Spectrum Aware Coarsening for Graph Partitioning

Christopher Brissette[1] George Slota[2]

[1]Rensselaer Polytechnic Institute

SIAM CSE23
Optimization, Design, and Machine Learning in the FASTMath SciDAC-5 Institute

March 2023

What is coarsening

Graph Coarsening: Find a representation of a graph $G = (V, E)$ with $|V| = n$, with a reduced dimension $G_c = (V_c, E_c)$ with $|V_c| = n_c < n$.

There are many methods to perform this task.

Method	Abbreviation	Objective
Heavy Weight [1]	HWC	$\underset{(u,v)\in E}{max} \left\{ \min \left\{ \frac{W_{uv}}{d_u}, \frac{W_{uv}}{d_v} \right\} \right\}$
Algebraic Distance [2]	ADC	$\min_{(u,v)\in E} \left\{ \|x_u^k - x_v^k\|_p \right\}$
Adjacency [4]	ASC	$\min_{(u,v)\in E} \left\{ \left\| \frac{W_{u:}}{d_u} - \frac{W_{v:}}{d_v} \right\|_1 \right\}$
Greedy Fiedler [3]	GFSC	—
METIS [5]	METIS	—

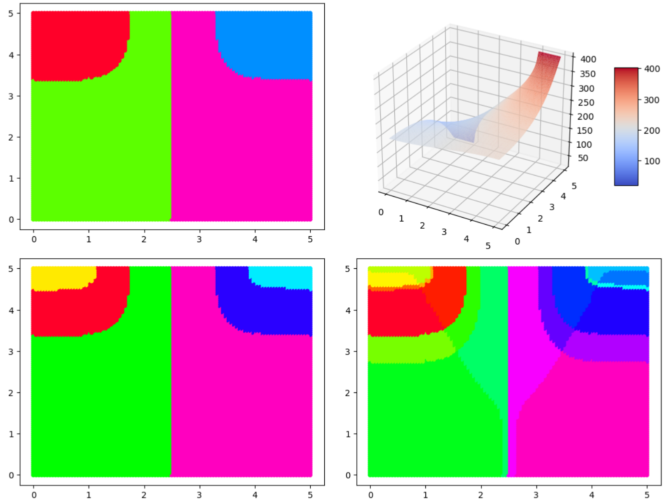
What is coarsening

There are many applications for graph coarsening.

- Reduced order modeling.
- Expediting partitioning.
- Multigrid methods for NLA.
- Even accelerating GNN training!

History

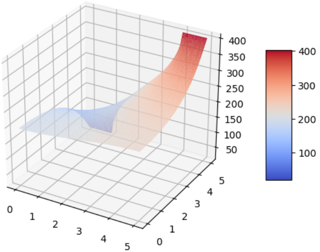
Started on this problem a couple years ago working on a project at Sandia with the goal of finding compact models for semiconductor devices.

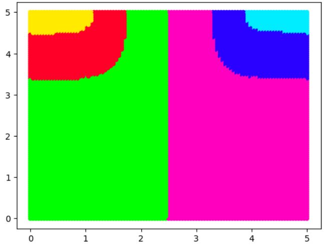


History

This was originally approached with function approximation in mind.

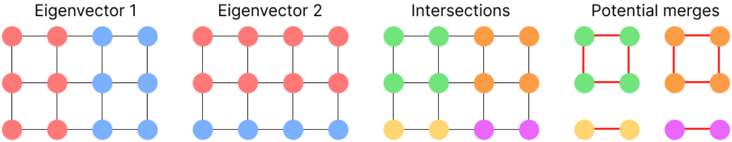
- Intuitively, we desired a partitioning of a mesh which would best approximate it when each partition had its own piece-wise constant value.
- Then from this, "learning" could happen to fill in the functions along partition boundaries.
- We adopted spectral methods for this.





NDC

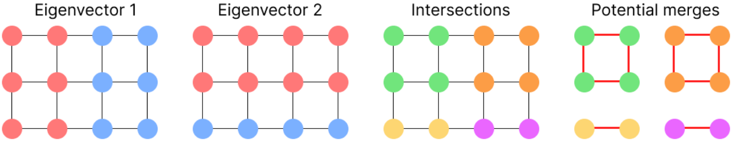
1. Compute the nodal domains of k eigenvectors on G, $\{x_i\}_{i \in [1..k]}$.
2. For each node $u \in V$, assign a vector of signs $\phi_u = [sign(x_0(u)), sign(x_1(u)), \cdots, sign(x_k(u))]$
3. Perform merges according to any chosen additional criteria between nodes $u, v \in V$ such that $\phi_u = \phi_v$.



NDC

Intuitively, the nodal domain coarsening method has some theoretical benefits.

- Any merged nodes have a similar spectral embeddings.
- Can be tacked onto any chosen coarsening method or metric to add global information into our coarsening.



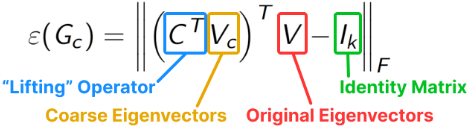
Results

We now compare several coarsening methods, and their spectral approximation properties on two suites of graphs. First, the **irregular dataset**, consisting of a mix of social networks, collaboration networks, road networks, email networks, and voting networks. Secondly the **mesh dataset** consisting of several mesh graphs.

Graph	Nodes	Edges	Type
CA-GrQc	5242	14496	Collaboration
CA-HepTh	9877	25998	Collaboration
email-Eu-core	986	332334	Email
facebook_combined	4039	88234	Social
lastfm_asia_edges	7624	27806	Social
Wiki-Vote	7115	103689	Voting
euroroad	1174	1417	Road
arenas-jazz	198	2277	Social
alligator	3208	9188	Mesh
cheburashka	6669	20001	Mesh
teapot	2259	6250	Mesh
woody	694	1960	Mesh
rocker-arm	10044	30132	Mesh

Eigenvector approximation

We first examine the eigenvector approximation properties of these methods. We do this by examining the inner product of their vector spaces.



$$\varepsilon(G_c) = \left\| \left(\boxed{C^T} \boxed{V_c} \right)^T \boxed{V} - \boxed{I_k} \right\|_F$$

“Lifting” Operator

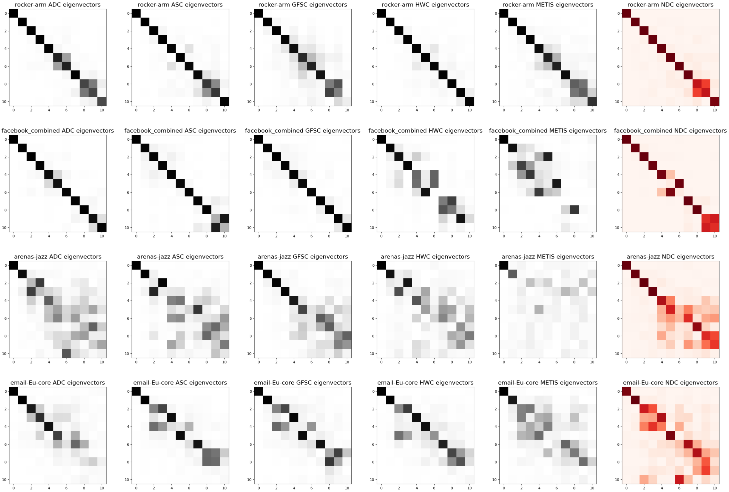
Identity Matrix

Coarse Eigenvectors

Original Eigenvectors

Intuitively, this captures the similarity in V_c and V by measuring the orthognality of their vectors. If orthognality is preserved, the inner product should resemble a diagonal matrix, and this norm should be small.

$\frac{n_c}{n} = 0.5$ coarsening eigenvector comparisons



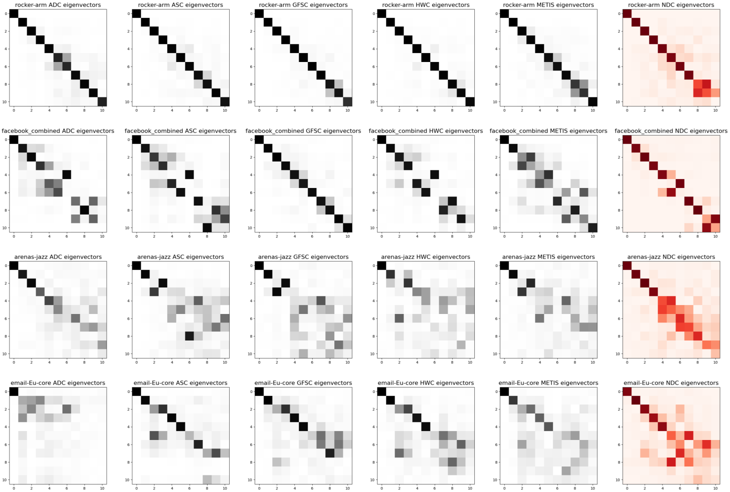
rocker-arm ADC eigenvectors
rocker-arm ASC eigenvectors
rocker-arm GFSC eigenvectors
rocker-arm HWC eigenvectors
rocker-arm METIS eigenvectors
rocker-arm NDC eigenvectors

facebook_combined ADC eigenvectors
facebook_combined ASC eigenvectors
facebook_combined GFSC eigenvectors
facebook_combined HWC eigenvectors
facebook_combined METIS eigenvectors
facebook_combined NDC eigenvectors

arenas-jazz ADC eigenvectors
arenas-jazz ASC eigenvectors
arenas-jazz GFSC eigenvectors
arenas-jazz HWC eigenvectors
arenas-jazz METIS eigenvectors
arenas-jazz NDC eigenvectors

email-Eu-core ADC eigenvectors
email-Eu-core ASC eigenvectors
email-Eu-core GFSC eigenvectors
email-Eu-core HWC eigenvectors
email-Eu-core METIS eigenvectors
email-Eu-core NDC eigenvectors

$\frac{n_c}{n} = 0.2$ coarsening eigenvector comparisons



rocker-arm ADC eigenvectors

rocker-arm ASC eigenvectors

rocker-arm GFSC eigenvectors

rocker-arm HWC eigenvectors

rocker-arm METIS eigenvectors

rocker-arm NDC eigenvectors

facebook_combined ADC eigenvectors

facebook_combined ASC eigenvectors

facebook_combined GFSC eigenvectors

facebook_combined HWC eigenvectors

facebook_combined METIS eigenvectors

facebook_combined NDC eigenvectors

arenas-jazz ADC eigenvectors

arenas-jazz ASC eigenvectors

arenas-jazz GFSC eigenvectors

arenas-jazz HWC eigenvectors

arenas-jazz METIS eigenvectors

arenas-jazz NDC eigenvectors

email-Eu-core ADC eigenvectors

email-Eu-core ASC eigenvectors

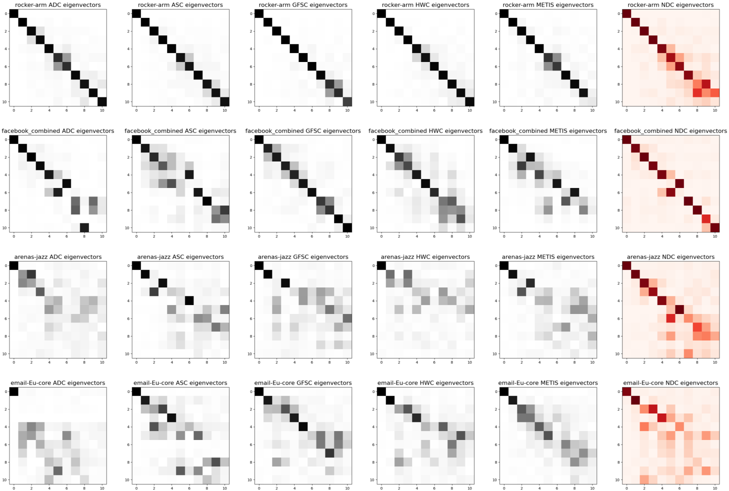
email-Eu-core GFSC eigenvectors

email-Eu-core HWC eigenvectors

email-Eu-core METIS eigenvectors

email-Eu-core NDC eigenvectors

$\frac{n_c}{n} = 0.1$ coarsening eigenvector comparisons



rocker-arm ADC eigenvectors

rocker-arm ASC eigenvectors

rocker-arm GFSC eigenvectors

rocker-arm HWC eigenvectors

rocker-arm METIS eigenvectors

rocker-arm NDC eigenvectors

facebook_combined ADC eigenvectors

facebook_combined ASC eigenvectors

facebook_combined GFSC eigenvectors

facebook_combined HWC eigenvectors

facebook_combined METIS eigenvectors

facebook_combined NDC eigenvectors

arenas-jazz ADC eigenvectors

arenas-jazz ASC eigenvectors

arenas-jazz GFSC eigenvectors

arenas-jazz HWC eigenvectors

arenas-jazz METIS eigenvectors

arenas-jazz NDC eigenvectors

email-Eu-core ADC eigenvectors

email-Eu-core ASC eigenvectors

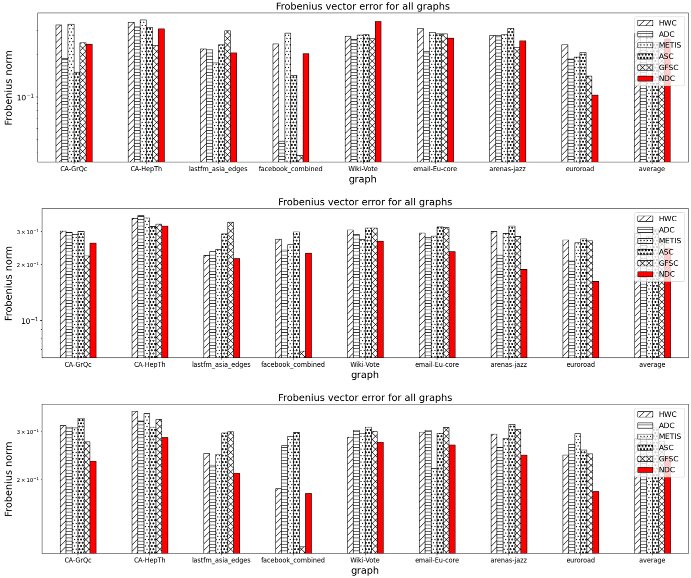
email-Eu-core GFSC eigenvectors

email-Eu-core HWC eigenvectors

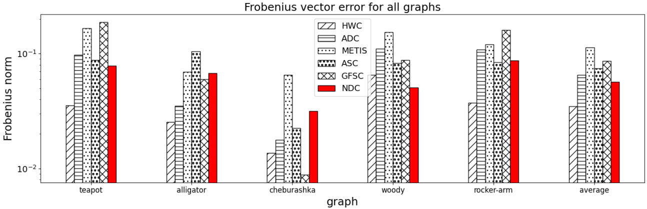
email-Eu-core METIS eigenvectors

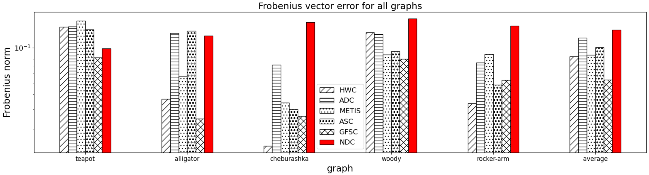
email-Eu-core NDC eigenvectors

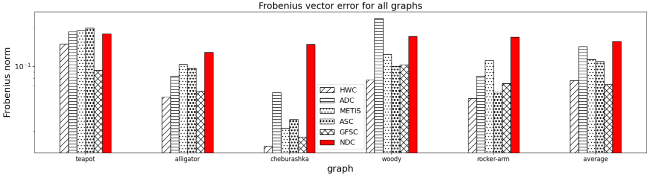
Eigenvector norm comparisons for **irregular dataset**.



Eigenvector norm comparisons for **mesh dataset**.







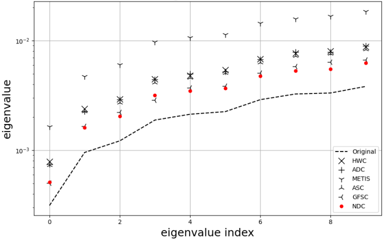
Eigenvalue approximation

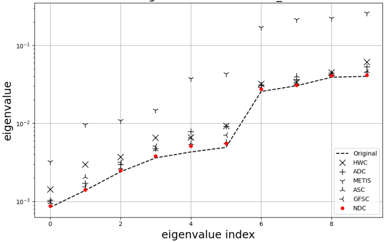
We now look at the eigenvalue approximation properties by examining the average difference in spectra between the coarsened graph and original graph.

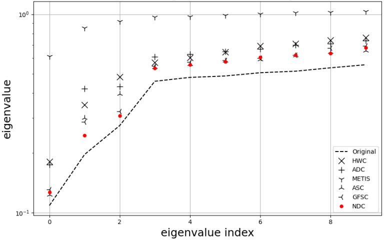
$$\varphi_k(G_c) = \frac{1}{k}\left(\sum_{i=0}^{k} |\lambda_i(G) - \lambda_i(G_c)|\right)$$

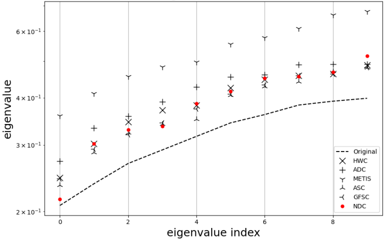
Intuitively this is measuring how "good" the spectral cuts of the coarse graph are in comparison to the spectral cuts of the original graph. Here we only condsider the k smallest eigenvalues because these control cut values. k is taken to be 10 for these experiments.

$\frac{n_c}{n} = 0.5$ coarsening eigenvalue comparisons

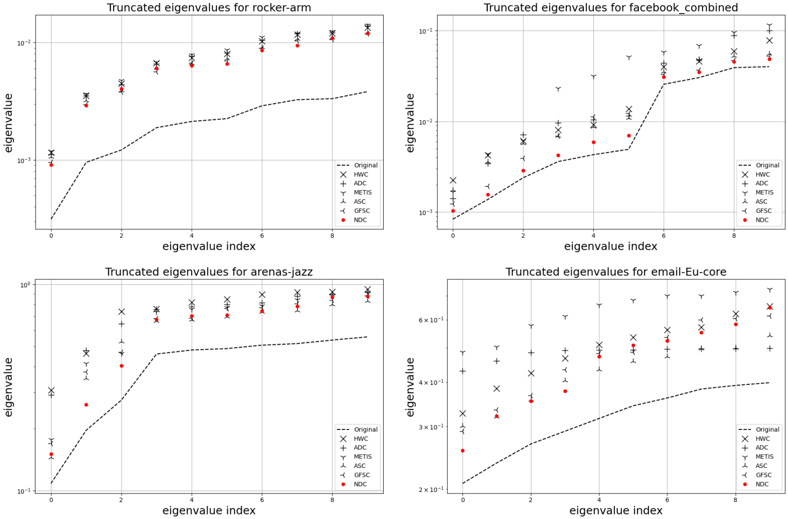

Truncated eigenvalues for rocker-arm


Truncated eigenvalues for facebook_combined

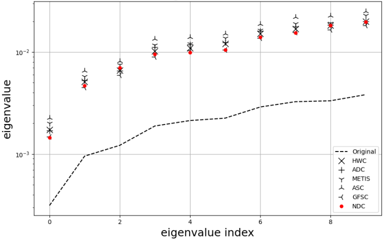

Truncated eigenvalues for arenas-jazz


Truncated eigenvalues for email-Eu-core

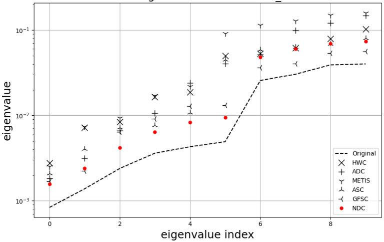
$\frac{n_c}{n} = 0.2$ coarsening eigenvalue comparisons



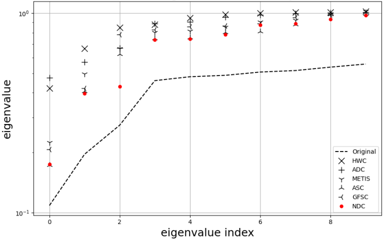
$\frac{n_c}{n} = 0.1$ coarsening eigenvalue comparisons



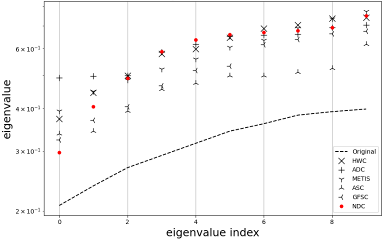
Truncated eigenvalues for rocker-arm



Truncated eigenvalues for facebook_combined

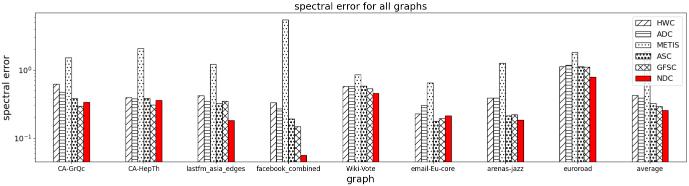


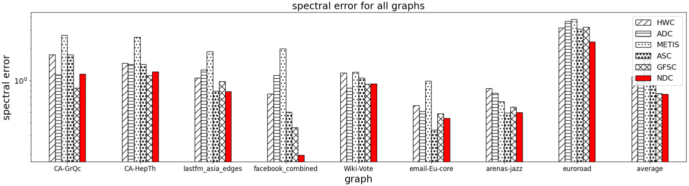
Truncated eigenvalues for arenas-jazz

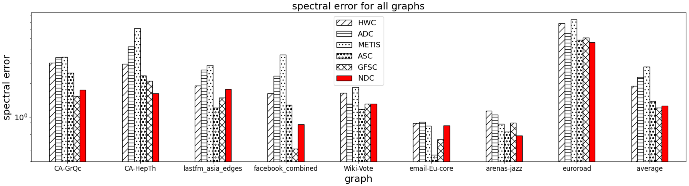


Truncated eigenvalues for email-Eu-core

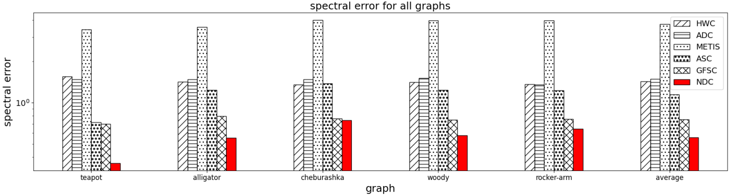
Eigenvalue comparisons for **irregular dataset**.

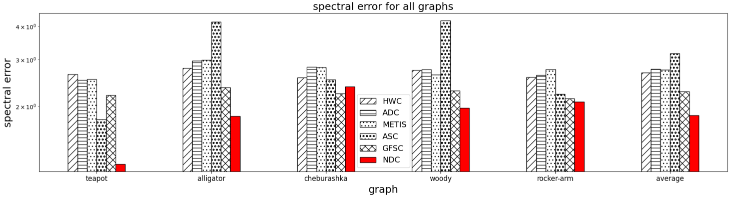


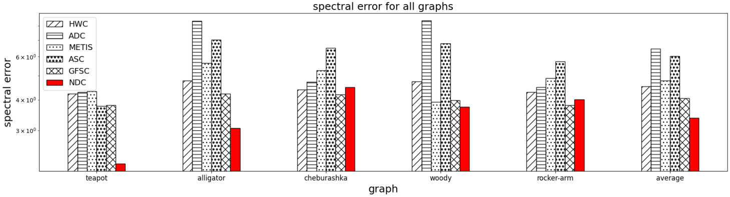




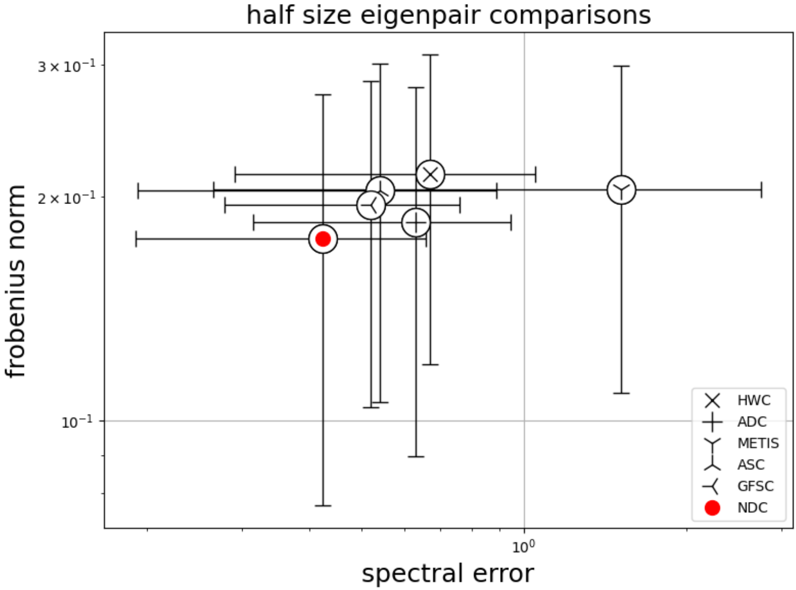
Eigenvalue comparisons for **mesh dataset**.



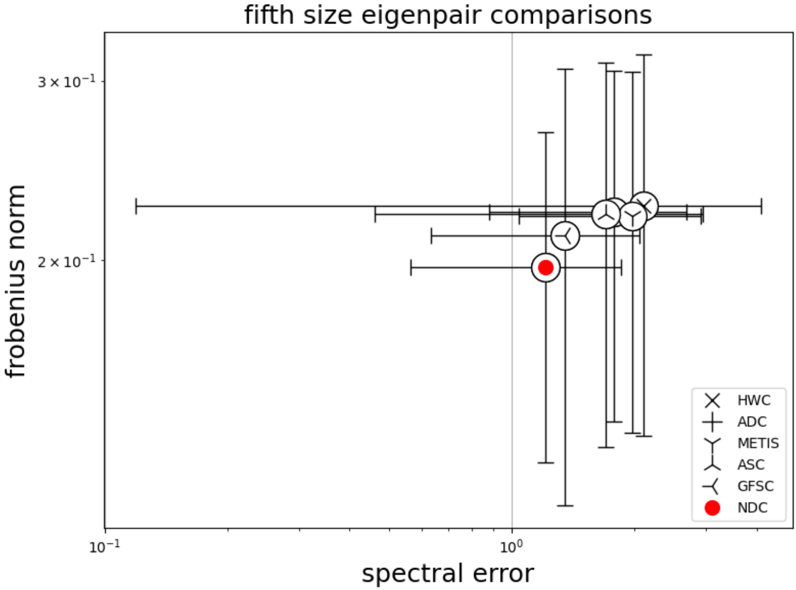




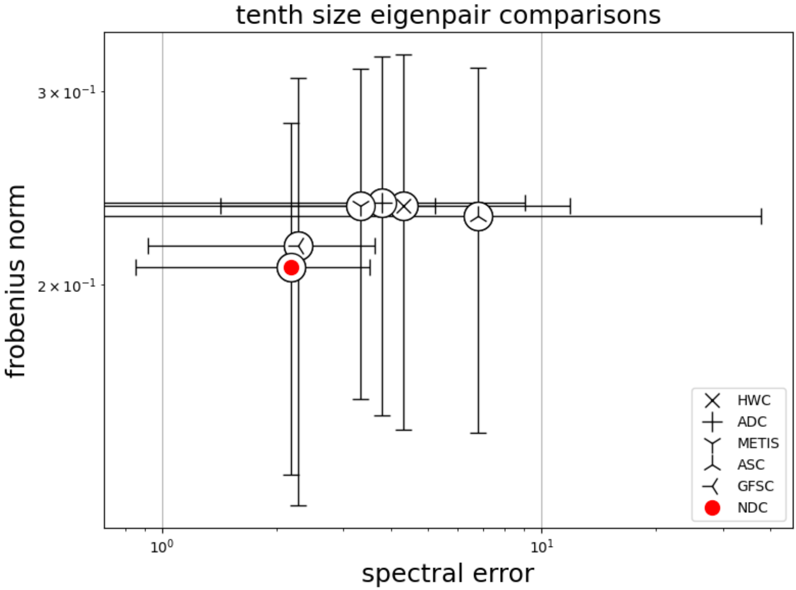
Vector versus value approximations for $\frac{n_c}{n} = 0.5$.



Vector versus value approximations for $\frac{n_c}{n} = 0.2$.



Vector versus value approximations for $\frac{n_c}{n} = 0.1$.



What about scalability

Nodal domain coarsening is an expensive operation.

- To explicitly compute the spectral pairs of a matrix is $O(n^3)$.
- Since we only need some vectors, it can be made faster with things like Krylov methods / power iteration.
- Can be made faster yet with approximation algorithms based on random walks / the heat equation / the wave equation.

For a sparse graph we can find the nodal domains in $O(mT + T^2 log(T))$ highly parallelizable work using the wave equation + a per-node Fourier transform in parallel.

What about scalability

For a sparse graph we can find the nodal domains in O($mT + T^2 log(T)$) highly parallelizable work using the wave equation + a per-node Fourier transform in parallel.

- **Pros**: It's fast. We are really only limited by how quickly we can perform SpMV.
- **Cons**: Incredibly memory intensive, and T relies on the mixing time of the graph. (Means the method lacks scalability)

Throw it all on a GPU

Timing results (seconds) for parallel agglomerative coarsening on a single nVidia A100 40GB GPU on 4092 threads.

graphs	**nodes**	**edges**	**SHWC**	**ASC**	**ADC**	**NDC**
CA-road	2.0e6	2.8e6	**0.66**	0.71	0.69	1.57
FLA-road	1.0e6	1.3e6	0.56	**0.51**	0.52	1.03
DBLP-collab	6.5e5	1.0e6	0.42	0.43	**0.39**	0.65
trimesh-mesh	1.9e6	9.4e6	**0.66**	0.77	0.75	1.61
Goog-web	8.7e5	5.1e6	**0.47**	1.6	0.51	0.92
BerkStan-comm	6.9e5	7.6e6	0.55	**37.99**	**0.54**	0.96

Next Steps

- There is much to still pursue relating to this research.
- The effect of matching versus agglommerative coarsening on spectral approximation properties.
- Explore alternative eigenvector approximation techniques for nodal domain coarsening.
- How does the quality of spectral approximation affect multigrid methods.

Thank you!

Contact: brissc@rpi.edu

Acknowledgements

- This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) Program through the FASTMath Institute under Contract No. DE-SC0021285 at the Rensselaer Polytechnic Institute, Troy NY.
- This work was also supported under Laboratory Directed Research and Development (LDRD) program, Project No. 218474, at Sandia National Laboratories. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.
- This presentation describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Summary and Thanks

- Showed that coarsening methods have drastically varying spectral approximation properties.
- Presented a coarsening method using nodal domains of eigenvectors of the graph Laplacian to restrict potential merges.
- We showed that this method does an excellent job preserving the eigenvalues of the matrix which are related to the optimal bisection.
- We further presented that this method, while running slower than competing coarsening methods, is still fast enough to find potential uses.

Thank you!

Contact: brissc@rpi.edu

[1] Jie Chen, Yousef Saad, and Zechen Zhang. “Graph coarsening: From scientific computing to machine learning”. In: *arXiv preprint arXiv:2106.11863* (2021).

[2] Jie Chen and Ilya Safro. “Algebraic distance on graphs”. In: *SIAM Journal on Scientific Computing* 33.6 (2011), pp. 3468–3490.

[3] Andy Huang et al. “Greedy Fiedler Spectral Partitioning for Data-driven Discrete Exterior Calculus.”. In: *AAAI Spring Symposium: MLPS*. 2021.

[4] Yu Jin, Andreas Loukas, and Joseph JaJa. “Graph coarsening with preserved spectral properties”. In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2020, pp. 4452–4462.

[5] George Karypis and Vipin Kumar. “METIS: A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices”. In: (1997).