GravSpec: Spectral-Gravitational flow on Density Manifolds for Deterministic Overlapping Community Detection

Swetha Balasubramanian[1], Pranab K. Muhuri[2]
Department of Computer Science and Engineering
South Asian University
New Delhi, India
[1]swethab@students.sau.ac.in, [2]pranabmuhuri@cs.sau.ac.in

***Abstract*—The discovery of overlapping community structures in large-scale networks remains a crucial aspect of graph mining. However, the existing methods face a fundamental trade-off between computational efficiency and structural representation. While traditional Density Peak Clustering (DPC) constitutes a robust center selection strategy, it is constrained by $O(N^2)$ complexity due to its pairwise distance calculations and limited topological awareness. In this paper we propose GravSpec, a sub-quadratic framework that reformulates DPC for non-Euclidean graph manifolds. GravSpec comprises two key ideas: 1) A spectral density estimation technique derived from Personalized PageRank and 2) A Jaccard-weighted gravitational flow that bounds the search space to the local adjacency manifold. We demonstrate that GravSpec runs in near-linear time complexity $O(M+NlogN)$ and in $O(M+N)$ space complexity. Results from extensive experiments on LFR benchmarks and massive real-world networks show that GravSpec outperforms its competitors. Notably, in high-noise environments where traditional methods degrade, our approach achieves up to a 34.8% improvement in Overlapping Normalized Mutual Information (ONMI) and superior structural resolution on large-scale graphs. while ensuring scalability for graphs containing millions of edges.**

***Index Terms*—Overlapping Community Detection, Density Peak clustering, Personalised PageRank, Manifold Learning, Harmonic Function on graphs**

I. Introduction

With the rapid growth of massive relational datasets, from biological interaction networks to global social graphs, overlapping community detection has transformed from a peripheral task to a core requirement. Real-world networks are inherently overlapping in nature, as nodes often participate in multiple functional roles. As the network sizes reach the tens of millions, the Scalability-Accuracy Gap makes most existing frameworks impractical.

Community detection is extensively researched. Contemporary approaches fall into four major categories, each with inherent limitations: 1) Partitioning and Spectral-based models focus on disjoint partitioning of the graph. Methods like Louvain [1] and Leiden [2] achieve near-linear modularity optimization but can't detect overlapping memberships of nodes. Spectral methods identify low-conductance cuts using the graph Laplacian but require a predefined community count and have a complexity of $O(N^3)$, limiting their applicability to large-scale graphs. 2) Overlapping-based methods: Label propagation based approaches such as COPRA [3] and SLPA [4] have near-linear time complexity but are unstable and are prone to epidemic community formation where a single label dominates the graph. Clique based methods for example CPM [5] and RaidB [6] can naturally detect overlapping communities, but they are highly sensitive to clique size threshold and are computationally intensive. Local expansion based approaches like OCLN [7] are efficient for finding communities based on seed nodes, but they fail to capture the global structure and are highly dependent on the quality of seeds selected. Fuzzy based approaches such as NeSiFC [8] effectively capture the soft memberships of network nodes, however their scalability is limited due to $O(n^2)$ time complexity. Methods like Bigclam [9] employ negative Matrix Factorization (NMF) to model latent community memberships. Despite their effectiveness, overlapping-based methods are parameter-sensitive and computationally expensive. 3) Gravity-based approaches (GCD) represent nodes as particles with "mass" (usually degree) and apply simulations of Newtonian gravitational forces ($F = G\frac{m_1 m_2}{r^2}$), leading to $O(N^2)$ due to all-pair interactions. GCD-based methods model nodes as celestial bodies, define "mass" based on degree centrality, and use "force" to establish relationships [10]. Since these methods depend on the degree to compute mass, they are affected by high-degree noise. These methods also have high computational complexity. 4) Density-peak-based (DPC) [11] methods have a clear advantage as they model community centers as local density peaks and identify communities based on density variations. Approaches such as DPOCD [12] extend DPC by calculating edge and point-link strengths to quantify the distance between nodes. DPC-based methods are either computationally intensive ($O(N^2)$) or rely on degree centrality, which fails to differentiate between a community core and a high-degree bridge node.

In this context we propose GravSpec, a framework that bridges the gap between scalability and structural resolution by combining the spectral structural awareness and scalable local propagation. The proposed approach is founded on the insight that community assignment can be modelled as Gravitational Flow on a spectral manifold. GravSpec has three

major innovations:

- **Spectral–structural integration:** GravSpec uses spectral density calculated using Personalized PageRank instead of degree-based density. The proposed strategy selects community peaks based on a node's stationary probability, thereby mitigating global structural influence and local connectivity noise.
- **Adjacency-Constrained Manifold:** We overcome the $O(N^2)$ bottleneck by defining Structural Distance δ using Jaccard-weighted adjacency. By confining parent search to the graph's adjacency manifold, we achieve a time complexity of $O(M + N \log N)$.
- **Adaptive Gravitational Propagation:** We propose an automated community seed selection mechanism based on a statistical γ-score ($\rho \times \delta$) followed by a membership flow phase that enables natural soft assignment of node memberships without requiring a predefined number of communities.

The remainder of this paper is organized as follows. Section II introduces the preliminaries, while Section III presents the GravSpec methodology. Section IV reports the experimental evaluation and discusses the results. Finally, Section V concludes the paper.

II. Preliminaries

Let $G = (V, E)$ be an undirected, unweighted graph where $|V|$ = Number of nodes and $|E|$ = Number of edges. Let A be the adjacency matrix and D be the diagonal degree matrix, where $d_{ii} = \sum_j A_{ij}$.

A. Personalized PageRank (PPR)

The PPR vector π_s for a seed node s represents the stationary distribution of a random walk that periodically restarts at s with probability $(1 - \alpha)$. It is the unique solution to $\pi_s = \alpha \pi_s P + (1 - \alpha)\mathbf{e}_s$ where $P = D^{-1}A$ is the transition matrix, α is the Damping factor and $\mathbf{e}_s$ is the indicator vector for node s. In GravSpec, we define the Spectral Density ρ_u as the aggregate importance of node u when the restart vector is uniform ($\mathbf{v} = \mathbf{1}/N$).

B. Jaccard Structural Similarity

To measure the "closeness" of two nodes within the graph topology, we utilize the Jaccard Coefficient. For nodes $u, v \in V$, the similarity $J(u, v)$ is defined as:

$$J(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|} \tag{1}$$

where $\Gamma(u) = \{w \in V \mid (u, w) \in E\} \cup \{u\}$ is the closed neighborhood of u. This metric captures the degree of local connectivity overlap, providing a more robust "distance" metric than simple edge weights.

Algorithm 1: GravSpec (Spectral-Gravitational Clustering)

Input: Graph $G = (V, E)$, Damping factor α, Overlap threshold τ, Sensitivity k
Output: Set of overlapping community assignments $\mathcal{C}$

// Step 1: Spectral Density Estimation
for *each node* $v_i \in V$ **do**
  Initialize local residue $r(v_i) = 1$, all other $r(v) = 0$
  Initialize local PPR $p(v) = 0$
  while $\max(r) > \epsilon$ **do**
    Pick node u where $r(u) > \epsilon$
    $p(u) \leftarrow p(u) + \alpha \cdot r(u)$
    for *each neighbor* v *of* u **do**
      $r(v) \leftarrow r(v) + (1 - \alpha) \cdot r(u)/d(u)$
    end
    $r(u) \leftarrow 0$
  end
  $\rho_i \leftarrow p(v_i)$
  //Density is the Self-Influence score
end
return ρ

// Step 2: Structural Gradient & Distance Calculation
foreach *node* $u \in V$ **do**
  Identify influencers
  $\Gamma^+(u) = \{v \in \text{Adj}(u) \mid \rho_v > \rho_u\}$
  if $\Gamma^+(u) \neq \emptyset$ **then**
    $\delta_u = \min_{v \in \Gamma^+(u)}(1 - \text{Jaccard}(u, v))$
    $\text{Parent}(u) = \text{argmin}_{v \in \Gamma^+(u)}(1 - \text{Jaccard}(u, v))$
  end
  else
    $\delta_u = \max_{v \in \text{Adj}(u)}(1 - \text{Jaccard}(u, v))$
    $\text{Parent}(u) = \text{None}$
  end
end

// Step 3: Automated Seed Selection Compute $\gamma_u = \rho_u \times \delta_u$ for all $u \in V$
Set threshold $T = \mu_\gamma + k\sigma_\gamma$
Define Seeds $S = \{u \in V \mid \gamma_u > T\}$
Assign each $s_i \in S$ a unique identity vector $\mathbf{m}_{s_i}$

// Step 4: Gravitational Membership Flow Sort V in descending order of density ρ
foreach *node* $u \in V \setminus S$ *(in sorted order)* **do**
  $M^u = \sum_{v \in \Gamma^+(u)} \omega_{uv} M^v$ where $\omega_{uv} = \frac{\text{Jaccard}(u,v)}{\sum_{w \in \Gamma^+(u)} \text{Jaccard}(u,w)}$
  if $M^u_c > \tau$ **then**
    Add u to community c
  end
end
return $\mathcal{C}$

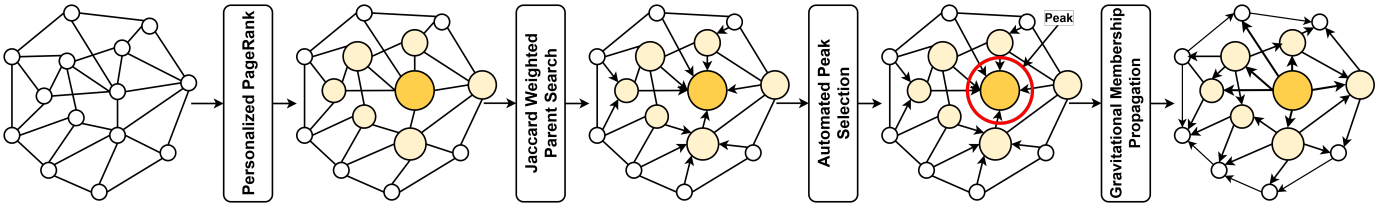

Fig. 1. Flowchart of the proposed algorithm.

C. The DPC Principle

The core of Density Peak Clustering relies on two variables for each node i:

- Local Density (ρ_i): A measure of the node's importance.
- Min-Dist (δ_i): The distance to the nearest node with a higher density:

$$\delta_i = \min_{j:\rho_j > \rho_i} \text{dist}(i,j) \tag{2}$$

Nodes that simultaneously exhibit high values ρ and δ are recognized as the Community Peaks.

III. Proposed Approach

To address the $O(N^2)$ bottleneck of density peak-based methods, GravSpec models the graph as a manifold, in which communities are formed due to a gravitational attraction "pull" toward spectral centers. Fig. 1 illustrates the flow of the proposed algorithm. In this section, we present the phases of the proposed approach, derive its time and space complexity and demonstrate its convergence.

A. Spectral Density Computation

A common limitation of graph-based algorithm is "degree-bias" where high-degree hubs are incorrectly identified as community centers "peaks". GravSpec addresses this limitation by computing Spectral Density ρ. A uniform Personalised Pagerank (PPR) is employed where the restart vector $\mathbf{v}$ is initialized to $1/N$. Node density ρ_u is defined as the probability of a random walker, after an infinite number of jumps and restarts, visiting node u. This ensures that a node's density is measured by its structural centrality within its community rather than by its mere degree. Thus, nodes in the dense core of a community attain a higher spectral score than the nodes in the sparsely connected bridge nodes at the boundaries.

B. Jaccard-Weighted Structural Gradient

Unlike classical DPC, which relies on all-pairs distance computations for calculating the distance δ to a higher-density neighbor, GravSpec uses a structural gradient to confine the search to the adjacency manifold. For each node i, the parent node Parent $P(i)$ is assigned by searching its immediate neighborhood $\Gamma(i)$. The parent is chosen as the neighbor j that offers the optimal topological path towards a higher-density region.

$$\delta_i = \min_{j \in \Gamma(i), \rho_j > \rho_i} (1 - J(i,j)) \tag{3}$$

where $J(i,j)$ is the Jaccard similarity between node i and j, equation 2 explains Jaccard similarity. If a node doesn't have a neighbour with higher density, it is a local peak. The δ of this local peak is initialized to the maximum observed value. This phase restructures the input undirected graph into a Directed Acyclic Graph (DAG), which captures the influence, with every edge pointing "uphill" toward a density peak.

C. Automated Peak Selection using γ-Outliers

In this, we calculate the γ-score of every node, which is defined as $\gamma_i = \rho_i \times \delta_i$. while nodes in the core of a community will have high ρ and nodes in the boundary (bridge nodes) will have high δ, only those nodes which are true peaks of a community will have high ρ and δ. Now we will employ a statistical filter based on the γ-distribution. The nodes with γ-score greater than $T = \mu_\gamma + 2\sigma_\gamma$ are selected as community seeds. In this phase, we truly adaptively calculate the number of communities, thus overcoming the manual filtration limitation of DPC-based methods and by automatically selecting the community seeds, this method eliminates the necessity for user-defined cluster numbers, a well-known drawback of spectral clustering.

D. Gravitational Membership Propagation

In this phase, we enable the membership to flow down the structural gradient, thus enabling soft overlapping memberships. For each seed node s, the membership vector M^s is initialized, where its own community index equals 1.0. All remaining nodes are then processed in descending order of density, where each node u acquires membership values from its influencers (higher-density neighbors), with contributions proportional to the Jaccard similarity between the nodes.

$$M_c^u = \frac{\sum_{v \in \text{Influencers}(u)} J(u,v) \cdot M_c^v}{\sum_{v \in \text{Influencers}(u)} J(u,v)} \tag{4}$$

A node u is assigned to community c if its membership value $M_c^u > \tau$, where τ denotes the overlap threshold. This approach enables nodes at the boundary of two (or more) "gravitational peaks" to naturally belong to both, effectively capturing the network's inherent overlapping structure. Algorithm 1 presents the pseudo-code of the proposed approach.

E. Time and space complexity Analysis

Time complexity: Now we will derive the time complexity of the proposed approach. Let $G = (V, E)$ be a graph with N nodes and M edges. In the first phase, we use the power-iteration method, in which the spectral density ρ converges

in T iterations (where in each iteration, sparse matrix-vector multiplication of the transition matrix P is performed) leading to $O(T{\cdot}M)$. From our observation, for most real-world graphs, $T < 50$ is sufficient for convergence. In the second phase, for each node u, we compute the Jaccard similarity $J(u, v)$ for its immediate neighbors $v \in \Gamma(u)$. In the worst case, the total cost of performing set intersections at all edges is $O(\sum_{u \in V} d_u^2)$. However, for sparse real-world networks, it is effectively $O(M \cdot \bar{d})$, where $\bar{d}$ is the average degree. In the third phase, we sort nodes according to ρ, which costs us $O(N \log N)$. In the fourth phase, we iterate over the nodes and, for each node u, update its membership based on its neighbours; for this, each edge is visited exactly once in the direction of the gradient. This costs us $O(M)$. When we combine the complexity, it sums upto: $O(T \cdot M + M \cdot \bar{d} + N \log N + M)$. Since T and $\bar{d}$ are small for sparse networks, we can concludes that GravSpec identifies overlapping communities in $O(M + N \log N)$ time, making the framework strictly sub-quadratic and suitable for large-scale networks.

Space Complexity: The Adjacency list requires $O(M + N)$. We store the density ρ, distance δ, and parent pointers, which occupy $O(N)$ space. We store the membership matrix using a sparse dictionary, which in the worst case of extreme overlap requires $O(c \cdot N)$, where c is the average number of communities per node (typically $c \ll N$). Thus, the space complexity of GravSpec is $O(M + N)$, making it scalable to graphs with millions of nodes.

F. Convergence and DAG Stability

Consider the induced graph $G_{grad} = (V, E_{grad})$, where a directed edge $(u, v) \in E_{grad}$ exists if and only if $(u, v) \in E$ and $\rho_v > \rho_u$, is a Directed Acyclic Graph (DAG). Let's assume for the sake of contradiction that there exists a cycle $v_1 \to v_2 \to \cdots \to v_k \to v_1$. Then, by definition of the edge directions, $\rho_{v_1} < \rho_{v_2} < \cdots < \rho_{v_k} < \rho_{v_1}$, which is a contradiction, as it implies $\rho_{v_1} < \rho_{v_1}$. Therefore, the graph contains no cycles. Now we prove that the Gravitational Propagation phase is guaranteed to converge in a single pass. Let the nodes be processed in the order $v_1, v_2, \ldots, v_n$ where $\rho(v_i) \geq \rho(v_{i+1})$. For a given node v_i, its membership M^{v_i} is determined solely by its higher-density neighbors. By this ordering, these neighbours are always v_j where $j < i$. Since v_j has been processed already and its membership was finalized M^{v_i} is calculated from stable values. By induction, every node's membership is finalized in one iteration.

G. Robustness of γ-Scoring

We justify the automated peak selection using Chebyshev's Inequality. For any distribution of γ, the probability that a node is a "false peak" decreases as its value deviates further from the mean. By setting the threshold at $\mu_\gamma + 2\sigma_\gamma$, only nodes in the top 5% of the structural core (peak) nodes are selected as community seeds. Thus, it filters out the noise-induced local maxima.

TABLE I
SPECIFIC PARAMETERS FOR LFR BENCHMARK GENERATION

Parameter Description	**Notation**	**Value**
Number of nodes	N	10^3 to 10^6
Mixing parameter	μ	$[0.1, 0.7]$
Fraction of overlapping nodes	on	$10\% - 50\%$
Overlapping memberships	om	$[2, 10]$
Average degree	k	20
Maximum degree	$maxk$	50
Degree exponent	τ_1	2
Community exponent	τ_2	1
Min community size	$minc$	20
Max community size	$maxc$	100

TABLE II
STATISTICS OF REAL-WORLD DATASETS

Dataset	**Acronym**	**Nodes**	**Edges**
Facebook-Ego	FE	4,039	88,234
Autonomous Systems	AS	6,474	13,233
Email-Enron	EE	36,692	183,831
Github-Musae	GM	37,700	289,003

IV. EXPERIMENTS AND RESULTS

In this section we will discuss the experimental setup and analyse the results. We used a diverse set of 4 real and 256 synthetic networks to test the proposed GravSpec against its competitors. We also use a diverse set of baseline and recently proposed algorithms to assess the performance of the proposed approach. Our goal is to test the robustness to topological noise, handling of high-degree overlaps, and scalability to million-node networks.

A. Experimental Setup

The details of the Synthetic networks generated by the LFR benchmark generator [13] and the real datasets used are given in Tables 1 and 2, respectively. We use the ONMI [14] and Omega Index [15] for synthetic data where the ground truth is available and for real-world datasets, we use Extended Modularity [16] and Conductance [17]. We evaluate the performance of GravSpec against a set of state-of-the-art and recently proposed algorithms: COPRA [3], SLPA [4], EdMot [18], OCLN [7], RaidB [6], and CoDeSEG [19].

B. Performance on Synthetic Networks

1) Robustness to Structural Noise: The primary advantage of GravSpec is its resilience to topological noise. As shown in the fig. 2, Gravspec maintains high structural integrity even when most algorithms in the literature fail. At $\mu = 0.7$, where 70% of the edges of a node are outside its community, GravSpec achieves an ONMI of 0.709, outperforming the next best competitor (SLPA, 0.525) by 34.8%. Methods like COPRA and RaidB fail in networks with high entropy because local voting mechanisms become unreliable. GravSpec mitigates this limitation by applying spectral density estimation as a global low-pass filter, thereby stabilising gravitational cores in the presence of topological noise.

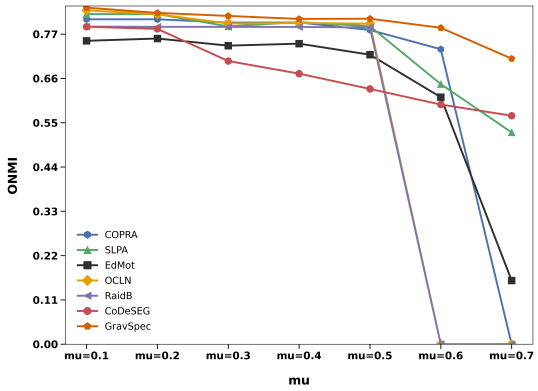

Fig. 2. Plot of the performance of algorithms with increasing noise.

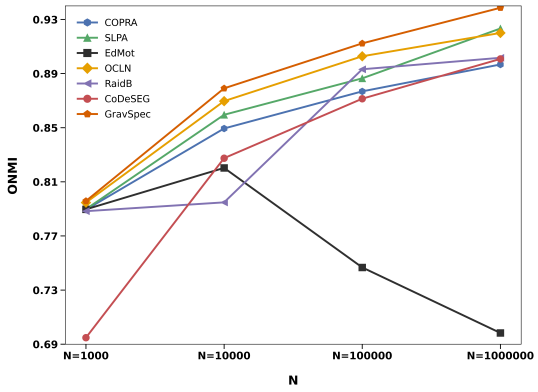

Fig. 3. Plot showing the performance of the algorithms with increasing network size.

2) Scalability across network sizes: As shown in the fig. 3, GravSpec's performance improves with increasing network size. Our $O(M+N\log N)$ framework was tested on networks from $N = 10^3$ to $N = 10^6$. GravSpec exhibits near-linear scaling, maintaining an ONMI > 0.93 and an Omega Index of ≈ 1.0 at the million-node scale, confirming that the sub-quadratic complexity does not come at the cost of accuracy.

3) Handling Deep Overlaps: The fig. 4 shows the performance of GravSpec and its competitors on networks with increasing overlaps. Notably, at om=10, the proposed GravSpec attains an ONMI value of 0.69, outperforming the second-best performer SLPA by 12.2%. Methods such as EdMot rely on edge motifs that become ineffective when node connectivity is spread across numerous communities. Since the proposed approach uses a continuous gravitational flow, allowing nodes to participate in multiple communities, it performs well in networks with deep overlaps.

4) Performance on Omega Index: The performance of the proposed approach and its competitors, evaluated using the omega index, is given in fig. 5. As shown, the proposed approach consistently outperforms its competitors across varying network sizes, densities, overlaps, and complexities.

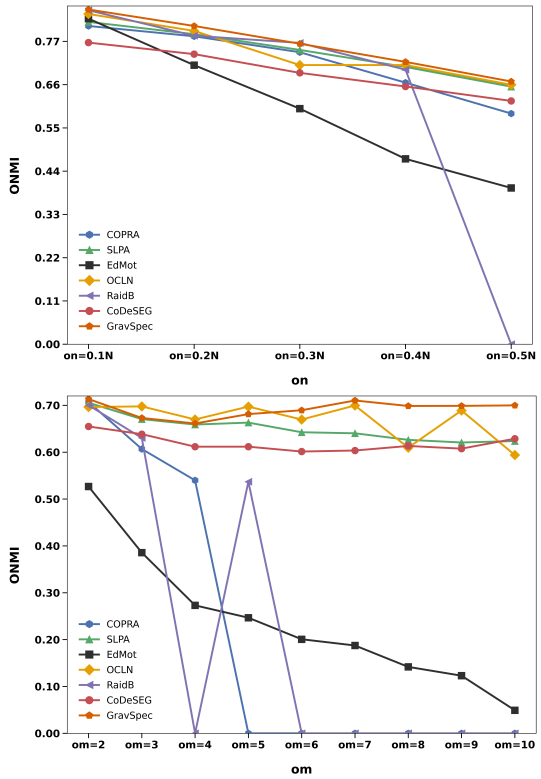

Fig. 4. Plot depicting the performance of algorithms with increasing network overlap.

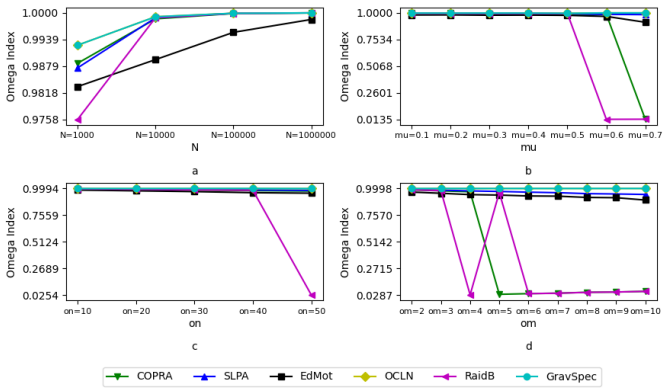

Fig. 5. Comparison of performance of the algorithms with varying network structure evaluated using Omega Index.

C. Performance on real-world networks

Fig.6 shows the performance of the proposed approach compared with competing methods on real-world graphs. The proposed approach consistently achieves high eQ and low conductance across large-scale real datasets. The superior performance of GravSpec can be attributed to three key factors: 1) Denoising using Spectral Density: By leveraging spectral density estimation, GravSpec effectively filters out transient local edges and prioritises nodes that are structurally central to the community manifold, resulting in a 34.8% performance improvement on noisy LFR benchmarks. 2) Structural

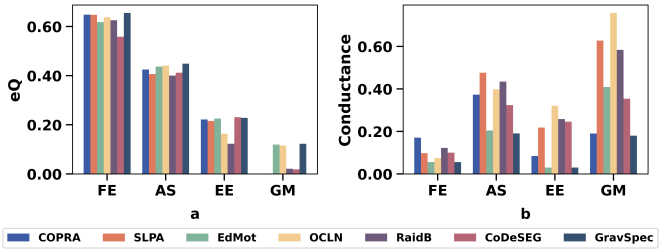

Fig. 6. Performance of the algorithm on large-scale real datasets.

Gradient: unlike label propagation methods, which assign equal influence to all neighbors, GravSpec employs a Jaccard-weighted structural gradient, ensuring that membership "flow" (propagation) is only towards structurally similar nodes and preventing leakage effects seen in OCLN and COPRA. 3) Manifold-Constrained Distance: GravSpec defines distances on the graph manifold, thereby avoiding the quadratic complexity of traditional Density Peak Clustering (DPC) while retaining the ability to find communities of arbitrary shapes and sizes.

V. Conclusion

In this paper, we introduced GravSpec, a new framework for overlapping community detection that achieves spectral-level accuracy with sub-quadratic complexity. By extending the Density Peak Clustering approach for complex networks, GravSpec effectively identifies communities of arbitrary shape and remains robust under topological noise, achieving up to 34.8% higher ONMI than competing methods. With a scalable $O(M + N \log N)$ complexity, GravSpec provides a practical and efficient tool for uncovering latent structures in large-scale networks. While GravSpec performs well, the current implementation is designed solely for static, undirected network structures. As future work, we would like to extend our approach to Dynamic Networks, in which community centres shift over time. A future extension could be incorporating higher-order motifs to stabilize density estimation in sparse graphs and automating the selection of membership thresholds using the eigenvalue spectrum.

References

[1] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, Oct. 2008. [Online]. Available: https://dx.doi.org/10.1088/1742-5468/2008/10/P10008

[2] V. A. Traag, L. Waltman, and N. J. van Eck, "From Louvain to Leiden: guaranteeing well-connected communities," *Scientific Reports*, vol. 9, no. 1, p. 5233, Mar. 2019. [Online]. Available: https://www.nature.com/articles/s41598-019-41695-z

[3] S. Gregory, "Finding overlapping communities in networks by label propagation," *New Journal of Physics*, vol. 12, no. 10, p. 103018, Oct. 2010. [Online]. Available: https://dx.doi.org/10.1088/1367-2630/12/10/103018

[4] J. Xie, B. K. Szymanski, and X. Liu, "SLPA: Uncovering Overlapping Communities in Social Networks via a Speaker-Listener Interaction Dynamic Process," in *2011 IEEE 11th International Conference on Data Mining Workshops*, Dec. 2011, pp. 344–349. [Online]. Available: https://ieeexplore.ieee.org/document/6137400

[5] G. Palla, I. Derenyi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature*, vol. 435, no. 7043, pp. 814–818, Jun. 2005, arXiv:physics/0506133. [Online]. Available: http://arxiv.org/abs/physics/0506133

[6] P. G. Sun, X. Wu, Y. Quan, and Q. Miao, "Rearranging 'indivisible' Blocks for Community Detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 6, pp. 6252–6263, Jun. 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9771068

[7] F. Cheng, C. Wang, X. Zhang, and Y. Yang, "A Local-Neighborhood Information Based Overlapping Community Detection Algorithm for Large-Scale Complex Networks," *IEEE/ACM Transactions on Networking*, vol. 29, no. 2, pp. 543–556, Apr. 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9288733

[8] U. K. Roy, P. K. Muhuri, and S. K. Biswas, "NeSiFC: Neighbors' Similarity-Based Fuzzy Community Detection Using Modified Local Random Walk," *IEEE Transactions on Cybernetics*, vol. 52, no. 10, pp. 10 014–10 026, Oct. 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9464160

[9] J. Yang and J. Leskovec, "Overlapping community detection at scale: a nonnegative matrix factorization approach," in *Proceedings of the sixth ACM international conference on Web search and data mining*, ser. WSDM '13. New York, NY, USA: Association for Computing Machinery, Feb. 2013, pp. 587–596. [Online]. Available: https://dl.acm.org/doi/10.1145/2433396.2433471

[10] J. Gomez, D. Dasgupta, and O. Nasraoui, "A New Gravitational Clustering Algorithm," in *Proceedings of the 2003 SIAM International Conference on Data Mining (SDM)*, ser. Proceedings. Society for Industrial and Applied Mathematics, May 2003, pp. 83–94. [Online]. Available: https://epubs.siam.org/doi/10.1137/1.9781611972733.8

[11] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks | Science," Jun. 2014. [Online]. Available: https://www.science.org/doi/10.1126/science.1242072

[12] L. Sun, T. Ye, J. Sun, X. Duan, and Y. Luo, "Density-Peak-Based Overlapping Community Detection Algorithm," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 4, pp. 1211–1223, Aug. 2022, conference Name: IEEE Transactions on Computational Social Systems. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9597479

[13] A. Lancichinetti and S. Fortunato, "Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities," *Physical Review E*, vol. 80, no. 1, p. 016118, Jul. 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.80.016118

[14] ——, "Normalized mutual information to evaluate overlapping community finding algorithms," https://www.researchgate.net/publication/51944343, 2009, accessed: May 4, 2025.

[15] A. Lancichinetti, S. Fortunato, and J. Kertész, "Detecting the overlapping and hierarchical community structure in complex networks," *New Journal of Physics*, vol. 11, no. 3, p. 033015, Mar. 2009. [Online]. Available: https://dx.doi.org/10.1088/1367-2630/11/3/033015

[16] M. Chen, T. Nguyen, and B. K. Szymanski, "On Measuring the Quality of a Network Community Structure," in *2013 International Conference on Social Computing*, Sep. 2013, pp. 122–127. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6693322

[17] "A Lower Bound for the Smallest Eigenvalue of the Laplacian." [Online]. Available: https://www.degruyterbrill.com/document/doi/10.1515/9781400869312-013/pdf?licenseType=restricted

[18] P.-Z. Li, L. Huang, C.-D. Wang, and J.-H. Lai, "EdMot: An Edge Enhancement Approach for Motif-aware Community Detection," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ser. KDD '19. New York, NY, USA: Association for Computing Machinery, Jul. 2019, pp. 479–487. [Online]. Available: https://doi.org/10.1145/3292500.3330882

[19] Y. Xian, P. Li, H. Peng, Z. Yu, Y. Xiang, and P. S. Yu, "Community detection in large-scale complex networks via structural entropy game," in *Proceedings of the ACM on Web Conference 2025*, 2025, pp. 3930–3941.