Approximation Schemes for Geometric Coverage Problems

Steven Chaplick
Lehrstuhl für Informatik I, Universität Würzburg, Germany
steven.chaplick@uni-wuerzburg.de
https://orcid.org/0000-0003-3501-4608

Minati De[1]
Department of Mathematics, Indian Institute of Technology Delhi, India
minati@maths.iitd.ac.in
https://orcid.org/0000-0002-1859-8800

Alexander Ravsky
Pidstryhach Institute for Applied Problems of Mechanics and Mathematics,
National Academy of Science of Ukraine, Lviv, Ukraine
alexander.ravsky@uni-wuerzburg.de

Joachim Spoerhase[2]
Lehrstuhl für Informatik I, Universität Würzburg, Germany; and
Institute of Computer Science, University of Wrocław, Poland
joachim.spoerhase@uni-wuerzburg
https://orcid.org/0000-0002-2601-6452

Abstract

In their seminal work, Mustafa and Ray [30] showed that a wide class of geometric *set cover (SC)* problems admit a PTAS via local search – this is one of the most general approaches known for such problems. Their result applies if a naturally defined "exchange graph" for two feasible solutions is planar and is based on subdividing this graph via a planar separator theorem due to Frederickson [17]. Obtaining similar results for the related *maximum coverage problem (MC)* seems non-trivial due to the hard cardinality constraint. In fact, while Badanidiyuru, Kleinberg, and Lee [4] have shown (via a different analysis) that local search yields a PTAS for two-dimensional real halfspaces, they only conjectured that the same holds true for dimension three. Interestingly, at this point it was already known that local search provides a PTAS for the corresponding set cover case and this followed directly from the approach of Mustafa and Ray.

In this work we provide a way to address the above-mentioned issue. First, we propose a *color-balanced* version of the planar separator theorem. The resulting subdivision approximates locally in each part the global distribution of the colors. Second, we show how this roughly balanced subdivision can be employed in a more careful analysis to strictly obey the hard cardinality constraint. More specifically, we obtain a PTAS for any "planarizable" instance of MC and thus essentially for all cases where the corresponding SC instance can be tackled via the approach of Mustafa and Ray. As a corollary, we confirm the conjecture of Badanidiyuru, Kleinberg, and Lee [4] regarding real halfspaces in dimension three. We feel that our ideas could also be helpful in other geometric settings involving a cardinality constraint.

2012 ACM Subject Classification Theory of computation → Packing and covering problems

Keywords and phrases balanced separators, maximum coverage, local search, approximation scheme, geometric approximation algorithms

[1] Partially supported by DST-INSPIRE Faculty Grant (DST-IFA-14-ENG-75)
[2] Partially supported by Polish National Science Centre grant 2015/18/E/ST6/00456

26th Annual European Symposium on Algorithms (ESA 2018).
Editors: Yossi Azar, Hannah Bast, and Grzegorz Herman; Article No. 17; pp. 17:1–17:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Digital Object Identifier 10.4230/LIPIcs.ESA.2018.17

Related Version A full version is available at `https://arxiv.org/abs/1607.06665`.

1 Introduction

The Maximum Coverage (MC) problem is one of the classic combinatorial optimization problems which is well studied due to its wealth of applications. Let U be a set of ground elements, $\mathcal{F} \subseteq 2^U$ be a family of subsets of U and k be a positive integer. The Maximum Coverage (MC) problem asks for a k-subset $\mathcal{F}'$ of $\mathcal{F}$ such that the number $|\bigcup \mathcal{F}'|$ of ground elements covered by $\mathcal{F}'$ is maximized.

Many real life problems arising from banking [13], social networks, transportation network [28], databases [22], information retrieval, sensor placement, security (and others) can be framed as an instance of MC problem. For example, the following are easily seen as MC problems: placing k sensors to maximize the number of covered customers, finding a set of k documents satisfying the information needs of as many users as possible [4], and placing k security personnel in a terrain to maximize the number of secured regions.

From the result of Cornuéjols [13], it is well known that the greedy algorithm is a $1-1/e$ approximation algorithm for the MC problem. Due to wide applicability of the problem, whether one can achieve an approximation factor better than $1-1/e$ was subject of research for a long period of time. From the result of Feige [16], it is known that if there exists a polynomial-time algorithm that approximates maximum coverage within a ratio of $1-1/e+\epsilon$ for some $\epsilon > 0$ then P = NP. Better results can, however, be obtained for special cases of MC. For example, Ageev and Sviridenko [1] show in their seminal work that their pipage rounding approach gives a factor $1-(1-1/r)^r$ for instances of MC where every element occurs in at most r sets. For constant r this is a strict improvement on $1-1/e$ but this bound is approached if r is unbounded. For example, pipage rounding gives a 3/4-approximation algorithm for Maximum Vertex Cover (MVC), which asks for a k-subset of nodes of a given graph that maximizes the number of edges incident on at least one of the selected nodes. Petrank [31] showed that this special case of MC is APX-hard.

In this paper, we study the approximability of MC in *geometric* settings where elements and sets are represented by geometric objects. Such problems have been considered before and have applications, for example, in information retrieval [4] and in wireless networks [15].

MC is related to the Set Cover problem (SC). For a given set U of ground elements and a family $\mathcal{F} \subseteq 2^U$ of subsets of U, this problem asks for a minimum cardinality subset of $\mathcal{F}$ which covers all the ground elements of U. This problem plays a central role in combinatorial optimization and in particular in the study of approximation algorithms. The best known approximation algorithm has a ratio of $\ln n$, which is essentially the best possible [16] under a plausible complexity-theoretic assumption. A lot of work has been devoted to beating the logarithmic barrier in the context of geometric set cover problems [6, 33, 8, 29]. Mustafa and Ray [30] introduced a powerful tool which can be used to show that a *local search* approach provides a PTAS for various geometric SC problems. Their result applies if a naturally defined *exchange graph* (whose nodes are the sets in two feasible solutions) is planar and is based on subdividing this graph via a planar separator theorem due to Frederickson [17]. In the same paper [30], they applied this approach to provide a PTAS for the SC problem when the family $\mathcal{F}$ consists of either a set of halfspaces in $\mathbb{R}^3$, or a set of disks in $\mathbb{R}^2$. Concurrently with Mustafa and Ray [30], Chan and Har-Peled [9] also introduced the exchange graph concept, but for the independent set problem. Subsequently, many results have been obtained using this technique for problems in geometric settings [10, 14, 19, 26]. Some of these works extend to cases where the exchange graph is not planar but admits a small-size separator [3, 7, 20, 21].

Beyond the context of SC, local search has also turned out to be a very powerful tool for other geometric problems but the analysis of such algorithms is usually non-trivial and highly tailored to the specific setting. Examples of such problems are Euclidean TSP, Euclidean Steiner tree, facility location, k-median [12]. In some recent breakthroughs, PTASs for the k-means problem in finite Euclidean dimension (and more general cases) via local search have been published [11, 18].

In this paper, we study the effectiveness of local search for geometric MC problems. In the general case, b-swap local search is known to yield a tight approximation ratio of $1/2$ [24]. However, for special cases such as geometric MC problems local search is a promising candidate for beating the barrier $1 - 1/e$. It seems, however, non-trivial to obtain such results using the technique of Mustafa and Ray [30]. In their analysis, each part of the subdivided planar exchange graph (see above) corresponds to a feasible candidate swap that replaces some sets of the local optimum with some sets of the global optimum and it is ensured that every element stays covered due to the construction of the exchange graph. It is moreover argued that if the global optimum is sufficiently smaller than the local optimum then one of the considered candidate swaps would actually reduce the size of the solution.

It is possible to construct the same exchange graphs also for the case of MC. However, the hard cardinality constraint given by input parameter k poses an obstacle. In particular, when considering a swap corresponding to a part of the subdivision, this swap might be infeasible as it may contain (substantially) more sets from the global optimum than from the local optimum. Another issue is that MC has a different objective function than SC. Namely, the goal is to maximize the number of covered elements rather than minimizing the number of used sets. Finally, while for SC all elements are covered by both solutions, in MC we additionally have elements that are covered by none or only one of the two solutions requiring a more detailed distinction of several types of elements.

In fact, subsequent to the work of Mustafa and Ray on SC [30], Badanidiyuru, Kleinberg, and Lee [4] studied geometric MC. They obtained fixed-parameter approximation schemes for MC instances for the very general case where the family $\mathcal{F}$ consists of objects with bounded VC dimension, but the running times are exponential in the cardinality bound k. They further provided APX-hardness for each of the following cases: set systems of VC-dimension 2, halfspaces in $\mathbb{R}^4$, and axis-parallel rectangles in $\mathbb{R}^2$. Interestingly, while they have shown that for MC instances where $\mathcal{F}$ consists of halfspaces in $\mathbb{R}^2$ local search can be used to provide a PTAS, they only conjecture that local search provides a PTAS when $\mathcal{F}$ consists of halfspaces in $\mathbb{R}^3$. This underlines the observation that it seems non-trivial to apply the approach of Mustafa and Ray to geometric MC problems as at that point a PTAS for halfspaces in $\mathbb{R}^3$ for SC was already known via the approach of Mustafa and Ray.

The difficulty of analyzing local search under the presence of a cardinality constraint is also known in other settings. For example, one of the main technical contributions of the recent breakthrough for the Euclidean k-means problem [11, 18] is that the authors are able to handle the hard cardinality constraint by the concept of so-called isolated pairs [11]. Prior to these works approximation schemes have only been known for bicriteria variants where the cardinality constraint may be violated or where there is no constraint but – analogously to SC – the cardinality contributes to the objective function [5].

1.1 Our Contribution

We show a way to cope with the above-mentioned issue with a cardinality constraint. We are able to achieve a PTAS for many geometric MC problems using the b-swap local search approach given in Algorithm 1. At a high level we follow the framework of Mustafa and Ray defining a planar (or more generally *f-separable* as formalized below) exchange graph and subdividing it into a number of small parts each of them corresponding to a candidate swap. As each part may be (substantially) imbalanced in terms of the number of sets of the global optimum and local optimum, respectively, a natural idea is to swap-in only a sufficiently small subset of the globally optimal sets. This idea alone is, however, not sufficient. Consider, for example, the case where each part contains either only sets from the local or only sets from the global optimum making it impossible to retrieve any feasible swap from considering the single parts. To overcome this difficulty, in Section 2, we prove in a first step a *color-balanced* version (Theorem 7) of the planar division theorem (Theorem 4 [17]). In this theorem, the input is a planar (or more generally *f-separable*) graph whose nodes are two-colored arbitrarily. The distinctions of our division theorem from the prior work, are that our division theorem guarantees that all parts have roughly the same size (rather than simply an upper limit on their size) and that the two colors are represented in each part in roughly the same ratio as in the whole graph. This balancing property allows us to address the issue of the above-mentioned infeasible swaps. In a second step, described in Section 3, we employ this roughly color-balanced subdivision to establish a set of perfectly balanced candidate swaps. We prove by a careful analysis (which turns out more intricate than for the SC case) that local search also yields a PTAS for the wide class *f-separable* MC problems (see Theorem 2). As a direct consequence, we obtain PTASs for essentially all cases of geometric MC problems (see (Theorem 3)) where the corresponding SC problem can be tackled via the approach of Mustafa and Ray. For example, this confirms the conjecture of Badanidiyuru, Kleinberg, and Lee [4] regarding halfspaces in $\mathbb{R}^3$. We also obtain PTASs for Maximum Dominating Set and Maximum Vertex Cover on f-separable and minor-closed graph classes (as formalized below) which, to the best of our knowledge, were not known before. We feel that our approach has the potential for further applications in similar cardinality constrained settings.

In the remainder of this section we formalize our main theorem (providing the needed definitions) and then present a set of applications of this theorem.

Definitions and the main theorem. For a number n, $[n]$ denotes the set $\{1, \ldots, n\}$. For a graph G, a subset S of $V(G)$ is an *α-balanced separator* when its removal breaks G into two collections of connected components such that each collection contains at most an α fraction of $V(G)$ where $\alpha \in [\frac{1}{2}, 1)$ and α is a constant. The size of a separator S is simply the number of vertices it contains. Let f be a non-decreasing (strictly) sublinear function, that is, $f(x) = O(x^{1-\delta})$ for some $\delta > 0$. A class of graphs that is closed under taking subgraphs is said to be *f-separable* if there is an $\alpha \in [\frac{1}{2}, 1)$ such that for any $n > 2$, an n-vertex graph in the class has an α-balanced separator whose size is at most $f(n)$. In what follows, whenever we discuss an f-separable graph classes we implicitly assume that the function f is non-decreasing and has the form $f(x) = O(x^{1-\delta})$ for some $\delta > 0$ – this is what we mean by *strictly* sublinear. Note that, by the Lipton-Tarjan separator theorem [27], planar graphs are a subclass of the $\sqrt{n}$-separable graphs. More generally, Alon, Seymour, and Thomas [2] have shown that every graph class characterized by a finite set of forbidden minors is also a subclass of the $(c \cdot \sqrt{n})$-separable graphs (here, the constant c depends on the size of the largest forbidden minor). In particular, from the graph minors theorem [32], every non-trivial

Algorithm 1: b-swap local search on an MC instance with ground set U, set family $\mathcal{S}$, and parameter k.

b-**LocalSearch**$(U, \mathcal{S}, k)$
$\mathcal{F} \leftarrow$ initialize with a greedy solution
while $\exists\ \mathcal{F}' \subseteq \mathcal{F}$, $\exists \mathcal{F}^* \subset \mathcal{S}$, *such that* $b = |\mathcal{F}^*| = |\mathcal{F}'|$ *and* $|\bigcup \mathcal{F}| < |\bigcup((\mathcal{F} \setminus \mathcal{F}') \cup \mathcal{F}^*)|$ **do**
  perform the swap, i.e., $\mathcal{F} \leftarrow (\mathcal{F} \setminus \mathcal{F}') \cup \mathcal{F}^*$
return $\mathcal{F}$

minor-closed graph class is a subclass of the $(c \cdot \sqrt{n})$-separable graphs (for some constant c). With this notion in mind, we define the concept of f-separable MC instances, then state our main theorem (the proof is given in Section 3).

▶ **Definition 1.** A class $\mathcal{C}$ of instances of MC is *f-separable* (or in particular *planarizable*) if for any two disjoint feasible solutions $\mathcal{F}$ and $\mathcal{F}'$ of any instance in $\mathcal{C}$ there exists an f-separable (planar) graph G with node set $\mathcal{F} \cup \mathcal{F}'$ with the following *exchange* property. For each element $u \in U$ that is covered both by $\mathcal{F}$ and $\mathcal{F}'$, there is an edge (S, S') in G with $S \in \mathcal{F}$ and $S' \in \mathcal{F}'$ with $u \in S \cap S'$.

▶ **Theorem 2.** *For any non-decreasing strictly sublinear function f, every f-separable class of MC instances (closed under removing elements and sets) admits a PTAS via Algorithm 1.*

Applications. We now describe several problems which are special instances of the MC problem. Then, in Theorem 3, we state several PTASs for each of these problems that can be obtained from our analysis of local search. As this latter part is essentially a direct consequence of Theorem 2, the details will be provided in the full version.

▶ **Problem 1.** *Let H be a set of ground elements, $\mathcal{S} \subseteq 2^H$ be a set of ranges and k be a positive integer. A range $S \in \mathcal{S}$ is hit by a subset H' of H if $S \cap H' \neq \emptyset$. The* Maximum Hitting *(MH) problem asks for a k-subset H' of H such that the number of ranges hit by H' is maximized.*

▶ **Problem 2.** *Let $G = (V, E)$ be a graph and k be a positive integer. A vertex $v \in V$ dominates itself and all its neighbors. The* Maximum Dominating *(MD) problem asks for a k-subset V' of V such that the number of vertices dominated by V' is maximized.*

▶ **Problem 3.** *Let T be a* 1.5D *terrain – an x-monotone polygonal chain in $\mathbb{R}^2$ consisting of a set of vertices $\{v_1, v_2, \ldots, v_m\}$ sorted in increasing order of their x-coordinate, and v_i and v_{i+1} are connected by an edge for all $i \in [m-1]$. For any two points $x, y \in T$, we say that y guards x if each point in $\overline{xy}$ lies above or on the terrain. Given finite sets $X, Y \subseteq T$ and a positive integer k, the* Maximum Terrain Guarding *(MTG) problem asks for a k-subset Y' of Y such that the number of points of X guarded by Y' is maximized.*

Let r be an even, positive integer. A set of regions in $\mathbb{R}^2$ (each bounded by a closed Jordan curve), is called *r-admissible* if for any two such regions q_1, q_2, the curves bounding them cross $s \leq r$ times for some even s and $q_1 \setminus q_2$ and $q_2 \setminus q_1$ are connected regions. A set of regions are called *pseudo-disks* if it is 2-admissible (e.g., a set of disks or squares).

The below consequences of Theorem 2 hold since the corresponding SC problem is known to be *planarizable* or by constructing the exchange graph as a minor of the input graph.

▶ **Theorem 3.** *Local search gives a PTAS for:*

(V) *the MVC problem on f-separable and subgraph-closed graph classes,*
(T) *the MTG problem.*

and the following classes of MC problems:
(C_1) *the set of ground elements is a set of points in $\mathbb{R}^3$, and the family of subsets is induced by a set of halfspaces in $\mathbb{R}^3$.*
(C_2) *the set of ground elements is a set of points in $\mathbb{R}^2$, and the family of subsets is induced by a set of convex pseudodisks (a set of convex objects where any two objects can have at most two intersections in their boundary).*

and the following MH problems:
(H_1) *the set of ground elements is a set of points in $\mathbb{R}^2$, and the set of ranges is induced by a set of r-admissible regions (this includes pseudodisks, same-height axis-parallel rectangles, circular disks, translates of convex objects).*
(H_2) *the set of ground elements is a set of points in $\mathbb{R}^3$, and the set of ranges is induced by a set of halfspaces in $\mathbb{R}^3$.*

and MD problems in each of the following graph classes:
(D_1) *intersection graphs of homothetic copies of convex objects (which includes arbitrary squares, regular k-gons, translated and scaled copies of a convex object).*
(D_2) *non-trivial minor-closed graph classes.*

2 Color-Balanced Divisions

In this section we provide the main tool (see Theorem 7) used to prove our main result (i.e., Theorem 2). We first describe a new subtle strengthening (see Lemma 5) of the standard division theorem on f-separable graph classes (see Theorem 4). This builds on the concept of $(r, f(r))$-divisions (in the sense of Henzinger et al. [23]) of graphs in an f-separable graph class. We then extend this strengthened division lemma by suitably aggregating the pieces of the partition to obtain a *two-color balanced* version (see Theorem 7). This result generalizes to more than two colors. However, as our applications stem from the two-colored version, we defer the generalization to the full version.

Fredericksonn [17] introduced the notion of an r-*division* of an n-vertex graph G, namely, a cover of $V(G)$ by $\Theta(\frac{n}{r})$ sets each of size $O(r)$ where each set has $O(\sqrt{r})$ *boundary* vertices, i.e., $O(\sqrt{r})$ vertices in common with the other sets. Frederickson showed that, for any r, every planar graph G has an r-division and that one can be computed in $O(n \log n)$ time. This result follows from a recursive application of the Lipton-Tarjan planar separator theorem [27]. This notion was further generalized by Henzinger et al. [23] to $(r, f(r))$-divisions[‡] where f is a function in $o(r)$ and each set has at most $f(r)$ vertices in common with the other sets. They noted that Frederickson's proof can easily be adapted to obtain an $(r, c \cdot f(r))$-division of any graph G from a subgraph-closed f-separable graph class – as formalized in Theorem 4. Note that we use an equivalent but slightly different notation than Frederickson and Henzinger et al. in that we consider the "boundary" vertices as a single separate set apart from the non-boundary vertices in each "region", i.e., our *divisions* are actually partitions of the vertex set. This allows us to carefully describe the number of vertices inside each "region".

‡ They use a more general notion of (r, s)-division but we need the restricted version as described here.

▶ **Theorem 4** ([17, 23]). *For any subgraph-closed f-separable class of graphs $\mathcal{G}$, there are constants c_1, c_2 such that every graph G in the class has an $(r, c_1 \cdot f(r))$-*division *for any r. Namely, for any $r \geq 1$, there is an integer $t \in \Theta(\frac{n}{r})$ such that V can be partitioned into $t+1$ sets $\mathcal{X}, V_1, \ldots, V_t$ where the following properties hold.*

(i) $N(V_i) \cap V_j = \emptyset$ *for each* $i \neq j$ *and* $\mathcal{X} = \bigcup_i N(V_i)$,

(ii) $|V_i \cup N(V_i)| \leq r$ *for each* i,

(iii) $|N(V_i)| \leq c_1 \cdot f(r)$ *for each* i *(thus,* $|\mathcal{X}| \leq \sum_{i=1}^{t} |N(V_i)| \leq c_2 \cdot \frac{f(r) \cdot n}{r}$*).*

Moreover, such a partition can be found in $O(g(n) \log n)$ time where $g(n)$ is the time required to find an f-separation in $\mathcal{G}$.

We obtain our color-balanced version of Theorem 4 via two steps. First, we strengthen the notion of $(r, f(r))$-divisions to *uniform* $(r, f(r))$-divisions. A *uniform* $(r, f(r))$*-division* is an $(r, f(r))$-division where the $\Theta(\frac{n}{r})$ sets have a *uniform* size. Namely, there are not only $O(r)$ many **internal** vertices (as in Theorem 4) but there are also $\Omega(r)$ many of them.

It is important to note that while this uniformity condition (i.e., that each region is not *too small* – see Lemma 5(ii)) has **not** been needed in the past[§], it is essential for our analysis of local search as applied to MC problems in the next section. Moreover, to the best of our knowledge, neither Frederickson's construction nor more modern constructions (e.g. [25]) of an r-division explicitly guarantee that the resulting r-division is uniform. To be specific, Frederickson's approach consists of two steps. The first step recursively applies the separator theorem until each region together with its boundary is "small enough". In the second step, each region where the boundary is "too large" is further divided. This is accomplished by applying the separator theorem to a weighted version of each such region where the boundary vertices are uniformly weighted and the non-boundary vertices are zero-weighted. Clearly, even a single application of this latter step may result in regions with $o(r)$ interior vertices. Modern approaches (e.g. [25]) similarly involve applying the separator theorem to weighted regions where boundary vertices are uniformly weighted and interior vertices are zero-weighted, i.e., regions which are *too small* are not explicitly avoided.

In the second step, we generalize the uniform $(r, f(r))$-divisions to *two-color uniform* $(r, f(r))$*-divisions* of a two-colored graph (the coloring need not be proper in the usual sense). A two-color uniform $(r, f(r))$-division of a two-colored graph is a uniform $(r, f(r))$-division where each set has the "same" proportion of each color class (this is formalized in Theorem 7).

The remainder of this section is outlined as follows. We will first show that for every f-separable graph class $\mathcal{G}$ there is a constant c such that every graph in $\mathcal{G}$ has a uniform $(r, c \cdot f(r))$-division (see Lemma 5). We then use this result to show that for every f-separable graph class $\mathcal{G}$ there is a constant c' such that every two-colored graph in $\mathcal{G}$ has a two-color uniform $(rq, c' \cdot q \cdot f(r))$-division for any q – see Theorem 7. Our proofs are constructive and lead to efficient algorithms that produce such divisions when there is a corresponding efficient algorithm to compute an f-separation.

To prove the first result, we start from a given $(r, f(r))$-division and "group" the sets carefully to obtain the desired uniformity. For the two-colored version, we start from a uniform $(r, f(r))$-division and again regroup the sets via a reformulation of the problem as a partitioning problem on two-dimensional vectors. Lemma 6 is used for this regrouping.

▶ **Lemma 5.** *Let $\mathcal{G}$ be a subgraph-closed f-separable graph class and G be a sufficiently large n-vertex graph in $\mathcal{G}$. There are constants r_0, x_0 (depending only on f) such that for any*

[§] E.g., to analyse local search for SC problems [30], or for fast algorithms to find shortest paths [23].

$r \in [r_0, \frac{n}{x_0}]$ there is an integer $t \in \Theta(\frac{n}{r})$ such that $V(G)$ can be partitioned into $t+1$ sets $\mathcal{X}, V_1, \ldots, V_t$ where c_1, c_2 are constants (depending only on f) with the following properties.

(i) *$N(V_i) \cap V_j = \emptyset$ for each $i \neq j$ and $\mathcal{X} = \bigcup_i N(V_i)$,*

(ii) *$\boxed{|V_i| \geq \frac{r}{2}}$¶ and $|V_i| \leq 2r$ for each i,*

(iii) *$|N(V_i)| \leq c_1 \cdot f(r)$ for each i (thus, $|\mathcal{X}| \leq \sum_{i=1}^{t} |N(V_i)| \leq \frac{c_2 \cdot f(r) \cdot n}{r}$).*

Moreover, such a partition can be found in $O(h(n) + n)$ time where $h(n)$ is the amount of time required to produce an $(r, f(r))$-division of G.

Proof. We pick the constants c_1', c_2' as obtained from Theorem 4 for our f-separable subgraph-closed graph class. We further pick r_0 such that it is divisible by 8 and $c^* = 1 - c_2' \cdot f(\frac{r_0}{8}) \cdot (\frac{r_0}{8})^{-1} > 0$. Now, let $x_0 = \frac{3}{c^*}$, and assume $r \in [r_0, \frac{n}{x_0}]$ in what follows.

Consider an $(\lfloor \frac{r}{8} \rfloor, c_1' \cdot f(\lfloor \frac{r}{8} \rfloor))$-division $\mathcal{U} = (\mathcal{X}, U_1, \ldots, U_\ell)$ as given by Theorem 4 – note: $|\mathcal{X}| \leq \frac{c_2' \cdot f(\lfloor \frac{r}{8} \rfloor) n}{\lfloor \frac{r}{8} \rfloor}$. We further define c_ℓ so that $\ell = c_\ell \cdot \frac{8 \cdot n}{r}$. We will partition $[\ell]$ into t sets $I_1, \ldots, I_t$ such that $(\mathcal{X}, V_1, \ldots, V_t)$ is a uniform $(r, c \cdot f(r))$-division $\mathcal{X}, V_1, \ldots, V_t$ where $V_i = \bigcup_{j \in I_i} U_j$. In order to describe the partitioning, we first observe some useful properties of $U_1, \ldots, U_\ell$ where, without loss of generality, $|U_1| \geq \cdots \geq |U_\ell|$. Let $n^* = \sum_{j=1}^{\ell} |U_j|$, and set $t = \lceil \frac{n^*}{r} \rceil$. Note that:

$$n^* = \sum_{j=1}^{\ell} |U_j| = n - |\mathcal{X}| \geq n \cdot \left(1 - \frac{c_2' \cdot f(\lfloor \frac{r}{8} \rfloor)}{\lfloor \frac{r}{8} \rfloor}\right).$$

From our choice of t, the average size of the sets V_i is $\frac{n^*}{t} \in (\frac{r}{1+\frac{r}{n^*}}, r]$. Additionally, $n^* \geq c^* \cdot n$, i.e., $c^* \leq \frac{n^*}{n}$. Thus, we have $r \leq \frac{n^*}{3}$, and the average size of our sets $|V_i|$ is in $[\frac{3r}{4}, r]$.

Notice that $\frac{\ell}{t} \leq c_\ell \cdot \frac{8 \cdot n}{r} \cdot (\frac{n^*}{r})^{-1} \leq \frac{8 c_\ell}{c^*}$. We build the sets I_i such that $|I_i| \leq 40 \cdot \frac{c_\ell}{c^*}$. This provides $|N(V_i)| \leq 40 \cdot \frac{c_\ell}{c^*} \cdot c_1' f(\lfloor \frac{r}{8} \rfloor) \in O(f(r))$. We build the sets I_i in two steps. In the first step we greedily fill the sets I_i according to the largest unassigned set U_j as follows. For each j^* from 1 to ℓ, we consider an index $i^* \in [t]$ where $|I_{i^*}| < 32 \cdot \frac{c_\ell}{c^*}$ and $|V_{i^*}|$ is minimized. If $|V_{i^*}| \leq \frac{n^*}{t}$, then we place j^* into I_{i^*}, that is, we replace V_{i^*} with $V_{i^*} \cup U_{j^*}$. Otherwise (there is no such index i^*), we proceed to step two (below). Before discussing step two, we first consider the state of the sets V_i at the moment when this greedy placement finishes. To this end, let j^* be the index of the first (i.e., the largest) U_j which has not been placed.

Claim 1: *If $|V_i| \leq \frac{n^*}{t}$ for every i, then each set U_j has been merged into some V_i and the V_i's satisfy the conditions of the lemma.*

First, suppose there is an unallocated set U_j. Since $|V_i| \leq \frac{n^*}{t}$ for each $i \in [t]$, our greedy procedure stopped due to having $|I_i| = 32 \cdot \frac{c_\ell}{c^*}$ for each $i \in [t]$. This contradicts the average size of the I_i's being $\frac{\ell}{t} \leq 8 \cdot \frac{c_\ell}{c^*}$. So, every set U_j must have been merged into some V_i. Thus, since $|V_i| \leq \frac{n^*}{t}$ and the average of the $|V_i|$'s is $\frac{n^*}{t}$, we have that for every $i \in [t]$, $|V_i| = \frac{n^*}{t}$. Moreover, for each $i \in [t]$, $|I_i| \leq 32 \frac{c_\ell}{c^*}$. Thus the V_i's satisfy the lemma.

Claim 2: *For every $i \in [t]$, $|V_i| \geq \frac{r}{2}$.*

Suppose some index i has $|V_i| < \frac{r}{2}$. Notice that, if $|I_i| < 32 \cdot \frac{c_\ell}{c^*}$, then for every $i' \in [t]$, $|V_{i'}| \leq |V_i| + \frac{r}{8} \leq \frac{3r}{4} \leq \frac{n^*}{t}$, i.e., contradicting Claim 1. Thus, $|I_i| = 32 \cdot \frac{c_\ell}{c^*}$ for each $i \in [t]$ where $|V_i| < \frac{r}{2}$. For each $i' \in [t], j' \in [\ell]$, let $I_{i'}^{j'}$ and $V_{i'}^{j'}$ be the states of $I_{i'}$ and $V_{i'}$(respectively) directly after index j' has been added to some set $I_{i''}$ by the greedy algorithm.

¶ This lower bound is the difference from the known Theorem 4.

We now let $\hat{j}$ be the largest index in I_i, and assume (without loss of generality) that for every $i' \in [t] \setminus \{i\}$, if $|V_{i'}^{\hat{j}}| < \frac{r}{2}$, then $I_{i'}^{\hat{j}} < 32 \cdot \frac{c_\ell}{c^*}$. Intuitively, i is the "first" index which attains $|I_i| = 32 \cdot \frac{c_\ell}{c^*}$ while still having $|V_i| < \frac{r}{2}$. Now, since $|I_i^{\hat{j}}| = 32 \cdot \frac{c_\ell}{c^*}$, and $|V_i^{\hat{j}}| < \frac{r}{2}$, we have $|U_{\hat{j}}| < r \cdot \frac{c^*}{64c_\ell}$. Thus, for every iteration $j > \hat{j}$, we have $|U_j| < r \cdot \frac{c^*}{64c_\ell}$. This means that after iteration $\hat{j}$, the number of unallocated vertices is strictly less than:

$$\sum_{j=\hat{j}}^{\ell} U_j < \ell \cdot r \cdot \frac{c^*}{64c_\ell} \leq t \cdot 8 \cdot \frac{c_\ell}{c^*} \cdot r \cdot \frac{c^*}{64c_\ell} = \frac{tr}{8}.$$

In particular, this means that on average each set V_i can grow by less than $\frac{r}{8}$. However, due to our choice of i, we see that for every $i' \in [t] \setminus \{i\}$, $|V_{i'}^{\hat{j}}| \leq |V_i^{\hat{j}}| + \frac{r}{8} < \frac{r}{2} + \frac{r}{8}$. This means that even if we allocate all the remaining vertices, the average size of our sets V_i will be strictly less than $\frac{3r}{4} \leq \frac{n^*}{t}$, i.e., providing a contradiction and proving Claim 2.

Claim 3: *If every $j \in [\ell]$ is placed into some I_i, the V_i's satisfy the conditions of the lemma.* First, note that $|I_i|$ is at most $32 \cdot \frac{c_\ell}{c^*}$, i.e., $|N(V_i)| \in O(f(r))$. By Claim 2, we see that $|V_i| \geq \frac{r}{2}$ for each $i \in [t]$. Additionally, from the greedy construction, we have that $|V_i| \leq \frac{n^*}{t} + \frac{r}{8}$. Thus, $|V_i| \in [\frac{r}{2}, \frac{9r}{8}] \subset [\frac{r}{2}, 2r]$.

We now describe the second step. By Claim 3, we assume there are unassigned sets U_j. By Claim 2, for every $i \in [t]$, $|V_i| \geq \frac{r}{2}$. Finally, by Claim 1, there is an index i' where $|V_{i'}| > \frac{n^*}{t}$. Thus, since we have $t = \lceil \frac{n^*}{r} \rceil$ sets which partition at most n^* elements, there must be some index i'' where $|V_{i''}| \leq \frac{n^*}{t}$ and $|I_{i''}| = 32 \cdot \frac{c_\ell}{c^*}$, i.e., $|U_{j^*}| \leq \frac{n^*}{t} \cdot (32 \cdot \frac{c_\ell}{c^*})^{-1} \leq \frac{r \cdot c^*}{32 \cdot c_\ell}$ where U_{j^*} is the largest unassigned set. Notice that there are at most $\ell \leq t \cdot 8 \cdot \frac{c_\ell}{c^*}$ indices which can be assigned and all the remaining sets contain at most $|U_{j^*}|$ vertices. If we spread these remaining U_j's uniformly throughout our V_i's, we will place at most $8 \cdot \frac{c_\ell}{c^*} \cdot |U_{j^*}| \leq \frac{r}{4}$ vertices into each V_i. Thus, for each $i \in [t]$, we have $|V_i| \leq \frac{n^*}{t} + \frac{r}{8} + \frac{r}{4} \leq 2r$. So, by uniformly assigning these remaining indices, we have $|I_i| \leq 40 \cdot \frac{c_\ell}{c^*}$, $|V_i| \in [\frac{r}{2}, 2r]$, and $|N(V_i)| \leq 40 \cdot \frac{c_\ell}{c^*} \cdot c_1' f(\lfloor \frac{r}{8} \rfloor) \in O(f(r))$, as needed.

We conclude with a brief discussion of the time complexity. First, we generate the $(\lfloor \frac{r}{8} \rfloor, c_1' f(\lfloor \frac{r}{8} \rfloor))$-division in $h(n)$ time. We then sort the sets $|U_1| \geq \ldots \geq |U_\ell|$ (this can be done in $O(n)$ time via bucket sort). In the next step we greedily fill the index sets – this takes $O(n)$ time. Finally, we place the remaining "small" sets uniformly throughout the V_i's – taking again $O(n)$ time. Thus, we have $O(h(n) + n)$ time in total. ◀

We now prove a technical lemma which, together with the previous lemma regarding uniform divisions, provides our uniform two-color balanced divisions (see Theorem 7).

▶ **Lemma 6.** *Let c and c' be positive constants, and $A = \{(a_1, b_1), \ldots, (a_n, b_n)\} \subseteq (\mathbb{Q} \cap [0, \infty))^2$ be a set of 2-dimensional vectors where $a_i + b_i \in [c', c]$ for each $i \in [n]$, and $\alpha \in [0, 1]$ such that $\sum_{i=1}^n a_i = \alpha \cdot \sum_{i=1}^n b_i$. Then:*

($\star$) *There is a permutation $p_1, \ldots, p_n$ of $[n]$ such that for any $1 \leq i \leq i' \leq n$, $|\sum_{j=i}^{i'} (a_{p_j} - \alpha \cdot b_{p_j})| \leq 2 \cdot c$; and*

($\star\star$) *For any positive integer q' such that $q'(q'+1) \leq n$ there is a partitioning of $[n]$ into subsets $I_1, \ldots, I_k$ such that for each $j \in [k]$:*

(i) *$q' \leq |I_j| \leq q' + 1$ (thus, $q'c' \leq \sum_{i \in I_j} a_i + b_i \leq q'c + c$),*

(ii) *$|\sum_{i \in I_j} (a_i - \alpha \cdot b_i)| \leq 2 \cdot c$.*

Moreover, the permutation $p_1, \ldots, p_n$ and partition can be computed in $O(n)$ time.

Proof. We first prove ($\star$). To this end, we partition $[n]$ into three sets $A_{>0}$, $A_{<0}$, and $A_{=0}$ according to whether the *weighted* difference $d_i = a_i - \alpha \cdot b_i$ is positive, negative, or 0 (respectively). Note that, $\sum_{i=1}^{n} d_i = 0$ and for each $i \in [n]$, $|d_i| \leq c$. We pick indices one-by-one from $A_{>0}$, $A_{<0}$, $A_{=0}$ to form the permutation.

We now construct a permutation $p_1, \ldots, p_n$ on the indices $[n]$ so that any consecutive subsequence S has $|\sum_{i\in S} d_{p_i}| \leq 2 \cdot c$. For notational convenience, for each $j \in [n]$, we use $\delta_{<j}$ to denote $\sum_{i=1}^{j-1} d_{p_i}$. We now pick the p_i's so that for each j, $|\delta_{<j}| \leq c$. We initialize $\delta_{<1} = 0$. For each j from 1 to n we proceed as follows. Assume that $|\delta_{<j}| \leq c$. We further assume that any index $i \in \{p_1, \ldots, p_{j-1}\}$ has been removed from the sets $A_{>0}$, $A_{<0}$, and $A_{=0}$. If $\delta_{<j}$ is negative, $A_{>0}$ must contain an index j^* since $\sum_{i\in[n]} d_i = 0$. Moreover, if we set $p_j = j^*$, we have $|\delta_{<j+1}| \leq c$ as needed (we also remove the index j^* from $A_{>0}$ at this point). Similarly, if $\delta_{<j}$ is positive, we pick any index j^* from $A_{<0}$, remove it from $A_{<0}$, and set $p_j = j^*$. Finally, when $\delta_{<j} = 0$, we take any index j^* from $A_{>0} \cup A_{<0} \cup A_{=0}$, remove it from $A_{>0} \cup A_{<0} \cup A_{=0}$, and set $p_j = j^*$. Thus, in all cases, $|\delta_{<j+1}| \leq c$. Notice that, for any $1 \leq j \leq j' \leq n$, we have $|\sum_{i=j}^{j'} d_{p_i}| = |\delta_{<j} - \delta_{<j'+1}| \leq |\delta_{<j}| + |\delta_{<j'+1}| \leq 2 \cdot c$ (thus, establishing ($\star$)).

We now prove ($\star\star$) using ($\star$). We partition $[n]$ to form the sets $I_1, \ldots, I_k$ by splitting $p_1, \ldots, p_n$ into k consecutive subsequences of almost equal size. Namely, since $\frac{n}{q'} - \frac{n}{q'+1} \geq 1$, we can pick a positive integer $k \in \left[\frac{n}{q'+1}, \frac{n}{q'}\right]$. Then $q'k \leq n \leq (q'+1)k$, so we can make the sets $I_1, \ldots, I_{n-q'k}$ with $q'+1$ indices each and the sets $I_{n-q'k+1}, \ldots, I_k$ with q' indices each by partitioning π into these sets in order. This is all easily accomplished in $O(n)$ time. ◀

We will now use Lemmas 5 and 6 to prove Theorem 7. In particular, for a given two-colored graph G where G belongs to an f-separable graph class, we first construct a uniform $(r, c \cdot f(r))$-division $(\mathcal{X}, V_1, \ldots, V_t)$ of G as in Lemma 5. From this division we can again carefully combine the V_i's to make new sets W_j where each W_j has roughly the same size and contains roughly the same proportion of each color class as occurring in G. This follows by simply imagining each region V_i of the uniform $(r, c \cdot f(r))$-division as a two-dimensional vector (according to its coloring) and then applying Lemma 6.

▶ **Theorem 7.** *Let $\mathcal{G}$ be a subgraph-closed f-separable graph class and G be a 2-colored n-vertex graph in $\mathcal{G}$ with color classes Γ_1, Γ_2 such that $|\Gamma_2| \geq |\Gamma_1|$. For any q and $r \ll n$ where r is suitably large, there is an integer $t \in \Theta(\frac{n}{q\cdot r})$ such that V can be partitioned into $t+1$ sets $\mathcal{X}, V_1, \ldots, V_t$ where c_1, c_2 are constants (depending only on f) and there is an integer $q' \in [q, 2q-1]$ satisfying the following properties.*

(i) $N(V_i) \cap V_j = \emptyset$ *for each* $i \neq j$ *and* $\mathcal{X} = \bigcup_i N(V_i)$*,*

(ii) $|V_i| \geq \frac{q'\cdot r}{2}$ *and* $|V_i| \leq 2 \cdot (q'+1) \cdot r$ *for each* i*,*

(iii) $|N(V_i)| \leq c_1 \cdot q \cdot f(r)$ *for each* i
(thus, $|\mathcal{X}| \leq \sum_{i=1}^{t} |N(V_i)| \leq \frac{c_2\cdot f(r)\cdot n}{r}$*),*

(iv) $\left||V_i \cap \Gamma_1| - \frac{|\Gamma_1|}{|\Gamma_2|} \cdot |V_i \cap \Gamma_2|\right| \leq 4 \cdot r$ *for each* i*.*

Moreover, such a partition can be found in $O(h(n) + n)$ time where $h(n)$ is the amount of time required to produce a uniform $(r, c \cdot f(r))$-division of G.

3 Proof of Theorem 2: PTAS for f-Separable Maximum Coverage

Recall that, we have an f-separable instance of MC where f is strictly sublinear. Our algorithm is based on local search. We fix a sufficiently large positive constant integer $b \geq 1$. Given an f-separable instance of MC, we pick an arbitrary initial solution $\mathcal{A}$. We check if it is possible to replace $2b^2 + 2b$ sets in $\mathcal{A}$ with the same number of sets from $\mathcal{F}$ so that the total number of elements covered is increased. We perform such a replacement (swap) as long as there is one. We stop if there is no profitable swap and output the resulting solution.

We will show that for sufficiently large b the above algorithm yields a $(1 - 8c_1 f(b)/b)$-approximate solution and that it runs in polynomial time (for constant b). Here, c_1 is the constant from Theorem 7. Setting b to be sufficiently large will complete this proof. Note that, if $c_1 < 1$, Theorem 7 also holds for $c_1 = 1$. Thus, it suffices to consider $c_1 \geq 1$.

Since each step increases the number of covered elements, the number of iterations of the above algorithm is at most $|U|$. In each iteration, we consider each of the $\binom{k}{b}\binom{|\mathcal{F}|}{b}$ potential b-swaps, and check whether it is an improvement. Therefore, the total running time of the algorithm is polynomial for constant b.

It remains to establish the performance guarantee. Let $\mathcal{O}$ be an optimum solution to the instance and let $\mathcal{A}$ be the (locally optimal) solution output by the algorithm. Let OPT, ALG denote the number of elements covered by $\mathcal{O}$, $\mathcal{A}$, respectively.

Suppose that ALG $< \left(1 - \frac{8c_1 f(b)}{b}\right)$ OPT. We will show that implies that there is a profitable swap (contradicting the local optimality of $\mathcal{A}$ and hence completing the proof).

We claim that it suffices to consider the case when $\mathcal{O}, \mathcal{A}$ are disjoint, which is justified as follows. Assume that $\mathcal{O} \cap \mathcal{A} \neq \emptyset$. We remove the sets in $\mathcal{O} \cap \mathcal{A}$ from $\mathcal{F}$ and all the *elements* covered by these sets from U. Moreover, we decrease k by $|\mathcal{O} \cap \mathcal{A}|$ and replace $\mathcal{O}$ with $\mathcal{O} \setminus \mathcal{A}$ and $\mathcal{A}$ with $\mathcal{A} \setminus \mathcal{O}$. Since our class of instances is closed under removing sets and elements, the resulting instance is still contained in the class. Moreover, $|\bigcup \mathcal{A}| < \left(1 - \frac{8c_1 f(b)}{b}\right) |\bigcup \mathcal{O}|$ (note that the number of elements covered by $\mathcal{A}$ and $\mathcal{O}$, respectively, decreases by the same amount as we remove the elements covered by $\mathcal{A} \cap \mathcal{O}$ from the instance). Finally, a feasible and profitable swap in the reduced instance implies that the same swap is also feasible and profitable in the original.

Therefore, we assume from now on that $\mathcal{A}$ and $\mathcal{O}$ are disjoint. Since our instance is f-separable, there exists an f-separable graph G with precisely $2k$ nodes for the two feasible solutions $\mathcal{O}$ and $\mathcal{A}$ with the properties stated in Definition 1.

We now apply Theorem 7 to G with color classes $\Gamma_1 = \mathcal{O}$ and $\Gamma_2 = \mathcal{A}$ and with parameters $r = b$ and $q = b$. Here, we assume that the constant b has been picked sufficiently large as required by Theorem 7. We further assume that the number $2k$ of nodes in the exchange graph is substantially larger than b as the problem can be solved to optimality in polynomial time for constant k. Since $|\mathcal{O}| = |\mathcal{A}| = k$, the two color classes in G are perfectly balanced. Let $\mathcal{A}_i = \mathcal{A} \cap V_i$, $\mathcal{O}_i = \mathcal{O} \cap V_i$, $N_i^{\mathcal{O}} = N(V_i) \cap \mathcal{O}$ and $\bar{\mathcal{O}}_i = \mathcal{O}_i \cup N_i^{\mathcal{O}}$ for any part V_i with $i \in [t]$ of the resulting subdivision of G. Note that every set in $\mathcal{O}$ is in $\bar{\mathcal{O}}_i$ for some $i \in [t]$.

The idea of the analysis is to consider for each $i \in [t]$ a nearly balanced (but possibly infeasible) *candidate swap* that replaces in $\mathcal{A}$ the sets $\mathcal{A}_i$ with $\bar{\mathcal{O}}_i$. We will first show that there exists a profitable candidate swap if ALG $< \left(1 - \frac{8c_1 f(b)}{b}\right)$ OPT. Second, we will show that even a *feasible* profitable swap can be obtained by adding only some of the sets in $\bar{\mathcal{O}}_i$ as the candidate swap was nearly balanced.

For technical reasons we will define a set Z of elements that we (temporarily) disregard from our calculations because they will remain covered and thus do not impact which of the sets in $\bar{\mathcal{O}}_i$ we will pick for the feasible swap. Let $Z_i = \bigcup(\mathcal{A} \setminus \mathcal{A}_i)$ be the set of elements that remain covered even if $\mathcal{A}_i$ is removed from $\mathcal{A}$ and let $Z = \bigcap_{i=1}^t Z_i$ be the set of elements that are covered by $\mathcal{A}$ but not exclusively by one of the $\mathcal{A}_i$. In particular, Z contains all elements in $\bigcup(\mathcal{X} \cap \mathcal{A})$. Let $L_i = \bigcup \mathcal{A}_i \setminus Z$ be the set of elements that are "lost" when removing the $\mathcal{A}_i$ from $\mathcal{A}$. Moreover, let $W_i = \bigcup \bar{\mathcal{O}}_i \setminus Z_i$ be the set of elements that are "won" when we add all the sets of $\bar{\mathcal{O}}_i$ after removing $\mathcal{A}_i$.

We claim that $\sum_{i=1}^{t} |L_i| \leq \text{ALG} - |Z|$. To this end, note that $Z \subseteq \bigcup \mathcal{A}$ and that the family $\{L_i\}_{i\in[t]}$ contains pairwise disjoint sets because all elements that are not exclusively covered by a single $\mathcal{A}_i$ are contained in Z and thus removed.

We further claim that $\sum_{i=1}^{t} |W_i| \geq \text{OPT} - |Z|$. To see this, note first that every element in Z contributes 0 to the left hand side and 0 or -1 to the right hand side. Every element covered by $\mathcal{O}$ but not by $\mathcal{A}$ contributes at least 1 to the left hand side and precisely 1 to the right hand side. Finally, consider an element u that is covered both by $\mathcal{A}$ and by $\mathcal{O}$ but does not lie in Z. Since $u \notin Z$ there is no $S \in \mathcal{X} \cap \mathcal{A}$ covering u. This also implies that u lies in a set $S \in \mathcal{A}_i$ for some *unique* $i \in [t]$. Thus, we even have $u \notin Z_i$. Since G is an exchange graph, there is some set $T \in \mathcal{O}$ with $u \in T$ and some set $S' \in \mathcal{A}$ with $u \in S'$ such that S' and T are adjacent in G. Since $u \notin Z_i$ we have $S' \in \mathcal{A}_i$. Since T is adjacent to $S' \in \mathcal{A}_i \subseteq V_i$, we have $T \in N(V_i)$. Since $T \in \mathcal{O}$, it follows that $T \in \mathcal{N}_i^{\mathcal{O}} \subseteq \bar{\mathcal{O}}_i$. Thus $u \in \bigcup \bar{\mathcal{O}}_i \setminus Z_i = W_i$. Hence u contributes at least 1 to the left hand side and precisely 1 to the right hand side of $\sum_{i=1}^{t} |W_i| \geq \text{OPT} - |Z|$, which shows the claim.

By $\text{OPT} > \text{ALG} \geq |Z|$, $\min\limits_{\substack{i\in[t]\\ |W_i|>0}} \dfrac{|L_i|}{|W_i|} \leq \dfrac{\sum_{i=1}^{t}|L_i|}{\sum_{i=1}^{t}|W_i|} \leq \dfrac{\text{ALG} - |Z|}{\text{OPT} - |Z|} \leq \dfrac{\text{ALG}}{\text{OPT}} < 1 - \dfrac{8c_1 f(b)}{b}$.

Hence, we pick $i \in [t]$ so that
$$\frac{|L_i|}{|W_i|} < 1 - \frac{8c_1 f(b)}{b} . \quad (1)$$

Recall that $c_1 \geq 1$ and assume that b is large enough so that $f(b) \geq 1$. Then by Properties (ii), (iii), and (iv) of Theorem 7 (respectively), we have that $|V_i| \geq b^2/2$, $|N(V_i)| \leq c_1 b \cdot f(b)$, and $||\mathcal{A}_i| - |\mathcal{O}_i|| \leq 4b$ (respectively). Now, since $|\mathcal{A}_i| + |\mathcal{O}_i| = |V_i|$, we have $|\bar{\mathcal{O}}_i| \leq \frac{1}{2}|V_i| + 2b + c_1 b \cdot f(b)$ and $|\mathcal{A}_i| \geq \frac{1}{2}|V_i| - 2b$. Hence

$$\frac{|\mathcal{A}_i|}{|\bar{\mathcal{O}}_i|} \geq \frac{\frac{1}{2}|V_i| - 2b}{\frac{1}{2}|V_i| + 2b + c_1 b \cdot f(b)} \geq \frac{(\frac{1}{2}|V_i| + 2b + 4c_1c_2 b \cdot f(b)) - 4b - 4c_1c_2 b \cdot f(b)}{\frac{1}{2}|V_i| + 2b + 4c_1c_2 b \cdot f(b)} \quad (2)$$
$$\overset{|V_i|\geq b^2/2}{\geq} 1 - \frac{8c_1 f(b)}{b} . \quad (3)$$

We are now ready to construct our feasible and profitable swap. We inductively define an order $S_1, \ldots, S_{|\bar{\mathcal{O}}_i|}$ on the sets in $\bar{\mathcal{O}}_i$ where we require that

$$S_j = \arg\max_{S\in\bar{\mathcal{O}}_i} \left| S \setminus \left(Z_i \cup \bigcup_{\ell=1}^{j-1} S_\ell \right) \right| \text{ for any } j = 1, \ldots, |\bar{\mathcal{O}}_i| \text{ where } S_1 \text{ maximizes } |S \setminus Z_i|.$$

Consider the following process of iteratively building a set W' starting with $W' = \emptyset$. Suppose that we add to W' the sets $(S_1 \setminus Z_i), \ldots, (S_{|\bar{\mathcal{O}}_i|} \setminus Z_i)$ in this order ending up with $W' = W_i$. In doing so, the incremental gain is monotonically non-increasing due to the definition of the order on $\bar{\mathcal{O}}_i$ and due to the submodularity of the objective function. Hence, for any prefix of the first j sets we have that
$$\left| \left(\bigcup_{\ell=1}^{j} S_\ell \right) \setminus Z_i \right| \geq \frac{j \cdot |W_i|}{|\bar{\mathcal{O}}_i|} . \quad (4)$$

Suppose that $|\bar{\mathcal{O}}_i| > |\mathcal{A}_i|$ (otherwise due to (1), we can just add all sets in $\bar{\mathcal{O}}_i$). Consider the swap where we replace the $|\mathcal{A}_i| \leq 1/2|V_i| + 2b \leq 2b^2 + 2b$ many sets $\mathcal{A}_i$ from the local optimum $\mathcal{A}$ with at most $|\mathcal{A}_i|$ many sets $\{S_1, \ldots, S_{|\mathcal{A}_i|}\}$ from $\bar{\mathcal{O}}_i$.

We now analyze how this swap affects the objective function value. Notice that, by removing the sets in $\mathcal{A}_i$, the objective function value drops by

$$|L_i| \overset{(1)}{<} \left(1 - \frac{8c_1 f(b)}{b}\right) \cdot |W_i| \overset{(4)}{\leq} \left(1 - \frac{8c_1 f(b)}{b}\right) \frac{|\bar{O}_i|}{|\mathcal{A}_i|} \left| \left(\bigcup_{\ell=1}^{|\mathcal{A}_i|} S_\ell \right) \setminus Z_i \right|$$

$$\overset{(3)}{\leq} \left| \left(\bigcup_{\ell=1}^{|\mathcal{A}_i|} S_\ell \right) \setminus Z_i \right| .$$

The right hand side of this inequality is the increase of the objective function due to adding the sets $\{S_1, \ldots, S_{|\mathcal{A}_i|}\}$ after removing the sets in $\mathcal{A}_i$.

Therefore the above described swap is feasible and also profitable and thus $\mathcal{A}$ is not a local optimum leading to a contradiction (and completing the proof of Theorem 2).

References

1 Alexander A. Ageev and Maxim Sviridenko. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *J. Comb. Optim.*, 8(3):307–328, 2004. doi:10.1023/B:JOCO.0000038913.96607.c2.

2 Noga Alon, Paul Seymour, and Robin Thomas. A separator theorem for nonplanar graphs. *J. of the American Mathematical Society*, 3:801–808, 1990. doi:10.1090/S0894-0347-1990-1065053-0.

3 Rom Aschner, Matthew J. Katz, Gila Morgenstern, and Yelena Yuditsky. Approximation schemes for covering and packing. In *7th Int. Workshop on Algorithms and Computation (WALCOM'13)*, pages 89–100, 2013.

4 Ashwinkumar Badanidiyuru, Robert Kleinberg, and Hooyeon Lee. Approximating low-dimensional coverage problems. In *Symp. Computational Geometry (SoCG'12)*, pages 161–170, 2012. doi:10.1145/2261250.2261274.

5 Sayan Bandyapadhyay and Kasturi R. Varadarajan. On variants of k-means clustering. In *32nd Symp. Computational Geometry (SoCG'16)*, pages 14:1–14:15, 2016. doi:10.4230/LIPIcs.SoCG.2016.14.

6 Hervé Brönnimann and Michael T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete & Computational Geometry*, 14(4):463–479, 1995. doi:10.1007/BF02570718.

7 Sergio Cabello and David Gajser. Simple PTAS's for families of graphs excluding a minor. *Discrete Applied Mathematics*, 189:41–48, 2015.

8 Timothy M. Chan, Elyot Grant, Jochen Könemann, and Malcolm Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *23rd ACM-SIAM Symp. Discrete Algorithms (SODA'12)*, pages 1576–1585, 2012. URL: http://dl.acm.org/citation.cfm?id=2095116.2095241.

9 Timothy M. Chan and Sariel Har-Peled. Approximation algorithms for maximum independent set of pseudo-disks. In *25th ACM Symp. Computational Geometry (SoCG'09)*, pages 333–340, 2009. doi:10.1145/1542362.1542420.

10 Timothy M. Chan and Sariel Har-Peled. Approximation algorithms for maximum independent set of pseudo-disks. *Discrete & Computational Geometry*, 48(2):373–392, 2012. doi:10.1007/s00454-012-9417-5.

11 Vincent Cohen-Addad, Philip N. Klein, and Claire Mathieu. Local search yields approximation schemes for k-means and k-median in Euclidean and minor-free metrics. In *57th IEEE Symp. Foundations of Computer Science (FOCS'16)*, pages 353–364, 2016. doi:10.1109/FOCS.2016.46.

12 Vincent Cohen-Addad and Claire Mathieu. Effectiveness of local search for geometric optimization. In *31st Symp. Computational Geometry (SoCG'15)*, pages 329–343, 2015. doi:10.4230/LIPIcs.SOCG.2015.329.

13 Gérard Cornuéjols, George L. Nemhauser, and Laurence A. Wolsey. Worst-case and probabilistic analysis of algorithms for a location problem. *Operations Research*, 28(4):847–858, 1980. doi:10.1287/opre.28.4.847.

14 Minati De and Abhiruk Lahiri. Geometric dominating set and set cover via local search. *CoRR*, abs/1605.02499, 2016. URL: http://arxiv.org/abs/1605.02499, arXiv:1605.02499.

15 Thomas Erlebach and Erik Jan van Leeuwen. Approximating geometric coverage problems. In *19th ACM-SIAM Symp. Discrete Algorithms (SODA'08)*, pages 1267–1276, 2008. URL: http://dl.acm.org/citation.cfm?id=1347082.1347220.

16 Uriel Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, 1998. doi:10.1145/285055.285059.

17 Greg N. Frederickson. Fast algorithms for shortest paths in planar graphs, with applications. *SIAM J. Comput.*, 16(6):1004–1022, 1987.

18 Zachary Friggstad, Mohsen Rezapour, and Mohammad R. Salavatipour. Local search yields a PTAS for k-means in doubling metrics. In *57th IEEE Symp. Foundations of Computer Science (FOCS'16)*, pages 365–374, 2016. doi:10.1109/FOCS.2016.47.

19 Matt Gibson and Imran A. Pirwani. Algorithms for dominating set in disk graphs: Breaking the log$n$ barrier - (extended abstract). In *18th European Symp. Algorithms (ESA'10)*, pages 243–254, 2010. doi:10.1007/978-3-642-15775-2_21.

20 Sathish Govindarajan, Rajiv Raman, Saurabh Ray, and Aniket Basu Roy. Packing and covering with non-piercing regions. In *24th European Symp. Algorithms (ESA'16)*, pages 47:1–47:17, 2016. doi:10.4230/LIPIcs.ESA.2016.47.

21 Sariel Har-Peled and Kent Quanrud. Approximation algorithms for polynomial-expansion and low-density graphs. In *23rd European Symp. Algorithms (ESA'15)*, pages 717–728, 2015.

22 Venky Harinarayan, Anand Rajaraman, and Jeffrey D. Ullman. Implementing data cubes efficiently. In *ACM SIGMOD Int. Conference on Management of Data (ICDM'96)*, pages 205–216, 1996. doi:10.1145/233269.233333.

23 Monika Rauch Henzinger, Philip N. Klein, Satish Rao, and Sairam Subramanian. Faster shortest-path algorithms for planar graphs. *J. Comput. Syst. Sci.*, 55(1):3–23, 1997. doi:10.1006/jcss.1997.1493.

24 R. B. O. Kerkkamp and Karen Aardal. A constructive proof of swap local search worst-case instances for the maximum coverage problem. *Oper. Res. Lett.*, 44(3):329–335, 2016. doi:10.1016/j.orl.2016.03.001.

25 Philip N. Klein, Shay Mozes, and Christian Sommer. Structured recursive separator decompositions for planar graphs in linear time. In *45th Symp. Theory of Computing (STOC'13)*, pages 505–514, 2013. doi:10.1145/2488608.2488672.

26 Erik Krohn, Matt Gibson, Gaurav Kanade, and Kasturi R. Varadarajan. Guarding terrains via local search. *J. of Computational Geometry*, 5(1):168–178, 2014. URL: http://jocg.org/index.php/jocg/article/view/128.

27 Richard J Lipton and Robert Endre Tarjan. A separator theorem for planar graphs. *SIAM J. on Applied Mathematics*, 36(2):177–189, 1979. doi:10.1137/0136016.

28 Steffen Mecke and Dorothea Wagner. Solving geometric covering problems by data reduction. In *12th European Symp. Algorithms (ESA'04)*, pages 760–771, 2004. doi:10.1007/978-3-540-30140-0_67.

29 Nabil H. Mustafa, Rajiv Raman, and Saurabh Ray. Quasi-polynomial time approximation scheme for weighted geometric set cover on pseudodisks and halfspaces. *SIAM J. Comput.*, 44(1):1650–1669, 2015. `doi:10.1137/14099317X`.

30 Nabil H. Mustafa and Saurabh Ray. Improved results on geometric hitting set problems. *Discrete & Computational Geometry*, 44(4):883–895, 2010.

31 Erez Petrank. The hardness of approximation: Gap location. *Comput. Complexity*, 4:133–157, 1994. `doi:10.1007/BF01202286`.

32 Neil Robertson and Paul D Seymour. Graph minors. xx. wagner's conjecture. *J. of Combinatorial Theory, Series B*, 92(2):325–357, 2004. `doi:10.1016/j.jctb.2004.08.001`.

33 Kasturi R. Varadarajan. Weighted geometric set cover via quasi-uniform sampling. In *42nd ACM Symp. Theory of Computing (STOC'10)*, pages 641–648, 2010. `doi:10.1145/1806689.1806777`.