The Parameterized Complexity of k-Biclique

Bingkai Lin *

Abstract

Given a graph G and a parameter k, the k-Biclique problem asks whether G contains a complete bipartite subgraph $K_{k,k}$. This is one of the most easily stated problems on graphs whose parameterized complexity has been long unknown. We prove that k-Biclique is $\mathbf{W}[\mathbf{1}]$-hard by giving an fpt-reduction from k-Clique to k-Biclique, thus solving this longstanding open problem.

Our reduction uses a class of bipartite graphs with a certain threshold property, which might be of some independent interest. More precisely, for positive integers n, s and t, we consider a bipartite graph $G = (A \dot\cup B, E)$ such that A can be partitioned into $A = V_1 \dot\cup V_2 \dot\cup, \cdots, \dot\cup V_n$ and for every s distinct indices $i_1, \cdots, i_s$, there exist $v_{i_1} \in V_{i_1}, \cdots, v_{i_s} \in V_{i_s}$ such that $v_{i_1}, \cdots, v_{i_s}$ have at least $t+1$ common neighbors in B; on the other hand, every $s+1$ distinct vertices in A have at most t common neighbors in B.

We prove that given such threshold bipartite graphs, we can construct an fpt-reduction from k-Clique to k-Biclique. Using the Paley-type graphs and Weil's character sum theorem, we show that for $t = (s+1)!$ and n large enough, such threshold bipartite graphs can be computed in polynomial time. One corollary of our reduction is that there is no $f(k) \cdot n^{o(k)}$ time algorithm to decide whether a graph contains a subgraph isomorphic to $K_{k!,k!}$ unless the Exponential Time Hypothesis (**ETH**) fails. We also provide a probabilistic construction with better parameters $t = \Theta(s^2)$, which indicates that k-Biclique has no $f(k) \cdot n^{o(\sqrt{k})}$-time algorithm unless 3-**SAT** with m clauses can be solved in $2^{o(m)}$-time with high probability. Besides the lower bound for exact computation of k-Biclique, our result also implies a dichotomy classification of the parameterized complexity of cardinality constraint satisfaction problems and the inapproximability of the maximum k-intersection problem.

*Department of Computer Science, The University of Tokyo. ERATO Kawarabayashi Large Graph Project.

1 Introduction

The Subgraph Isomorphism is a basic problem in algorithms and graph theory. Due to its generality, we do not expect it to have a polynomial time algorithm. However, this does not rule out the possibility that there exist efficient algorithms to solve this problem on some special class of graphs. For example, it is well known that whether G is a subgraph of H can be decided in $f(|G|) \cdot |H|^{O(tw(G))}$ time using the color-coding technique in [2], where $tw(G)$ denotes the tree-width of G and f is a computable function. Hence, if $\mathbf{C}$ is a class of graphs with tree-width bounded by some constant, the subgraph isomorphism problem with $G \in \mathbf{C}$ is fixed parameter tractable, and this is believed to be optimal. In [16], Martin Grohe conjectured that the subgraph embedding problem with $G \in \mathbf{C}$ is $\mathbf{W}[\mathbf{1}]$-hard if and only if $\mathbf{C}$ has unbounded tree-width. Under the assumption of $\mathbf{FPT} \neq \mathbf{W}[\mathbf{1}]$, this would imply that there is no $f(k) \cdot |H|^{O(1)}$-time algorithm to decide whether H 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. In other words, we can not prove Grohe's conjecture without answering the parameterized complexity of k-Biclique. Although k-Biclique is believed to be $\mathbf{W}[\mathbf{1}]$-hard, despite many attempts[6, 10, 15, 21], no **FPT**-reduction from k-Clique to k-Biclique has previously been found. Let us not fail to mention that a polynomial reduction is given in [19], however, since such reduction requires the size of the clique instance to be $|V(G)|/2$, it is not an fpt-reduction.

A possible line of attack 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 $\phi(u)$ have the same color; (2) if u and v are adjacent in G, then $\phi(u)$ and $\phi(v)$ are adjacent in H. It is not hard to see that Partitioned Subgraph Isomorphism problem on the graph class $\mathbf{C}$ is $\mathbf{W}[\mathbf{1}]$-hard if $\mathbf{C}$ has unbounded tree-width[16]. 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[22]. Unfortunately, this approach does not work for k-Biclique because any bipartite graph has a homomorphism to any of its edges.

Therefore, resolving the complexity of k-Biclique would significantly improve our understanding of the Subgraph Isomorphism problem. In addition, k-Biclique also has connections with the cardinality constraints satisfaction problem. Bulatov and Marx obtained a trichotomy classification of the parameterized complexity of the constraint satisfaction problem with cardinality constraints(CCSP) in [8]. They showed that for any set of relations closed under substitution of constants, CCSP with the relations restricted in Γ(denoted as CCSP(Γ)) is fixed parameterized tractable, Biclique-hard or **W**[**1**]-hard. By the well known dichotomy conjecture of Feder and Vardi, it is reasonable to believe that CCSP(Γ) is either **FPT** or **W**[**1**]-hard. Thus giving further incentive for the study of k-Biclique.

We remark that the parameterized complexity of k-biclique has received heavy attention from the parameterized complexity community[4, 8, 14, 16, 17]. It is the first problem on the "most infamous" list(page 677) in a new text book[11] by Downey and Fellows. "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!"

In the rest of this section, we state our main results with some further applications and corollaries.

1.1 Our Results

Theorem 1.1. *For any n-vertex graph G and positive integer k with $n^{\frac{6}{k+6}} > (k+6)!$, we can compute a graph G' in $O(n^{18})$-time such that G' contains a $K_{k',k'}$ if and only if G contains a K_k, where $k' = \Theta(k!)$.*

Corollary 1.1. *k-Biclique is **W**[**1**]-hard.*

Theorem 1.2. *For any n-vertex graph G and positive integer k with $k \geq 3$ and $n^{\frac{1}{(k+1)k^4}} > 2k^{4k^2+k+3}$, we can compute a random graph G' in $O(n^6)$-time such that, with probability at least $\frac{9}{10}$, G' contains a K_{k^2,k^2} if and only if G contains a K_k.*

The core of our reduction is the construction of a bipartite graph $H = (A \dot\cup B, E)$ with a (ℓ, h)-threshold property: every $k+1$ distinct vertices in A have at most ℓ common neighbors in B; while there exist many sets of k distinct vertices in A having at least h common neighbors in B, where $\ell < h$. An explicit construction of similar threshold bipartite graphs has been given in [5], in which they show that a certain fraction of k distinct vertices in A have this property(see Lemma 3.7 of [5]). Our contribution is proving that we can partition A into several sets and guarantee that for any k distinct sets, it is possible to choose one vertex from each set, the resulting k vertices satisfying the property.

1.2 Lower Bound for Computing k-Biclique

One corollary of our main results is the lower bound for exact computation of k-Biclique under the well-known **ETH**-conjecture made by Impagliazzo, Paturi and Zane [18]:

Conjecture 1.2. (Exponential Time Hypothesis) 3-**SAT** *cannot be solved in time $2^{o(m)}$, where m is the number of clauses in the input formula.*

The result in [9] implies that for any instance C of 3-**SAT** with m clauses, we can construct an instance (G, k) of k-Clique in $2^{o(m)}$-time such that C is an yes-instance of 3-**SAT** if and only if G contains a k-Clique. If the k-Clique problem has $f(k)\cdot n^{o(k)}$-time algorithm, we can solve such 3-**SAT** instance in $2^{o(m)}$-time. That is: Assuming **ETH**, k-Clique problem has no $f(k) \cdot n^{o(k)}$-time algorithm for any computable function f. With Theorem 1.1, we have the following lower bound: Assuming **ETH**, there is no $f(k) \cdot n^{o(k)}$-time algorithm to decide whether a given graph with order n contains a subgraph isomorphic to $K_{k!,k!}$.

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. If we assume a stronger version of **ETH**, then Theorem 1.2 yields a better lower bound for k-Biclique:

Corollary 1.3. *Unless m-clause 3-**SAT** can be solved in $2^{o(m)}$-time with high probability, there is no $f(k) \cdot n^{o(\sqrt{k})}$ algorithm for any computable function f to decide whether a given graph with order n contains a subgraph isomorphic to $K_{k,k}$.*

1.3 Maximum k-Intersection Problem

In our reduction from k-Clique to k-Biclique, we actually prove that

Theorem 1.3. *For an n-vertex graph G and a positive integer k with $\lceil n^{\frac{6}{k+6}} \rceil > (k+6)!$, let k' be the minimum integer such that $6 \mid k'+1$ and $k' \geq k$, let $s = \binom{k'}{2}$,*

we can compute a bipartite graph $H = (A \dot\cup B, E)$ 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.*

This gap allows us to deduce an inapproximation result for the Maximum k-Intersection Problem:

MAXIMUM k-INTERSECTION PROBLEM	
Input:	A family of sets $\{S_1, S_2, \cdots, S_n\}$ with $S_i \subseteq [n]$ and a number k .
Parameter:	k.
Problem:	Find k sets $S_{i_1}, \cdots, S_{i_k}$ with maximum $\lvert S_{i_1} \cap \cdots \cap S_{i_k} \rvert$

It is not hard to see that, our reduction implies

COROLLARY 1.4. *Assuming* **FPT** $\neq$ **W**[**1**]*, there is no $f(k) \cdot n^{O(1)}$-time algorithm approximating Maximum k-Intersection Problem with n^ϵ-approximation ratio for $\epsilon < \frac{6}{\sqrt{k}+1}$.*

The polynomial time inapproximability of Maxinum k-Intersection has been proved in [25] basing on the inapproximability of Maxinum Edge Biclique [3].

1.4 Cardinality Constraints Satisfaction Problem Fix a domain 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 $\mathbf{v} = v_1 v_2 \cdots v_r$, an assignment $\tau : V \to D$ satisfies a constraint $\langle \mathbf{v}, R \rangle$ if $(\tau(v_1), \cdots, \tau(v_r)) \in R$. 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 [8], Andrei A. Bulatov and Dániel Marx introduced two parameterized versions of CSP. More specifically, they assume that the domain contain a "free" value, say 0 and other non-zero values, which are "expensive". The goal is find an assignment with 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. 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.

For CSP with cardinality constraints, the situation is strange. An simple observation shows that the k-BICLIQUE problem can be express as a CCSP instance. Without lose 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 $C = \{\langle (v_1, v_2), R \rangle \mid v_1 v_2 \in E(G), R = \{(0,0), (1,0), (0,2)\}\}$, 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 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 [8], we finally obtain a dichotomy theorem for the parameterized complexity of CCSP(Γ):

THEOREM 1.4. *For every finite Γ closed under substitution of constants,* CCSP(Γ) *is either* **FPT** *or* **W**[**1**]*-hard.*

Organization of the Paper. The main idea of the reduction is presented in Section 3 after introducing the class of threshold bipartite graphs. To complete the reduction, we provide efficient constructions of the bipartite graph with threshold property. A probabilistic construction is given in Section 4, while the explicit construction can be found in Section 5. The explicit construction uses the Paley-type graph defined in [5] and a generalization of Lemma 3.8 in [5], whose proof is given in the Appendix. Finally, we discuss some interesting topics and open questions in Section 6.

2 Preliminaries

We use $\mathbb{N}$, $\mathbb{N}^+$ and $\mathbb{C}$ to denote the sets of nonnegative integers, positive integers and complex numbers respectively. For any number $n \in \mathbb{N}^+$, let $[n] := \{1, \ldots, n\}$. For any real numbers a, b, we use the notation $a \pm b$ to denote the numbers between $a - b$ and $a + b$. For any prime power $q = p^t$, $GF(q)$ is the Galois field with size q, $GF^\times(q)$ is the multiplicative group of $GF(q)$. For every set S we use $|S|$ to denote its size. Moreover, for any $t \in \mathbb{N}^+$, we let $\binom{S}{t}$ be the set of all t-element subsets

of S.

2.1 Parameterized Complexity We denote the alphabet $\{0,1\}$ by Σ and identify problems Q with subsets of Σ^*. A *parameterized problem* is a pair (Q,κ) consisting of a classical problem $Q \subseteq \Sigma^*$ and a polynomial time computable *parameterization* $\kappa : \Sigma^* \to \mathbb{N}$. For example, the parameterized clique problem is defined in the form:

p-CLIQUE	
Input:	A graph G and a positive integer k.
Parameter:	k.
Problem:	Does G contains a subgraph isomorphic to K_k?

An algorithm $\mathbb{A}$ is an fpt-*algorithm with respect to a parameterization* κ if for every $x \in \Sigma^*$ the running time of $\mathbb{A}$ on x is bounded by $f(\kappa(x))\cdot|x|^{O(1)}$ for a computable function $f : \mathbb{N} \to \mathbb{N}$. A parameterized problem is fixed-parameter tractable (or **FPT** for short) if it has an fpt-algorithm.

Let (Q,κ) and (Q',κ') be two parameterized problems. An fpt-*reduction* from (Q,κ) to (Q',κ') is a mapping $R : \Sigma^* \to \Sigma^*$ such that:

1. For every $x \in \Sigma^*$ we have $x \in Q$ if and only if $R(x) \in Q'$.
2. R is computable by an fpt-algorithm with respect to k;
3. There is a computable function $g : \mathbb{N} \to \mathbb{N}$ such that $\kappa'(R(x)) \le g(\kappa(x))$ for all $x \in \Sigma^*$.

A fpt-reduction is linear if $k' = O(k)$. We write $(Q,\kappa) \le^{\text{fpt}} (Q',\kappa')$ if there is an fpt-reduction from (Q,κ) to (Q',κ'); $(Q,\kappa) \equiv^{\text{fpt}} (Q',\kappa')$ if $(Q,\kappa) \le^{\text{fpt}} (Q',\kappa')$ and $(Q',\kappa') \le^{\text{fpt}} (Q,\kappa)$. Suppose $(Q,\kappa) \le^{\text{fpt}} (Q',\kappa')$, it is easy to see that if (Q',κ') is **FPT**, then so is (Q,κ); in particular, if p-CLIQUE $\le^{\text{fpt}} (Q,\kappa)$, then it follows that (Q,κ) is $\mathbf{W}[\mathbf{1}]$-hard (for the definition of $\mathbf{W}[\mathbf{1}]$-hardness, see [12, 14]). Obviously, if $(Q',\kappa') \le^{\text{fpt}} (Q,\kappa)$ and (Q',κ') is $\mathbf{W}[\mathbf{1}]$-hard, then so is (Q,κ).

2.2 Graphs Every graph $G = (V,E)$ is determined by a nonempty vertex set$V$ and an edge set $E \subseteq \binom{V}{2}$. Every nonempty subset $S \subseteq V(G)$ induces a subgraph $G[S]$ with the vertex set S and the edge set $E(G[S]) := \binom{S}{2} \cap E(G)$. And $G[S]$ is a *clique* in G, if for every distinct $u,v \in S$ we have $\{u,v\} \in E(G)$. A clique with k vertices is denoted as K_k or k-clique. A graph $G = (V,E)$ is *bipartite* if V admits a partition into two classes such that every edge has its ends in different classes. A *complete bipartite graph* or *biclique* is a bipartite graph such that every two vertices from different partition classes are adjacent. We use the notation $K_{s,t}$ to denote the complete bipartite graph with s vertices on one side and t vertices on the other side. In the bipartite graph $G = (A \dot\cup B, E)$, for $\mathbf{v} \subseteq A$, let $\Gamma(\mathbf{v}) = \{u \in B \mid \forall v \in \mathbf{v}, vu \in E\}$.

3 Reduction

We first define (s,t)-BICLIQUE, an imbalanced version of BICLIQUE. Then we prove that (s,t)-BICLIQUE and k-BICLIQUE are equivalent under linear fpt-reductions. Hence, to prove Theorem 1.1, we only need to prove Theorem 1.3. To this end, we introduce the threshold graphs. Theorem 1.3 then follows by the reduction in Lemma 3.3 and the efficient construction of threshold graphs given in Lemma 3.4. Also, Theorem 1.2 follows in analogy with Theorem 1.3, but calling on Lemma 4.4, a probabilistic analog to Lemma 3.4. Lemma 3.5 and Lemma 4.4 are proved in Section 4 and 5.

(s,t)-BICLIQUE	
Input:	A *bipartite* graph $G = (A \dot\cup B, E)$ and two positive integers s,t.
Parameter:	$s+t$.
Problem:	Find a $K_{s,t}$ in G with the left s vertices in A and the right t vertices in B.

LEMMA 3.1. *k-BICLIQUE $\equiv^{\text{fpt}}$ (s,t)-BICLIQUE and the reductions of both directions are linear.*

Proof. We need to check two directions:

1. k-BICLIQUE $\le^{\text{fpt}}$ (s,t)-BICLIQUE: given a k-BICLIQUE instance (G,k), construct a bipartite graph $B(G) = (A \dot\cup B, E)$, with A and B are two copies of $V(G)$ and $E = \{\{u,v\} \mid u \in A, v \in B, uv \in E(G)\}$. It is routine to check that $K_{k,k} \subseteq G \iff K_{k,k} \subseteq B(G)$, so $B(G)$ with $s := k, t := k$ is an instance of k-BICLIQUE$_{s,t}$;
2. (s,t)-BICLIQUE $\le^{\text{fpt}}$ k-BICLIQUE: suppose (G,s,t) is an instance of (s,t)-BICLIQUE, where $G = (A \dot\cup B, E)$ and $s \le t$. Construct a new bipartite graph G' by adding $t-s$ vertices into A and connect all of these new vertices with vertices in B. Then G' contains a $K_{t,t}$ iff G contains a $K_{s,t}$ with s vertices in A and t vertices in B.

DEFINITION 3.2. ((n,k,ℓ,h)-THRESHOLD PROPERTY) *Suppose $h > \ell$, a bipartite graph $G = (A \dot\cup B, E)$ with a partition $A = V_1 \dot\cup V_2 \dot\cup \cdots \dot\cup V_n$ 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.

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

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

$$\exists \boldsymbol{v} \in V_{i_1} \times \cdots \times V_{i_k}, |\Gamma(\boldsymbol{v})| \geq h$$

LEMMA 3.3. (REDUCTION) *Given an (n, k, ℓ, h)-threshold bipartite graph F. Let $s = \binom{k}{2}$. For any n vertices graph G, we can construct a new graph $H = (A \dot\cup B, E)$ in $n^{O(1)}$-time, such that:*

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

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

Proof. Suppose G is a graph with n vertices, our goal is to construct a bipartite graph $H = (A \dot\cup B, E)$ satisfying (H1) and (H2).

Let $V(G) = \{v_1, \cdots, v_n\}$, $F = (A' \dot\cup B', E') = ((V_1 \dot\cup V_2 \dot\cup \cdots \dot\cup V_n) \cup B', E')$. We associate to each V_i a vertex $v_i \in V(G)$ with the same index i. Let $\iota : A' \to V(G)$ be the function that for each $u \in V_i$, $\iota(u) = v_i$.

Then we construct the bipartite graph $H = (A \dot\cup B, E)$ with:

- $A = \{\{u_1, u_2\} \mid u_1, u_2 \in A', \{\iota(u_1), \iota(u_2)\} \in E(G)\}$;
- $B = B'$;
- $E = \{\{e, v\} \mid \{u_1, u_2\} = e \in A, v \in B, u_1 v \in E', u_2 v \in E'\}$.

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

1. If $K_k \subseteq G$, let us say $\{v_{a_1}, \cdots, v_{a_k}\}$ induces a K_k in G, then by (T2), there exists $u_{a_i} \in V_{a_i} (\forall i \in [k])$ such that $\{u_{a_1}, \cdots, u_{a_k}\}$ has at least h common neighbors in B', let $X = \{u_{a_1}, \cdots, u_{a_k}\}$ and $Y = \Gamma(X)$, we have $|X| = k$ and $|Y| \geq h$. Let $E_X = \binom{X}{2}$, since $\{\iota(u_{a_i}), \iota(u_{a_j})\} = \{v_{a_i}, v_{a_j}\} \in E(G)$ 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 $K_k \nsubseteq G$ but $\exists \mathbf{v} \in \binom{A}{s}$, s.t. $|\Gamma(\mathbf{v})| \geq \ell + 1$. Let $E_X = \mathbf{v} \subseteq A$, $Y = \Gamma(\mathbf{v}) \subseteq B$. We have $|E_X| = s$ and $|Y| \geq \ell + 1$. Consider $X = \{u \in A' \mid \exists e \in E_X\, u \in e\}$. By the definition of the edge set E, in the graph F, $Y \subseteq \Gamma(X)$. Since $|Y| = \ell + 1$ and F contains no $K_{k+1,\ell+1}$, 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 $u_1, u_2 \in X$, $\{u_1, u_2\} \in A \iff \{\iota(u_1), \iota(u_2)\} \in E(G)$. It follows that $\{\iota(u) \mid u \in X\}$ induces a K_k in G, this is impossible.

By Lemma 3.3, to prove Theorem 1.3, we only need to compute the threshold bipartite graphs efficiently. Our main technical lemma is:

LEMMA 3.4. *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 computed in $O(n^{18})$-time.*

Proof. [of Theorem 1.3] 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 3.4 on n and k', we obtain a $(n, k', (k'+1)!, \lceil (n+1)^{\frac{6}{k'+1}} \rceil)$-threshold bipartite graph. The result then follows from Lemma 3.3.

Theorem 1.1 can be easily deduced from Theorem 1.3 and Lemma 3.1. To prove Theorem 1.2, we show:

LEMMA 3.5. *For $k, h, n \in \mathbb{N}$ with $k \geq 3$, $h = k^2$ and $n^{\frac{2}{(k+1)k^{2h}}} > 2k^{k+1}h^{2h+1}$, we can compute in $O(n^6)$-time a bipartite random graph satisfying the $(n, k, h-1, h)$ threshold property with probability at least $\frac{9}{10}$.*

4 Probabilistic Construction

The Erdős-Rényi random graph $ER(n, p)$ is constructed on n vertices by joining every distinct pair of vertices independently with an edge with probability p. An interesting property of these random graphs is that there is a parameter $\texttt{thres}(H) = |V(H)|/|E(H)|$ such that if a graph H is *balanced* (i.e. every subgraph H' of H has $\texttt{thres}(H') \geq \texttt{thres}(H)$.), then for $p \gg n^{-\texttt{thres}(H)}$, $ER(n, p)$ contains a subgraph isomorphic to H with high probability; and for $p \ll n^{-\texttt{thres}(H)}$, $ER(n, p)$ contains no subgraph isomorphic to H with high probability (See [1] Chapter 4.4).

This suggests that we may construct the threshold bipartite graph defined in Section 3 using random graph. For $n \in \mathbb{N}$ and $p \in [0,1]$, define a bipartite random graph $G(n,p) = (A \dot\cup B, E)$ with $|A| = |B| = n$ and every pair of vertices $u \in A$ and $v \in B$ is joined by an edge with probability p, randomly and independently. We will show that with high probability $G(n,p)$ satisfies the $(n^\gamma, k, h-1, h)$-threshold property for some constant $\gamma \in (0,1)$. To bound the probability of $G(n,p)$ containing a subgraph $K_{k+1,h}$, we need the following lemma, which is a simple consequence of Markov's Inequality.:

LEMMA 4.1. *Let X be a nonnegative integral valued random variable, then* $\Pr[X > 0] \le E[X]$.

Let $p_\epsilon = n^{-\frac{(k+1+h+\epsilon)}{(k+1)h}}$, the value of ϵ will be determined later. It follows that:

LEMMA 4.2. $\Pr[K_{k+1,h} \subseteq G(n,p_\epsilon)] \le n^{-\epsilon}$.

Proof. Let X be the number of $K_{k+1,h}$ in $G(n,p)$, then

$$\begin{aligned} E[X] &= \binom{n}{k+1} \cdot \binom{n}{h} \cdot p_\epsilon^{(k+1)h} \\ &\le n^{(k+1+h)} \cdot n^{-(k+1+h+\epsilon)} \\ &= n^{-\epsilon} \end{aligned}$$

We have $\Pr[X > 0] \le E[X] \le n^{-\epsilon}$.

Hence, when $\epsilon > 0$, $n \to \infty$, $G(n,p_\epsilon)$ contains no $K_{k+1,h}$ with high probability.

Suppose $V_1, V_2, \cdots, V_k$ are k disjoint subsets of A and for each $i \in [k]$, $|V_i| = n^\alpha$, where $\alpha \in (0,1)$ is a constant. Let X_α be the number of $K_{k,h}$ in $G(n,p_\epsilon)$ with the restriction that each $V_i (i \in [s])$ contains exactly one vertex from the left side of such $K_{k,h}$. It is easy to see that:

$$\begin{aligned} E[X_\alpha] &= n^{\alpha k} \binom{n}{h} \cdot p_\epsilon^{kh} \\ &\ge n^{\alpha k} \cdot \frac{n^h}{h^h} \cdot p_\epsilon^{kh} \\ &= \frac{1}{h^h} \cdot n^{[\alpha k + h - \frac{k(k+1+h+\epsilon)}{(k+1)}]} \\ &= \frac{1}{h^h} \cdot n^{[\frac{h-(1-\alpha)k(1+k)-k\epsilon}{k+1}]} \end{aligned}$$

Let $\epsilon = \frac{1}{k}$ and $h = (1-\alpha)k(1+k)+2$, then $E[X_\alpha] = \Theta(n^{\frac{1}{1+k}})$. As n goes large, $E[X_\alpha] \to \infty$. Of course, $E[X_\alpha] \to \infty$ does not mean that $Pr[X_\alpha > 0] \to 1$. By the Chebyshev's Inequality, $Pr[X = 0]$ is upper bounded by:

THEOREM 4.1. (THEOREM 4.3.1 IN [1]) $\Pr[X = 0] \le \frac{Var[X]}{E[X]^2}$.

To show that $Pr[X_\alpha = 0]$ is very close to zero, we need to prove that $Var[X_\alpha]$ is $o(E[X_\alpha]^2)$. This can be easily deduced from the fact that $K_{k,h}$ is balanced(See [1] Chapter 4.4), however, since we want to upper bound the probability of $G(n,p_\epsilon)$ does not satisfy (T2), we need to show a slightly stronger result saying that $Var[X_\alpha]$ is $O(E[X_\alpha]^2) \cdot n^{-\Omega(1)}$.

Let $V_1 \times V_2 \times \cdots \times V_k = \{S_1, \cdots, S_\ell\}$, $\binom{B}{h} = \{T_1, \cdots, T_r\}$, where $\ell = n^{\alpha k}$ and $r = \binom{n}{h}$. We can rewrite X_α as $X_\alpha = \sum_{i\in[\ell], j\in[r]} X_{S_i,T_j}$, where X_{S_i,T_j} is the indicator random variable for event $A_{i,j} = [T_j \subseteq \Gamma(S_i)]$. Denote $(i,j) \sim (i',j')$ for $i,i' \in [\ell], j,j' \in [r]$ if $(i,j) \ne (i',j')$ and A_{ij}, $A_{i'j'}$ are not independent. Let $\Delta^* = \sum_{(i,j)\sim(i',j')} Pr[A_{ij}|A_{i'j'}]$, then $Var[X_\alpha] \le (1+\Delta^*)E[X_\alpha]$ and it is not hard to see that $(i,j) \sim (i',j')$ if and only if $|S_i \cap S_{i'}| > 0$, $|T_j \cap T_{j'}| > 0$ and $(i,j) \ne (i',j')$ (See the discussion in Chapter 4.3 of [1]). Then

$$\begin{aligned} \Delta^* &= \sum_{(i,j)\sim(i',j')} Pr[A_{ij}|A_{i'j'}] \\ &= \sum_{\substack{i\in[k],j\in[h] \\ i+j<k+h}} \binom{k}{i}\binom{h}{j} n^{\alpha(k-i)} \binom{n}{h-j} p_\epsilon^{(kh-ij)} \\ &\le k^k h^h \sum_{\substack{i\in[k],j\in[h] \\ i+j<k+h}} n^{\alpha(k-i)} n^{(h-j)} p_\epsilon^{(kh-ij)} \\ &\le k^k h^{2h} \sum_{\substack{i\in[k],j\in[h] \\ i+j<k+h}} E[X_\alpha] n^{-i\alpha-j} p_\epsilon^{-ij} \\ &= k^k h^{2h} \sum_{\substack{i\in[k],j\in[h] \\ i+j<k+h}} E[X_\alpha] n^{-i\alpha-j+ij\frac{(k+1+h+\epsilon)}{(k+1)h}} \\ &= k^k h^{2h} \sum_{\substack{i\in[k],j\in[h] \\ i+j<k+h}} E[X_\alpha] n^{\frac{ij}{(k+1)h}[-\frac{\alpha(k+1)h}{j} - \frac{(k+1)h}{i} + (k+1+h+\epsilon)]} \\ &\le k^k h^{2h} \sum_{\substack{i\in[k],j\in[h] \\ i+j<k+h}} E[X_\alpha] n^{\frac{ij}{(k+1)h}[-\alpha(k+1) - (1+\frac{1}{k})h + (k+1+h+\epsilon)]} \\ &\le k^{k+1} h^{2h+1} E[X_\alpha] n^{-\frac{1}{k(k+1)h}} \end{aligned}$$

We have

LEMMA 4.3. *For $n^{\frac{1}{(k+1)k^2h}} > 2k^{k+1}h^{2h+1}$,* $\Pr[X_\alpha = 0] \le n^{-\frac{1}{k^2h}}$.

Proof.

$$\begin{aligned}\Pr[X_\alpha = 0] &\le \frac{Var[X_\alpha]}{E[X_\alpha]^2} \\ &\le \frac{(1+\Delta^*)}{E[X_\alpha]} \\ &\le 2k^{k+1}h^{2h+1}n^{-\frac{1}{k(k+1)h}} \\ &\le n^{-\frac{1}{k(k+1)h}+\frac{1}{(k+1)k^2h}} \\ &= n^{-\frac{1}{k^2h}}\end{aligned}$$

Now suppose $U_1, \cdots, U_k$ are k disjoint subsets of A with $|U_i| = n^\beta (i \in [k])$, where $\alpha < \beta < 1$. We know that each U_i can be further partitioned into $U_i = V_{i1} \dot\cup \cdots \dot\cup V_{im}$ with $m = n^{\beta-\alpha}$ and for all $j \in [m]$, $|V_{ij}| = n^\alpha$. Let X_β be the number of $K_{k,h}$ in $G(n, p_\epsilon)$ such that each U_i contains exactly one vertex from the left side of such $K_{k,h}$ and for $j \in [m]$, $X_{\beta,j}$ be the number of $K_{k,h}$ in $G(n, p_\epsilon)$ such that for each $i \in [k]$, $V_{i,j}$ contains exactly one vertex from the left side of such $K_{k,h}$. It is not hard to see that $\Pr[X_{\beta,j} = 0] = \Pr[X_\alpha = 0]$, and for any distinct $j, j' \in [m]$, $X_{\beta,j}$ and $X_{\beta,j'}$ are independent. It follows that:

$$\begin{aligned}\Pr[X_\beta = 0] &\le Pr[X_{\beta,1} = 0, \cdots, X_{\beta,m} = 0] \\ &= Pr[X_\alpha = 0]^m \\ &\le n^{-(\frac{n^{\beta-\alpha}}{k^2h})}\end{aligned}$$

Given a bipartite random graph $G(n, p_\epsilon) = (A \dot\cup B, E)$, we partition A into $n' = n^{1-\beta}$ sets $A = U_1 \dot\cup \cdots \dot\cup U_{n'}$ with $|U_i| = n^\beta$. Then the probability that $G(n, p_\epsilon)$ with such partition does not satisfy (T2) for parameter $(n', k, h-1, h)$ is bounded by

$$\Pr[G(n, p_\epsilon) \text{ does not satisfy } (T2)] \le n^{(1-\beta)k} n^{-(\frac{n^{\beta-\alpha}}{k^2h})}$$

It follows that

$$\begin{aligned}&\Pr[G(n, p_\epsilon) \text{ does not satisfy T1 or T2}] \\ &\le n^{-\epsilon} + n^{(1-\beta)k-(\frac{n^{\beta-\alpha}}{k^2h})}\end{aligned}$$

So when $n \to \infty$, $G(n, p_\epsilon)$ is an $(n', k, h-1, h)$ threshold bipartite graph with high probability. We have

LEMMA 4.4. *For any $0 < \alpha < \beta < 1, \epsilon = \frac{1}{k}$, and $n^{\frac{1}{(k+1)k^2h}} > 2k^{k+1}h^{h+1}$, $G(n, n^{-\frac{(k+1+h+\epsilon)}{(k+1)h}})$ satisfies the $(n^{1-\beta}, k, h-1, h)$ threshold property with probability at least $1 - n^{-\epsilon} - n^{(1-\beta)k-(\frac{n^{\beta-\alpha}}{k^2h})}$.*

Proof. [of Lemma 3.5] Let $\alpha = \frac{k+2}{k(k+1)}$, we have $h = k^2 = (1-\alpha)k(1+k)+2$. When $k \ge 3$, we have $\alpha < \frac{1}{2}$, let $\beta = \frac{1}{2}$, $\theta = \frac{1}{1-\beta} = 2$, for $n^{\frac{2}{(k+1)k^2h}} > 2k^{k+1}h^{2h+1}$, the random graph $G(n^\theta, n^{-\theta\frac{(k+1+h+\epsilon)}{(k+1)h}})$ satisfies the $(n, k, h-1, h)$ threshold property with probability at least $1 - n^{-\frac{2}{k}} - n^{k-2(\frac{n^{\beta-\alpha}}{k^4})} \ge 1 - 2n^{-\frac{2}{k}} > \frac{9}{10}$. It is not hard to see that such random graph can be generated in $O(n^6)$-time by a probabilistic Turing machine, hence proving Lemma 3.5.

5 Explicit Construction

DEFINITION 5.1. (PALEY-TYPE GRAPH) *For any prime power $q = p^t$ and $d \mid q-1$, $G(q, d) := (A \dot\cup B, E)$ is a Paley-type bipartite graph with*

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

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 5.2. *$G(q, d)$ can be computed in $O(q^3)$ time.*

The Paley-type graphs have many nice properties, the following one is proved in [20, 5]:

THEOREM 5.1. (THEOREM 5.1 IN [5]) *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 5.3. (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 $a_1, \cdots, a_k$ be distinct elements in $GF^\times(q)$, g be the generator of $GF^\times(q)$, for each $i \in [s]$, denote $V_i := \{g^{i+s}, g^{i+2s}, \cdots, g^{i+sr}\}$, 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 5.3 generalizes Lemma 3.8 in [5] by restricting the solutions to any subset $V_j (j \in [s])$. If we set $s = 1$, then we obtain Lemma 3.8 in [5]. The intuition behind Lemma 5.3 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 5.3.

LEMMA 5.4. *For any $p, r, s, t \in \mathbb{N}^+$ with p is prime, $\frac{s}{p-1} + 1 \leq \sqrt{p^{t+1}}$ and $p^{t+1} - 1 = rs$. Let g be the generator of $GF^{\times}(p^{t+1})$, for each $i \in [s]$, denote $V_i := \{g^{i+s}, g^{i+2s}, \cdots, g^{i+sr}\}$. Then in the Paley-type bipartite graph $G(p^{t+1}, p-1) = (A \dot\cup B, E)$, for any t distinct indices $a_1, a_2, \cdots, a_t \in [s]$, there exist $\boldsymbol{v} \in V_{a_1} \times \cdots \times V_{a_t}$, such that $|\Gamma(\boldsymbol{v})| \geq p$.*

Proof. Fix t distinct indices $a_1, a_2, \cdots, a_t \in [s]$. Consider the sets $S = V_{a_1} \times \cdots \times V_{a_t}$ and $\Gamma\langle S\rangle = \{\{\mathbf{v}, u\} \mid \mathbf{v} \in S, u \in B, u \in \Gamma(\mathbf{v})\}$. Since $\frac{s}{p-1} + 1 \leq \sqrt{p^{t+1}}$, apply Lemma 5.3 with $q \leftarrow p^{t+1}$ $d \leftarrow p-1$ $k \leftarrow 1$, each elements in $GF^{\times}(p^{t+1})$ has at least

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

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

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

In the construction of the threshold bipartite graphs, we also use the famous Bertrand's Postulate from number theory, whose proof can be found in [23, 13].

Proof. [of Lemma 3.4] 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 $s \geq n$ 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 $V_i := \{g^{i+s}, g^{i+2s}, \cdots, g^{i+rs}\}$, 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 computed 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 5.1, $G(p^{k+1}, p-1)$ 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 5.4 with $t \leftarrow k$, we have for any k distinct indices $a_1, a_2, \cdots, a_k \in [s]$, there exist $v_{a_i} \in V_{a_i}$ $(\forall i \in [k])$ such that $v_{a_1}, \cdots, v_{a_k}$ 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 a $(n, k, (k+1)!, \lceil (n+1)^{\frac{1}{\ell}} \rceil)$ threshold bipartite graph.

6 Conclusions

In Section 4, we have seen that with high probability the bipartite random graph $G(n, n^{-\frac{(s+t+\epsilon)}{st}})$ for $s \leq t$ contains no subgraph isomorphic to $K_{s,t}$. Notice that such graph also has nearly $n^{(2-\frac{1}{s}-\frac{1}{t}-O(\frac{1}{st}))}$ number of edges. In extremal graph theory, the famous Zarankiewicz problem asks for $K_{s,t}$-free graphs with $\Omega(n^{(2-\frac{1}{s})})$ edges. As far as we know, the explicit construction for $s > 3$ is rare[7]. It seems that $h \geq \Omega(k^2)$ is required in the probabilistic construction of $(n, k, h-1, h)$-threshold bipartite graph. Does any $(n, k, h-1, h)$-threshold bipartite graph G exists for $h = \Theta(k)$ and $|G| = n^{O(1)}$?

It is still open whether there exists any $f(k) \cdot n^{o(k)}$-time algorithm solving k-BICLIQUE. Our reduction causes a quadratic blow-up of the parameter. Even if the (n, k, k^2, k^2+1)-threshold bipartite graph can be computed in deterministic fpt time, we could only show that k-BICLIQUE has no $f(k) \cdot n^{o(\sqrt{k})}$ algorithm under **ETH**. 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[22]. 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 [22] 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.

Notice that the class of bipartite graphs with threshold property allows us to distinguish every s vertices from $s+1$ vertices in some way. Can we exploit this property to prove the hardness of the subgraph isomorphic problem on other graph classes?

7 Acknowledgments

The authors would like to thank Yijia Chen, Hiroshi Imai and the anonymous reviewers for their valuable comments and suggestions to improve the paper.

Appendix: Proof of the Intersection Lemma

Some definitions:

DEFINITION 7.1. (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 7.2. *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 7.3. (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 7.1. (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 [24], page 43, Theorem 2C')

REMARK 7.4. *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 7.5. *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 5.3] 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:

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

We only need to estimate S. Using Lemma 7.5, 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

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

References

[1] N. Alon and J. H. Spencer. *The probabilistic method*. John Wiley & Sons, 2008.

[2] N. Alon, R. Yuster, and U. Zwick. Color-coding. *J. ACM*, 42(4):844–856, 1995.

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

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

[5] L. Babai, A. Gál, J. Kollár, L. Rónyai, T. Szabó, and A. Wigderson. Extremal bipartite graphs and superpolynomial lower bounds for monotone span programs. In *STOC*, pages 603–611, 1996.

[6] D. Binkele-Raible, H. Fernau, S. Gaspers, and M. Liedloff. Exact exponential-time algorithms for finding bicliques. *Information Processing Letters*, 111(2):64–67, 2010.

[7] P. V. Blagojević, B. Bukh, and R. Karasev. Turán numbers for $K_{s,t}$-free graphs: Topological obstructions and algebraic constructions. *Israel Journal of Mathematics*, 197(1):199–214, 2013.

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

[9] J. Chen, X. Huang, I. A. Kanj, and G. Xia. Linear FPT reductions and computational lower bounds. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 212–221. ACM, 2004.

[10] J.-F. Couturier and D. Kratsch. Bicolored independent sets and bicliques. *Information Processing Letters*, 112(8):329–334, 2012.

[11] F. M. R. Downey, Rodney G. *Fundamentals of Parameterized Complexity*. Springer, 2013.
[12] R. Downey and M. Fellows. *Parameterized Complexity*. Springer-Verlag, 1999.
[13] P. Erdős. A Theorem of Sylvester and Schur. *Journal of the Indian Mathematical Society*, s1-9 (4):282C288, 1934.
[14] J. Flum and M. Grohe. *Parameterized Complexity Theory (Texts in Theoretical Computer Science. An EATCS Series)*. Springer-Verlag New York, Inc., 2006.
[15] S. Gaspers, D. Kratsch, and M. Liedloff. On independent sets and bicliques in graphs. *Algorithmica*, 62(3-4):637–658, 2012.
[16] M. Grohe. The complexity of homomorphism and constraint satisfaction problems seen from the other side. *J. ACM*, 54(1), 2007.
[17] J. Hastad, A. Krokhin, and D. Marx. The Constraint Satisfaction Problem: Complexity and Approximability (Dagstuhl Seminar 12451). *Dagstuhl Reports*, 2(11):1–19, 2013.
[18] R. Impagliazzo, R. Paturi, and F. Zane. Which problems have strongly exponential complexity? In *Foundations of Computer Science, 1998. Proceedings. 39th Annual Symposium on*, pages 653–662. IEEE, 1998.
[19] D. S. Johnson. The NP-completeness column: An ongoing guide. *Journal of Algorithms*, 8(3):438–448, 1987.
[20] J. Kollár, L. Rónyai, and T. Szabó. Norm-graphs and bipartite Turán numbers. *Combinatorica*, 16(3):399–406, 1996.
[21] K. Kutzkov. An exact exponential time algorithm for counting bipartite cliques. *Information Processing Letters*, 112(13):535–539, 2012.
[22] D. Marx. Can you beat treewidth? In *Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on*, pages 169–179. IEEE, 2007.
[23] S. Ramanujan. A proof of Bertrand's postulate. *Journal of the Indian Mathematical Society*, 11:181C182, 1919.
[24] W. M. Schmidt. *Equations over Finite Fields An Elementary Approach(Lecture Notes in Mathematics Volume 536)*. Springer Berlin Heidelberg, 1976.
[25] E. C. Xavier. A note on a maximum k-subset intersection problem. *Information Processing Letters*, 112(12):471–472, 2012.