

# Quantum-Assisted Correlation Clustering

Antonio Macaluso<sup>1\*</sup>, Supreeth Mysore Venkatesh<sup>1,2\*</sup>, Diego Arenas<sup>3</sup>,  
Matthias Klusch<sup>1</sup>, Andreas Dengel<sup>1,3</sup>

<sup>1</sup>*German Research Center for Artificial Intelligence (DFKI), Saarbruecken, Germany*

<sup>2</sup>*University of Kaiserslautern-Landau (RPTU), Kaiserslautern, Germany*

<sup>3</sup>*German Research Center for Artificial Intelligence (DFKI), Kaiserslautern, Germany*

{antonio.macaluso, diego.arenas,  
matthias.klusch, andreas.dengel}@dfki.de  
supreeth.mysore@rptu.de

**Abstract**—This work introduces a hybrid quantum-classical method to correlation clustering, a graph-based unsupervised learning task that seeks to partition the nodes in a graph based on pairwise agreement and disagreement. In particular, we adapt GCS-Q [1], a quantum-assisted solver originally designed for coalition structure generation, to maximize intra-cluster agreement in signed graphs through recursive divisive partitioning. The proposed method encodes each bipartitioning step as a quadratic unconstrained binary optimization problem, solved via quantum annealing. This integration of quantum optimization within a hierarchical clustering framework enables handling of graphs with arbitrary correlation structures, including negative edges, without relying on metric assumptions or a predefined number of clusters. Empirical evaluations on synthetic signed graphs and real-world hyperspectral imaging data demonstrate that, when adapted for correlation clustering, GCS-Q outperforms classical algorithms in robustness and clustering quality on real-world data and in scenarios with cluster size imbalance. Our results highlight the promise of hybrid quantum-classical optimization for advancing scalable and structurally-aware clustering techniques in graph-based unsupervised learning.

**Index Terms**—Correlation Clustering, Quantum Annealing, Graph Partitioning, Unsupervised Learning

## I. INTRODUCTION

Correlation clustering (CC) [2] is a graph-based paradigm that extends traditional clustering methods by directly modeling pairwise relationships. Unlike approaches such as  $k$ -means or hierarchical clustering, which require data to be embedded in a metric space and rely on geometric proximity, CC operates on a weighted graph where edges represent real-valued affinities. Importantly, these weights can be negative, capturing dissimilarities and more general forms of conflict or disagreement between data points. As a result, it can naturally capture non-metric, asymmetric, or context-dependent relations between objects, making it more flexible in modeling diverse types of data, including those where distances are ill-defined or unavailable. In social network analysis [3], CC enables the detection of communities by grouping individuals based on patterns of affinity or antagonism. In recommendation systems [4], it supports collaborative filtering by accounting for both

agreement and disagreement in user preferences. Applications also span natural language processing [2], where CC helps group texts based on semantic relationships such as synonymy and antonymy, as well as in bioinformatics [5], where it has been used to identify groups of genes with similar patterns of variation in gene expression data, distinguishing co-expressed from anti-correlated gene profiles. Notably, CC plays a role in Earth Observation (EO), particularly in the analysis of high-dimensional hyperspectral imagery, where it helps identify clusters within subspaces that are not necessarily aligned with the coordinate axes [6]. In this case, CC can detect redundant spectral bands, reducing dimensionality and improving interpretability without relying on strict axis-aligned projections.

Classical methods for CC typically rely on heuristics and assumptions about the underlying data structure to achieve computational efficiency [2], [7]. These assumptions, such as spherical clusters in  $k$ -means or connectivity in spectral clustering, guide the algorithm’s behavior but may not hold in all real-world scenarios. As a result, there is often a trade-off between the quality of the clustering solution and the efficiency of the heuristic used to obtain it. While these approximations allow algorithms to scale to large datasets, they may also lead to suboptimal or biased partitions when the true data distribution deviates from the assumed model.

To overcome these limitations, it is useful to revisit CC from a more general, graph-theoretic perspective—one that naturally connects to the problem of coalition structure generation (CSG) in induced subgraph games (ISGs) [8]. In ISGs, each node corresponds to an agent, and the value of a coalition is defined as the sum of the weights of the edges within the subgraph it induces. When correlation clustering is applied to a signed or weighted graph to maximize intra-cluster agreement—namely, the total weight of edges within clusters—the problem becomes equivalent to identifying an optimal coalition structure in an ISG. This formulation also aligns with the broader task of community detection [9], [10]. Both frameworks aim to partition the graph to maximize internal cohesion, and an optimal solution in one domain translates directly to the other.

However, the equivalence holds only under specific conditions. CC becomes distinct from CSG in ISGs when the

This work has been partially funded by the German Ministry for Education and Research (BMB+F) in the project QAIAC-QAI2C under grant 13N17167.

\* These authors contributed equally to this work.

objective is to minimize disagreement rather than maximize agreement, when edge weights are limited to binary labels without magnitude, or when the context lacks game-theoretic considerations such as strategic behavior or fairness. In contrast, CSG in ISGs focuses on identifying partitions that maximize total value (social welfare) and negotiating stable and fair payoff allocations among rational agents. These stability considerations are formalized through solution concepts such as the Core, the Shapley value, and the Kernel [11].

In this paper, we adapt GCS-Q [1], a hybrid quantum annealing-based solver originally developed for CSG in ISGs, to address the task of correlation clustering on weighted graphs. By leveraging GCS-Q’s ability to maximize intra-cluster edge weights, we reinterpret the coalition value maximization objective as an agreement-based correlation clustering criterion. We benchmark the adapted algorithm against several classical clustering methods, including  $k$ -means, Partitioning Around Medoids (PAM), Hierarchical, and Spectral clustering (for a comprehensive description of all clustering methods see [12]), across a wide range of settings. First, evaluating performance on simulated graph-based datasets with varying cluster size distributions to enable a rigorous assessment of each method’s ability to recover meaningful partitions, particularly in the presence of substantial imbalance. Cluster size variation is quantified using the Gini index, which captures configurations from nearly uniform to highly skewed. Complementary experiments on EO datasets are conducted to demonstrate the stability and effectiveness of the approach compared to classical methods on real-world data.

## II. RELATED WORKS

The problem of Correlation Clustering focuses on partitioning a dataset based on pairwise relationships between data points, whether they are similar (positively correlated) or dissimilar (negatively correlated), without requiring an explicit geometric embedding or metric assumptions. This stands in contrast to traditional distance-based clustering methods, which rely on predefined distance functions and often fail to distinguish between different types of correlations [7], [12]–[14]. The CC problem is known to be NP-hard [15], and among classical approaches, spectral clustering [16] has emerged as the most effective method for correlation-based tasks [17]–[19]. By constructing a similarity graph and analyzing the spectral properties (eigenvalues and eigenvectors) of its Laplacian, spectral clustering captures the global structure of the data, including indirect relationships that are often critical in correlation graphs. This enables the identification of coherent groups of positively correlated points, even when they are not spatially adjacent in the original feature space.

A significant limitation shared by most classical clustering methods is the need to specify the number of clusters  $k$  in advance. This imposes a reliance on model selection heuristics or costly cross-validation procedures, where multiple values of  $k$  must be evaluated to identify the optimal partitioning. Such trial-and-error approaches hinder scalability and interpretability,

especially in exploratory settings where the true number of clusters is unknown.

Recently, quantum approaches have been proposed to improve clustering by offloading specific computational routines to quantum hardware. In particular, gate-based quantum algorithms [20]–[22] offer theoretical speedups in fault-tolerant settings, but empirical evaluations are limited to very small datasets and depend heavily on classical pre- and post-processing. Furthermore, these methods do not introduce fundamentally new clustering strategies, but focus on accelerating existing algorithms. Quantum annealing-based approaches to graph partitioning typically reformulate the entire clustering problem as a quadratic unconstrained binary optimization (QUBO) tailored to specific applications [23], but are not scalable and generalizable.

## III. METHODOLOGY

Our proposed method follows the principles of hierarchical divisive clustering [1], [14], a top-down approach that starts with a single cluster containing all nodes and recursively partitions it into smaller, more coherent clusters. Unlike agglomerative strategies that rely on local distances and pairwise merges, divisive methods operate by globally optimizing a splitting criterion at each step. This framework is particularly well-suited for correlation clustering, where the objective is to group similar nodes and separate dissimilar ones, based on edge weights representing pairwise agreement/disagreement. The *GCS-Q algorithm* [1], originally developed for coalition structure generation in induced subgraph games, naturally fits this paradigm. We adapt it here for correlation clustering by reinterpreting its coalition value maximization objective as the task of maximizing *intra-cluster agreement*.

Given a weighted undirected graph  $G = (V, E, w)$ , where  $w_{ij} \in \mathbb{R}$  reflects the affinity between nodes  $i$  and  $j$ , the goal is to find a partition  $\mathcal{C} = \{C_1, \dots, C_k\}$  that maximizes:

$$\max \sum_{i,j \in S} w_{ij} \cdot \mathbb{I}[x_i = x_j]. \quad (1)$$

Here,  $x_i \in \{0, 1\}$  is a binary variable indicating the cluster assignment of node  $i$ . For a weighted graph with both positive and negative edge weights  $w_{ij}$ , Eq. (1) corresponds to the *maximize-agreement* variant of CC and is equivalent to minimizing the total weight of edges cut across clusters. Therefore, the goal is to maximize intra-cluster agreement, where positively weighted edges connect nodes within the same cluster, and negatively weighted edges span different clusters. To formalize this, we note that the indicator function for agreement between node labels can be expressed as:

$$\mathbb{I}[x_i = x_j] = 1 - (x_i + x_j - 2x_i x_j), \quad (2)$$

which evaluates to 1 when  $x_i = x_j$ , and 0 otherwise. Using this identity, the agreement maximization objective becomes:

$$\max_{x \in \{0,1\}^n} \sum_{i < j} w_{ij} (1 - x_i - x_j + 2x_i x_j). \quad (3)$$

Dropping the constant term (which does not affect the optimization), this formulation is fully compatible with QUBO solvers and equivalent in objective value to the disagreement minimization version used in GCS-Q. It provides a clear and implementable pathway for optimization on quantum annealing to perform correlation clustering based on agreement maximization.

Importantly, a key distinction between GCS-Q [1] and classical hierarchical divisive clustering lies in the splitting strategy. Traditional divisive methods typically rely on heuristic or local distance-based criteria (e.g., edge betweenness or spectral cuts), which may fail to identify globally optimal partitions. Moreover, these classical approaches generally assume metric properties and non-negative similarities, making them ill-suited for settings involving signed relationships or negative correlations. In contrast, the adaptation of GCS-Q inherits the advantages of the original formulation [1] and explicitly optimizes a global objective over the current subgraph at each iteration using combinatorial search guided by quantum annealing. This enables GCS-Q to evaluate all possible binary bipartitions simultaneously, thereby avoiding the suboptimal greedy decisions often made by classical methods. As a result, GCS-Q is expected to outperform classical divisive approaches, particularly when applied to graphs with both positive and negative edge weights. Moreover, since GCS-Q does not rely on an explicit notion of distance, it is also expected to outperform its classical counterpart when clustering is driven by general correlation structures rather than spatial proximity.

The complete procedure is summarized in Algorithm 1, where the original game-theoretic stopping criterion in [1] is replaced by one based on intra-cluster agreement.

---

**Algorithm 1** GCS-Q for Correlation Clustering

---

**Require:** Weighted graph  $G = (V, E, w)$   
**Ensure:** Clustering  $\mathcal{C} = \{C_1, C_2, \dots, C_k\}$

```

1:  $\mathcal{C} \leftarrow \emptyset$ , Queue  $\leftarrow \{V\}$ 
2: while Queue is not empty do
3:    $S \leftarrow \text{Queue.pop}()$ ,  $G_S \leftarrow \text{subgraph of } S$ 
4:   Solve QUBO:  $\max \sum_{i,j \in S} w_{ij} \cdot \mathbb{I}[x_i = x_j]$ 
5:    $C \leftarrow \{i \in S \mid x_i = 1\}$ ,  $\bar{C} \leftarrow S \setminus C$ 
6:   if  $\text{cut}(C, \bar{C}) \leq 0$  or  $C = \emptyset$  or  $\bar{C} = \emptyset$  then
7:      $\mathcal{C} \leftarrow \mathcal{C} \cup \{S\}$ 
8:   else
9:     Queue  $\leftarrow \text{Queue} \cup \{C, \bar{C}\}$ 
10:  end if
11: end while
12: return  $\mathcal{C}$ 
```

---

#### IV. EXPERIMENTS

To evaluate the effectiveness and robustness of our proposed GCS-Q-based correlation clustering approach, we conduct a comprehensive set of experiments on both synthetic and real-world datasets. The experimental design aims to assess performance relative to classical methods under diverse structural

conditions, including varying cluster size distributions and real-world configurations, using both internal (modularity) and external (Normalized Mutual Information) evaluation metrics where applicable.

##### A. Settings

1) *Datasets:* In this work, we evaluate clustering performance using two types of datasets: synthetically generated signed graphs with known clustering labels and four real-world hyperspectral image datasets.

a) *Synthetic Signed Graph Dataset:* We generated three classes of synthetic signed graphs varying structural and distributional conditions. Each class is represented as a weighted adjacency matrix (Fig. 2), where edge weights denote pairwise correlation coefficients between nodes. The graphs differ in the clarity of their block structure, reflecting the degree of separation and the relative sizes of the underlying clusters. To simulate varying levels of cluster size disparity, we systematically adjusted the Gini index across the three types of datasets. The Gini index, ranging from 0 (perfectly uniform cluster sizes) to 1 (complete inequality), provides a quantitative measure of imbalance in cluster membership. This variation enables a controlled assessment of how different clustering algorithms respond to increasingly skewed distributions. All generated graphs are balanced [24] in the sense that their vertices can be divided into two clusters such that all intra-cluster edges are positive and all inter-cluster edges are negative. Our synthetic data generalizes this notion to multiple clusters, maintaining internal coherence through positive correlations and ensuring negative separation across clusters. This guarantees that clusters are well-separated, allowing for a clean evaluation of clustering performance under ideal yet structurally diverse conditions. The three matrices shown in Fig. 2 are examples of generated graphs with 170 nodes, which represents the practical upper limit for fully connected graphs that can currently be embedded on available quantum annealing hardware. However, it is important to note that assuming a certain degree of sparsity in the graph structure allows QUBO formulations to scale to larger problem instances [25].

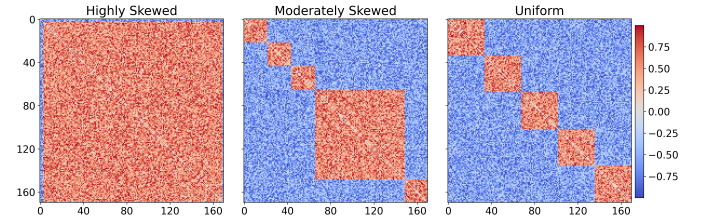


Fig. 2: Correlation matrices of three synthetic signed graphs used in our experiments. The examples correspond to a setting with 5 clusters. Warmer colors indicate strong positive correlations (intra-cluster), while cooler tones represent negative correlations (inter-cluster). The corresponding Gini index and Cluster Size Ratio for each matrix are reported in Table I.



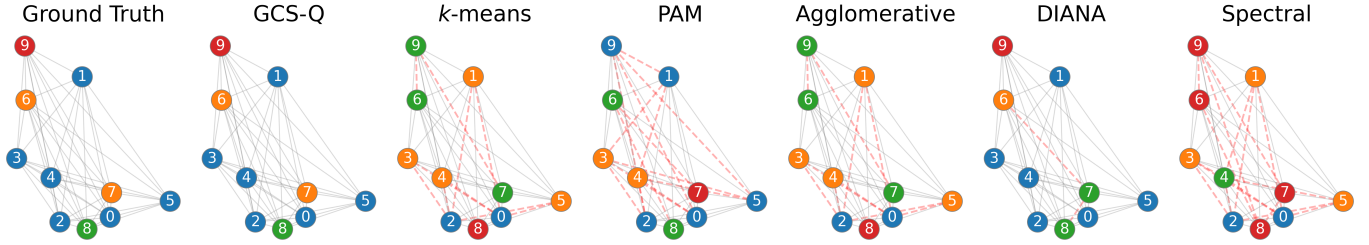


Fig. 1: The ground truth clusters and the outputs of various algorithms for a balanced graph with 10 nodes and edge weights in the range  $[-1, 1]$  representing correlations. Red dashed edges highlight inconsistent assignments: a red edge between nodes in the same cluster indicates a negative correlation, while a red edge between nodes in different clusters indicates a positive correlation. In other words, red dashed lines mark violations of clustering agreement. All other edges are shown in grey, either because they are neutral or consistent with the clustering. These inconsistencies are usually minimized by GCS-Q and are more frequent in classical methods.

Metric	High Skewed	Moderately Skewed	Uniform
Gini Index	0.05	0.70	0.80
Cluster Size Ratio	166.00	3.95	1.00

TABLE I: Gini index and cluster size ratio (largest cluster size divided by smallest) for the three synthetic graphs in Fig. 2. The low Gini index case shows extreme cluster size disparity, with most nodes concentrated in a single cluster.

*b) Hyperspectral Image Dataset:* We also test our methods on real-world data drawn from publicly available hyperspectral remote sensing scenes provided by the Universidad del País Vasco repository<sup>1</sup>. In this experiment, we perform correlation clustering on hyperspectral datasets (Indian Pines, Salinas, Pavia University, and Kennedy Space Center (KSC)) to identify redundant spectral bands [6], [26]. The data consists of hundreds of narrow, continuous spectral bands capturing detailed information across the electromagnetic spectrum. We calculate the Pearson correlation coefficients between bands, forming a correlation matrix, and use this matrix to perform clustering, grouping bands with high correlation to reduce data redundancy for further analysis [18], [27].

2) *Methods:* To evaluate the performance of GCS-Q, we compare several representative clustering methods. *k*-means is a classical distance-based algorithm that partitions data by minimizing intra-cluster variance, assuming a Euclidean embedding. PAM is a more robust alternative to *k*-means that selects actual data points as cluster centers, making it less sensitive to outliers and more effective for non-spherical clusters. Spectral clustering leverages the eigenstructure of a similarity graph to detect globally coherent clusters and is particularly suited for correlation-based tasks. Hierarchical clustering builds a nested tree of clusters based on pairwise similarities; in our study, we employ both agglomerative (bottom-up) and divisive (top-down, DIANA) variants to capture different structural perspectives.

3) *Evaluation Metrics:* To assess the quality of the clustering results, we employ two complementary metrics aiming

at capturing different aspects of clustering performance and providing a more comprehensive evaluation.

*a) Normalized Mutual Information (NMI):* This is an external validation metric that quantifies the similarity between the clustering output and a known ground-truth partition. It measures the mutual dependence between the predicted labels and the true labels. NMI is normalized to lie in the range  $[0, 1]$ , where 0 indicates no mutual information (i.e., independent assignments), and 1 indicates a perfect match between the cluster assignments and the ground truth.

*b) Modularity:* In practical situations, clustering algorithms are often applied to data for which the true cluster labels are unknown. In this context, the modularity metric [28] is essential for assessing the quality of clustering by comparing the observed fraction of intra-cluster edge weights to the expected fraction under a null model of random edge assignment. Given a division of a graph into clusters, modularity is defined as

$$Q = \sum_i (e_{ii} - a_i^2), \quad (4)$$

where  $e_{ii}$  is the fraction of edge weight connecting nodes within cluster  $i$ , and  $a_i$  is the fraction of all edge weights connected to nodes in cluster  $i$ . A higher value of  $Q$  indicates that the clustering captures significant structure in the graph, i.e., nodes are more positively connected within clusters than between them. In correlation clustering, this helps assess whether the clustering correctly groups positively correlated points while separating negatively correlated ones.

## B. Results

Figure 3 illustrates the performance of several clustering algorithms in terms of NMI across different distributional settings—namely, highly skewed, moderately skewed, and uniform distributions—and for varying cluster counts ( $k = 5, 10, 20$ ). Importantly, we provide the correct number of clusters to the classical methods to ensure a fair comparison, even though this information is typically unavailable in real-world scenarios.

Among the evaluated methods, GCS-Q consistently achieves strong performance, maintaining high NMI scores

<sup>1</sup>[https://www.ehu.es/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](https://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes)

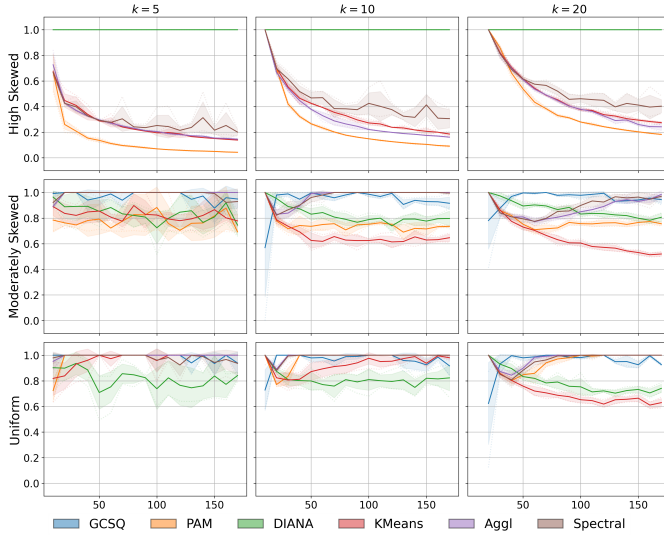


Fig. 3: NMI scores across synthetic signed graphs with increasing cluster size disparity.

across all scenarios. While minor drops are observed in moderately skewed and uniform cases, GCS-Q remains competitive and exhibits remarkable stability compared to classical alternatives. In highly skewed scenarios, DIANA and GCS-Q produce comparable results, which is expected: when  $k - 1$  clusters contain singletons, DIANA’s greedy strategy of isolating one point per iteration aligns well with the underlying structure. However, in all other conditions, GCS-Q clearly outperforms DIANA, owing to its more comprehensive strategy that evaluates all possible bipartitions at each recursive step, as opposed to DIANA’s limited single-point exclusions.

Spectral clustering, though a state-of-the-art method for correlation clustering under balanced conditions, struggles with strong cluster size imbalances. In contrast, only GCS-Q and DIANA exhibit resilience under such distortions. Interestingly, in moderately skewed distributions with small  $k$  (e.g.,  $k = 5, 10$ ), spectral clustering performs well and often matches GCS-Q’s results. However, as  $k$  increases (e.g.,  $k = 20$ ), GCS-Q decisively outperforms all classical baselines, including spectral clustering. Finally, in uniform distributions, GCS-Q continues to perform competitively. Although other methods improve in this more balanced setting, GCS-Q remains one of the most robust performers across all configurations, reinforcing its versatility and generalization capabilities.

Overall, these results confirm that GCS-Q offers a robust and high-performing alternative to classical clustering algorithms, particularly in challenging tasks involving signed graphs with significant cluster size disparity. Importantly, GCS-Q does not require the number of clusters  $k$  to be provided in advance, as it automatically stops when no further bipartition increases intra-cluster agreement. Its ability to maintain high NMI across diverse structural conditions underscores its suitability for practical applications involving complex graph data.

To evaluate the performance of clustering methods in real-world scenarios, we apply correlation clustering to four hyperspectral remote sensing datasets—Indian Pines, Salinas, Pavia University, and KSC—sourced from the Universidad del País Vasco repository. A preliminary analysis was performed to estimate the appropriate number of clusters for the classical methods. Specifically, the spectral gap criterion [29] was used for spectral clustering, while the silhouette score [30] was applied to the other methods.

Figure 4 presents the modularity scores achieved by each method across the four hyperspectral datasets. GCS-Q consistently obtains the highest modularity values, demonstrating its ability to identify strongly coherent spectral clusters. This advantage is particularly pronounced on the KSC, PaviaU and Salinas datasets, which are characterized by higher spectral complexity. Among classical methods, DIANA performs competitively and shows stable performance across all datasets, occasionally approaching GCS-Q’s performance. In contrast, other classical approaches such as PAM,  $k$ -means, and agglomerative hierarchical clustering tend to yield lower modularity scores, indicating less effective partitioning of the spectral correlation structure.

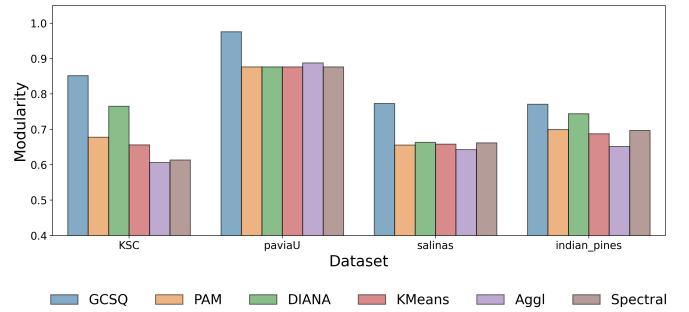


Fig. 4: Modularity scores for clustering hyperspectral bands across four real-world remote sensing datasets. GCS-Q consistently achieves superior modularity, indicating more coherent grouping of highly correlated spectral bands.

## V. CONCLUSION

In this work, we proposed a novel adaptation of GCS-Q, a quantum-assisted solver originally developed for coalition structure generation, to address the problem of correlation clustering on signed graphs. By formulating the clustering task as a maximize-agreement objective over a weighted graph, we leveraged the power of quantum annealing to solve a sequence of QUBO problems that recursively identify meaningful partitions. Our method follows a divisive hierarchical strategy, avoiding the need to predefine the number of clusters and improving global optimization at each step.

Through extensive experiments on both synthetic signed graphs and real hyperspectral datasets, we demonstrated the robustness and superior clustering quality of GCS-Q, especially under challenging conditions such as high cluster size disparity. On synthetic data, GCS-Q achieved competitive NMI scores across varying Gini index levels, demonstrating

superior performance and robustness to skewed cluster size distributions. On real hyperspectral EO data, GCS-Q consistently outperformed classical baselines in terms of modularity, indicating better separation of correlated spectral bands.

These results confirm the viability of hybrid quantum-classical optimization for graph-based optimization, even when using noisy quantum annealing devices such as the D-Wave Advantage employed in this study. This encourages further exploration of quantum-assisted algorithms for unsupervised learning tasks. Future work includes expanding the experimental evaluation of GCS-Q, exploring alternative gate-based quantum solvers [31], [32], and applying the method to moderately sparse, large-scale real-world graph data beyond the remote sensing domain.

## REFERENCES

- [1] S. M. Venkatesh, A. Macaluso, and M. Klusch, "Gcs-q: Quantum graph coalition structure generation," in *International Conference on Computational Science*. Springer, 2023, pp. 138–152.
- [2] N. Bansal, A. Blum, and S. Chawla, "Correlation clustering," *Machine learning*, vol. 56, pp. 89–113, 2004.
- [3] P. Doreian and A. Mrvar, "A partitioning approach to structural balance," *Social networks*, vol. 18, no. 2, pp. 149–168, 1996.
- [4] M. Charikar, V. Guruswami, and A. Wirth, "Clustering with qualitative information," *Journal of Computer and System Sciences*, vol. 71, no. 3, pp. 360–383, 2005.
- [5] A. Bhattacharya and R. K. De, "Divisive correlation clustering algorithm (dcca) for grouping of genes: detecting varying patterns in expression profiles," *bioinformatics*, vol. 24, no. 11, pp. 1359–1366, 2008.
- [6] A. Mehta and O. Dikshit, "Spcas assisted correlation clustering of hyperspectral imagery," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-8, pp. 111–116, 2014. [Online]. Available: <https://isprs-annals.copernicus.org/articles/II-8/111/2014/>
- [7] N. Ailon and E. Liberty, "Correlation clustering revisited: The "true" cost of error minimization problems," in *International Colloquium on Automata, Languages, and Programming*. Springer, 2009, pp. 24–36.
- [8] X. Deng and C. H. Papadimitriou, "On the complexity of cooperative solution concepts," *Mathematics of operations research*, vol. 19, no. 2, pp. 257–266, 1994.
- [9] M. Zitnik, R. Sosič, and J. Leskovec, "Prioritizing network communities," *Nature communications*, vol. 9, no. 1, p. 2544, 2018.
- [10] K. E. Avrachenkov, A. Y. Kondratev, V. V. Mazalov, and D. G. Rubanov, "Network partitioning algorithms as cooperative games," *Computational social networks*, vol. 5, pp. 1–28, 2018.
- [11] E. Elkind and J. Rothe, "Cooperative game theory," *Economics and computation: an introduction to algorithmic game theory, computational social choice, and fair division*, pp. 135–193, 2016.
- [12] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009.
- [13] M. C. Thrun, "Distance-based clustering challenges for unbiased benchmarking studies," *Scientific reports*, vol. 11, no. 1, p. 18988, 2021.
- [14] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 2, no. 1, pp. 86–97, 2012.
- [15] M. Levorato, R. Figueiredo, Y. Frota, and L. Drummond, "Evaluating balancing on social networks through the efficient solution of correlation clustering problems," *EURO Journal on Computational Optimization*, vol. 5, no. 4, pp. 467–498, 2017.
- [16] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, pp. 395–416, 2007.
- [17] L. Ding, C. Li, D. Jin, and S. Ding, "Survey of spectral clustering based on graph theory," *Pattern Recognition*, p. 110366, 2024.
- [18] H. Çukur, H. Binol, F. S. Uslu, Y. Kalaycı, and A. Bal, "Cross correlation based clustering for feature selection in hyperspectral imagery," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, 2015, pp. 232–236.
- [19] F. McSherry, "Spectral partitioning of random graphs," in *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*. IEEE, 2001, pp. 529–537.
- [20] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum algorithms for supervised and unsupervised machine learning," *arXiv preprint arXiv:1307.0411*, 2013.
- [21] E. Aïmeur, G. Brassard, and S. Gambs, "Quantum speed-up for unsupervised learning," *Machine Learning*, vol. 90, pp. 261–287, 2013.
- [22] I. Kerenidis, J. Landman, A. Luongo, and A. Prakash, "q-means: A quantum algorithm for unsupervised machine learning," *Advances in neural information processing systems*, vol. 32, 2019.
- [23] A. Kalra, F. Qureshi, and M. Tisi, "Portfolio asset identification using graph algorithms on a quantum annealer," *Available at SSRN 3333537*, 2018.
- [24] S. Aref and M. C. Wilson, "Balance and frustration in signed networks," *Journal of Complex Networks*, vol. 7, no. 2, pp. 163–189, 2019.
- [25] S. M. Venkatesh, A. Macaluso, M. Nuske, M. Klusch, and A. Dengel, "Q-seg: Quantum annealing-based unsupervised image segmentation," *IEEE Computer Graphics and Applications*, 2024.
- [26] W. Wang, W. Wang, and H. Liu, "Correlation-guided ensemble clustering for hyperspectral band selection," *Remote Sensing*, vol. 14, no. 5, 2022. [Online]. Available: <https://www.mdpi.com/2072-4292/14/5/1156>
- [27] A. Ghorbanian and A. M. and, "An unsupervised feature extraction method based on band correlation clustering for hyperspectral image classification using limited training samples," *Remote Sensing Letters*, vol. 9, no. 10, pp. 982–991, 2018. [Online]. Available: <https://doi.org/10.1080/2150704X.2018.1500723>
- [28] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical review E*, vol. 69, no. 2, p. 026113, 2004.
- [29] L. Zelnik-Manor and P. Perona, "Self-tuning spectral clustering," *Advances in neural information processing systems*, vol. 17, 2004.
- [30] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.
- [31] S. M. Venkatesh, A. Macaluso, and M. Klusch, "Quacs: Variational quantum algorithm for coalition structure generation in induced subgraph games," in *Proceedings of the 20th ACM International Conference on Computing Frontiers*, ser. CF '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 197–200. [Online]. Available: <https://doi.org/10.1145/3587135.3592192>
- [32] S. M. Venkatesh, A. Macaluso, M. Nuske, M. Klusch, and A. Dengel, "Qubit-efficient variational quantum algorithms for image segmentation," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, 2024, pp. 450–456.