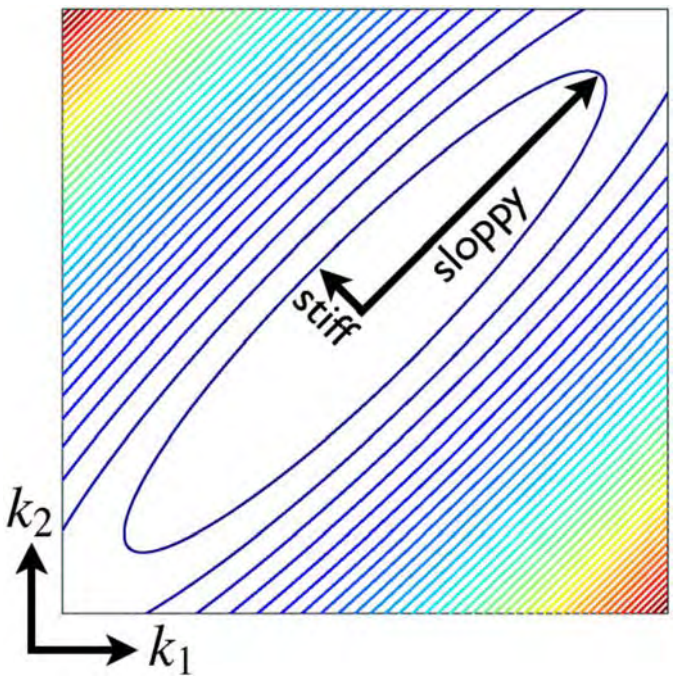
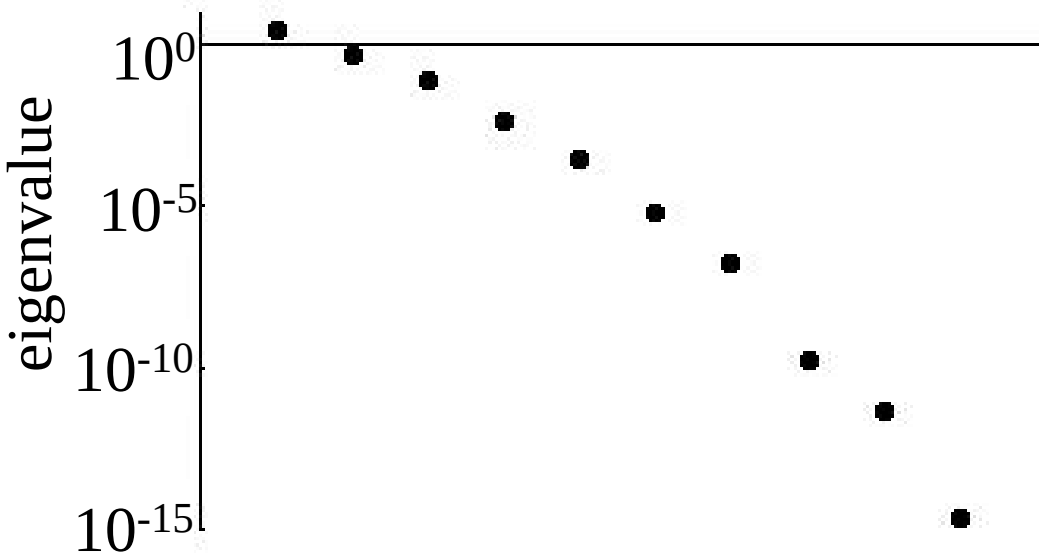
Cornell University
Laboratory of
Atomic and Solid State Physics

Origins of sloppiness in biological networks

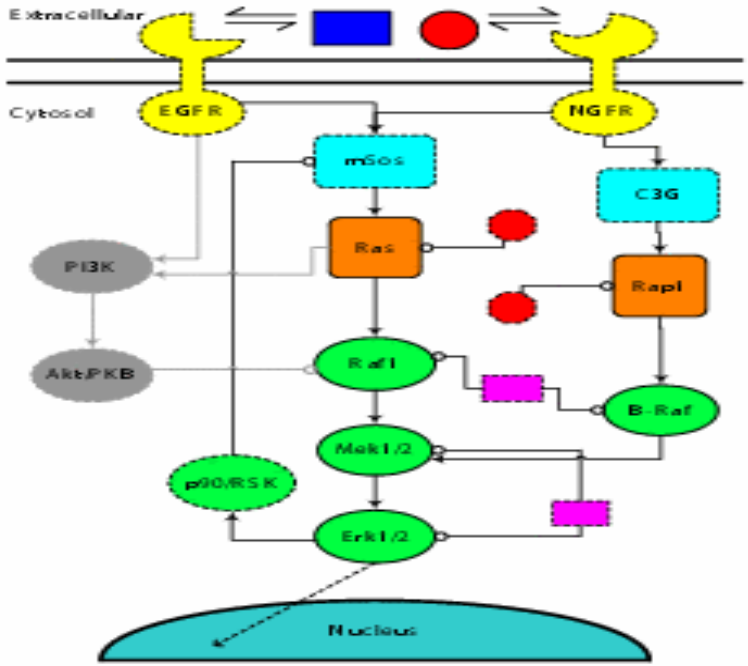
Joshua Waterfall

DOE CSGF Meeting
June 20, 2006





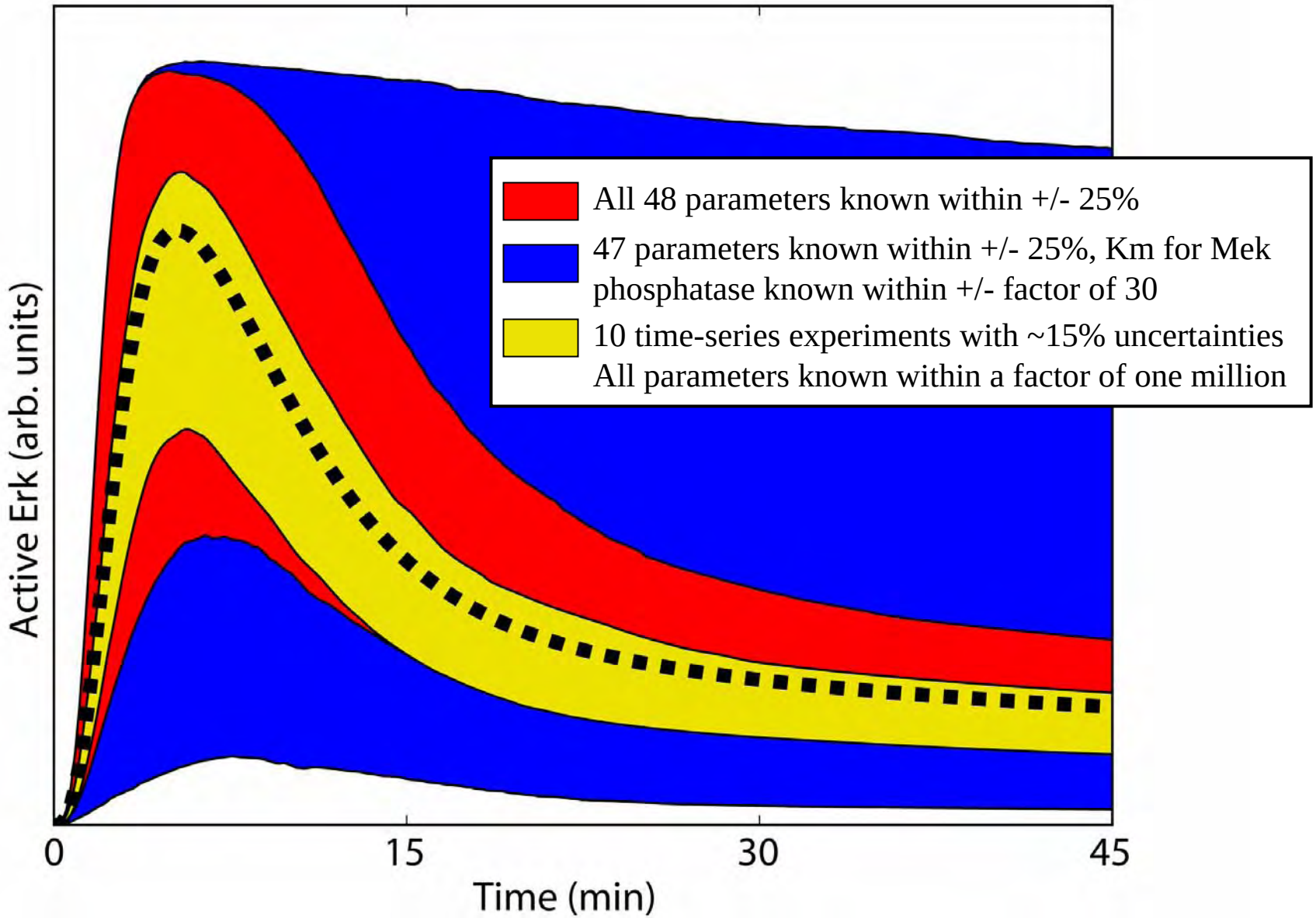
Biological networks



Growth factor signaling in PC12 cells

$$\frac{d[SOSActive]}{dt} = \frac{k_{EGF}[boundEGFR][SOSInactive]}{[SOSInactive] + K_{mEGF}}$$

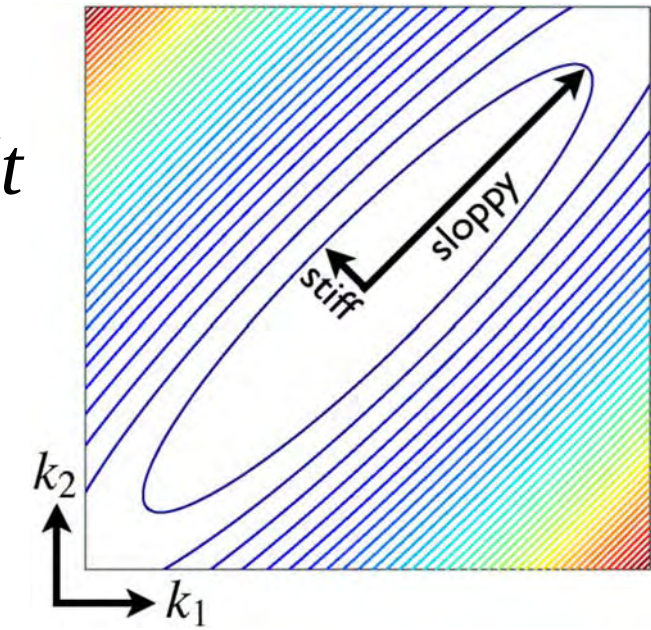
Prediction for Erk activation by EGF with PI3K inhibitor

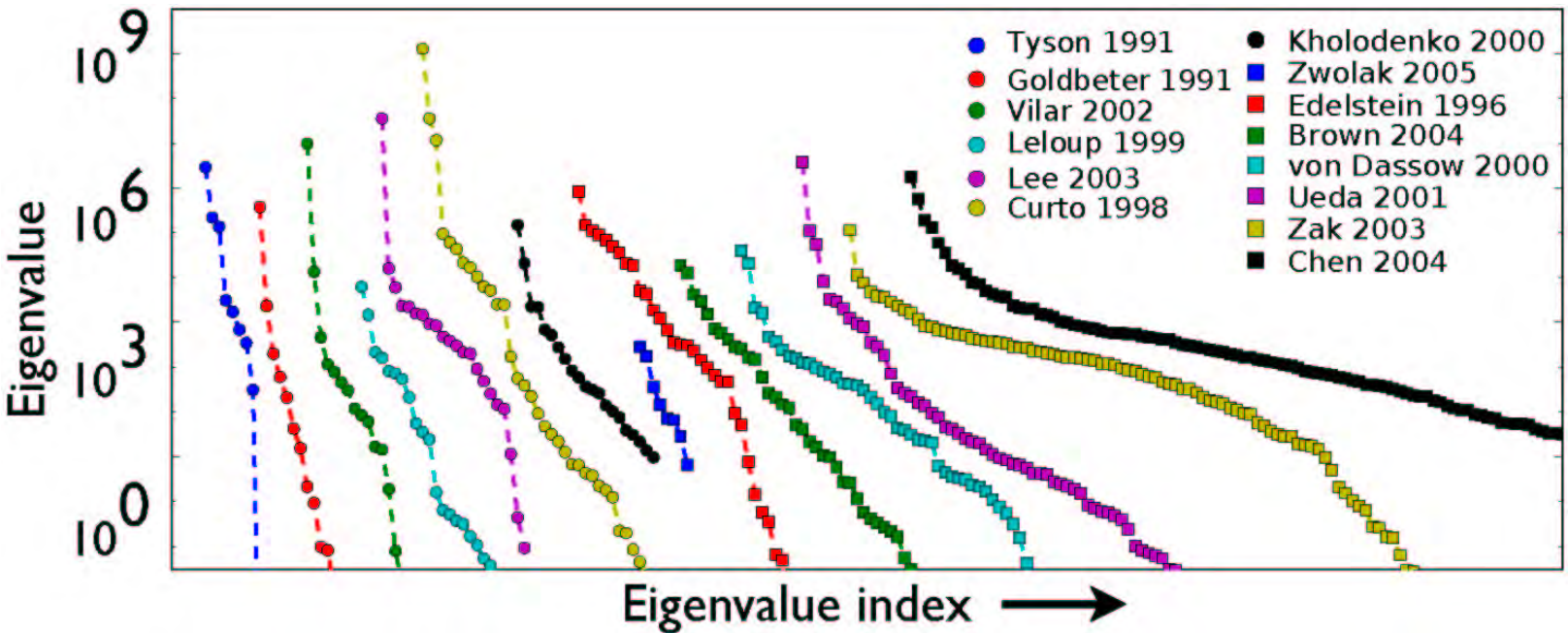


Parameter Fluctuations and Systems Fluctuations

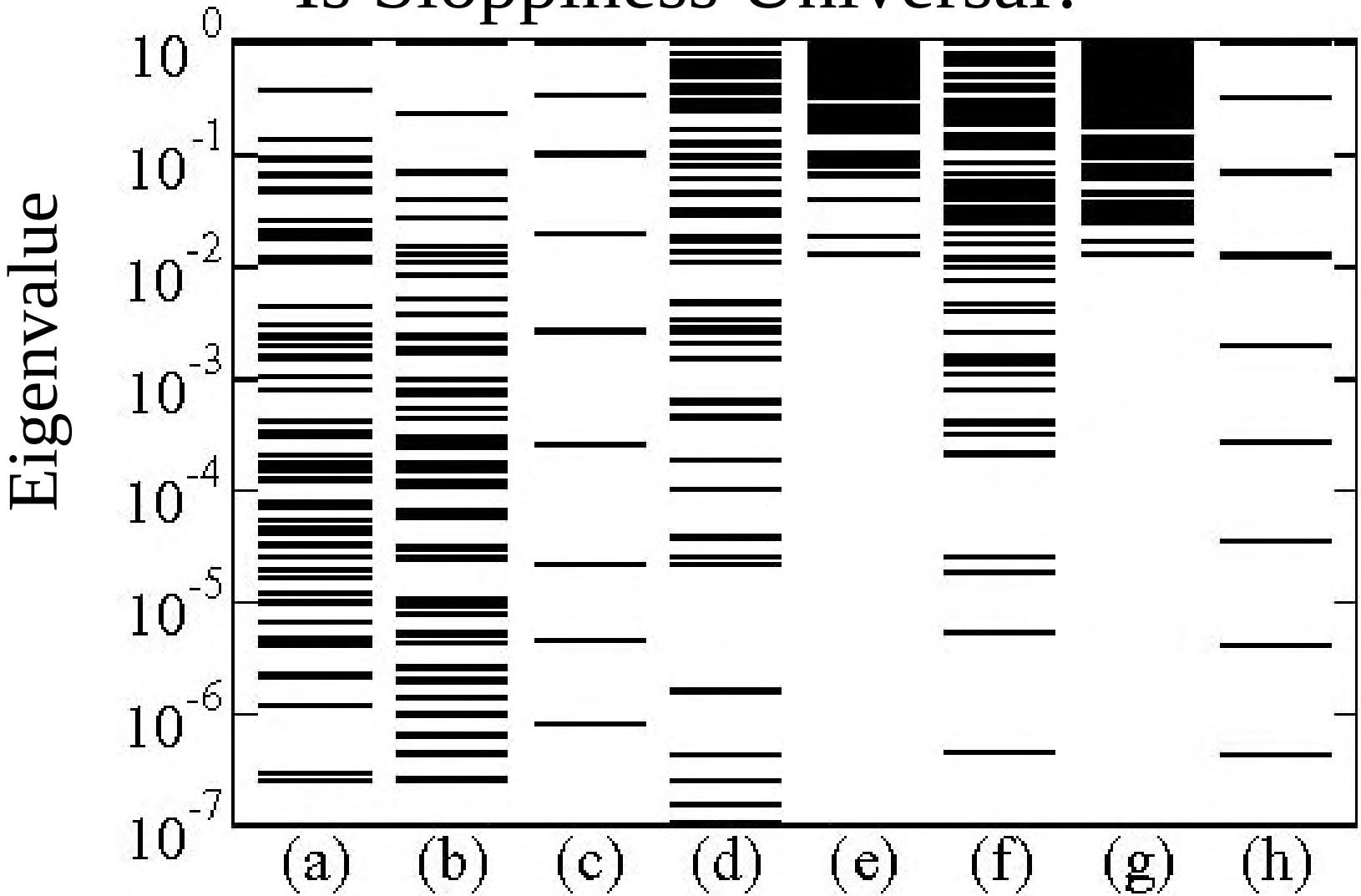
$$\chi^2(\theta) = \sum_{\text{species}} \frac{1}{T}\int_0^T \frac{(y(\theta,t) - y(\theta_0,t))^2}{\sigma^2}\,dt$$

$$H_{ij} = \frac{\partial^2 \chi^2}{\partial \log\theta_i \partial \log\theta_j}$$





Is Sloppiness Universal?



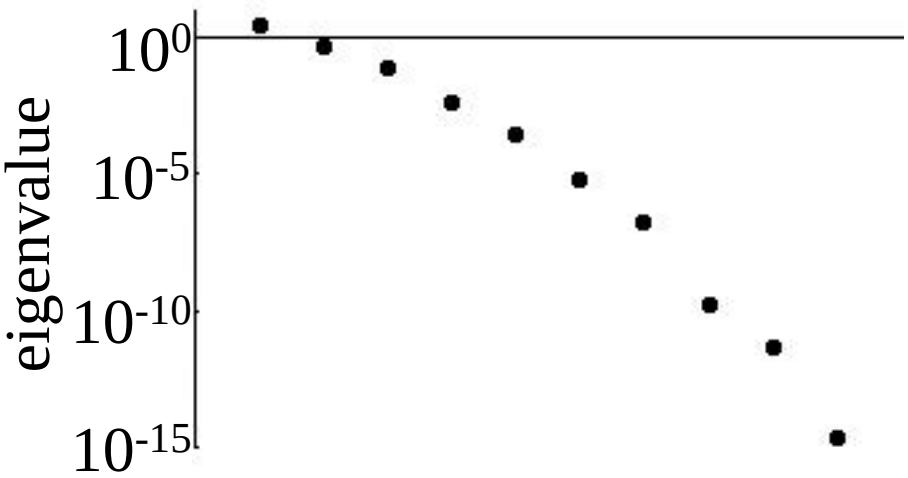
(a) Growth factor signaling in PC12 cells, (b) Quantum Monte Carlo (Jastrow parameters), (c) radioactivity from pharmaceutical isotopes, (d) 48 random exponentials from log uniform(± 25.), (e) random symmetric matrix (GOE), (f) products of GOE random matrices, (g) Wishart matrix, (h) fitting monomials (Hilbert matrix)

Sums of exponentials

$$y(x) = \sum_{i=0}^{n} e^{-\gamma_i x}$$

$$H_{ij} = \frac{2\gamma_i \gamma_j}{(\gamma_i + \gamma_j)^3}$$

SLOPPY



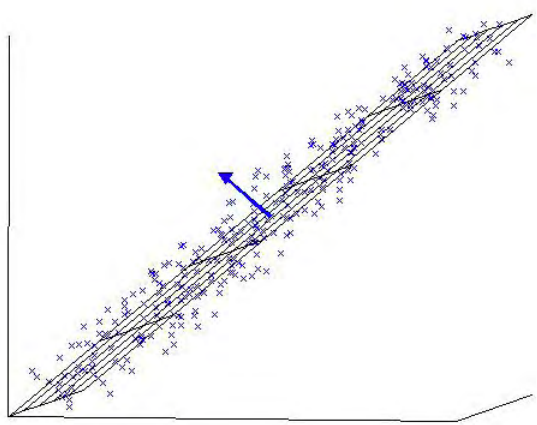
Linear Correlations

$$a(\vec{p}) = \vec{n} \cdot \vec{p}$$

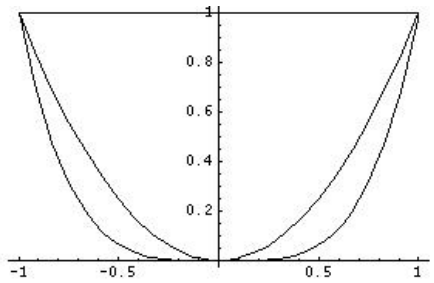
Hessian is covariance matrix of data.
Wishart statistics for eigenvalues.

$$H = \sum_i \vec{p}_{(i)} \vec{p}^{\,T}_{(i)}$$

NOT SLOPPY

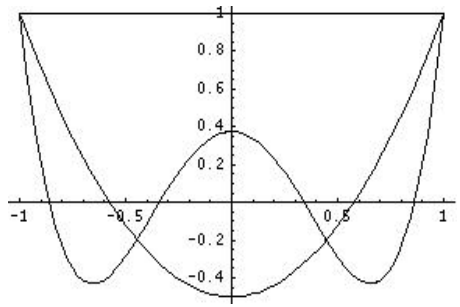


Sums of monomials



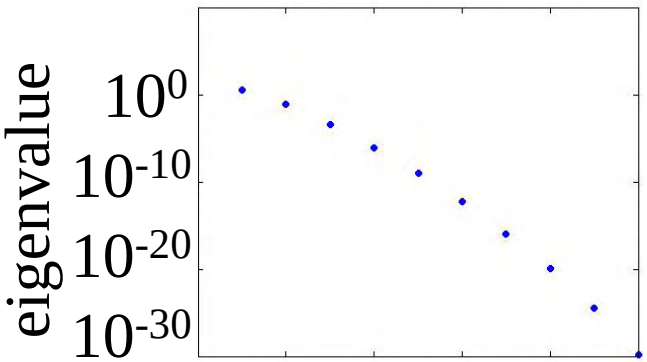
$\sqrt{|H|}$ = volume change under basis transformation

Sums of Shifted Legendre polynomials



Hessian = Hilbert matrix

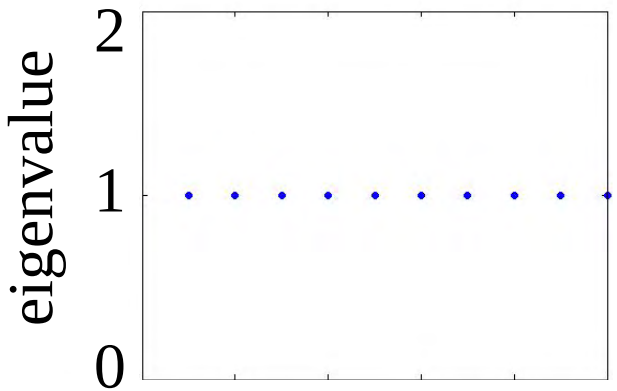
$$H_{ij} = \frac{2}{i + j + 1}$$



SLOPPY

Hessian = Identity matrix

$$H_{ij} = \delta_{ij}$$



NOT SLOPPY

The Vandermonde Ensemble

Assumptions for proof:

i. Parameters are nearly degenerate.
ii. Cost function is sum of squares of residuals.
iii. Each residual is a symmetric function of all the parameters.

Biological network properties:

i. Dozens to hundreds of biochemical parameters in limited range.
ii. χ^2 measure of system behavior variation.
iii. Similar biochemical reactions repeated throughout network.

$$\theta_i = \theta_0 + \varepsilon_i \qquad C(\vec{\theta}) = \sum_k r_k^2(\{m_i\}) \qquad m_i = \sum_j \varepsilon_j^i$$

$$J_{ij} = \frac{\partial r_i}{\partial \theta_j} = \sum_k \frac{\partial r_i}{\partial m_k} k\varepsilon_j^{k-1} = A_{ik}V_{kj}$$

$$H = J^T J = V^T A^T A V$$

$$V = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_N \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_1^d & \varepsilon_2^d & \cdots & \varepsilon_N^d \end{bmatrix}$$

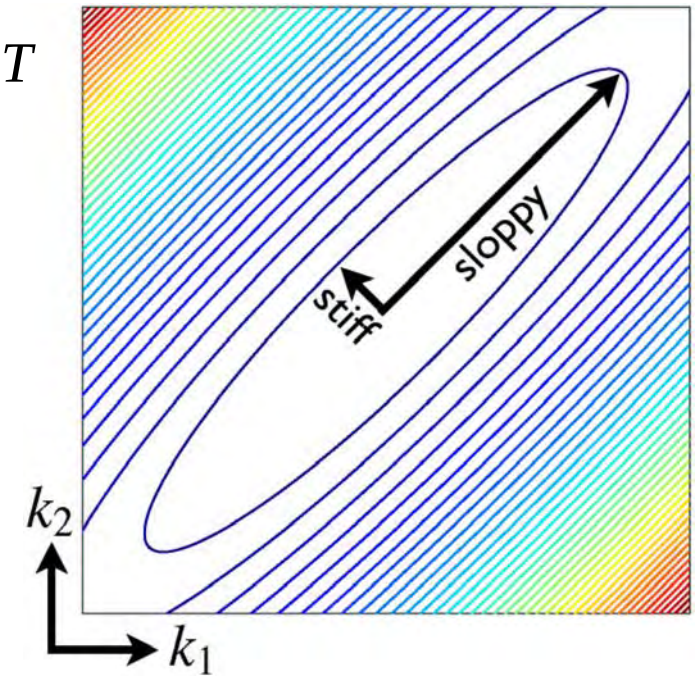
$$\det(V) = \prod_{i<j} (\varepsilon_i - \varepsilon_j) \propto \varepsilon^{N(N-1)/2}$$

The Vandermonde Ensemble

$$H = J^T J = V^T A^T A V = W \tilde{H} W^T$$

$$\Sigma \quad U^T A^T A U \quad \Sigma \quad = \quad \tilde{H}$$

$$\begin{bmatrix} \varepsilon^0 & & \\ & \ddots & \\ & & \varepsilon^{n-1} \end{bmatrix} \begin{bmatrix} & & \\ & \blacksquare & \\ & & \end{bmatrix} \begin{bmatrix} \varepsilon^0 & & \\ & \ddots & \\ & & \varepsilon^{n-1} \end{bmatrix} = \begin{bmatrix} & & \\ & /// & \\ & & \end{bmatrix}$$



$$\tilde{H}_{ij} = O(\varepsilon^{i+j-2})$$

$$\tilde{\lambda}^{(i)} = O(\varepsilon^{2(i-1)})$$

$$\tilde{v}_j^{(i)} = \delta_{ij} + O(\varepsilon^{|i-j|})$$

$$\lambda^{(i)} = O(\varepsilon^{2(i-1)})$$

$$v_j^{(i)} = W_{ij} + O(\varepsilon)$$

Conclusions

• Large biological networks are governed by a small number of combinations of parameters.

• The origin of this sloppiness is the **redundant behavior** of the constituent biochemical reactions.

• We have found a **new class of "sloppy" models**, which includes those of biological networks, and are defined by being a Vandermonde transformation of a random symmetric matrix.

Acknowledgements

Fergal Casey
Ryan Gutenkunst
Kevin Brown
Christopher Myers
James Sethna
Veit Elser
Piet Brouwer