arXiv:1110.1360v1 [cs.DS] 6 Oct 2011

Polynomial integrality gaps for strong SDP relaxations of Densest k-subgraph

Aditya Bhaskara* Moses Charikar† Venkatesan Guruswami ‡
Aravindan Vijayaraghavan§ Yuan Zhou ¶

Abstract

The Densest k-subgraph problem (i.e. find a size k subgraph with maximum number of edges), is one of the notorious problems in approximation algorithms. There is a significant gap between known upper and lower bounds for Densest k-subgraph: the current best algorithm gives an $\approx O(n^{1/4})$ approximation, while even showing a small constant factor hardness requires significantly stronger assumptions than P $\neq$ NP. In addition to interest in designing better algorithms, a number of recent results have exploited the conjectured hardness of Densest k-subgraph and its variants. Thus, understanding the approximability of Densest k-subgraph is an important challenge.

In this work, we give evidence for the hardness of approximating Densest k-subgraph within polynomial factors. Specifically, we expose the limitations of strong semidefinite programs from SDP hierarchies in solving Densest k-subgraph. Our results include:

- A lower bound of $\Omega\big(n^{1/4}/\log^3 n\big)$ on the integrality gap for $\Omega(\log n/\log\log n)$ rounds of the Sherali-Adams relaxation for Densest k-subgraph. This also holds for the relaxation obtained from Sherali-Adams with an added SDP constraint. Our gap instances are in fact Erdös-Renyi random graphs.
- For every $\epsilon > 0$, a lower bound of $n^{2/53-\epsilon}$ on the integrality gap of $n^{\Omega(\epsilon)}$ rounds of the Lasserre SDP relaxation for Densest k-subgraph, and an $n^{\Omega_\epsilon(1)}$ gap for $n^{1-\epsilon}$ rounds. Our construction proceeds via a reduction from random instances of a certain Max-CSP over large domains.

In the absence of inapproximability results for Densest k-subgraph, our results show that beating a factor of $n^{\Omega(1)}$ is a barrier for even the most powerful SDPs, and in fact even beating the best known $n^{1/4}$ factor is a barrier for current techniques.

Our results indicate that approximating Densest k-subgraph within a polynomial factor might be a harder problem than Unique Games or Small Set Expansion, since these problems were recently shown to be solvable using $n^{\epsilon^{\Omega(1)}}$ rounds of the Lasserre hierarchy, where ϵ is the completeness parameter in Unique Games and Small Set Expansion.

*Department of Computer Science, Princeton University, and Center for Computational Intractability. Supported by Supported in part by NSF under CCF 0832797 and AF 0916218. Email: bhaskara@cs.princeton.edu.

†Department of Computer Science, Princeton University, and Center for Computational Intractability. Supported in part by NSF under CCF 0832797 and AF 0916218. Email: moses@cs.princeton.edu.

‡Computer Science Department, Carnegie Mellon University. Research supported in part by a Packard Fellowship and NSF CCF 1115525. Email: guruswami@cmu.edu.

§Department of Computer Science, Princeton University, and Center for Computational Intractability. Supported in part by NSF under CCF 0832797 and AF 0916218. Email: aravindv@cs.princeton.edu.

¶Computer Science Department, Carnegie Mellon University. Supported in part by US-Israel BSF grant 2008293, MSR-CMU Center for Computational Thinking, and NSF CCF 1115525. Email: yuanzhou@cs.cmu.edu.

1 Introduction

The densest k-subgraph problem takes as input a graph $G(V, E)$ on n vertices and a parameter k, and asks for a subgraph of G on at most k vertices having the maximum number of edges. While it is a fundamental graph optimization problem and arises in several applications (community detection in social networks, identifying protein families and molecular complexes in protein-protein interaction networks, etc), there is a huge gap between the best approximation algorithm and the known inapproximability results. The current best approximation algorithm due to [BCC+10] gives $O(n^{1/4+\varepsilon})$-factor approximation algorithm which runs in time $n^{O(1/\varepsilon)}$ for any constant $\varepsilon > 0$. On the inapproximability side, [Fei02] initially showed a small constant factor inapproximability for Densest k-subgraph using the random 3-SAT assumption. [Kho04] used quasi-random PCPs to rule out a PTAS. More recently, [RS10, AAM+10] used more non-standard assumptions to rule out any constant factor approximation algorithms.

While only constant factor approximations have been ruled out, it is commonly believed that Densest k-subgraph is much harder to approximate even on average (for a natural distribution on hard instances). Recently, average-case hardness assumptions based on the hardness of "planted" versions of Densest k-subgraph were used for public key cryptography [ABW08] and in showing that financial derivates can be fraudulently priced without detection [ABBG10]. Given the interest in Densest k-subgraph from both the algorithms and the complexity point of view, developing a better understanding of the problem is an important challenge for the field.

In this work, we study lift-and-project relaxations for Densest k-subgraph. Lift-and-project methods are systematic iterative procedures to obtain sequences of increasingly stronger mathematical programming relaxations for an integer optimization problem (e.g. Lovász-Schrijver [LS91], Sherali-Adams [SA90] and Lasserre [Las01]. See the survey by Laurent [Lau03] for a comparison). Typically, the relaxation obtained after r levels of these strengthenings can be solved in $n^{O(r)}$ time. A number of recent papers have studied the strength and limitations of such relaxations as a basis for designing approximation algorithms for various problems [ABLT06, CMM09, Chl06, CS06, dVM07, GMPT07, GMT09, KS09, RS09, Sch08, STT07a, STT07b, Tul09] (see the recent survey by Chlamtac and Tulsiani [CT11]). In most cases of approximation algorithms that use strengthened LP and SDP relaxations, such relaxations can be obtained from a few levels of such lift-and-project procedures. In fact, the $O(n^{1/4+\varepsilon})$ approximation algorithm of [BCC+10] for Densest k-subgraph uses a linear programming relaxation which is weaker than that obtained from $O(1/\varepsilon)$ levels of the Sherali-Adams hierarchy.[1] [BCC+10] also show that the integrality gap becomes $O(n^{1/4-\varepsilon})$ after $n^{O(\varepsilon)}$ levels of the Sherali Adams LP hierarchy.

Our results. In this paper, we first study lift-and-project relaxations for Densest k-subgraph obtained from the Sherali-Adams hierarchy. We show that even $\Omega\left(\frac{\log n}{\log \log n}\right)$ levels of the Sherali-Adams relaxation have an integrality gap of $\tilde{\Omega}(n^{1/4})$. Then, we turn to the Lasserre hierarchy for Densest k-subgraph. We show an integrality gap of polynomial ratio (n^{ϵ}, for small enough constant ϵ) for almost linear ($n^{1-O(\epsilon)}$) levels of the Lasserre relaxation. If we only aim at an integrality gap for polynomial (n^{ϵ}) levels of the Lasserre relaxation, the ratio of the gap can be as large as $n^{2/53-O(\epsilon)}$.

Our gap instances are actually (Erdös - Renyi) random graph instances $\mathcal{G}(n, p)$ and random bipartite graphs under a special distribution – hence, we show that natural distributions of instances are integrality gap instances with high probability.

[1] [BCC+10] also gives a purely combinatorial algorithm that does not use a linear program.

We note that prior results exhibiting gap instances for lift-and-project relaxations do so for problems that are *already known* to be hard to approximate under some suitable assumption; based on this hardness result, one would *expect* lift-and-project relaxations to have an integrality gap that matches the inapproximability factor.

Our gap constructions for Densest k-subgraph in this paper are a rare exception to this trend, as the integrality gaps we show are substantially stronger than the (very weak) hardness bounds known for the problem. In fact, we are only aware of the following examples where a polynomial-round Lasserre integrality gap stronger than the corresponding NP-hardness result is known : Max K-CSP, K-coloring [Tul09], Balanced Separator and Uniform Sparsest Cut [GSZ11]. In the first two cases, NP-hardness results that are not that far from the gaps are known [ST00, Kho02] and for Max K-CSP a matching Unique-Games hardness is also known [ST09]. For the other two problems, constant factor integrality gaps were shown for linear number of rounds of Lasserre hierarchy [GSZ11]. Again, while these problems are not known to be APX-hard, under the conjectured intractability of Small Set Expansion, they are known to be hard to approximate within any constant factor.

In the absence of inapproximability results for Densest k-subgraph, our results show that beating a factor of $n^{\Omega(1)}$ is a barrier for even the most powerful SDPs, and in fact even beating the best known $n^{1/4}$ factor is a barrier for current techniques. These results are perhaps indicative of the hardness of approximating Densest k-subgraph within $n^{\Omega(1)}$ factors.

Relation to the Small Set Expansion and Unique Games. A problem related to Densest k-subgraph is the Small Set Expansion (SSE) problem, which has received a lot of recent attention due to strong connections to the Unique Games conjecture [RS10]. One way to state the SSE conjecture [RS10] (which is known to imply the Unique Games conjecture) is as follows: for all $\epsilon > 0$, there exists δ, D (think of D as a constant), such that the following problem is not polynomial-time solvable:

Definition 1.1 (The Gap-SSE problem)**.** *Given a D-regular instance $G(V, E)$ with $k = \delta n$, the Gap-SSE problem is to distinguish between the following two cases.*

- ***Yes case.*** *There exists a subgraph of k vertices with average degree at least $(1 - \varepsilon)D$.*
- ***No case.*** *All subgraphs of k vertices have average degree at most εD.*

Clearly, Densest k-subgraph is hard to approximate within any constant factor, assuming the Small Set Expansion conjecture. On the other hand, our results indicate that approximating Densest k-subgraph even within a polynomial factor may be a harder problem than Unique Games or Small Set Expansion, because these problems were recently shown to be solvable using $n^{\epsilon^{\Omega(1)}}$ rounds of the Lasserre hierarchy, where ϵ is the completeness parameter in Unique Games and Small Set Expansion [BRS11, GS11].

2 Preliminaries

2.1 Notation.

We introduce some notation which will be used throughout the paper. $G = (V, E)$ refers to a graph which is an instance of the Densest k-subgraph problem on n vertices, and k refers to the size of

the subgraph we are required to output. For an induced subgraph $H \subseteq G$, we denote by $d(H)$ the average degree (or *density* of H). For a vertex v in subgraph H, we will denote by $\Gamma_H(v)$ the set of neighbors of v in H (the suffix will be dropped when $H = G$).

The phrase "with high probability" will mean: with probability $1 - \frac{1}{p(n)}$, for any polynomial $p(n)$.

2.2 The relaxation hierarchies for Densest k-subgraph.

We will be concerned with the SDP relaxations derived from the Sherali-Adams and Lasserre hierarchies for Densest k-subgraph. As in other lift-and-project schemes, a feasible solution to r levels of these hierarchies satisfies the condition that for any set of r vertices, it defines a valid distribution over integral solutions for these vertices – in particular, the integrality gap becomes 1 after n levels. Further, the relaxations given by r levels of the Sherali-Adams and Lasserre hierarchies can be solved in $n^{O(r)}$ time. We are interested in the integrality gap of r levels of these relaxations for Densest k-subgraph. Refer to [CT11] for a more comprehensive comparison of these relaxation hierarchies.

2.2.1 The Sherali-Adams LP hierarchy.

The Sherali-Adams hierarchy starts with a simple LP relaxation of a $\{0, 1\}$ integer program, and obtains a sequence of successively tighter relaxations with more levels. The natural LP relaxation for Densest k-subgraph (LP1 in Figure 1) [SW98, FS97] has variables $\{x_i\}$ to denote if vertex i belongs to the solution, and edge variables $\{x_{ij}\}_{(i,j)\in E(G)}$ to denote if both i, j are in the subgraph. This LP has an integrality gap of $\Omega(\frac{n}{k})$ ([FKP01, FS97]).

Natural LP (LP1):	**Min. degree LP (LP2):**
$\max \sum_{(i,j)\in E(G)} x_{ij}$	$\max \quad d$
$\text{s.t.} \sum_{i\in V} x_i \leq k$	$\text{s.t.} \sum_{i\in V} x_i \leq k, \quad$ and
$\forall i, j \in V, \quad 0 \leq x_{ij} \leq x_i \leq 1$	$\exists \{x_{ij} \mid i, j \in V\}$ s.t.
	$\forall i \in V \quad \sum_{j\in\Gamma(i)} x_{ij} \geq d x_i$
	$\forall i, j \in V \quad x_{ij} = x_{ji}$
	$\forall i, j \in V \quad 0 \leq x_{ij} \leq x_i \leq 1$

Figure 1: Two Linear Programming relaxations for Densest k-subgraph

For our integrality gaps, we will in fact start with a stronger basic (first-level) linear program (LP2 in Figure 1) which is equivalent upto a factor of 2 (see [BCC+10]). Intuitively, it tries to find a k-subgraph H where the *minimum degree* d_H is maximized. An LP hierarchy obtained from this min. degree LP (LP2) was in fact used by [BCC+10] to obtain their approximation algorithm. [2]

[2]While the program as stated is not linear, we guess the degree d and consider the feasibility linear program that

Let us consider strengthening this LP by considering r levels of the Sherali-Adams hierarchy (SA_r, shown in Figure 2). In the *lifted* LP, the variable x_S is supposed to capture whether every vertex in S belongs to the chosen k-subgraph (i.e., $x_S = \prod_{i \in S} x_i$). Further if we take two sets S, S' of $\leq r$ vertices, the local distributions induced by a feasible solution (using the inclusion-exclusion constraints), agree on the variables in the intersection $S \cap S'$. We follow the notation established in [CT11] while defining the hierarchy.

2.2.2 The mixed hierarchy (Sherali-Adams + SDP).

The mixed hierarchy (also refered to as $SA+$) imposes an additional SDP constraint on top of the Sherali-Adams LP relaxation. In particular, it asks for the values x_{ij} to come from vector inner products i.e. the matrix $X = (x_{ij})$ is p.s.d. Most known algorithms which proceed by rounding a relaxation obtained from an SDP hierarchy [Chl06, CS06, BRS11] work with this mixed hierarchy [3].[RS09, KS09] and [GMT09] considered this hierarchy and obtained integrality gaps for Unique Games and approximation-resistant CSPs.

$$\begin{aligned}
&\max \quad d, \text{ s.t.} \\
&\exists \{x_S \mid S \subseteq V, |S| \leq r\} \text{ s.t. } x_\emptyset = 1 \text{ and} \\
\forall S, T \subseteq V \text{ s.t } |S| + |T| \leq r : \\
&\sum_{i \in V} \sum_{J \subseteq T} (-1)^{|J|} x_{S \cup J \cup \{i\}} \leq k \sum_{J \subseteq T} (-1)^{|J|} x_{S \cup J} \qquad (1) \\
\forall i \in V \quad &\sum_{j \in \Gamma(i)} \sum_{J \subseteq T} (-1)^{|J|} x_{S \cup J \cup \{i,j\}} \geq d \sum_{J \subseteq T} (-1)^{|J|} x_{S \cup J \cup \{i\}} \qquad (2) \\
&0 \leq \sum_{J \subseteq T} (-1)^{|J|} x_{S \cup J} \leq 1 \qquad (3)
\end{aligned}$$

Figure 2: Sherali-Adams LP relaxation (r levels) for Densest k-subgraph: SA_r

One level of the mixed hierarchy for Densest k-subgraph gives the SDP relaxation introduced in [FS97, SW98]. [BCC+10] show that the mixed hierarchy performs better than log-density based arguments (which are captured by just the LP hierarchy) in a *planted* model.[4] It is interesting in this light to obtain integrality gaps for mixed hierarchy.

2.2.3 The Lasserre hierarchy.

The Lasserre hierarchy produces a sequence of SDP relaxations which are stronger than the Sherali-Adams and the mixed hierarchies. As in [CT11], the r-level Lasserre SDP for Densest k-subgraph introduces a vector $\boldsymbol{U}_S$ for each subset $S \subseteq V$ with $|S| \leq r$ (Figure 3). The intended solution sets $\boldsymbol{U}_S = \boldsymbol{U}_\emptyset$ if every vertex in S belongs to the densest k-subgraph, and $\boldsymbol{U}_S = \boldsymbol{0}$ otherwise. The vector lengths $\|\boldsymbol{U}_S\|^2$ correspond to valid LP values x_S for the Sherali-Adams relaxation presented above.

is obtained.

[3] [GS11] is an exception and seems to need a relaxation given by the Lasserre hierarchy.

[4] In particular, the problem of detecting if dense k-subgraph is planted in a random graph or not, in the parameter range $D < n^{1/2}$.

$$\max \sum_{(u,v)\in E} \left\|\boldsymbol{U}_{\{u,v\}}\right\|^2$$

such that

$$\begin{aligned}
&\langle \boldsymbol{U}_{S_1}, \boldsymbol{U}_{S_2}\rangle \geq 0 \text{ for all } S_1, S_2\\
&\langle \boldsymbol{U}_{S_1}, \boldsymbol{U}_{S_2}\rangle = \langle \boldsymbol{U}_{S_3}, \boldsymbol{U}_{S_4}\rangle \text{ when } S_1 \cup S_2 = S_3 \cup S_4\\
&\sum_{v\in V} \langle \boldsymbol{U}_{\{v\}}, \boldsymbol{U}_S\rangle \leq k \left\|\boldsymbol{U}_S\right\|^2 \text{ for all } S\\
&\left\|\boldsymbol{U}_\emptyset\right\|^2 = 1
\end{aligned}$$

Figure 3: Lasserre hierarchy (r levels) for Densest k-subgraph

Remark 2.1. *As in Section 2.2.1, we can write an SDP which tries to find the k-subgraph of largest induced minimum degree d. This can be captured by the SDP constraint (analogous to (2))*

$$\forall S, \forall u \in V, \quad \sum_{v\in\Gamma(u)} \langle \boldsymbol{U}_{\{u,v\}}, \boldsymbol{U}_S\rangle \geq d \cdot \langle \boldsymbol{U}_{\{u\}}, \boldsymbol{U}_S\rangle, \tag{4}$$

However, we show in Section 4.3 that our integrality gaps also hold for the Lasserre hierarchy defined by this SDP. We refer to the SDP with constraint (4) as the Min degree Lasserre SDP *.*

3 Integrality Gap for the Sherali-Adams hierarchy

In what follows L will denote the number of levels of the hierarchy we will consider.

Theorem 3.1. *Let $L \leq \frac{\log n}{10 \log\log n}$. The integrality gap of SA_L is at least $\Omega\left(\frac{n^{1/4}}{L \log^2 n}\right)$.*

To prove Theorem 3.1, we present instances G where the relaxation has a solution with value $d = \Omega(n^{1/4}/L)$, while the *integer optimum*, i.e., the largest density of a k-subgraph in G is only $O(\log^2 n)$. It will be notationally convenient to construct gaps for $L/2$ levels.

3.1 The instance.

We in fact give a distribution over instances, and prove that the desired gap holds with high probability. The instances we consider are $\mathcal{G}(n,p)$ random graphs with $p = n^{-1/2}\log n$ (thus the expected degree of each vertex is $D = n^{1/2}\log n$). The parameter k is chosen to be $n^{1/2}$. An easy calculation shows that in any k subgraph, the density (and hence the min-degree) is at most $O(\log^2 n)$ (see full version or [BCC+10, FKP01]). The meat of the argument is thus to show that there exists an LP solution to $\mathrm{SA}_{L/2}$ (Equations (1)-(3)) of value $d = \Omega(n^{1/4}/L)$ even for L of the order $\log n/\log\log n$.

The following are the properties of the distribution $\mathcal{G}(n,p)$ (with above parameters) we will truly be using [see Section A for proofs]. Any graph with these properties admits the solution to $\mathrm{SA}_{L/2}$ which we describe.

1. Every vertex has degree between $(n^{1/2}\log n)/2$ and $2n^{1/2}\log n$.

2. Any two vertices i, j have at least one common neighbor and has at most $O(\log^2 n)$ common neighbours.

3.2 Feasible solution.

Before formally giving the x_S values, we give intuition as to what they *ought to* be. First, we start out setting $x_i = n^{-1/2}$ (equal for all vertices, since $\sum_i x_i \leq k = n^{1/2}$ and no vertex is special). Next, suppose $S \subset V$ with $i \in S$ and think of $d \approx n^{1/4}$. Now (2) implies that $\sum_{j \in \Gamma(i)} x_{S \cup j} \geq n^{1/4} x_S$. Further from (1), we obtain $\sum_{j \in V} x_{S \cup j} \leq n^{1/2} x_S$. Thus we conclude that $x_{S \cup j}$ must be roughly $n^{-1/4} x_S$ for $j \in \Gamma(S)$, while for $j \notin \Gamma(S)$, it should be only $n^{-1/2} x_S$. Now consider $T \subset V$ which span a tree: we could imagine starting with one vertex and adding vertices one by one (each added vertex is a neighbour of the previous ones), and thus conclude that x_T is roughly $n^{-(|T|+1)/4}$ (since $x_i = n^{-1/2}$ to begin with). Now let S be an arbitrary set of vertices and consider a tree $T \supseteq S$: by monotonicity (a corollary of (3)), $x_S \geq x_T$, and since this is true for *every* such T, we need to set x_S to be at least $n^{-(\mathsf{st}(S)+1)/4}$, where $\mathsf{st}(S)$ is the number of vertices (size) in the minimum Steiner tree of S.

These, with additional 'dampening' factors (L-terms), are precisely the values we will set. More precisely we consider the solution

$$x_S = n^{-\frac{1}{4} \cdot (\mathsf{st}(S)+1)} \cdot L^{-|S|}, \tag{5}$$

where $\mathsf{st}(S)$, as above, is the size of the minimum Steiner tree of S. Thus for instance $x_i = n^{-1/2}/L$, while $x_{\{i,j\}} = 1/(n^{3/4} L^2)$ when $(i, j) \in E$ and $1/(nL^2)$ otherwise (the latter is because there is a path of length-2 between any $i, j \in G$ with high probability).

Let us fix $L \leq \log n / (10 \log \log n)$. We now show that the LP solution presented above is feasible for $\mathrm{SA}_{L/2}$ with high probability. The following lemma is useful in simplifying the analysis: it implies that we need to only consider $T = \emptyset$ while showing that the LP solution satisfies constraints (1) and (2). This is where the 'dampening' factors come into play.

Lemma 3.2. *Let S, T be disjoint subsets of V of size at most t and x_S be the solution described above. Then*

$$x_S \geq \sum_{J \subseteq T} (-1)^{|J|} x_{S \cup J} \geq \frac{x_S}{2}$$

Proof. One property of the assignment (5) is that $x_{S \cup i} \leq x_S / L$ for $i \notin S$. Further all the x_S are ≥ 0, and thus in the sum above, the term corresponding to $J \subseteq T$ contibutes positively when $|J|$

is even and negatively otherwise. Hence,

$$\begin{aligned}
&\sum_{J\subseteq T}(-1)^{|J|}x_{S\cup J}\\
&\geq \sum_{\ell=0}^{\ell=\lfloor\frac{|T|}{2}\rfloor}\sum_{\substack{J\subseteq T\\|J|=2\ell}}\left(x_{S\cup J}-\sum_{i\in T\setminus S}x_{S\cup J\cup\{i\}}\right)\\
&\geq \sum_{\ell=0}^{\lfloor\frac{|T|}{2}\rfloor}\sum_{\substack{J\subseteq T\\|J|=2\ell}}x_{S\cup J}(1-|T|/L)\\
&\geq \sum_{\ell=0}^{\lfloor\frac{|T|}{2}\rfloor}\frac{1}{2}\cdot\sum_{\substack{J\subseteq T\\|J|=2\ell}}x_{S\cup J}\geq\frac{x_S}{2}\ \text{(since } |T|\leq L/2)
\end{aligned}$$

A similar proof shows the upper bound, since the $x_{S\cup\{i\}}$ terms for $i\in T$ dominate the contributions of $x_{S\cup J}$ for $|J|>1$. □

Corollary 3.3. *In checking feasibility, it suffices to check* (1) *and* (2) *with* $T=\emptyset$.

Proof. Lemma 3.2 allows us to 'remove' the $\sum_{J\subseteq T}$ on both sides of the equations (and set $T=\emptyset$) by losing a factor of 2. Since we allow constant slack, the claim follows. □

We refer to the constraints (1) and (2) as the *size* and the *density* constraints respectively, because the former says that we should pick only a k-subgraph, and the latter says the minimum degree (density) is at least d. The assignment we described allows us to prove the density constraint easily.

Lemma 3.4. *(Density Constraint) The* x_S *described above satisfy constraints* (2).

Proof. Let $S\subset V$ and $i\in S$. We need to check that $\sum_{j\in\Gamma(i)}x_{S\cup j}\geq\frac{n^{1/4}}{L}\cdot x_S$. It is easy to see that for every $j\in\Gamma(i)$, $\mathsf{st}(S\cup j)\leq\mathsf{st}(S)+1$, and thus $x_{S\cup j}\geq\frac{n^{-1/4}}{L}\cdot x_S$ (the L term is due to the dependence on $|S|$ in (5)). Since there are at least $n^{1/2}\log n/2$ terms in the LHS, the inequality follows. □

3.3 The Size Constraint and Minimum Steiner trees in $\mathcal{G}(n,p)$.

By the above corollary, it suffices to check (noting $k=n^{1/2}$) that

$$\sum_{i\in V}x_{S\cup i}\leq n^{1/2}x_S\qquad\text{for all } S\subset V, |S|<t. \tag{6}$$

We show this by proving that $\mathsf{st}(S\cup i)\geq\mathsf{st}(S)+2$ for *most* $i\in V$, in particular we bound the number of exceptions (lemmas below state the precise bounds). This then implies that (6) holds.

We start with some basic facts (and notation) about Minimum Steiner trees (minST) of $S(\subset V)$ in $\mathcal{G}(n,p)$, with our parameters. We will refer to the vertices in S as the *terminals*, and the rest

of the vertices in a minST as the *non-terminals*. First, the minST must have all its leaves to be terminals. Further, since every two vertices in G have a path of length two, we must have $\mathsf{st}(S) \leq 2|S| - 1$ for all S. This helps us bound the number of *tree structures* the minST of S can have. We define this formally.

Given $S \subset V$, a *tree structure* for S is a tree T along with a mapping $g : V(S) \to V(T)$ which is one-one (not necessarily onto). The vertices in T without an inverse image in S are called *internal vertices* and the rest are also called *fixed vertices*. A tree structure for S is *valid* if it is possible to 'fill in' the internal vertices with distinct vertices from V such that all the edges in the tree are also present in G. [The relation to Steiner trees is apparent – the internal vertices are the Steiner vertices]. Given an internal vertex in T, the vertices of G which take that *position* in some valid 'filling in' are called the set of *candidates* for that position.

Before we get to the lemmas, we note that the number of tree structures for S of size $\leq 2|S|$ is at most $(2|S|)^{2|S|}$ (this is just by a naïve bound using the number of trees). Let us now bound the number of $i \in V$ for which $\mathsf{st}(S \cup i) \leq \mathsf{st}(S) + 1$.

Lemma 3.5. *Let $S \subset V$ and T be a tree structure for a min Steiner tree of S (so the leaves of T are elements of S). Then the number of candidates for each of the* positions *in T is at most* $(\log n)^{2|S|}$.

Proof. The proof is by induction on the size of S. The base case $|S| = 1$ is trivial. Assume the result for all tree structures of sets of size $\leq |S| - 1$. Now consider S. We may assume that T has at least one non-terminal, as otherwise there is nothing to prove.

First, note that there exists a vertex $u \in T$ which is adjacent to at most one non-leaf vertex in T. This is because deleting all the leaves in T gives a tree (which is not empty as there is at least one non-terminal in T), and a leaf in this tree our required u. If u is a terminal, we could remove the leaves attached to u (thus obtaining a subset S' of the terminals), and the remaining tree structure would be a valid min Steiner tree for S'. Further, the set of non-terminals is precisely the same, and thus the inductive hypothesis implies the claim for S. Thus suppose u is a non-terminal.

If the degree of u (in T) is 2, then u has precisely one leaf attached to it (call it ℓ). Consider the tree T' obtained by removing u, ℓ, and let $S' = S \setminus \ell$. Now T' is a min Steiner tree for S' (if not, we could consider use this smaller tree for S' along with a path of length 2 to ℓ to obtain a smaller minimum steiner tree for S). If b is the vertex in T' attached to u, there are at most $(\log n)^{2|S|-2}$ candidates for b, by Induction Hypothesis. For each candidate b, the number of candidate u is only $O(\log^2 n)$, and since ℓ is a terminal. Thus the number of candidates for u is at most $(\log n)^{2|S|}$. The rest of the non-terminals in T are also present in T', and this gives the result.

If degree$(u) > 2$, then there are at least two leaves attached to u, thus the number of candidates for u is only $\log^2 n$. Consider one candidate x for u. Let T' be the tree obtained by removing all the leaves attached to u (thus u is now a leaf), and S' be $S \cup x$ minus the set of leaves attached to u. Now T' is a min Steiner tree structure for S' (otherwise we can obtain a smaller tree for S). Thus by the inductive hypothesis, the number of candidates for any internal vertex in T' is at most $(\log n)^{2|S|-2}$. Since there are only $\log^2 n$ of the x's, it follows that the total #(candidates) for an internal vertex is at most $(\log n)^{2|S|}$.

This completes the proof, by induction. □

An easy corollary is the following.

Corollary 3.6. *Let $S \subset V$. There are at most $(2|S| \log n)^{2|S|}$ vertices i such that $\mathsf{st}(S \cup i) = \mathsf{st}(S)$.*

Proof. Each such i must be the internal vertex of some min Steiner tree for S, and there are at most $(2|S|)^{2|S|}$ tree structures. Lemma 3.5 now implies the claim. □

Lemma 3.7. *Let $S \subset V$. There are at most $(4|S| \log n)^{4|S|} \times n^{1/2}$ vertices i such that $\mathsf{st}(S \cup i) = \mathsf{st}(S) + 1$.*

Proof. Let i be such a vertex. First, note that if there exists a min Steiner tree for $S \cup i$ with i as a leaf, we are done. This is because removing i gives a min Steiner tree for S, and thus i is a neighbour of an internal vertex in a min Steiner tree for S. Thus by Corollary 3.6 there are only $(2|S| \log n)^{2|S|} \times (2n^{1/2} \log n)$ such i.

Thus suppose that the min Steiner tree for $S \cup i$ has i as an internal vertex. We will prove the bound as follows: we consider a tree structure T of size $\mathsf{st}(S)+1$ with leaves being terminals from S; then we show that the number of candidates for any fixed position in T is at most $(2|S| \log n)^{2|S|} n^{1/2}$. This suffices, because the number of choices of tree structures adds an additional factor of $(2|S|)^{2|S|}$.

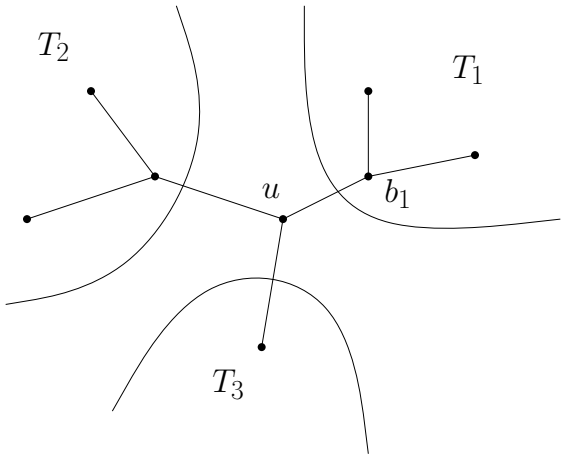


Figure 4: An example

Let us consider a structure T as above, and a position u. Since u is not a leaf, it has degree at least 2. Let the degree be d, and let $T_1, \ldots, T_d$ be the subtrees of T formed by removing u (see figure ...). Now if for some i, T_i is the min Steiner tree for the terminals in T_i, we are done, because then, each candidate for u must be neighbour of an internal vertex in the tree, and by Corollary 3.6 there are only $\sqrt{n} \times (2|S| \log n)^{2|S|}$ candidates. Thus for *each* i, T_i must have a strictly smaller tree T_i'. Let the vertex in T_1 connected to u be called b_1. Now construct a new tree as follows: leave T_1 intact, and replace $T_2, \ldots, T_d$ by $T_2', \ldots, T_d'$; connect b_1 to $T_2', \ldots, T_d'$ using paths of length 2. The number of edges in the new tree is now at most $|T| - d - (d-1) + 2(d-1)$. The first term is the original cost, followed by removal of u, followed by the decrease by using T_i' as opposed to T_i, followed by the cost of adding length-2 paths.

Thus the new tree has cost at most $|T| - 1$, and thus it is optimal for S! Further, u is adjacent to b_1 which is an internal vertex, and thus the number of candidates is bounded by the desired quantity. □

Putting things together. Consider the sum $\sum_{i \in V} x_{S \cup i}$. Corollary 3.6 implies that there are at most $(L \log n)^L$ terms which contribute a value x_S/L. Lemma 3.7 implies that there are at most

$n^{1/2} \cdot (2L \log n)^{2L}$ terms which contribute a value $x_S/(n^{1/4}L)$. Thus if we pick $(2L \log n)^{2L} < n^{1/4}$, we have the bound that the sum is at most $n^{1/2}x_S$, as desired.

Thus we have verified each of the constraints (1)-(3). This completes the proof of Theorem 3.1.

3.4 Gaps for the mixed hierarchy (SA+).

Consider the relaxation SA_t described in (1)-(2), along with the constraint: $Z = (x_{ij})_{1 \le i,j \le n} \succeq 0$. The solution considered earlier (Equation (5)) turns out to also satisfy this PSD condition with high probability. The entries of Z are

$$Z_{ij} = \begin{cases} n^{-1/2}/L \text{ if } i = j \\ n^{-\frac{3}{4}}/L^2 \text{ if } (i,j) \in E(G) \\ n^{-1}/L^2 \quad \text{otherwise} \end{cases}$$

Thus we have

$$Z = \frac{1}{L} \cdot \Big[\frac{1}{nL} J + \frac{1}{n^{1/2}} I + \frac{1}{n^{3/4}L} A\Big],$$

where A is the adjacency matrix of G. Now A is a $\mathcal{G}(n,p)$ matrix with $p = n^{-1/2} \log n$. Thus the least eigenvalue is at least $-2\sqrt{np(1-p)}$ with high probability (by the Semicircle law). This is at least $-4n^{1/4}(\log n)^{1/2}$. Thus we have $A + 4n^{1/4}\sqrt{\log n} I \succeq 0$. Using the fact that $J \succeq 0$, we obtain that $Z \succeq 0$.

This shows that adding an SDP constraint at the *first level* does not give us any additional power – the relaxation obtained after $\Omega(\frac{\log n}{\log \log n})$ levels also has an integrality gap of $\tilde{\Omega}(n^{1/4})$.

Question 3.8. *Does Theorem 3.1 hold for $L \gg \log n$?*

We conjecture that even $L \approx n^{\varepsilon}$ levels does not reduce the integrality gap substantially. We need a different approach (involving a better argument for bounding the number of trees) to extend the arguments above to this range of L.

4 Integrality Gap for the Lasserre hierarchy

In this section, we show a gap instance with arbitrary large constant ratio for linear-round Lasserre relaxation, and a gap instance with n^{ε} ratio for $n^{1-O(\varepsilon)}$-round Lasserre relaxation (Theorem 4.7). We also aim at maximizing the ratio of a polynomial-round Lasserre gap instance, getting a ratio of $\Omega(n^{2/53-\varepsilon})$ (Theorem 4.8).

Our construction is based on a variant of Tulsiani's gap instance for Max K-CSP [Tul09] – we extend the parameter range of Tulsiani's instance. Then we convert the Max K-CSP instance to a constraint-variable graph and duplicate the variable vertices, which is our gap instance for Densest k-subgraph. Note that the gap for Max K-CSP problem is indeed a set of random instances. The vector solution from Lasserre gap for Max K-CSP will help us exhibit a good Lasserre vector solution for Densest k-subgraph. We finally use the structure of random instances of Max K-CSP to show the soundness holds with high probability.

Now, let us proceed to the first step, the gap instance for Max K-CSP.

4.1 Lasserre Gap for Max K-CSP from [Tul09].

We start by defining the Max K-CSP problem.

Definition 4.1. *Let $C \subseteq \mathbb{F}_q^K$ be a q-ary linear code of block length K.*

1. *An instance Φ of Max K-CSP(C) is a set of constraints $C_1, C_2, \cdots, C_m$ where each constraint C_i is over a K-tuple $T_i = (x_{i_1}, x_{i_2}, \cdots, x_{i_K})$, and is of the form $(x_{i_1} + b_1^{(i)}, x_{i_2} + b_2^{(i)}, \cdots, x_{i_K} + b_K^{(i)}) \in C$ for some $b^{(i)} \in \mathbb{F}_q^K$.*
2. *A random instance of Max K-CSP(C) is sampled by choosing each constraint C_i independently, where we sample K variables without replacement from $[n]$ to get $T_i = (x_{i_1}, x_{i_2}, \cdots, x_{i_K})$ and $b^{(i)}$ is chosen from $\mathbb{F}_q^K$ uniformly.*

The following theorem is an extension of the main theorem in [Tul09], showing that polynomial-round Lasserre relaxation cannot refute random Max K-CSP with high probability.

Theorem 4.2. *If C is the dual code of a distance $2\delta \geq 3$ code (in terms of number of coordinates, not fractional distance), for every $10 \leq K < n^{1/2}$, if $n^{\kappa-1} \leq \eta \leq 1/(10^8 \cdot (\beta K^{2\delta+0.75})^{1/(\delta-1)})$ for some $\kappa > 0$, then for large enough n, a random instance Φ of Max K-CSP(C) over $m = \beta n$ constraints and n variables, with probability $1-o(1)$, admits a perfect solution for the SDP relaxation obtained by $\eta n/16$ rounds of the Lasserre hierarchy, i.e. there are vectors $\boldsymbol{V}_{(S,\alpha)}$ for all $S \subseteq [n]$ with $|S| \leq \eta n/16$ and all $\alpha : S \to \mathbb{F}_q$, such that*

- *the value of the solution is perfect: $\sum_{i=1}^m \sum_{\alpha: T_i \to \mathbb{F}_q} C_i(\alpha) \left\|\boldsymbol{V}_{(T_i,\alpha)}\right\|^2 = m$;*
- *$\langle \boldsymbol{V}_{(S_1,\alpha_1)}, \boldsymbol{V}_{(S_2,\alpha_2)} \rangle \geq 0$ for all $S_1, S_2, \alpha_1, \alpha_2$;*
- *$\langle \boldsymbol{V}_{(S_1,\alpha_1)}, \boldsymbol{V}_{(S_2,\alpha_2)} \rangle = 0$ if $\alpha_1(S_1 \cap S_2) \neq \alpha_2(S_1 \cap S_2)$;*
- *$\langle \boldsymbol{V}_{(S_1,\alpha_1)}, \boldsymbol{V}_{(S_2,\alpha_2)} \rangle = \langle \boldsymbol{V}_{(S_3,\alpha_3)}, \boldsymbol{V}_{(S_4,\alpha_4)} \rangle$ for all $S_1 \cup S_2 = S_3 \cup S_4$ and $\alpha_1 \circ \alpha_2 = \alpha_3 \circ \alpha_4$;*
- *$\left\|\boldsymbol{V}_{(\emptyset,\emptyset)}\right\|^2 = 1$ and $\sum_{j \in \mathbb{F}_q} \left\|\boldsymbol{V}_{(\{i\},\{x_i \to j\})}\right\|^2 = 1$ for all $i \in [n]$.*

Note that Theorem 4.2 extends the original theorem of [Tul09] to the regime where K might be superconstant (even poly(n)). The proof of Theorem 4.2 follows the proof in Tulsiani's paper, with the following changes.

Observation 4.3. *By the first property of the solution given in Theorem 4.2, we know that for every $i \in [m]$, we have $\sum_{\alpha: T_i \to \mathbb{F}_q} C_i(\alpha) \left\|\boldsymbol{V}_{T_i,\alpha}\right\|^2 = 1$, and therefore $\sum_{\alpha: T_i \to \mathbb{F}_q} C_i(\alpha) \boldsymbol{V}_{(T_i,\alpha)} = \boldsymbol{V}_{(\emptyset,\emptyset)}$.*

Recall that Tulsiani showed that, if the constraint-variable graph of a Max K-CSP(C) instance has very high left-expansion, then the Lasserre SDP admits a perfect solution for it. Formally, the following lemma is (implicitly) shown in [Tul09].

Lemma 4.4 ([Tul09])**.** *Given a Max K-CSP(C) instance, if every set of constraints of cardinality $s \leq r$ involves more than $(K-\delta)s$ variables (where 2δ is the distance of the dual code of C), and if $4\delta \leq K$, then there is a perfect solution for the SDP relaxation obtained by $r/16$ rounds of the Lasserre hierarchy.*

Hence, we only need to prove the following lemma which shows that the constraint-variable graph still has very high left-expansion, even when a constraint might involve superconstant many variables (i.e. the left degree might be superconstant).

Lemma 4.5. *Given β, η, K as in Theorem 4.2, with probability $1 - o(1)$, for all $2 \leq s \leq \eta n$, every set of s constraints involves more than $(K - \delta)s$ variables.*

A similar lemma can be found in [Tul09] (Lemma A.1), which only deals with constant K. We need a more refined argument for superconstant K, which is in Section 4.4.

4.2 The Lasserre gap for Densest k-subgraph.

The gap instance is reduced from the gap instance for Max K-CSP in Theorem 4.2. Let C be the dual code of a $[K, K - t, 2\delta]_q$ code as used in Theorem 4.2, where K is the block length, $(K - t)$ is the dimension, and $2\delta \geq 3$ is the distance of the code. Such a code has size $|C| = q^t$, and is very sparse for small enough t. For $1000 < q$ and $K > q^2$, we let $\beta = (40q^{t+2} \ln q)/K$, and do the following reduction.

Given a Max K-CSP(C) instance Φ with $m = \beta n$ constraints and n variables. Let $G_\Phi = (L_\Phi, R_\Phi, E_\Phi)$ be the bipartite graph with $m|C|$ left vertices and nq right vertices. For every constraint C_i and every partial assignment to variables in the corresponding tuple T_i which satisfies the constraint C_i, we introduce a left vertex. For every variable x_i and its corresponding assignment, we introduce a right vertex. Formally,

$$\begin{aligned} L_\Phi &= \{(C_i, \alpha) | i \in [m], \alpha : T_i \to \mathbb{F}_q, C_i(\alpha) = 1\}, \\ R_\Phi &= \{(x_j, \alpha) | j \in [n], \alpha : \{x_j\} \to \mathbb{F}_q\}. \end{aligned}$$

We connect a left vertex (C_i, α) and right vertex (x_j, α') when $x_j \in T_i$ and α' is consistent with α, i.e.

$$E_\Phi = \{\{(C_i, \alpha), (x_j, \alpha')\} | (C_i, \alpha) \in L_\Phi, x_j \in T_i, \alpha'(x_j) = \alpha(x_j)\}.$$

Now we define the final graph $G'_\Phi = (L_\Phi, R'_\Phi, E'_\Phi)$ in which we want to find a dense k-subgraph where $k = 2m$. We take β copies of the right vertices in R_Φ to get R'_Φ. To get E'_Φ, we connect a left vertex $u \in L_\Phi$ and a right vertex $v \in R'_\Phi$ if u is connected to v's corresponding vertex in R_Φ in E_Φ. The graph G'_Φ has $N = m|C| + \beta nq = O(nq^{2t+2} \ln q/K)$ vertices.

In our analysis of the reduction, we need a q-ary linear code $\mathcal{C}$ that has a small constant distance (but no less than 3), small block length (but more than q), and very high dimension. Thus, we instantiate the code $\mathcal{C}$ with Generalized BCH codes given by the following.

Lemma 4.6 (Generalized BCH Codes)**.** *For every prime tower q, and integer $2\delta \geq 3$, there are q-ary linear codes of block length $K = q^2 - 1$, dimension $(K - 4\delta + 3)$, and distance at least 2δ.*

We include a simple proof of Lemma 4.6 as follows.

Proof. Let γ be a primitive element of $\mathbb{F}_{q^2}$. Let $D = 2\delta$ for notational ease. We construct the following code

$$\begin{aligned} \tilde{C} = \{(c_1, c_2, \cdots, c_{q^2-1}) \in \mathbb{F}_q^{q^2-1} | c(1) = c(\gamma) = c(\gamma^2) = \cdots = c(\gamma^{D-2}) = 0, \\ \text{where } c(X) = c_1 X + c_2 X^2 + c_3 X^3 + \cdots + c_{q^2-1} X^{q^2-1}\}. \end{aligned}$$

We first show that the distance of $\tilde{C}$ is at least D. Since $\tilde{C}$ is a linear code, we only need to show that every non-zero codeword has weight at least D.

We show the contrapositive statement : the only codeword of weight at most $D-1$ is $\mathbf{0}$. For every codeword of weight at most $D-1$, suppose the non-zero entries are in the set $\{c_{i_1}, c_{i_2}, c_{i_3}, \cdots, c_{i_{D-1}}\}$, we have

$$\begin{array}{rcrcrcrcr}
c_{i_1} & + & c_{i_2} & + & c_{i_3} & + & \cdots & + & c_{i_{D-1}} & =0 \\
\gamma^{i_1} c_{i_1} & + & \gamma^{i_2} c_{i_2} & + & \gamma^{i_3} c_{i_3} & + & \cdots & + & \gamma^{i_{D-1}} c_{i_{D-1}} & =0 \\
\gamma^{2i_1} c_{i_1} & + & \gamma^{2i_2} c_{i_2} & + & \gamma^{2i_3} c_{i_3} & + & \cdots & + & \gamma^{2i_{D-1}} c_{i_{D-1}} & =0 \\
 & & & & \vdots & & & & & \\
\gamma^{(D-2)i_1} c_{i_1} & + & \gamma^{(D-2)i_2} c_{i_2} & + & \gamma^{(D-2)i_3} c_{i_3} & + & \cdots & + & \gamma^{(D-2)i_{D-1}} c_{i_{D-1}} & =0
\end{array}$$

Note that the coefficients form a Vandermonde matrix (which has full rank). Therefore we have $c_{i_1} = c_{i_2} = c_{i_3} = \cdots = c_{i_{D-1}} = 0$, i.e. the codeword is $\mathbf{0}$.

Now we show that the dimension of $\tilde{C}$ is at least $(K - 2D + 3)$. Note that each constraint $c(\gamma^i) = 0 (1 \leq i \leq D-2)$ can be implemented by 2 linear constraints in $\mathbb{F}_q$ (since $\gamma^i \in \mathbb{F}_{q^2}$), while the constraint $c(1) = 0$ is indeed a linear constraint in $\mathbb{F}_q$. Therefore, we need at most $2(D-2)+1 = 2D-3$ linear constraints for $\tilde{C}$, i.e. the dimension of $\tilde{C}$ is at least $(K - 2D + 3)$.

Finally, if the dimension of $\tilde{C}$ is more than $(K - 2D + 3)$, we can take a linear subspace of $\tilde{C}$ of dimension $(K - 2D + 3)$, while the distance of the subspace code is no less than the distance of $\tilde{C}$. □

4.3 Analysis.

We get a family of gap instances G'_Φ parameterized by $q > 1000$ and $2\delta \geq 3$ (using Lemma 4.6). We obtain our two main results of this section by picking appropriate parameters for code C as follows. To get lasserre integrality gaps for $N^{1-O(\epsilon)}$ levels , we show the following by setting the distance $2\delta = 3$.

Theorem 4.7. *For every $1000 < q < N^\epsilon$ (where ϵ is an absolute small constant), there is a gap instance of ratio $\Omega(q)$ for $N/q^{O(1)}$-level Lasserre SDP. The same construction also works for the Min degree Lasserre SDP, when $q = \Omega(\log n)$ and $q < N^\epsilon$.*

We now aim at getting a gap instance of ratio N^ϵ for polynomial-round Lasserre SDP, where ϵ is maximized. By setting $q = n^\gamma$ for some small constant $\gamma > 0$, the distance $2\delta = 4$, and optimizing the other parameters, we obtain the following (refer to section 4.3 for details)

Theorem 4.8. *For small enough $\kappa > 0$, there is a gap instance of ratio $N^{2/53-O(\kappa)}$ for the N^κ-round Min degree Lasserre SDP.*

The two theorems follow because of Theorem 4.2, Lemma 4.9, Lemma 4.11 (completeness) and Lemma 4.12 (soundness). In the completeness case, we will use our r-level Lasserre solution for Max K-CSP to show that the Lasserre SDP after $R = r/K$ levels of the hierarchy has value at least $\beta m K$. In the soundness case, we show that with probability $1 - o(1)$, the graph G'_Φ does not have any $2m$-subgraph of value more than $17/q$ times the SDP value (Lemma 4.12). Therefore, the graph G'_Φ is a gap instance of ratio $\Omega(q)$ for R-round Lasserre SDP. We proceed by first proving these lemmas.

Completeness.

Lemma 4.9. *If the* Max K-CSP(C) *instance* Φ *admits a perfect solution for* r*-round Lasserre SDP relaxation, then the* r/K*-round Lasserre SDP relaxation for the* Densest k-subgraph *instance* G'_Φ *has a solution of value* βmK*.*

Proof. For any set $S \subseteq L_\Phi \cup R'_\Phi$, suppose the left vertices included in S are

$$(C_{i_1}, \alpha_1), (C_{i_2}, \alpha_2), \cdots, (C_{i_{r_1}}, \alpha_{r_1}),$$

and the right vertices included in S are

$$(x_{j_1}, \alpha'_1), (x_{j_2}, \alpha'_2), \cdots, (x_{j_{r_2}}, \alpha'_{r_2}),$$

where $r_1 + r_2 \leq r/K$. Let

$$S' = T_{i_1} \cup T_{i_2} \cup \cdots \cup T_{i_{r_1}} \cup \{x_{j_1}\} \cup \{x_{j_2}\} \cup \cdots \cup \{x_{j_{r_2}}\}.$$

We have $|S'| \leq Kr_1 + r_2 \leq r$. If all the partial assignments α_i's and α'_i's are consistent to each other (i.e. there are not two of them assigning the same variable to different values), we can define

$$\alpha = \alpha_1 \circ \alpha_2 \circ \cdots \alpha_{r_1} \circ \alpha'_1 \circ \alpha'_2 \circ \cdots \alpha'_{r_2},$$

and let $\boldsymbol{U}_S = \boldsymbol{V}_{(S',\alpha)}$, or we let $\boldsymbol{U}_S = \boldsymbol{0}$.

Observation 4.10. *For every* $S \subseteq L_\Phi \cup R'_\Phi$*, we have* $\langle \boldsymbol{V}_{(\emptyset,\emptyset)}, \boldsymbol{U}_S \rangle = \|\boldsymbol{U}_S\|^2$*.*

Proof. If $\boldsymbol{U}_S = \boldsymbol{V}_{(S',\alpha)}$ for some S', α, we have $\langle \boldsymbol{V}_{(\emptyset,\emptyset)}, \boldsymbol{U}_S \rangle = \langle \boldsymbol{V}_{(\emptyset,\emptyset)}, \boldsymbol{V}_{(S',\alpha)} \rangle = \left\| \boldsymbol{V}_{(S',\alpha)} \right\|^2 = \|\boldsymbol{U}_S\|^2$. If $\boldsymbol{U}_S = \boldsymbol{0}$, we have $\langle \boldsymbol{V}_{(\emptyset,\emptyset)}, \boldsymbol{U}_S \rangle = \|\boldsymbol{U}_S\|^2 = 0$. □

We can check that all the Lasserre constraints are satisfied.

- For two sets S_1, S_2, either at least one of the vectors $\boldsymbol{U}_{S_1}, \boldsymbol{U}_{S_2}$ is $\boldsymbol{0}$ (therefore their inner-product is 0), or $\boldsymbol{U}_{S_1} = \boldsymbol{V}_{S'_1,\alpha_1}, \boldsymbol{U}_{S_2} = \boldsymbol{V}_{S'_2,\alpha_2}$ for some $S'_1, S'_2, \alpha_1, \alpha_2$ and $\langle \boldsymbol{U}_{S_1}, \boldsymbol{U}_{S_2} \rangle = \langle \boldsymbol{V}_{S'_1,\alpha_1}, \boldsymbol{V}_{S'_2,\alpha_2} \rangle \geq 0$.
- For any S_1, S_2, S_3, S_4 such that $S_1 \cup S_2 = S_3 \cup S_4$, either the set of partial assignments in $S_1 \cup S_2 = S_3 \cup S_4$ are consistent to each other, in which case we have $\boldsymbol{U}_{S_1 \cup S_2} = \boldsymbol{U}_{S_3 \cup S_4} = \boldsymbol{V}_{S,\alpha}$ where S is the union of all the variables included in $S_1 \cup S_2$ and α is the concatenation of the partial assignments in $S_1 \cup S_2$; or we have $\boldsymbol{U}_{S_1 \cup S_2} = \boldsymbol{U}_{S_3 \cup S_4} = \boldsymbol{0}$.

- For every set S, we have

$$\begin{aligned}
&\sum_{v\in L_\Phi\cup R'_\Phi}\langle \boldsymbol{U}_{\{v\}},\boldsymbol{U}_S\rangle\\
=&\sum_{(C_i,\alpha)\in L_\Phi}\langle \boldsymbol{U}_{\{(C_i,\alpha)\}},\boldsymbol{U}_S\rangle+\sum_{(x_j,\alpha)\in R'_\Phi}\langle \boldsymbol{U}_{\{(x_j,\alpha)\}},\boldsymbol{U}_S\rangle\\
=&\sum_{(C_i,\alpha)\in L_\Phi}\langle \boldsymbol{V}_{\{(T_i,\alpha)\}},\boldsymbol{U}_S\rangle+\sum_{(x_j,\alpha)\in R'_\Phi}\langle \boldsymbol{V}_{(\{x_j\},\alpha)},\boldsymbol{U}_S\rangle\\
=&\sum_{i=1}^{m}\sum_{\alpha:T_i\to\mathbb{F}_q}C_i(\alpha)\langle \boldsymbol{V}_{\{(T_i,\alpha)\}},\boldsymbol{U}_S\rangle+\beta\sum_{j=1}^{n}\sum_{\alpha:\{x_j\}\to\mathbb{F}_q}\langle \boldsymbol{V}_{(\{x_j\},\alpha)},\boldsymbol{U}_S\rangle\\
=&\sum_{i=1}^{m}\langle V_{(\emptyset,\emptyset)},U_S\rangle+\beta\sum_{j=1}^{n}\sum_{\alpha:\{x_j\}\to\mathbb{F}_q}\langle V_{(\emptyset,\emptyset)},U_S\rangle\\
=&(m+\beta n)\left\|U_S\right\|^2=2m\left\|U_S\right\|^2,
\end{aligned}$$

  where the third last equality is because of Observation 4.3, and the second last equality is because of Observation 4.10.

- Finally, we have $\|\boldsymbol{U}_\emptyset\|^2=\left\|\boldsymbol{V}_{(\emptyset,\emptyset)}\right\|^2=1$.

Now, we calculate the value of the solution

$$\begin{aligned}
&\sum_{(u,v)\in E'_\Phi}\left\|\boldsymbol{U}_{\{u,v\}}\right\|^2\\
=&\beta\sum_{(u,v)\in E_\Phi}\left\|\boldsymbol{U}_{\{u,v\}}\right\|^2\\
=&\beta\sum_{i=1}^{m}\sum_{\alpha:T_i\to\mathbb{F}_q,C_i(\alpha)=1}\sum_{x_j\in T_i}\left\|\boldsymbol{U}_{\{(C_i,\alpha),(x_j,\alpha_{|\{x_j\}})\}}\right\|^2\\
=&\beta\sum_{i=1}^{m}\sum_{\alpha:T_i\to\mathbb{F}_q,C_i(\alpha)=1}K\left\|\boldsymbol{V}_{(T_i,\alpha)}\right\|^2\\
=&\beta\sum_{i=1}^{m}K=\beta mK.
\end{aligned}$$

□

If we add the constraint (4), we can still get a good SDP solution for the Min degree Lasserre SDP with high probability, as long as q is superconstant.

Lemma 4.11. *For $q=\Omega(\log n)$, with probability $1-o(1)$, this vector solution also satisfies the added constraint (4) with $d=\beta K/2$, i.e., for every set S, for each vertex u, we have*

$$\forall S,\forall u\in V,\quad \sum_{v\in\Gamma(u)}\langle \boldsymbol{U}_{\{u,v\}},\boldsymbol{U}_S\rangle\geq \beta K/2\cdot\langle \boldsymbol{U}_{\{u\}},\boldsymbol{U}_S\rangle.$$

Proof. For each left vertex (C_i, α), we have

$$\begin{aligned}
&\sum_{v\in\Gamma((C_i,\alpha))} \langle \boldsymbol{U}_{\{(C_i,\alpha),v\}}, \boldsymbol{U}_S\rangle \\
=&\beta \sum_{x_j\in T_i} \langle \boldsymbol{U}_{\{(C_i,\alpha),(x_j,\alpha_{|\{x_j\}})\}}, \boldsymbol{U}_S\rangle \\
=&\beta \sum_{x_j\in T_i} \langle \boldsymbol{U}_{\{(C_i,\alpha)\}}, \boldsymbol{U}_S\rangle = \beta K \langle \boldsymbol{U}_{\{(C_i,\alpha)\}}, \boldsymbol{U}_S\rangle.
\end{aligned}$$

For each right vertex (x_j, α'), we have

$$\begin{aligned}
&\sum_{v\in\Gamma((x_j,\alpha'))} \langle \boldsymbol{U}_{\{(x_j,\alpha'),v\}}, \boldsymbol{U}_S\rangle \\
=& \sum_{i:T_i\ni x_j} \sum_{\alpha:T_i\to\mathbb{F}_q, C_i(\alpha)=1, \alpha(x_j)=\alpha'(x_j)} \langle \boldsymbol{U}_{\{(x_j,\alpha'),(C_i,\alpha)\}}, \boldsymbol{U}_S\rangle \\
=& \sum_{i:T_i\ni x_j} \sum_{\alpha:T_i\to\mathbb{F}_q, C_i(\alpha)=1, \alpha(x_j)=\alpha'(x_j)} \langle \boldsymbol{U}_{\{(C_i,\alpha)\}}, \boldsymbol{U}_S\rangle \\
=& \sum_{i:T_i\ni x_j} \sum_{\alpha:T_i\to\mathbb{F}_q, \alpha(x_j)=\alpha'(x_j)} \langle \boldsymbol{U}_{\{(C_i,\alpha)\}}, \boldsymbol{U}_S\rangle,
\end{aligned}$$

where the last equality is because we know that $\boldsymbol{U}_{\{(C_i,\alpha)\}} = \boldsymbol{0}$ when $C_i(\alpha) \neq 1$. By the property of Lasserre vectors, we know that for each $i \in [m]$,

$$\sum_{\alpha:T_i\to\mathbb{F}_q, \alpha(x_j)=\alpha'(x_j)} \boldsymbol{U}_{\{(C_i,\alpha)\}} = \boldsymbol{U}_{\{(x_j,\alpha')\}},$$

therefore

$$\sum_{v\in\Gamma((x_j,\alpha'))} \langle \boldsymbol{U}_{\{(x_j,\alpha'),v\}}, \boldsymbol{U}_S\rangle = \sum_{i:T_i\ni x_j} \langle \boldsymbol{U}_{\{(x_j,\alpha')\}}, \boldsymbol{U}_S\rangle.$$

For $q = \Omega(\log n)$, the expected number of constraints containing x_j is $\beta K = \Omega((\log n)^{t+2}) = \Omega(\log n)$, by our choice of β. Therefore, by Chernoff bound and union bound, with probability $1 - o(1)$, for all x_j, there are at least $\beta K/2$ constraints containing x_j, and for every S, x_j and α', we have

$$\sum_{v\in\Gamma((x_j,\alpha'))} \langle \boldsymbol{U}_{\{(x_j,\alpha'),v\}}, \boldsymbol{U}_S\rangle \geq \beta K/2 \cdot \langle \boldsymbol{U}_{\{(x_j,\alpha')\}}, \boldsymbol{U}_S\rangle.$$

□

Soundness.

Now, we show that random instances of Max K-CSP give rise to graphs G'_φ whose $2m$-sized subgraphs have density $O(\beta K/q)$. Note that the large alphabet size q allows us to get a much larger gap than we would starting from random AND instances [Fei02]. This allows us some slack in the size of the subgraphs we need to argue about.

For C the dual of a $[K, K-t, 2\delta]_q$ code, we prove the following soundness lemma.

Lemma 4.12. *When $\beta \geq (40q^{t+2}\ln q)/K$, for a random* Max K-CSP(C) *instance Φ, with probability $1-o(1)$, any subgraph of G'_Φ obtained by choosing $2m$ left vertices and $2m$ right vertices contains at most $17\beta mK/q$ edges, and therefore any $2m$-subgraph of G'_Φ contains at most $17\beta mK/q$ edges.*

Note that G'_φ was constructed by taking β copies of the right bipartition and replicating the edges. To prove Lemma 4.12, we only need to prove the following lemma.

Lemma 4.13. *Suppose that $q > 1000, K > q^2/2, t \leq 10$. When $\beta \geq (40q^{t+2}\ln q)/K$, for a random* Max K-CSP(C) *instance Φ, with probability $1-o(1)$, any subgraph of G_Φ obtained by choosing $2m$ left vertices and $2n$ right vertices contains at most $17mK/q$ edges.*

Proof of Lemma 4.12 from Lemma 4.13. We only need to prove once there is a $2m \times 2m$ subgraph of G'_Φ with t edges, there is a $2m \times 2n$ subgraph of G_Φ with at least t/β edges. Fix $2m$ left vertices in G'_Φ, to maximize the number of edges in the subgraph, we need to select the $2m$ right vertices with most edges connected to the chosen $2m$ left vertices. Since any two right vertices G'_Φ corresponding to the same right vertex in G_Φ have the same set of neighbors, there is an densest $2m \times 2m$ subgraph H' of G'_Φ that, for any two such vertices, chooses either both or neither of them. Now we define an subgraph H of G_Φ that contains the same $2m$ left vertices. It contains a right vertex if any copy of the vertex is contained in H'. H contains $2m/\beta = 2n$ vertices, and it is easy to see that there are (at least) t/β edges in H. □

We proceed by fixing a set of $2n$ vertices R on the right. Lemma 4.13 follows from the following claim by a standard union bound over all possible choices of R.

Claim 4.14. *Recall that $G_\Phi = (L_\Phi, R_\Phi, E_\Phi)$. Suppose that $q > 1000, K > q^2/2, t \leq 10$. Fix a subset $R \subseteq R_\Phi$ (note that R_Φ is the same for all the instances Φ of n variables), $|R| = 2n$, the probability (over choice of Φ) that there does not exist a subset $L \subseteq L_\Phi$ of size $2m$ such that the number of edges in the induced subgraph by $L \cup R$ is more than $17mK/q$, is at least $1 - \exp(-mK/(10q^{t+2}))$.*

Proof of Lemma 4.13 from Claim 4.14. Since there are only $\binom{qn}{2n} \leq \exp(2n(\ln q + 1))$ choices of R, by a union bound, with probability at least

$$\begin{aligned} &1 - \exp(2n(\ln q + 1)) \cdot \exp(-mK/(10q^{t+2})) \\ &\qquad = 1 - \exp(2n(\ln q + 1) - \beta nK/(10q^{t+2})), \end{aligned}$$

there is no $2m \times 2n$ subgraph of G_Φ containing more than $17mK/q$ edges. The probability becomes $1 - o(1)$ when $\beta = (40q^{t+2}\ln q)/K$. □

Proof. (Proof of Claim 4.14) First, we show that with high probability, a constraint C_i is "poorly satisfied". That is, none of the left vertices corresponding to a constraint C_i has more than $\Omega(K/q)$ neighbors in R – this number is roughly $1/q$ times the corresponding value in completeness case. We prove this in the following two steps.

Step 1. Fix a subset $R \subseteq R_\Phi$, $|R| = 2n$, for each variable x_j, let $\deg(x_j)$ be the number of vertices in R that corresponding to x_j, i.e. let $\deg(x_j) = |R \cap \{(x_j, \alpha) | \alpha : \{x_j\} \to \mathbb{F}_q\}|$. For a subset of variables $T \subseteq \{x_1, x_2, \cdots, x_n\}$, let $\deg(T) = \sum_{x_j \in T} \deg(x_j)$. We call T *good* if the average degree of variables in T is not more than 4, i.e. $\deg(T) \leq 4|T|$.

For a random T with $|T| = K$, note that the expected degree $\mathbf{E}[\deg(T)] = 2K$. Therefore, by Hoeffding's inequalities for sampling without replacement (Theorem 1 and Theorem 4 in [Hoe62]), we have

$$\Pr[T \text{ is not good}] = \Pr[\deg(T) > 4K] < \exp(-\ln(4/e) \cdot 2K/q) < \exp(-K/(2q)).$$

Step 2. Again, fix $R \subseteq R_\Phi$, $T \subseteq \{x_1, x_2, \cdots, x_n\}$, for a codeword α on coordinates in T, i.e. $\alpha : T \to \mathbb{F}_q$, let $\mathrm{agr}_T(\alpha, R) = |\{(x_j, \alpha_{|x_j})|x_j \in T\} \cap R|$. For a constraint C_i, say it is *poorly satisfied* if for all $\alpha : T \to \mathbb{F}_q$ such that $C_i(\alpha) = 1$, we have $\mathrm{agr}_{T_i}(\alpha, R) \leq 8K/q$.

Recall that to sample a random constraint C_i, we first sample a random K-tuple T_i, and a random shifting function $b^{(i)}$. Note that for a fixed $\alpha : T \to \mathbb{F}_q$, and a fixed T_i that is good, when we take a random shifting function $b^{(i)} : T_i \to \mathbb{F}_q$, we have $\mathbf{E}_{b^{(i)}:T_i \to \mathbb{F}_q}[\mathrm{agr}_{T_i}(\alpha - b^{(i)}, R)] = \deg(T_i)/q = 4K/q$, therefore, by standard Chernoff bound, for a fixed codeword $\alpha \in C$, the probability that α makes C_i not poorly satisfied is bounded from above by

$$\Pr_{b^{(i)}}[\mathrm{agr}_{T_i}(\alpha - b^{(i)}, R) > 8K/q] < \exp(-\ln(4/e) \cdot 4K/q) < \exp(-K/q).$$

Since there are $|C| = q^t \leq q^{10}$ codewords, by a union bound, for $K > q^2/2$ and $q > 1000$, we have

$$\Pr[C_i \text{ is not poorly satisfied}|T_i \text{ is good}] < q^{10} \cdot \exp(-K/q) < \exp(-K/(2q)).$$

In all, we have

$$\begin{aligned}
&\Pr[C_i \text{ is poorly satisfied}] \\
\geq &\Pr[C_i \text{ is poorly satisfied}|T_i \text{ is good}] \cdot \Pr[T \text{ is good}] \\
> &(1 - \exp(-K/q))(1 - \exp(-K/(2q))) \\
> &1 - \exp(-K/(3q)).
\end{aligned}$$

Now, again, by standard Chernoff bound, we have

$$\begin{aligned}
&\Pr[|\{C_i|C_i \text{ is not poorly satisfied}\}| > m/(q \cdot |C|)] \\
< &(e \cdot |C| \cdot q \cdot \exp(-K/(3q)))^{m/(q \cdot |C|)} \\
\leq &\left(e \cdot q^{10} \cdot q \cdot \exp(-K/(3q))\right)^{m/(q \cdot |C|)} \\
< &\exp(-K/(10q))^{m/(q \cdot |C|)} \\
= &\exp(-mK/(10q^{t+2})).
\end{aligned}$$

By the calculation above we know that with probability at least $1 - \exp(-mK/(10q^{t+2}))$, there are at most $m/(q \cdot |C|)$ constraints that are not poorly satisfied.

For each left vertex $(C_i, \alpha) \in L_\Phi$, if C_i is poorly satisfied, we know there are at most $8K/q$ edges from (C_i, α) to R. If C_i is not poorly satisfied, there are at most K edges to R_Φ – this upperbound also applies to R.

Therefore, with probability at least $1 - \exp(-mK/(10q^{t+2}))$, any set of $2m$ left vertices has at most $2m \cdot 8K/q + m/(q \cdot |C|) \cdot |C| \cdot K \leq 17mK/q$ edges connected to R. □

We now complete the proofs of the main theorems in this section.

Proof of Theorem 4.7.

By combining Theorem 4.2, Lemma 4.9, Lemma 4.11 (completeness), and Lemma 4.12 (soundness) we see that with probability $1 - o(1)$, the graph G'_Φ provides a $\Omega(q)$ integrality for the number of levels R given by

$$\begin{aligned} R &= \Omega\left(\frac{n}{K(\beta K^{2\delta+0.75})^{1/(\delta-1)}}\right) \\ &= \Omega\left(\frac{N}{q^{2t+2}\ln q \cdot (\beta K^{2\delta+0.75})^{1/(\delta-1)}}\right) \\ &= \Omega\left(\frac{N}{K^{(2\delta-0.25)/(\delta-1)} q^{(2t+2)+(t+2)/(\delta-1)} \mathrm{poly}\log q}\right). \\ &\qquad\qquad \text{(recall that } \beta = (40q^{t+2}\ln q)/K) \end{aligned}$$

Recall that $K = q^2 - 1$. By setting $K = q^2 - 1$ and $2\delta = 3$, we verify that the theorem holds. □

Proof of Theorem 4.8.

Let $q = n^\gamma$, since $N = O(nq^{2t+2}\ln q/K) = O(nq^{2t}\ln q)$, ratio of the gap due to Lemma 4.9 and Lemma 4.12 is

$$\Omega(q) = \Omega(N^{\gamma/(1+2t\gamma+o(1))}) = \Omega(N^{\gamma/(1+(8\delta-6)\gamma+o(1))}). \qquad \text{(note that } t = 4\delta - 3)$$

This means that

$$\epsilon = \frac{\gamma}{1+(8\delta-6)\gamma+o(1)} = \frac{1}{8\delta-6} - \frac{1+o(1)}{(1+(8\delta-6)\gamma+o(1))(8\delta-6)}.$$

Note that when $2\delta \geq 3$ is fixed, ϵ is maximized when γ is maximized.

The number of rounds (due to Theorem 4.2 and Lemma 4.9) is

$$\begin{aligned} R &= \Omega\left(\frac{N}{K^{(2\delta-0.25)/(\delta-1)} q^{(2t+2)+(t+2)/(\delta-1)} \mathrm{poly}\log q}\right) \\ &= \Omega\left(\frac{N}{n^{2\gamma(2\delta-0.25)/(\delta-1)} \cdot n^{\gamma((8\delta-4)+(4\delta-1)/(\delta-1))+o(1)}}\right) \\ &\qquad\qquad \text{(by } K = \Theta(q^2),\ q = n^\gamma,\ t = 4\delta-3) \\ &= \Omega\left(\frac{N}{n^{\gamma(8\delta+4+6.5/(\delta-1))+o(1)}}\right) \\ &= \Omega\left(N^{1-\frac{\gamma(8\delta+4+6.5/(\delta-1))+o(1)}{1+o(1)+\gamma(8\delta-6)}}\right) \end{aligned}$$

For very small $\kappa > 0$, to get a gap instance for $N^{\Omega(\kappa)}$-round Lasserre, we need

$$\begin{aligned} &1 - \frac{\gamma(8\delta+4+6.5/(\delta-1))+o(1)}{1+o(1)+\gamma(8\delta-6)} \geq \Omega(\kappa) \\ \Rightarrow &1+o(1)+\gamma(8\delta-6) - (\gamma(8\delta+4+6.5/(\delta-1))+o(1)) \geq \Omega(\kappa) \\ \Rightarrow &1 - \gamma(10 - 6.5(\delta-1)) \geq \Omega(\kappa) \\ \Rightarrow &\gamma \leq \frac{1-\Omega(\kappa)}{10-6.5(\delta-1)} \end{aligned}$$

Let $\gamma = \frac{1-O(\kappa)}{10+6.5/(\delta-1)}$, we have

$$\begin{aligned}\epsilon =& \frac{1}{8\delta-6} - \frac{1+o(1)}{(1+(8\delta-6)\gamma+o(1))(8\delta-6)}\\ =& \frac{1}{8\delta-6} - \frac{1+o(1)}{(1+(8\delta-6)\frac{1-O(\kappa)}{10+6.5/(\delta-1)}+o(1))(8\delta-6)}\\ =& \frac{1}{8\delta-6} - \frac{1}{(1+\frac{(8\delta-6)}{10+6.5/(\delta-1)})(8\delta-6)} - O(\kappa).\end{aligned}$$

When $2\delta = 4$, we get the maximized value $\epsilon = 2/53 - O(\kappa)$. □

4.4 Expansion for random Max K-CSP instances.

In this section, we prove Lemma 4.5, restated as follows.

Lemma 4.5 (restated). *Given β, η, K as in Theorem 4.2, with probability $1 - o(1)$, for all $2 \le s \le \eta n$, every set of s constraints involves more than $(K-\delta)s$ variables.*

Proof. Fix $2 \le s \le \eta n$, let us upperbound the probability that there is a set of s constraints containing at most $(K-\delta)s$ variables. Since there are $\binom{\beta n}{s}$ such sets, the probability is at most

$$\begin{aligned}&\binom{\beta n}{s} \Pr[\text{the first } s \text{ constraintscontain at most } (K-\delta)s \text{ variables}]\\ =&\binom{\beta n}{s} \sum_{i=1}^{(K-\delta)s} \Pr[\text{the first } s \text{ constraints contain exactly } i \text{ variables}].\end{aligned}$$

Fix a set T of i variables, let $p(s,i)$ be the number of s-tuples $(T_1, T_2, \cdots, T_s)$ where for each $1 \le j \le s$, T_j is a set of K variables, such that $\cup_{1\le j\le s} T_j = T$. We have

$$\Pr[\text{the first } s \text{ constraints contain exactly } i \text{ variables}] = \binom{n}{i} \cdot p(i,s) \Big/ \binom{n}{K}^s.$$

To upperbound $p(i,s)$, we view the way to enumerating valid $(T_1, T_2, \cdots, T_s)$ as, to choose a multiset of Ks variables (each one from T) so that each element in T appears at least once in the multiset, then view each element in the multiset as a distinct element, and distribute these Ks elements to s sets, in a balanced way. Note that in this way, we are able to enumerate all the valid s-tuples (although some of them might be enumerated more than once). Since there are at most $\binom{Ks-1}{i-1} < \binom{Ks}{i}$ valid multisets, we have

$$p(i,s) \le \binom{Ks}{i} (Ks)!/(K!)^s.$$

Therefore, we have

$$\begin{aligned}&\binom{\beta n}{s} \Pr[\text{the first } s \text{ constraintscontain at most } (K-\delta)s \text{ variables}]\\ =&\binom{\beta n}{s} (Ks)! \cdot (K!)^{-s} \binom{n}{K}^{-s} \sum_{i=1}^{(K-\delta)s} \binom{n}{i}\binom{Ks}{i},\end{aligned}$$

Note that when $K^2 s < \delta n$ and $i \le (K-\delta)s$, we have $i < nKs/(n+Ks)$ (since $i \le Ks(1-\delta/K) \le Ks/(1+\delta/K) = nKs/(n+\delta n/K) < nKs/(n+Ks)$), and therefore

$$\frac{\binom{n}{i}\binom{Ks}{i}}{\binom{n}{i-1}\binom{Ks}{i-1}} = \frac{(n-i)(Ks-i)}{i^2} > 1 \qquad (\Leftarrow (n-i)(Ks-i) > i^2 \Leftarrow nKs > (n+Ks)i),$$

therefore the function $\binom{n}{i}\binom{Ks}{i}$ is increasing when $i \le (K-\delta)s$, therefore

$$\begin{aligned}
&\binom{\beta n}{s}(Ks)! \cdot (K!)^{-s}\binom{n}{K}^{-s}\sum_{i=1}^{(K-\delta)s}\binom{n}{i}\binom{Ks}{i} \\
\le&\binom{\beta n}{s}(Ks)! \cdot (K!)^{-s}\binom{n}{K}^{-s} \cdot Ks \cdot \binom{n}{(K-\delta)s}\binom{Ks}{(K-\delta)s} \\
=&\binom{\beta n}{s}(Ks)! \cdot (K!)^{-s}\binom{n}{K}^{-s} \cdot Ks \cdot \binom{n}{(K-\delta)s}\binom{Ks}{\delta s}.
\end{aligned}$$

for $K \le n^{1/2}$, we use the fact that $\binom{n}{K} \ge (n-K)^K/K! \ge n^K/3/((K/e)^K \cdot (5\sqrt{K})) = (en/K)^K/(15\sqrt{K})$ (since by Stirling's formula, we have $K! \le 5\sqrt{K}(K/e)^K$), and again use the fact that $\sqrt{2\pi K}(K/e)^K \le K! \le 5\sqrt{K}(K/e)^K$, we bound the expression above by

$$\begin{aligned}
&\left(\frac{e\beta n}{s}\right)^s \frac{5\sqrt{Ks}(Ks/e)^{Ks}}{(\sqrt{2\pi K}(K/e)^K)^s} \cdot \left(15\sqrt{K}\left(\frac{K}{en}\right)^K\right)^s \cdot Ks \cdot \left(\frac{en}{(K-\delta)s}\right)^{(K-\delta)s}\left(\frac{eK}{\delta}\right)^{\delta s} \\
\le&5(Ks)^{1.5} \cdot \left(\frac{15e\beta s^{\delta-1}K^{K+\delta}}{\sqrt{2\pi}n^{\delta-1}(K-\delta)^{K-\delta}\delta^\delta}\right)^s \\
\le&5(Ks)^{1.5} \cdot \left(\frac{15e^{1+\delta}\beta s^{\delta-1}K^{2\delta}}{\sqrt{2\pi}n^{\delta-1}\delta^\delta}\right)^s
\end{aligned}$$

For $2 \le s \le \ln^2 n$, since $n^{\kappa-1} \le 1/(10^8 \cdot (\beta K^{2\delta+0.75})^{1/(\delta-1)})$, we have $\beta^2 K^{4\delta+1.5}/n^{2(\delta-1)} \le n^{-(2\delta-1)\kappa}$, we have

$$\begin{aligned}
&5(Ks)^{1.5} \cdot \left(\frac{15e^{1+\delta}\beta s^{\delta-1}K^{2\delta}}{\sqrt{2\pi}n^{\delta-1}\delta^\delta}\right)^s \\
\le&5(Ks)^{1.5} \cdot \left(\frac{15e^{1+\delta}\beta s^{\delta-1}K^{2\delta}}{\sqrt{2\pi}n^{\delta-1}\delta^\delta}\right)^2 \\
\le&\frac{5 \cdot 15e^{1+\delta}s^{\delta-1}}{\sqrt{2\pi}\delta^\delta n^{2(\delta-1)\kappa}} \le O(n^{-(\delta-1)\kappa}).
\end{aligned}$$

For $\ln^2 n < s \le \eta n$, since $\eta \le 1/(10^8 \cdot (\beta K^{2\delta+0.75})^{1/(\delta-1)})$, we get $\eta \le 1/(10^8 \cdot (\beta K^{2\delta})^{1/(\delta-1)})$, and

further we have $\beta K^{2\delta}\eta^{\delta-1} \le \delta^\delta/(100\cdot 15e^{1+\delta}/\sqrt{2\pi})$ for all $\delta > 5/4$. Therefore,

$$\begin{aligned}
&5(Ks)^{1.5}\cdot\left(\frac{15e^{1+\delta}\beta s^{\delta-1}K^{2\delta}}{\sqrt{2\pi}n^{\delta-1}\delta^\delta}\right)^s\\
\le&5(Ks)^{1.5}\left(\frac{s^{\delta-1}}{100(\eta n)^{\delta-1}}\right)^s\\
\le&5\cdot\left(\frac{s^{\delta-1}(Ks)^{(1.5/\ln^2 n)}}{100(\eta n)^{\delta-1}}\right)^s\\
\le&5\cdot\left(\frac{2}{100}\right)^s. \qquad\qquad \text{(by } s\le\eta n \text{ and } Ks\le n^2\text{)}
\end{aligned}$$

Now, we upperbound probability that there exists a set of constraints of size $s \le \eta n$ involving at most $(K-\delta)s$ variables by

$$\begin{aligned}
&\sum_{s=2}^{\eta n} 5(Ks)^{1.5}\cdot\left(\frac{15e^{1+\delta}\beta s^{\delta-1}K^{2\delta+0.5}}{n^{\delta-1}\delta^\delta}\right)^s\\
=&\sum_{s=2}^{\ln^2 n} 5(Ks)^{1.5}\cdot\left(\frac{15e^{1+\delta}\beta s^{\delta-1}K^{2\delta+0.5}}{n^{\delta-1}\delta^\delta}\right)^s+\sum_{s=\ln^2 n+1}^{\eta n} 5(Ks)^{1.5}\cdot\left(\frac{15e^{1+\delta}\beta s^{\delta-1}K^{2\delta+0.5}}{n^{\delta-1}\delta^\delta}\right)^s\\
\le&\sum_{s=2}^{\ln^2 n} O(n^{-\kappa/2})+\sum_{s=\ln^2 n+1}^{\eta n} 5\cdot(1/50)^s = o(1).
\end{aligned}$$

□

5 Conclusion

In this paper, we show integrality gap lower bounds of $\Omega(n^{1/4}/\log^3 n)$ for $\Omega(\log n/\log\log n)$ levels of the Sherali-Adams+ SDP relaxation, and $\Omega(n^{2/53-\epsilon})$ for $n^{\Omega(\epsilon)}$ levels of the Lasserre SDP relaxation for the Densest k-subgraph problem.

The gap instances for SA+ SDP are actually (Erdös-Renyi) random graph instances $\mathcal{G}(n,p)$. We believe these instances should give $\Omega(n^{1/4-\epsilon})$ gaps for even stronger relaxations – in particular higher levels of the Sherali-Adams hierarchy, with stronger SDP constraints. The sub-exponential time algorithms for Densest k-subgraph in [BCC+10] imply that the integrality gap becomes $O(n^{1/4-\varepsilon})$ after $n^{O(\varepsilon)}$ levels of an LP hierarchy which is weaker than the Sherali-Adams hierarchy. In fact, these sub-exponential time algorithms were inspired by attempts to construct integrality gap lower bounds for many levels (polynomial in n). It would be interesting to close this gap by obtaining matching integrality gap lower bounds for $n^{\Omega(\varepsilon)}$ levels. As a further goal, one might also hope to combine the techniques used in both parts of this paper, to get $\Omega(n^{1/4-\epsilon})$ gaps for polynomial levels of the Lasserre hierarchy.

References

[ABBG10] Sanjeev Arora, Boaz Barak, Markus Brunnermeier, and Rong Ge. Computational complexity and information asymmetry in financial products. In *Proceedings of the First Symposium on Innovations in Computer Science (ICS)*, 2010. 1

[ABW08] Benny Applebaum, Boaz Barak, and Avi Wigderson. Public key cryptography from different assumptions. Preliminary version as cryptology eprint report 2008/335 by Barak and Wigderson, 2008. 1

[ABLT06] Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis. Proving Integrality Gaps without Knowing the Linear Program. Theory of Computing, vol. 2, pp. 19–51, 2006. 1

[ALN+06] Sanjeev Arora, László Lovász, Ilan Newman, Yuval Rabani, Yuri Rabinovich, and Santosh Vempala. Local versus global properties of metric spaces. Proceedings of the 17th ACM-SIAM symposium on Discrete algorithms, pp. 41–50, 2006.

[AAM+10] Noga Alon, Sanjeev Arora, Rajsekar Manokaran, Dana Moshkovitz, and Omri Weinstein. On the inapproximability of the densest k-subgraph problem. Manuscript, 2011. 1

[BCC+10] Aditya Bhaskara, Moses Charikar, Eden Chlamtac, Uriel Feige, and Aravindan Vijayaraghavan. Detecting high log-densities: an $o(n^{1/4})$ approximation for densest k-subgraph. In *Proceedings of the 42nd ACM symposium on Theory of computing*, STOC '10, pages 201–210, New York, NY, USA, 2010. ACM. 1, 3, 4, 5, 22

[BRS11] Boaz Barak, Prasad Raghavendra, David Steurer. Rounding Semidefinite Programming Hierarchies via Global Correlation. To appear in FOCS '11. 2, 4

[CMM07] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. *Local Global Tradeoffs in Metric Embeddings.* Proceedings of the 48th Symposium on Foundations of Computer Science, pp. 713–723, 2007.

[CMM09] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. *Integrality gaps for Sherali-Adams relaxations.* Proceedings of the 41st ACM Symposium on Theory of Computing, pp. 283–292, 2009. 1

[Chl06] Eden Chlamtac. *Approximation Algorithms Using Hierarchies of Semidefinite Programming Relaxations.* Proceedings of the 47th Symposium on Foundations of Computer Science, pp. 691-701, 2006. 1, 4

[CS06] Eden Chlamtac and Gyanit Singh. *Improved Approximation Guarantees through Higher Levels of SDP Hierarchies.* Proceedings of the 11th International Workshop, APPROX 2008, pp. 49–62. 1, 4

[CT11] Eden Chlamtac and Madhur Tulsiani. Convex relaxations and integrality gaps. In *To appear in Handbook on Semidefinite, Cone and Polynomial Optimization.*, http://www.cs.princeton.edu/ chlamtac/sdpchapter.pdf, 2011. 1, 3, 4

[dVM07] Wenceslas Fernandez de la Vega and Claire Kenyon-Mathieu. *Linear Programming Relaxations of Maxcut.* Proceedings of the 18th Symposium on Discrete Algorithms, pp. 53–61, 2007. 1

[Fei02] Uriel Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the 34th annual ACM Symposium on Theory of Computing (STOC'02)*, pages 534–543. ACM Press, 2002. 1, 16

[FKP01] Uriel Feige, Guy Kortsarz, and David Peleg. The dense k-subgraph problem. *Algorithmica*, 29(3):410–421, 2001. 3, 5

[FS97] Uriel Feige and Michael Seltser. On the densest k-subgraph problems. Technical Report CS97-16, Weizmann Institute of Science, Rehovot, Israel, 1997. 3, 4

[GMPT07] Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis. *Integrality gaps of* $2-o(1)$ *for vertex cover SDPs in the Lovász–Schrijver hierarchy.* Proceedings of the 48th Symposium on Foundations of Computer Science, pp. 702–712, 2007. 1

[GMT09] Konstantinos Georgiou, Avner Magen, Madhur Tulsiani. Optimal Sherali-Adams Gaps from Pairwise Independence. Proceedings of APPROX 2009: 125-139 1

[GS11] Venkatesan Guruswami, Ali Kemal Sinop. Lasserre Hierarchy, Higher Eigenvalues, and Approximation Schemes for Quadratic Integer Programming with PSD Objectives. To appear in FOCS '11. 2, 4

[GSZ11] Venkatesan Guruswami, Ali Kemal Sinop, Yuan Zhou. Constant Factor Lasserre Gaps for Graph Partitioning Problems . Manuscript, 2011. 2

[Hoe62] Wassily Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, Vol. 58, No. 301 (Mar., 1963), pp. 13–30. 18

[Kho02] Subhash Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *Proceedings of the 42nd Annual IEEE Symposium on the Foundations of Computer Science (FOCS'02)*, pages 600–609, 2002. 2

[Kho04] Subhash Khot. Ruling out PTAS for graph min-bisection, densest subgraph and bipartite clique. In *Proceedings of the 44th Annual IEEE Symposium on the Foundations of Computer Science (FOCS'04)*, pages 136–145, 2004. 1

[KS09] Subhash Khot and Rishi Saket. Sdp integrality gaps with local ℓ_1-embeddability. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 565–574, Washington, DC, USA, 2009. IEEE Computer Society. 1, 4

[Las01] Jean B. Lasserre. *Global Optimization with Polynomials and the Problem of Moments.* SIAM J. Optim. vol. 11, no. 3, pp. 796–817, 2001. 1

[Lau03] Monique Laurent. *A comparison of the Sherali–Adams, Lovász–Schrijver and Lasserre relaxations for* $0-1$ *programming.* Mathematics of Operations Research, vol. 28, pp. 470–496, 2003. 1

[LS91] László Lovász and Alexander Schrijver. *Cones of matrices and set-functions and 0-1 optimization.* SIAM J. Optim., vol. 1, pp. 166–190, 1991. 1

[RS09] Prasad Raghavendra and David Steurer. Integrality gaps for strong sdp relaxations of unique games. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 575–585, Washington, DC, USA, 2009. IEEE Computer Society. 1, 4

[RS10] Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *STOC '10: Proceedings of the 42nd ACM symposium on Theory of computing*, pages 755–764, New York, NY, USA, 2010. ACM. 1, 2

[SA90] Hanif D. Sherali and Warren P. Adams. *A Hierarchy of Relaxations between the Continuous and Convex Hull Representations for Zero-One Programming Problems.* SIAM J. Discrete Math.,, vol. 3, pp. 411–430, 1990. 1

[Sch08] Grant Schoenebeck. *Linear Level Lasserre Lower Bounds for Certain k-CSPs.* Proceedings of the 49th Symposium on Foundations of Computer Science, pp. 593–602, 2008. 1

[ST00] Alex Samorodnitsky and Luca Trevisan. *A PCP characterization of NP with optimal amortized query complexity.* Proceedings of the 32nd ACM Symposium on Theory of Computing, 2000. 2

[ST09] Alex Samorodnitsky and Luca Trevisan. *Gowers Uniformity, Influence of Variables, and PCPs.* SIAM J.Comput. 39(1):323-360 (2006). 2

[STT07a] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. *Tight Integrality Gaps for Lovász–Schrijver LP Relaxations of Vertex Cover and Max Cut.* Proceedings of the 39th Symposium on Theory of Computing, pp. 302–310, 2007. 1

[STT07b] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. *A Linear Round Lower Bound for Lovász–Schrijver SDP Relaxations of Vertex Cover.* Proceedings of the 22nd Conference on Computational Complexity, pp. 205–216, 2007. 1

[SW98] Anand Srivastav and Katja Wolf. Finding dense subgraphs with semidefinite programming. In *Proceedings of the International Workshop on Approximation Algorithms for Combinatorial Optimization (APPROX)*, pages 181–191, 1998. 3, 4

[Tul09] Madhur Tulsiani. CSP gaps and reductions in the lasserre hierarchy. In *Proceedings of the 41st ACM symposium on Theory of computing*, pages 303–312, 2009. 1, 2, 10, 11, 12

[GMT09] Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani", Optimal Sherali-Adams Gaps from Pairwise Independence. In *Proceedings of the International Workshop on Approximation Algorithms for Combinatorial Optimization (APPROX)*, pages 125–139, 2009. 4

A Random graph properties

We prove that the properties used in our gap construction hold for $\mathcal{G}(n,p)$, with $p = n^{-1/2}(\log n)^{1/2}$. These properties are listed in Section 3.1. In what follows fix p to be the value above. As mentioned in Section 2.1, the phrase "with high probability" (w.h.p.) refers to 'with probability at least $1 - \frac{1}{q(n)}$', where $q(n)$ is an arbitrary polynomial in n (sometimes there will be a constant depending on the polynomial).

Lemma A.1. *Every vertex of G has degree between $(n^{1/2}\log n)/2$ and $2n^{1/2}\log n$ w.h.p.*

Proof. Let $u \in V$. The degree $d(u)$ (as a random variable) is the sum of n i.i.d. Bernoulli random variables each having parameter $p = n^{-1/2} \log n$. The expected value is thus $n^{1/2} \log n$. This is $\gg \log n$, and thus by Chernoff bounds, the probability that $\Pr[|d(u) - n^{1/2} \log n| > t] \leq e^{-t^2/4np(1-p)} < \frac{1}{nq(n)}$, for any polynomial $q(n)$. Taking union bound gives the claim. □

Lemma A.2. *Every pair of vertices in G have at most $2 \log^2 n$ common neighbours w.h.p.*

Proof. Let $u, v \in V$. Let X_i be a random variable which is an indicator for $i \in \Gamma(u) \cap \Gamma(v)$. In $\mathcal{G}(n,p)$, we have $\mathbf{E}\, X_i = p^2 = \log^2 n/n$. Thus $\mathbf{E}\, |\Gamma(u) \cap \Gamma(v)| = np^2 = \log^2 n$. Thus the probability that it is $> 2 \log^2 n$ is at most $e^{-\log^2 n/4}$. Taking union bound over all u, v, we obtain that this is smaller than any polynomial. □

Lemma A.3. *Every pair of vertices have at least one common neighbour w.h.p.*

Proof. As above, consider some u, v; we have $\mathbf{E}\, |\Gamma(u) \cap \Gamma(v)| = np^2 = \log^2 n$. Thus $\Pr[|\Gamma_u \cap \Gamma(v)| < \log n] \leq e^{-\log^2 n/4}$ (since we can use Chernoff bounds as long as the expectation $\gg \log n$). Taking union bound again implies the result. □

Lemma A.4. *No induced subgraph on $n^{1/2}$ vertices has density $> 5 \log n$ w.h.p.*

Proof. Let $S \subseteq V$ of size $n^{1/2}$. Then $\mathbf{E}\, E(S,S) = \binom{n^{1/2}}{2} \cdot p = n^{1/2} \log n/2$. Further the variance of this quantity is $\binom{n^{1/2}}{2} p(1-p) < n^{1/2} \log n$. Thus by Chernoff bound,

$$\Pr[|E(S,S) - \mathbf{E}\, E(S,S)| > t] \leq e^{-t^2/4n^{1/2} \log n}.$$

Picking $t = 4n^{1/2} \log n$, the probability upper bound is $e^{-4n^{1/2} \log n}$. Thus we can take a union bound over all the $\binom{n}{n^{1/2}}$ subsets S. This proves the claim. □