Approximating Maximum Cut on Interval Graphs and Split Graphs Beyond Goemans-Williamson

Jungho Ahn
Inha University, Incheon, South Korea

Ian DeHaan
University of Michigan, Ann Arbor, MI, USA

Eun Jung Kim
KAIST and IBS, Daejon, South Korea
CNRS, Paris, France

Euiwoong Lee
University of Michigan, Ann Arbor, MI, USA

Abstract

We present a polynomial-time $(\alpha_{GW} + \varepsilon)$-approximation algorithm for the Maximum Cut problem on interval graphs and split graphs, where $\alpha_{GW} \approx 0.878$ is the approximation guarantee of the Goemans-Williamson algorithm and $\varepsilon > 10^{-34}$ is a fixed constant. To attain this, we give an improved analysis of a slight modification of the Goemans-Williamson algorithm for graphs in which triangles can be packed into a constant fraction of their edges. We then pair this analysis with structural results showing that both interval graphs and split graphs either have such a triangle packing or have maximum cut close to their number of edges. We also show that, subject to the Small Set Expansion Hypothesis, there exists a constant $c > 0$ such that there is no polyomial-time $(1 - c)$-approximation for Maximum Cut on split graphs.

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis

Keywords and phrases Maximum cut, graph theory, interval graphs, split graphs

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2025.20

Category APPROX

Related Version *Full Version*: `https://arxiv.org/abs/2507.10436`

Funding *Jungho Ahn*: Supported by Leverhulme Trust Research Project Grant RPG-2024-182.
Eun Jung Kim: Supported by Institute for Basic Science (IBS-R029-C1) and National Research Foundation, Korea (RS-2025-00563533).
Euiwoong Lee: Supported in part by NSF grant CCF-2236669.

1 Introduction

Given a graph $G = (V, E)$, the Maximum Cut problem asks for a subset of vertices $S \subseteq V$ that maximizes the number of edges with exactly one endpoint contained in S. In this paper, we study the approximability of the Maximum Cut problem on interval graphs and split graphs.

A graph $G = (V, E)$ is *interval* if there exists a collection of intervals on the real line $\{\mathcal{I}_v\}_{v \in V}$ such that $uv \in E$ if and only if $\mathcal{I}_u \cap \mathcal{I}_v \neq \emptyset$. See Figure 1 for an example. Interval graphs are used in the field of biology, where they model natural phenomena such as DNA and food webs [13]. They have also been used in the study of register allocation, where vertices correspond to variables and intervals correspond to "live ranges" [11]. Finally, they have numerous desirable theoretical properties and have arisen as a natural class of graphs to design algorithms for. For example, it is shown in [5] that a certain variant of the graph homomorphism problem is polynomial-time solvable if and only if the label graph is interval.

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2025).
Editors: Alina Ene and Eshan Chattopadhyay; Article No. 20; pp. 20:1–20:22
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

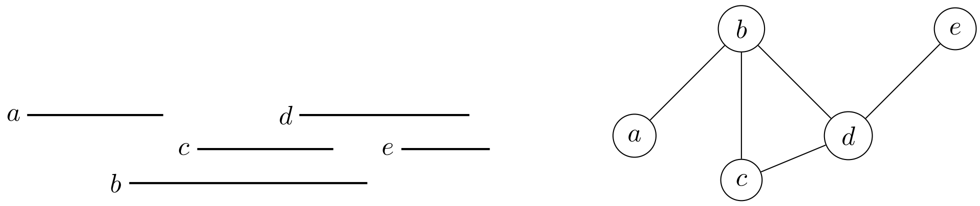

Figure 1 An interval graph. The left figure shows the interval representation. The right figure shows the resulting graph.

A graph $G = (V, E)$ is *split* if there exists a partition of $V = K \sqcup I$ such that $G[K]$, the graph induced by K, is a clique and $G[I]$, the graph induced by I, is independent. See Figure 2 for an example. Split graphs and interval graphs are both important subclasses of *chordal graphs*, which themselves are a subclass of *perfect graphs*. In particular, split graphs are often the "simplest" subclass of perfect graphs in which problems are difficult to approximate. Thus, considering split graphs is a natural first step when attempting to characterize a problem on chordal or perfect graphs. In this paper, we show that, assuming the Small Set Expansion Hypothesis, there is some constant $c > 0$ such that there is no polynomial-time $(1-c)$-approximation for Maximum Cut on split graphs. To our knowledge, this is the first known hardness of approximation result for Maximum Cut on perfect graphs.

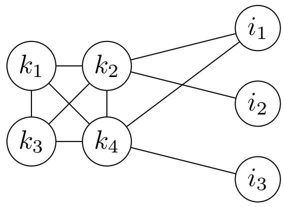

Figure 2 A split graph with $K = \{k_1, k_2, k_3, k_4\}$ and $I = \{i_1, i_2, i_3\}$.

Maximum Cut is one of the original 21 problems shown to be NP-Complete by Karp [14] and has long been a staple problem among algorithm researchers. The seminal work of Goemans and Williamson shows that there is a polynomial-time $(\alpha_{GW} \approx 0.878)$-approximation algorithm for Maximum Cut on all graphs [8]. The optimality of this result was open until 2007, when Khot et al. showed that, subject to the Unique Games Conjecture, there is no polynomial-time approximation algorithm for Maximum Cut with an approximation ratio better than α_{GW} [15].

Remarkably, there is relatively little known about Maximum Cut when the input graph is restricted to graphs from some structured class. Even for subclasses of perfect graphs, where problems such as Independent Set and Chromatic number admit polynomial-time algorithms based on semidefinite programming, Maximum Cut, another flagship application of semidefinite programming, remains mostly unexplored. In particular the extremely well-structured classes of interval graphs and split graphs, two important subclasses of perfect graphs, had no known approximation algorithm with a ratio better than α_{GW} prior to this work. Our main results provide the first improved approximation for these classes of graphs since the work of Goemans and Williamson.

▶ **Theorem 1.** *There is a polynomial-time $(\alpha_{GW} + 10^{-34})$-approximation algorithm for Maximum Cut on interval graphs.*

Table 1 Known results for Maximum Cut. Our work appears in bold. Results marked with † are subject to the Unique Games Conjecture. Results marked with ⋆ are subject to the Small Set Expansion Hypothesis under randomized reductions.

	lower bound	upper bound
General graphs	α_{GW} [8]	$\alpha_{GW}^{\dagger}$ [15]
Degree $\leq d$ graphs	$\alpha_{GW} + \Omega(\frac{1}{d^2 \log d})$ [12]	$\alpha_{GW} + \mathcal{O}(\frac{1}{\sqrt{d}})^{\dagger}$ [19]
Interval graphs	$\boldsymbol{\alpha_{GW} + 10^{-34}}$	NP-Complete [1]
Split graphs	$\boldsymbol{\alpha_{GW} + 10^{-16}}$	$\mathbf{1 - c^{\star}}$
Planar graphs	1 [10]	
Line graphs	1 [9]	

▶ **Theorem 2.** *There is a polynomial-time* $(\alpha_{GW} + 10^{-16})$*-approximation algorithm for Maximum Cut on split graphs.*

On the hardness side, it was shown by Bodlander and Jansen in 2000 that Maximum Cut on split graphs is NP-Complete [3]. Maximum Cut on interval graphs has seen further attention lately - it was shown only recently that Maximum Cut on interval graphs is NP-Complete [1]. Further work refined this hardness for interval graphs which have at most 4 interval lengths [4], and later at most 2 interval lengths [2]. Hardness for unit interval graphs - interval graphs with only 1 interval length - remains open. We remark that none of these hardness results imply any hardness of approximation. Our final main result is that, subject to the Small Set Expansion Hypothesis of [17], Maximum Cut on split graphs is hard to approximate to some constant factor.

▶ **Theorem 3.** *There exists a constant $c > 0$ such that there is no polynomial-time $(1 - c)$-approximation algorithm for Maximum Cut on split graphs, subject to the Small Set Expansion Hypothesis under randomized reductions.*

1.1 Our Techniques

1.1.1 Analysis of the Perturbed Goemans-Williamson Algorithm

Our starting point is the Goemans-Williamson algorithm [8]. This algorithm first solves the following semidefinite program (SDP)

$$\textbf{maximize} \left\{ \frac{1}{2} \sum_{uv \in E} (1 - x_u \cdot x_v) \mid x_v \in \mathbb{S}^{|V|-1} \ \forall\, v \in V \right\}.$$

That is, the algorithm maps each vertex to a unit vector in a way that maximizes the sum of $1 - x_u \cdot x_v = 1 - \cos\theta_{uv}$, where θ_{uv} is the angle between x_u and x_v. Next, the algorithm samples a random Gaussian vector $r \sim \mathcal{N}(0,1)^{|V|}$, creates the set $S = \{v \mid r \cdot x_v \geq 0\}$, and returns the cut defined by S. This step is equivalent to sampling a random hyperplane and taking S to be the vertices on one side of this hyperplane. It is a straightforward calculation to see that the probability of an edge $e \in E$ being cut by S is equal to θ_e/π. Thus, the approximation guarantee of the Goemans-Williamson algorithm is equal to

$$\alpha_{GW} := \min_{\theta \in [0,\pi]} \frac{2}{\pi} \frac{\theta}{1 - \cos\theta} \approx 0.878.$$

Define $\theta_c := \text{argmin}_{\theta \in [0,\pi]} \frac{\theta}{1-\cos\theta} \approx 134°$ to be the "critical angle" at which this ratio is minimized. We plot the performance guarantee over all angles in Figure 3. Any edge $e \in E$ with $\theta_e \neq \theta_c$ has approximation ratio strictly better than α_{GW}. Intuitively, if at least a constant fraction of edges have angle bounded away from θ_c, we should expect the Goemans-Williamson algorithm to achieve a better approximation ratio. Unfortunately, this is not exactly true.

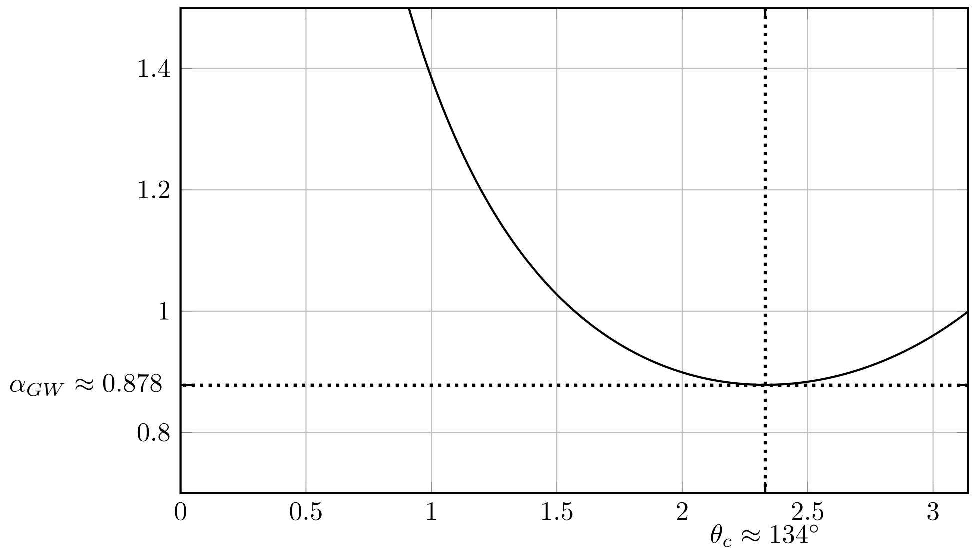

Figure 3 Plot of $\frac{2}{\pi}\frac{\theta}{1-\cos(\theta)}$ for $\theta \in (0, \pi)$.

Consider the graph P_3, the path on 3 vertices. Suppose the SDP is solved suboptimally, so that the first edge has angle θ_c and the second edge has angle 0. Even though half of the edges have angle bounded away from θ_c, the expected size of the cut from rounding is still only α_{GW} times the optimal value of the SDP. To handle situations like this, we introduce a "perturbed" version of the Goemans-Williamson algorithm. In this perturbed algorithm, any vertex $v \in V$ with $|r \cdot x_v| < \eta$ for some small η will instead be included in S with probability $\frac{1}{2}$. The motivation behind this perturbation is to consider what happens when every x_v is "moved" by a small amount. If x_u and x_v are close together, then they are likely to move further apart after this perturbation. Inversely, if x_u and x_v are far apart, they are likely to move closer together after this perturbation.

Indeed, in Lemma 17, we show that the perturbed algorithm is more likely to cut any edge $e \in E$ with $\theta_e < \pi/2$, and less likely to cut any edge with $\theta_e > \pi/2$. Moreover, the further away θ_e is from $\pi/2$, the more the results of the perturbed algorithm vary from the unperturbed algorithm. Thus, if there are many edges with angle near 0, but few edges with angle near π, the perturbed algorithm will achieve an approximation ratio above α_{GW}, see Lemma 19. On the other hand, if there are many edges with angle near π, then the Goemans-Williamson algorithm itself will achieve an approximation ratio above α_{GW}, see Lemma 15. Thus, if we take the maximum result of the perturbed and unperturbed Goemans-Williamson algorithm, we can ignore the case of having many 0 edges which contribute little to the optimal. That is, as shown in Lemma 21, we can indeed assume that if at least a constant fraction of edges have angle bounded away from θ_c, the Goemans-Williamson algorithm will achieve an approximation ratio above α_{GW}.

Consider a single triangle $T = \{uv, vw, wu\}$. A straightforward analysis shows that

$$\theta_{uv} + \theta_{vw} + \theta_{wu} \leq 360°,$$

regardless of the values of x_u, x_v, x_w. Thus, there is some edge $e \in T$ with $\theta_e \leq 120° < \theta_c$. That is, for every triangle in our graph, we can expect at least one of its edges to have an angle bounded away from θ_c. This leads us into Section 2, where we show that every interval graph and split graph either have a large edge-disjoint triangle packing, or have a cut with at least $0.9|E|$ edges.

1.1.2 Finding an Edge-Disjoint Triangle Packing

Suppose $G = (K \cup I, E)$ is a split graph. If at least $0.1|E|$ of its edges are contained in the clique K, then we can pack triangles almost perfectly into those edges. Otherwise, we have that at least $0.9|E|$ edges are crossing between K and I, and so we have found a cut with $0.9|E| \geq 0.9 \cdot \mathrm{mc}(G)$ edges, where $\mathrm{mc}(G)$ is the size of a maximum cut of G. Thus, we find ourselves in a "win-win" situation, where G either has an edge-disjoint triangle packing on a constant fraction of its edges, or G is nearly bipartite, and we can find a large cut. This analysis leads directly to an improved approximation for Maximum Cut on split graphs.

It turns out that the same "win-win" structural result also holds for interval graphs, which we prove in Theorem 5. To show this, we employ a marking scheme, where each vertex is classified as either "small" or "large" based on how it interacts with the rest of the graph. We first show in Lemma 8 that the number of edges between large vertices is low compared to $|E|$. The situation for small vertices is more complicated. We show in Lemmas 9 and 10 that one of the following conditions must always hold:

1. G has a bridge; or
2. there is a clique C in G such that the sum of degrees $\sum_{v \in C} d_v$ is not much more than $|C|^2$; or
3. the number of edges between small vertices is low compared to $|E|$.

If Condition 1 holds, we can essentially delete the bridge; because adding a bridge to a bipartite graph does not create an odd cycle, we can always add back in the bridge after finding a cut in the rest of the graph. If Condition 2 holds and $|C| \geq 3$, then we can pack edge-disjoint triangles into C and delete C from the graph. Because the sum of degrees is not much more than $|C|^2$, we have packed triangles into at least a constant fraction of the deleted edges. The case of $|C| \leq 2$ is more tricky, as one cannot pack triangles into one or two vertices, and must be handled separately. If Condition 3 holds, then most of the edges in G go between small and large vertices. If we have already packed triangles into a constant fraction of edges via Condition 2, then we are done. Otherwise, "most" edges are remaining in the graph, and we have found a cut on "most" of these remaining edges. For the right definition of "most," this is a large cut of at least $0.9|E|$ edges.

1.1.3 Hardness of Approximating Maximum Cut on Split Graphs

To show that Maximum Cut is hard to approximate on split graphs, we start from the following hardness result that follows using standard techniques from [18]. Assuming the Small Set Expansion Hypothesis holds under randomized reductions, for any sufficiently small $\varepsilon > 0$, there is no polynomial time algorithm that, given a graph $G = (V, E)$, can distinguish between the following two cases.

1. There exists a cut $S \subseteq V$ with $|S| \approx 0.5|V|$ such that $|\delta(S)| \leq \mathcal{O}(\varepsilon|E|)$.
2. For all cuts $S \subseteq V$ with $|S| \leq 0.5|V|$, either $|S| \leq 0.2|V|$ or $|\delta(S)| \geq \Omega(\sqrt{\varepsilon}|E|)$.

That is, in Case 1, there is a cut with nearly half the vertices that cuts very few edges. In Case 2, all cuts with at least a constant fraction of vertices must cut many more edges. We transform G into a split graph $G' = (K \cup I, E')$ by turning V into a clique and setting $K := V$, setting $I := \{v_e \mid e \in E\}$, and adding an edge from v_e to each endpoint of e. Any cut of G can have at most $0.25|V|^2$ edges from those edges contained in K, and at most $2|E|$ edges from those going between I and K. In Case 1, there exists a cut that does cut nearly $0.25|V|^2 + 2|E|$ edges. In Case 2, any cut of G either cuts at most $0.2|V|^2$ edges from those edges contained in K, or at most $(2 - \Omega(\sqrt{\varepsilon}))|E|$ from those edges crossing between I and K. After ensuring that $|E| \approx |V|^2$, this shows that Maximum Cut is hard to approximate on split graphs.

1.2 Preliminaries

1.2.1 Graphs

We consider only finite graphs in this paper. Apart from Section 4, all considered graphs will also be simple and unweighted. When the subject graph G is clear from context, we will use $V := V(G)$ to refer to the vertex set of G, $E := E(G)$ to refer to the edge set of G, and $n := |V|$ to refer to the number of vertices in G. Let $S \subseteq V$ be a subset of vertices of G. We define $G[S] := (S, \{uv \in E \mid u, v \in S\})$ as the subgraph of G induced by S and $E[S] := E(G[S])$ as the set of edges in this subgraph. We define $\delta_G(S) := \{uv \in E \mid u \in S, v \notin S\}$ to be the cut induced by S and $\mathrm{mc}(G) := \max_{S' \subseteq V}\{|\delta_G(S')|\}$ to be the maximum cut size of G. We will often omit G and write only $\delta(S)$ when the graph G is clear from context. Let $v \in V$ be any vertex of G. We define $\delta(v) := \delta(\{v\})$ as the set of edges adjacent to v, $N(v) := \{u \in V \mid uv \in E\}$ as the set of neighbors of v, and $d_v := |N(v)|$ as the degree of v. We define a triangle of G as a set of edges $\{uv, vw, wu\} \subseteq E$ forming a triangle. We say that G has a *triangle packing* of size t if there exist t edge-disjoint triangles $T_1, T_2, \dots, T_t \subseteq E$.

1.2.2 Gaussians

We define $\mathcal{N}(0,1)$ as the Gaussian distribution with mean 0 and variance 1. Moreover, we define $\mathcal{N}(0,1)^n$ as the n-dimensional Guassian distribution, where a sampled vector $v \sim \mathcal{N}(0,1)^n$ has $v_i \sim \mathcal{N}(0,1)$ for each $i \in [n]$ and v_i is independent of v_j for $i \neq j$. We make use of the fact that sampling a vector in this way is equivalent to sampling a random direction; that is, after normalizing, v becomes a uniformly random unit vector in n-dimensional space. In particular, this means that $\mathcal{N}(0,1)^n$ is symmetric up to rotations, which we will exploit frequently. We will also make use of the following lemma, which says that $\mathcal{N}(0,1)$ is roughly equivalent to a uniform distribution close to 0.

▶ **Lemma 4.** *For a randomly sampled $r \sim \mathcal{N}(0,1)$, we have that*

$$\frac{x}{2} \leq \mathbb{P}[|r| \leq x] \leq x$$

for all $x \in [0,1]$.

The proof of Lemma 4 follows from direct calculation and is not instructive, so we omit it.

2 Triangle Packing and Maximum Cut Tradeoff

This section is devoted to proving the following structural result.

▶ **Theorem 5.** *If $G = (V, E)$ is an interval graph, then G either has a triangle packing of size $10^{-8}|E|$ or has a cut of size at least $0.9|E|$ that can be found in polynomial time.*

2.1 Warmup: Tradeoff for Split Graphs

As a warmup, we prove the following structural result that is essentially equivalent to Theorem 5, except it is for split graphs instead of interval graphs.

▶ **Theorem 6.** *If $G = (V, E)$ is a split graph, then G either has a triangle packing of size $0.01|E|$ or has a cut of size at least $0.9|E|$ that can be found in polynomial time.*

Before we prove this, we need the following helpful lemma.

▶ **Lemma 7.** *The complete graph K_n on $n \geq 3$ vertices has an edge-disjoint triangle packing of size $\frac{|E[K_n]|}{10} = \frac{n(n-1)}{20}$.*

Edge-disjoint triangle packings in complete graphs have been studied in the literature before, with [6] giving an optimal bound. Lemma 7 is not very close to the optimal bound, but it is sufficient for our purposes and substantially easier to prove.

Proof. Label the vertices of K_n with numbers $0, 1, 2, \dots, n-1$. Label each triangle $\{uv, vw, wu\}$ with $(u + v + w)$ modulo n. Fix any edge $uv \in E(K_n)$. Then we have that, for each possible triangle label, there is only one triangle involving uv with that label. So we may take all the triangles of any specific label and find an edge-disjoint triangle packing. By the pigeon hole principle, some label has at least $\binom{n}{3}/n$ triangles. Therefore, K_n has a triangle packing of size $\binom{n}{3}/n \geq \frac{|E[K_n]|}{10}$, as wanted. ◀

With Lemma 7 in hand, we can now prove Theorem 6.

Theorem 6. Write $V = K \sqcup I$ where K and I are the clique and independent portions of G, respectively. Then we can partition $E = \delta(K) \sqcup E[K]$. If $|\delta(K)| \geq 0.9|E|$, then we are done, as we have constructed a cut of size at least $0.9|E|$ in polynomial time. Otherwise, we have that $|E[K]| \geq 0.1|E|$. Now, recalling that $G[K]$ is the complete graph on $|K|$ vertices, we use Lemma 7 to find that $G[K]$ (and thus G) has an edge-disjoint triangle packing of size at least $\frac{|E[K]|}{10} \geq 0.01|E|$. ◀

2.2 Finding a Cut

The proof of Theorem 5 maintains a similar flavor as that of Theorem 6. While the proof of Theorem 6 either finds a large cut or one large clique to pack triangles into, we need to repeatedly apply Lemma 7 during the proof of Theorem 5, as there may be no single large-enough clique. That is, at each stage, we either identify a clique that we may pack triangles into and remove while being careful to bound the number of non-clique edges we remove, or we conclude there is a large cut.

Given an interval graph $G = (V, E)$, we proceed by partitioning the vertices into "small" and "large" vertices $V = \mathcal{S} \sqcup \mathcal{L}$. At each step, we will either certify that the implied cut is large $|\delta(\mathcal{S})| \geq 0.9|E|$, or find a way to expand a triangle packing on edges in $E[\mathcal{S}]$.

For any $t \in \mathbb{R}$, define $B_t := \{v \in V \mid t \in \mathcal{I}_v\}$ to be the "bag" of vertices whose intervals occupy position t. For a vertex $v \in V$, define $\mathcal{B}_v := \min_{t \in \mathcal{I}_v} |B_t|$ to be the size of the smallest bag in v's interval. Let T be a constant we will decide the value of later. We say that v is "small" if $d_v \leq T \cdot \mathcal{B}_v$ and "large" otherwise. In other words, if at least a constant fraction of v's edges "come from" $\mathcal{B}_v$, then v is small. We define $\mathcal{S}$ as the set of small vertices and $\mathcal{L}$ as the set of large vertices.

▶ **Lemma 8.** $|E[\mathcal{L}]| \leq 8T^{-1} \cdot |E|$.

Proof. Define an ordering $\prec$ on $\mathcal{L}$ by $u \prec v$ if $\mathcal{B}_u < \mathcal{B}_v$. If $\mathcal{B}_u = \mathcal{B}_v$ we say $u \prec v$ if the leftmost point of $\mathcal{I}_u$ is to the left of the leftmost point of $\mathcal{I}_v$. For $u, v \in V$ with $\mathcal{B}_u = \mathcal{B}_v$ and equal leftmost point, we break ties arbitrarily to extend $\prec$ into a total ordering.

Fix any $v \in \mathcal{L}$. We will bound the number of edges $uv \in E[\mathcal{L}]$ with $u \prec v$ as a function of d_v. Consider any $u \in \mathcal{L} \cap N(v)$ such that $\mathcal{I}_u$ does not intersect either the leftmost point or rightmost point of $\mathcal{I}_v$. Then we must have that $\mathcal{I}_u \subseteq \mathcal{I}_v$ and thus $\mathcal{B}_u \geq \mathcal{B}_v$. Combined with the fact that the leftmost bag of $\mathcal{I}_u$ is to the right of the leftmost bag of $\mathcal{I}_v$ this implies that $v \prec u$. Thus, we need only consider $u \in \mathcal{L} \cap N(v)$ such that $\mathcal{I}_u$ intersects either the leftmost or rightmost bag of $\mathcal{I}_v$.

Let P_v be the set of $u \in \mathcal{L} \cap N(v)$ such that $\mathcal{I}_u$ contains the leftmost point of $\mathcal{I}_v$. Let $L_v \subseteq P_v$ be the first $\min\{|P_v|, \mathcal{B}_v\}$ of these neighbors sorted by increasing leftmost point. Similarly, let $R_v \subseteq P_v$ be the last $\min\{|P_v|, \mathcal{B}_v\}$ of these neighbors sorted by increasing rightmost point. Now notice that all $u \in P_v \setminus (L_v \cup R_v)$ have $\mathcal{B}_u > \mathcal{B}_v$ and thus $v \prec u$. This is because every point of $\mathcal{I}_u$ either intersects all of L_v or all of R_v, which have size at least $\mathcal{B}_v$ assuming $P_v \setminus (L_v \cup R_v)$ is non-empty. See Figure 4 for an illustration. Thus, the number of vertices $u \in P_v$ such that $u \prec v$ is at most $|L_v \cup R_v| \leq 2\mathcal{B}_v$. We can similarly bound the number of vertices $u \in \mathcal{L} \cap N(v)$ such that $\mathcal{I}_u$ contains the rightmost point of $\mathcal{I}_v$ and $u \prec v$ by $2\mathcal{B}_v$.

Putting these bounds together with the fact that $v \in \mathcal{L}$, we can bound the total number of edges $uv \in E[\mathcal{L}]$ with $u \prec v$ by $4\mathcal{B}_v \leq 4T^{-1} \cdot d_v$. Thus, we find that $|E[\mathcal{L}]| \leq \sum_{u \in \mathcal{L}} 4T^{-1} \cdot d_u \leq 8T^{-1} \cdot |E|$. ◀

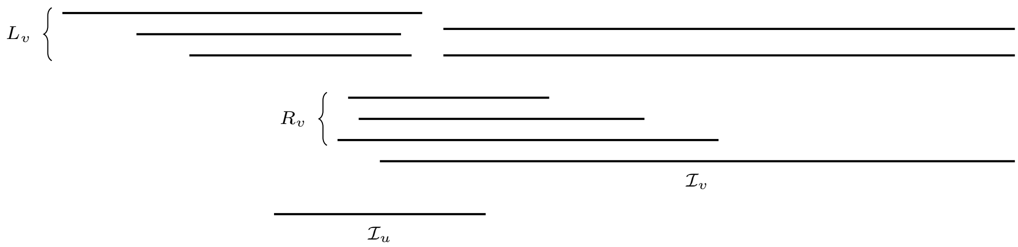

Figure 4 Representation of $\mathcal{I}_u, \mathcal{I}_v, L_v, R_v$ with $\mathcal{B}_v = 3$. Every point of $\mathcal{I}_u$ is contained in either all of L_v or all of R_v.

Unlike with large vertices, we cannot unconditionally bound $|E[\mathcal{S}]|$. We instead introduce a condition that, if unsatisfied, will allow us to make progress towards a triangle packing.

▶ **Lemma 9.** *If, for some $\varepsilon > 0$, all $t \in \mathbb{R}$ have $|B_t \cap \mathcal{S}| \leq \max\{1, \varepsilon \cdot |B_t|\}$, then $|E[\mathcal{S}]| \leq 4\varepsilon \cdot |E|$.*

Proof. Fix any non-isolated $v \in \mathcal{S}$. Let t denote the leftmost point of $\mathcal{I}_v$. Let $P_v \subseteq \mathcal{S} \cap N(v)$ denote the set of small neighbors of v whose intervals contain t. Note that $\sum_{u \in \mathcal{S}} |P_u| \geq |E[\mathcal{S}]|$. By the definition of B_t, we can bound $|P_v| = |B_t \cap \mathcal{S}| - 1 \leq \varepsilon \cdot |B_t|$. Additionally, we have that $d_v \geq |B_t| - 1 \geq |B_t|/2$ and thus $|P_v| \leq 2\varepsilon \cdot d_v$. Now we can bound $|E[\mathcal{S}]|$ by iterating over all small vertices $u \in \mathcal{S}$

$$|E[\mathcal{S}]| \leq \sum_{u \in \mathcal{S}} |P_u| \leq \sum_{u \in \mathcal{S}} 2\varepsilon \cdot d_u \leq 4\varepsilon \cdot |E|.$$ ◀

2.3 Building a Triangle Packing

▶ **Lemma 10.** *If, for some* $\varepsilon > 0$ *and* $t \in \mathbb{R}$, $|B_t \cap \mathcal{S}| \geq \max\{2, \varepsilon \cdot |B_t|\}$, *then either*

1. *we can pack at least* $\frac{\varepsilon}{30T} \cdot \sum_{u \in B_t \cap \mathcal{S}} d_u$ *edge-disjoint triangles into* $\bigcup_{u \in B_t \cap \mathcal{S}} \delta(u)$ *or*
2. G *has a bridge.*

Proof. Suppose $|B_t \cap \mathcal{S}| \geq 3$. Then by Lemma 7, we can pack at least $|B_t \cap \mathcal{S}|^2/30$ edge-disjoint triangles into $B_t \cap \mathcal{S}$. By the definition of $\mathcal{S}$, we have that

$$\sum_{u \in B_t \cap \mathcal{S}} d_u \leq \sum_{u \in B_t \cap \mathcal{S}} T \cdot \mathcal{B}_u \leq |B_t \cap \mathcal{S}| \cdot T \cdot |B_t| \leq |B_t \cap S|^2 \cdot T \cdot \varepsilon^{-1}.$$

This fulfills Condition 1.

Now suppose $|B_t \cap \mathcal{S}| = 2$. Label $B_t \cap \mathcal{S} = \{v_1, v_2\}$. If $N(v_1) \cap N(v_2) = \emptyset$, then v_1v_2 is a bridge of G and thus Condition 2 is fulfilled. Otherwise, let $u \in N(v_1) \cap N(v_2)$. We can pack a single triangle into the edges v_1v_2, uv_1, uv_2. Further, we have that $\varepsilon \cdot |B_t| \leq 2$, and so $d_{v_1} + d_{v_2} \leq \frac{4T}{\varepsilon}$. This fulfills Condition 1. ◀

This leads us to Algorithm 1. We will iteratively find a clique to pack triangles into, delete the clique, and continue. If we reach a point where we have deleted at least $0.01|E|$ edges, then we can finish as we have packed triangles into a constant fraction of the edges. Otherwise, if we run out of cliques to pack triangles into, we certify that we have an almost-complete cut on the remaining graph, which still contains most of the original edges of G. Finally, whenever we identify a bridge, we can simply delete it from the graph and add it to our final cut, as bridges can always be added to any cut.

Algorithm 1 IntervalMaxCut($G = (V, E), T, \varepsilon$).

```
G_0 ← G, 𝒯 ← ∅, A ← ∅, i ← 0
while |𝒯| ≤ 0.01|E| do
    label V(G_i) = 𝒮 ∪ ℒ as defined
    if G_i has a bridge e then
        A ← A ∪ {e}
        G_{i+1} ← (V(G_i), E(G_i) \ {e})
    else if some t ∈ ℝ has |B_t ∩ 𝒮| ≥ max{2, ε · |B_t|} then
        𝒯 ← 𝒯 ∪ ⋃_{u∈B_t∩𝒮} δ_{G_i}(u)
        G_{i+1} ← G_i[V(G_i) \ (B_t ∩ 𝒮)]
    else
        return cut δ(𝒮) ∪ A
    end if
    i ← i + 1
end while
return PerturbGW(G)
```

Note that Algorithm 1 runs in polynomial time. There are at most $|E|$ iterations of the while loop, as each iteration either returns or removes at least one edge.

▶ **Lemma 11.** *If* IntervalMaxCut(G, T, ε) *returns the set* $\delta(\mathcal{S}) \cup A$ *from within the while loop, then* $|\delta(\mathcal{S}) \cup A| \geq 0.99 \cdot (1 - 4\varepsilon - 8T^{-1}) \cdot |E|$ *and* $\delta(\mathcal{S}) \cup A$ *is a valid cut.*

Proof. Suppose that Algorithm 1 returns on the ith iteration of the while loop. By Lemmas 8 and 9, we have that $|\delta(\mathcal{S})| \geq (1 - 4\varepsilon - 8T^{-1}) \cdot |E_i|$ edges. Note that $E = E_i \cup A \cup \mathcal{T}$ and due to the while loop condition, we have that $|\mathcal{T}| \leq 0.01|E|$. Thus, $|\delta(\mathcal{S}) \cup A| \geq 0.99 \cdot (1 - 4\varepsilon - 8T^{-1}) \cdot |E|$.

To see that $\delta(\mathcal{S}) \cup A$ is a valid cut, note that adding a bridge to a bipartite graph cannot introduce an odd cycle. Thus, we can iteratively add each edge of A to $\delta(\mathcal{S})$ without invalidating our cut. ◀

▶ **Lemma 12.** *If* IntervalMaxCut(G) *exits the while loop, then G has an edge-disjoint triangle packing of size at least $\frac{\varepsilon}{3000T}|E|$.*

Proof. At the conclusion of the while loop, we have that $|\mathcal{T}| \geq 0.01 \cdot |E|$. Due to Lemma 10, each time we expand $\mathcal{T}$ by x edges, we can pack an additional $\frac{\varepsilon}{30T}x$ edge-disjoint triangles into $\mathcal{T}$. By iterating this process, there exists an edge-disjoint triangle packing of size at least $\frac{\varepsilon}{30T}|\mathcal{T}| \geq \frac{\varepsilon}{3000T}|E|$ in G. ◀

Now set $\varepsilon = 0.01$ and $T = 200$. Lemmas 11 and 12 imply that either G has a cut of size at least $0.99 \cdot 0.92|E| > 0.9|E|$ or an edge-disjoint triangle packing of size at least $\frac{1}{6} \cdot 10^{-7}|E| > 10^{-8}|E|$, as wanted.

2.4 No Tradeoff for Chordal Graphs

In this subsection, we will show that there is no equivalent tradeoff for chordal graphs as there are for interval graphs and split graphs. Thus, improving upon α_{GW} for chordal graphs will likely require new algorithmic insights.

▶ **Theorem 13.** *For all $c > 0$, there exists a chordal graph $G = (V, E)$ such that $mc(G) < \alpha_{GW}|E|$ and G has no edge-disjoint triangle packing of size $c|E|$.*

Proof. Suppose we are given a chordal graph $G' = (V', E')$ with $|V'| < c|E'|$ and no triangle packing of size $c|E'|$. We will show later how to obtain such a G'. Construct $G = (V, E)$ by setting $V := V' \cup \{w\} \cup \{x_e, y_e \mid e \in E'\}$ and $E := E' \cup \{x_e u, y_e v \mid e = uv \in E'\} \cup \{wv \mid v \in V \setminus \{w\}\}$. That is, for each edge e, we attach one fresh vertex to each endpoint of the edge, and create a "universal" vertex w connected to all vertices in the graph.

We first note that G is chordal, as it is obtained from a chordal graph G' by iteratively adding simplicial vertices. By inspection, any triangle $T \not\subseteq E'$ must contain at least one edge from $\{wv \mid v \in V'\}$. Thus, the maximum number of edge-disjoint triangles in G is at most $|V'| + c|E'| < 2c|E'| < c|E|$. For each edge $e = uv \in E'$, G contains a 5-cycle $wx_e, x_e u, uv, vy_e, y_e w$, and so G contains $|E'|$ edge-disjoint 5-cycles. Thus, we have that $\mathrm{mc}(G) \leq |E| - |E'|$. Note that $|E| = 5|E'| + |V'| \leq 6|E'|$, so $\mathrm{mc}(G) \leq \frac{5}{6}|E| < \alpha_{GW}|E|$.

It remains to show that a chordal graph G' with the desired properties exists. We give the following interval graph construction for G'. Select k large enough. Create $2^k - 1$ vertices in a "segment-tree" pattern as follows. In the first layer, create one vertex with interval $(0, 1)$. In the second layer, create two vertices with intervals $(0, 0.5)$ and $(0.5, 1)$ respectively. In the third layer, create four vertices with intervals $(0, 0.25), (0.25, 0.5), (0.5, 0.75)$ and $(0.75, 1)$ respectively. Then iterate this process for k total layers, see Figure 5 for an illustration.

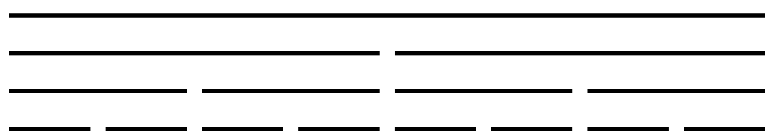

Figure 5 The interval representation of G' for $k = 4$.

We have that $|V'| = 2^k - 1$ and $|E'| \geq 2^{k-1} \cdot (k - 1) > |V'| \cdot \frac{\log_2 |V'|}{2}$ by counting only the edges adjacent to the bottom layer. Thus, for k sufficiently large, we have that $|V'| < c|E'|$.

To show that G' has no edge-disjoint triangle packing of size $c|E'|$, it is sufficient to show that $\mathrm{mc}(G') > (1 - \frac{c}{3}) \cdot |E'|$, as each triangle results in at least one un-cuttable edge. Consider the cut of G' obtained by taking the bottom t layers as one side of the cut. The number of edges in this cut is $(2^k - 2^{k-t}) \cdot (k - t)$. Note also that $|E'| \leq 2^k \cdot k$. Therefore, we have that $\mathrm{mc}(G') \geq \frac{(2^k - 2^{k-t}) \cdot (k-t)}{2^k \cdot k} |E'|$. As k tends to infinity, the term $\frac{(2^k - 2^{k-t}) \cdot (k-t)}{2^k \cdot k}$ tends to $1 - 2^{-t}$. So, by setting t to be a sufficiently large constant based on c and letting k be sufficiently large based on t, we have that $\mathrm{mc}(G') > (1 - \frac{c}{3}) \cdot |E'|$ and so G' has no edge-disjoint triangle packing of size $c|E'|$. Thus, G' fulfills all the conditions we needed to produce G. ◀

3 Analysis of the Perturbed Goemans-Williamson Algorithm

Our second main contribution is an improved approximation for Maximum Cut in graphs with large triangle packings. Given a fixed triangle $\{uv, vw, wu\}$, it is impossible for all angles $\theta_{uv}, \theta_{vw}, \theta_{wu}$ to be equal or very close to the critical angle θ_c. However, it is possible for $\theta_{uv} = \theta_{vw} = \theta_c$ to achieve the critical angle and have $\theta_{wu} = 0$. In this case, despite the entire graph being a single triangle, the Goemans-Williamson rounding algorithm will not perform beyond its worst-case guarantee. This is because the contribution of the edge wu to the objective function is 0, and so rounding it "better" does not actually increase our expected value.

To grapple with this issue, we introduce the "Perturbed Goemans-Williamson Algorithm." Intuitively speaking, this algorithm randomly "perturbs" each vector slightly. We will see that edges with near-zero angle stand to gain much more in this perturbation process than any other edges have to lose besides those with an angle of nearly π. Thus, any SDP solution with many near-zero angle edges and few near-π angle edges can be rounded with a guarantee better than α_{GW}.

Before presenting the algorithm, we must first define the semidefinite program from which we will round a solution.

$$\textbf{maximize:} \quad \frac{1}{2} \sum_{uv \in E} (1 - x_u \cdot x_v) \qquad \textbf{(SDP-GW)}$$

$$\textbf{subject to:} \quad x_v \in \mathbb{S}^{n-1} \quad \forall\, v \in V$$

Algorithm 2 PerturbGW($G = (V, E), \eta$).

Solve **SDP-GW** and obtain optimal solution $\{x_v^*\}_{v \in V}$.
Sample a random n-dimensional vector $r \sim \mathcal{N}(0, 1)^n$.
for all $v \in V$ **do**
  $s_v \leftarrow \mathrm{sign}(r \cdot x_v^*)$.
  if $|r \cdot x_v^*| \geq \eta$ **then**
    $s'_v \leftarrow s_v$.
  else
    Uniformly sample $s'_v \sim \{-1, 1\}$.
  end if
end for
$S \leftarrow \{v \mid s_v = 1\}, S' \leftarrow \{v \mid s'_v = 1\}$.
Return $\mathrm{argmax}_{\{\delta(S), \delta(S')\}}(|\delta(S)|, |\delta(S')|)$.

Note that Algorithm 2 runs in polynomial time, as solving SDPs and sampling from a Gaussian distribution can be made to run in polynomial time.

▶ **Theorem 14.** *If G has an edge-disjoint triangle packing of size at least $t|E|$, then*

$$\left|\textsc{Perturb}GW(G,\eta := \frac{t^2}{10^4})\right| \geq (\alpha_{GW} + 10^{-10}t^3)\cdot mc(G).$$

Theorem 14, when combined with Theorem 5 and Theorem 6, immediately shows that there is a polynomial-time $(\alpha_{GW} + 10^{-34})$-approximation for Maximum Cut on interval graphs and a polynomial-time $(\alpha_{GW} + 10^{-16})$-approximation for Maximum Cut on split graphs.

Note that $\delta(S)$ is the result of running the original Goemans-Williamson algorithm, so the result of Algorithm 2 is immediately at least $\mathbb{E}[|\delta(S)|] \geq \alpha_{GW}\cdot \mathrm{mc}(G)$.

For an edge $e \in E$, let $C_e := \mathbb{I}[e \in \delta(S)]$ and $C'_e := \mathbb{I}[e \in \delta(S')]$ be the random variables indicating that e is cut by S and S', respectively. Define $\theta_e := \arccos(x^*_u \cdot x^*_v)$ as the angle between u and v. For $\theta \in [0,\pi]$, let

$$E_\theta := \{e \in E \mid |\theta_e - \theta| \leq \sqrt{\eta}\}$$

be the set of edges with angle "close to" θ, where η is a parameter of Algorithm 2. We first deal with the case where there are many edges in E_π.

▶ **Lemma 15.** *If $\eta \leq 0.01$ and $|E_\pi| \geq \eta^{3/2}|E|$, then $\mathbb{E}[|\delta(S)|] \geq (\alpha_{GW} + 10^{-2}\eta^{3/2})\cdot mc(G)$.*

Proof. Let SDP^* equal the value of **SDP-GW** at optimal solution $\{x^*_v\}_{v\in V}$. Consider any $e = uv \in E$. Recall that the contribution of e to SDP^* is $\frac{1-x^*_u\cdot x^*_v}{2} = \frac{1-\cos\theta_e}{2}$. Also, by simple calculation, we have that $\mathbb{P}[e \in \delta(S)] = \frac{\theta_e}{\pi}$. We calculate

$$\begin{aligned}
\mathbb{E}[|\delta(S)|] &= \sum_{e\in E}\frac{\theta_e}{\pi}\\
&= \sum_{e\in E}\frac{1-\cos\theta_e}{2}\cdot\frac{2\theta_e}{\pi(1-\cos\theta_e)}\\
&= \sum_{e\in E\setminus E_\pi}\frac{1-\cos\theta_e}{2}\cdot\frac{2\theta_e}{\pi(1-\cos\theta_e)} + \sum_{e\in E_\pi}\frac{1-\cos\theta_e}{2}\cdot\frac{2\theta_e}{\pi(1-\cos\theta_e)}\\
&\geq \alpha_{GW}\sum_{e\in E\setminus E_\pi}\frac{1-\cos\theta_e}{2} + (\alpha_{GW}+10^{-2})\sum_{e\in E_\pi}\frac{1-\cos\theta_e}{2}\\
&= \alpha_{GW}\cdot SDP^* + 10^{-2}\sum_{e\in E_\pi}\frac{1-\cos\theta_e}{2}\\
&\geq \alpha_{GW}\cdot SDP^* + 10^{-2}\cdot\frac{|E_\pi|}{|E|}\cdot SDP^*.
\end{aligned}$$

The first inequality is by calculating $\frac{2}{\pi}\frac{\theta_e}{1-\cos\theta_e} \geq \alpha_{GW} + 10^{-2}$ for $\theta_e > \pi - \sqrt{0.01}$. The second inequality follows from the fact that $1-\cos\theta_e$ is increasing on $[0,\pi]$. The lemma now follows from the fact that $SDP^* \geq \mathrm{mc}(G)$. ◀

Now we show that for edges not in E_π, S' is not much worse than S. For technical reasons, our lemma statement also excludes edges in E_0. We will see later that for all edges with angle at most $\frac{\pi}{2}$, S' is no worse than S. Additionally, for edges with angle very close to 0, S' is substantially better than S.

▶ **Lemma 16.** *For all $e \notin E_0 \cup E_\pi$, we have that $\mathbb{P}[C'_e] \geq \mathbb{P}[C_e] - 10\eta^{3/2}$.*

Proof. For any vertex $v \in V$, let $R_v := \mathbb{I}[|r \cdot x_v^*| < \eta]$ be the random variable indicating that s'_v is randomly selected. Fix any $e = uv \notin E_0 \cup E_\pi$. We first note that $\mathbb{P}[C'_e \mid \neg R_u \wedge \neg R_v] = \mathbb{P}[C_e \mid \neg R_u \wedge \neg R_v]$ and $\mathbb{P}[C'_e \mid R_u \vee R_v] = \frac{1}{2}$.

We now turn our attention to $\mathbb{P}[C_e \mid R_v]$. Note that x_u^* and x_v^* lie on a single plane through the origin, so by symmetry of $\mathbb{S}^{n-1}$, we can assume without loss of generality that $x_v^* = (1, 0, \ldots, 0)$ and $x_u^* = (x, y, 0, \ldots, 0)$ for $y \geq 0$. Label the random variable $r = (r_1, r_2, \ldots, r_n)$. By symmetry of output, we can assume without loss of generality that $r_1 \geq 0$. Note that these assumptions imply that R_v is independent of the value of r_2.

Suppose that $|r_2| > \frac{\eta}{y}$. Then we have that $|r_2 \cdot y| > \eta > |r_1 \cdot x|$, and so $\text{sign}(r \cdot x_u^*) = \text{sign}(r_2)$. That is, the sign of $r \cdot x_u^*$ is entirely determined by the sign of r_2. This implies that, when $|r_2| > \frac{\eta}{y}$, $C_e = \mathbb{I}[\text{sign}(r_2) < 0]$. Thus, by independence of R_v and the value of r_2,

$$\mathbb{P}[C_e \mid R_v \wedge |r_2| > \frac{\eta}{y}] = \mathbb{P}[\text{sign}(r_2) < 0 \mid R_v \wedge |r_2| > \frac{\eta}{y}] = \mathbb{P}[\text{sign}(r_2) < 0] = \frac{1}{2}.$$

We then apply Lemma 4 to bound

$$\mathbb{P}[C_e \mid R_v] \leq \frac{1}{2} + \mathbb{P}[|r_2| \leq \frac{\eta}{y} \mid R_v] \leq \frac{1}{2} + \frac{\eta}{y}.$$

Let $M := \max\{\mathbb{P}[C_e \mid R_u], \mathbb{P}[C_e \mid R_v]\} \leq \frac{1}{2} + \frac{\eta}{y}$. We wish to bound $\mathbb{P}[C_e \mid R_u \vee R_v] \leq M + \frac{4\eta}{y}$. To this end, we manipulate probabilities

$$\begin{aligned}
\mathbb{P}[C_e \wedge (R_u \vee R_v)] &\leq \mathbb{P}[C_e \wedge R_u] + \mathbb{P}[C_e \wedge R_v] \\
&= \mathbb{P}[C_e \mid R_u]\mathbb{P}[R_u] + \mathbb{P}[C_e \mid R_v]\mathbb{P}[R_v] \\
&\leq M(\mathbb{P}[R_u] + \mathbb{P}[R_v]) \\
&= M(\mathbb{P}[R_u \vee R_v] + \mathbb{P}[R_u \wedge R_v]).
\end{aligned}$$

Now we find

$$\mathbb{P}[C_e \mid R_u \vee R_v] = \frac{\mathbb{P}[C_e \wedge (R_u \vee R_v)]}{\mathbb{P}[R_u \vee R_v]} \leq M + \frac{\mathbb{P}[R_u \wedge R_v]}{\mathbb{P}[R_u \vee R_v]}.$$

To bound the numerator of the error term, first consider $\mathbb{P}[R_u \mid R_v]$. Recall that $R_u = \mathbb{I}[|r_1 x + r_2 y| < \eta]$ and $R_v = \mathbb{I}[|r_1| < \eta]$. Thus, assuming R_v holds, in order for R_u to hold, we must have that $|r_2 y| < 2\eta$. Using Lemma 4, we can bound

$$\mathbb{P}[R_u \mid R_v] \leq \mathbb{P}[|r_2 y| \leq 2\eta \mid R_v] = \mathbb{P}[|r_2| \leq \frac{2\eta}{y}] \leq \frac{2\eta}{y}.$$

and using Lemma 4 again,

$$\mathbb{P}[R_u \wedge R_v] = \mathbb{P}[R_u \mid R_v]\mathbb{P}[R_v] \leq \frac{2\eta^2}{y}.$$

Applying this with yet another application of Lemma 4 gives

$$\frac{\mathbb{P}[R_u \wedge R_v]}{\mathbb{P}[R_u \vee R_v]} \leq \frac{4\eta}{y}.$$

This yields the desired result, that $\mathbb{P}[C_e \mid R_u \vee R_v] \leq \frac{1}{2} + \frac{5\eta}{y}$.

Putting it all together, we find that

$$\begin{aligned}\mathbb{P}[C_e] &= \mathbb{P}[C_e \mid \neg R_u \wedge \neg R_v] \cdot \mathbb{P}[\neg R_u \wedge \neg R_v] + \mathbb{P}[C_e \mid R_u \vee R_v] \cdot \mathbb{P}[R_u \vee R_v] \\ &\leq \mathbb{P}[C_e \mid \neg R_u \wedge \neg R_v] \cdot \mathbb{P}[\neg R_u \wedge \neg R_v] + (\frac{1}{2} + \frac{5\eta}{y}) \cdot \mathbb{P}[R_u \vee R_v].\end{aligned}$$

and using Lemma 4,

$$\begin{aligned}\mathbb{P}[C'_e] &= \mathbb{P}[C'_e \mid \neg R_u \wedge \neg R_v] \cdot \mathbb{P}[\neg R_u \wedge \neg R_v] + \mathbb{P}[C'_e \mid R_u \vee R_v] \cdot \mathbb{P}[R_u \vee R_v] \\ &= \mathbb{P}[C_e \mid \neg R_u \wedge \neg R_v] \cdot \mathbb{P}[\neg R_u \wedge \neg R_v] + \frac{1}{2} \cdot \mathbb{P}[R_u \vee R_v] \\ &\geq \mathbb{P}[C_e] - \frac{5\eta}{y} \cdot \mathbb{P}[R_u \vee R_v] \\ &\geq \mathbb{P}[C_e] - \frac{10\eta^2}{y}.\end{aligned}$$

Noting that $y = \sin\theta_e \geq \sqrt{\eta}$ because $e \notin E_0 \cup E_\pi$ completes the proof of the lemma. ◀

Lemma 16 allows us to bound the perturbation loss on all angles sufficiently bounded away from 0 and π. However, we will not be able to guarantee that the angles we consider are sufficiently bounded away from 0 to properly utilize Lemma 16. Thus, we strengthen Lemma 16 to show that perturbation does not cause *any* loss on angles below $\frac{\pi}{2}$.

▶ **Lemma 17.** *For all $e \in E$ with $\theta_e \leq \frac{\pi}{2}$, we have that $\mathbb{P}[C'_e] \geq \mathbb{P}[C_e]$.*

Proof. Take any $e = uv \in E$ such that $\theta_e \leq \frac{\pi}{2}$. As in the proof of Lemma 16, restrict to two dimensions, rotate, and reflect so we may assume $x_v^* = (1, 0, \ldots, 0), x_u^* = (x, y, 0, \ldots, 0)$ for $y \geq 0$, and $r_1 \geq 0$. Due to our assumption on θ_e, we have that $x = \cos\theta_e \geq 0$. As earlier, define the event $R_w := \mathbb{I}[|r \cdot x_w^*| < \eta]$ for $w \in V$. As in the proof of Lemma 16, our main task is to bound $\mathbb{P}[C_e \mid R_v]$. However this time, we will show $\mathbb{P}[C_e \mid R_v] \leq \frac{1}{2}$.

The event R_v is equivalent to $\mathbb{I}[|r_1| < \eta]$. Since $x, y \geq 0$, if C_e happens, then we must have $r_2 \leq 0$, so

$$\mathbb{P}[C_e|R_v] = \mathbb{P}[r_1 x + r_2 y < 0 \mid r_1 \in [0, \eta)] \leq \mathbb{P}[r_2 \leq 0] = \frac{1}{2}.$$

We have that

$$R_u = \mathbb{I}[|r_1 x + r_2 y| < \eta] = \mathbb{P}[r_2 y \in (-\eta - r_1 x, \eta - r_1 x)].$$

Recall that $x \geq 0$ and $r_1 \geq 0$, so $|-\eta - r_1 x| > |\eta - r_1 x|$. Thus, the event R_u contains a larger portion of the negative space than the positive space and

$$\mathbb{P}[r_2 y < 0 \mid \neg R_u] \leq \frac{1}{2}.$$

Recalling that $y > 0$, and r_1 and r_2 are independent, we can now calculate

$$\begin{aligned}\mathbb{P}[C_e \mid R_v \wedge \neg R_u] &= \mathbb{P}[r_1 x + r_2 y < 0 \mid r_1 \in [0, \eta) \wedge \neg R_u] \\ &\leq \mathbb{P}[r_2 < 0 \mid r_1 \in [0, \eta) \wedge \neg R_u] \\ &\leq \mathbb{P}[r_2 < 0 \mid \neg R_u] \\ &\leq \frac{1}{2}.\end{aligned}$$

and by symmetry, $\mathbb{P}[C_e \mid R_v \wedge \neg R_u] \leq \frac{1}{2}$. A similar computation as that in the proof of Lemma 16 completes the proof. ◀

▶ **Lemma 18.** *For all $e \in E$, we have that $\mathbb{P}[C'_e] \geq \frac{\eta}{4}$.*

Proof. Take any $e = uv \in E$ and calculate, using Lemma 4,

$$\mathbb{P}[C'_e] \geq \mathbb{P}[C'_e \mid R_v] \cdot \mathbb{P}[R_v] \geq \frac{1}{2} \cdot \frac{\eta}{2}.$$ ◀

Now notice that if $\theta_e < \frac{\pi}{4}\eta$, then $\mathbb{P}[C'_e] \geq \frac{\eta}{4} > \frac{\theta_e}{\pi} = \mathbb{P}[C_e]$. Thus, if a substantial fraction of E has very small angle, then $\delta(S')$ will be noticeably larger than $\delta(S)$. To this end, define

$$E_z := \{e \in E : \theta_e \leq \frac{\pi}{8}\eta\}.$$

Note that for each $e \in E_z$, we have $\mathbb{P}[C_e] \leq \frac{\pi}{8}\eta \cdot \frac{1}{\pi} = \frac{\eta}{8} \leq 2\mathbb{P}[C'_e]$.

▶ **Lemma 19.** *If $|E_z| \geq 96\sqrt{\eta}|E|$ and $|E_\pi| < \eta^{3/2}|E|$, then*

$$\mathbb{E}[|\delta(S')|] \geq \mathbb{E}[|\delta(S)|] + \eta^{3/2}|E| \geq (\alpha_{GW} + \eta^{3/2}) \cdot mc(G).$$

Proof. Calculate, using Lemmas 16, 17, and 18

$$\begin{aligned}
\mathbb{E}[|\delta(S')|] &= \sum_{e \in E_z} \mathbb{P}[C'_e] + \sum_{e \in E_\pi} \mathbb{P}[C'_e] + \sum_{e \in E \setminus (E_\pi \cup E_z)} \mathbb{P}[C'_e] \\
&\geq \sum_{e \in E_z} \frac{\eta}{4} + \sum_{e \in E \setminus (E_\pi \cup E_z)} \mathbb{P}[C_e] - 10\eta^{3/2} \\
&\geq \sum_{e \in E_z} (\mathbb{P}[C_e] + \frac{\eta}{8}) + \sum_{e \in E \setminus (E_\pi \cup E_z)} \mathbb{P}[C_e] - 10\eta^{3/2} \\
&\geq \mathbb{E}[|\delta(S)|] + \frac{\eta}{8}|E_z| - 10\eta^{3/2}|E| - |E_\pi| \\
&\geq \mathbb{E}[|\delta(S)|] + \eta^{3/2}|E|.
\end{aligned}$$ ◀

We have shown that if either E_z or E_π contain a constant fraction of E, then we can improve upon Goemans-Williamson, so we may assume from now on that both of these sets have negligible size. Now we will utilize the fact that G has an edge-disjoint triangle packing of size $t \cdot |E|$ to show that there are many edges not near the critical angle θ_c. Define the set

$$E' = \{e \in E \mid \theta_e \leq \frac{2\pi}{3}\} \setminus E_z.$$

▶ **Lemma 20.** $|E'| \geq t \cdot |E| - |E_z|$.

Proof. Fix any triangle $T = \{uv, vw, wu\}$. We can bound the sum of angles $\theta_{uv} + \theta_{vw} + \theta_{uw} \leq 2\pi$, and so at least one $e \in T$ has $\theta_e \leq \frac{2\pi}{3}$. Thus, in any triangle T, we have that $T \cap (E' \cup E_z)$ is non-empty. Because G has $t \cdot |E|$ edge-disjoint triangles, there are at least $t \cdot |E|$ edges in $(E' \cup E_z)$. Subtracting out those edges in E_z completes the proof. ◀

We note that $\min_{\theta \in [0, \frac{2\pi}{3}]} \frac{2}{\pi} \frac{\theta}{1 - \cos\theta} = \frac{8}{9} > \alpha_{GW} + 0.01$, which motivates the following lemma.

▶ **Lemma 21.** *If $t \geq 97\sqrt{\eta}$ and $|E_z| \leq 96\sqrt{\eta}|E|$, then*

$$\mathbb{E}[|\delta(S)|] \geq (\alpha_{GW} + 10^{-4}\eta^{3/2}) \cdot mc(G).$$

Proof. The assumptions on t and $|E_z|$ imply, through Lemma 20, that $|E'| \geq \sqrt{\eta}|E|$. We calculate

$$\begin{aligned}
\mathbb{E}[|\delta(S)|] &= \sum_{e\in E'} \frac{\theta_e}{\pi} + \sum_{e\in E\setminus E'} \frac{\theta_e}{\pi} \\
&\geq \frac{8}{9} \sum_{e\in E'} \frac{1-\cos\theta_e}{2} + \alpha_{GW} \sum_{e\in E\setminus E'} \frac{1-\cos\theta_e}{2} \\
&\geq \alpha_{GW} \cdot SDP^* + 0.01 \sum_{e\in E'} \frac{1-\cos\theta_e}{2} \\
&\geq \alpha_{GW} \cdot SDP^* + 10^{-3} \sum_{e\in E'} \frac{\theta_e}{\pi} \\
&\geq \alpha_{GW} \cdot SDP^* + 10^{-4} \cdot \eta \cdot |E'| \\
&\geq \alpha_{GW} \cdot SDP^* + 10^{-4} \cdot \eta^{3/2} \cdot |E| \\
&\geq (\alpha_{GW} + 10^{-4} \cdot \eta^{3/2}) \cdot \mathrm{mc}(G).
\end{aligned}$$ ◀

Now we can recall and prove Theorem 14.

▶ **Theorem 14.** *If G has an edge-disjoint triangle packing of size at least $t|E|$, then*

$$\left|\textsc{PerturbGW}(G, \eta := \frac{t^2}{10^4})\right| \geq (\alpha_{GW} + 10^{-10}t^3) \cdot mc(G).$$

Proof. Note that the definition of η implies $\eta < 0.01$ and $t > 97\sqrt{\eta}$. If $|E_\pi| \geq \eta^{3/2}|E|$, then apply Lemma 15 to find that

$$\mathbb{E}[|\delta(S)|] \geq (\alpha_{GW} + 10^{-2}\eta^{3/2}) \cdot \mathrm{mc}(G) = (\alpha_{GW} + 10^{-8}t^3) \cdot \mathrm{mc}(G).$$

If $|E_\pi| < \eta^{3/2}|E|$ and $|E_z| \geq 96\sqrt{\eta}|E|$, then apply Lemma 19 to find that

$$\mathbb{E}[|\delta(S')|] \geq (\alpha_{GW} + \eta^{3/2}) \cdot \mathrm{mc}(G) = (\alpha_{GW} + 10^{-6}t^3) \cdot \mathrm{mc}(G).$$

Finally, if $|E_\pi| < \eta^{3/2}|E|$ and $|E_z| < 96\sqrt{\eta}|E|$, then apply Lemma 21 to find that

$$\mathbb{E}[|\delta(S')|] \geq (\alpha_{GW} + 10^{-4}\eta^{3/2}) \cdot \mathrm{mc}(G) = (\alpha_{GW} + 10^{-10}t^3) \cdot \mathrm{mc}(G).$$ ◀

4 Hardness of Approximating Maximum Cut on Split Graphs

In this section, we show that, subject to the Small Set Expansion Hypothesis, Maximum Cut on split graphs is hard to approximate to some constant. Our starting point is hardness of finding a small balanced cut on weighted graphs. Given a graph $G = (V, E)$ with weights $w_e \in [0, 1]$ for $e \in E$, we define

$$\mu(S) := \frac{\sum_{u\in S} d_u}{\sum_{v\in V} d_v}$$

as the *normalized set size* of $S \subseteq V$. Here, $d_u := \sum_{e\in\delta(u)} w_e$ is defined as the *weighted* degree of u. In an unweighted regular graph, we have that $\mu(S) = \frac{|S|}{|V|}$. Also define

$$\Phi(S) := \frac{\sum_{e\in\delta(S)} w_e}{\sum_{u\in S} d_u}$$

as the *edge expansion* of $S \subseteq V$. In an unweighted d-regular graph, we have that $\Phi(S) = \frac{|\delta(S)|}{d|S|}$.

▶ **Lemma 22** (Corollary 3.6 of [18]). *There is a constant $c_1 > 0$ such that for any sufficiently small $\epsilon > 0$, it is SSE-hard to distinguish between the following two cases for a weighted graph $G = (V, E)$:*

- ***Yes:*** *There exists $S \subseteq V$ such that $\mu(S) = \frac{1}{2}$ and $\Phi(S) \leq 2\epsilon$.*
- ***No:*** *Every $S \subseteq V$ with $\mu(S) \in [\frac{1}{10}, \frac{1}{2}]$ satisfies $\Phi(S) \geq c_1\sqrt{\epsilon}$.*

To utilize this hardness, we first create an unweighted instance as follows.

▶ **Lemma 23.** *There is a constant $c_2 > 0$ such that for any sufficiently small $\epsilon, \eta > 0$, it is SSE-hard to distinguish between the following two cases for an unweighted, non-simple graph $G = (V, E)$ with $|E| \leq |V|^5$:*

- ***Yes:*** *There exists $S \subseteq V$ such that $\mu(S) \in [\frac{1}{2} - \eta, \frac{1}{2} + \eta]$ and $\Phi(S) \leq 3\epsilon$.*
- ***No:*** *Every $S \subseteq V$ with $\mu(S) \in [\frac{1}{10} + \eta, \frac{1}{2}]$ satisfies $\Phi(S) \geq c_2\sqrt{\epsilon}$.*

Proof. We begin with a gap instance $G = (V, E)$ of the form in Lemma 22. We will create a weighted instance G' with the same vertex and edge set such that each edge has weight in $\{1, 2, \ldots, n^3\}$. Duplicating edges according to their weight will then yield the desired statement for unweighted graphs.

First, we may assume without loss of generality that the edges in G are scaled such that the sum of degrees is $\sum_{v \in V} d_v = n^3$. This is because multiplying the weight of all edges by an equal constant factor does not change the results of μ or Φ. Now produce the weights $\{w'_e\}_{e \in E}$ by setting $w'_e := \lfloor w_e \rfloor$. Note in particular that this implies $w'_e \leq |V|^3$ for all $e \in E$. Define μ', Φ', and d'_v analogously to μ, Φ, and d_v, except for weights $\{w'_e\}_{e \in E}$. Note that $w_e \geq w'_e > w_e - 1$ for all $e \in E$ and so $d_v \geq d'_v > d_v - n$ for all $v \in V$.

Suppose that G is in the **Yes** case of Lemma 22, and let $S \subseteq V$ have $\mu(S) = \frac{1}{2}$ and $\Phi(S) \leq 2\epsilon$. Then we calculate

$$\mu'(S) = \frac{\sum_{u \in S} d'_u}{\sum_{v \in V} d'_v} \geq \frac{\sum_{u \in S} d_u - n|S|}{\sum_{v \in V} d_v} = \mu(S) - \frac{n|S|}{n^3} \geq \frac{1}{2} - \frac{1}{n}.$$

Thus, for sufficiently large n, we have that $\mu'(S) \geq \frac{1}{2} - \eta$. By a similar calculation, we have that $\mu'(V \setminus S) \geq \mu(V \setminus S) - \eta = \frac{1}{2} - \eta$ and so $\mu'(S) = 1 - \mu'(V \setminus S) \leq \frac{1}{2} + \eta$. We now aim to show that $\Phi'(S) \leq 3\epsilon$. We calculate

$$\begin{aligned}\Phi'(S) &= \frac{\sum_{e \in \delta(S)} w'_e}{\sum_{u \in S} d'_u} \leq \frac{\sum_{e \in \delta(S)} w_e}{\sum_{u \in S} d_u - n|S|} \leq \frac{\sum_{e \in \delta(S)} w_e}{(1 - 2/n)\sum_{u \in S} d_u} \\ &= \frac{1}{1 - 2/n}\Phi(S) \leq \frac{1}{1 - 2/n} 2\epsilon.\end{aligned}$$

Here, the second inequality uses the fact that $\mu(S) = \frac{1}{2}$ and so $\sum_{u \in S} d_u = \frac{n^3}{2}$. Thus, for sufficiently large n, we have that $\Phi'(S) \leq 3\epsilon$. This establishes that the **Yes** case of G maps to the **Yes** case of G'.

Suppose that G is in the **No** case of Lemma 22. Then any $S \subseteq V$ with $\mu(S) \leq \frac{1}{2}$ either has $\mu(S) < \frac{1}{10}$ or $\Phi(S) \geq c_1\sqrt{\epsilon}$. Take any $S \subseteq V$. If $\mu(S) < \frac{1}{10}$, then $\mu'(S) \leq \mu(S) + \eta < \frac{1}{10} + \eta$. So consider the case in which $\mu(S) \geq \frac{1}{10}$. Then we calculate

$$\begin{aligned}\Phi'(S) &= \frac{\sum_{e \in \delta(S)} w'_e}{\sum_{u \in S} d'_u} \geq \frac{\sum_{e \in \delta(S)} w_e - n^2}{\sum_{u \in S} d_u} = \Phi(S) - \frac{n^2}{\sum_{u \in S} d_u} \\ &\geq \Phi(S) - \frac{10}{n} \geq c_1\sqrt{\epsilon} - \frac{10}{n}.\end{aligned}$$

If we set $c_2 = \frac{c_1}{2}$, then for sufficiently large n, $\Phi'(S) \geq c_2\sqrt{\epsilon}$. This establishes that the **No** case of G maps to the **No** case of G'. ◀

Our final reduction to Maximum Cut requires our graph to be regular and $|E| \in \mathcal{O}(n^2)$, so before proceeding, we must first take a standard sparsification step.

▶ **Lemma 24.** *There is a constant $c_3 > 0$ such that for any sufficiently small $\epsilon, \zeta > 0$, it is SSE-hard under randomized reductions to distinguish between the following two cases for an unweighted, non-simple graph $G = (V, E)$ with $|E| \in \Theta(|V|^2)$ and maximum degree at most $(1 + \zeta)d$, where d is the minimum degree:*

Yes: *There exists $S \subseteq V$ such that $\mu(S) \in [\frac{1}{2} - \zeta, \frac{1}{2} + \zeta]$ and $\Phi(S) \leq 4\epsilon$.*

No: *Every $S \subseteq V$ with $\mu(S) \in [\frac{1}{10} + \zeta, \frac{1}{2}]$ satisfies $\Phi(S) \geq c_3\sqrt{\epsilon}$.*

Proof. We begin with a gap instance $G = (V, E)$ of the form in Lemma 23. We randomly create $G' = (V', E')$ as follows. For each vertex $u \in V$, create d_u copies and add them to G'. That is, we set $V' = \{u_1, u_2, \ldots, u_{d_u} \mid u \in V\}$. Let $g : V' \to V$ map copies $u_i \in V'$ to their original vertex $u \in V$. For each edge $uv \in E$ and copies $u' \in g^{-1}(u), v' \in g^{-1}(v)$, let $p_{uv} := \frac{R}{d_u d_v}$ for a parameter R we will adjust later. Now add $\lfloor p_{uv} \rfloor$ copies of $u'v'$ to E' and randomly add a final edge with probability $p_{uv} - \lfloor p_{uv} \rfloor$. Note that the expected number of $u'v'$ edges is equal to p_{uv}.

Let us consider the properties of G'. We have that $|V'| = \sum_{v \in V} d_v = 2|E|$ and

$$\mathbb{E}[|E'|] = \sum_{uv \in E} p_{uv} \cdot d_u \cdot d_v = R|E|.$$

Let q_e be the random variable denoting the number of edges in G' "produced" from $e \in E$. By a standard application of Chernoff bounds and the union bound over $2^{V'}$ and E, we can show, for any constant $\alpha > 0$, that

1. $(1 - \alpha)\mathbb{E}[|\delta_{G'}(S)|] \leq |\delta_{G'}(S)| \leq (1 + \alpha)\mathbb{E}[|\delta_{G'}(S)|]$ for all $S \subseteq V'$ and
2. $(1 - \alpha)R \leq q_e \leq (1 + \alpha)R$ for all $e \in E$

with high probability, assuming $R \in \Omega(n)$. In particular, item 1 implies that $d'_v \in [(1 - \alpha)R, (1 + \alpha)R]$ for all $v \in V'$, where d'_v is the degree of v in G'. Suppose from now on that items 1 and 2 are both true. Now suppose G is a **Yes** instance of Lemma 23. That is, there is a $S \subseteq V$ such that $\mu(S) \in [\frac{1}{2} - \eta, \frac{1}{2} + \eta]$ and $\Phi(S) \leq 3\epsilon$. Let $S' := g^{-1}(S)$, and μ' be defined for G' as μ is defined for G. Then we calculate

$$\mu'(S') = \frac{\sum_{u \in S} d'_u}{\sum_{v \in V} d'_v} \geq \frac{(R - \alpha)|S'|}{(R + \alpha)|V'|} = \frac{R - \alpha}{R + \alpha} \frac{\sum_{u \in S} d_u}{\sum_{v \in V} d_v} = \frac{R - \alpha}{R + \alpha}\mu(S) \geq \frac{R - \alpha}{R + \alpha}(\frac{1}{2} - \eta).$$

Applying the same calculation for $\mu'(V' \setminus S')$ yields the upper bound

$$\mu'(S') = 1 - \mu'(V \setminus S') \leq 1 - \frac{R - \alpha}{R + \alpha}(\frac{1}{2} - \eta).$$

For sufficiently small $\alpha > 0$, this implies that $\mu'(S') \in [\frac{1}{2} - \zeta, \frac{1}{2} + \zeta]$. Let Φ' be defined for G' as Φ for G. We calculate

$$\Phi'(S') = \frac{|\delta_{G'}(S')|}{\sum_{u \in S'} d_u} \leq \frac{(R + \alpha)|\delta_G(S)|}{(R - \alpha)\sum_{u \in S} d_u} = \frac{R + \alpha}{R - \alpha}\Phi(S) \leq \frac{R + \alpha}{R - \alpha}3\epsilon.$$

For sufficiently small $\alpha > 0$, we get the upper bound $\Phi'(S') \leq 4\epsilon$. This establishes that the **Yes** case of G maps to the **Yes** case of G' with high probability.

Now suppose that G is a **No** instance of Lemma 23. Consider any $S' \subseteq V'$ with $\mu'(S') \in [\frac{1}{10} + \zeta, \frac{1}{2}]$. Let $f \in [0, 1]^V$ be a vector defined by $f_v = \frac{|S' \cap g^{-1}(v)|}{d_v}$. That is, f indicates the proportion of each original vertex selected by S'. We calculate

$$\begin{aligned}\Phi'(S') &= \frac{|\delta_{G'}(S')|}{\sum_{v\in V'} d'_v} \geq \frac{(1-\alpha)\mathbb{E}[|\delta_{G'}(S')|]}{(R+\alpha)|V'|} \\ &= \frac{1-\alpha}{R+\alpha}\frac{\sum_{uv\in E} p_{uv}\cdot((1-f_u)d_u f_v d_v + f_u d_u(1-f_v)d_v)}{|V'|} \\ &= \frac{R(1-\alpha)}{R+\alpha}\frac{\sum_{uv\in E}(1-f_u)f_v + f_v(1-f_u)}{|V'|}\end{aligned}$$

Let $S \subseteq V$ be a random variable sampled by including $v \in S$ with probability f_v. Then we have that

$$\mathbb{E}[\Phi(S)] = \frac{\mathbb{E}[|\delta_G(S)|]}{\sum_{v\in V} d_v} = \frac{\sum_{uv\in E}(1-f_u)f_v + f_u(1-f_v)}{|V'|} \leq \frac{R+\alpha}{R(1-\alpha)}\Phi'(S').$$

We can apply Chernoff bounds on $|\delta_G(S)|$ to find that $\Phi(S) < (1+\alpha)\mathbb{E}[\Phi(S)]$ with high probability. We additionally calculate

$$\mathbb{E}[\mu(S)] = \frac{\sum_{u\in V} f_u d_u}{\sum_{v\in V} d_v} = \frac{|S'|}{|V'|} \geq \frac{R-\alpha}{R+\alpha}\mu(S) \geq \frac{R-\alpha}{R+\alpha}(\frac{1}{10}+\zeta).$$

By setting $\alpha, \eta > 0$ sufficiently small and applying Chernoff bounds, this implies that $\mathbb{E}[\mu(S)] > \frac{1}{10} + \eta$ with probability at least $\frac{1}{2}$. Thus, by the probabilistic method, there exists some $S \subseteq V$ such that $\mu(S) > \frac{1}{10} + \eta$ and $\Phi(S) < (1+\alpha)\mathbb{E}[\Phi(S)]$. Applying Lemma 23, we find that

$$\Phi'(S') \geq \frac{R(1-\alpha)}{R+\alpha}\mathbb{E}[\Phi(S)] > \frac{R(1-\alpha)}{(R+\alpha)(1+\alpha)}\Phi(S) \geq \frac{R(1-\alpha)}{(R+\alpha)(1+\alpha)}c_2\sqrt{\epsilon}.$$

So, set $c_3 := \frac{R(1-\alpha)}{(R+\alpha)(1+\alpha)}c_2$. This establishes that the **No** case of G maps to the **No** case of G' with high probability. ◀

We now convert the language of μ and Φ to the simpler language of cardinalities of sets.

▶ **Lemma 25.** *There exists a constant $c' > 0$ such that for all sufficiently small $\varepsilon, \eta > 0$, it is SSE-hard under randomized reductions to distinguish between the following cases for an unweighted graph $G = (V, E)$ with $|E| \in \Theta(|V|^2)$:*

Yes: *There exists a cut $S \subseteq V$ with $(\frac{1}{2}-\eta)|V| \leq |S| \leq \frac{|V|}{2}$ such that $|\delta(S)| \leq 10\varepsilon|E|$.*

No: *For all cuts $S \subseteq V$ with $|S| \leq \frac{|V|}{2}$, either $|S| \leq \frac{|V|}{5}$ or $|\delta(S)| \geq c'\sqrt{\varepsilon}|E|$.*

Proof. We begin with a gap instance $G = (V, E)$ of the form in Lemma 24. We will not need to modify this instance to produce our desired result. Suppose that G is a **Yes** instance of Lemma 24. Then there is some $S \subseteq V$ with $\mu(S) \in [\frac{1}{2}-\zeta, \frac{1}{2}+\zeta]$ and $\Phi(S) \leq 4\epsilon$. Using that every vertex $v \in V$ has degree $d_v \leq (1+\zeta)d$, where d is the minimum degree of G, we can bound

$$\mu(S) = \frac{\sum_{u\in S} d_u}{\sum_{v\in V} d_v} \leq \frac{|S|(1+\zeta)d}{|V|d}.$$

Thus, $|S| \geq \frac{1/2-\zeta}{1+\zeta}|V|$. For sufficiently small $\zeta > 0$, we can then bound below $|S| \geq (\frac{1}{2}-\eta)|V|$. Because $\mu(S) = \mu(V \setminus S)$, we also bound $|V \setminus S| \geq (\frac{1}{2}-\eta)|V|$. If $|S| > \frac{|V|}{2}$, then swap S with $V \setminus S$, noting that $\delta(S) = \delta(V \setminus S)$. This fulfills the $(\frac{1}{2}-\eta)|V| \leq |S| \leq \frac{|V|}{2}$ condition. Similarly, we can bound

$$\Phi(S) = \frac{|\delta(S)|}{\sum_{u\in S} d_u} \geq \frac{|\delta(S)|}{|V|d(1+\epsilon)}.$$

With the bound of $\Phi(S) \leq 4\epsilon$ from Lemma 24, this implies that

$$|\delta(S)| \leq 4\epsilon|V|d(1+\zeta) \leq 10\epsilon|E|$$

for sufficiently small $\zeta > 0$. Thus, G being a **Yes** instance of Lemma 24 implies the **Yes** conditions of this lemma.

Suppose that G is a **No** instance of Lemma 24. Consider any $S \subseteq V$. Suppose first that $\mu(S) < \frac{1}{10} + \zeta$. Then we have that

$$\mu(S) = \frac{\sum_{u \in S} d_u}{\sum_{v \in V} d_v} \geq \frac{|S|d}{|V|(1+\zeta)d},$$

and so $|S| \leq (\frac{1}{10} + \zeta)(1+\zeta)|V|$. For ζ sufficiently small, this implies $|S| \leq \frac{|V|}{5}$. Suppose otherwise, that $\mu(S) \geq \frac{1}{10} + \zeta$. Then we have that

$$\Phi(S) = \frac{|\delta(S)|}{\sum_{u \in S} d_u} \leq \frac{|\delta(S)|}{|V|d},$$

and so $|\delta(S)| \geq 3d\sqrt{\epsilon}|V| \geq \frac{6}{1+\zeta}\sqrt{\epsilon}|E|$. Setting $c' := \frac{6}{1+\zeta}$ finishes the proof. ◀

We now reduce from gap instances of the type in Lemma 25 to show hardness for Maximum Cut on split graphs.

▶ **Theorem 3.** *There exists a constant $c > 0$ such that there is no polynomial-time $(1-c)$-approximation algorithm for Maximum Cut on split graphs, subject to the Small Set Expansion Hypothesis under randomized reductions.*

Proof. We reduce from a graph $G = (V, E)$ of the form in Lemma 25 to a (simple) split graph $G' = (V' = K \cup I, E')$ as follows. Let the clique portion of G' be $K := V$, and let the independent portion of G' be $I := \{v_e\}$. We define the edge set of G' as $E' := \{uv_e \mid u \in V, e \in u\} \cup \{uw \mid u, w \in V\}$. That is, we place a copy of V in the clique portion of G', and place one vertex for each edge of E in the independent set portion, each connected to its two endpoints in V.

Note that in a split graph, the maximum cut is defined solely by its intersection with the clique portion K of the graph, as the decision for vertices in the independent set portion I can be made greedily. Define $\delta'(S)$ for $S \subseteq K = V$ to be the edges in the maximum cut of G' defined by S, breaking ties arbitrarily.

Suppose that G is a **Yes** instance as defined in Lemma 25. Then by direct counting, we have that, in G',

$$\mathrm{mc}(G') \geq |\delta'(S)| \geq (0.25 - \eta^2)|V|^2 + (2 - 10\varepsilon)|E|.$$

Define $\omega := (0.25 - \eta^2)|V|^2 + (2 - 10\varepsilon)|E|$. Suppose instead that G is a **No** instance and consider any $S \subseteq V$ with $|S| \leq 0.5|V|$. We will show that, with the right choice of ε and η, $|\delta'(S)| \leq (1-c)\omega$ for some constant $c > 0$. If $|S| \leq \frac{|V|}{5}$, then

$$|\delta'(S)| \leq \frac{|V|^2}{5} + 2|E|.$$

Because $|E| \in \Theta(|V|^2)$, and $\frac{1}{5} < 0.25 - \eta^2$ for sufficiently small η, this implies that $|\delta'(S)| < (1-c)\omega$ for some constant $c > 0$ when η and ε are sufficiently small. Otherwise, if $|S| \geq \frac{|V|}{5}$, then we must have that $|\delta_G(S)| \geq c'\sqrt{\varepsilon}|E|$. Then

$$|\delta'(S)| \leq 0.25|V|^2 + (2 - c'\sqrt{\varepsilon})|E|.$$

As above, because $2 - c'\sqrt{\varepsilon} < 2 - 10\varepsilon$ for sufficiently small ε, we have that $|\delta'(S)| < (1-c)\omega$ for some constant $c > 0$ when η and ε are sufficiently small. Thus, in the **No** case, we have that $\mathrm{mc}(G') \leq (1-c)\omega$. Therefore, it is SSE-hard to distinguish between $\mathrm{mc}(G') \geq \omega$ and $\mathrm{mc}(G') \leq (1-c)\omega$. This implies SSE-hardness of approximating Maximum Cut to a factor of $1-c$ on split graphs. ◀

It is critical that Lemma 25 allows for non-simple graphs. If G were simple, then our reduction in the proof of Theorem 3 results in the relation $\mathrm{mc}(G') = \mathrm{mc}(G^c) + 2|E|$, where $G^c = (V, E^c)$ is the complement of G. This is proved explicitly in [3]. Then, if $|E| \in \omega(|E^c|)$, the value $2|E|$ is a good estimate for $\mathrm{mc}(G')$. If $|E| \in \mathcal{O}(|E^c|)$, then $|E^c| \in \Omega(|V|^2)$. In this case, G^c is a dense graph and so there is a PTAS for $\mathrm{mc}(G^c)$ [7, 16]. In particular, these cases imply that there is a PTAS for $\mathrm{mc}(G')$ whenever G' is a 2-split graph (i.e., all vertices in I have degree 2) without any "duplicated" vertices of I that have the same neighborhood.

References

1. Ranendu Adhikary, Kaustav Bose, Satwik Mukherjee, and Bodhayan Roy. Complexity of maximum cut on interval graphs. *Discrete & Computational Geometry*, 70(2):307–322, 2023. doi:10.1007/S00454-022-00472-Y.
2. Alexey Barsukov and Bodhayan Roy. Maximum cut on interval graphs of interval count two is np-complete. *arXiv preprint*, 2022. arXiv:2203.06630.
3. Hans L Bodlaender and Klaus Jansen. On the complexity of the maximum cut problem. *Nordic Journal of Computing*, 7(1):14–31, 2000.
4. Celina MH de Figueiredo, Alexsander A de Melo, Fabiano S Oliveira, and Ana Silva. Maximum cut on interval graphs of interval count four is np-complete. *Discrete & Computational Geometry*, 71(3):893–917, 2024. doi:10.1007/S00454-023-00508-X.
5. Tomas Feder and Pavol Hell. List homomorphisms to reflexive graphs. *Journal of Combinatorial Theory, Series B*, 72(2):236–250, 1998. doi:10.1006/JCTB.1997.1812.
6. Tomás Feder and Carlos Subi. Packing edge-disjoint triangles in given graphs. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 19, page 13, 2012.
7. W Fernandez De La Vega. Max-cut has a randomized approximation scheme in dense graphs. *Random Structures & Algorithms*, 8(3):187–198, 1996. doi:10.1002/(SICI)1098-2418(199605)8:3\%3C187::AID-RSA3\%3E3.0.CO;2-U.
8. Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995. doi:10.1145/227683.227684.
9. Venkatesan Guruswami. Maximum cut on line and total graphs. *Discrete applied mathematics*, 92(2-3):217–221, 1999. doi:10.1016/S0166-218X(99)00056-6.
10. Frank Hadlock. Finding a maximum cut of a planar graph in polynomial time. *SIAM Journal on Computing*, 4(3):221–225, 1975. doi:10.1137/0204019.
11. Laurie J Hendren, Guang R Gao, Erik R Altman, and Chandrika Mukerji. A register allocation framework based on hierarchical cyclic interval graphs. In *Compiler Construction: 4th International Conference, CC'92 Paderborn, FRG, October 5–7, 1992 Proceedings 4*, pages 176–191. Springer, 1992. doi:10.1007/3-540-55984-1_17.
12. Jun-Ting Hsieh and Pravesh K Kothari. Approximating max-cut on bounded degree graphs: Tighter analysis of the fkl algorithm. *arXiv preprint arXiv:2206.09204*, 2022. doi:10.48550/arXiv.2206.09204.
13. John R. Jungck and Rama Viswanathan. Chapter 1 - graph theory for systems biology: Interval graphs, motifs, and pattern recognition. In Raina S. Robeva, editor, *Algebraic and Discrete Mathematical Methods for Modern Biology*, pages 1–27. Academic Press, Boston, 2015. doi:10.1016/B978-0-12-801213-0.00001-0.

14 Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103, 1972. doi:10.1007/978-1-4684-2001-2_9.

15 Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM Journal on Computing*, 37(1):319–357, 2007. doi:10.1137/S0097539705447372.

16 Claire Mathieu and Warren Schudy. Yet another algorithm for dense max cut: go greedy. In *SODA*, volume 8, pages 176–182, 2008. URL: http://dl.acm.org/citation.cfm?id=1347082.1347102.

17 Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 755–764, 2010. doi:10.1145/1806689.1806792.

18 Prasad Raghavendra, David Steurer, and Madhur Tulsiani. Reductions between expansion problems. In *2012 IEEE 27th Conference on Computational Complexity*, pages 64–73. IEEE, 2012. doi:10.1109/CCC.2012.43.

19 Luca Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 453–461, 2001. doi:10.1145/380752.380839.