

In vitro Philosophers: Chemical Reaction Networks that do Machine Learning

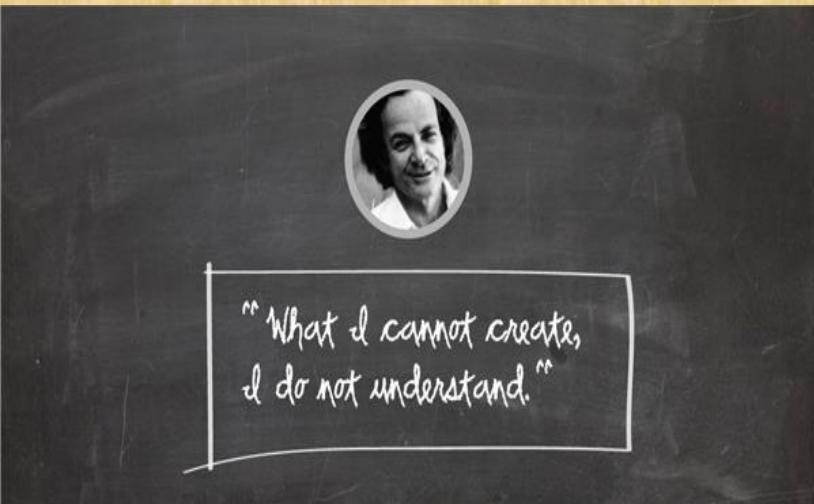
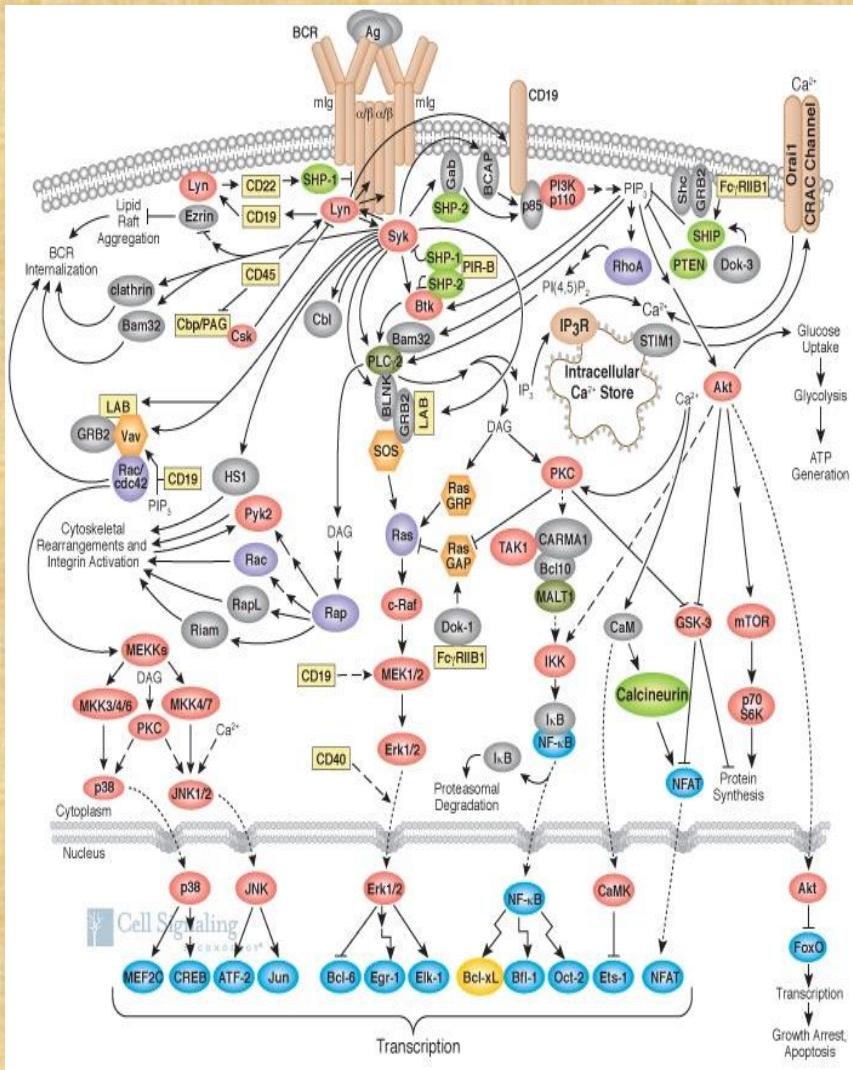
Manoj Gopalkrishnan
IIT Bombay
manoj.gopalkrishnan@gmail.com

Neutrophil chasing a bacterium

[David Rogers, Vanderbilt, 1950s]



Biochemical Reaction Networks



Molecular Programming Project

Computer science and engineering has mastered complexity for electronic computation – can we do the same for engineering molecular devices and systems?

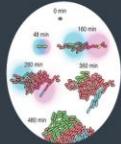
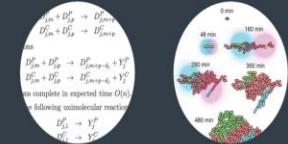
[Learn More](#)



Example works



NUPACK:
nucleic acid package



Leaderless deterministic
chemical reaction networks

Chemical implementation of neural networks and Turing machines

ALLEN HJELMFELT[†], EDWARD D. WEINBERGER[†], AND JOHN ROSS[‡]

Biomolecular implementation of linear I/O systems

K. Oishi E. Klavins

Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades

Lulu Qian¹, Erik Winfree^{1,2,3,*}

¹*Bioengineering, California Institute of Technology, Pasadena, CA 91125, USA.*

²*Computer Science, California Institute of Technology, Pasadena, CA 91125, USA.*

³*Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125, USA.*

✉ To whom correspondence should be addressed. E-mail: winfree@caltech.edu

– Hide authors and affiliations

Science 03 Jun 2011:
Vol. 332, Issue 6034, pp. 1196-1201
DOI: 10.1126/science.1200520



Letter

Synthetic analog computation in living cells

Ramiz Daniel, Jacob R. Rubens, Rahul Sarpeshkar & Timothy K. Lu

Nature **497**, 619–623 (30 May 2013)
doi:10.1038/nature12148

Received: 10 October 2012
Accepted: 03 April 2013
Published online: 15 May 2013

Message Passing Inference with Chemical Reaction Networks

Nils Napp
Wyss Institute for Biologically Inspired Engineering
Harvard University
Cambridge, MA 02138
nnapp@wyss.harvard.edu

Ryan Prescott Adams
School of Engineering and Applied Sciences
Harvard University
Cambridge, MA 02138
rpa@seas.harvard.edu

Computing Algebraic Functions with Biochemical Reaction Networks

H. J. Buisman, Huub M. M. ten Eikelder, Peter A. J. Hilbers, Anthony M. L. Liekens
 Published 2009 in Artificial Life

Chemical Reaction Network Designs for Asynchronous Logic Circuits

Authors

[Authors and affiliations](#)

Luca Cardelli, Marta Kwiatkowska, Max Whitby 

Computation with finite stochastic chemical reaction networks

Authors

[Authors and affiliations](#)

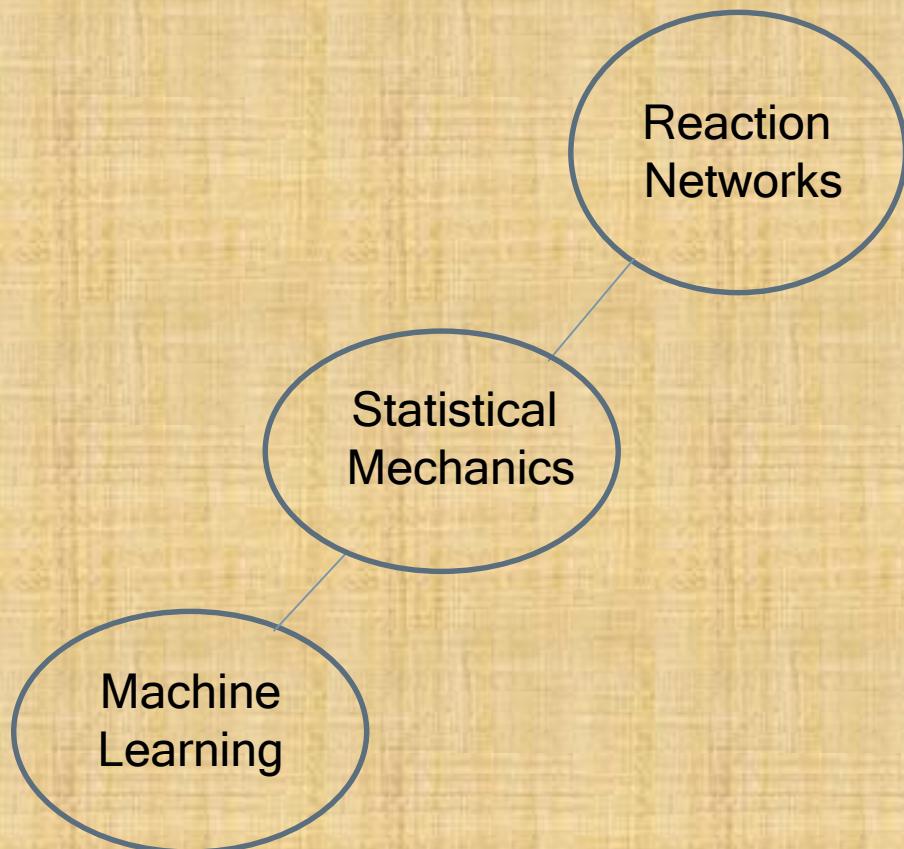
David Soloveichik , Matthew Cook, Erik Winfree, Jehoshua Bruck

Computing with Reaction Networks

- Turing machines
- Boolean circuits
- Linear input/ output systems
- Analog electronic circuits
- Algebraic functions
- Marginals of a joint distribution
- Asynchronous computing
- Distributed computing
- Register machines, ...

Find: A computational model **natural** to reaction networks

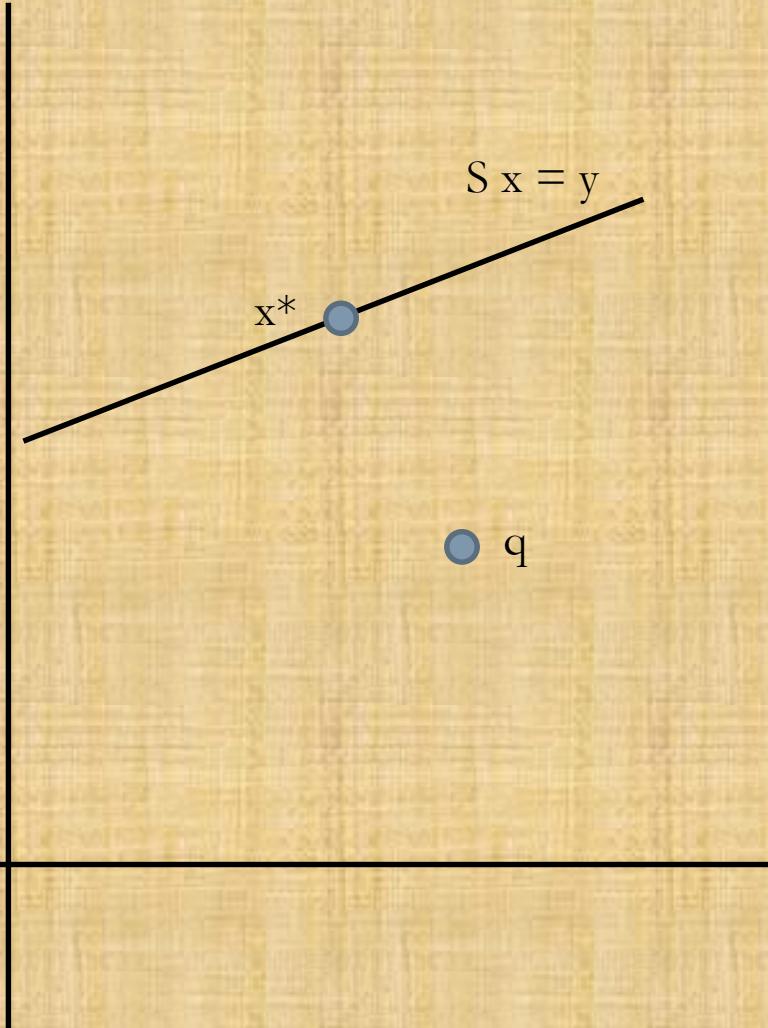
- **Explain** biological reality
- **Efficient** use of resources, cheap to implement
- **Exploit, extend** statistical mechanics



EXPECTATION MAXIMIZATION THE EM ALGORITHM

E(xponential)-projection

$$D(p||q) := \sum_i p_i \log p_i - p_i - p_i \log q_i + q_i$$



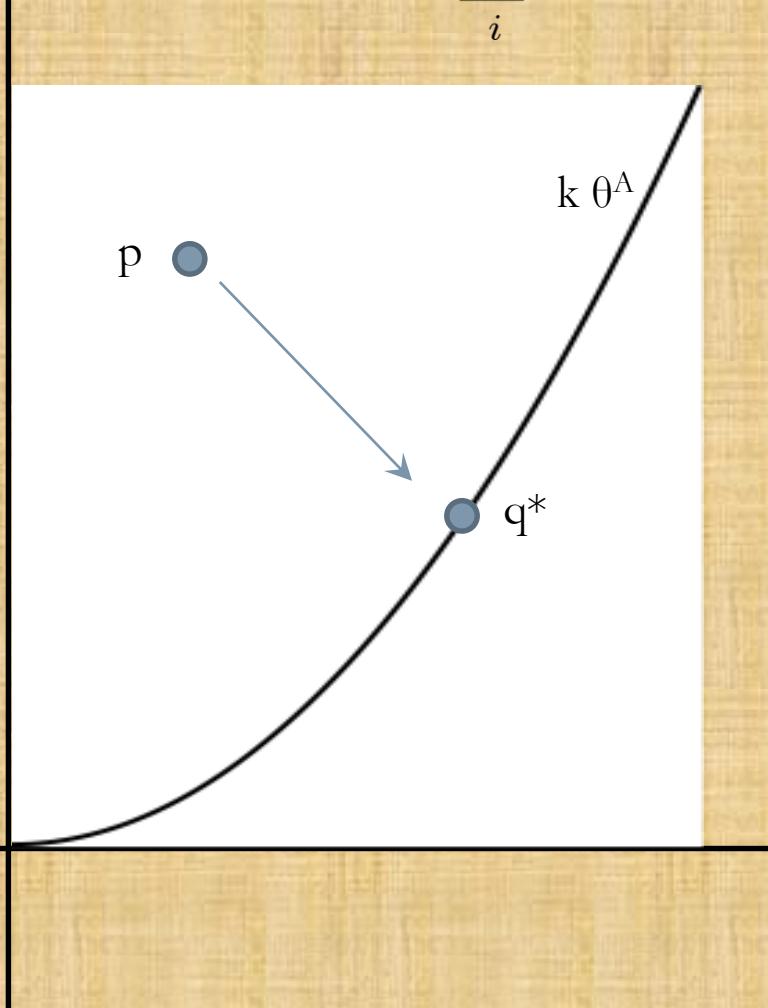
- Given: $S x = y$
- $x^* := \arg \min_{x: Sx=y} D(x||q)$



Rationalism: Leibnitz

M(ixture)-projection

$$D(p||q) := \sum_i p_i \log p_i - p_i - p_i \log q_i + q_i$$



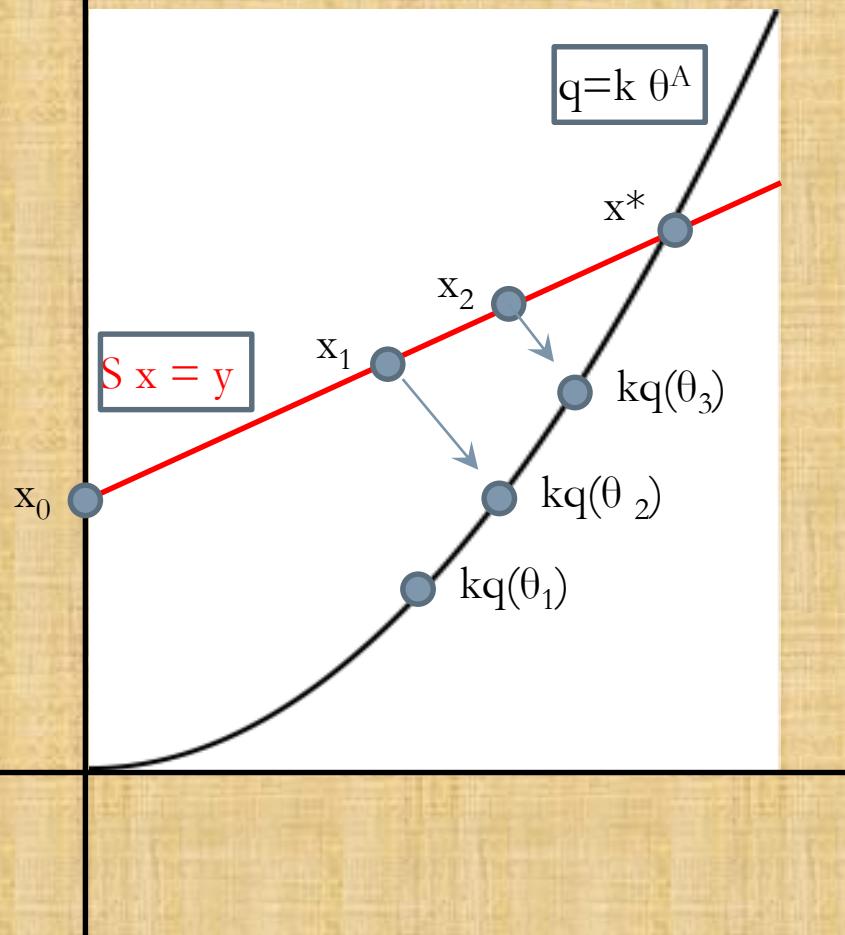
- $q(\theta) = k \theta^A$
- $\theta^* := \arg \min_{\theta} D(p||q(\theta))$
- $q^* := q(\theta^*)$



Empiricism: David Hume 12

Expectation Maximization (EM) Algorithm

$$D(p||q) := \sum_i p_i \log p_i - p_i - p_i \log q_i + q_i$$



- $q(\theta) = k\theta^A$
- Observation y
- Find

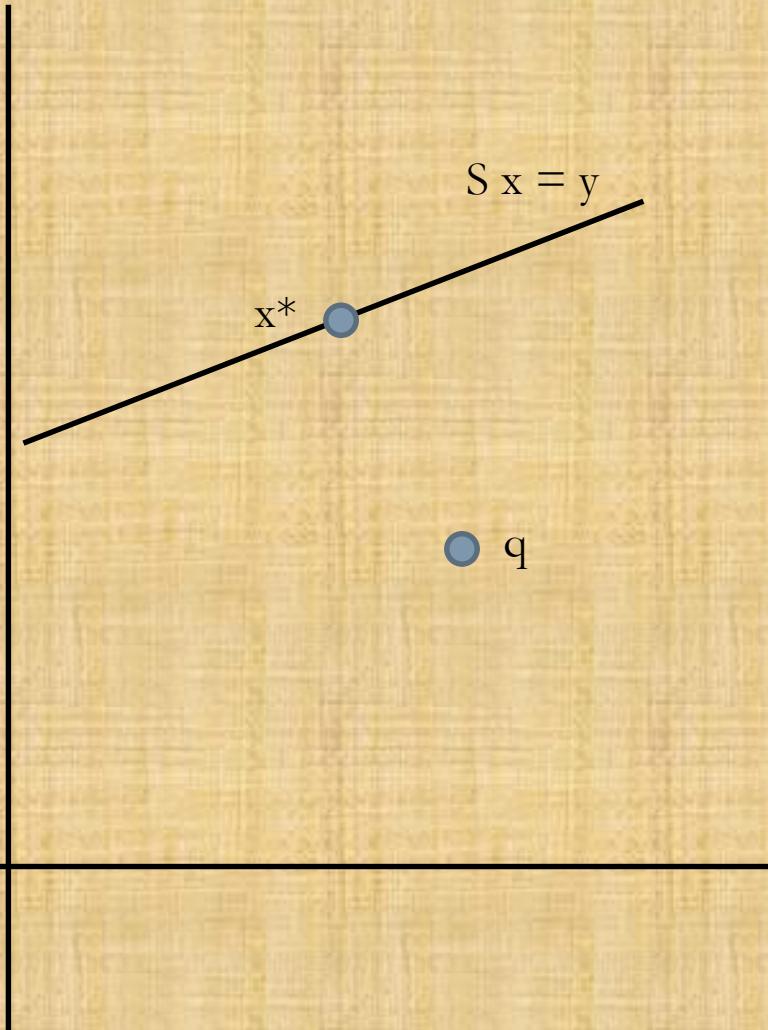
$$(x^*, \theta^*) = \arg \min_{x: Sx=y, \theta} D(x||q(\theta))$$



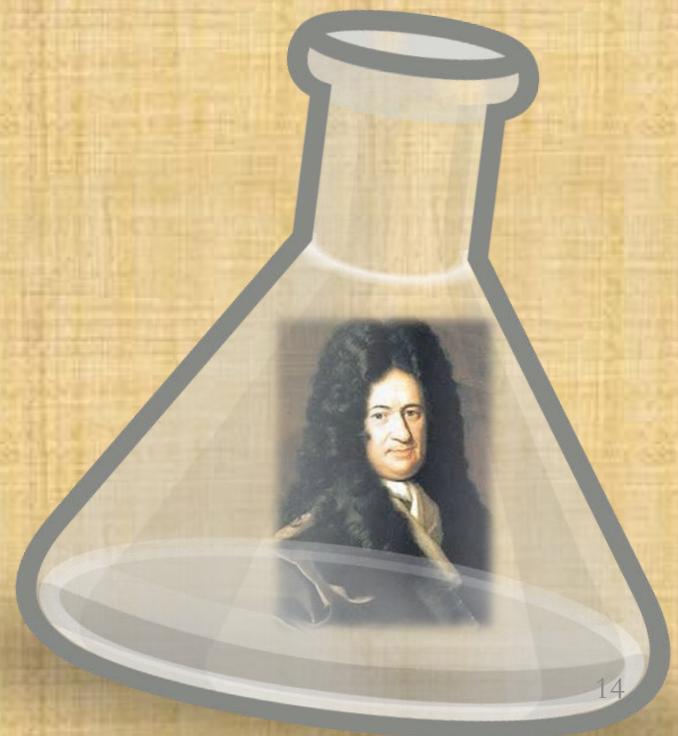
Synthesis: Immanuel Kant

E(xponential)-projection

$$D(p||q) := \sum_i p_i \log p_i - p_i - p_i \log q_i + q_i$$



- Given: $S x = y$
- $x^* := \arg \min_{x: Sx=y} D(x||q)$

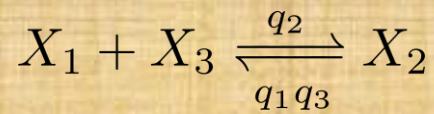


In vitro E-projection

$$x^* := \arg \min_{x: Sx=y} D(x||q)$$

- **Matrix S**
- Compute basis of $\ker(S)$
- Each basis element → reversible reaction
 - Detailed Balance at q
- **Deterministic mass-action**
 - Set $x(0)$ s.t. $S x(0) = y$
- **Thm** $x^* = \lim_{t \rightarrow \infty} x(t)$ exists and is the required E-projection.

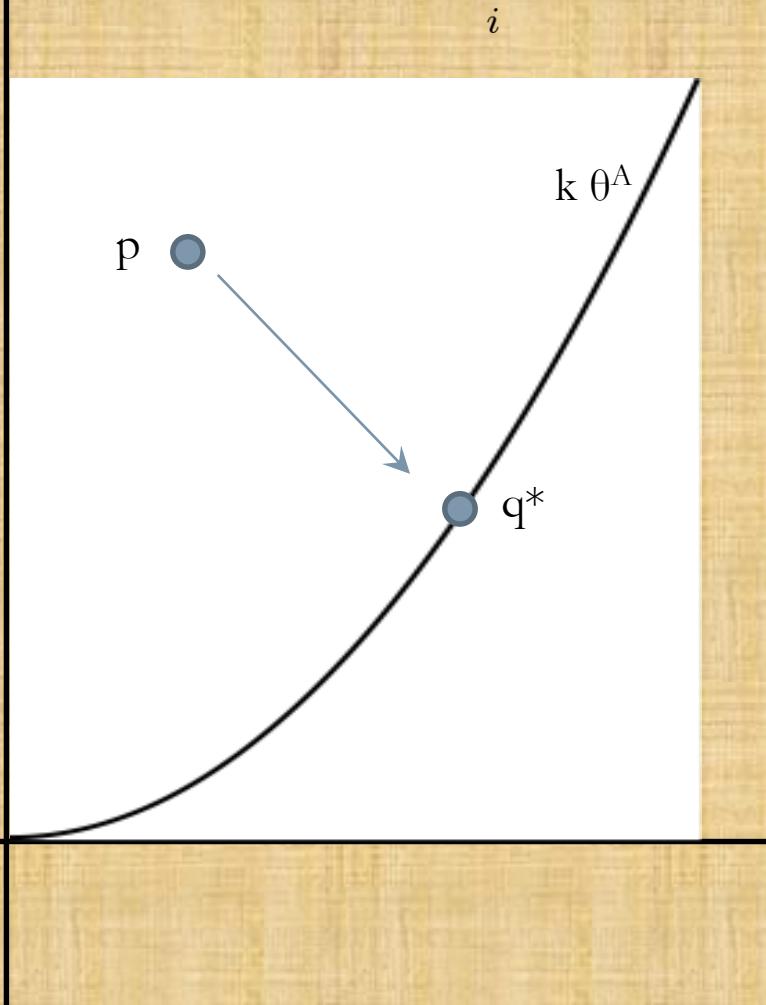
$$\ker \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$



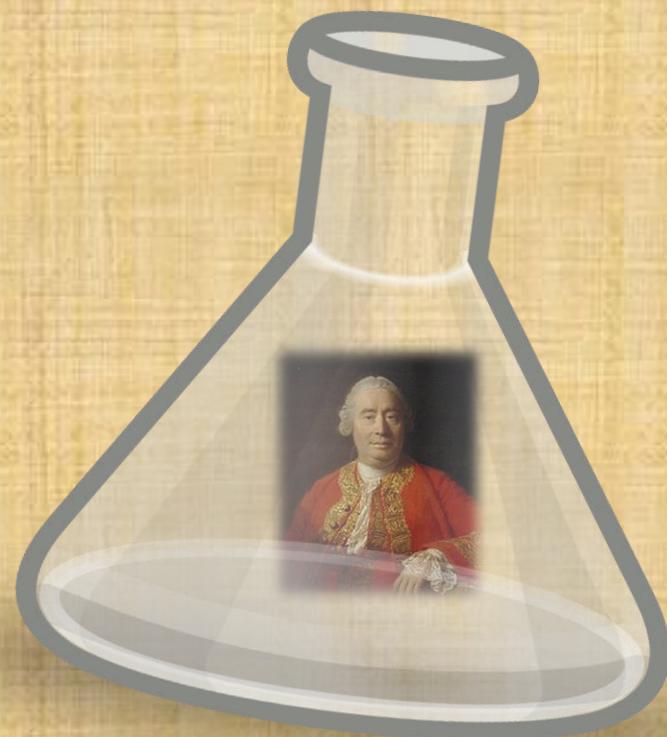
$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = (q_2 x_1 x_3 - q_1 q_3 x_2) \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$$

M(ixture)-projection

$$D(p||q) := \sum_i p_i \log p_i - p_i - p_i \log q_i + q_i$$



- $q(\theta) = k \theta^A$
- $\theta^* := \arg \min_{\theta} D(p||q(\theta))$
- $q^* := q(\theta^*)$



In vitro M-projection

$$\theta^* := \arg \min_{\theta} D(p || k\theta^A)$$

- Matrix A
- Initial conditions
- Each column of A => 2 irreversible reactions.
- Apply mass-action kinetics
- **Thm:** $\frac{\dot{\theta}_i}{\theta_i} = -\frac{\partial D(p || k\theta^A)}{\partial \theta_i}$

Hence $D(p | | k\theta^A)$ decreases along solution trajectories and the dynamics converges to a local minimum. If the reaction network is detailed balanced, then dynamics converges to the global minimum and computes M-projection

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

$$p = (p_1, p_2, p_3)$$

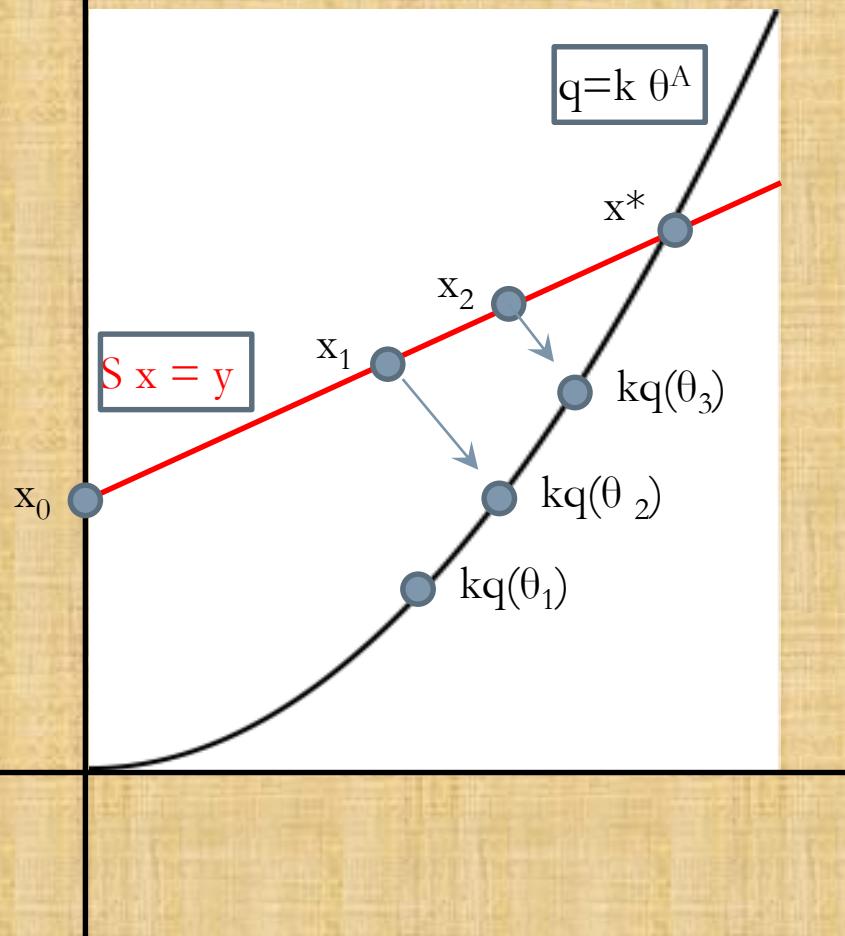
$$2\theta_1 \xrightleftharpoons[p_1]{k_1} 0, \quad \theta_1 + \theta_2 \xrightleftharpoons[p_2]{k_2} 0, \quad 2\theta_2 \xrightleftharpoons[p_3]{k_3} 0$$

$$\begin{aligned}\dot{\theta}_1 &= -2k_1\theta_1^2 + 2p_1 - k_2\theta_1\theta_2 + p_2 \\ \dot{\theta}_2 &= -2k_3\theta_2^2 + 2p_3 - k_2\theta_1\theta_2 + p_2\end{aligned}$$

$$\theta(\infty) = \theta^*$$

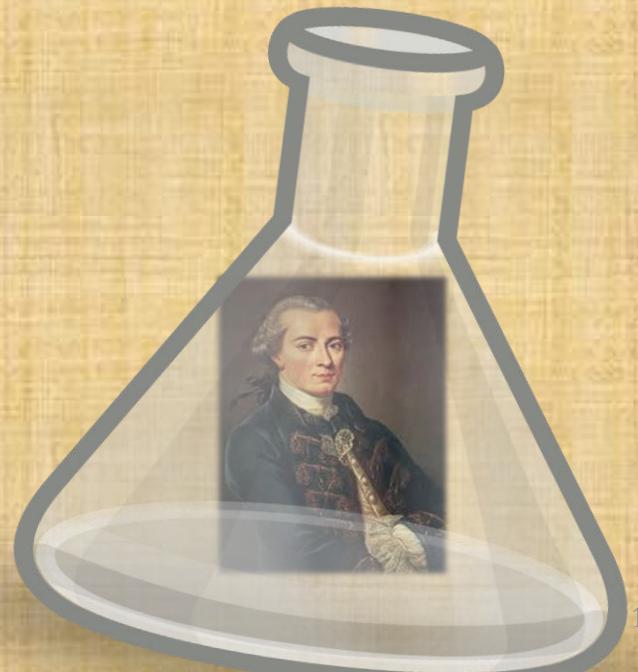
Expectation Maximization (EM) Algorithm

$$D(p||q) := \sum_i p_i \log p_i - p_i - p_i \log q_i + q_i$$



- $q(\theta) = k\theta^A$
- Observation y
- Find

$$(x^*, \theta^*) = \arg \min_{x: Sx=y, \theta} D(x||q(\theta))$$



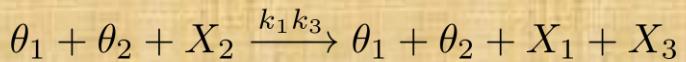
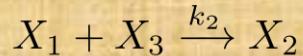
In vitro EM-algorithm

$$(x^*, \theta^*) = \arg \min_{x: Sx=y, \theta} D(x || k\theta^A)$$

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

- Compute basis for $\ker S$
- Each basis element \Rightarrow pair of reactions
- Each column of $A \Rightarrow$ 2 irreversible reactions
- Convert data into initial conditions
- Apply mass-action kinetics
- **Thm:** $D(x || k\theta^A)$ decreases along solution trajectories and the dynamics converges to a local minimum.

$$\ker \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$



$$2\theta_1 \xrightarrow{k_1} 0, \quad \theta_1 + \theta_2 \xrightarrow{k_2} 0, \quad 2\theta_2 \xrightarrow{k_3} 0,$$

$$X_1 \xrightarrow{1} X_1 + 2\theta_1, \quad X_2 \xrightarrow{1} X_2 + \theta_1 + \theta_2, \quad X_3 \xrightarrow{1} X_3 + 2\theta_2$$

$$x(0) = (u_1, u_2, u_3), \theta(0) = (0, 0)$$

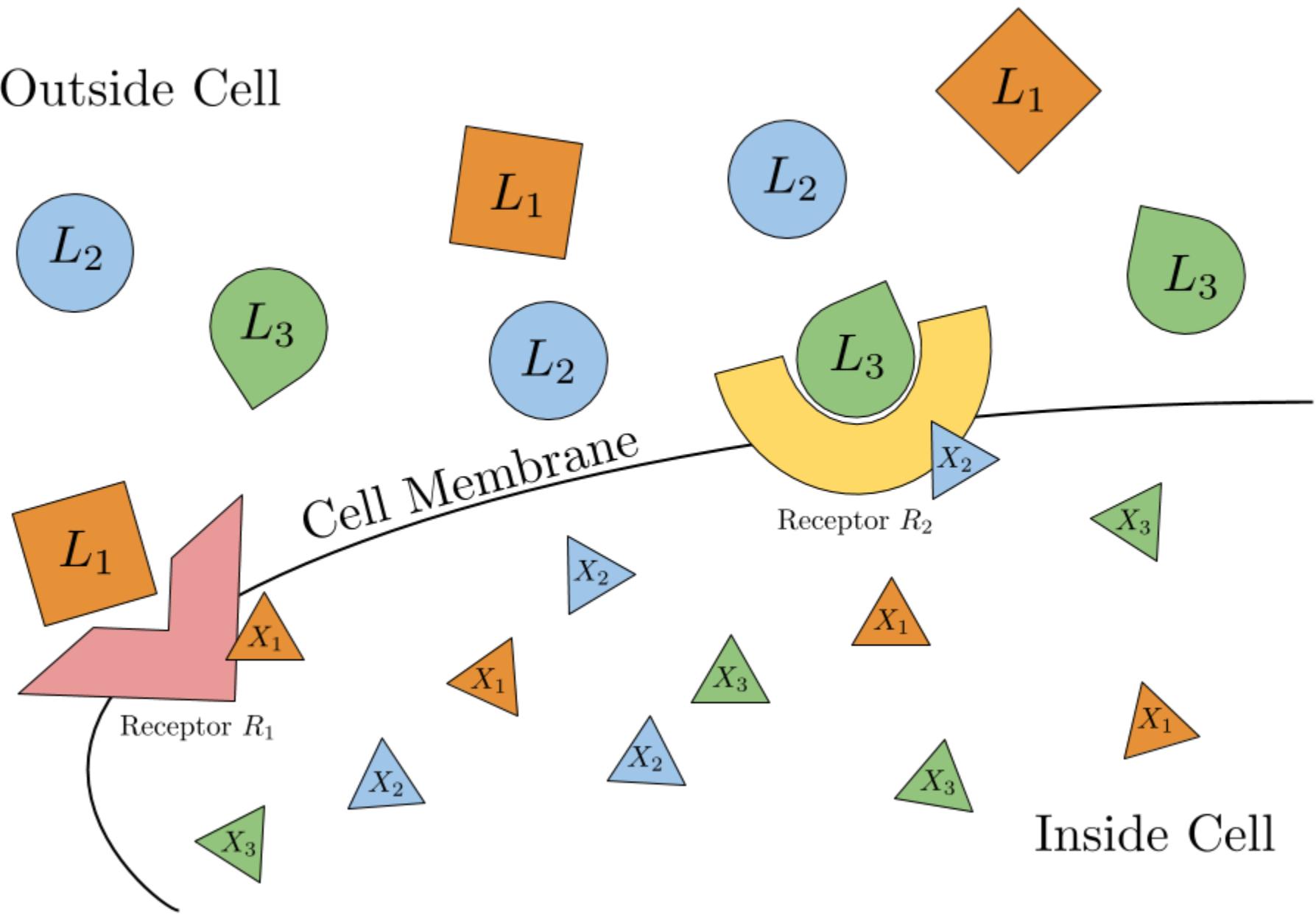
$$\dot{X}_1 = \dot{X}_3 = -\dot{X}_2 = -k_2 X_1 X_3 + k_1 k_3 X_2 \theta_1 \theta_2$$

$$\dot{\theta}_1 = -2k_1 \theta_1^2 + 2X_1 - k_2 \theta_1 \theta_2 + X_2$$

$$\dot{\theta}_2 = -2k_3 \theta_2^2 + 2X_3 - k_2 \theta_1 \theta_2 + X_2$$

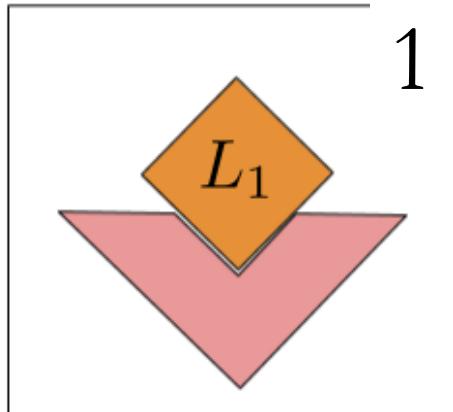
$$(x(\infty), \theta(\infty)) = (x^*, \theta^*)$$

Outside Cell



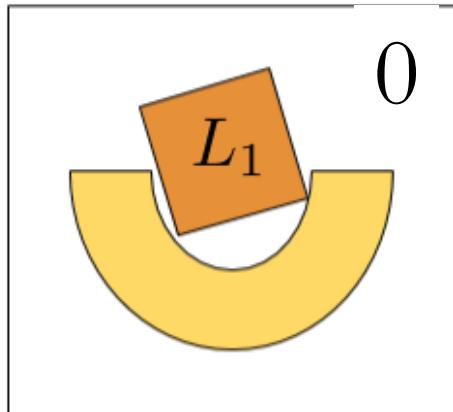
Inside Cell

Receptor R_1



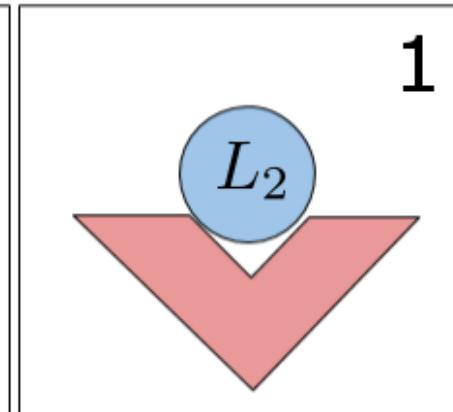
1

Receptor R_2

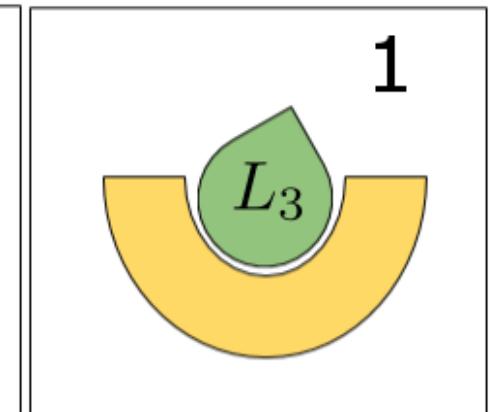


0

Ligand L_1



1



Ligand L_2

Ligand L_3

Sensitivity Matrix

Cellular sensing as an EM problem

- Hypothesis: “State of the world” parameters

$$\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_m$$

- Hypothesis: Ligand concentrations

$$q_1, q_2, \dots, q_j, \dots, q_n \quad \text{where} \quad q_j = k_j \prod_{i=1}^m \theta_i^{a_{ij}}$$

- Observation: $y = S x$ where ligand numbers in unit aliquot are $x_1, x_2, \dots, x_j, \dots, x_n$
- Question: Best hypothesis θ^* ?

Conclusions and Future Work

- Reaction network for EM algorithm
- Hidden Markov Models
- Boltzmann machine with learning rule
- Limitation: Dimension explosion
- Opportunity: Explaining cell signalling networks
- Chemical controllers
- Implementations with DNA molecules

References

1. Gopalkrishnan, Manoj. "A scheme for molecular computation of maximum likelihood estimators for log-linear models." *International Conference on DNA-Based Computers*. Springer International Publishing, 2016.
2. MV Virinchi, A Behera, M Gopalkrishnan, (2017, September). **A stochastic molecular scheme for an artificial cell to infer its environment from partial observations**. In *International Conference on DNA-Based Computers* (pp. 210-231). Springer, Cham.
3. Poole, W., Ortiz-Munoz, A., Behera, A., Jones, N. S., Ouldridge, T. E., Winfree, E., & Gopalkrishnan, M. (2017, September). **Chemical Boltzmann Machines**. In *International Conference on DNA-Based Computers* (pp. 210-231). Springer, Cham.

Information Projection

1. Information Geometry and its applications, Shun-ichi Amari, Springer 2016
2. Csiszár, Imre, and Paul C. Shields. "Information theory and statistics: A tutorial." *Foundations and Trends® in Communications and Information Theory* 1.4 (2004): 417-528.

EM algorithm

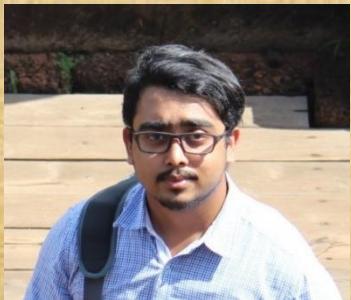
1. Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the royal statistical society. Series B (methodological)* (1977): 1-38.

Reaction Network Theory

1. Gopalkrishnan, Manoj. "Catalysis in reaction networks." *Bulletin of mathematical biology* 73.12 (2011): 2962-2982.
2. Anderson, D. F., Craciun, G., Gopalkrishnan, M., & Wiuf, C. (2015). Lyapunov functions, stationary distributions, and non-equilibrium potential for reaction networks. *Bulletin of mathematical biology*, 77(9), 1744-1767.

Thermodynamics of computation references

- Gopalkrishnan, Manoj. "A Cost/Speed/Reliability Tradeoff to Erasing." *Entropy* 18.5 (2016): 165.
- **Designing the optimal bit: balancing energetic cost, speed and reliability,** Abhishek Deshpande, Manoj Gopalkrishnan, Thomas E. Ouldridge, Nick S. Jones, Proc. R. Soc. A 2017 473 20170117
- Gopalkrishnan, Manoj, et al. "Bayesian view of single-qubit clocks, and an energy versus accuracy tradeoff." *Physical Review A* 96.3 (2017): 032339.
- Gopalkrishnan, Manoj. "The hot bit I: The Szilard-Landauer correspondence." *arXiv preprint arXiv:1311.3533* (2013).



Viswa Virinchi



Abhishek Behera

Thank You