Bicovering: Covering Edges With Two Small Subsets of Vertices

Amey Bhangale*[1], Rajiv Gandhi†[2], Mohammad T. Hajiaghayi‡[3], Rohit Khandekar[4], and Guy Kortsarz§[5]

1 **Department of Computer Science, Rutgers University, New Brunswick, USA**
amey.bhangale@rutgers.edu
2 **Department of Computer Science, Rutgers University, Camden, USA**
rajivg@camden.rutgers.edu
3 **Department of Computer Science, University of Maryland, College Park, USA**
hajiagha@cs.umd.edu
4 **KCG holdings Inc., Jersey City, USA**
rkhandekar@gmail.com
2 **Department of Computer Science, Rutgers University, Camden, USA**
guyk@camden.rutgers.edu

Abstract

We study the following basic problem called Bi-Covering. Given a graph $G(V, E)$, find two (not necessarily disjoint) sets $A \subseteq V$ and $B \subseteq V$ such that $A \cup B = V$ and that every edge e belongs to either the graph induced by A or to the graph induced by B. The goal is to minimize $\max\{|A|, |B|\}$. This is the most simple case of the Channel Allocation problem [Gandhi et. al, Networks, 2006]. A solution that outputs $V, \emptyset$ gives ratio at most 2. We show that under the similar *Strong* Unique Game Conjecture by [Bansal-Khot, FOCS, 2009] there is no $2 - \epsilon$ ratio algorithm for the problem, for any constant $\epsilon > 0$.

Given a bipartite graph, Max-bi-clique is a problem of finding largest $k \times k$ complete bipartite sub graph. For Max-bi-clique problem, a constant factor hardness was known under random 3-SAT hypothesis of Feige [Feige, STOC, 2002] and also under the assumption that $\mathrm{NP} \not\subseteq \cap_{\epsilon>0} \mathsf{BPTIME}(2^{n^\epsilon})$ [Khot, SIAM J. on Comp., 2011]. It was an open problem in [Ambühl et. al., SIAM J. on Comp., 2011] to prove inapproximability of Max-bi-clique assuming weaker conjecture. Our result implies similar hardness result assuming the Strong Unique Games Conjecture.

On the algorithmic side, we also give better than 2 approximation for Bi-Covering on numerous special graph classes. In particular, we get 1.876 approximation for Chordal graphs, exact algorithm for Interval Graphs, $1 + o(1)$ for Minor Free Graph, $2 - 4\delta/3$ for graphs with minimum degree δn, $2/(1+\delta^2/8)$ for δ-vertex expander, 8/5 for Split Graphs, $2 - (6/5) \cdot 1/d$ for graphs with minimum constant degree d etc. Our algorithmic results are quite non-trivial. In achieving these results, we use various known structural results about the graphs, combined with the techniques that we develop tailored to getting better than 2 approximation.

1998 ACM Subject Classification G.2.2 Graph Algorithms

Keywords and phrases Bi-covering, Unique Games, Max Bi-clique.

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.6

* Research supported in part by NSF grant CCF-1253886.
† Supported in part by NSF grant number 1218620.
‡ Supported in part by NSF CAREER award CCF-1053605, NSF BIGDATA grant IIS-1546108, NSF AF:Medium grant CCF-1161365, DARPA GRAPHS/AFOSR grant FA9550-12-1-0423, and another DARPA SIMPLEX grant.
§ Supported in part by NSF grant number 1218620 and by NSF grant 1540547.

43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016).
Editors: Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi;
Article No. 6; pp. 6:1–6:12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

We study the Bi-Covering problem - Given a graph $G(V, E)$, find two (not necessarily disjoint) sets $A, B \subseteq V$ such that $A \cup B = V$ and that every edge $e \in E$ belongs to either the graph induced by A or to the graph induced by B. The goal is to minimize $\max\{|A|, |B|\}$.

The problem we study is closely related to the problem of *Channel Allocation* which was studied in [10]. The Channel Allocation Problem can be described as follows: there is a universe of topics, a fixed number of channels and a set of requests where each request is a subset of topics. The task is to send a subset of topics through each channel such that each request is satisfied by set of topics from one of the channel i.e. for every request there must exists at least one channel such that the set of topics present in that channel is a superset of the set of topics from the request. Of course, one can achieve this task trivially by sending all topics through one channel. But, the optimization version of Channel Allocation Problem asks for a way to satisfy all the request by minimizing the maximum number of topics sent through a channel.

Any *connected* undirected graph $G(V, E)$ on n vertices and m edges along with an integer k can be viewed as a special case of channel allocation problem - The set of topics is a set of n vertices, each edge represents a request, where the requested set of topics corresponding to an edge is a pair of its endpoints and the number of channels is k. If we fix the number of channels to $k = 2$ then the optimization problem exactly corresponds to the Bi-Covering problem. Specifically, the optimization problem asks for two subsets A and B of V minimizing $\max\{|A|, |B|\}$ such that $A \cup B = V$ and every edge is totally contained in a graph induced by either A or B.

2 Our Results

Getting 2 approximation for Bi-Covering problem is trivial (by setting A = B = V). We show that Bi-Covering problem is hard to approximate within any factor strictly less than 2 assuming a strong Unique Games Conjecture (UGC) similar to the one in [5] (see Conjecture 12).

▶ **Theorem 1.** *Let $\epsilon > 0$ be any small constant. Assuming a strong Unique Games Conjecture (Conjecture 12), given a graph $G(V, E)$, it is* NP*-hard to distinguish between following two cases:*
1. *G has* Bi-Covering *of size at most $(1/2 + \epsilon)|V|$.*
2. *Any* Bi-Covering *of G has size at least $(1 - \epsilon)|V|$.*

In particular, it is NP-hard (assuming strong UGC) to approximate Bi-Covering *within a factor $2 - \epsilon$ for every $\epsilon > 0$.*

Given this structural hardness result, we get a $\frac{3}{2} - \epsilon$ hardness of Bi-Covering restricted to bipartite graphs by transforming a hard instance from Theorem 1 into a bipartite graph in a natural way (getting a $\frac{3}{2}$-approximation is easy - given a bipartite graph on X and Y with $|X| \geq |Y|$, one can take arbitrary partition X into two equal sized parts X_1 and X_2 and set the Bi-Covering to be $X_1 \cup Y$ and $X_2 \cup Y$).

▶ **Theorem 2.** *Assuming the strong Unique Games Conjecture, for every $\epsilon > 0$,* Bi-Covering *is* NP*-hard to approximate within a factor $\frac{3}{2} - \epsilon$ for bi-partite graphs.*

Our Theorem 1 implies hardness result for the following well known problem:

Max-Bi-Clique problem is as follows:

Input: A bipartite graph $G(X, Y, E)$ with $|X| = |Y| = n$.
Output: Find largest k such that there exists two subsets $A \subseteq X, B \subseteq Y$ of size k and the graph induced on (A, B) is a complete bipartite graph.

Inapproximability of Max-Bi-Clique problem has been studied extensively [1, 6, 8, 13]. Feige[8] showed that using an assumption of average case hardness of 3SAT instance, Max-Bi-Clique cannot be approximated within any constant factor in polynomial time (and hence within n^δ for some $\delta > 0$ using known amplification technique [1, 6]). Feige-Kogan [9] showed that assuming $SAT \notin DTIME(2^{n^{3/4+\epsilon}})$ there is no $2^{(\log n)^\delta}$ approximation for Max-Bi-Clique. They also showed that it is NP-hard to approximate Max-Bi-Clique within any constant factor assuming Max$-$Clique (finding a maximum sized clique in a graph) does not have a $n/2^{c\sqrt{\log n}}$-approximation. Khot [13] later proved a similar inapproximability result but assuming NP $\nsubseteq \cap_{\epsilon>0}$ BPTIME(2^{n^ϵ}) using a *quasi-random* PCP. It is an important open problem to extend similar hardness results based on weaker complexity assumptions [2]. In particular, it is still not known if UGC implies a constant factor hardness for Max-Bi-Clique. A straightforward corollary from Theorem 1 (see 4.2.2) implies that we get similar hardness results for Max-Bi-Clique based on Conjecture 12.

▶ **Corollary 3.** *Assuming strong Unique Games Conjecture, it is* NP*-hard to approximate* Max-Bi-Clique *within any constant factor.*

As mentioned above, the hardness factor can be boosted to n^δ for some $\delta > 0$ using known techniques. (such as described in [1, 6])

UGC and strong UGC

Unique games conjecture so far helped in understanding the tight inapproximability factors of many problems including, but not limited to, Vertex Cover [14], optimal algorithm for every Max-CSP[16], Ordering CSPs[11], characterizing strong approximation resistance of CSPs[15] etc. The inherent difficulty in showing hardness results assuming Unique Games Conjecture for the problems that we study is that we need some kind of expansion property on the unique games instance which it lacks. It is shown that Unique Games are easy when the constraint graph is an expander[4]. In general, in [3] it is shown that Unique Games are easy when a normalized adjacency matrix of a constraint graph has very few eigenvalues close to 1. So the natural direction is to modify the unique games instance to get some expansion property but weak enough so that it is not tractable by the techniques of [4], [3]. A similar Strong Unique Games Conjecture, which has a *weak expansion property*, has been used earlier in [5] and [17] to show inapproximability results for minimizing weighted completion time on a single machine with precedence constraints and minimizing makespan in precedence constrained scheduling on identical machines respectively. Our result adds another interesting implication of Unique Games Conjecture with weak expansion property, namely inapproximability of Max-Bi-Clique and Bi-Covering. We hope that our results will help motivate study of Strong Unique Games Conjecture and ultimately answering the question about its equivalence to the Unique Games Conjecture.

Algorithmic Results

We give better than 2 approximation for Bi-Covering on numerous special graph classes.

Graph types. A δ*-vertex expander* is a graph so that for every S of size $|S| \leq n/2$, $N_1(S) \geq \delta n$, where $N_1(S)$ is the set of neighbors of S not in S. A *chordal graph* is a graph that does not contain a cycle of size at least 4 as an induced subgraph. A *split graph* is a graph whose vertex set is a union of a Clique and and independent set, with arbitrarily connections between the clique and the independent set.

A minor of a graph is any subgraph G' that can be derived from G by contracting and removing edges. A *minor free graph* is a graph that does not contain some constant size graph H as a minor.

An *interval graph* is the intersection graph of a family of intervals on the real line. It has one vertex for each interval in the family, and an edge between every pair of vertices corresponding to intervals that intersect.

The algorithmic results can be summarized in the following theorem.

▶ **Theorem 4.** *The* Bi-Covering *problem admits polynomial time algorithms that attain the following ratios (Graph type: approximation ratio):*
1. *Chordal graphs :* 1.876.
2. *Interval Graphs: exact $O(n^5)$ time algorithm.*
3. *Minor Free Graph:* $1 + o(1)$.
4. *Graph with minimum degree δn:* $2 - 4\delta/3$.
5. *δ-vertex expander Graph:* $2/(1 + \delta^2/8)$.
6. *Split Graphs :* $8/5$.
7. *Graphs with minimum degree d:* $2 - (6/5) \cdot 1/d$.

Our algorithms are quite non-trivial. Most of our algorithmic results relies on the fact that if we can find two disjoint sets each of size at least ϵn with no edges in between, then this itself gives $2 - \epsilon$ approximation. To get better bound on ϵ in some special cases we use known theorems related to the structural results of graphs, size of separator, lower bound on independent set size etc. In some of the cases, we create a bipartite graph from a given graph instance and show that the vertex cover in the bipartite graph is small. We then use the bound on the size of vertex cover to find a better bi-covering of the edges in a graph.

3 Organization

In Section 4, we prove the main inapproximability of Bi-Covering and related problems. We refer to the full version of the paper for algorithmic results.

4 Inapproximability of Bi-Covering

The Bi-Covering problem is:

Input: A graph $G(V, E)$
Output: Two subsets $A, B \subseteq V$ such that $A \cup B = V$ and every edge $(u, v) \in E$ either $\{u, v\} \subseteq A$ or $\{u, v\} \subseteq B$. Minimize $\max\{|A|, |B|\}$.

The optimal value of a Bi-Covering on instance $G(V, E)$ is always at least $|V|/2$ and hence getting a 2-approximation for this problem is *trivial* by setting $A = V$ and $B = \emptyset$. In order to beat 2-approximation, one should be able to solve the following weaker problem.

Problem

For small enough $\epsilon > 0$, given an undirected graph $G(V, E)$, distinguish between the following two cases:

1. There exists two disjoint sets $A, B \subseteq V$, $|A|, |B| \geq (1/2 - \epsilon)|V|$ such that there are no edges between A and B.
2. There exists no two disjoint sets $A, B \subseteq V$ $|A|, |B| \geq \epsilon|V|$ such that there are no edges between A and B.

In this section, we show that it is UG-Hard to distinguish between (1) and (2) for any constant $\epsilon > 0$ proving Theorem 1.

4.1 Preliminaries

Let q be any prime for convenience. We are interested in space of functions from $\mathbb{F}_q^n$ to $\mathbb{R}$. Define inner product on this space as $\langle f, g \rangle = \frac{1}{q^n} \sum_{x \in \mathbb{F}_q^n} f(x)g(x)$. Let ω_q be the q^{th} root of unity. For a vector $\alpha \in \mathbb{F}_q^n$, we will denote α_i the i^{th} coordinate of vector α. The *Fourier decomposition* of a function $f : \mathbb{F}_q^n \to \mathbb{R}$ is given as

$$f(x) = \sum_{\alpha \in \mathbb{F}_q^n} \hat{f}(\alpha)\chi_\alpha(x)$$

where $\chi_\alpha(x) := \omega_q^{\langle \alpha, x \rangle}$ and a *Fourier coefficient* $\hat{f}(\alpha) := \langle f, \chi_\alpha \rangle$.

▶ **Definition 5** (Symmetric Markov Operator). Symmetric Markov operator on $\mathbb{F}_q$ can be thought of as a random walk on an undirected graph with the vertex set $\mathbb{F}_q$. It can be represented as a $q \times q$ matrix T where (i, j) th entry is the probability of moving to vertex j from i.

▶ **Definition 6.** For a symmetric Markov operator T, let $1 = \lambda_0 \geq \lambda_1 \geq \lambda_2 \ldots \geq \lambda_{q-1}$ be the eigenvalues of T in a non-increasing order. The spectral radius of T, denoted by $r(T)$, is defined as:

$$r(T) = \max\{|\lambda_1|, |\lambda_{q-1}|\}.$$

For a Markov operator T the condition $r(T) < 1$ is equivalent to saying that the induced regular graph (self-loop allowed) on $\mathbb{F}_q$ is non-bipartite and connected.

For T as above, we also define a Markov operator $T^{\otimes n}$ on $[q]^n$ in a natural way i.e applying a Markov operator $T^{\otimes n}$ to $x \in [q]^n$ is same as applying the Markov operator T on each x_i independently. Note that if T is symmetric then $T^{\otimes n}$ is also symmetric and $r(T^{\otimes n}) = r(T)$.

▶ **Definition 7** (Influence). Let $f : \mathbb{F}_q^n \to \mathbb{R}$ be a function. the influence of the $i'th$ variable on f, denoted by $\mathbf{Inf}_i(f)$ is defines as:

$$\mathbf{Inf}_i(f) = \mathbb{E}[\mathbf{Var}_{x_i}[f(x)|x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]]$$

where $x_1, \ldots, x_n$ are uniformly distributed. In terms of Fourier coefficients, it has the following formula:

$$\mathbf{Inf}_i(f) = \sum_{\alpha_i \neq 0} \hat{f}(\alpha)^2.$$

The low-level (level k) influence of $i'th$ variable is defined as:

$$\mathbf{Inf}_i^{\leq k}(f) = \sum_{\alpha_i \neq 0, |\alpha| \leq k} \hat{f}(\alpha)^2.$$

where $|\alpha|$ is the number of non-zero co-ordinates in α.

We will need the following Gaussian stability measure in our analysis:

▶ **Definition 8.** Let $\phi : \mathbb{R} \to [0,1]$ be the cumulative distribution function of the standard Gaussian random variable. For a parameter $\rho, \mu, \nu \in [0,1]$, we define the following two quantities:

$$\underline{\Gamma}_\rho(\mu,\nu) = \Pr[X \le \phi^{-1}(\mu), Y \ge \phi^{-1}(1-\nu)]$$

$$\overline{\Gamma}_\rho(\mu,\nu) = \Pr[X \le \phi^{-1}(\mu), Y \le \phi^{-1}(\nu)]$$

where X and Y are two standard Gaussian variables with covariance ρ.

We are now ready to state the invariance principle from [7] that we need for our reduction.

▶ **Theorem 9** ([7])**.** *Let T be a symmetric Markov operator on $\mathbb{F}_q$ such that $\rho = r(T) < 1$. Then for any $\tau > 0$ there exists $\delta > 0$ and $k \in \mathbb{N}$ such that if $f, g : \mathbb{F}_q^n \to [0,1]$ are two functions satisfying*

$$\min(\mathbf{Inf}_i^{\le k}(f), \mathbf{Inf}_i^{\le k}(g)) \le \delta$$

for all $i \in [n]$, then it holds that

$$\langle f, T^{\otimes n} g\rangle \ge \underline{\Gamma}_\rho(\mu,\nu) - \tau$$

where $\mu = \mathbb{E}[f]$, $\nu = \mathbb{E}[g]$.

Our hardness result is based on a variant of Unique Games conjecture. First, we define what the Unique game is:

▶ **Definition 10** (Unique Game)**.** An instance $G = (U, V, E, [L], \{\pi_e\}_{e\in E})$ of the Unique Game constraint satisfaction problem consists of a bi-regular bipartite graph (U, V, E), a set of alphabets $[L]$ and a permutation map $\pi_e : [L] \to [L]$ for every edge $e \in E$. Given a labeling $\ell : U \cup V \to [L]$, , an edge $e = (u, v)$ is said to be satisfied by ℓ if $\pi_e(\ell(v)) = \ell(u)$.

G is said to be *at most δ-satisfiable* if every labeling satisfies at most a δ fraction of the edges.

The following is a conjecture by Khot [12] which has been used to prove many *tight* inapproximability results.

▶ **Conjecture 11** (Unique Games Conjecture [12])**.** *For every sufficiently small $\delta > 0$ there exists $L \in \mathbb{N}$ such that the following holds. Given a an instance $\mathcal{G} = (U, V, E, [L], \{\pi_e\}_{e\in E})$ of* Unique Game *it is NP-hard to distinguish between the following two cases:*

- *YES case: There exist an assignment that satisfies at least $(1-\delta)$ fraction of the edges.*
- *NO case: Every assignment satisfies at most δ fraction of the edge constraints.*

Our hardness results are based on the following stronger conjecture which is similar to the one in Bansal-Khot [5]. We refer readers to [5] for more discussion on comparison between these two conjectures.

▶ **Conjecture 12** (Strong Unique Games Conjecture)**.** *For every sufficiently small $\delta, \gamma, \eta > 0$ there exists $L \in \mathbb{N}$ such that the following holds: Given an instance $\mathcal{G} = (U, V, E, [L], \{\pi_e\}_{e\in E})$ of* Unique Game *which is bi-regular, it is* NP*-hard to distinguish between the following two cases:*

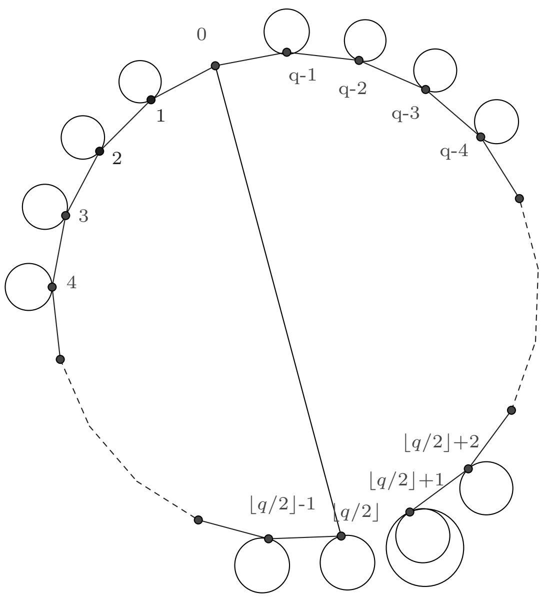

Figure 1 Gadget.

- *YES case: There exist sets $V' \subseteq V$ such that $|V'| \geq (1-\eta)|V|$ and an assignment that satisfies all edges connected to V'.*
- *NO case: Every assignment satisfies at most γ fraction of the edge constraints. Moreover, the instance satisfies the following expansion property. For every set $S \subseteq V$, $|S| = \delta|V|$, we have $|\Gamma(S)| \geq (1-\delta)|U|$, where $\Gamma(S) := \{u \in U \mid \exists v \in S \texttt{s.t.}(u,v) \in E\}$.*

▶ Remark. We would like to point out that the above conjecture differs from the one in [5] in the completeness case. In [5], the Yes instance has a guarantee that there exists sets $V' \subseteq V, U' \subseteq U$ with $|V'| \geq (1-\eta)|V|, |U'| \geq (1-\eta)|U|$ such that all edges between V' and U' are satisfied.

4.2 $(2-\epsilon)$-inapproximability

In order to prove the $(2-\epsilon)$ hardness, we first start with a *dictatorship test* that we will use as a gadget in the actual reduction.

4.2.1 Dictatorship Test

We design a dictatorship test for the problem Bi-Covering. We are interested in functions $f : \mathbb{F}_q^n \to \mathbb{R}$. f is called a dictator if it is of the form $f(x_1, x_2, \ldots, x_n) = x_i$ for some $i \in [n]$.

4.2.1.1 Dictatorship gadget

For convenience, we will let $q > 2$ be any prime number for the description of the dictatorship gadget. Let $G(\mathbb{F}_q, \mathcal{E})$ be a 3-regular graph on $\mathbb{F}_q$ (where we identify the elements of $\mathbb{F}_q$ by $\{0, 1, \ldots, q-1\}$) with self loops as shown in figure 1:

It is constructed as follows : Take a cycle on $0, 1, 2, \ldots, q-1, 0$, then add a self loop to every vertex except to the vertex 0. Remove the edge $(\lfloor q/2 \rfloor, \lfloor q/2 \rfloor + 1)$, add an edge

$(0, \lfloor q/2 \rfloor)$. Finally, to make it 3-regular, add a self loop to the vertex $\lfloor q/2 \rfloor + 1$. This completes the description of graph G. Since the graph G is connected and non-bipartite, the symmetric Markov operator T defined by the random walk in G has $r(T) < 1$. One crucial thing about G is that it has two large disjoint subsets of vertices, namely $\{1, 2, \ldots, \lfloor q/2 \rfloor\}$ and $\{\lfloor q/2 \rfloor + 1. \lfloor q/2 \rfloor + 2, \ldots, q-1\}$, with no edges in between.

Consider the vertex set $V = \mathbb{F}_q^R$ for some constant R. We will construct a graph H on V as follows : $(x, y) \in (\mathbb{F}_q^R)^2$ forms an edge in H iff they satisfy the following condition:

$$\forall i \in [R], (x_i, y_i) \in \mathcal{E},$$

x is adjacent to y iff $T^{\otimes R}(x \leftrightarrow y) \neq 0$.

4.2.1.2 Completeness

Let $f : \mathbb{F}_q^R \to \mathbb{R}$ be any dictator, say i^{th} dictator i.e. $f(x) = x_i$. By letting set A to be $f^{-1}(0) \cup f^{-1}(1) \cup \ldots \cup f^{-1}(\lfloor q/2 \rfloor)$ and set B to be $f^{-1}(0) \cup f^{-1}(\lfloor q/2 \rfloor + 1) \cup f^{-1}(\lfloor q/2 \rfloor + 2) \cup \ldots \cup f^{-1}(q-1)$, it can be seen easily that there is no edge between sets $A \setminus B$ and $B \setminus A$. More precisely,

$$A \setminus B = \{x \in \mathbb{F}_q^R \quad | \quad x_i \in \{1, 2, \ldots, \lfloor q/2 \rfloor\}$$
$$B \setminus A = \{y \in \mathbb{F}_q^R \quad | \quad y_i \in \{\lfloor q/2 \rfloor + 1, \lfloor q/2 \rfloor + 2, \ldots, q-1\}\}$$

By the property of Markov operator $T^{\otimes R}$, (x, y) are not adjacent if $(x_i, y_i) \notin \mathcal{E}$ for some $i \in [R]$. Hence, there are no edges between $A \setminus B$ and $B \setminus A$. Thus, the optimal value is at most

$$\frac{1}{|V|} \cdot \max\{|A|, |B|\} = \frac{1}{2} + \frac{1}{2q}.$$

4.2.1.3 Soundness

Let $A, B \subseteq V$ such that $A \cup B = V$ and $f, g : \mathbb{F}_q^R \to \{0, 1\}$ be the indicator functions of sets $A \setminus B$ and $B \setminus A$ respectively. Suppose $|A \setminus B| = \epsilon |V|$ and $|B \setminus A| = \epsilon |V|$ for some $\epsilon > 0$ and that there are no edges in between $A \setminus B$ and $B \setminus A$. We will show that in this case, f and g must have a common influential co-ordinate. Since, there are no edges between these sets, we have

$$\mathop{\mathbb{E}}_{\substack{x \sim \mathbb{F}_q^R, \\ y \sim T^{\otimes R}(x)}} [f(x) g(y)] = \langle f, T^{\otimes R} g \rangle = 0 \,.$$

For the application of Invariance principle, Theorem 9, in our case we have $\mathbb{E}[f] = \mathbb{E}[g] = \epsilon > 0$ and $\rho = r(T) < 1$. Thus, for small enough $\tau := \tau(\rho, \epsilon) > 0$,

$$\underline{\Gamma}_\rho(\epsilon, \epsilon) - \tau > 0.$$

We can now apply Theorem 9 to conclude that there exists $i \in [R]$ and $k \in \mathbb{N}$ independent of R such that

$$\min(\mathbf{Inf}_i^{\leq k}(f), \mathbf{Inf}_i^{\leq k}(g)) \geq \delta,$$

for some $\delta(\tau) > 0$. Hence, unless f and g have a common influential co-ordinate, $\frac{1}{|V|} \cdot \max\{|A|, |B|\} \geq 1 - \epsilon$. Thus, the optimum value is at least $1 - \epsilon$

4.2.2 Actual Reduction

The above dictatorship test for large enough q can be composed with the Unique Games instance having some stronger guarantee (Conjecture 12) in a straightforward way that gives $(2-\epsilon)$ hardness for every constant $\epsilon > 0$ assuming UGC. Details as follows:

Let $\mathcal{G} = (U, V, E, [L], \{\pi_e\}_{e\in E})$ be the given instance of Unique Game with parameters $\delta < \frac{\epsilon}{4}, \gamma, \eta > 0$ from Conjecture 12 . We replace each vertex $v \in V$ by a block of q^L vertices, namely by a hypercube $[q]^L$. We will denote this block by $[v]$. As defined in the dictatorship test, let G be the graph on $\mathbb{F}_q$ and T be the induced symmetric Markov operator. For every pair of edges $e_1(u, v_1)$ and $e_2(u, v_2)$ in $\mathcal{G}$, we will add the following edges between $[v_1]$ and $[v_2]$: Let π_1 and π_2 be the permutation constraint associated with e_1 and e_2 respectively. $x \in [v_1]$ and $y \in [v_2]$ are connected by an edge iff $T^{\otimes L}((x \circ \pi_1^{-1}) \leftrightarrow (y \circ \pi_2^{-1})) \neq 0$ (where $(x \circ \pi^{-1})_i = x_{\pi^{-1}(i)}$ for all $i \in [L]$) i.e. for every $i \in [L]$, $x_{\pi_1^{-1}(i)}$ and $y_{\pi_2^{-1}(i)}$ are connected by an edge in graph G. This completes the description of a graph. Let's denote this graph by H.

▶ **Lemma 13** (Completeness). *If there exists an assignment to vertices in $\mathcal{G}$ that satisfies all edges connected to $(1-\eta)$ fraction of vertices in V then H has a* Bi-Covering *of size at most $(1-\eta)(1/2 + 1/2q) + \eta$.*

Proof. Fix a labeling ℓ such that for at least $(1-\eta)$ fraction of vertices in V in $\mathcal{G}$, all edges attached to them are satisfied. Suppose X be the set of remaining η fraction of vertices of V in $\mathcal{G}$. For every vertex $v \in V$, consider the following two partitions of $[v]$:

$$A_v = \{x \in [q]^L : x_{\ell(v)} \in \{1, \dots, \lfloor q/2 \rfloor\}\}$$
$$B_v = \{x \in [q]^L : x_{\ell(v)} \in \{\lfloor q/2 \rfloor\} + 1, \lfloor q/2 \rfloor\} + 2, \dots, q\}\}$$
$$C_v = \{x \in [q]^L : x_{\ell(v)} = 0\}$$

Let $A = \cup_{v\in V}(A_v \cup C_v) \cup_{z\in X} [z]$ and $B = \cup_{v\in V}(B_v \cup C_v) \cup_{z\in X} [z]$. The claim is that this is the required edge separating sets. To see this, consider any vertex pair (a, b) such that $a \in A \setminus B$ and $b \in B \setminus A$. We need to show that (a, b) must not be adjacent in H. Suppose $a \in [v_1]$ and $b \in [v_2]$. If v_1 and v_2 don't have a common neighbor then clearly, there is no edge between a and b. Suppose they have a common neighbor u and let $e_1 = (u, v_1)$ and $e_2 = (u, v_2)$ be the edges and π_1 and π_2 be the associated permutation constraints. Since $X \subseteq A \cap B$, $v_1, v_2 \notin X$. Hence ℓ satisfies all constraints associated with v_1 and v_2. In particular, $\pi_1(\ell(v_1)) = \pi_2(\ell(v_2)) =: j$ for some $j \in [L]$. Since $a \in A_{v_1}$, we have $a_{\pi_1^{-1}(j)} = a_{\ell(v_1)} \in \{1, \dots, \lfloor q/2 \rfloor\}\}$. Similarly, $b_{\pi_2^{-1}(j)} \in \{\lfloor q/2 \rfloor\} + 1, \lfloor q/2 \rfloor\} + 2, \dots, q\}$. By the construction of edges in H, a and b are not adjacent.

For any v, $|A_v \cup C_v| = |B_v \cup C_v| = (\frac{1}{2} + \frac{1}{2q})q^L$. Thus,

$$|A| = |B| \le \left(\eta + (1-\eta)\left(\frac{1}{2} + \frac{1}{2q}\right)\right)|V|q^L . \quad ◀$$

▶ **Lemma 14** (Soundness). *For every constant $\epsilon > 0$, there exists a constant γ such that, if $\mathcal{G}$ is at most γ-satisfiable then H has* Bi-Covering *of size at least $1-\epsilon$.*

Proof. Suppose for contradiction, there exists an Bi-Covering of size at most $(1-\epsilon)$. This means there exists two disjoint sets X, Y of size at least ϵ fraction of vertices in H such that there are no edges in between X and Y. Let X^* be the set of vertices in $v \in V$ such that $[v] \cap X \ge \frac{\epsilon}{2}|[v]|$. Similarly, Y^* be the set of vertices in $v \in V$ such that $[v] \cap Y \ge \frac{\epsilon}{2}|[v]|$. By simple averaging argument, $|X^*| \ge \frac{\epsilon}{2}|V|$ and $|Y^*| \ge \frac{\epsilon}{2}|V|$.

▶ **Lemma 15.** *The total fraction of edges connected to X^* whose other end point is in $\Gamma(X^*) \cap \Gamma(Y^*)$ is at least $\frac{1}{2}$.*

Proof. Let $\mathcal{G}$ has left-degree d_1 and right-degree d_2. We have $d_1 = \frac{d_2|V|}{|U|}$. Suppose the claim is not true, then at least $\frac{1}{2}$ fraction of edges have there endpoint in $U \setminus \Gamma(Y^*)$. As, $|U \setminus \Gamma(Y^*)| \leq \delta|U|$, the average degree of a vertex in $U \setminus \Gamma(Y^*)$ is at least $\frac{(1/2)d_2|X^*|}{\delta|U|} \geq \frac{(d_2/2)\cdot(\epsilon/2)|V|}{\delta|U|}$ which is greater than d_1 as $\epsilon > 4\delta$. ◀

For $v \in X^* \cup Y^*$, let $f_v : [q]^L \to \{0,1\}$ be the indicator function of a set $[v] \cap (X \cup Y)$. Define the following label set for $v \in X^* \cup Y^*$ for some $\tau' > 0$ and $k \in \mathbb{N}$:

$$\mathcal{F}(v) := \{i \in [L] \mid \mathbf{Inf}_i^{\leq k}(f_v) \geq \tau'\}.$$

We have $|\mathcal{F}(v)| \leq \frac{\tau'}{k}$ as $\sum_i \mathbf{Inf}_i^{\leq k}(f_v) \leq k$.

▶ **Lemma 16.** *There exists a constant $\tau' := \tau'(q,\epsilon)$ and $k := k(q,\epsilon)$ such that for every $u \in U$ and edges $e_1(u,v), e_2(u,w)$ such that $v \in X^*$ and $w \in Y^*$, we have*

$$\pi_{e_1}(\mathcal{F}(v)) \cap \pi_{e_2}(\mathcal{F}(w)) \neq \emptyset.$$

Proof. As there are no edges between X and Y, we have

$$\mathop{\mathbb{E}}_{\substack{(x\circ\pi_{e_1}^{-1})\sim\mathbb{F}_q^L,\\ (y\circ\pi_{e_2}^{-1})\sim T^{\otimes L}(x\circ\pi_{e_1}^{-1})}} [f_v(x\circ\pi_{e_1}^{-1}) f_w(y\circ\pi_{e_2}^{-1})] = 0.$$

By the soundness analysis of the dictatorship test, it follows that there exists $i \in [L]$ such that

$$\min(\mathbf{Inf}^{\leq k}_{\pi_{e_1}^{-1}(i)}(f_v), \mathbf{Inf}^{\leq k}_{\pi_{e_2}^{-1}(i)}(f_w)) \geq \tau',$$

for some τ', k as a function of q and ϵ. Thus, $i \in \pi_{e_1}(\mathcal{F}(v))$ and $i \in \pi_{e_2}(\mathcal{F}(w))$. ◀

◀

4.2.2.1 Labeling

Fix τ' and k from Lemma 16. We now define a labeling ℓ to vertices in $X^* \subseteq V$ and in $\Gamma(X^*) \cap \Gamma(Y^*) \subseteq U$ as follows: For a vertex $v \in X^*$ set $\ell(v)$ to be an uniformly random label from $\mathcal{F}(v)$. For $u \in \Gamma(X^*) \cap \Gamma(Y^*)$, select an arbitrary neighbor w of u in Y^* and set $\ell(u)$ to be an uniformly random label from the set $\pi_{(u,w)}(\mathcal{F}(w))$ of size at most $\frac{k}{\tau'}$. Fix an edge (u,v) such that $u \in \Gamma(X^*) \cap \Gamma(Y^*)$ and $v \in X^*$. By Lemma 16, for any $w \in Y^*$ since $\pi_{(u,w)}(\mathcal{F}(w)) \cap \pi_{(u,v)}(\mathcal{F}(v)) \neq \emptyset$, The probability that the edge is satisfied by the randomized labeling is at least $\left(\frac{\tau'}{k}\right)^2$. Thus in expectation, at least $\left(\frac{\tau'}{k}\right)^2$ fraction of edges between X^* and $\Gamma(X^*) \cap \Gamma(Y^*)$ are satisfied. By Lemma 15, at least $\frac{1}{2}$ fraction of edges connected to X^* are in between X^* and $\Gamma(X^*) \cap \Gamma(Y^*)$. Finally using bi-regularity, this labeling satisfies at least $\frac{1}{2}\frac{\epsilon}{2}\left(\frac{\tau'}{k}\right)^2$ fraction of edges in $\mathcal{G}$. Setting $\gamma < \frac{1}{2}\frac{\epsilon}{2}\left(\frac{\tau'}{k}\right)^2$ completes the proof. ◀

Proof of Theorem 1

The proof follows from Lemma 13, Lemma 14 and Conjecture 12.

Proof of Theorem 2

Given an input as a bipartite graph, there is a trivial 3/2 approximation for Bi-Covering - Take set A to be the union of a smaller part and half of the larger bi partition and B to be union of smaller part and remaining half of the larger part. It is easy to see these two sets A and B satisfy the property of being a Bi-Covering. As $\max\{|A|, |B|\} \le \frac{3}{4}|V|$, this is a $\frac{3}{2}$ approximation as OPT is at least $\frac{|V|}{2}$.

The $\frac{3}{2} + \epsilon$ inapproximability follows easily from the above $(2 - \epsilon)$ inapproximability for the general case. The reduction is as follows: Let $G(V, E)$ be the given instance of a Bi-Covering. Construct a natural bipartite graph G' between $V \times V$ where (i, j) forms an edge if $(i, j) \in E$ (or $(j, i) \in E$). Fix a small enough constant $\epsilon > 0$. It is easy to see that if G has a solution of fractional size $1/2 + \epsilon$ then so does G'. Next, if there are sets A' and B' where $\frac{1}{2|V|}\max\{|A'|, |B'|\} \le \frac{3}{4} - \epsilon$ which satisfy the Bi-Covering property, we have $\frac{1}{2|V|}|A' \setminus B'| = \frac{1}{2|V|}(2|V| - |B'|) \ge 1 - (\frac{3}{4} - \epsilon) = \frac{1}{4} + \epsilon$ and similarly $\frac{1}{2|V|}|B' \setminus A'| \ge \frac{1}{4} + \epsilon$. Note that $A' \setminus B'$ and $B' \setminus A'$ are two disjoint sets whose size of union is at least $(1 + 2\epsilon)|V|$. Thus, we can find two sets, say X' and Y' (namely X' is intersection of $A' \setminus B'$ with left part of the bipartite graph and Y' is the intersection of $B' \setminus A'$ with right part) of size at least $\epsilon|V|$ each, where X' is from left side and Y' is from right side with no edges in between. We now think of X' and Y' as a subset of V. Let $Z = X' \cap Y'$. Partition Z into Z_1 and Z_2 of equal sizes. Take $X = Z_1 \cup (X' \setminus Y')$ and $Y = Z_2 \cup (Y' \setminus X')$. It is now easy to verify that there are no edges in between X and Y in G and $\frac{1}{|V|}\min\{|X|, |Y|\} \ge \frac{\epsilon}{2}$. Hence, if we can find a solution of fractional cost $\frac{3}{4} - \epsilon$ in G' in polynomial time then we can also find a solution of fractional cost $1 - \frac{\epsilon}{2}$ in G in polynomial time and this gives a polynomial time algorithm with approximation factor $2 - \frac{\epsilon}{2}$ for small enough constant $\epsilon > 0$. As Bi-Covering is UG hard to approximate within $(2 - \epsilon)$ for all $\epsilon > 0$ for general graph, this gives a $\frac{3}{2} + \epsilon$ hardness for Bi-Covering in bipartite graph.

Proof of Corollary 3

We prove it by giving reduction from Bi-Covering. Let $G(V, E)$ be the given instance of Bi-Covering. Construct a bipartite graph H between $V \times V$ where (i, j) forms an edge if $(i, j) \notin E$. Fix a small enough constant $\epsilon > 0$. In one direction, if G has a Bi-Covering of fractional size at most $(1/2 + \epsilon)$ then H' contains a $(1/2 - \epsilon)|V| \times (1/2 - \epsilon)|V|$ bipartite clique. In other direction, if H' has a bipartite clique of size $2\epsilon|V| \times 2\epsilon|V|$ then let X' and Y' be the subset of vertices from left and right side of bipartite clique. As before, let $Z = X' \cap Y'$ and Z_1 and Z_2 be the partition of Z of equal size. Let $X = (X' \setminus Y') \cup Z_1$ and $Y = (Y' \setminus X') \cup Z_2$. It follows that $|X|, |Y|$ is at least $\epsilon|V|$ and are disjoint viewed as a subset of V. Also, there are no edges between X and Y. Therefore, $V \setminus X$ and $V \setminus Y$ each of size at most $(1 - \epsilon)|V|$ gives a Bi-Covering of G. Thus, Theorem 1 implies that it is hard to distinguish between Bi$-$Clique of size $(1/2 - \epsilon)|V|$ and $\epsilon|V|$ which completes the proof of corollary.

Acknowledgements. We would like to thank anonymous reviewers for their comments which significantly helped in improving the presentation of this paper.

References

1 Noga Alon, Uriel Feige, Avi Wigderson, and David Zuckerman. Derandomized graph products. *Computational Complexity*, 5(1):60–75, 1995.

2 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.

3 Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. In *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*, pages 563–572. IEEE, 2010.

4 Sanjeev Arora, Subhash A Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K Vishnoi. Unique games on expanding constraint graphs are easy. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 21–28. ACM, 2008.

5 Nikhil Bansal and Subhash Khot. Optimal long code test with one free bit. In *Foundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on*, pages 453–462. IEEE, 2009.

6 Thang Nguyen Bui and Lisa C. Strite. An ant system algorithm for graph bisection. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO'02, pages 43–51, 2002.

7 Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009.

8 Uriel Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 534–543. ACM, 2002.

9 Uriel Feige and Shimon Kogan. *Hardness of approximation of the balanced complete bipartite subgraph problem*, 2004.

10 Rajiv Gandhi, Samir Khuller, Aravind Srinivasan, and Nan Wang. Approximation algorithms for channel allocation problems in broadcast networks. *Networks*, 47(4):225–236, 2006.

11 Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar. Beating the random ordering is hard: Every ordering csp is approximation resistant. *SIAM Journal on Computing*, 40(3):878–914, 2011.

12 Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775. ACM, 2002.

13 Subhash Khot. Ruling out ptas for graph min-bisection, dense k-subgraph, and bipartite clique. *SIAM Journal on Computing*, 36(4):1025–1071, 2006.

14 Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within 2-ε. *Journal of Computer and System Sciences*, 74(3):335–349, 2008.

15 Subhash Khot, Madhur Tulsiani, and Pratik Worah. A characterization of strong approximation resistance. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 634–643. ACM, 2014.

16 Prasad Raghavendra. Optimal algorithms and inapproximability results for every csp? In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 245–254. ACM, 2008.

17 Ola Svensson. Conditional hardness of precedence constrained scheduling on identical machines. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 745–754. ACM, 2010.