

Packing Compact Subgraphs with Applications to Districting

Ho-Lin Chen 

National Taiwan University, Taipei, Taiwan

Po-Yu Chou 

National Taiwan University, Taipei, Taiwan

Prathamesh Dharangutte 

Rutgers University, Piscataway, NJ, USA

Jie Gao 

Rutgers University, Piscataway, NJ, USA

Shang-En Huang 

National Taiwan University, Taipei, Taiwan

Fang-Yi Yu 

George Mason University, Fairfax, VA, USA

Abstract

Packing disjoint subgraphs in a given graph is a fundamental problem with many applications. Motivated by political districting, we focus on connected subgraphs that are compact (e.g., having constant radius from a single center vertex) and that satisfy additional composition requirements, such as a minimum population/weight threshold or balanced weight types (e.g., political affiliations). We aim to maximize coverage by disjoint districts that meet these requirements.

In this work, we present new results that substantially improve the previously known bounds on balanced star districts for planar and minor-free graphs [33]. In particular, we improve the approximation factor from $O(\log n)$ to $O(1)$ for packing balanced star districts using the exact same algorithm, but with a refined analysis. We also extend the results beyond planar graphs to minor-free graphs and an even broader family of graphs of bounded expansion. Additionally, we obtain an $O(1)$ approximation for packing radius- k districts (with a constant k) in planar and apex-minor-free graphs. In order to get a $(1 + \varepsilon)$ approximation on the max coverage, we show that this can be achieved if we allow a slight relaxation of the balancedness parameters (by a factor that can be made arbitrarily close to 1), for bounded radius- k districts on planar and apex-minor-free graphs.

We show that all of these results can also be obtained if we enforce a minimum weight threshold for each district as the composition requirement, rather than balancedness. We present various results on hardness and hardness of approximation for this variant, by graph and district types.

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis; Theory of computation → Graph algorithms analysis

Keywords and phrases Approximation algorithms, algorithmic fairness

Digital Object Identifier 10.4230/LIPIcs.FORC.2026.10

Related Version *Full Version:* <https://arxiv.org/abs/2604.09522>

Funding *Ho-Lin Chen:* Supported by NSTC grant 113-2221-E-002-204-MY3.

Po-Yu Chou: Supported by NSTC grant 114-2222-E-002-004-MY3.

Prathamesh Dharangutte: Supported by NSF grants IIS-2229876, DMS-2220271, CNS2515159, DMS-2311064, and CCF-2118953.

Jie Gao: Supported by NSF grants IIS-2229876, DMS-2220271, CNS2515159, DMS-2311064, and CCF-2118953.

Shang-En Huang: Supported by NSTC grant 114-2222-E-002-004-MY3.



© Ho-Lin Chen, Po-Yu Chou, Prathamesh Dharangutte, Jie Gao, Shang-En Huang, and Fang-Yi Yu; licensed under Creative Commons License CC-BY 4.0

7th Symposium on Foundations of Responsible Computing (FORC 2026).

Editor: Huijia (Rachel) Lin; Article No. 10; pp. 10:1–10:25

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

We consider a general family of districting problems, formulated as packing compact, connected subgraphs, with the objective of maximizing coverage. Given a graph $G = (V, E)$ with vertex set V of $|V| = n$ and edge set E . Each vertex v represents an atomic region with integer *weight* w (e.g., the population counts in the region, or reward). A district is a *connected* subgraph on a vertex set $S \subseteq V$ in G , and two districts must be *vertex-disjoint*. We aim to maximize the total weight of the vertices covered by districts.

Depending on the application scenarios, a district needs to meet additional requirements. One commonly seen requirement is that any two vertices in the same district are not very far, that is, the district is relatively *compact*, or simply not too elongated. This is a desirable property that often appears in discussions of political districting. Further, there may be additional requirements on the population composition or budgeting constraints. We discuss the two dimensions separately.

Compactness. To characterize this requirement, we consider two possible ways of limiting distances between vertices in district S : *diameter- k* or *radius- k* , for a constant k . A district S meets the *diameter- k* constraint if every two vertices in S are within k hops from each other. A district S has *radius- k* if there is a center vertex $v \in S$ and all other vertices of S are within k hops from v . Further, we can consider the diameter or radius in either the *strong* notion – within the induced subgraph on S , or the *weak* notion – in the original graph G . Notice that the weak notion is hereditary: for a subgraph G' with weak diameter k , any subgraph of G' continues to have weak diameter at most k . But this is obviously not always true for the strong diameter or radius.

Composition. In addition, the districts may also need to satisfy additional properties that involve the composition of the vertices, such as:

- *Balancedness*: in political districting there are two types of weights $w_1(v)$ and $w_2(v)$ for each vertex v , with $w(v) = w_1(v) + w_2(v)$. For any vertex subset S and each $i \in \{1, 2\}$ we define $w_i(S) = \sum_{v \in S} w_i(v)$. We call S *c-balanced* if $\min(w_1(S), w_2(S)) \geq (w_1(S) + w_2(S))/c$. That is, the two types of weights are comparable in size.
- *Weight threshold*: a district is valid if $w(S) \geq B$ for a threshold B . This is a natural property that requires all districts to be significant in weight/size. Notice that we consider a lower bound type of constraint. The upper bound version is often trivial – for example, taking all feasible singleton districts is the optimal solution.

1.1 Motivation

Packing disjoint subgraphs is a fundamental problem that has many applications. The most notable applications are in political districting, in which connectivity, contiguity and compactness is highly preferred and sometimes enforced by state law [4]. Furthermore, the *c*-balanced property guarantees that each political group has a decent amount of representation, which can be viewed as compliance with the Voting Rights Act of 1965 [41]. The weight threshold is another natural property such that each district is significant in size/population. In general, the districting problem can also be applied to the allocation of other types of resources and services [13, 51].

Earlier work of [33] considered the *c*-balanced districting problem, which is one case of the districting problem defined earlier. [33] provided various hardness results and approximation algorithms for different families of graphs. There are two major directions left open in [33].

First, most of the positive results (without additional assumptions on the districts or on very special graphs such as complete graphs or trees) are only for *star* districts, which are (strong) radius-1 subgraphs. It was not clear how the results could be extended to higher, constant values of radius- k . Second, and more importantly, one of the most prominent applications in political districting considers planar graphs (e.g., the geographical subdivision into townships and counties). The best approximation factor for packing c -balanced star districts in a planar graph in [33] is $O(\log n)$ without a matching lower bound. Resolving this issue by either developing a constant-factor approximation or proving a stronger hardness result is a major open question.

1.2 Related Work

Our problem is connected to computational (re)districting for schools and elections, which dates back to the 1960s [45]. Since then, extensive work (see [9]) has studied them as an algorithmic problem, typically focusing on specific constraints such as balance or compactness. For balancedness, recent work introduces vote-band metrics [27], which require a certain fraction of votes to fall within a specified range (e.g., 45-55%) for competitive elections. Subsequently, [19, 33] introduce the c -balanced district problem. Other research focuses on optimizing balance scores directly [40]. Regarding compactness, one line of work treats it as a transportation cost [2, 38, 22], such notions also relate to the fair clustering problem [12, 18, 49, 17]. Other research focuses on optimizing compactness scores [8, 50, 53, 48] or using Voronoi or power diagrams with some variant of k -means [67, 24, 25, 39]. Besides the optimization approach, another popular method uses sampling to generate a distribution over districts and create a collection of district plans for selection [5, 21, 28, 59, 16].

Finally, several papers take a fair division approach [54, 64, 26]. The problem is quite different, however, as fairness in this context is defined concerning parties (types) and the number of seats they would win (i.e., the number of districts where they would have a majority) compared to other districts, rather than the internal composition of individual districts.

Besides its applications in political districting, packing compact subgraphs from a given family, such as paths, cycles, and stars, is a fundamental problem in its own right. Let $\mathcal{H}$ be a family of graphs. A $\mathcal{H}$ -packing of a graph G is a set of vertex disjoint subgraphs such that each subgraph is isomorphic to some element of $\mathcal{H}$. These types of problems have been studied in many papers with various hardness and approximation results, such as packing stars [63, 6, 56, 68, 66], induced stars [52], constant sized graphs [57, 14, 55], cycles and

■ **Table 1** A summary of new approximation results on the threshold and balanced districting problem, $\delta, \varepsilon > 0$ are constants. New results are highlighted in bold.

Graph Type	District Type	Results
Planar	star (strong radius-1)	$O(\log n)$ -approx [33]
	strong radius-k	$O(1)$-approx (Theorem 4)
	weak/strong radius-k	$(1 + \varepsilon)$-approx, δ-relaxation (Theorem 22)
Apex-Minor-Free	strong radius-k	$O(1)$-approx (Theorem 4)
	weak/strong radius-k	$(1 + \varepsilon)$-approx, δ-relaxation (Theorem 22)
H -Minor-Free	star (strong radius-1)	$O(h^2 \log n)$ -approx, $h = H $ [33]
	star (strong radius-1)	$O(1)$-approx (Theorem 5)
Bounded Expansion	star (strong radius-1)	$O(1)$-approx (Theorem 5)

paths [44, 47, 66, 10], cliques and bicliques [46, 42, 43], and low-diameter spanning trees [20]. Another line of work [1, 36, 35, 34, 58, 69] explores covering and partitioning graph with k -clubs, as a generalization to the clique cover problem. $S \subseteq V$ is a k -club in graph G if S induces a subgraph in G with diameter at most k (cliques are k -club for $k = 1$). k -club with at least t vertices is called a (t, k) -club. The Minimum k -club cover problem asks for minimum number of k -clubs that cover all vertices in the graph and [35, 69] explore hardness and approximation algorithms for different values of k . The Maximum Disjoint (t, k) -club covering problem asks for vertex disjoint (t, k) -clubs that cover maximum number of vertices in the graph G . [36, 58] study this problem for $k = 2, 3$ on general and bipartite graphs. Note that none of the above work consider the composition properties of the subgraphs. Thus, the results are only marginally relevant to this paper.

1.3 New Results

In this paper, we resolve both problems left open in [33], significantly expanding the theoretical understanding of this problem. Our results are summarized in Table 1.

$O(1)$ -approximation. Compared to the $O(\log n)$ -approximation of packing star districts on a planar graph in [33], we improved the results in two directions. First, we show an algorithm (Theorem 5) with a *constant* approximation factor for packing c -balanced *star* districts in graphs of *bounded expansion*, which is a concept of graph family that are “everywhere sparse” introduced by Nešetřil and Ossona de Mendez [62, 61, 60]. Graphs with bounded expansion include minor-free graphs (and planar graphs) and graphs of bounded degree, as well as graphs that appear in complex network models such as the configuration model and the Chung–Lu model [31]. This improves the approximation factor from $O(\log n)$ to $O(1)$ for a much bigger family of graphs. Second, we show a *constant* approximation for packing c -balanced or threshold districts of *bounded strong radius* on *planar graphs and apex-minor-free* graphs (Theorem 4). Both results also apply to the threshold districts (instead of c -balanced ones for the composition requirement).

$(1+\varepsilon)$ -approximation. Next, we wonder what can be said if we want a $(1+\varepsilon)$ approximation for the district packing problem. We show that PTAS can be obtained for packing *bounded weak radius* districts on *planar and apex-minor-free graphs*, if we allow a small relaxation of c -balanced property to $c(1+\delta)$ -balanced, for any $\delta > 0$ (Theorem 22). The same holds for the weight-threshold version if we relax the threshold B to $B(1-\delta)$.

Hardness results. Finally, we complement the positive results with hardness results in Appendix C. For the majority of the positive results, we consider compactness as defined by a bounded radius k . For maximum packing of subgraphs of bounded *diameter* (both weak and strong) on general graphs, we show that even finding *one* district with weight threshold is NP-hard (Theorem 6.6 in full version) by reducing it from the max clique problem. Additionally, the separation oracle for the LP formulation relies on finding a violating district, which itself is a feasible district; implying that implementing the separation oracle for districts of bounded diameter in general graphs is NP-hard. For maximum packing of districts with a weight threshold, we show that it is NP-hard by reduction from the maximum independent set problem (Theorem 25). Specifically, we show it is NP-hard to approximate to a factor better than $n^{1/2-\delta}$ for $\delta > 0$ on general graphs, and the maximum packing problem remains NP-hard even on planar graphs. This hardness holds even when we require the subgraphs to have bounded radius, with $k = 1$ (stars). The problem is trivial for a complete graph but remains NP-hard on a tree by a reduction from the Knapsack problem.

Lastly, the c -balanced property considers two types of weights; if three or more weights are considered, the problem of packing compact graphs is hard to approximate within any factor (see Theorem 6.5 in full version). We will now formally define the problems and then provide an overview of the challenges and new techniques to obtain the results.

2 Preliminary

2.1 Problem Statement

Consider a connected undirected graph $G = (V, E)$ with $n = |V|$ vertices and $m = |E|$ edges. We define a general district packing problem. The input graph is associated with an *objective weight* function $w : V \rightarrow \mathbb{Z}_{\geq 0}$ and potentially a list of *feature weight* functions $w_1, \dots, w_d : V \rightarrow \mathbb{Z}_{\geq 0}$. These feature weights can be independent of each other as well as of objective weights. A *district* is a subset of vertices $T \subseteq V$ such that the induced subgraph $G[T]$ is connected and satisfies specific validity constraints: *compactness* (constraints on topology) and *composition* (constraints on feature weights). A *(partial) districting* solution is a collection of vertex-disjoint districts $\mathcal{T} = \{T_1, \dots\}$. We remark that finding a complete partition of the graph into vertex-disjoint districts (i.e., $\bigcup_{T \in \mathcal{T}} T = V$) may not be possible; rather, we seek a packing of disjoint districts with a maximum total objective weight.

For the compactness constraints, besides connectivity, we consider radius or diameter bounds. Specifically, the distance $d_G(u, v)$ between vertices u and v in G is the number of edges on the shortest path between u and v . A district T has bounded *strong radius* k if there exists a vertex $\gamma \in T$ such that for all other vertices $u \in T$, $d_{G[T]}(u, \gamma) \leq k$. Similarly, T has bounded *weak radius* k if there exists a vertex $\gamma \in V$ such that for all other vertices $u \in T$, $d_G(u, \gamma) \leq k$. A districting $\mathcal{T}$ satisfies strong (or weak) radius bound k if every $T \in \mathcal{T}$ has bounded strong (or weak) radius k . A district T has bounded *strong diameter* if all pairs of vertices u, v in T have distance in the induced subgraph at most k , i.e., $d_{G[T]}(u, v) \leq k$. It has bounded *weak diameter* if for every pair of vertices $u, v \in T$, their distance in the original graph G is at most k , i.e., $d_G(u, v) \leq k$.

Composition constraints impose limits on the feature weights of each district. First, we can extend our weight to a single district T and a districting solution $\mathcal{T}$. For any weight function (representing either the objective w or a feature w_i), the district weight is $w(T) = \sum_{v \in T} w(v)$, and the total weight of the districting solution $\mathcal{T}$ is $w(\mathcal{T}) = \sum_{T \in \mathcal{T}} w(T)$. We consider two types of composition requirements:

Balanced Districting. Given feature weights w_1, w_2 and a balancedness parameter $c \geq 2$, a district T is *c-balanced* [33] if:

$$\min\{w_1(T), w_2(T)\} \geq \frac{1}{c} (w_1(T) + w_2(T)). \quad (1)$$

A districting solution $\mathcal{T}$ is *c-balanced* if every district $T \in \mathcal{T}$ satisfies Eq. (1). For bicriteria approximation, we may also refer to a δ -relaxed *c-balanced* condition (where $0 < \delta < 1$):

$$\min\{w_1(T), w_2(T)\} \geq \frac{1}{c(1 + \delta)} (w_1(T) + w_2(T)). \quad (2)$$

Threshold Districting. Given a feature weight w_1 and a threshold parameter $B \geq 0$, a district T is a *B-threshold* district if $w_1(T) \geq B$. A districting solution $\mathcal{T}$ is *B-threshold* if every district $T \in \mathcal{T}$ is a *B-threshold* district.

Similarly, a δ -relaxed B -threshold district is defined by the condition: $w_1(T) \geq B(1 - \delta)$. We will use S, T or U to denote districts and $\mathcal{S}$ for the set of valid districts (e.g., the set of c -balanced districts with strong radius bound k).

Approximate and Bicriteria Solutions. Given a graph G , weights $w, \{w_i\}$, and constraint parameters (k and c or B), our goal is to find a districting solution $\mathcal{T}$ that maximizes the total objective weight $w(\mathcal{T})$ and every district $T \in \mathcal{T}$ satisfies the (compactness and composition) constraints. We say a districting solution $\mathcal{T}$ is an α -approximation if every district in $\mathcal{T}$ satisfies the exact constraints and

$$\alpha \cdot w(\mathcal{T}) \geq w(\mathcal{T}_{\text{OPT}}),$$

where $\mathcal{T}_{\text{OPT}}$ is an optimal solution satisfying the exact constraints. We also consider *bicriteria* (α, δ) -approximations. In this setting, the solution $\mathcal{T}$ satisfies the weight approximation guarantee ($\alpha \cdot w(\mathcal{T}) \geq w(\mathcal{T}_{\text{OPT}})$) but is permitted to satisfy the δ -relaxed composition constraints. For instance, in the c -balanced districting problem with strong radius bound k :

- An α -approximation requires that each district in $\mathcal{T}$ is strictly c -balanced and has strong radius at most k .
- A bicriteria (α, δ) -approximation allows the districts in $\mathcal{T}$ to be δ -relaxed c -balanced (i.e., balanced with parameter $c(1 + \delta)$), while still comparing the solution against the strictly c -balanced optimum $\mathcal{T}_{\text{OPT}}$.

Graph Types. A graph $G = (V, E)$ is *planar* if there exists an embedding of all vertices to the Euclidean plane such that all edges can be drawn without intersections other than the endpoints. A *face* of a planar embedding is a connected region separated by the embedded edges. G is said to be *outerplanar* if there exists an embedding of G such that there is a face containing all vertices. Often, this face is assumed to be the outer face. A graph H is said to be a *minor* of G if H is isomorphic to the graph obtained by a sequence of vertex deletions, edge contractions, and edge deletions from G . We say that G is *H -minor-free* if G does not have H as its minor. A graph is an *apex graph* if there is a vertex whose removal results in a planar graph. A graph G is *apex-minor free* if it does not have some apex graph H as a minor. See [37, 29] for more discussions on apex-minor-free graphs. We now formally define bounded expansion graphs.

► **Definition 1.** Let $G = (V, E)$ be an undirected (multi-)graph and let $f : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$ be a non-decreasing function. We say that G belongs to the class of f -bounded expansion graphs if the following holds. Let $(V_1, V_2, \dots, V_{n'})$ be a partition of V , where each part V_i induces a connected graph. Let k be the maximum diameter of the induced subgraphs $G[V_i]$ for all i . Consider the contracted graph H from G where each part V_i is contracted to a single vertex. Then, for any subgraph of $H' \subseteq H$, the density of H' satisfies $\frac{2|E(H')|}{|V(H')|} \leq f(k)$.

At a high-level, a bounded expansion graph G keeps a bounded density (and thus bounded degeneracy) if one contracts bounded diameter subgraphs all at once in G .

2.2 Algorithmic Framework

We first examine the districting problem from a bird's-eye view and explain how the technical tools introduced in this work can be applied more generally.

Linear Program Formulation. Recall that a candidate district is represented by a set of vertices $S \subseteq V$ as defined in Section 2.1. We use $\mathcal{S}$ as the set of all candidate districts that meet the requirements. This set $\mathcal{S}$ can be of potentially exponential size. For each candidate district $S \in \mathcal{S}$, we define a variable $x_S \in \{0, 1\}$ indicating whether or not this district is chosen. Due to the disjointness of districts, among all the districts that cover the same vertex v , only one of them can take a value of 1. Denote by $w(S)$ the weight of S . The integer program is defined as

$$\begin{aligned} & \text{maximize } \sum_S w(S)x_S \\ & \text{subject to } \forall v \in V, \sum_{S \ni v} x_S \leq 1 \\ & \quad \forall S, x_S \in \{0, 1\} \end{aligned} \tag{3}$$

To solve an integer solution, we use two steps. First, solve for an (approximate) solution to the LP problem. Then round the fractional solution to an integer solution. We discuss the two problems separately. To relax this integer program to a fractional solution, we allow x_S to take real values in $[0, 1]$ and define $x'_S = x_S \cdot w(S)$, $x'_S \in [0, w(S)]$. Rewriting the LP using x'_S , the primal and dual LP problems are as follows.

Primal: Maximizing $\sum_S x'_S$, subject to $\sum_{S: v \in S} \frac{x'_S}{w(S)} \leq 1$ for all $v \in V$; $x'_S \geq 0$ for all S .
Dual: Minimizing $\sum_v y'_v$, subject to $\sum_{v \in S} \frac{y'_v}{w(S)} \geq 1$ for all S ; $y'_v \geq 0$ for all v .

Solving the LP problem is straightforward if we are given a set of polynomially many valid districts. One scenario in which this could happen is when a districting already exists on G and, for a new districting solution, only minor changes to the previous solution are acceptable. In general, the number of potential districts could be exponentially large, even if we consider only stars (e.g., under the compactness restriction). We cannot explicitly write the variables and constraints in the LP and must rely on a separation oracle. For the dual LP problem, a *strongly violating dual constraint* is a district S such that S is valid and for the dual variables values of vertices in S and $0 < \varepsilon < 1$, $\sum_{v \in S} y'_v < (1 - \varepsilon)w(S)$. We formulate the task of implementing a separation oracle as follows:

Input: Graph G , dual variables $\{y'_v\}$, weight function $w : V \rightarrow \mathbb{Z}_{\geq 0}$, implicit set $\mathcal{S}$ of all valid districts.
Goal: Either returns a violating district $S \in \mathcal{S}$ such that its weight $w(S)$ is at least half to the maximum weight among all violating districts, and

$$\sum_{v \in S} y'_v < (1 - \varepsilon/2)w(S); \tag{4}$$

or reports that *all* districts $S \in \mathcal{S}$ satisfy $\sum_{v \in S} y'_v \geq (1 - \varepsilon)w(S)$.

Randomized Rounding. Once we have an (approximate) fractional solution to the LP, we use the same randomized rounding as in [33] to find an integer solution I : sort all districts with non-zero x -values in decreasing order of their weights; in this order a district S is selected with probability proportional to x_S , if it does not overlap with any districts already selected in the integer solution. This rounding procedure introduces another approximation factor of $O(\tau)$, where τ is called the *correlation ratio* and bounds the correlation term $\sum_{A, B \in \mathcal{S}, A \cap B \neq \emptyset} x_A x_B \leq \tau \sum_{S \in \mathcal{S}} x_S$.

We state this lemma from [33] for completeness. Suppose $\mathcal{S}_{\geq \delta}$ is the set of districts from the fractional LP solution that have x -value at least δ and $w(I)$ is the total weight of the integer solution I after rounding. The expectation is on the random coin flips in the rounding process.

► **Lemma 2** (Lemma B.5 in [33]). *Suppose that there exists a $\tau \in \mathbb{R}_{>0}$ such that for all $\delta > 0$,*

$$\sum_{A, B \in \mathcal{S}_{\geq \delta}, A \cap B \neq \emptyset} x_A x_B \leq (\tau/2) \cdot \sum_{S \in \mathcal{S}_{\geq \delta}} x_S,$$

then $E[w(I)] \geq \frac{1}{2\tau} \sum_S w(S) \cdot x_S$.

3 Technical Overview

This section summarizes the main algorithmic ideas and the core technical challenges underlying Sections 4 and B. The new progress requires addressing two major challenges. First, we need to identify the candidate districts that meet the specified requirements (connectivity, compactness, balancedness, or weight threshold). Next, we choose a disjoint subset of the candidate districts with maximum total weight. The first issue is related to *feasibility*, and the second issue regards *disjointness* or *packing*.

At a high level, our positive results rely on two complementary approaches. Section 4 develops a refined analysis of a linear programming (LP) relaxation combined with randomized rounding, yielding constant-factor approximations on minor-free and bounded-expansion graph classes. Section B strengthens the result for apex-minor-free graphs by combining Baker’s layering technique with a new dynamic programming (DP) framework on bounded-treewidth graphs, leading to a PTAS under relaxed parameters.

3.1 Refined Rounding Analysis (Section 4)

Recall that x_S is the primal variable in the LP formulation for a district S , $x_S \in [0, 1]$. We consider districts of strong radius k . The LP framework allows addressing feasibility and packing in an orthogonal manner. The separation oracle helps to find a feasible district that substantially improves the LP solution, and the randomized rounding manages packing.

The performance of the randomized rounding procedure is governed by a second-order correlation term $\sum_{A \cap B \neq \emptyset} x_A x_B$, which measures the extent to which overlapping districts interfere during rounding. Given a graph G and $\mathcal{S}$ (a collection of connected subgraphs of G), we say that $G, \mathcal{S}$ has *correlation ratio* τ if $\sum_{S: v \in S} x_S \leq 1$, and $\sum_{A \cap B \neq \emptyset} x_A x_B \leq \tau \sum_S x_S$, for all v and $\{x_S\}$. Lemma 2 gives an $O(\tau)$ approximation guarantee if we can establish any bound on τ . In particular, in Lemma 3 we show that if G is H -minor-free for some H of size h , then τ can be bounded by a function $q(h, k)$ that depend solely on h and k . Establishing this bound yields a constant-factor approximation. The main technical difficulty is that intersecting districts may overlap in highly structured ways. Even when each district has a strong radius k , a fixed district A could, a priori, intersect with many other districts B . A naïve charging argument could therefore lead to an $O(n)$ blow-up.¹

To overcome this, we exploit the structural sparsity of minor-free graphs. Specifically, an H -minor-free graph has bounded *degeneracy* – there is an orientation of the edges such that the out-degree of any vertex is bounded by $O(h\sqrt{\log h})$ [65, 3]. Using this property, we

¹ For example, consider A being a star of $\Omega(n)$ vertices, with each leaf intersecting distinct districts. In this case, the terms x_A involved in the correlation sum are at least $\Omega(n)x_A$. This shows some districts will be charged for $\Omega(n)$ times in a naïve charging argument.

design a probabilistic charging scheme based on a sequence of random edge orientations and contractions. For each intersecting pair (A, B) , we identify a carefully structured “intersecting pivot” along a canonical path between their centers. We then show that with probability $\Omega_{h,k}(1)$, the pair can be charged to one endpoint district in a way that ensures:

- Each intersecting pair is charged with constant probability.
- Each district receives only $O_{h,k}(1)$ total charge.

This argument combines bounded degeneracy, controlled path contractions, and a forward-trajectory encoding that limits the number of distinct pivots that can be associated with a fixed district. The result is a constant bound on the correlation term, completing the rounding analysis. The analysis here is independent of how the LP fractional solution is computed and thus can be combined with any algorithm or implementation that produces an (approximate) fractional solution.

A similar argument extends to bounded-expansion graphs (Section 4.3), since such graphs have constant “density” everywhere and also enjoy bounded degeneracy. With a slightly modified contraction, the correlation term is bounded by a factor that depends only on the density parameter. We remark that the randomized rounding algorithm we use is fairly standard [23, 15], but the analysis of its performance is new and may find further applications.

The new analysis of randomized rounding, combined with the separation oracle for star districts in [33], immediately yields an $O(1)$ -approximation algorithm for packing c -balanced or B -threshold star districts in planar, minor-free, and bounded-expansion graphs.

In order to handle districts of a general radius k with $k > 1$, we also need to extend the separation oracle for star districts in [33] to accommodate a general radius k (Appendix A). To do that, we formulate a vector-valued subset sum problem on a connected compact subgraph. Similar to the standard subset sum, we trim the solution space by keeping a succinct representation of the subsets that approximate every possible solution with a multiplicative factor and maximize a linear score. This new framework can be applied to both c -balanced districts and threshold districts, as well as to a generic connected knapsack problem (see [32]). For the connected subgraph sum problem, we present an FPTAS on graphs of *bounded treewidth*. On a planar graph or an apex-minor-free graph, the k -hop neighborhood of any center vertex has bounded treewidth [37, 11]. Thus, we can enumerate every possible center vertex and apply the separation oracle in each bounded radius- k neighborhood. This provides an efficient separation oracle for planar and apex minor-free graphs.

3.2 PTAS for Apex Minor-Free Graphs (Appendix B)

Section B improves the approximation guarantee for apex-minor-free graphs using a different paradigm with two elements: the Baker’s layering approach [7] for planar and apex minor-free graphs [30] and a PTAS for *subgraph packing* for graphs of bounded treewidth.

We explain the main idea for planar graphs. The Baker’s approach considers a BFS from an arbitrary root and removes the vertices on the layers $= i \bmod t$, for some $0 \leq i \leq t - 1$. With a fixed i , this will partition the graph into connected components, each with at most t consecutive layers of the BFS. Thus, each piece is t -outerplanar and thus has treewidth at most $3t - 1$ [11].

Suppose, for now, we have a magic algorithm $\mathcal{A}$ that solves compact subgraph packing in a graph of bounded treewidth. We apply algorithm $\mathcal{A}$ for each piece, such that all obtained districts from the connected components combined form a feasible packing solution for the original graph G . If we try this for all possible values of i between 0 and $t - 1$, and take

the maximum solution, this is a $(1 + \varepsilon)$ approximation to the optimal with $t = O(k/\varepsilon)$. The reason is that in an optimal solution, only the subgraphs that are “influenced” by the removed layers are destroyed. These vertices are within $O(k)$ hops from the removed layers since each compact subgraph has radius k . By the pigeonhole principle, at least one of the t candidate solutions keeps all but $O(k/t) = O(\varepsilon)$ fraction of the optimal weight.

The central technical task is therefore to find the algorithm $\mathcal{A}$ that solves the problem of packing compact districts on bounded-treewidth graphs. Unlike the separation oracle problem (which asks for a *single* feasible district), packing requires simultaneously selecting *multiple* disjoint connected subgraphs subject to either balancedness or threshold constraints. This introduces a significantly more complex state space.

We develop a dynamic programming scheme over a tree decomposition that encodes partial district configurations within each bag. The DP must simultaneously track the connectivity information of partially formed districts and the cumulated weight (for threshold constraints) or the weight balance (for c -balanced districts). Our DP mechanism has to relax the parameter c slightly (for the c -balanced districts) or the weight threshold B (for the threshold districts) by a factor.

4 $O(1)$ -approximation on Planar Graphs and Beyond

4.1 Constant Ratio for Minor-Free Graphs

In this section, we briefly describe the linear programming approach for the district packing problem, together with the randomized rounding framework based on the support variables of an optimal linear program solution. The approximation guarantee of this framework relies on bounding certain second-order terms involving the support variables. In particular, throughout this section we aim to prove the following key technical Lemma 3 for minor-free graphs and Theorem 13 for bounded expansion graphs.

► **Lemma 3.** *For all $k \geq 1$ and $h \geq 1$, there exists $q(h, k) = h^{O(k^2)}$ such that the following holds. Let G be an H -minor-free graph with $|H| \leq h$, $\mathcal{S}$ be the collection of connected subgraphs with strong radius k , and $x : V \rightarrow \mathbb{R}_{\geq 0}$ is a nonnegative weight function satisfying $\sum_{S \in \mathcal{S}: S \ni v} x_S \leq 1$ for all v . Then,*

$$\sum_{A, B \in \mathcal{S} \text{ and } A \cap B \neq \emptyset} x_A x_B \leq q(h, k) \cdot \sum_{S \in \mathcal{S}} x_S.$$

Combining the above lemma with the linear programming formulation and the randomized rounding framework yields the following results.

► **Theorem 4.** *There exists a polynomial-time algorithm that computes an $O(1)$ -approximate solution for packing c -balanced or threshold districts with bounded strong radius on planar graphs and apex-minor-free graphs.*

► **Theorem 5.** *There exists a polynomial-time algorithm that computes an $O(1)$ -approximate solution for packing c -balanced or threshold star districts on H -minor-free graphs and bounded-expansion graphs.*

We prove Theorem 4 and Theorem 5 at the end of this section. The high level approach of proving Lemma 3 is to partition the pairs of intersecting districts into different cases, and give the upper bound to each of the case. To start with the simplest case, we first introduce the following notions of *center vertices* and *intersecting pivots*.

Center Vertices and Intersecting Pivots. For each district $S \in \mathcal{S}$, we designate a *center* vertex $c_S \in S$ to the district, where every vertex in S is within distance k from c_S . If there are multiple choices, c_S can be chosen arbitrarily. Note that since S has a strong radius at most k , there is at least one center vertex in S .

For each pair of districts (A, B) that intersect, we further define the following terms.

► **Definition 6.** Let $A, B \in \mathcal{S}$ be two districts with $A \cap B \neq \emptyset$. Let $P_{A,B}$ be an arbitrarily chosen (undirected) path that satisfies the following conditions:

1. $P_{A,B}$ is a path connecting c_A and c_B . If $c_A = c_B$ then $P_{A,B}$ is a single vertex.
2. There exists an intersecting pivot $i_{A,B} \in P_{A,B} \cap A \cap B$.
3. The length of the subpath between c_A and $i_{A,B}$, and the length of the subpath between c_B and $i_{A,B}$ are both at most k .
4. All vertices on the path between c_A and $i_{A,B}$ are in A , and all vertices on the path between $i_{A,B}$ and c_B are in B .

We require $P_{A,B} = P_{B,A}$ and $i_{A,B} = i_{B,A}$.

Now, with the definition of the intersecting pivot, we are able to partition the set of intersecting district pairs (A, B) into two types – an “easy” case where $i_{A,B} \in \{c_A, c_B\}$, or otherwise a “non-trivial” case.

For the easy case, without loss of generality, we may assume $i_{A,B} = c_A$. In this case, we bound the total correlation using the primal constraints:

$$\sum_{A \cap B \neq \emptyset \text{ and } i_{A,B} = c_A} x_A x_B \leq \sum_{A \in \mathcal{S}} x_A \sum_{B \ni c_A} x_B \leq \sum_{S \in \mathcal{S}} x_S$$

Therefore, it remains to bound the correlation sum for all intersecting districts (A, B) , where c_A, c_B and $i_{A,B}$ are all distinct vertices.

4.2 A Probabilistic Method for Challenging Case

The analysis of the easy case does not simply generalize to the non-trivial case where $i_{A,B}, c_A$ and c_B are all distinct. One potentially convincing reason is that if we fix the district A , all the intersecting pivots $i_{A,B}$ might be distinct, which implies that the sum over all intersecting districts $\sum_{B: A \cap B \neq \emptyset} x_A x_B$ could be as large as $|A| \cdot x_A = \Omega(n) \cdot x_A$ in the worst case.

Fortunately, one can utilize the bounded degeneracy property of H -minor-free graphs for a fixed minor H of size h . A d -degenerate graph is an undirected graph in which every subgraph has at least one vertex of degree at most d . The *degeneracy* of a graph is the smallest value d for which it is d -degenerate. The following fact bounds the degeneracy of H -minor-free graphs.

► **Lemma 7** ([65, 3]). Let $G = (V, E)$ be an H -minor free graph for a fixed graph H of h vertices, then the degeneracy of G is $d = O(h\sqrt{\log h})$.

With this property, we are able to design a randomized charging scheme, such that for each pair of intersecting districts (A, B) , with probability $\Omega_{h,k}(1)$, the cost $x_A x_B$ will be charged to either x_A or x_B and simultaneously each x_A or x_B will be charged by at most $O_{h,k}(1)$ times. This completes the proof by the standard probabilistic method. We summarize the above discussion into the following lemma, which directly leads to Lemma 3.

► **Lemma 8.** Let $G = (V, E)$ be an H -minor-free graph (where H has h vertices) and $\mathcal{S}$ be a set of subgraphs with strong radius k . Then, there exists a randomized algorithm $\mathcal{R}$ that takes $(G, \mathcal{S}, \{x_S\})$ as input, and output a list $\mathcal{O}$ of ordered tuples $(A, B, j) \in \mathcal{S} \times \mathcal{S} \times V$, such that the following holds.

10:12 Packing Compact Subgraphs with Applications to Districting

- (1) $j \in B$.
- (2) For each district $A \in \mathcal{S}$, the set $J_A := \{j \mid (A, B, j) \in \mathcal{O}\}$ contains at most $d^{k+1} \binom{2k}{k}$ distinct vertices.
- (3) For any intersecting pair (A, B) , with probability at least $1/(d+1)^{k^2}$, either $(A, B, j) \in \mathcal{O}$ or $(B, A, j') \in \mathcal{O}$, for some $j \in B$ or some $j' \in A$.

Before proving Lemma 8, we first show that Lemma 8 implies the correctness of Lemma 3, the main technical lemma of this section.

Proof of Lemma 3. By item (3) in Lemma 8, we obtain the expected bound:

$$\mathbb{E}_{\mathcal{R}} \left[\sum_{(A,B,j) \in \mathcal{O}} x_A x_B \right] \geq \frac{1}{(d+1)^{k^2}} \sum_{A \cap B \neq \emptyset} x_A x_B.$$

On the other hand, by items (1) and (2) we have

$$\sum_{(A,B,j) \in \mathcal{O}} x_A x_B = \sum_{A \in \mathcal{S}} x_A \sum_{j \in J_A} \left(\sum_{B: (A,B,j) \in \mathcal{O}} x_B \right) \leq d^{k+1} \binom{2k}{k} \sum_{S \in \mathcal{S}} x_S.$$

By the probabilistic method argument, there must exist a set of intersecting pairs $\mathcal{O}'$ that satisfies both inequalities, which completes the proof of Lemma 3 with $q(h, k) = (d+1)^{k^2} d^{k+1} \binom{2k}{k} = h^{O(k^2)}$. $\blacktriangleleft$

The rest of the section devotes to proving Lemma 8. We first describe the randomized algorithm $\mathcal{R}$ below.

The Algorithm $\mathcal{R}$. The algorithm consists of two parts: in the first part, the algorithm randomly produces a sequence of minors of G : $G_0 = G, G_1, \dots, G_k$ via a sequence of random contractions, where k is the radius/diameter constraint. In the second part, the algorithm considers each intersecting district pair (A, B) , obtains some vertex $j \in V$, and decides if the tuple (A, B, j) will be added to the output list and counted toward the correlation sum. Note that due to asymmetry, (A, B) and (B, A) are considered separately.

Part 1. The algorithm initializes a copy $G_0 \leftarrow G$ of the input graph. Run the following procedure in k rounds. In each round $r = 1, 2, \dots, k$, the algorithm computes an arbitrary orientation $\vec{G}_{r-1}$ from G_{r-1} , where each vertex has out-degree at most d . The algorithm then assigns a randomly chosen neighbor (or not choosing at all) for each vertex, denoted as the variable $\text{pick}_r[v] \in V \cup \{\perp\}$. Specifically, for each vertex $v \in V$, each out-going neighbor $v' \in \text{out}(v)$ will be assigned to $\text{pick}_r[v]$ with probability $\frac{1}{d+1}$, and for the remaining probability $1 - |\text{out}(v)|/(d+1)$, the algorithm chooses nothing (sets $\text{pick}_r[v] \leftarrow \perp$) for v .

Finally, at the end of each round, the algorithm forms the (undirected) contracted graph G_r by contracting v toward $\text{pick}_r[v]$ all at once. The contraction naturally induces a partition of $V(G_{r-1})$, where each part is contracted to a single vertex and there is at most one vertex v in the part with $\text{pick}_r[v] = \perp$. If such vertex v exists, we name the contracted vertex in G_r to be v . Otherwise, we do not need this vertex anymore so they can be completely removed from G_r . We emphasize that this naming convention is important to our analysis.

Part 2. The algorithm computes an orientation $\vec{G}_k$ from the latest contracted graph G_k . For each pair of intersecting district (A, B) with $i_{A,B}$ different from c_A and c_B , the algorithm checks if the following event $E_{A,B}$ holds:

- The vertices c_A, c_B are never contracted. That is, $\text{pick}_r[c_A] = \text{pick}_r[c_B] = \perp$ for all r .
- $\text{pick}_r[i_{A,B}] = \perp$ for all $r = 1, 2, \dots, k-1$. Note that we do not restrict $\text{pick}_k[i_{A,B}]$.
- For every other vertex v on $P_{A,B}$ except $c_A, c_B, i_{A,B}$, whenever $\text{pick}_r[v] \neq \perp$, we must have $\text{pick}_r[v] \in P_{A,B}$. In other words, if a vertex on $P_{A,B}$ participates in a contraction, the vertex contracts “along the path” and shortens the path.
- There are three possible values for $\text{pick}_k[i_{A,B}]$ (contract to the left, contract to the right, or stay.) Define j' and j to be the two vertices on $P_{A,B}$ closest to $i_{A,B}$ that are not contracted in the end, such that $j' \in A$ and $j \in B$. In particular, if $\text{pick}_k[i_{A,B}] = \perp$ we let $j' = j = i_{A,B}$.

The event $E_{A,B}$ requires that, if $j' \neq j$, then on the last orientation $\vec{G}_k$ there must be a directed edge from j' to j . (That is, swap the role of A and B when needed.)

- Let $P'_{A,B}$ be the contracted path after the last round. Note that $j \in P'_{A,B}$ and $j \neq c_A$. Let j'' be the neighbor of j on $P'_{A,B}$ toward c_A . Let us define the term *forward trajectory* on the union of oriented graphs $G^* := \vec{G}_0 \cup \vec{G}_1 \cup \dots \cup \vec{G}_{k-1}$, where each directed edge is associated with the round index. A forward trajectory is then a path P^* on G^* such that the round indices is non-decreasing.

The event $E_{A,B}$ requires the existence of a *forward trajectory* P_A^* from c_A to j'' with $V(P_A^*) \subseteq V(P_{A,B})$. If $E_{A,B}$ holds, the algorithm adds (A, B, j) to the output list $\mathcal{O}$.

Analysis. We verify each item in the statement of Lemma 8 as follows.

► **Lemma 9.** *If $(A, B, j) \in \mathcal{O}$, then $j \in B$.*

Proof. This is guaranteed by Item 4 of the event $E_{A,B}$. ◀

► **Lemma 10.** *For each district $A \in \mathcal{S}$ the set J_A contains at most $d^{k+1} \binom{2k}{k}$ distinct vertices.*

Proof. Each forward trajectory can be uniquely encoded with a sequence of “next round” or “go to the i -th neighbor”, where the “next round” appears exactly k times and the “go to” appears for at most k times (this is because the radius of A is at most k). So there are at most $d^k \binom{2k}{k}$ vertices of $j' \in A$.

By Item 4 of the event $E_{A,B}$, there is a directed edge going from j' to j in the last oriented graph $\vec{G}_k$. This bounds the size of J_A to be at most $d^{k+1} \binom{2k}{k}$. ◀

► **Lemma 11.** *The probability of an intersecting pair (A, B) appearing in the output list (in either the form (A, B, j) or (B, A, j')) is at least $1/(d+1)^{k^2}$.*

Proof. For any r , $0 \leq r \leq k$, let $P_{A,B}^{(r)}$ be the contracted path of $P_{A,B}$ in G_r . Let $\vec{P}_{A,B}^{(r)}$ be the associated sequence of directed edges on $\vec{G}_r$ after the orientation.

It suffices to observe that, for rounds $r = 1, 2, \dots, k-1$, if the oriented sequence of edges $P_{A,B}^{(r-1)}$ do not belong to one of the following 9 cases, with probability at least $1/(d+1)^k$, the path between c_A and $i_{A,B}$ and the path between $i_{A,B}$ and c_B will become shorter. These 9 cases are:

1. $c_A \rightarrow u \leftarrow i_{A,B} \rightarrow v \leftarrow c_B$
2. $c_A \rightarrow u \leftarrow i_{A,B} \leftrightarrow c_B$ (2 cases)
3. $c_A \leftrightarrow i_{A,B} \rightarrow v \leftarrow c_B$ (2 cases)
4. $c_A \leftrightarrow i_{A,B} \leftrightarrow c_B$ (4 cases)

By the observation, after $k - 1$ rounds, the oriented sequence of $P_{A,B}^{(k-1)}$ on $\vec{G}_{k-1}$ must be in one of the 9 cases stated above (we call them *directed-minimal* paths). Now, consider the last round k . For each of the 9 cases, there is at least one scenario (i.e., with probability at least $1/(d+1)^5$) such that $E_{A,B}$ or $E_{B,A}$ holds, which completes the proof. $\blacktriangleleft$

Proof of Lemma 8. It follows from Lemma 9, Lemma 10, and Lemma 11. $\blacktriangleleft$

4.3 Bounded Expansion Graphs

While the correlation ratio is constant on H -minor-free graphs (Lemma 3), it has an $\Omega(\sqrt{n})$ lower bound on general graphs [33]. This dichotomy naturally raises the question of how broadly the constant bound persists, and whether one can obtain a smooth tradeoff between these two regimes. In this subsection, we take a step in this direction by considering bounded expansion graphs [60, 61, 62] (refer Definition 1), a class that formalizes the notion of *everywhere sparsity*. To better demonstrate our results, we first introduce a recursively defined sequence $\{d_r^f\}_{r=0}^\infty$.

► **Definition 12.** Let $f : \mathbb{N} \rightarrow \mathbb{R}_{\geq 1}$ be any increasing function. Define $d_0^f := f(0)$. For each integer $r \geq 1$, define

$$d_r^f := f\left(2^r \cdot d_0^f d_1^f \cdots d_{r-1}^f\right).$$

If the context is clear we will omit f and use d_r to denote d_r^f . We remark that when f is a polynomial, d_r^f is doubly exponential in r .

► **Theorem 13.** Let G be an f -bounded expansion graph. Let $\mathcal{S}$ be a collection of districts with a strong radius bound k . Let $\{x_S\}$ be any feasible fractional solution for packing $\mathcal{S}$. Then, we have:

$$\sum_{A \cap B \neq \emptyset} x_A x_B \leq \left(4^k d_k^{2k^2+k}\right) \cdot \binom{2k}{k} \cdot \sum_{S \in \mathcal{S}} x_S.$$

It is straightforward to check that (from Definition 12), whenever k is a constant, despite being enormous, the approximation ratio is still a constant.

Algorithm $\mathcal{R}$ in Lemma 8 does not directly imply Theorem 13, because in each contraction which produces G_r from G_{r-1} , the diameter of each contracted vertex subset is not bounded. Fortunately, with a slight tweak to algorithm $\mathcal{R}$, we are able to guarantee the diameter on each contraction, thereby proving Theorem 13.

Proof of Theorem 13. We first describe the modified algorithm $\mathcal{R}'$, and then provide the modified analysis that is analogous to Lemma 10 and Lemma 11.

The Modified Algorithm $\mathcal{R}'$. In each round r of contraction, the algorithm first computes an *acyclic* orientation $\vec{G}_{r-1}$ via the greedy ordering algorithm² for computing the degeneracy d_{r-1} . This also computes a $(d_{r-1} + 1)$ -vertex-coloring for G with colors $\{1, 2, \dots, d_{r-1} + 1\}$. Then, the algorithm computes a random shuffle σ to the colors. Each vertex v obtains a color $c(v) \in \{1, 2, \dots, d_{r-1} + 1\}$.

² The algorithm repeatedly identifies a vertex with the minimum current degree, orients all incident edges as outgoing edges, and then remove this vertex and all incident edges until the graph become empty.

The algorithm then assigns the value $pick_r[v]$ for each vertex v with a slightly modified distribution: for each neighbor u of v , if $c(u) < c(v)$, assigns $pick_r[v] = u$ with probability $1/(d_{r-1} + 1)$. Otherwise, set $pick_r[v] = \perp$.

The rest of the algorithm is the same as algorithm $\mathcal{R}$.

The Analysis. To complete the proof of Theorem 13, it suffices to show two statements, analogous to Lemma 10 and Lemma 11. First, we analyze contracted components' diameter. Consider any round r of contraction and any vertex v . The sequence formed by successively applying $pick_r[\cdot]$ has length at most $2d_{r-1}$ on G_{r-1} . Therefore, by induction, the diameter of each contracted component of G at round r for the shallow minor is at most $f(2^r d_0 d_1 \cdots d_{r-1}) = d_r$.

Let (A, B) be any intersecting pair and $P_{A,B}$ be a connecting path from c_A to c_B . To count the number of forward trajectories, we simply observe that the outdegrees for each vertex in any contracted graphs after all k rounds is d_k . Any forward trajectories takes at most k steps hopping forward and k steps advancing to the next round. Thus the size of J_A in the output list can be bounded by $d_k^k \binom{2k}{k}$ as desired.

Suppose the contracted path $P_{A,B}^{(r)}$ at round r has length at least 5. So none of the nine directed-minimal cases occur. In this case, one of the partial path between c_A and $i_{A,B}$ or between c_B and $i_{A,B}$ must have at least length three. There must be an edge $x \rightarrow y$ in the oriented graph that does not invalidate the event $E_{A,B}$ (or $E_{B,A}$) whenever $pick_r[x] = y$. Now, the probability that the algorithm assigns $pick_r[x] = y$ is no longer $1/(d_{r-1} + 1)$. In fact, since shuffling colors on $\vec{G}_{r-1}$ is independent to the initial assignment to $pick_r[x]$, we know that $\Pr(c(y) < c(x)) = 1/2$. Hence, we conclude that with probability (A, B) appearing in the output list (in either form (A, B, j) or (B, A, j')) is at least

$$\prod_{r=1}^k \left(\frac{1}{2(d_{r-1} + 1)^k} \right)^2 = 4^{-k} \cdot \prod_{r=1}^k (d_{r-1} + 1)^{-2k} \geq 4^{-k} d_k^{-2k^2},$$

and the statement follows. $\blacktriangleleft$

► **Remark.** Although Theorem 13 applies to strong radius districts, it does not immediately yield an efficient algorithm for our packing problem. The main obstacle lies in algorithmically extracting a polynomial-sized support from the linear program. In particular, employing either the multiplicative weight update method or the ellipsoid method necessitates an efficient implementation of the separation oracle. However, designing such an oracle appears to be surprisingly challenging, and we currently do not have a solution, even for everywhere sparse graphs.

On the other hand, when every vertex neighborhood has bounded treewidth (as is the case for *apex-minor-free graphs*) an efficient implementation of the separation oracle becomes possible. In the next subsection, we present such a separation oracle.

5 Conclusion and Open Problems

In this paper, we investigated several variants of packing compact subgraphs into a larger graph under balancedness/threshold composition constraints. We extend the work of [33], who focused only star districts, by studying districts of bounded radius.

Our work provides algorithmic results for two variants of the problem. On one hand, we design an $O(1)$ -approximation algorithm under exact c -balanced and B -threshold districting requirements on apex-minor-free graphs. This is done by solving a linear program and then

applying the rounding framework. On the other hand, we obtain a PTAS on apex-minor-free graphs by relaxing the constraints, allowing $(1 + \delta)c$ -balancedness and $(1 - \delta)B$ threshold parameters, for any $\delta > 0$. This is achieved by dynamic programming on tree decomposition with Baker’s method.

Interestingly, we do not immediately see a clean path to get the best of the two worlds – a PTAS on apex-minor-free graphs with exact c -balanced or B threshold constraints. The rounding framework leaves an $\Omega(1)$ integrality gap, which seems very hard to overcome (as discussed in [33]).

On the other hand, for the dynamic programming approach, The main bottleneck appears to be maintaining intermediate states that encode multiple partial districts with exact composition requirements simultaneously. Our algorithm requires “state blow-up” and “trimming” procedures at each DP transition step. The trimming step maintains an invariant that specifically holds for only one partial district in the description of Equation (7). Similar to the 3-color-balancedness hardness result (Theorem 6.5 in full version), the current trimming algorithm does not extend naturally to guarantee Equation (7). Clearly, if we choose not to trim at all, the size of the state space grows exponentially with n , thereby yielding an intractable algorithm.

However, this problem is not as dire as the 3-color case. One example is that we can obtain a PTAS for *grid graphs* by applying Baker’s idea *twice*: we split the grid into $t^2 = O(k/\varepsilon) \times O(k/\varepsilon)$ square subgrids and consider all t^2 possible shifts. Since each connected component has only a constant number $\Theta(t^2) = O_{k,\varepsilon}(1)$ of vertices, there is a constant-time algorithm for obtaining the optimal solution on each piece, leading to a PTAS overall. We are therefore optimistic about the following conjecture.

► **Conjecture 14.** *There exists a PTAS for packing c -balanced or B -threshold districts on apex-minor-free graphs.*

Efficient Separation Oracles for Various Compact Subgraph Types. Our current rounding framework handles exact balancedness constraints only for districts of bounded *strong radius*. In contrast, the relaxed PTAS relies on districts of bounded weak radius.

► **Open Question 15.** *Is the correlation ratio bounded by a constant for bounded weak-radius districts?*

Beyond Apex-Minor-Free Graphs. Our current techniques rely on dynamic programming on bounded-treewidth graphs. Extending these results to broader graph classes will likely require fundamentally new ideas. In particular, for obtaining similar approximation guarantees for broader graph classes, such as bounded-expansion graphs.

► **Open Question 16.** *Does there exist an efficient separation oracle for bounded strong-radius districts on bounded-expansion graphs?*

Graphs Parametrized by Worst-Case Correlation Ratio. In this paper, as well as in [15, 33], we have seen rounding algorithms whose approximation guarantee is closely related to the correlation ratio. Roughly speaking, this correlation ratio describes how districts satisfying a certain composition requirement interact with one another in the worst-case vertex-weight assignment. From a purely graph-theoretical perspective, it gives a flavor of per-graph-class analysis of approximation ratios. For example, in Section 4 we have shown that in planar graphs and bounded expansion graphs have a constant worst-case correlation ratio. In [33] it shows an $\Omega(\sqrt{n})$ lower bound ratio on general graphs and a $\Omega(1)$ lower bound ratio on planar

graphs. It is not hard to verify (using balanced separators) that on constant treedepth graphs the ratio is $O(1)$ as well and on bounded treewidth and pathwidth graphs the ratio is $O(\log n)$, but we do not know if this polylogarithmic bound is tight. We are interested in connecting this parameter to other graph parameter and classes in a reversed way, perhaps leading to universal optimal (per-graph topology) analysis on approximation ratios. Finally, we wonder whether there are any interesting observations that can be made from the parametrization of this worst-case correlation ratio.

► **Open Question 17.** *Are there interesting results when defining the graph classes according to the correlation ratio?*

References

- 1 Nesrine Abbas and Lorna Stewart. Clustering bipartite and chordal graphs: Complexity, sequential and parallel algorithms. *Discrete applied mathematics*, 91(1-3):1–23, 1999.
- 2 Maxwell Allman, Itai Ashlagi, Irene Lo, Juliette Love, Katherine Mentzer, Lulabel Ruiz-Setz, and Henry O’Connell. Designing school choice for diversity in the San Francisco unified school district. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, EC ’22, pages 290–291, New York, NY, USA, 2022. Association for Computing Machinery. doi:10.1145/3490486.3538271.
- 3 Noga Alon, Michael Krivelevich, and Benny Sudakov. Complete minors and average degree: A short proof. *J. Graph Theory*, 103(3):599–602, 2023. doi:10.1002/JGT.22937.
- 4 Micah Altman. Traditional districting principles: judicial myths vs. reality. *Social Science History*, 22(2):159–200, 1998.
- 5 Micah Altman and Michael P McDonald. Bard: Better automated redistricting. *Journal of Statistical Software*, 42:1–28, 2011.
- 6 Maxim Babenko and Alexey Gusakov. New exact and approximation algorithms for the star packing problem in undirected graphs. In Thomas Schwentick and Christoph Dürr, editors, *28th International Symposium on Theoretical Aspects of Computer Science (STACS 2011)*, volume 9 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 519–530, Dagstuhl, Germany, 2011. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.STACS.2011.519.
- 7 Brenda S Baker. Approximation algorithms for NP-complete problems on planar graphs. *J. ACM*, 41(1):153–180, January 1994. doi:10.1145/174644.174650.
- 8 Assaf Bar-Natan, Lorenzo Najt, and Zachary Schutzman. The gerrymandering jumble: map projections permute districts’ compactness scores. *Cartography and Geographic Information Science*, 47(4):321–335, 2020.
- 9 Amariah Becker and Justin Solomon. Redistricting algorithms. In Moon Duchin and Olivia Walch, editors, *Political Geometry: Rethinking Redistricting in the US with Math, Law, and Everything In Between*, pages 303–340. Springer International Publishing, Cham, 2022. doi:10.1007/978-3-319-69161-9_16.
- 10 Matthias Bentert, Fedor V. Fomin, Petr A. Golovach, Tuukka Korhonen, William Lochet, Fahad Panolan, M. S. Ramanujan, Saket Saurabh, and Kirill Simonov. Packing short cycles. *ACM Trans. Algorithms*, 22(1):8:1–8:35, 2026. doi:10.1145/3765285.
- 11 Hans L Bodlaender. A partial k -arboretum of graphs with bounded treewidth. *Theor. Comput. Sci.*, 209(1):1–45, 1998. doi:10.1016/S0304-3975(97)00228-4.
- 12 Matteo Böhm, Adriano Fazzzone, Stefano Leonardi, and Chris Schwiegelshohn. Fair clustering with multiple colors. arXiv 2002.07892.
- 13 Victor Bucarey, Fernando Ordóñez, and Enrique Bassaletti. Shape and balance in police districting. In *Applications of location analysis*, pages 329–347. Springer, 2015.
- 14 Alberto Caprara and Romeo Rizzi. Packing triangles in bounded degree graphs. *Inf. Process. Lett.*, pages 175–180, 2002. doi:10.1016/S0020-0190(02)00274-0.

- 15 Timothy M. Chan and Sarel Har-Peled. Approximation algorithms for maximum independent set of pseudo-disks. *Discret. Comput. Geom.*, 48(2):373–392, 2012. doi:10.1007/S00454-012-9417-5.
- 16 Harry Chen, Kamesh Munagala, and Govind S. Sankar. Balanced spanning tree distributions have separation fairness. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2337–2402, 2026. doi:10.1137/1.9781611978971.84.
- 17 Anshuman Chhabra, Karina Masalkovaitė, and Prasant Mohapatra. An overview of fairness in clustering. *IEEE Access*, pages 130698–130720, 2021. doi:10.1109/ACCESS.2021.3114099.
- 18 Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 5036–5044, 2017.
- 19 Gabriel Chuang, Oussama Hanguir, and Clifford Stein. Drawing Competitive Districts in Redistricting. In Guy N. Rothblum, editor, *5th Symposium on Foundations of Responsible Computing (FORC 2024)*, volume 295 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:22, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.FORC.2024.7.
- 20 Julia Chuzhoy, Merav Parter, and Zihan Tan. On packing low-diameter spanning trees. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, Saarbrücken, Germany (Virtual Conference), July 8–11, 2020*, volume 168 of *LIPIcs*, pages 33:1–33:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ICALP.2020.33.
- 21 Carmen Cirincione, Thomas A Darling, and Timothy G O’Rourke. Assessing South Carolina’s 1990s congressional districting. *Political Geography*, 19(2):189–211, 2000.
- 22 S Clarke and J Surkis. An operations research approach to racial desegregation of school systems. *Socio-Economic Planning Sciences*, 1(3):259–272, 1968.
- 23 Kenneth L. Clarkson. Applications of random sampling in computational geometry, ii. In *Proceedings of the Fourth Annual Symposium on Computational Geometry*, SCG ’88, pages 1–11, New York, NY, USA, 1988. Association for Computing Machinery. doi:10.1145/73393.73394.
- 24 Vincent Cohen-Addad, Philip N Klein, and Neal E Young. Balanced power diagrams for redistricting. *arXiv preprint*, 2017. arXiv:1710.03358.
- 25 Vincent Cohen-Addad, Philip N. Klein, and Neal E. Young. Balanced centroidal power diagrams for redistricting. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL ’18, pages 389–396, New York, NY, USA, 2018. Association for Computing Machinery. doi:10.1145/3274895.3274979.
- 26 Jessica De Silva, Brady Gales, Bryson Kagy, and David Offner. An analysis of a fair division protocol for drawing legislative districts. *arXiv preprint*, 2018. arXiv:1811.05705.
- 27 Daryl DeFord, Moon Duchin, and Justin Solomon. A computational approach to measuring vote elasticity and competitiveness. *Statistics and Public Policy*, 7(1):69–86, 2020.
- 28 Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A family of Markov chains for redistricting. *Harvard Data Science Review*, 3(1):3, 2021.
- 29 Erik D. Demaine and MohammadTaghi Hajiaghayi. Equivalence of local treewidth and linear local treewidth and its algorithmic applications. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA ’04, pages 840–849, USA, 2004. Society for Industrial and Applied Mathematics.
- 30 Erik D Demaine, MohammadTaghi Hajiaghayi, and Ken-ichi Kawarabayashi. Algorithmic graph minor theory: Decomposition, approximation, and coloring. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS’05)*, pages 637–646. IEEE, 2005.
- 31 Erik D. Demaine, Felix Reidl, Peter Rossmanith, Fernando Sánchez Villaamil, Somnath Sikdar, and Blair D. Sullivan. Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs. *Journal of Computer and System Sciences*, 105:199–241, 2019. doi:10.1016/j.jcss.2019.05.004.

- 32 Palash Dey, Sudeshna Kolay, and Sipra Singh. Knapsack: Connectedness, path, and shortest-path. In *Lecture Notes in Computer Science*, Lecture notes in computer science, pages 162–176, Cham, 2024. Springer Nature Switzerland. doi:10.1007/978-3-031-55601-2_11.
- 33 Prathamesh Dharangutte, Jie Gao, Shang-En Huang, and Fang-Yi Yu. Hardness and approximation algorithms for balanced districting problems. In Mark Bun, editor, *6th Symposium on Foundations of Responsible Computing, FORC 2025, Stanford University, CA, USA, June 4-6, 2025*, LIPIcs, pages 4:1–4:24. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025. doi:10.4230/LIPIcs.FORC.2025.4.
- 34 Riccardo Dondi and Manuel Lafond. On the tractability of covering a graph with 2-clubs. *Algorithmica*, 85(4):992–1028, 2023. doi:10.1007/S00453-022-01062-3.
- 35 Riccardo Dondi, Giancarlo Mauri, Florian Sikora, and Italo Zoppis. Covering a graph with clubs. *J. Graph Algorithms Appl.*, 23(2):271–292, 2019. doi:10.7155/JGAA.00491.
- 36 Riccardo Dondi, Giancarlo Mauri, and Italo Zoppis. On the tractability of finding disjoint clubs in a network. *Theoretical Computer Science*, 777:243–251, 2019. doi:10.1016/J.TCS.2019.03.045.
- 37 David Eppstein. Diameter and treewidth in minor-closed graph families. *Algorithmica*, 27(3):275–291, June 2000. doi:10.1007/S004530010020.
- 38 Allen D Franklin and Ernest Koenigsberg. Computed school assignments in a large district. *Operations Research*, 21(2):413–426, 1973. doi:10.1287/OPRE.21.2.413.
- 39 Roland G Fryer Jr and Richard Holden. Measuring the compactness of political districting plans. *The Journal of Law and Economics*, 54(3):493–535, 2011.
- 40 Nabeel Gillani, Doug Beeferman, Christine Vega-Pourheydarian, Cassandra Overney, Pascal Van Hentenryck, and Deb Roy. Redrawing attendance boundaries to promote racial and ethnic diversity in elementary schools. *Educational Researcher*, 52(6):348–364, 2023.
- 41 J Gerald Hebert. *The Realist’s Guide to Redistricting: Avoiding the Legal Pitfalls*. American Bar Association, 2010.
- 42 Pavol Hell and David Kirkpatrick. Packings by cliques and by finite families of graphs. *Discrete Math.*, 49(1):45–59, March 1984. doi:10.1016/0012-365X(84)90150-X.
- 43 Pavol Hell and David Kirkpatrick. Packings by complete bipartite graphs. *SIAM J. Algebr. Discrete Methods*, pages 199–209, 1986.
- 44 Pavol Hell, David Kirkpatrick, Jan Kratochvíl, and Igor Kříž. On restricted two-factors. *SIAM J. Discrete Math.*, pages 472–484, 1988.
- 45 Sidney Wayne Hess, JB Weaver, HJ Siegfeldt, JN Whelan, and PA Zitlau. Nonpartisan political redistricting by computer. *Operations Research*, 13(6):998–1006, 1965.
- 46 Ian Holyer. The NP-completeness of some edge-partition problems. *SIAM J. Comput.*, pages 713–717, 1981. doi:10.1137/0210054.
- 47 Chien-Chung Huang, Mathieu Mari, Claire Mathieu, Kevin Schewior, and Jens Vygen. An approximation algorithm for fully planar edge-disjoint paths. *SIAM J. Discrete Math.*, 35(2):752–769, January 2021. doi:10.1137/20M1319401.
- 48 Matt Jacobs and Olivia Walch. A partial differential equations approach to defeating partisan gerrymandering. *arXiv preprint*, 2018. arXiv:1806.07725.
- 49 Xinrui Jia, Kshiteej Sheth, and Ola Svensson. Fair colorful k -center clustering. In *Integer Programming and Combinatorial Optimization*, pages 209–222. Springer International Publishing, 2020. doi:10.1007/978-3-030-45771-6_17.
- 50 Hai Jin. *Spatial optimization methods and system for redistricting problems*. PhD thesis, University of South Carolina, 2017.
- 51 Jörg Kalcsics, Stefan Nickel, and Michael Schröder. Towards a unified territorial design approach—applications, algorithms and GIS integration. *Top*, 13:1–56, 2005.
- 52 Alexander K Kelmans. Optimal packing of induced stars in a graph. *Discrete Math.*, 173(1):97–127, 1997. doi:10.1016/S0012-365X(96)00121-5.
- 53 Myung Jin Kim. *Optimization approaches to political redistricting problems*. PhD thesis, The Ohio State University, 2011.

- 54 Zeph Landau, Oneil Reid, and Ilona Yershov. A fair division solution to the problem of redistricting. *Social Choice and Welfare*, 32(3):479–492, 2009. doi:10.1007/S00355-008-0336-6.
- 55 Shoham Letzter, Abhishek Methuku, and Benny Sudakov. Packing subgraphs in regular graphs, 2026. arXiv:2509.26180.
- 56 Mengya Li and Wensong Lin. On star family packing of graphs. *RAIRO Oper. Res.*, pages 2129–2140, 2021. doi:10.1051/R0/2021096.
- 57 Martin Loeb and Svatopluk Poljak. Efficient subgraph packing. *J. Combin. Theory Ser. B*, pages 106–121, 1993.
- 58 Angelo Monti and Blerina Sinimeri. Disjoint covering of bipartite graphs with s-clubs. In Rastislav Kráľovic and Vera Kurková, editors, *SOFSEM 2025: Theory and Practice of Computer Science - 50th International Conference on Current Trends in Theory and Practice of Computer Science, SOFSEM 2025, Bratislava, Slovak Republic, January 20-23, 2025, Proceedings, Part II*, volume 15539 of *Lecture Notes in Computer Science*, pages 198–210. Springer, 2025. doi:10.1007/978-3-031-82697-9_15.
- 59 Elle Najt, Daryl DeFord, and Justin Solomon. Complexity and geometry of sampling connected graph partitions. *arXiv preprint*, 2019. arXiv:1908.08881.
- 60 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion I. decompositions. *Eur. J. Comb.*, pages 760–776, 2008.
- 61 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion II. algorithmic aspects. *Eur. J. Comb.*, pages 777–791, 2008.
- 62 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion III. restricted graph homomorphism dualities. *Eur. J. Comb.*, pages 1012–1024, 2008.
- 63 Qi Ning. On the star packing problem. *Ann. N. Y. Acad. Sci.*, 576(1):411–416, December 1989.
- 64 Wesley Pegden, Ariel D Procaccia, and Dingli Yu. A partisan districting protocol with provably nonpartisan outcomes. *arXiv preprint*, 2017. arXiv:1710.08781.
- 65 Andrew Thomason. The extremal function for complete minors. *J. Comb. Theory B*, 81(2):318–338, 2001. doi:10.1006/JCTB.2000.2013.
- 66 Santiago Valdés Ravelo and Flávio K. Miyazawa. Packing sets of paths, stars and triangles: Tractability and approximability. In Daya Gaur and Rogers Mathew, editors, *Algorithms and Discrete Applied Mathematics*, pages 247–257, Cham, 2025. Springer Nature Switzerland.
- 67 James B Weaver and Sidney W Hess. A procedure for nonpartisan districting: Development of computer techniques. *Yale Law Journal*, 73(2):288–308, 1963.
- 68 Wenying Xi, Wensong Lin, and Yuquan Lin. Packing 2- and 3-stars into cubic graphs. *Appl. Math. Comput.*, 460(128287):128287, January 2024. doi:10.1016/J.AMC.2023.128287.
- 69 Jack Zhang, Lucas Silveira, Hamidreza Validi, Logan Smith, Austin Buchanan, and Illya V Hicks. Partitioning a graph into low-diameter clusters, 2024.

A Separation Oracle for Districts of Bounded Radius

Here, we outline how to implement separation oracles to search for a weakly violating connected subgraph. The main idea is to formulate the problem as a *vector-valued connected subgraph sum* problem where each vertex has a vector that captures its objective weight, features, and dual variables. The proofs are deferred to the full version.

Given an undirected graph $G = (V, E)$ where each node $v \in V$ has a vector-valued weight $\mathbf{w}(v) \in \mathbb{Z}_{\geq 0}^d$, the *Connected Subgraph Sum* problem defines the set W^* of all possible total weights achievable by a connected and compact subgraphs of G :

$$W^* = \left\{ \mathbf{w}(T) = \sum_{v \in T} \mathbf{w}(v) : \text{the induced subgraph } G[T] \text{ is connected and compact} \right\}. \quad (5)$$

We define the set of feasible weight vectors W^* only based on the topological constraints (connectivity and compactness). The composition constraints (e.g., balancedness or thresholds) are encoded in the vector values and handled subsequently via the linear score function ℓ and the domination idea.

As the set W^* can be exponentially large, explicitly computing and verifying composition constraints in Equation (5) is infeasible. Instead, we propose the (ℓ, ε) -Trimmed Connected Subgraph Sum problem. This seeks a succinct subset that approximates every feasible weight vector multiplicatively, while maximizing a specific linear score ℓ , which encodes the necessary information to satisfy the composition requirements.

(ℓ, ε) -TRIMMED CONNECTED SUBGRAPH SUM

Input: A graph $G = (V, E)$ with a vector-valued weight function $\mathbf{w} : V \rightarrow \mathbb{Z}_{\geq 0}^d$, an error parameter $\varepsilon \geq 0$, and a linear function $\ell : \mathbb{Z}_{\geq 0}^d \rightarrow \mathbb{Z}$.

Goal: Compute a subset $W \subseteq \mathbb{Z}_{\geq 0}^d$ such that $W \subseteq W^*$ and W is an (ℓ, ε) -trimming of W^* . That is, for every $\mathbf{z} = (z_1, \dots, z_d) \in W^*$, there exists a representative $\mathbf{z}' = (z'_1, \dots, z'_d) \in W$ satisfying:

$$e^{-\varepsilon} \leq \frac{z'_i}{z_i} \leq e^{\varepsilon} \quad \text{for all } i = 1, \dots, d, \text{ and} \quad (6)$$

$$\ell(\mathbf{z}') \geq \ell(\mathbf{z}). \quad (7)$$

We say that $\mathbf{z}'$ ε -approximates $\mathbf{z}$ if condition (6) holds, and that $\mathbf{z}'$ ℓ -dominates $\mathbf{z}$ if condition (7) holds. Crucially, this problem is computationally tractable on graphs of bounded treewidth. The main result of this section is an FPTAS for the trimmed connected subgraph sum problem, which serves as the tool for our separation oracles.

► **Lemma 18.** *Given $G = (V, E)$, a radius bound k , vector-valued weights $\mathbf{w} : V \rightarrow \mathbb{Z}_{\geq 0}^d$, an error parameter $\varepsilon > 0$, and a linear function ℓ , there is an algorithm for (ℓ, ε) -trimmed connected subgraph sum with strong radius bound that has running time*

$$O\left(n^{2+2d} \varepsilon^{-2d} (k+1)^{2(tw+1)} 2^{O(tw \log tw)} (\ln R)^{2d}\right),$$

where n is the number of vertices and tw is the treewidth of G , and $R = \max_i \sum_{v \in V} w_i(v)$.

Armed with this efficient algorithm, we now demonstrate how the separation oracles for our districting problems can be reduced to instances of the trimmed connected subgraph sum problem.

► **Lemma 19** (Separation Oracle for c -Balanced Districts). *Consider the c -balanced districting problem where each vertex v has an objective weight $w(v)$ and feature weights $w_1(v), w_2(v)$. The separation oracle seeks a district $S_{\max}$ maximizing objective weight subject to the balance condition and a violated dual constraint so that*

$$S_{\max} \in \arg \max \{w(S') : y(S') < (1 - \varepsilon)w(S') \text{ and } S' \text{ is } c\text{-balanced}\}.$$

This can be reduced to solving two instances of the trimmed connected subgraph sum problem. We define the vector-valued weight as $\mathbf{w} = (w_1, w_2, w, y)$ and use the linear functions:

$$\ell_1(\mathbf{z}) = (c-1)z_1 - z_2 \quad \text{and} \quad \ell_2(\mathbf{z}) = (c-1)z_2 - z_1$$

By finding an (ℓ_1, ε) -trimmed and (ℓ_2, ε) -trimmed subset, we can identify a weakly violating district S that satisfies $y(S) < (1 - \frac{\varepsilon}{2})w(S)$, approximates the optimal weight $w(S) \geq \frac{1}{2}w(S_{\max})$, and satisfies the balance condition Equation (1).

► **Lemma 20** (Separation Oracle for Threshold Districts). *Consider the B -threshold districting problem with objective weight $w(v)$ and feature weight $w_1(v)$ subject to $w_1(S) \geq B$. The separation oracle additionally has a dual variable y and seeks a district maximizing objective weight subject to the threshold and dual constraints.*

We can reduce this to solving the trimmed connected subgraph sum problem with vector weights $\mathbf{w} = (w_1, w, y)$ and the linear function $\ell_{th}(\mathbf{z}) = z_1$. The resulting set allows us to compute a weakly violating district satisfying $y(S) < (1 - \frac{1}{2}\varepsilon)w(S)$, approximate objective weight $w(S) \geq \frac{1}{2}w(S_{\max})$, and the B -threshold constraint.

► **Example 21** (Connected Knapsack Problem [32]). Given a graph G , where each vertex has a value $\alpha(v)$ and a cost $\beta(v)$, and a global budget B , the goal is to find a connected subgraph U maximizing $\alpha(U)$ subject to $\beta(U) \leq B$. This is exactly recovered by setting $\mathbf{w}(v) = (\alpha(v), \beta(v))$ and solving for the appropriate approximation with $\ell(\alpha, \beta) = -\beta$.

Proofs to Theorem 4 and Theorem 5

Finally, we conclude this section by describing how to apply separation oracles provided in the previous section. In particular, the new separation oracle leads to $O(1)$ -approximation algorithms for bounded strong radius districts on apex-minor-free graphs.

Proof of Theorem 4. To apply Lemma 18 (with the reduction adapted from Lemma 19 on $d = 4$ and Lemma 20 on $d = 3$) in the separation oracle, the algorithm enumerates all vertices $v \in V$ as centers and considers the subgraph induced by vertices within distance k of v . As G is apex-minor-free, each such subgraph has bounded treewidth, which makes Lemma 18 applicable, and therefore the theorem follows. ◀

Proof of Theorem 5. We adopt the implementation of the separation oracle from [33], enumerating each vertex as a star center and invoking an approximate knapsack algorithm. The theorem follows from combining the analysis of Section 4.3. ◀

B PTAS for δ -Relaxed Districting on Apex-Minor-Free Graphs

In this section, we show that allowing relaxation to the composition requirements helps us obtain PTAS to the districting problem on bounded treewidth graphs. Combining this with the standard Baker's approach [7], we obtain PTAS for δ -relaxed districting on planar and apex minor-free graphs.

► **Theorem 22.** *There exists a polynomial-time algorithm that computes an $(1 + \varepsilon)$ -approximate solution for packing δ -relaxed c -balanced or threshold districts with bounded strong or weak radius on planar graphs and apex-minor-free graphs.*

We first state the result that obtains a $(1 + \varepsilon)$ -approximate solution for packing δ -relaxed districts on bounded treewidth graphs. The proof employs dynamic programming, combining Lemma 18 with the techniques of [33] (see Appendix B in full version).

► **Lemma 23.** *There exists a polynomial-time algorithm that computes an $(1 + \varepsilon)$ -approximate solution for packing δ -relaxed c -balanced or B -threshold districts with bounded strong or weak radius on graphs of bounded treewidth.*

Using Lemma 23 as a black box, we prove the main result for δ -relaxed districting.

Proof of Theorem 22. The algorithm is as follows. Start with an arbitrary vertex r as the root and construct a BFS tree and let L_j be the j^{th} level which are vertices at hop distance j from r . Now, for an integer $t > 0$, and shift $i \in \{0, \dots, t-1\}$, remove vertices in level $i \bmod t$, and let $G^{(i)}$ be the induced graph on remaining vertices. Notice that each connected component in $G^{(i)}$ is fully contained in some $(t-1)$ consecutive levels of the BFS tree. For planar or apex-minor-free G , note that each connected component in $G^{(i)}$ now has bounded treewidth [37] (specifically, $O(t)$). Use the algorithm from Lemma 23 with parameter $\varepsilon/2$ and solve the problem on each connected component separately and let the union of these solutions for a fixed i be $\mathcal{T}_i$. Since districts cannot use vertices from different levels, taking union gives us a valid districting. Return the solution with largest weight $w(\mathcal{T}_i)$ across all i .

Now, we show that setting $t = O(k/\varepsilon)$ suffices. Let $\mathcal{T}^*$ be the optimal solution on G with $w(\mathcal{T}^*) = \text{OPT}$. A radius k district in $\mathcal{T}^*$ is valid in some connected component in $G^{(i)}$ if no vertex in the district (or the path considered by the optimal solution for weak radius setting) lies in the deleted levels. Hence, for at most $2k+1$ different shifts (i values), this particular district (or the path from the center vertex to some vertex in the district) is not valid. Now, for shift i , let S_i be the set of districts from $\mathcal{T}^*$ that intersect with some deleted layer. The optimal districting in $G^{(i)}$ has weight $\geq \text{OPT} - w(S_i)$. Now consider the total weight lost across different shifts. We have

$$\sum_{i=0}^{t-1} w(S_i) \leq (2k+1) \sum_{T \in \mathcal{T}^*} w(T) \leq (2k+1)\text{OPT}.$$

By pigeonhole principle, for $t = 2(2k+1)/\varepsilon$, there is some i^* such that

$$w(S_{i^*}) \leq \frac{2k+1}{t} \text{OPT} \leq \frac{\varepsilon}{2} \text{OPT}$$

As a result, for this particular shift i^* , the optimal districting in $G^{(i^*)}$ has weight at least $\text{OPT} - w(S_{i^*}) \geq (1 - \frac{\varepsilon}{2}) \text{OPT}$. Combining this with $(1 - \varepsilon/2)$ approximation on $G^{(i^*)}$ (Lemma 23), we have that the solution has weight at least $(1 - \frac{\varepsilon}{2})^2 \text{OPT} \geq (1 - \varepsilon) \text{OPT}$. ◀

C Hardness Results

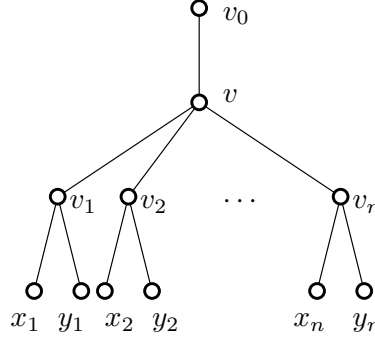
In this section we prove our hardness results. Note that the hardness for maximum packing of c -balanced districts were shown in [33]. Here, we show the hardness of packing under the weight threshold condition.

Hardness of Packing Compact Subgraphs with Weight Threshold

► **Theorem 24.** *The B -threshold districting problem with (weak or strong) radius- k bounded districts is NP-hard when G is a tree.*

Proof. We first focus on star districts and show a reduction from the Knapsack Problem. Given a knapsack problem with n items, where item i has weight w_i and utility u_i , we want to select items such that utility is at least U , and minimize the total weight of the selected one. This is the dual version of the standard knapsack problem which fixes the cap of the total weight and maximizes utility, and is still NP-hard. Without loss of generality, we can assume that $u_i > 1$ and $w_i < 1/n$. This can be achieved by scaling up all utility values and the utility goal U and scale down the weights. It does not change the problem structure.

Now we describe the instance for the districting problem. Consider a tree rooted at v with children of v as $v_1, v_2, \dots, v_n$ corresponding to the items. Let $w(v) = 1$, and $B = 2U$. Node v_i has two children x_i and y_i . Define $w(x_i) = B$, $w(y_i) = w_i$, $w(v_i) = u_i$. Further, we add a separate neighbor/child v_0 of v with $w(v_0) = U - 1$. This is detailed in Figure 1.



■ **Figure 1** Instance for districting on the tree from a given knapsack instance.

Now we consider what are the possible candidate districts that have a total weight at least B . First we can define a district centered at v , including v_0 and some set of vertices in $\{v_1, \dots, v_n\}$. The population of v and v_0 is $U = B/2$, so we must include additional vertices from $\{v_i\}$ to gain additional population of at least U . On the other hand, we can also define a district centered at v_i , which includes x_i, y_i (and possibly v). However, we cannot have both a district centered at v including v_i and a district centered at v_i , as we want disjoint districts. If v_i is included in the district at v , then we can still define a district with a solo vertex x_i (of population B), but y_i cannot be included in any district.

Thus the main decision to make is whether we can include a district centered at v or not. To do so, we select a subset of vertices in $\{v_1, \dots, v_n\}$ which has a total weight at least U . For these ones selected, the corresponding vertices y_i will not be included anywhere – causing a loss of w_i from the potential districts centered at v_i . However, since $\sum_i w_i < 1$, it would always be more beneficial to create a district at v whenever possible. Specifically, the total coverage of the two cases is

- If we do not have a district in v , the optimal solution is $1 + nB + \sum_i u_i + \sum_i w_i$. In the above, the 1 comes from some district at v_i that can include v .
- If we do include a district in v , which includes v_0 and a subset of vertices $S \subseteq \{v_1, \dots, v_n\}$, then the solution is $B/2 + nB + \sum_i u_i + \sum_i w_i - \sum_{i \in S} w_i$.

In other words, the best solution is to try to include a district at v and the set S of children included identifies the items that have the smallest total weight yet have utility to be at least U . Notice that this is precisely the optimal solution for the knapsack problem. Next, we note that the same reduction extends to radius- k districts, for any fixed $k \geq 1$, by replacing every edge of the above tree by a path of length exactly k , and assigning weight 0 to all newly introduced subdivision vertices. The resulting graph is still a tree, and all original districts retain the same weights. The same proof goes through by noting that on a tree, the weak-radius and strong-radius notions are equivalent for connected districts. ◀

Next, we show hardness for general graphs.

► **Theorem 25.** *The B -threshold districting problem with (weak or strong) radius- k bounded districts is NP-hard when G is a bounded degree graph or a planar graph. Moreover, on general graphs, it is NP-hard to approximate better than $O(n^{1/2-\delta})$ for any $\delta > 0$. The results hold even when each subgraph is required to be a star (radius 1).*

Proof. We show hardness and hardness of approximation for $k = 1$ by using the same reduction from maximum independent set as in [33]. Specifically, given a graph $G = (V, E)$, the maximum independent set problem finds a subset of vertices $S \subseteq V$ with maximum

cardinality such that no vertices in S have edges between them. We now construct a graph G' that starts with G . For each edge (u, v) in G , split it into two edges in G' – create a new vertex p_{uv} which connects to u and v and remove the edge (u, v) . In addition, suppose the maximum degree in G is Δ . For a vertex v in G' (corresponding to a vertex in G) whose degree $d(v)$ is less than Δ , create $\Delta - d(v)$ additional neighbors (called filler nodes). We then set $w(v) = 1$ for each vertex in G' , $B = \Delta + 1$ and $k = 1$. Without loss of generality, we can assume that $\Delta \geq 3$ and thus $B \geq 4$. This means that any district with weight at least B in G' must be at a vertex which corresponds to vertices in G . Further, this district will need to use all Δ neighbors in order to have enough total weight. This means that any two districts in G' cannot have their centers to be adjacent in G . Thus any independent set S in G can be translated to a collection star districts in G' , with each district centered on a vertex in S with all its Δ neighbors. Similarly, any districting solution will produce a set of vertices in V that are independent. The total weight of the optimal districting problem of G' is $(\Delta + 1)|S|$ where S is the maximum independent set in G . Similar to the argument in [33], since maximum independent set cannot be approximated by a factor of $n^{1-\delta}$ for any $\delta > 0$, this means that our districting problem cannot be approximated better than $n^{1/2-\delta}$ in a general graph for any $\delta > 0$, unless $P = NP$. Further, it is NP-hard to solve the districting problem for graphs of bounded degree or on planar graphs, since the maximum independent set for these graphs remains NP-hard.

To extend the hardness for radius- k districts, we simply replace the edges we add in constructing G' with paths of length k . So, instead of connecting u and p_{uv} with a single edge, we connect them with a path of length k (i.e., with $k - 1$ new vertices on the path from u to p_{uv}). We do the same for v and p_{uv} and for each vertex v with their respective $\Delta - d(v)$ filler neighbors. We set $B = 1 + k\Delta$. A valid district must be centered at a vertex v corresponding to vertices in the original graph G and include Δ paths from v , each of length k . Following the same arguments completes the proof. ◀

► **Remark 26.** The reduction in Theorem 25 also works for the relaxed version of B -threshold districting. This is because for fixed δ and radius- k districts, we can always choose B such that $\Delta < (1 - \delta)B \leq \Delta + 1$, which forces the B -threshold districts to be the exact same as described in the proof.