Bicovering: Covering edges with two small subsets of vertices

Amey Bhangale* Rajiv Gandhi† Mohammad T. Hajiaghayi‡ Rohit Khandekar§
Guy Kortsarz¶

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 [14]. A solution that outputs $V, \emptyset$ gives ratio at most 2. We show that under a similar *Strong* Unique Game Conjecture by Bansal and Khot [6] 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 [11] and also under the assumption that NP $\nsubseteq \cap_{\epsilon>0} \mathsf{DTIME}(2^{n^{\epsilon}})$ [18]. It was an open problem in [3] 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.

*Department of Computer Science, Rutgers University, New Brunswick. Research supported in part by NSF grant CCF-1253886 and NSF award 1218620 amey.bhangale@rutgers.edu

†Department of Computer Science, Rutgers University, Camden. Supported in part by NSF grant number 1218620 and NSF award 1433220.rajivg@camden.rutgers.edu

‡Department of Computer Science, Maryland. 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. hajiagha@cs.umd.edu

§KCG holdings Inc., USA rkhandekar@gmail.com

¶Department of Computer Science, Rutgers University, Camden. Supported in part by NSF grant number 1218620 and by NSF grant 1540547. guyk@camden.rutgers.edu

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 [14]. 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 [6] (see Conjecture 2).

Theorem 1. *Let $\epsilon > 0$ be any small constant. Assuming a strong Unique Games Conjecture (Conjecture 2), 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 [2, 7, 11, 18]. Feige [11] 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 [2, 7]). Feige-Kogan [13] 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 [18] later proved a similar inapproximability result but assuming NP $\nsubseteq \cap_{\epsilon>0} \mathsf{BPTIME}(2^{n^\epsilon})$ using a *quasi-random* PCP. and the assumption can be weakened to NP $\nsubseteq \cap_{\epsilon>0} \mathsf{DTIME}(2^{n^\epsilon})$ using a deterministic reduction from [8]. It is an important open problem to extend similar hardness results based on weaker complexity assumptions [3]. 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 2.

Corollary 1. *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 [2, 7])

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 [19], optimal algorithm for every Max-CSP [21], Ordering CSPs [16], characterizing strong approximation resistance of CSPs [20] 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 [5]. In general, in [4] 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 [5], [4]. A similar Strong Unique Games Conjecture, which has a *weak expansion property*, has been used earlier in [6] and [22] 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 $|S| \leq n/2$, $N_1(S) \geq \delta|S|$, 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 3. *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 rely 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 bounds 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.

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 the 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 1 (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 2. *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 3 (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:*

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

where $x_1, \dots, x_n$ 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-degree (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 4. *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:*

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

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

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

We are now ready to state the invariance principle from [10] that we need for our reduction.

Theorem 4 ([10])**.** *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*

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

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

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

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 5 (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 [17] which has been used to prove many *tight* inapproximability results.

Conjecture 1 (Unique Games Conjecture [17]). *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 [6]. We refer readers to [6] for more discussion on comparison between these two conjectures.

Conjecture 2 (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:*

- *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 $|N(S)| \geq (1-\delta)|U|$, where $N(S) := \{u \in U \mid \exists v \in S \texttt{ s.t. } (u, v) \in E\}$.*

Remark 1. *We would like to point out that the above conjecture differs from the one in [6] in the completeness case. In [6], 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. To the best of our knowledge, it is not known if the above conjecture implies the one in [6] or the other way around. Also, it is not know if* Unique Games Conjecture *implies any of the two conjectures.*

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 $i \in [n]$.

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

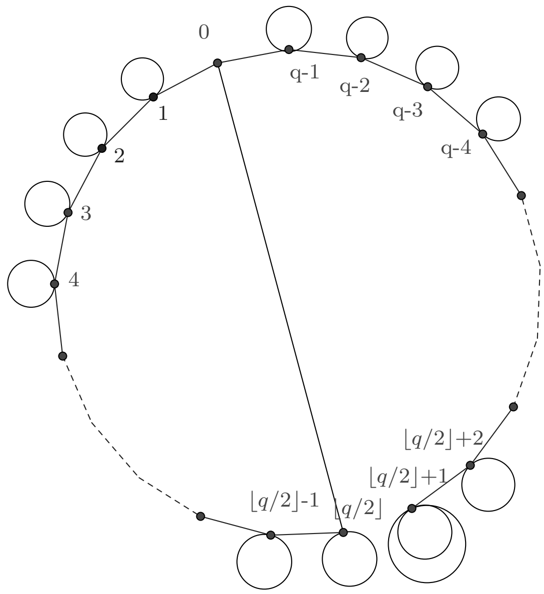

Figure 1: Gadget

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:

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

In other words, x is adjacent to y iff $T^{\otimes R}(x \leftrightarrow y) \neq 0$, where $T^{\otimes R}$ is a coordinate-wise random walk on $\mathbb{F}_q^R$ with respect to T.

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,

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

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

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

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 4, 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 4 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 2) 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 2 . 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 1 (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 partitions of $[v]$:

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

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,

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

□

Lemma 2 (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 $[v] \cap X \geq \frac{\epsilon}{2}|[v]|$. Similarly, Y^* be the set of vertices in $v \in V$ such that $[v] \cap Y \geq \frac{\epsilon}{2}|[v]|$. By simple averaging argument, $|X^*| \geq \frac{\epsilon}{2}|V|$ and $|Y^*| \geq \frac{\epsilon}{2}|V|$.

Lemma 3. *The total fraction of edges connected to X^* whose other end point is in $N(X^*) \cap N(Y^*)$ 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 $U \setminus N(Y^*)$. As, $|U \setminus N(Y^*)| \leq \delta|U|$, the average degree of a vertex in $U \setminus N(Y^*)$ 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 4. *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 $i \in \pi_{e_2}(\mathcal{F}(w))$. $\square$

$\square$

Labeling: Fix τ' and k from Lemma 4. We now define a labeling ℓ to vertices in $X^* \subseteq V$ and in $N(X^*) \cap N(Y^*) \subseteq U$ as follows: For a vertex $v \in X^*$ set $\ell(v)$ to be an uniformly random label from $\mathcal{F}(v)$. For $u \in N(X^*) \cap N(Y^*)$, 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 $u \in N(X^*) \cap N(Y^*)$ and $v \in X^*$. By Lemma 4, 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 $N(X^*) \cap N(Y^*)$ are satisfied. By Lemma 3, at least $\frac{1}{2}$ fraction of edges connected to X^* are in between X^* and $N(X^*) \cap N(Y^*)$. 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. $\square$

Proof of Theorem 1: The proof follows from Lemma 1, Lemma 2 and Conjecture 2.

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 $\max\{|A|, |B|\} \leq \frac{3}{4}|V|$, 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 $\frac{1}{2|V|}\max\{|A'|, |B'|\} \leq \frac{3}{4}-\epsilon$ which satisfy the Bi-Covering property, we have $\frac{1}{2|V|}|A' \setminus B'| = \frac{1}{2|V|}(2|V| - |B'|) \geq 1 - (\frac{3}{4} - \epsilon) = \frac{1}{4} + \epsilon$ and similarly $\frac{1}{2|V|}|B' \setminus A'| \geq \frac{1}{4} + \epsilon$. 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 $\frac{1}{|V|}\min\{|X|, |Y|\} \geq \frac{\epsilon}{2}$. 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 1: 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.

5 Better Approximation Algorithms for special graphs

We denote the input graph by $G(V, E)$ throughout this section. We start with basic notations which we use in our algorithms.

5.1 Notations for algorithms

For a given graph $G(V, E)$, let $N(v)$ denote the vertices joined to v (its neighbors). The number of neighbors of v is denoted by $deg(v)$. For a set S let $N(S)$ be the union all vertices that are joined to at least one vertex in S. Let $N_1(S)$ be the set of vertices *not in* S that have at least one neighbor in S. For a collection of numbers X, let $s(X)$ be the sum of the numbers in X. Given a family of sets $\mathcal{S} = \{S_1, S_2, \ldots, S_t\}$, define $un(\mathcal{S}) := \bigcup_{S_i \in \mathcal{S}} S$. The optimum covering for the BI-COVERING instance at hand is denoted by A^*, B^* and we denote $S^* = A^* \cap B^*$

Our problem is related to the Vertex Separator problem which we define as follows:

Definition 6. *A* Vertex Separator *in a graph G is a set S so that after S is removed no connected component has more than $2n/3$ vertices*

We start with some easy lemmas.

Lemma 5. *If $|A^*| \le n/2 - \epsilon \cdot n$ or $|B^*| \le n/2 - \epsilon \cdot n$ then returning $V, \emptyset$ gives a $2/(1 + 2\epsilon)$ ratio.*

Proof. Say that $|B^*| \le n/2 - \epsilon \cdot n$. Then as $|A^*| + |B^*| \ge n$ we get that $|A^*| \ge n/2 + \epsilon \cdot n$. Returning $V, \emptyset$ gives ratio $2/(1 + 2\epsilon)$. □

Lemma 6. *If $|S^*| \ge 2\epsilon \cdot n$ returning $V, \emptyset$ gives an approximation ratio of $2/(1 + 2\epsilon)$*

Proof. Note that $|A^*| + |B^*| = |(V \setminus S^*)| + 2|S^*| \ge (1 + 2\epsilon)n$. Thus either $|A^*| \ge n/2 + \epsilon \cdot n/2$, or $|B^*| \ge 1/2 + \epsilon \cdot n/2$. We return the trivial solution $V, \emptyset$. Those the solution $V, \emptyset$ the ratio derived is at most

$$\frac{1}{1/2 + \epsilon} = \frac{2}{1 + 2\epsilon}.$$

□

The next two lemmas are also used several times.

Lemma 7. *Let C be a Clique. Then either $C \subseteq A^*$ or $C \subseteq B^*$.*

Proof. Note that $V \setminus A^* = B^* \setminus S^*$. If $C \nsubseteq A^*$ it means that there is a vertex $u \in C \cap (B^* \setminus S^*)$. If $C \nsubseteq B^*$ it means that there is a vertex $v \in C \cap (A^* \setminus S^*)$. However, as u, v belong to the clique u and v are neighbors. This gives an edge between $A^* \setminus S^*$ and $B^* \setminus S^*$ contradicts Lemma 8. $\square$

Lemma 8. *If $A, B \subseteq V$ is a feasible solution to* Bi-Covering *and $S = A \cap B$ then there are no edges from $A \setminus S$ to $B \setminus S$.*

Proof. As $A \setminus S$ and $B \setminus S$ are *disjoint*, an edge (u, v) between a vertex v in $A \setminus S$ and a vertex $u \in B \setminus S$ can satisfy $u, v \in A$ or $u, v \in B$ contradicting the feasibility of the solution. $\square$

Lemma 9. *Say that we can find in polynomial time two* disjoint *sets A', B' with at least $\epsilon \cdot n$ vertices each, so that A', B' share no edges, then* Bi-Covering *admits a $2 - 2\epsilon$ ratio algorithms.*

Proof. Set $S \leftarrow V \setminus (A' \cup B')$, $A = S \cup A'$ and $B = S \cup B'$ getting a feasible solution. Indeed, edges inside A', B', S are clearly covered. Since $A = A' \cup S$, $B = B' \cup S$ edges between A' and S are covered by A and edges between $B' \cup S$ are covered by B. Finally, we are given the property that there are no edges between A' and B'.

Note that A', B' are disjoint and so $A' \cap S = B' \cap S = \emptyset$. Thus $A' \cap B = \emptyset$ and $B' \cap A = \emptyset$. Given the size of A', B', we get that $|A|, |B| \leq (1-\epsilon) \cdot n$. Thus the value of ours solution is $(1-\epsilon) \cdot n$ versus $n/2$ for the optimum. The ratio is $2 - 2\epsilon$. $\square$

The following Procedure is the one which attains the promised ratio:

Procedure Big (A',B'):

1. Let $S \leftarrow V \setminus A' \cup B'$
2. Return $A = A' \cup S, B = B' \cup S$

5.2 Basic definitions and tools

Say that we remove a set C of vertices and get connected components $H_1, H_2, \ldots, H_p$. A 2-Covering of the components $H_1, \ldots, H_p$ are two collections of components X and Y (namely, X either contains a whole component or none of the vertices of the component and the same holds for Y) so that in addition $X \cap Y = \emptyset$ and $X \cup Y = \{H_1, \ldots, H_p\}$. A minimum covering by two sets minimizes $max\{|un(X)|, |un(Y)|\}$. We call this minimization problem a 2-Cover problem.

Lemma 10. *An optimal solution for the 2-Cover problem can be found in polynomial time.*

Proof. The *Subset Sum* problem is, given a set of n input numbers $T = \{x_1, x_2, \ldots, x_p\}$ and a number Q, decide if there is a subset $T' \subseteq T$ that sums to Q. In our case the $x_i = |H_i|$. Thus $\max_i x_i \leq n$ and therefore there exists an exact polynomial time solution for subset sum in this case (the weights are bounded by polynomial in n). For every number Q between 1 and $|\bigcup H_i|$, check if there is a collection of connected components X so that $|un(X)| = Q$. For any feasible X check the value of the solution $un(X), V \setminus un(X)$. Output the best solution over all Q. Clearly this is the optimum solution. $\square$

5.3 A central algorithmic tool

Given two disjoint sets C_1, C_2 of vertices, the bipartite graph $\mathcal{B}(C_1, C_2, E')$ *that corresponds to* C_1, C_2 is defined as follows: Make C_1 one side of the bipartite graph and remove all edges internal to C_1. Make C_2 the other side of the bipartite graph and remove all edges internal to C_2. The edges E' are all edges with one endpoint in C_1 and the other in C_2.

We present a simple procedure that takes two disjoint subsets X, Y of V and return a solution to the bicovering instance on G based on the bipartite graph $\mathcal{B}(X, Y, E')$ and on removing a vertex cover of the bipartite graph.

1. Compute $\mathcal{B}(X, Y, E')$
2. Compute the minimum size vertex cover D of $\mathcal{B}$.
   /* As the graph is bipartite the minimum vertex cover can be found in polynomial time */
3. Return $A = X \cup D \cup (V \setminus X \cup Y),\ B = Y \cup D \cup (V \setminus X \cup Y)$.

Figure 2: Procedure Vertex-Cover(X, Y)

Lemma 11. *If $|X \setminus D| \geq \epsilon \cdot n$ and $|Y \setminus D| \geq \epsilon n$, then the procedure Vertex-Cover returns a solution of ratio $2 - 2\epsilon$.*

Proof. Note that there are no edges between $X \setminus D$ and $Y \setminus D$ because a vertex cover was removed. As $A \setminus B = X \setminus D$ and $B \setminus A = Y \setminus D$, by Lemma 9 the ratio of the returned solution is $2 - 2\epsilon$. □

6 Algorithmic results

The following theorems are derived directly from the techniques we develop.

Theorem 5. *1. The* Bi-Covering *problem admits a $2 - 1/(\bar{d} + 1)$ ratio for a graph G with $O(n)$ edges with $\bar{d}$ the (constant) average degree of G.*

2. The Bi-Covering *problem admits a $4/3 + o(1)$ ratio for any graph that has a separators of size $o(n/\sqrt{\log n})$.*

We prove these theorems along with Theorem 3 in the following sections.

6.1 Chordal graphs

In this section, we provide a polynomial time algorithm with approximation ratio at most 1.876 for the Bi−Covering problem on Chordal graph. Set $\epsilon = 1/16$.

We use the following theorem is due to [15].

Theorem 6. *Every n-vertex chordal graph $G(V, E)$ contains a polynomially computable maximal clique C, so that if the vertices of C are removed, any connected components in the graph induced by the non deleted vertices has at most $n/2$ vertices.*

We are now ready to give our algorithm for chordal graphs.

1. The clique separator of G from Theorem 6 is denoted by C. Denote $H = V \setminus C$. Without loss of generality, let H_1 be largest connected component.
/*Note that $V \setminus C$ decomposes to a collection of connected component $H_1, \ldots, H_q$ so that $H = \bigcup_{i=1}^{q} H_i$.*/
2. If $|H_1| \leq \epsilon n$, apply Algorithm 2-Cover (from Lemma 10) on $H_1, H_2, \ldots, H_q$. Let A, B be the partition. Return $A \cup C, B \cup C$.
3. If $|\bigcup_{i=2}^{n} H_i| \geq \epsilon \cdot n$ Apply algorithm $Big(H_1, \bigcup_{i=2}^{q} H_i)$
/* Note that H_1 and $\bigcup_{i=2}^{q} H_i$ share no edges */
4. Else, Apply Algorithm Vertex-Cover on H and C and return the solution.

Figure 3: Algorithm Chordal

Analysis:

Lemma 12. *If $|C| \geq n/2 + \epsilon \cdot n$ returning $V, \emptyset$ derives a $2/(1+2\epsilon)$ ratio. For the choice of $\epsilon = 1/16$, this ratio is strictly smaller than 1.876.*

Proof. As $C \subseteq A^*$, this means that $|A^*| \geq n/2 + \epsilon \cdot n$ and thus the ratio is less than 1.876. □

From previous lemma, we may assume that unless there is a constant ratio smaller than 1.876 the following holds:

(a) W.l.o.g. $C \subseteq A^*$.

(b) $|A^*| \geq n/2 - \epsilon n$ and $|B^*| \geq n/2 - \epsilon n$, and

(c) $|S^*| \leq 2\epsilon n$.

(d) $|H| \geq n/2 - \epsilon n$.

Note what the resulting ratios are if one of the above statements does not hold. If $|A^*| \leq n/2 - \epsilon n$ or $B^* \leq n/2 - \epsilon \cdot n$ by Lemma 5 returning $V, \emptyset$ gives a ratio of $2/(1 + 2\epsilon) < 1.876$ for $\epsilon = 1/16$. If $|S| \geq 2 \cdot \epsilon n$ the a ratio is the same $2/(1 + 2\epsilon) < 1.876$. (d) follows as $H = V \setminus C$ and by Lemma 12 as $H \cup C = V$.

Lemma 13. *If $|H_1| \leq \epsilon \cdot n$, Algorithm 2-Cover (that is applied in this case by Algorithm Chordal) returns a solution of ratio at most at most $2 - 2\epsilon < 1.876$ for $\epsilon = 1/16$.*

Proof. We define a possibly sub-optimal solution for the 2-cover problem. The 2-Cover solution is an optimum one and may be only better. Initiate $A \leftarrow \emptyset$. As longs as $|A| \leq \epsilon \cdot n$ add an arbitrary H_i to A. Note that for every i, $|H_i| \leq \epsilon n$ because H_1 is the largest connected component. This implies that when we stop, $|A| \leq 2\epsilon n$. By Assumption (d), $H \geq n/2 - \epsilon n$. Therefore $H - |A| \geq n/2 - \epsilon n - 2\epsilon n = n/2 - 3\epsilon n > \epsilon n$ for $\epsilon = 1/16$. As A and $H \setminus A$ share no edges, because they are a collection of connected components, and both are of size at least ϵn, by Lemma 9, a ratio of $2 - 2\epsilon n$ applies. For $\epsilon = 1/16$ this ratio is less than 1.876. □

Lemma 14. *If* $|\bigcup_{i=2}^{q} H_i| > \epsilon n$, *then* Bi−Covering *admits a* $2 - 2\epsilon < 1.876$ *ratio.*

Proof. In this case Algorithm *Chordal* applies Algorithm *Big*. By Lemma 13, we may assume that $|H_1| \geq \epsilon n$. Therefore we have two sets H_1 and $|\bigcup_{i=2}^{q} H_i|$, both of size at lest ϵn and share no edges, (because they are connected component) a ratio of $2 - 2\epsilon < 1.876$ follows from Lemma 9. □

Claim 1. $|C| \geq n/2 - \epsilon n$ *and* $|H| \leq n/2 + \epsilon n$.

Proof. Note that $|H_1| \leq n/2$ by Theorem 6 as H_1 is a connected component resulting after C is removed. Note that $|V \setminus C| = H = H_1 \cup \bigcup_{i=2}^{q} H_i \leq n/2 + \epsilon n$. The last inequality follows because $|H_1| \leq n/2$ and by Lemma 14. This gives ratio less than 1.876. Thus, $|C| \geq n - |H| \geq n/2 - \epsilon n$ and $|H| = V - |C| \leq n/2 + \epsilon n$. □

Claim 2. *We now have following conditions:*

1. $C \cap B^* \subseteq S^*$.
2. $B^* \subseteq S^* \cup (H \cap B^*)$.
3. $|(H \cap B^*| \geq n/2 - 3\epsilon n$.
4. $|H \setminus B^*| + |S^*| \leq 6\epsilon n$.
5. *The sets* $(H \cap B^*) \setminus S^*$ *and* $C \setminus S^*$ *share no edges.*

Proof.

1. As $C \subseteq A^*$, $C \cap B^* \subseteq A^* \cap B^* = S^*$.
2. As $V = C \cup H$, $B^* = B^* \cap V = (B^* \cap C) \cup (B^* \cap H) \subseteq S^* \cup (H \cap B^*)$. The last inequality follows since $C \subseteq A^*$ and so $(B^* \cap C) \subseteq S^*$.
3. By 2 above, $B^* \subseteq S^* \cup (H \cap B^*)$. Thus $|B^*| \leq |S^*| + |H \cap B^*|$. By Assumption (b), $|B^*| \geq n/2 - \epsilon n$. and therefore $|H \cap B^*| \geq |B^*| - |S^*| \geq n/2 - 3\epsilon n$.
4. By Claim 1, $|H| \leq n/2 + \epsilon n$. By 3., $|H \cap B^*| \geq n/2 - 3\epsilon \cdot n$. Thus $|H \setminus B^*| = |H| - |H \cap B^*| \leq n/2 + \epsilon \cdot n - n/2 + 3\epsilon \cdot n = 4\epsilon n$. Therefore, $|H \setminus B^*| + |S^*| \leq 6\epsilon \cdot n$.
5. We know that $C \subseteq A^*$ and thus $C \setminus S^* \subseteq A^* \setminus S^*$. Also $(H_1 \cap B^*) \setminus S^* \subseteq B^* \setminus S^*$. The claim follows from Lemma 8.

□

Lemma 15. *Algorithm Chordal returns a solution of ratio less than* 1.876.

Proof. Algorithm Chordal applies procedure $VertexCover$ on $\mathcal{B}(C, H)$. Let D be the vertex cover. By Claim 1, $|C| \geq n/2 - \epsilon n$. By Assumption (d), $|H| \geq n/2 - \epsilon n$.

Let $\mathcal{B}(H, C, E')$ as the bipartite graph corresponding to H and C. By Claim 2 (5), $(H \setminus B^*) \cup S^*$ is a Vertex Cover of $\mathcal{B}$. Hence, by Claim 2 (4), the graph $\mathcal{B}(H, C, E')$ has a vertex cover of size at most $6 \cdot \epsilon \cdot n$.

Therefore, $|H \setminus D|, |C \setminus D| \geq n/2 - 7\epsilon n = \epsilon n$. The last inequality follows by the setting of $\epsilon = 1/16$. By Lemma 11, the ratio of $2 - 2\epsilon < 1.876$ follows. □

6.2 Exact solution for interval graphs

In this section, we give an exact solution for interval graph using dynamic programming. For the dynamic programming we have the following definition of a state. For every time t we maintain two sets A, B of vertices that are a partial solution, namely sharing no edges.

Definition 7. *A state for an ending time t contains the following information. We have a partial solution A, B. We need to carry:*

1. *The size of A*
2. *The size of B*
3. *For a time t, we remember the interval in $A \setminus B$ whose ending time is maximal in $A \setminus B$ among intervals that end at time t or earlier.*
4. *The size of $A \cap B$*

We maintain an example of a solution for each one of the states.

Say that A_1, B_1 and A_2, B_2 have the same state with respect to a time t. Let V_t be all the intervals that end at or before time t. Let $P_1 = V_t - (A_1 \cup B_1)$. Let $P_2 = V_t - (A_2 \cup B_2)$.

Claim 3. $|P_1| = |P_2|$.

Proof. We start by showings that $|A_1 \cup B_1| = |A_2 \cup B_2|$. Note that $|A_1 \cup B_1| = |A_1| + |B_1| - |A_1 \cap B_1|$. Since the two solutions have the same state: $|A_1| = |B_1|$, $|A_2| = |B_2|$ and $|A_1 \cap B_1| = |A_2 \cap B_1|$, which implies that $|A_1 \cup B_1| = |A_2 \cup B_2|$. $|P_1| = |V_t| - |A_1 \cup B_1| = |V_t| - |A_2 \cup B_2| = |P_2|$. The claim follows. □

We assume the solution is maximal for inclusion. In such case

Lemma 16. *For every solutions of a certain state and time t, there is a unique way to extend the solution to legal solutions for time t. Moreover, if A_1, B_1 and A_2, B_2 have the same state, the value of the two extensions is equal.*

Proof. Let P_1, P_2 be defined as above. Note that there is a vertex in P_1 that has a neighbor in A_1, and there is a vertex in P_1 that is a neighbor of a vertex in B_1. Otherwise we can extend one of A_1, B_1 to larger sets contradicting the maximality assumption. For example if I is an interval in P_1 with no neighbors in A_1, we may add I to A_1 and the solution is still feasible. The same goes for P_2, A_2, B_2. Thus, the vertices of P_1 must be in the intersection of A_1, B_1. Otherwise we get a contradiction to Lemma 8. Thus the only legal way to extend A_1, B_1 to a legal solution is by setting, $A_1 \cup P_1, B_1 \cup P_1$. The same applies for A_2, B_2, P_2. The unique extension is $A_2 \cup P_2, B_1 = B_2 \cup P_2$.

We now show that the extended solution A_1, B_1 has the same size as the extended solution of A_2, B_2. As $P_1 \cap A_1 = \emptyset$ and $P_1 \cap B_1 = \emptyset$, and the same claim holds for A_2, B_2, P_2, by Claim 3 $|P_1| = |P_2|$. Thus $|A_1 \cup P_1| = |A_1 + |P_1| = |A_2| + |P_2| = |A_2 \cup P_2|$. Namely, the two (unique) extensions have equal value. □

We deal with intervals by increasing finishing times.

Definition 8. *A state A, B is* extendable *if it can be extended to an optimum solution by adding intervals*

Clearly we may assume that the finishing time are pairwise distinct.

Lemma 17. *Say that we have an extendable state with time t'. Assume that t is the next lowest finishing time and let I be the unique interval to end at time t. Then we can find a collection of states so that at least one of them is extendable and contains I.*

Proof. We prove this by induction, with the base case being time 0. In time 0 the solution is the empty set and can be extended to any optimal solution. Say that we computed an extendable solution for time t'. Consider the next to end interval I, and let t be its ending time. Thus no interval ends strictly between t' and t. Let I_1, be the interval in B, for time t' with maximum finishing time (note that this information is part of a state). By the induction hypothesis the choice until time t' is extendable to an optimum solution. We now produce states with I. If I starts at time strictly before time t', then I intersects $I_1 \in B$, and it cant be that $I \in A \setminus B$. Otherwise we can add I to A only. In the same way we can check if I can be added to B only. Thus we get 4 cases.

1. If I can not be added to $A \setminus B$ nor to $B \setminus A$, add I to $A \cap B$ and set size $|A \cap B| \leftarrow |A \cap B| + 1$. Also increase $|A|$ and $|B|$ by 1 The interval I becomes the last to end in A and B with respect to time t.

2. If I can be added to A but not to B then we are forced to add I too A only. In this case $|A|$ grows up by 1, but $|B|$ and $|A \cap B|$ stay the same. Also the last to end interval in B for time t is the same one that is last to end for time t'. However, update I to be the last interval of A to end at time t

3. The case that I can be added to B only is treated in a symmetric way.

4. If I can be added both to A and to B there are 3 new states. One in which we add I to A only. One in which we add I to B only and one for which we add I to $A \cap B$. Updating the states is done as above.

□

Lemma 18. *There exists an $O(n^5)$ exact algorithm for the* Bi-Covering *problem on interval graphs.*

Proof. Consider all states for the highest finishing time t. By the definition of a state, there are at most n^5 states. By lemma 17, one of the states $A, B, V \setminus (A \cup B)$ is extendable. However since there are no more intervals, by Lemma 16 the only way to extend A, B to a solution is to add $V \setminus (A \cup B)$ to both A and B. By Lemma 16 our solution will be optimal.

To achieve the above time we assign a unique integral keys to every state. We build a perfect Hash function so that the time for insert and search is $O(1)$ in the worst case (see [9]). When we check if we should extend a leaf, we first check in worst case time $O(1)$ if the "new" states do not already appear in the tree. The number of inserts into the tree is $O(n^5)$ therefore the running time for inserting states is $O(n^5)$. $\square$

6.3 Minor free graphs

In this section, we give the $1 + o(1)$ ratio approximation algorithm for *Minor free graphs*. We need the following theorem from [1].

Theorem 7. *Every subgraph $G'(V', E')$ with n' vertices of a minor free graph G has a separator of size $O(\sqrt{n'})$. Furthermore, this separator can be found in polynomial time in n'.*

In this section, we set ϵ to be any function of n so that $O(1/\epsilon^2) = o(\sqrt{n})$. For example $1/\epsilon$ can be $\log^* n/n^{1/4}$.

Define the sets $\mathcal{S}$ and $\mathcal{L}$ as follows: In the algorithm, we maintain a collection $\mathcal{S}$ of components that belongs to one of two types. First, $\mathcal{S}$ contains some separators of some larger components. The second type are connected components with at least $\epsilon \cdot n$ vertices. The set $\mathcal{L}$ does not contain separators, and contains all connected components of size less than $\epsilon \cdot n$. The algorithm for minor free graphs is given in Figure 4. We now proceed with the analysis.

1. Initialize $\mathcal{S} \leftarrow V$, $\mathcal{L} \leftarrow \emptyset$ and a separator tree $\mathcal{T}$ with V as a root.
2. **While** $\mathcal{S}$ contains a connected component with more that $\epsilon \cdot n$ vertices **do:**
   (a) Let S be a connected components in $\mathcal{S}$ such that $|S| > \epsilon n$.
   (b) Let S_C be the separator of $G(S)$
   (c) Remove S from $\mathcal{S}$ and add S_C to $\mathcal{S}$. Replace S by S_C in $\mathcal{T}$.
   (d) Add every connected component in $G(S) \setminus S_C$ with at most $\epsilon \cdot n$ vertices to $\mathcal{L}$ and the rest of the components to $\mathcal{S}$. Also, add all the connected components in $G(S) \setminus S_C$ as children to node S_C in $\mathcal{T}$.
3. Set $A' \leftarrow \emptyset$
4. While $|A'| < n/2$ pick a connected component in $\mathcal{L}$ and add it to A'
5. Set $A = A' \cup un(\mathcal{S})$ and set $B = (V \setminus A) \cup un(\mathcal{S})$.
6. Return A, B

Figure 4: Algorithm Minor-Free

Analysis: The following claim follows by definition.

Claim 4. *At the end $\mathcal{S}$ contains only the separators used. The set $\mathcal{L}$ contains all the tree leaves, of the separation tree, each containing less than $\epsilon \cdot n$ vertices. Thus, $un(\mathcal{S}) \cup un(\mathcal{L}) = V$.*

We now bound the height of the separators tree.

Claim 5. *The height of the separators tree is at most $O(1/\epsilon)$.*

Proof. Since the size of the separators in level $i \geq 0$ is $\sqrt{n} \cdot (2/3)^i$, clearly there exists a large enough constant c so that the height of the separators tree is bounded by

$$c \cdot \ln\left(\frac{n}{\epsilon \cdot n}\right) = c \cdot \ln(1/\epsilon),$$

Using the inequality $\ln(1/\epsilon) \leq 1/\epsilon - 1$, the claim follows. □

Claim 6. *$un(\mathcal{S}) = o(n)$.*

Proof. To count the number of vertices in $un(\mathcal{S})$ we bound it by the number of non leaf separators at level i, times $O(\sqrt{n})$ times the number of levels. This applies as every separator has size $O(\sqrt{n})$.

By definition and induction it follows that the separators of level $i \geq 0$ have size bounded by $\sqrt{n} \cdot (2/3)^i$.

We now bound the possible number of non-leaf components at level i. We claim that the number of non leaf connected components in level i is at most $1/\epsilon$. To bound the number of separators in level i, we consider level $i+1$. Since we are talking on non-leaves in level i, each such S in level i has a child in level $i+1$. By definition the parent $p(S)$ of S has at least $\epsilon \cdot n$ vertices, for otherwise the parent would have been a leaf. Consider a maximal set Q of separators level $i+1$ so that the parents of the separators are pairwise distinct. If S, S' have two different parents $p(S), p(S')$, clearly $p(S) \cap p(S') = \emptyset$. Thus, the collection of leaves of a vertex at level level i, contributes $\epsilon \cdot n$ new vertices to level i. By disjointness, the number of non leaf components at level i is at most $1/\epsilon$.

Each level contains at most $1/\epsilon$ separators each of size $O(\sqrt{n})$ and the height is bounded by $O(1/\epsilon)$ thus: $un(\mathcal{S}) = O(1/\epsilon) \cdot 1/\epsilon \cdot \sqrt{n})$. By the definition of ϵ, the number of separators in the tree is $O(1/\epsilon^2 \cdot \sqrt{n}) = o(\sqrt{n}) \cdot \sqrt{n}) = o(n)$. Thus $un(\mathcal{S}) = o(n)$. □

Lemma 19. *The algorithm runs in time polynomial in n and the approximation ratio is $1 + o(1)$.*

Proof. The running time is dominated by the computation of separators and computes at most n separators. Since computing a separator requires time polynomial in n, the claim about running time follows.

By Claim 4, at the end of the algorithm, $un(\mathcal{S}) \cup un(\mathcal{L}) = V$. By Claim 6, $|un(\mathcal{L})| = n - o(n)$. Furthermore, by definition $\mathcal{L}$ has pairwise disjoint connected components, each of size at most ϵn. The sets is $\mathcal{L}$ share no edges because these sets are leaves in the separator tree and thus were separated by their least common ancestor in the tree. We start adding to $A' \leftarrow \emptyset$ and then iteratively adding to A' connected components of $\mathcal{L}$ as long as $|A'| \leq n/2$. Consider the moment a leaf $L \in \mathcal{L}$ is added and the number of vertices in A' goes becomes at least $n/2$. As the last component added is of size $|S| \leq \epsilon n$ (because all components in $\mathcal{L}$ have size less than $\epsilon \cdot n$) and by definition $|A'| \leq n/2$. Therefore, $|A'| \leq n/2 + \epsilon n$. Note that $A = A' \cup \mathcal{S}$. By Claim 6 $|A| \leq n/2 + \epsilon n + o(n)$ By the algorithm, $|A| \geq n/2$ and so, $|B| \leq n/2$. Thus the maximum size set between $|A|$ and $|B|$ is $|A|$. The approximation ratio is bounded by

$$\frac{n/2 + \epsilon n + o(n)}{n/2} = 1 + 2 \cdot \epsilon + o(1) = 1 + o(n)$$

for a chosen ϵ as claimed. $\square$

The approximation applies as special cases to planar, and bounded genus graphs, since these graphs are minor free.

6.4 Dense graphs

This section deals with graphs $G(V, E)$ which has minimum degree $\delta \cdot n$ for some constant $\delta > 0$.

1. **For** every pairs of vertices a and b so that $(a, b) \notin E$ **do:**
   (a) Let $S = N(a) \cap N(b)$.
   (b) Create a bipartite graph $\mathcal{B}(N(a) \setminus S, N(b) \setminus S, E')$ by removing edges inside $N(a) \setminus S$ and $N(b) \setminus S$ and joining $a \in N(a) \setminus S$ to $b \in N(b) \setminus S$ if $(a, b) \in E$.
   (c) Compute in polynomial time (using flow) the minimum Vertex Cover D of the graph of $\mathcal{B}$.
   (d) Set $A \leftarrow N(a) \setminus (S \cup D)$ and $B = N(b) \setminus (S \cup D)$.
2. Output the best $Big(A, B)$ over all pairs a, b.

Figure 5: Algorithm Dense

Analysis: Let the optimum solution be A^*, B^* and let $S^* = A^* \cap B^*$. Assume that

$$|S^*| \leq \epsilon \cdot n \tag{1}$$

for some ϵ as a function of δ which we will fix later. A pair a, b of vertices is good pair if $a \in A^* \setminus S^*$, and $b \in B^* \setminus S^*$. Note that a good pair exists unless $A^* \subseteq S^*$ or $B^* \subseteq S^*$. If one of these two cases holds, $opt = n$ and $(V, \emptyset)$ is an optimal solution. Also recall by Claim 8, that $(a, b) \notin E$.

Claim 7. *Let a, b be a good pair. Then $N(a) \subseteq A^*$ and $N(b) \subseteq B^*$*

Proof. We show the first containment. The proof of the second containment is identical. The complement set of A^* is the set $V \setminus A^* = B^* \setminus S^*$. If the sets $N(a)$ and $B^* \setminus S^*$ intersect, as $a \in A^* \setminus S^*$ there is an edge between of internal vertex of $A^* \setminus S^*$ and a vertex of $B^* \setminus S^*$ which contradicts Lemma 8. $\square$

From now on we assume that a and b are the good pair and hence

$$N(a) \subseteq A^*, \quad N(b) \subseteq B^* \tag{2}$$

Let $S = N(a) \cap N(b)$.

Claim 8. $|S \cup D| \le \epsilon n$.

Proof. First note that $S \subseteq S^*$. This follows as $N(a) \subseteq A^*$ and $N(b) \subseteq B^*$ as shown in (2), $S = (N(a) \cap N(b)) \subseteq (A^* \cap B^*) = S^*$. The set $S^* \setminus S$ is one of the vertex covers of $\mathcal{B}$ and hence the minimum vertex cover D has size at most $|S^* \setminus S| = |S^*| - |S|$. Thus, $|S \cup D| \le |S^*|$ which is at most ϵn by our assumption. $\square$

Lemma 20. *The approximation ratio is* $2 - 4\delta/3$.

Proof. The approximation ratio is $\max\{2/(1+2\epsilon), 2 - 2(\delta - \epsilon)\}$, because the case that $|S^*| > \epsilon \cdot n$, gives $2/(1+2\epsilon)$ ratio. Otherwise if $|S^*| \le \epsilon \cdot n$, then as $|N(a)| = |N(b)| \ge \delta \cdot n$ and $|S \cup D| \le \epsilon n$, the number of vertices in $N(a) \setminus (S \cup D)$ and $N(b) \setminus (S \cup D)$ is at least $(\delta - \epsilon) \cdot n$. The $2 - 2(\delta - \epsilon)$ ratio then follows from Lemma 9. If we put $\epsilon = \delta/3$, the approximation ratio we get is $2 - 4\delta/3$. $\square$

6.5 Expander graphs

Say that G is a vertex expander with parameter δ for some constant δ. Let A^*, B^* be the optimum solution. By Lemma 5 we may assume that $|A^*| \ge n/2 - \delta \cdot n/8$.

Lemma 21. $|N_1(A^*)| \ge \delta n/4$

Proof. The worst case its may be that $|A^*| = n/2 - \delta \cdot n/8$. See Lemma 5. As the graph is an expander, $|N_1(A^*)| \ge \delta n/2 - \delta^2 n/8$. By definition of an expander $\delta \le 2$. Thus $|N_1(A^*)| \ge \delta n/2 - \delta \cdot n/4 = \delta \cdot n/4$. The above follows by replacing one of the δ in the expression $\delta^2 n/8$ by 2. $\square$

Lemma 22. *Returning* $V, \emptyset$ *gives ratio at most* $2/(1 + \delta^2/8)$ *for* Bi-Covering *on expanders.*

Proof. By definition $N_1(A^*) \cap A^* = \emptyset$. Thus $N_1(A^*) \subset S^* \cup (B^* \setminus S^*)$. If half of $N_1(A^*)$ belongs to S then $S^* \ge \delta \cdot n/8$. By Claim 6 the resulting ratio is $2/(1+\delta/8) = 16/(8+\delta)$ which is a constant less than 2. Else we may assume that $|(B^* \cap N_1(A^*)) \setminus S^*| \ge \delta \cdot n/8$. Let $D = (B^* \cap N_1(A^*)) \setminus S^*$. Note that all the neighbors of of D belong to S^*. Otherwise there is an edge between a vertex in $D \subseteq B^* \setminus S^*$ and $A^* \setminus S^*$ which gives a contradiction. Recall that $|D| \ge \delta \cdot n/8$. Since the graph is an expander, in the worse case $N_1(D) \ge \delta^2 n/8$. Since $N_1(D) \subseteq S^*$, $|S^*| \ge \delta^2 \cdot n/8$. Thus a ratio of $2/(1 + \delta^2/8)$ follows from Lemma 1. $\square$

6.6 Bounded degree graphs

Let the maximum degree in G is d for some constant d. We assume that $d \ge 3$. Indeed, if $d = 2$ the graph is a collection of paths and cycles and the problem can be solve optimally on such graphs.

Lemma 23. *The ratio is at most* $2 - 6 \cdot /(5d)$

Proof. Note that $N_1(A') \le 3n/5$ because $|A'| \le 3n/(5d)$ and the maximum degree is d. The size of V' is at least $n - 3n/(5d) - 3n/5 = 2n/5 - 3n/(5d) \ge 3n/(5d)$. The last inequality follows as $d \ge 3$. This means that we can select a set B' of size at least $3n/5$ that has no edges to A' of the same size as A'. By Lemma 9, the ratio is $2 - 6/(5 \cdot d)$. Note that this ratio is better than $1 - 1/d$. $\square$

1. Set $S = \emptyset$. Choose a set A' of $3n/5d$ vertices and add $N_1(A')$ to S
2. Let $V' \leftarrow V \setminus (S \cup A')$
3. Select a set $B' \subseteq V'$ of size $3 \cdot n/(5 \cdot d)$and add $N_1(B')$ in S
4. Set $A \leftarrow A' \cup S,\ B \leftarrow B' \cup S$ and output (A, B)

Figure 6: Algorithm Bounded-Degree

6.7 Sparse graphs

The case of sparse graphs is the case $|E| = O(n)$. Thus the average degree $\bar{d}$ is constant. This is a weaker condition than maximum degree d hence we get a worse ratio.

Lemma 24. *The* BI-COVERING *problem admits a polynomial time* $2-1/(\bar{d}+1)$ *ratio approximation on graphs with* $O(n)$ *edges with average degree* $\bar{d}$.

Proof. Applying the Turán's theorem [23], there exists an independent set I' of size $n/(\bar{d}+1)$ that can be found in polynomial time. We partition the independent set to two independent sets I_1 and I_2 of size $n/2(\bar{d}+1)$. and I_2 of size $n/2(\bar{d}+1)$. Put I_1 in A and I_2 in B. We now apply Lemma 9 with $\epsilon = 1/(2(\bar{d}+1))$. The ratio resulting is $2 - 1/(\bar{d}+1)$ which is a constant less than 2 as $\bar{d}$ is a constant by itself. □

6.8 Split graphs

A split graph is composed of a clique C and an independent set I with arbitrary edges among the two. The algorithm is at follows:

1. Set $\epsilon = 1/8$.
2. If the size of the independent set is I at least $2\epsilon \cdot n$ let A', B' be disjoint halves of I. Set $A \leftarrow A' \cup C$ and $B \leftarrow B' \cup C$ (with C the clique)
3. Else return $V, \emptyset$

Lemma 25. *The approximation ratio of the above algorithm is at most* $8/5$ *for split graphs.*

Proof. The case that $|I| \geq 2 \cdot \epsilon \cdot n$, by Lemma 9, the ratio is $2 - 2\epsilon$. If $|I| \leq 2\epsilon n$ and $|S^*| \geq \epsilon n$, by Lemma 6 the ratio is $2/(1 + 2\epsilon)$. Thus the remaining case is that $I \leq 2\epsilon \cdot n$ and $|S^*| \leq \epsilon n$. By Lemma 7, we may assume without loss of generality that that $C \subseteq A^*$. Thus, $C \cap (B^* \setminus S^*) = \emptyset$. Thus $|B^*| \leq |S^*| + |I| \leq 3\epsilon n$. This implies that $|A^*| \geq n - 3\epsilon n$. In this case returning $V, \emptyset$ gives ratio $1/(1 - 3\epsilon)$. Consider $\max\{1/(1 - 3\epsilon), 2/(1 + 2\epsilon), 2 - 2\epsilon\}$ Choosing $\epsilon = 1/8$ gives ratio at most $8/5$. □

6.9 Graph that contain a separator of size $o(n/\sqrt{\log n})$

Lemma 26. *If G has a separator S and $|S| = o(n/\sqrt{\log n})$ then the* Bi-Covering *problem admits a $4/3 + o(1) < 2$ ratio.*

Proof. In [12] an $O(\sqrt{\log n})$ approximation algorithm is given for the Min Size Vertex Separator problem. As S is a vertex separator, clearly the approximation algorithm returns a solution of size at most $O(\sqrt{\log n}) \cdot |S| = o(n)$, and maximum connected component of size $2n/3$. Thus, adding S to the connected component, gives $2n/3(1 + o(1))$ maximum size of every set. Dividing by the best possible optimum of value $opt = n/2$, the claim follows. $\square$

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

References

[1] N. Alon, P. Seymour, and R. Thomas. A separator theorem for graphs with an excluded minor and its applications. In *Proceedings of the Twenty-second Annual ACM Symposium on Theory of Computing*, STOC '90, pages 293–299, New York, NY, USA, 1990. ACM.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] Qi Cheng and Daqing Wan. A deterministic reduction for the gap minimum distance problem:[extended abstract]. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 33–38. ACM, 2009.

[9] T. Cormen, C. Stein, R Rivest, and C. Leiserson. *Introduction to Algorithms.* McGraw-Hill Higher Education, 2nd edition, 2001.

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

[11] 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.

[12] Uriel Feige, MohammadTaghi Hajiaghayi, and James R. Lee. Improved approximation algorithms for minimum weight vertex separators. *SIAM J. Comput.*, 38(2):629–657, 2008.

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

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

[15] J. Gilbert, D. Rose, and A. Edenbrant. A separator theorem for chordal graphs. *Siam J. on Discrete Methods*, 5:306–313, 1984.

[16] 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.

[17] 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.

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

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] Paul Turán. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok (in Hungarian)*, 48:436–452, 1941.