Graphical Learning Optimization and Dimensionality Reduction with Geometric Multi-Resolution Analysis

Felicia Schenkelberg
Dartmouth College
felicia.schenkelberg.th@dartmouth.edu

Allison Gunby-Mann
Dartmouth College
allison.mann.th@dartmouth.edu

Emma Graham
Dartmouth College
emma.graham.th@dartmouth.edu

Shuoxuan Li
Carnegie Mellon University
shuoxuanli@cmu.edu

Peter Chin
Dartmouth College
pc@dartmouth.edu

***Abstract*—This paper employs Geometric Multi-Resolution Analysis (GMRA) as a technique for dimensionality reduction and explores its impact on high-dimensional graphical learning tasks. The burgeoning surge in data collection practices, driven by technological advancements across diverse domains, has resulted in an influx of datasets wherein the number of features significantly exceeds the number of observations—a paradigm characteristic of high-dimensional datasets. Analyzing such high-dimensional datasets presents immediate challenges owing to the intricacies of dataset complexity as well as the wealth of information encapsulated within each data point. GMRA exploits redundant representations in such high-dimensional datasets, embedding the high-dimensional data into an intrinsic, underlying lower-dimensional structure. This process aims to preserve essential features while reducing dimensionality and facilitate analysis by mitigating the computational complexities associated with analyzing high-dimensional datasets. This paper proposes a novel application of Geometric Multi-Resolution Analysis to dimensionality reduction in graph embeddings. Empirically, its efficacy is validated by its performance in computing the intrinsic, underlying lower-dimensional structure for a comprehensive set of graph learning tasks, including node classification, edge classification, link prediction, anomaly detection, and graph clustering.**

***Index Terms*—graph embedding, dimensionality reduction, machine learning**

I. Introduction

In the rapidly evolving fields of finance, marketing, medicine, and more, technological advancements have significantly altered the landscape of data collection, leading to the emergence of high-dimensional datasets where the number of features often surpasses the number of observations. This shift challenges traditional statistical methods designed predominantly for low-dimensional spaces, where the number of observations exceeds the number of features. In high-dimensional settings, classical techniques such as least squares regression struggle with issues including the bias-variance trade-off and the risk of over-fitting, as they cannot efficiently manage the complexity introduced by the vast number of features.

Further complicating the analysis of high-dimensional data is the issue of multicollinearity, where variables can often be expressed as linear combinations of others, leading to uncertainty in determining which features genuinely impact the outcome. This uncertainty makes it difficult to ascertain optimal regression coefficients and necessitates a cautious approach in model evaluation, emphasizing the importance of using independent test sets or cross-validation over traditional metrics like p-values or R statistics, which can be misleading in such contexts.

Addressing these challenges, dimensionality reduction emerges as a crucial technique. It simplifies high-dimensional data into a more manageable, lower-dimensional representation, allowing for easier analysis without the direct consideration of associated labels. This paper focuses on the application of Geometric Multi-Resolution Analysis (GMRA), a sophisticated tool for dimensionality reduction that exploits redundant representations in data. By embedding high-dimensional datasets into their intrinsic lower-dimensional structures, this technique not only preserves essential features but also reduces computational complexities, facilitating a more efficient analysis of complex datasets. Empirically, we employ GMRA and effectively compute the intrinsic, underlying lower-dimensional structure of a selection of datasets and evaluate its consequential impact on a comprehensive gamut of graph learning tasks, including but not limited to node classification, edge classification, link prediction, anomaly detection, and graph clustering.

II. Background

A. Intrinsic Dimensions and Manifold Projection

The development of computational methods that capture data's intrinsic geometry has revolutionized the analysis of high-dimensional data. Manifold-learning algorithms typically assume a finite set of data points drawn randomly from a smooth t-dimensional manifold with a metric defined by

geodesic distance. These points are then embedded into a high-dimensional input space with Euclidean metric, resulting in the input data points. Linear manifold learning, a subset of manifold learning, focuses on linear dimensionality reduction. It views data observed in a high-dimensional space as potentially close to a lower-dimensional linear manifold, with the intrinsic dimensionality of the manifold assumed to be much smaller than the data dimensionality. Identifying such linear manifolds is akin to the classical statistics problem of linear dimensionality reduction, often achieved through projection methods like principal component analysis (PCA) [1], a widely used technique in this domain. When the manifold is linear, data can be projected onto a linear combination of dictionary columns in low-dimensional space using techniques like singular value decomposition (SVD) [2]. However, many datasets exist on nonlinear manifolds, requiring more sophisticated techniques for dimensionality reduction and feature extraction.

B. Geometric Multi-Resolution Analysis

Geometric Multi-Resolution Analysis (GMRA) provides a powerful framework for estimating and representing these nonlinear manifolds within high-dimensional spaces. This analytical method excels with complex datasets, such as point clouds or graph-structured data, which are not adequately addressed by linear methods. The hierarchical dissection produced by GMRA's tree decomposition results in finer groups of points with shared local topology at deeper layers of the tree. Affine approximations are formed for each subset, with basis functions computed via singular value decomposition (SVD) of subset covariance, capturing the local area's tangent planes. Difference operators, representing changing scales, are expressed as wavelet bases orthogonal to the current scale's basis function. This method effectively maps points to their approximate locations while managing error specific to scaling. This study aims to apply GMRA in the analysis of graphs and networks, leveraging its ability to uncover hierarchical structures within the data, as seen advantageous in previous work [3] [4] [5].

Neural Networks are known for their ability to identify patterns in data that adhere to low-dimensional manifolds, a feature prominently displayed in digital image processing tasks like the MNIST dataset. Expanding upon these insights, we apply GMRA to graph and network data to explore both supervised and unsupervised learning paradigms.

The GMRA construction involves a multi-scale nested decomposition of the dataset into partitions arranged in a dyadic tree structure, as detailed in 2012 by Allard et al. [3]. Each node of the tree represents a subset of the data at a certain scale, and the subsets at each scale collectively partition the dataset. The GMRA algorithm can be summarized as follows:

1. Insert data points into a cover tree where each node is a point and an edge is drawn if they are within euclidean distance of 2^s of each other, where s is the scale at each level.

2. Create a dyadic tree structure using the cover tree as input by greedily clustering points within a radius, with the radius shrinking at lower levels of the tree to generate levels of local geometry.

3. For each cluster in the tree, compute a d-dimensional affine approximation of the data points, creating a wavelet tree with local linear approximations to the manifold.

4. Construct low-dimensional affine difference operators that encode the differences between approximations at consecutive scales [3].

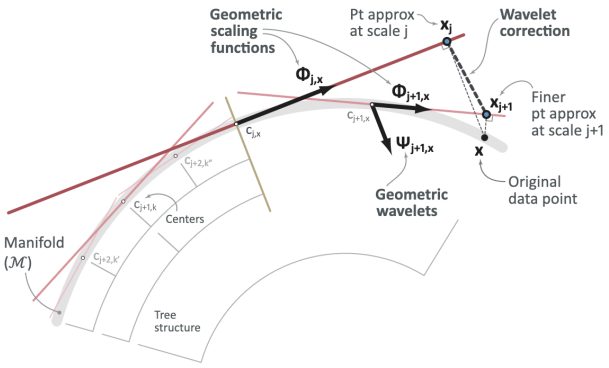

Fig. 1. An illustration of the geometric wavelet decomposition, step 3 in the GMRA algorithmic process. The d-dimensional linear approximations are given by the red tangent planes to the manifold ($\mathcal{M}$). [3]

The basis functions for these approximations, known as geometric scaling functions and wavelets, are determined by SVD on the covariance of the data points within each subset, capturing local tangent spaces [6]. These functions allow for the hierarchical representation of data, with successive refinements at each scale that provide a more accurate approximation of the underlying manifold.

III. Applications of GMRA

A. Signal Processing

The application of GMRA in signal processing, particularly with the MusicNet dataset, presents a novel approach to representing and analyzing complex musical recordings. Since MusicNet consists of classical music recordings with over one million note-tracking labels, the dataset poses a significant challenge due to its high dimensionality. However, by leveraging GMRA, it becomes possible to compute low-dimensional representations of these recordings, facilitating more efficient processing and analysis.

In the MusicNet experiments, GMRA was applied to six recordings, each sampled at 11 kHz, reducing data from 1-10 million dimensions to a 50-dimensional space. This reduction enabled the efficient representation of the recordings while preserving essential musical features. The compressed data was then trained on a shallow, two-hidden-layer neural network. As a result, GMRA shows competitive performance with significantly fewer learnable parameters in comparison to state-of-the-art models [3].

The application of GMRA in signal processing aligns with the broader context of harmonic analysis and efficient signal representations. Traditionally, linear representations in terms

of dictionaries of atoms have been studied extensively in harmonic analysis, with a focus on achieving sparsity or concentration of coefficients. Various dictionaries, including Fourier-like bases, wavelets, ridgelets, and curvelets, have been developed to provide optimal representations for different classes of signals or operators. The trend towards over-complete dictionaries, such as frames and libraries of dictionaries, allows for non-unique representations and enables fast transforms through multi-scale organization. GMRA contributes to this landscape by offering a geometrically motivated approach to signal representation, complementing existing techniques and advancing the field of signal processing.

B. Digital Image Processing

The application of GMRA in the context of the MNIST dataset showcased its potential in image processing and pattern recognition tasks. By generating low-dimensional representations for MNIST images, GMRA effectively reduced the dimensionality of the dataset while preserving essential geometric structures. In particular, the use of GMRA produces embeddings as compact as 11 dimensions, allowing for a more efficient representation of the handwritten digits as opposed to the original 784 dimensions.

Through experiments conducted with GMRA embeddings on the MNIST dataset, notable insights have been uncovered [7]. One significant finding is the ability to train neural networks with significantly fewer parameters while maintaining competitive performance. By embedding MNIST images into an 11-dimensional space and utilizing shallow neural network architectures, the study achieved impressive results, demonstrating the effectiveness of GMRA in reducing model complexity without sacrificing accuracy.

Furthermore, the exploration of GMRA's application in image reconstruction provides valuable insights into its capabilities. By reconstructing MNIST images from their low-dimensional embeddings, GMRA showcases its ability to capture essential features of the handwritten digits while achieving efficient representations. This opens up avenues for further research in refining GMRA-based methods for image reconstruction and representation learning, with potential applications across various domains beyond MNIST. The ability to capture essential features of handwritten digits through low-dimensional embeddings showcases the potential of GMRA in broader applications in pattern recognition and data analysis within the field of machine learning [3] [4] [8].

C. Remote Sensing Image Fusion

Beyond these examples, GMRA is integral to the fusion of remote sensing images. It is a key component among the multi-resolution analysis (MRA) and multi-geometric analysis (MGA) tools used to combine information from different sources into a coherent image [9]. GMRA's effectiveness in this domain is crucial for enhancing the detail and quality of information obtained from earth observation technologies, reflecting its significant impact on the field of remote sensing. The versatility of GMRA across these applications demonstrates its robustness in extracting and preserving valuable data characteristics. Whether applied to music, images, graphs, or satellite imagery, GMRA enhances our ability to analyze complex datasets, offering a pathway to more insightful and efficient data processing.

IV. Method

In this section we describe our experimental method to demonstrate the impact of GMRA embeddings on the performance of various graph tasks compared to the original, unreduced embeddings.

The experimental procedure involved several key steps. First, to make the problem compatible with GMRA, we computed embeddings from the induced graphs of datasets. Then we conducted baseline experiments on these embeddings using off-the-shelf models without employing any dimensionality reduction techniques, allowing us to establish a control for comparison. Subsequently, we applied the GMRA algorithm to the embeddings and extracted the reduced embeddings from the wavelet tree. Finally, we ran the same baseline experiments on these dimensionally reduced embeddings.

Our evaluation extended beyond traditional classification tasks to encompass a broader spectrum of graph-based analyses. We delved into tasks such as link prediction, anomaly detection, graph clustering, and visualization, seeking a comprehensive understanding of the efficacy of GMRA in enhancing graph learning across different domains and tasks. We aim to gain insights into the strengths and limitations of GMRA as a dimensionality reduction technique for graph-based machine learning.

A. Initial Graph Embedding

For the initial graph embedding step, we experimented with GraphSage [10] and node2vec [11] embeddings to generate the point clouds representing our input graphs. Node2vec is an embedding technique based on random walks. GraphSage is a more modern embedding technique based on a graph neural network (GNN); it aggregates neighbor features into an embedding. We used event counts as the node feature vectors for preliminary experiments. Under the application of GMRA, the GraphSage embeddings displayed a significant dimension reduction compared to node2vec due to the large amount of redundancy in the embeddings. There were also many duplicate embeddings, which indicates that some network structure was lost during dimensionality reduction. For that reason, the majority of the experimental results are done with the node2vec embeddings as seeds.

B. Iterative Batch Processing

In order to mitigate the expense of inserting many nodes into a cover tree at once, we implemented an option for batch processing, which was not in the original GMRA workflow. Each point in the dataset is processed and inserted into the cover tree data structure sequentially, maintaining the tree's integrity throughout the insertion process. We begin by initializing the

cover tree with parameters such as maximum scale and base value. The insert method handles point insertion by iterating through each point in the dataset, either initializing the tree with the first dataset or appending subsequent dataset to the existing structure. When inserting a point, we first determine whether the new point becomes the root node or is inserted as a child of an existing node. It is important to note that the point's position in the tree is determined by calculating its distance from nodes at different scales and selecting an appropriate parent based on distance criteria. After each insertion, the tree's structure is updated to maintain correctness. The results we experimentally found showed exact replication between points inserted in batches and points inserted all at once, but the batch processing allows for generating results at intermediate steps or adding in new points later in an online process without having to recompute the entire cover tree.

C. Embedding Extraction

The process of extracting embeddings from a wavelet tree-based data structure is systematic. Initially, GMRA calculates difference operators across tree levels to enable query initiation from the root, embedding retrieval, and subsequent traversal with updates using linear difference operators. We then navigate the tree to pinpoint nodes at a consistent depth where all nodes share the same dimensionality. This step ensures coherence in the extracted embeddings. Finally, the information from these identified nodes, encompassing basis vectors, indices, and scaling factors is aggregated to construct the embeddings matrix representing the intrinsic, lower-dimensional features associated with each dataset point at the chosen scale.

D. Inverse Wavelet Reconstruction

We implemented the reconstruction of lower-dimensional vector embeddings to higher-dimensional vector embeddings using first generation wavelet transforms (FGWT) focusing specifically on images from the MNIST dataset for visualization and verification purposes (see Fig. 2). The algorithm computes wavelet coefficients for each data point in the lower-dimensional input, which involves traversing the wavelet tree structure and determining the coefficients at different scales. Before applying the FGWT, the wavelet tree structure must be constructed, defining the hierarchical arrangement of wavelet coefficients and bases. This step is essential as it facilitates the efficient calculation of wavelet coefficients based on input data coefficients, scaling bases, and wavelet bases. The core of the reconstruction process involves inversely traversing the wavelet tree from leaf nodes to the root. At each scale, projections are accumulated to reconstruct the original data. This process ensures that the lower-dimensional vector embeddings are accurately reconstructed.

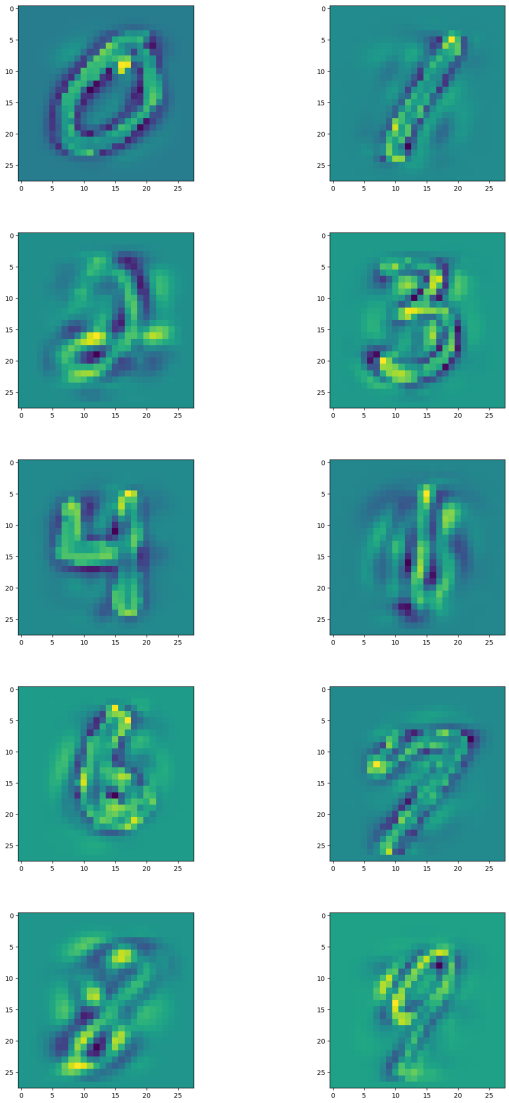

Fig. 2. This figure demonstrates the bidirectional verification method applied to MNIST digits [0-9]. By computing the inverse of the GMRA reduced embeddings in higher-dimensional spaces, we ensure precise representations of data in both datasets, affirming the accuracy of calculations.

V. Datasets

The LANL 2015 Comprehensive Multi-Source Cyber Security Events dataset spans 57 days of log files from five different sources within the Los Alamos National Laboratory's internal corporate network [12]. It encompasses both normal activity and a redteam malicious campaign.

Two datasets from StellarGraph were employed. The Cora dataset consists of 2708 scientific publications categorized into seven distinct classes. Its citation network is comprised of 5429 links, and each publication is represented by a binary word vector indicating the presence or absence of words from a 1433-word dictionary.

The CiteSeer dataset consists of 3312 scientific publications across six classes. Its citation network has 4732 links, though only 4715 are used as 17 of them involve publications absent from the dataset. Similar to Cora, each publication in CiteSeer is described by a binary word vector reflecting word presence or absence from a dictionary comprising 3703 unique words.

VI. Experiments

This research investigates the application of GMRA in graph and network processing, particularly with the StellarGraph and LANL datasets, presenting a novel approach to representing and analyzing complex graphs and networks. These datasets pose a significant challenge due to their high dimensionality. However, by leveraging GMRA, it becomes possible to compute low-dimensional representations of these graphs and networks, facilitating more efficient processing and analysis.

We undertook an exploration into the utilization of GMRA within graph and network processing, encompassing both supervised and unsupervised learning tasks.

A. Classification

1) Node Classification: Node classification is a prevalent machine learning task in graph data analysis. While node classification may seem akin to standard supervised classification, it differs significantly due to the interdependent nature of graph nodes. Unlike independent and identically distributed (i.i.d.) data points assumed in standard supervised learning, nodes in a graph are interconnected, necessitating modeling of these interdependencies. Successful node classification approaches capitalize on these connections, leveraging concepts like homophily (nodes sharing attributes with neighbors) and structural equivalence (nodes with similar local structures having similar labels). These concepts, along with heterophily (nodes connecting preferentially to those with different labels), guide the construction of node classification models that capture node relationships rather than treating them as independent data points.

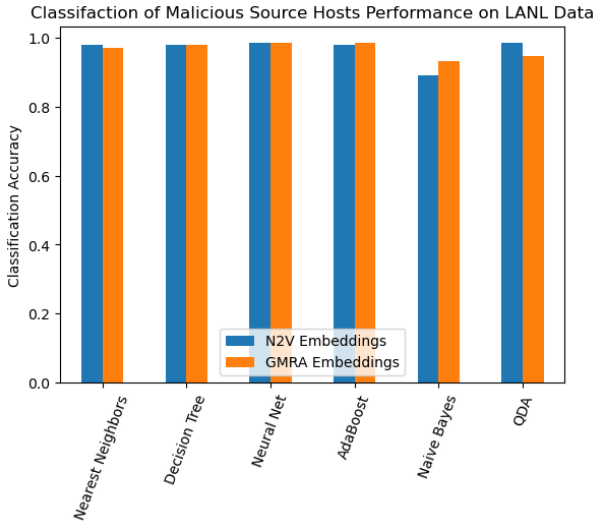

Fig. 3. Comparative Iterative Analysis of Classification Methods on the LANL Dataset Using Original and GMRA-Reduced Embedding Vectors. The graph depicts a comparative study of classification methods applied iteratively to the LANL dataset, utilizing both the original feature vectors and the reduced embedding vectors obtained through GMRA. The nodes range from 64 in the initial graph to 8908 in the final graph, showcasing the scalability and performance impact of dimensionality reduction techniques.

TABLE I
Node Classification on StellarGraph Datasets

Dataset	Original Dimension	Original Accuracy	Reduced Dimension	Reduced Accuracy
Cora	128	0.7045	16	0.7068
CiteSeer	128	0.7553	13	0.7547

Interpretation of Results: Results for LANL are shown in Fig. 3 and results for Stellargraph are in Table I. The accuracy score was computed using scikit-learn metrics accuracy score. The comparable accuracy scores between the original and GMRA reduced datasets indicate consistent model performance across both sets of datasets. We reduced the LANL dataset embedding vectors from 256 dimensions to their intrinsic 31 dimensions. A similar reduction was applied to the stellargraph learned vector embeddings from the Cora and CiteSeer datasets, both initially having 128 dimensions. However, our analysis indicates that Cora's intrinsic dimensions are 16, while CiteSeer's are 13. The high accuracy scores indicate accurate predictions for the majority of samples in both datasets, signifying a strong generalization of the model to both the original and GMRA reduced datasets. See Fig. 4 for a visual comparison of node embeddings with and without GMRA.

2) Edge Classification: Edge classification involves predicting attributes or labels associated with the edges of a graph. Unlike node classification, which focuses on predicting attributes of individual nodes, edge classification aims to understand the relationships between nodes by predicting properties specific to the connections between them. This task is essential in various domains such as social network analysis, where edges represent relationships between individuals, and predicting attributes on these edges can provide insights into the nature of connections, such as friendship strength or interaction frequency. Edge classification models can be utilized for tasks like link prediction, where the goal is to predict whether a connection will form between two nodes in the future based on existing graph topology and attributes. Additionally, edge regression models can be employed to predict continuous attributes on edges, enabling applications like predicting the strength of interactions or the likelihood of transactions between entities in a network.

TABLE II
Edge Classification on StellarGraph Datasets

Dataset	Original Dimension	Original Accuracy	Reduced Dimension	Reduced Accuracy
Cora	128	0.93	13	0.90

Interpretation of Results: Results are shown in Table II. The accuracy score was computed using scikit-learn metrics accuracy score. The comparable accuracy scores between the original and GMRA reduced datasets indicate consistent model performance across both datasets. For this experiment, we

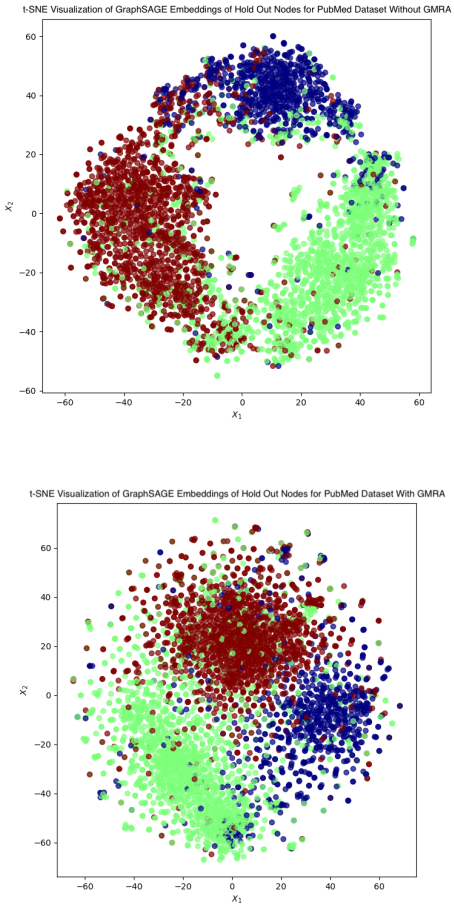

Fig. 4. Comparison of GraphSAGE Classification Results on the PubMedDiabetes Dataset Using Original and GMRA-Reduced Embedding Vectors. This figure presents a comparative analysis of the PubMedDiabetes dataset before and after GMRA dimensionality reduction. Node embeddings were learned and classified using the GraphSAGE algorithm. The visual representation utilizes t-SNE [13] from the scikit-learn manifold library, further reducing both sets of embeddings for effective plotting.

reduced embedding vectors from 128 dimensions to its lower intrinsic dimensions. The high accuracy scores reflect accurate predictions for the majority of samples in both datasets, indicating a robust model generalization to both the original and GMRA reduced datasets.

3) Link Prediction: Link prediction, also known as graph completion or relational inference, is a fundamental task in machine learning applied to graph data. It addresses scenarios where information about relationships between nodes is incomplete or missing entirely. The primary goal of link prediction is to infer information about a node based on its relationship with other nodes in the graph. The complexity of link prediction varies depending on the graph's nature; while simple graphs like social networks may rely on basic heuristics, more complex multi-relational graphs such as biomedical knowledge graphs require sophisticated reasoning and inference strategies. Link prediction blurs traditional machine learning boundaries, being both supervised and unsupervised, and necessitates domain-specific inductive biases. It encompasses various variants, including predictions within a single graph and predictions across multiple disjoint graphs.

TABLE III
LINK PREDICTION ON STELLARGRAPH DATASETS

Dataset	Original Dimension	Original ROC AUC	Reduced Dimension	Reduced ROC AUC
Cora	128	0.9667	13	0.9792

Interpretation of Results: Results are shown in Table III. The ROC AUC score was calculated to assess the performance of the logistic regression classifier (LogisticRegressionCV) from scikit-learn for the binary classification. The comparable scores between the original and GMRA reduced datasets indicate consistent model performance across both sets of datasets. For this experiment, we were able to reduce the Cora dataset embedding vectors from 128 dimensions to an intrinsic 13 dimensions, as demonstrated above. The high accuracy scores indicate accurate predictions for the majority of samples in both datasets, reflecting effective learning, improved generalization, and model stability across both the original and GMRA reduced datasets.

B. Anomaly Detection

Anomaly detection in graphs involves identifying irregularities or deviations from expected patterns within graph structures. These anomalies can manifest in various forms, such as the presence or absence of vertices or edges, the existence of anomalous subgraphs like near-stars (vertices with sparse connections among neighbors) or near-cliques (densely connected subgraphs), heavy vicinity (where most edge weight is concentrated on a few edges), and dominant heavy links (edges with exceptionally large weights compared to others). In applications like fraud detection, anomaly detection plays a crucial role in uncovering suspicious activities camouflaged as legitimate transactions to evade detection. Algorithms for graph-based anomaly detection often employ breadth-first search and minimum description length (MDL) principles to extract normative graph patterns that best compress the graph in MDL terms. These algorithms analyze changes in normative patterns, approximate the probability of anomalous additions, and investigate sub-patterns to detect anomalous absences, providing a comprehensive approach to identifying graph anomalies across various contexts.

Interpretation of Results: Results are shown in Table IV. The accuracy score was computed using scikit-learn metrics accuracy score. The comparable accuracy scores between the original and GMRA reduced datasets indicate consistent model performance across both sets of datasets. For this experiment, we were able to reduce the LANL dataset embedding vectors from 256 dimensions to an intrinsic 31 dimensions. The high accuracy scores indicate accurate predictions for the majority of samples in both datasets, signifying a strong generalization of the model to both the original and GMRA reduced datasets.

TABLE IV
ANOMALY DETECTION ON LANL DATASET

Method	Original Dimension	Original Accuracy	Reduced Dimension	Reduced Accuracy
Robust Covariance	256	0.9511	31	0.9481
One-Class SVM	256	0.9490	31	0.9463
Isolation Forest	256	0.9479	31	0.9461
Local Outlier Factor	256	0.9517	31	0.9468

C. Graph Clustering

Graph clustering aims to group vertices in a graph into clusters where vertices within each cluster exhibit strong connections while having sparse connections with vertices outside the cluster. It generally refers to a subset of vertices that are densely interconnected. Cluster density can be measured by the density of the induced subgraph within the cluster, while connectivity outside the cluster is assessed by the size of the graph cut that removes the cluster from the graph. Unlike node classification and relation prediction, which involve inferring missing information in graph data akin to supervised learning, community detection is the graph analogue of unsupervised clustering. For example, in a collaboration graph where researchers are connected if they co-authored a paper, community detection seeks to identify clusters representing research areas, institutions, or other demographic factors. Real-world applications of community detection include identifying functional modules in genetic networks and uncovering fraudulent user groups in financial transaction networks.

TABLE V
GRAPH CLUSTERING ON STELLARGRAPH DATASETS

Dataset	Original Dimension	Original Accuracy	Reduced Dimension	Reduced Accuracy
CiteSeer	128	0.7891	14	0.7701

Interpretation of Results: Results are shown in Table V. The accuracy score was calculated using scikit-learn metrics accuracy score. The similar accuracy scores between the original and GMRA reduced datasets suggest stable model performance across both datasets. In this experiment, we successfully reduced the CiteSeer dataset embedding vectors from 128 dimensions to an intrinsic 14 dimensions, as demonstrated above. The high accuracy scores indicate accurate predictions for the majority of samples in both datasets, indicating robust model generalization to both the original and GMRA reduced datasets.

VII. CONCLUSION

In this paper we have proposed a novel technique of applying GMRA for dimensionality reduction in graph embeddings. By demonstrating comparable performance to its higher-dimensional counterparts for various classical graph tasks including node classification, edge classification, link prediction, anomaly detection, and graph clustering, it validates the efficacy of GMRA and underscores its efficiency in reducing computational complexity without sacrificing predictive power for downstream processes.

This study proves the potential of leveraging dimensionality reduction techniques like GMRA to address the challenges posed by high-dimensional data in graph based tasks. Additionally, it sheds light on several avenues for future exploration and enhancement:

1) **Fine-tuning GMRA Parameters:** Investigation into optimal parameter settings of GMRA, such as the choice of clustering function and hyperparameters based on the specific graph task, may further improve the algorithm's accuracy.
2) **Exploring Additional Graph Tasks:** Such examples potentially include graph classification and cross network alignment to provide evidence for broader applicability.
3) **Improving Scalability:** Currently, the GMRA algorithm is inefficient on graphs with a very large number of nodes due to the cover tree algorithm. There is an opportunity here for improving scalability of the graph via optimization or investigating alternative clustering techniques.
4) **Interpretability and Visualization:** Developing techniques to interpret and visualize the transformed embeddings obtained through GMRA could provide insights into the underlying structures and relationships present within complex network data.

ACKNOWLEDGMENT

Distribution Statement A (Approved for Public Release, Distribution Unlimited): The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. This Research was developed with funding from the Defense Advanced Research Projects Agency (DARPA).

REFERENCES

[1] K. Pearson, "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901. [Online]. Available: https://doi.org/10.1080/14786440109462720

[2] G. W. Stewart, "On the early history of the singular value decomposition," *SIAM Review*, vol. 35, no. 4, pp. 551–566, 1993. [Online]. Available: https://doi.org/10.1137/1035134

[3] W. K. Allard, G. Chen, and M. Maggioni, "Multi-scale geometric methods for data sets ii: Geometric multi-resolution analysis," *Applied and computational harmonic analysis*, vol. 32, no. 3, pp. 435–462, 2012.

[4] G. Chen, M. Iwen, S. Chin, and M. Maggioni, "A fast multiscale framework for data in high-dimensions: Measure estimation, anomaly detection, and compressive measurements," in *2012 Visual Communications and Image Processing*, 2012, pp. 1–6.

[5] Y. Wang, G. Chen, and M. Maggioni, "High-dimensional data modeling techniques for detection of chemical plumes and anomalies in hyperspectral images and movies," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4316–4324, 2016.

[6] D. N. Tran and S. P. Chin, "Geometric multi-resolution analysis based classification for high dimensional data," in *Cyber Sensing 2014*, I. V. Ternovskiy and P. Chin, Eds., vol. 9097. SPIE, 2014, pp. 132 – 139, backup Publisher: International Society for Optics and Photonics. [Online]. Available: https://doi.org/10.1117/12.2063316

[7] H. Le, A. Wood, S. Dandekar, and P. Chin, "Neural network optimization with biologically inspired low-dimensional manifold learning," in *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, 2021, pp. 8–13.

[8] A. Wood, "Accelerating machine learning with memory management and persistent memory," Ph.D. dissertation, 2023.

[9] P. Mangalraj, V. Sivakumar, and S. e. a. Karthick., "A review of multi-resolution analysis (mra) and multi-geometric analysis (mga) tools used in the fusion of remote sensing images," *Circuits, Systems, and Signal Processing*, vol. 39, 06 2020.

[10] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 1025–1035.

[11] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 855–864. [Online]. Available: https://doi.org/10.1145/2939672.2939754

[12] A. D. Kent, "Comprehensive, multi-source cyber-security events data set," 5 2015. [Online]. Available: https://www.osti.gov/biblio/1179829

[13] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, no. 86, pp. 2579–2605, 2008. [Online]. Available: http://jmlr.org/papers/v9/vandermaaten08a.html

[14] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for machine learning*. Cambridge University Press, 2020.

[15] W. Jia, M. Sun, J. Lian, and S. Hou, "Feature dimensionality reduction: a review," 2022.

[16] A. Belkina, C. Ciccolella, and R. e. a. Anno, "Automated optimized parameters for t-distributed stochastic neighbor embedding improve visualization and analysis of large datasets," *Nature Communications*, 2019.

[17] A. Paranjape, A. R. Benson, and J. Leskovec, "Motifs in temporal networks," 2017.