# The Many Applications of Eigenvalues

David Bindel

7 Feb 2019

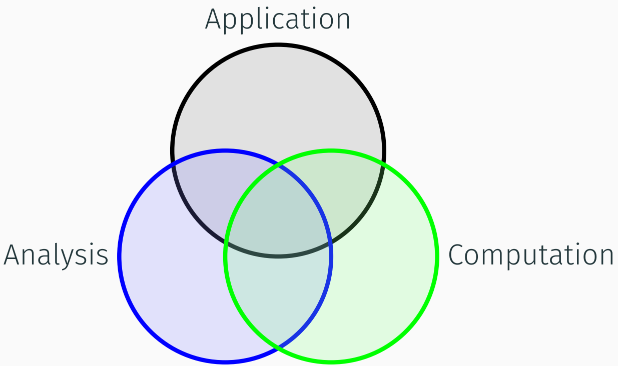
# My Goals for Today


`bindel@cornell.edu`

Jones 317 (mostly until mid-May)


- Show how applied math happens (to me at Cornell).
- Convince you that eigenvalue problems are fun!
- Get you to talk to me, read slides, read papers, etc. (And maybe apply to Cornell for grad school!)

# The Computational Science & Engineering Picture




- MEMS
- Fusion
- Networks
- Systems

- Linear algebra
- Approximation theory
- Symmetry + structure
- Optimization

- HPC / cloud
- Simulators
- Solvers
- Frameworks

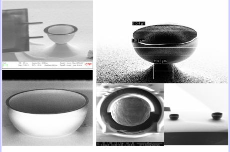
# Today: Eigenvalue Problems

*My super power is turning everything you show me into an eigenvalue problem.*

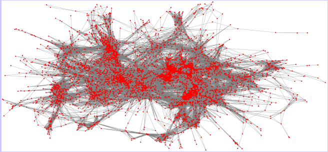
*— Me (at every new grad student lunch)*

# Why Eigenvalue Problems?

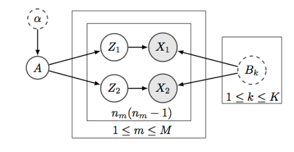
## Dynamics



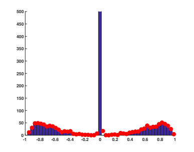
## Optimization



## Data approximation




## Densities and invariants



# Why Eigenvalue Problems?

Dynamics: $\frac{du}{dt} = Au$ or $u(k+1) = Au(k)$

Optimization: minimize $x^T A x$ s.t. $x^T x = 1$

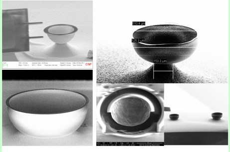
Data approximation: minimize $\|A - XY^T\|_F^2$

Invariants: $\forall$ analytic $f : \mathbb{C} \to \mathbb{C}$, compute $\mathrm{tr}(f(A))$

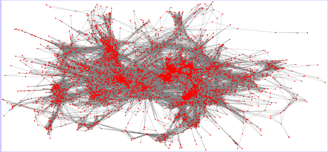
All these perspectives are connected!

# Why Eigenvalue Problems?

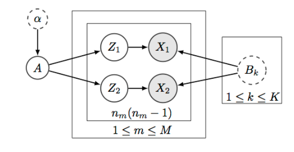
## Dynamics



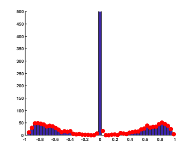
## Optimization

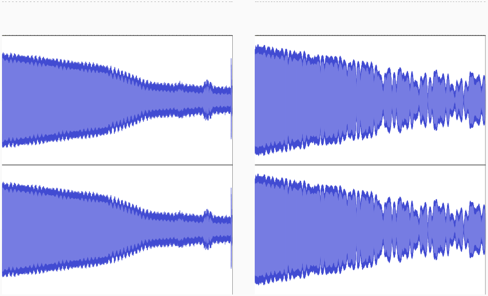


## Data approximation



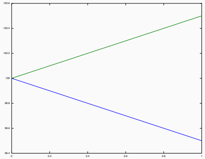

## Densities and invariants

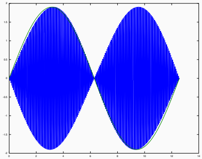


“On the beats in the vibrations of a revolving cylinder or bell”
by G. H. Bryan, 1890

# The Beat Goes On





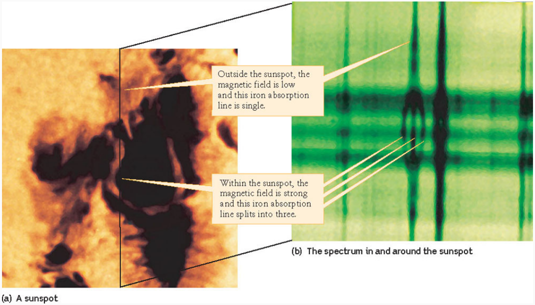
Free vibrations in a rotating frame (simplified):

$$\ddot{\mathbf{q}} + 2\beta\Omega\mathbf{J}\dot{\mathbf{q}} + \omega_0^2\mathbf{q} = 0, \qquad \mathbf{J} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Eigenvalue problem: $\left(-\omega^2\mathbf{I} + 2i\omega\beta\Omega\mathbf{J} + \omega_0^2\right) q = 0.$

Solutions: $\omega \approx \Omega_0 \pm \beta\Omega. \implies$ beating $\propto \Omega$!

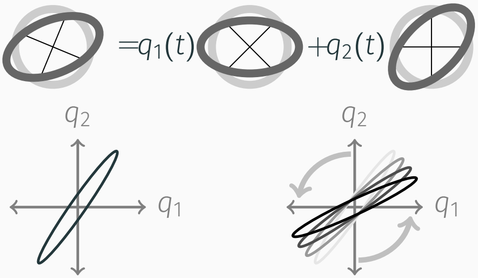
# Bryan, Zeeman, Stark, ...




This is a common picture:

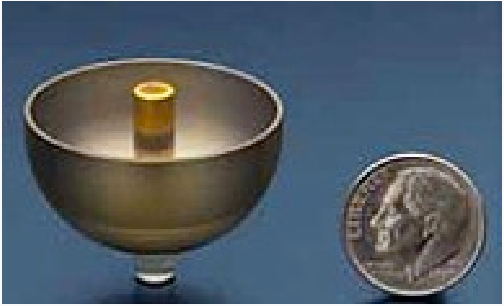
- Symmetry leads to degenerate modes
- Perturbations split (some) degeneracies

## A General Picture



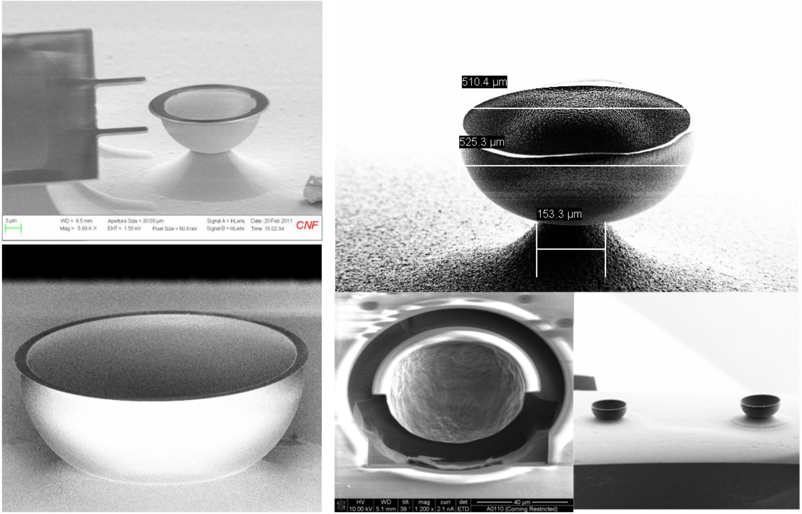

$$\begin{bmatrix} q_1(t) \\ q_2(t) \end{bmatrix} \approx \begin{bmatrix} \cos(-\beta\Omega t) & -\sin(-\beta\Omega t) \\ \sin(-\beta\Omega t) & \cos(-\beta\Omega t) \end{bmatrix} \begin{bmatrix} q_1^0(t) \\ q_2^0(t) \end{bmatrix}.$$

# Foucault in Solid State

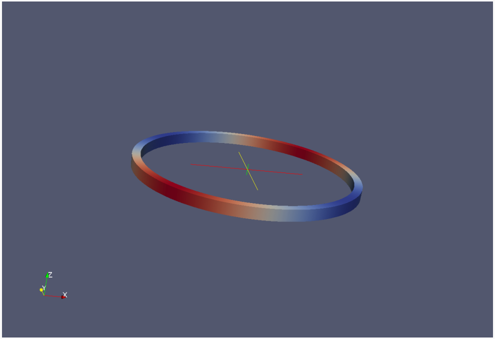


Northrup-Grummond HRG
(developed c. 1965–early 1990s)

# A Smaller Application (Cornell)



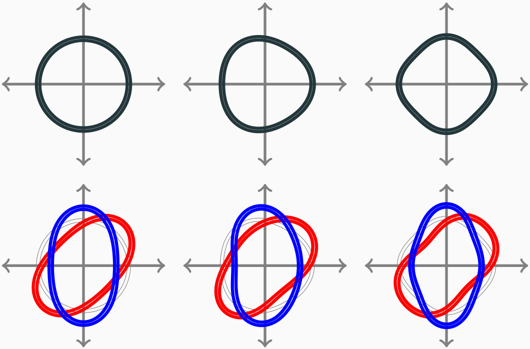
# The Perturbation Picture




Perturbations split degenerate modes:

- Coriolis forces (good)
- Imperfect fab (bad, but physical)
- Discretization error (non-physical)

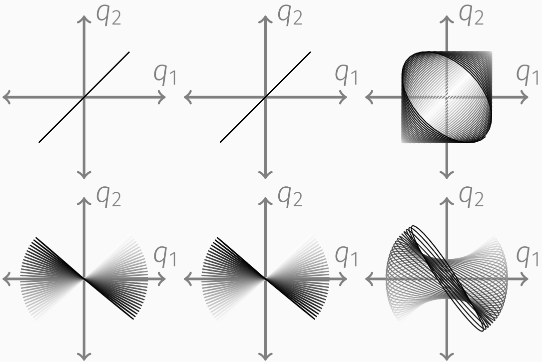
# Analyzing Imperfections



Basic framework:

- Represent geometry and imperfections in Fourier series
- Treat imperfections as perturbations

# Analyzing Imperfections




Payoff:

- Quantitative: Fast and accurate "2.5D" simulations
- Qualitative: *Selection rules* identify "dangerous" imperfections

# More

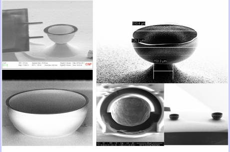
Yilmaz and Bindel
"Effects of Imperfections on Solid-Wave Gyroscope Dynamics"
Proceedings of IEEE Sensors 2013, Nov 3–6.

Or talk to me about:

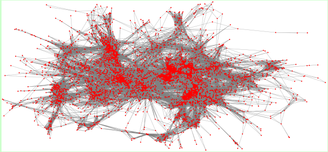
- Damping, radiation, and nonlinear eigenproblems in MEMS
- Nonlinear dynamics in MEMS (ongoing!)

# Why Eigenvalue Problems?

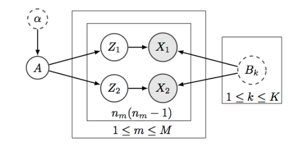
## Dynamics



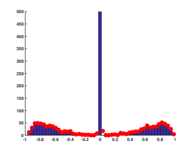
## Optimization

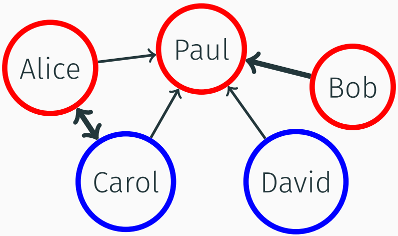


## Data approximation




## Densities and invariants




Alice
Paul
Bob
Carol
David

# Modeling Opinion Formation

A basic model:

- A fixed *intrinsic* opinion $s_i$
- A variable *expressed* opinion $x_i$
- Equilibrium $x_i = \text{argmin}_{z_i} c_i(z_i)$, where

$$c_i(z_i) \equiv (s_i - z_i)^2 + \sum_{j \in N(i)} w_{ij}(z_i - x_j)^2$$

- Define a *social cost* $c(z) = \sum_i c_i(z_i)$

# From Networks to Numerical Linear Algebra

**Methodology:** Graph problem $\mapsto$ linear algebra problem.

| | |
|---|---|
| Nash equilibrium: | $(L + I)x = s$ |
| Social optimum: | $(A + I)y = s$ |
| Cost at equilibrium: | $c(x) = s^T C s$ |
| Optimal social cost: | $c(y) = s^T B s$ |

*Price of anarchy* is a ratio of quadratics:

$$\mathrm{PoA}(s) = \frac{c(x)}{c(y)} = \frac{s^T C s}{s^T B s}$$

# Enter eigenvalues

Given

$$\mathrm{PoA}(s) = \frac{s^T C s}{s^T B s}$$

Maximize by setting gradient to zero:

$$\nabla_s \mathrm{PoA}(s) = \frac{2}{s^T B s}\left[Cs - \mathrm{PoA}(s)\, Bs\right] = 0$$

Find worst case through a *a generalized eigenvalue problem*:

$$Cs_* = \lambda B s_*$$

# How this happened

- Sigal Oren: Jon Kleinberg and I are working on this problem, he suggested you might have some insight [explains]. So why is PoA always bounded by 9/8 for symmetric networks?
- DB: OK
    - PoA is a generalized eigenvalue.
    - Matrices are $B = p(L)$ and $C = q(L)$
    - Eigs are $p(\mu)/q(\mu)$ for $\mu$ an eig of $L$
    - $p(\mu)/q(\mu)$ has a max of 9/8 for $\mu \geq 0$.
- SO: Great, thanks! [Exit office]
- — Ten minutes pass —
- SO (knocks): So what about nonsymmetric networks?

# More

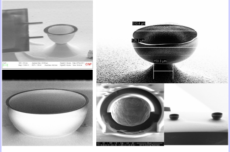
Bindel, Kleinberg, Oren
"How Bad is Forming Your Own Opinion?"
Games and Economic Behavior, vol 92, pp. 248–265, 2015.

Or talk to me about:

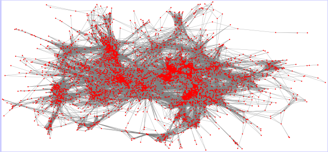
- Similar bounds for 3D image reconstruction!
- Spectral methods for community detection
- Fast parameterized PageRank computations

# Why Eigenvalue Problems?

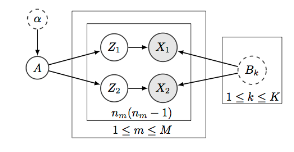
## Dynamics



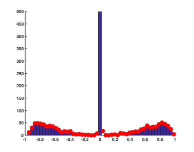
## Optimization

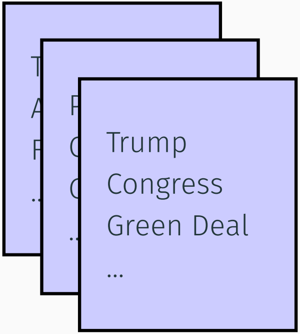


## Data approximation

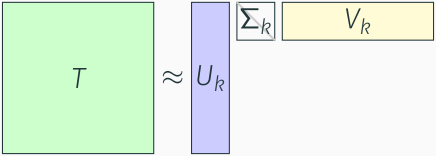



## Densities and invariants




Trump
Congress
Green Deal
...

# Old idea: Latent Semantic Indexing




- Documents as a word count vectors (“bag of words”)
- Reweight to account for frequency (tf-idf)
- Compute *singular value decomposition* and truncate
  - Gives best rank $k$ approximation to $T$
- Cluster words/docs via $U_k$ and $V_k$
  - Rows for similar documents are similar
  - “Blurs out” related terms (car/automobile)
- But hard to interpret rows of $U_k$ / cols of $V_k$
  - May have negative entries, not normalized

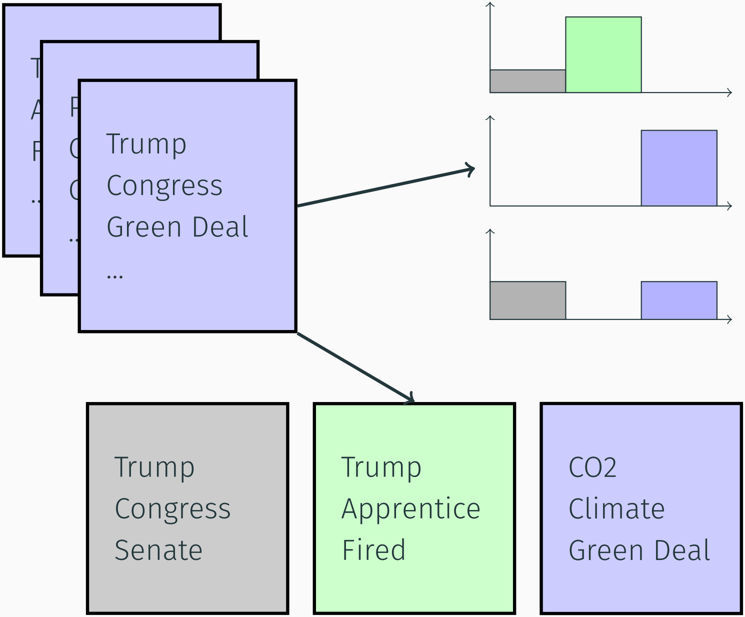
# Latent Dirichlet Allocation (LDA)

A *generative* model for documents:

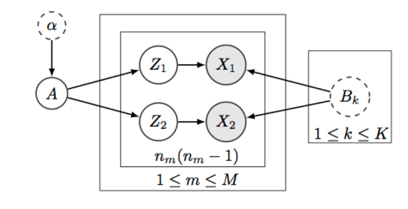
- Topics are distributions over words
- Documents involve distribution over topics
- Generate document by picking topic, then word from topic

Goal: Jointly determine topic and document distributions.

# Topic Modeling and LDA



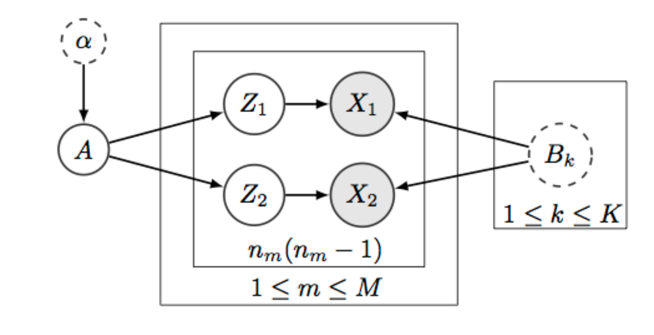
# Beyond LDA




Ex: "A Practical Algorithm for Topic Modeling with Provable Guarantees." Arora *et al*, ICML 2013

- Work with *word co-occurrence statistics* (topics only)
- Assume *anchor words* for each topic
- Much faster than MCMC-based LDA training (NLA-based)
- Provable guarantees *with enough data from model*

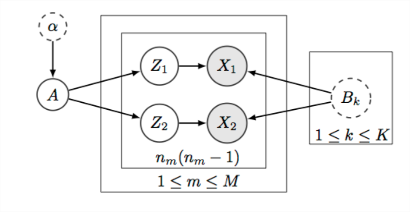
# Beyond LDA




But — this is not how we write documents!

- Co-occurrence may not behave as model predicts
- Result: sometimes funky topics for real data

# Rectification

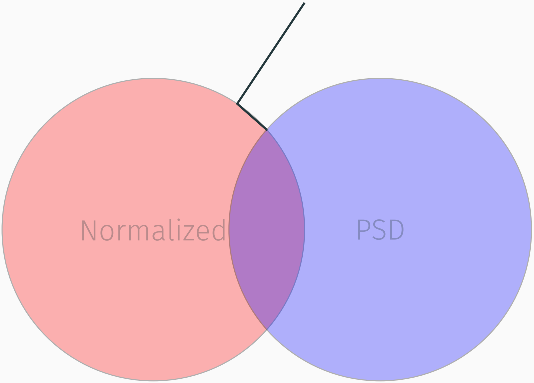



Idea: Enforce co-occurrence structure under model

- Should represent probability (non-negative, sums to 1)
- Should be low rank and positive semi-definite

Algorithm: Alternating projections

# Alternating projections




- Alternate PSD-rank-$k$ and normalized matrix projections
- PSD-rank-$k$ projection by partial eigendecomposition
- Can compute fast using only matrix-vector products
- Run inference on the resulting matrix

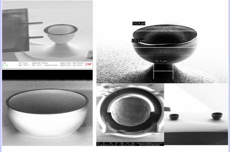
# More

Lee, Bindel, and Mimno,
"Robust Spectral Inference for Joint Stochastic Matrix Factorization," NIPS 2015

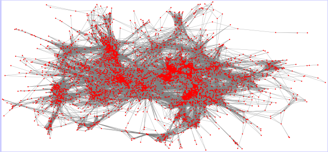
- Still some ongoing work in this direction!
- Moontae Lee is now faculty at the UIC business school

# Why Eigenvalue Problems?

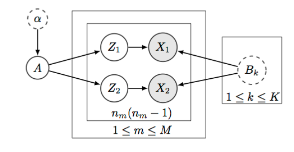
## Dynamics



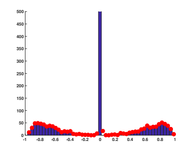
## Optimization



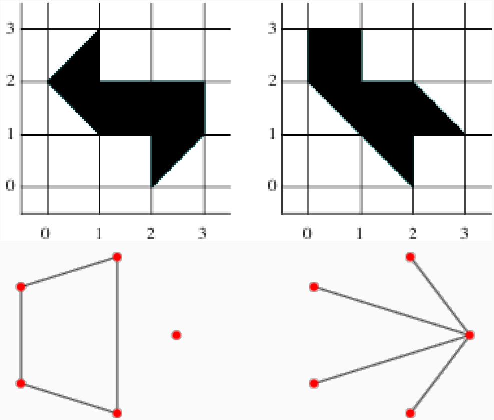
## Data approximation




## Densities and invariants

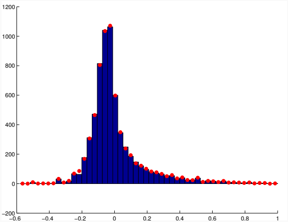


# Chapter 4: Can One Hear the Shape of a Drum?




*"You mean, if you had perfect pitch could you find the shape of a drum."* — Mark Kac (quoting Lipmann Bers)
American Math Monthly, *1966*

# Another Perspective: Density of States




Spectra define a *generalized function* (a *density*):

$$\operatorname{tr}(f(H)) = \int f(\lambda)\mu(\lambda)\,dx = \sum_{j=k}^{N} f(\lambda_k)$$

where $f$ is an analytic test function. Smooth to get a picture: a *spectral histogram* or *kernel density estimate*.

# A Bestiary of Matrices

- Adjacency matrix: $A$
- Laplacian matrix: $L = D - A$
- Unsigned Laplacian: $L = D + A$
- **Random walk matrix**: $P = AD^{-1}$ (or $D^{-1}A$)
- **Normalized adjacency**: $\bar{A} = D^{-1/2}AD^{-1/2}$
- **Normalized Laplacian**: $\bar{L} = I - \bar{A} = D^{-1/2}LD^{-1/2}$
- Modularity matrix: $B = A - \frac{dd^T}{2n}$
- Motif adjacency: $W = A^2 \odot A$

All have examples of co-spectral graphs

… through spectrum uniquely identifies *quantum graphs*

# Example: Estrada Index

Consider

$$\text{tr}(\exp(\alpha A)) = \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} \cdot (\text{\# closed random walks of length } k)\,.$$

- Global measure of connectivity in a graph.
- Can clearly be computed via DoS.
- Generalizes to other weights.

# Heat Kernels

DoS information equivalent to looking at the *heat kernel trace*:

$$h(s) = \mathrm{tr}(\exp(-sH)) = \mathcal{L}[\mu](s)$$

Use $H = LD^{-1}$ (continuous time random walk generator) $\implies$ $h(s)/N = P$(self-return after time $s$ from uniform start).

# Power Moments

DoS information equivalent to looking at the *power moments*:

$$\mathrm{tr}(H^j).$$

Natural interpretation for matrices associated with graphs:

- $A$: number of length $k$ cycles.
- $\bar{A}$ or $P$: return probability for $k$-step random walk (times $N$).
- $L$: ??

# Local DoS

*Local* DoS $\nu_k(x)$: symmetric case with $H = Q\Lambda Q^T$,

$$\int f(x)\nu_k(x)\,dx = f(H)_{kk} = e_k^T Q f(\Lambda) Q^T e_k$$

$$\nu_k(x) = \sum_{j=1}^{n} q_{kj}^2\,\delta(x - \lambda_j)$$

DoS is sum of local densities of states:

$$\mu(x) = \sum_{k=1}^{n} \nu_k(x)$$

# LDoS Information

Can compute common *centrality measures* with LDoS

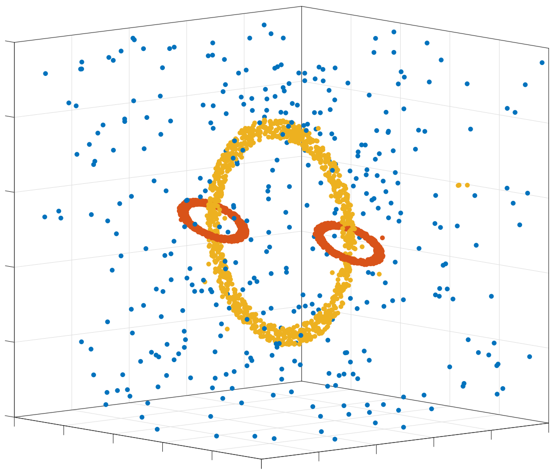
- Estrada centrality: $\exp(\gamma A)_{kk}$
- Resolvent centrality: $\left[(I - \gamma \bar{A})^{-1}\right]_{kk}$

Some motifs associated with localized eigenvectors:

- Chief example: Null vectors of $\bar{A}$ supported on leaves.
- Use LDoS + topology to find motifs?

What else?

# LDoS and Clustering



# Phase Retrieval in Graph Reconstruction

Reconstruct graph from *fully resolved* LDoS at all nodes?

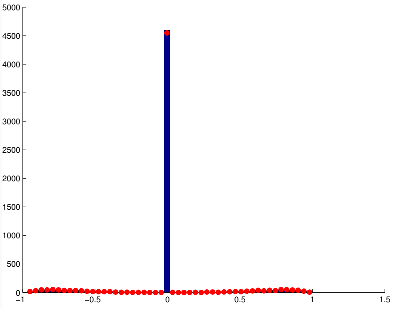
- Assume $H = Q\Lambda Q^T$
- No multiple eigenvalues $\implies$ know $|Q|$ and $\Lambda$
- Can we recover signs in $Q$?

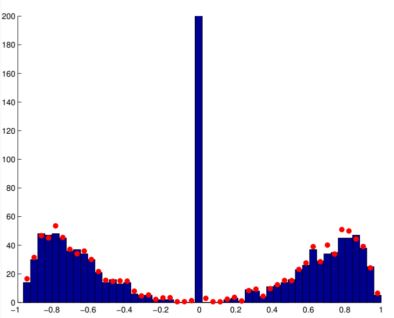
Feels a little like phase retrieval...

# Computing the (L)DoS?

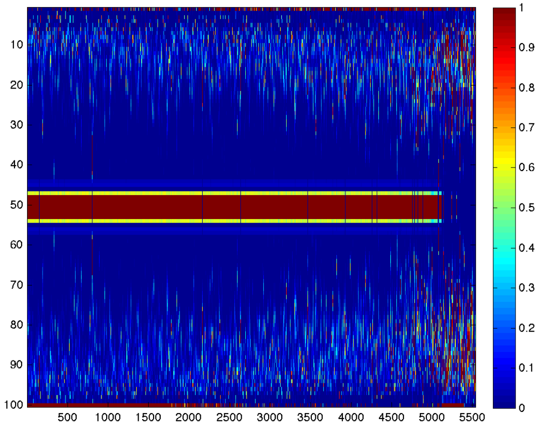
- *Kernel Polynomial Method* (KPM) from physics
  - Expand density of $H$ in a (dual) Chebyshev series
  - Coefficients look like $d_j = \mathrm{tr}(T_j(H))$
  - Use stochastic trace estimation for fast traces
  - Filtering to kill Gibbs oscillations
- Other related methods (e.g. Golub-Meurant GQL)
- Got into this by knowing KPM and a chat with David Gleich!
- Some additional tricks for graph case
- Not enough time for details – let's look at pictures!

# Erdos


5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
−1
−0.5
0
0.5
1
1.5

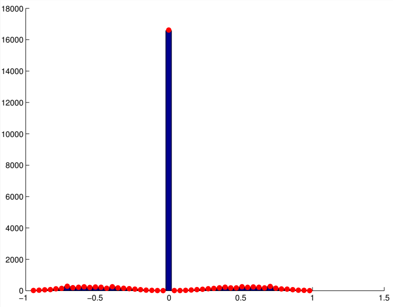

200
180
160
140
120
100
80
60
40
20
0
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1

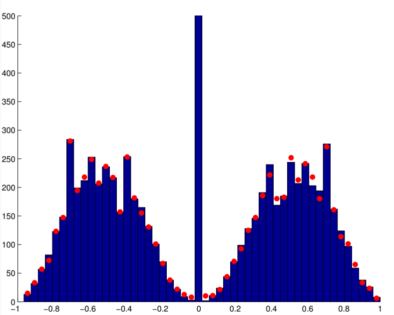
# Erdos (local)



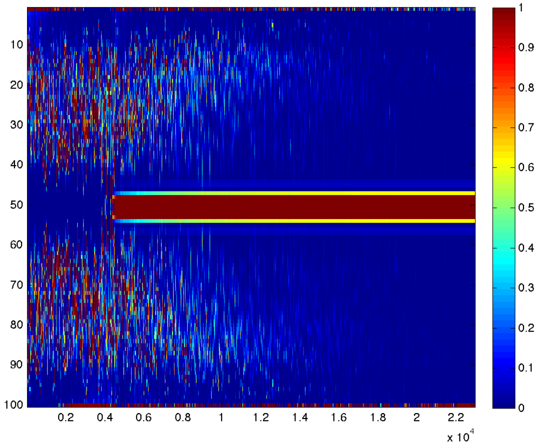
10
20
30
40
50
60
70
80
90
100
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

# Internet topology

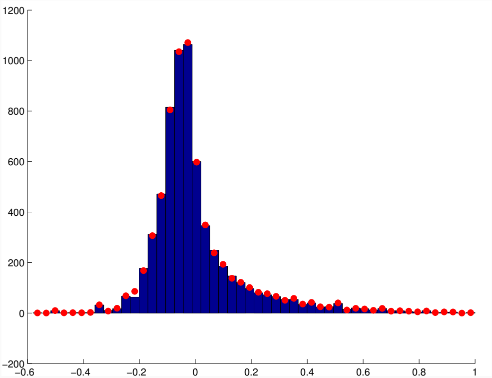


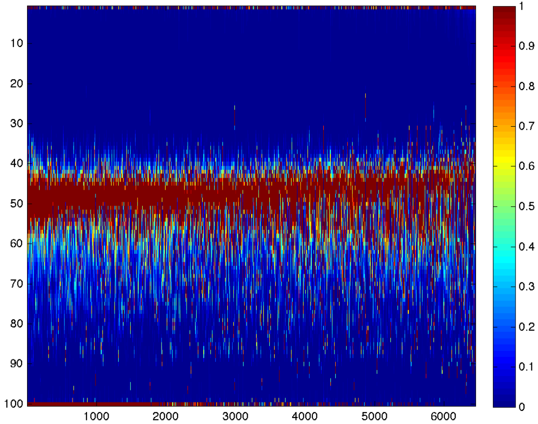
# Internet topology (local)



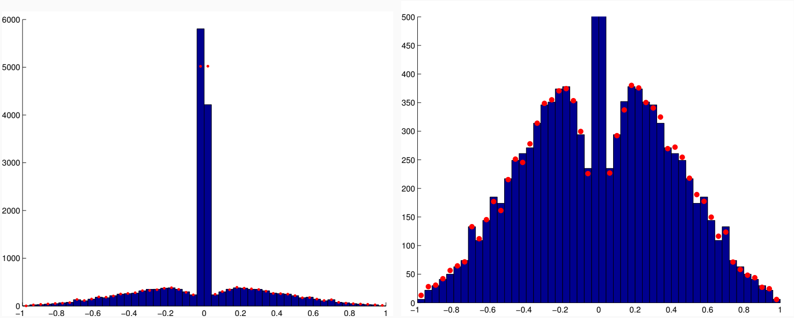
# Marvel characters



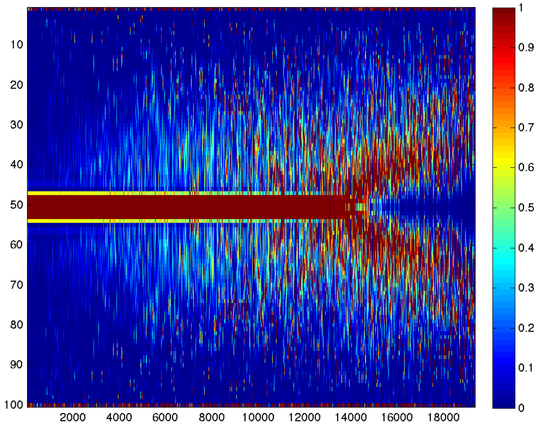
# Marvel characters (local)



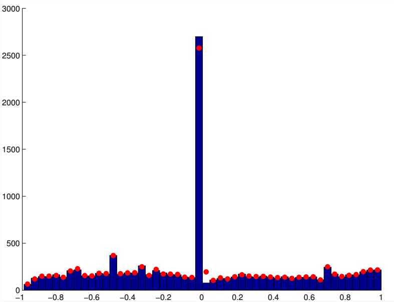
# Marvel comics



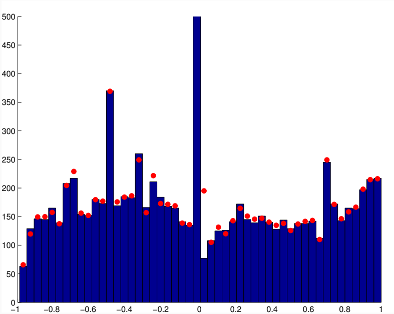
# Marvel comics (local)



# PGP

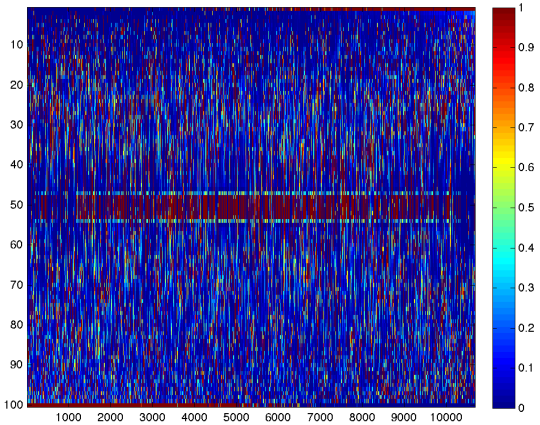


3000
2500
2000
1500
1000
500
0
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1

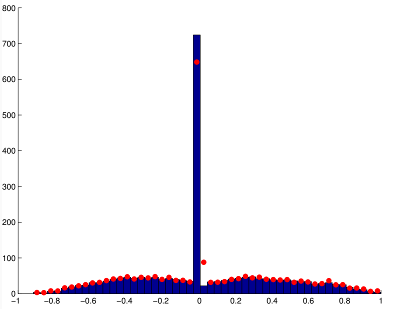
500
450
400
350
300
250
200
150
100
50
0
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1

# PGP (local)

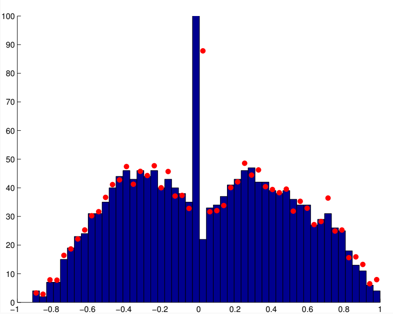


10
20
30
40
50
60
70
80
90
100
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

# Yeast

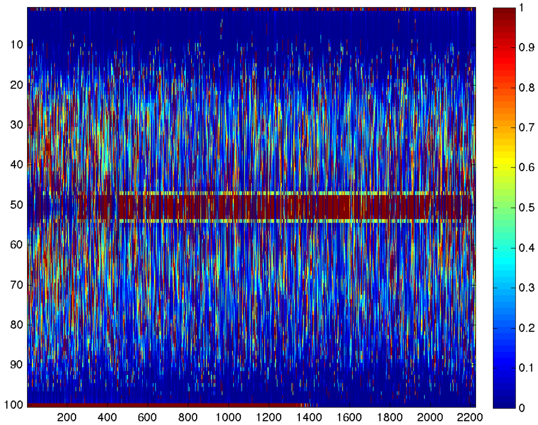


800
700
600
500
400
300
200
100
0
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1

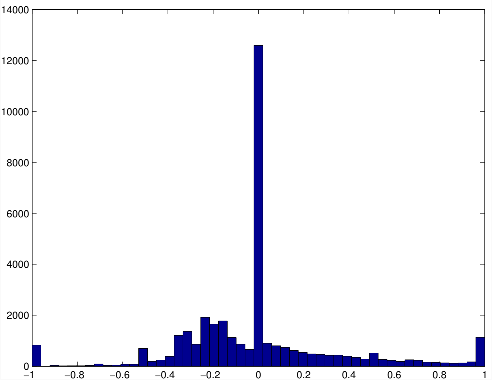
100
90
80
70
60
50
40
30
20
10
0
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1

# Yeast (local)

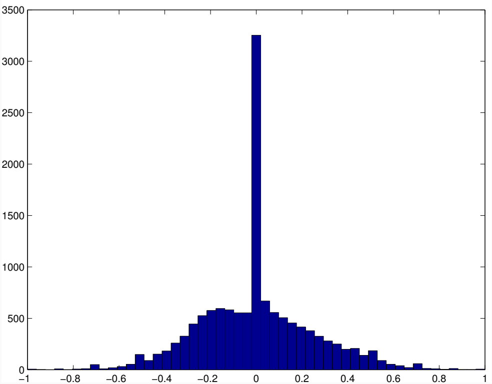


10
20
30
40
50
60
70
80
90
100
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

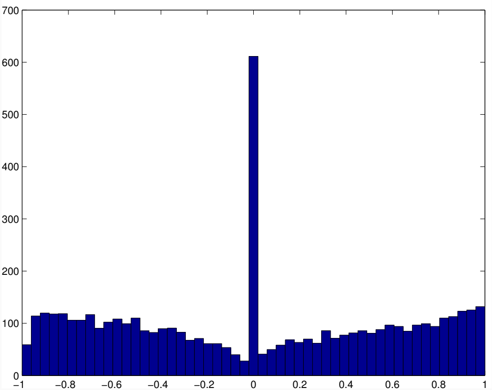
# Enron emails (SNAP)



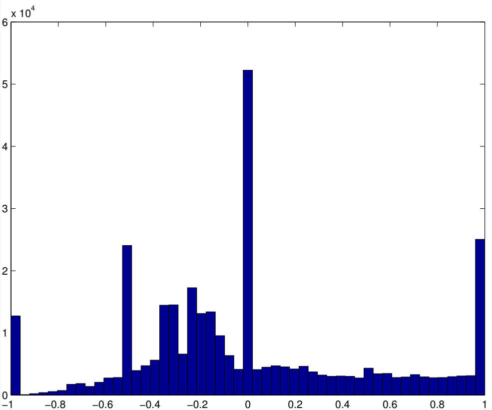
# Reuters911 (Pajek)



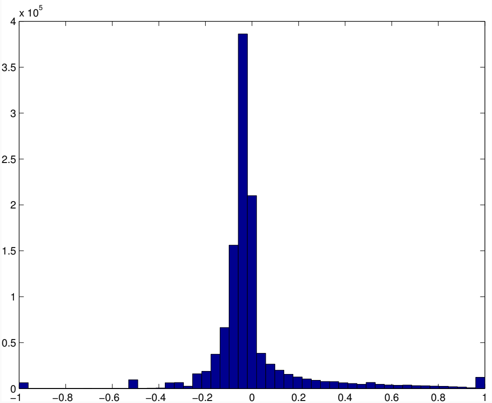
# US power grid (Pajek)



# DBLP 2010 (LAW)

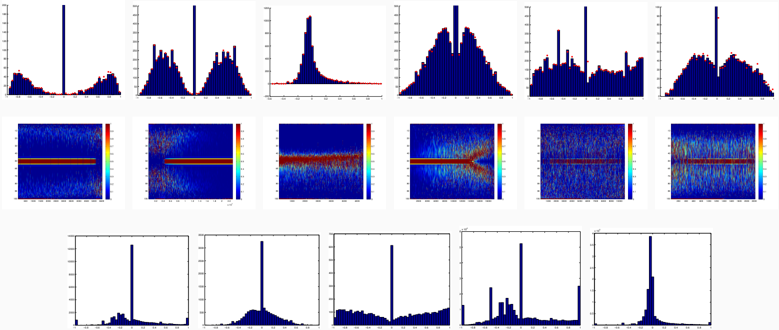



$N = 326186$, $nnz = 1615400$, 80 s (1000 moments, 10 probes)




$N = 1139905$, $nnz = 113891327$, 2093 s (1000 moments, 10 probes)

# What Do You Hear?



# For more...

http://www.cs.cornell.edu/bindel

bindel@cornell.edu

Jones 317 (mostly until mid-May)