Bicovering: Covering edges with two small subsets of vertices

Amey Bhangale\textsuperscript{1}, Rajiv Gandhi\textsuperscript{2}, Mohammad T. Hajiaghayi\textsuperscript{3}, Rohit Khandekar\textsuperscript{4}, and Guy Kortsarz\textsuperscript{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, Maryland, USA hajiagha@cs.umd.edu
4 KCG holdings Inc., USA rkhandekar@gmail.com
5 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 [13]. A solution that outputs $V, \emptyset$ gives ratio at most 2. We show that under the 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 [10] and also under the assumption that $\text{NP} \not\subset \text{BPTIME}(2^{n^{\epsilon}})$ [17]. 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 $\delta n$, $2/(1 + \delta^2/8)$ for $\delta$-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.

\textsuperscript{*} Research supported in part by NSF grant CCF-1253886.
\textsuperscript{1} Supported in part by NSF grant number 1218620.
\textsuperscript{2} Supported in part by NSF CAREER award 1650120, NSF grant CCF-161626, ONR YIP award N000141110662, and DARPA/AFOSR grant FA9550-12-1-0423.
\textsuperscript{3} Supported in part by NSF grant number 1218620 and by NSF grant 1540547

© Amey Bhangale, Rajiv Gandhi, Mohammad T. Hajiaghai, Rohit Khandekar and Guy Kortsarz; licensed under Creative Commons License CC-BY

Conference title on which this volume is based on.
Editors: Billy Editor and Bill Editors; pp. 1–24
Leibniz International Proceedings in Informatics
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 [13]. 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 approximated within any factor strictly less than $2$ assuming a strong Unique Games Conjecture by [6] (see Conjecture 12).

> **Theorem 1.** Let $\epsilon > 0$ be any small constant. Assuming a strong Unique Games Conjecture of [6], 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 $3/2 - \epsilon$ 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 $3/2$-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 $3/2 - \epsilon$ 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, 10, 17]. Feige [10] 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 $n^\delta$ for some $\delta > 0$ using known amplification technique [2, 7]). Feige-Kogan [12] showed that assuming $SAT \notin \text{DTIME}(2^{n^{1/4+\epsilon}})$ 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 $n^{1/2^{\sqrt{\log n}}}$ approximation. Khot [17] later proved a similar inapproximability result but assuming $NP \subseteq \text{BPTIME}(2^\epsilon n)$ using a quasi-random PCP. 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 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 $n^\delta$ 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 [18], optimal algorithm for every Max-CSP [20], Ordering CSPs [15], characterizing strong approximation resistance of CSPs [19] 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]. The **Strong Unique Games Conjecture**, which has a weak expansion property, has been used earlier in [6] and [21] 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 **Strong Unique Games Conjecture**, 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. See section A.4 for the definition of graph classes. 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):

2. Interval graphs: exact $O(n^5)$ time algorithm.
3. Minor Free Graph: $1 + o(1)$.
4. Graph with minimum degree $\delta$: $2 - 4\delta/3$.
5. $\delta$-vertex expander Graph: $2/(1 + \delta^2/8)$.
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 $\epsilon n$ with no edges in between, then this itself gives $2 - \epsilon$ approximation (see Lemma 26). To get better bound on $\epsilon$ 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. In Section A we present notations and tools required for our approximation algorithms. Finally, in section B we present our approximation algorithms for special graph classes.

### 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 $A, B \subseteq V$, $|A|, |B| \geq 1/2 - \epsilon$ such that there are no edges between $A$ and $B$.
2. There exists no two disjoint sets $A, B \subseteq V$, $|A|, |B| \geq \epsilon$ 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. 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 $\omega_q$ be the $q^{th}$ root of unity. For a vector $\alpha \in \mathbb{F}_q^n$, we will denote $\alpha_i$ the $i^{th}$ coordinate of vector $\alpha$. 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 $1 = \lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{q-1}$ 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 $\text{Inf}_i(f)$ is defined as:

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

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

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

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

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

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

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:

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

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

where $X$ and $Y$ are two standard Gaussian variables with covariance $\rho$. 
We are now ready to state the invariance principle from [9] that we need for our reduction.

> **Theorem 9** ([9]). Let $T$ be a symmetric Markov operator on $\mathbb{F}_q^n$ 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

$$\min_{i} (\text{Inf}_{i}^{g,k}(f), \text{Inf}_{i}^{g,k}(g)) \leq \delta$$

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

$$\langle f, T^{\otimes n} g \rangle \geq \sum_{i} (\mu, \nu) - \tau$$

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 $G = (U, V, E, \pi_e, \pi_e)_{e \in E}$ 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 $\ell$ if $\pi_e(\ell(v)) = \ell(u)$.

$G$ is said to be at most $\delta$-satisfiable if every labeling satisfies at most a $\delta$ fraction of the edges.

The following is a conjecture by Khot [16] which has been used to prove many tight inapproximability results.

> **Conjecture 11** (Unique Games Conjecture [16]). For every sufficiently small $\delta > 0$ there exists $L \in \mathbb{N}$ such that the following holds. Given an instance $G = (U, V, E, \pi_e)_{e \in E}$ 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 $\delta$ fraction of the edge constraints.

Our hardness results are based on the following stronger conjecture by Bansal-Khot [6]. We refer readers to [6] for more discussion on comparison between these two conjectures.

> **Conjecture 12** (Strong Unique Games Conjecture [6]). For every sufficiently small $\delta, \gamma, \eta > 0$ there exists $L \in \mathbb{N}$ such that the following holds: Given an instance $G = (U, V, E, \pi_e)_{e \in E}$ 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 $\gamma$ fraction of the edge constraints. Moreover, the instance satisfies the following expansion property. For every set $S \subseteq V$ with $|S| = \delta|V|$, we have $|\Gamma(S)| \geq (1 - \delta)|V|$, where $\Gamma(S) := \{u \in S | \exists v \in V : (u, v) \in E\}$.

### 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 $f : \mathbb{F}_q^n \rightarrow \mathbb{R}$. $f$ 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:

Let $q > 2$ be any prime number. Let $G(\mathbb{F}_q, E)$ be a 3-regular graph on $\mathbb{F}_q$ with self loops as shown below:

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:

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

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

4.2.1.2 Completeness:

Let $f : \mathbb{F}_q^R \rightarrow \mathbb{R}$ 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 $B \setminus A$. 
8 Bicovering: Covering edges with two small subsets of vertices

More precisely,
\[ A \setminus B = \{ x \in \mathbb{F}_q^R \mid x_i \in \{1, 2, \ldots, [q/2]\} \}, \]
\[ B \setminus A = \{ y \in \mathbb{F}_q^R \mid y_i \in ([q/2] + 1, [q/2] + 2, \ldots, q - 1) \} \]

By the property of Markov operator \( T^{\otimes R} \), \((x, y)\) are not adjacent if \((x_i, y_i) \notin E\) 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
\[
\mathbb{E}_{x \sim \mathbb{F}_q^R, y \sim T^{\otimes R}(x)} [f(x)g(y)] = \langle f, T^{\otimes R} g \rangle = 0.
\]

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

We can now apply Theorem 9 to conclude that there exists \(i \in [R]\) and \(k \in \mathbb{N}\) independent of \(R\) such that
\[
\min\{\mathsf{Inf}_i^{\leq k}(f), \mathsf{Inf}_i^{\leq k}(g)\} \geq \delta,
\]
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 \textsc{Unique Game} with parameters \(\delta < \frac{1}{4}, \gamma, \eta > 0\) from Conjecture 12. We replace each vertex \(v \in V\) by a block of \(q^k\) 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 \(\pi_1\) and \(\pi_2\) 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 \(T^L((x \circ \pi_1^{-1}) \leftrightarrow (y \circ \pi_2^{-1})) \neq 0\) (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 \(H\).

\textbf{Lemma 13 (Completeness).} \textit{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 $\ell$ such that for at least $(1 - \eta)$ fraction of vertices $v$ in $V$ in $G$, all edges attached to them are satisfied. Suppose $X$ be the set of remaining $\eta$ fraction of vertices of $V$ in $G$. For every vertex $v \in V$, consider the following two partitions of $[v]$:

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

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

$$C_v = \{x \in [q]^L : x_{\ell(v)} = 0\}$$

Let $A = \cup_{v \in V} (A_v \cup C_v) \cup_{\in \in X} [z]$ and $B = \cup_{v \in V} (B_v \cup C_v) \cup_{\in \in X} [z]$. 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 $\pi_1$ and $\pi_2$ be the associated permutation constraints. Since $X \subseteq A \setminus B$, $v_1, v_2 \notin X$. Hence $\ell$ 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 $a_{\pi_1^{-1}(j)} = a_{\ell(v_1)} \in \{1, \ldots, [q/2]\}$. Similarly, $b_{\pi_2^{-1}(j)} \in \{[q/2] + 1, [q/2] + 2, \ldots, q\}$. By the construction of edges in $H$, $a$ and $b$ are not adjacent.

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

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

* Lemma 14 (Soundness). For every constant $\epsilon > 0$, there exists a constant $\gamma$ such that, if $G$ is at most $\gamma$-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 sets $X, Y$ of size at least $\epsilon$ 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 15. The total fraction of edges connected to $X^*$ whose other end point is in $\Gamma(X^*) \cap \Gamma(Y^*)$ is at least $\frac{\epsilon}{2}$.

Proof. Let $G$ has left-degree $d_1$ and right-degree $d_2$. We have $d_1 = \frac{d_1|V|}{|U|}$. 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 $\frac{(1/2)|V|X^*}{|U|} \geq \frac{(d_2/2)(\epsilon/2)|V|}{|U|}$ which is greater than $d_1$ as $\epsilon > 4\delta$.

For $v \in X^* \cup Y^*$, let $f_v : [q]^L \rightarrow \{0, 1\}$ 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 $k \in \mathbb{N}$:

$$F(v) := \{i \in [L] \mid \inf_{k}^{\ell} (f_v) \geq \tau'\}$$

We have $|F(v)| \leq \frac{\tau'}{k}$ as $\sum F_i^{\ell} (f_v) \leq k$.

* 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 $\pi_{e_1}(F(v)) \cap \pi_{e_2}(F(w)) \neq \emptyset$. 
Bicovering: Covering edges with two small subsets of vertices

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

$$
\mathbb{E}_{(x \sim \pi_{e_1}^{-1}(y) \sim \pi_{e_2}^{-1})} [f_e(x \circ \pi_{e_1}^{-1})f_w(y \circ \pi_{e_2}^{-1})] = 0
$$

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

$$
\min(\text{Inf}_{\pi_{e_1}(v)}^{\ell}(f_v), \text{Inf}_{\pi_{e_2}(v)}^{\ell}(f_w)) \geq \tau',
$$

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

4.2.2.1 Labeling:

Fix $\tau'$ and $k$ from Lemma 16. We now define a labeling $\ell$ 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 $F(v)$. 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 $\pi(u,w)(F(w))$ of size at most $\frac{k}{2}$. 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 $\pi(u,w)(F(w)) \cap \pi(u,v)(F(v)) \neq \emptyset$, The probability that the edge is satisfied by the randomized labeling is at least $\left(\frac{\tau'}{2}\right)^2$. Thus in expectation, at least $\frac{1}{2} \left(\frac{\tau'}{2}\right)^2$ 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 $\frac{1}{2} \left(\frac{\tau'}{2}\right)^2$ fraction of edges in $G$. Setting $\gamma < \frac{1}{2} \left(\frac{\tau'}{2}\right)^2$ 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 that $A$ and $B$ satisfy the property of being a Bi-Covering. As $\max(|A|, |B|) \leq \frac{1}{2}|V|$, 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 $\frac{1}{2|V|}|A' \cup B'| \geq \frac{3}{4} + \epsilon$ and $\frac{1}{2|V|}|B' \setminus A'| \geq \frac{1}{4} + \epsilon$. Thus, we can find two sets, say $X'$ and $Y'$, 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 $\frac{1}{1|V|} \min(|X|, |Y|) \geq \frac{1}{2}$. 


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.

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.

References

In this section, we present notations and preliminaries which are required for approximation algorithms for Bi-Covering for several classes of graphs on which we get strictly better than 2 approximation.

A.1 Preliminaries

Recall, the Bi-Covering problem is:

**Input:** A graph $G(V,E)$

**Output:** Two subsets $A, B \subseteq V$ so that every edge $(u,v)$ either $\{u,v\} \subseteq A$ or $\{u,v\} \subseteq B$. Minimize $\max\{|A|, |B|\}$.

For a given graph $G$, let $N(v)$ denote the vertices joined to $v$ (its neighbors). The number of neighbors of $v$ is denoted by $deg(v)$.

> **Definition 17.** For a set $S$ let $N(S)$ be $S$ 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 $X$.

> **Definition 18.** Given a set $S = \{S_i\}$, define $un(S) := \bigcup_{S \in S} S$. 

---

A Better Approximation Algorithms for special graphs

In this section, we present notations and preliminaries which are required for approximation algorithms for Bi-Covering for several classes of graphs on which we get strictly better than 2 approximation.
Definition 19. The optimum partition for the Bi-Covering instance at hand is denoted by $A^*, B^*$ and we denote $S^* = A^* \cap B^*$

Lemma 20. If $|A^*| \leq n/2 - \epsilon \cdot n$ or $|B^*| \leq n/2 - \epsilon \cdot n$ then returning $V, \emptyset$ gives a $2/(1 + 2\epsilon)$ ratio.

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

A.2 Basic definitions and tools

Definition 21. Say that we remove a set $C$ 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 $\max\{|\text{un}(X)|, |\text{un}(Y)|\}$

Lemma 22. 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 $\max x_i \leq n$ and there exists an exact polynomial time solution for subset sum. For every number $Q$ between 1 and $|\bigcup H_i|$, check if there is a collection of connected components $X$ so that $|\text{un}(X)| = Q$. For any feasible $X$ check the value of the solution $\text{un}(X), V \setminus \text{un}(X)$. Output the best solution over all $Q$. Clearly this is the optimum solution.\hfill ●

Definition 23. The above algorithm is called the 2-Cover algorithm.

Let $A, B$ be a feasible solution for the Bi-Covering problem and let $S = A \cap B$

Lemma 24. 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.\hfill ●

Our problem is related to the Vertex Separator problem.

Definition 25. 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

The next lemma is used several times in the rest of the paper.

Lemma 26. 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 $2 - 2\epsilon$.\hfill ●

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 \(A = A' \cup S, B = B' \cup D\)

> **Lemma 27.** If \(|S^*| \geq 2\varepsilon \cdot n\) returning \(V, \emptyset\) gives an approximation ratio of \(2/(1 + 2\varepsilon)\)

**Proof.** Note that \(|A^*| + |B^*| = (|V \setminus S^*| + 2|S^*|) \geq (1 + 2\varepsilon)n\). Thus either \(|A^*| \geq n/2 + \varepsilon \cdot n/2\), or \(|B^*| \geq 1/2 + \varepsilon \cdot n/2\). We return the trivial solution \(V, \emptyset\). Those the solution \(V, \emptyset\) the ratio derived is at most
\[
\frac{1}{1/2 + \varepsilon} = \frac{2}{1 + 2\varepsilon}.
\]

The next lemma is also used several times.

> **Lemma 28.** Let \(C\) be a Clique. Then either \(C \subseteq A^*\) or \(C \subseteq B^*\).

**Proof.** Note that \(V \setminus A^* = B^* \setminus S^*\). If \(C \not\subseteq A^*\) it means that there is a vertex \(u \in C \cap (B^* \setminus S^*)\). If \(C \not\subseteq B^*\) 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 24.

### A.3 A central algorithmic tool

> **Definition 29.** Given two disjoint sets \(C, C_1\), the bipartite graph \(B(C, C_1, E')\) that corresponds to \(C, C_1\) is defined as follows. Make \(C\) one side of the bipartite graph and remove all edges internal to \(C\). Make \(C_1\) the other side of the bipartite graph and remove all edges internal to \(C_1\). The edges \(E'\) are all edges with one endpoint in \(C\) and the other in \(C_1\).

We present a simple procedure that takes two disjoint sets \(X, Y\) and return a solution based on their bipartite graph and on removing a vertex cover of the bipartite graph.

**Procedure Vertex-Cover**(X,Y):

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

> **Lemma 30.** If \(|X \setminus D| \geq \varepsilon \cdot n\) and \(|Y \setminus D| \geq \varepsilon n\), Algorithm Vertex-Cover returns a solution of ratio \(2 - 2\varepsilon\)

**Proof.** Note that there are no edges between \(X \setminus D\) and \(Y \setminus D\) because a vertex cover was removed. Thus by Lemma 26 the ratio of the returned solution is \(2 - 2\varepsilon\).

### A.4 Graph types

A \(\delta\)-vertex expander is a graph so that for every \(S\) of size \(|S| \leq n/2\), \(N_1(S) \geq \delta n\). 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.

## B Algorithmic results

We restate Theorem 4 for convenience here.

> **Theorem 31.** The Bi-Covering problem admits polynomial time algorithms that attain the following ratios (Graph type: approximation ratio):

2. Interval Graphs: exact $O(n^5)$ time algorithm.
3. Minor Free Graph: $1 + o(1)$.
4. Graph with minimum degree $\delta n$: $2 - 4\delta / 3$.
5. $\delta$-vertex expander Graph: $2/(1 + \delta^2 / 8)$.
7. Graphs with minimum degree $\delta$: $2 - (6/5) \cdot 1/d$.

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

> **Theorem 32.** 1. The Bi-Covering problem admits a $2 - 1/c$ ratio for any perfect graph whose chromatic number is some constant $c$.
2. 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 $G$.
3. 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 in the following sections.

### B.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 [14].

> **Theorem 33.** 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.

> **Definition 34.** The clique separator of $G$ is denoted by $C$. Denote $H = V \setminus C$. Note that $V \setminus C$ decomposes to a collection of connected components $H_1, \ldots, H_q$ so that $H = \bigcup_{i=1}^{q} H_i$. Without loss of generality, let $H_1$ be largest connected component.

**Algorithm Chordal:**

1. If $|H_1| \leq \epsilon n$, apply Algorithm 2-Cover (from Definition 23) on $H_1, H_2, \ldots, H_q$. Let $A, B$ be the partition. Return $A \cup C, B \cup C$.
2. If $|\bigcup_{i=2}^{q} H_i| \geq \epsilon \cdot n$ Apply algorithm $Biq(H_1, \bigcup_{i=2}^{q} H_i)$ /* Note that $H_1$ and $\bigcup_{i=2}^{q} H_i$ share no edges */
3. Else, Apply Algorithm Vertex-Cover on $H$ and $C$ and return the solution.

**Analysis:** Let the optimum partition be $A^*, B^*$ with $S^* = A^* \cap B^*$. Set $\epsilon = 1/16$.

- **Lemma 35.** 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. \hfill $\blacksquare$

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) $|A^*| \geq n/2 - \epsilon \cdot n$ and $|B^*| \geq n/2 - \epsilon$, and
(c) $|S^*| \leq 2\epsilon \cdot n$.
(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 20 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 ratio is the same $2/(1 + 2\epsilon) < 1.876$. (d) follows as $H = V \setminus C$ and by Lemma 35 as $H \cup C = V$.

- **Lemma 36.** 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. Initialize $A \leftarrow \emptyset$. As longs as $|A| \leq \epsilon \cdot n$ add an arbitrary $H_i$ to $A$. Note that for every $i$, $|H_i| \leq \epsilon \cdot n$ because $H_1$ is the largest connected component. This implies that when we stop, $|A| \leq 2\epsilon n$. By Assumption (d), $H \geq n/2 - \epsilon \cdot n$. Therefore $H - |A| \geq n/2 - \epsilon \cdot n - 2\epsilon n = n/2 - 3\epsilon \cdot n > \epsilon n$ 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 $\epsilon \cdot n$, by Lemma 26, a ratio of $2 - 2\epsilon n$ applies. For $\epsilon = 1/16$ this ratio is less than 1.876. \hfill $\blacksquare$

- **Lemma 37.** If $\bigcup_{i=1}^{n} H_i > \epsilon \cdot n$ Bi-Covering admits a $2 - 2\epsilon < 1.876$ ratio.

**Proof.** In this case Algorithm Chordal applies Algorithm Big. By Lemma 36, we may assume that $|H_1| \geq \epsilon n$. Therefore we have two sets $H_1$ and $\bigcup_{i=1}^{n} H_i$, both of size at least $\epsilon \cdot n$ and share no edges, (because they are connected component) a ratio of $2 - 2\epsilon < 1.876$ follows for Lemma 26. \hfill $\blacksquare$

- **Lemma 38.** $|C| \geq n/2 - \epsilon \cdot n$ and $|H| \leq n/2 + \epsilon n$.

**Proof.** Note that $|H_1| \leq n/2$ by Theorem 33 as $H_1$ is a connected component resulting after $C$ is removed. Note that $|V \setminus C| = H = H_1 \cup \bigcup_{i=2}^{n} H_i \leq n/2 + \epsilon n$. The last inequality follows because $|H_1| \leq n/2$ and by Lemma 37. This gives ratio less than 1.876. Thus $|C| \geq n - |H| \geq n/2 - \epsilon \cdot n$. And $|H| = V - |C| \leq n/2 + \epsilon \cdot n$. \hfill $\blacksquare$

- **Lemma 39.** 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 Lemma 38, \(|H| \leq n/2 + \epsilon n\). By 3., \(|H \cap B^*| \geq n/2 - 3\epsilon n\). Thus \(|H \setminus B^*| = |H| - |H \cap B^*| \leq n/2 + \epsilon \cdot n - n/2 + 3\epsilon \cdot n = 4\epsilon n\). Therefore, \(|H \setminus B^*| + |S^*| \leq 6\epsilon n\).
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 24.

Define \(B(H, C)\) as the bipartite graph corresponding to \(H\) and \(C\).

**Lemma 40.** The graph \(B(C, H)\) has a vertex cover of size at most \(6 \cdot \epsilon \cdot n\).

Proof. By Lemma 39 (5), \((H \setminus B^*) \cup S^*\) is a Vertex Cover of \(B\). The claim follows by Lemma 39 (4). The optimal vertex cover may be even smaller.

**Lemma 41.** Algorithm Chordal returns a solution of ratio less than 1.876.

Proof. Algorithm Chordal applies procedure VertexCover on \(B(C, H)\). Let \(D\) be the vertex cover. By Lemma 38, \(|C| \geq n/2 - \epsilon n\). By Assumption (d), \(|H| \geq n/2 - \epsilon n\). Therefore by Lemma 40 \(|H \setminus D|, |C \setminus D| \geq n/2 - 7\epsilon = \epsilon n\). The last inequality follows by the setting of \(\epsilon = 1/16\). By Lemma 30, the ratio of 2 - 2\(\epsilon < 1.876\) follows.

**B.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 \(A_t, B_t\) that are a partial solution, namely sharing no edges.

**Definition 42.** 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 \(B\) among intervals that end at time \(t\) or earlier.
4. The interval that belongs to \(B \setminus A\) that is the last to end before or at time \(t\).
5. 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 \(P_1 = V_t \setminus (A_1 \cup B_1)\). Let \(P_2 = V_t \setminus (A_2 \cup B_2)\).

**Lemma 43.** \(|P_1| = |P_2|\).

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 \(|A_1 \cap B_1| = |A_2 \cap B_2|\), 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 44.** 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 \( A \) and \( B \) 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 \( A \), and there is a vertex in \( P_1 \) that is a neighbor of a vertex in \( B \). Otherwise we can extend one of \( A, B \) to larger sets contradicting the maximality assumption. For example if \( I \) is an interval in \( P_1 \) with no neighbors in \( A \), we may add \( I \) to \( A \) 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 24. 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 Lemma 43 \( |P_1| = |P_2| \). 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 45.** 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 46.** 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 \( I_1 \), 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 \( I \in A \cup B \). 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 \( A \backslash B \) nor to \( B \backslash A \), 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 47. 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 46, one of the states $A, B, V \setminus (A \cup B)$ is extendable. However since there are no more intervals, by Lemma 44 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 44 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 [8]).

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 $O(n^5)$.

B.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 48. 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 $\epsilon$ to be any function of $n$ so that $O(1/\epsilon^2) = o(\sqrt{n})$. For example $1/\epsilon$ can be $\log^{*} n/n^{1/4}$.

Definition 49. The set $S$ and $L$. In the algorithm we maintain a collection $S$ of components that belongs to one of two types. First, $S$ contains some separators of some larger components. The second type are connected components with at least $\epsilon \cdot n$ vertices. The set $L$ does not contain separators, and contains all connected components of size less than $\epsilon \cdot n$.

Definition 50. A basic step is as follows:

1. Find a connected components $S \in S$ that has at least $\epsilon \cdot n$ vertices (if any).
2. Identify the separator $S_C$ of $S$
3. Let $H_1, H_2, \ldots$ be the connected components of $S \setminus S_C$.
4. Each connected component $H_i$ with less than $\epsilon \cdot n$ vertices is added into $L$. Other components are added to $S$. The separator $S_C$ is added to $S$ and the component $S$ is deleted from $S$.

Definition 51. The separators tree: The above process creates in a natural way a tree of separators.

1. Initial step The set $V$ is the root. Replace $V$ by its separator $V_C$. Compute the connected component $V \setminus S_C$
2. Every connected component in $V \setminus V_C$ becomes a child of $V_C$.
3. General step: The following is done as long as there is a connected components $S \in S$ with at least $\epsilon \cdot n$ vertices.
   a. Replace $S$ by its separator $S_C$.
   b. All connected components of $S \setminus S_C$ are designated to be children of $S_C$.

The following lemma follows by definition.

Lemma 52. At the end $S$ contains only the separators used. The set $L$ contains all the tree leaves, of the separation tree, each containing less than $\epsilon \cdot n$ vertices. Thus, $\text{un}(S) \cup \text{un}(L) = V$. 

Algorithm Minor-Free:

1. $S \leftarrow V$. $L \leftarrow \emptyset$.
2. While $S$ contains a connected component with more than $\epsilon \cdot n$ vertices do:
   a. Let $S$ be a connected components in $S$ such that $|S| > \epsilon n$.
   b. Let $S_C$ be the separator of $G(S)$
   c. Remove $S$ from $S$ and add $S_C$ to $S$.
   d. Add every connected component in $G(S)\setminus S_C$ with at most $\epsilon \cdot n$ vertices to $L$ and the rest of the components to $S$.
3. Set $A' \leftarrow \emptyset$
4. While $|A'| < n/2$ pick a connected component in $L$ and add it to $A'$
5. Set $A = A' \cup un(S)$
6. Set $B = (V \setminus A) \cup un(S)$
7. Return $A, B$

Analysis:

> Lemma 53. The separators of level $i \geq 0$ have size bounded by $\sqrt{n} \cdot (2/3)^i$.

Proof. The proof follows by induction from the definition.

We now bound the possible number of non-leaf components at level $i$.

> Lemma 54. The number of non leaf connected components in level $i$ is at most $1/\epsilon$.

Proof. 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 $1/\epsilon$.

We now bound the height of the separators tree.

> Lemma 55. 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.

To count the number of vertices in $S$ 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 $O(\sqrt{n})$.

> Lemma 56. $un(S) = o(n)$

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


Lemma 57. The algorithm runs in time polynomial in \( n \).

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 follows.

Lemma 58. The approximation ratio is \( 1 + o(1) \).

Proof. By Lemma construction, at the end of the algorithm, \( u_n(S) \cup u_n(L) = V \). By Lemma 56 \( |u_n(L)| = n - o(n) \). Furthermore, by definition \( L \) has pairwise disjoint connected components, each of size at most \( \epsilon n \). The sets are \( L \) 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 \( L \) as long as \( |A'| \leq n/2 \). Consider the moment a leaf \( L \in L \) 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 \( L \) have size less than \( \epsilon n \)) and by definition \( |A'| \leq n/2 \). Therefore, \( |A'| \leq n/2 + \epsilon n \). Note that \( A = A' \cup S \). By Claim 56 \( |A| \leq n/2 + \epsilon n + o(n) \) 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 \( \epsilon \) as claimed.

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

B.4 High degree vertices

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

Algorithm Dense:

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 \( B(N(a) \setminus S, N(b) \setminus S, E') \) 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 \( B \).
   d. Set \( A \leftarrow N(a) \cup D \) and \( B = N(b) \cup D \).
2. Output the best \( A,B \) over all pairs \( a,b \)

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 \( \epsilon \) as a function of \( \delta \) which we will fix later.
Bicovering: Covering edges with two small subsets of vertices

Definition 59. A pair \( a, b \) is good pair if \( a \in A^* \backslash S^* \), and \( b \in B^* \backslash S^* \).

Note that a good pair exists unless \( A^* \subset S^* \) or \( B^* \subset S^* \). If one of these two cases holds, \( opt = n \) and \((V, Q)\) is an optimal solution. Also recall by Claim 24, that \((a, b) \notin E\).

Lemma 60. Let \( a, b \) be a good pair. Then \( N(a) \subset A^* \) and \( N(b) \subset B^* \).

\[ N(a) = A^*, \quad N(b) = B^* \quad (2) \]

Proof. One possible vertex cover is \( S = N(a) \cap N(b) \).

Lemma 61. \( S \subset S^* \) and \( |S| \leq \epsilon \cdot n \).

Proof. As \( N(a) \subset A^* \) and \( N(b) \subset B^* \) as shown in (2), \( S = (N(a) \cap N(b)) \subset (A^* \cap B^*) = S^* \).

Thus \( |S| \leq \epsilon n \) with the latter inequality following from (1).

Lemma 62. There is a set \( S^\prime \) of size at most \( \epsilon \cdot n \) whose removal creates two sets \( A^\prime \backslash S^\prime \) and \( B^\prime \backslash S^\prime \), that share no edges.

Proof. The removal of \( S^* \) leaves no edges between \( A^* \backslash S^* \) and \( B^* \backslash S^* \). As \( N(a) \subset A^* \) and \( N(b) \subset B^* \), the removals of \( S^* \) leaves no edges between \( N(a) \backslash S^* \) and \( N(b) \backslash S^* \). The lemma follows from the bound on the size of \( S^* \).

Lemma 63. The Vertex Cover \( D \) found has size at most \( \epsilon \cdot n \).

Proof. One possible vertex cover is \( S^* \). The claim follows.

Lemma 64. The approximation ratio is \( 2 - 4\delta/3 \).

Proof. The approximation ratio is \( \max\{1/(1+2\epsilon), 2-\delta-\epsilon\} \), because the case that \( |S^*| > \epsilon \cdot n \), gives \( 1/(1+2\epsilon) \) ratio otherwise if \( |S^*| \leq \epsilon \cdot n \) then as \( |N(a)| = |N(b)| \geq \delta \cdot n \), the number of vertices in \( N(a) \backslash S^* \) and \( N(b) \backslash S^* \) is at least \( (\delta - \epsilon) \cdot n \). The above term is minimized for \( \epsilon = \delta/3 \). Thus the approximation ratio is \( 2 - 4\delta/3 \).

B.5 Expander graphs

Say that \( G \) is a vertex expander with parameter \( \delta \) for some constant \( \delta \). Let \( A^*, B^* \) be the optimum solution. By Lemma 20 we may assume that \( |A^*| \geq n/2 - \delta \cdot n/8 \).

Lemma 65. \( |N_1(A^*)| \geq \delta n/4 \)

Proof. The worst case its may be that \( |A^*| = n/2 - \delta \cdot n/8 \). See Lemma 20. As the graph is an expander, \( |N_1(A^*)| \geq \delta n/2 - \delta^2 n/8 \). By definition of an expander \( \delta \leq 2 \). Thus \( |N_1(A^*)| \geq \delta n/2 - \delta \cdot n/4 = \delta \cdot n/4 \). The above follows by replacing one of the \( \delta \) in the expression \( \delta^2 n/8 \) by \( 2 \).

Lemma 66. Returning \( V, Q \) gives ratio at most \( 2/(1+\delta^2/8) \) for Bi-Covering on expanders.
Proof. By definition \( N_1(A^*) \cap A^* = \emptyset \). Thus \( N_1(A) \subseteq S^* \cup (B^* \setminus S) \). If half of \( N_1(A^*) \) belongs to \( S \) then \( S \geq \delta \cdot n/8 \). By Claim 27 the resulting ratio is \( 2/(1 + \delta/8) = 16/(8 + \delta) \) which is a constant less than 2. Else we may assume that \( |(B^* \cap N_1(A^*)) \setminus S^*| \geq \delta \cdot n/8 \). 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 \( |D| \geq \delta \cdot n/8 \). Since the graph is an expander, in the worse case \( N_1(D) \geq \delta^2 n/8 \). Since \( N_1(D) \subseteq S^* \), \( |S^*| \geq \delta^2 \cdot n/8 \). Thus a ratio of \( 2/(1 + \delta^2/8) \) follows from Lemma 1.

B.6 Bounded degree graphs

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

Algorithm Bounded-Degree:

1. Choose a set \( A' \) of \( 3n/5d \) vertices.
2. Add \( N_1(A') \) to \( S \).
3. Let \( V' \leftarrow V \setminus (S \cup A') \).
4. Select a set \( B' \subseteq V' \) of size \( 3 \cdot n/(5 \cdot d) \).
5. Put \( N_1(B') \) in \( S \).
6. Set \( A \leftarrow A' \cup S \), \( B \leftarrow B' \cup S \).
7. Output \((A, B)\).

> Lemma 67. The ratio is at most \( 2 - 6/\left(5d\right) \).

Proof. Note that \( N_1(A') \leq 3n/5 \) because \( |A'| \leq 3n/(5d) \) and the maximum degree is \( d \). The size of \( V' \) is at least \( n - 3n/(5d) - 3n/5 = 2n/5 - 3n/(5d) \geq 3n/(5d) \). The last inequality follows as \( d \geq 3 \). 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 26, the ratio is \( 2 - 6/(5 \cdot d) \). Note that this ratio is better than \( 1 - 1/d \).

B.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 weak condition than maximum degree \( d \) hence we get a worse ratio.

> Lemma 68. The Bi-Covering problem admits a polynomial time \( 1 - 1/(\bar{d} + 1) \) ratio approximation on graphs with \( O(n) \) edges with average degree \( \bar{d} \).

Proof. Applying the Turán’s theorem [22], 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 26 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.

B.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 as follows:
1. 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).

2. Else return $V, \emptyset$

> Lemma 69. 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 26, the ratio is $2 - 2\epsilon$. If $|I| \leq 2\epsilon n$ and $|S^*| \geq \epsilon n$, by Lemma 27 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 28, we may assume without loss of generality 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 $8/5$.

B.9 Graph that contain a separator of size $o(n/\sqrt{\log n})$

> Lemma 70. 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 [11] 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.