Network Density of States

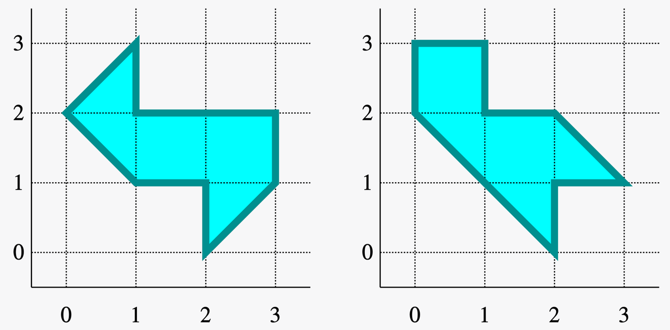
Kun Dong[1] Austin R. Benson[2] David Bindel[2]

Applied Math[1], CS[2]

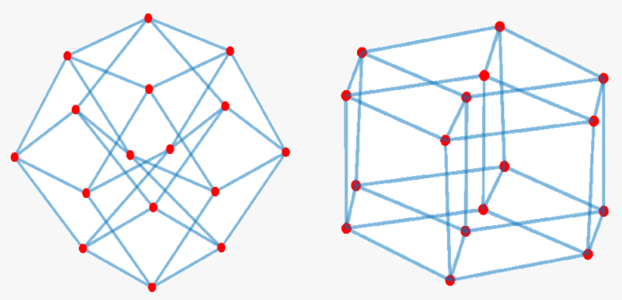
Cornell University

Background

- Weyl's Law connects the geometric properties of a manifold with the asymptotic spectral distribution of Laplace-Beltrami operator
- Eigenvalues alone do not uniquely identify an object, but what can we learn?
  - How do we compute spectrum for large real-world graphs?
  - What geometric properties can we infer from graph spectrum?
  - Do non-extremal eigenpairs contain valuable information?



(a) Co-spectral Domains



(b) Co-spectral Graphs

"Can you hear the shape of a drum?" - Mark Kac

Density of States

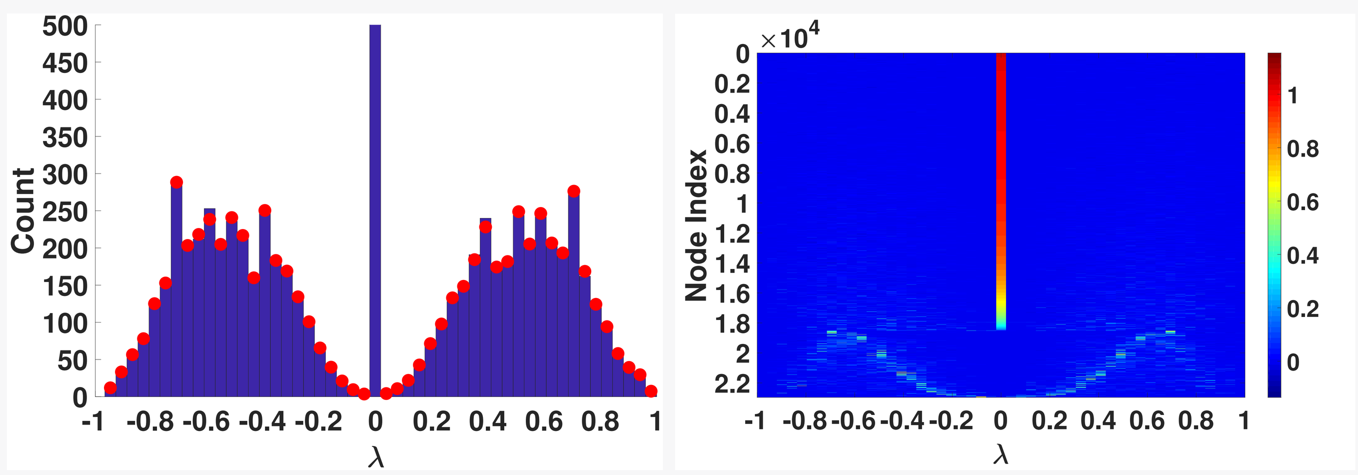
- Given a symmetric graph matrix $H \in \mathbb{R}^{N\times N}$, such as normalized adjacency
- It has an eigendecomposition $H = Q\Lambda Q^T$
- $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_N)$, and $Q = [q_1, \cdots, q_N]$ orthogonal
- Spectral density, also known as *density of states* (DOS), is

$$\mu(\lambda) = \frac{1}{N}\sum_{i=1}^{N}\delta(\lambda - \lambda_i)$$

- Pointwise density of states (PDOS) is

$$\mu_k(\lambda) = \mu(\lambda; e_k) = \sum_{i=1}^{N} q_i(k)^2 \delta(\lambda - \lambda_i)$$

- Integrate these distributions over equal sized intervals for spectral histogram
  - On the left: Blue bins - actual count; red dot - approximated count
  - On the right: Vertically stacked PDOS spectral histogram. Color for heights



Approximating Density of States with the Kernel Polynomial Method (KPM)

- Chebyshev series: Fourier cosine series with a change of variable

$$T_m(\cos(\theta)) = \cos(m\theta)$$

- Chebyshev polynomials can also be defined with the recurrence

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{m+1}(x) = 2xT_m(x) - T_{m-1}(x)$$

- Expand DOS in the dual Chebyshev basis

$$T_m^*(x) = w(x)T_m(x), \qquad w(x) = \frac{2}{\pi(1+\delta_{0m})\sqrt{1-x^2}}$$

$$\mu(\lambda) = \sum_{m=0}^{\infty} d_m T_m^*(\lambda), \qquad d_m = \int_{-1}^{1} T_m(\lambda)\mu(\lambda)d\lambda = \frac{1}{N}\text{tr}(T_m(H))$$

- Similarly for PDOS: $\mu_k(\lambda) = \sum_{m=0}^{\infty} d_{mk} T_m^*(\lambda), \quad d_{mk} = T_m(H)_{kk}$
- Alternatives: Gauss Quadrature and Lanczos (GQL).

Stochastic Estimators for Approximating Chebyshev Moments

- For a probe vector z with i.i.d entries of mean 0 and variance 1,

$$\mathbb{E}[z^T f(H)z] = \text{tr}(f(H)), \quad \mathbb{E}[z \odot Hz] = \text{diag}(f(H))$$

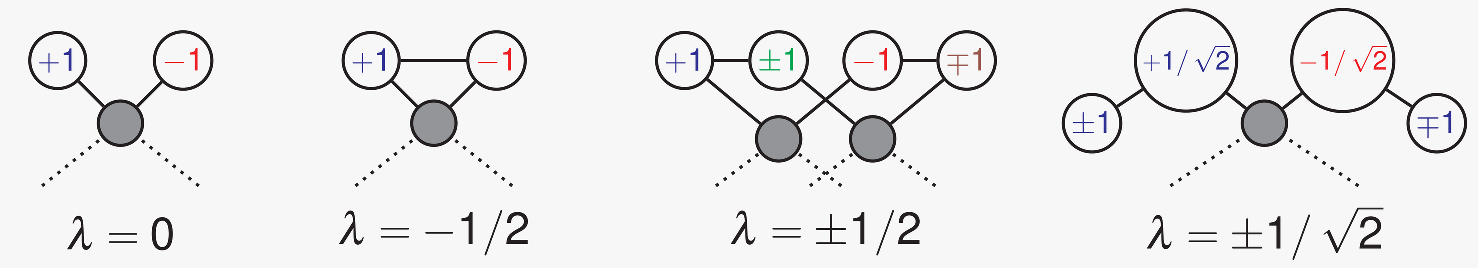
- Choose N_z independent probe vector Z_j's, then

$$\text{tr}(f(H)) \approx \frac{1}{N_z}\sum_{j=1}^{N_z} Z_j^T f(H) Z_j, \quad \text{diag}(f(H)) \approx \frac{1}{N_z}\sum_{j=1}^{N_z} Z_j \odot f(H) Z_j$$

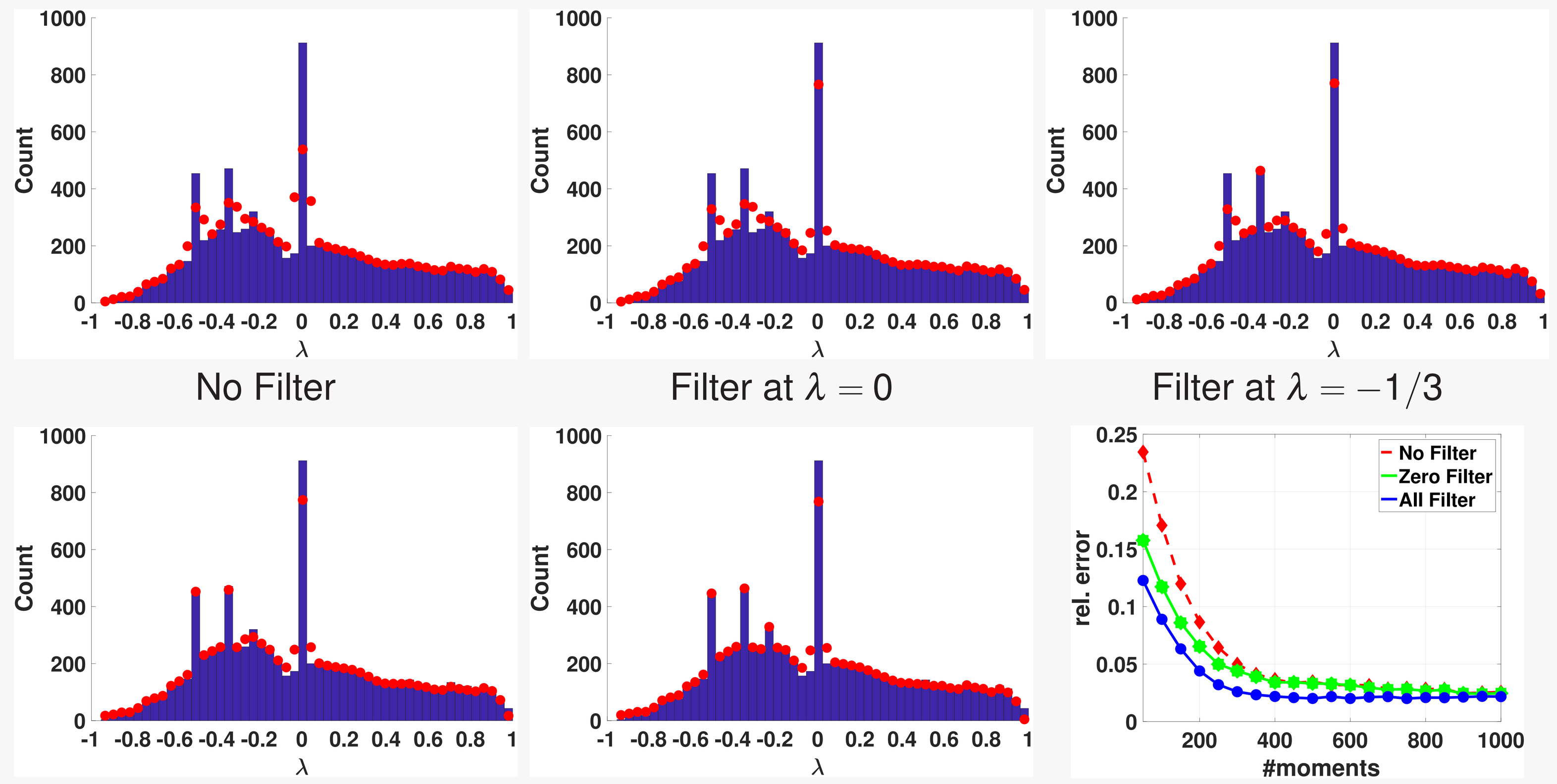
- Use the three-term recurrence to efficiently compute $T_m(H)z$

Efficiently Removing DOS Spikes Coming From Motifs

- Commonly observe spikes in spectrum from eigenvalues of high multiplicity
- They can usually be attributed to local symmetry groups



- Collapse the graph onto the quotient space of these localized eigenvectors
- Example: High Energy Physics Theory Collaboration Network

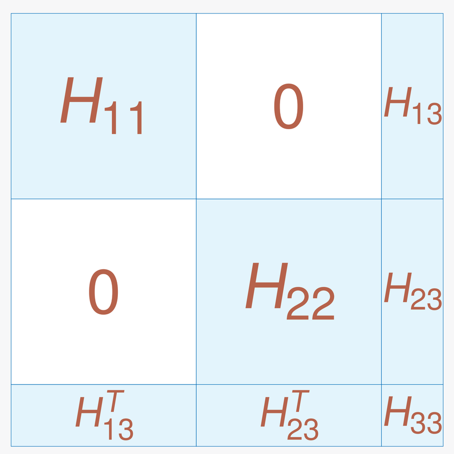


No Filter; Filter at $\lambda = 0$; Filter at $\lambda = -1/3$; Filter at $\lambda = -1/2$; Filter at $\lambda = -1/4$; Relative Error

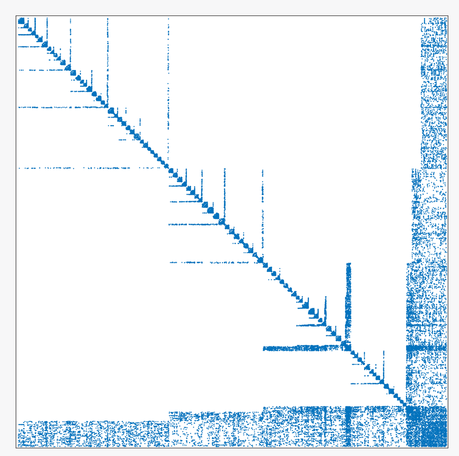
Nested Dissection (ND)

- Previous methods rely on Monte Carlo sampling, which converges at $\mathcal{O}(\frac{1}{\sqrt{N_z}})$
- Alternative: an exact method using nested dissection
- If we can find vertex separators that bisect the graph, the recurrence becomes

$$T_{m+1}(H)_{11} = 2H_{11}T_m(H)_{11} - T_{m-1}(H)_{11} + 2H_{13}T_m(H)_{31}.$$



Vertex Separation



Nested Dissection Ordering

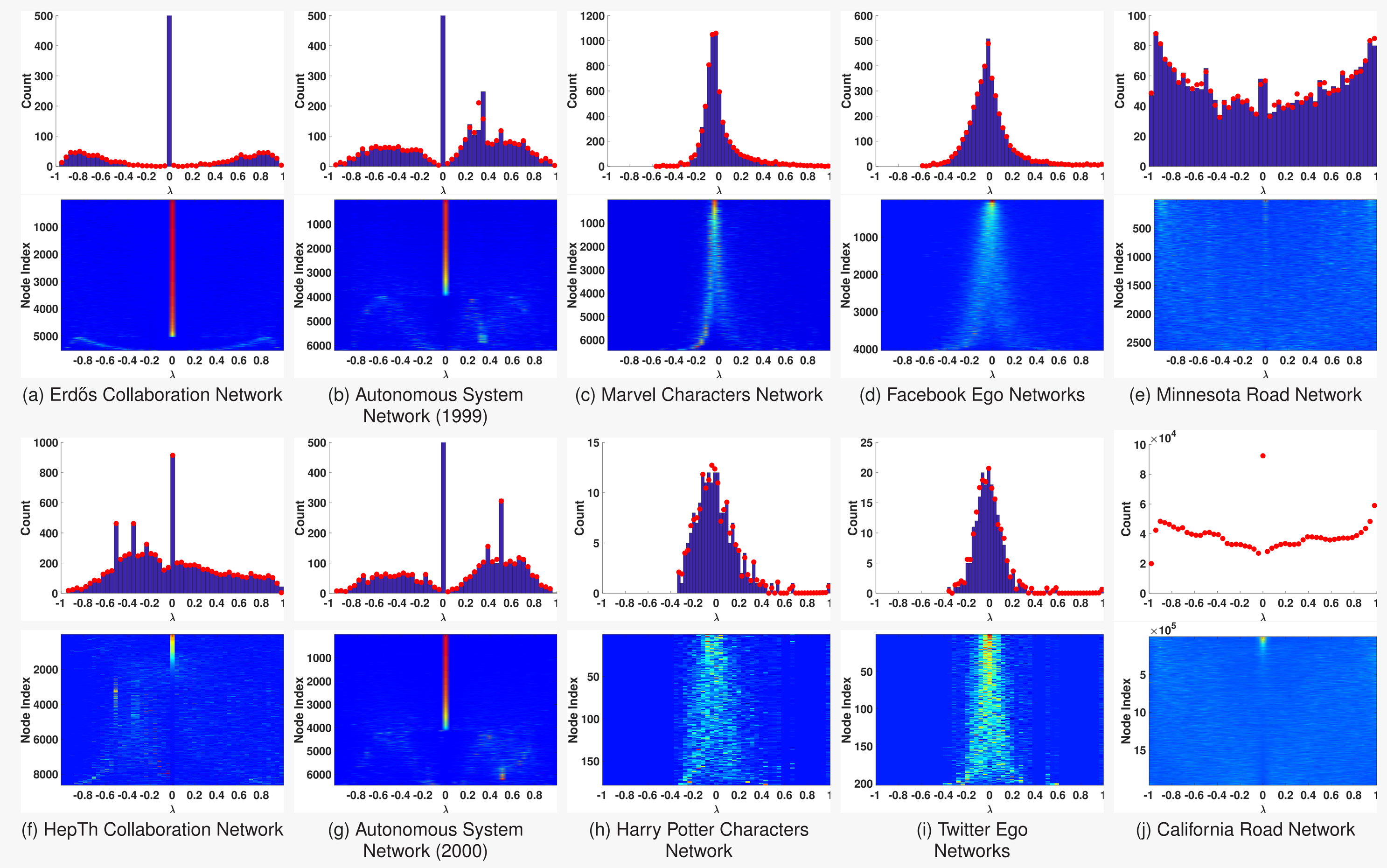
- Hierarchical partition to avoid off-diagonal computation and storage
- First traverse in pre-order for the separator blocks, then fill in the leaf blocks

Comparison between Our Methods

	Exact	ND	KPM		
Network	Time[s]	Time[s]	Time[s]	Rel-L_2	W_1
Musm	0.9	0.9	0.2	3.7e-3	2.3e-2
Yeast	1.8	1.9	0.3	4.3e-3	1.8e-2
Erdős	13.6	4.7	1.9	3.0e-4	4.8e-3
PGP	60.9	11.1	7.9	9.4e-4	2.7e-3
MvlComic	1007.3	57.2	43.0	2.9e-4	1.0e-3
Caida	5915.5	43.2	50.1	1.5e-4	1.0e-4

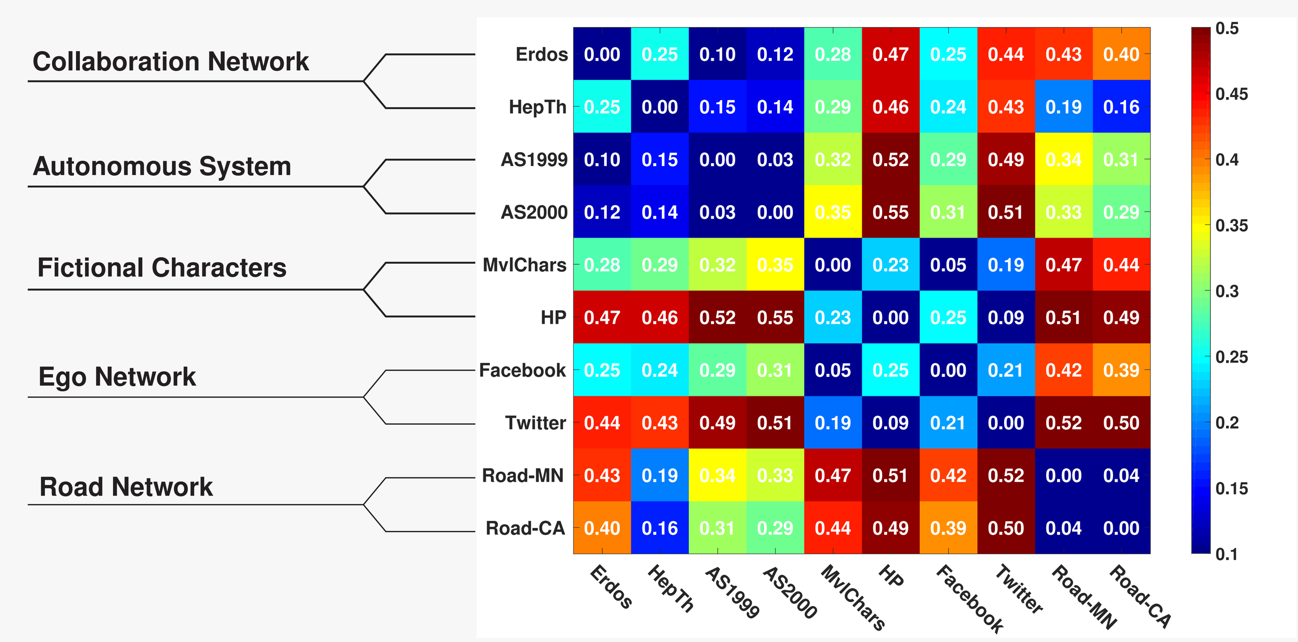
- All methods compute 50 Chebyshev moments
- KPM uses $N/8$ probe vectors
- Rel-L_2 is on the moments; W_1 is the Wasserstein distance on spectral histogram

Gallery of DOS/PDOS



Computed using KPM (500 moments, 20 probe vectors).

DOS Wasserstein Distance Similarity



Computation Time

KPM Single Core Average Computation Time Per Moment				
Network	# Nodes	# Edges	Avg. Deg.	Time (s)
Enron	36,692	183,831	10.02	0.046
Gplus	107,614	13,673,453	254.12	1.133
Amazon	334,863	925,872	5.53	0.628
Neuron	1,018,524	24,735,503	48.57	9.138
RoadNetCA	1,965,206	2,766,607	2.82	2.276
Orkut	3,072,441	117,185,083	76.28	153.7
LiveJournal	3,997,962	34,681,189	17.35	14.52
Friendster	65,608,366	1,806,067,135	55.06	1,017

- Datasets from Stanford Network Analysis Project
- Single Intel Xeon E5 v3 CPU at 2.30GHz with 200GB memory
- Give ample opportunities for parallel computation

Discussion

- We develop practical spectral density computation for real-world networks
- DOS and PDOS facilitate comparisons among graphs and models
- Implementation available at: `https://github.com/kd383/NetworkDOS`