Parameterized Algorithms for Colored Clustering

Leon Kellerhals[1], **Tomohiro Koana**[1], **Pascal Kunz**[1,2], **and Rolf Niedermeier**[1]

[1] Technische Universität Berlin, Algorithmics and Computational Complexity, Berlin, Germany
[2] Humboldt-Universität zu Berlin, Algorithm Engineering, Berlin, Germany
{leon.kellerhals, tomohiro.koana, p.kunz.1}@tu-berlin.de

Abstract

In the COLORED CLUSTERING problem, one is asked to cluster edge-colored (hyper-)graphs whose colors represent interaction types. More specifically, the goal is to select as many edges as possible without choosing two edges that share an endpoint and are colored differently. Equivalently, the goal can also be described as assigning colors to the vertices in a way that fits the edge-coloring as well as possible.

As this problem is NP-hard, we build on previous work by studying its parameterized complexity. We give a $2^{\mathcal{O}(k)} \cdot n^{\mathcal{O}(1)}$-time algorithm where k is the number of edges to be selected and n the number of vertices. We also prove the existence of a kernel of size $\mathcal{O}(k^{5/2})$, resolving an open problem posed in the literature. We consider parameters that are smaller than k, the number of edges to be selected, and r, the number of edges that can be deleted. Such smaller parameters are obtained by considering the difference between k or r and some lower bound on these values. We give both algorithms and lower bounds for COLORED CLUSTERING with such parameterizations. Finally, we settle the parameterized complexity of COLORED CLUSTERING with respect to structural graph parameters by showing that it is W[1]-hard with respect to both vertex cover number and tree-cut width, but fixed-parameter tractable with respect to slim tree-cut width.

Introduction

Graph clustering is one of the most fundamental tasks in analyzing data that captures interactions between entities. The idea is that if two vertices are in the same cluster, then the corresponding entities are similar in terms of their interactions. Typical approaches lead to meaningful clusterings whenever the edges model interactions of the same type, possibly with weights. Those approaches are not, however, designed to deal with data that captures interactions of different types. There are several settings in which the clusters should capture similarity not only in terms of interactions, but in terms of types of interactions. For instance, brain coactivation graphs capture which brain regions are active or inactive at the same time when exposed to certain types of stimuli (Crossley et al. 2013). In the Drug Abuse Warning Network,[1] interactions describe combinations of drugs taken by patients prior to an ER visit. Other settings with similar interaction categorization by type include coauthorship networks (categorized by publication venue) and copurchasing networks (categorized by type of purchase).

Angel et al. (2016) introduced an approach to finding such category-sensitive clusters, which we will call COLORED CLUSTERING (CC): Given an edge-colored graph, the goal is to color the vertices in a way that maximizes the number of *stable* edges: edges whose endpoints are both assigned the same color as the edge. Angel et al. (2016) proved the problem to be NP-hard and gave approximation algorithms as well as tractable special cases for the problem. Their approximation algorithm was improved by Ageev and Kononov (2014) and Alhamdan and Kononov (2019). Cai and Leung (2018) studied the parameterized complexity of the problem, giving FPT algorithms with respect to the number k of stable edges[2] and to the number r of unstable edges. Amburg, Veldt, and Benson (2020) introduced the problem under a different name and for hypergraphs. Veldt (2022) continued the study of the problem on hypergraphs. In several of the examples listed above, it is sensible to model the interactions as hypergraphs.

Related work. Another clustering model which also captures edge colors is chromatic correlation clustering (Bonchi et al. 2015), in which the goal is to cluster the vertices such that the number of edge modifications (additions, deletions, recolorings) to obtain a disjoint set of monochromatic cliques is minimized. This approach penalizes missing edges, that is, edges that need to be added such that each cluster becomes a clique. Chromatic correlation clustering generalizes correlation clustering (Bansal, Blum, and Chawla 2004) and thus is NP-hard even for one color. COLORED CLUSTERING does not penalize such missing edges and becomes more tractable: it is polynomial-time solvable for two colors. Closely related to graph clustering in edge-colored graphs is the field of clustering multi-layer graphs (Mucha et al. 2010; Chen et al. 2018). In that scenario, the layers do not relate to a cluster type. Another related problem is the MAXIMUM EDGE-COLORABLE SUBGRAPH (Agrawal et al. 2022), in which the goal is to prop-

[1] https://www.samhsa.gov/data/data-we-collect/dawn-drug-abuse-warning-network

[2] Roughly speaking, an FPT algorithm with respect to k has a running time of $f(k) \cdot n^{\mathcal{O}(1)}$, where k is the parameter and n is the instance size.

erly color as many edges of a graph as possible with a fixed number of colors. The difference to CC is that the colors of the edges are not prescribed, but may be freely chosen. Finally, hypergraph clustering has been studied extensively (Agarwal, Branson, and Belongie 2006; Papa and Markov 2007; Gleich, Veldt, and Wirth 2018; Fukunaga 2019; Li, Puleo, and Milenkovic 2020).

Our contributions. We continue the study of the parameterized complexity of COLORED CLUSTERING on graphs and on hypergraphs. For the parameter k, the number of stable edges, we improve on the FPT result due to Cai and Leung (2018) by giving a single-exponential time algorithm. We also prove that COLORED CLUSTERING admits a polynomial kernel for k, thus answering an open question by Cai and Leung (2018). Both results translate to hypergraphs with constant-sized edges. As the problem is FPT with respect to both the number of stable and unstable edges, we consider above guarantee parameterizations. We introduce the concept in the corresponding section and show that it can be used to obtain fixed-parameter algorithms for parameters that are smaller than the number of stable edges and a parameter smaller than the number of unstable edges. Again, both results can be lifted to work for hypergraphs. We complement these results with hardness proofs for above-guarantee parameters that are even slightly smaller. We also consider structural graph parameters. The problem is hard even for fairly large structural parameters: We show that there is presumably no FPT algorithm for the parameters vertex cover number and tree-cut width. However, the problem is FPT when parameterized by the slim tree-cut width. Finally, we close a gap in the classical tractability of COLORED CLUSTERING by proving that the problem is NP-hard on graphs even if every vertex has degree at most three and on hypergraphs even if every vertex has degree at most two.

Preliminaries

Graphs and problem definition. For standard graph terminology, we refer to Diestel (2017). A *hypergraph* $G = (V, E)$ consists of a finite vertex set V and an edge set $E \subseteq 2^V$. It is a *graph* if $|e| = 2$ for all $e \in E$. Let $v \in V$ be a vertex. We denote its set of *incident* edges by $\delta_G(v) := \{e \in E \mid v \in e\}$ and its *neighborhood* by $N_G(v) := (\bigcup_{e \in \delta(v)} e) \setminus \{v\}$. The *degree* of v is $\deg_G(v) := |\delta_G(v)|$. For an edge set $F \subseteq E$, let $\deg_{G,F}(v) = |\delta(v) \cap F|$. Let C be a finite set of colors. An *edge coloring* is a function $\ell \colon E \to C$ and a *vertex coloring* a function $f \colon V \to C$. The *chromatic degree* of $v \in V$ is $\deg_G^X(v) := \{c \in C \mid \exists e \in \delta(v) \colon \ell(e) = c\}$ For any $c \in C$, the color-c degree of $v \in V$ is $\deg_{G,c}(v) := |\{e \in \delta(v) \mid \ell(e) = c\}|$. We drop the subscript $\cdot_G$ whenever it is clear from the context. A set of edges $F \subseteq E$ in a hypergraph $G = (V, E)$ with an edge coloring $\ell \colon E \to C$ is *stable* if $\ell(e) = \ell(e')$ for all $e, e' \in F$ with $e \cap e' \neq \emptyset$. This leads to the computational problem of selecting a largest set of stable edges:

COLORED (HYPERGRAPH) CLUSTERING (CC/CHC)
Input: A (hyper)-graph $G = (V, E)$, an edge coloring $\ell \colon E \to C$, and an integer $k \in \mathbb{N}$.
Question: Does G contain a stable edge set F of size at least k?

Sometimes it is more convenient to express stability using vertex colorings. An edge $e \in E$ is *stable* under the vertex coloring $f \colon V \to C$, if $f(v) = \ell(e)$ for all $v \in E$ and unstable otherwise.

One can transform any stable edge set F into a vertex coloring $f_F \colon V \to C$ by setting:

$$f_F(v) := \begin{cases} \ell(e), & \text{if } v \in e \in F, \\ \bot, & \text{if } v \text{ is not incident to any edges in } F, \end{cases}$$

where $\bot \in C$ is an arbitrary default color. We note that if $F \subseteq E$ is stable, then the edges in F are stable under f_F.

Parameterized complexity. Let Σ be a finite alphabet. A *parameterized problem* $L \subseteq \Sigma^* \times \mathbb{N}$ is a subset of all instances (x, κ) in $\Sigma^* \times \mathbb{N}$, and κ is the *parameter*. A parameterized problem L is (i) *fixed-parameter tractable* (or contained in the class FPT) if there is an algorithm that decides L in $f(\kappa) \cdot |x|^{O(1)}$ time, (ii) contained in the class XP if there is an algorithm that decides L in $|x|^{f(\kappa)}$ time, and (iii) *para-NP-hard* if L is NP-hard for any constant value of the parameter, where $f \colon \mathbb{N} \to \mathbb{N}$ is any computable function that only depends on the parameter. Note that FPT $\subseteq$ XP. For running time bounds, we use the $\mathcal{O}^*$ notation which hides factors that are polynomial in the input size. If a parameterized problem is W[1]-*hard*, then it is presumably not in FPT, and if it is para-NP-hard, then it is not in XP (unless P = NP). A *kernel* for L is a polynomial-time algorithm that takes the instance (x, κ) and outputs a second instance (x', κ') such that (i) $(x, \kappa) \in L \iff (x', \kappa') \in L$ and (ii) $|(x', \kappa')| \leq f(\kappa)$ for a computable function f. The *size* of the kernel is f. For further details, we refer to the standard literature (Cygan et al. 2015; Downey and Fellows 2013).

Parameterizing by Number of Stable Edges

Cai and Leung (2018) showed that COLORED CLUSTERING is FPT with respect to the maximum number k of stable edges. Their algorithm uses color-coding and can be derandomized to yield a deterministic algorithm with running time $k^{2k+\mathcal{O}(\log k)}(n+m)$. In the following, we will improve on this running time and give a single-exponential algorithm for CC parameterized by k. Cai and Leung asked if CC has a kernel that is polynomial in k. We will show that the problem does, indeed, admit a kernel of size $\mathcal{O}(k^{5/2})$.

Single-Exponential Time Algorithm

Our single-exponential time algorithm is a parameterized reduction to WEIGHTED EXACT COVER, defined as follows.

Weighted Exact Cover	
Input:	A universe U, a family $\mathcal{S}$ of nonempty subsets of U, a weight function $w\colon \mathcal{S} \to \mathbb{N}$, and integers s, W.
Question:	Is there a subfamily $\mathcal{S}'$ of pairwise disjoint subsets with $\lvert\bigcup_{S\in\mathcal{S}'} S\rvert = s$ such that $\sum_{S\in\mathcal{S}'} w(S) \geq W$?

Using the fact that Weighted Exact Cover can be solved in $\mathcal{O}(2.851^s|S|\cdot|U|\log^2|U|)$ time (Goyal et al. 2015; Shachnai and Zehavi 2017), we prove the following.

Theorem 1. Colored Clustering *can be solved in* $\mathcal{O}^*(2^{\mathcal{O}(k)})$ *time.*

Proof. We provide a parameterized reduction to Weighted Exact Cover. Let $(G = (V,E), \ell, k)$ be an instance of Colored Clustering. For each color $c \in C$ denote by G^c the subgraph spanned by the edges of color c (note that G^c does not contain isolated vertices). Let q_c be the number of connected components in G^c and let $Q^c_1, Q^c_2, \ldots, Q^c_{q_c} \subseteq V$ be the connected components. If $q_c \geq k$ for any color c or $G^c[Q^c_p]$ has at least k edges for any c and $p \in [q_c]$, then this is a `yes`-instance. So, assume the contrary. We create an instance $(U, \mathcal{S}, s := 2k, W := k)$ as follows. Set $U := V \uplus \{x_1, \ldots, x_s\}$. For each color $c \in C$, for each component Q^c_p, $p \in [q_c]$, and for each nonempty subset $X \subseteq Q^c_p$, we add X to $\mathcal{S}$ and set $w(X) := |\{\{u,v\} \in E \mid u,v \in X, \ell(\{u,v\}) = c\}|$, that is, the number of edges of color c in the graph induced by X. We also add the singletons $\{x_1\}, \ldots, \{x_s\}$ to $\mathcal{S}$ and set $w(\{x_i\}) := 0$.

Suppose that the instance of CC is a `yes`-instance, that is, there exists a stable edge set $F \subseteq E$ of size at least k. Then, for each connected component $X \subseteq V$ of $G_F = (V, F)$ of size at least two, there is a set $X \in \mathcal{S}$ whose weight is the number of edges in $G_F[X]$, since the edges of $G_F[X]$ have the same color. These connected components are pairwise disjoint. We create $\mathcal{S}'$ by adding all of these sets X and, if $|\bigcup_{S\in\mathcal{S}'} S| < s$, an appropriate number of singletons $\{x_i\}$. Thus, $\mathcal{S}'$ covers exactly s elements and its weight is at least W.

Conversely, suppose that there exists a subfamily $\mathcal{S}'$ with weight at least W and $|\bigcup_{S\in\mathcal{S}'} S| = s$. Each subset in $\mathcal{S}'$ except for the singletons corresponds to a set V' of vertices such that, by construction, $G[V']$ contains $w(V')$ edges, all having the same color. Hence, setting $f(v) = c$ for each $v \in V'$ yields a solution for the CC instance with at least k stable edges.

As for the running time, note that the instance of Weighted Exact Cover contains at most $|C| \cdot k \cdot 2^k + 2k$ sets in $\mathcal{S}$ because each component Q^c_i has at most k vertices. Using the aforementioned algorithm yields a running time of $\mathcal{O}(2.851^s \cdot (|C|\cdot k \cdot 2^k + 2k)(|V| + 2k)\log^2(|V| + 2k)) = \mathcal{O}(16.26^k \cdot k^2 \log^2 k \cdot |V| \log^2 |V|)$. $\square$

The above result can be generalized to hypergraphs with maximum edge size d. Since a connected component of a hypergraph with at most $k-1$ edges may have at most $d(k-1)$ vertices, the instance of Weighted Exact Cover has at most $|C| \cdot k \cdot 2^{dk} + dk$ sets.

Corollary 2. Colored Hypergraph Clustering *can be solved in* $\mathcal{O}^*(2^{\mathcal{O}(d^2k)})$ *time.*

A Polynomial Kernel

We next show that Colored Clustering admits a kernel of size polynomial in the number k of stable edges, answering an open question posed by Cai and Leung (2018).

We begin with the observation that a matching is stable:

Reduction rule 1. If G has a matching of size k, then return `yes`.

Let M be any maximal matching and let $S := \bigcup_{e\in M} e$. Then, S is a vertex cover of size at most $2k$ by Reduction Rule 1. Thus, it remains to bound the size of the independent set $I := V \setminus S$. It can be arbitrarily large at this point; in fact, we will show that CC is W[1]-hard when parameterized by the vertex cover number, so it is presumably not possible to bound the size of I in terms of just S.

We use another simple observation that a set of monochromatic edges is stable:

Reduction rule 2. If there are k edges of the same color, then return `yes`.

In view of Reduction Rule 2, we only need to bound the number of colors. We can achieve this by bounding the chromatic degree of every vertex.

Reduction rule 3. If there is a vertex $v \in V$ of chromatic degree at least $2k + 1$, then delete all edges with color c in $\delta(v)$, where c is the least frequent color in $\delta(v)$.

Lemma 3. *Reduction Rule 3 is correct.*

Proof. Let $G' = (V, E')$ be the resulting graph after applying Reduction Rule 3. If $F \subseteq E'$ with $|F| \geq k$ is stable in G', then it is also stable in G.

Conversely, suppose that $F \subseteq E$ with $|F| = k$ is stable in G. We show that there exists an edge set $F' \subseteq E'$ of size k which is stable in G'. If $f_F(v) \neq c$, then F is stable in G since we have only deleted edges of color c incident to v. Suppose that $f_F(v) = c$. Since there are at least $2k$ other colors among the edges incident to v, and $|\bigcup_{e\in F} e| \leq 2k$, there exists a color $c' \neq c$ such that the edges incident to v with color c' are vertex-disjoint to the edges in F (except for v). As c is the least frequent color among the edges incident to v, we obtain a desired solution F by replacing the edges of color c with those of color c'. $\square$

At this point, we can show that the resulting graph has at most $\mathcal{O}(k^3)$ edges. Recall that S is a vertex cover, i.e., every edge is incident to S. Thus, there are at most $(2k+1)\cdot|S| \in \mathcal{O}(k^2)$ colors that appear at least once. By Reduction Rule 2, each color appears at most k times in the graph by Reduction Rule 2, and consequently, we have $\mathcal{O}(k^3)$ edges.

We will show how to improve upon this using a meet-in-the-middle argument. For this, we denote for a vertex v and color c the set of neighbors connected by edges of color c by $N_c(v)$. Intuitively speaking, if we have "many" vertices v in S such that there are "many" colors c appearing "many" times in $\delta(v)$, then we can construct a stable set of size k.

Reduction rule 4. Let $T \subseteq S$ be the set of vertices v with $|\{c \in C \mid |N_c(v) \setminus S| \geq 2k^{1/2}\}| \geq 2k^{1/2}$. If $|T| \geq k^{1/2}$, then return `yes`.

Lemma 4. *Reduction Rule 4 is correct.*

Proof. Assume without loss of generality that $t := |T| = k^{1/2}$ and let $T := \{v_1, \ldots, v_t\}$. We show that the following greedy algorithm finds a stable set $F = F_t$ of size k: We begin with $F_0 := \emptyset$. We will construct F_q for increasing $q \in [t]$ such that $|F_q| = \sum_{p=1}^{q} 2k^{1/2} - (p-1)$. Let $V_{F_{q-1}} := (\bigcup_{e \in F_{q-1}} e) \setminus S$ be the set of endpoints of F_{q-1} that are outside of S. For increasing $q \in [t]$ let $c := \arg\max_{c' \in C} |N_{c'}(v_q) \setminus (S \cup V_{F_{q-1}})|$. We claim that $|N_c(v_q) \setminus (S \cup V_{F_{q-1}})| \geq 2k^{1/2} - (q-1)$. Note that $|V_{F_{q-1}}| \leq |F_{q-1}| \leq 2(q-1)k^{1/2}$. As there are $2k^{1/2}$ colors $c \in C$ such that $|N_c(v_q) \setminus S| \geq 2k^{1/2}$, we have by pigeonhole principle that there exists a color c with $|N_c(v_q) \cap V_{F_{q-1}}| \leq |V_{F_{q-1}}|/(2k^{1/2}) = q-1$. Hence, $|N_c(v_q) \setminus S| - |N_c(v_q) \cap V_{F_{q-1}}| \geq 2k^{1/2} - (q-1)$. Now, let F_q be formed by F_{q-1} and $2k^{1/2} - (q-1)$ arbitrary edges between v_q and $N_c(v_q) \setminus (S \cup V_{F_{q-1}})$. Finally, let $F := F_t$. We have $|F| = \sum_{q \in [t]} |F_q| - |F_{q-1}| = \sum_{q \in [t]} 2k^{1/2} - (q-1) \geq \sum_{q \in [t]} k^{1/2} = k$. □

After applying Reduction Rule 4, we have "not too many" vertices such that there are "many" colors c appearing "many" times in $\delta(v)$, which results in a smaller kernel.

Theorem 5. COLORED CLUSTERING *admits a kernel of size* $\mathcal{O}(k^{5/2})$.

Proof. We apply Reduction Rules 1 to 4 exhaustively and delete all isolated vertices. We show that the resulting graph has size $\mathcal{O}(k^{5/2})$.

Let T be as specified in Reduction Rule 4. We can assume that $|T| \leq k^{1/2}$ and thus there are at most $\mathcal{O}(k^2 \cdot |T|) = \mathcal{O}(k^{5/2})$ incident to T. Now consider a vertex $v \in S \setminus T$. Note that it suffices to bound the number of edges one of whose endpoint is in S, since there are at most $\mathcal{O}(k^2)$ edges with both endpoints in S. Let X be the set of colors c such that $|N_c(v) \setminus S| \geq 2k^{1/2}$ and Y be the set of colors $c \notin X$ such that $N_c(v) \setminus S \neq \emptyset$. By the definition of T, we have at most $2k^{1/2}$ colors $c \in X$, amounting to at most $k \cdot |X| = \mathcal{O}(k^{3/2})$ edges incident to v of color $c \in X$. Moreover, there are at most $\mathcal{O}(k^{3/2})$ edges incident to v of color in Y since at most $k^{1/2}$ edges of $\delta(v)$ have color $c \in Y$ and $|Y| \in \mathcal{O}(k)$. Thus, there are at most $\mathcal{O}(k^{5/2})$ vertices and edges in the graph since all isolated vertices have been deleted. □

This result generalizes to hypergraphs with maximum edge size d. From Reduction Rule 1 we obtain a vertex cover of size at most dk. Reduction Rule 2 translates immediately. We can adapt Reduction Rule 3 to bound the chromatic degree by dk. As observed above, this already yields a kernel.

Corollary 6. COLORED HYPERGRAPH CLUSTERING *admits a kernel of size* $\mathcal{O}(d^2k^3)$.

Note that the meet-in-the middle argument does not translate to hypergraphs as in Lemma 4 we implicitly assume each edge to have at most one endpoint outside S.

Above-Guarantee Parameters

As shown by Cai and Leung (2018), COLORED CLUSTERING can be solved in time $\mathcal{O}^*(1.2783^r)$, where $r := |E| - k$ is the number of unstable edges, by reduction to VERTEX COVER. The recently proposed $\mathcal{O}^*(1.2530^k)$ time algorithm for VERTEX COVER (Harris and Narayanaswamy 2022) can be used to improve this running time. There is a line of research that seeks to improve on such algorithms by introducing smaller parameters that take a lower bound on the parameter and then use the difference between the traditional parameter and the lower bound as a smaller parameter. This approach has become known as "above guarantee" parameterization (see, e.g., (Gutin and Mnich 2022; Kellerhals, Koana, and Kunz 2022; Mahajan and Raman 1999)).

One above-guarantee parameterization for COLORED CLUSTERING can be obtained by using the conflict graph of an edge-colored graph. The *conflict graph* $\partial(G, \ell)$, introduced by Angel et al. (2016), of a graph $G = (V, E)$ with an edge coloring $\ell\colon E \to C$ contains a vertex for every edge in G. Two vertices in $\partial(G, \ell)$ are adjacent if the corresponding edges in G have different colors and share a vertex. Formally, $\partial(G, \ell) = (E, \partial E)$, where $\partial E = \{\{e, e'\} \subseteq E \mid e \cap e' \neq \emptyset \text{ and } \ell(e) \neq \ell(e')\}$. Any stable edge set in G corresponds to an independent set in $\partial(G, \ell)$ and consequently the minimum number of unstable edges in G is the vertex cover number of $\partial(G, \ell)$. It follows that any lower bound on the size of a minimum vertex cover in $\partial(G, \ell)$ is a lower bound on r. VERTEX COVER is FPT when parameterized above the size of a maximum matching or above the minimum fractional solution of a standard linear program for VERTEX COVER (Lokshtanov et al. 2014; Razgon and O'Sullivan 2009). These results can also be used to solve COLORED CLUSTERING parameterized above the corresponding lower bounds on $\partial(G, \ell)$.

Above Degree-Based Lower Bound

We will consider a different lower bound on the number r of edges that must be unstable in any vertex coloring of a graph. If we consider the edges incident to a particular vertex v, all of these edges save the edges of one color are unstable. Hence, the number of unstable edges incident to v is at least $\deg(v) - \max_{c \in C} \deg_c(v)$. To obtain a lower bound on the total number of edges that must be unstable in a graph, we can add up this value over all vertices in the graph, but we must divide the sum by two to account for the fact that we may count edges twice in this way. This lower bound can be formalized as follows. For any graph $G = (V, E)$, define

$$\rho(G, \ell) := \frac{1}{2} \sum_{v \in V} \big(\deg(v) - \max_{c \in C} \deg_c(v)\big).$$

Observe that for any `yes`-instance $(G = (V, E), \ell, k = |E| - r)$ of COLORED CLUSTERING, we have $r \geq \rho(G, \ell)$.

Hence, in the following we will consider the parameterized complexity of COLORED CLUSTERING with respect to the parameter $r-\rho(G,\ell)$. Unfortunately, it turns out that this problem is para-NP-hard in general (Theorem 9). However, our finding is that if we consider the parameter $r-\rho'(G,\ell)$ obtained from a smaller, yet still tight, lower bound, COLORED CLUSTERING becomes FPT:

$$\rho'(G,\ell) := \frac{1}{2}\sum_{v\in V}\min\left(\deg(v)-\max_{c\in C}\deg_c(v), \tfrac{1}{2}\deg(v)\right).$$

By definition, $\rho(G,\ell) \geq \rho(G',\ell)$ and hence $r \geq \rho'(G,\ell)$ for any yes-instance of COLORED CLUSTERING. We show that COLORED CLUSTERING is FPT for the parameter $r-\rho'(G,\ell)$ by proving that $\rho'(G,\ell)$ is at most the optimal value of the LP relaxation of VERTEX COVER on the conflict graph. This allows us to use the FPT algorithm for VERTEX COVER parameterized by the solution size minus the optimal value of the LP relaxation (Lokshtanov et al. 2014).

Theorem 7 (★[3]). COLORED CLUSTERING *is* FPT *with respect to* $r-\rho'(G,\ell)$.

We remark that this leads to an FPT algorithm for the smaller parameter $r-\rho(G,\ell)$ if the maximum chromatic degree is two. Note that COLORED CLUSTERING remains NP-hard with this restriction (Angel et al. 2016).

Corollary 8 (★). COLORED CLUSTERING *is FPT with respect to* $r-\rho(G,\ell)$ *if the chromatic degree of every vertex is at most two.*

We complement our positive results by showing that when each vertex may be incident to edges of three (or more) colors, the problem is para-NP-hard with respect to $r-\rho(G,\ell)$. The proof is by a reduction from MONOTONE ONE-IN-THREE SAT. It utilizes the fact that, if $r-\rho(G,\ell)=0$, then a vertex coloring must make all edges in one of the most frequent colors incident to a vertex stable.

Theorem 9 (★). COLORED CLUSTERING *is* NP*-hard even if* $r-\rho(G,\ell)=0$ *and the chromatic degree of every vertex is at most three.*

Hypergraphs. We can lift the algorithm in Theorem 7 to hypergraphs with maximum edge size d, albeit with a smaller lower bound

$$\rho_H(G,\ell) := \frac{1}{d}\sum_{v\in V}\min\left(\deg(v)-\max_{c\in C}\deg_c(v), \tfrac{1}{2}\deg(v)\right).$$

Corollary 10 (★). COLORED HYPERGRAPH CLUSTERING *is FPT with respect to* $r-\rho_H(G,\ell)$.

Above Matching-Based Lower Bounds

We will now consider matchings as lower bounds on k. A *matching* in a graph $G=(V,E)$ is a set of edges $M\subseteq E$ such that $e\cap e'\neq\emptyset$ for any two distinct $e,e'\in M$. Let $M(G)$ denote the size of a maximum matching in a graph G. As we noted in Reduction Rule 1, any matching is trivially stable, which implies the following:

Observation 11. *If* $M(G)\geq k$, *then* $(G=(V,E),\ell,k)$ *is a* yes*-instance.*

[3]Proofs of claims marked with ★ are deferred to the full version.

A matching M in G is *induced* if there are no $e,e'\in M$ and $v\in e, v'\in e'$ such that v and v' are adjacent. Let $I(G)$ denote the size of a maximum induced matching in a graph G. Of course, $I(G)\leq M(G)$, implying that if $I(G)\geq k$, then $(G=(V,E),\ell,k)$ is a yes-instance. We will consider the parameters $k-M(G)$ and $k-I(G)$. Finding a maximum induced matching is NP-hard (Stockmeyer and Vazirani 1982), so we assume that such a matching is given as part of the input, when dealing with the latter parameter.

Our main result concerning parameterizations of COLORED CLUSTERING above matching-based lower bounds is that the problem is FPT with respect to $(k-I(G))+|C|$ (Theorem 13). We also show that the problem is XP with respect to $k-I(G)$ (Corollary 14). We will also show that this is, in a sense, the best one can do. We will prove that COLORED CLUSTERING is NP-hard for $k-M(G)=1$ and $|C|=19$ (Theorem 16), implying that there is not even XP algorithm for the parameter $(k-M(G))+|C|$, unless P = NP. Moreover, we prove that CC is W[1]-hard with respect to $k-I(G)$ (Theorem 17), meaning that there is no FPT algorithm for $k-I(G)$, unless FPT = W[1].

We write $\deg_F(e) := \deg_F(u)+\deg_F(v)$ for an edge $e=\{u,v\}$. We will use the following observation.

Observation 12. *Let* $G=(V,E)$ *be a graph with a matching* M. *For an edge subset* $F\subseteq E$, *it holds that* $2|F| = \sum_{v\in V}\deg_F(v) = \sum_{v\in X}\deg_F(v)+\sum_{e\in M}\deg_F(e)$, *where* $X := V\setminus\bigcup_{e\in M} e$ *is the set of unmatched vertices.*

First, we present an FPT algorithm for $(k-I(G))+|C|$.

Theorem 13. *Given a maximum induced matching,* COLORED CLUSTERING *can be solved in* $\mathcal{O}^*(|C|^{\mathcal{O}(k-I(G))})$ *time.*

Proof. Suppose that we are a given yes-instance $(G=(V,E),\ell,k)$ of COLORED CLUSTERING along with an induced matching M. Among possibly many solutions, our goal is to find a stable set F of size k such that $|F\setminus M|\leq|F'\setminus M|$ for any stable set of size k (in other words, F contains as many edges of M as possible). This restriction on F will play a central role in the algorithm. We will first show that $|F\setminus M|\leq 2(k-|M|)$.

We first show that each edge $e\in M\setminus F$ intersects at least two edges in $F\setminus M$. Suppose not, that is, there is an edge $e\in M\setminus F$ that intersects at most one edge in F. Let f be that edge, or let $f\in F$ be arbitrary if e intersects no edge in F. Then $F' := (F\setminus\{f\})\cup\{e\}$ is stable and $|F'\setminus M| > |F\setminus M|$, a contradiction.

So assume that each edge $e\in M\setminus F$ intersects at least two edges in $F\setminus M$. As M is an induced matching, every edge in $F\setminus M$ intersects at most one edge in M. Thus, $|F\setminus M|\geq 2|M\setminus F|$, which yields

$$\begin{aligned}|F\setminus M| &\leq |F\setminus M| + (|F\setminus M| - 2|M\setminus F|)\\ &= 2(|F|-|F\cap M|-|M\setminus F|) = 2(k-|M|).\end{aligned}$$

To obtain an FPT algorithm, we use the color coding technique. We say that a vertex coloring $f\colon V\to C$ is *good* for F if $F\setminus M$ is stable under f. We color the vertices of G independently and uniformly at random, that is we assign color $c\in C$ to each vertex V with probability $|C|^{-1}$. Then, the vertex $v \in \bigcup_{e \in F \setminus M} e$ receives the color $f_F(v)$ with probability $|C|^{-1}$. This implies that the probability that f is good is at least $|C|^{-|\bigcup_{e \in F \setminus M} e|} = |C|^{-2|F \setminus M|} \geq |C|^{-4(k-|M|)}$. Hence, we obtain a good coloring of G with constant probability by repeating the coloring $|C|^{\mathcal{O}(k-|M|)}$ times. Given a good coloring f of G, we obtain a coloring f' with at least k stable edges in polynomial time as follows. For every $e = \{u, v\} \in M$, we check whether at least two edges in $\delta(u) \cup \delta(v)$ are stable under f. If so, let $f'(u) := f(u)$ and $f'(v) := f(v)$. If not, set $f'(u) := f'(v) := \ell(e)$.

We must show that, if G contains a stable edge set of size at least k, then algorithm finds a coloring f' with at least k stable edges with constant probability. Let F be a stable edge set of size at least k that maximizes the number of edges in M. We will show that, if f is good for F, then $F \setminus M$ is stable under f' and $|\{e \in E \setminus (M \cup F) \mid e \text{ is stable under } f'\}| \geq |\{e \in M \cap F \mid e \text{ is unstable under } f'\}|$. This implies that at least $|F| = k$ edges are stable under f'. First, every $e \in F \setminus M$ is stable under f. Moreover, by the observation above, if e intersects $e' \in M \setminus F$, then there must be a second edge $e'' \in F \setminus F$. The edge e'' must also be stable under f. It follows that f and f' agree on the endpoints of e and, hence, e is stable under f'. Now consider $e = \{u, v\} \in M \cap F$. If e is unstable under f', then there must be at least two edges $e', e'' \in \delta(u) \cup \delta(v)$ that are stable under f. Because M is an induced matching, it follows that e is the only edge in M that intersects e' or e''. Hence, f and f' agree on the endpoints of e' and e'', implying that these two edges are stable under f'. Hence, for each $e \in M \cap F$ that is unstable under f', there are at least two edges in $E \setminus (M \cup F)$ that are stable under f'. ☐

The bound $|F \setminus M| \leq 2(k - |M|)$ in the proof of Theorem 13 actually implies that COLORED CLUSTERING is XP with respect to the parameter $k - |M|$:

Corollary 14 (★). *Given a maximum induced matching M,* COLORED CLUSTERING *can be solved in $n^{\mathcal{O}(k-I(G))}$ time.*

Observe that the above definition of induced matchings carries over to hypergraphs. The bound on $|F \setminus M|$ also holds on hypergraphs with maximum edge size d, and the two results above carry over. As $|\bigcup_{e \in F \setminus M} e| \leq d|F \setminus M| \leq 2d(k - |M|)$ in hypergraphs, the probability that the coloring f is good is $|C|^{-2d(k-|M|)}$ and we have to adjust the number of repetitions accordingly.

Corollary 15. *Given an induced matching M of size $I(G)$,* COLORED HYPERGRAPH CLUSTERING *can be solved in $\mathcal{O}^*(|C|^{\mathcal{O}(d(k-I(G)))})$ time and in $n^{\mathcal{O}(k-I(G))}$ time.*

Next, we will show that this XP algorithm for $k - I(G)$ cannot be improved to an XP algorithm for $k - M(G)$, unless P = NP. This suggests that the assumption that there is no edge connecting endpoints of two edges of M is imperative in the XP algorithm of Corollary 14. We will show that COLORED CLUSTERING is NP-hard when $k - M(G) = 1$ and $|C| = 19$ by a reduction from the NP-hard (Wu 2012) problem TWO DISJOINT MONOCHROMATIC PATHS, in which we are given an edge-bicolored graph and two terminal pairs and are asked to find two

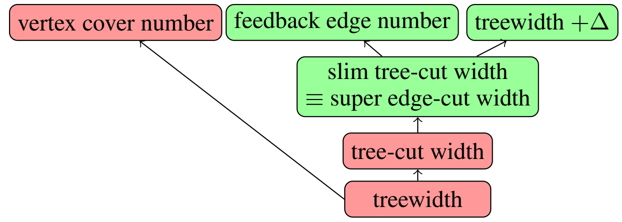

Figure 1: A Hasse diagram of graph parameters. COLORED CLUSTERING is FPT for those highlighted in green and W[1]-hard with respect to those in red. There is an arrow from p to q if there is a function f such that $f(p(G)) \geq q(G)$ for every graph G.

vertex-disjoint monochromatic paths to connect each pair. The reduction uses Observation 12 and is based on the following idea: We create a graph with a perfect matching in which there are two matching edges which can have degree three in a solution while all other matching edges can have degree at most two. A solution must then connect those two particular matching edges with a kind of path that can be translated into two disjoint monochromatic paths in the input instance.

Theorem 16 (★). COLORED CLUSTERING *is* NP*-hard even if $k - M(G) = 1$ and $|C| \leq 19$.*

Finally, we will show that the XP algorithm for $k - I(G)$ most likely also cannot be improved to an FPT algorithm. The proof is a reduction from MULTICOLORED CLIQUE.

Theorem 17 (★). COLORED CLUSTERING *is* W[1]*-hard with respect to $k - I(G)$.*

Structural Parameters

In the following, we will classify the parameterized complexity of COLORED CLUSTERING with respect to structural graph parameters. Our results are summarized in Figure 1. The problem is W[1]-hard with respect to vertex cover number, which rules out FPT algorithms for many graph parameters. Additionally, the problem is also W[1]-hard with respect to tree-cut width. On the positive side, it is FPT for the slim tree-cut width, a parameter that was very recently introduced for the express purpose of dealing with problems that are hard for tree-cut width and treewidth (Ganian and Korchemna 2022). Hence, COLORED CLUSTERING can be added to a list of problems compiled by Ganian and Korchemna (2022) that are hard for tree-cut width, but FPT for slim tree-cut width.

We start with a more general observation. A graph class $\mathcal{C}$ is *monotone* if $G \in \mathcal{C}$ and $H \subseteq G$ implies $H \in \mathcal{C}$. For a graph $G = (V, E)$, let $d^e_{\mathcal{C}}(G) := \min_{E' \subseteq E, G - E' \in \mathcal{C}} |E'|$ denote the *edge deletion distance* of G to $\mathcal{C}$. COLORED CLUSTERING is FPT with respect to $d^e_{\mathcal{C}}(G)$ if we are given the deletion set because we can guess which edges in the deletion set are stable and then solve a resulting instance in $\mathcal{C}$.

Proposition 18 (★). *Let $\mathcal{C}$ be a monotone graph class on which* COLORED CLUSTERING *is polynomial-time solv-*

able. Then, CC *is* FPT *with respect to* $d^e_{\mathcal{C}}$ *if a minimum edge deletion set to* $\mathcal{C}$ *is given in the input.*

Proposition 18 directly implies that COLORED CLUSTERING is fixed-parameter tractable with respect to feedback edge number (CC is polynomial-time solvable on forests) and with respect to the number of edges that do not have one of the two most frequent colors (CC is polynomial-time solvable for $|C| = 2$). For vertex deletion distance, a similar statement is not true, since the problem is W[1]-hard with respect to vertex cover number as we will show in Theorem 20. It is also easy to prove that COLORED CLUSTERING is FPT with respect to the joint parameterization by the treewidth of the input graph and the maximum number of colors incident to any vertex. This can be shown by a standard dynamic program on the tree decomposition by iterating over all colorings of each bag. Since the maximum number of colors incident to any vertex is at most the maximum degree, it follows that CC is also FPT with respect to treewidth plus maximum degree.

We will now consider the parameter parameter *slim tree-cut width*, which is smaller than feedback edge number and treewidth plus maximum degree. We show that COLORED CLUSTERING is fixed-parameter tractable with respect to slim tree-cut width, by considering the asymptotically equivalent parameter *super edge-cut width*. Let $G = (V, E)$ be a graph and $T = (V, E')$ a tree on the same vertex set. For any $v \in V$, the *local feedback edge number* at v is

$$\text{lfe}(G, T, v) := |\{\{u, w\} \in E \setminus E' \mid v \text{ is on the unique } u\text{-}w\text{-path in } T\}|,$$

that is, $\text{lfe}(G, T, v)$ counts the number of edges that are not in T such that the unique path in T which connects the endpoints visits v. Note that possibly $u = v$ or $w = v$. The *local feedback edge number* of (G, T) is $\text{lfe}(G, T) := \max_{v \in V} \text{lfe}(G, T, v)$. The *super edge-cut width* (Ganian and Korchemna 2022) of G is

$$\text{secw}(G) := 1 + \min_{T \text{ is a tree on } V} \text{lfe}(G, T).$$

Ganian and Korchemna (2022) showed that there is an algorithm with running time $\mathcal{O}(f(k) \cdot |G|^{\mathcal{O}(1)})$ that given $k \in \mathbb{N}$ and a graph $G = (V, E)$ either outputs a tree T on V such that $\text{lfe}(G, T) \leq \mathcal{O}(k^6)$ or correctly determines that $\text{secw}(G) > k$. We will not formally define slim tree-cut width, but since it is asymptotically equivalent to super edge-cut width, it is sufficient to consider the latter. COLORED CLUSTERING is fixed-parameter tractable with respect to super edge-cut width by a bottom-up dynamic programming algorithm on T.

Theorem 19 (★). COLORED CLUSTERING *is fixed-parameter tractable with respect to super edge-cut width and slim tree-cut width.*

In contrast to local feedback edge number, COLORED CLUSTERING is W[1]-hard with respect to tree-cut width, which is a lower bound for local feedback edge number, and for vertex cover number, which is incomparable to local feedback edge number. These two hardness results, in particular the one for vertex cover number, rule out FPT algorithms for many other structural graph parameters, including treewidth, treedepth, etc.

The *vertex cover number* of a graph is the size of a smallest vertex cover. Tree-cut width is a parameter introduced fairly recently by Wollan (2015). The algorithmic uses of this parameter were systematically investigated by Ganian, Kim, and Szeider (2015). For the definition of tree-cut width we refer to the full version of our article and to their work.

Theorem 20 (★). COLORED CLUSTERING *on bipartite graphs is* W[1]*-hard with respect to both vertex cover number and tree-cut width.*

Cai and Leung (2018) showed that COLORED CLUSTERING is NP-hard, even on planar, bipartite graphs with the maximum degree at most four and only three colors. We conclude by strengthening this result and showing that COLORED CLUSTERING is also NP-hard on cubic graphs.

Theorem 21 (★). COLORED CLUSTERING *is NP-hard even if every vertex has degree at most three and* $|C| = 5$*.*

For hypergraphs, one can show NP-hardness in an even more restricted setting:

Theorem 22 (★). COLORED HYPERGRAPH CLUSTERING *is* NP*-hard even if every vertex has degree at most two, every edge contains at most three vertices, and* $|C| = 3$*.*

COLORED HYPERGRAPH CLUSTERING is polynomial-time solvable if the maximum degree is at most 1 or if $|C| \leq 2$. In the first case, the conflict graph is empty and, in the second case, it is bipartite.

Conclusion

Our results in many ways complete and extend the picture of the parameterized complexity of COLORED CLUSTERING initiated by Cai and Leung (2018). For the parameterization by the number of stable edges, we have given an improved algorithm and a polynomial kernel. We have also initiated the study of strictly smaller parameters than both the number of stable edges k and the number of unstable edges r. Finally, we gave a picture of the problem's parameterized complexity for structural graph parameters. We conclude by listing a few open problems and avenues for further research:

- Can the kernel (Theorem 1) be improved to size $\mathcal{O}(k^2)$?
- Are there other natural (tight) lower bounds for r or k, besides the degree-based and matching based lower bounds we considered, that yield fixed-parameter algorithms?
- One issue with the structural parameters we have considered is that they are oblivious to the complexity introduced by the edge colors. It may be useful to consider structural parameters that explicitly take the structure of the coloring into account. Such parameters were studied by Morawietz et al. (2022). Unfortunately, most of those parameters are smaller than the number of colors and, therefore, of little use in the context of COLORED CLUSTERING. However, Proposition 18 implies that CC is FPT with respect to $m_{>2}$ (in Morawietz et al.'s terminology), and the NP-hardness of CC on tricolored graphs implies para-NP-hardness with respect to $m_{>3}$. It may be an interesting challenge to develop color-sensitive structural parameters that lead to useful FPT algorithms for COLORED CLUSTERING.

Acknowledgments

The second author is supported by the DFG Project DiPa, NI 369/21. The third author is supported by the DFG Research Training Group 2434 "Facets of Complexity".

We thank Till Fluschnik (TU Clausthal) and Klaus Heeger (TU Berlin) for a fruitful discussion which led to the NP-hardness for cubic graphs (Theorem 21).

This paper is dedicated to Rolf, our co-author, colleague, and advisor. Rolf's tremendous contributions to computer science, particularly to parameterized algorithmics, will be dearly missed. The computer science community will build on the foundations he has laid.

References

Agarwal, S.; Branson, K.; and Belongie, S. J. 2006. Higher order learning with graphs. In Cohen, W. W.; and Moore, A. W., eds., *Proceedings of the 23rd International Conference on Machine Learning (ICML)*, volume 148, 17–24. ACM.

Ageev, A.; and Kononov, A. 2014. Improved Approximations for the Max k-Colored Clustering Problem. In *Proceedings of the 12th International Workshop on Approximation and Online Algorithms (WAOA)*, 1–10. Springer.

Agrawal, A.; Kundu, M.; Sahu, A.; Saurabh, S.; and Tale, P. 2022. Parameterized Complexity of Maximum Edge Colorable Subgraph. *Algorithmica*, 84: 3075–3100.

Alhamdan, Y. M.; and Kononov, A. 2019. Approximability and inapproximability for maximum k-edge-colored clustering problem. In *Proceedings of the 14th International Computer Science Symposium in Russia (CSR)*, 1–12. Springer.

Amburg, I.; Veldt, N.; and Benson, A. R. 2020. Clustering in graphs and hypergraphs with categorical edge labels. In *Proceedings of the Web Conference 2020 (WWW)*, 706–717. ACM / IW3C2.

Angel, E.; Bampis, E.; Kononov, A.; Paparas, D.; Pountourakis, E.; and Zissimopoulos, V. 2016. Clustering on k-edge-colored graphs. *Discrete Applied Mathematics*, 211: 15–22.

Bansal, N.; Blum, A.; and Chawla, S. 2004. Correlation Clustering. *Machine Learning*, 56(1-3): 89–113.

Bonchi, F.; Gionis, A.; Gullo, F.; Tsourakakis, C. E.; and Ukkonen, A. 2015. Chromatic Correlation Clustering. *ACM Transactions on Knowledge Discovery from Data*, 9(4): 34:1–34:24.

Cai, L.; and Leung, O. Y. 2018. Alternating Path and Coloured Clustering. arXiv:1807.10531.

Chen, J.; Molter, H.; Sorge, M.; and Suchý, O. 2018. Cluster Editing in Multi-Layer and Temporal Graphs. In *Proceedings of the 29th International Symposium on Algorithms and Computation (ISAAC)*, volume 123, 24:1–24:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

Crossley, N. A.; Mechelli, A.; Vértes, P. E.; Winton-Brown, T. T.; Patel, A. X.; Ginestet, C. E.; McGuire, P.; and Bullmore, E. T. 2013. Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, 110(28): 11583–11588.

Cygan, M.; Fomin, F. V.; Kowalik, L.; Lokshtanov, D.; Marx, D.; Pilipczuk, M.; Pilipczuk, M.; and Saurabh, S. 2015. *Parameterized Algorithms*. Springer.

Diestel, R. 2017. *Graph Theory*. Springer, 5th edition.

Downey, R. G.; and Fellows, M. R. 2013. *Fundamentals of Parameterized Complexity*. Springer.

Fukunaga, T. 2019. LP-based pivoting algorithm for higher-order correlation clustering. *Journal of Combinatorial Optimization*, 37(4): 1312–1326.

Ganian, R.; Kim, E. J.; and Szeider, S. 2015. Algorithmic Applications of Tree-Cut Width. In *Proceedings of the 40th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, 348–360. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

Ganian, R.; and Korchemna, V. 2022. Slim Tree-Cut Width. In *Proceedings of the 17th International Symposium on Parameterized and Exact Computation (IPEC)*, 15:1–15:18.

Gleich, D. F.; Veldt, N.; and Wirth, A. 2018. Correlation Clustering Generalized. In *Proceedings of the 29th International Symposium on Algorithms and Computation (ISAAC)*, volume 123, 44:1–44:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

Goyal, P.; Misra, N.; Panolan, F.; and Zehavi, M. 2015. Deterministic parameterized algorithms for matching and packing problems. *SIAM Journal on Discrete Mathematics*, 29(4): 1815–1836.

Gutin, G.; and Mnich, M. 2022. A Survey on Graph Problems Parameterized Above and Below Guaranteed Values. arXiv:2207.12278.

Harris, D. G.; and Narayanaswamy, N. S. 2022. A faster algorithm for Vertex Cover parameterized by solution size. arXiv:2205.08022.

Kellerhals, L.; Koana, T.; and Kunz, P. 2022. Vertex Cover and Feedback Vertex Set Above and Below Structural Guarantees. In *Proceedings of the 17th International Symposium on Parameterized and Exact Computation (IPEC)*, 19:1–19:14.

Li, P.; Puleo, G. J.; and Milenkovic, O. 2020. Motif and Hypergraph Correlation Clustering. *IEEE Transactions on Information Theory*, 66(5): 3065–3078.

Lokshtanov, D.; Narayanaswamy, N. S.; Raman, V.; Ramanujan, M. S.; and Saurabh, S. 2014. Faster parameterized algorithms using linear programming. *ACM Transactions on Algorithms*, 11(2): 1–31.

Mahajan, M.; and Raman, V. 1999. Parameterizing above Guaranteed Values: MaxSat and MaxCut. *Journal of Algorithms*, 31(2): 335–354.

Morawietz, N.; Grüttemeier, N.; Komusiewicz, C.; and Sommer, F. 2022. Refined Parameterizations for Computing Colored Cuts in Edge-Colored Graphs. *Theory of Computing Systems*, 66(5): 1019–1045.

Mucha, P. J.; Richardson, T.; Macon, K.; Porter, M. A.; and Onnela, J.-P. 2010. Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science*, 328(5980): 876–878.

Papa, D. A.; and Markov, I. L. 2007. Hypergraph Partitioning and Clustering. In Gonzalez, T. F., ed., *Handbook of Approximation Algorithms and Metaheuristics*. Chapman and Hall/CRC.

Razgon, I.; and O'Sullivan, B. 2009. Almost 2-SAT is fixed-parameter tractable. *Journal of Computer and System Sciences*, 75(8): 435–450.

Shachnai, H.; and Zehavi, M. 2017. Parameterized Algorithms for Graph Partitioning Problems. *Theory of Computing Systems*, 61: 721–738.

Stockmeyer, L. J.; and Vazirani, V. V. 1982. NP-completeness of some generalizations of the maximum matching problem. *Information Processing Letters*, 15(1): 14–19.

Veldt, N. 2022. Optimal LP Rounding and Fast Combinatorial Algorithms for Clustering Edge-Colored Hypergraphs. arXiv:2208.06506.

Wollan, P. 2015. The structure of graphs not admitting a fixed immersion. *Journal of Combinatorial Theory Series B*, 110: 47–66.

Wu, B. Y. 2012. On the maximum disjoint paths problem on edge-colored graphs. *Discrete Optimization*, 9(1): 50–57.