GTCD: Gravity-Tension based 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*—Uncovering communities in large-scale networks remains a fundamental challenge due to the tradeoff between computationally expensive global optimization methods and efficient but unstable local heuristic-based methods. In this paper, we propose GTCD (Gravity-Tension based overlapping Community Detection), a deterministic framework that models the community detection problem as a discrete gradient flow over a density-based manifold. By treating community centers as local density peaks or "gravitational anchors", the proposed GTCD moves beyond modularity-based methods, using Topological Tension to determine the membership of nodes. We quantify the divergence of local flow pointers in a node's neighborhood, which provides a geometry-based strategy for multi-community assignment which evades the resolution limit problem faced by global optimization methods. GTCD achieves $O(m + n \log n)$ time complexity, thereby scaling to million-node networks on standard hardware. Extensive experiments performed on LFR benchmarks and diverse large-scale real-world networks demonstrate that GTCD outperforms state-of-the-art community detection methods in both accuracy and quality.**

***Index Terms*—Overlapping Community Detection, Density Peak clustering, Graph Manifold Learning, Discrete Gradient Flow, Topological Data Analysis, Discrete Morse Theory, Scalar Density Fields**

I. Introduction

As data volumes continue to grow exponentially, the need for community detection in large-scale networks has intensified. Identifying these community structures is key to identifying meaningful structure in complex relational data [1]. Overlapping community structure is a fundamental property of complex systems, in which nodes may participate in multiple functional groups, from molecular interaction networks to large-scale social networks [2]. Although the domain of community detection is extensively researched, there still exists a tradeoff between efficient local heuristics and computationally expensive global optimization [3] [4]. The literature in the domain can be primarily classified into four categories: Spectral and Matrix Factorisation-based methods that can truly represent the network topology but suffer from $O(n^2)$ or $O(n^3)$ complexities, thus inapplicable to million-node networks. Modularity-based algorithms are highly efficient, but they are affected by the resolution limit, which merges small, dense communities into a single large community. Label Propagation-based methods that are computationally efficient but non-deterministic and lack a global objective, which leads to unstable community boundaries, making them sensitive to parameters. Density-based methods detect communities by identifying peaks through local maxima. Though these methods were accurate and intuitive, their reliance on all-pairs distances leads to $O(n^2)$ complexity, limiting their applicability to massive-scale networks [5] [6]. In this paper, we introduce GTCD (Gravity-Tension Community Detection), which models a graph as a discrete density manifold in which nodes naturally "flow" toward local maxima. This approach is based on the intuition that communities emerge around density peaks, which are gravitational anchors whose structural influence is exerted on their neighborhoods [7]. Our primary contribution is the way GTCD models overlap. Unlike methods existing in the literature, the proposed GTCD models overlap as Topological Tension. This tension emerges from the nodes at the boundaries between basins of attraction. This is due to the competing gravitational "forces exerted" by the neighbouring nodes. GTCD views overlapping nodes as structural bridges which are being influenced by the competing gravitational centers "peaks". GTCD is a parameter-light, deterministic framework which has a computational complexity of $O(n \log n)$.

This paper is structured as follows: Section II reviews literature related to the proposed approach, Section III introduces preliminaries and definitions of concepts used in our approach. Section IV outlines the methodology of the proposed GTCD framework. Section V evaluates the framework through extensive experiments on synthetic and real-world networks. Section VI concludes the paper.

II. Related Works

The domain of community detection has evolved from basic graph partitioning to detecting complex overlapping structures. We classify the literature related to our framework into three distinct categories as follows:

A. Label Propagation based approaches

Label-propagation-based methods such as COPRA [8], SLPA [9] and LaProFC [10] are computationally efficient but yield variable results across runs. This makes them unsuitable for applications that require stable, reproducible results.

B. Higher-order (motifs) based methods

These methods focus on higher-order Motifs (e.g., triangles, cliques) rather than simple edges [11]. Methods like CPM [2], Edmot [12] and RaidB [13] model communities as entities built on specific sub-structures. While these approaches are effective at detecting communities efficiently, they suffer from a "combinatorial explosion" as the size of the motifs increases. Since the proposed GTCD utilises Local Density Peaks, it can naturally identify areas where dense motifs accumulate, thus, providing a motif-aware result explicitly searching the sub-graph.

C. Density-Peak-Based Methods

The Density-Peak clustering (DPC) [14], proposed by Rodriguez and Laio, was very accurate in detecting the density peaks, but due to its all-pairs distance computation, it has $O(n^2)$ complexity. GTCD replaces this global distance computational overhead with Discrete Gradient Flow, achieving $O(n \log n)$ while maintaining the deterministic strengths of peak-based methods.

GTCD (Gravity-Tension based Overlapping Community Detection) is designed to address the above-mentioned challenges faced by the existing methods in the literature. The GTCD method evades the resolution limit problem faced by Modularity-based methods by using Local Density Peaks. Since the gradient flow is determined by local topology rather than global optimization, this enables GTCD to identify small, high-density overlapping communities. Most overlapping approaches depend on stochastic processes (e.g., Label Propagation) or iterative matrix optimization (e.g., BigClam). These methods are either unstable across different runs or computationally heavy. GTCD introduces the concept of Topological Tension, which models the overlap problem into a deterministic geometric measurement. By quantifying the "conflict" of flow pointers, GTCD provides a reproducible, stable, and instantaneous assignment of overlapping nodes. Since GTCD interprets the minimum distance computation procedure of the DPC-based algorithm as a local gradient problem, it restricts the search for density peaks to the immediate neighborhood, thus effectively reducing the complexity from $O(n^2)$ to $O(m + n \log n)$. This strategy ensures that the effectiveness of DPC is achieved in a much lower time complexity. In today's world of "hyper-scale" graphs, it is crucial for algorithms to not just be accurate but achieve "Hardware-Aware Scalability". GTCD bridges this gap in the literature.

III. Preliminaries

In this section, we will discuss the formal notation and the geometric properties of the graph manifold [15]. Let $G = (V, E, W)$ be a weighted graph where $V = \{v_1, v_2, \ldots, v_n\}$ represents the set of nodes and $E \subseteq V \times V$ the set of edges. The weight function $W : E \to \mathbb{R}^+$ defines the strength of interaction. For unweighted graphs, $W(u, v) = 1$.

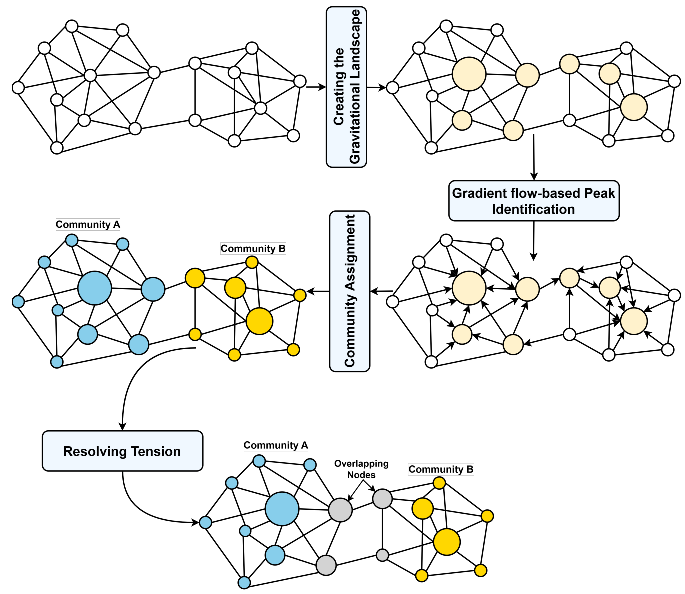

Fig. 1. Flowchart of the proposed algorithm.

A. Graph Representation and Discrete Density

Definition 1 - Discrete Density Field: Let $\rho : V \to \mathbb{R}$ represent a scalar field that represents the topological mass of nodes in the graph manifold. The immediate structural connectivity is used to compute the density of a node v_i:

$$\rho(v_i) = \text{deg}(v_i) + \sum_{v_j \in \mathcal{N}(v_i)} \omega(v_i, v_j) \tag{1}$$

where $\mathcal{N}(v_i)$ is the immediate local neighborhood of node v_i, and $W(v_i, v_j)$ is the edge weight. For unweighted graphs, this naturally simplifies to $\rho(v_i) = 2 \cdot \text{deg}(v_i)$. In our framework, this ρ creates the Gravitational Landscape of the graph [14].

B. Discrete Gradient flow

In Euclidean space, gradient ascent directs points toward density maxima. We extend this principle to the graph domain following the principles of Discrete Morse Theory [16]:

Definition 2 - Steepest Ascent Operator (Flow Operator): The flow operator $\Phi : V \to V$ is a mapping that assigns each node to its "local leader":

$$\Phi(v_i) = \text{argmax}_{v_j \in \{v_i\} \cup \mathcal{N}(v_i)} \rho(v_j) \tag{2}$$

By iteratively applying Φ, A Flow Path

$$\mathcal{S} = \{v, \Phi(v), \Phi(\Phi(v)), \ldots\} \tag{3}$$

is defined. This path will eventually terminate at a fixed point v^* where $\Phi(v^*) = v^*$.These points are the density peaks "gravitational anchors" [14].

Definition 3: Basins of Attraction: All nodes $\{v \in V\}$ such that their ascent paths converge at the same fixed point v^*, form a basin of attraction [17]. In our proposed approach, GTCD, the fixed point where the accent paths converge are the density peak, and the node's respective basin is its primary community core [2].

C. Concept of Gravitational Anchors and Structural Tension

Definition 4: Gravitational Anchor (Topological Peak): A node $v^* \in V$ is a gravitational anchor if and only if it is a fixed point of the steepest ascent operator Φ, i.e. $\Phi(v^*) = v^*$ Let $\Phi^{(k)}(v)$ be the k-fold composition of the flow operator Φ. For any node $v \in V$, there exists a finite t such that $\Phi^{(t)}(v) = v^*$. We formulate the peak assignment function as $P : V \to \mathcal{C}$ (where $\mathcal{C}$ is the set of all anchors) as:

$$P(v) = \lim_{t\to\infty} \Phi^{(t)}(v) \tag{4}$$

This mapping ensures that every node in the graph is assigned to the "catchment area" of exactly one gravitational anchor deterministically.

Definition 5: Topological Tension: In a network with ideal communities, all the neighbours of a node v belong to the same basin of attraction (community). The divergence in the "flow" destinations of node v's neighbours defines the Topological Tension (T). For a node v in the density manifold "located" between multiple gravitational basins, the tension $\mu(v, k)$ for node v arising due to a community core c_k is the fraction of its neighborhood flowing towards that core:

$$\mu(v, k) = \frac{|u \in \mathcal{N}(v) \mid P(u) = c_k|}{|\mathcal{N}(v)|} \tag{5}$$

IV. Proposed Approach

The GTCD framework is designed on the idea that community structure is a fundamental property of a Network. Hence, we allow the latent structures (communities) to reveal themselves instead of enforcing partitions. Fig. 1 explains the steps of GTCD. The algorithm consists of three phases:

A. Creating the Gravitational Landscape (Density Estimation)

In the first phase, we model the flat graph into a mountain range with varying densities. Each node $v \in V$ is considered as a particle of "mass", which is determined by its local connectivity. We then define a Scalar Density Field ρ on the graph using eq. 1. The density of every node v_i is an indicator of its local influence. In this density field, "high altitude" indicates that it corresponds to the core "peak" of a community, while the nodes in the border (sparse boundaries) between them are located in the "valleys". This representation of Local density ensures that even small communities are identified, which are often merged into larger ones: termed as the resolution limit [18].

B. The Gradient flow (Peak Identification)

After constructing the density landscape, we induce motion within the "nodes". Each node "climbs" towards its local peak by tracing the "steepest" path dictated by the density field, i.e. Nodes converge to their local maxima by following the steepest density gradient. This step produces a directed acyclic graph, where each node points to its neighbour with the "steepest" density variation. We define this using a Flow Operator Φ for every node v_i as given in eq. 2. In density "plateaus" where a node has two or more neighbors with exact same maximal density, the tie is strictly broken by selecting the neighbour with the lowest unique lexicographical Node ID. This ensures a deterministic gradient ascent over the graph manifold.

Each node is assigned a pointer to its denser neighbor, forming a directed forest of paths. Every path terminates at a topological peak ($\Phi(v^*) = v^*$), which is its own denser neighbor. This node is the peak that anchors the community. This naturally partitions the network into basins of attraction, with each node assigned to a unique core.

The strictly localized, non-diffusive nature of the discrete gradient flow enables the proposed approach to avoid the risk of "hub-dominance" where high-degree nodes exert an artificial global pull. Since the flow operator Φ restricts the movement to a node's immediate neighborhood, the gravitational pull

Algorithm 1: GTCD: Gravity-Tension Community Detection

```
Input: Graph G = (V, E), Weight function W,
       Tension threshold ε
Output: Set of overlapping communities 𝒞

// Phase I: Gravitational Field Construction
for each node v ∈ V do
    ρ(v) ← deg(v) + Σ_{u∈𝒩(v)} ω(u, v);
end

// Phase II: Discrete Gradient Flow
for each node v ∈ V do
    Parent(v) ← argmax_{u∈{v}∪𝒩(v)} ρ(u);
end
// Path Compression to find Gravitational Anchors
//(Peaks)
for each node v ∈ V do
    Peak(v) ← FindRoot(Parent, v);
end

// Phase III: Resolving Structural Tension
Initialize 𝒞 as an empty mapping;
for each node v ∈ V do
    Counts ← Frequency table of {Peak(u) | u ∈
     𝒩(v)};
    D ← |𝒩(v)|;
    for each peak Ω_k in Counts do
        // Calculate localized membership (Tension)
         μ(v, k) ← Counts[Ω_k]/D;
        if μ(v, k) ≥ ε then
            Add v to Community C_k;
        end
    end
    if v has not been assigned to any C_k then
        Add v to C_Peak(v);
    end
end
return 𝒞;
```

exerted by macro-hubs on distant nodes is zero. Further if a dense small community is separated from a larger hub by a structural "valley" (a bridging region of low-density nodes), the "flow" can not cross the valley, thus causing it to converge at the local density peak. Hence, the proposed GTCD preserves the small communities and evades the resolution limit [18].

C. Resolving Structural Tension (Overlap Detection)

In this final phase, we resolve the "tension" in the boundary region. Most nodes in the community (interior nodes) flow towards a single peak, whereas nodes on the community's borders (in the basins) experience Topological Tension. Imagine a node v whose neighbors are divided: half point toward Peak A, and half toward Peak B. This node is caught in a structural tug-of-war. We quantify this as the Membership Vector μ:

$$\mu(v,k) = \frac{|\{u \in \mathcal{N}(v) : \text{Peak}(u) = \Omega_k\}|}{|\mathcal{N}(v)|} \quad (6)$$

Where Ω_k is the k-th density peak. The proposed approach is designed to embrace this tension. If the influence of a peak k on node v exceeds a threshold ϵ, the node is assigned to that community. This overlap is a measurable geometric state where a node bridges two gravitational fields. This allows GTCD to identify multi-community membership with $O(1)$ lookup time after the initial flow.

D. Time and Space complexity Analysis

In this section, we will derive the time and space complexity of the proposed GTCD approach, proving its efficiency for large scale community detection.

Time Complexity: In phase I, for the Density Field Construction, we calculated the scalar density ρ, which requires a single pass over the edge list to aggregate neighbor weights, which costs $O(m)$ where $m = |E|$. In sparse real-world networks where $m \ll n^2$, this is strictly linear with respect to the number of edges. Then, in phase II, where Gradient Flow and Path Compression are computed, we search for a local maximum in each node's immediate neighborhood. Finding the "leader" for all nodes takes $O(m)$ total. The subsequent path compression (to map every node to its ultimate peak) uses a Union-Find-inspired logic. For a graph with a maximum tree depth of h, the complexity is $O(n \cdot \alpha(n))$, where α is the inverse Ackermann function, which is effectively constant for all practical purposes. In the final phase, overlapping nodes are assigned based on boundary tension. The tension calculation requires checking the peak assignments of a node's neighbors. Since each edge is visited exactly twice, the complexity is $O(m)$. Summing these phases we obtain the final time complexity of our proposed GTCD: $O(m+n\log n)$. In sparse real-world networks where $m \approx n$, GTCD achieves $O(n \log n)$ time complexity. This proves that GTCD has near-linear time complexity.

Space Complexity: For Adjacency representation, we use a Compressed Sparse Row (CSR) format, requiring $O(n+m)$ space. The arrays for density (ρ), flow pointers (P), and peak assignments (L) each need $O(n)$ space. The total space complexity is $O(n+m)$. This means that even for a graph with 10^6 nodes and 10^7 edges, GTCD requires approximately 120 MB of RAM for its core structure

TABLE I
SPECIFIC PARAMETERS FOR LFR BENCHMARK GENERATION

Parameter Description	Notation	Value
Number of nodes	N	10^3 to 10^7
Mixing parameter	μ	$[0.1, 0.8]$
Fraction of overlapping nodes	on	$10\% - 50\%$
Overlapping memberships	om	$[2, 8]$
Average degree	k	15
Maximum degree	$maxk$	50
Degree exponent	τ_1	2
Community exponent	τ_2	1
Min community size	$minc$	$[10, 50]$
Max community size	$maxc$	$[50, 500]$

TABLE II
STATISTICS OF REAL-WORLD DATASETS

Dataset	Acronym	Nodes	Edges
Twitch Gamers	TG	1,68,114	6,797,557
DBLP	DB	3,17,080	1,049,866
Amazon	AM	3,34,863	9,25,872
Hyves	HY	14,02,611	27,77,419

V. EXPERIMENTAL RESULTS AND ANALYSIS

We have conducted extensive experiments on the proposed GTCD across a comprehensive suite of synthetic and real-world benchmarks, comparing its results with a diverse set of state-of-the-art algorithms from the literature. The results are as follows:

A. Experimental Setup

Datasets: For synthetic datasets we used the LFR benchmarks graphs [19] to evaluate the performance of the proposed approach and its competitors. The parameters used to generate the graphs are given in Table I. The details of the large-scale real-world datasets used for validation are given in Table II.

Baselines: We compare the proposed GTCD against a diverse set of five high-performance algorithms. As already discussed COPRA [8] and SLPA [9] are two standard label propagation based approaches. EdMot [12] and RaidB [13] exploit higher order motifs to detect communities while OCLN [20] is a local expansion-based algorithm.

Metrics: Diverse metrics were employed to validate the results of the community detection algorithms. ONMI [21] was used to evaluate the accuracy of the methods while eQ [22] and Conductance [23] were employed to assess the quality of communities detected.

B. Results on LFR networks

1) Robustness to network noise: As shown in Fig. 2, GTCD remains the most stable performer for μ up to 0.5. Even at $\mu = 0.8$ (extreme noise), it maintains a significantly higher ONMI than EdMot and OCLN. This robustness stems from the Deterministic Path Compression. LPA-based methods such as COPRA or SLPA are easily "distracted" by noise

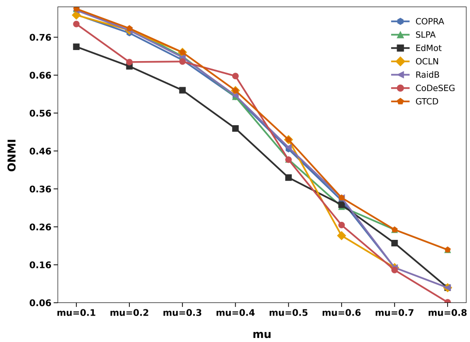

Fig. 2. Plot of the performance of algorithms with increasing noise.

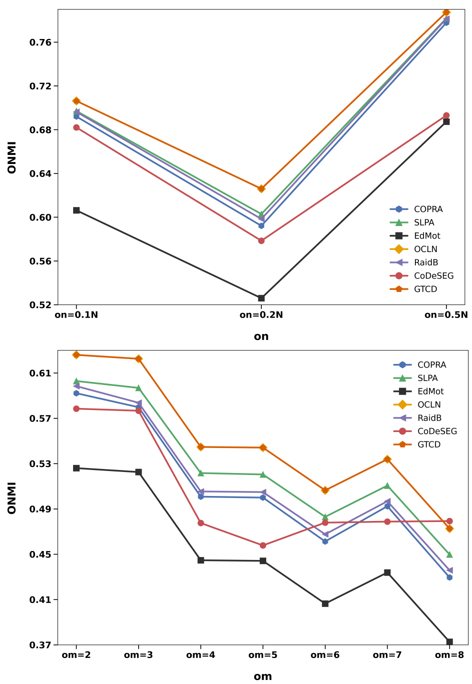

Fig. 3. Plot depicting the performance of algorithms with increasing network overlap.

edges, leading to label flips between communities. Whereas in GTCD, a noise edge only affects the flow if it points to a neighbor with higher density. Since community cores (peaks) inherently have the highest degree and weight density, noise edges in the periphery typically flow back toward the core, effectively self-correcting the community assignment.

2) Handling High-Overlap: GTCD consistently performs better as the number of memberships per node increases. Even at om=8, where a node can belong to 8 communities, the proposed approach outperforms its competitors as shown in Fig. 3. With more nodes participating in overlaps, the proposed GTCD performs better even at 50% overlap. This is the direct result of the Topological Tension metric. Most overlapping algorithms try to "spread" a node's membership across multiple centres using probabilistic weights, which dilute the signal. GTCD's Tension identifies overlap as a structural conflict.

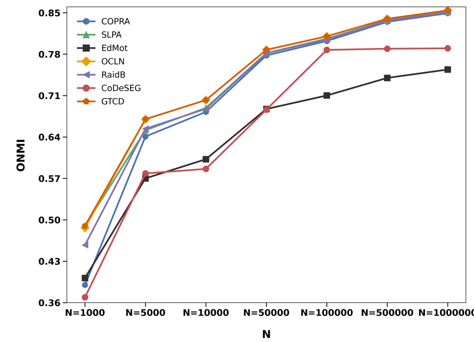

Fig. 4. Plot showing the performance of the algorithms with increasing network size.

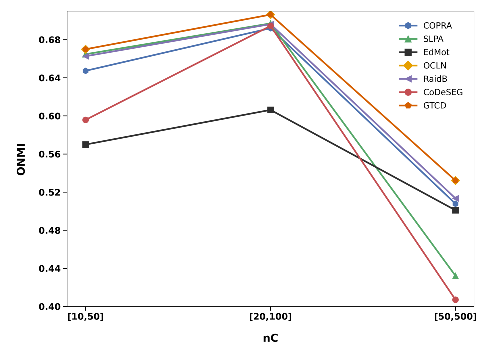

Fig. 5. Comparison of performance of the algorithms with varying number of communities evaluated using ONMI.

Even if a node is pulled by 8 different peaks, GTCD identifies these as 8 distinct "vectors of tension", thus enabling high accuracy at extreme networks with ($o_m = 8$) and ($o_n = 0.5N$).

3) Scalability and Network Size: As can be inferred from fig. 4, the ONMI score of GTCD increases with an increase in network size. It consistently outperforms its competitors even at $N = 10^6$. This performance is due to the Density Field Concentration. In large-scale networks, the local density peaks become more statistically significant and distinct. GTCD's Local Gradient Flow is unaffected by global network volume, it treats a million-node graph as a collection of local manifolds, ensuring that the "gravitational pull" of a community core remains sharp regardless of the total number of nodes, leading to stable results irrespective of the scale of the datasets.

As shown in fig. 5, the proposed approach consistently outperforms its competitors across network structures and community numbers.

C. Results on large-scale real datasets

We can see in fig. 6 that GTCD achieved the highest eQ on the large-scale real datasets. This is because GTCD can accurately identify high-density cores and, by anchoring nodes to local density peaks, ensures that communities are built around them. Further, GTCD includes nodes in a community only if they are within its gravitational basin, resulting in more internally cohesive clusters. The proposed approach achieved a low conductance across datasets, indicating that the communi-

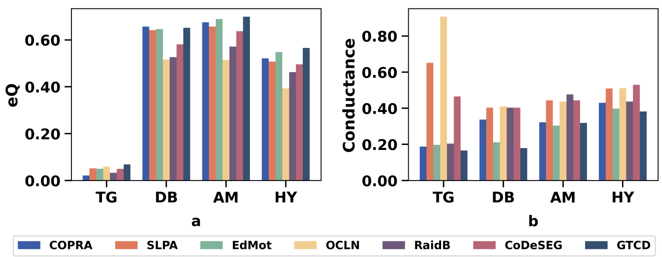

Fig. 6. Performance of the algorithm on large-scale real datasets.

ties detected by GTCD are well separated from the rest of the network. This is because the Gradient Flow naturally stops at the boundaries (valleys) where density drops. It sharply separates communities at the point where the topological pull switches direction, leading to the remarkably clean boundaries reflected in these conductance scores

VI. Conclusion

In this paper, we propose GTCD (Gravity-Tension Community Detection), which models community detection as a discrete gradient flow on a graph manifold. To bridge the accuracy-scalability tradeoff, the proposed GTCD provides a principled approach to uncovering latent structures in a network in near-linear time. Our experimental results across million-node synthetic benchmarks and diverse large-scale real-world datasets show that the "Gravitational" modelling of community and anchoring communities to local density peaks not only achieves superior accuracy in detecting small, overlapping clusters but also ensures stability. The introduction of Topological Tension provides a robust, parameter-light mechanism for detecting overlapping membership. Thus, it converts the problem of overlap detection into a measurable geometric conflict. As data volumes continue to grow exponentially in scale and complexity, the need for algorithms that are both accurate and efficient becomes a necessity. GTCD bridges this gap, offering neural-level precision at an algorithmic cost.

References

[1] M. E. J. Newman, "Detecting community structure in networks," *The European Physical Journal B*, vol. 38, no. 2, pp. 321–330, Mar. 2004. [Online]. Available: https://doi.org/10.1140/epjb/e2004-00124-y

[2] 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

[3] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3, pp. 75–174, Feb. 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370157309002841

[4] M. A. Javed, M. S. Younis, S. Latif, J. Qadir, and A. Baig, "Community detection in networks: A multidisciplinary review," *Journal of Network and Computer Applications*, vol. 108, pp. 87–111, Apr. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804518300560

[5] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics Reports*, vol. 659, pp. 1–44, Nov. 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0370157316302964

[6] D. Jin, Z. Yu, P. Jiao, S. Pan, D. He, J. Wu, P. S. Yu, and W. Zhang, "A Survey of Community Detection Approaches: From Statistical Modeling to Deep Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 1149–1170, Feb. 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9511798

[7] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, Jul. 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1063520306000546

[8] 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

[9] 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

[10] U. K. Roy and P. K. Muhuri, "Label propagation-based membership degree computation towards a computationally efficient fuzzy community detection approach," *IEEE Transactions on Knowledge and Data Engineering*, 2025.

[11] "Higher-order organization of complex networks," *Science*. [Online]. Available: https://www.science.org/doi/10.1126/science.aad9029

[12] 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

[13] 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

[14] 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

[15] M. Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, Jun. 2003. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6789755

[16] R. Forman, "A User's Guide To Discrete Morse Theory," *Sém. Lothar. Combin.*, vol. 48, Dec. 2001.

[17] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 3rd ed. New York: Chapman and Hall/CRC, Jan. 2024.

[18] S. Fortunato and M. Barthélemy, "Resolution limit in community detection," *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 36–41, Jan. 2007. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0605965104

[19] 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

[20] 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

[21] A. Lancichinetti and S. Fortunato, "Normalized mutual information to evaluate overlapping community finding algorithms," https://www.researchgate.net/publication/51944343, 2009, accessed: May 4, 2025.

[22] 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

[23] Z. Lu, J. Wahlström, and A. Nehorai, "Community Detection in Complex Networks via Clique Conductance," *Scientific Reports*, vol. 8, no. 1, p. 5982, Apr. 2018. [Online]. Available: https://www.nature.com/articles/s41598-018-23932-z