On Parameterized Inapproximability of Several Optimization Problems
（幾つかの最適化問題のパラメータ化近似不可能性）

by

Lin Bingkai
林 冰凱

A Doctor Thesis
博士論文

Submitted to
the Graduate School of the University of Tokyo
on December 11, 2015
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Information Science and Technology
in Computer Science

Thesis Supervisor: Hiroshi Imai　今井 浩
Professor of Computer Science

ABSTRACT

In this thesis, we study the parameterized inapproximability of several optimization problems. Approximation algorithms and parameterized complexity are two powerful methods to deal with hard computational problems. Approximation algorithm finds a solution that is close to the optimum in polynomial time. While in the area of parameterized complexity we consider problems with an additional parameter k and design algorithms that output the exact solution in $f(k) \cdot |x|^{O(1)}$-time on input an instance x. There exist problems that do not admit polynomial time approximation algorithms within certain ratios if NP $\neq$ P. The celebrated PCP theorem is the main technique to prove hardness results of polynomial time approximation. However, it does not rule out the existence of parameterized approximation algorithms. For example, an important open question is whether there is an $f(k) \cdot |G|^{O(1)}$-time algorithm that can find a dominating set with $2k$ vertices for G given that G has a dominating set with k vertices. One of the main contributions of this thesis is to refute the existence of such an algorithm under FPT $\neq$ W[1], a stardard hypothesis in parameterized complexity theory. Our starting point is the parameterized inapproximability of Maximum k-Subset Intersection.

Given a collection $\mathcal{F} = \{S_1, S_2, \cdots, S_n\}$ of subsets over a finite set with m elements, the goal of Maximum k-Subset Intersection is to select k distinct subsets from $\mathcal{F}$ such that their intersection size is as large as possible. The decision version of this problem is related to the k-Biclique problem, which asks if an input graph contains a complete bipartite subgraph with k vertices in each side. A longstanding open question in parameterized complexity is whether there exist any $f(k) \cdot n^{O(1)}$-time algorithms for k-Biclique. This thesis also gives a negative answer to this question assuming FPT $\neq$ W[1].

The core result in this thesis is to provide a gap-producing fpt-reduction from k-Clique to Maximum k-Subset Intersection. More precisely, we construct a set family $\mathcal{F}$ on input a graph G and a small positive integer k in polynomial time such that if G contains a subgraph isomorphic to K_k then there exist $s := \binom{k}{2}$ sets in $\mathcal{F}$ with intersection size no less than $n^{\Theta(1/k)}$, otherwise every s sets in $\mathcal{F}$ have intersection size at most $O(k!)$.

Then we derive the parameterized inapproximability of dominating set problem based on the hardness approximation result of Maximum k-Subset Intersection. Significantly, our hardness approximation result does not rely on the PCP theorem.

Finally, we consider the problem of finding the maximum clique whose edges use at most k distinct colors on input a multigraph with colored edges, which we call Maximum k-Edge-Color Clique. We show that if the input graph has unbounded number of multi-edges between its every two vertices, then this problem does not admit fpt-approximation algorithms within any computable ratio function $\rho(k)$ assuming FPT $\neq$ W[1]. We also point out that the fpt-inapproximability of Maximum k-Edge-Color Clique on simple graphs is related to the fpt-inapproximability of k-Clique.

Acknowledgements

First of all, I am grateful to my supervisor Hiroshi Imai for his support and advices during my PhD program. I would like to thank Yijia Chen, the coauthor of the paper that make up one part of this thesis, for many valuable communications. I also want to thank Benjamin Rossman for bringing the probabilistic method into my attention and the enlightened discussions in NII.

I would like to thank the organizers and lecturers of the school on parameterized algorithms and complexity held in August 2014, at Conference Center in Bedlewo, Poland. The key idea of the gap reduction for MIN-DOMINATING-SET was inspired by Dániel Marx's slide for the lecture of W[1]-hardness in this summer school. More precisely, the idea occurred to me when I was reviewing the parameterized reduction from Multicolored Independent Set to Dominating Set in page 8 of this slide in April 2015.

I am very honored to have Prof. Reiji Suda as the chair, and Prof. Naoki Kobayashi, Prof. Seiya Imoto, Prof. Masami Hagiya, Prof. Masashi Sugiyama, Prof. Kunihiko Sadakane, Prof. Yijia Chen as the members of the committee for this thesis. I am grateful to them for taking the time to examine this thesis.

Thanks all the members of Imai Lab and the JST ERATO Kawarabayashi Large Graph Project.

Finally, I would like to thank the love and support of my family.

Scientists want to show that
things that don't look alike are really the same.
That is one of their innermost Freudian motivations.
In fact, that is what we mean by understanding.

Gian-Carlo Rota

Contents

1 Introduction **1**
- 1.1 Parameterized Inapproximability . . . 2
- 1.2 Our Contributions . . . 4
  - 1.2.1 W[1]-hardness of k-Biclique . . . 5
- 1.3 Organization of this Thesis . . . 8

2 Preliminaries **9**
- 2.1 Basic Notations . . . 9
  - 2.1.1 Sets, Numbers and Functions . . . 9
  - 2.1.2 Graphs . . . 10
- 2.2 Computational Complexity . . . 11
  - 2.2.1 Complexity Classes and Reductions . . . 11
  - 2.2.2 Hypothesis . . . 12
- 2.3 Optimization Problems . . . 12
  - 2.3.1 Parameterized Approximation . . . 13
  - 2.3.2 Gap-producing fpt-reductions . . . 14
- 2.4 Probability . . . 15
- 2.5 Finite Field . . . 16
- 2.6 Color-Coding . . . 17
  - 2.6.1 (n, k)-Universal Sets . . . 18

3 (n, k, ℓ, h)-Threshold Property **20**
- 3.1 Definition . . . 20
- 3.2 Probabilistic Construction . . . 22
  - 3.2.1 Estimate $\Pr[G(n, p)$ does not satisfy (T1)$]$ . . . 22
  - 3.2.2 Estimate $\Pr[G(n, p)$ does not satisfy (T2)$]$ . . . 22
  - 3.2.3 Proof of Lemma 3.1.3 . . . 25
- 3.3 Explicit Construction . . . 25
  - 3.3.1 Using Paley-Type Graphs . . . 25
  - 3.3.2 Proof of the Intersection Lemma . . . 27
- 3.4 Some Extremal Questions . . . 30

4 **Maximum k-Subset Intersection** **32**
4.1 Introduction . . . 32
4.2 The Gap-producing Reduction . . . 34
4.2.1 Lower Bounds under ETH . . . 37
4.3 Open Questions . . . 38

5 **Minimum Dominating Set** **41**
5.1 Introduction . . . 41
5.1.1 Our Results . . . 42
5.1.2 Overview of Our Reduction . . . 44
5.2 The Case $\rho < 3/2$ . . . 45
5.3 The Constant-Inapproximbility of k-Dominating-Set . . . 49
5.3.1 Proof of Theorem 5.3.1 . . . 49
5.3.2 Soundness . . . 52
5.4 Some Consequences . . . 56
5.5 Conclusions . . . 57

6 **Maximum k-Edge-Color Clique** **58**
6.1 Introduction . . . 58
6.1.1 Main Results . . . 60
6.2 Inapproximability of Max-k-EdgeCol-Clique . . . 61
6.3 From Max-k-EdgeCol-Clique to k-Clique . . . 63
6.3.1 The Construction. . . . 64
6.3.2 Completeness. . . . 64
6.3.3 Soundness . . . 65
6.4 Multi-graphs with Few Multi-edges . . . 68

7 **Conclusions and Future Work** **69**
7.1 Questions on Parameterized Inapproximability . . . 69

References **72**

List of Figures

4.1 Reduction from Finding K_3 to Max 3-Subset Intersection 34
4.2 The Reduction of Theorem 4.2.1 35
4.3 A 3×4 grid. 39

5.1 Overview of the Reduction 44
5.2 2-gap fpt-reduction for Min Weighted Dominating Set 45

6.1 Max-2-Edge-Color Clique 60

Chapter 1

Introduction

One theme of computational complexity theory is to distinguish between computational problems that are efficient solvable and those that are intractable, under some well-accepted hypothesis. In classical computational complexity theory, a problem is considered to be tractable if it can be solved by algorithms that are guaranteed to terminate in a number of steps bounded by a polynomial in the length of its input. The theory of NP-completeness and the polynomial-time reduction [Coo71, Kar72, Lev73] allow researchers to classify almost all the computational problems into P and NP-hard. Under the hypothesis that P $\neq$ NP, those NP-hard problems can not be solved in polynomial time. Such a classification could be further refined if we define "*efficiently solvable*" in a more subtle way.

In parameterized complexity theory [DF99, FG06, Nie06, DF13, CFK⁺16], we consider problems with a parameter k. For an input with length n and small parameter k, algorithms with running-time in $2^k \cdot n^{O(1)}$ can still be considered efficient. In general, we say a problem is *fixed-parameter tractable* or in FPT if it can be solved in $f(k) \cdot |x|^{O(1)}$-time on input x with parameter k for some computable function f. The theory of parameterized complexity leads to a refined analysis of the complexity class. Among those problems that are NP-hard, some are showed to be in FPT, e.g. k-Vertex-Cover. Some turn out to be W[1]-hard (analogous to NP-hard), e.g. k-Clique. The hypothesis FPT $\neq$ W[1] implies that every W[1]-hard problem has no fpt-algorithm.

Many computational problems have their corresponding optimization versions. In an optimization problem, the goal is to find a feasible solution with maximum or minimum value of cost. Another natural relaxation is to require the algorithm to produce a solution that is close to the optimum one. For minimization (maximization, respectively) problems, a c-approximation algorithm always returns a solution whose cost is at most c (at least $1/c$, respectively) times the optimum. There exist optimization problems with some hardness factors such that no polynomial time algorithms can achieve the approximate ratios below these factors unless P = NP. For example, the hardness factor for Min-Vertex-Cover is $O(1)$ [Hås01, DS02], and $O(n^{\varepsilon})$ for Max-Clique [FGL⁺96, AS98, ALM⁺98, Hås96, Kho01, Zuc06], $O(\log n)$ for Min-Dominating-Set [LY94, RS97, Fei98, AMS06, DS14]. To prove such hardness approximation results, we need stronger reductions which create gaps in the optimal cost values. The celebrated PCP-theorem [AS98, ALM⁺98, Din07] makes such gap-producing reductions possible. See [Tre04] for a survey.

Theorem 1.0.1 (PCP Theorem: Hardness of approximation view)**.** *For a 3SAT instance φ, let $\mathrm{OPT}(\varphi)$ be the maximum fraction of satisfied clauses in φ. There exist $\rho < 1$ and a polynomial time reduction that on input a 3SAT instance φ of size n, it produces a 3SAT instance η, with size $|\eta| = n \cdot (\log n)^{O(1)}$ such that*

$$\mathrm{OPT}(\varphi) = 1 \Rightarrow \mathrm{OPT}(\eta) = 1 \tag{1.1}$$

$$\mathrm{OPT}(\varphi) < 1 \Rightarrow \mathrm{OPT}(\eta) < \rho \tag{1.2}$$

Of course, there is no reason to restrict approximation algorithms to those running in polynomial-time. For optimization problems that are hard to approximate under $\mathrm{P} \neq \mathrm{NP}$, we want to know if it is possible to achieve better approximation ratios using algorithms with moderate exponential running time. This thesis aims for proving hardness approximation results in a parameterized setting.

For instance, let us consider MIN-DOMINATING-SET. Let $\gamma(G)$ denote the size of minimum dominating set of G. One of the main problems that we address is

Question 1.0.2. *Is there any algorithm $\mathbb{A}$ such that on input a graph G, $\mathbb{A}$ outputs a dominating set of G with size $c \cdot \gamma(G)$ in time $f(\gamma(G)) \cdot |G|^{O(1)}$ for some computable function f and constant c?*

To prove an optimization problem does not have fpt-approximation algorithms, we should use a reduction that is not only gap-producing but also fpt. For example, suppose we want to show that MIN-DOMINATING-SET has no 2-approximation fpt-algorithms unless $\mathrm{FPT} = \mathrm{W}[1]$. It suffices to find an algorithm $\mathbb{A}$ such that for every graph G and $k \in \mathbb{N}$, $\mathbb{A}(G, k)$ outputs a graph G' and k' satisfying:

- (1) if G contains a k-clique, then $\gamma(G') \leq k'$.
- (2) if G contains no k-clique, then $\gamma(G') > 2k'$.
- (3) $k' \leq g(k)$ for some computable function g.
- (4) $\mathbb{A}(G, k)$ is computable in time $f(k) \cdot |G|^{O(1)}$.

Conditions (1) and (2) require the reduction to produce a gap larger than two. Conditions (3) and (4) ensure that the reduction is fpt. The previous works based on the PCP theorem always cause the value of k' to be polynomial in $|G'|$. Hence they do not satisfy Condition (3). It seems that we need new techniques to answer Question 1.0.2.

1.1 Parameterized Inapproximability

To see what kind of techniques are helpful in proving parameterized inapproximability. We review the previous work in this direction.

Let us start with results exploiting the *non-monotonicity* of the problems. A minimization problem is called *monotone* if the superset of a solution is also a solution. Similarly, a maximization problem is *anti-monotone* if a subset of its solution is also a solution. For example, a superset of a dominating set for a graph is still a dominating set. However, with an additional condition that requires the dominating set to be independent, the problem of finding minimum *independent dominating set* is no longer monotone. In [DFMR08], Downey et al. showed that the *independent dominating set* problem has no fpt approximation with any approximation ratio. Another example where parameterized inapproximability is showed using the non-monotonicity is the weighted satisfiability problem of CNF-formulas. In [CGG07] it is shown that the weighted satisfiability problem of CNF-formulas have no fpt approximation of any possible ratio.

Inapproximability results based on non-monotonicity are somewhat unsatisfying. As Marx pointed out in [Mar08], "it can happen that the optimum is k, and every feasible solution has cost k, which makes approximation equivalent to finding an optimum solution. The inapproximability proofs in these examples tell us more about the hardness of finding exact solutions than about the hardness of approximation". What is more, many natural minimization problems, e.g. Min-Dominating-Set, are monotone.

So let us turn our attention to monotone problems. In [AR01], Alekhnovich and Razborov showed that there is no fpt 2-approximation algorithm for Weighted Monotone Circuit Satisfiability, unless the class W[P] from the hierarchy of parameterized problems is fixed-parameter tractable by randomized algorithms with one-sided error. In [EGG08], Eickmeyer et al. derandomized Alekhnovich and Razborov's inapproximability result and at the same time strengthened it. In fact, they proved that the weighted monotone circuit satisfiability problem has no fixed-parameter tractable approximation algorithm with polylogarithmic approximation ratio unless FPT = W[P]. Finally, Marx strengthened this result significantly in [Mar13] by proving that the weighted satisfiability problem is not fpt approximable for circuits of depth 4 without negation gates, unless FPT = W[2]. Recently, Cristina Bazgan et al. [BCNS14] gave a reduce from Monotone Circuit Satisfiability to Target Set Selection in polynomial time. They showed that for any functions f and ρ, the Target Set Selection problem cannot be approximated within a factor of $\rho(k)$ in $f(k) \cdot n^{O(1)}$ time, unless FPT = W[P]. Let us not fail to mention that Marx's result relies on the *k-perfect family of hashing functions* in [AYZ95]. This technique also plays an important role in many reductions of this thesis.

Another approach to parameterized inapproximability is to work on stronger hypothesis. The Linear PCP conjecture claims that it is possible to achieve $|\eta| = O(n)$ in Theorem 1.0.1. Assuming the Linear PCP conjecture and ETH, Edouard Bonnet et al. [BEKP13], showed that there is no r-approximation algorithm for Min-Dominating-Set running in time $f(k) \cdot n^{o(k)}$ for some $r > 1$. In the same paper, they also proved that independent set has no constant approximation fpt-algorithm, again under the Linear PCP conjecture and ETH. In [CHK13, HKK13] it is proved that assuming ETH there is no $c\sqrt{\log \gamma(G)}$-approximation algorithm for the set cover problem, with running time $2^{O(\gamma(G)^{(\log \gamma(G))^d})}|G|^{O(1)}$, where c and d are some appropriate constants. With the additional

Projection Game Conjecture due to [Mos12] and some of its further strengthening, the authors of [CHK13, HKK13] are able to even rule out $\gamma(G)^c$-approximation algorithms with running time almost doubly exponential in terms of $\gamma(G)$. In addition, they also prove that unless NP $\subseteq$ SUBEXP, for every $1 > \delta > 0$, there exists a constant $F(\delta) > 0$ such that Clique has no approximation with ratio $k^{1-\delta}$ in $2^{k^F} \cdot n^{O(1)}$-time. The drawback of the results in [CHK13, HKK13] is that the dependence of the running time on parameters is not an arbitrary computable function. Under the assumption $\mathrm{ETH_{nu}}$, Yijia Chen, Kord Eickmeyer and Jörg Flum [CEF12] showed that for every $d \in \mathbb{N}$ there is a $\rho > 1$ such that k-Clique has no parameterized approximation algorithm with approximation ratio ρ and running time $f(k) \cdot n^d$ for some function $f : \mathbb{N} \to \mathbb{N}$. Assuming Deg-2-sat is not fixed parameter tractable, Subhash Khot and Igor Shinkar [KS16] proved that k-Clique can not be fpt-approximated to any constant ratio.

1.2 Our Contributions

The main contributions of the thesis is to present gap-producing fpt-reductions from k-Clique to several parameterized optimization problems, hence proving the parameterized inapproximability of these problems under FPT $\neq$ W[1]. The results of this thesis are from two papers [Lin15, CL15] and some unpublished observations.

– **Parameterized Inapproximability of Max-k-Subset-Intersection.** Assuming FPT $\neq$ W[1], there is no fpt-algorithm to distinguish whether an input set family contains k sets with intersection size no less than $n^{\Theta(1/\sqrt{k})}$ or every k sets of this family has intersection size at most $O(k!)$ (Theorem 4.2.1). Assuming ETH, for every constant $c \geq 2$, there is no $f(k) \cdot n^{o(\sqrt[c]{k})}$-time algorithm to approximate Max-k-Subset-Intersection to ratio below $n^{\Theta(1/\sqrt[c]{k})}$ (Corollary 4.2.8).

– **Parameterized Inapproximability of Min-Dominating-Set.** Assuming FPT $\neq$ W[1], for any constant $c \geq 1$, there is no fpt-algorithm to distinguish whether an input graph can be dominated by k vertices or every dominating set of it contains at least ck vertices (Theorem 5.1.1). Assuming ETH, there is *no* fpt-algorithm which on every input graph G outputs a dominating set of size at most $\sqrt[4+\varepsilon]{\log{(\gamma(G))}} \cdot \gamma(G)$ for every $0 < \varepsilon < 1$ (Theorem 5.1.3).

– **Parameterized Inapproximability of Max-k-EdgeCol-Clique and Its Relation to k-Clique** Unless FPT = W[1], there is no fpt-algorithm to find a k-edge-color clique with size no less than $1/\rho(k)$ times the maximum ones on input a multi-graph with colored edges (Theorem 6.1.2). Parameterized inapproximability of Max-k-EdgeCol-Clique on simple graphs implies k-Clique does not admit constant fpt-approximation algorithms (Theorem 6.1.3). Under some conditions, it is possible to turn a multi-graph into a simple graph without changing the size of its maximum k-edge-color clique (Theorem 6.1.5).

Main Technique Our main technical of showing inapproximability is to use the gap of set intersections. To create a gap for set intersections, we introduce a definition of threshold property and provide two constructions of graphs satisfying such property in polynomial time. Intuitively, for $h > \ell$, a bipartite graph $T = (A \dot\cup B, E)$ satisfying the (n, k, ℓ, h)-threshold property if every $k+1$ vertices from A have at most ℓ common neighbors while there are many k vertices from A with h common neighbors. Such a gap between the cardinality of common neighbors of k-vertex sets and $(k+1)$-vertex sets allows us to construct gap-producing fpt-reductions.

- **Probabilistic Construction** We can construct bipartite random graphs satisfying the (n, k, ℓ, h)-threshold property with probability at least $\frac{9}{10}$ in polynomial time for $\ell = O(k^2) < h < n^{\Theta(1/k)}$. (Lemma 3.1.3)
- **Explicit Construction.** We can construct bipartite graphs satisfying the (n, k, ℓ, h)-threshold property in polynomial time for $\ell = O(k!) < h < n^{\Theta(1/k)}$. (Lemma 3.1.4)

1.2.1 W[1]-hardness of k-Biclique

As we have mentioned, the decision version of Max-k-Subset-Intersection is related to the problem of testing whether an input graph contains a subgraph (not necessarily induced) isomorphic to a complete bipartite graph $K_{k,k}$.

k-Biclique	
Instance:	$k \in \mathbb{N}^+$ and a graph G.
Parameter:	k.
Problem:	Decide if G contains a subgraph isomorphic to $K_{k,k}$.

It is not difficult to see that k-Biclique restricted to bipartite graphs can be interpreted as finding k distinct vertices from one side of the bipartite graph with at lest k vertices in their common neighbors. The k-Biclique problem is NP-hard [Joh87]. Furthermore, Max-Biclique, which is the optimization version of k-Biclique, does not admit polynomial time approximation algorithms to ratio n^{ε} if SAT does not have a $2^{n^{\varepsilon'}}$-time probabilistic algorithm [FK04, Kho06, AMS11].

Max-Biclique	
Instance:	$k \in \mathbb{N}$ and a graph G.
Solution:	a complete bipartite subgraph of G that isomorphic to $K_{k,k}$.
Cost:	k.
Goal:	max.

Whether there exist $f(k) \cdot n^{O(1)}$-time algorithms solving k-Biclique for any computable function f has received heavy attention from the parameterized complexity community

[ALR12, BM11, FG06, Gro07, HKM13]. It is the first problem on the "most infamous" list (page 677) in a new text book [DF13].

> "Almost everyone considers that this problem should obviously be W[1]-hard, and... it is rather an embarrassment to the field that the question remains open after all these years!".

Despite many attempts [BRFGL10, CK12, GKL12, Kut12], no fpt-reduction from k-Clique to k-Biclique has previously been found. Another contribution of this thesis is to give a negative answer to the question above by proving the W[1]-hardness of k-Biclique (Corollary 4.2.6). As a consequence, we obtain a dichotomy theorem of cardinality constraints satisfaction problem.

Dichotomy for Cardinality Constraint Satisfaction Problem

Fix a domain $D = \{0, 1, 2, \dots, d\}$. An instance of the constraint satisfaction problem (CSP) is a pair $I = (V, C)$, where V is a set of variables and C is a set of constraints. Each constraint of C can be written as $\langle \mathbf{v}, R \rangle$, where R is an r-ary relation on D for some positive integer r and $\mathbf{v} = (v_1, v_2, \dots, v_r)$, an assignment $\tau : V \to D$ satisfies a constraint $\langle \mathbf{v}, R \rangle$ if $(\tau(v_1), \dots, \tau(v_r)) \in R$. The goal is to find an assignment $\tau : V \to D$ satisfying all the constraints in C. In the research of complexity of CSP, we usually fix a set of relation Γ, and denote CSP(Γ) the CSP problem in which all the relations of the constraints are in Γ.

It is well-known that many hard problems including satisfiability and graph coloring can be expressed under the CSP framework, hence solving constraint satisfaction problems is NP-hard. One way to cope with this NP-hard problem is to introduce a parameter and consider the parameterized version of such problem. In [BM11], Andrei A. Bulatov and Dániel Marx introduced two parameterized versions of CSP. More specifically, they assume that the domain contains a "free" value 0 and other non-zero values, which are "expensive". The goal is find an assignment with a limited number of variables assigning expensive values. One way to reflect this goal is to take the number of nonzero values used in an assignment as parameter, which leads to the definition of the CSP with size constraints (OCSP); another more refined way is to prescribe how many variables have to be assigned each particular nonzero value, this leads to the definition of CSP with cardinality constraints (CCSP). They provide a complete characterization of the fixed-parameter tractable cases of OCSP(Γ) and show that all the remaining problems are W[1]-hard.

OCSP(Γ)

Instance:	A CSP(Γ) instance $I = (V, C)$.
Parameter:	k.
Problem:	find a satisfying assignment for I with k variables assigned nonzero.

CCSP(Γ)		
	Instance:	A CSP(Γ) instance $I = (V, C)$ with an additional cardinality constraint $\pi : D \setminus \{0\} \to \mathbb{N}$.
	Parameter:	$k = \sum_{i \in D\setminus\{0\}} \pi(i)$.
	Problem:	find a satisfying assignment for I with $\pi(i)$ variables assigned i ($\forall i \in D \setminus \{0\}$).

For CSP with cardinality constraints, the situation is strange. A simple observation shows that the k-Biclique problem can be expressed as a CCSP instance. Without loss of generality, consider the k-Biclique on bipartite graph, let $D = \{0, 1, 2\}$, for any bipartite graph G, we construct a CCSP instance with $V = V(G)$ and $C = \{\langle(v_1, v_2), R\rangle \mid v_1v_2 \in E(G), R = \{(0, 0), (1, 0), (0, 2)\}\}$, then we ask for an assignment $\tau : V \to D$ with k variables assigning 1 and k variables assigning 2. It is easy to check that for a bipartite graph G, if and only if the corresponding CCSP instance has such an assignment, then the bipartite complement $\bar{G}$ of G contains a $K_{k,k}$. Therefore, without settling the parameterized complexity of k-Biclique, they can only show that CCSP(Γ) is fixed-parameter tractable, Biclique-hard or W[1]-hard. Combining our result and Theorem 1.2 in [BM11], we finally obtain a dichotomy theorem for the parameterized complexity of CCSP(Γ):

Theorem 1.2.1. *For every finite* Γ *closed under substitution of constants,* CCSP(Γ) *is either* FPT *or* W[1]*-hard.*

Towards a Dichotomy for Subgraph Embedding

Let $\mathcal{C}$ be a class of graphs. k-Biclique can be formulated as a special case of subgraph embedding defined as follows.

p-Emb($\mathcal{C}$)		
	Instance:	G is a graph and $H \in \mathcal{C}$.
	Parameter:	$\|H\|$.
	Problem:	Decide whether H is a subgraph of G.

To give a complete characterization of the fixed-parameter tractable cases of p-Emb($\mathcal{C}$) is a big open question.

It is well known that whether H is a subgraph of G can be decided in $f(|H|) \cdot |G|^{O(tw(H))}$ time using the color-coding technique in [AYZ95], where $tw(H)$ denotes the tree-width of H and f is a computable function. Hence, if $\mathcal{C}$ is a class of graphs with tree-width bounded by some constant, the subgraph isomorphism problem with $H \in \mathcal{C}$ is fixed parameter tractable. In [Gro07], Martin Grohe conjectured that p-Emb is W[1]-hard if and only if $\mathcal{C}$ has unbounded tree-width. Under the assumption of FPT $\neq$ W[1], this implies that there is no $f(k) \cdot |G|^{O(1)}$-time algorithm to decide whether G contains a subgraph isomorphic to $K_{k,k}$, because the class of balanced complete bipartite graphs $\{K_{k,k} \mid k \in \mathbb{N}\}$ has unbounded tree-width.

A possible approach is to consider the Partitioned Subgraph Isomorphism problem, in which each vertex of the smaller graph G has a distinct color and the vertices of H are partitioned into $|V(G)|$ subsets, each set is corresponding to one color. The problem is to find an injective mapping φ from $V(G)$ to $V(H)$ such that: (1) for all $u \in V(G)$, u and $\varphi(u)$ have the same color; (2) if u and v are adjacent in G, then $\varphi(u)$ and $\varphi(v)$ are adjacent in H. It is already known that Partitioned Subgraph Isomorphism problem on the graph class $\mathcal{C}$ is W[1]-hard if $\mathcal{C}$ has unbounded tree-width [Gro07]. An interesting fact is that if the graph G has no homomorphism to any of its proper induced subgraphs, then the colored and uncolored version of Subgraph Isomorphism of G are equivalent [Mar07]. Unfortunately, this approach does not work for k-Biclique because any bipartite graph has a homomorphism to any of its edges.

In summary, proving the W[1]-hardness of k-Biclique is one step towards a dichotomy for Subgraph Embedding. It remains to see if we can extend our technique to more general graph classes. On the other hand, in [LRR14], Yuan Li, Alexander Razborov and Benjamin Rossman introduced a parameter κ. They showed that the average-case AC^0 complexity for deciding whether G contains a subgraph H is $|G|^{\Theta(\kappa(H))}$. Surprisingly, there exist graph H whose tree-width is k but $\kappa(H) = O(1)$ (See Remark 2.7 of [LRR14]). An interesting question is to give an $n^{O(1)}$-time algorithm to test whether $H \subseteq G$ for some bipartite H with $\kappa(H) = O(1)$ and $tw(H) = k$.

1.3 Organization of this Thesis

After introducing some preliminaries in Chapter 2, we start with the definition of threshold property and two efficient constructions in Chapter 3. In Chapter 4, we prove the parameterized inapproximability result of Max-k-Subset-Intersection. Using this result, we derive the constant inapproximability of parameterized dominating set problem in Chapter 5. The content of Chapter 6 is about a new optimization problem which arises in our research on the inapproximability of the parameterized clique problem. Chapter 7 contains open questions for future work.

Chapter 2

Preliminaries

We first introduce some basic notations. Then we review the necessary background concerning computational complexity and optimization algorithms. Finally, we introduce the mathematical tools that are useful for the rest of the thesis.

2.1 Basic Notations

2.1.1 Sets, Numbers and Functions

Sets $\emptyset$ denotes the *empty set*. For two sets A, B and $k \in \mathbb{N}^+$, we define the following notations.

- $A \times B := \{(a, b) \mid a \in A, b \in B\}$.
- $A \dot\cup B$ is the union of A and B where A and B are two disjoint sets.
- $|A|$ is the cardinality of A.
- $A^k := \{(a_1, a_2, \ldots, a_k) \mid \forall i \in [k], a_i \in A\}$.
- $\binom{A}{k} := \{\{a_1, a_2, \ldots, a_k\} \mid \forall i \in [k], a_i \in A, \forall i \neq j \in [k], a_i \neq a_j\}$.

In other words, $\binom{A}{k}$ is the set of all k-element subsets of A, A^k is the set of all k-tuple of A. As a consequence, if A contains n elements, then $|\binom{A}{k}| = \binom{n}{k}$, and $|A^k| = n^k$. For $\boldsymbol{v} = (v_1, v_2, \ldots, v_k) \in A^k$ and $i \in [k]$, $\boldsymbol{v}(i) := v_i$. $\{0,1\}^* := \bigcup_{n \in \mathbb{N}} \{0,1\}^n$ is the set of all strings over the alphabet $\{0,1\}$. For $x \in \{0,1\}^*$, $|x| := n$ if and only if $x \in \{0,1\}^n$.

Numbers $\mathbb{N}$ denotes the set of all natural numbers, i.e., $\mathbb{N} := \{0, 1, 2, \ldots\}$. $\mathbb{N}^+ := \mathbb{N} \setminus \{0\}$ is the set of all positive natural numbers. For each $n \in \mathbb{N}^+$, let $[n] := \{1, 2, \ldots, n\}$ be the finite set of integer between 1 and n. We use $\mathbb{R}$ and $\mathbb{C}$ to denote the sets of real numbers and complex numbers. For $a, b \in \mathbb{R}$, we use $a \pm b$ to denote the set of real numbers that are between a and b, i.e., $a \pm b := \{x \in \mathbb{R} \mid a \leq x \leq b\}$.

Functions For every function $f : A \to B$ and a subset $S \subseteq A$, $f(S) := \{f(x) \mid x \in A\}$. Similarly, for $\boldsymbol{v} = (v_1, v_2, \dots, v_k) \in A^k$, $f(\boldsymbol{v}) = (f(v_1), f(v_2), \dots, f(v_k))$.

Asymptotic Notation. Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions. We treat $O(g)$, $\Omega(g)$, $o(g)$ and $\Theta(g)$ as sets of functions. However, the statement $f \in O(g)$ is usually written as "$f = O(g)$".

(1) $f = O(g)$ if there exists some positive constant c such that $|f(x)| \leq c|g(x)|$;

(2) $f = \Omega(g)$ if $g \in O(f)$;

(3) $f = o(g)$ if for every positive constant ε there exists a constant N such that $f(n) \leq \varepsilon g(n)$ for all $n > N$;

(4) $f = \Theta(g)$ if there exists two positive constants c_1, c_2 such that $c_1 g(x) \leq f(x) \leq c_2 g(x)$.

In particular, $O(1)$ is the class of constants. $n^{O(1)}$ is the class of functions of n^c for any constant c.

2.1.2 Graphs

A *simple graph* $G := (V, E)$ contains a vertex set V and an edge set $E \subseteq \binom{V}{2}$. We also use $V(G)$ and $E(G)$ to denote the vertex set and edge set of G. Fix a graph G, for two vertices u, v in $V(G)$, we say u and v are adjacent if $\{u, v\} \in E(G)$.

For every $v \in V(G)$, $N^G(v) := \{u \mid u \in V(G), \{u, v\} \in E(G)\}$. Sometimes we simply omit the superscript G in $N^G(v)$ and write $N(v)$. We also abbreviate edges $\{u, v\}$ by uv. For $S \subseteq V(G)$, let $\Gamma(S) := \{u \mid \forall v \in S, u \in N^G(v)\}$ be the set of common neighbors of all vertices in S. Similarly, for $\boldsymbol{v} := (v_1, v_2, \dots, v_k) \in V(G)^k$, we also use $\Gamma(\boldsymbol{v})$ to denote the common neighbors of vertices in all coordinates of $\boldsymbol{v}$.

A graph H is a subgraph of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. For $X \subseteq V(G)$, $G[X] := (X, E(G) \cap \binom{X}{2})$ is the subgraph of G induced by X. Every two vertex u, v in X are adjacent in $G[X]$ if and only if there are adjacent in G.

A graph G is *isomorphic* to another graph H if there exists a one to one function $f : V(G) \to V(H)$ such that for all $u, v \in V(G)$, $uv \in E(G)$ if and only if $f(u)f(v) \in E(H)$. We say G contains a H-subgraph or $H \subseteq G$ if there exists a subgraph H' of G such that H is isomorphic to H'.

A *dominating set* of G is a set $D \subseteq V(G)$ such that every vertex in G is either in D or adjacent to some vertex in D. $\gamma(G)$ is the minimum value of the size of dominating set of G.

For every $k \in \mathbb{N}^+$, a *k-clique* K_k is a graph with k vertices and every pair of vertices are adjacent.

For every $s, t \in \mathbb{N}^+$, a *biclique* $K_{s,t}$ is a bipartite graph with s vertices in its left side and t vertices in its right side and every two vertices from different side are adjacent. We call $K_{s,t}$ a *balanced biclique* or *k-biclique* if $s = t = k$.

A *multi-graph* $G := (V, E, C)$ contains a vertex set V, a color set C and an edge set $E \subseteq \binom{V}{2} \times C$. We say G contains a subgraph K_k or $K_k \subseteq G$, if there exist k vertices $v_1, v_2, \ldots, v_k$ in V such that for different $i, j \in [k]$, $(\{v_i, v_j\}, c) \in E$ for some color $c \in C$. For any k colors $c_1, c_2, \ldots, c_k \in C$, let $G_{\{c_1, c_2, \ldots, c_k\}} := (V, E', C)$, where $E' = \Big\{(\{u, v\}, c) \in E \;\Big|\; c \in \{c_1, c_2, \ldots, c_k\}\Big\}$. For any $f \in \mathbb{N}$, we say G is a *f-multi-graph* if for all $v, u \in V(G)$, there are at most f multi-edges between v, u. For any multi-graph $G = (V, E, C)$, let

$$\mathsf{CC}(G, k) := \max\left\{h \;\middle|\; h \in \mathbb{N} \text{ and there exists } I \in \binom{C}{k} \text{ such that } K_h \subseteq G_I\right\}.$$

2.2 Computational Complexity

2.2.1 Complexity Classes and Reductions

We identify the $\{0,1\}$-strings and the objects encoded by it. A *decision problem* L can be viewed as a subset of $\{0,1\}^*$. For example, the problem k-Clique consists of all $x \in \{0,1\}^*$ which encodes a graph G and an integer k such that G contains a clique with k vertices.

In the classical computational complexity theory, there are two important classes of problems, namely P and NP. We say $L \in \mathrm{P}$ if there exists a *Turing machine* $M : \{0,1\}^* \to \{0,1\}$ such that for all $x \in \{0,1\}^*$, $M(x) = 1$ iff $x \in L$ and the running time of $M(x)$ is $|x|^{O(1)}$. And $L \in \mathrm{NP}$ if there exists a Turing machine $M : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ with two inputs such that for all $x \in \{0,1\}^*$, $x \in L$ if and only if there exists $w \in \{0,1\}^{|x|^{O(1)}}$ such that $M(x, w) = 1$ and the running time of $M(x, w)$ is $|x|^{O(1)}$.

In parameterized complexity, we have an additional parameter function $\kappa : \{0,1\}^* \to \mathbb{N}$ for each problem. The class FPT is a set of parameterized problems such that for every $(L, \kappa) \in \mathrm{FPT}$, we can decide whether $x \in L$ in $f(\kappa(x)) \cdot |x|^{O(1)}$-time for some computable function f. We call an algorithm fpt if its running-time is in $f(\kappa(x)) \cdot |x|^{O(1)}$ on input an instance x parameterized by κ.

An important concept in complexity theory is *reduction*, which is a mapping from one problem A to another problem B showing that B is at least as hard as A. There two main types of reductions, *Turing reduction* and *many-one reduction*. We only use the many-one reduction in this thesis. For two problems A and B, a many-one reduction from A to B is a mapping $R : \{0,1\}^* \to \{0,1\}^*$ such that $x \in \{0,1\}^*$, $R(x) \in B \iff x \in A$. Obviously, if $R(x)$ can be computed in polynomial time, then $B \in \mathrm{P}$ implies $A \in \mathrm{P}$. R is an fpt-reduction from a parameterized problem (L_1, κ_1) to another parameterized problem (L_2, κ_2) if

- for all $x \in \{0,1\}^*$, $R(x) \in L_2 \iff x \in L_1$;
- $R(x)$ can be computed in $f(\kappa_1(|x|)) \cdot |x|^{O(1)}$-time;
- $\kappa_2(R(x)) \leq g(\kappa_1(x))$ for some computable function g.

We write $(L_1, \kappa_1) \leq_{\text{fpt}} (L_2, \kappa_2)$ if there is an fpt-reduction from (L_1, κ_1) to (L_2, κ_2). Of course, if $(L_1, \kappa_1) \leq_{\text{fpt}} (L_2, \kappa_2)$, then $(L_2, \kappa_2) \in \text{FPT}$ implies $(L_1, \kappa_1) \in \text{FPT}$.

Two important graph problems are believed not to be in FPT.

k-Dominating-Set	
Instance:	A graph G and $k \in \mathbb{N}$.
Parameter:	k.
Problem:	Decide whether G has a dominating set of size at most k.

k-Clique	
Instance:	A graph G and $k \in \mathbb{N}$.
Parameter:	k.
Problem:	Decide whether G has a clique of size at most k.

In fact, k-Dominating-Set is complete for the parameterized complexity class W[2], the second level of the W-hierarchy. The parameterized clique problem k-Clique is W[1] complete. The W-hierarchy can be defined using constant depth circuit. For more detail, we refer to [DF99, FG06, CFK$^+$16]. Here we define W[1] and W[2] in a pragmatic way. That is W[1] contains all the problems that can be fpt-reduced to k-Clique and W[2] contains all the problems that can be fpt-reduced to k-Dominating-Set.

2.2.2 Hypothesis

All the hardness results of this thesis are conditional, i.e., they are based on some standard hypothesis in computational complexity theory. The main hypothesis we work on is FPT $\neq$ W[1], which is equivalent to state that there is no $f(k) \cdot |G|^{O(1)}$-time algorithm to solve k-Clique.

We also need a stronger hypothesis made by Impagliazzo, Paturi and Zane [IPZ98].

Conjecture 2.2.1 (Exponential Time Hypothesis)**.** *n-variable* 3-**SAT** *cannot be solved in time* $2^{o(n)}$.

We use the following lower bound for k-Clique under ETH.

Theorem 2.2.2 ([CHKX04])**.** *Assumming* ETH *there is no $f(k) \cdot n^{o(k)}$-time algorithm for k-Clique.*

2.3 Optimization Problems

An *optimization problem O* is a 4-tuple $(I, \text{sol}, \text{cost}, \text{goal})$ where

(1) I is the set of instances.

(2) For an instance $x \in I$, $\mathrm{sol}(x)$ is a set of feasible solutions of x.

(3) For $x \in I$ and $y \in \mathrm{sol}(x)$, $\mathrm{cost}(x, y) \in \mathbb{N}$.

(4) $\mathrm{goal} \in \{\max, \min\}$.

Example 2.3.1. *Consider the following optimization problem.*

MAX-CLIQUE	
Instance:	A graph G.
Solution:	A subset $K \subseteq V(G)$ such that $G[K]$ is a clique.
Cost:	$\lvert K\rvert$.
Goal:	max.

In this case, goal $=$ max *and* I *is the set of all graphs. For every graph* G, $\mathrm{sol}(G)$ *is the collection of subsets of* $V(G)$ *which induces a clique in* G*. For* $K \in \mathrm{sol}(G)$, $\mathrm{cost}(G, K) = |K|$.

The task of an optimization problem is to find a feasible solution $y \in \mathrm{sol}(x)$ for an input x with optimal cost value $\mathrm{OPT}(x)$, i.e., $\mathrm{OPT}(x) = \mathrm{cost}(x, y) = \mathrm{goal}\{\mathrm{cost}(y', x) \mid y' \in \mathrm{sol}(x)\}$. For $\rho \geq 1$, $\mathbb{A}$ is an ρ-approximation algorithm for O if for all $x \in I$, $\frac{\mathrm{cost}(x,\mathbb{A}(x))}{\mathrm{OPT}(x)} \leq \rho$ when goal $=$ min and $\frac{\mathrm{OPT}(x)}{\mathrm{cost}(x,\mathbb{A}(x))} \leq \rho$ when goal $=$ max.

An optimization problem O is NP-optimization if for all $y \in \mathrm{sol}(x)$, $|y| \leq |x|^{O(1)}$ and there is a polynomial time algorithm that decides whether $y' \in \mathrm{sol}(x)$ for all $y' \in \{0, 1\}^{|x|^{O(1)}}$ and the value of $\mathrm{cost}(x, y)$ can be computed in polynomial time. All the optimization problems we consider in this thesis are NP-optimization.

2.3.1 Parameterized Approximation

Parameterized complexity and approximation algorithm can be combined in several ways [Mar08]. In this thesis, we only consider the following two versions:

- **Approximation parameterized by cost.** In this case, we use the cost of the optimal solution as parameter. Using the definition of [CGG07], $\mathbb{A}$ is an fpt-approximation algorithm with ratio $\rho \geq 1$ if for every input (x, k) with $\mathrm{sol}(x) \neq \emptyset$ satisfying

$$\begin{cases} \mathrm{OPT}(x) \geq k & \text{if goal} = \max, \\ \mathrm{OPT}(x) \leq k & \text{if goal} = \min, \end{cases} \tag{2.1}$$

  $\mathbb{A}$ outputs $y \in \mathrm{sol}(x)$ in running time $f(k) \cdot |x|^{O(1)}$ such that

$$\begin{cases} \mathrm{cost}(x, y) \geq \frac{k}{\rho} & \text{if goal} = \max, \\ \mathrm{cost}(x, y) \leq \rho \cdot k & \text{if goal} = \min, \end{cases} \tag{2.2}$$

  If (x, k) does not satisfy (2.1), the output of $\mathbb{A}$ can be arbitrary. For MAX-CLIQUE and MIN-DOMINATING-SET, we consider approximation algorithms parameterized by cost.

– **Approximation with instance parameters.** In this case, the parameter k and $\mathrm{OPT}(x)$ need not to satisfy the condition (2.1). We try to find an approximation algorithm whose running time is $f(k) \cdot |x|^{O(1)}$. For Max-k-Subset-Intersection and Max-k-EdgeCol-Clique, we consider approximation algorithms parameterized with instance parameters k.

2.3.2 Gap-producing fpt-reductions

We need stronger reductions for parameterized optimization problems. Suppose we want to prove an maximization problem O with parameter function κ is hard to approximate to ratio ρ under FPT $\neq$ W[1]. It suffices to find an fpt-reduction R from k-Clique to O with two additional conditions that

(1) if G contains a k-clique, then $\mathrm{OPT}(R(G,k)) \geq \rho \cdot g(G,k)$,

(2) if G contains no k-clique, then $\mathrm{OPT}(R(G,k)) < g(G,k)$,

where g is a computable function.

Remark 2.3.2. *If the optimization problem O is parameterized by cost, then g is a computable function that depends only on k.*

Similarly, we can define gap-producing fpt-reductions for minimization problems. If O is a minimization problem, then the fpt-reduction R should also satisfy the following two conditions:

(1) if G contains a k-clique, then $\mathrm{OPT}(R(G,k)) \leq g(G,k)$.

(2) if G contains no k-clique, then $\mathrm{OPT}(R(G,k)) > \rho \cdot g(G,k)$.

For example, a gap-producing fpt-reduction from k-Clique to Min-Dominating-Set for ratio 2 is a reduction R such that on input every graph G and k, it constructs a graph H in $f(k) \cdot |G|^{O(1)}$-time satisfying

– if G contains a k-clique, then $\gamma(H) \leq g(k)$,

– if G contains no k-clique, then $\gamma(H) > 2g(k)$.

Given that there is a gap-producing fptreduction from k-Clique to an optimization problem O_1, we can prove the inapproximability of another optimization problem O_2 by providing a gap-preserving fpt-reduction from O_1 to O_2.

Gap-preserving fpt-reductions

Depending on whether O_1 and O_2 are minimization or maximization, there are four cases of gap-preserving reduction. We define gap-preserving fpt-reduction from a maximization problem O_1 to a minimization problem O_2. The other cases are similar.

For $\alpha, \beta > 1$, $g_1, g_2, : \mathbb{N} \to \mathbb{N}$, and two parameter function $\kappa_1, \kappa_2 : \{0,1\}^* \to \mathbb{N}$. An fpt-reduction from O_1 to O_2 is gap-preserving if on input every instance $x_1 \in O_1$, it constructs an instance x_2 satisfying the following property.

- $\mathrm{OPT}(x_1) \geq \alpha \cdot g_1(x_1, \kappa(x_1)) \Rightarrow \mathrm{OPT}(x_2) \leq g_2(x_1, \kappa_1(x_1))$.
- $\mathrm{OPT}(x_1) < g_1(x_1, \kappa(x_1)) \Rightarrow \mathrm{OPT}(x_2) > \beta \cdot g_2(x_1, \kappa_1(x_1))$.

2.4 Probability

In this thesis, we deal exclusively with probability spaces $(\Omega, \Pr)$ where Ω contains finite samples. Pr is a nonnegative function from Ω to $[0,1]$ such that $\sum_{x\in\Omega} \Pr[x] = 1$.

A *random variable* over $(\Omega, \Pr)$ is a function $X : \Omega \to \mathbb{R}$. The *expectation* and *variance* of a random variable are defined as follows.

- $\mathrm{E}[X] := \sum_{x\in\Omega} X(x) \cdot \Pr[x]$.
- $\mathrm{Var}[X] := \sum_{x\in\Omega} \mathrm{E}[X(x)^2] - \mathrm{E}[X(x)]^2$.

An *event* E can be treated as a subset of Ω. The probability of E is defined as $\Pr[E] := \sum_{x\in E} \Pr[x]$. For example, given a random variable X, the event "$X > 0$" can be regarded as a set $E := \{x \in \Omega \mid X(x) > 0\}$. Hence $\Pr[X > 0] = \sum_{x\in\Omega, X(x)>0} \Pr[x]$.

One important methodology we learn from Erdős's paper [Erd59] on graph theory and probability is that to prove some graph with a certain property exists, it suffices to demonstrate that the probability of such graph is positive in some probability space.

Example 2.4.1. *Suppose we want to show that there exists a graph which contains many K_k-subgraphs but has no $K_{k+1,k+1}$-subgraph.*

For $n \in \mathbb{N}$ and $p : \mathbb{N} \to [0,1]$. The random graph $\mathcal{G}(n,p)$ *is a probability space where Ω is the set of all graphs with vertex set $[n]$. $p(n)$ is the probability that two vertices are adjacent. For each graph $G \in \Omega$, $\Pr(G) = p(n)^{|E(G)|}(1-p(n))^{\binom{n}{2}-|E(G)|}$.*

For a graph H, let $X_H : \Omega \to \mathbb{N}$ be the random variable over $\mathcal{G}(n,p)$ such that for each $G \in \Omega$ $X_H(G)$ is the number of H-subgraphs in G.

For $\varepsilon > 0$, it is easy to check that:

- *If $p(n) = n^{-\frac{|V(H)|+\varepsilon}{|E(H)|}}$, then $\mathrm{E}[X_H] \in O(n^{-\varepsilon})$.*
- *If $p(n) = n^{-\frac{|V(H)|-\varepsilon}{|E(H)|}}$, then $\mathrm{E}[X_H] \in O(n^{\varepsilon})$.*

Observe that $\frac{2k+2}{(k+1)^2} < \frac{2k}{k(k-1)}$. *Let* $\delta \in (\frac{2k+2}{(k+1)^2}+\varepsilon, \frac{2k}{k(k-1)}-\varepsilon)$ *for some small positive* ε. *Then we consider the random graph* $\mathcal{G}(n, n^{-\delta})$ *and the random variables* X_{K_k} *and* $X_{K_{k+1,k+1}}$ *over it. If we could show that* $\Pr[X_{K_k} < n^{\varepsilon}] + \Pr[X_{K_{k+1,k+1}} > n^{-\varepsilon}] < 1$, *then there exists* $G \in \mathcal{G}(n, n^{-\delta})$ *satisfying* $X_{K_k}(G) \geq n^{\varepsilon}$ *and* $X_{K_{k+1,k+1}}(G) \leq n^{-\varepsilon} < 1$. *Hence* G *has* n^{ε} K_k*-subgraphs but contains no* $K_{k+1,k+1}$*-subgraph.*

To give upper bounds for formulas in the forms $\Pr[X > \alpha]$ and $\Pr[X < \alpha]$, we need the following tools.

Theorem 2.4.2 (Markov's Inequality)**.** *Let* $X \geq 0$ *be a random variable , then*

$$\Pr[X \geq \alpha] \leq \frac{\mathrm{E}[X]}{\alpha}. \tag{2.3}$$

Theorem 2.4.3 (Chebyshev's Inequality)**.** *For any real* $\lambda > 0$,

$$\Pr[|X - \mathrm{E}[X]| \geq \lambda] \leq \frac{\mathrm{Var}[X]}{\lambda^2}. \tag{2.4}$$

2.5 Finite Field

A finite field $(\mathbb{F}, \cdot, +)$ consists of a set of finite elements $\mathbb{F}$ and two operations, namely multiplication $\cdot$ and addition $+$. $\mathbb{F}$ contains two distinct elements $\boldsymbol{1}$ and $\boldsymbol{0}$. The elements in $\mathbb{F}$ and the two operations obey the following rules.

- Rules about addition:
  - for every $u, v, w \in \mathbb{F}$, $u + v + w = u + (v + w)$;
  - for every $v \in \mathbb{F}$, $v + \boldsymbol{0} = v$;
  - for every $u, v \in \mathbb{F}$, $u + v = v + u$;
  - for every $v \in \mathbb{F}$, there exists a unique $u \in \mathbb{F}$ such that $u + v = \boldsymbol{0}$. Such u is written as $-v$. We can define subtraction by $x - v = x + (-v)$.
- Rules about multiplication:
  - for every $u, v, w \in \mathbb{F}$, $u \cdot v \cdot w = u \cdot (v \cdot w)$;
  - for every $v \in \mathbb{F}$, $v \cdot \boldsymbol{0} = \boldsymbol{0}$;
  - for every $v \in \mathbb{F}$, $v \cdot \boldsymbol{1} = v$;
  - for every $u, v \in \mathbb{F}$, $u \cdot v = v \cdot u$;
  - for every $v \in \mathbb{F} \setminus \{\boldsymbol{0}\}$, there exists a unique $u \in \mathbb{F}$ such that $u \cdot v = \boldsymbol{1}$. Such u can be regarded as $v^{-\boldsymbol{1}}$. We can define division for nonzero v by $x/v = x \cdot v^{-\boldsymbol{1}}$.
- Distributivity: for every $u, v, w \in \mathbb{F}$, $u \cdot (v + w) = u \cdot v + u \cdot w$.

Since $\boldsymbol{1} \in \mathbb{F}$ and $\mathbb{F}$ is close under addition, the elements $\boldsymbol{1}$, $\boldsymbol{1}+\boldsymbol{1}$, …, $\boldsymbol{1}+\boldsymbol{1}+\ldots+\boldsymbol{1}$ are all in $\mathbb{F}$. To ease the notation, for $n \in \mathbb{N}$ and $x \in \mathbb{F}$, let $nx := \sum_{i=1}^{n} x$ be the sum of n identical elements x in $\mathbb{F}$. Of course if $n = 0$, then $nx := \boldsymbol{0}$.

Note that $\mathbb{F}$ is finite. There must exist distinct $n, m \in \mathbb{N}$ such that $m\boldsymbol{1} = n\boldsymbol{1}$. Thus there exists a minimum $p \in \mathbb{N}^+$ such that $p\boldsymbol{1} = \boldsymbol{0}$, which is called the *characteristic* of this field. It is a well-known fact that the characteristic of a finite field is prime.

Fact 2.5.1. *The characteristic of every finite field* $\mathbb{F}$ *is prime.*

Example 2.5.2. *For every prime p, we have a finite field $GF(p) := \{0, 1, 2, \ldots, p-1\}$ with p elements. For $a, b \in GF(p)$, $a + b$ (resp. $a \cdot b$) is the remainder of the integer addition (resp. multiplication) of a and b divided by p. The existence of inverse for every nonzero element under operation $\cdot$ comes from the fact that for every $1 < a < p$, there exists $1 < b < p$ such that $ab = 1 \mod p$. Obviously, $a^{-1} = b$.*

Example 2.5.3. *For every $d \in \mathbb{N}^+$, suppose $f(x) := c_0 + c_1 x + \ldots + c_d x^d$ is a polynomial with degree d, which means that $c_d \neq 0$. Let $GF(p^d) := \{a_0 + a_1 x + \ldots + a_{d-1} x^{d-1} \mid a_i \in GF(p)\}$ be a set with p^d polynomials. For $g, h \in GF(p^d)$, define $g + h$ (resp. $g \cdot h$) as the remainder of the polynomial addition (resp. multiplication) of g and h divided by f. Suppose $f(x)$ is* irreducible, *i.e. there do not exist non-constant polynomials $f_1(x)$ and $f_2(x)$ such $f(x) = f_1(x) f_2(x)$. It follows that for every nonzero $g(x) \in GF(p^d)$, there exists a unique nonzero $h(x) \in GF(p^d)$ such that $gh + ef = 1$ for some polynomial e. We have $g^{-1} = h$. Thus $GF(p^d)$ is a finite field.*

To construct a finite field with p^d elements, we need to find an irreducible polynomial with degree d over $GF(p)$.

Fact 2.5.4. *There exist irreducible polynomials of every degree over every prime field.*

Given d and p, the problem of finding an irreducible polynomial with degree d over the field $GF(p)$ is a fundamental problem. In [AL86], the authors give a deterministic algorithm that outputs an irreducible polynomial with degree d in $(d \cdot \log p)^{O(1)}$ steps, assuming extended Riemann hypothesis. To remove the need for extended Riemann hypothesis in this algorithm is still an open question. In this thesis, we do not need such $(d \cdot \log p)^{O(1)}$-time algorithm. A algorithm in $p^{O(d)}$ steps suffices. It is easy to see that we can enumerate every polynomial f with degree d and check whether f can be divide by other polynomial g using $p^{O(d)}$ arithmetic operations. For more details we refer the reader to [Shp99].

2.6 Color-Coding

We introduce the color-coding technique [AYZ95] which can be used to reduce connected subgraph isomorphic problem to its colored version.

Lemma 2.6.1 ([AYZ95]). *For every $n, k \in \mathbb{N}$ there is a family $\Lambda_{n,k}$ of polynomial time computable functions from $[n]$ to $[k]$ such that for every k-element subset X of $[n]$, there is an $h \in \Lambda_{n,k}$ such that $h(X) = [k]$. Moreover, $\Lambda_{n,k}$ can be computed in time $2^{O(k)} \cdot n^{O(1)}$.*

An immediate application of Lemma 2.6.1 is: when we try to find a reduction to a problem P from a problem of deciding whether a graph G contains a small connected subgraph H, we can assume there is a coloring $\varphi : V(G) \to V(H)$ and find a reduction to our target problem P from the problem of deciding whether a graph G contains a subgraph X isomorphic to H and the vertices of $V(X)$ get distinct colors under φ.

2.6.1 (n, k)-Universal Sets

To construct the family of functions in Lemma 2.6.1, we use a tool called (n, k)-*universal set*.

Definition 2.6.2 ((n, k)-universal set). *Let $n, k \in \mathbb{N}^+$. A set $U \subseteq \{0, 1\}^n$ is an (n, k)-*universal set *if for all k distinct indices $i_1, i_2, \ldots, i_k \in [n]$,*

$$\Big| \{(\mathbf{v}(i_1), \mathbf{v}(i_2), \ldots, \mathbf{v}(i_k)) | \mathbf{v} \in U\} \Big| = 2^k$$

The following lemma can be deduced from Theorem 10.20 and Proposition 10.19 of [Juk11], which provided an elegant construction of (n, k)-Universal Sets using Paley-type graphs and Weil's character sum theorem.

Lemma 2.6.3. *For $k2^k < n$. We can compute an (n, k)-universal set in $n^{O(1)}$-time.*

Lemma 2.6.4. *Let $k, n, m \in \mathbb{N}$ with*

$$(k \cdot \log m) 2^{k \cdot \log m} < \sqrt{n \cdot \log m}.$$

Then in time $(n \log m)^{O(1)}$ we can construct a set $\mathcal{C}$ of functions $[n] \to [m]$ of size $n \cdot \log m$ such that for every $1 \le i_1 < i_2 < \ldots < i_k \le n$ and every $c : [k] \to [m]$ there is a function $\mathbf{c} : [n] \to [m]$ in $\mathcal{C}$ with

$$\mathbf{c}(i_j) = c(j)$$

for all $j \in [k]$.

Proof: We use Lemma 2.6.3 with parameters

$$k \leftarrow k \cdot \log m \quad \text{and} \quad n \leftarrow n \cdot \log m$$

to construct a $(n \cdot \log m, k \cdot \log m)$-universal set $\mathcal{C} \subseteq \{0, 1\}^{n \cdot \log m}$. It is easy to see that every string $\mathbf{c} \in \{0, 1\}^{n \cdot \log m}$ can be understood as a function $\mathbf{c} : [n] \to [m]$ by dividing $\mathbf{c}$ into n blocks, each of length $\log m$. The desired property of $\mathcal{C}$ as a set of functions follows directly from the $(n \cdot \log m, k \cdot \log m)$-universality. $\square$

Using the tool of (n, k)-universal set, we can obtain a stronger version of Lemma 2.6.1.

Lemma 2.6.5. *For $k, n \in \mathbb{N}$, we can construct a set $\Phi := \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$, where $\varphi_i : [n] \to [k]$ in $f(k) \cdot n^{O(1)}$-time such that for all $S \in \binom{[n]}{k}$ and $\psi : S \to [k]$, there exists $\varphi \in \Phi$ such that for all $s \in S$, $\varphi(s) = \psi(s)$.*

Proof: If $(k \cdot \log k)2^{k \cdot \log k} \geq \sqrt{n \cdot \log k}$, then $n \leq g(k)$ for some computable function g. We can construct the set Φ by enumerating all possible function from $[n]$ to $[k]$ in time $n^k = O(g(k)^k)$, which is still in FPT. If $(k \cdot \log k)2^{k \cdot \log k} < \sqrt{n \cdot \log k}$, then by Lemma 2.6.4, we are done. $\square$

Chapter 3

(n, k, ℓ, h)-Threshold Property

In this chapter, we aim to answer the question "where does the gap come from?". To this end, we introduce a threshold property for bipartite graphs. This is followed by two efficient constructions of bipartite graphs satisfying such a property.

3.1 Definition

The key step in proving results of hardness approximation is to produce a gap in the reduction. The PCP-theorem in its hardness of approximation version creates a gap for the maximum fraction of satisfiable clauses of a SAT instance. We note that the maximum fraction of satisfiable clauses is a *global* notion. It involves all the clauses in the instance of SAT problem. Of course, one can derive gaps for *local* objects like clique using the PCP-theorem. However, the reduction always causes the size of the clique to depend on the cardinality of the graph, hence is not fpt.

We make a new approach for proving hardness of approximation by using a class of graphs with a threshold property defined as follows.

Definition 3.1.1 ((n, k, ℓ, h)-threshold property)**.** *For $h > \ell$, a bipartite graph $T = (A \dot\cup B, E)$ with a partition $A = V_1 \dot\cup V_2 \dot\cup \ldots \dot\cup V_n$ satisfy the (n, k, ℓ, h)-threshold property if:*

(T1) Every $k+1$ distinct vertices in A have at most ℓ common neighbors in B, i.e.

$$\forall V \in \binom{A}{k+1}, |\Gamma(V)| \leq \ell.$$

(T2) For every k distinct indices $\{i_1, i_2, \ldots, i_k\} \in \binom{[n]}{k}$, there exist $v_{i_1} \in V_{i_1}, \ldots, v_{i_k} \in V_{i_k}$ such that $v_{i_1}, \ldots, v_{i_k}$ have at least h common neighbors in B, i.e.

$$\exists v_{i_1} \in V_{i_1}, \ldots, v_{i_k} \in V_{i_k}, |\Gamma(\{v_{i_1}, \ldots, v_{i_k}\})| \geq h.$$

Observed that for $h \gg \ell$, there is a gap between the sizes of common neighbors of k-vertex set and $(k+1)$-vertex set. In Chapter 4, we will see that such a gap allows us to conduct reductions which perform local transformations and also produce a gap.

Example 3.1.2. *Let $G := (A \dot\cup R, E)$ be a bipartite graph with $A := [n]$, $B := \binom{[n]}{k}$ and for every $v \in A$ and $u \in B$,*

$$\{v,u\} \in E \iff v \in u$$

Then for every k-vertex set $V \in \binom{A}{k} = \binom{[n]}{k}$, there exists exactly one vertex $u \in B = \binom{[n]}{k}$ such that $V = u$. It is easy to verify that for every $v \in V$, $\{v,u\} \in E(G)$. On the other hand, every $u \in B$ has at most k neighbors according to the definition, which means that every $k+1$ vertices in A have no common neighbors. The graph G satisfies the $(n,k,0,1)$-threshold property with each V_i contains exactly one vertex $i \in [n]$. We note that G contains $\Omega(n^k)$ vertices.

The definition of threshold property was inspired by the work from a remarkable paper [BGK$^+$96], in which the authors gave an explicit construction of a class of graphs satisfying (T1) and

- $\star$ At least a $\frac{1}{2\ell-1}$ fraction of the sets $V \in \binom{A}{k}$ have h common neighbors.

for $\ell = (k+1)!$ and $h = n^{\Theta(1/k)}$. (see Theorem 3.6 and Lemma 3.7 of [BGK$^+$96]) We replace the property $\star$ by (T2) because in the reduction based on graphs with threshold property, we need the k-vertex set with h common neighbors to appear in a predictable way. (The corresponding relation between our notation and that in [BGK$^+$96] is: $k+1 \leftrightarrow t$, $\ell \leftrightarrow s$, $A \leftrightarrow V_1$.)

Another more intuitive way of defining the threshold property is to replace (T2) with

– (T2') Every k distinct vertices in A have at least h common neighbors in B.

We can see that (T2') is a special case of (T2) when $|V_i| = 1$ for all $i \in [n]$. We did not choose (T2'), even though it looks simpler, because we do not know how to construct a graph satisfying (T1) and (T2') in $f(k) \cdot n^{O(1)}$ time (see the Remark 3.2.6). However, with (T2), we can construct graphs satisfying the threshold property in polynomial time.

In the following sections, we give two efficient constructions.

Lemma 3.1.3 (Probabilistic Construction)**.** *For $k, \ell, h, n \in \mathbb{N}$ with $n \geq \max\{2(k+1)^2, 20\}$, $\ell = 2k^2 + 4k - 1$ and $\ell < h \leq n^{\frac{1}{4(k+1)}}$, we can construct in polynomial time a bipartite random graph satisfying the (n,k,ℓ,h) threshold property with probability at least $\frac{9}{10}$.*

Lemma 3.1.4 (Explicit Construction)**.** *For $k, n \in \mathbb{N}^+$ with $k = 6\ell - 1$ for some $\ell \in \mathbb{N}^+$ and $\lceil (n+1)^{\frac{6}{k+1}} \rceil > (k+1)!$, a bipartite graph with the $(n,k,(k+1)!,\lceil (n+1)^{\frac{6}{k+1}} \rceil)$-threshold property can be constructed in $O(n^{18})$-time.*

3.2 Probabilistic Construction

For $k, n \in \mathbb{N}^+$ and $p \in [0,1]$, we define a probability space $\mathcal{G}(n,p) = (\Omega, p)$, where Ω is the set of all bipartite random graphs $G = (A \,\dot\cup\, B, E)$ with $|A| = |B| = n^2$ and every pair of vertices $u \in A$ and $v \in B$ is joined by an edge with probability p, randomly and independently. Furthermore we partition A into n subsets $A = V_1 \,\dot\cup\, V_2 \,, \ldots, \dot\cup\, V_n$ with $|V_i| = n$ for each $i \in [n]$. We use $G(n,p)$ to denote the random variable which is the identity function on Ω. Hence, for every $G \in \Omega$, $\Pr[G(n,p) = G] = \Pr(G)$. We will show that with high probability $G(n,p)$ satisfies the (n,k,ℓ,h)-threshold property for $\ell = 2k^2 + 4k - 1 < h \leq n^{\frac{1}{4(k+1)}}$ and $p = n^{-\frac{2(k+\ell+3)}{(k+1)(\ell+1)}}$.

3.2.1 Estimate $\Pr[G(n,p)$ does not satisfy (T1)$]$

To bound the probability of $G(n,p)$ containing a subgraph isomorphic to $K_{k+1,h}$, we need the following lemma, which is a simple consequence of Markov's Inequality:

Lemma 3.2.1. *Let X be a nonnegative integral random variable, then* $\Pr[X > 0] \leq E[X]$.

Lemma 3.2.2. *With probability at most n^{-2}, $G(n,p)$ does not satisfy (T1).*

Proof: Let X be the number of $K_{k+1,\ell+1}$-subgraphs in $G(n,p)$ with the left $k+1$ vertices in A and the other $\ell+1$ vertices in B. Then

$$E[X] = \binom{n^2}{k+1} \cdot \binom{n^2}{\ell+1} \cdot p^{(k+1)(\ell+1)} \leq n^{2(k+1+\ell+1)} \cdot n^{-2(k+\ell+3)} = n^{-2} \tag{3.1}$$

We have $\Pr[X > 0] \leq E[X] \leq n^{-2}$. It follows from the definition that:

$$\Pr[G(n,p) \text{ does not satisfy (T1)}] \leq \Pr[X > 0] \leq n^{-2}.$$

Hence, when $n \to \infty$, $G(n,p)$ satisfies the first condition of threshold property with high probability.

3.2.2 Estimate $\Pr[G(n,p)$ does not satisfy (T2)$]$

Recall that $\Gamma(S)$ is the common neighbors of all vertices in S. For $S \in \binom{A}{k}$, let

$$X_S := \left| \left\{ T \in \binom{B}{h} \middle| T \subseteq \Gamma(S) \right\} \right| \tag{3.2}$$

In other words, X_S denotes the number of $K_{k,h}$-subgraphs in $G(n,p)$ whose left side vertex set is S.

Lemma 3.2.3. *If $h < n^{\frac{1}{4(k+1)}}$ then* $\Pr[X_S = 0] \leq n^{-\frac{1}{2(k+1)}}$.

Proof: By the Chebyshev's Inequality, $\Pr[X_S = 0] \leq \frac{\mathrm{Var}[X_S]}{E[X_S]^2}$. To give an upper bound for $\Pr[X_S = 0]$, we need to estimate $E[X_S]$ and $Var[X_S]$. It is easy to see that

$$E[X_S] = \binom{n^2}{h} \cdot p^{kh} \tag{3.3}$$

For $T \in \binom{B}{h}$, Let $X_{s,t}$ be a random variable that $X_{S,T} = 1$ if $S \cup T$ forms a complete bipartite subgraph, and $X_{S,T} = 0$ otherwise. It follows that:

$$\begin{aligned}
&\mathrm{Var}[X_S] \\
=&E[X_S^2] - E[X_S]^2 \\
=&E[(\sum_{T\in\binom{B}{h}} X_{S,T})^2] - E[X_S]^2 \\
=&\sum_{T,T'\in\binom{B}{h}} E[X_{S,T}X_{S,T'}] - E[X_S]^2 \\
=&\sum_{i=0}^{h} \sum_{T,T'\in\binom{B}{h},|T\cap T'|=i} X_{S,T}X_{S,T'} \Pr[X_{S,T} = 1, X_{S,T'} = 1] - E[X_S]^2 \\
=&\sum_{i=0}^{h} \binom{n^2}{h}\binom{n^2-h}{h-i}\binom{h}{i} \cdot p^{2hk-ik} - E[X_S]^2 \\
\leq&\sum_{i=1}^{h} \binom{n^2}{h}\binom{n^2-h}{h-i}\binom{h}{i} \cdot p^{2hk-ik} \quad (\text{using } \binom{n^2-h}{h} \leq \binom{n^2}{h} \text{ and (3.3)}) \\
\leq&\binom{n^2}{h}\binom{n^2}{h} p^{2hk} \sum_{i=1}^{h} \frac{\binom{n^2-h}{h-i}\binom{h}{i}p^{-ik}}{\binom{n^2}{h}} \\
\leq&E[X_S]^2 \sum_{i=1}^{h} \frac{\binom{n^2-h}{h-i}\binom{h}{i}p^{-ik}}{\binom{n^2}{h}} \quad (\text{using (3.3)}) \\
\leq&E[X_S]^2 \sum_{i=1}^{h} h^{2i}n^{-2i}p^{-ik} \quad (\text{using } \binom{n^2-h}{h-i}\binom{n^2}{i} \leq \binom{n^2}{h}\binom{h}{i} \text{ and } \binom{h}{i}n^{2i} \leq \binom{n^2}{i}h^i) \\
\leq&E[X_S]^2 \sum_{i=1}^{h} n^{-2i[1-\frac{k(k+\ell+3)}{(k+1)(\ell+1)}-\frac{1}{4(k+1)}]} \quad (\text{using } h \leq n^{\frac{1}{4(k+1)}} \text{ and } p = n^{-\frac{2(k+\ell+3)}{(k+1)(\ell+1)}}) \\
\leq&E[X_S]^2 \sum_{i=1}^{h} n^{\frac{-i}{2(k+1)}} \quad (\text{using } \ell = 2k^2+4k-1) \\
\leq&E[X_S]^2 \cdot n^{-\frac{1}{2(k+1)}}
\end{aligned}$$

Applying the Chebyshev's Inequality, we obtain $\Pr[X_S = 0] \leq n^{-\frac{1}{2(k+1)}}$.

Remark 3.2.4. *In the above deduction, we use $\ell = 2k^2 + 4k - 1$ to show that*

$$1 - \frac{k(k+\ell+3)}{(k+1)(\ell+1)} - \frac{1}{4(k+1)} > \frac{1}{4(k+1)}$$

We can see that if $1 - \frac{k(k+\ell+3)}{(k+1)(\ell+1)} - \frac{1}{4(k+1)} = \alpha$, then $Var[X_S] \leq E[X_S]^2 \cdot n^{-2\alpha}$. Using the the Chebyshev's Inequality will give us $\Pr[X_S = 0] \leq n^{-2\alpha}$. Since our goal is to show that $\Pr[X_S = 0]$ is very small, we need to guarantee that $\alpha > 0$, which implies we must choose $\ell \geq \Omega(k^2)$.

Lemma 3.2.5. *If $n > 2(1+k)^2$ then with probability at most n^{-1}, $G(n,p)$ does not satisfy the second condition of threshold graph.*

Proof: For $I = \{a_1, a_2, \dots, a_k\} \in \binom{[n]}{k}$, consider the subsets $V_{a_1}, V_{a_2}, \dots, V_{a_k}$ in the partition of A. For each $i \in [k]$, suppose $V_{a_i} = \{v_{i1}, v_{i2}, \dots, v_{in}\}$. Denote by Y_I the number of $K_{k,h}$-subgraph in $G(n,p)$ with the restriction that each $V_{a_i} (i \in [k])$ contains exactly one vertex from the left side of such $K_{k,h}$-subgraph. For each $j \in [n]$ let $S_j := \{v_{1j}, v_{2j}, \dots, v_{kj}\}$ and X_{S_j} be the number of h-vertex sets in $\Gamma(S_j)$ as defined in (3.2). We note that for distinct j and j', $S_j \cap S_{j'} = \emptyset$. It is easy to see that:

$$\begin{aligned}
\Pr[Y_I = 0] \leq & \Pr[\forall j \in [n], X_{S_j} = 0] && (3.4)\\
= & \prod_{j=1}^{n} \Pr[X_{S_j} = 0] && (3.5)\\
\leq & n^{-\frac{n}{2(k+1)}} && (3.6)
\end{aligned}$$

$G(n,p)$ does not satisfy the second condition of threshold graph if there exists $I \in \binom{[n]}{k}$ such that $Y_I = 0$. By the union bound:

$$\begin{aligned}
& \Pr[G(n,p) \text{ does not satisfy (T2)}] && (3.7)\\
\leq & \sum_{I \in \binom{[n]}{k}} \Pr[Y_I = 0] && (3.8)\\
\leq & n^{k - \frac{n}{2(k+1)}} && (3.9)\\
\leq & n^{-1} && (3.10)
\end{aligned}$$

Remark 3.2.6. *In step 3.5, we use the fact that each V_i contains n vertices. Suppose we replace (T2) with (T2') in the definition of threshold property, then each V_i contains exactly one vertex. In this case, we can only bound $\Pr[Y_I = 0]$ by $n^{-\frac{1}{2(k+1)}}$. In consequence, we obtain $n^{k-\frac{1}{2(k+1)}}$ in step 3.9 instead of $n^{k-\frac{n}{2(k+1)}}$, which is insufficient* *for us to deduce*

$$\Pr[G(n,p) \textit{ does not satisfy (T2')}] \le n^{-1}$$

3.2.3 Proof of Lemma 3.1.3

Choose n large enough such that $n > 2(1+k)^2$ and $n > 20$, then from Lemma 3.2.3 and Lemma 3.2.5 we can deduce:

$$\Pr[G(n,p) \text{ does not satisfy T1 or T2}] \le n^{-2} + n^{-1} \le 1/10 \tag{3.11}$$

Thus $G(n,p)$ is an (n,k,ℓ,h) threshold bipartite graph with probability larger than 9/10.

3.3 Explicit Construction

3.3.1 Using Paley-Type Graphs

Definition 3.3.1 (Paley-type Graphs). *For any prime power $q = p^t$ and $d \mid q-1$, a Paley-type bipartite graph $G(q,d) := (A \dot\cup B, E)$ is defined as follows.*

Vertices $A = B = GF(q)^{\times}$*;*

Edges $\forall x \in A, y \in B, xy \in E \iff (x+y)^{\frac{q-1}{d}} = 1$.

It is a well-known fact that for any prime power $q = p^t$, there exists a finite field $\mathbb{F}_q$ with q elements and $\mathbb{F}_q = \mathbb{F}_p[X]/(f)$, where f is an irreducible polynomial over $\mathbb{F}_p$ with degree t. Such irreducible polynomial can be found by brute-force search. It is not hard to see that:

Lemma 3.3.2. *$G(q,d)$ can be constructed in $O(q^3)$ time.*

The Paley-type graphs have many nice properties, the following one is proved in [KRS96, BGK⁺96]:

Theorem 3.3.3 (Theorem 5.1 in [BGK⁺96]). *The graph $G(p^t, p-1)$ contains no subgraph isomorphic to $K_{t,t!+1}$.*

Therefore, the graph $G(p^t, p-1)$ satisfies (T1) for $k \leftarrow t-1$ and $\ell \leftarrow t!$, our next step is to show that it also satisfies (T2) for a proper choice of parameter values. To this end, we prove:

Lemma 3.3.4 (Intersection). *For any $d,k,r,s \in \mathbb{N}^+$ and prime power q with $q-1 = rs$, $d \mid q-1$ and $\sqrt{q} \ge \frac{sk}{d} + 1$. Let $a_1, \dots, a_k$ be distinct elements in $GF^{\times}(q)$, g be the generator of $GF^{\times}(q)$, for each $i \in [s]$, denote $V_i := \{g^{i+s}, g^{i+2s}, \dots, g^{i+sr}\}$, then for any $j \in [s]$, the number of solutions $x \in V_j$ to the system of equations $(a_i + x)^{\frac{q-1}{d}} =$*

$1(\forall i \in [k])$ is in $\frac{q}{sd^k} \pm k\sqrt{q}$.

Lemma 3.3.4 generalizes Lemma 3.8 in [BGK+96] by restricting the solutions to any subset $V_j (j \in [s])$. If we set $s = 1$, then we obtain Lemma 3.8 in [BGK+96]. The intuition behind Lemma 3.3.4 is that the solutions of $(a_i + x)^{\frac{q-1}{d}} = 1$ distribute "randomly": the equation $(a_i+x)^{\frac{q-1}{d}} = 1$ has $\frac{q-1}{d}$ solutions, we may say that a random generated element $x \in GF^\times(q)$ satisfies this equation with probability $\frac{1}{d}$, hence x satisfies the system of equations $(a_i + x)^{\frac{q-1}{d}} = 1(\forall i \in [k])$ with probability $\frac{1}{d^k}$. Since V_j contains $\frac{1}{s}$ elements of $GF^\times(q)$, we expect the number of solutions $x \in V_j$ to the system of equations $(a_i + x)^{\frac{q-1}{d}} = 1(\forall i \in [k])$ is dominated by $\frac{q}{sd^k}$, and $k\sqrt{q}$ is the error term. We postpone the proof of Lemma 3.3.4.

Lemma 3.3.5. *For any $p, r, s, t \in \mathbb{N}^+$ with p is prime, $\frac{s}{p-1} + 1 \le \sqrt{p^{t+1}}$ and $p^{t+1} - 1 = rs$. Let g be the generator of $GF^\times(p^{t+1})$, for each $i \in [s]$, denote $V_i := \{g^{i+s}, g^{i+2s}, \dots, g^{i+sr}\}$. Then in the Paley-type bipartite graph $G(p^{t+1}, p-1) = (A \dot\cup B, E)$, for any t distinct indices $a_1, a_2, \dots, a_t \in [s]$, there exist $\boldsymbol{v} \in V_{a_1} \times \dots \times V_{a_t}$, such that $|\Gamma(\boldsymbol{v})| \ge p$.*

Proof: Fix t distinct indices $a_1, a_2, \dots, a_t \in [s]$. Consider the sets $S := V_{a_1} \times \dots \times V_{a_t}$ and $\Gamma\langle S\rangle := \{\{\mathbf{v}, u\} \mid \mathbf{v} \in S, u \in B, u \in \Gamma(\mathbf{v})\}$. Since $\frac{s}{p-1} + 1 \le \sqrt{p^{t+1}}$, apply Lemma 3.3.4 with $q \leftarrow p^{t+1} \quad d \leftarrow p-1 \quad k \leftarrow 1$, each elements in $GF^\times(p^{t+1})$ has at least

$$\frac{p^{t+1}}{s(p-1)} - p^{\frac{t+1}{2}} \ge \frac{p^t}{s} + \frac{p^{t-1}}{s} - p^{\frac{t+1}{2}} \ge \frac{p^t}{s} + p^{\frac{t+1}{2}} - p^{\frac{t+1}{2}} = \frac{p^t}{s}$$

neighbors in each V_{a_i}. Thus $|\Gamma\langle S\rangle| \ge (\frac{p^t}{s})^t(p^{t+1} - 1)$; on the other hand, $|S| = (\frac{p^{t+1}-1}{s})^t$, by the pigeonhole principle, there exists $\mathbf{v} \in S$ such that

$$|\Gamma(\mathbf{v})| \ge \frac{|\Gamma\langle S\rangle|}{|S|} \ge \frac{(\frac{p^t}{s})^t(p^{t+1}-1)}{(\frac{p^{t+1}-1}{s})^t} = \frac{p^{t^2}}{(p^{t+1}-1)^{t-1}} \ge \frac{p^{t^2}}{p^{t^2-1}} \ge p$$

In the construction of the threshold bipartite graphs, we need the famous Bertrand's Postulate from number theory, whose proof can be found in [Ram19, Erd34].

Proof: [of Lemma 3.1.4] For any positive integer n and $k = 6\ell - 1$, by Bertrands's Postulate, we can choose an arbitrary prime p between $\lceil (n+1)^{\frac{1}{\ell}} \rceil$ and $2\lceil (n+1)^{\frac{1}{\ell}} \rceil$, then we construct the Paley-type graph $G(p^{k+1}, p-1) = (A \dot\cup B, E)$. Let $s = p^\ell - 1$, we have $s \ge n$ and $p^{k+1} - 1 = p^{6\ell} - 1 = sr$, where $r = (p^{2\ell} + p^\ell + 1)(p^{3\ell} + 1)$. For each $i \in [s]$, denote $V_i := \{g^{i+s}, g^{i+2s}, \dots, g^{i+rs}\}$, where g is the generator of $GF^\times(p^{k+1})$. It is easy to see that the graph $G(p^{k+1}, p-1)$ including the partition of its vertices set can be constructed in $O(p^{3(k+1))} = O(n^{18})$. We only need to check $G(p^{k+1}, p-1)$ satisfies (T1) and (T2) for parameter n, k, $\ell \leftarrow (k+1)!$ and $h \leftarrow \lceil (n+1)^{6/(k+1)} \rceil$.

By Theorem 3.3.3, $G(p^{k+1}, p-1)$ contains no subgraph isomorphic to $K_{k+1,(k+1)!+1}$, i.e. every $k+1$ distinct vertices in A have at most $(k+1)!$ common neighbors in B. Thus $G(p^{k+1}, p-1)$ satisfies (T1).

Since $\frac{s}{p-1}+1=\frac{p^{\ell}-1}{p-1}+1\leq p^{3\ell}=p^{\frac{k+1}{2}}$, apply Lemma 3.3.5 with $t\leftarrow k$, we have for any k distinct indices $a_1,a_2,\ldots,a_k\in[s]$, there exist $v_{a_i}\in V_{a_i}$ $(\forall i\in[k])$ such that $v_{a_1},\ldots,v_{a_k}$ have at least $p\geq\lceil(n+1)^{\frac{1}{\ell}}\rceil>(k+1)!$ common neighbors in B.

Finally, since $s\geq n$, $G(p^{k+1},p-1)$ is an $(n,k,(k+1)!,\lceil(n+1)^{\frac{1}{\ell}}\rceil)$ threshold bipartite graph.

3.3.2 Proof of the Intersection Lemma

Definition 3.3.6 (Character)**.** *A character of a finite field $GF(q)$ is a function $\chi: GF(q)\to\mathbb{C}$ satisfying the following conditions:*

1 $\chi(0)=0$;

2 $\chi(1)=1$;

3 $\forall a,b\in GF(q),\chi(ab)=\chi(a)\chi(b)$

Remark 3.3.7. *Since for all $x\in GF^{\times}(q)$, $x^{q-1}=1$, we have $\chi(x)^{q-1}=\chi(x^{q-1})=1$. That is χ maps all the elements in $GF^{\times}(q)$ to the roots of $z^{q-1}=1$ in $\mathbb{C}$.*

Definition 3.3.8 (Order)**.** *A character χ of a finite field $GF(q)$ has* order *d if d is the minimal positive integer such that $\forall a\in GF(q)^{\times}$, $\chi(a)^d=1$.*

Theorem 3.3.9 (A. Weil)**.** *Let $GF(q)$ be a finite field, χ a character of $GF(q)$ and $f(x)$ a polynomial over $GF(q)$ if:*

1 *The order of χ is d;*

2 *$f(x)\neq c\cdot(g(x))^d$ for any polynomial g over $GF(q)$ and $c\in GF(q)$;*

3 *The number of distinct roots of f in the algebraic closure of $GF(q)$ is s.*

then

$$|\sum_{x\in GF(q)}\chi(f(x))|\leq(s-1)\sqrt{q}$$

(See [Sch76], page 43, Theorem 2C')

Remark 3.3.10. *It is well known that the expected translation distance after n-step random walk in 2-dimension space is about $\sqrt{n}$. By the character sum theorem, we can see that the values of $f(x)$ for $x\in GF(q)$ distribute randomly to some extent.*

Suppose g is the generator of $GF(q)$, where q is a prime power and $q-1=rs(s,r\in\mathbb{N})$, let $V_i:=\{g^{i+s},g^{i+2s},\cdots,g^{i+rs}\}\ l(i\in[s])$. It is obvious that $GF^{\times}(q)=V_1\cup V_1\cdots\cup V_s$ and $\forall i\in[s],|V_i|=r$. With these notations, we have:

Lemma 3.3.11. *Suppose f is a function from $GF(q)$ to $\mathbb{C}$, then $\forall i \in [s]$,*

$$\sum_{z \in V_i} f(z) = \frac{1}{s} \sum_{x \in GF^{\times}(q)} f(g^i x^s)$$

Proof: For any element $z = g^{i+js} \in V_i (j \in [r])$, consider the set

$$X_j := \{x \in GF^{\times}(q) \mid g^i x^s = g^{i+js}\}.$$

It is easy to check that $X_j = \{g^{j+r}, \cdots, g^{j+sr}\}$, i.e. there are exactly s element x in $GF^{\times}(q)$ such that $g^i x^s = z$ for each $z \in V_i$. Thus $\sum_{z \in V_i} f(z) = \frac{1}{s} \sum_{x \in GF^{\times}(q)} f(g^i x^s)$.

Proof: [of Lemma 3.3.4] Let $\omega \in \mathbb{C}$ be the primitive d^{th} root of unity and g be a generator of the multiplicative group $GF^{\times}(q)$, define a function $\chi : GF(q) \to \mathbb{C}$ as:

1 $\chi(0) = 0$;

2 for $g^\ell \in GF^{\times}(q)$ set $\chi(g^\ell) = \omega^\ell$.

Then:

i χ is a character of $GF(q)$. Because $\chi(g^a \cdot g^b) = \omega^{a+b} = \chi(g^a)\chi(g^b)$ and $\chi(1) = \chi(g^{q-1}) = w^{q-1} = 1$ since $d \mid q-1$;

ii The order of χ is d. Observed that $\chi(g)^n = \chi(g^n) = 1 \iff \omega^n = 1 \iff d \mid n$, the order of χ is $\geq d$; on the other hand, for all $z = g^{i_z} \in GF(q)^{\times}$, $\chi(z)^d = \chi(g^{i_z d}) = \omega^{d i_z} = 1$, so the order of χ is $\leq d$;

iii $\chi(x) = 1 \iff x^{\frac{q-1}{d}} = 1$. Suppose $x = g^i$ and notice that $g^\ell = 1 \iff q-1 \mid \ell$, it follows that $1 = x^{\frac{q-1}{d}} = g^{\frac{i(q-1)}{d}} \iff q-1 \mid \frac{i(q-1)}{d} \iff d \mid i \iff \omega^i = 1 \iff \chi(x) = \chi(g^i) = 1$.

By iii, $(a_i + x)^{\frac{q-1}{d}} = 1 \iff \chi(a_i + x) = 1$, let

$$X := \{x \in V_j \mid \forall i \in [k], \chi(x + a_i) = 1\}$$

Recall that $a \pm b$ denotes the set of real number between $a - b$ and $a + b$, our goal is to show that $|X| \in \frac{q}{sd^k} \pm k\sqrt{q}$.

Consider a polynomial $h \colon \mathbb{C} \to \mathbb{C}$ with $h(z) = \frac{z^d - 1}{z - 1} = 1 + z + \cdots + z^{d-1}$, then:

$h(1) = d$;

$h(\omega^i) = 0$, for $i = 1, 2, \cdots, d-1$;

$h(0) = 1$.

Let $H(x) = \prod_{i=1}^{k} h(\chi(a_i + x))$, then:

if $x \in X$, then $H(x) = d^k$;

if $x = -a_i$ for some $i \in [k]$ and $\chi(x + a_{i'}) = 1(\forall i' \in [k], i' \neq i)$, then $H(x) = d^{k-1}$;

otherwise $H(x) = 0$

Now consider the sum $S := \sum_{x \in V_j} H(x)$, we have:

$$|X|d^k \leq S \leq |X|d^k + kd^{k-1}$$

We only need to estimate S. Using Lemma 3.3.11, we can rewrite S as

$$\begin{aligned} S &= \sum_{x \in V_j} H(x) \\ &= \frac{1}{s} \sum_{x \in GF^{\times}(q)} H(g^j x^s) \\ &= \frac{1}{s} [\sum_{x \in GF(q)} H(g^j x^s) - H(0)] \end{aligned}$$

Expand the product in $H(g^j x^s)$:

$$\begin{aligned} &\sum_{x \in GF(q)} H(g^j x^s) \\ =& \sum_{x \in GF(q)} \prod_{i=1}^{k} h(\chi(a_i + x^s g^j)) \\ =& \sum_{x \in GF(q)} \prod_{i=1}^{k} [1 + \chi(a_i + x^s g^j) + \cdots + \chi(a_i + x^s g^j)^{d-1}] \\ =& \sum_{x \in GF(q)} \sum_{\psi \in \{0, \cdots, d-1\}^k} \chi(f_\psi(x)) \\ =& q + \sum_{\psi \in \{0, \cdots, d-1\}^k \setminus \{0\}^k} \sum_{x \in GF(q)} \chi(f_\psi(x)) \end{aligned}$$

Where $\psi \in \{0, 1, \cdots, d-1\}^k$ is a function from $[k]$ to $\{0, \cdots, d-1\}$ and $f_\psi(x) := \prod_{i=1}^{k}(a_i + x^s g^j)^{\psi(i)}$.

To invoke Weil's theorem on the character sum $\sum \chi(f_\psi(x))$ for any $\psi \in \{0, \cdots, d-1\}^k \setminus \{0\}^k$, we need to check:

(1) The order of χ is d, this is done in the previous discussion;

(2) $f_\psi(x) \neq c \cdot (g(x))^d$ for any polynomial g over $GF(q)$ and $c \in GF(q)$. It suffices to show that any solution of $f_\psi(x)$ in the algebraic closure of $GF(q)$ has multiplicity $\leq d-1$. Let $f_{ij}(x) = a_i + x^s g^j$, notice that the derivative of $f_{ij}(x)$ is $f'_{ij}(x) = sg^j x^{s-1}$, we claim that all the roots of $f_{ij}(x)$ have multiplicity 1, otherwise $f_{ij}(x)$ and $f'_{ij}(x)$

have a common root, then $sa_i = 0$. This is impossible because $q-1 = sr$ implies $rsa_i = -a_i \neq 0$; on the other hand, for distinct $i, i' \in [k]$, $f_{ij}(x)$ and $f_{i'j}(x)$ do not share a common root because $a_i \neq a_{i'}$. It follows that each root of f_ψ has multiplicity $\leq d-1$.

(3) f_ψ has at most ks distinct roots in the algebraic closure field of $GF(q)$.

By Weil's theorem

$$|\sum_{x\in GF(q)} \chi(f_\psi(x))| \leq (ks-1)\sqrt{q},$$

So

$$\begin{aligned}|S + \frac{H(0)}{s} - \frac{q}{s}| &= \frac{1}{s} \sum_{\psi\in\{0,\cdots,d-1\}^k\setminus\{0\}^k} \sum_{x\in GF(q)} \chi(f_\psi(x)) \\ &\leq \frac{d^k}{s}(ks-1)\sqrt{q}\end{aligned}$$

Finally, notice that $H(0) \leq d^k$ and $\sqrt{q} > \frac{sk}{d} + 1$, we have

$$\begin{aligned}|X| &\in \frac{S}{d^k} \pm \frac{k}{d} \\ &\subseteq \frac{q - H(0) \pm (ks-1)d^k\sqrt{q}}{sd^k} \pm \frac{k}{d} \\ &\subseteq \frac{q}{sd^k} \pm (k\sqrt{q} + \frac{k}{d} + \frac{1}{s} - \frac{\sqrt{q}}{s}) \\ &\subseteq \frac{q}{sd^k} \pm k\sqrt{q}\end{aligned}$$

3.4 Some Extremal Questions

In Remark 3.2.4, we observe that $\ell = \Omega(k^2)$ is necessary for the probabilistic construction of graphs satisfying the threshold property. A natural question is

Question 3.4.1. *Is it possible to construct an $n^{O(1)}$-vertex graph satisfying the (n, k, ℓ, h)-threshold for every large n and $\ell = o(k^2)$?*

In Remark 3.2.6, we discuss why we use (T2) instead of (T2') in the definition of threshold property. It seems that the answer to the following question is negative.

Question 3.4.2. *Is there any $n^{O(1)}$-vertex bipartite graph $G := (A \dot\cup B, E)$ satisfying*

- *(T1) for all $V \in \binom{A}{k+1}$, V has at most ℓ common neighbors.*
- *(T2') for all $V \in \binom{A}{k}$, V has at least h common neighbors.*

where $h > \ell > 0$.

In both questions above, we require the vertex number to be in $n^{O(1)}$, otherwise a trivial solution is to introduce h new vertices for every $V \subset \binom{[n]}{k}$ and make them adjacent to every vertex in V. The graph constructed in this way has $\Theta(n^k)$ vertices but satisfies (T1) and (T2') and the parameter ℓ is equal to zero. (See Example 3.1.2.)

In Question 3.4.2, if $\binom{n}{h} < \binom{n}{k}$, we can show that no such graph exists. By (T1), every k-vertex subset of A has at least h common neighbors. Since $\binom{n}{h} < \binom{n}{k}$, by the pigeonhole principle, there must exist two distinct k-vertex subsets V_1 and V_2 of A such that V_1 and V_2 share h common neighbors. This implies that $V_1 \cup V_2$ has $h > \ell$ common neighbors. However $|V_1 \cup V_2| \geq k + 1$, which contradicts (T2').

To ask more questions with such flavor, a definition would bring us a lot of convenience.

Definition 3.4.3. *For bipartite graphs $G := (A(G) \cup B(G), E(G))$ and $S := (A(S) \cup B(S), E(S))$, let*

$$\boldsymbol{lcount}(G, S) := \left| \left\{ X \in \binom{A(G)}{|A(S)|} \;\middle|\; S \subseteq G[X \cup B(G)] \right\} \right|$$

Assume that G contains n vertices in each side. Then Question 3.4.2 and Question 3.4.1 can be formulated as follows.

Question 3.4.4. *Does $\boldsymbol{lcount}(G, K_{k+1,\ell}) = 0 \Rightarrow \boldsymbol{lcount}(G, K_{k,\ell+1}) = n^{o(k)}$ for every $\ell = o(k^2)$?*

Question 3.4.5. *Does $\boldsymbol{lcount}(G, K_{k,\ell}) = \binom{n}{k} \Rightarrow \boldsymbol{lcount}(G, K_{k+1,h}) > 0$ for any $h, \ell \in \mathbb{N}^+$?*

A more general question in this direction is

Question 3.4.6. *Let S and T be two bipartite graphs. What is the maximum value of $\boldsymbol{lcount}(G, S)$ over all the bipartite graphs G with $\boldsymbol{lcount}(G, T) = 0$?*

The pursuit of the answer to Question 3.4.6 is beyond the scope of this thesis. I believe that it is not an easy task.

Chapter 4

Maximum k-Subset Intersection

In this chapter, we consider an optimization problem whose goal is to find a maximum k-subset intersection on input a family of subsets over $[n]$. This is equivalent to finding a k-vertex set in a bipartite graph with largest common neighbors, i.e. a maximum biclique with k vertices on its left side. The main result of this chapter is to give a gap-producing fpt-reduction from k-Clique to this problem. The reduction heavily relies upon bipartite graphs with the threshold property. In the introduction, we give a high level overview of the main ideas for this reduction. This is followed by the main result and some of its consequences. Our methodology is to exploit the cardinality gap between the common neighbors of k-vertex set and $(k+1)$-vertex set in graphs satisfying the threshold property. Using the fact that a k-vertex subgraph is a clique if and only if its edge number is equal to s, we are able to encode the k-clique in an input graph G by s subsets with large intersection size in the set family $\mathcal{F}$. Finally, we end this chapter by giving some related open questions.

4.1 Introduction

Given a collection $\mathcal{F} = \{S_1, S_2, \dots, S_n\}$ of subsets over $[m]$, where $m = n^{O(1)}$. The goal of Maximum k-Subset Intersection is to select k distinct subsets from $\mathcal{F}$ such that their intersection size is as large as possible.

Max-k-Subset-Intersection	
Instance:	k and a set family $\mathcal{F} = \{S_1, S_2, \dots, S_n\}$ over $[m]$, where $m = n^{O(1)}$.
Solution:	k distinct subsets $S_{j_1}, S_{j_2}, \dots, S_{j_k}$ from $\mathcal{F}$.
Cost:	the intersection size $\lvert S_{j_1} \cap \dots \cap S_{j_k} \rvert$.
Goal:	max.

Example 4.1.1. *Suppose* $m = 6$, $\mathcal{F} = \{\{1,2,3\},\{2,3,4\},\{2,3,5\},\{2,4,6\}\}$ *and* $k = 3$. *We can see that* $n = 4$. *There are* $\binom{n}{3} = \binom{4}{3} = 4$ *possibilities of choosing* 3 *subsets from* $\mathcal{F}$.

1 $\{1,2,3\},\{2,3,4\},\{2,3,5\}$

2 $\{1,2,3\},\{2,3,4\},\{2,4,6\}$

3 $\{2,3,4\},\{2,3,5\},\{2,4,6\}$

4 $\{1,2,3\},\{2,3,5\},\{2,4,6\}$

The maximum cardinality intersection of three subsets from $\mathcal{F}$ is

$$\{2,3\} = \{1,2,3\} \cap \{2,3,4\} \cap \{2,3,5\}$$

The polynomial time inapproximability of Maximum k-Subset Intersection has been proved in [Xav12] based on the inapproximability of Maximum Edge Biclique [AMS11].

However, it does not rule out approximate algorithms in $f(k) \cdot n^{O(1)}$-time. The main result of this chapter is that we can construct $\mathcal{F}$ and s on input a graph G and a positive integer k in polynomial time such that $s = \binom{k}{2}$ and

- if G contains a subgraph isomorphic to K_k then $\mathcal{F}$ has an s-intersection with size no less than $n^{\Theta(1/k)}$;
- if G contains no K_k-subgraph then every s-intersection of $\mathcal{F}$ has size at most $(k+1)!$.[1]

In other words, unless FPT = W[1], there are no $f(s) \cdot n^{O(1)}$-time algorithms to approximate the Maximum s-Subset Intersection within ratio $o(n^{1/\sqrt{s}})$.

We emphasize that our result of hardness approximation does not rely on the PCP-theorem. Instead, we exploit the gap between the sizes of the common neighbors of k-vertex sets and $(k+1)$-vertex sets in graphs satisfying the threshold property. Here we give a high level overview of the main ideas of our reduction for Maximum s-Subset Intersection. Suppose we can construct a collection of sets $\mathcal{T} = \{S_1, S_2, \dots, S_n\}$ for any positive integers k, n and $h > \ell$ (e.g. $h = n^{1/k}$ and $\ell = (k+1)!$) such that

(T1) every $k+1$ distinct subsets in $\mathcal{T}$ has intersection size at most ℓ;

(T2') every k distinct subsets in $\mathcal{T}$ has intersection size at least h.

Here we use (T2') instead of (T2) to simplify the explanation. Then on input a graph G with $V(G) = [n]$, we construct a new family of subsets $\mathcal{F} := \{S_{\{ij\}} := S_i \cap S_j \mid ij \in E(G)\}$.

If G has a k-clique, say $\{1,2,\dots,k\}$ forms a K_k-subgraph in G, then by (T2') $|\bigcap_{i\in[k]} S_i| \geq h$. It follows that $\{S_{\{ij\}} \mid \{i,j\} \in \binom{[k]}{2}\}$ are $s = \binom{k}{2}$ distinct subsets in $\mathcal{F}$ with intersection cardinality at least h; on the other hand, if G contains no k-clique then every s distinct subsets in $\mathcal{F}$ must come from at least $k+1$ distinct subsets in $\mathcal{T}$, by (T1) these subsets have intersection size at most ℓ.

It remains to construct the set family $\mathcal{T}$ efficiently for some proper parameters k, n, h and ℓ. However, the condition (T2') is too strong. We make a relaxation by considering a family with n' ($n < n' = n^{O(1)}$) subsets and partitioning $[n']$ into n disjoint subsets $[n'] = I_1 \dot\cup I_2, \dots, I_n$. Then we replace (T2') with

[1] However, in this case, there may exist $s-1$ distinct subsets in $\mathcal{F}$ with intersection size $n^{\Theta(1/k)}$.

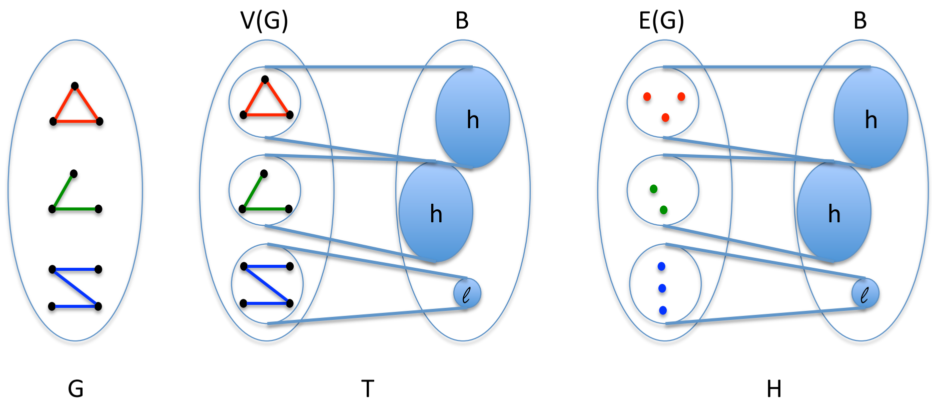

Figure 4.1: Reduction from Finding K_3 to Max 3-Subset Intersection

(T2) for any k distinct indices $a_1, a_2, \ldots, a_k \in [n]$, there exists $b_1 \in I_{a_1}, b_2 \in I_{a_2} \ldots, b_k \in I_{a_k}$ such that $|\bigcap_{i \in [k]} S_{b_i}| \geq h$.

With little effort, we adapt our reduction to the set family satisfying (T1) and (T2). In the previous chapter, we have seen that a bipartite graph in which the family of neighbor sets for all vertices satisfying (T1) and (T2) can be computed in polynomial time. This complete our reduction.

4.2 The Gap-producing Reduction

Theorem 4.2.1 (Main). *We can construct a bipartite graph $H = (A \dot\cup B, E)$ on input an n-vertex graph G and a positive integer k with $\lceil n^{\frac{6}{k+6}} \rceil > (k+6)!$ in $O(n^{18})$-time such that:*

(1) if $K_k \subseteq G$, then there are s vertices in A with at least $\lceil n^{\frac{6}{k'+1}} \rceil$ common neighbors in B;

(2) if $K_k \nsubseteq G$, then every s vertices in A have at most $(k'+1)!$ common neighbors in B.

where $s = \binom{k'}{2}$ and k' is the minimum integer such that $6 \mid k'+1$ and $k' \geq k$.

To prove the main theorem, we need the following reduction lemma.

Lemma 4.2.2 (reduction). *On input an n-vertex simple graph G and an (n, k, ℓ, h)-threshold bipartite graph T, we can construct a new graph $H = (A \dot\cup B, E)$ in $n^{O(c)}$-time, such that:*

(H1) if $K_k \subseteq G$, then $\exists V \in \binom{A}{s}$, $|\Gamma(V)| \geq h$.

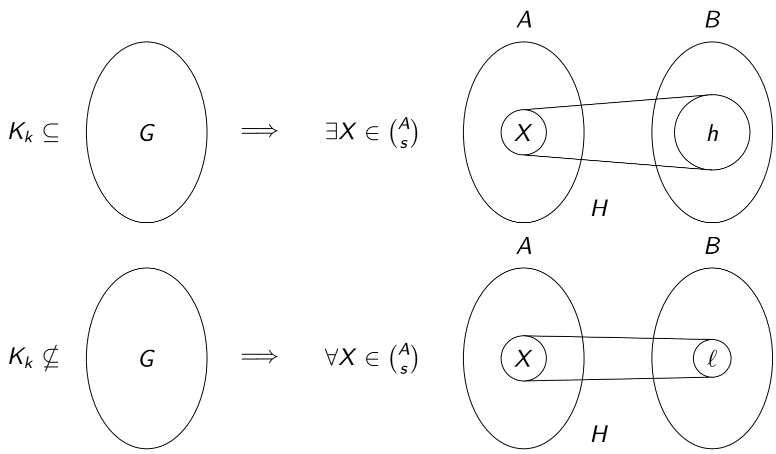

Figure 4.2: The Reduction of Theorem 4.2.1

(H2) if $K_k \nsubseteq G$, then $\forall V \in \binom{A}{s}$, $|\Gamma(V)| \leq \ell$.

where $s := \binom{k}{2}$.

Proof: Suppose $V(G) = [n]$ and $A(T) = V_1 \dot\cup V_2 \dot\cup \ldots \dot\cup V_n$. Our goal is to construct a bipartite graph $H = (A \dot\cup B, E)$ satisfying (H1) and (H2). We associate to each V_i a vertex $i \in V(G)$. Let $\iota : A(T) \to V(G)$ be the function that for each $u \in V_i$, $\iota(u) = i$. Then we construct the bipartite graph $H = (A \dot\cup B, E)$ as follows.

$$A := \{\{u_1, u_2\} \mid u_1, u_2 \in A(T), \{\iota(u_1), \iota(u_2)\} \in E(G)\};$$

$$B := B(T);$$

$$E := \{\{e, v\} \mid \{u_1, u_2\} = e \in A, v \in B, u_1 v \in E(T), u_2 v \in E(T)\}.$$

We show that H satisfies (H1) and (H2):

(1) If $K_k \subseteq G$, let us say $\{a_1, \ldots, a_k\}$ induces a K_k in G. Then by (T2), there exists $u_{a_i} \in V_{a_i} (\forall i \in [k])$ such that $\{u_{a_1}, \ldots, u_{a_k}\}$ has at least h common neighbors in $B(T)$. Let $X := \{u_{a_1}, \ldots, u_{a_k}\}$ and $Y := \Gamma(X)$, we have $|X| = k$ and $|Y| \geq h$. Let $E_X := \binom{X}{2}$. Since $\{\iota(u_{a_i}), \iota(u_{a_j})\} = \{a_i, a_j\} \in E(G)$ for all distinct $i, j \in [k]$, we have $E_X \subseteq A$, hence for all $e \in E_X$ and $v \in Y$, $\{e, v\} \in E$. So $E_X \dot\cup Y$ induces a complete bipartite subgraph in H. It follows that H satisfies (H1) because $|E_X| = \binom{|X|}{2} = \binom{k}{2} = s$ and $|Y| \geq h$;

(2) If there exists $V \in \binom{A}{s}$ with $|\Gamma(V)| \geq \ell + 1$, we need to show $K_k \subseteq G$. Let $E_X := V$, $Y := \Gamma(V)$. We have $|E_X| = s$ and $|Y| \geq \ell + 1$. Consider $X := \{u \in A(T) \mid \exists e \in E_X \, u \in e\}$. By the definition of the edge set E, in the graph T, $Y \subseteq \Gamma(X)$. Since $|Y| = \ell + 1$ and $\forall X \in \binom{A(T)}{k+1}$, $|\Gamma(X)| \leq \ell$, we have $|X| \leq k$; on the other hand, it is not hard to see that $E_X \subseteq \binom{X}{2}$, hence $|E_X| = \binom{k}{2}$ implies $|X| > k - 1$. Thus $|X| = k$ and for any distinct $u_1, u_2 \in X$, $\{u_1, u_2\} \in A$. Recall that $\{u_1, u_2\} \in A \iff \{\iota(u_1), \iota(u_2)\} \in E(G)$ and G is simple. It follows that $\{\iota(u) \mid u \in X\}$ induces a K_k in G.

□

Remark 4.2.3. *Inspection of the proof of Lemma 4.2.2 shows that the graph $H = (A \dot\cup B, E)$ created by our reduction also satisfies the following property. For $0 < \beta < \alpha \leq 1$.*

- *if G contains a k-vertex subgraph with $\alpha \cdot s$ edges, then $\exists V \in \binom{A}{\alpha s}$, $|\Gamma(V)| \geq h$.*
- *if every k-vertex subgraph of G has at most $\beta \cdot s$ edges, then $\forall V \in \binom{A}{\beta s+1}$, $|\Gamma(V)| \leq \ell$.*

In other words, if there is no fpt*-algorithm to distinguish between graphs with a k-vertex subgraph containing αs edges and graphs without k-vertex subgraph containing $(\beta s + 1)$ edges, then k-*Biclique *and hence k-*Clique *cannot be* fpt*-approximated to ratio $\frac{\alpha}{\beta}$.*

Lemma 4.2.2 can be extended as follows. For $\Delta : \mathbb{N} \to \mathbb{N}^+$, we duplicate every vertex of A in the bipartite graph H $\Delta(k)$ times. Then we obtain a more general reduction lemma.

Lemma 4.2.4. *For $\Delta : \mathbb{N} \to \mathbb{N}^+$. On input an n-vertex simple graph G and an (n, k, ℓ, h)-threshold bipartite graph T, we can construct a new graph $H = (A \dot\cup B, E)$ in $\Delta(k) \cdot n^{O(1)}$-time, such that:*

- *if $K_k \subseteq G$, then $\exists V \in \binom{A}{\Delta(k)\cdot s}$, $|\Gamma(V)| \geq h$.*
- *if $K_k \nsubseteq G$, then $\forall V \in \binom{A}{\Delta(k)\cdot s}$, $|\Gamma(V)| \leq \ell$.*

Now we are ready to prove Theorem 4.2.1.

Proof: [of Theorem 4.2.1] Given G and k, let k' be the minimum integer such that $k' \geq k$ and $6 \mid k' + 1$, we have $k' \leq k + 5$. Then we add a new clique with $k' - k$ vertices into G and connect them with every vertex in G. It is easy to see that the new graph contains a k'-clique if and only if G contains a k-clique. Since $\lceil n^{\frac{6}{k+6}} \rceil > (k + 6)!$, we have $\lceil n^{\frac{6}{k'+1}} \rceil > (k' + 1)!$. Apply Lemma 3.1.4 on n and k', we obtain an $(n, k', (k' + 1)!, \lceil (n + 1)^{\frac{6}{k'+1}} \rceil)$-threshold bipartite graph. The result then follows from Lemma 4.2.2. □

From Theorem 4.2.1, we obtain an inapproximation result for the Maximum k-Subset Intersection Problem immediately.

Corollary 4.2.5. *Assuming* FPT $\neq$ W[1]*, there is no* $f(k) \cdot n^{O(1)}$*-time algorithm approximating Maximum k-Subset Intersection within* n^{ε}*-approximation ratio for* $\varepsilon < \frac{6}{\sqrt{k}+1}$.

To see that Theorem 4.2.1 implies the W[1]-hardness of k-Biclique. Let $t = (k' + 1)! + 1$. We add $(t - s)$ vertices to H and make them adjacent to every vertex in B. It is easy to check that the resulting graph contains a $K_{t,t}$ if and only if the original graph G contains a K_k.

Corollary 4.2.6. k-Biclique *is* W[1]*-hard.*

4.2.1 Lower Bounds under ETH

More refined lower bounds can be obtained if we assume ETH. From Theorem 2.2.2 and Theorem 4.2.1, we can deduce

Corollary 4.2.7. *Assuming* ETH*, there is no* $f(k) \cdot n^{o(\sqrt{k})}$*-time algorithm approximating Maximum k-Subset Intersection within* n^{ε}*-approximation ratio for* $\varepsilon < \frac{6}{\sqrt{k}+1}$.

Using Lemma 4.2.4, we obtain a inapproximability result with a trade-off between the running time of the algorithm and the approximation ratio.

Corollary 4.2.8. *Assuming* ETH*, for every constant* $c \geq 2$ *there is no* $f(k) \cdot n^{o(\sqrt[c]{k})}$*-time algorithm approximating Maximum k-Subset Intersection within* n^{ε}*-approximation ratio for* $\varepsilon < \frac{6}{\sqrt[c]{k}+1}$.

An interesting question is to find a *linear* fpt-reduction from k-Clique to k-Biclique, that is given G and k, computing a new graph G' in $f(k) \cdot n^{O(1)}$-time such that $K_k \subseteq G$ if and only if $K_{k',k'} \subseteq G'$, where $k' = ck$ for some constant c. The existence of such reduction would imply that k-Biclique has no $f(k) \cdot n^{o(k)}$-time algorithm under the ETH. However, since our reduction causes a quadratic blow-up of the size of solution, $k' = \binom{k}{2}$ is the best we may achieve. We note that by Theorem 4.2.1, we can get $k' = \Omega(k!)$. Using the probabilistic method, we have:

Theorem 4.2.9. *For any n-vertex graph G and positive integer k with* $n \geq \max\{2(k+1)^2, 20\}$*,* $\ell = 2k^2 + 4k - 1 < n^{\frac{1}{4(k+1)}}$*, we can construct a random graph* G' *in polynomial time such that, with probability at least* $\frac{9}{10}$*,* G' *contains a* $K_{\ell+1,\ell+1}$ *if and only if* G *contains a* K_k.

We can see Theorem 4.2.9 follows from Lemma 3.1.3.

Consider a randomized version of ETH which states that there are no randomized algorithms with two-sided error such that on input an instance of 3-**SAT** decide if it is satisfiable or not correctly with probability larger than 1/2 in $2^{o(n)}$-time. For more detail we refer to [CFK+16]. Then Theorem 4.2.9 yields a better lower bound for k-Biclique:

Corollary 4.2.10. *Under the randomized* ETH*, there is no* $f(k) \cdot n^{o(\sqrt{k})}$ *algorithm to decide whether a given graph contains a subgraph isomorphic to* $K_{k,k}$.

4.3 Open Questions

In Corollary 4.2.8, we obtained a hardness result with a trade-off between the approximation ratio and the running-time of the algorithm for MAX-k-SUBSET-INTERSECTION. This suggests a question.

Question 4.3.1. *Given two function $h, \ell : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ with $h(n,k) \geq \ell(n,k)$, what is the minimum value $\delta(k)$ such that there is an $f(k) \cdot n^{\delta(k)}$-time algorithm that on input an n-vertex bipartite graph $G = (A \dot\cup B, E)$ distinguishes*

- *whether there exists $S \in \binom{A}{k}$ such that $|\Gamma(S)| \geq h(n,k)$;*
- *or every $S \in \binom{A}{k}$ has $|\Gamma(S)| \leq \ell(n,k)$.*

The following question is relevant.

Question 4.3.2. *Fix two function $h, \ell : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ with $h(n,k) \geq \ell(n,k)$, what is the minimum value of $\delta(k) := \frac{k\varepsilon_1(k)}{\delta_1(k)}$ such that $\textbf{lcount}(G, K_{\delta_1(k),h(n,k)}) > n^{\varepsilon_1(k)}$ implies $\textbf{lcount}(G, K_{k,\ell(n,k)}) > 0$?*

The minimum value of $\delta(k)$ is linked to the running time of the algorithm in Question 4.3.1. We can enumerate every $\delta_1(k)$-vertex subset of $A(G)$ with $h(n,k)$ common neighbors. If the number of such subsets exceed $n^{\varepsilon_1(k)}$, then there are k vertices in $A(G)$ with more than $\ell(n,k)$ common neighbors. Otherwise, we only need to check $n^{\frac{k\varepsilon_1(k)}{\delta_1(k)}} = n^{\delta(k)}$ cases.

In particular, by Lemma 6.3.4, we have that for $0 < \varepsilon < 1$ if

$$\textbf{lcount}(G, K_{1,n^\varepsilon}) \cdot n^\varepsilon > \textbf{lcount}(G, K_{1,n^\varepsilon})^{1-1/t} n + t\textbf{lcount}(G, K_{1,n^\varepsilon}),$$

then $\textbf{lcount}(G, K_{2,t}) > 0$. For instance, suppose $(1-\varepsilon)t < 1$ and $n^\varepsilon \gg t$, then we obtain

$$\textbf{lcount}(G, K_{1,n^\varepsilon}) > n^{(1-\varepsilon)t} \Rightarrow \textbf{lcount}(G, K_{2,t}) > 0$$

It is possible to distinguish graphs containing two vertices with n^ε common neighbors and graphs without $K_{2,t+1}$-subgraph in time $n^{2(1-\varepsilon)t} = o(n^2)$.

The following question aims to derandomize Theorem 4.2.9.

Question 4.3.3. *Give an explicit construction of graphs satisfying the (n, k, k^2, k^2+1)-threshold property in $f(k) \cdot n^{O(1)}$-time.*

A more aggressive one is

Question 4.3.4. *Give an explicit construction of graphs satisfying the $(n, k, k^2, n^{\Theta(1/k)})$-threshold property in $f(k) \cdot n^{O(1)}$-time.*

According to the construction in Chapter 5, we can see that solving Question 4.3.4 leads to a better inapproximability result of MIN-DOMINATING-SET under ETH.

Note that even if the bipartite graphs with (n, k, k^2, k^2+1)-threshold property can be constructed in deterministic fpt time, we could only show that k-BICLIQUE has no $f(k) \cdot n^{o(\sqrt{k})}$ algorithm under ETH, since our reduction causes a quadratic blow-up of the parameter. The following question is still open.

Question 4.3.5. *Is there any $f(k) \cdot n^{o(k)}$-time algorithm solving k-BICLIQUE?*

A possible way to avoid such quadratic blow-up of the parameter is to do reduction from the Partition Subgraph Isomorphism, in which the number of edge is treated as parameter [Mar07]. However, we can only reduce the Partition Subgraph Isomorphism of a smaller graph G with v-vertex to the k-BICLIQUE problem with $k = \binom{v}{2}$. The hardness result in [Mar07] states that if Partitioned Subgraph Isomorphism can be solved in $f(G) \cdot n^{o(|E(G)|/\log|E(G)|)}$, then ETH fails. In this statement, $|E(G)| = \Theta(|V(G)|)$, we still can not avoid the quadratic blow-up of parameter.

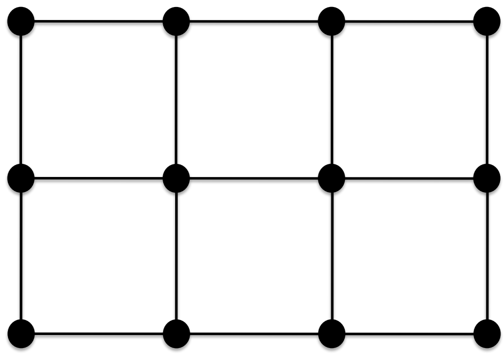

Figure 4.3: A 3×4 grid.

We have shown the W[1]-hardness of k-BICLIQUE. Can we extend our technique to other graph classes with unbounded tree-width? Another natural graph class with unbounded treewidth is *grid*.

Definition 4.3.6 ($G_{r\times c}$)**.** *For $c, r \in \mathbb{N}^+$, a $r \times c$-grid is a graph $G := (V, E)$ with*

- *$V := [r] \times [c]$;*
- *For $(i, j), (i', j') \in V$, $\{(i, j), (i', j')\} \in E$ iff $|i - i'| + |j - j'| = 1$.*

k-GRID	
Instance:	$k \in \mathbb{N}^+$ and a graph G.
Parameter:	k.
Problem:	Decide if G contains a subgraph isomorphic to a $G_{k\times k}$.

Question 4.3.7. *Is there any* $f(k) \cdot n^{O(1)}$*-time algorithm solving* k-GRID*?*

Chapter 5

Minimum Dominating Set

The main result of this chapter is to prove the constant inapproximability of the parameterized minimum dominating set problem. To help readability, we first prove that the dominating set problem is not fpt approximable with ratio smaller than 3/2 in Section 5.2. In the case of the clique problem, once we have inapproximability for a particular constant ratio, it can be easily improved to any constant by gap-amplification via graph products. But dominating sets for general graph products are notoriously hard to understand (see e.g. [KZ96]). So to prove the main result, Section 5.3 presents a modified reduction which contains a tailor-made graph product. Section 5.4 discusses some consequences of our results. We conclude in Section 5.5.

5.1 Introduction

The dominating set problem, or equivalently the set cover problem, was among the first problems proved to be NP-hard [Kar72]. Moreover, it has been long known that the greedy algorithm achieves an approximation ratio $\approx \ln n$ [Joh74, Ste74, Lov75, Chv79, Sla97]. And after a sequence of papers (e.g. [LY94, RS97, Fei98, AMS06, DS14]), this is proved to be best possible. In particular, Raz and Safra [RS97] showed that the dominating set problem cannot be approximated with ratio $c \cdot \log\, n$ for some constant $c \in \mathbb{N}$ unless P = NP [RS97]. Under a stronger assumption NP $\not\subseteq$ DTIME $\left(n^{O(\log\log n)}\right)$ Feige proved that no approximation within $(1-\varepsilon)\ln n$ is feasible [Fei98]. Finally Dinur and Steuer established the same lower bound assuming only P $\neq$ NP [DS14]. However, it is important to note that the approximation ratio $\ln n$ is measured in terms of the size of an input graph G, instead of $\gamma(G)$, i.e., the size of its minimum dominating set. As a matter of fact, the standard examples for showing the $\Theta(\log\, n)$ greedy lower bound have constant-size dominating sets. Thus, the size of the greedy solutions cannot be bounded by any function of $\gamma(G)$. So the question arises whether there is an approximation algorithm $\mathbb{A}$ that always outputs a dominating set whose size can be bounded by $\rho(\gamma(G)) \cdot \gamma(G)$, where the function $\rho : \mathbb{N} \to \mathbb{N}$ is known as the approximation ratio of $\mathbb{A}$. The constructions in [Fei98, AMS06] indeed show that we can rule out $\rho(x) \leq \ln x$. To the best of our knowledge, it is not known whether this bound is tight. For instance, it is still conceivable that there is a polynomial time algorithm that always outputs a dominating set of size at most $2^{2^{\gamma(G)}}$.

Other than looking for approximate solutions, parameterized complexity [DF99, FG06, Nie06, DF13, CFK$^+$16] approaches the dominating set problem from a different perspective. With the expectation that in practice we are mostly interested in graphs with relatively small dominating sets, algorithms of running time $2^{\gamma(G)} \cdot |G|^{O(1)}$ can still be considered efficient. Unfortunately, it turns out that the parameterized dominating set problem is complete for the second level of the so-called W-hierarchy [DF95], and thus fixed-parameter intractable unless FPT = W[2]. So one natural follow-up question is whether the problem can be approximated in fpt-time. More precisely, we aim for an algorithm with running time $f(\gamma(G)) \cdot |G|^{O(1)}$ which always outputs a dominating set of size at most $\rho(\gamma(G)) \cdot \gamma(G)$. Here, $f : \mathbb{N} \to \mathbb{N}$ is an arbitrary computable function. The study of parameterized approximability was initiated in [CH10, CGG07, DFM06]. Compared to the classical polynomial time approximation, the area is still in its very early stage with few known positive and even less negative results.

5.1.1 Our Results

We prove that any constant-approximation of the parameterized dominating set problem is W[1]-hard.

Theorem 5.1.1. *For any constant $c \in \mathbb{N}$ there is* no fpt-*algorithm $\mathbb{A}$ such that on every input graph G the algorithm $\mathbb{A}$ outputs a dominating set of size at most $c \cdot \gamma(G)$, unless* FPT = W[1] *(which implies that the exponential time hypothesis (*ETH*) fails).*

In the above statement, clearly we can replace "fpt-algorithm" by "polynomial time algorithm," thereby obtaining the classical constant-inapproximability of the dominating set problem. But let us mention that our result is not comparable to the classical version, even if we restrict ourselves to polynomial time tractability. The assumption FPT $\neq$ W[1] or ETH is apparently much stronger than P $\neq$ NP, and in fact ETH implies NP $\not\subseteq$ DTIME $\left(n^{O(\log\log n)}\right)$ used in aforementioned Feige's result. But on the other hand, our lower bound applies even in case that we know in advance that a given graph has no large dominating set.

Corollary 5.1.2. *Let $\beta : \mathbb{N} \to \mathbb{N}$ be a nondecreasing and unbounded computable function. Consider the following promise problem.*

MIN-DOMINATING-SET$_\beta$	
Instance:	A graph $G = (V, E)$ with $\gamma(G) \leq \beta(\lvert V \rvert)$.
Solution:	A dominating set D of G.
Cost:	$\lvert D \rvert$.
Goal:	min.

Then MIN-DOMINATING-SET$_\beta$ *does not admit a polynomial time constant approximation algorithm, unless* FPT = W[1].

The proof of Theorem 5.1.1 is crucially built on Theorem 4.2.1. We exploit the gap created in its hardness reduction. In the known proofs of the classical inapproximability of the dominating set problem, one always needs the PCP theorem in order to have such a gap, which makes those proofs highly non-elementary. More importantly, it can be verified that reductions based on the PCP theorem produce instances with optimal solutions of relatively large size, e.g., a graph $G = (V, E)$ with $\gamma(G) \geq |V|^{\Theta(1)}$. This is inevitable, since otherwise we might be able to solve every NP-hard problem in sub-exponential time. As an example, if it is possible to reduce an NP-hard problem to the approximation of $\textsc{Min-Dominating-Set}_\beta$ for $\beta(n) = \log\log\log n$, then by brute-force searching for a minimum dominating set, we are able to solve the problem in time $n^{O(\log\log\log n)}$. It implies $\mathrm{NP} \subseteq \mathrm{DTIME}\left(n^{O(\log\log\log n)}\right)$. Because of this, Corollary 5.1.2, and hence also Theorem 5.1.1, is unlikely provable following the traditional approach.

Using a result of Chen et.al. [CHKX04] the lower bound in Theorem 5.1.1 can be further sharpened.

Theorem 5.1.3. *Assume* ETH *holds. Then there is* no fpt-*algorithm which on every input graph G outputs a dominating set of size at most* $\sqrt[4+\varepsilon]{\log(\gamma(G))} \cdot \gamma(G)$ *for every* $0 < \varepsilon < 1$.

Related Work

The existing literature on the dominating set problem is vast. The most relevant to our work is the classical approximation upper and lower bounds as explained in the beginning. But as far as the parameterized setting is concerned, what was known is rather limited.

Downey et al. proved that there is no *additive* approximation of the the parameterized dominating set problem [DFMR08]. In the same paper, they also showed that the independent dominating set problem has no fpt approximation with any approximation ratio. Recall that an independent dominating set is a dominating set which is an independent set at the same time. With this additional requirement, the problem is no longer *monotone*, i.e., a superset of a solution is not necessarily a solution. Thus it is unclear how to reduce the independent dominating set problem to the dominating set problem by an approximation-preserving reduction.

In [CHK13, HKK13] it is proved under ETH that there is no $c\sqrt{\log\gamma(G)}$-approximation algorithm for the dominating set problem[1] with running time $2^{O(\gamma(G)^{(\log\gamma(G))^d})}|G|^{O(1)}$, where c and d are some appropriate constants. With the additional *Projection Game Conjecture* due to [Mos12] and some of its further strengthening, the authors of [CHK13, HKK13] are able to even rule out $\gamma(G)^c$-approximation algorithms with running time almost doubly exponential in terms of $\gamma(G)$. Clearly, these lower bounds are against far better approximation ratio than those of Theorem 5.1.1 and Theorem 5.1.3, while the drawback is that the dependence of the running time on $\gamma(G)$ is not an arbitrary computable function.

[1]The papers actually address the set cover problem, which is equivalent to the dominating set problem as mentioned in the beginning.

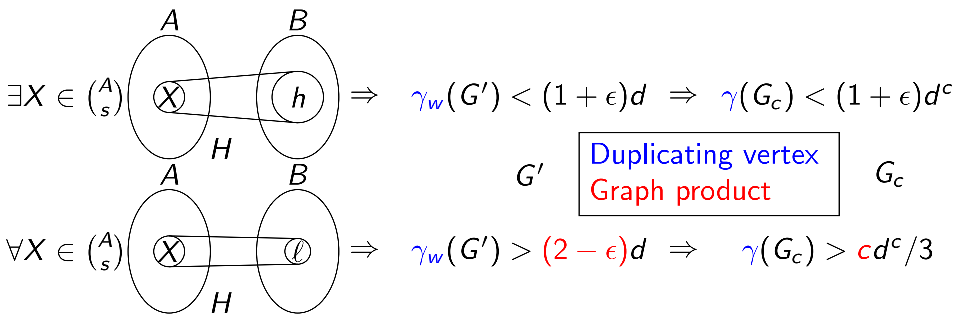

Figure 5.1: Overview of the Reduction

The dominating set problem can be understood as a special case of the weighted satisfiability problem of CNF-formulas, in which all literals are positive. The weighted satisfiability problems for various fragments of propositional logic formulas, or more generally circuits, play very important roles in parameterized complexity. In particular, they are complete for the W-classes. In [CGG07] it is shown that they have no fpt approximation of any possible ratio, again by using the non-monotonicity of the problems. Marx strengthened this result significantly in [Mar13] by proving that the weighted satisfiability problem is not fpt approximable for circuits of depth 4 without negation gates, unless FPT = W[2]. Our result can be viewed as an attempt to improve Marx's result to depth-2 circuits, although at the moment we are only able to rule out fpt approximations with constant ratio.

5.1.2 Overview of Our Reduction

To give a brief overview of our reduction, let us consider the weighted version of the minimum dominating set problem.

MIN WEIGHTED DOMINATING SET	
Instance:	A graph G and $w : V(G) \to \mathbb{R}$.
Solution:	$D \subseteq V(G)$ is a dominating set of G.
Cost:	$\sum_{v \in D} w(v)$.
Goal:	min.

For a given weight function w, let $\gamma_w(G)$ be the minimum value of weighted dominating set for G.

The intuition of our reduction is illustrated in Figure 5.1. First, we provide an fpt-reduction from MAX-k-SUBSET-INTERSECTION to minimum weighted dominating set which creates a gap close to 2. Then we improve this reduction using two methods: duplicating the vertices and taking graph product.

In Section 5.2, we obtain a 3/2-gap fpt-reduction for the unweighted version by duplicating the vertices. In Section 5.3, we increase the gap to any constant using the method of graph product.

We end this section with an informal description of the fpt-reduction for minimum weighted dominating set which creates a gap close to 2.

First, we choose a large $d \in \mathbb{N}^+$ with $\varepsilon\sqrt{d} > \ell$ and partition B into d disjoint subsets. Then for each $i \in [d]$, we introduce a set of vertices W_i and add edges between A and W_i, W_i and B_i such that there are two ways to dominate the vertices in W_i:

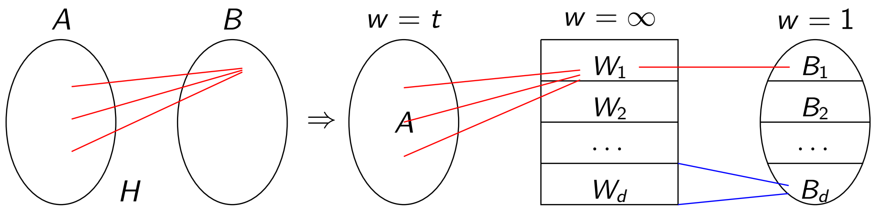

Figure 5.2: 2-gap fpt-reduction for Min Weighted Dominating Set

– Using one vertex in B_i and its s neighbors in A.

– Using two vertices in B_i.

Finally, we choose $t \in \mathbb{N}^+$ with $t = d^{1-\frac{1}{2s}}$ and give the vertices in A weight t. On the other hand, we assign to all the vertices in W infinite weight and to vertices in B weight 1.

If $\exists X \in \binom{A}{s}\ \Gamma(X) \geq d$, then we can choose this s-vertex set X and its d common neighbors as the dominating set. The weight of this dominating set is $st + d$. Choosing d large enough, we have that $st + d \leq (1 + \varepsilon)d$ by $t = o(d)$.

If $\forall X \in \binom{A}{s}\ \Gamma(X) \leq \ell$. Let D be the dominating set. Then

– Either there are $(1 - \varepsilon)$ fraction of $i \in [d]$, $|D \cap B_i| \geq 2$. In this case, the weight of D is $w(D) \geq (1 - \varepsilon)2d \approx 2d$;

– Or for εd distinct $i \in [d]$, we use a vertex $v_i \in B_i \cap D$ and its s neighbors in $D \cap A$ to dominate W_i. Assume that $w(D) \leq 2d$, otherwise we are done. It follows that $|D \cap A| \leq 2d/t$. We have that there are εd vertices in $D \cap B$, each having at least s neighbors in $|D \cap A| \leq 2d/t = O(d^{1/2s})$. There are at most $(2d/t)^s = O(d^{1/2})$ distinct s-tuples of $D \cap A$. By the pigeonhole principle, at least $\varepsilon d^{1/2} > \ell$ vertices in $D \cap B$ must be adjacent to the same s-tuples of $D \cap A$. This contradicts to the fact that every s-vertex set in A has at most ℓ common neighbors in B.

5.2 The Case $\rho < 3/2$

As the first illustration of how to use the gap created in Theorem 4.2.1, we show in this section that k-Dominating-Set cannot be fpt approximated within ratio $< 3/2$. This serves as a stepping stone to the general constant-inapproximability of the problem.

Theorem 5.2.1. *Let $\rho < 3/2$. Then there is no* fpt *approximation of the parameterized dominating set problem achieving ratio ρ unless* FPT = W[1].

Proof: We fix some $\varepsilon, \delta \in \mathbb{R}$ with $0 < \varepsilon < 1$, $0 < \delta < 1/2$, and

$$\frac{3/2 - \delta}{1 + \varepsilon} > \rho. \tag{5.1}$$

Let G be a graph with n vertices and $k \in \mathbb{N}$ a parameter. We set $s := \binom{k}{2}$,

$$d := \left\lceil \frac{s}{\varepsilon} \right\rceil^{2s}, \quad \text{and} \quad t := \left\lceil \left(\frac{1}{2} - \delta\right) \cdot d^{1-1/2s} \right\rceil.$$

As a consequence, when k and n are sufficiently large, we have

$$st < \varepsilon d, \quad \left(\frac{1}{2} - \delta\right) \cdot \frac{d}{t} \le \sqrt[2s]{d}, \quad (k+1)! < 2\delta\sqrt{d} - 1, \quad \text{and} \quad d \le \lceil n^{\frac{6}{k+1}} \rceil. \tag{5.2}$$

By Theorem 4.2.1 (and the preprocessing) we can compute in fpt-time a bipartite graph $H_0 = (A_0 \dot\cup B_0, E_0)$ such that:

- if $K_k \subseteq G$, then there are s vertices in A_0 with d common neighbors in B_0;
- if $K_k \nsubseteq G$, then every s vertices in A_0 have at most $(k+1)!$ common neighbors in B_0.

Then using the color-coding in Lemma 2.6.1, again in fpt-time, we construct two function families $\Lambda_A := \Lambda_{|A_0|,s}$ and $\Lambda_B := \Lambda_{|B_0|,d}$ such that

- for every s-element subset $X \subseteq A_0$ there is an $h \in \Lambda_A$ with $h(X) = [s]$;
- for every d-element subset $Y \subseteq B_0$ there is an $h \in \Lambda_B$ with $h(Y) = [d]$.

Define the bipartite graph $H = \Big(A(H) \dot\cup B(H), E(H)\Big)$ by

$$A(H) := A_0 \times \Lambda_A \times \Lambda_B, \qquad B(H) := B_0 \times \Lambda_A \times \Lambda_B$$
$$E(H) := \left\{ \Big\{(u, h_1, h_2), (v, h_1, h_2)\Big\} \,\middle|\, u \in A_0,\ v \in B_0,\ h_1 \in \Lambda_A,\ h_2 \in \Lambda_B, \text{ and } \{u, v\} \in E_0 \right\}.$$

Moreover, define two colorings $\alpha : A(H) \to [s]$ and $\beta : B(H) \to [d]$ by

$$\alpha(u, h_1, h_2) := h_1(u) \quad \text{and} \quad \beta(v, h_1, h_2) := h_2(v).$$

It is straightforward to verify that

(H1) if $K_k \subseteq G$, then there are s vertices of distinct α-colors in $A(H)$ with d common neighbors of distinct β-colors in $B(H)$;

(H2) if $K_k \nsubseteq G$, then every s vertices in $A(H)$ have at most $(k+1)!$ common neighbors in $B(H)$.

Now from H, α, and β we construct a new graph $G' = \Big(V(G'), E(G')\Big)$ as follows. First, its vertex set is defined by

$$V(G') := B(H) \dot\cup \Big\{x_i, y_i \,\Big|\, i \in [d]\Big\} \dot\cup C \dot\cup W,$$

where

$$C := A(H) \times [t] \quad \text{and} \quad W := \left\{ w_{b,j,i} \,\middle|\, b \in B(H), i \in [t], j \in [s] \right\}.$$

Moreover, G' contains the following types of edges.

(E1) $\{b, b'\} \in E(G')$ with $b, b' \in B(H)$, $b \neq b'$, and $\beta(b) = \beta(b')$ $\Big($i.e., all vertices in $B(H)$ with the same color under β form a clique in $G'\Big)$.

(E2) Let $b \in B(H)$ and $c := \beta(b)$. Then $\{x_c, b\}, \{y_c, b\} \in E(G')$.

(E3) Let $b, b' \in B(H)$ with $\beta(b) = \beta(b')$ and $b \neq b'$. Then $\Big\{w_{b,j,i}, b'\Big\} \in E(G')$ for every $i \in [t]$ and $j \in [s]$.

(E4) $\Big\{(a, i), w_{b,j,i}\Big\} \in E(G')$ for every $\{a, b\} \in E(H)$, $j = \alpha(a)$ and $i \in [t]$.

(E5) Let $a, a' \in A(H)$ with $a \neq a'$ and $i \in [t]$. Then $\Big\{(a, i), (a', i)\Big\} \in E(G')$.

To ease presentation, for every $c \in [d]$ we set

$$B_c := \Big\{b \in B(H) \;\Big|\; \beta(b) = c\Big\} \cup \{x_c, y_c\}.$$

Claim 1. If D is a dominating set of G', then $D \cap B_c \neq \emptyset$ for every $c \in [d]$.

Proof of the claim. We observe that every x_c is only adjacent to vertices in B_c. $\dashv$

Claim 2. If G contains a k-clique, then $\gamma(G') < (1 + \varepsilon)d$.

Proof of the claim. By (H1) the bipartite graph H has a $K_{s,d}$ biclique K with $\alpha(A(H) \cap K) = [s]$ and $\beta(B(H) \cap K) = [d]$. It is then easy to verify that

$$\Big(B(H) \cap K\Big) \,\dot\cup\, \Big((A(H) \cap K) \times [t]\Big)$$

is a dominating set of G', whose size is $d + s \cdot t < (1 + \varepsilon)d$ by (5.2). $\dashv$

Claim 3. If G contains no k-clique, then every s-vertex set of $A(H)$ has at most $(k+1)! < 2\delta\sqrt{d} - 1$ common neighbors in $B(H)$.

Claim 4. If G contains no k-clique, then

$$\gamma(G') > \left(\frac{3}{2} - \delta\right) \cdot d.$$

Proof of the claim. Let D be a dominating set of G'. By Claim 1 we have $D \cap B_c \neq \emptyset$ for every $c \in [d]$. Define

$$e := \left|\Big\{c \in [d] \;\Big|\; |D \cap B_c| \geq 2\Big\}\right|.$$

If $e > (1/2 - \delta) \cdot d$ then $|D| > d + e > (3/2 - \delta) \cdot d$ and we are done.

So let us consider $e \leq (1/2 - \delta) \cdot d$ and without loss of generality $|D \cap B_c| = 1$ for every $c \leq (1/2 + \delta) \cdot d$. Fix such a c and assume $D \cap B_c = \{b_c\}$. Recall $x_c, y_c \in V(G')$

are not adjacent to any vertex outside B_c, and there is no edge between them, thus $b_c \in B_c \setminus \{x_c, y_c\} = \Big\{b \in B(H) \;\Big|\; \alpha(b) = c\Big\}$. Let

$$W_1 := \Big\{w_{b_c,j,i} \;\Big|\; i \in [t],\ j \in [s], \text{ and } c \le (1/2 + \delta) \cdot d\Big\} \subseteq W.$$

(E3) implies that every $w_{b_c,j,i} \in W_1$ is not dominated by any vertex in $D \cap \bigcup_{c \in [d]} B_c$. Therefore, it has to be dominated by or included in $D \cap (C \cup W)$.

If $|D \cap W_1| > (1/2 - \delta) \cdot d$, then again we are done. So suppose $|D \cap W_1| \le (1/2 - \delta) \cdot d$. Without loss of generality let

$$W_2 := \Big\{w_{b_c,j,i} \;\Big|\; i \in [t],\ j \in [s], \text{ and } c \le 2\delta d\Big\} \subseteq W_1$$

and assume $W_2 \cap D = \emptyset$. Thus W_2 has to be dominated by $D \cap C$. For later purpose, let

$$Y := \Big\{b_c \;\Big|\; c \le 2\delta d\Big\}.$$

Obviously, $|Y| \ge 2\delta d - 1$.

Again we only need to consider the case $|D \cap C| \le (1/2 - \delta) \cdot d$. Recall $C = A(H) \times [t]$. Thus there is an $i \in [t]$ such that

$$\Big|D \cap \Big(A(H) \times \{i\}\Big)\Big| \le \Big(\frac{1}{2} - \delta\Big) \cdot \frac{d}{t}.$$

Let $X := \Big\{a \in A(H) \;\Big|\; (a, i) \in D\Big\}$, and in particular, $|X| \le (1/2 - \delta) \cdot d/t$. Since W_2 is dominated by $D \cap C$, we have for all $b \in Y$ and $j \in [s]$ there exists $a \in X$ such that $\Big\{(a, i), w_{b,j,i}\Big\} \in E(G')$, which means that $\{a, b\} \in E(H)$ and $\alpha(a) = j$. It follows that in the graph H every vertex of Y has at least s neighbors in X. Recall that $(1/2-\delta)\cdot d/t \le \sqrt[2s]{d}$ by (5.2). There are at most $\sqrt{d}$ different types of s-vertex sets in X, i.e.,

$$\left|\binom{X}{s}\right| \le \binom{(1/2 - \delta) \cdot d/t}{s} \le \left(\sqrt[2s]{d}\right)^s = \sqrt{d}.$$

By the pigeonhole principle, there exists an s-vertex set of $X \subseteq A(H)$ having at least $|Y|/\sqrt{d} \ge 2\delta\sqrt{d} - 1$ common neighbors in $Y \subseteq B(H)$, which contradicts Claim 3. $\dashv$.

Claim 2 and Claim 4 indeed imply that there is an fpt-reduction from the clique problem to the dominating set problem which creates a gap great than

$$\frac{3/2 - \delta}{1 + \varepsilon}.$$

So if there is a ρ-approximation of the dominating set problem, by (5.1) we can decide the clique problem in fpt time. $\square$

5.3 The Constant-Inapproximibility of k-Dominating-Set

Theorem 5.1.1 is a fairly direct consequence of the following theorem.

Theorem 5.3.1 (Main)**.** *There is an algorithm $\mathbb{A}$ such that on input a graph G, $k \geq 3$, and $c \in \mathbb{N}$ the algorithm $\mathbb{A}$ computes a graph G_c such that*

(i) if $K_k \subseteq G$, then $\gamma(G_c) < 1.1 \cdot d^c$;

(ii) if $K_k \nsubseteq G$, then $\gamma(G_c) > c \cdot d^c/3$,

where $d = (30 \cdot c^2 \cdot (k+1)^2)^{4 \cdot k^3 + 3c}$. Moreover the running time of $\mathbb{A}$ is bounded by $f(k, c) \cdot |G|^{O(c)}$ for a computable function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$.

Proof: [of Theorem 5.1.3] Suppose for some $\varepsilon > 0$ there is an fpt-algorithm $\mathbb{A}(G)$ which outputs a dominating set for G of size at most $\sqrt[4+\varepsilon]{\log(\gamma(G))} \cdot \gamma(G)$. Of course we can further assume that $\varepsilon < 1$. Then on input a graph G and $k \in \mathbb{N}$, let

$$c := \left\lceil k^{1-\varepsilon/5} \right\rceil = o(k) \quad \text{and} \quad d := \left(30 \cdot c^2 \cdot (k+1)^2\right)^{4 \cdot k^3 + 3c}.$$

We have

$$\sqrt[4+\varepsilon]{\log(1.1 \cdot d^c)} = O\left(\sqrt[4+\varepsilon]{c \cdot k^3 \cdot \log k}\right) = o\left(k^{\frac{4}{4+\varepsilon}}\right) = o(c).$$

By Theorem 5.3.1, we can construct a graph G_c with properties (i) and (ii) in time

$$f(k, c) \cdot |G|^{O(c)} = h(k) \cdot |G|^{o(k)}$$

for an appropriate computable function $h : \mathbb{N} \to \mathbb{N}$. Thus, G contains a clique of size k if and only if $\mathbb{A}(G_c)$ returns a dominating set of size at most

$$1.1 \cdot d^c \cdot \sqrt[4+\varepsilon]{\log(1.1 \cdot d^c)} = o(c \cdot d^c) < \frac{c \cdot d^c}{3},$$

where the inequality holds for sufficiently large k $\Big($and hence sufficiently large $c \cdot d^c\Big)$.

Therefore we can determine whether G contains a k-clique in time $g(k) \cdot |G|^{o(k)}$ for some computable $g : \mathbb{N} \to \mathbb{N}$. This contradicts a result in Chen et.al. [CHKX04, Theorem 4.4] under ETH. $\square$

5.3.1 Proof of Theorem 5.3.1

We start by showing a variant of Theorem 4.2.1.

Theorem 5.3.2. *Let $\Delta \in \mathbb{N}^+$ be a constant and $d : \mathbb{N}^+ \to \mathbb{N}^+$ a computable function. Then there is an* fpt*-algorithm that on input a graph G and a parameter $k \in \mathbb{N}$ with $6 \mid k+1$ constructs a bipartite graph $H = \Big(A(H) \,\dot\cup\, B(H), E(H)\Big)$ together with two colorings*

$$\alpha : A(H) \to [\Delta s] \quad \textit{and} \quad \beta : B(H) \to [d(k)]$$

such that:

(H1) if $K_k \subseteq G$, then there are Δs vertices of distinct α-colors in $A(H)$ with $d(k)$ common neighbors of distinct β-colors in $B(H)$;

(H2) if $K_k \not\subseteq G$, then every $\Delta(s-1)+1$ vertices in $A(H)$ have at most $(k+1)!$ common neighbors in $B(H)$,

where $s = \binom{k}{2}$.

Proof: Let G be a graph with n vertices and $k \in \mathbb{N}$. Assume without loss of generality

$$\left\lceil n^{\frac{6}{k+6}} \right\rceil > (k+6)! \quad \text{and} \quad \left\lceil n^{\frac{6}{k+1}} \right\rceil \geq d(k).$$

By Theorem 4.2.1 we can construct in polynomial time a bipartite graph $H_0 = (A_0 \dot\cup B_0, E_0)$ such that for $s := \binom{k}{2}$:

- if $K_k \subseteq G$, then there are s vertices in A_0 with at least $d(k)$ common neighbors in B_0;
- if $K_k \not\subseteq G$, then every s vertices in A_0 have at most $(k+1)!$ common neighbors in B_0.

Define

$$A_1 := A_0 \times [\Delta],\ B_1 := B_0, \text{ and } E_1 := \Big\{\{(u,i),v\} \;\Big|\; (u,i) \in A_0 \times [\Delta],\, v \in B_0, \text{ and } \{u,v\} \in E_0\Big\}.$$

It is easy to verify that in the bipartite graph $(A_1 \dot\cup B_1, E_1)$

- if $K_k \subseteq G$, then there are Δs vertices in A_1 with at least $d(k)$ common neighbors in B_2;
- if $K_k \not\subseteq G$, then every $\Delta(s-1)+1$ vertices in A_1 have at most $(k+1)!$ common neighbors in B_1.

Applying Lemma 6.3.2 on

$$\big(n \leftarrow |A_1|, k \leftarrow \Delta s\big) \quad \text{and} \quad \big(n \leftarrow |B_1|, k \leftarrow d(k)\big)$$

we obtain two function families $\Lambda_A := \Lambda_{|A_1|,\Delta s}$ and $\Lambda_B := \Lambda_{|B_1|,d(k)}$ with the stated properties. Finally the desired bipartite graph H is defined by $\Big((A_1 \times \Lambda_A \times \Lambda_B) \dot\cup (B_1 \times \Lambda_A \times \Lambda_B), E)\Big)$ with

$$E := \Big\{\big\{(u,h_1,h_2),(v,h_1,h_2)\big\} \;\Big|\; u \in A_1,\, v \in B_1,\, h_1 \in \Lambda_A,\, h_2 \in \Lambda_B, \text{ and } \{u,v\} \in E_1\Big\}$$

and the colorings

$$\alpha(u,h_1,h_2) := h_1(u) \quad \text{and} \quad \beta(v,h_1,h_2) := h_2(v). \qquad \square$$

Setting the parameters. Let $\Delta := 2$. Recall that $k \geq 3$, $s = \binom{k}{2} \geq 3$, and $c \in \mathbb{N}^+$. We first define

$$d := d(k) := \left(30 \cdot c^2 \cdot (k+1)^2\right)^{4 \cdot k^3 + 3c}.$$

It is easy to check that:

(i) $d^{\frac{1}{2} - \frac{1}{2s}} > c \cdot s^c \left(= c \cdot \binom{k}{2}^c\right)$.

(ii) $d > \left(3(k+1)!\right)^{2s}$.

(iii) $d > (10\Delta s \cdot c^2)^{2\Delta s}$.

Then let

$$t := c \cdot d^{c - \frac{1}{2\Delta s}}.^2 \tag{5.3}$$

From (ii), (iii), and (5.3) we conclude

$$\Delta sct < 0.1 \cdot d^c, \quad \frac{c \cdot d^c}{3t} \leq \sqrt[2\Delta s]{d}, \quad \text{and } (k+1)! < \frac{\sqrt[2s]{d}}{3}. \tag{5.4}$$

Moreover by (i) and $\Delta = 2$ we have

$$c \cdot d^c + c\Delta^c s^c d^{c - \frac{1}{2} + \frac{1}{2s}} < 2\Delta^c d^c. \tag{5.5}$$

Construction of G_c. We invoke Theorem 5.3.2 to obtain $H = (A \dot\cup B, E)$, α, and β. Then we construct a new graph $G_c = \left(V(G_c), E(G_c)\right)$ as follows. First, the vertex set of G_c is given by

$$V(G_c) := \bigcup_{\boldsymbol{i} \in [d]^c} V_{\boldsymbol{i}} \dot\cup C \dot\cup W,$$

where

$$V_{\boldsymbol{i}} := \left\{\boldsymbol{v} \in B^c \;\middle|\; \beta(\boldsymbol{v}) = \boldsymbol{i}\right\} \quad \text{for every } \boldsymbol{i} \in [d]^c,$$

$$C := A \times [c] \times [t], \quad \text{and} \quad W := \left\{w_{\boldsymbol{v},\boldsymbol{j},i} \;\middle|\; \boldsymbol{v} \in V_{\boldsymbol{i}} \text{ for some } \boldsymbol{i} \in [d]^c,\ \boldsymbol{j} \in [\Delta s]^c \text{ and } i \in [t]\right\}.$$

Moreover, G_c contains the following types of edges.

(E1) For each $\boldsymbol{i} \in [d]^c$, $V_{\boldsymbol{i}}$ forms a clique.

(E2) Let $\boldsymbol{i} \in [d]^c$ and $\boldsymbol{v}, \boldsymbol{v}' \in V_{\boldsymbol{i}}$. If for all $\ell \in [c]$ we have $\boldsymbol{v}(\ell) \neq \boldsymbol{v}'(\ell)$ then $\{w_{\boldsymbol{v},\boldsymbol{j},i}, \boldsymbol{v}'\} \in E(G_c)$ for every $i \in [t]$ and $\boldsymbol{j} \in [\Delta s]^c$.

(E3) Let $i \in [t]$. Then $\left\{(u, \ell, i), w_{\boldsymbol{v},\boldsymbol{j},i}\right\} \in E(G_c)$ if $\{u, \boldsymbol{v}(\ell)\} \in E$ and $\boldsymbol{j}(\ell) = \alpha(u)$.

[2]Here, we assume $d^{c - \frac{1}{2\Delta s}}$ is an integer. Otherwise, let $d \leftarrow d^{2\Delta s}$ which maintains (i)– (iii).

(E4) Let $u, u' \in A(H)$ with $u \neq u'$, $\ell \in [c]$, and $i \in [t]$. Then $\Big\{(u, \ell, i), (u', \ell, i)\Big\} \in E(G_c)$.

Theorem 5.3.1 then follows from the completeness and the soundness of this reduction.

Lemma 5.3.3 (Completeness)**.** *If G contains k-clique, then $\gamma(G_c) < 1.1d^c$.*

Lemma 5.3.4 (Soundness)**.** *If G contains no k-clique then $\gamma(G_c) > c \cdot d^c/3$.*

We first show the easier completeness.

Proof: [of Lemma **??**] By (H1) in Theorem 5.3.2, if G contains a subgraph isomorphic to K_k, then the bipartite graph H has a $K_{\Delta s,d}$-subgraph K such that $\alpha(A \cap K) = [\Delta s]$ and $\beta(B \cap K) = [d]$. Let

$$D := (B \cap K)^c \,\dot\cup\, \Big((A \cap K) \times [c] \times [t]\Big).$$

Obviously, $|D| = d^c + \Delta sct < 1.1 \cdot d^c$ by (5.4). And (E1) and (E4) imply that D dominates every vertex in C and every vertex in $V_{\boldsymbol{i}}$ for all $\boldsymbol{i} \in [d]^c$.

To see that D also dominates W, let $w_{\boldsymbol{v},\boldsymbol{j},i}$ be a vertex in W. First consider the case where $\boldsymbol{v}(\ell) \notin B \cap K$ for all $\ell \in [c]$. Since $\beta\Big((B \cap K)^c\Big) = [d]^c$, there exists a vertex $\boldsymbol{v'} \in (B \cap K)^c$ with $\beta(\boldsymbol{v'}) = \beta(\boldsymbol{v})$ and $\boldsymbol{v}(\ell) \neq \boldsymbol{v}'(\ell)$ for all $\ell \in [c]$. Then $w_{\boldsymbol{v},\boldsymbol{j},i}$ is dominated by $\boldsymbol{v'}$ because of (E2).

Otherwise assume $\boldsymbol{v}(\ell) \in B \cap K$ for some $\ell \in [c]$, then $A \cap K \subseteq N^H(\boldsymbol{v}(\ell)) = \Big\{u \in A \;\Big|\; \{u, \boldsymbol{v}(\ell)\} \in E\Big\}$. There exists a vertex $u \in A \cap K$ such that $\alpha(u) = \boldsymbol{j}(\ell)$ and $\Big\{\boldsymbol{v}(\ell), u\Big\} \in E$. By (E3), $w_{\boldsymbol{v},\boldsymbol{j},i}$ is adjacent to (u, ℓ, i). □

5.3.2 Soundness

Lemma 5.3.5. *Suppose $c, \Delta, t \in \mathbb{N}^+$ and $\Delta < t$. Let $V \subseteq [t]^c$. If there exists a function $\theta : V \to [c]$ such that for all $i \in [c]$ we have*

$$\left|\Big\{\mathbf{v}(i) \;\Big|\; \mathbf{v} \in V \text{ and } \theta(\mathbf{v}) = i\Big\}\right| \leq t - \Delta, \tag{5.6}$$

then $|V| \leq t^c - \Delta^c$.

Proof: When $c = 1$, we have $|V| \leq t - \Delta$ by (5.6). Suppose the lemma holds for $c \leq n$ and consider $c = n + 1$. Given $V \subseteq [t]^{n+1}$ and θ, let

$$C_{n+1} := \Big\{\boldsymbol{v}(n+1) \;\Big|\; \boldsymbol{v} \in V \text{ and } \theta(\boldsymbol{v}) = n + 1\Big\}.$$

By (5.6), $|C_{n+1}| \leq t - \Delta$. If $|C_{n+1}| < t - \Delta$, we add $\Big(t - \Delta - |C_{n+1}|\Big)$ arbitrary integers from $[t] \setminus C_{n+1}$ to C_{n+1}. So we have $|C_{n+1}| = t - \Delta$. Let $A := \Big\{\boldsymbol{v} \in V \;\Big|\; \boldsymbol{v}(n+1) \in C_{n+1}\Big\}$ and $B := V \setminus A$. It follows that

$$|A| \leq (t - \Delta)t^{c-1}, \tag{5.7}$$

$\left|\left\{\boldsymbol{v}(n+1) \;\middle|\; \boldsymbol{v} \in B\right\}\right| \leq \Delta$, and $\theta(\boldsymbol{v}) \in [c-1]$ for $\boldsymbol{v} \in B$. Let

$$V' := \left\{(v_1, v_2, \ldots, v_n) \;\middle|\; \exists v_{n+1} \in [t], (v_1, v_2, \ldots, v_n, v_{n+1}) \in B\right\}.$$

We define a function $\theta' : V' \to [c-1]$ as follows. For all $\boldsymbol{v'} \in V'$, choose $\boldsymbol{v} \in B$ with the minimum $\boldsymbol{v}(c)$ such that for all $i \in [c-1]$ it holds $\boldsymbol{v'}(i) = \boldsymbol{v}(i)$. By the definition of V', such a $\boldsymbol{v}$ must exist, and we let $\theta'(\boldsymbol{v'}) := \theta(\boldsymbol{v})$. By (5.6), $\left|\left\{\boldsymbol{v'}(i) \;\middle|\; \boldsymbol{v'} \in V' \text{ and } \theta'(\boldsymbol{v'}) = i\right\}\right| \leq t-\Delta$ for all $i \in [c-1]$. Applying the induction hypothesis, we get $|V'| \leq t^{c-1} - \Delta^{c-1}$. Obviously,

$$|B| \leq \Delta|V'| \leq \Delta t^{c-1} - \Delta^c. \tag{5.8}$$

From (5.7) and (5.8), we deduce that $|V| = |A| + |B| \leq (t-\Delta)t^{c-1} + \Delta t^{c-1} - \Delta^c \leq t^c - \Delta^c$. $\square$

We are now ready to prove the soundness of our reduction.

Proof: [of Lemma 5.3.4] Let D be a dominating set of G_c. Define

$$a := \left|\left\{\boldsymbol{i} \in [d]^c \;\middle|\; |D \cap V_{\boldsymbol{i}}| \geq c+1\right\}\right|.$$

If $a > d^c/3$, then $|D| \geq (c+1)a > c \cdot d^c/3$ and we are done.

So let us consider $a \leq d^c/3$. Thus, the set

$$I := \left\{\boldsymbol{i} \in [d]^c \;\middle|\; |D \cap V_{\boldsymbol{i}}| \leq c\right\}$$

has size $|I| \geq 2d^c/3$. Let $\boldsymbol{i} \in I$ and assume that $D \cap V_{\boldsymbol{i}} = \left\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{c'}\right\}$ for some $c' \leq c$. We define a $\boldsymbol{v_i} \in V_{\boldsymbol{i}}$ as follows. If $c' = 0$, we choose an arbitrary $\boldsymbol{v_i} \in V_{\boldsymbol{i}}$.[3] Otherwise, let

$$\boldsymbol{v_i}(\ell) := \begin{cases} \boldsymbol{v}_\ell(\ell) & \text{for all } \ell \in [c']; \\ \boldsymbol{v}_1(\ell) & \text{for all } c' < \ell \leq c. \end{cases}$$

Obviously, $\beta(\boldsymbol{v_i}) = \boldsymbol{i}$.

(E2) implies that for every $\boldsymbol{j} \in [\Delta s]^c$ and every $i \in [t]$, the vertex $w_{\boldsymbol{v_i},\boldsymbol{j},i}$ is not dominated by $D \cap V_{\boldsymbol{i}}$. Observe that $w_{\boldsymbol{v_i},\boldsymbol{j},i}$ cannot be dominated by other $D \cap V_{\boldsymbol{i'}}$ with $\boldsymbol{i'} \neq \boldsymbol{i}$ either, by (E2) and (E3). Therefore every vertex in the set

$$W_1 := \left\{w_{\boldsymbol{v_i},\boldsymbol{j},i} \;\middle|\; \boldsymbol{i} \in I, \, \boldsymbol{j} \in [\Delta s]^c, \text{ and } i \in [t]\right\}$$

is *not* dominated by $D \cap \bigcup_{\boldsymbol{i} \in [d]^c} V_{\boldsymbol{i}}$. As a consequence, W_1 has to be dominated by or included in $D \cap (C \cup W)$.

[3]Since the coloring β is obtained by the color-coding used in the proof of Theorem 5.3.2, for every $b \in [d]$ it holds that $\{v \in B \mid \beta(v) = b\} \neq \emptyset$, hence $V_{\boldsymbol{i}} \neq \emptyset$.

If $|D \cap W_1| > c \cdot d^c/3$, then again we are done. So suppose $|D \cap W_1| \le c \cdot d^c/3$ and let $W_2 := W_1 \setminus D$. It follows that W_2 has to be dominated by $D \cap C$. Once again we only need to consider the case $|D \cap C| \le c \cdot d^c/3$, and hence there is an $i' \in [t]$ such that

$$\left|D \cap \Big(A \times [c] \times \{i'\}\Big)\right| \le \frac{c \cdot d^c}{3t}. \tag{5.9}$$

Then we define

$$Z := \Big\{w_{\boldsymbol{v},\boldsymbol{j},i} \in W_2 \;\Big|\; i = i'\Big\} = \Big\{w_{\boldsymbol{v_i},\boldsymbol{j},i'} \;\Big|\; \boldsymbol{i} \in I,\, \boldsymbol{j} \in [\Delta s]^c, \text{ and } w_{\boldsymbol{v_i},\boldsymbol{j},i'} \notin D\Big\}.$$

So Z has to be dominated by $D \cap C$, and in particular those vertices of the form $(u, \ell, i') \in D \cap C$. Moreover,

$$|Z| \ge \Delta^c s^c |I| - |D \cap W_1| \ge \Delta^c s^c |I| - c \cdot d^c/3. \tag{5.10}$$

Our next step is to upper bound $|Z|$. To that end, let

$$X := \Big\{u \in A \;\Big|\; (u, \ell, i') \in D \text{ for some } \ell \in [c]\Big\}.$$

Thus Z is dominated by those vertices (u, ℓ, i') with $u \in X$. And by (5.9)

$$|X| \le \frac{c \cdot d^c}{3t}.$$

Set

$$Y := \Big\{v \in B \;\Big|\; \big|N^H(v) \cap X\big| > \Delta(s-1)\Big\}.$$

Recall that $c \cdot d^c/(3t) \le \sqrt[2\Delta s]{d}$ by (5.4). Hence X has at most $\sqrt{d}$ different subsets of size $\Delta(s-1)+1$, i.e.,

$$\left|\binom{X}{\Delta(s-1)+1}\right| \le |X|^{\Delta(s-1)+1} \le |X|^{\Delta s} \le \sqrt{d}.$$

We should have

$$|Y| \le \sqrt{d} \cdot (k+1)! \le \frac{d^{\frac{1}{2}+\frac{1}{2s}}}{3}, \tag{5.11}$$

where the second inequality is by (5.4). Otherwise, by the pigeonhole principle, there exists a $(\Delta(s-1)+1)$-vertex set of $X \subseteq A(H)$ having at least $|Y|/\sqrt{d} > (k+1)!$ common neighbors in $Y \subseteq B(H)$. However, if G contains no k-clique, then by (H2) every $\big(\Delta(s-1)+1\big)$-vertex set of $A(H)$ has at most $(k+1)!$ common neighbors in $B(H)$, and we obtain a contradiction.

Let

$$\begin{aligned} Z_1 :=& \Big\{w_{\boldsymbol{v},\boldsymbol{j},i'} \in Z \;\Big|\; \text{there exists an } \ell \in [c] \text{ with } \boldsymbol{v}(\ell) \in Y\Big\} \quad \Big(\subseteq Z\Big) \\ =& \Big\{w_{\boldsymbol{v_i},\boldsymbol{j},i'} \;\Big|\; \boldsymbol{i} \in I,\, \boldsymbol{j} \in [\Delta s]^c,\, w_{\boldsymbol{v_i},\boldsymbol{j},i'} \notin D, \text{ and there exists an } \ell \in [c] \text{ with } \boldsymbol{v_i}(\ell) \in Y\Big\} \end{aligned}$$

and

$$Z_2 := Z \setminus Z_1 = \Big\{w_{\boldsymbol{v_i},\boldsymbol{j},i'} \;\Big|\; \boldsymbol{i} \in I,\, \boldsymbol{j} \in [\Delta s]^c,\, w_{\boldsymbol{v_i},\boldsymbol{j},i'} \notin D, \text{ and } \boldsymbol{v_i}(\ell) \notin Y \text{ for all } \ell \in [c]\Big\}.$$

Moreover, let $I_1 := \{\boldsymbol{i} \in I \mid \text{there exists a } w_{\boldsymbol{v_i},\boldsymbol{j},i'} \in Z_1\}$. From the definition, we can deduce that

$$\text{for all } \boldsymbol{i} \in I_1 \text{ there exists an } \ell \in [c] \text{ such that } \boldsymbol{i}(\ell) \in \beta(Y).$$

Then $|I_1| \le c|Y|d^{c-1}$ and hence

$$|Z_1| \le |I_1|\Delta^c s^c \le c|Y|d^{c-1}\Delta^c s^c.$$

To estimate $|Z_2|$, let us fix an $\boldsymbol{i} \in I$ and thus fix the tuple $\boldsymbol{v_i} \in B^c$, and consider the set

$$J_{\boldsymbol{i}} := \Big\{ \boldsymbol{j} \in [\Delta s]^c \;\Big|\; w_{\boldsymbol{v_i},\boldsymbol{j},i'} \in Z_2 \Big\}.$$

Recall that Z is dominated by those vertices (u, ℓ, i') with $u \in X$, so for every $\boldsymbol{j} \in J_{\boldsymbol{i}}$ the vertex $w_{\boldsymbol{v_i},\boldsymbol{j},i'}$ is adjacent to some (u, ℓ, i') in the dominating set D with $u \in X$. Moreover, for every $\ell \in [c]$, in the original graph H the vertex $\boldsymbol{v_i}(\ell) \in B$ has at most $\Delta(s-1)$ neighbors in X, by the fact that $\boldsymbol{v_i}(\ell) \notin Y$ and our definition of the set Y.

Define a function $\theta : J_{\boldsymbol{i}} \to [\Delta s]$ such that for each $\boldsymbol{j} \in J_{\boldsymbol{i}}$, if $w_{\boldsymbol{v_i},\boldsymbol{j},i'}$ is adjacent to a vertex $(u, \ell, i') \in D$ with $u \in X$, then $\theta(\boldsymbol{j}) = \ell$. As argued above, such a (u, ℓ, i') must exist, and if there are more than one such, choose an arbitrary one.

Let $\boldsymbol{j} \in J_{\boldsymbol{i}}$ and $\ell := \theta(\boldsymbol{j})$. By (E3), in the graph H the vertex $\boldsymbol{v_i}(\ell)$ is adjacent to some vertex $u \in X$ with $\alpha(u) = \boldsymbol{j}(\ell)$. It follows that for each $\ell \in [c]$ we have

$$\left|\Big\{ \boldsymbol{j}(\ell) \;\Big|\; \boldsymbol{j} \in J_{\boldsymbol{i}} \text{ and } \theta(\boldsymbol{j}) = \ell \Big\}\right| \le \left|\Big\{\alpha(u) \;\Big|\; u \in X \text{ adjacent to } \boldsymbol{v_i}(\ell)\Big\}\right| \le \Delta(s-1).$$

Applying Lemma 5.3.5, we obtain

$$\Big|J_{\boldsymbol{i}}\Big| \le \Delta^c s^c - \Delta^c.$$

Then

$$\Big|Z_2\Big| = \sum_{\boldsymbol{i} \in I} \Big|J_{\boldsymbol{i}}\Big| \le |I|(\Delta^c s^c - \Delta^c).$$

By (5.10) and the definition of Z_1 and Z_2, we should have

$$\Delta^c s^c |I| - c \cdot d^c/3 \le |Z| = |Z_1| + |Z_2| \le c|Y|d^{c-1}\Delta^c s^c + |I|(\Delta^c s^c - \Delta^c).$$

That is,

$$c \cdot d^c/3 + c|Y|d^{c-1}\Delta^c s^c \ge \Delta^c |I| \ge 2\Delta^c d^c/3.$$

Combined with (5.11), we have

$$c \cdot d^c + c\Delta^c s^c d^{c-\frac{1}{2}+\frac{1}{2s}} \ge 2\Delta^c d^c,$$

which contradicts the equation (5.5). $\square$

5.4 Some Consequences

Proof: [of Corollary 5.1.2] Let $c \in \mathbb{N}^+$, and assume that $\mathbb{A}$ is a polynomial time algorithm which on input a graph $G = (V, E)$ with $\gamma(G) \leq \beta(|V|)$ outputs a dominating set D with $|D| \leq c \cdot \gamma(G)$. Without loss of generality, we further assume that given $0 \leq k \leq n$ it can be tested in time $n^{O(1)}$ whether $k > c \cdot \beta(n)$.

Now let G be an arbitrary graph. We first simulate $\mathbb{A}$ on G, and there are three possible outcomes of $\mathbb{A}$.

– $\mathbb{A}$ does not output a dominating set. Then we know $\gamma(G) > \beta(|V|)$. So in time

$$2^{O(|V|)} \leq 2^{O(\beta^{-1}(\gamma(G)))}$$

we can exhaustively search for a minimum dominating set D of G.

– $\mathbb{A}$ outputs a dominating set D_0 with $|D_0| > c \cdot \beta(|V|)$. We claim that again $\gamma(G) > \beta(|V|)$. Otherwise, the algorithm $\mathbb{A}$ would have behaved correctly with

$$|D_0| \leq c \cdot \gamma(G) \leq c \cdot \beta(|V|).$$

So we do the same brute-force search as above.

– $\mathbb{A}$ outputs a dominating set D_0 with $|D_0| \leq c \cdot \beta(|V|)$. If $|D_0| > c \cdot \gamma(G)$, then

$$c \cdot \beta(|V|) \geq |D_0| > c \cdot \gamma(G), \quad \text{i.e.,} \ \beta(|V|) > \gamma(G),$$

which contradicts our assumption for $\mathbb{A}$. Hence, $|D_0| \leq c \cdot \gamma(G)$ and we can output $D := D_0$.

To summarize, we can compute a dominating set D with $|D| \leq c \cdot \gamma(G)$ in time $f(\gamma(G)) \cdot |G|^{O(1)}$ for some computable $f : \mathbb{N} \to \mathbb{N}$. This is a contradiction to Theorem 5.1.1. □

Now we come to the approximability of the monotone circuit satisfiability problem.

MONOTONE-CIRCUIT-SAT	
Instance:	A monotone circuit C.
Solution:	A satisfying assignment S of C.
Cost:	The weight of $|S|$.
Goal:	min.

Recall that a Boolean circuit C is *monotone* if it contains no negation gates; and the *weight* of an assignment is the number of inputs assigned to 1.

As mentioned in the Introduction, Marx showed [Mar13] that MONOTONE-CIRCUIT-SAT has no fpt approximation with any ratio ρ for circuits of depth 4, unless FPT = W[2].

Corollary 5.4.1. *Assume* FPT $\neq$ W[1]. *Then* MONOTONE-CIRCUIT-SAT *has no constant fpt approximation for circuits of depth 2.*

Proof: This is an immediate consequence of Theorem 5.1.1 and the following well-known approximation-preserving reduction from Monotone-Circuit-Sat to Min-Dominating-Set. Let $G = (V, E)$ be a graph. We define a circuit

$$C(G) = \bigwedge_{v \in V} \bigvee_{\{u,v\} \in E} X_u.$$

There is a one-one correspondence between a dominating set in G of size k and a satisfying assignment of $C(G)$ of weight k. $\square$

Remark 5.4.2. *Of course the constant ratio in Corollary 5.4.1 can be improved according to Theorem 5.1.3.*

5.5 Conclusions

We have shown that k-Dominating-Set has *no* fpt approximation with any constant ratio, and in fact with a ratio slightly super-constant. The immediate question is whether the problem has fpt approximation with *some* ratio $\rho : \mathbb{N} \to \mathbb{N}$, e.g., $\rho(k) = 2^{2^k}$. We tend to believe that it is not the case.

Remark 5.5.1. *Of course the constant ratio in Corollary 5.4.1 can be improved according to Theorem 5.1.3.*

Question 5.5.2. *Does* Monotone-Circuit-Sat *admit any fpt approximation algorithm for some computable ratio function $\rho(k)$?*

Chapter 6

Maximum k-Edge-Color Clique

In the Maximum k-Edge-Color Clique problem, the task is to find a maximum clique whose edges use at most k distinct colors in a multi-graph G with colored edges. As far as we are aware, this problem has received no attention in the literature. Our study of this problems is motivated by its connection to the parameterized inapproximability of k-Clique. We show that there is no fpt-algorithm that can approximate Maximum k-Edge-Color Clique to any computable ratio function $\rho(k)$ on **mutigraphs** unless FPT = W[1]. We also prove that the inapproximability of Maximum k-Edge-Color Clique on **simple graphs** implies the constant inapproximability of k-Clique. Finally, we point out a possible approach to transfer a multi-graphs into a simple graph without changing the cardinality of its maximum k-edge-color clique.

6.1 Introduction

Let us start with the question of proving the constant inapproximability of the parameterized clique problem. In order to show that k-Clique has no ρ-approximation fpt-algorithm for some ratio $\rho > 1$ under FPT $\neq$ W[1], it suffices to find an algorithm $\mathbb{A}$ such that on input an instance (G, k) of k-Clique, $\mathbb{A}$ outputs a graph G' and an integer k' in $f(k) \cdot |G|^{O(1)}$-time. Moreover, G' satisfies the following properties.

- if G contains a k-clique, then G' contains a $\rho k'$ clique.
- if G contains no k-clique, then G' contains no k' clique.
- $k' = g(k)$ for some computable function g.

Informally, by Theorem 4.2.1, we know how to construct a bipartite graph $H = (A(H) \dot\cup B(H), E(H))$ from a graph G and a small integer k in polynomial time satisfying the following properties.

- (H1) if G contains a k-clique, then H contains a $K_{\binom{k}{2}, n^{1/k}}$-subgraph;
- (H2) if G contains no k-clique, then H contains no $K_{\binom{k}{2}, (k+1)!+1}$-subgraph.

A simple idea is to start with an operation that transforms a $K_{\binom{k}{2},n^{1/k}}$ into a large clique but turns a $K_{\binom{k}{2},(k+1)!}$ into a small clique. Note that in the case (H2) of Theorem 4.2.1, there may exist a $K_{\binom{k}{2}-1,n^{1/k}}$-subgraph in the graph H. This means that our desired operation should also transform a $K_{\binom{k}{2}-1,n^{1/k}}$-subgraph into a small clique.

Let $s := \binom{k}{2}$. Our approach is motivated by an observation that it is possible to construct a family of edge sets $\mathcal{E} := \{E_1, E_2, \ldots, E_n\}$ over a vertex set $V := [n]$ satisfying the following properties.

- for every $I \in \binom{n}{s-1}$, the graph $G_I := (V, \bigcup_{i \in I} E_i)$ contains no $(2^{s-1}+1)$-clique.
- for a random $I \in \binom{n}{s}$, with high probability the graph $G_I := (V, \bigcup_{i \in I} E_i)$ contains a 2^s-clique.

This observation promoted us to do the following reduction. On input a bipartite graph $H := (A(H) \dot\cup B(H), E(H))$ with $A(H) = [n]$, we associate each vertex u in $A(H)$ with a set of edge $E_u \subseteq \binom{N^H(u)}{2}$ and hence obtain a family of edge sets $\mathcal{E} := \{E_1, E_2, \ldots, E_n\}$ where each E_i is a subset of $\binom{B(H)}{2}$. For any $I \subseteq [n]$, let G_I be the graph with $V(G_I) := B(H)$ and $E(G_I) := \bigcup_{i \in I} E_i$. Choosing the edge set E_u for each $u \in A(H)$ properly, we may expect to prove something as follows.

- if H contains a $K_{s,2^s}$, then there exists $I \in \binom{A(H)}{s}$ such that G_I contains a 2^s-clique;
- if H contains no $K_{s,(k+1)!+1}$, then for all $I \in \binom{A(H)}{s}$, G_I contains no $(2^{s-1}+O(k!))$-clique.

The discussion above provides the intuition of a gap-preserving reduction from maximum subset intersection to an optimization problem whose goal is to find a maximum clique in a graph G_I with $V(G_I) := V$ and $E(G_I) := \bigcup_{i \in I} E_i$ for some $I \in \binom{n}{k}$, on input a family of edge sets $\mathcal{E} := \{E_1, E_2, \ldots, E_n\}$ over $V := [n]$. Recall the definition of multi-graph $G = (V, E, C)$ in Chapter 2. The family of edge sets $\mathcal{E}$ together with V can be treated as a multi-graph $G := (V, E, C)$ with $C := [|\mathcal{E}|]$ and $E := \bigcup_{i \in C} E_i \times \{i\}$.

Formally, we consider the following parameterized optimization problems.

Max-k-EdgeCol-Clique	
Instance:	A multi-graph $G = (V, E, C)$ and a positive integer k.
Solution:	A clique $K \subseteq G_I$ for some $I \in \binom{C}{k}$.
Cost:	$|K|$.
Goal:	max.

Example 6.1.1. *For example, the multi-graph G in Figure 6.1 has a 2-edge-color clique with three vertices. The edges of this clique use two color, red and blue. Obviously, it is the maximum 2-edge-color clique of this graph.*

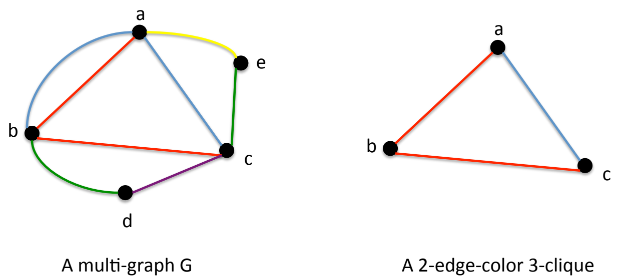

Figure 6.1: Max-2-Edge-Color Clique

6.1.1 Main Results

Theorem 6.1.2. *For any computable function $\rho(k)$.* MAX-k-EDGECOL-CLIQUE *has no $\rho(k)$-approximation* fpt-*algorithm, unless* FPT = W[1].

In a multi-graph $G = (V, E, C)$, there may exist $|C|$ edges between two vertices in V. For $f \in \mathbb{N}$, recall again that f-multi-graphs is a special case of multi-graphs such that every two vertices has at most f edges. We consider MAX-k-EDGECOL-CLIQUE restricted to 1-multi-graphs.

MAX-k-EDGECOL-CLIQUE$_1$	
Instance:	A 1-multi-graph $G = (V, E, C)$ and a positive integer k.
Solution:	A clique $K \subseteq G_I$ for some $I \in \binom{C}{k}$.
Cost:	$\lvert K \rvert$.
Goal:	max.

The following theorem is an attempt to provide a gap-preserving fpt-reduction to k-CLIQUE from MAX-k-EDGECOL-CLIQUE, although at the moment we are only able to do this for MAX-k-EDGECOL-CLIQUE$_1$.

Theorem 6.1.3. *If* MAX-k-EDGECOL-CLIQUE$_1$ *has no $4^{(k+1)^k}$ approximation* fpt-*algorithm, then k-*CLIQUE *cannot be* fpt-*approximated to any constant ratios.*

Theorem 6.1.3 and Theorem 6.1.2 suggest a question:

Question 6.1.4. *Is it possible to turn a general multi-graph into a 1-multi-graph without changing the size its maximum k-edge-color clique significantly?*

To ease notation, for any multi-graph $G = (V, E, C)$, let

$$\mathsf{CC}(G, k) := \max\left\{h \;\middle|\; h \in \mathbb{N} \text{ and there exists } I \in \binom{C}{k} \text{ such that } K_h \subseteq G_I\right\}.$$

As a partial answer to Question 6.1.4, we prove

Theorem 6.1.5. *Let $f, h : \mathbb{N} \to \mathbb{N}$ be two computable functions. Assume for every $k \in \mathbb{N}$ the computable function f satisfies*

$$f(k) > \binom{h(k)}{2}.$$

Then, there is an fpt*-algorithm $\mathbb{A}$ that on input $k \in \mathbb{N}$ and an $n^{1/f(k)}$-multi-graph G with $\mathsf{CC}(G, k) \leq h(k)$, it outputs a 1-multi-graph G' satisfying*

$$\mathsf{CC}(G, k) = \mathsf{CC}(G', k)$$

6.2 Inapproximability of Max-k-EdgeCol-Clique

In this section, we establish the parameterized inapproximability of the maximum edge-colored clique problem on multi-graphs.

Theorem 6.2.1. *There is an algorithm $\mathbb{A}$ such that on input an n-vertex graph G and a positive integer k with $\lceil n^{\frac{6}{k+1}} \rceil > (k+1)!$ and $6 \mid k+1$, it output a multi-graph G' in time $f(k) \cdot |G|^{O(1)}$ satisfying the following properties.*

(1) if G contains a k-clique, then G' contains an s-edge-colored clique with sized d^s;

(2) if G contains no k-clique, then every s'-edge-color clique in G' has at most $\binom{s'}{s-1}d^{s-1} + \binom{s'}{s}(k+1)!$ vertices.

where $s = \binom{k}{2}$ and $d \in \mathbb{N}$ with $d^{s-1} > (k+1)!$ and $\lceil n^{\frac{6}{k+1}} \rceil \geq d^s$.

Obviously, Theorem 6.1.2 is a corollary of the above theorem.

The Construction Given a graph G and an integer k, we first compute the graph $H = (A \dot\cup B, E(H))$ as stated in Theorem 4.2.1. Choose $d \in \mathbb{N}$, $s = \binom{k}{2}$ with $d^{s-1} > (k+1)!$. For $\alpha : A \to [s]$ and $\beta : B \to [d]^s$, we construct a multi-graph $G'_{\alpha,\beta,s,d} = (V, E, C)$, where

- $V = B$;
- $C = A$;
- For $u_1, u_2 \in V, c \in C$, $(\{u_1, u_2\}, c) \in E$ if:

(E1) $\beta(u_1)(\alpha(c)) \neq \beta(u_2)(\alpha(c))$;[1]

(E2) $\{u_1, c\}, \{u_2, c\} \in E(H)$;

Using the color-coding technique, we have:

Lemma 6.2.2. *There is an* fpt*-algorithm* $\mathbb{A}$ *such that on input a bipartite graph* $H = (A \dot\cup B, E)$ *and two parameters* s *and* t*, it constructs two class of functions* $\mathscr{A}_{A,s} \subseteq [s]^A$ *and* $\mathscr{B}_{B,t} \subseteq [t]^B$ *such that for every* $C \in \binom{A}{s}$*,* $D \in \binom{B}{t}$ *there exist* $\alpha \in \mathscr{A}_{A,s}$ *and* $\beta \in \mathscr{B}_{B,q}$ *with*

$$\alpha(C) = [s],\ \beta(D) = [t].$$

Lemma 6.2.3 (Completeness)**.** *If* $K_k \subseteq G$*, then there exist* $\alpha \in \mathscr{A}_{A,s}$ *and* $\beta \in \mathscr{B}_{B,d^s}$ *such that* $G'_{\alpha,\beta,s,d}$ *contains an* s*-edge-colored clique with size* d^s*.*

Proof: If $K_k \subseteq G$, then there exist an s-vertex set $C' = \{c_1, \ldots, c_s\}$ in $A(H)$ and a d^s-vertex set V' in $B(H)$ such that $V' \subseteq \Gamma(C')$. By Lemma 6.2.2, there exist $\alpha \in \mathscr{A}_{A,s}$ and $\beta \in \mathscr{B}_{B,d^s}$ with $\alpha(C') = [s]$, $\beta(V') = [d]^s$.

For all $v_i, v_j \in V'$ with $v_i \neq v_j$, by $|\beta(V')| = d^s = |V'|$, we should have that $\beta(v_i) \neq \beta(v_j)$. Then there must exist an index $\ell \in [s]$ such that $\beta(v_i)(\ell) \neq \beta(v_j)(\ell)$. Since $\alpha(C') = [s]$, we can find a $c \in C'$ such that $\alpha(c) = \ell$.

It is easy to verify that v_i, v_j and c satisfy

$$\beta(v_i)(\alpha(c)) \neq \beta(v_j)(\alpha(c)). \tag{6.1}$$

On the other hand, recall that $V' \subseteq \Gamma(C')$, which implies

$$\{v_i, c\}, \{v_j, c\} \in E(H). \tag{6.2}$$

(6.1) and (6.2 actually mean that $(\{v_i, v_j\}, c) \in E(G'_{\alpha,\beta,s,d})$.

We conclude that V' is a clique in $G'_{\alpha,\beta,s,d}$, its edges only use colors in C'. Thus V' is a s-edge-color clique with size d^s. □

Lemma 6.2.4 (Soundness)**.** *Let* $s' \geq s$*. If* $K_k \nsubseteq G$*, then every* s'*-edge-color clique in* $G'_{\alpha,\beta,s,d}$ *has at most* $\binom{s'}{s-1} d^{s-1} + \binom{s'}{s}(k+1)!$ *vertices.*

Proof: Let $W \subseteq C$ be an s'-color set such that there exists a set X with

$$\forall u, v \in X, \exists c \in W, s.t. \Big(\{u, v\}, c) \in E(G'_{\alpha,\beta,s,d}\Big). \tag{6.3}$$

For any u in X, let

$$C(u) := \Big\{ c \in W \;\Big|\; \exists v \in X, (\{u, v\}, c) \in E(G'_{\alpha,\beta,s,d}) \Big\}. \tag{6.4}$$

[1] $\beta(u_i)$ is a vector with s elements, $\beta(u_i)(\alpha(v))$ is the $\alpha(v)$-th element of this vector

Let $Y = \left\{u \in X \;\middle|\; |C(u)| \leq s-1\right\}$. Then we have:

Claim: $|Y| \leq \binom{s'}{s-1} d^{s-1}$.

Proof of the Claim: Let $\mathcal{W}_{s-1} := \binom{W}{s-1}$. Because $|W| = s'$, there are $\binom{s'}{s-1}$ distinct $(s-1)$-vertex subsets of W, i.e., $|\mathcal{W}_{s-1}| = \binom{s'}{s-1}$. For every $u \in Y$, $|C(u)| \leq s-1$, there exists $C' \in \mathcal{W}_{s-1}$ such that $C(u) \subseteq C'$. If $|Y| > \binom{s'}{s-1} d^{s-1}$, then by the pigeonhole principle, there exist $C' \in \mathcal{W}_{s-1}$ and $Y' \subseteq Y$ such that $|Y'| > d^{s-1}$ and $\forall u \in Y', C(u) \subseteq C'$.

By the pigeonhole principle again, there must exist two vertices $u, u' \in Y'$ such that $\forall c \in C', \beta(u_1)(\alpha(c)) = \beta(u_2)(\alpha(c))$. According to (E1) there is no $c \in C'$ such that $(\{u, u'\}, c) \in E(G'_{\alpha,\beta,s,d})$. This contradicts (6.3) and (6.4). $\dashv$

Let $Z = \left\{u \in X \;\middle|\; |C(u)| \geq s\right\}$. By (E2), every vertex in $C(u)$ is adjacent to u in the graph H. Since $K_k \nsubseteq G$, every s-vertex set in $A(H)$ has at most $(k+1)!$ common neighbors in $B(H)$. We have

$$|Z| \leq \binom{s'}{s}(k+1)!.$$

Therefore $|X| \leq \binom{s'}{s-1} d^{s-1} + \binom{s'}{s}(k+1)!$. $\square$

6.3 From Max-k-EdgeCol-Clique to k-Clique

In this section, we present a gap-preserving reduction from the maximum edge-colored clique problem on 1-multi-graphs to the clique problem. Recall that a 1-multi-graph is a simple graph with colored edges.

Theorem 6.3.1. *Let $\rho : \mathbb{N} \to \mathbb{N}$. On input a 1-multi-graph $G = (V, E, C)$ and $k, \ell \in \mathbb{N}$, we can construct a simple graph $\mathcal{H}$ in $f(k,\ell) \cdot |G|^{O(1)}$-time such that:*

- *if $\mathsf{CC}(G, k) = h$, then $\mathcal{H}$ contains a hq-clique;*
- *if $\mathsf{CC}(G, k) = \ell$, then $\mathcal{H}$ contains no clique with more than $\frac{2hq}{\rho(k)}$ vertices.*

where $h = \ell\rho(k)^{(k+1)^k}$, $q = 2\rho(k)(k+1)^k$.

Let $\rho = 4$. If k-Clique can be fpt-approximated with ratio $2 = \rho/2$, then the parameterized edge-colored clique problem on 1-multi-graphs can be fpt-approximated with ratio

$$\frac{h}{\ell} = 4^{(k+1)^k}.$$

Hence we have Theorem 6.1.3.

6.3.1 The Construction.

Let $G = \Big(V(G), E(G), C(G)\Big)$ be an edge-colored *simple* graph and $k \in \mathbb{N}$ with $6 \mid k+1$. Moreover, let $n := |C|$ and $h \in \mathbb{N}$. We assume without loss of generality that $C = [n]$, n is large enough and G has no clique of size $h+1$.

We apply Lemma 3.1.4 and obtain a bipartite graph $G_{n,k} = \Big(A \,\dot\cup\, B, E(G_{n,k})\Big)$ with the $\Big(n, k, (k+1)!, n^{\lceil 6/(k+1)\rceil}\Big)$-threshold property, in particular

$$A = \dot\bigcup\nolimits_{i\in[n]} V_i(A).$$

For every $q \in \mathbb{N}$ and $\beta : B \to [q]$ we construct a graph $H = H\Big(G, G_{n,k}, \beta\Big)$ as follows:

(1) $V(H) := V(G) \times B$

(2) $\Big\{(u_1, b_1), (u_2, b_2)\Big\} \in E(H)$ if $u_1 = u_2$ and $\beta(b_1) \neq \beta(b_2)$ or all of following conditions are satisfied:

   (a) $u_1 \neq u_2$

   (b) $b_1 = b_2 \vee \beta(b_1) \neq \beta(b_2)$

   (c) $\Big(\{u_1, u_2\}, c\Big) \in E(G)$ for some color $c \in C(G)$ and $\{a, b_1\}, \{a, b_2\} \in E(G_{n,k})$ for some $a \in V_c(A)$

The following lemma is essentially the Lemma 2.6.1.

Lemma 6.3.2. *We can construct in fpt-time on input a set B and a parameter q a class of functions $\mathscr{B}_{B,q} \subseteq [q]^B$ such that for every $D \in \binom{B}{q}$ there is a $\beta \in \mathscr{B}_{B,q}$ with*

$$\beta(D) = [q].$$

From G, $G_{n,k}$ and $\mathscr{B}_{B,q}$, we construct our target graph $\mathcal{H}$ as:

$$\mathcal{H} := \dot\bigcup\nolimits_{\beta\in\mathscr{B}_{B,q}} H\Big(G, G_{n,k}, \beta\Big)$$

6.3.2 Completeness.

Lemma 6.3.3 (completeness)**.** *If G has a k-edge-colored clique of size h, then for every $q \leq n^{\lceil 6/(k+1)\rceil}$ there is a $\beta \in \mathscr{B}_{B,q}$ such that $H = H\Big(G, G_{n,k}, \beta\Big)$ has a clique of size*

$$h \cdot q.$$

Proof: Let $c_1, c_2, \ldots, c_k$ be k colors in $C(G)$ such that $G_{\{c_1,c_2,\ldots,c_k\}}$ contains a h-clique with vertex set $\{u_1, u_2, \ldots, u_h\}$. We have that

$$\forall i \neq j \in [h], \exists \ell \in [k], (\{u_i, u_j\}, c_\ell) \in E(G) \tag{6.5}$$

By (T2) there exist $a_1 \in V_{c_1}, a_2 \in V_{c_2}, \ldots, a_k \in V_{c_k}$ and $b_1, b_2, \ldots, b_q \in B$ such that

$$\forall i \in [k], j \in [q], \{a_i, b_j\} \in E(G_{n,k}). \tag{6.6}$$

By Lemma 6.3.2, there exists $\beta \in \mathscr{B}_{B,q}$ such that

$$\beta(\{b_1, b_2, \ldots, b_q\}) = [q]. \tag{6.7}$$

Claim. The vertex set

$$K := \{(u_i, b_j) \mid i \in [h], j \in [q]\}$$

forms a K_{hq}-subgraph in $H = H(G, G_{n,k}, \beta)$.

Proof of Claim. Let (u_{i_1}, b_{j_1}) and (u_{i_2}, b_{j_2}) be two distinct vertices in K, where $i_1, i_2 \in [h]$, $j_1, j_2 \in [q]$.

If $i_1 = i_2$, then we should have $b_{j_1} \neq b_{j_2}$. By (6.7), $\beta(b_{j_1}) \neq \beta(b_{j_2})$. According to the definition of $E(H)$, $\{(u_{i_1}, b_{j_1}), (u_{i_2}, b_{j_2})\} \in E(H)$.

Now assume $i_1 \neq i_2$. By (6.5), there exists some color $c_\ell \in C(G)$ such that $(\{u_i, u_j\}, c_\ell) \in E(G)$. By (6.6), there exists a $a_\ell \in V_{c_\ell}$ with $\{a_\ell, b_{j_1}\}, \{a_\ell, b_{j_2}\} \in E(G_{n,k})$). Note that if $b_{j_1} \neq b_{j_2}$, then $\beta(b_{j_1}) \neq \beta(b_{j_2})$. We conclude that $\{(u_{i_1}, b_{j_1}), (u_{i_2}, b_{j_2})\} \in E(H)$ $\dashv$.

Thus H contains a K_{hq}-subgraph. We are done. $\square$

6.3.3 Soundness

The following lemma is a well-known result in extremal graph theory. It is linked to the famous Zarankiewicz's problem. More detail can be found in Section 2.4 of [Juk11]. For the convenience of the reader, we provide a proof based on the double counting argument.

Lemma 6.3.4. *Let $s, t \in \mathbb{N}$ and $T = \big(U \dot\cup V, E(T)\big)$ be a bipartite graph such that every s vertices in U have at most t common neighbors in V with $|V| \geq t$. Then*

$$\Big|E(T)\Big| \leq s^{1/t}|U|^{1-1/t}|V| + t|U|. \tag{6.8}$$

Proof: Let

$$S := \left\{ \Big(u, \{v_1, \ldots, v_t\}\Big) \;\middle|\; u \in U \text{ and } \{v_1, \ldots, v_t\} \in \binom{N(u)}{t} \right\}.$$

Then

$$\sum_{u \in U} \binom{\deg(u)}{t} = |S| \leq s\binom{|V|}{t}. \tag{6.9}$$

Observe that $\Big|E(T)\Big| = \sum_{u \in U} \deg(u)$. Following the notation in Section 2.4 of [Juk11], we set $f(x) := \binom{x}{t}$ if $x \geq t$, and $f(x) := 0$ otherwise. Then f is convex, Jensen's Inequality yields

$$\sum_{u\in U}\binom{\deg(u)}{t} \geq \sum_{u\in U} f(\deg(u)) \geq |U| f(\sum_{u\in U}\frac{\deg(u)}{|U|}) = |U|\cdot f(\frac{|E(T)|}{|U|}).$$

If $|E(T)| \leq |U|t$, then (6.8) holds. Otherwise $|E(T)| > |U|t$, then we have that

$$|S| \geq |U| f(\frac{|E(T)|}{|U|}) = |U|\binom{\left|E(T)\right|/|U|}{t}.$$

Combined with (6.9), we obtain:

$$|U|\binom{\left|E(T)\right|/|U|}{t} \leq s\binom{|V|}{t} \tag{6.10}$$

Using $\frac{(n-k+1)^k}{k!} \leq \binom{n}{k} \leq \frac{n^k}{k!}$, then (6.10) implies:

$$|U|\cdot(|E(T)/|U|-t+1)^t \leq s(|V|)^t$$

Hence,

$$\left|E(T)\right| \leq s^{1/t}|U|^{1-1/t}|V| + (t-1)|U|.$$ □

Lemma 6.3.5 (Soundness)**.** *For any $\rho : \mathbb{N} \to \mathbb{N}$. If G has no k-edge-colored clique of size ℓ with*

$$\ell = \frac{h}{\rho(k)^{(k+1)^k}}, \tag{6.11}$$

then for every $\beta : B \to [q]$, $H = H\Big(G, G_{n,k}, \beta\Big)$ has no clique of size larger than

$$\frac{2h\cdot q}{\rho(k)}.$$

where $q = 2\rho(k)(k+1)^k$.

Proof: Let

$$X = \Big\{(u_1, b_1), \ldots, (u_m, b_m)\Big\} \tag{6.12}$$

be a clique of size

$$m > \frac{2h\cdot q}{\rho(k)} \geq 2h(k+1)^k \tag{6.13}$$

in H. We define a bipartite graph $T = \Big(U \dot\cup V, E(T)\Big)$ with

$$U := \{u_i \mid i \in [m]\},\ V := \{b_i \mid i \in [m]\},\ \text{and}\ \ E(T) := \Big\{\{u_i, b_i\} \Bigm| i \in [m]\Big\}. \tag{6.14}$$

By the definition of H, it is easy to see that U is a clique in G, and therefore $|U| \leq h$. And recall $\beta : B \to [q]$, thus $|V| \leq q$ by $\beta(b_i) \neq \beta(b_j)$ for every $1 \leq i < j \leq m$.

We should have that $|V| \geq (k+1)^k$, otherwise $m \leq h|V| < h(k+1)^k$, which contradicts (6.13). Now we apply Lemma 6.3.4 to T with

$$t \leftarrow (k+1)^k \quad s \leftarrow \ell$$

and hence by (6.11)

$$\left(\frac{h}{s}\right)^{\frac{1}{(k+1)^k}} = \rho(k). \tag{6.15}$$

Thus $\left|E(T)\right| = m > 2\frac{h \cdot q}{\rho(k)}$ implies, by (6.15),

$$\frac{\left|E(T)\right|}{2} = \rho(k) \cdot \left(\frac{s}{h}\right)^{\frac{1}{(k+1)^k}} \frac{\left|E(T)\right|}{2} > s^{\frac{1}{(k+1)^k}} h^{1-\frac{1}{(k+1)^k}} q \geq s^{1/t}|U|^{1-1/t}|V|.$$

On the other hand, if $m \leq 2|U|t$ then $m \leq 2ht$, which contradicts (6.13). Therefore, we have

$$|E(T)| = m \geq s^{1/t}|U|^{1-1/t}|V| + t|U|.$$

That is, (6.8) in Lemma 6.3.4 does not hold. It follows that, in the graph T, there are s vertices in U with more than t common neighbors in V. Without loss of generality, assume they are $u_1, \ldots, u_s$ and $b_1, \ldots, b_t$, respectively.

Note $\{u_1, \ldots, u_s\}$ is an s-clique in G. We collect its corresponding edge colors by

$$K := \left\{c \;\middle|\; \left(\{u_i, u_j\}, c\right) \in E(G) \text{ for } 1 \leq i < j \leq s\right\}. \tag{6.16}$$

The goal is to show $|K| \leq k$.

Claim: For every $c \in K$ and every $\ell \in [t]$ we have $\{c, b_\ell\} \in E(G_{n,k})$.

Proof of the claim: By (6.16) there exist some distinct $i, j \in [s]$ with $\left(\{u_i, u_j\}, c\right) \in E(G)$. Take an arbitrary $\ell' \in [t]$ with $\ell' \neq \ell$. Since all $u_1, \ldots, u_s$ are adjacent to all $b_1, \ldots, b_t$ in T, we have from (6.14)

$$(u_i, b_\ell) \ \text{ and } \ (u_j, b_{\ell'})$$

are two vertices in the clique X in H defined in (6.12). Then by the definition of H, we can conclude $\{c, b_\ell\} \in E(G_{n,k})$, since every edge in G has a unique color. $\dashv$

Recall $G_{n,k}$ satisfies the $\left(n, k, (k+1)!, n^{\lceil 6/(k+1) \rceil}\right)$-threshold property. Therefore, if $|K| \geq k$, then (T1) implies that all vertices in K have at most $(k+1)! < t$ common neighbors, which contradicts the above claim. $\square$

6.4 Multi-graphs with Few Multi-edges

In this section, we present a proof of Theorem 6.1.5.

Proof: [of Theorem 6.1.5] Let $n, k \in \mathbb{N}$ and G be a $n^{1/f(k)}$-multi-graph with n vertices. We set $h := h(k)$. For $f(k) > \binom{h}{2}$ and n sufficiently large, we have that

$$\frac{\binom{h}{2}\log n}{f(k)} n^{\binom{h}{2}/f(k)} < \sqrt{\frac{\binom{n}{2}\log n}{f(k)}}, \tag{6.17}$$

Then we can apply Lemma 2.6.4 with parameters

$$k \leftarrow \binom{h}{2}, \ n \leftarrow \binom{n}{2}, \text{ and } m \leftarrow n^{1/f(k)}$$

to obtain a set $\mathcal{C}$ of functions $\left[\binom{n}{2}\right] \to \left[n^{1/f(k)}\right]$.

Now for every $\mathbf{c} \in \mathcal{C}$ we construct a simple edge-colored graph $G_{\mathbf{c}}$ from the multi-graph $G = (V, E, C)$ as follows:

(1) As $|V| = n$, we can assume $V = [n]$.

(2) For every pair $\{u, v\} \in \binom{[n]}{2}$ we order the edges between u and v. By the assumption, there are at most m such edges.

(3) Note $\{u, v\}$ can be uniquely identified with a number in $\left[\binom{n}{2}\right]$ which is the domain of $\mathbf{c}$. Then in $G_{\mathbf{c}}$ we only keep the $\mathbf{c}(\{u, v\})$th edge from G with all the others removed. Of course, it might happen that there is no $\mathbf{c}(\{u, v\})$th edge. Consequently, there is no edge between u and v in $G_{\mathbf{c}}$.

Then we let

$$H := \dot{\bigcup}_{\mathbf{c}\in\mathcal{C}} G_{\mathbf{c}},$$

which can be constructed in fpt-time. Again H is simple. Obviously $\mathsf{CC}(H, k) \le \mathsf{CC}(G, k)$. Suppose G has a k-edge-colored clique K of size at most h. There exists $\mathbf{c} \in \mathcal{C}$ such that for every $\{u, v\} \in \binom{K}{2}$, $(\{u, v\}, \mathbf{c}(\{u, v\})) \in E(G)$. It is routine to verify that $G_{\mathbf{c}}$ and hence H contains a k-edge-colored clique with size $|K|$. Thus $\mathsf{CC}(G, k) \le h$ implies that $\mathsf{CC}(H, k) = \mathsf{CC}(G, k)$. □

Chapter 7

Conclusions and Future Work

There are two kinds of hardness results in computational complexity. One states that to solve some problem enumeration is inevitable. The other states that it is impossible to approximate some problem to a certain ratio in an efficient way. This thesis is a step in the direction of proving hardness results which combine these two perspectives.

We have developed a tool that allows us to prove some results of this kind. By exploiting the threshold property of a certain class of bipartite graphs, we have established the parameterized inapproximability of MAX-k-SUBSET-INTERSECTION, MIN-DOMINATING-SET, and MAX-k-EDGECOL-CLIQUE.

This is not an end of the story. It remains to see if it is possible to combine our technique with the deep PCP-theorem. The ultimate goal is to find parameters that determine the hardness of approximation for the problems. For certain problems, we expect to obtain hardness results with trade-off between the running-time of algorithms and the approximation ratios.

To motivate further research, we end this thesis by listing some open questions.

7.1 Questions on Parameterized Inapproximability

k-Clique One of the the open problems in parameterized complexity is whether k-CLIQUE admits any constant fpt-approximation algorithms. There is evidence suggesting that the answer to this question is negative. In [Ros08], Benjamin Rossman showed that with high probability constant-depth circuits of size $O(n^t)$ can not distinguish between a random graph G with edge probabilities $n^{-\alpha}$ where $\alpha \leq \frac{1}{2t-1}$ and the graph obtained by planting a k-clique into G. Since for $\alpha > \frac{2}{k'-1}$, the random graph G does not contain a $K_{k'}$ with high probability. For $t = k'/4$, there exists α satisfying $\alpha \leq \frac{1}{2t-1}$ and $\alpha > \frac{2}{k'-1}$. In particular, let $k' = k/2$, we obtain a corollary that constant-depth circuits of size $O(n^{k/8})$ can not distinguish whether a graph contains a K_k or contains no $K_{k/2}$.

In Chapter 6, we prove the following results.

(1) MAX-k-EDGECOL-CLIQUE on multi-graphs has no fpt-approximation algorithm to any ratio $\rho(k)$ unless FPT = W[1].

(2) If Max-k-EdgeCol-Clique on simple graphs does not admit fpt-approximation algorithms to ratio $\rho(k)$, then k-Clique has no constant fpt-approximation algorithms.

To rule out the constant fpt-approximation for k-Clique, it suffices to solve the following question.

Question 7.1.1. *On input a multi-graph G and $h \in \mathbb{N}$, construct a simple graph G' with colored edges in* fpt*-time such that*

– *if $\mathsf{CC}(G, k) \geq h$, then $\mathsf{CC}(G', k') \geq h'$;*

– *if $\mathsf{CC}(G, k) \leq \frac{h}{\rho}$, then $\mathsf{CC}(G', k') \leq \frac{h'}{4^{(k'+1)^{k'}}}$.*

Another approach is promoted by the observation that the constant inapproximability of k-Clique can be derived from the hardness approximation of Densest-k-Subgraph (Remark 4.2.3).

Densest-k-Subgraph. Densest-k-Subgraph is a classical optimization problem. On input a graph G, it asks for a k-vertex subgraph of G with the most edges. In [KP93], a polynomial time approximation for Densest-k-Subgraph with ratio $n^{7/18}$ was presented. The approximation ratio was improved to $n^{1/3}$ in [FPK01] and $O(n^{1/4})$ in [BCC+10]. On the other hand, it remains a major open problem to prove Densest-k-Subgraph is NP-hard to approximated to any constant ratio.

Densest-k-Subgraph	
Instance:	k and a graph G.
Solution:	k-vertex subsets $K \in \binom{V(G)}{k}$.
Cost:	$\frac{\lvert E(G) \cap \binom{K}{2} \rvert}{\binom{k}{2}}$.
Goal:	max.

The result in [BKRW15] rules out an additive PTAS for Densest k-Subgraph up to ETH. Notice that this problem has a natural parameter k, we may ask whether it has a constant fpt-approximation algorithm.

Question 7.1.2. *Assuming* ETH*, does* Densest-k-Subgraph *admit an* fpt*-approximation algorithm to any constant ratio?*

Set Cover In the decision version of Set Cover problem, we are given $k \in \mathbb{N}$ and a family of subset over a finite set U with n elements, the goal is to select k subsets whose union covers U. This problem has two optimization versions. One is to keep the goal unchanged and look for a minimum number of subsets that can cover U. Let us call such version Min-Set-Cover. Another version, which we call Max-k-Set-Cover, is to select k subsets whose union covers as many elements as possible.

The constant inapproximability of Min-Dominating-Set can be easily transfer to Min-Set-Cover. The immediate question is whether the problem has fpt approximation with *some* ratio $\rho : \mathbb{N} \to \mathbb{N}$. On the other hand, the reduction in Chapter 5 does not apply to Max-k-Set-Cover, which can be approximated to $\frac{e}{e-1}$ by the greedy algorithm. In [Fei98], it is shown that Max-k-Set-Cover cannot be approximated within a ratio of $\frac{e}{e-1} - \varepsilon$, unless P = NP. The question is, can we achieve better ratios by using some moderate exponential time approximation algorithms.

Question 7.1.3. *Given $\delta(k)$, what is the minimum $\varepsilon(k)$ such that there is an $f(k) \cdot n^{\delta(k)}$-time algorithm that approximate Max-k-Set-Cover to $1 + \varepsilon(k)$.*

References

[AL86] Leonard M Adleman and Hendrik W Lenstra. Finding irreducible polynomials over finite fields. In *Proceedings of the eighteenth annual ACM symposium on Theory of computing*, pages 350–355. ACM, 1986.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998.

[ALR12] Aistis Atminas, Vadim V. Lozin, and Igor Razgon. Linear time algorithm for computing a small biclique in graphs without long induced paths. In *SWAT*, pages 142–152, 2012.

[AMS06] Noga Alon, Dana Moshkovitz, and Shmuel Safra. Algorithmic construction of sets for k-restrictions. *ACM Transactions on Algorithms*, 2(2):153–177, 2006.

[AMS11] Christoph Ambühl, Monaldo Mastrolilli, and Ola Svensson. Inapproximability results for maximum edge biclique, minimum linear arrangement, and sparsest cut. *SIAM Journal on Computing*, 40(2):567–596, 2011.

[AR01] Michael Alekhnovich and Alexander Razborov. Resolution is not automatizable unless W[P] is tractable. In *FOCS*, page 210. IEEE, 2001.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM (JACM)*, 45(1):70–122, 1998.

[AYZ95] Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. *J. ACM*, 42(4):844–856, 1995.

[BCC+10] Aditya Bhaskara, Moses Charikar, Eden Chlamtac, Uriel Feige, and Aravindan Vijayaraghavan. Detecting high log-densities: an $O(n^{1/4})$ approximation for densest k-subgraph. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 201–210. ACM, 2010.

[BCNS14] Cristina Bazgan, Morgan Chopin, André Nichterlein, and Florian Sikora. Parameterized inapproximability of target set selection and generalizations. In *Language, Life, Limits*, pages 11–20. Springer, 2014.

[BEKP13] Edouard Bonnet, Bruno Escoffier, Eun Jung Kim, and Vangelis Th. Paschos. On subexponential and fpt-time inapproximability. In *Parameterized and Exact Computation*, pages 54–65. Springer, 2013.

[BGK+96] László Babai, Anna Gál, János Kollár, Lajos Rónyai, Tibor Szabó, and Avi Wigderson. Extremal bipartite graphs and superpolynomial lower bounds for monotone span programs. In *STOC*, pages 603–611, 1996.

[BKRW15] Mark Braverman, Young Kun Ko, Aviad Rubinstein, and Omri Weinstein. Eth hardness for densest-k-subgraph with perfect completeness. *arXiv preprint arXiv:1504.08352*, 2015.

[BM11] Andrei A. Bulatov and Dániel Marx. Constraint satisfaction parameterized by solution size. In *ICALP (1)*, pages 424–436, 2011.

[BRFGL10] Daniel Binkele-Raible, Henning Fernau, Serge Gaspers, and Mathieu Liedloff. Exact exponential-time algorithms for finding bicliques. *Information Processing Letters*, 111(2):64–67, 2010.

[CEF12] Yijia Chen, Kord Eickmeyer, and Jörg Flum. The exponential time hypothesis and the parameterized clique problem. In *Parameterized and Exact Computation*, pages 13–24. Springer, 2012.

[CFK+16] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer International Publishing, 1 edition, 2016.

[CGG07] Yijia Chen, Martin Grohe, and Magdalena Grüber. On parameterized approximability. *Electronic Colloquium on Computational Complexity (ECCC)*, 14(106), 2007.

[CH10] Liming Cai and Xiuzhen Huang. Fixed-parameter approximation: Conceptual framework and approximability results. *Algorithmica*, 57(2):398–412, 2010.

[CHK13] Rajesh Chitnis, MohammadTaghi Hajiaghayi, and Guy Kortsarz. Fixed-parameter and approximation algorithms: A new look. In *Parameterized and Exact Computation - 8th International Symposium, IPEC 2013, Sophia Antipolis, France, September 4-6, 2013, Revised Selected Papers*, pages 110–122, 2013.

[CHKX04] Jianer Chen, Xiuzhen Huang, Iyad A. Kanj, and Ge Xia. Linear FPT reductions and computational lower bounds. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 212–221, 2004.

[Chv79] V. Chvátal. A greedy heuristic for the set-covering problem. *Mathematics of Operations*, 4(3):233 – 235, 1979.

[CK12] Jean-François Couturier and Dieter Kratsch. Bicolored independent sets and bicliques. *Information Processing Letters*, 112(8):329–334, 2012.

[CL15] Yijia Chen and Bingkai Lin. The constant inapproximability of the parameterized dominating set problem. *CoRR*, abs/1511.00075, 2015.

[Coo71] Stephen A Cook. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158. ACM, 1971.

[DF95] Rodney G. Downey and Michael R. Fellows. Fixed-parameter tractability and completeness I: basic results. *SIAM Journal on Computing*, 24(4):873–921, 1995.

[DF99] Rodney G. Downey and Michael R. Fellows. *Parameterized Complexity.* Springer-Verlag, 1999.

[DF13] Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity.* Texts in Computer Science. Springer, 2013.

[DFM06] Rodney G. Downey, Michael R. Fellows, and Catherine McCartin. Parameterized approximation problems. In *Parameterized and Exact Computation, Second International Workshop, IWPEC 2006, Zürich, Switzerland, September 13-15, 2006, Proceedings*, pages 121–129, 2006.

[DFMR08] Rodney G. Downey, Michael R. Fellows, Catherine McCartin, and Frances Rosamond. Parameterized approximation of dominating set problems. *Information Processing Letter*, 109(1):68–70, 2008.

[Din07] Irit Dinur. The PCP theorem by gap amplification. *Journal of the ACM (JACM)*, 54(3):12, 2007.

[DS02] Irit Dinur and Shmuel Safra. The importance of being biased. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 33–42. ACM, 2002.

[DS14] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 624–633, 2014.

[EGG08] Kord Eickmeyer, Martin Grohe, and Martin Gruber. Approximation of natural W[P]-complete minimisation problems is hard. In *Computational Complexity, 2008. CCC'08. 23rd Annual IEEE Conference on*, pages 8–18. IEEE, 2008.

[Erd34] Paul Erdős. A Theorem of Sylvester and Schur. *Journal of the Indian Mathematical Society*, s1-9(4):282–288, 1934.

[Erd59] Paul Erdős. Graph theory and probability. *canad. J. Math*, 11:34G38, 1959.

[Fei98] Uriel Feige. A threshold of ln n for approximating set cover. *The Journal of ACM*, 45(4):634–652, 1998.

[FG06] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Springer, 2006.

[FGL+96] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM (JACM)*, 43(2):268–292, 1996.

[FK04] Uriel Feige and Shimon Kogan. Hardness of approximation of the balanced complete bipartite subgraph problem. *Dept. Comput. Sci. Appl. Math., Weizmann Inst. Sci., Rehovot, Israel, Tech. Rep. MCS04-04*, 2004.

[FPK01] Uriel Feige, David Peleg, and Guy Kortsarz. The dense k-subgraph problem. *Algorithmica*, 29(3):410–421, 2001.

[GKL12] Serge Gaspers, Dieter Kratsch, and Mathieu Liedloff. On independent sets and bicliques in graphs. *Algorithmica*, 62(3-4):637–658, 2012.

[Gro07] Martin Grohe. The complexity of homomorphism and constraint satisfaction problems seen from the other side. *J. ACM*, 54(1), 2007.

[Hås96] Johan Håstad. Clique is hard to approximate within $\mathrm{n}^{\text{1-epsilon}}$. In *37th Annual Symposium on Foundations of Computer Science, FOCS '96, Burlington, Vermont, USA, 14-16 October, 1996*, pages 627–636, 1996.

[Hås01] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4):798–859, 2001.

[HKK13] MohammadTaghi Hajiaghayi, Rohit Khandekar, and Guy Kortsarz. Fixed parameter inapproximability for Clique and SetCover in time super-exponential in OPT. *CoRR*, abs/1310.2711, 2013.

[HKM13] Johan Hastad, Andrei Krokhin, and Dániel Marx. The Constraint Satisfaction Problem: Complexity and Approximability (Dagstuhl Seminar 12451). *Dagstuhl Reports*, 2(11):1–19, 2013.

[IPZ98] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? In *Foundations of Computer Science, 1998. Proceedings. 39th Annual Symposium on*, pages 653–662. IEEE, 1998.

[Joh74] David S. Johnson. Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9(3):256–278, 1974.

[Joh87] David S Johnson. The NP-completeness column: An ongoing guide. *Journal of Algorithms*, 8(3):438–448, 1987.

[Juk11] Stasys Jukna. *Extremal combinatorics: with applications in computer science.* Springer, 2011.

[Kar72] Richard M. Karp. Reducibility among combinatorial problems. In *Proceedings of a symposium on the Complexity of Computer Computations, held March 20-22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York.*, pages 85–103, 1972.

[Kho01] Subhash Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pages 600–609. IEEE, 2001.

[Kho06] Subhash Khot. Ruling out PTAS for graph min-bisection, dense k-subgraph, and bipartite clique. *SIAM Journal on Computing*, 36(4):1025–1071, 2006.

[KP93] Guy Kortsarz and David Peleg. On choosing a dense subgraph. In *Foundations of Computer Science, 1993. Proceedings., 34th Annual Symposium on*, pages 692–701. IEEE, 1993.

[KRS96] János Kollár, Lajos Rónyai, and Tibor Szabó. Norm-graphs and bipartite Turán numbers. *Combinatorica*, 16(3):399–406, 1996.

[KS16] Subhash Khot and Igor Shinkar. On hardness of approximating the parameterized clique problem. In *ITCS 2016*, 2016.

[Kut12] Konstantin Kutzkov. An exact exponential time algorithm for counting bipartite cliques. *Information Processing Letters*, 112(13):535–539, 2012.

[KZ96] Sandi Klavžar and Blaž Zmazek. On a Vizing-like conjecture for direct product graphs. *Discrete Mathematics*, 156(1-3):243–246, 1996.

[Lev73] Leonid A Levin. Universal sequential search problems. *Problemy Peredachi Informatsii*, 9(3):115–116, 1973.

[Lin15] Bingkai Lin. The parameterized complexity of k-biclique. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 605–615, 2015.

[Lov75] L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, 13(4):383 – 390, 1975.

[LRR14] Yuan Li, Alexander A. Razborov, and Benjamin Rossman. On the AC0 complexity of subgraph isomorphism. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 344–353, 2014.

[LY94] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *The Journal of ACM*, 41(5):960–981, 1994.

[Mar07] Dániel Marx. Can you beat treewidth? In *Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on*, pages 169–179. IEEE, 2007.

[Mar08] Dániel Marx. Parameterized complexity and approximation algorithms. *The Computer Journal*, 51(1):60–78, 2008.

[Mar13] Dániel Marx. Completely inapproximable monotone and antimonotone parameterized problems. *J. Comput. Syst. Sci.*, 79(1):144–151, 2013.

[Mos12] Dana Moshkovitz. The projection games conjecture and the NP-hardness of *ln* n-approximating Set-Cover. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 15th International Workshop, APPROX 2012, and 16th International Workshop, RANDOM 2012, Cambridge, MA, USA, August 15-17, 2012. Proceedings*, pages 276–287, 2012.

[Nie06] Rolf Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 2006.

[Ram19] S. Ramanujan. A proof of Bertrand's postulate. *Journal of the Indian Mathematical Society*, 11:181–182, 1919.

[Ros08] Benjamin Rossman. On the constant-depth complexity of k-clique. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 721–730. ACM, 2008.

[RS97] Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, El Paso, Texas, USA, May 4-6, 1997*, pages 475–484, 1997.

[Sch76] Wolfgang M. Schmidt. *Equations over Finite Fields An Elementary Approach(Lecture Notes in Mathematics Volume 536)*. Springer Berlin Heidelberg, 1976.

[Shp99] Igor Shparlinski. *Finite Fields: Theory and Computation*. Springer Netherlands, 1999.

[Sla97] Petr Slavík. A tight analysis of the greedy algorithm for set cover. *Journal of Algorithms*, 25(2):237 – 254, 1997.

[Ste74] S. K. Stein. Two combinatorial covering theorems. *Journal of Combinatorial Theory, Series A*, 16(3):391–397, 1974.

[Tre04] Luca Trevisan. Inapproximability of combinatorial optimization problems. *The Computing Research Repository*, 2004.

[Xav12] Eduardo C Xavier. A note on a maximum k-subset intersection problem. *Information Processing Letters*, 112(12):471–472, 2012.

[Zuc06] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 681–690. ACM, 2006.