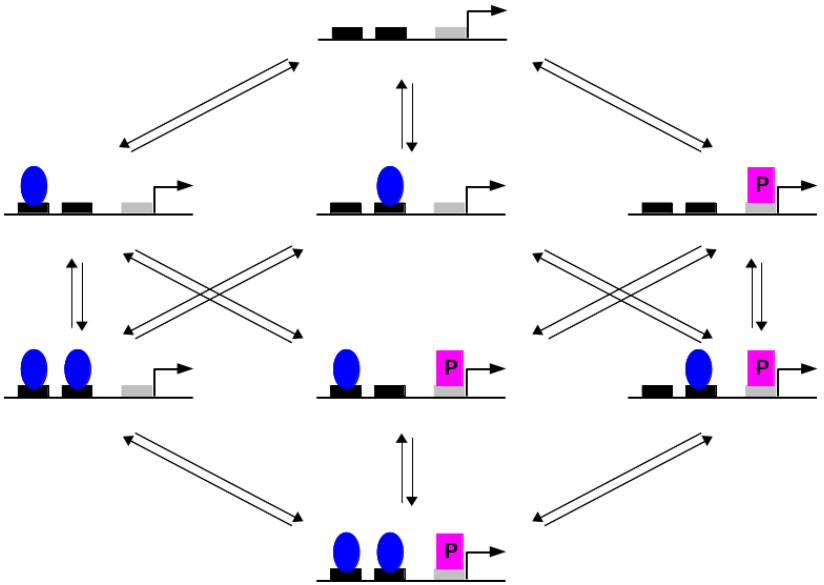


Thomson, Gunawardena, **Nature** 460:274-7 2009; Xu, Gunawardena, **J Theor Biol** 311:139-52 2012; Dasgupta, Croll, Owen, Vander Heiden, Locasale, Alon, Cantley, Gunawardena, **J Biol Chem** 289:13010-25 2014; Nam, Gyori, Amethyst, Bates, Gunawardena, **PLoS Comp Biol** 16:e1007573 2020

markov processes

external (“slow”) components form “reservoirs”

graph is the infinitesimal generator of a Markov process



$$\ell(i \to j) = \lim_{\Delta t \to 0} \frac{\Pr(X(t+\Delta t) = j \mid X(t) = i)}{\Delta t}$$

$$\frac{d\mathbf{p}}{dt} = \mathcal{L}(G).\mathbf{p}$$

$$\mathbf{p}_1 + \cdots + \mathbf{p}_n = 1$$

master equation

Mirzaev, Gunawardena, “*Laplacian dynamics on general graphs*", **Bull Math Biol** 75:2118-49 2013

steady states and the MTT

provided G is strongly connected $\quad \ker \mathcal{L}(G) = \langle\, \rho(G) \,\rangle$

matrix-tree theorem (MTT) gives

$$\rho_i(G) = \sum_{T \in \Theta_i(G)} \left(\prod_{j \to k \in T} \ell(j \to k) \right)$$

$\Theta_i(G) =$ spanning trees of G rooted at microstate i

steady-state probabilities of microstates

$$\mathbf{p}_i^* = \frac{\rho_i(G)}{\rho_1(G) + \cdots + \rho_N(G)}$$

at thermodynamic equilibrium

G must be **reversible** – if $i \to j$ then also $j \to i$

detailed balance – for any pair of reversible edges $i \leftrightharpoons j$

$$\mathbf{p}_i^* \ell(i \to j) = \mathbf{p}_j^* \ell(j \to i)$$

$$\ker \mathcal{L}(G) = \langle\, \mu(G) \,\rangle$$

$$1 = i_1 \leftrightharpoons i_2 \leftrightharpoons \cdots \leftrightharpoons i_{k-1} \leftrightharpoons i_k = i$$

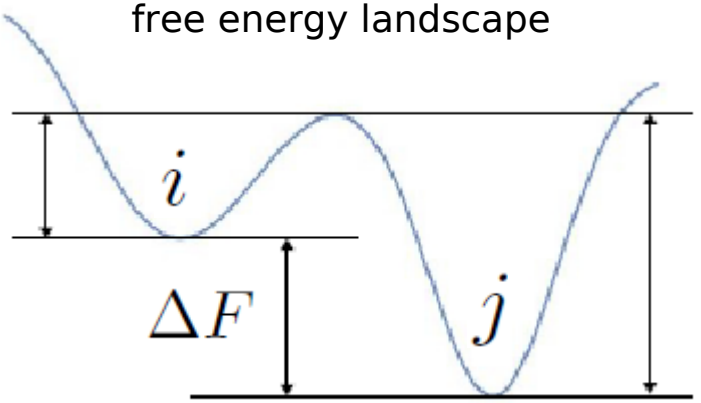
$$\mu_i(G) = \left(\frac{\ell(i_1 \to i_2)}{\ell(i_2 \to i_1)}\right) \cdots \left(\frac{\ell(i_{k-1} \to i_k)}{\ell(i_k \to i_{k-1})}\right)$$

path independence

equilibrium statistical mechanics

$$i \leftrightharpoons j$$

free energy landscape



$$\frac{\ell(i \to j)}{\ell(j \to i)} = \exp\left(\frac{\Delta F}{k_B T}\right)$$

partition function for the grand canonical ensemble

$$\mathbf{p}_i^* = \frac{\mu_i(G)}{\mu_1(G) + \cdots + \mu_N(G)}$$

e pluribus unum

Kirchhoff 1848

GRAPH THEORY
Tutte 1948

COMPUTER SCIENCE

ECONOMICS
Bott &
Mayberry 1954

ELECTRICAL ENGINEERING

MARKOV PROCESSES

ENZYME KINETICS
King &
Altman 1956

LARGE DEVIATION THEORY
Wentzell & Freidlin 1970

QUANTUM FIELD THEORY

BIOPHYSICS
Terrell Hill 1966

NON-EQUILIBRIUM STATISTICAL MECHANICS
Schnakenberg 1976

CHEMICAL PHYSICS

what is distinctive about the linear framework?

the graph is a mathematical entity

which allows the structure of a system to be rigorously specified

and theorems proved which rise above molecular complexity

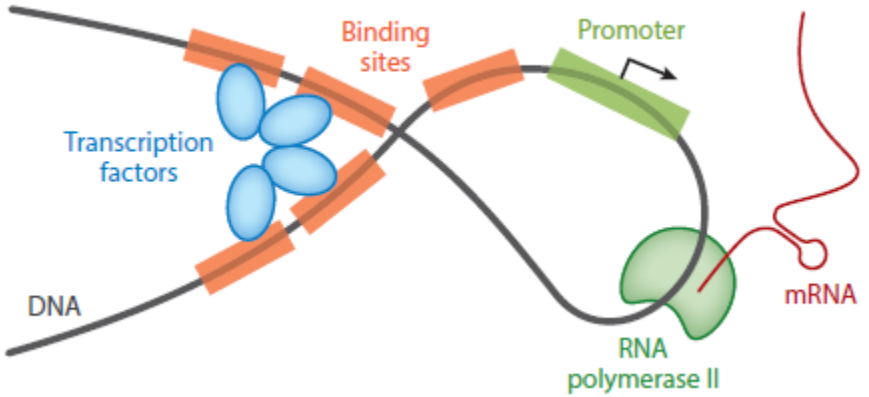
Wong, Dutta, Chowdhury, Gunawardena, "*Structural conditions on complex networks for the Michaelis-Menten input-output response*", **PNAS** 115:9738-43 2018

The Michaelis–Menten (MM) formula arose to explain simple enzyme behavior. It has since been found to describe input–output responses in several other biological contexts. Its ubiquity has been surprising and poorly understood. Here, we use the graph-based "linear framework" to show how the MM formula arises whenever appropriate structural conditions are satisfied, both at thermodynamic equilibrium and when energy is being dissipated. These conditions are based on separating the graph into two parts and constraining how and where the input variable appears. The conditions do not depend on parameter values and allow many of the details to be arbitrary. This explains the ubiquity of the MM formula and substantially generalizes previous results.

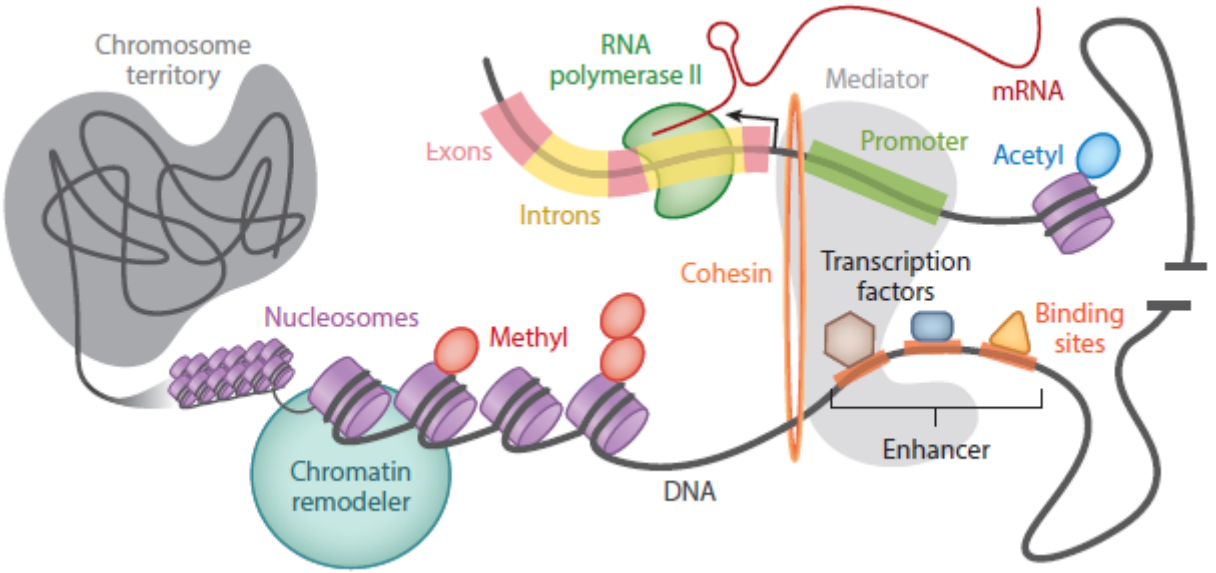
- hopfield and his barrier
- the linear framework
- gene regulation
- the hill function as hopfield barrier
- non-equilibrium challenges

vast increase in regulatory complexity

eubacteria

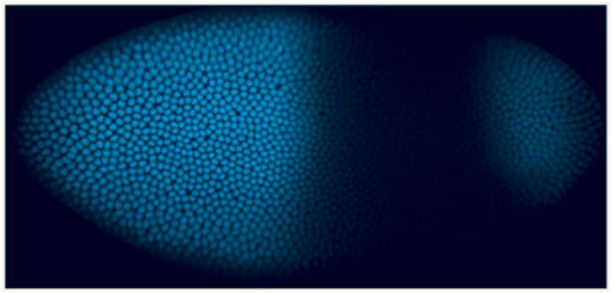


eukaryotes

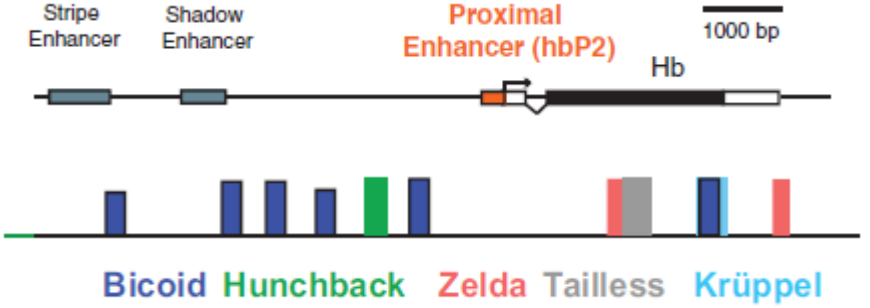


Wong, Gunawardena "*Gene regulation in and out of equilibrium*", **Annu Rev Biophys** 49:199-226 2020

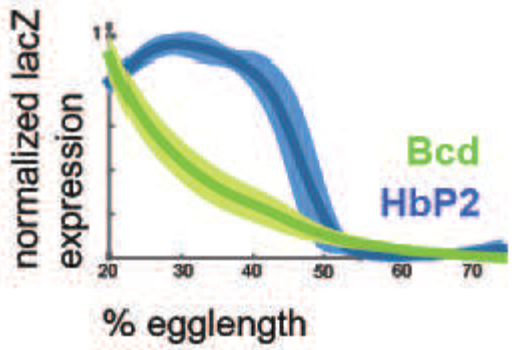
regulation of hb by Bcd in Drosophila



Hb expression

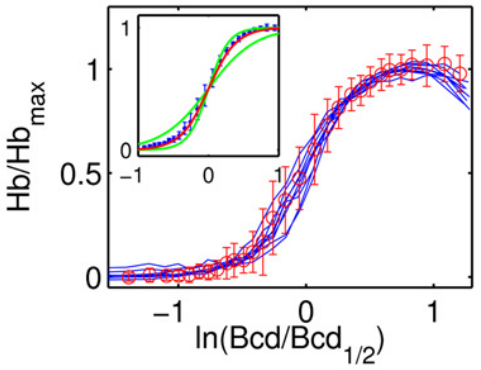


hb P2 enhancer



Park, Estrada, Johnson, Vincent, Ricci-Tam, Bragdon, Shulgina, Cha, Wunderlich, Gunawardena, DePace, "*Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity*, **eLife** 8:e41266 2019.

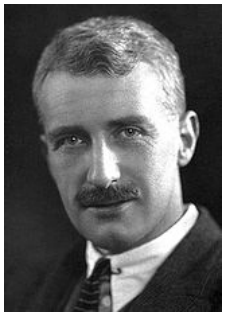
Hill functions measure sharpness of regulation



$$\frac{[\mathrm{Hb}]}{[\mathrm{Hb}]_{max}} \approx \frac{x^5}{1+x^5}$$

$$x = [\mathrm{Bcd}]/[\mathrm{Bcd}]_{0.5}$$

Gregor, Tank, Wieschaus, Bialek, *"Probing the limits to positional information"*, **Cell** 130:153-64 2007;



1886-1977

$$\frac{x^h}{1+x^h}$$

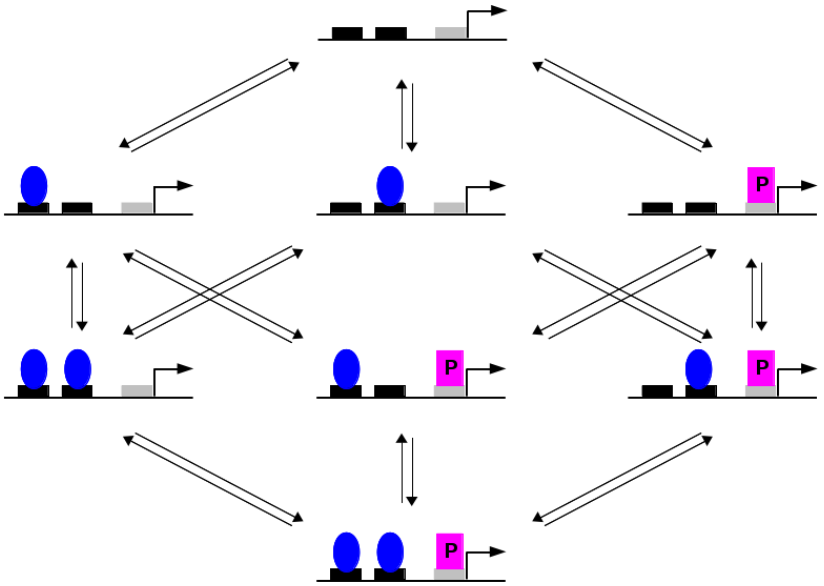
Hill function with coefficient h

Hill, *"The combinations of haemoglobin with oxygen and with carbon dioxide"*, **Biochem J** 7:471-80 1913

Engel, "*A hundred years of the Hill equation*", **Biochem J** doi:10.1042/BJ20131164 2013;
Weiss, "*The Hill equation revisited: uses and abuses*", **FASEB J** 11:835-41 1997.

- hopfield and his barrier
- the linear framework
- gene regulation
- the hill function as hopfield barrier
- non-equilibrium challenges

gene-regulation function

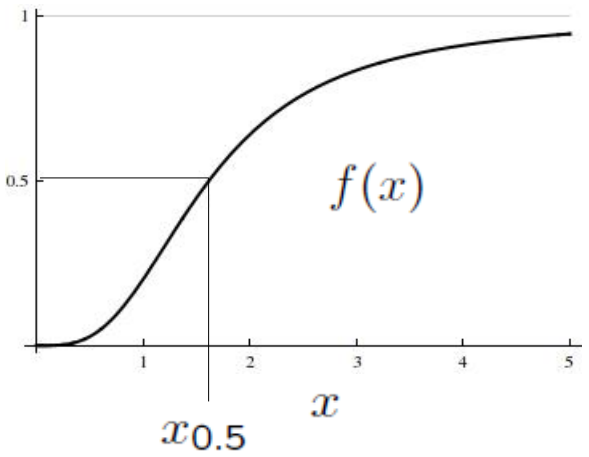


transcription factor L

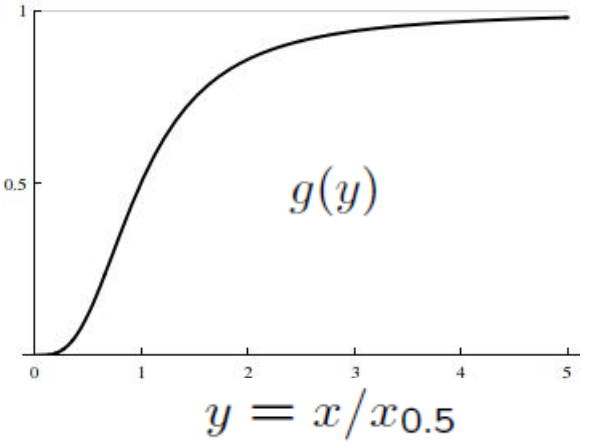
P RNA polymerase (RNAP)

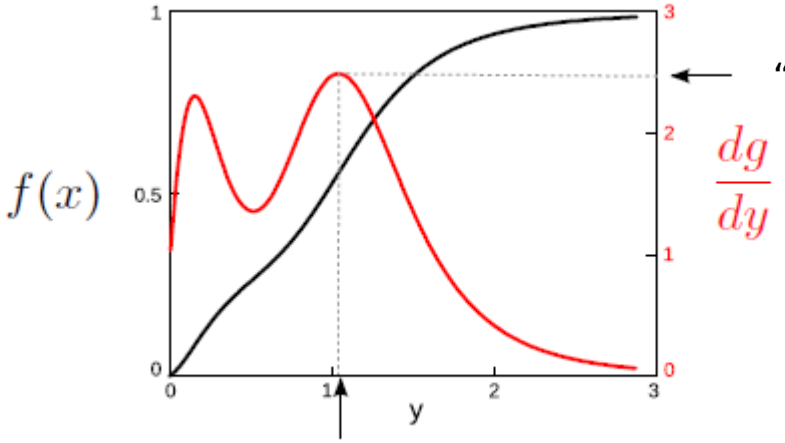
gene regulation function (GRF) – $f(x) = \sum_{i \,|\, \text{RNAP is bound}} \mathbf{p}_i^* \qquad (x = [L])$

measuring sharpness



normalisation





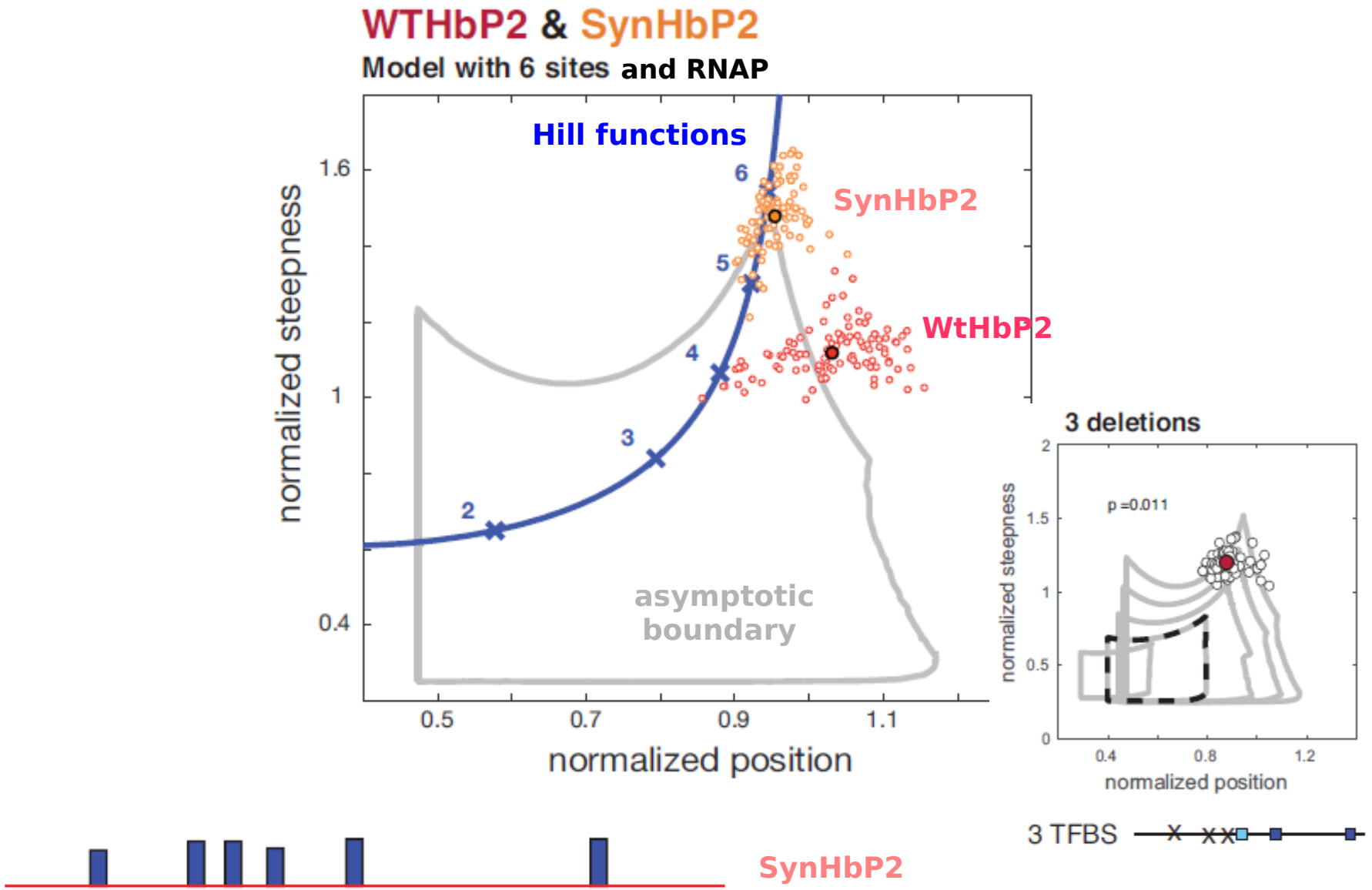
$f(x)$

$\frac{dg}{dy}$

“steepness”

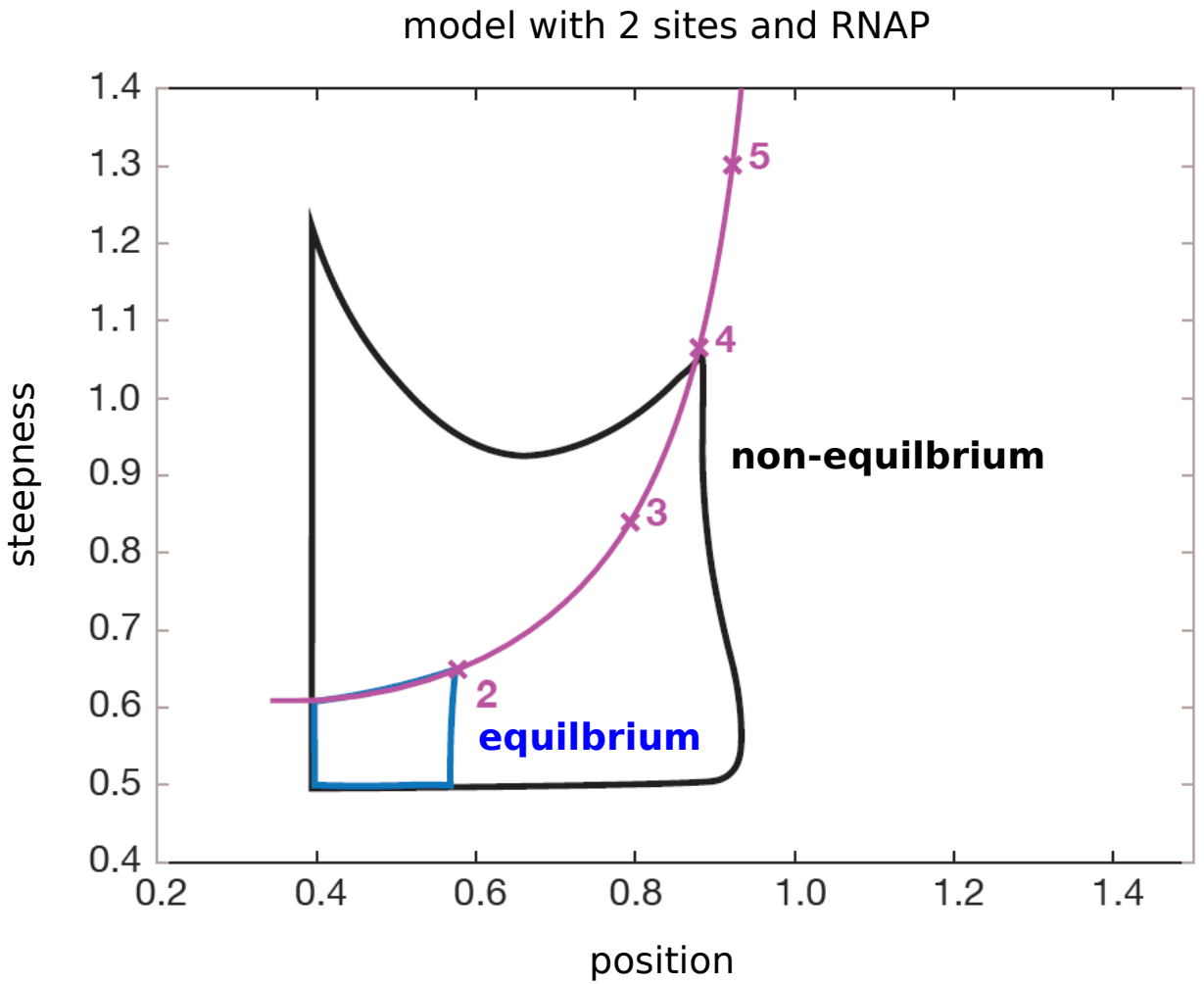
“position”

position-steepness at thermodynamic equilibrium



Park, et al, **eLife** 8:e41266 2019.

and away from thermodynamic equilibrium

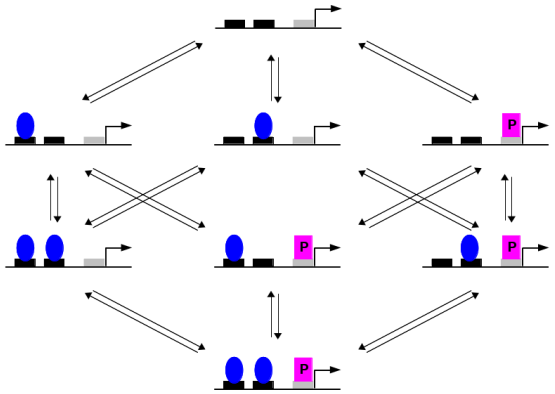


- hopfield and his barrier
- the linear framework
- gene regulation
- the hill function as hopfield barrier
- non-equilibrium challenges

the darkness away from equilibrium

$$\rho_i(G) = \sum_{T \in \Theta_i(G)} \left(\prod_{j \to k \in T} \ell(j \to k) \right)$$

global parameter dependence: steady-state probabilities depend on all system parameters, even when detailed balance is broken at only one edge



1 TF site + P	4 spanning trees
2 TF sites + P	384 spanning trees
3 TF sites + P	**42,467,328** spanning trees

combinatorial explosion: non-equilibrium physics has typically been studied by approximating to a small number of dominant spanning trees

lack of physical interpretation – we cant see the wood for the trees!

but there is some light ...

Cetiner, Gunawardena, "*Reformulating non-equilibrium steady states and generalised Hopfield discrimination*", **arXiv** doi:10.1101/456640 2020

minimal path from i to 1 $$P : i = i_1 \rightleftharpoons i_2 \rightleftharpoons \cdots \rightleftharpoons i_k = 1$$

local detailed balance $$S(P) = \log\left[\frac{\ell(i_1 \to i_2)}{\ell(i_2 \to i_1)} \cdots \frac{\ell(i_{k-1} \to i_k)}{\ell(i_k \to i_{k-1})}\right]$$

at thermodynamic equilibrium – $$\mathbf{p}_i^* \propto \exp(-S(P))$$

away from equilibrium – $$\boxed{\mathbf{p}_i^* \propto \langle \exp(-S(P)) \rangle_{\mathcal{A}(T)}}$$

 # distinct minimal path entropies scales with number of "energetic edges"

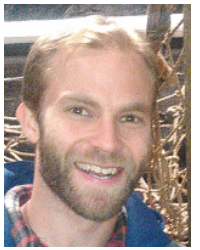
with special thanks to

Angela DePace

Jeehae Park

Javier Estrada

Felix Wong



john Biddle

Kate Shulgina


John Biddle


Pencho Yordanov


Dan Lu


David Croll


Matt Thomson


Kate Shulgina

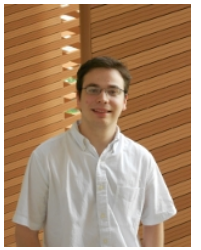

Ugur Cetiner


Rosa Martinez-Corral


Chris Nam

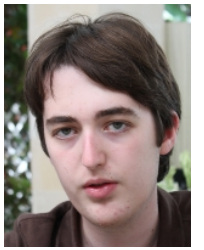

Sabina Haque


Felix Wong


Joseph Dexter

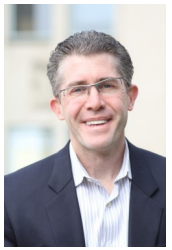

Inom Mirzaev


Max Nguyen


Jeremy Owen


Angela DePace


Galit Lahav


Neil Kelleher

With thanks to HMS, Armenise, NSF, NIH, CDP@MIT, HFSP, Novartis