Algorithmica
https://doi.org/10.1007/s00453-022-01072-1

Maximizing Coverage While Ensuring Fairness: A Tale of Conflicting Objectives

Abolfazl Asudeh[1] · Tanya Berger-Wolf[2] · Bhaskar DasGupta[1] · Anastasios Sidiropoulos[1]

Received: 17 December 2020 / Accepted: 16 November 2022

Abstract
Ensuring fairness in computational problems has emerged as a *key* topic during recent years, buoyed by considerations for equitable resource distributions and social justice. It *is* possible to incorporate fairness in computational problems from several perspectives, such as using optimization, game-theoretic or machine learning frameworks. In this paper we address the problem of incorporation of fairness from a *combinatorial optimization* perspective. We formulate a combinatorial optimization framework, suitable for analysis by researchers in approximation algorithms and related areas, that incorporates fairness in maximum coverage problems as an interplay between *two* conflicting objectives. Fairness is imposed in coverage by using coloring constraints that *minimizes* the discrepancies between number of elements of different colors covered by selected sets; this is in contrast to the usual discrepancy minimization problems studied extensively in the literature where (usually two) colors are *not* given *a priori* but need to be selected to minimize the maximum color discrepancy of *each* individual set. Our main results are a set of randomized and deterministic approximation algorithms that attempts to *simultaneously* approximate both fairness and coverage in this framework.

✉ Bhaskar DasGupta
bdasgup@uic.edu

Abolfazl Asudeh
asudeh@uic.edu

Tanya Berger-Wolf
berger-wolf.1@osu.edu

Anastasios Sidiropoulos
sidiropo@uic.edu

[1] Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA

[2] Department of Computer Science and Engineering, Ohio State University, Columbus, OH 43210, USA

Published online: 07 December 2022

Keywords Fairness · Maximum coverage · Combinatorial optimization · Approximation algorithms

1 Introduction

In this paper we introduce and analyze a combinatorial optimization framework capturing two *conflicting* objectives: *optimize* the main objective while trying to ensure that the selected solution is as *fair* as possible. We illustrate the framework with the following *simple* graph-theoretic illustration. Consider the graph G of 10 nodes and 18 edges as shown in Fig. 1 where each edge is colored from one of $\chi = 3$ colors (red, blue or green) representing three different attributes. Suppose that we want to select *exactly* $k = 3$ nodes that maximize the number of edges they "cover" subject to the "fairness" constraint that the proportion of red, blue and green edges in the selected edges are the *same*. An optimal solution is shown in Fig. 1 by the solid black nodes u_1, u_2, u_3 covering 6 edges; Fig. 1 also shows that the solution is quite different from what it would have been (the yellow corner nodes v_1, v_2, v_3 covering 11 edges) *if* the fairness constraint was absent. A simple consequence of the analysis of our algorithms for a more general setting is that, assuming that there exists *at least* one feasible solution and assuming k is large enough, we can find a randomized solution to this fair coverage problem for graphs where we select *exactly* k nodes, cover at least 63% of the optimal number of edges on an average and, for *every* pair of colors, with high probability the ratio of the number of edges of these two colors among the selected edges is $O(1)$.

In this paper we consider this type of problem in more general settings. Of course, in the example in Fig. 1 (and in general) there is nothing special about requiring that the proportion of red, blue and green edges in the covered edges should be *exactly* equal as opposed to a *pre-specified* unequal proportion. For example, we may also require that the proportion of edges of different colors in our solution should mimic that in the entire graph, *i.e.*, in Fig. 1 among the covered edges the proportion of red, blue and green edges should be q_1, q_2 and q_3 where $q_1 = 1/6$, $q_2 = 1/3$, and $q_3 = 1/2$. Our algorithms will work with *easy* modifications for any *constant* values of q_1, q_2 and q_3.

1.1 Different Research Perspectives in Ensuring Fairness

Theoretical investigations of ensuring fairness in computation can be pursued from many perspectives. We briefly comment on a few of them.

One line of research dealing with the goal of ensuring fairness uses the *optimization* framework, *i.e.*, we model the problem as an optimization problem with precisely defined fairness constraints. This is a common framework used by researchers in combinatorial and graph-theoretic algorithms, such as research works that involve designing exact or approximation algorithms, investigating fixed-parameter tractability issues or proving inapproximability results. In this paper we use such a framework. Fairness is imposed in coverage by using coloring constraints that minimize the dis-

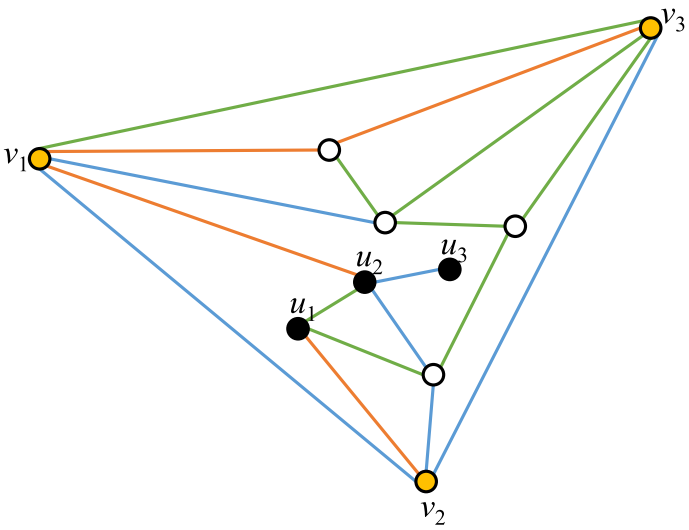

Fig. 1 A simple illustration of fairness in maximum coverage problems for graphs

crepancies between different colors among elements covered by selected sets; this is in contrast to the usual discrepancy minimization problems studied extensively in the literature [12] where the (usually two) colors are *not* given a priori but need to be selected to *minimize* the maximum color discrepancy of *each* individual set.

A second line of research dealing with fairness involves machine learning frameworks. Even though it is a relatively new research area, there is already a large body of research dealing with ensuring fairness in machine learning algorithms by preprocessing the data used in the algorithms, optimization of statistical outcomes with appropriate fairness criteria and metric during the training, or by post-processing the answers of the algorithms [25, 50, 51].

A third line of research dealing with fairness involves game theoretic frameworks. For example, developments of solutions for fair ways of sharing transferable utilities in cooperative game-theoretic environments have given rise to interesting concepts such as Shapley values and Rabin's fairness model. We refer the reader to the excellent textbook in algorithmic game theory by Nisan *et al.* [39] for further details on these research topics.

Yet another more recent line of research dealing with fairness involve applying fairness criteria in the context of clustering of points in a *metric space* under k-means objective, k-median or other ℓ_p-norm objectives [9, 13]. The assumption of an underlying metric space allows the development of efficient algorithms for these frameworks.

2 Fair Maximum Coverage: Notations, Definitions and Related Concepts

The *Fair Maximum Coverage* problem with χ colors is defined as follows. We are given a universe $\mathscr{U} = \{u_1, \ldots, u_n\}$ of n elements, a weight function $w : \mathscr{U} \mapsto \mathbb{R}$ assigning

a non-negative weight to every element, a color function $\mathscr{C} : \mathscr{U} \mapsto \{1, \ldots, \chi\}$ assigning a color to every element, a collection of m sets $\mathscr{S}_1, \ldots, \mathscr{S}_m \subseteq \mathscr{U}$, and a positive integer k. A collection of k distinct subsets, say $\mathscr{S}_{i_1}, \ldots, \mathscr{S}_{i_k}$, with the set of "covered" elements $\bigcup_{j=1}^{k} \mathscr{S}_{i_j}$ containing p_i elements of color i is considered a *valid solution*[1] provided $p_i = p_j$ for all i and j. The objective is to *maximize* the sum of weights of the covered elements. More explicitly, our problem is defined as follows:

Problem name:	Fair Maximum Coverage with χ colors (FMC(χ, k))
Input:	• universe $\mathscr{U} = \{u_1, \ldots, u_n\}$ • (element) weight function $w : \mathscr{U} \mapsto \mathbb{R}^+ \cup \{0\}$ • (element) color function $\mathscr{C} : \mathscr{U} \mapsto \{1, \ldots, \chi\}$ • sets $\mathscr{S}_1, \ldots, \mathscr{S}_m \subseteq \mathscr{U}$ • integer $k > 0$
Valid solution:	collection of k distinct subsets $\mathscr{S}_{i_1}, \ldots, \mathscr{S}_{i_k}$ satsifying $\forall i, j \in \{1, \ldots, \chi\}$: $p_i \stackrel{\text{def}}{=} \left\vert \{u_\ell \mid u_\ell \in \bigcup_{j=1}^{k} \mathscr{S}_{i_j} \text{ and } \mathscr{C}(u_\ell) = i\} \right\vert$ = $p_j \stackrel{\text{def}}{=} \left\vert \{u_\ell \mid u_\ell \in \bigcup_{j=1}^{k} \mathscr{S}_{i_j} \text{ and } \mathscr{C}(u_\ell) = j\} \right\vert$
Objective:	*maximize* $\sum_{u_\ell \in \bigcup_{j=1}^{k} \mathscr{S}_{i_j}} w(u_\ell)$

We denote FMC(χ, k) by just FMC when χ and k are clear from the context. In the sequel, we will distinguish between the following two versions of the problem:

(i) *unweighted* FMC in which $w(u_\ell) = 1$ for all $\ell \in \{1, \ldots, n\}$ and thus the objective is to maximize the *number* of elements covered, and
(ii) *weighted* FMC in which $w(u_\ell) \geq 0$ for all $\ell \in \{1, \ldots, n\}$.

For the purpose of stating and analyzing algorithmic performances, we define the following notations and natural parameters associated with an instance of FMC(χ, k):

- $a \in \{2, 3, \ldots, n\}$ denotes the maximum of the *cardinalities* (number of elements) of all sets.
- $f \in \{1, 2, \ldots, m\}$ denotes the *maximum* of the frequencies of all elements, where the *frequency* of an element is the number of sets in which it belongs.
- OPT denotes the optimal objective value of the given instance of FMC.
- $\mathsf{OPT}_\#$ denotes the *number of covered elements* in an optimal solution of the given instance of FMC. For weighted FMC, if there are multiple optimal solutions then $\mathsf{OPT}_\#$ will the maximum number of elements covered *among* these optimal solutions. Note that $\mathsf{OPT} = \mathsf{OPT}_\#$ for unweighted FMC. The reason we need to consider $\mathsf{OPT}_\#$ separately from OPT for weighted FMC is because the coloring constraints are tied to $\mathsf{OPT}_\#$ whereas the optimization objective is tied to OPT.

[1] For a more general version of the problem we are given χ "color-proportionality constants" $q_1, \ldots, q_\chi \in (0, 1]$ with $q_1 + \cdots + q_\chi = 1$, and a valid solution must satisfy $p_i / p_j = q_i / q_j$ for all i and j. As we mentioned already, with suitable modifications our algorithms will work with similar asymptotic performance guarantee for any *constant* values of $q_1, \ldots, q_\chi$, but to simplify exposition we will assume the simple requirement of $q_1 = \cdots = q_\chi$ in the sequel.

- The performance ratios of many of our algorithms are expressed using the function $\rho(\cdot)$:

$$\rho(x) \stackrel{\text{def}}{=} \left(1 - 1/x\right)^x$$

  Note that $\rho(x) < \rho(y)$ for $x > y > 0$ and $\rho(x) > 1 - \mathbf{e}^{-1}$ for all $x > 0$.

For NP-completeness results, if the problem is trivially in NP then we will *not* mention it. To analyze our algorithms in this paper, we have used several standard mathematical equalities or inequalities which are listed explicitly below for the convenience of the reader:

$$\forall\, x \in [0,1] \,:\, \mathbf{e}^{-x} \geq 1 - x \tag{1}$$

$$\forall\, x \,:\, \mathbf{e}^{-x} = 1 - x + (x^2/2)\mathbf{e}^{-\xi} \;\text{ for some }\; \xi \in [0, x] \tag{2}$$

$$\forall\, \alpha_1, \ldots, \alpha_q \geq 0 \,:\, \left(\tfrac{1}{q}\textstyle\sum_{j=1}^{q} \alpha_j\right)^q \geq \prod_{j=1}^{q} \alpha_j \tag{3}$$

$$\forall\, x \in [0,1]\ \forall\, y \geq 1 \,:\, 1 - \left(1 - \tfrac{x}{y}\right)^y \geq \left(1 - \left(1 - \tfrac{1}{y}\right)^y\right) x \tag{4}$$

2.1 Three Special Cases of the General Version of FMC

In this subsection we state three important special cases of the general framework of FMC.

Fair maximum *k*-node coverage or NODE-FMC

This captures the scenario posed by the example in Fig. 1. We are given a connected undirected edge-weighted graph $G = (V, E)$ where $w(e) \geq 0$ denotes the weight assigned to edge $e \in E$, a color function $\mathscr{C} : E \mapsto \{1, \ldots, \chi\}$ assigning a color to every edge, and a positive integer k. A node v is said to *cover* an edge e if e is incident on v. A collection of k nodes $v_{i_1}, \ldots, v_{i_k}$ covering p_i edges of color i for each i is considered a valid solution provided $p_i = p_j$ for all i and j. The objective is to *maximize* the sum of weights of the covered edges. It can be easily seen that this is a special case of FMC by using the *standard* translation from node cover to set cover, *i.e.*, the edges are the set of elements, and corresponding to every node v there is a set containing the edges incident on v. Note that for this special case $f = 2$ and a is equal to the maximum node-degree in the graph.

Segregated FMC or SEGR-FMC

Segregated FMC is the special case of FMC when all the elements in any set have the *same* color, *i.e.*,

$$\forall\, j \in \{1, \ldots, m\}\, \forall\, u_p, u_q \in \mathscr{S}_j \,:\, \mathscr{C}(u_p) = \mathscr{C}(u_q)$$

Another equivalent way of describing SEGR- FMC is as follows. Let C_j be the set of all elements colored j for $j \in \{1, \ldots, \chi\}$ in the given instance of SEGR- FMC. Let the notation 2^A denote the power set for any set A. Then, SEGR- FMC is the special case when $\mathscr{S}_j \in \bigcup_{j=1}^{\chi} 2^{C_j}$ holds for all $j \in \{1, \ldots, m\}$. A simple example of an instance of SEGR- FMC with $n = 6$, $m = 10$ and $\chi = 2$ is shown below:

$$\begin{aligned} \mathscr{U} &= \{u_1, u_2, u_3, u_4, u_5, u_6\} \\ \mathscr{C}(u_1) = \mathscr{C}(u_2) = \mathscr{C}(u_3) = \mathscr{C}(u_4) &= 1, \quad \mathscr{C}(u_5) = \mathscr{C}(u_6)) = 2 \\ \mathscr{S}_1 = \{u_1, u_2, u_3\}, \quad \mathscr{S}_2 = \{u_2, u_3, u_4\}, &\quad \mathscr{S}_3 = \{u_5, u_6\}, \quad \mathscr{S}_4 = \{u_6\} \\ w(u_1) = w(u_2) = 7, \quad w(u_3) = 9, &\quad w(u_4) = w(u_5) = w(u_6) = 1 \end{aligned}$$

Even though computing an exact solution of SEGR- FMC is still NP-complete, it is much easier to approximate (see Sect. 11.1). From our application point of view as discussed in Sect. 3, this may for example model cases in which city neighborhoods are segregated in some manner, *e.g.*, racially or based on income.

Δ-balanced FMC or Δ-BAL-FMC

Δ-balanced FMC is the special case of FMC when the number of elements of each color in a set are within an *additive* range of Δ, *i.e.*,

$$\begin{aligned} &\forall\, j \in \{1, \ldots, m\} \forall\, p \in \{1, \ldots, \chi\} : \\ &\max\left\{1, \left\lfloor \frac{|\mathscr{S}_j|}{\chi} \right\rfloor - \Delta\right\} \leq \left|u_\ell \,|\, (u_\ell \in \mathscr{S}_j) \wedge (\mathscr{C}(u_\ell) = p)\right| \leq \left\lceil \frac{|\mathscr{S}_j|}{\chi} \right\rceil + \Delta \end{aligned}$$

Similar to SEGR- FMC, it is much easier to approximate Δ-BAL- FMC for small Δ (see Sect. 11.2).

Geometric FMC or GEOM-FMC

In this *unweighted* "geometric" version of FMC, the elements are points in $[0, \Delta]^d$ for some Δ and some constant $d \geq 2$, the sets are unit radius balls in $\mathbb{R}^d$, and the distributions of points of different colors are given by χ Lipschitz-bounded measures. More precisely, the distribution of points of color i is given by a probability measure μ_i supported on $[0, \Delta]^d$ with a C-Lipschitz density function[2] for some $C > 0$ that is upper-bounded by 1. Given a set of k unit balls $\mathscr{B}_1, \ldots, \mathscr{B}_k \subset \mathbb{R}^d$, the number of points p_i of color i covered by these balls is given by $\mu_i\big(\bigcup_{i=1}^{k} \mathscr{B}_i\big)$, and the total number of points covered by these balls is given by $\sum_{i=1}^{\chi} \mu_i\big(\bigcup_{i=1}^{k} \mathscr{B}_i\big)$. This variant has an *almost* optimal polynomial-time approximation algorithm for *fixed* d and under some *mild* assumption on OPT (see Sect. 12).

[2] A function $f : \Delta \mapsto \mathbb{R}$ for some subset Δ of real numbers is C-Lipschitz provided $|f(x) - f(y)| \leq C\,|x - y|$ for all real numbers $x, y \in \Delta$.

3 Sketch of Application Scenarios

FMC and its variants are *core* abstractions of many data-driven societal domain applications. We present three *diverse* categories of applications and highlight the *real-world* fairness issues addressed by our problem formulations (leaving other applications in the cited references).

Service/Facility Allocation

One of the most common data based policy decisions is assigning services/facilities across different places, *e.g.*, placing schools [33], bus stops, or police/fire stations, choosing a few hospitals for specific medical facilities or services, or deciding where to put cell-phone towers. Of course, a major objective in such assignments is to serve the *maximum* number of people (*i.e.*, maximize the *coverage*). Unfortunately, historical discriminations, such as *redlining* [29], through their long drawn-out effects of manifestations in different aspects of public policy are *still* hurting the minorities. As a result, blindly optimizing for maximum coverage biases the assignment against *equitable* distribution of services. Below are examples of two *real* cases that further underline the importance of fairness while maximizing coverage:

- *Bike sharing:* As more and more cities adopt advanced transportation systems such as bike-sharing, concerns such as equity and fairness arise with them [49]. For instance, according to [23] the bike-sharing network at NYC neglects many low-income neighborhoods and communities of color while giving the priority to well-to-do neighborhoods. Here the location of bike stations (or bikes) determines the set of people that will have access to the service, perpetuating the unhealthy cycle of lack of transportation, movement, *etc*.
- *Delivery services for online shopping:* Online shopping has by now gained a major share of the shopping market. Platforms such as Amazon provide services such as same-day delivery to make e-shopping even more convenient to their customers. While Amazon's main aim is to maximize the number of customers covered by this service, by not considering fairness it demonstrably *failed* to provide such services for predominantly black communities [28, 45].

Data Integration

Combining multiple data sources to augment the power of any individual data source is a popular method for data collection. Naturally the main objective of data integration is to collect ("cover") a maximum number of data points. However, failing to include an adequate number of instances from minorities, known as *population bias*, in datasets used for training machine learning models is a major reason for model unfairness [4, 40]. For example, image recognition and motion detection services by Google [38] and HP [43] with a reasonable overall performance failed to tag/detect African Americans since their training datasets did not include enough instances from this minority group. While solely optimizing for coverage may result in biased datasets, considering fairness for integration may help remove population bias.

Targeted advertisement

Targeted advertising is popular in social media. Consider a company that wants to target its "potential customers". To do so, the company needs to select a set of features (such as "single" or "college student") that specify the groups of users to be targeted. Of course, the company wants to maximize coverage over its customers. However, solely optimizing for coverage may result in incidents such as racism in the Facebook advertisements [2] or sexism in the job advertisements [27]. Thus, a desirable goal for the company would be to select the keywords such that it provides fair coverage over users of diverse demographic groups.

4 Review of Prior Related Works

To the best of our knowledge, FMC in its full generalities has *not* been separately investigated before. However, there are several prior lines of research that conceptually intersect with FMC.

Maximum *k*-set coverage and *k*-node coverage problems

The maximum k-set coverage and k-node coverage problems are the same as the FMC and NODE- FMC problems, respectively, without element colors and without coloring constraints. These problems have been extensively studied in the algorithmic literature, *e.g.*, see [1, 16, 26] for k-set coverage and [3, 17, 20–22, 24, 35] for k-node coverage. A summary of these results is as follows:

k-set coverage: The best approximation algorithm for k-set coverage is a deterministic algorithm that has an approximation ratio of $\max\left\{\rho(f),\ \rho(k)\right\} > 1 - 1/\mathbf{e}$ [1, 26]. On the inapproximability side, assuming $\mathsf{P} \neq \mathsf{NP}$ an asymptotically optimal inapproximability ratio of $1 - 1/\mathbf{e} + \varepsilon$ (for any $\varepsilon > 0$) is known for any polynomial-time algorithm [16].

k-node coverage: The best approximation for k-node coverage is a randomized algorithm that has an approximation ratio of 0.7504 with high probability [17, 24]. On the inapproximability side, k-node coverage is NP-complete even for bipartite graphs [3], and cannot be approximated within a ratio of $1 - \varepsilon$ for some (small) constant $\varepsilon > 0$ [31, 42]. More recently, Manurangsi [35] provided a semidefinite programming based approximation algorithm with an approximation ratio of 0.92, and Austrin and Stankovic [6] used the results in [5] to provide an *almost* matching upper bound of $0.929 + \varepsilon$ (for any $\varepsilon > 0$) on the approximation ratio of any polynomial time algorithm assuming the *unique games conjecture* is true. There is also a significant body of prior research on the fixed parameter tractability issues for the k-node coverage problem: for example, k-node coverage is *unlikely* to allow an FPT algorithm as it is $W[1]$-hard [20], but Marx designed an FPT *approximation scheme* in [36] whose running time was subsequently improved in Gupta, Lee and Li in [21, 22].

However, the coloring constraints make FMC *fundamentally different* from the maximum set or node coverage problems. Below we point out some of the significant

aspects of these differences. For comparison purposes, for an instance of FMC let $\mathsf{OPT}_{\text{coverage}}$ denote the objective value of an optimal solution for the corresponding maximum k-set coverage problem for this instance by ignoring element colors and coloring constraints.

Existence of a feasible solution: For the maximum k-set coverage problem, a feasible solution trivially exists for any k. However, a valid solution for FMC(χ, k) may *not* exist for some or all k even if $\chi = 2$ and in fact our results (Lemma 1) show that even deciding if there exists a valid solution is NP-complete. The NP-completeness result holds even if $f = 1$ (*i.e.*, the sets are mutually *disjoint*); note that if $f = 1$ then it is *trivial* to compute an optimal solution to the maximum k-set coverage problem. That is why for algorithmic purposes we will assume the existence of at least *one* feasible solution[3] and for showing computational hardness results we will show the existence of at least one trivial feasible solution.

Number of covered elements: The number of covered elements and the corresponding selected sets in an optimal solution in FMC can differ vastly from that in the maximum k-set coverage problem on the same instance. The reason for the discrepancy is because in FMC one may need to select fewer covered elements to satisfy the coloring constraints.

Exactly k sets vs. at most k sets: For the maximum k-set coverage problem any solution trivially can use exactly k sets and therefore there is *no* change to the solution space whether the problem formulation requires exactly k sets or at most k sets. However, the corresponding situation for FMC is different since it may be non-trivial to convert a feasible solution containing $k' < k$ sets to one containing exactly k sets because of the coloring constraints.

Discrepancy minimization problems

Informally, the discrepancy minimization problem for set systems (MIN- DISC) is orthogonal to unweighted FMC. Often MIN- DISC is studied in the context of *two* colors, say red and blue, and is defined as follows. Like unweighted FMC we are given m sets over n elements. However, unlike FMC element colors are *not* given *a priori* but the goal is to color every element red or blue to minimize the maximum *discrepancy* over all sets, where the discrepancy of a set is the absolute difference of the number of red and blue elements it contains. The Beck-Fiala theorem [8] shows that the discrepancy of any set system is at most $2f$, Spencer showed in [44] that the discrepancy of any set system is $O(\sqrt{n \log(2m/n)})$, Bansal provided a randomized polynomial time algorithm achieving Spencer's bound in [7], and a deterministic algorithm with similar bound was provided in [34]. On the lower bound side, it is possible to construct set systems such that the discrepancy is $\Omega(\sqrt{n})$ [12]. For generalization of the formulation to *more* than two colors and corresponding results, see for example [14, 15, 47].

[3] Actually, our LP-relaxation based algorithms require only the existence of a feasible *fractional* solution but we cannot say anything about the approximation ratio in the absence of a feasible integral solution.

Maximization of non-decreasing submodular set functions with linear inequality constraints

Kulik *et al.* [30] provided approximation algorithms for maximizing a non-decreasing submodular set function subject to multiple linear **inequality** constraints over the elements. Unfortunately, because the linear constraints in SEGR- FMC are **equality** constraints, SEGR- FMC *cannot* be put in the framework of [30] and the approximation algorithms in [30] do not directly apply to SEGR- FMC.

5 Summary of Our Contribution

5.1 Feasibility Hardness Results

Obviously FMC (*resp.*, NODE- FMC) obeys all the inapproximability results for the maximum k-set coverage (*resp.*, k-node coverage) problem. We show in Lemma 1 that determining feasibility of FMC instances is NP-complete even under very restricted parameter values; the proofs cover (or can be easily modified to cover) all the special cases of FMC investigated in this paper. However, our subsequent algorithmic results show that even the existence of *one* feasible solution gives rise to non-trivial approximation bounds for the objective and the coloring constraints.

5.2 Algorithmic Results

A summary of our algorithmic results is shown in Table 1. Based on the discussion in the previous section, all of our algorithms assume that at least one feasible solution for the FMC instance exists.

6 Organization of the Paper and Proof Structures

The rest of the paper is organized as follows.

- In Sect. 7 we present our result in Lemma 1 on the computational hardness of finding a feasible solution of FMC.
- Based on the results in Sect. 7, we need to make some minimal assumptions and need to consider appropriate approximate variants of the coloring constraints. They are discussed in Sect. 8 for the purpose of designing (deterministic or randomized) approximation algorithms.
- In Sect. 9 we design and analyze our LP-relaxation based randomized approximation algorithms for FMC. In particular, in Theorem 1 we employ two different LP-relaxations of FMC and combine three randomized rounding analyses on them to get an approximation algorithm whose approximation qualities depend on the range of relevant parameters.

Table 1 A summary of our algorithmic results. The $O(\cdot)$ notation is used when constants are irrelevant or not precisely calculated. Precise definitions of (deterministic) ε-approximation, randomized ε-approximation and strong randomized ε-approximation appear in Sect. 8

Problem name	Algorithm name & type	Approximation ratio	Coloring constraints approximation: ε-approximation $\varepsilon =$	randomized ε-approximation $\varepsilon =$	strong randomized ε-approximation $\varepsilon =$	Parameter and other restrictions (if any): $\mathsf{OPT}_{\#}$	χ	Other	Theorem
FMC	ALG-LARGE-OPT# $\mathscr{RALG}$	$\rho(f)$	N/A	$3.16f$	$O(f)$	$\Omega(\chi\sqrt{n}\log\chi)$	–	N/A	Theorem 1
	ALG-MEDIUM-OPT# $\mathscr{RALG}$	$\rho(f)$	N/A	$3.16f$	$O(f^2)$	$\Omega(a\chi\log\chi)$ and $O(\chi\sqrt{n\log\chi})$	–	N/A	Theorem 1
	ALG-SMALL-OPT# $\mathscr{RALG}$	$\rho(f)$	N/A	$3.16f$	$O(f^2\sqrt{a\,\chi\mathsf{OPT}_{\#}})$	–	$O(\max\{1, \frac{\log n}{\log m}\})$	N/A	Theorem 1
	ALG-ITER-ROUND $\mathscr{DALG}$	$1/f$	$O(f^2)$	N/A	N/A	–	$O(1)$	At most $k+\frac{\chi-1}{2}$ sets	Theorem 3
		$1/f$	$O(f^2+\chi^2 f)$	N/A	N/A	–	–	At most $k+\chi-1$ sets	Theorem 3
NODE- FMC	ALG-ITER-ROUND $\mathscr{DALG}$	1/2	$4+4\chi$	N/A	N/A	–	$O(1)$	At most $k+\frac{\chi-1}{2}$ sets	Theorem 2

Table 1 continued

Problem name	Algorithm name & type	Approximation ratio	Coloring constraints approximation: ε-approximation $\varepsilon =$	Coloring constraints approximation: randomized ε-approximation $\varepsilon =$	Coloring constraints approximation: strong randomized ε-approximation $\varepsilon =$	Parameter and other restrictions (if any): $\mathrm{OPT}_{\#}$	Parameter and other restrictions (if any): χ	Parameter and other restrictions (if any): Other	Theorem
		1/2	$4+2\chi+4\chi^2$	N/A	N/A	–	–	At most $k+\chi-1$ sets	Theorem 2
Other results: same as FMC with $f=2$									
SEGR- FMC	ALG-GREED-PLUS $\mathscr{DA}$ [g]	ρ	2	N/A	N/A	–	–	At most k sets	Theorem 4
Δ-BAL- FMC	ALG-GREEDY $\mathscr{DA}$ [g]	ρ	$O(\Delta f)$	N/A	N/A	–	–	N/A	Proposition 2
GEOM- FMC	ALG-GEOM $\mathscr{RA}$ [g]	$1-O(\delta)$	N/A	N/A	$1+\delta$	–	–	$d=O(1)$ $\frac{\mathrm{OPT}}{\chi}\geq\eta$	Theorem 5

$\mathscr{DA}$ [g]: deterministic algorithm. $\mathscr{RA}$ [g]: randomized algorithm. N/A: not applicable. –: unrestricted. $\rho(x)=\left(1-1/x\right)^x>1-1/\mathbf{e}$. $\rho=\max\{\rho(f),\ \rho(k)\}>1-1/\mathbf{e}$. δ: any constant in the range (0,1]. η: any constant strictly greater than 1

  - Parts of the algorithm and analysis specific to the three algorithms ALG-LARGE-OPT$_\#$, ALG-MEDIUM-OPT$_\#$ and ALG-SMALL-OPT$_\#$ are discussed in Sects. 9.4, 9.5 and 9.6, respectively.
  - Proposition 1 in Sect. 9.7 shows that the dependence of the coloring constraint bounds in Theorem 1$(e)(i)$–(ii) on f cannot be completely eliminated by better analysis of our LP-relaxations even for $\chi = 2$.

- In Sect. 10 we provide polynomial-time deterministic approximations of FMC via iterated rounding of a new LP-relaxation. Our approximation qualities depend on the parameters f and χ.
  - For better understanding, we first prove our result for the special case NODE-FMC of FMC in Theorem 2 (Sect. 10.1) and later on describe how to adopt the same approach for FMC in Theorem 3 (Sect. 10.2).
  - The proofs for both Theorem 2 and Theorem 3 are themselves divided into two parts depending on whether $\chi = O(1)$ or not.
- In Sect. 11 we provide deterministic approximation algorithms for two special cases of FMC, namely SEGR- FMC and Δ-BAL- FMC.
- Sect. 12 provides the deterministic approximation for GEOM- FMC.

Our proofs are structured as follows. A complex proof is divided into subsections corresponding to logical sub-divisions of the proofs and the algorithms therein. Often we provide some informal intuitions behind the proofs (including some intuition about why other approaches may not work, if appropriate) before describing the actual proofs.

7 Computational Hardness of Finding a Feasible Solution of FMC

We show that determining if a given instance of FMC has even one feasible solution is NP-complete even in very restricted parameter settings. The relevant parameters of importance for FMC is a, f and χ; Lemma 1 shows that the NP-completeness result holds even for very small values of these parameters.

Lemma 1 *Determining feasibility of an instance of* FMC *of n elements is NP-complete even with the following restrictions:*

- *the instances correspond to the unweighted version,*
- *the following combinations of maximum set-size a, frequency f and number of colors χ are satisfied:*

  (a) $f \in \{1, 3\}$, *all but one set contains exactly* 3 *elements and all* $\chi \geq 2$,
  (b) *the instances correspond to* NODE- FMC *(which implies* $f = 2$*),* $a = O(\sqrt{n})$, *and all* $\chi \geq 2$, *or*
  (c) $f = 1$, $a = 3$ *and* $\chi = n/3$.

  Moreover, the following assertions also hold:

- *The instances of* FMC *generated in* (a) *and* (b) *actually are instances of* SEGR- FMC.

- *For the instances of* FMC *generated in* (c), $OPT_{\#} = \chi = n/3$ *and, assuming* $P \neq NP$, *there is no polynomial time approximation algorithm that has either a finite approximation ratio or satisfies the coloring constraints in the* ε*-approximate sense* (*cf. Eq.* (5)) *for any finite* ε.

A proof of Lemma 1 appears in the appendix.

Remark 1 It may be tempting to conclude that an approach similar to what is stated below using the k-set coverage problem as a "black box" may make the claims in Lemma 1 completely obvious. We simply take any hard instance of k-set coverage and equi-partition the universe arbitrarily into χ many color classes and let this be the corresponding instance of FMC. Using a suitable standard reductions of NP-hardness the k-set coverage problem, if there is a feasible solution of FMC then the k sets trivially cover the entire universe and thus trivially satisfy the color constraints but otherwise one may (incorrectly) claim that no k sets cover the universe and so the fairness constraints cannot be satisfied. Additionally, one may be tempted to argue that if one takes a k-set coverage problem instance with any additional structure (*e.g.*, bounded occurrence of universe elements) then the coloring does not affect this additional property at all and hence the property is retained in the k-set coverage problem instance with color constraints.

However, such a generic reduction will fail because it is incorrect and because it will not capture all the special parameter restrictions imposed in Lemma 1. For example:

- Even though the k sets may not cover the entire universe, it is still possible that they may satisfy the color constraints. For example, consider the following instance of the k-set coverage problem:

$$\mathscr{U} = \{u_1, u_2, u_3, u_4, u_5, u_6\},\ k = 2,\ \mathscr{S}_1 = \{u_1, u_2\},\ \mathscr{S}_2 = \{u_3, u_4\},\ \mathscr{S}_3 = \{u_5, u_6\}$$

  Suppose we select the equi-partition $\{u_1, u_3, u_5\}, \{u_2, u_4, u_6\}$, thus setting $\mathscr{C}(u_1) = \mathscr{C}(u_3) = \mathscr{C}(u_5) = 1$ and $\mathscr{C}(u_2) = \mathscr{C}(u_4) = \mathscr{C}(u_6) = 2$. Then any two selected sets will satisfy the coloring constraints.
- Consider the requirements of the FMC instances in part (c). Since $f = 1$, every element occurs in exactly one set and thus the set systems for the k-set coverage problem form a partition of the universe. Such an instance cannot be a hard instance of k-set coverage since it admits a trivial polynomial time solution: sort the sets in non-decreasing order of their cardinalities and simply take the first k sets.

8 Relaxing Coloring Constraints for Algorithmic Designs

Based on Lemma 1 we need to make the following *minimal* assumptions for the purpose of designing approximation algorithms with finite approximation ratios:

(i) We assume the existence of at least one feasible solution for the given instance of FMC.

(ii) We assume that $\mathsf{OPT}_{\#}$ is sufficiently large compared to χ, *e.g.*, $\mathsf{OPT}_{\#} \geq c\chi$ for some large constant $c > 1$.

Lemma 1 and the example in Fig. 1 also show that satisfying the color constraint *exactly* (*i.e.*, requiring p_i/p_j to be exactly equal to 1 for all i and j) need to be relaxed for the purpose of designing efficient algorithms since non-exact solutions of FMC may not satisfy these constraints exactly. We define an (*deterministic*) ε-*approximate coloring* of FMC (for some $\varepsilon \geq 1$) to be a coloring that satisfies the coloring constraints in the following manner:

$$\textbf{Deterministic}\varepsilon-\textbf{approximatecoloring}: \quad \forall\, i, j \in \{1, \ldots, \chi\} : \; p_i \leq \varepsilon\, p_j \tag{5}$$

Note that (5) automatically implies that $p_i \geq p_j/\varepsilon$ for all i and j. Thus, in our terminology, a 1-approximate coloring corresponds to satisfying the coloring constraints exactly. Finally, if our algorithm is randomized, then the p_j's could be a random values, and then we will assume that the relevant constraints will be satisfied *in expectation* or *with high probability* in an appropriate sense. More precisely, (5) will be modified as follows:[4]

Randomized ε-approximate coloring:

$$\forall\, i, j \in \{1, \ldots, \chi\} \; : \; \mathbb{E}[p_i] \leq \varepsilon\, \mathbb{E}[p_j] \tag{5$'$}$$

Randomized strong ε-approximate coloring:

$$\bigwedge_{i,j \in \{1,\ldots,\chi\}} \left(\Pr\left[p_i \leq \varepsilon\, p_j \right] \right) \geq 1 - o(1) \tag{5$''$}$$

Unless otherwise stated explicitly, our algorithms will select exactly k sets.

9 LP-Relaxation Based Randomized Approximation Algorithms for FMC

If k is a constant then we can solve FMC(χ, k) *exactly* in polynomial (*i.e.*, $O(n^k)$) time by exhaustive enumeration, so we assume that k is at least a sufficiently large constant. In this section we will employ two slightly different LP-relaxations of FMC and combine three randomized rounding analyses on them to get an approximation algorithm whose approximation qualities depend on the range of various relevant parameters. The combined approximation result is stated in Theorem 1. *In the proof of this theorem no serious attempt was made to optimize most constants since we are mainly interested in the asymptotic nature of the bounds, and to simplify exposition constants have been over-estimated to get nice integers*. In the statement of Theorem 1 and in its proof we will refer to the three algorithms corresponding to the two LP-relaxations as ALG-SMALL-OPT$_{\#}$, ALG-MEDIUM-OPT$_{\#}$ and ALG-LARGE-OPT$_{\#}$.

[4] We do not provide a bound on $\mathbb{E}[p_i/p_j]$ since $p_i/p_j = \infty$ when $p_j = 0$ and p_j may be zero with a strictly positive probability, and for arbitrary χ selecting a set individually for each to avoid this situation in our randomized algorithms may select too many sets.

Theorem 1 *Suppose that the instance of* FMC(χ, k) *has n elements and m sets. Then, we can design three randomized polynomial-time algorithms* ALG-SMALL-OPT$_\#$, ALG-MEDIUM-OPT$_\#$ *and* ALG-LARGE-OPT$_\#$ *with the following properties:*

(a) *All the three algorithms select k sets (with probability 1).*
(b) *All the three algorithms are randomized $\rho(f)$-approximation for* FMC*, i.e., the expected total weight of the selected elements for both algorithms is at least $\rho(f) > 1 - 1/e$ times OPT.*
(c) *All the three algorithms satisfy the randomized ε-approximate coloring constraints (cf. Inequality (5′)) for $\varepsilon = O(f)$, i.e., for all $i, j \in \{1, \ldots, \chi\}$, $\frac{\mathbb{E}[p_i]}{\mathbb{E}[p_j]} \leq \frac{2f}{\rho(f)} < 3.16f$.*
(d) *The algorithms satisfy the strong randomized ε-approximate coloring constraints (cf. Eq. (5″)), i.e., $\bigwedge_{i,j\in\{1,\ldots,\chi\}} \left(\Pr\left[p_i \leq \varepsilon p_j\right]\right) \geq 1 - o(1))$ for the values of ε, $\mathsf{OPT}_\#$ and χ as shown below:*

	ε	Range of $\mathsf{OPT}_\#$	Range of χ	Algorithm
(*i*)	$O(f)$	$\Omega(\chi\sqrt{n}\log\chi)$	unrestricted	ALG-LARGE-OPT$_\#$
(*ii*)	$O(f^2)$	$\Omega(a\chi\log\chi)$ and $O(\chi\sqrt{n\log\chi})$	unrestricted	ALG-MEDIUM-OPT$_\#$
(*iii*)	$O(f^2\sqrt{a\,\chi\,\mathsf{OPT}_\#})$	unrestricted	$\chi = O(\max\{1, \frac{\log n}{\log m}\})$	ALG-SMALL-OPT$_\#$

Remark 2 Note that the high-probability $\varepsilon = O(f)$ bound in Theorem 1(d)(i) is asymptotically *the same* as the "ratio of expectation" bound in Theorem 1(c).

Remark 3 The dependence on f of the bounds for ε in Theorem 1(d)(i)–(ii) can be contrasted with the Beck-Fiala theorem in discrepancy minimization that shows that the discrepancy of any set system is at most $2f$.

Remark 4 Consider the special case NODE- FMC with $\chi = O(1)$: for this case $f = 2$ and a is equal to the maximum node-degree $\mathsf{deg}_{\max}$ in the graph. The bounds in Theorem 1(d)(i)–(ii) for this special case imply a $O(1)$-approximation of color constraints unless $\mathsf{OPT}_\#$ is not sufficiently large compared to $\mathsf{deg}_{\max}$. To illustrate the bound for smaller $\mathsf{OPT}_\#$ in Theorem 1(d)(iii), if $\mathsf{OPT}_\# = \mathsf{deg}_{\max}^{(1/2)-\varepsilon}$ for some $\varepsilon > 0$ then the approximation bound of the coloring constraints is $O(\mathsf{deg}_{\max}^{1-\varepsilon})$.

A proof of Theorem 1 is discussed in the remaining subsections of this section. The following notations will be used uniformly throughout the proof.

- $x_1, \ldots, x_n \in \{0, 1\}$ and $y_1, \ldots, y_m \in \{0, 1\}$ are the usual *indicator* variables for the elements $u_1, \ldots, u_n$ and the sets $\mathcal{S}_1, \ldots, \mathcal{S}_m$, respectively, Their values in an optimal solution of the LP-relaxation under consideration will be denoted by $x_1^*, \ldots, x_n^*$ and $y_1^*, \ldots, y_m^*$, respectively.
- C_j is the set of all elements colored j for $j \in \{1, \ldots, \chi\}$ in the given instance of FMC.
- $\mathsf{OPT}_{\text{frac}}$ is the optimum value of the objective function of the LP-relaxation under consideration.

maximize	$\sum_{i=1}^{n} w(u_i)x_i$	
subject to	$x_j \leq \sum_{u_j \in \mathscr{S}_\ell} y_\ell$	for $j = 1, \ldots, n$
	$\sum_{\ell=1}^{m} y_\ell = k$	
	$0 \leq x_j \leq 1$	for $j = 1, \ldots, n$
	$0 \leq y_\ell \leq 1$	for $\ell = 1, \ldots, m$

Fig. 2 A well-known LP-relaxation of the element-weighted maximum k-set coverage problem

9.1 An Obvious Generalization of LP-Relaxation of Maximum k-Set Coverage Fails

It is well-known that the LP-relaxation for the element-weighted maximum k-set coverage problem as shown in Fig. 2 followed by a suitable deterministic or randomized rounding provides an optimal approximation algorithm for the problem (*e.g.*, see [1, 37]). A straightforward way to extend this LP-relaxation is to add the following $\chi(\chi - 1)/2$ additional constraints, one corresponding to each pair of colors:

$$\sum_{u_\ell \in C_i} x_\ell = \sum_{u_\ell \in C_j} x_\ell \text{ for } i, j \in \{1, \ldots, \chi\}, i < j$$

Unfortunately, this may not lead to a ε-approximate coloring (*cf.* Eq. (5)) for any non-trivial ε as the following example shows. Suppose our instance of an unweighted FMC(2, 2) has four sets $\mathscr{S}_1 = \{u_1\}$, $\mathscr{S}_2 = \{u_2, \ldots, u_{n-2}\}$, $\mathscr{S}_3 = \{u_{n-1}\}$ and $\mathscr{S}_4 = \{u_n\}$ with the elements u_1 and u_{n-1} having color 1 and all other elements having color 2. Clearly the solution to this instance consists of the sets $\mathscr{S}_3$ and $\mathscr{S}_4$ with $\mathsf{OPT} = 2$. On the other hand, the fractional solution $y_1^* = y_2^* = x_1^* = x_2^* = 1$ and all remaining variables being zero is also an optimal solution of the LP-relaxation, but any rounding approach that does not change the values of zero-valued variables in the fractional solution must necessarily result in an integral solutions in which $p_2/p_1 = n - 3$. The example is easily generalized for arbitrary k.

9.2 ALG-LARGE-OPT$_\#$: Strengthening LP-Relaxation via Additional Inequalities

One problem that the LP-relaxation in Fig. 2 faces when applied to FMC is the following. We would like *each* element-indicator variable x_j to satisfy $x_j^* = \min\left\{1, \sum_{u_j \in \mathscr{S}_\ell} y_\ell^*\right\}$ in an optimum solution of the LP, but this may not be true as shown in the simple example in the previous section. Past researchers have corrected this kind of situation by introducing extra valid inequalities that hold for any solution to the problem but restrict the feasible region of the LP. For example, Carnes and Shmoys in [10] and Carr *et al.* in [11] introduced a set of additional inequalities, which they called the KC (Knapsack Cover) inequalities, to strengthen the integrality gaps of certain types of capacitated covering problems. Following their ideas, we add the extra $O(fn)$ "covering inequalities" which are satisfied by *any* integral solution of the LP:

$$x_j \geq y_\ell \text{ for } j = 1, \ldots, n, \ \ell = 1, \ldots, m, \text{ and } u_j \in \mathscr{S}_\ell$$

maximize	$\sum_{i=1}^{n} w(u_i)x_i$	
subject to	$x_j \leq \sum_{u_j \in \mathscr{S}_\ell} y_\ell$	for $j = 1, \ldots, n$
	$\sum_{\ell=1}^{m} y_\ell = k$	
	$x_j \geq y_\ell$	for $j = 1, \ldots, n$, $\ell = 1, \ldots, m$, and $u_j \in \mathscr{S}_\ell$
	$\sum_{i=1}^{n} x_i = \mathsf{OPT}_{\#}$	
	$\sum_{u_\ell \in C_i} x_\ell = \sum_{u_\ell \in C_j} x_\ell$	for $i, j \in \{1, \ldots, \chi\}$, $i < j$
	$0 \leq x_j \leq 1$	for $j = 1, \ldots, n$
	$0 \leq y_\ell \leq 1$	for $\ell = 1, \ldots, m$

Fig. 3 A LP-relaxation for ALG-LARGE-OPT$_{\#}$ with $n + m$ variables and $O\left(fn + \chi^2\right)$ constraints

In addition, we adjust our LP-relaxation in the following manner. Since $\mathsf{OPT}_{\#}$ is an integer from $\{\chi, 2\chi, \ldots, (\lfloor n/\chi \rfloor)\chi\}$, we can "guess" the correct value of $\mathsf{OPT}_{\#}$ by running the algorithm for *each* of the $\lfloor n/\chi \rfloor$ possible value of $\mathsf{OPT}_{\#}$, consider those solutions that maximized its objective function and select that one among these solutions that has the largest value of $\mathsf{OPT}_{\#}$. Thus, we may assume that our LP-relaxation knows the value of $\mathsf{OPT}_{\#}$ *exactly*, and we add the following extra equality:

$$\sum_{i=1}^{n} x_i = \mathsf{OPT}_{\#}$$

The resulting LP-relaxation is shown in its completeness in Fig. 3 for convenience.

9.3 A General Technique for Obtaining Joint High-Probability Statement

Suppose that our randomized LP-relaxation based algorithm guarantees that $\Pr\left[p_i/p_j > \varepsilon\right] < 1/(c\chi^2)$ for some constant $c \geq 3$ independently for all $i, j \in \{1, \ldots, \chi\}$. Then

$$\bigwedge_{i,j \in \{1,\ldots,\chi\}} \Pr\left[p_i \leq \varepsilon\, p_j\right] = 1 - \bigvee_{i,j \in \{1,\ldots,\chi\}} \Pr\left[p_i > \varepsilon\, p_j\right]$$
$$\geq 1 - \sum_{i,j \in \{1,\ldots,\chi\}} \Pr\left[p_i \geq \varepsilon\, p_j\right] > 1 - (\chi^2)\frac{1}{c\chi^2} = 1 - \frac{1}{c} = c'$$

To boost the success probability, we repeat the randomized rounding $c' \ln n$ times, compute the quantity $\sigma = \max_{i,j \in \{1,\ldots,\chi\},\, p_j \neq 0}\{p_i/p_j\}$ in each iteration, and output the solution in that iteration that resulted in the minimum value of σ. It then follows that the selected solution satisfies the strong randomized ε-approximate coloring constraints since $\bigwedge_{i,j \in \{1,\ldots,\chi\}} \Pr\left[p_i \leq \varepsilon\, p_j\right] \geq 1 - \left(1/c'\right)^{c' \ln n} > 1 - 1/n^2$.

9.4 ALG-LARGE-OPT$_{\#}$: Further Details and Relevant Analysis

For our randomized rounding approach, we recall the following result from [46].

Fact 1 [46] *Given numbers $p_1, \ldots, p_r \in [0,1]$ such that $\ell = \sum_{i=1}^{r} p_i$ is an integer, there exists a polynomial-time algorithm that generates a sequence of integers $X_1, \ldots, X_r$ such that* ***(a)*** *$\sum_{i=1}^{r} X_i$ with probability 1,* ***(b)*** *$\Pr[X_i = 1] = p_i$ for all $i \in \{1, \ldots, r\}$, and* ***(c)*** *for any real numbers $\alpha_1, \ldots, \alpha_r \in [0,1]$ the sum $\sum_{i=1}^{r} \alpha_i X_i$ satisfies standard Chernoff bounds.*

We round $y_1^*, \ldots, y_m^*$ to $y_1^+, \ldots, y_m^+$ using the algorithm mentioned in Fact 1; this ensures $\sum_{\ell=1}^{m} y_\ell^+ = \sum_{\ell=1}^{m} y_\ell^* = k$ resulting in selection of *exactly* k sets. This proves the claim in ***(b)***. We round $x_1^*, \ldots, x_n^*$ to $x_1^+, \ldots, x_n^+$ in the following way: for $j = 1, \ldots, n$, if $u_j \in \mathscr{S}_\ell$ for some $y_\ell^+ = 1$ then set $x_j^+ = 1$.

Proof of (b) Our proof of ***(c)*** is similar to that for the maximum k-set coverage and is included for the sake of completeness. Note that $x_j^+ = 0$ if and only if $y_\ell^+ = 0$ for every set $\mathscr{S}_\ell$ containing u_j and thus:

$$\mathbb{E}[x_j^+] = \Pr[x_j^+ = 1] = 1 - \Pr[x_j^+ = 0] = 1 - \prod_{u_j \in \mathscr{S}_\ell} \Pr[y_\ell^+ = 0] = 1 - \prod_{u_j \in \mathscr{S}_\ell} (1 - y_\ell^*)$$

$$\geq 1 - \left(\frac{\sum_{u_j \in \mathscr{S}_\ell} (1 - y_\ell^*)}{f_j}\right)^{f_j} = 1 - \left(1 - \frac{\sum_{u_j \in \mathscr{S}_\ell} y_\ell^*}{f_j}\right)^{f_j} \geq 1 - \left(1 - \frac{\sum_{u_j \in \mathscr{S}_\ell} y_\ell^*}{f}\right)^{f} \quad (6)$$

where we have used inequality (3). If $\sum_{u_j \in \mathscr{S}_\ell} y_\ell^* \geq 1$ then obviously (6) implies $\mathbb{E}[x_j^+] \geq \rho(f) \geq \rho(f) x_j^*$. Otherwise, $x_j^* \leq \sum_{u_j \in \mathscr{S}_\ell} y_\ell^* < 1$ and then by (4) we get

$$\mathbb{E}[x_j^+] \geq 1 - \left(1 - \frac{\sum_{u_j \in \mathscr{S}_\ell} y_\ell^*}{f}\right)^{f} > 1 - \left(1 - \frac{x_j^*}{f}\right)^{f} \geq \rho(f) x_j^* \quad (7)$$

This implies our bound since

$$\mathbb{E}[\textstyle\sum_{i=1}^{n} w(u_i) x_i^+] = \displaystyle\sum_{i=1}^{n} w(u_i) \mathbb{E}[x_i^+] \geq \sum_{i=1}^{n} w(u_i)\, \rho(f) x_j^*$$
$$= \rho(f) \mathsf{OPT}_{\mathrm{frac}} \geq \rho(f) \mathsf{OPT}$$

□

Proof of (c) Note that inequalities (1) and (2) imply $1 - x \leq \mathbf{e}^{-x} \leq 1 - x + (x^2/2)$ for all $x \in [0,1]$. In particular, the following implication holds:

$$\forall\, c > 1\ \forall\, x \in [0, (2/c^2)(c-1)] :\ 1 - x \geq 1 - cx + (c^2 x^2/2) \geq \mathbf{e}^{-cx} \quad (8)$$

We estimate an upper bound on $\mathbb{E}[x_j^+]$ in terms of x_j^* in the following manner:

Case 1: $\exists\, \boldsymbol{\ell}$ such that $\boldsymbol{u_j \in \mathscr{S}_\ell}$ and $\boldsymbol{y_\ell^* > 1/2}$. Thus, $x_j^* \geq y_\ell^* > 1/2$, and $\mathbb{E}[x_j^+] \leq 1 \leq 2x_j^*$.

Case 2: $\boldsymbol{y^*_\ell \le 1/2}$ for every index $\boldsymbol{\ell}$ satisfying $\boldsymbol{u_j \in \mathscr{S}_\ell}$. Note that $x^*_j \ge \left(\sum_{u_j \in \mathscr{S}_\ell} y^*_\ell\right)/f$, and setting $c = 2$ in inequality (8) we get $1 - x \ge \mathbf{e}^{-2x}$ for all $x \in [0, 1/2]$. Now, standard calculations show the following:

$$\begin{aligned}\mathbb{E}[x^+_j] &= 1 - \prod_{u_j \in \mathscr{S}_\ell} (1 - y^*_\ell) \le 1 - \prod_{u_j \in \mathscr{S}_\ell} \mathbf{e}^{-2y^*_\ell} = 1 - \mathbf{e}^{-2\sum_{u_j \in \mathscr{S}_\ell} y^*_\ell} \\ &= 1 - \mathbf{e}^{-2fx^*_j} \le 1 - (1 - 2fx^*_j) = 2fx^*_j\end{aligned}$$

Combining all the cases and using (7), it follows that $\rho(f)x^*_j \le \mathbb{E}[x^+_j] \le \min\{1, 2fx^*_j\}$. Recall that $\sum_{u_\ell \in C_j} x^*_\ell = \mathsf{OPT}_\#/\chi > 0$ for every $j \in \{1, \ldots, \chi\}$. Since $\mathbb{E}[p_j] = \mathbb{E}[\sum_{u_\ell \in C_j} x^+_\ell] = \sum_{u_\ell \in C_j} \mathbb{E}[x^+_\ell]$, we get the following bounds for all $j \in \{1, \ldots, \chi\}$:

$$\begin{aligned}\rho(f)\frac{\mathsf{OPT}_\#}{\chi} &= \sum_{u_\ell \in C_j} \rho(f)x^*_\ell \le \mathbb{E}[p_j] \\ &= \sum_{u_\ell \in C_j} \mathbb{E}[x^+_\ell] \le \sum_{u_\ell \in C_j} (2f)x^*_\ell = 2f\frac{\mathsf{OPT}_\#}{\chi} \qquad (9)\end{aligned}$$

which gives the bound $\frac{\mathbb{E}[c_i]}{\mathbb{E}[c_j]} \le \frac{2f}{\rho(f)}$ for all $i, j \in \{1, \ldots, \chi\}$. □

Proofs of (d)(i) via Doob martingales Note that the random variables $x^+_1, \ldots, x^+_n$ may *not* be pairwise independent since two distinct elements belonging to the same set are correlated, and consequently the random variables $p_1, \ldots, p_\chi$ also may *not* be pairwise independent. Indeed in the worst case an element-selection variable may be correlated to $(a-1)f$ other element-selection variables, thereby ruling out straightforward use of Chernoff-type tail bounds.

For sufficient large $\mathsf{OPT}_\#$, this situation can be somewhat remedied by using Doob martinagales and Azuma's inequality by finding a suitable ordering of the element-selection variables conditional on the rounding of the set-selection variables. We assume that the reader is familiar with basic definitions and results for the theory of martingales (*e.g.*, see [37, Section 4.4]). Fix an arbitrary ordered sequence $y^+_1, \ldots, y^+_m$ of the set-indicator variables. Call an element-indicator variable x^+_i "settled" at the t^{th} step if and only if $\cup_{u_i \in \mathscr{S}_j}\{y_j\} \nsubseteq \{y^+_1, \ldots, y^+_{t-1}\}$ and $\cup_{u_i \in \mathscr{S}_j}\{y_j\} \subseteq \{y^+_1, \ldots, y^+_t\}$. The elementary event in our underlying sample space $\boldsymbol{\Omega}$ are all possible 2^n assignments of 0-1 values to the variables $x^+_1, \ldots, x^+_n$. For each $t \in \{1, \ldots, m\}$, let V_t be the subset of element-selection variables whose values are settled at the t^{th} step, let π_t be an arbitrary ordering of the variables in V_t, and let us relabel the element-indicator variable names so that $x^+_1, x^+_2, \ldots, x^+_n$ is the ordering of all element-selection variables given by the ordering $\pi_1, \ldots, \pi_m$. For each $t \in \{0, 1, \ldots, m\}$ and each $w_1, \ldots, w_t \in \{0, 1\}$, let $\mathbf{B}_{w_1,\ldots,w_t}$ denote the event that $y^+_j = w_j$ for $j \in \{1, \ldots, t\}$. Let $x^+_1, x^+_2, \ldots, x^+_{q_t}$ be the union of set of all q_t element-indicator variables that are settled at the i^{th} step over all $i \in \{1, \ldots, t\}$, and suppose that the event $\mathbf{B}_{w_1,\ldots,w_t}$ induces the following assignment of values to the element-indicator variables: $x^+_1 = b_1, \ldots, x^+_{q_t} = b_{q_t}$ for some

$b_1, \ldots, b_{q_t} \in \{0,1\}$. Define the block $\mathbf{B}'_{w_1,\ldots,w_t} \subseteq \boldsymbol{\Omega}$ induced by the event $\mathbf{B}_{w_1,\ldots,w_t}$ as $\mathbf{B}'_{w_1,\ldots,w_t} = \{b_1, \ldots, b_{q_t} r_{q_t+1} \ldots r_n \mid r_{q_t+1}, \ldots, r_n \in \{0,1\}\}$. Letting $\mathbb{F}_t$ be the σ-field generated by the partition of $\boldsymbol{\Omega}$ into the blocks $\mathbf{B}'_{w_1,\ldots,w_t}$ for each $w_1, \ldots, w_t \in \{0,1\}$, it follows that $\mathbb{F}_0, \mathbb{F}_1, \ldots, \mathbb{F}_m$ form a filter for the σ-field $(\boldsymbol{\Omega}, 2^{\boldsymbol{\Omega}})$. Suppose that $\mathscr{U}$ contains $n_i > 0$ elements of color i, let $x^+_{\alpha_1}, \ldots, x^+_{\alpha_{n_i}}$ be the ordered sequence of the element-selection variables for elements of color i determined by the subsequence of these variables in the ordering $x^+_1, \ldots, x^+_n$, and suppose that $x^+_{\alpha_j}$ was settled at the $t^{\text{th}}_{\alpha_j}$ step. Let $X^i = \sum_{j=1}^{n_i} x^+_{\alpha_j}$, and define the Doob martingale sequence $X_0, X_1, \ldots, X_{n_i}$ where $X_0 = \mathbb{E}[X^i] = \sum_{j=1}^{n_i} \mathbb{E}[x^+_{\alpha_j}]$, and $X_\ell = \mathbb{E}[X^i \mid y^+_1, \ldots, y^+_{t_{\alpha_\ell}}]$ for all $\ell \in \{1, \ldots, n_i\}$. Since $n_i < n$, $X_{n_i} = X^i$ and $|X_\ell - X_{\ell-1}| \leq 1$ for for all $\ell \in \{1, \ldots, n_i\}$, by Azuma's inequality (for any $\Delta > \mathbf{e}$),

$$
\begin{aligned}
&\Pr\left[\left|\textstyle\sum_{j=1}^{n_i} x^+_{\alpha_j} - \sum_{j=1}^{n_i} \mathbb{E}[x^+_{\alpha_j}]\right| \geq 3\sqrt{\ln \Delta}\sqrt{n}\right] \\
&\leq \Pr\left[\left|\textstyle\sum_{j=1}^{n_i} x^+_{\alpha_j} - X_0\right| \geq 3\sqrt{\ln \Delta}\sqrt{n_i}\right] \\
&= \Pr\left[|X_{n_i} - X_0| \geq 3\sqrt{\ln \Delta}\sqrt{n_i}\right] \leq \mathbf{e}^{-4\ln\Delta} = \Delta^{-4} \\
&\Rightarrow \Pr\left[\textstyle\sum_{j=1}^{n_i} \mathbb{E}[x^+_{\alpha_j}] - 3\sqrt{\ln \Delta}\sqrt{n} < \sum_{j=1}^{n_i} x^+_{\alpha_j} < \sum_{j=1}^{n_i} \mathbb{E}[x^+_{\alpha_j}] + 3\sqrt{\ln \Delta}\sqrt{n}\right] \\
&> 1 - \Delta^{-4}
\end{aligned}
$$

Recall from (9) that $(1 - 1/\mathbf{e})\frac{\mathsf{OPT}_\#}{\chi} < \sum_{j=1}^{n_i} \mathbb{E}[x^+_{\alpha_j}] \leq 2f\frac{\mathsf{OPT}_\#}{\chi}$, and the inequality $\sum_{j=1}^{n_i} \mathbb{E}[x^+_{\alpha_j}] \geq 6\sqrt{\ln \Delta}\sqrt{n}$ is therefore satisfied provided $\mathsf{OPT} \geq 6\big(1 - 1/\mathbf{e}\big)^{-1}(\sqrt{n \ln \Delta})\chi$. Setting $\Delta = 2\chi$ and remembering that $p_i = \sum_{j=1}^{n_i} x^+_{\alpha_j}$, we get the following bound for any $i \in \{1, \ldots, \chi\}$:

$$
\Pr\left[\tfrac{1-1/\mathbf{e}}{2}\left(\frac{\mathsf{OPT}_\#}{\chi}\right) < p_i < 3f\left(\frac{\mathsf{OPT}_\#}{\chi}\right)\right] > 1 - (1/16)\chi^{-4}
$$

and therefore $\Pr\left[\frac{p_i}{p_j} \leq \frac{6}{1-1/\mathbf{e}} f\right] > 1 - (1/8)\chi^{-4}$. Thus, it follows that

$$
\begin{aligned}
\bigwedge_{i,j\in\{1,\ldots,\chi\}} \Pr\left[p_i \leq \tfrac{6}{1-1/\mathbf{e}}\, p_j\right] &= 1 - \bigvee_{i,j\in\{1,\ldots,\chi\}} \Pr\left[p_i > \tfrac{6}{1-1/\mathbf{e}}\, p_j\right] \\
&\geq 1 - \sum_{i,j\in\{1,\ldots,\chi\}} \Pr\left[p_i \geq \tfrac{6}{1-1/\mathbf{e}}\, p_j\right] > 1 - (\chi^2)\tfrac{1}{8\chi^4} \geq \tfrac{1}{32}
\end{aligned}
$$

This implies our claim in $(d)(i)$ using the technique in Sect. 9.3. □

9.5 ALG-MEDIUM-OPT$_\#$: Details and Proofs of Relevant Claims in (*a*)–(*c*) and (*d*)(*ii*)

ALG-MEDIUM-OPT$_\#$: **idea behind the modified LP-relaxation and approach**

maximize	$\sum_{i=1}^{n} w(u_i)x_i$	
subject to	$x_j \leq \sum_{u_j \in \mathscr{S}_\ell} y_\ell$	for $j = 1, \ldots, n$
	$\sum_{\ell=1}^{m} y_\ell = k$	
	$x_j \geq y_\ell$	for $j = 1, \ldots, n$, $\ell = 1, \ldots, m$, and $u_j \in \mathscr{S}_\ell$
	$\sum_{i=1}^{n} x_i = \mathsf{OPT}_{\#}$	
	$\sum_{i=1}^{m} v_{i,j} y_i = \mathfrak{h}_j \mathsf{OPT}_{\#}$	for $j \in \{1, \ldots, \chi\}$
	$\sum_{u_\ell \in C_i} x_\ell = \sum_{u_\ell \in C_j} x_\ell$	for $i, j \in \{1, \ldots, \chi\}$, $i < j$
	$0 \leq x_j \leq 1$	for $j = 1, \ldots, n$
	$0 \leq y_\ell \leq 1$	for $\ell = 1, \ldots, m$
	$1/\chi \leq \mathfrak{h}_j \leq f/\chi$	for $j = 1, \ldots, \chi$

Fig. 4 A modified LP-relaxation for ALG-MEDIUM-OPT$_{\#}$ with with $n + m + \chi$ variables and $O\left(fn + \chi^2\right)$ constraints

A limitation of ALG-LARGE-OPT$_{\#}$ is that we could not use Fact 1 of Srinivasan to the fullest extent. Although Fact 1 guaranteed that the set-indicator variables are negatively correlated and hence Chernoff-type tail bounds can be applied to them due to the result by Panconesi and Srinivasan [41], our coloring constraints are primarily indicated by element-indicator variables which depend *implicitly* on the set-indicator variables. In fact, it is not difficult to see that the element-indicator variables are *not* negatively correlated in the sense of [41, 46][5] even if the set-indicator variables are negatively correlated.

Our idea is to remedy the situation by expressing the coloring constraints also by set-indicator variables and use the element-indicator variables to *implicitly* control the set-indicator variables in these coloring constraints. This will also necessitate using additional variables.

A modification of the LP-relaxation in Fig. 3

To begin, we quantify the number of elements of different colors in a set $\mathscr{S}_i$ using the following notation: for $j \in \{1, \ldots, \chi\}$, let $v_{i,j}$ be the number of elements in $\mathscr{S}_i$ of color j. Note that $0 \leq v_{i,j} \leq a$. Fix an optimal integral solution of FMC(χ, k) covering $\mathsf{OPT}_{\#}$ elements and a color value j, and consider the following two quantities: $\mathscr{A} = \sum_{u_\ell \in C_j} x_\ell$ and $\mathscr{B} = \sum_{i=1}^{m} v_{i,j}\, y_i$. Note that $\mathscr{A} = \frac{\mathsf{OPT}_{\#}}{\chi}$, $\mathscr{B} \in \{k, k+1, \ldots, ak\}$, and $\mathscr{A} \leq \mathscr{B} \leq f\mathscr{A}$ by definition of f. Thus, $\mathscr{B} = \mathfrak{h}_j \mathsf{OPT}_{\#}$ is satisfied by a $\mathfrak{h}_j$ that is a rational number from the set $\left\{\frac{1}{\chi}, \frac{1}{\chi} + \frac{1}{\mathsf{OPT}_{\#}}, \frac{1}{\chi} + \frac{2}{\mathsf{OPT}_{\#}}, \ldots, \frac{f}{\chi} - \frac{1}{\mathsf{OPT}_{\#}}, \frac{f}{\chi}\right\}$. We will use the LP-relaxation in Fig. 3 with χ additional variables $\mathfrak{h}_1, \ldots, \mathfrak{h}_\chi$ and the following additional constraints

$$\begin{array}{ll} \sum_{i=1}^{m} v_{i,j}\, y_i = \mathfrak{h}_j \mathsf{OPT}_{\#} & \text{for } j \in \{1, \ldots, \chi\} \\ 1/\chi \leq \mathfrak{h}_j \leq f/\chi & \text{for } j = 1, \ldots, \chi \end{array}$$

For reader's convenience, the new LP-relaxation in its entirety is shown in Fig. 4.

[5] The binary random variables $z_1, \ldots, z_r \in \{0, 1\}$ are called negatively correlated in [41, 46] if the following holds: $\forall\, I \subseteq \{1, \ldots, r\} : \Pr\left[\wedge_{i \in I}(z_i = 0)\right] \leq \prod_{i \in I} \Pr\left[z_i = 0\right]$ and $\Pr\left[\wedge_{i \in I}(z_i = 1)\right] \leq \prod_{i \in I} \Pr\left[z_i = 1\right]$

Analysis of the modified LP-relaxation

By our assumption on $\mathfrak{h}_j$'s, the LP-relaxation has a feasible solution. We use the same randomized rounding procedure (using Fact 1) as in Sect. 9.4 for ALG-LARGE-OPT$_\#$. The proofs for parts (b)–(d) are the same as before since all prior relevant inequalities are still included. Thus, we concentrate on the proof of $(e)(ii)$. A crucial thing to note is the following simple observation:

> Consider the sum $\Delta = \sum_{i=1}^{m} v_{i,j}\, y_i$ for any assignment of values $y_1, \ldots, y_m \in \{0, 1\}$. Then, the number of elements covered by the sets corresponding to those variables that are set to 1 is between Δ/f and Δ.

Fix a color j. Let $\mathscr{K}_j = \mathfrak{h}_j \mathsf{OPT}_\#$, $\alpha_i = \frac{v_{i,j}}{a}$ and consider the summation $\mathscr{L}_j = \sum_{i=1}^{m} \alpha_i\, y_i^+$. Since $\alpha_i \in [0, 1]$ for all i, by Fact 1 we can apply standard Chernoff bounds [37] for $\mathscr{L}_j$. Note that $\mathbb{E}[\mathscr{L}_j] = \sum_{i=1}^{m} \alpha_i\, y_i^* = \frac{\mathscr{K}_j}{a}$. Assuming $\mathscr{K}_j \geq 16a \ln \chi$, we get the following for the tail-bounds:

$$\Pr\left[\sum_{i=1}^{m} v_{i,j}\, y_i^+ > 5\mathscr{K}_j\right] = \Pr\left[\mathscr{L}_j > 5(\mathscr{K}_j/a)\right] < 2^{-6\mathscr{K}_j/a} \leq 2^{-96\ln\chi} < \chi^{-96}$$
$$\Pr\left[\sum_{i=1}^{m} v_{i,j}\, y_i^+ < \mathscr{K}_j/2\right] = \Pr\left[\mathscr{L}_j < \frac{\mathscr{K}_j/a}{2}\right] < \mathbf{e}^{-\frac{\mathscr{K}_j}{8a}} \leq \mathbf{e}^{-2\ln\chi} = \chi^{-2}$$

Remember that $p_j = \sum_{u_\ell \in C_j} x_\ell^+$ is the random variable denoting the number of elements of color j selected by our randomized algorithm. Since $\frac{1}{f}\sum_{i=1}^{m} v_{i,j}\, y_i^+ \leq c_j \leq \sum_{i=1}^{m} v_{i,j}\, y_i^+$ we get

$$\Pr\left[p_j > 5\mathscr{K}_j\right] \leq \Pr\left[\sum_{i=1}^{m} v_{i,j}\, y_i^+ > 5\mathscr{K}_j\right] < \chi^{-96},$$
$$\Pr\left[p_j < \frac{\mathscr{K}_j}{2f}\right] \leq \Pr\left[\sum_{i=1}^{m} v_{i,j}\, y_i^+ < \frac{\mathscr{K}_j}{2}\right] < \chi^{-2}$$

Note that $\frac{\mathsf{OPT}_\#}{\chi} \leq \mathscr{K}_j = \mathfrak{h}_j \mathsf{OPT}_\# \leq \frac{f\,\mathsf{OPT}_\#}{\chi}$, and therefore $1/f \leq \frac{\mathscr{K}_i}{\mathscr{K}_j} \leq f$ for any two $i, j \in \{1, \ldots, \chi\}$. Let $\mathscr{E}_j$ be the event defined as $\mathscr{E}_j \stackrel{\text{def}}{=} \frac{\mathscr{K}_j}{2f} \leq c_j \leq 5\mathscr{K}_j$. Then for any two $i, j \in \{1, \ldots, \chi\}$ we get

$$\begin{aligned}
\Pr\left[\frac{p_i}{p_j} \leq 10f^2\right] &\geq \Pr\left[\mathscr{E}_i \wedge \mathscr{E}_j\right] = 1 - \Pr\left[\overline{\mathscr{E}_i} \vee \overline{\mathscr{E}_j}\right] \geq 1 - \Pr\left[\overline{\mathscr{E}_i}\right] - \Pr\left[\overline{\mathscr{E}_j}\right] \\
&\geq 1 - \Pr\left[p_i > 5\mathscr{K}_i\right] - \Pr\left[p_i < \mathscr{K}_i/(2f)\right] - \Pr\left[p_j > 5\mathscr{K}_j\right] - \Pr\left[p_j < \mathscr{K}_j/(2f)\right] \\
&> 1 - 2\left(\chi^{-96} + \chi^{-2}\right) \qquad (10)
\end{aligned}$$

The assumption of $\mathscr{K}_j \geq 16a \ln \chi$, is satisfied provided $\mathsf{OPT}_\# \geq 16a\chi \ln \chi$. (10) implies our claim in $(d)(ii)$ using the technique in Sect. 9.3.

9.6 ALG-SMALL-OPT$_{\#}$: Details and Proofs of Relevant Claims in (*a*)–(*c*) and (*d*)(*iii*)

Another modification of the LP-relaxation in Fig. 3

Note that for this case $\chi = O(1)$. Fix an optimal solution for our instance of FMC. Let $\mathscr{A}_i$ be the collection of those sets that contain *at least* one element of color i and let $Z_i = \sum_{\mathscr{S}_\ell \in \mathscr{A}_i} y_\ell$ indicate the number of sets from $\mathscr{A}_i$ selected in an integral solution of the LP; obviously $Z_i \geq 1$. We consider two cases for Z_i depending on whether it is at most $5 \ln \chi$ or not. We cannot know *a priori* whether $Z_i \leq 5 \ln \chi$ or not. However, for our analysis it suffices if we can guess *just one* set belonging to Z_i correctly. We can do this by trying out all relevant possibilities exhaustively in the following manner. Let $\Psi = \{1, \ldots, \chi\}$ be the set of indices of all colors. For each of the $2^\Psi - 1$ subsets Ψ' of Ψ, we "guess" that $Z_i \leq 5 \ln \chi$ if and only if $i \in \Psi'$. Of course, we still do not know one set among these $5 \ln \chi$ subset for each such i, so we will exhaustively try out each of the at most $|\mathscr{A}_i| \leq m$ sets for each i. For every such choice of Ψ' and every such choice of a set $\mathscr{S}_{i_{\Psi'}} \in \mathscr{A}_i$ for each $i \in \Psi'$, we perform the following steps:

- Select the sets $\mathscr{S}_{i_{\Psi'}}$ and their elements for each $i \in \Psi'$ Set the variables corresponding to these sets and elements to 1 in the LP-relaxation in Fig. 3, *i.e.*, set $y_{i_{\Psi'}} = 1$ and $x_j = 1$ for every $i \in \Psi' \mathscr{S}_{i_{\Psi'}}$ and $j \in$. Remove any constraint that is already satisfied after the above step.
- Add the following (at most χ) constraints to the LP-relaxation:

$$\sum_{(u_\ell \in C_i) \wedge (u_\ell \in \mathscr{S}_j)} y_j > 5 \ln \chi \text{ for} i \notin \Psi'$$

Note that the total number of iterations that is needed is at most $O((2m)^\chi)$, which is polynomial provided $\chi = O\big(\max\{1, \frac{\log n}{\log m}\}\big)$.

Analysis of the modified LP-relaxation

We now analyze that iteration of the LP-relaxation that correctly guesses the value of OPT$_{\#}$, the subset $\Psi' \subseteq \Psi$ and the sets $\mathscr{S}_{i_{\Psi'}} \in \mathscr{A}_i$ for each $i \in \Psi'$. As already mentioned elsewhere, the random variables $x_1^+, \ldots, x_n^+$ may *not* be pairwise independent since two distinct elements belonging to the same set are correlated, and consequently the random variables $p_1, \ldots, p_\chi$ also may *not* be pairwise independent. For convenience, let $\mu_i = \mathbb{E}[x_i^+]$ and $\mathscr{E}_i$ denote the event $\mathscr{E}_i \equiv x_i^+ = 1$; note that $(1 - \mathbf{e}^{-1}) x_i^* < \Pr[\mathscr{E}_i] = \mu_i \leq \min\{1, 2 f x_i^*\}$. We first calculate a bound on $\text{cov}(x_i^+, x_j^+)$ for all $i \neq j$ as follows. If x_i^+ and x_j^+ are independent then of course $\text{cov}(x_i^+, x_j^+) = 0$, otherwise

$$\begin{aligned} &- \min\{\mu_i, \mu_j\} \leq -\Pr[\mathscr{E}_i] \Pr[\mathscr{E}_j] \leq \Pr[\mathscr{E}_i \wedge \mathscr{E}_j] - \Pr[\mathscr{E}_i] \Pr[\mathscr{E}_j] = \mathbb{E}[x_i^+ x_j^+] - \mu_i \mu_j \\ &= \text{cov}(x_i^+, x_j^+) \leq \Pr[\mathscr{E}_i \wedge \mathscr{E}_j] - \mu_i \mu_j \leq \min\{\mu_i, \mu_j\} - \mu_i \mu_j < \min\{\mu_i, \mu_j\} \end{aligned}$$

giving the following bounds:

$$- \min\{\rho(f) x_i^*, \rho(f) x_j^*\} \leq \text{cov}(x_i^+, x_j^+) \leq \min\{2 f x_i^*, 2 f x_j^*, 1\} \qquad (11)$$

For notational convenience, let $\mathscr{D}_{i,j} = \{\ell \mid u_i, u_j \in \mathscr{S}_\ell,\ j \neq i\}$ be the indices of those sets in which both the elements u_i and u_j appear, and let $\mathscr{D}_i = \cup_{j=1}^{n} \mathscr{D}_{i,j}$. Note that $|\mathscr{D}_{i,j}| \leq f$, $|\mathscr{D}_i| \leq (a-1)f$, and the random variable x_i^+ is independent of all x_j^+ satisfying $j \notin \mathscr{D}_i$. Using this observation and (11), for any $i \in \{1, \ldots, n\}$ we get

$$\sum_{j=1}^{n} \operatorname{cov}(x_i^+, x_j^+) \leq \sum_{j \in \mathscr{D}_i} \big(\min\{\mu_i,\ \mu_j\}\big) \leq |\mathscr{D}_i|\,\mu_i \leq af\mu_i \leq \min\{2af^2 x_i^*,\ af\}$$

$$\sum_{j=1}^{n} \operatorname{cov}(x_i^+, x_j^+) \geq -\sum_{j \in \mathscr{D}_i} \min\{\mu_i,\ \mu_j\} \geq -\min\Big\{|\mathscr{D}_i|\,\mu_i,\ \rho(f)\sum_{j \in \mathscr{D}_i} x_j^*\Big\} \quad (12)$$

$$\geq -\min\Big\{(a-1)f\,\mu_i,\ \rho(f)\sum_{j \in \mathscr{D}_i} x_j^*\Big\} > -af x_i^* \quad (13)$$

The above bounds can be used to bound the total pairwise co-variance between elements in two same or different color classes as follows. Consider two color classes C_i and C_j ($i = j$ is allowed). Then,

$$\sum_{u_r \in C_i} \sum_{u_s \in C_j} \operatorname{cov}(x_r^+, x_s^+) \leq \sum_{u_r \in C_i} \sum_{j=1}^{n} \operatorname{cov}(x_r^+, x_j^+) \leq \sum_{u_r \in C_i} \min\{2af^2 x_r^*,\ af\}$$

$$= \min\Big\{2af^2 \textstyle\sum_{u_r \in C_i} x_r^*,\ af|C_i|\Big\} \leq \min\Big\{2af^2 \frac{\mathsf{OPT}_\#}{\chi},\ afn\Big\} \quad (14)$$

$$\sum_{u_r \in C_i} \sum_{u_s \in C_j} \operatorname{cov}(x_r^+, x_s^+) \geq -\sum_{u_r \in C_i} \sum_{j \in \mathscr{D}_r} \min\{\mu_r,\ \mu_j\} > -\sum_{u_r \in C_i} af x_r^* = -af\frac{\mathsf{OPT}_\#}{\chi} \quad (15)$$

For calculations of probabilities of events of the form "$p_i > \Delta p_j$", we first need to bound the probability of events "$p_j = 0$" for $j \in \{1, \ldots, \chi\}$. If $j \in \Psi'$ then $\Pr[p_j = 0] = 0$ since at least one set containing an element of color j is always selected. Otherwise, $\sum_{\mathscr{S}_\ell \in \mathscr{A}_j} y_\ell^* > 5 \ln \chi$, and $p_j = 0$ if and only if $y_\ell^+ = 0$ for *every* $\mathscr{S}_\ell \in \mathscr{A}_j$. This gives us the following bound for $j \notin \Psi'$:

$$\Pr[p_j = 0] = \prod_{\mathscr{S}_\ell \in \mathscr{A}_j} \Pr[y_\ell^+ = 0] = \prod_{\mathscr{S}_\ell \in \mathscr{A}_j} (1 - y_\ell^*) \leq \prod_{\mathscr{S}_\ell \in \mathscr{A}_j} \mathbf{e}^{-y_\ell^*}$$
$$= \mathbf{e}^{-\sum_{\mathscr{S}_\ell \in \mathscr{A}_j} y_\ell^*} \leq \mathbf{e}^{-5\ln\chi} = \chi^{-5}$$

Combining both cases, we have $\Pr[p_j = 0] \leq 1/\chi^5$ for all j.

We now can calculate the probabilities of events of the form "$p_i > \Delta p_j$" for $\Delta = \Delta_1 + \Delta_2 \geq 1$, $\Delta_1, \Delta_2 \geq 0$ and $i, j \in \{1, \ldots, \chi\}$ as follows:

$$\begin{aligned}
\Pr\left[p_i \geq \Delta p_j\right] &\leq \Pr\left[p_j = 0\right] + \Pr\left[p_i \geq \Delta p_j \mid p_j \geq 1\right] \\
&\leq \chi^{-5} + \Pr\left[p_i \geq \Delta_1 p_j + \Delta_2 \mid p_j \geq 1\right] \\
&= \chi^{-5} + \frac{\Pr\left[(p_i \geq \Delta_1 p_j + \Delta_2) \wedge (p_j \geq 1)\right]}{\Pr\left[p_j \geq 1\right]} < \chi^{-5} + \frac{\Pr\left[p_i \geq \Delta_1 p_j + \Delta_2\right]}{1 - \Pr\left[p_j = 0\right]} \\
&< \chi^{-5} + \frac{\Pr\left[p_i \geq \Delta_1 p_j + \Delta_2\right]}{1 - \chi^{-5}} \qquad (16)
\end{aligned}$$

For a real number $\zeta > 0$, let $\delta_{i,j} = p_i - \zeta p_j$. We have the following bound on $\mathbb{E}[\delta_{i,j}]$ for all $\zeta \geq 3f$:

$$\mathbb{E}[\delta_{i,j}] = \mathbb{E}[p_i] - \zeta\, \mathbb{E}[p_j] \leq 2f \frac{\mathsf{OPT}_\#}{\chi} - \zeta\, \rho(f) \frac{\mathsf{OPT}_\#}{\chi} < 0$$

Therefore, using Chebyshev's inequality we get (for all $\zeta \geq 3f$ and $\lambda > 1$):

$$\begin{aligned}
\Pr\left[p_i \geq \zeta p_j + \lambda\sqrt{\mathrm{var}(\delta_{i,j})}\right] &= \Pr\left[\delta_{i,j} \geq \lambda\sqrt{\mathrm{var}(\delta_{i,j})}\right] \\
&< \Pr\left[\left|\, \delta_{i,j} - \mathbb{E}[\delta_{i,j}]\,\right| > \lambda\sqrt{\mathrm{var}(\delta_{i,j})}\right] \leq 1/\lambda^2 \qquad (17)
\end{aligned}$$

Using (17) in (16) with $\Delta_1 = \zeta$, $\Delta_2 = \lambda\sqrt{\mathrm{var}(\delta_{i,j})}$ and $\lambda = 10\chi$ we get

$$\begin{aligned}
\Pr\left[p_i < \left(\zeta + 10\chi\sqrt{\mathrm{var}(\delta_{i,j})}\right) p_j\right] &= 1 - \Pr\left[p_i \geq \left(\zeta + 10\chi\sqrt{\mathrm{var}(\delta_{i,j})}\right) p_j\right] \\
&> 1 - \chi^{-5} - \tfrac{\chi^{-100}}{1-\chi^{-5}} > 1 - \chi^{-4} \qquad (18)
\end{aligned}$$

We now calculate a bound on $\mathrm{var}(\delta_{i,j})$ using (14) and (15) as follows:

$$\begin{aligned}
\mathrm{var}(\delta_{i,j}) &= \mathrm{var}(p_i - \zeta p_j) = \mathrm{var}\Big(\sum_{u_\ell \in C_i} x_\ell^+ + \sum_{u_\ell \in C_j} (-\zeta x_\ell^+)\Big) \\
&= \sum_{u_\ell \in C_i} \mathrm{var}\big(x_\ell^+\big) + \zeta^2 \sum_{u_\ell \in C_j} \mathrm{var}\big(x_\ell^+\big) + \sum_{u_r, u_s \in C_i, r \neq s} \mathrm{cov}(x_r^+, x_s^+) + \sum_{u_r, u_s \in C_j, r \neq s} \mathrm{cov}(-\zeta x_r^+, -\zeta x_s^+) \\
&\quad + \sum_{u_r \in C_i, u_s \in C_j} \mathrm{cov}(x_r^+, -\zeta x_s^+) \\
&\leq \sum_{u_\ell \in C_i} \mu_\ell + \zeta^2 \sum_{u_\ell \in C_j} \mu_\ell + 2af^2 \tfrac{\mathsf{OPT}_\#}{\chi} + \zeta^2 \sum_{u_r, u_s \in C_j, r \neq s} \mathrm{cov}(x_r^+, x_s^+) - \zeta \sum_{u_r \in C_i, u_s \in C_j} \mathrm{cov}(x_r^+, x_s^+) \\
&\leq \mathbb{E}[c_i] + \zeta^2 \mathbb{E}[c_j] + 2af^2 \tfrac{\mathsf{OPT}_\#}{\chi} + 2\zeta^2 af^2 \tfrac{\mathsf{OPT}_\#}{\chi} + \zeta af \tfrac{\mathsf{OPT}_\#}{\chi} \\
&\leq 2f \tfrac{\mathsf{OPT}_\#}{\chi} + \zeta^2 2f \tfrac{\mathsf{OPT}_\#}{\chi} + 2af^2 \tfrac{\mathsf{OPT}_\#}{\chi} + \zeta^2 2af^2 \tfrac{\mathsf{OPT}_\#}{\chi} + \zeta af \tfrac{\mathsf{OPT}_\#}{\chi} \leq 4\zeta^2 af^2 \tfrac{\mathsf{OPT}_\#}{\chi} \\
&\Rightarrow \sqrt{\mathrm{var}(\delta_{i,j})} \leq 2\zeta\sqrt{a} f \sqrt{\mathsf{OPT}_\#/\chi} \qquad (19)
\end{aligned}$$

Setting $\zeta = 3f$ and using (19) in (18) we get $\Pr\left[c_i < \left(3f + 60\sqrt{a} f^2 \sqrt{\mathsf{OPT}_\# \chi}\right) c_j\right] > 1 - \chi^{-4}$. This implies our claim in $(d)(iii)$ using the technique in Sect. 9.3.

9.7 Limitations of Our LP-Relaxation: "A Gap of Factor *f*" for Coloring Constraints

The coloring constraint bounds in Theorem 1(*e*)(*i*)–(*ii*) depend on f or f^2 only. It is natural to ask as a possible first direction of improvement whether this dependence can be eliminated or improved by better analysis of our LP-relaxations. Proposition 1 shows that this may *not* be possible even for $\chi = 2$ unless one uses a significantly different LP-relaxation for FMC(χ, k).

Proposition 1 *There exists optimal non-integral solutions of* FMC *with the following property: any rounding approach that does not change the values of zero-valued variables in the fractional solution must necessarily result in an integral solutions in which the color constraints differ by at least a factor of* f.

Proof We will show our result for the LP-relaxation in Fig. 3; proofs for other modified versions of this LP-relaxation are similar. Consider $\alpha \gg 1$ disjoint collections of sets and elements of the following type: for $j \in \{1, \ldots, \alpha\}$, the j^{th} collection consists of a set of $\alpha + 1$ elements $\mathscr{U}^j = \{u_1^j, \ldots, u_{\alpha+1}^j\}$ with $\mathscr{C}(u_1^j) = 1$ and $\mathscr{C}(u_2^j) = \cdots = \mathscr{C}(u_{\alpha+1}^j) = 2$, and the $\alpha + 1$ sets $\mathscr{S}_1^j, \ldots, \mathscr{S}_{\alpha+1}^j$ where $\mathscr{S}_i^j = \mathscr{U}^j \backslash \{u_i^j\}$ for $i \in \{1, \ldots, \alpha + 1\}$ (note that each element u_i^j is in exactly α sets). Add to these collections the additional $2\alpha + 2$ elements u_1^ℓ, u_2^ℓ with $\mathscr{C}(u_1^\ell) = 1$ and $\mathscr{C}(u_2^\ell) = 2$, and the $\alpha + 1$ sets $\mathscr{S}^\ell = \{u_1^\ell, u_2^\ell\}$ for $\ell \in \{\alpha + 1, \ldots, 2\alpha + 1\}$. Note that for our created instance $f = \alpha$. Consider the following two different solutions of the LP-relaxation:

(1) For a non-integral solution, let $y_1^j = \cdots = y_{\alpha+1}^j = 1/\alpha$, let $x_1^j = 1$ and let $x_2^j = \cdots = x_{\alpha+1}^j = 1/\alpha$ for $j \in \{1, \ldots, \alpha\}$, and set all other variables to zero. This results in a solution with summation of set variables being $\alpha + 1$ (*i.e.*, $\alpha + 1$ sets are selected non-integrally), and summation of element variables being $2\alpha + 2$ (*i.e.*, $2\alpha + 2$ elements are selected non-integrally). Moreover, the summation of element variables with the color value of 1 is precisely the same as summation of element variables with the color value of 2 since both are equal to $\alpha + 1$.

(2) For an integral solution, let $y^\ell = x_1^\ell = x_2^\ell = 1$ for $\ell \in \{\alpha + 1, \ldots, 2\alpha + 1\}$. This also results in a solution in which $\alpha + 1$ sets are selected, the number of elements covered is $2\alpha + 2$ and the number of elements of each color is $\alpha + 1$.

The crucial things to note here is that the two above solutions are disjoint (*i.e.*, non-zero variables in one solution are zero in the other and *vice versa*), and thus any rounding approach for the solution in **(1)** that does not change values of the zero-valued variables results in an integral solution in which the number of elements of color 2 is f times the number of elements of color 1. □

10 A Tale of Fewer Colors: Deterministic Approximation for FMC When χ is "not too large"

In this section we provide polynomial-time deterministic approximations of FMC via the *iterated rounding* technique for LP-relaxations. We assume that the reader is familiar with the basic concepts related to this approach as described, for example, in [32].

Our approximation qualities will depend on the parameters f and χ and the coloring constraint bounds are interesting only if χ is not too large, *e.g.*, no more than, say, poly-logarithmic in n. *For better understanding of the idea, we will first consider the special case* Node- fmc *of* Fmc *for which* $f = 2$, *and later on describe how to adopt the same approach for arbitrary* f. As per the proof of Theorem 1 (see Sect. 9.2) we may assume we know the value of $\mathsf{OPT}_{\#}$ exactly. A main ingredient of the iterated rounding approach is the following "rank lemma".

Fact 2 *(Rank lemma)*[32, Lemma 2.1.4] *Consider any convex polytope* $P \stackrel{def}{=} \{\boldsymbol{x} \in \mathbb{R}^n \mid \mathbf{A}_j \boldsymbol{x} \geq b_j \text{ for } j \in \{1, \ldots, m\}, \boldsymbol{x} \geq \mathbf{0}\}$ *for some* $\mathbf{A}_1, \ldots, \mathbf{A}_m \in \mathbb{R}^n$ *and* $(b_1, \ldots, b_m)^T \in \mathbb{R}^m$. *Then the following property holds for every extreme-point for* P*: the number of any maximal set of linearly independent tight constraints* (*i.e., constraints satisfying* $\mathbf{A}_j \boldsymbol{x} = b_j$ *for some* j) *in this solution equals the number of non-zero variables.*

10.1 Approximating Node-fmc

Theorem 2 *We can design a deterministic polynomial-time approximation algorithm* Alg-iter-round *for* Node- fmc *with the following properties:*

(a) *The algorithm selects* τ *nodes where* $\tau \leq \begin{cases} k + \frac{\chi - 1}{2}, & \text{if } \chi = O(1) \\ k + \chi - 1, & \text{otherwise} \end{cases}$

(b) *The algorithm is a* $\frac{1}{2}$*-approximation for* Node- fmc, *i.e., the total weight of the selected elements is at least* $\mathsf{OPT}/2$.

(c) *The algorithm satisfies the* ε*-approximate coloring constraints* (*cf. Inequality* (5)) *as follows:*

$$\text{for all } i, j \in \{1, \ldots, \chi\}, \ \frac{p_i}{p_j} < \begin{cases} 4 + 4\chi, & \text{if } \chi = O(1) \\ 4 + 2\chi + 4\chi^2, & \text{otherwise} \end{cases}$$

We discuss the proof in the rest of this section. Let $G = (V, E)$ be the given graph, and let $\mathsf{deg}(v)$ denote the degree of node v. Assume that G has no isolated nodes.

10.1.1 The Case of $\chi = O(1)$

Since the problem can be exactly solved in polynomial time by exhaustive enumeration if k is a constant, we can assume k is at least a sufficiently large integer, *e.g.*, assume that $k > 10\chi$.

Initial preprocessing

To begin, we "guess" $\chi + 1$ nodes, say $v_1, \ldots, v_{\chi+1} \in V$ with $\mathsf{deg}(v_1) \leq \mathsf{deg}(v_2) \leq \cdots \leq \mathsf{deg}(v_{\chi+1})$ such that there exists an optimal solution contains these $\chi + 1$ nodes with the following property: "the remaining $k - (\chi + 1)$ nodes in the solution have degree *at most* $\mathsf{deg}(v_1)$". Since there are at most $\binom{n}{\chi+1} = n^{O(1)}$ choices for such $\chi + 1$ nodes, we can try them out in an *exhaustive* fashion. Thus, we only need to analyze that run of our algorithm where the our guess is *correct*. Once

these $\chi+1$ nodes have been selected, we will use the following sets of nodes in $\widehat{V}$ and "incidence-indexed" edges $\widehat{E}$ as input to our algorithm (note that an edge $e \stackrel{\text{def}}{=} \{u, v\}$ may appear as two members (e, u) and (e, v) in $\widehat{E}$ if both u and v are in $\widehat{V}$):

$$\widehat{V} = V \setminus \big(\{v_1, \ldots, v_{\chi+1}\} \cup \{v \mid \text{degree of } v \text{ in } G \text{ is strictly larger than } \mathsf{deg}(v_1)\}\big)$$
$$\widehat{E} = \{(e, u) \mid \big(e \stackrel{\text{def}}{=} \{u, v\} \in E\big) \bigwedge \big(v \notin \{v_1, \ldots, v_{\chi+1}\}\big) \bigwedge \big(u \in \widehat{V}\big)\}$$

Fix an optimal solution $V_{\text{opt}} \subseteq V$ that includes the nodes $v_1, \ldots, v_{\chi+1}$. We next make the following parameter adjustments:

- We update an *estimate* for p_j (the number of edges of color j covered by the optimal solution) from its initial value of $\mathsf{OPT}_{\#}/\chi$ in the following manner. Let μ_j be the number of edges of color j incident on at least one of the nodes in $\{v_1, \ldots, v_{\chi+1}\}$. Consider the quantity $\widehat{q}_j = \sum_{u \in V_{\text{opt}} \setminus \{v_1, \ldots, v_{\chi+1}\}} \big|\big\{e \stackrel{\text{def}}{=} \{u, v\} \in E \mid (v \notin \{v_1, \ldots, v_{\chi+1}\}) \wedge (\mathscr{C}(e) = j)\big\}\big|$. Note that $\sum_{i=1}^{\chi+1} \mathsf{deg}(v_i) \leq 2\mathsf{OPT}_{\#}$ and $\widehat{q}_j$ is an integer in the set $\left\{\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_j, \frac{\mathsf{OPT}_{\#}}{\chi} - \mu_j + 1, \ldots, 2\big(\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_j\big)\right\} \subseteq \{0, 1, 2, \ldots, n\}$ since any edge can be covered by either one or two nodes. Note that there are at most $\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_j + 1 \leq \frac{\mathsf{OPT}_{\#}}{\chi} \leq \frac{n}{\chi}$ possible number of integers values that each $\widehat{q}_j$ may take. Since $\chi = O(1)$, we can try out all possible combinations of $\widehat{q}_j$ values over *all* colors in polynomial time since $(n/\chi)^{\chi} = n^{O(1)}$. Thus, we henceforth assume that we know the *correct* value of $\widehat{q}_j$ for each $j \in \{1, \ldots, \chi\}$. Note that $\widehat{q}_j \geq 0$ since our guess is correct.
- Update k (the number of nodes to be selected) by subtracting $\chi + 1$ from it, and call the new value $\widehat{k}$.
- Update C_i (the set of edges of color i) to be the set of edges *in* $\widehat{E}$ that are of color i.

Yet another LP-relaxation

Let $|\widehat{V}| = \widehat{n}$ and $|\widehat{E}| = \widehat{m}$. We will start with an initial LP-relaxation of Node- fmc on $\widehat{G}$ which will be iteratively modified by our rounding approach. Our LP-relaxation is the following modified version of the LP-relaxation in Fig. 3.

- There is a node indicator variable y_v for every node $v \in \widehat{V}$ and an edge indicator variable $x_{e,u}$ for every edge $(e, u) \in \widehat{E}$; thus we have $\widehat{n} + \widehat{m}$ variables in total.
- Constraints of the form "$x_j \geq y_\ell$" and "$x_j \leq \sum_{u_j \in \mathscr{S}_\ell} y_\ell$" in Fig. 3 are removed now and instead replaced by at most two constraints $x_{e,u} = y_u$ if $y_u \in \widehat{V}$ and $x_{e,v} = y_v$ if $y_v \in \widehat{V}$. This is done so that we can apply the rank lemma in a meaningful way.
- Note that the quantity $\sum_{u \in \widehat{V}} \sum_{e \stackrel{\text{def}}{=} \{u,v\} \in C_i} x_{e,u}$ for each color i is the integer $\widehat{q}_i$ mentioned before.
- To maximize the parameter ranges over which our algorithm can be applied, we replace the $\binom{\chi}{2}$ constraints in Fig. 3 of the form "$\sum_{u_\ell \in C_i} x_\ell = \sum_{u_\ell \in C_j} x_\ell$ for $i, j \in \{1, \ldots, \chi\}, i < j$" by the χ constraints $\sum_{u \in \widehat{V}} \sum_{e \stackrel{\text{def}}{=} \{u,v\} \in C_i} x_{e,u} = q'_i$ for $i \in \{1, \ldots, \chi\}$.

	maximize	$\psi = \sum_{u \in \widehat{V}} \sum_{(e,u) \in \widehat{E}} w(e) x_{e,u}$	
	subject to		
(1)		$x_{e,u} = y_u$	for all $u \in \widehat{V}$ and $(e,u) \in \widehat{E}$
(2)		$\sum_{v \in V_1} y_v = \widehat{k}$	
(3)		$\sum_{u \in \widehat{V}} \sum_{e \overset{\text{def}}{=} \{u,v\} \in C_i} x_{e,u} = \widehat{q}_i$	for all $i \in \{1, \ldots, \chi\}$
(4)		$0 \le x_{e,u} \le 1$	for all $(e,u) \in \widehat{E}$
(5)		$0 \le y_v \le 1$	for all $v \in \widehat{V}$

Fig. 5 The initial LP-relaxation $\mathscr{L} = \mathscr{L}^{(0)}$ for the graph $\widehat{G}$ used in Theorem 2. The iterated rounding approach will successively modify the LP to create a sequence $\mathscr{L}^{(1)}, \mathscr{L}^{(2)}, \ldots$ of LP's

The entire initial LP-relaxation $\mathscr{L}$ for $\widehat{G}$ is shown in Fig. 5 for convenience. Note that the number of constraints in lines **(1)**–**(3)** of Fig. 5 is *exactly* $\widehat{m} + \chi + 1$.

Details of iterated rounding

We will use the variable $t \in \{0, 1, 2, \ldots, n\}$ to denote the iteration number of our rounding, with $t = 0$ being the situation before any rounding has been performed, and we will use a "*superscript* (t)" for the relevant quantities to indicate their values or status *after* the t^{th} iteration of the rounding, *e.g.*, $\widehat{n}^{(0)} = \widehat{n}$ and $\widehat{n}^{(1)}$ is the value of $\widehat{n}$ after the first iteration of rounding. Our iterated rounding algorithm ALG-ITER-ROUND in high level details is shown in Fig. 6, where the following notation is used for brevity for a node $u \in \widehat{V}$:

$$Z_u^{\text{variables}} = \{y_u\} \bigcup \left\{ x_{e,u} \mid (e,u) \in \widehat{E} \right\}$$

For concise analysis of our algorithm, we will use the following notations:

- $W^{\chi+1}$ is the sum of weights of all the edges incident to one or more nodes from the set of nodes $\{v_1, \ldots, v_{\chi+1}\}$.
- $w(X) = \sum_{x_{u,e} \in X} w(e)$ for a subset of variable $X \subseteq \{x_{e,u} \mid (e,u) \in \widehat{E}\}$.
- $W_{\text{ALG}}^{(t)} = w(\widehat{X}_{\text{sol}}^{(t)})$ is the *sum* of weights of the edges whose variables are in $\widehat{X}_{\text{sol}}^{(t)}$ (thus, for example, $W_{Alg}^{(0)} = 0$).
- $\text{OPT}_{\text{frac}}^{(t)}$ is the optimum value of the objective function of the LP-relaxation $\mathscr{L}^{(t)}$ during the t^{th} iteration of rounding.
- $\widehat{p}_i^{(t)}$ is the number of edges of color i selected by ALG-ITER-ROUND *up to and including* the t^{th} iteration of rounding.
- t^{final} is the value of t in the *last* iteration of rounding.

Lemma 2 ALG-ITER-ROUND *terminates after at most n iterations and selects at most* $k + \frac{\chi - 1}{2}$ *nodes.*

Proof For finite termination, it suffices to show that at least one of the three cases in ALG-ITER-ROUND always applies. Consider the first iteration, say when $t = \alpha$, when neither Case 1 nor Case 2 applies. Note that this also implies that $x_{e,u} \notin \{0, 1\}$ for any variable $x_{e,u}$ in $\text{LP}^{(\alpha)}$ since otherwise the variable y_u in $\text{LP}^{(\alpha)}$ will be either 0 or 1 via the equality constraint $y_u = x_{e,u}$ and one of Case 1 or Case 2 will apply. Thus

$(*$ initialization $*)$
$t \leftarrow 0 \quad \widehat{V}_{\text{sol}} \leftarrow \emptyset; \quad \widehat{X}_{\text{sol}} \leftarrow \emptyset; \quad \text{var-count}_{\text{remaining}} \leftarrow \widehat{n} + \widehat{m}; \quad \widehat{r} \leftarrow \widehat{m}; \quad \widehat{n} \leftarrow \widehat{n};$
$(*$ iterations of rounding $*)$
while $(\text{var-count}_{\text{remaining}} \neq 0)$ **do**
  $t \leftarrow t + 1$
  find an *extreme-point* optimal solution of objective value $\mathsf{OPT}_{\text{frac}}^{(t-1)}$ for the LP $\mathscr{L}^{(t-1)}$ (*cf.* Fig. 5)
  begin cases
    Case 1: there exists a variable y_u in the solution such that $y_u = 0$
      $\text{var-count}_{\text{remaining}} \leftarrow \text{var-count}_{\text{remaining}} - \left|Z_u^{\text{variables}}\right|; \ \widehat{n} \leftarrow \widehat{n} - 1$
      $\widehat{r} \leftarrow \widehat{r} - \left|\{x_{e,u} \,|\, x_{e,u} \in Z_u^{\text{variables}}\}\right|; \ \widehat{X}_{\text{sol}} \leftarrow \widehat{X}_{\text{sol}} \cup \{x_{e,u} \,|\, x_{e,u} \in Z_u^{\text{variables}}\}$
      remove the variables in $Z_u^{\text{variables}}$ from $\mathscr{L}^{(t-1)}$, and delete or update the constraints and the objective function to reflect the removal of variables
    Case 2: there exists a variable y_u in the solution such that $y_u = 1$
      $\widehat{V}_{\text{sol}} \leftarrow \widehat{V}_{\text{sol}} \cup \{u\}; \ \text{var-count}_{\text{remaining}} \leftarrow \text{var-count}_{\text{remaining}} - \left|Z_u^{\text{variables}}\right|$
      $\widehat{k} \leftarrow \widehat{k} - 1; \ \widehat{n} \leftarrow \widehat{n} - 1; \ \widehat{r} \leftarrow \widehat{r} - \left|\{x_{e,u} \,|\, x_{e,u} \in Z_u^{\text{variables}}\}\right|$
      $\widehat{X}_{\text{sol}} \leftarrow \widehat{X}_{\text{sol}} \cup \{x_{e,u} \,|\, x_{e,u} \in Z_u^{\text{variables}}\}$
      remove the variables in $Z_u^{\text{variables}}$ from $\mathscr{L}^{(t-1)}$, and delete or update the constraints and the objective function to reflect the removal of variables
      $\forall i \in \{1, \ldots, \chi\} : \widehat{q}_i \leftarrow \widehat{q}_i - \left|\{x_{e,u} \,|\, x_{e,u} \in Z_u^{\text{variables}} \text{ and } \mathscr{C}(e) = i\}\right|$
      subtract the value $\sum_{x_{e,u} \in Z_u^{\text{variables}}} w(e) x_{e,u}$ from the objective function $\psi^{(t-1)}$
    Case 3: $1 \leq \widehat{n} \leq \chi + 1$
      let $y_{u_1}, \ldots, y_{u_{\widehat{n}'}}$ be the remaining non-zero $1 \leq \widehat{n}' \leq \widehat{n}/2$ node indicator variables
      $\text{var-count}_{\text{remaining}} \leftarrow 0; \ \widehat{k} \leftarrow 0; \ \widehat{n} \leftarrow 0; \ \widehat{r} \leftarrow 0$
      $\forall i \in \{1, \ldots, \chi\} : \widehat{q}_i \leftarrow \widehat{q}_i - \left|\{x_{e,u} \,|\, x_{e,u} \in Z_u^{\text{variables}} \text{ and } \mathscr{C}(e) = i\}\right|$
      $\widehat{V}_{\text{sol}} \leftarrow \widehat{V}_{\text{sol}} \cup \{u_1, \ldots, u_{\widehat{n}'}\}; \ \widehat{X}_{\text{sol}} \leftarrow \widehat{X}_{\text{sol}} \cup \{x_{e,u_j} \,|\, x_{e,u_j} \in Z_{u_j}^{\text{variables}}, j \in \{1, \ldots, \widehat{n}'\}\}$
  end cases
end while
$\mathsf{OPT}_{\text{frac}} \leftarrow 0$
return $\widehat{V}_{\text{sol}}$

Fig. 6 Pseudo-code of the iterated rounding algorithm ALG-ITER-ROUND used in Theorem 2. $\widehat{V}_{\text{sol}}^{(t^{\text{final}})}$ is the set of nodes selected in our solution

the total number of non-zero variables is $\widehat{n}^{(\alpha)} + \widehat{r}^{(\alpha)}$. Since the constraints in lines **(4)**–**(5)** of Fig. 5 are not strict constraints now (*i.e.*, not satisfied with equalities), the total number of *any* maximal set of strict constraints is at most the total number of constraints in lines **(1)**–**(3)** of Fig. 5, *i.e.*, at most $\widehat{r}^{(\alpha)} + \chi + 1$. By the rank lemma (Fact 2) $\widehat{r}^{(\alpha)} + \chi + 1 \geq \widehat{n}^{(\alpha)} + \widehat{r}^{(\alpha)} \equiv \widehat{n}^{(\alpha)} \leq \chi + 1$, which implies Case 3 applies and the algorithm terminates.

We now prove the bound on the number of selected sets. The value of $\widehat{k}$ decreases by 1 every time a new node is selected in Case 2 and remains unchanged in Case 1 where no node is selected. In the very last iteration involving Case 3, since G has no isolated nodes the number of node indicator variables is at least the number of edge indicator variables, implying $\widehat{n}' \leq \widehat{n}/2 \leq \frac{\chi+1}{2}$. Since $\widehat{k}^{(t^{\text{final}}-1)} \geq 1$, the total number of nodes selected is at most $k + (\widehat{n}' - 1) \leq k + \frac{\chi-1}{2}$. □

Lemma 3 *The sum of weights Γ of the edges selected by* ALG-ITER-ROUND *is at least $OPT/2$.*

Proof Let $W_{Alg-W}^{(t)} = W_{Alg}^{(t)} - W^{\chi+1}$, and $\mathsf{OPT}_{-W} = \mathsf{OPT} - W^{\chi+1}$. The proof of Lemma 2 shows that Case 3 of ALG-ITER-ROUND is executed only when $t = t^{\text{final}}$.

Thus, the details of ALG-ITER-ROUND in Fig. 6 imply the following sequence of assertions:

(i) $\mathsf{OPT}_{\text{frac}}^{(0)} \geq \mathsf{OPT}_{-W}$ and $\mathsf{OPT}_{\text{frac}}^{(t)} = \mathsf{OPT}_{\text{frac}}^{(t-1)} - \left(w(\widehat{X}_{\text{sol}}^{(t)}) - w(\widehat{X}_{\text{sol}}^{(t-1)})\right)$ for $t \in \{1, \ldots, t^{\text{final}} - 1\}$. Since the variables $x_{e_{u,j}} \in \widehat{X}_{\text{sol}}^{(t^{\text{final}})} \setminus \widehat{X}_{\text{sol}}^{(t^{\text{final}}-1)}$ are at most 1, we have

$$\begin{aligned}\mathsf{OPT}_{\text{frac}}^{(t^{\text{final}})} &= \mathsf{OPT}_{\text{frac}}^{(t^{\text{final}}-1)} - \sum_{x_{e_{u,j}} \in \widehat{X}_{\text{sol}}^{(t^{\text{final}})} \setminus \widehat{X}_{\text{sol}}^{(t^{\text{final}}-1)}} w(e) x_{e_{u,j}} \geq \mathsf{OPT}_{\text{frac}}^{(t^{\text{final}}-1)} \\ &\quad - \sum_{x_{e_{u,j}} \in \widehat{X}_{\text{sol}}^{(t^{\text{final}})} \setminus \widehat{X}_{\text{sol}}^{(t^{\text{final}}-1)}} w(e) \\ &= \mathsf{OPT}_{\text{frac}}^{(t^{\text{final}}-1)} - \left(w(\widehat{X}_{\text{sol}}^{(t^{\text{final}})}) - w(\widehat{X}_{\text{sol}}^{(t^{\text{final}}-1)})\right)\end{aligned}$$

Using the fact that $\mathsf{OPT}_{\text{frac}}^{(t^{\text{final}})} = 0$, we can therefore unravel the recurrence to get

$$\mathsf{OPT}_{\text{frac}}^{(t^{\text{final}})} \geq \mathsf{OPT}_{\text{frac}}^{(0)} - w(\widehat{X}_{\text{sol}}^{(t^{\text{final}})}) \Rightarrow w(\widehat{X}_{\text{sol}}^{(t^{\text{final}})}) \geq \mathsf{OPT}_{\text{frac}}^{(0)} \geq \mathsf{OPT}_{-W} \quad (20)$$

(ii) $W_{Alg-W}^{(0)} = 0$ and $W_{Alg-W}^{(t)} = W_{Alg-W}^{(t-1)} + \left(w(\widehat{X}_{\text{sol}}^{(t)}) - w(\widehat{X}_{\text{sol}}^{(t-1)})\right)$ for $t \in \{1, \ldots, t^{\text{final}}\}$. Using (20) we can unravel the recurrence we get

$$W_{Alg-W}^{(t^{\text{final}})} = w(\widehat{X}_{\text{sol}}^{(t^{\text{final}})}) \geq \mathsf{OPT}_{\text{frac}}^{(0)} \geq \mathsf{OPT}_{-W}$$

Noting that an edge $e \stackrel{\text{def}}{=} \{u, v\}$ can contribute the value of $w(e)$ *twice* in $W_{Alg}^{(t^{\text{final}})}$ corresponding to the two variables $x_{e,u}$ and $x_{e,v}$, the total weight Γ of selected edges in our solution is at least

$$\Gamma \geq W^{\chi+1} + \frac{1}{2} W_{Alg}^{(t^{\text{final}})} \geq \frac{W^{\chi+1} + W_{Alg}^{(t^{\text{final}})}}{2} \geq \frac{W^{\chi+1} + \mathsf{OPT}_{-W}}{2} = \frac{\mathsf{OPT}}{2} \square$$

□

Our proof of Theorem 2 is therefore completed once we prove the following lemma.

Lemma 4 *For all* $i, j \in \{1, \ldots, \chi\}$ $\frac{\widehat{p}_i^{(t^{final})}}{\widehat{p}_j^{(t^{final})}} \leq 4 + 4\chi$.

Proof When $t = t^{\text{final}}$ Case 3 applies and, since the variables $x_{e_{u,j}} \in \widehat{X}_{\text{sol}}^{(t^{\text{final}})} \setminus \widehat{X}_{\text{sol}}^{(t^{\text{final}}-1)}$ are at most 1, $\widehat{q}_i^{(t^{\text{final}})} \leq 0$ and consequently $\widehat{q}_i^{(t^{\text{final}}-1)} - \widehat{q}_i^{(t^{\text{final}})} \geq \widehat{q}_i^{(t^{\text{final}}-1)}$. Noting that an edge $e \stackrel{\text{def}}{=} \{u, v\}$ can contribute *twice* in the various $\widehat{q}_i^{(t)}$'s corresponding to the

two variables $x_{e,u}$ and $x_{e,v}$ and remembering that $\widehat{q}_i^{(0)} = \widehat{q}_i$, we get

$$\widehat{p}_i^{(t^{\text{final}})} \geq \frac{1}{2}\Big(\sum_{t=1}^{t^{\text{final}}}\big(\widehat{q}_i^{(t-1)} - \widehat{q}_i^{(t)}\big)\Big) + \mu_i \geq \frac{1}{2}\Big(\sum_{t=1}^{t^{\text{final}}-1}\big(\widehat{q}_i^{(t-1)} - \widehat{q}_i^{(t)}\big) + \widehat{q}_i^{(t^{\text{final}}-1)}\Big)$$
$$+ \mu_i = \frac{\widehat{q}_i^{(0)}}{2} + \mu_i$$
$$= \frac{\widehat{q}_i}{2} + \mu_i \geq \frac{\mathsf{OPT}_{\#}}{2\chi} - \frac{\mu_i}{2} + \mu_i = \frac{\mathsf{OPT}_{\#}}{2\chi} + \frac{\mu_i}{2}$$

We can get an upper bound on $\widehat{p}_i^{(t^{\text{final}})}$ by getting an upper bound on $\widehat{q}_i^{(t^{\text{final}}-1)} - \widehat{q}_i^{(t^{\text{final}})}$ in the following manner. Consider the $\widehat{n}^{(t^{\text{final}})'} < \widehat{n}^{(t^{\text{final}})} \leq \chi+1$ nodes $u_1, \ldots, u_{\widehat{n}^{(t^{\text{final}})'}}$ in Case 3. By choice of the nodes $v_1, \ldots, v_{\chi+1}$ of degrees $\mathsf{deg}(v_1), \ldots, \mathsf{deg}(v_{\chi+1})$, respectively, the number of edges incident on u_i is at most $\mathsf{deg}(v_i)$ for all $i \in \{1, \ldots, \widehat{n}^{(t^{\text{final}})'}\}$. Thus, we get $\widehat{q}_i^{(t^{\text{final}}-1)} - \widehat{q}_i^{(t^{\text{final}})} \leq \sum_{j=1}^{\chi+1} \mathsf{deg}(v_j) \leq 2\mathsf{OPT}_{\#}$, and consequently

$$\widehat{p}_i^{(t^{\text{final}})} \leq \sum_{t=1}^{t^{\text{final}}}\big(\widehat{q}_i^{(t-1)} - \widehat{q}_i^{(t)}\big) \leq \sum_{t=1}^{t^{\text{final}}-1}\big(\widehat{q}_i^{(t-1)} - \widehat{q}_i^{(t)}\big) + 2\mathsf{OPT}_{\#} = \widehat{q}_i + 2\mathsf{OPT}_{\#}$$
$$\leq 2\left(\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_i\right) + 2\mathsf{OPT}_{\#} = (2+2\chi)\frac{\mathsf{OPT}_{\#}}{\chi}$$

Thus, for all $i, j \in \{1, \ldots, \chi\}$ we have

$$\frac{\widehat{p}_i^{(t^{\text{final}})}}{\widehat{p}_j^{(t^{\text{final}})}} \leq \frac{(2+2\chi)\frac{\mathsf{OPT}_{\#}}{\chi}}{\frac{\mathsf{OPT}_{\#}}{2\chi} + \frac{\mu_j}{2}} < 4 + 4\chi \square$$

□

10.1.2 The Case of Arbitrary χ

As stated below, there are two steps in the previous algorithm that cannot be executed in polynomial time when χ is *not* a constant:

(1) We cannot guess the $\chi + 1$ nodes $v_1, \ldots, v_{\chi+1}$ in polynomial time. Instead, we guess only one node v_1 such that there exists an optimal solution contains v_1 with the following property: "the remaining $k - 1$ nodes in the solution have degree *at most* $\mathsf{deg}(v_1)$".

(2) We cannot guess the exact value of $\widehat{q}_i$ by exhaustive enumeration and therefore we cannot use the χ constraints "$\sum_{u \in \widehat{V}} \sum_{e \stackrel{\text{def}}{=} \{u,v\} \in C_i} x_{e,u} = \widehat{q}_i$" in line **(3)** of the LP-relaxation in Fig. 5 anymore. However, note that it still holds that $\widehat{q}_i$ is an

integer in the set $\left\{\frac{\mathsf{OPT}_\#}{\chi} - \mu_i, \frac{\mathsf{OPT}_\#}{\chi} - \mu_i + 1, \ldots, 2\left(\frac{\mathsf{OPT}_\#}{\chi} - \mu_i\right)\right\}$. Thus, instead we use the 2χ constraints

$$\textbf{(3)} \quad \frac{\mathsf{OPT}_\#}{\chi} - \mu_i \le \sum_{u \in \widehat{V}} \sum_{e \stackrel{\text{def}}{=} \{u,v\} \in C_i} x_{e,u} = \widehat{q}_i \le 2\left(\frac{\mathsf{OPT}_\#}{\chi} - \mu_i\right) \quad \text{for all } i \in \{1, \ldots, \chi\}$$

We need modifications of the bounds in the previous proof to reflect these changes as follows:

- We make some obvious parameter value adjustments such as: $\widehat{V} = V \backslash \{v_1\}$, $\widehat{E} = E \backslash \{\{u, v_1\} \mid \{u, v_1\} \in E\}$, $\widehat{k}^{(0)} = k - 1$, $\mu_i \le \mathsf{deg}(v_1)$ for all i.
- The number of constraints in lines **(1)**–**(3)** of Fig. 5 is now $\widehat{m} + 2\chi + 1$.
- The condition in Case 3 of Fig. 6 is now $1 \le \widehat{n} \le 2\chi + 1$.
- In Lemma 2, we select at most $\left\lfloor k + \frac{2\chi - 1}{2} \right\rfloor = k + \chi - 1$ nodes.
- The calculations for the upper bound for $\widehat{p}_i^{(t^{\text{final}})}$ in Lemma 4 change as follows. By choice of the node v_1 of degree $\mathsf{deg}(v_1)$, the number of edges incident on u_i is at most $\mathsf{deg}(v_1)$ for all $i \in \{1, \ldots, \widehat{n}^{(t^{\text{final}})'}\}$. This now gives $\widehat{q}_i^{(t^{\text{final}}-1)} - \widehat{q}_i^{(t^{\text{final}})} \le (2\chi + 1)\mathsf{deg}(v_1) \le (2\chi + 1)\mathsf{OPT}_\#$, and therefore $\widehat{p}_i^{(t^{\text{final}})} \le \widehat{q}_i + (2\chi + 1)\mathsf{OPT}_\# \le 2\left(\frac{\mathsf{OPT}_\#}{\chi} - \mu_i\right) + (2\chi + 1)\mathsf{OPT}_\# < (2 + \chi + 2\chi^2)\frac{\mathsf{OPT}_\#}{\chi}$. This gives us the following updated bound:

$$\frac{\widehat{p}_i^{(t^{\text{final}})}}{\widehat{p}_j^{(t^{\text{final}})}} \le \frac{(2 + \chi + 2\chi^2)\frac{\mathsf{OPT}_\#}{\chi}}{\frac{\mathsf{OPT}_\#}{2\chi} + \frac{\mu_j}{2}} < 4 + 2\chi + 4\chi^2$$

10.2 The General Case: Approximating FMC

Theorem 3 (generalizing Theorem 2 for FMC) *We can design a deterministic polynomial-time approximation algorithm for* FMC *with the following properties:*

(a) *The algorithm selects τ sets where* $\tau \le \begin{cases} k + \frac{\chi - 1}{2}, & \text{if } \chi = O(1) \\ k + \chi - 1, & \text{otherwise} \end{cases}$

(b) *The algorithm is a $1/f$-approximation for* NODE- FMC*, i.e., the total weight of the selected elements is at least OPT/f.*

(c) *The algorithm satisfies the ε-approximate coloring constraints (cf. Inequality (5)) as follows:*

$$\text{for all } i, j \in \{1, \ldots, \chi\}, \ \frac{p_i}{p_j} < \begin{cases} O(\min\{\chi^2 f, \ \chi f^2\}), & \text{if } \chi = O(1) \\ O(f^2 + \chi^2 f), & \text{otherwise} \end{cases}$$

The proof of Theorem 3 is a suitable modified version of the proof of Theorem 2. We point out the important alterations that are needed.

General Modifications

- Nodes and edges now correspond to sets and elements, respectively, incidence of an edge on a node corresponds to membership of an element in a set, and degree of a node corresponds to number of elements in a set.
- There is a set indicator variable y_j for every element $\mathscr{S}_j \in \widehat{V}$. For every element $(u_i, \mathscr{S}_j) \in \widehat{E}$, there is an element indicator variable $x_{i,j}$ and a constraint $x_{i,j} = y_j$.
- Now $\sum_{i=1}^{\chi+1} |\mathscr{S}_i| \leq \min\{\chi, f\}\mathsf{OPT}_{\#}$ since any element in any one of the sets from $\mathscr{S}_1, \ldots, \mathscr{S}_{\chi+1}$ can appear in at most $\min\{\chi, f\}$ other sets in the collection of sets $\mathscr{S}_1, \ldots, \mathscr{S}_{\chi+1}$. Also, $\widehat{q}_j$ is an integer in the set $\left\{\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_j, \frac{\mathsf{OPT}_{\#}}{\chi} - \mu_j + 1, \ldots, f\left(\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_j\right)\right\} \subseteq \{0, 1, 2, \ldots, n\}$ since any element can appear in at most f sets.
- An element u_i appearing in $f_i \leq f$ sets, say sets $\mathscr{S}_1, \ldots, \mathscr{S}_{f_i}$, can now contribute the value of $w(u_i)$ at most $f_i \leq f$ times in $W_{Alg}^{(t^{\text{final}})}$ corresponding to the f_i variables $x_{i,1}, \ldots, x_{i,f_i}$. Thus, we get a $1/f$-approximation to the objective function.

Modifications Related to $\chi = O(1)$ case

- An element u_i appearing in $f_i \leq f$ sets, say sets $\mathscr{S}_1, \ldots, \mathscr{S}_{f_i}$, can now contribute at most $f_i \leq f$ times in the various $\widehat{q}_i^{(t)}$'s corresponding to the f_i variables $x_{i,1}, \ldots, x_{i,f_i}$. This modifies the relevant inequality for $\widehat{p}_i^{(t^{\text{final}})}$ as follows:

$$\widehat{p}_i^{(t^{\text{final}})} \geq \frac{\widehat{q}_i^{(0)}}{f} + \mu_i \geq \frac{\mathsf{OPT}_{\#}}{f\chi} - \frac{\mu_i}{f} + \mu_i > \frac{\mathsf{OPT}_{\#}}{f\chi}$$

$$\widehat{p}_i^{(t^{\text{final}})} \leq \widehat{q}_i + \sum_{i=1}^{\chi+1} |\mathscr{S}_i| \leq f\left(\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_i\right) + \min\{\chi, f\}\mathsf{OPT}_{\#} < \min\{\chi^2 + f, 2\chi f\}\frac{\mathsf{OPT}_{\#}}{\chi}$$

$$\frac{\widehat{p}_i^{(t^{\text{final}})}}{\widehat{p}_j^{(t^{\text{final}})}} \leq \frac{\min\{\chi^2 + f, 2\chi f\}\frac{\mathsf{OPT}_{\#}}{\chi}}{\frac{\mathsf{OPT}_{\#}}{f\chi}} \leq \min\{\chi^2 f + f^2, 2\chi f^2\} = O(\min\{\chi^2 f, \chi f^2\})$$

Modifications related to the arbitrary χ case

- The calculations for the upper bound for $\widehat{p}_i^{(t^{\text{final}})}$ in Lemma 4 change as follows.

$$f\left(\frac{\mathsf{OPT}_{\#}}{\chi} - \mu_i\right) + (2\chi + 1)\mathsf{OPT}_{\#} < \left(f + \chi + 2\chi^2\right)\frac{\mathsf{OPT}_{\#}}{\chi} < \left(f + 3\chi^2\right)\frac{\mathsf{OPT}_{\#}}{\chi}$$

This gives the final bound of $\frac{\widehat{p}_i^{(t^{\text{final}})}}{\widehat{p}_j^{(t^{\text{final}})}} \leq f^2 + 3\chi^2 f = O(f^2 + \chi^2 f)$.

11 Approximation Algorithms for Two Special Cases of FMC

For approximating these special cases of FMC, which are still NP-complete, we will be specific about the various constants and will try to provide approximation algorithms with as tight a constant as we can. For this section, let $\rho = \max\{\rho(f),\ \rho(k)\}$. Note that $\rho > 1 - 1/\mathbf{e}$.

11.1 SEGR-FMC: Almost Optimal Deterministic Approximation with "at most" k Sets

Note that Lemma 1 shows that finding a feasible solution is NP-complete even for unweighted SEGR- FMC with $\chi = 2$. Further inapproximability results for SEGR- FMC are stated in Remark 5.

Theorem 4 *There exists a polynomial-time deterministic algorithm* ALG-GREED-PLUS *that, given an instance of unweighted* SEGR- FMC (χ, k) *outputs a solution with the following properties:*

(a) *The number of selected sets is at most* k.
(b) *The approximation ratio is at least* $\rho > 1 - 1/\boldsymbol{e}$.
(c) *The coloring constraints are* 2*-approximately satisfied* (*cf. (5)*)*, i.e.,*

$$\forall\, i, j \in \{1, \ldots, \chi\} :\ p_i \leq 2\, p_j$$

Remark 5 Based on the $(1-1/\mathbf{e})$-inapproximability result of Feige in [16] for the maximum k-set coverage problem, it is not difficult to see the two constants in Theorem 4, namely ρ and 2, cannot be improved beyond $1-1/\mathbf{e}+\varepsilon$ and $(1 - 1/\mathbf{e})^{-1}+\varepsilon \approx 1.58+\varepsilon$, respectively, for any $\varepsilon > 0$ and all $\chi \geq 2$ assuming P $\neq$ NP.

Remark 6 The "at most k sets" part of the proof arises in the following steps of the algorithm. Since we cannot know k_r exactly, we can only assume $\widehat{k_r} \leq k_r$ since it is possible that the algorithm for the maximum k-set coverage also covers at least $\rho \frac{\mathsf{OPT}_{\#}}{\chi}$ elements for some $k < k_r$. Secondly, even if we have the guessed the correct value of k_r, the algorithm for the maximum k_r-set coverage may cover more than $2\rho \frac{\mathsf{OPT}_{\#}}{\chi}$ elements, and thus we have to "un-select" some of the selected sets to get the desired bounds (the proof shows that sometimes we may have to un-select all but one set). The following example shows that a solution that insists on selecting exactly k sets may need to select sets all of which are not in our solution. Consider the following instance of unweighted FMC$(1, \ell)$: $\mathscr{U} = \{u_1, \ldots, u_n\}$, $\ell = n/2$, $\mathscr{S}_1 = \{u_1, \ldots, u_{n/2}\}$, and $\mathscr{S}_{j+1} = \{u_{(n/2)+j}\}$ for $j = 1, \ldots, n/2$. Our algorithm will select the set $\mathscr{S}_1$ whereas any solution that selects exactly ℓ sets must select the sets $\mathscr{S}_2, \ldots, \mathscr{S}_{(n/2)+1}$.

Proof We reuse the notations, terminologies and bounds shown in the proof of Theorem 1 as needed. Let $\mathscr{U}_1, \ldots, \mathscr{U}_\chi$ be the partition of the universe based on the color of the elements, *i.e.*, $\mathscr{U}_r = \{u_\ell \mid \mathscr{C}(u_\ell) = r\}$ for $r \in \{1, \ldots, \chi\}$. By the definition of SEGR- FMC every set contains elements from exactly one such partition and thus, after renaming the sets and elements for notational convenience, we may set assume that

our collection $\mathscr{S}_1, \ldots, \mathscr{S}_m$ of m sets is partitioned into χ collection of sets, where the r^{th} collection (for $r \in \{1, \ldots, \chi\}$) contains the sets $\mathscr{S}^r_1, \ldots, \mathscr{S}^r_{m_r}$ over the universe $\mathscr{U}_r = \{u_1, \ldots, u_{n_r}\}$ of n_r elements such that $\sum_{r=1}^{\chi} m_r = m$ and $\sum_{r=1}^{\chi} n_r = n$. For $r \in \{1, \ldots, \chi\}$ and any ℓ let $\text{FMC}_r(1, \ell)$ be the unweighted $\text{FMC}(1, \ell)$ problem defined over the universe $\mathscr{U}_r$ and the collection of sets $\mathscr{S}^r_1, \ldots, \mathscr{S}^r_{m_r}$. The following observation holds trivially.

Unweighted SEGR- FMC (χ, k) has a valid solution covering $\ell \in \{\chi, 2\chi, \ldots, \lfloor n/\chi \rfloor \chi\}$ elements if and only if *(i)* for each $r \in \{1, \ldots, \chi\}$, $\text{FMC}_r(1, k_r)$ has a valid solution covering ℓ/χ elements for some $k_r > 0$, and *(ii)* $\sum_{r=1}^{\chi} k_r = k$.

The above observation suggests that we can guess the value of $\mathsf{OPT}_{\#}$ by trying out all possible values of ℓ just like the algorithms in Theorem 1, and for each such value of ℓ we can solve χ *independent* FMC instances and combine them to get a solution of the original SEGR- FMC instance. Although we cannot possibly solve the $\text{FMC}_r(1, k_r)$ problems exactly, appropriate approximate solutions of these problems do correspond to a similar approximate solution of SEGR- FMC (χ, k) as stated in the following observation:

Suppose that for each $r \in \{1, \ldots, \chi\}$ we have a solution $\mathscr{S}^r_{i_1}, \ldots, \mathscr{S}^r_{i_{\widehat{k_r}}} \subseteq \mathscr{U}_r$ of $\text{FMC}_r(1, k_r)$ with the following properties (for some $\eta_1 \leq 1$ and $\eta_2 \geq 1$): *(i)* $\eta_1(\ell/\chi) \leq |\cup_{p=1}^{\widehat{k_r}} \mathscr{S}^r_{i_p}| \leq \eta_2(\ell/\chi)$, and *(ii)* $\widehat{k_r} \leq k_r$. Then, the collection of sets $\left\{\mathscr{S}^r_{i_{\ell_r}} \mid \ell_r \in \{1, \ldots, \widehat{k_r}\},\ r \in \{1, \ldots, \chi\}\right\}$ outputs a solution of SEGR-FMC (χ, k) with the following properties: *(a)* the number of selected sets is at most k, *(b)* the number of elements covered is at least $\eta_1 \ell$, and *(c)* for any pair $i, j \in \{1, \ldots, \chi\}$, $p_i/p_j \leq \eta_2/\eta_1$.

By the above observation, to prove our claim it suffices if we can find a solution for $\text{FMC}_r(1, k_r)$ for any r with $\ell = \mathsf{OPT}_{\#}$, $\eta_1 = \rho$ and $\eta_2 = 2\ \rho$. For convenience, we will omit the *superscript* r from the set labels while dealing with $\text{FMC}_r(1, k_r)$. Remove from consideration any sets from $\mathscr{S}_1, \ldots, \mathscr{S}_{m_r}$ that contains more than ℓ/χ elements, and consider the standard (unweighted) maximum k-set coverage problem, that ignores constraint (i) of the above observation, on these remaining collection of sets $\mathscr{T}$ over the universe $\mathscr{U}_r$. Since we have guessed the correct value of ℓ, there is at least one valid solution and thus the following assertions hold: (I) there exists a set of k_r sets that covers $\frac{\mathsf{OPT}_{\#}}{\chi}$ elements, and (II) $|\mathscr{T}| \geq k_r$.

Let ν_k denote the *maximum* number of elements that can be covered by selecting k sets from $\mathscr{T}$. There are the following two well-known algorithms for the maximum k-set coverage problem both of which select exactly k sets: the greedy algorithm covers at least $\rho(k)\nu_k$ elements [16, Proposition 5.1], where the pipage-rounding algorithm (based on the LP-relaxation in Fig. 2) covers at least $\rho(f)\nu_k$ elements [1]. Note that we do *not* know the exact value of k_r and we *cannot* guess by enumerating every possible k_r values for every $r \in \{1, \ldots, \chi\}$ in polynomial time. To overcome this obstacle, we use the following steps.

- We run both the algorithms for maximum k-set coverage for $k = 1, 2, \ldots$ until we find the first (*smallest*) index $\widehat{k_r} \leq k_r$ such that the better of the two algorithms cover *at least* $\max\{\rho(\widehat{k_r}),\ \rho(f)\}\frac{\mathsf{OPT}_\#}{\chi} \geq \rho\frac{\mathsf{OPT}_\#}{\chi}$ elements.
- Suppose that this algorithm selects the $\widehat{k_r}$ sets (after possible re-numbering of set indices) $\mathscr{S}_1, \ldots, \mathscr{S}_{\widehat{k_r}}$, where we have ordered the sets such that for every $j \in \{2, 3, \ldots, \widehat{k_r}\}$ the number of elements covered by $\mathscr{S}_j$ and *not* covered by any of the sets $\mathscr{S}_1, \ldots, \mathscr{S}_{j-1}$ is *at least* as many as the number of elements covered by $\mathscr{S}_\ell$ and not covered by any of the sets $\mathscr{S}_1, \ldots, \mathscr{S}_{j-1}$ for *any* $\ell > j$. Remember that $\max_{j\in\{1,\ldots,\widehat{k_r}\}}\{|\mathscr{S}_j|\} \leq \mathsf{OPT}_\#/\chi$. Let j be the *smallest* index such $|\cup_{\ell=1}^{j-1}\mathscr{S}_j| < \rho\frac{\mathsf{OPT}_\#}{\chi}$ but $|\cup_{\ell=1}^{j}\mathscr{S}_j| \geq \rho\frac{\mathsf{OPT}_\#}{\chi}$. We have the following cases.
  - If $|\mathscr{S}_j| \geq \rho\frac{\mathsf{OPT}_\#}{\chi}$ then we select $\mathscr{S}_j$ as our solution since $\rho\frac{\mathsf{OPT}_\#}{\chi} \leq |\mathscr{S}_j| \leq \frac{\mathsf{OPT}_\#}{\chi} < 2\ \rho\frac{\mathsf{OPT}_\#}{\chi}$.
  - Otherwise $|\mathscr{S}_j| < \rho\frac{\mathsf{OPT}_\#}{\chi}$ and in this case we select the $j \leq \widehat{k_r} \leq k_r$ sets $\mathscr{S}_1, \ldots, \mathscr{S}_j$ in our solution since $\rho\frac{\mathsf{OPT}_\#}{\chi} \leq |\cup_{\ell=1}^{j}\mathscr{S}_j| \leq 2\ \rho\frac{\mathsf{OPT}_\#}{\chi}$. □

11.2 Δ-BAL-FMC: Improved Deterministic Approximation

Proposition 2 *There exists a polynomial-time deterministic algorithm* ALG-GREEDY *that, given an instance of unweighted* Δ-BAL- FMC(χ, k) *outputs a solution with the following properties:*

(a) *The number of selected sets is* (*exactly*) k.
(b) *The approximation ratio is at least* $\rho > 1 - 1/\boldsymbol{e}$.
(c) *The coloring constraints are* $O(\Delta f)$*-approximately satisfied* (*cf. (5)*)*, i.e.,* $\forall\, i, j \in \{1, \ldots, \chi\} :\ p_i/p_j \leq (2 + 2\Delta)f$.

Proof As already mentioned in the proof of Theorem 4 and elsewhere, there is a deterministic polynomial-time algorithm for the maximum k-set coverage problem with an approximation ratio of ρ. For the given instance of Δ-BAL- FMC(χ, k), we run this algorithm (ignoring element colors) selecting k sets, say $\mathscr{S}_1, \ldots, \mathscr{S}_k$. Obviously, the total weight of all the elements covered in the selected solution is at least ρ OPT. Let $\alpha^+ = \sum_{i=1}^{k}\lceil|\mathscr{S}_i|/\chi\, y\rceil + \Delta$ and $\alpha^- = \sum_{i=1}^{k}\max\left\{1,\ \lfloor|\mathscr{S}_i|/\chi\rfloor - \Delta\right\}$. Note that $k \leq \alpha^- \leq \alpha^+ \leq \alpha^- + (2\Delta + 1)k$. Since each of the sets in the solution is balanced, an upper bound for the number p_i of elements of color i in the solution is given by $p_i \leq \alpha^+$. Also note that by definition of f we have $p_i \geq \frac{\alpha^-}{f}$. It thus follows that for any i and j we have $p_i/p_j \leq f \times \frac{\alpha^+}{\alpha^-} \leq (2 + 2\Delta)f$. □

12 Approximating GEOM-FMC via Randomized Shifting

We refer the reader to textbooks such as [48] for a general overview of the randomized shifting technique (textbook [48] illustrates the technique in the context of Euclidean travelling salesperson problem).

Theorem 5 *For any constant $0 < \varepsilon < 1$, we can design a randomized algorithm* ALG-GEOM *for* GEOM- FMC *with the following properties:*

(a) ALG-GEOM *runs in $O((\Delta/d)^d 2^{(Cd/\varepsilon)^{O(d)}} k^d)$ time.*
(b) ALG-GEOM *satisfies the following properties with probability $1-o(1)$ (cf. Inequality (5′)–(5″)):*

- *The algorithm covers at least $(1 - O(\varepsilon))(\mathsf{OPT} - \varepsilon\chi)$ points.*
- *The algorithm satisfies the $(1 + \varepsilon)$-approximate coloring constraints (cf. Inequality (5″)), i.e., for all $i, j \in \{1, \ldots, \chi\}$, $p_i \le (1 + \varepsilon)p_j$.*

Remark 7 In many geometric applications, the dimension parameter d is a *fixed* constant. For this case, ALG-GEOM runs in polynomial time, and moreover, under the *mild* assumption of $\frac{\mathsf{OPT}}{\chi} \ge \eta$ for *some* constant $\eta > 1$, ALG-GEOM covers at least $(1 - O(\varepsilon))\mathsf{OPT}$ points, *i.e.*, under these conditions ALG-GEOM behaves like a randomized polynomial-time approximation scheme.

Proof Fix an optimal solution having k unit balls $\mathscr{B}_1^*, \ldots, \mathscr{B}_k^* \subset \mathbb{R}^d$, such that for all $i, j \in \{1, \ldots, \chi\}$, $\mu_i(\mathscr{B}^*) = \mu_j(\mathscr{B}^*)$, where $\mathscr{B}^* = \bigcup_{i=1}^k \mathscr{B}_i^*$. Thus, we need to show that our algorithm ALG-GEOM computes in $(\Delta/d)^d 2^{(Cd/\varepsilon)^{O(d)}} k^d$ time a set of unit balls $B_1, \ldots, B_k \subset \mathbb{R}^d$ such that the following assertions hold with probability $1 - o(1)$ (where $B = \bigcup_{i=1}^k B_i$):

$$\sum_{i=1}^{\chi} \mu_i(B) > (1 - O(\varepsilon)) \sum_{i=1}^{\chi} (\mu_i(\mathscr{B}^*) - \varepsilon)$$
$$\forall i, j \in \{1, \ldots, \chi\} : \mu_i(B) \le (1 + \varepsilon)\mu_j(B)$$

Set $L = 8d/\varepsilon$. Let $G \subset \mathbb{R}^d$ be an axis-parallel grid such that every connected component of $\mathbb{R}^d \setminus G$ is an open d-dimensional hypercube isometric to $(0, L)^d$. In other words, G is the union of d infinite families of axis-parallel $(d-1)$-dimensional hyperplanes, spaced apart by L in each orthonormal direction. Let $\alpha \in [0, L)$ be chosen uniformly at random, and let $G' = G + \alpha$ be the *random* translation of G by α, *i.e.*,

$$(p_1 + \alpha, \ldots, p_d + \alpha) \in G' \equiv (p_1, \ldots, p_d) \in G$$

Let F is the set of indices of all balls $\mathscr{B}_i^*$ that has a non-empty intersection with the randomly shifted grid G', *i.e.*,

$$F = \{i \in \{1, \ldots, k\} : \mathscr{B}_i^* \cap G' \neq \emptyset\}$$

Let $\mathscr{B}^{*,F} = \bigcup_{i \in F} \mathscr{B}_i^*$. Any point $p \in \mathbb{R}^d$ is contained in $\mathscr{B}^{*,F}$ only if it is contained in some unit ball intersecting G'. Therefore, $p \in \mathscr{B}^{*,F}$ only if it is at distance at most 2 from G' (in other words, $\mathscr{B}^{*,F}$ is contained in the 2-neighborhood of G'). The probability that any particular point p is at distance at most 2 from any family of

parallel randomly shifted hyperplanes in G' is exactly $4/L$. By the union bound over all dimensions, $\Pr[p \in \mathscr{B}^{*,F}] \leq 4d/L$. Therefore, by the linearity of expectation,

$$\mathbb{E}[\mu_i(\mathscr{B}^{*,F})] = \sum_{p \in \bigcup_{i=1}^{k} \mathscr{B}_i^*} \Pr\left[p \in \mathscr{B}^{*,F}\right] \mu_i(p) \leq \frac{4d}{L} \mu_i(\mathscr{B}^*)$$

Consequently, by Markov's inequality, we get

$$\Pr\left[\mu_i(\mathscr{B}^{*,F}) \geq \tfrac{8d}{R} \mu_i(\mathscr{B}^*)\right] \leq 1/2$$

Set $\delta = 2^{-\Theta(d)}\varepsilon/C$. Let $B_1^{**}, \ldots, B_k^{**}$ be the collection of unit balls in $\mathbb{R}^d$ obtained as follows. For each $i \in \{1, \ldots, k\} \setminus F$, obtain a unit ball by translating $\mathscr{B}_i^*$ such that its center has coordinates that are *integer multiples* of δ, *i.e.*, it is an element of the dilated integer lattice $\delta \cdot \mathbb{Z}^d$. For every $i \in F$, we obtain a unit ball by picking an arbitrary ball obtained for some $j \in \{1, \ldots, k\} \setminus F$ as described above. Essentially, the new solution $B_1^{**}, \ldots, B_k^{**}$ is missing all the balls that intersect G', and rounds every other ball so that its center is contained in some integer lattice. In this construction, each ball $\mathscr{B}_i^*$, with $i \notin F$, gets translated by at most some distance $\sqrt{d}\delta$. Since for all $j \in \{1, \ldots, \chi\}$, μ_j is C-Lipschitz, it follows that

$$|\mu_j(\mathscr{B}_i^*) - \mu_j(B_i^{**})| \leq \mathrm{vol}(\mathscr{B}_i^*)\, \sqrt{d}\, \delta\, C \leq 2^{\Theta(d)} \delta C$$

Letting $B^{**} = \bigcup_{i=1}^{k} B_i^{**}$, we get that for all $i \in \{1, \ldots, \chi\}$,

$$|\mu_j(\mathscr{B}^*) - \mu_j(B^{**})| \leq k 2^{\Theta(d)} \delta C$$

Let $\mathscr{I}$ be the set of connected components of $[0, \Delta]^d \setminus G'$. We refer to the elements of $\mathscr{I}$ as *cells*. For each $A \in \mathscr{I}$, we enumerate the set, $\mathscr{S}_A$, of all possible subsets of at most k unit balls with centers in $A \cap (\delta \cdot \mathbb{Z}^d)$. There are at most $(L/\delta)^d$ lattice points in A, and thus there are at most $2^{(L/\delta)^d}$ such subsets of unit balls. Since $|\mathscr{I}| \leq (\lceil \Delta/L \rceil)^d$, it follows that this enumeration takes $O((\Delta/L)^d\, 2^{(L/\delta)^d})$ time.

For each enumerated subset $J \in \mathscr{S}_A$ of unit balls, we record the vector

$$\left(|J|, \frac{\varepsilon}{k}\left\lfloor \mu_1(X)\frac{k}{\varepsilon}\right\rfloor, \ldots, \frac{\varepsilon}{k}\left\lfloor \mu_k(X)\frac{k}{\varepsilon}\right\rfloor\right)$$

where $X = \bigcup_{Y \in J} Y$. There are at most $(2^{O(d)}k/\varepsilon)^d$ such vectors for each cell in $\mathscr{I}$. Via standard dynamic programming, we can inductively compute all possible sums of vectors such that we pick at most one vector from each cell, and the total sum of the first coordinate, *i.e.*, the number of unit balls, is at most k. This can be done in $O((\Delta/L)^d 2^{(L/\delta)^d} (2^{O(d)}k/\varepsilon)^d)$ time. For the correct choice of vectors that corresponds to the solution B^{**}, the sum of the vectors we compute is correct up to an additive factor of ε on each coordinate. This means that we compute a solution $B_1, \ldots, B_k$, with the following property:

$$\sum_{i=1}^{k} \mu_i(B) \geq (1-\varepsilon) \sum_{i=1}^{\chi} \mu_i(B^{**}) \geq (1-\varepsilon)\Big[\sum_{i=1}^{\chi} \big(\mu_i(\mathscr{B}^*) - 2^{\Theta(d)}\delta C - \mu_i(\mathscr{B}^{*,F})\big)\Big]$$

$$\geq \big(1-\varepsilon-(8d/L)\big)\Big[\sum_{i=1}^{\chi} \big(\mu_i(\mathscr{B}^*) - 2^{\Theta(d)}\delta C\big)\Big] = \big(1-2\varepsilon\big)\Big[\sum_{i=1}^{\chi} \big(\mu_i(\mathscr{B}^*) - \varepsilon\big)\Big]$$

with probability at least 1/2. Repeating the algorithm $O(\log n)$ times and returning the best solution found, results in the high-probability assertion, which concludes the proof. □

13 Conclusion and Open Problems

In this paper we formulated a natural combinatorial optimization framework for incorporating fairness issues in coverage problems and provided a set of approximation algorithms for the general version of the problem as well as its special cases. Of course, it is possible to design other optimization frameworks depending on the particular application in hand, and we encourage researchers to do that. Below we list some future research questions related to our framework:

Eliminating the gap of factor f in LP-relaxation: As noted in Sect. 9.7, all of our LP-relaxations incur a gap of factor f in the coloring constraints while rounding. It seems non-trivial to close the gap using additional linear inequalities while preserving the same approximation ratio. However, it may be possible to improve the gap using SDP-relaxations.

Primal-dual schema: Another line of attack for the FMC problems is via the primal-dual approach [48]. For example, can the primal-dual approach for partial coverage problem by Gandhi, Khuller and Srinivasan [18] be extended to FMC? A key technical obstacle seems to center around effective interpretation of the dual of the coloring constraints. Our iterated rounding approach was able to go around this obstacle but the case when $\chi = \omega(1)$ may be improvable.

Fixed parameter tractability: As mentioned in Sect. 4 fixed-parameter tractability issues for k-node coverage have been investigated by prior researchers such as Marx [36] and Gupta, Lee and Li in [21, 22]. It would be interesting to extend these results to NODE- FMC.

Generalizing to non-decreasing submodular set objective functions: The proofs and proof techniques in this paper do not generalize to the case when the objective function for our FMC problems is a (more general) non-decreasing submodular set function. It would be interesting to devise new algorithmic techniques and proofs for this more general case. Approximation algorithms for such generalizations for the standard maximum k-set coverage problem (see Sect. 4) were provided in [30].

Acknowledgements We thank the editor and reviewers for their helpful comments.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Proof of Lemma 1

(a) We describe the proof for $\chi = 2$; generalization to $\chi > 2$ is obvious. The reduction is from the *Exact Cover by 3-sets* (X3C) problem which is defined as follows. We are given an universe $\mathscr{U}' = \{u_1, \ldots, u_{n'}\}$ of n' elements for some n' that is a multiple of 3, and a collection of n' subsets $\mathscr{S}_1, \ldots, \mathscr{S}_{n'}$ of $\mathscr{U}$ such that $\bigcup_{j=1}^{n'} \mathscr{S}_j = \mathscr{U}$, every element of $\mathscr{U}'$ occurs in exactly 3 sets and $|\mathscr{S}_j| = 3$ for $j = 1, \ldots, n'$. The goal is to decide if there exists a collection of $n'/3$ (disjoint) sets whose union is $\mathscr{U}'$. X3C is known to be NP-complete [19]. Given an instance $\langle \mathscr{U}', \mathscr{S}_1, \ldots, \mathscr{S}_{n'} \rangle$ of X3C as described, we create the following instance $\langle \mathscr{U}, \mathscr{S}_1, \ldots, \mathscr{S}_{n'+1}, k \rangle$ of FMC$(2, k)$:

(i) The universe is $\mathscr{U} = \{u_1, \ldots, u_{n'}\} \cup \{u_{n'+1}, \ldots, u_{2n'}\}$ (and thus $n = 2n'$),
(ii) $w(u_j) = 1$ for $j = 1, \ldots, 2n'$,
(iii) the sets are $\mathscr{S}_1, \ldots, \mathscr{S}_{n'}$ and a new set $\mathscr{S}_{n'+1} = \{u_{n'+1}, \ldots, u_{2n'}\}$,
(iv) the coloring function is given by $\mathscr{C}(u_j) = \begin{cases} 1, \text{ if } 1 \le j \le n' \\ 2, \text{ otherwise} \end{cases}$, and
(v) $k = \frac{n'}{3} + 1 = \frac{n}{6} + 1$.

Clearly, every element of $\mathscr{U}$ occurs in no more than 3 sets and all but the set $\mathscr{S}_{n'+1}$ contains exactly 3 elements. The proof is completed once the following is shown:

($*$) *the given instance of X3C has a solution if and only if the transformed instance of* FMC$(2, 1 + n/6)$ *has a solution.*

A proof of ($*$) is easy: since the set $\mathscr{S}_{n'+1}$ must appear in any valid solution of FMC, a solution $\mathscr{S}_{i_1}, \ldots, \mathscr{S}_{i_{n'/3}}$ of X3C corresponds to a solution $\mathscr{S}_{i_1}, \ldots, \mathscr{S}_{i_{n'/3}}, \mathscr{S}_{n'+1}$ of FMC$(2, k)$ and *vice versa*.
(b) The proof is similar to that in (a) but now instead of X3C we reduce the node cover problem for cubic (*i.e.*, 3-regular) graphs (VC$_3$) which is defined as follows: given a cubic graph $G = (V, E)$ of n' nodes and $3n'/2$ edges and an integer k', determine if there is a set of k' nodes that cover all the edges. VC_3 is known to be NP-complete even if G is planar [19]. For the translation to an instance of FMC$(2, k)$, edges of G are colored with color 1, we add a *new* connected component $\mathscr{K}_{(3n'/2)+1}$ to G that is a complete graph of $(3n'/2) + 1$ nodes with every edge having color 2, transform this to the set-theoretic version of FMC using the standard transformation from node cover to set cover and set $k = k' + 1$; note that $n = 3n'/2 + \binom{(3n'/2)+1}{2} = \Theta((n')^2)$ and $a = 3n'/2 = O(\sqrt{n})$. To complete the proof, note that any feasible solution for the FMC$(2, k)$ instance must contain *exactly* one node from $\mathscr{K}_{(3n'/2)+1}$ covering $3n'/2$ edges and therefore the solution for the edges with color 1 must correspond to a node cover in G (and *vice versa*).
(c) We given a different reduction from X3C. Given an instance $\langle \mathscr{U}', \mathscr{S}_1, \ldots, \mathscr{S}_{n'} \rangle$ of X3C as in (a), we create the following instance $\langle \mathscr{U}, \mathscr{T}_1, \ldots, \mathscr{T}_{n'}, k \rangle$ of FMC(n', k):

(i) For every set $\mathscr{S}_i = \{u_{i_1}, u_{i_2}, u_{i_3}\}$ of X3C we have three elements $u^i_{i_1}, u^i_{i_2}, u^i_{i_3}$ and a set $\mathscr{T}_i = \{u^i_{i_1}, u^i_{i_2}, u^i_{i_3}\}$ in FMC (and thus $n = 3n'$, $a = 3$ and $f = 1$),

(ii) $w(u^i_{i_j}) = 1$ and $\mathscr{C}(u^i_{i_j}) = i_j$ for $i \in \{1, \ldots, n'\}$, $j \in \{1, 2, 3\}$ (and thus $\chi = n' = n/3$),
(iii) $k = n'/3 = n/9$.

The proof is completed by showing the given instance of X3C has a solution if and only if the transformed instance of FMC$(n/3, n/9)$ has a solution. This can be shown as follows. We include the set $\mathscr{T}_i$ in the solution for FMC if and only if the set $\mathscr{S}_i$ is in the solution for X3C. For any valid solution of X3C and every $j \in \{1, \ldots, n'\}$ the element $u_j \in \mathscr{U}'$ appears in *exactly one* set, say $\mathscr{S}_\ell = \{u_{\ell_1}, u_{\ell_2}, u_{\ell_3}\}$, of X3C where one of the elements, say u_{ℓ_1}, is u_j. Then, the solution of FMC contains exactly one element, namely the element $u^\ell_{\ell_1}$, of color $\ell_1 = j$. Conversely, given a feasible solution of FMC with at most $k \leq n/9$ sets, first note that if $k < n/9$ then the total number of colors of various elements in the solution is $3k < n'$ and thus the given solution is not valid. Thus, $k = n/9$ and therefore the solution of X3C contains $n/9 = n'/3$ sets. Now, for every color j the solution of FMC contains a set, say $\mathscr{T}_\ell = \{u^\ell_{\ell_1}, u^\ell_{\ell_2}, u^\ell_{\ell_3}\}$, containing an element of color j, say the element $u^\ell_{\ell_1}$. Then $\ell_1 = j$ and the element u_j appears in a set in the solution of X3C. To see that remaining claims about the reduction, there is no solution of FMC that includes at least one element of every color and that is not a solution of X3C.

References

1. Ageev, A., Sviridenko, M.: Pipage rounding: a new method of constructing algorithms with proven performance guarantee. J. Comb. Optim. **8**(3), 307–328 (2004)
2. Angwin, J., Parris, T.: Facebook lets advertisers exclude users by race, Propublica, October 28, 2016. http://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race
3. Apollonio, N., Simeone, B.: The maximum vertex coverage problem on bipartite graphs. Discrete Appl. Math. **165**, 37–48 (2014)
4. Asudeh, A., Jin, Z., Jagadish, H.V.: Assessing and remedying coverage for a given dataset. In: 35th Annual IEEE International Conference on Data Engineering, pp. 554–565 (2019)
5. Austrin, P., Khot, S., Safra, M.: Inapproximability of vertex cover and independent set in bounded degree graphs. Theory Comput. **7**(1), 27–43 (2011)
6. Austrin, P., and Stankovic, A.: Global cardinality constraints make approximating some max-2-CSPs harder. In: Achlioptas, D., Végh, L.A. (eds.) Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, vol. 145, pp. 24:1–24:17. Leibniz International Proceedings in Informatics (2019)
7. Bansal, N.: Constructive algorithms for discrepancy minimization. In: 51st Annual IEEE Symposium on Foundations of Computer Science, pp. 3–10 (2010)
8. Beck, J., Fiala, T.: "Integer-making" theorems. Discrete Appl. Math. **3**(1), 1–8 (1981)
9. Bera, S., Chakrabarty, D., Flores, N., Negahbani, M.: Fair Algorithms for clustering. In: Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 32, pp. 4954–4965 (2019)
10. Carnes, T., Shmoys, D.: Primal-dual schema for capacitated covering problems. In: Lodi, A., Panconesi, A., Rinaldi, G. (eds.) Integer Programming and Combinatorial Optimization, Lecture Notes in Computer Science, vol. 5035, pp. 288–302. Springer, Berlin (2008)
11. Carr, R.D., Fleischer, L.K., V. Leung, J., Phillips, C.A.: Strengthening integrality gaps for capacitated network design and covering problems. In: 11th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 106–115 (2000)
12. Chazelle, B.: The Discrepancy Method: Randomness and Complexity. Cambridge University Press, Cambridge (2000)

13. Chierichetti, F., Kumar, R., Lattanzi, S., Vassilvitskii, S.: Fair clustering through fairlets. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 30, pp. 5029–5037 (2017)
14. Doerr, B.: Discrepancy in different numbers of colors. Discrete Math. **250**, 63–70 (2002)
15. Doerr, B., Srivastav, A.: Approximation of multi-color discrepancy. In: Hochbaum, D., Jansen, K., Rolim, J.D.P., Sinclair, A. (eds.) Randomization, Approximation and Combinatorial Optimization, Lecture Notes in Computer Science, vol. 1671, pp. 39–50. Springer, Berlin (1999)
16. Feige, U.: A threshold of $\ln n$ for approximating set cover. J. ACM **45**(4), 634–652 (1998)
17. Feige, U., Langberg, M.: Approximation algorithms for maximization problems arising in graph partitioning. J. Algorithms **41**, 174–211 (2001)
18. Gandhi, R., Khuller, S., Srinivasan, A.: Approximation algorithms for partial covering problems. J. Algorithms **53**(1), 55–84 (2004)
19. Garey, M.R., Johnson, D.S.: Computers and Intractability—A Guide to the Theory of NP-Completeness. W. Freeman, H., & Co., San Francisco (1979)
20. Guo, J., Niedermeier, R., Wernicke, S.: Parameterized complexity of vertex cover variants. Theory Comput. Syst. **41**(3), 501–520 (2007)
21. Gupta, A., Lee, E., Li, J.: An FPT algorithm beating 2 approximation for k-cut. In: 29th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 2821–2837 (2018)
22. Gupta, A., Lee, E., Li, J.: Faster exact and approximate algorithms for k-cut. In: 59th Annual IEEE Symposium on Foundations of Computer Science, pp. 113–123 (2018)
23. Guse, C.: Citi Bike neglects poor NYC neighborhoods and communities of color: report, New York Daily News, July 10, 2019. https://bit.ly/3aEB4rz
24. Han, Q., Ye, Y., Zhang, H., Zhang, J.: On approximation of max-vertex-cover. Eur. J. Oper. Res. **143**, 342–355 (2002)
25. Hardt, M., Price, E., Srebro, N.: Equality of opportunity in supervised learning. In: 30th International Conference on Neural Information Processing Systems, pp. 3323–3331 (2016)
26. Hochbaum, D.S.: Approximating covering and packing problems: set cover, vertex cover, independent set, and related problems. In: Hochbaum, D.S. (ed.) Approximation Algorithms for NP-Hard Problems, pp. 94–143. PWS Publishing Company, Boston (1997)
27. Hunter, A.: Job ads show sexism still prevalent in most industries, HRD, February 22 (2019) (bit.ly/38eom11)
28. Ingold, D., Soper, S.: Amazon Doesn't Consider the Race of Its Customers. Should It?, Bloomberg, April 21 (2016)
29. Jan, T.: Redlining was banned 50 years ago. It's still hurting minorities today, Washington Post, March 28 (2018)
30. Kulik, A., Shachnai, H., Tamir, T.: Maximizing submodular set functions subject to multiple linear constraints. In: 20th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 545–554 (2009)
31. Langberg, M.: Approximation algorithms for maximization problems arising in graph partitioning, M. Sc. Thesis, Weizmann Institute of Science (1998)
32. Lau, L.C., Ravi, R., Singh, M.: Iterative Methods in Combinatorial Optimization. Cambridge University Press, Cambridge (2011)
33. Lee, J., Lubienski, C.: The impact of school closures on equity of access in Chicago. Educ. Urban Soc. **49**(1), 53–80 (2017)
34. Levy, A., Ramadas, H., Rothvoss, T.: Deterministic discrepancy minimization via the multiplicative weight update method. In: Eisenbrand, F., Koenemann, J. (eds.) Integer Programming and Combinatorial Optimization, Lecture Notes in Computer Science, vol. 10328, pp. 380–391. Springer, Berlin (2017)
35. Manurangsi, P.: A note on max k-vertex cover: faster FPT-AS, smaller approximate kernel and improved approximation. In: 2nd Symposium on Simplicity in Algorithms, pp. 15:1–15:21 (2019)
36. Marx, D.: Parameterized complexity and approximation algorithms. Comput. J. **51**(1), 60–78 (2008)
37. Motwani, R., Raghavan, P.: Randomized Algorithms. Cambridge University Press, Cambridge (1995)
38. Mulshine, M.: A major flaw in Google's algorithm allegedly tagged two black people's faces with the word 'gorillas', Business Insider, July 1 (2015)
39. Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V.: Algorithmic Game Theory. Cambridge University Press, Cambridge (2007)
40. Olteanu, A., Castillo, C., Diaz, F., Kiciman, E.: Social data: biases, methodological pitfalls, and ethical boundaries. Front. Big Data **2**, 13 (2019)

41. Panconesi, A., Srinivasan, A.: Randomized distributed edge coloring via an extension of the Chernoff–Hoeffding bounds. SIAM J. Comput. **26**(2), 350–368 (1997)
42. Petrank, E.: The hardness of approximation: gap location. Comput. Complex. **4**, 133–157 (1994)
43. Simon, M.: HP looking into claim webcams can't see black people, CNN, December 23 (2009)
44. Spencer, J.: Six standard deviations suffice. Trans. Am. Math. Soc. **289**, 679–706 (1985)
45. Spencer, S.: Amazon to Bring Same-Day Delivery to Roxbury After Outcry, Bloomberg, April 26 (2016)
46. Srinivasan, A.: Distributions on level-sets with applications to approximation algorithms. In: 42nd Annual IEEE Symposium on Foundations of Computer Science, pp. 588–597 (2001)
47. Srivastav, A.: Multicolour discrepancies. Comb. Probab. Comput. **12**, 365–399 (2003)
48. Vazirani, V.: Approximation Algorithms. Springer, Berlin (2001)
49. Yan, A., Howe, B.: Fairness in practice: a survey on equity in urban mobility. Data Eng. Bull. **42**(4), 49 (2019)
50. Zemel, R., Wu, Y., Swersky, K., Pitassi, T., Dwork, C.: Learning fair representations. In: 30th International Conference on Machine Learning, PMLR, vol. 28(3), pp. 325–333 (2013)
51. Zhang, B.H., Lemoine, B., Mitchell, M.: Mitigating unwanted biases with adversarial learning. In: AAAI/ACM Conference on AI, Ethics, and Society, pp. 335–340 (2018)