# SPECTRUM CONSISTENT COARSENING APPROXIMATES EDGE WEIGHTS*

CHRISTOPHER BRISSETTE†, ANDY HUANG ‡, AND GEORGE M. SLOTA§

**Abstract.** Finding coarse representations of large graphs which preserve particular features is important in many fields of study, including clustering, numerical approximation, and the creation of reduced order models. Likewise, preserving spectral properties of the original graph during coarsening is also of particular interest for several domains of study. Our contributions are two fold. First, we generalize previous work on coarsening graphs while preserving eigenvalues of the normalized Laplacian by merging nodes with similar adjacencies, and we show that a similar analysis can be done in the case of the combinatorial Laplacian. We additionally show that when the lifted graph of a coarsening spectrally approximates the original graph, the difference between the edge weights of the graph and the edge weights of the lift depend only on the quality of spectral approximation and the strength of connectivity of the graph. It is then shown that in the case of weighted regular graphs the difference between the edge weights of the graph and the edge weights of the lift are bounded purely by the quality of spectral approximation.

**Key words.** spectral graph theory, graph coarsening, graph approximation, eigenvalues

**AMS subject classifications.** 05C22, 05C50, 68R10

## 1. Introduction.

Graph coarsening has been a long standing field of study since Gabriel Kron's 1939 work [11] creating reduced order models of electrical networks. While this original work was focused mainly on coarsening for applications to electrical networks, it has spurred a great deal of study in several different disciplines, such as machine learning and scientific computing [3]. In machine learning and scientific computing, graph coarsening is often used as a pre-processing step for clustering or partitioning. An example of this is the METIS algorithm which partitions a coarsened graph before performing a series of refinement steps [10]. Recently, some work [12, 8] has put forward methods for coarsening graphs while preserving portions of the eigenspace relating to the graph Laplacian and normalized graph Laplacian [4]. There are many benefits to this sort of analysis for preserving broader functional behavior of a network. As opposed to preserving another metric, such as approximate cut values, preserving a portion of the Laplacian eigenspace also preserves a portion of the behavior of the discrete heat and wave equations on the graph.

It is well known in the continuous case that the Laplacian operator can confer a significant amount of information about geometry [21, 9, 6]. While not all information can be retrieved [6], one may wonder what similar methods may reveal in the discrete case. Fortunately, a great deal of geometric information is conveyed through the spectrum of graphs as well [18, 19, 15, 13, 4]. One may ideally wish for there to be a way to uniquely determine a graph and its automorphisms [5] by its spectrum. Unfortunately, a complete characterization of which graphs can be determined by their spectrum is an open problem, and most classes of graphs that are known to be classified by their spectrum are either small or rather simple in their structure. However, some geometric quantities can be found. For instance, we can "hear" the volume of a graph, defined to be the sum of it's degrees, simply by adding together each eigenvalue of the graph Laplacian. Additionally, we can "hear" an approximation of the optimal conductance cut in a graph by considering the first nontrivial eigenvalue [4].

The aim of this manuscript is to formalize and prove a statement similar to the following. If the spectrum of a graph $G$ is close to that of its coarsened representation $G_c$, then the edge weights of $G$ can be closely determined from those of $G_c$. The utility of such a statement is perhaps best explored through the geometry of data. Assume there exists a set of $k$ points $\{a_i\}_{i\in[1..k]}$, $a_i \in \mathbb{R}^n$. Form a graph from these points using a

*

**Funding:** This work is funded 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 paper 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. This work was additionally 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.

†Rensselaer Polytechnic Intitute, Troy, NY (brissc@rpi.edu).

‡Sandia National Labs, Albuquerque, NM (ahuang@sandia.gov).

§Rensselaer Polytechnic Intitute, Troy, NY (slotag@rpi.edu).

concave or convex weighting function $w_{ij} = f(\|a_i - a_j\|_2)$. Given the ability to approximate these weights within some bound given a coarsening $G_c$, it follows that the pairwise distances between nodes are also approximated within some bound given the same coarsening. For such a weighting scheme, this implies that coarsenings which closely preserve eigenvalues also closely preserve nodal embeddings in $\mathbb{R}^n$ to within some perturbation of a rigid transformation. This is discussed further in the final discussion section. For now, we begin towards this goal by defining the coarsening of a graph with respect to a nodal partition.

Definition 1.1. *Consider a weighted graph $G = (V, W)$ and a partition of its nodes into $k$ disjoint sets, $P = \{V_1, \cdots, V_k\}$, $V = V_1 \cup V_2 \cup \cdots \cup V_k$. The* ***coarsened graph of*** $G$ ***with respect to*** $P$*, $G_c$, is the loopy weighted graph given by collapsing each of these partitions to a single node $\{\nu_1, \cdots, \nu_k\}$. The adjacency matrix elements are given by $W^c_{\nu_i\nu_j} = \sum_{u\in V_i}\sum_{v\in V_j} W_{uv}$. For brevity, we will often leave out the explicit partition $P$, and instead we refer to $G_c$ simply as the coarsening of $G$.*

This is the interpretation provided in Loukas [12] and Jin et al. [8] and allows for coarsening can be expressed as a product of matrices $W_c = SWS^T$ for a coarsening matrix $S$. This definition retains the sum of weighted degrees within and between partitions, as well as the total sum of weighted degrees of the graph. One should note that the coarsened graph will have fewer eigenvalues than in the original graph. Comparing the spectra becomes difficult in this instance. For this reason, Loukas considers a truncated spectrum that is cut to the dimension of the coarsened graph. We instead follow Jin et al.'s later work, where the original and coarsened graphs are compared through a structure called the lift, which extends the spectrum of the coarsened graph to the correct dimension.

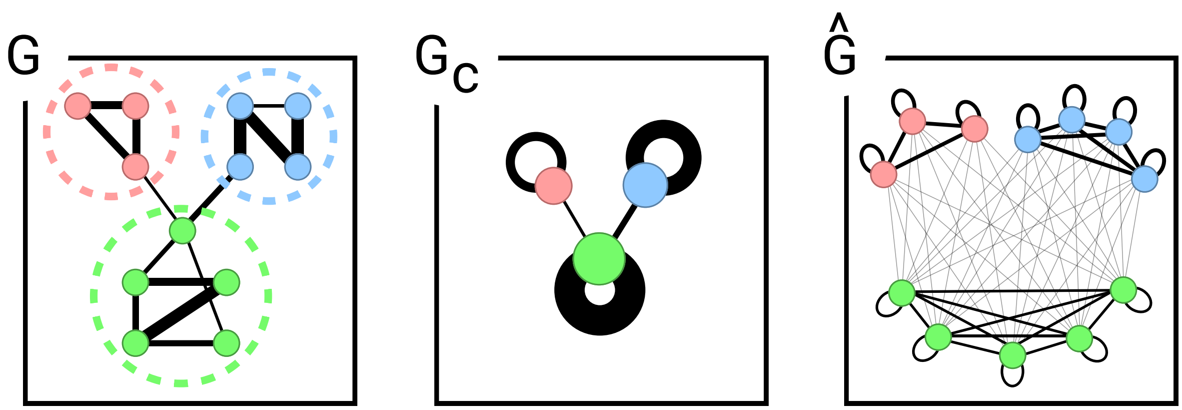


Fig. 1: **Coarsening and lifting**: A visualization of the coarsening and lifting process. In the figure, the relative thickness of an edge positively correlates with the edge weight. Note how in the shift from $G$ to $G_c$ the graph gains self-loops. Additionally, after lifting the coarsened graph $G_c$ to $\hat{G}$, the weights within and between partitions become evenly distributed.

Definition 1.2. *Consider a coarsening $G_c$ of graph $G = (V, W)$ with respect to nodal partition $P = \{V_1, \cdots, V_k\}$. We call $\hat{G} = (\hat{V}, \hat{W})$ the* ***lift of*** $G$ ***with respect to*** $P$*, where $|\hat{V}| = |V|$. The adjacency matrix elements are given by $\hat{W}_{uv} = W^c_{\nu_i\nu_j}/(|V_i||V_j|)$ where $u \in V_i$ and $v \in V_j$. For brevity, we will often assume a partition $P$ with associated coarsening $G_c$ and simply refer to $\hat{G}$ as the lift of $G$.*

As previously mentioned, the lift is useful because the sorted eigenvalues of the lift $\hat{G}$ align with those of the original graph $G$. Before continuing we define the graph Laplacian as $L = D - W$, and the normalized Laplacian as $\mathcal{L} = D^{-1/2}LD^{-1/2}$, where $D$ is the diagonal degree matrix of $G$, and $W$ is the weighted adjacency matrix. We now define notions which will be useful when comparing the structure of the original graph $G$ with that of the lift $\hat{G}$. This begins with the notion of $\sigma$-connectedness.

Definition 1.3. *Consider a graph $G = (V, W)$ and a nodal partition $P = \{V_1, \cdots, V_k\}$. The weighted adjacency of $G$ can be written as $W = W(C) + W(R)$. Here $W(C)$ is a block-diagonal matrix of $k$ disconnected weighted adjacencies corresponding to the $k$ induced subgraphs of $G$ given by the entries in $P$. The matrix*

*$W(R)$ is the adjacency of a weighted k-partite graph on the same partitions. Then for $\|W(R)\|_1 \leq \frac{\sigma}{2}$, we call the graph* ***σ-connected****.*

$$L = \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1k} \\ L_{21} & L_{22} & \cdots & L_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ L_{k1} & L_{k2} & \cdots & L_{kk} \end{bmatrix} = \begin{bmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_k \end{bmatrix} + \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1k} \\ R_{21} & R_{22} & \cdots & R_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ R_{k1} & R_{k2} & \cdots & R_{kk} \end{bmatrix}$$

We will also refer to Laplcian $C$ associated with $W(C)$ as the **core Laplacian**, and the Laplcian $R$ associated with $W(R)$ as the **ambient Laplacian**. The associated graphs for these matrices will be referred to as the **core structure** $\mathcal{C}$, and **ambient structure** $\mathcal{R}$ respectively. There are several noteworthy properties of the core and ambient structures of a graph $G$ given the nodal partition $P$. First, the adjacency matrix $W_c$ for a coarsening of $G$ with respect to partition $P$ is equal to $W(C)_c + W(R)_c$, where $W(C)_c$ and $W(R)_c$ are the coarsened adjacency matrices of the core and ambient structures with respect to the same partition. Additionally, the lift $\hat{W}$ is equal to $\hat{W}(C) + \hat{W}(R)$, where $\hat{W}(C)$ and $\hat{W}(R)$ are the lifted adjacency matrices of the core and ambient structures with respect to the partition $P$. Note that all graphs are $\sigma$-connected for some value $\sigma$. Because of this, all graphs can be broken into a core and ambient structure, where the ambient structure defines perturbations in the Laplacian of the core structure $C$. This is largely how this definition is used as the paper progresses.

Definition 1.4. *We call a weighted graph $G = (V, W)$* ***δ-complete*** *if all the weights in the graph are $W_{uv} = \frac{\delta}{N}$.*

The definition of a $\delta$-complete graph generalizes the concept of the complete graph in the unweighted case to one in the weighted case. It is worth noting that $\delta$-complete graphs have the same normalized Laplacian spectrum as complete graphs. Additionally, the nontrivial eigenvalues of the combinatorial Laplacian for $\delta$-complete graphs are all equal to $\delta$. Graphs which are $\delta$-complete form the natural building blocks of the lift of the core structure $\hat{\mathcal{C}}$. With this, all the requisite language is defined.

**1.1. Notation.** Graph $G = (V, W)$ is assumed to be weighted, and $|V| = N$, $|W| = M$. The variable $\epsilon$ will refer to a real number such that $\epsilon > 0$. Additionally we will be discussing many eigenvalues of different matrices. The ordered eigenvalues of the Laplacians of $G$ and $\hat{G}$ will be denoted as $\lambda_i$ and $\hat{\lambda}_i$ respectively. Similarly the ordered eigenvalues of the adjacency matrices $W$ and $\hat{W}$ associated with $G$ and $\hat{G}$ will be denoted $\omega_i$ and $\hat{\omega}_i$ respectively. Additionally we will concern ourselves with the normalized Laplacians $\mathcal{L}$ and $\hat{\mathcal{L}}$. The eigenvalues of these will be denoted $\eta_i$ and $\hat{\eta}_i$ respectively. The eigenvalues of the core Laplacian $C$ with $k$ connected components will be denoted $\mu_i(k)$, where $(k)$ denotes membership within the assumed partition $P = \{V_1, \cdots, V_k\}$. the associated eigenvalues of the lifted core Laplacian $\hat{C}$ will be denoted by $\hat{\mu}_i(k)$.

Weighted adjacency matrices will be denoted by $W$ and $\hat{W}$ respectively. Individual adjacencies between nodes $u, v \in V$ will be denoted by $W_{uv}$, and $\hat{W}_{uv}$ will denote the adjacency between $u, v \in \hat{V}$. Additionally $M_{i:}$ and $M_{:i}$ represent the $i^{th}$ and column respectively for an arbitrary matrix $M$. The degree of any node $u \in V$ will be denoted by $d_u$. Similarly, $\hat{d}_u$ will denote the same for $u \in \hat{V}$. These degrees show up in the diagonal degree matrices $D$ and $\hat{D}$. Additionally, when considering eigenvalues of induced subgraphs with respect to some partition $P = \{V_1, \cdots, V_k\}$, they will be expressed as $\lambda_i(j)$ where $j \in [1..k]$ denotes set membership within an element of $P$. This notation extends to degrees, as well as all other associated eigenvalues. Finally $vol(H)$, for some subgraph $H$, denotes the sum of weighted degrees within the subgraph.

**2. Spectrum Consistent Coarsening.** We first present a method for spectrum consistent coarsening of a graph $G$ with respect to the combinatorial Laplacian. This is in contrast to the work in Jin et al. [8] using normalized Laplacians; however, the proof method is incredibly similar. The idea behind this method is simple. Two nodes may be merged if their rows in the adjacency matrix are approximately linearly dependent. This linear dependence is evaluated by computing the 1-norm of the difference between rows in the adjacency matrix, and merging the rows with the smallest 1-norm difference. This is then iterated to the users desired level of coarsening. The proof follows the same format as Proposition 4.2 in Jin et al.

Theorem 2.1 (Spectrum consistent coarsening). *For a graph $G$ with all self-loops having the same weight, if it is coarsened by combining nodes $u, v \in V$, then $|\lambda_i - \hat{\lambda}_i| \leq \frac{3\epsilon}{2}$ if $\|W_{u:} - W_{v:}\|_1 \leq \epsilon$.*

*Proof.* The proof follows similarly to that in Jin et al. We consider entries of the lifted adjacency matrix.

$$\hat{W}_{ij} = \begin{cases} \frac{W_{uu}+W_{uv}+W_{vu}+W_{vv}}{4} & \text{if } i,j \in \{u,v\} \\ \frac{W_{uj}+W_{vj}}{2} & \text{if } i \in \{u,v\} \text{ and } j \notin \{u,v\} \\ \frac{W_{iu}+W_{iv}}{2} & \text{if } i \notin \{u,v\} \text{ and } j \in \{u,v\} \\ W_{ij} & \textit{else} \end{cases}$$

As noted in the original citation, this then means that the degrees of the lifted nodes will be as follows.

$$\hat{d}_i = \begin{cases} \frac{d_u+d_v}{2} & \text{if } i \in \{u,v\} \\ d_i & \textit{else} \end{cases}$$

There are now a Laplacian $L = D - W$ and a lifted Laplacian $\hat{L} = \hat{D} - \hat{W}$, and we wish to know the difference between these $E = L - \hat{L} = D - \hat{D} + \hat{W} - W$ as to apply Weyl's inequality.

$$\hat{W}_{ij} - W_{ij} = \begin{cases} \frac{W_{uu}+W_{uv}+W_{vu}+W_{vv}}{4} - W_{ij} & \text{if } i,j \in \{u,v\} \\ \frac{W_{uj}+W_{vj}}{2} - W_{ij} & \text{if } i \in \{u,v\} \text{ and } j \notin \{u,v\} \\ \frac{W_{iu}+W_{iv}}{2} - W_{ij} & \text{if } i \notin \{u,v\} \text{ and } j \in \{u,v\} \\ 0 & \textit{else} \end{cases}$$

$$D_{ii} - \hat{D}_{ii} = \begin{cases} d_i - \frac{d_u+d_v}{2} & \text{if } i \in \{u,v\} \\ 0 & \textit{else} \end{cases}$$

Because $\|W_{u:} - W_{v:}\|_1 \leq \epsilon$, it is also true that $|d_u - d_v| \leq \epsilon$ by the triangle inequality. Therefore, without loss of generality, $\frac{d_u+d_v}{2} \leq d_u + \frac{\epsilon}{2}$ meaning $|d_u - \frac{d_u+d_v}{2}| \leq \frac{\epsilon}{2}$. Therefore, to prove the lemma, only considerations for the difference in the adjacency matrices remain. For this, two cases need to be analyzed, including the case where $i \in \{u,v\}$ and the case where $i \notin \{u,v\}$. Without loss of generality, take $i = u$, then in the former case above the following is true.

$$\begin{aligned} \|\hat{W}_{i:} - W_{i:}\|_1 &= \sum_{j \in V} |\hat{W}_{uj} - W_{uj}| \\ &= \left|\frac{W_{uu}+W_{uv}+W_{vu}+W_{vv}-4W_{uu}}{4}\right| + \left|\frac{W_{uu}+W_{uv}+W_{vu}+W_{vv}-4W_{uv}}{4}\right| \\ &+ \sum_{j \notin \{u,v\}} \left|\frac{W_{uj}+W_{vj}-2W_{uj}}{2}\right| \\ &= \frac{1}{4}|W_{uv}+W_{vu}+W_{vv}-3W_{uu}| + \frac{1}{4}|W_{uu}+W_{vv}-2W_{uv}| + \frac{1}{2}\sum_{j \notin \{u,v\}} |W_{vj} - W_{uj}| \\ &\leq \frac{3}{4}|W_{uu}-W_{uv}| + \frac{3}{4}|W_{vv}-W_{uv}| + \frac{1}{4}|W_{uu}-W_{vv}| + \frac{1}{2}\sum_{j \notin \{u,v\}} |W_{vj} - W_{uj}| \\ &\leq |W_{uu}-W_{uv}| + |W_{vv}-W_{uv}| + \frac{1}{2}\sum_{j \notin \{u,v\}} |W_{vj} - W_{uj}| \\ &\leq \|W_{u:} - W_{v:}\|_1 \leq \epsilon \end{aligned}$$

Now we prove a similar result when $i \notin \{u,v\}$.

$$\begin{aligned} \|\hat{W}_{i:} - W_{i:}\|_1 &= \sum_{j \in V} |\hat{W}_{ij} - W_{ij}| \\ &= \frac{1}{2}|W_{iu} - W_{iv}| + \frac{1}{2}|W_{iv} - W_{iu}| \\ &= |W_{ui} - W_{vi}| \\ &\leq \|W_{u:} - W_{v:}\|_1 \leq \epsilon \end{aligned}$$

Then $\|E\|_1 = \|\hat{L} - L\|_1 = \|\hat{D} - D + W - \hat{W}\|_1 \leq \frac{\epsilon}{2} + \epsilon = \frac{3\epsilon}{2}$. Our lemma then follows immediately from Weyl's inequality. ☐

**2.1. Discussion.** We note that this bound can be repeated in succession $m$ times, and if each successive coarsening has an L1 difference less than or equal to $\epsilon$, then the spectral gap we obtain at the end between our graphs is less than or equal to $\frac{3m\epsilon}{2}$. We show an example of this coarsening performed on a $\sigma$-connected graph according to the criteria of theorem 2.1 in figure 2.

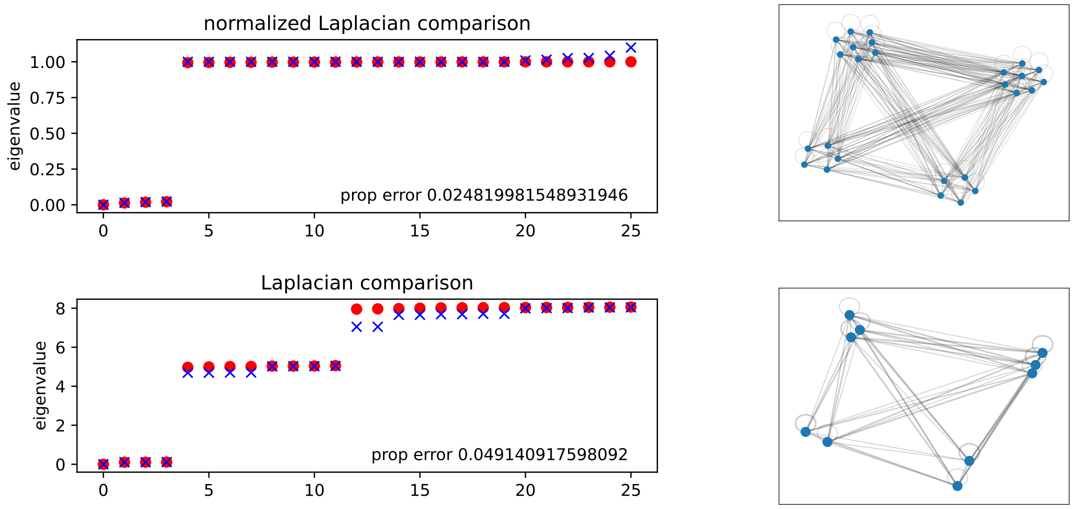


Fig. 2: **A coarsening example:** An example of $\sigma$-connected graph is pictured in the top right of the figure and a greedily coarsened representation is shown beneath it. This graph was coarsened according to the criteria in Theorem 2.1, using $\epsilon \leq 0.1$ as a maximum threshold. The red-dots denote eigenvalues of the original graph, and the blue crosses denote the eigenvalues of the lift after coarsening. Both the spectrum of the normalized Laplacian and the spectrum of the combinatorial Laplacian are shown along with the proportional L1 error for each of them as $\|\Lambda - \hat{\Lambda}\|_1 / \|\Lambda\|_1$, where $\Lambda$ and $\hat{\Lambda}$ are vectors containing the sorted eigenvalues of $G$ and $\hat{G}$ respectively.

When coarsening with respect to the normalized Laplacian, the eigenvalues of the coarsened graph are all eigenvalues of the lifted graph. This makes comparisons between the lift and the original graph meaningful, because they bound the spectral behavior of the coarse graph. This is not true in the case of the combinatorial Laplacian. However, coarsening with respect to the combinatorial Laplacian achieves multiple objectives.

THEOREM 2.2. *Given a graph $G = (V, W)$, if two nodes $u, v \in V$ are such that $\|W_{u:} - W_{v:}\|_1 \leq \epsilon$, then $\|W_{u:}/d_u - W_{v:}/d_v\|_1 \leq 2\epsilon/max\{d_u, d_v\}$.*

*Proof.* First note that $\|W_{u:} - W_{v:}\|_1 \leq \epsilon$ implies $|d_u - d_v| \leq \epsilon$. Then without loss of generality the following is true.

$$\begin{aligned} \left\| \frac{W_{u:}}{d_u} - \frac{W_{v:}}{d_v} \right\|_1 &= \left\| \frac{d_v W_{u:} - d_u W_{v:}}{d_u d_v} \right\|_1 \\ &\leq \left\| \frac{W_{u:} - W_{v:}}{d_v} \right\|_1 + \frac{\epsilon}{d_v} \leq \frac{2\epsilon}{d_v} \end{aligned}$$

Because $d_v$ was chosen without loss of generality, $\|W_{u:}/d_u - W_{v:}/d_v\|_1 \leq 2\epsilon/max\{d_u, d_v\}$ is true, proving the theorem. ☐

As shown in Theorem 2.2, the bound $\epsilon$ given in combinatorial Laplacian coarsening enforces a related bound with regards to normalized Laplacian coarsening. That is, combinatorial Laplacian coarsening coarsens with respect to both objectives at once. This is not guaranteed with normalized Laplacian coarsening. Furthermore, as is discussed throughout this manuscript, relations between nodes in the fine graph are preserved when coasening with respect to the combinatorial Laplacian. This has potential applications to data mining and clustering tasks.

We now present a comparison between the normalized objective and algebraic distance. AMG methods are extensions of classical multigrid methods [16] in scientific computing, which rely on successive coarsenings of a linear operator to efficiently solve a linear system. The algebraic distances used in relaxation based AMG methods rely on sampling a set of test vectors [14] $\{x^i\}_{i\in[1..k]}$ and computing $\chi^i = L_{rw}x^i$, where $L_{rw}$ is the random walk normalized Laplacian [4]. A distance metric between nodes is computed using the output of these vectors. One such metric is the following.

$$\alpha_{uv} = \max_{i\in[1..k]} |\chi^i_u - \chi^i_v| \tag{2.1}$$

Intuitively, this quantifies the linear dependence between rows of the random walk Laplacian $L_{rw}$. In Jin et al. the linear dependence between rows of the random walk Laplacian is also considered. In the former case, linear dependence is probed by test vectors, whereas it is quantified by the 1-norm difference in the latter case. Using the bound in Jin et al. one can also bound the distance metric $\alpha_{uv}$. Assuming $\|x^i\|_2 = 1$ for all $i \in [1..k]$, then the following is true.

$$\begin{aligned} &\left\| \frac{W_{u:}}{d_u} - \frac{W_{v:}}{d_v} \right\|_1 \leq \epsilon \\ \Rightarrow &\left| \left( \frac{W_{u:}}{d_u} - \frac{W_{v:}}{d_v} \right) x^i \right| \leq \epsilon \|x^i\|_2 \leq \epsilon \\ \Rightarrow &\ \alpha_{uv} \leq \epsilon \end{aligned}$$

One may construct a similar algebraic distance using the combinatorial Laplacian. In this case, the same argument can be made using the bound given in Theorem 2.1. A very similar bounding argument may be made for another algebraic distance metric used in AMG methods.

$$\beta_{uv} = \sum_{i\in[1..k]} \left(\chi^k_u - \chi^k_v\right)^2$$

If the bound from Jin et al. is assumed, one knows $\beta_{uv} \leq k\epsilon^2$ due to a similar argument as before. An identical argument can be made for an algebraic distance using the combinatorial Laplacian when the bound given in Theorem 2.1 is used. In this sense, the bound from Theorem 2.1 can be interpreted as a stronger criteria than those of relaxation based algebraic multigrid. To achieve the same spectral bound using an algebraic metric such as $\alpha_{uv}$ or $\beta_{uv}$, one requires $k = N$ linearly independent test vectors at minimum. This implies the work complexity of relaxation based AMG needs to be $O(MN)$ to achieve a similar spectral bound to Theorem 2.1. Alternatively, computing a sparse norm between adjacency vectors for every edge in the graph requires only $O(N\langle d^2 \rangle)$ where $\langle d^2 \rangle$ is the second moment of the graph's degree distribution [2]. While it may be possible to define a spectral bound given algebraic distances between nodes, to the authors' knowledge no such bound exists [3]. Figure 3 compares coarsening heuristics on three different graphs from the Koblenz Network Collection[1]. In Figure 3 one can see that, despite not having the same guarantees as the criteria in Theorem 2.1, coarsening with respect to algebraic distance does perform well for spectral approximation. Both methods perform better than heavy weight matching, which is a popular coarsening method attempting to maximize $\frac{w_ij}{d_i d_j}$ for each merge [3].

**3. Edge Weight Approximation for General Graphs.** We now proceed with the proof of the main result of the manuscript. The details behind the main result of this paper are that every graph is $\sigma$-connected

[1] http://konect.cc/networks/

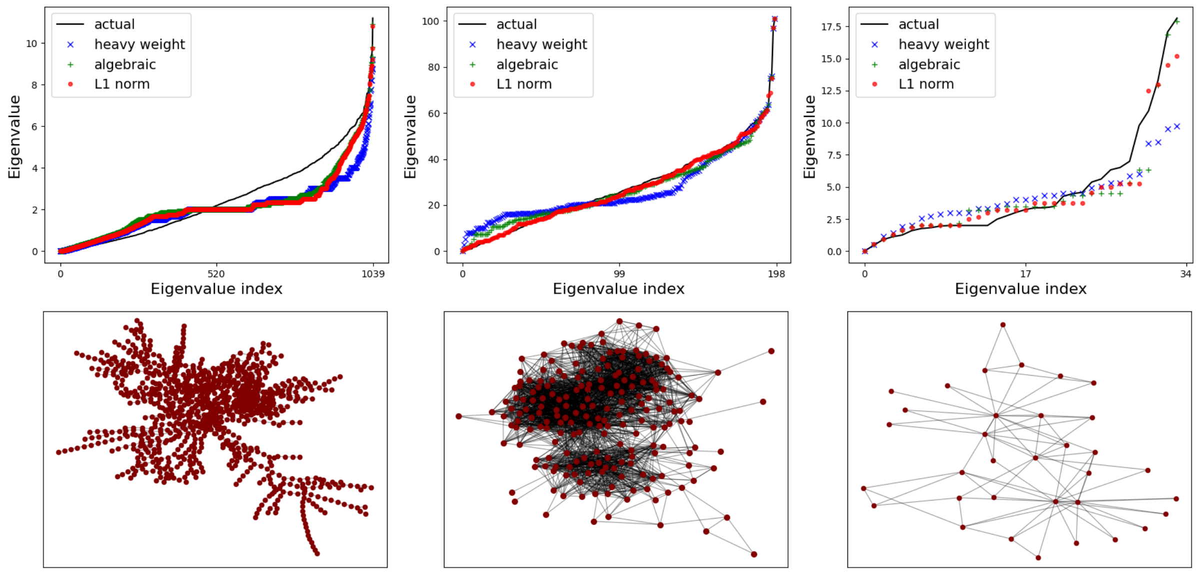


Fig. 3: **Coarsening method comparison:** Three graphs were coarsened using three different heuristics to half-size and then the spectra of their lifts were compared with the original. From left to right, the graphs are Euroroad, Jazz Musicians, and the Zachary Karate Club. All of these were collected from the Koblenz Network Collection. The bottom row of the figure displays the original graphs. For heavy weight matching the edge $(u, v)$ corresponding to the highest value of $\frac{W_{uv}}{d_u d_v}$ was contracted at each step. For algebraic distance, the edge corresponding to the minimum of Equation 2.1 was contracted. Finally for the L1 method, the edge minimizing the criteria in Theorem 2.1 was contracted. It should additionally be noted that 20 test vectors were used for computing algebraic distances.

for some value $\sigma$ and the Laplacian can be expressed as the sum of two independent graph Laplacians. These Laplacians are defined by the core Laplacian $C$ and the ambient Laplacian $R$. Comparisons can then be made between the core structures of the original graph and the lift, viewing $R$ and $\hat{R}$ as perturbation matrices. In this way the degrees of individual nodes may be bounded according to the difference in spectra, and the parameter $\sigma$. This allows for the application of discrepancy bounds [4] to various subgraphs and, most importantly, pairs of nodes. These bounds are then used to bound weight differences accordingly. A visualization of the dependencies between results can be seen in Figure 4.

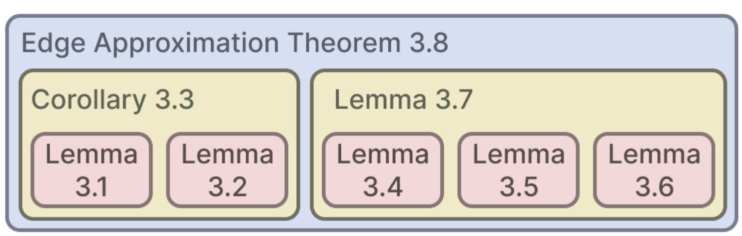


Fig. 4: **Result dependencies:** The dependencies between results in this section are shown. The boxes are numbered with their respective results, and any boxes nested within them represent sub-results which are used to prove the larger corollary, lemma, or theorem.

Lemma 3.1. *Given a $\sigma$-connected graph $G = (V, W)$ and its approximated lifted graph $\hat{G} = (\hat{V}, \hat{W})$, say $|\lambda_i - \hat{\lambda}_i| \le \epsilon$ for all $i \in [1..N]$. Then $|\mu_i - \hat{\lambda}_i| \le \epsilon + \sigma$ for all $i \in [1..N]$.*

*Proof.* Since $G$ is $\sigma$-connected, its Laplacian structure can be written as the sum of two separate Laplacians, $L = C + R$. Note $\|R\|_2 \leq \|R\|_1 \leq \sigma$. Using this, we directly apply Weyl's inequality [1] to get $|\lambda_i - \mu_i| \leq \sigma$. Then, because $|\lambda_i - \hat{\lambda}_i| \leq \epsilon$, $|\mu_i - \hat{\lambda}_i| \leq \epsilon + \sigma$. ☐

In simplifying language, Lemma 3.1 states that the eigenvalues of the core Laplacian $C$ closely approximate eigenvalues of the lift $\hat{G}$ when the eigenvalues of $\hat{G}$ closely approximate the eigenvalues of $G$.

LEMMA 3.2. *If $G = (V, W)$ is a $\sigma$-connected graph, then its lift $\hat{G} = (\hat{V}, \hat{W})$ is also $\sigma$-connected.*

*Proof.* Note that lifting preserves the sum of edge weights within partitions and the sum of edge weights between partitions. Therefore, if $L$ is expressed in terms of its core and ambient structures, the lifts of both of these structures may be independently considered. For any node $u \in V_i$ in $\hat{\mathcal{R}}$, the degree $\hat{d_u} = \frac{1}{|V_i|}\sum_{v \in V_i} d_v$. The following relationship then holds true.

$$\max_{v \in V} |\hat{d_v}| \leq \max_{u \in V} |d_u|$$

This implies $\|\hat{R}\|_1 \leq \|R\|_1 \leq \sigma$ proving our lemma. ☐

Lemma 3.2 in conjunction with Lemma 3.1 allows for direct comparison between the spectra of the core Laplacians $C$ and $\hat{C}$.

COROLLARY 3.3. *Given a $\sigma$-connected graph $G = (V, W)$ and its approximated lifted graph $\hat{G} = (\hat{V}, \hat{W})$, respectively, say $|\lambda_i - \hat{\lambda}_i| \leq \epsilon$ for all $i \in [1..N]$. Then $|\mu_i - \hat{\mu}_i| \leq \epsilon + 2\sigma$.*

*Proof.* This follows immediately from the fact that $\hat{G}$ is $\sigma$-connected from Lemma 3.2, and then applying Lemma 3.1. ☐

Corollary 3.3 states that the core-structure of a graph and its lift have similar eigenvalues when $G$ and $\hat{G}$ have similar eigenvalues. Each independent core sub-Laplacian $\hat{C}(i)$ of our lift $\hat{G}$ is $\delta_i$-complete, where $\delta_i$ is the average degree within $\mathcal{C}(i)$. This implies $\hat{\mu}_j(i) = \delta_i$ for all nontrivial eigenvalues. Therefore, full information of the degrees and spectra of every $\hat{\mathcal{C}}(i)$ are known. In conjunction with Corollary 3.3, this will allow for comparison between the degrees of partitions of the core structures $\mathcal{C}(i)$ and $\hat{\mathcal{C}}(i)$.

LEMMA 3.4. *If all the nontrivial eigenvalues of the Laplacian $L$ of a connected graph $G = (V, W)$ lie within the bounds $\delta - \epsilon \leq \lambda_i \leq \delta + \epsilon$, with $\delta = \frac{Vol(G)}{N}$, then $|d_i - \delta| \leq 4\epsilon$ for all $i \in [1..N]$. Here $d_i = \sum_{j \in [1..N]} W_{ij}$ is the degree of the node $i \in V$.*

*Proof.* Consider the vector $e_{ij} = \frac{1}{\sqrt{2}}(e_i - e_j)$, where $e_i, e_j$ are the unit vectors with value zero everywhere except for a one in the $i^{th}$ and $j^{th}$ element, respectively. Note that $e_{ij}^T \perp \mathbf{1}$, meaning $\|Le_{ij}\|_2$ cannot be arbitrarily small. Instead, it is bounded below and above by $\lambda_2 \leq \|Le_{ij}\|_2 \leq \lambda_N$. By assumption, $\delta - \epsilon \leq \lambda_i \leq \delta + \epsilon$ for all nontrivial eigenvalues. This means $\delta - \epsilon \leq e_{ij}^T L e_{ij} \leq \delta + \epsilon$ must be true due to $e_{ij}$ having unit length $\|e_{ij}\|_2 = 1$. By writing out $e_{ij}^T L e_{ij}$ explicitly, one gets that $2(\delta - \epsilon) \leq (d_i + d_j + 2W_{ij}) \leq 2(\delta + \epsilon)$. From this, the following must be true.

$$2N(N-1)(\delta - \epsilon) \leq \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} d_i + d_j + 2W_{ij} \leq 2N(N-1)(\delta + \epsilon)$$

$$2N(N-1)(\delta - \epsilon) \leq \sum_{i=1}^{N} (N-1)d_i + (Vol(G) - d_i) + 2d_i \leq 2N(N-1)(\delta + \epsilon)$$

$$2N(N-1)(\delta - \epsilon) \leq 2NVol(G) \leq 2N(N-1)(\delta + \epsilon)$$

$$\Rightarrow \frac{1-N}{N}\epsilon \leq \delta + \frac{1-N}{N}\delta \leq \frac{N-1}{N}\epsilon$$

$$\Rightarrow (1-N)\epsilon \leq \delta \leq (N-1)\epsilon$$

The inequality $\delta \leq (N-1)\epsilon$ can now be used to prove our lemma. The proof follows similarly to the previous inequalities, however now the outer sum is removed.

$$
\begin{aligned}
2(N-1)(\delta-\epsilon) &\leq \sum_{j=1, j\neq i}^{N} d_i + d_j + 2W_{ij} \leq 2(N-1)(\delta+\epsilon) \\
2(N-1)(\delta-\epsilon) &\leq (N-1)d_i + (Vol(G) - d_i) + 2d_i \leq 2(N-1)(\delta+\epsilon) \\
2(N-1)(\delta-\epsilon) &\leq Nd_i + Vol(G) \leq 2(N-1)(\delta+\epsilon) \\
\Rightarrow 2\frac{(N-1)}{N}(\delta-\epsilon) &\leq d_i + \delta \leq 2\frac{(N-1)}{N}(\delta+\epsilon) \\
\Rightarrow \frac{1-N}{N}\epsilon &\leq d_i - \delta + \frac{2\delta}{N} \leq 2\frac{N-1}{N}\epsilon \\
\Rightarrow -2\epsilon - \frac{2(\delta-\epsilon)}{N} &\leq d_i - \delta \leq 2\epsilon - \frac{2(\epsilon+\delta)}{N} \\
\Rightarrow -2\epsilon - 2\epsilon &\leq d_i - \delta \leq 2\epsilon \\
\Rightarrow |d_i - \delta| &\leq 4\epsilon
\end{aligned}
$$

The second to last line comes as an immediate consequence of the previous inequality $\delta \leq (N-1)\epsilon$, and proves our lemma. □

Lemma 3.4 allows for statements to be made about the degrees of nodes in $G$ based on the average degrees of partitions. This completes one of two major building blocks for the final edge approximation theorem. Before proving the next lemma we state a weighted version of Theorem 5.1 in Chung and Graham [4], noting that the original proof provided does not change in the case of weights.

Lemma 3.5 (Chung. 5.1). *Suppose $X, Y$ are two subsets of the vertex set $V$ of a graph $G$. Then,*

$$
\left| \sum_{x\in X, y\in Y} W_{xy} - \frac{vol(X)vol(Y)}{vol(G)} \right| \leq \bar{\lambda}\sqrt{vol(X)vol(Y)}
$$

*where $\bar{\lambda} = \max_{i\neq 1} |1-\eta_i|$. Here $\{\eta_i\}_{i\in[2..N]}$ are eigenvalues of the normalized Laplacian $\mathcal{L}(G) = D^{-\frac{1}{2}}L(G)D^{-\frac{1}{2}}$.*

Ideally, this theorem could be applied directly to a $\sigma$-connected graph $G$ to bound the difference between the edge weights of $G$ and $\hat{G}$. In order to apply lemma 3.5, an approximation of $\bar{\lambda}$ is required.

Lemma 3.6. *Assume graph $G$ is coarsened to a single node and then lifted to $\hat{G}$. If $|\lambda_i - \hat{\lambda}_i| \leq \epsilon$ then $|\eta_i - \hat{\eta}_i| \leq min\{h_p, 1\}$ where $h_p = \frac{5p}{1-2p}$ and $p = \frac{\epsilon}{\delta} < \frac{1}{4}$.*

*Proof.* Begin by noting that the random-walk normalized Laplacian $\mathcal{L}_{rw} = D^{-\frac{1}{2}}\mathcal{L}D^{\frac{1}{2}} = D^{-1}L$ has the same eigenvalues as the normalized Laplacian. It is true from Lemma 3.4 that $|d_v - \delta| \leq 4\epsilon$. This implies the following bounds on the eigenvalues of $D^{-1}$.

$$
\frac{1}{\delta + 4\epsilon} \leq \lambda_i(D^{-1}) \leq \frac{1}{\delta - 4\epsilon}
$$

This implies that the nontrivial eigenvalues $\{\eta_i\}_{i\in[2..N]}$ of $D^{-1}L$ lie in the following bounds.

$$
\begin{aligned}
\frac{\delta-\epsilon}{\delta+4\epsilon} &\leq \eta_i \leq \frac{\delta+\epsilon}{\delta-4\epsilon} \\
\Rightarrow \frac{1-p}{1+4p} &\leq \eta_i \leq \frac{1+p}{1-4p} \\
\Rightarrow \frac{(1-p)-(1+4p)}{(1+4p)} &\leq \eta_i - 1 \leq \frac{(1+p)-(1-4p)}{(1-4p)} \\
\Rightarrow \frac{-5p}{(1+4p)} &\leq \eta_i - 1 \leq \frac{5p}{(1-4p)}
\end{aligned}
$$

For $0 \leq p < \frac{1}{4}$, $\frac{5p}{(1-4p)} > \frac{5p}{(1+4p)}$. Additionally, $\hat{\eta}_i = 1$ implying the following and completing the proof.

$$|\eta_i - \hat{\eta}_i| \leq min\left\{\frac{5p}{1-4p}, 1\right\} \qquad \square$$

Lemma 3.6 provides a bound on $\bar{\lambda}$ which, in conjunction with Lemma 3.5, may be used to prove that the differences between edge weights in $G$ and $\hat{G}$ remain bounded within partitions. This bound does require $p = \frac{\epsilon}{\delta}$ to be rather small, however, since $\frac{5p}{1-4p} \to \infty$ as $p \to \frac{1}{4}$. In fact, $p = \frac{1}{9}$ is where this bound becomes devoid of useful information, since $|\eta_i - \hat{\eta}_i| \leq 1$ by virtue of this being a difference of normalized Laplacian eigenvalues. We can now prove a useful discrepancy bound for the case $p < \frac{1}{9}$.

Lemma 3.7. *Let there be a weighted graph $G = (V, W)$. Additionally, consider the lift of the graph $\hat{G} = (\hat{V}, \hat{W})$, which comes from first coarsening $G$ to a single node. This implies $\hat{G}$ is $\delta$-complete with $\delta = \frac{Vol(G)}{N}$. If $|\lambda_j - \hat{\lambda}_j| \leq \epsilon$ for all $j \in [1..N]$, then $|W_{uv} - \hat{W}_{uv}| \leq min\{h_p, 1\}(\delta + 4\epsilon) + \frac{4\epsilon}{N}(p+1)$ where $h_p = \frac{5p}{1-4p}$ and $p = \frac{\epsilon}{\delta} < \frac{1}{9}$.*

*Proof.* We directly apply Lemma 3.5 by considering $X = u$ to be a single node and $Y = v$ to be a single node.

$$\begin{aligned}\left|W_{uv} - \frac{d_u d_v}{vol(G)}\right| &\leq \bar{\lambda}\sqrt{d_u d_v} \\ &\leq \frac{5p}{1-4p}(\delta + 2\epsilon)\end{aligned}$$

The right hand side follows from Lemma 3.4 and Lemma 3.6. Going forward, $\frac{3p}{1-2p}$ will be denoted by $h_p$. Note that $vol(G) = vol(\hat{G}) = N\delta$.

$$\begin{aligned}\frac{\delta^2 - 4\delta\epsilon + 4\epsilon^2}{N\delta} &\leq \frac{d_u d_v}{N\delta} \leq \frac{\delta^2 + 4\delta\epsilon + 8\epsilon^2}{N\delta} \\ \Rightarrow \frac{4p\epsilon - 4\epsilon}{N} &\leq \frac{d_u d_v}{N\delta} - \frac{\delta}{N} \leq \frac{4p\epsilon + 4\epsilon}{N} \\ \Rightarrow -\frac{4p\epsilon + 4\epsilon}{N} &\leq \frac{d_u d_v}{N\delta} - \frac{\delta}{N} \leq \frac{4p\epsilon + 4\epsilon}{N} \\ \Rightarrow \left|\frac{d_u d_v}{N\delta} - \frac{\delta}{N}\right| &\leq \frac{4\epsilon}{N}(p+1)\end{aligned}$$

Using this, we can refine our statement further, thus proving our lemma.

$$\left|W_{uv} - \frac{\delta}{N}\right| = \left|W_{uv} - \hat{W}_{uv}\right| \leq min\{h_p, 1\}(\delta + 2\epsilon) + \frac{4\epsilon}{N}(p+1) \qquad \square$$

Using Lemma 3.7, the main theorem is ready to be proven.

Theorem 3.8 (Edge Approximation). *Let $G$ be a $\sigma$-connected graph with respect to the partition $P = \{V_1, \cdots, V_k\}$, and let $\hat{G}$ be the lift of $G$ with respect to that partition. Additionally, assume $|V_i|$ is large for each $i \in [1..k]$. If $|\lambda_i - \hat{\lambda}_i| \leq \epsilon$, the difference between in-partition weights is bounded by $|W_{uv} - \hat{W}_{uv}| \leq min\{h_q(j), 1\}(\delta(j) + 2(\epsilon + 2\sigma)) + \frac{4(\epsilon+2\sigma)}{N}(q(j)+1)$ for $u, v \in V_j$. Additionally, between-cluster weights are bounded by $|W_{uv} - \hat{W}_{uv}| \leq \sigma$ where $u \in V_i$ and $v \in V_j$ where $i \neq j$. Here, $h_q(j) = \frac{5q(j)}{1-4q(j)}$ and $q(j) = \frac{\epsilon+2\sigma}{\delta_j}$.*

*Proof.* From the definition of $\sigma$-connected, the Laplacian $R$ is such that $\|R\|_1 \leq \sigma$. This bounds the maximum value of the matrix, implying that $|W_{uv} - \hat{W}_{uv}| \leq \sigma$ for $u \in V_i$ and $v \in V_j$ where $i \neq j$. For in-partition weights, first note that from Corollary 3.3, the eigenvalues of $C(j)$ and the eigenvalues of $\hat{C}(j)$ are bounded such that $|\mu_i(j) - \hat{\mu}_i(j)| \leq \epsilon + 2\sigma$. By using this as the error term in Lemma 3.7, one obtains the following bound: $|W_{uv} - \hat{W}_{uv}| \leq min\{h_q(j), 1\}(\delta(j) + 2(\epsilon + 2\sigma)) + \frac{4(\epsilon+2\sigma)}{N}(q(j)+1)$, which proves the theorem. $\square$

**3.1. Discussion.** Theorem 3.8 states that as the difference in spectrum $|\lambda_i - \hat{\lambda}_i|$ approaches zero for all $i \in [1..N]$, the difference in the weights of all edges depends only on the connectivity between subgraphs in the partition $P$. As a consequence of this, one can in a sense "hear" the shape of the original graph, given a coarsened graph $G_c$ whose lift $\hat{G}$ spectrally approximates it. In practice, this bound is only practical for graphs which do not occur in general applications. To observe why, assume that a simple graph $G$ is coarsened to $G_c$ with respect to some partitioning $P = \{V_1, \cdots, V_k\}$. Further assume that within any partition $V_i$ there are two nodes which are not adjacent. Because all nodes within the same partition are adjacent in the lift, the maximum edge-weight difference is bounded below by $\frac{\delta_i}{N}$. This minimum upper bound exists regardless of the spectral properties of $G$ and $\hat{G}$. Furthermore, in most real world graphs, $\sigma$ is relatively large and the resulting bound in Theorem 3.8 is dominated by the $\sigma$ term in the expression, often leading to bounds larger than the largest degree in the graph. This implies that meaningful uses of the upper bound in Theorem 3.8 may generally be restricted to weighted graphs where every node is adjacent to every other, and there are small weights between partitions. While not generally found in social-science, or scientific computing applications, such graphs are used in practice for image segmentation [17, 22, 7], and data mining tasks [20]. These methods use weighting schemes based on the distance between nodes to define similarities in arbitrary data. One common weight function is $w_{uv} = exp\{\|r_u - r_v\|_2^2/\Theta\}$ where $r_u, r_v$ are the embeddings of $u, v$ in $\mathbb{R}^N$, and $\Theta$ is a positive constant. Given this, or other similar weighting schemes, one can bound the distance between $r_u$ and $r_v$. This implies that by bounding edge weights between the graph $G$ and it's lift $\hat{G}$, one is simultaneously preserving the distances between these nodal embeddings. However, the bound in theorem 3.8 is only usable in the most well clustered of test cases, and requires further refinement before being usable in application.

**4. Edge Weight Approximation for Weighted Regular Graphs.** We briefly turn our attention to a special case where the spectrum fully determines the properties of graph connectivity. This is in the case of weighted regular graphs where $d_i = d$ for all nodes $i \in V$ and some positive real number $d$. For this purpose we will instead examine the adjacency matrices $W$ and $\hat{W}$. Coarsening as defined in definition 1.1 may be expressed as a matrix product $W = SWS^T$ for a coarsening matrix $S$ discussed in further detail in Loukas [12]. Additionally the lifting operation can be expressed as the pseudo-inverse of this operation, given by $\hat{W} = P^\dagger PW(P^\dagger P)^T$. The matrix $PP^\dagger = \Pi$ has a simple form given in both Loukas [12] and Jin et al. [8]. Given a partition $P = \{V_1, \cdots, V_k\}$ each element $\Pi_{ij} = \frac{1}{|V_r|}$ for $i, j \in V_r$, otherwise $\Pi_{ij} = 0$. One can easily check that this coincides with our definition of coarsening.

This matrix relation between the original and lifted adjacencies allows for a powerful theorem to be proven.

Theorem 4.1. *For a weighted adjacency matrix $W$ and lifted adjacency matrix $\hat{W} = \Pi W \Pi$, if $|\omega_i - \hat{\omega}_i| \leq \gamma$ for all $i \in [1..N]$, then $\|W - \hat{W}\|_F^2 \leq N\gamma\,(2\|W\|_2 + \gamma)$ where $\|\cdot\|_F$ is the Frobenius norm.*

*Proof.* We begin by breaking the Frobenius norm into individual traces.

$$\|W - \Pi W \Pi\|_F^2 = Tr((W - \Pi W \Pi)^2) = Tr(W^2) + Tr(\hat{W}^2) - 2Tr(W\hat{W}) = Tr(W^2) - Tr(\hat{W}^2)$$

Additionally note the following.

$$\begin{aligned} &|\omega_i - \hat{\omega}_i| \leq \gamma \\ \Rightarrow\ &|\omega_i + \hat{\omega}_i| \leq \gamma + 2|\omega_i| \\ \Rightarrow\ &|\omega_i^2 - \hat{\omega}_i^2| \leq \gamma^2 + 2\gamma|\omega_i| \end{aligned}$$

From here, each trace is considered independently, with the intent of upper bounding $\|W - \Pi W \Pi\|_F^2$.

$$\begin{aligned} Tr(\hat{W}) - Tr(\hat{W}^2) &= \sum_{i \in V} (\omega_i^2 - \hat{\omega}_i^2) \\ &\leq \sum_{i \in V} (\gamma^2 + 2\gamma|\omega_i|) \\ \Rightarrow \|W - \hat{W}\|_F^2 &\leq N\gamma(\gamma + 2\|W\|_2) \end{aligned}$$

□

Theorem 4.1 states that, preserving the spectrum of the adjacency matrix while coarsening is sufficient to preserve all edge weight information. This is a far stronger statement than the one proposed in Theorem 3.8. However, this is only applicable when the adjacency spectrum is preserved, not necessarily the Laplacian since the two spectra are not directly related for general graphs. In the case of weighted regular graphs it is easy to check that these are one in the same since, for a weighted regular graph with degree $d$, $\lambda_i = d - \omega_i$. Unfortunately this is not true for most graphs. Using this theorem 4.1 in the general case will require bounding $|\omega_i - \hat{\omega}_i| \leq f(\epsilon)$ for some function $f(\cdot)$ where $|\lambda_i - \lambda| \leq \epsilon$ for all $i \in [1..N]$. This remains an open problem.

**5. Closing remarks.** The contributions of this manuscript have been twofold. A result originally derived by Jin et al. [8] was generalized to the case of the combinatorial Laplacian. We showed that, by using this result, one can closely preserve the spectrum of the graph Laplacian while performing graph coarsening. Additionally it was shown that the suggested coarsening criteria implies bounds on algebraic distances between nodes of the same graph. A comparison between coarsening methods was also presented. The latter half of the manuscript studied how closely the edge weights of a graph's lift approximate those of the original graph under an assumption that their Laplacian spectra are close. A sufficiently tight bound would guarantee that arbitrary data sets in $\mathbb{R}^n$ imbued with a graph structure could be coarsened while preserving their relative embeddings in $\mathbb{R}^n$. This is a novel question with potential applications to image segmentation and data mining. Unfortunately the bound proven relies on the connectivity of the graph and is unlikely to be useful in real world applications. It was then shown that, in the case of weighted regular graphs the connectivity of the graph does not require consideration, and a spectral approximation provides an edge weight approximation. Various avenues for extensions and branching research exist.

One obvious path for future research is to diminish the bound provided in Theorem 3.8. The proof for the theorem relies heavily on a discrepancy bound which is particularly loose. By circumventing this, perhaps with a more sophisticated extension to Lemma 3.4, one may be able to significantly tighten this bound. As an extension of this, removing the dependency on $\sigma$ is important for applicability. In practice $\sigma$ will be too large for this bound to be useful to practitioners. One avenue for exploring this may be to relate spectral differences between the adjacency and coarsened adjacency with those of the Laplacian and coarsened Laplacian, and then apply Theorem 4.1. Additionally, there are several interesting questions one may ask about the effects of coarsening arbitrary data sets. For instance, if the spectrum between a graph $G$ and it's lift are close, how close are their edge weights on average? This is answered in a special case by Theorem 4.1, but is not known in general. This question is significantly less restrictive than the one presented in this paper, however it still provides insight into the effects of coarsening on node embeddings. As for extending the results discussed in section 2, while it was shown that the coarsening criteria in Theorem 2.1 implies a bound on algebraic distances, a result in the opposite direction would be preferable. Algebraic distances are cheap to compute for small numbers of test vectors, and if there were a reasonable guarantee on the accuracy of the spectrum, they would be preferable to the criteria presented in this paper. Such a bound would likely be probabilistic for $k < N$ due to the fact that ensuring linear dependence between nodes using the algebraic distance requires the test vectors to span $\mathbb{R}^N$.

## REFERENCES


[1] R. Bhatia, *Perturbation bounds for matrix eigenvalues*, SIAM, 2007.

[2] C. Brissette, A. Huang, and G. Slota, *Parallel coarsening of graph data with spectral guarantees*, arXiv preprint arXiv:2204.11757, (2022).

[3] J. Chen, Y. Saad, and Z. Zhang, *Graph coarsening: From scientific computing to machine learning*, arXiv preprint arXiv:2106.11863, (2021).

[4] F. R. Chung and F. C. Graham, *Spectral graph theory*, no. 92, American Mathematical Soc., 1997.

[5] C. Godsil and G. F. Royle, *Algebraic graph theory*, vol. 207, Springer Science & Business Media, 2001.

[6] C. Gordon, D. L. Webb, and S. Wolpert, *One cannot hear the shape of a drum*, Bulletin of the American Mathematical Society, 27 (1992), pp. 134–138.

[7] H. Jia, S. Ding, X. Xu, and R. Nie, *The latest research progress on spectral clustering*, Neural Computing and Applications, 24 (2014), pp. 1477–1486.

[8] Y. Jin, A. Loukas, and J. JaJa, *Graph coarsening with preserved spectral properties*, in International Conference on Artificial Intelligence and Statistics, PMLR, 2020, pp. 4452–4462.

[9] M. Kac, *Can one hear the shape of a drum?*, The american mathematical monthly, 73 (1966), pp. 1–23.

[10] G. Karypis and V. Kumar, *Metis: A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices*, (1997).
[11] G. Kron, *Tensor analysis of networks*, New York, (1939).
[12] A. Loukas, *Graph reduction with spectral and cut guarantees.*, J. Mach. Learn. Res., 20 (2019), pp. 1–42.
[13] M. H. McIlwain, *Can you hear the size of the vertices? An inverse spectral problem of Laplacians on weighted graphs*, Rice University, 1998.
[14] D. Ron, I. Safro, and A. Brandt, *Relaxation-based coarsening and multiscale graph organization*, Multiscale Modeling & Simulation, 9 (2011), pp. 407–423.
[15] T. Sahai, A. Speranzon, and A. Banaszuk, *Hearing the clusters of a graph: A distributed algorithm*, Automatica, 48 (2012), pp. 15–24.
[16] J. W. Thomas, *Numerical partial differential equations: conservation laws and elliptic equations*, vol. 33, Springer Science & Business Media, 2013.
[17] F. Tung, A. Wong, and D. A. Clausi, *Enabling scalable spectral clustering for image segmentation*, Pattern Recognition, 43 (2010), pp. 4069–4076.
[18] E. R. Van Dam and W. H. Haemers, *Which graphs are determined by their spectrum?*, Linear Algebra and its applications, 373 (2003), pp. 241–272.
[19] E. R. Van Dam and W. H. Haemers, *Developments on spectral characterizations of graphs*, Discrete Mathematics, 309 (2009), pp. 576–586.
[20] M. J. Zaki, W. Meira Jr, and W. Meira, *Data mining and analysis: fundamental concepts and algorithms*, Cambridge University Press, 2014.
[21] S. Zelditch, *Eigenfunctions of the Laplacian on a Riemannian manifold*, vol. 125, American Mathematical Soc., 2017.
[22] S. Zeng, R. Huang, Z. Kang, and N. Sang, *Image segmentation using spectral clustering of gaussian mixture models*, Neurocomputing, 144 (2014), pp. 346–356.