

Inapproximability of Maximum Biclique Problems, Minimum k -Cut and Densest At-Least- k -Subgraph from the Small Set Expansion Hypothesis*

Pasin Manurangsi[†]

UC Berkeley

May 11, 2017

Abstract

The *Small Set Expansion Hypothesis* is a conjecture which roughly states that it is NP-hard to distinguish between a graph with a small subset of vertices whose (edge) expansion is almost zero and one in which all small subsets of vertices have expansion almost one. In this work, we prove conditional inapproximability results for the following graph problems based on this hypothesis:

- *Maximum Edge Biclique*: given a bipartite graph G , find a complete bipartite subgraph of G that contains maximum number of edges. We show that, assuming the Small Set Expansion Hypothesis and that $\text{NP} \not\subseteq \text{BPP}$, no polynomial time algorithm approximates Maximum Edge Biclique to within a factor of $n^{1-\varepsilon}$ of the optimum for every constant $\varepsilon > 0$.
- *Maximum Balanced Biclique*: given a bipartite graph G , find a balanced complete bipartite subgraph of G that contains maximum number of vertices. Similar to Maximum Edge Biclique, we prove that this problem is inapproximable in polynomial time to within a factor of $n^{1-\varepsilon}$ of the optimum for every constant $\varepsilon > 0$, assuming the Small Set Expansion Hypothesis and that $\text{NP} \not\subseteq \text{BPP}$.
- *Minimum k -Cut*: given a weighted graph G , find a set of edges with minimum total weight whose removal partitions the graph into (at least) k connected components. For this problem, we prove that it is NP-hard to approximate it to within $(2 - \varepsilon)$ factor of the optimum for every constant $\varepsilon > 0$, assuming the Small Set Expansion Hypothesis.
- *Densest At-Least- k -Subgraph*: given a weighted graph G , find a set S of at least k vertices such that the induced subgraph on S has maximum *density*, which is defined as the ratio between the total weight of the edges and the number of the vertices. We show that this problem is NP-hard to approximate to within $(2 - \varepsilon)$ factor of the optimum for every constant $\varepsilon > 0$, assuming the Small Set Expansion Hypothesis.

The ratios in our inapproximability results are essentially tight since trivial algorithms give n -approximation to both Maximum Edge Biclique and Maximum Balanced Biclique and polynomial time 2-approximation algorithms are known for Minimum k -Cut [SV95] and Densest At-Least- k -Subgraph [And07, KS09].

Our first two results are proved by combining a technique developed by Raghavendra, Steurer and Tulsiani [RST12] to avoid locality of gadget reductions with a generalization of Bansal and Khot's long code test [BK09] whereas our last two results are shown via elementary reductions.

*An extended abstract of this work will appear at ICALP 2017 under a different title [Man17b].

[†]Email: pasin@berkeley.edu. This material is based upon work supported by the National Science Foundation under Grants No. CCF 1540685 and CCF 1655215

1 Introduction

Since the PCP theorem was proved two decades ago [ALM⁺98, AS98], our understanding of approximability of combinatorial optimization problems has grown enormously; tight inapproximability results have been obtained for fundamental problems such as Max-3SAT [Hås01], Max Clique [Hås96] and Set Cover [Mos15, DS14]. Yet, for other problems, including Vertex Cover and Max Cut, known NP-hardness of approximation results come short of matching best known algorithms.

Khot's introduction of the Unique Games Conjecture (UGC) [Kho02] propelled another wave of development in hardness of approximation that saw many of these open problems resolved (see e.g. [KR08, KKMO07]). Alas, some problems continue to elude even attempts at proving UGC-hardness of approximation. For a class of such problems, the failure stems from the fact that typical reductions are local in nature; many reductions from unique games to graph problems could produce disconnected graphs. If we try to use such reductions for problems that involve some forms of expansion of graphs (e.g. Sparsest Cut), we are out of luck.

One approach to overcome the aforementioned issue is through the *Small Set Expansion Hypothesis* (SSEH) of Raghavendra and Steurer [RS10]. To describe the hypothesis, let us introduce some notations. Throughout the paper, we represent an undirected edge-weighted graph $G = (V, E, w)$ by a vertex set V , an edge set E and a weight function $w : E \rightarrow \mathbb{R}_{\geq 0}$. We call G *d-regular* if $\sum_{v:(u,v) \in E} w(u, v) = d$ for every $u \in V$. For a d -regular weighted graph G , the edge expansion $\Phi(S)$ of $S \subseteq V$ is defined as

$$\Phi(S) = \frac{E(S, V \setminus S)}{d \min\{|S|, |V \setminus S|\}}$$

where $E(S, V \setminus S)$ is the total weight of edges across the cut $(S, V \setminus S)$. The small set expansion problem $\text{SSE}(\delta, \eta)$, where η, δ are two parameters that lie in $(0, 1)$, can be defined as follows.

Definition 1 (SSE(δ, η)) *Given a regular edge-weighted graph $G = (V, E, w)$, distinguish between:*

- (Completeness) *There exists $S \subseteq V$ of size $\delta|V|$ such that $\Phi(S) \leq \eta$.*
- (Soundness) *For every $S \subseteq V$ of size $\delta|V|$, $\Phi(S) \geq 1 - \eta$.*

Roughly speaking, SSEH asserts that it is NP-hard to distinguish between a graph that has a small non-expanding subset of vertices and one in which all small subsets of vertices have almost perfect edge expansion. More formally, the hypothesis can be stated as follows.

Conjecture 2 (SSEH [RS10]) *For every $\eta > 0$, there is $\delta = \delta(\eta) > 0$ such that $\text{SSE}(\delta, \eta)$ is NP-hard.*

Interestingly, SSEH not only implies UGC [RS10], but it is also equivalent to a strengthened version of the latter, in which the graph is required to have almost perfect small set expansion [RST12].

Since its proposal, SSEH has been used as a starting point for inapproximability of many problems whose hardnesses are not known otherwise. Most relevant to us is the work of Raghavendra, Steurer and Tulsiani (henceforth RST) [RST12] who devised a technique that exploited structures of SSE instances to avoid locality in reductions. In doing so, they obtained inapproximability of Min Bisection, Balanced Separator, and Minimum Linear Arrangement, which are not known to be hard to approximate under UGC.

1.1 Maximum Edge Biclique and Maximum Balanced Biclique

Our first result is adapting RST technique to prove inapproximability of Maximum Edge Biclique (MEB) and Maximum Balanced Biclique (MBB). For both problems, the input is a bipartite graph. The goal for the former is to find a complete bipartite subgraph that contains as many edges as possible whereas, for the latter, the goal is to find a balanced complete bipartite subgraph that contains as many vertices as possible.

Both problems are NP-hard. MBB was stated (without proof) to be NP-hard in [GJ79, page 196]; several proofs of this exist such as one provided in [Joh87]. For MEB, it was proved to be NP-hard more recently by Peeters [Pee03]. Unfortunately, much less is known when it comes to approximability of both problems. Similar to Maximum Clique, folklore algorithms give $O(n / \text{polylog } n)$ approximation ratio for both MBB and

MEB, and no better algorithm is known. However, not even NP-hardness of approximation of some constant ratio is known for the problems. This is in stark contrast to Maximum Clique for which strong inapproximability results are known [Hås96, Kho01, KP06, Zuc07]. Fortunately, the situation is not completely hopeless as the problems are known to be hard to approximate under stronger complexity assumptions.

Feige [Fei02] showed that, assuming that random 3SAT formulae cannot be refuted in polynomial time, both problems¹ cannot be approximated to within n^ε of the optimum in polynomial time for some $\varepsilon > 0$. Later, Feige and Kogan [FK04] proved $2^{(\log n)^\varepsilon}$ ratio inapproximability for both problems for some $\varepsilon > 0$, assuming that $\text{3SAT} \notin \text{DTIME}(2^{n^{3/4+\delta}})$ for some $\delta > 0$. Moreover, Khot [Kho06] showed, assuming $\text{3SAT} \notin \text{BPTIME}(2^{n^\delta})$ for some $\delta > 0$, that no polynomial time algorithm achieves n^ε -approximation for MBB for some $\varepsilon > 0$. Ambühl et al. [AMS11] subsequently built on Khot's result and showed a similar hardness for MEB. Recently, Bhangale et al. [BGH⁺16] proved that both problems are hard to approximate to within² $n^{1-\varepsilon}$ factor for every $\varepsilon > 0$, assuming a certain strengthened version of UGC and $\text{NP} \neq \text{BPP}$. In addition, while not stated explicitly, the author's recent reduction for Densest k -Subgraph [Man17a] yields $n^{1/\text{polyloglog } n}$ ratio inapproximability for both problems under the Exponential Time Hypothesis [IPZ01] ($\text{3SAT} \notin \text{DTIME}(2^{o(n)})$) and this ratio can be improved to $n^{f(n)}$ for any $f \in o(1)$ under the stronger Gap Exponential Time Hypothesis [Din16, MR17] (no $2^{o(n)}$ time algorithm can distinguish a fully satisfiable 3SAT formula from one which is only $(1 - \varepsilon)$ -satisfiable for some $\varepsilon > 0$); these ratios are better than those in [FK04] but worse than those in [Kho06, AMS11, BGH⁺16].

Finally, it is worth noting that, assuming the Planted Clique Hypothesis [Jer92, Kuč95] (no polynomial time algorithm can distinguish between a random graph $\mathcal{G}(n, 1/2)$ and one with a planted clique of size $\Omega(\sqrt{n})$), it follows (by partitioning the vertex set into two equal sets and delete all the edges within each partition) that Maximum Balanced Biclique cannot be approximated to within $\tilde{O}(\sqrt{n})$ ratio in polynomial time. Interestingly, this does not give any hardness for Maximum Edge Biclique, since the planted clique has only $O(n)$ edges, which less than that in a trivial biclique consisting of any vertex and all of its neighbors.

In this work, we prove strong inapproximability results for both problems, assuming SSEH:

Theorem 3 *Assuming SSEH, there is no polynomial time algorithm that approximates MEB or MBB to within $n^{1-\varepsilon}$ factor of the optimum for every $\varepsilon > 0$, unless $\text{NP} \subseteq \text{BPP}$.*

We note that the only part of the reduction that is randomized is the gap amplification via randomized graph product [BS92, Blu91]. If one is willing to assume only that $\text{NP} \neq \text{P}$ (and SSEH), our reduction still implies that both are hard to approximate to within any constant factor.

Only Bhangale et al.'s result [BGH⁺16] and our result achieve the inapproximability ratio of $n^{1-\varepsilon}$ for every $\varepsilon > 0$; all other results achieve at most n^ε ratio for some $\varepsilon > 0$. Moreover, only Bhangale et al.'s reduction and ours are candidate NP-hardness reductions, whereas each of the other reductions either uses superpolynomial time [FK04, Kho06, AMS11, Man17a] or relies on an average-case assumption [Fei02]. It is also worth noting here that, while both Bhangale et al.'s result and our result are based on assumptions which can be viewed as stronger variants of UGC, the two assumptions are incomparable and, to the best of our knowledge, Bhangale et al.'s technique does not apply to SSEH. A discussion on the similarities and differences between the two assumptions can be found in Appendix C.

Along the way, we prove inapproximability of the following hypergraph bisection problem, which may be of independent interest: given a hypergraph $H = (V_H, E_H)$ find a bisection³ (T_0, T_1) of V_H such that the number of uncut hyperedges is maximized. We refer to this problem as *Max UnCut Hypergraph Bisection* (MUCHB). Roughly speaking, we show that, assuming SSEH, it is hard to distinguish a hypergraph whose optimal bisection cuts only ε fraction of hyperedges from one in which every bisection cuts all but ε fraction of hyperedges:

Lemma 4 *Assuming SSEH, for every $\varepsilon > 0$, it is NP-hard to, given a hypergraph $H = (V_H, E_H)$, distinguish between the following two cases:*

¹While Feige only stated this for MBB, the reduction clearly works for MEB too.

²In [BGH⁺16], the inapproximability ratio is only claimed to be n^ε for *some* $\varepsilon > 0$. However, it is not hard to see that their result in fact implies $n^{1-\varepsilon}$ factor hardness of approximation as well.

³ (T_0, T_1) is a bisection of V_H if $|T_0| = |T_1| = |V_H|/2$, $T_0 \cap T_1 = \emptyset$ and $V_H = T_0 \cup T_1$.

- (Completeness) *There is a bisection (T_0, T_1) of V_H s.t. $|E_H(T_0)|, |E_H(T_1)| \geq (1/2 - \varepsilon)|E_H|$.*
- (Soundness) *For every set $T \subseteq V_H$ of size at most $|V_H|/2$, $|E_H(T)| \leq \varepsilon|E_H|$.*

Here $E_H(T) \triangleq \{e \in E_H \mid e \subseteq T\}$ denotes the set of hyperedges that lie completely inside of the set $T \subseteq V_H$.

Our result above is similar to Khot's quasi-random PCP [Kho06]. Specifically, Khot's quasi-random PCP can be viewed as a hardness for MUCHB in the setting where the hypergraph is d -uniform; roughly speaking, Khot's result states that it is hard (if $\text{3SAT} \notin \bigcap_{\delta > 0} \text{BPTIME}(2^{n^\delta})$) to distinguish between a d -uniform hypergraph where $1/2^{d-2}$ fraction of hyperedges are uncut in the optimal bisection from one where roughly $1/2^{d-1}$ fraction of hyperedges are uncut in any bisection. Note that the latter is the fraction of uncut hyperedges in random hypergraphs and hence the name "quasi-random". In this sense, Khot's result provides better soundness at the expense of worse completeness compared to Theorem 4.

1.2 Minimum k -Cut

In addition to the above biclique problems, we prove an inapproximability result for the Minimum k -Cut problem, in which a weighted graph is given and the goal is to find a set of edges with minimum total weight whose removal partitions the graph into (at least) k connected components. The Minimum k -Cut problem has long been studied. When $k = 2$, the problem can be solved in polynomial time simply by solving Minimum $s - t$ cut for every possible pairs of s and t . In fact, for any fixed k , the problem was proved to be in P by Goldschmidt and Hochbaum [GH94], who also showed that, when k is part of the input, the problem is NP-hard. To circumvent this, Saran and Vazirani [SV95] devised two simple polynomial time $(2 - 2/k)$ -approximation algorithms for the problem. In the ensuing years, different approximation algorithms [NR01, ZNI01, RS02, XCY11] have been proposed for the problem, none of which are able to achieve an approximation ratio of $(2 - \varepsilon)$ for some $\varepsilon > 0$. In fact, Saran and Vazirani themselves conjectured that $(2 - \varepsilon)$ -approximation is intractible for the problem [SV95]. In this work, we show that their conjecture is indeed true, if the SSEH holds:

Theorem 5 *Assuming SSEH, it is NP-hard to approximate Minimum k -Cut to within $(2 - \varepsilon)$ factor of the optimum for every constant $\varepsilon > 0$.*

Note that the problem was claimed to be APX-hard in [SV95]. However, to the best of our knowledge, the proof has never been published and no other inapproximability is known.

1.3 Densest At-Least- k -Subgraph

Our final result is a hardness of approximating the Densest At-Least- k -Subgraph (DALkS) problem, which can be stated as follows. Given a weighted graph, find a subset S of at least k vertices such that the induced subgraph on S has maximum *density*, which is defined as the ratio between the total weight of edges and the number of vertices. The problem was first introduced by Andersen and Chellapilla [AC09] who also gave a 3-approximation algorithm for the problem. Shortly after, 2-approximation algorithms for the problem were discovered by Andersen [And07] and independently by Khuller and Saha [KS09]. We show that, assuming SSEH, this approximation guarantee is essentially the best we can hope for:

Theorem 6 *Assuming the Small Set Expansion Hypothesis, it is NP-hard to approximate Densest At-Least- k -Subgraph to within $(2 - \varepsilon)$ factor of the optimum for every constant $\varepsilon > 0$.*

To the best of our knowledge, no hardness of approximation for DALkS was explicitly proved before. We remark that DALkS is a variant of the Densest k -Subgraph (DkS) problem, which is the same as DALkS except that the desired set S must have size exactly k . DkS has been extensively studied dating back to the early 90s [KP93, FS97, SW98, FL01, FKP01, AHI02, Fei02, Kho06, GL09, RS10, BCC⁺10, AAM⁺11, BCV⁺12, BKRW17, Man17a]. Despite these considerable efforts, its approximability is still wide open. In particular, even though lower bounds have been shown under stronger complexity assumptions [Fei02, Kho06, RS10, AAM⁺11, BKRW17, Man17a] and for LP/SDP hierarchies [BCC⁺10, Man15, CMMV17], not even constant factor NP-hardness of approximation for DkS is known. On the other hand, the best polynomial time algorithm for DkS achieves only $O(n^{1/4+\varepsilon})$ -approximation [BCC⁺10]. Since any inapproximability result for DALkS translates directly to DkS, even proving some constant factor NP-hardness of approximating DALkS would advance our understanding of approximability of DkS.

2 Inapproximability of Minimum k -Cut

We now proceed to prove our main results. Let us start with the simplest: Minimum k -Cut.

Proof of Theorem 5. The reduction from SSE(δ, η) to Minimum k -Cut is simple; the graph G remains the input graph for Minimum k -Cut and we let $k = \delta n + 1$ where $n = |V|$.

Completeness. If there is $S \subseteq V$ of size δn such that $\Phi(S) \leq \eta$, then we partition the graph into k groups where the first group is $V \setminus S$ and each of the other groups contains one vertex from S . The edges cut are the edges across the cut $(S, V \setminus S)$ and the edges within the set S itself. The total weight of edges of the former type is $d|S|\Phi(S) \leq \eta d|S|$ and that of the latter type is at most $d|S|/2$. Hence, the total weight of edges cut in this partition is at most $(1/2 + \eta)d|S| = (1/2 + \eta)\delta dn$.

Soundness. Suppose that, for every $S \subseteq V$ of size δn , $\Phi(S) \geq 1 - \eta$. Let $T_1, \dots, T_k \subseteq V$ be any k -partition of the graph. Assume without loss of generality that $|T_1| \leq \dots \leq |T_k|$. Let $A = T_1 \cup \dots \cup T_i$ where i is the maximum index such that $|T_1 \cup \dots \cup T_i| \leq \delta n$.

We claim that $|A| \geq \delta n - \sqrt{n}$. To see that this is the case, suppose for the sake of contradiction that $|A| < \delta n - \sqrt{n}$. Since $|A \cup T_{i+1}| > \delta n$, we have $|T_{i+1}| > \sqrt{n}$. Moreover, since $A = T_1 \cup \dots \cup T_i$, we have $i \leq |A| < \delta n - \sqrt{n}$. As a result, we have $n = |T_1 \cup \dots \cup T_k| \geq |T_{i+1} \cup \dots \cup T_k| \geq (k-i)|T_i| > \sqrt{n} \cdot \sqrt{n} = n$, which is a contradiction. Hence, $|A| \geq \delta n - \sqrt{n}$.

Now, note that, for every $S \subseteq V$ of size δn , $\Phi(S) \geq 1 - \eta$ implies that $E(S) \leq \eta d\delta n/2$ where $E(S)$ denote the total weight of all edges within S . Since $|A| \leq \delta n$, we also have $E(A) \leq \eta d\delta n/2$. As a result, the total weight of edges across the cut $(A, V \setminus A)$, all of which are cut by the partition, is at least

$$d|A| - \eta d\delta n \geq (1 - \eta)d\delta n - d\sqrt{n} = \left(1 - \eta - \frac{1}{\delta\sqrt{n}}\right)\delta dn.$$

For every sufficiently small constant $\varepsilon > 0$, by setting $\eta = \varepsilon/20$ and $n \geq 100/(\varepsilon^2\delta^2)$, the ratio between the two cases is at least $(2 - \varepsilon)$, which concludes the proof of Theorem 5. \square

3 Inapproximability of Densest At-Least- k -Subgraph

We next prove our inapproximability result for Densest At-Least- k -Subgraph, which is also very simple. For this reduction and the subsequent reductions, it will be more convenient for us to use a different (but equivalent) formulation of SSEH. To state it, we first define a variant of SSE(δ, η) called SSE(δ, η, M); the completeness remains the same whereas the soundness is strengthened to include all S of size in $\left[\frac{\delta|V|}{M}, \delta|V|M\right]$.

Definition 7 (SSE(δ, η, M)) Given a regular edge-weighted graph $G = (V, E, w)$, distinguish between:

- (Completeness) There exists $S \subseteq V$ of size $\delta|V|$ such that $\Phi(S) \leq \eta$.
- (Soundness) For every $S \subseteq V$ with $|S| \in \left[\frac{\delta|V|}{M}, \delta|V|M\right]$, $\Phi(S) \geq 1 - \eta$.

The new formulation of the hypothesis can now be stated as follows.

Conjecture 8 For every $\eta, M > 0$, there is $\delta = \delta(\eta, M) > 0$ such that SSE(δ, η, M) is NP-hard.

Raghavendra et al. [RST12] showed that this formulation is equivalent to the earlier formulation (Conjecture 2); please refer to Appendix A.2 of [RST12] for a simple proof of this.

Proof of Theorem 6. Given an instance $G = (V, E, w)$ of SSE(δ, η, M), we create an input graph $G' = (V', w')$ for Densest At-Least- k -Subgraph as follows. V' consists of all the vertices in V and an additional vertex v^* . The weight function w' remains the same as w for all edges in V whereas v^* has only a self-loop⁴ with weight $d\delta n/2$. In other words, $E' = E \cup \{v^*\}$ and $w'(v^*) = d\delta n/2$. Finally, let $k = 1 + \delta n$ where $n = |V|$.

⁴If we would like to avoid self-loops, we can replace v^* with two vertices v_1^*, v_2^* with an edge of weight $d\delta n/2$ between them.

Completeness. If there is $S \subseteq V$ of size δn such that $\Phi(S) \leq \eta$, consider the set $S' = S \cup \{v^*\}$. We have $|S'| = k$ and the density of S' is $(d\delta n/2 + E(S))/k$ where $E(S)$ denote the total weight of edges within S . This can be written as

$$\left(\frac{\delta n}{k}\right)\left(\frac{d\delta n/2 + E(S)}{\delta n}\right) = \left(\frac{\delta n}{k}\right)\left(\frac{d}{2} + \left(\frac{1 - \Phi(S)}{2}\right)d\right) \geq \frac{d\delta n(1 - \eta/2)}{k}.$$

Soundness. Suppose that $\Phi(S) \geq 1 - \eta$ for every $S \subseteq V$ of size $|S| \in [\delta n/M, \delta nM]$. Consider any set $T' \subseteq V$ of size at least k . Let $T = T' \setminus \{v^*\}$ and let $E(T)$ denote the total weight of edges within T . Observe that the density of S is at most $(d\delta n/2 + E(T))/|T'|$. Let us consider the following two cases.

1. $|T| \leq \delta nM$. In this case, $\Phi(T) \geq 1 - \eta$ and we have

$$\frac{d\delta n/2 + E(T)}{|T'|} \leq \frac{d\delta n}{2k} + \left(\frac{1 - \Phi(T)}{2}\right) \leq \frac{d\delta n}{2k} + d\eta/2 = \frac{d\delta n (\frac{1}{2} + \frac{\eta}{2} (1 + \frac{1}{\delta n}))}{k} \leq \frac{d\delta n (1/2 + \eta)}{k}.$$

2. $|T| > \delta nM$. In this case, we have

$$\frac{d\delta n/2 + E(T)}{|T'|} < \frac{d}{2M} + \frac{d}{2} = \frac{d\delta n}{k} \left(\frac{1}{2} + \frac{1}{2M}\right) \left(1 + \frac{1}{\delta n}\right) \leq \frac{d\delta n (\frac{1}{2} + \frac{1}{\delta n} + \frac{1}{M})}{k}.$$

Hence, in both cases, the density of T' is at most $d\delta n (1/2 + \max\{\eta, \frac{1}{\delta n} + \frac{1}{M}\})/k$.

For every sufficiently small constant $\varepsilon > 0$, by picking $\eta = \varepsilon/20, M = 40/\varepsilon$ and $n \geq 800/(\varepsilon\delta)$, the ratio between the two cases is at least $(2 - \varepsilon)$, concluding the proof of Theorem 6. \square

4 Inapproximability of MEB and MBB

Let us now turn our attention to MEB and MBB. First, note that we can reduce MUCHB to MEB/MMB by just letting the two sides of the bipartite graph be E_H and creating an edge (e_1, e_2) iff $e_1 \cap e_2 = \emptyset$. This immediately shows that Lemma 4 implies the following:

Lemma 9 *Assuming SSEH, for every $\delta > 0$, it is NP-hard to, given a bipartite graph $G = (L, R, E)$ with $|L| = |R| = n$, distinguish between the following two cases:*

- (Completeness) G contains $K_{(1/2-\delta)n, (1/2-\delta)n}$ as a subgraph.
- (Soundness) G does not contain $K_{\delta n, \delta n}$ as a subgraph.

Here $K_{t,t}$ denotes the complete bipartite graph in which each side contains t vertices.

We provide the full proof of Lemma 9 in Appendix A. We also note that Theorem 3 follows from Lemma 9 by gap amplification via randomized graph product [BS92, Blu91]. Since this has been analyzed before even for biclique [Kho06, Appendix D], we defer the full proof to Appendix B.

We are now only left to prove Lemma 4; we devote the rest of this section to this task.

4.1 Preliminaries

Before we continue, we need additional notations and preliminaries. For every graph $G = (V, E, w)$ and every vertex v , we write $G(v)$ to denote the distribution on its neighbors weighted according to w . Moreover, we sometimes abuse the notation and write $e \sim G$ to denote a random edge of G weighted according to w .

While our reduction can be understood without notation of unique games, it is best described in a context of unique games reductions. We provide a definition of unique games below.

Definition 10 (Unique Game (UG)) *A unique game instance $\mathcal{U} = (\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W}), [R], \{\pi_e\}_{e \in \mathcal{E}})$ consists of an edge-weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$, a label set $[R] = \{1, \dots, R\}$, and, for each $e \in \mathcal{E}$, a permutation $\pi_e : [R] \rightarrow [R]$. The goal is to find an assignment $F : V \rightarrow [R]$ such that $\text{val}_{\mathcal{U}}(F) \triangleq \Pr_{(u,v) \sim G}[\pi_{(u,v)}(F(u)) = F(v)]$ is maximized; we call an edge (u, v) such that $\pi_{(u,v)}(F(u)) = F(v)$ satisfied.*

Khot's UGC [Kho02] states that, for every $\varepsilon > 0$, it is NP-hard to distinguish between a unique game in which there exists an assignment satisfying at least $(1 - \varepsilon)$ fraction of edges from one in which every assignment satisfies at most ε fraction of edges.

Finally, we need some preliminaries in discrete Fourier analysis. We state here only few facts that we need. We refer interested readers to [O'D14] for more details about the topic.

For any discrete probability space Ω , $f : \Omega^R \rightarrow [0, 1]$ can be written as $\sum_{\sigma \in [| \Omega |]^R} \hat{f}(\sigma) \phi_\sigma$ where $\{\phi_\sigma\}_{\sigma \in [| \Omega |]^R}$ is the product Fourier basis of $L^2(\Omega^R)$ (see [O'D14, Chapter 8.1]). The degree- d influence on the j -th coordinate of f is $\text{infl}_j^d(f) \triangleq \sum_{\sigma \in [| \Omega |]^R, \sigma_j \neq 1, \# \sigma \leq d} \hat{f}^2(\sigma)$ where $\#\sigma \triangleq |\{i \in [R] \mid \sigma_i \neq 1\}|$. It is well known that $\sum_{j=1}^R \text{infl}_j^d(f) \leq d$ (see [MOO10, Proposition 3.8]).

We also need the following theorem. It follows easily⁵ from the so-called “It Ain’t Over Till It’s Over” conjecture, which is by now a theorem [MOO10, Theorem 4.9].

Theorem 11 ([MOO10]) *For any $\beta, \varepsilon_T, \gamma > 0$, there exists $\kappa > 0$ and $t, d \in \mathbb{N}$ such that, if any functions $f_1, \dots, f_t : \Omega^R \rightarrow \{0, 1\}$ where Ω is a probability space whose probability of each atom is at least β satisfy⁶*

$$\forall i \in [t], \mathbb{E}_{x \in \Omega^R} [f_i(x)] \leq 0.99 \quad \text{and} \quad \forall j \in [R], \forall 1 \leq i_1 \neq i_2 \leq t, \min\{\text{infl}_j^d(f_{i_1}), \text{infl}_j^d(f_{i_2})\} \leq \kappa,$$

then

$$\Pr_{x \in \Omega^R, D \sim S_{\varepsilon_T}(R)} \left[\bigwedge_{i=1}^t f_i(C_D(x)) \equiv 1 \right] < \gamma$$

where $D \sim S_{\varepsilon_T}(R)$ is a random subset of $[R]$ where each $i \in [R]$ is included independently w.p. ε_T , $C_D(x) \triangleq \{x' \mid x'_{[R] \setminus D} = x_{[R] \setminus D}\}$ and $f_i(C_D(x)) \equiv 1$ is short for $\forall x' \in C_D(x), f_i(x') = 1$.

4.2 Bansal-Khot Long Code Test and A Candidate Reduction

Theorem 11 leads us nicely to the Bansal-Khot long code test [BK09]. For UGC hardness reductions, one typically needs a *long code test* (aka dictatorship gadget) which, on input $f_1, \dots, f_t : \{0, 1\}^R \rightarrow \{0, 1\}$, has the following properties:

- (*Completeness*) If $f_1 = \dots = f_t$ is a long code⁷, the test accepts with large probability.
- (*Soundness*) If f_1, \dots, f_t are balanced (i.e. $\mathbb{E} f_1 = \dots = \mathbb{E} f_t = 1/2$) and are “far from being a long code”, then the test accepts with low probability.

A widely-used notion of “far from being a long code”, and one we will use here, is that the functions do not share a coordinate with large low degree influence (i.e. for every $j \in [R]$ and every $i_1 \neq i_2 \in [t]$, at least one of $\text{infl}_j^d(f_{i_1}), \text{infl}_j^d(f_{i_2})$ is small),

Bansal-Khot long code test works by first picking $x \sim \{0, 1\}^R$ and $D \sim S_{\varepsilon_T}(R)$. Then, test whether f_i evaluates to 1 on the whole $C_D(x)$. This can be viewed as an “algorithmic” version of Theorem 11; specifically, the theorem (with $\Omega = \{0, 1\}$) immediately implies the soundness property of this test. On the other hand, it is obvious that, if $f_1 = \dots = f_t$ is a long code, then the test accepts with probability $1/2 - \varepsilon_T$.

Bansal and Khot used this test to prove tight hardness of approximation of Vertex Cover. The reduction is via a natural composition of the test with unique games. Their reduction also gives a candidate reduction from UG to MUCHB, which is stated below in Figure 1.

As is typical for gadget reductions, for $T \subseteq V_H$, we view the indicator function $f_u(x) \triangleq \mathbf{1}[(u, x) \in T]$ for each $u \in \mathcal{V}$ as the intended long code. If there exists an assignment ϕ to the unique game instance that satisfies nearly all the constraints, then the bisection corresponding to $f_u(x) = x_{\phi(u)}$ cuts only small fraction of edges, which yields the completeness of MUCHB.

⁵For more details on how this version follows from there, please refer to [Sve13, page 769].

⁶0.99 could be replaced by any constant less than one; we use it to avoid introducing more parameters.

⁷A long code is simply j -junta (i.e. a function that depends only on the x_j) for some $j \in [R]$.

Input: A unique game $(\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W}), [R], \{\pi_e\}_{e \in E})$ and parameters $\ell \in \mathbb{N}$ and $\varepsilon_T \in (0, 1)$.

Output: A hypergraph $H = (V_H, E_H)$.

The vertex set V_H is $\mathcal{V} \times \{0, 1\}^R$ and the hyperedges are distributed as follows:

- Sample $u \sim \mathcal{V}$ and sample $v_1 \sim \mathcal{G}(u), \dots, v_\ell \sim \mathcal{G}(u)$.
- Sample $x \sim \{0, 1\}^R$ and a subset $D \sim S_{\varepsilon_T}(R)$.
- Output a hyperedge $e = \{(v_p, x') \mid p \in [\ell], x' \in C_D(x)\}$.

Figure 1: A Candidate Reduction from UG to MUCHB

As for the soundness, we want to decode an UG assignment from $T \subseteq V_H$ of size at most $|V_H|/2$ which contains at least ε fraction of hyperedges. In terms of the tests, this corresponds to a collection of functions $\{f_u\}_{u \in \mathcal{V}}$ such that $\mathbb{E}_{u \sim \mathcal{V}} \mathbb{E}_{x \sim \{0, 1\}^R} f_u(x) \leq 1/2$ and the Bansal-Khot test on f_{v_1}, \dots, f_{v_t} passes with probability at least ε where v_1, \dots, v_t are sampled as in Figure 1. Now, if we assume that $\mathbb{E}_x f_u(x) \leq 0.99$ for all $u \in \mathcal{V}$, then such decoding is possible via a similar method as in [BK09] since Theorem 11 can be applied here.

Unfortunately, the assumption $\mathbb{E}_x f_u(x) \leq 0.99$ does not hold for an arbitrary $T \subseteq V_H$ and the soundness property indeed fails. For instance, imagine the constraint graph \mathcal{G} of the starting UG instance consisting of two disconnected components of equal size; let $\mathcal{V}_0, \mathcal{V}_1$ be the set of vertices in the two components. In this case, the bisection $(\mathcal{V}_0 \times \{0, 1\}^R, \mathcal{V}_1 \times \{0, 1\}^R)$ does not even cut a single edge! This is regardless of whether there exists an assignment to the UG that satisfies a large fraction of edges.

4.3 RST Technique and The Reduction from SSE to MUCHB

The issue described above is common for graph problems that involves some form of expansion of the graph. The RST technique [RST12] was in fact invented to specifically circumvent this issue. It works by first reducing SSE to UG and then exploiting the structure of the constructed UG instance when composing it with a long code test; this allows them to avoid extreme cases such as one above. There are four parameters in the reduction: $R, k \in \mathbb{N}$ and ε_V, β . Before we describe the reduction, let us define additional notations:

- Let $G^{\otimes R}$ denote the R -tensor graph of $G = (V, E, w)$; the vertex set of $G^{\otimes R}$ is V^R and, for every $A, B \in V^R$, the edge weight between A, B is the product of $w(A_i, B_i)$ in G for all $i \in [R]$.
- For each $A \in V^R$, $T_V(A)$ denote the distribution on V^R where the i -th coordinate is set to A_i with probability $1 - \varepsilon_V$ and is randomly sampled from V otherwise.
- Let $\Pi_{R,k}$ denote the set of all permutations π 's of $[R]$ such that, for each $j \in [k]$, $\pi(\{R(j-1)/k + 1, \dots, Rj/k\}) = \{R(j-1)/k + 1, \dots, Rj/k\}$.
- Let $\{0, 1, \perp\}_\beta$ denote the probability space such that the probability for $0, 1$ are both $\beta/2$ and the probability for \perp is $1 - \beta$.

The first step of reduction takes an $\text{SSE}(\delta, \eta, M)$ instance $G = (V, E, w)$ and produces a unique game $\mathcal{U} = (\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W}), [R], \{\pi_e\}_{e \in E})$ where $\mathcal{V} = V^R$ and the edges are distributed as follows:

1. Sample $A \sim V^R$ and $\tilde{A} \sim T_V(A)$.
2. Sample $B \sim G^{\otimes R}(\tilde{A})$ and $\tilde{B} \sim T_V(B)$.
3. Sample two random $\pi_A, \pi_B \sim \Pi_{R,k}$.
4. Output an edge $e = (\pi_A(A), \pi_B(B))$ with $\pi_{(A,B)} = \pi_B \circ \pi_A^{-1}$.

Here ε_V is a small constant, k is large and R/k should be think of as $\Theta(1/\delta)$. When there exists a set $S \subseteq V$ of size $\delta|V|$ with small edge expansion, the intended assignment is to, for each $A \in V^R$, find the first block $j \in [k]$ such that $|A(j) \cap S| = 1$ where $A(j)$ denotes the multiset $\{A_{R(j-1)/k+1}, \dots, A_{Rj/k}\}$ and let $F(A)$ be the coordinate of the vertex in that intersection. If no such j exists, we assign $F(A)$ arbitrarily. Note that, since $R/k = \Theta(1/\delta)$, $\Pr[|A(j) \cap S| = 1]$ is constant, which means that only $2^{-\Omega(k)}$ fraction of vertices are assigned arbitrarily. Moreover, it is not hard to see that, for the other vertices, their assignments rarely violate constraints as ε_V and $\Phi(S)$ are small. This yields the completeness. In addition, the soundness was shown in [RS10, RST12], i.e., if every $S \subseteq V$ of size $\delta|V|$ has near perfect expansion, no assignment satisfies many constraints in \mathcal{U} (see Lemma 16).

The second step is to reduce this UG instance to a hypergraph $H = (V_H, E_H)$. Instead of making the vertex set $V^R \times \{0, 1\}^R$ as in the previous candidate reduction, we will instead make $V_H = V^R \times \Omega^R$ where $\Omega = \{0, 1, \perp\}_\beta$ and β is a small constant. This does not seem to make much sense from the UG reduction standpoint because we typically want to assign which side of the bisection $(A, x) \in V_H$ is in according to $x_{F(A)}$ but $x_{F(A)}$ could be \perp in this construction. However, it makes sense when we view this as a reduction from SSE directly: let us discard all coordinates i 's such that $x_i = \perp$ and define $A(j, x) \triangleq \{A_i \mid i \in \{R(j-1)/k+1, \dots, Rj/k\} \wedge x_i \neq \perp\}$. Then, let $j^*(A, x) \triangleq \min\{j \mid |A(j, x) \cap S| = 1\}$ and let $i^*(A, x)$ be the coordinate in the intersection between $A(j^*(A, x), x)$ and S , and assign (A, x) to $T_{x_{i^*(A, x)}}$. (If $j^*(A, x)$ does not exist, then assign (A, x) arbitrarily.)

Observe that, in the intended solution, the side that (A, x) is assigned to does not change if (1) A_i is modified for some $i \in [R]$ s.t. $x_i = \perp$ or (2) we apply some permutation $\pi \in \Pi_{R,k}$ to both A and x . In other words, we can “merge” two vertices (A, x) and (A', x') that are equivalent through these changes together in the reduction. For notational convenience, instead of merging vertices, we will just modify the reduction so that, if (A, x) is included in some hyperedge, then every (A', x') reachable from (A, x) by these operations is also included in the hyperedge. More specifically, if we define $M_x(A) \triangleq \{A' \in V^R \mid A'_i = A_i \text{ for all } i \in [R] \text{ such that } x_i \neq \perp\}$ corresponding to the first operation, then we add $\pi(A', x)$ to the hyperedge for every $A' \in M_x(A)$ and $\pi \in \Pi_{R,k}$. The full reduction is shown in Figure 2.

Input: An edge-weighted graph G with vertex set V and parameters $R, k, \ell \in \mathbb{N}$ and $\varepsilon_T, \varepsilon_V, \beta \in (0, 1)$.
Output: A hypergraph $H = (V_H, E_H)$.
 $V_H \triangleq V^R \times \Omega^R$ where $\Omega \triangleq \{0, 1, \perp\}_\beta$ and the hyperedges are distributed as follows:
• Sample $A \sim V^R$ and $\tilde{A}^1, \dots, \tilde{A}^\ell \sim T_V(A)$.
• Sample $B^1 \sim G^{\otimes R}(\tilde{A}^1), \dots, B^\ell \sim G^{\otimes R}(\tilde{A}^\ell)$ and $\tilde{B}^1 \sim T_V(B^1), \dots, \tilde{B}^\ell \sim T_V(B^\ell)$
• Sample $x \in \Omega^R$ and a subset $D \sim S_{\varepsilon_T}(R)$.
• Output a hyperedge $e = \{\pi(B', x') \mid p \in [\ell], \pi \in \Pi_{R,k}, x' \in C_D(x), B' \in M_{x'}(\tilde{B}^p)\}$.

Figure 2: Reduction from SSE to Max UnCut Hypergraph Bisection

Note that the test we apply here is slightly different from Bansal-Khot test as our test is on $\Omega = \{0, 1, \perp\}_\beta$ instead of $\{0, 1\}$ used in [BK09]. Another thing to note is that now our vertices and hyperedges are weighted, the vertices according to the product measure of $V^R \times \Omega^R$ and the edges according to the distribution produced from the reduction. We write μ_H to denote the measure on the vertices, i.e., for $T \subseteq V^R \times \{0, 1, \perp\}^R$, $\mu_H(T) = \Pr_{A \sim V^R, x \sim \Omega^R}[(A, x) \in T]$, and we abuse the notation $E_H(T)$ and use it to denote the probability that a hyperedge as generated in Figure 2 lies completely in T . We note here that, while the MUCHB as stated in Lemma 4 is unweighted, it is not hard to see that we can go from weighted version to unweighted by copying each vertex and each edge proportional⁸ to their weights.

The advantage of this reduction is that the vertex “merging” makes gadget reduction non-local; for instance, it is clear that even if the starting graph V has two connected components, the resulting hypergraph is now connected. In fact, Raghavendra et al. [RST12] show a much stronger quantitative bound. To state this, let us consider any $T \in V_H$ with $\mu_H(T) = 1/2$. From how the hyperedges are defined, we can assume w.l.o.g. that, if $(A, x) \in T$, then $\pi(A', x) \in T$ for every $A' \in M_x(A)$ and every $\pi \in \Pi_{R,k}$. Again, let $f_A(x) \triangleq \mathbb{1}[(A, x) \in T]$. The following bound on the variance of $\mathbb{E}_x f_A(x)$ is implied by the proof of Lemma 6.6 in [RST12]:

$$\mathbb{E}_{A \sim V^R} \left(\mathbb{E}_{x \sim \Omega^R} f_A(x) - 1/2 \right)^2 \leq \beta.$$

The above bound implies that, for most A 's, the mean of f_A cannot be too large. This will indeed allow us to ultimately apply Theorem 11 on a certain fraction of the tuples $(\tilde{B}^1, \dots, \tilde{B}^\ell)$ in the reduction, which leads to an UG assignment with non-negligible value.

⁸Note that this is doable since we can pick $\beta, \varepsilon_V, \varepsilon_T$ to be rational.

4.4 Completeness

In the completeness case, we define a bisection similar to that described above. This bisection indeed cuts only a small fraction of hyperedges; quantitatively, this yields the following lemma.

Lemma 12 *If there is a set $S \subseteq V$ such that $\Phi(S) \leq \eta$ and $|S| = \delta|V|$ where $\delta \in \left[\frac{k}{10\beta R}, \frac{k}{\beta R}\right]$, then there is a bisection (T_0, T_1) of V_H such that $E_H(T_0), E_H(T_1) \geq 1/2 - O(\varepsilon_T/\beta) - O(\eta\ell/\beta) - O(\varepsilon_V\ell/\beta) - 2^{-\Omega(k)}$ where $O(\cdot)$ and $\Omega(\cdot)$ hide only absolute constants.*

Proof. Suppose that there exists $S \subseteq V$ of size $|S| = \delta|V|$ where $\left[\frac{k}{10\beta R}, \frac{k}{\beta R}\right]$ and $\Phi(S) \leq \eta$. For $A \in V^R, x \in \{0, 1, \perp\}^R$, we will use the following notations throughout this proof:

- For $j \in [k]$, let $W(A, x, j)$ denote the set of all coordinates i in j -th block such that $x_i \neq \perp$ and $A_i \in S$, i.e., $W(A, x, j) = \{i \in \{R(j-1)/k+1, \dots, Rj/k\} \mid A_i \in S \wedge x_i \neq \perp\}$.
- Let $j^*(A, x)$ denote the first block j with $|W(A, x, j)| = 1$, i.e., $j^*(A, x) = \min\{j \in [k] \mid |W(A, x, j)| = 1\}$. Note that if such block does not exist, we set $j^*(A, x) = -1$.
- Let $i^*(A, x)$ be the only element in $W(A, x, j^*(A, x))$. If $j^*(A, x) = -1$, let $i^*(A, x) = -1$.

To define T_0, T_1 , we start by constructing $T'_0 \subseteq T_0$ and $T'_1 \subseteq T_1$ as follows: assign each $(A, x) \in V_H$ such that $j^*(A, x) \neq -1$ to $T_{x_{i^*(A, x)}}$. Finally, we assign the rest of the vertices arbitrarily to T_0 and T_1 in such a way that $\mu_H(T_0) = \mu_H(T_1)$. Since $T'_0 \subseteq T_0, T'_1 \subseteq T_1$, it suffices to show the desired bound for $E_H(T'_0), E_H(T'_1)$. Due to symmetry, it suffices to bound $E_H(T'_0)$. Recall that $E_H(T'_0) = \Pr[e \subseteq T'_0]$ where e is generated as detailed in Figure 2.

To compute $E_H(T'_0)$, it will be most convenient to make a block-by-block analysis. In particular, for each block $j \in [k]$, we define G_j to denote the event that $j^*(B', x') = j$ for some $(B', x') \in e$. We will be interested in bounding the following conditional probabilities:

- $c_1 \triangleq \Pr[j^*(A, x) = j \mid j^*(A, x) > j-1]$
- $c_2 \triangleq \Pr[G_j \mid j^*(A, x) > j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1}]$
- $c_3 \triangleq \Pr[e \not\subseteq T'_0 \mid j^*(A, x) = j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1}]$

Here and throughout the proof, $e, A, \tilde{A}^1, \dots, \tilde{A}^\ell, B^1, \dots, B^\ell, \tilde{B}^1, \dots, \tilde{B}^\ell$ are as sampled by our reduction in Figure 2. Note also that it is clear that c_1, c_2, c_3 do not depend on j .

Before we bound c_1, c_2, c_3 , let us see how these probabilities can be used to bound $E_H(T'_0)$.

$$\begin{aligned}
\Pr_{e \sim E_H}[e \subseteq T'_0] &\geq \sum_{j=1}^k \Pr_{e \sim E_H}[e \subseteq T'_0 \wedge j^*(A, x) = j] \\
&\geq \sum_{j=1}^k \Pr_{e \sim E_H}[e \subseteq T'_0 \wedge j^*(A, x) = j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1}] \\
&= \sum_{j=1}^k \Pr_{e \sim E_H}[e \subseteq T'_0 \mid j^*(A, x) = j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1}] \Pr[j^*(A, x) = j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1}] \\
&= (1 - c_3) \sum_{j=1}^k \Pr[j^*(A, x) = j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1}] \\
&= (1 - c_3) \sum_{j=1}^k \Pr[j^*(A, x) = j] \Pr[\neg G_1 \wedge \dots \wedge \neg G_{j-1} \mid j^*(A, x) = j]. \tag{1}
\end{aligned}$$

The probability that $j^*(A, x) = j$ is in fact simply

$$\Pr[j^*(A, x) = j] = \Pr[j^*(A, x) = j \mid j^*(A, x) > j-1] \prod_{q=1}^{j-1} \Pr[j^*(A, x) \neq q \mid j^*(A, x) > q-1] = c_1(1 - c_1)^{j-1}.$$

Moreover, $\Pr[\neg G_1 \wedge \dots \wedge \neg G_{j-1} \mid j^*(A, x) = j]$ can be written as

$$\Pr[\neg G_1 \wedge \dots \wedge \neg G_{j-1} \mid j^*(A, x) = j] = \prod_{q=1}^{j-1} \Pr[\neg G_q \mid j^*(A, x) = j \wedge \neg G_1 \wedge \dots \wedge \neg G_{q-1}] = (1 - c_2)^{j-1}.$$

Plugging these two back into (1), we have

$$\begin{aligned} E_H(T'_0) &\geq (1 - c_3) \sum_{j=1}^k c_1 (1 - c_1)^{j-1} (1 - c_2)^{j-1} \\ &\geq c_1 (1 - c_3) \sum_{j=1}^k (1 - c_1 - c_2)^{j-1} \\ &= c_1 (1 - c_3) \left(\frac{1 - (1 - c_1 - c_2)^k}{c_1 + c_2} \right) \\ &= (1 - c_3) (1 - (1 - c_1 - c_2)^k) \left(\frac{1}{1 + c_2/c_1} \right) \\ &\geq (1 - c_3) (1 - (1 - c_1)^k) (1 - c_2/c_1) \\ &\geq 1 - c_3 - (1 - c_1)^k - c_2/c_1. \end{aligned} \tag{2}$$

With (2) in mind, we will proceed to bound c_1, c_2, c_3 . Before we do so, let us state two inequalities that will be useful: for every $i \in [R], p \in [\ell]$, we have

$$\Pr[\tilde{B}_i^p \in S \mid A_i \notin S] \leq 2\varepsilon_V \delta + 2\eta \delta \tag{3}$$

and

$$\Pr[\tilde{B}_i^p \notin S \mid A_i \in S] \leq 2\varepsilon_V + \eta. \tag{4}$$

The first inequality comes from the fact that, for \tilde{B}_i^p to be in S when $A_i \notin S$, at least one of the following events must occur: (1) $A_i \neq \tilde{A}_i^p$ and $\tilde{A}_i^p \in S$, (2) $B_i^p \neq \tilde{B}_i^p$ and $\tilde{B}_i^p \in S$, (3) $(\tilde{A}_i^p, \tilde{B}_i^p) \in (V \setminus S) \times S$. Each of first two occurs with probability $\varepsilon_V \delta$ whereas the last event occurs with probability at most $\eta \delta / (1 - \delta) \leq 2\eta \delta$. On the other hand, for the second inequality, at least one of the following events must occur: (1) $A_i \neq \tilde{A}_i^p$, (2) $B_i^p \neq \tilde{B}_i^p$, (3) $(\tilde{A}_i^p, \tilde{B}_i^p) \in S \times (V \setminus S)$. Each of first two occurs with probability ε_V whereas the last event occurs with probability at most η .

Bounding c_1

To compute c_1 , observe that $\Pr[j^*(A, z) = j \mid j^*(A, z) > j-1]$ is the probability that, for exactly one i in the j -th block, $A_i \in S$ and $x_i \neq \perp$. For a fixed i , this happens with probability $\beta \delta$. Hence, $c_1 = (R/k) \beta \delta (1 - \beta \delta)^{R/k-1}$. Since $\delta \in \left[\frac{k}{10\beta R}, \frac{k}{\beta R} \right]$, we can conclude that c_1 is simply a constant (i.e. $c_1 \in [10^{-5}, 0.5]$).

Bounding c_2

We next bound c_2 . If $j^*(A, x) > j$, we know that $|W(A, x, j)| \neq 1$. Let us consider two cases:

1. $W(A, x, j) = \emptyset$. Observe that, if G_j occurs, then there exist $p \in [\ell]$ and $i \in \{R(j-1)/k+1, \dots, Rj/k\}$ such that $\tilde{B}_i^p \in S$, and $x_i \neq \perp$ or $i \in D$. For brevity, below we denote the conditional event $j^*(A, x) > j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1} \wedge W(A, x, j) = \emptyset$ by E . By union bound, our observation gives the following bound.

$$\Pr[G_j \mid E] \leq \sum_{i=R(j-1)/k+1}^{Rj/k} \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \wedge (x_i \neq \perp \vee i \in D) \mid E]$$

$$= \sum_{i=R(j-1)/k+1}^{Rj/k} (\Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \wedge x_i \neq \perp \mid E] + \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \wedge x_i = \perp \wedge i \in D \mid E]) \quad (5)$$

We can now bound the first term by

$$\begin{aligned} \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \wedge x_i \neq \perp \mid E] &\leq \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \mid x_i \neq \perp \wedge E] \\ &\leq \sum_{p=1}^{\ell} \Pr[\tilde{B}_i^p \in S \mid x_i \neq \perp \wedge E] \\ (\text{Since } W(A, j) = \emptyset) &= \sum_{p=1}^{\ell} \Pr[\tilde{B}_i^p \in S \mid A_i \notin S] \\ (\text{From (3)}) &\leq \ell(2\varepsilon_V \delta + 2\eta\delta). \end{aligned} \quad (6)$$

Consider the other term in (5). We can rearrange it as follows.

$$\begin{aligned} \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \wedge x_i = \perp \wedge i \in D \mid E] &= \varepsilon_T \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \wedge x_i = \perp \mid i \in D \wedge E] \\ &\leq \varepsilon_T \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \mid x_i = \perp \wedge i \in D \wedge E] \\ &\leq \varepsilon_T (\Pr[A_i \in S \mid x_i = \perp \wedge i \in D \wedge E] + \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \mid A_i \notin S \wedge x_i = \perp \wedge i \in D \wedge E]) \\ &= \varepsilon_T (\Pr[A_i \in S \mid x_i = \perp] + \Pr[\exists p \in [\ell], \tilde{B}_i^p \in S \mid A_i \notin S]) \\ (\text{From (3)}) &\leq \varepsilon_T (\delta + 2\varepsilon_V \delta \ell + 2\eta\delta \ell). \end{aligned} \quad (7)$$

Combining (5), (6) and (7) and from $\delta \leq \frac{k}{\beta R}$, we have

$$\Pr[G_j \mid E] \leq 1/(\beta\delta)(2\varepsilon_V \delta \ell + 2\eta\delta \ell + \varepsilon_T \delta + 2\varepsilon_V \delta \ell + 2\eta\delta \ell) \leq O(\varepsilon_T/\beta) + O(\eta\ell/\beta) + O(\varepsilon_V \ell/\beta).$$

2. $|W(A, x, j)| > 1$. Let i_1^* and i_2^* be two different (arbitrary) elements of $W(A, x, j)$. Again, for convenient, we use E to denote the conditional event $j^*(A, x) > j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1} \wedge \{i_1^*, i_2^*\} \subseteq W(A, x, j)$. Now, let us first split $\Pr[G_j \mid E]$ as follows.

$$\begin{aligned} \Pr[G_j \mid E] &\leq \Pr[i_1^* \in D] + \Pr[i_2^* \in D] + \Pr[G_j \mid E \wedge i_1^*, i_2^* \notin D] \\ &= 2\varepsilon_T + \Pr[G_j \mid E \wedge i_1^*, i_2^* \notin D]. \end{aligned} \quad (8)$$

Observe that, when $i_1^*, i_2^* \notin D$, $x'_{i_1^*}, x'_{i_2^*} \neq \perp$ for every $x' \in C_D(x)$. Hence, for G_j to occur, there must be $p \in [\ell]$ such that at least one of $\tilde{B}_{i_1^*}^p, \tilde{B}_{i_2^*}^p$ is not in S . In other words,

$$\begin{aligned} \Pr[G_j \mid E \wedge i_1^*, i_2^* \notin D] &\leq \Pr[\exists p \in [\ell], \tilde{B}_{i_1^*}^p \notin S \vee \tilde{B}_{i_2^*}^p \notin S \mid E \wedge i_1^*, i_2^* \notin D] \\ &\leq \sum_{p=1}^{\ell} (\Pr[\tilde{B}_{i_1^*}^p \notin S \mid E \wedge i_1^*, i_2^* \notin D] + \Pr[\tilde{B}_{i_2^*}^p \notin S \mid E \wedge i_1^*, i_2^* \notin D]) \\ (\text{From } i_1^*, i_2^* \in W(A, x, j)) &= \sum_{p=1}^{\ell} (\Pr[\tilde{B}_{i_1^*}^p \notin S \mid A_{i_1^*} \in S] + \Pr[\tilde{B}_{i_2^*}^p \notin S \mid A_{i_2^*} \in S]) \\ (\text{From (4)}) &\leq 2\varepsilon_V \ell + \eta \ell. \end{aligned} \quad (9)$$

Combining this with (8), we have $\Pr[G_j \mid E] \leq O(\varepsilon_T) + O(\eta\ell) + O(\varepsilon_V \ell)$.

As a result, we can conclude that c_2 is at most $O(\varepsilon_T/\beta) + O(\eta\ell/\beta) + O(\varepsilon_V \ell/\beta)$.

Bounding c_3

Finally, let us bound c_3 . First, note that the probability that $x_{i^*(A,x)} = 1$ is $1/2$ and that the probability that $i^*(A,x) \in D$ is ε_T . This means that

$$c_3 \leq 1/2 + \varepsilon_T + \Pr[e \notin T'_0 \mid E].$$

where E is the event $x_{i^*(A,x)} = 0 \wedge i^*(A,x) \notin D \wedge j^*(A,x) = j \wedge \neg G_1 \wedge \dots \wedge \neg G_{j-1}$.

Moreover, since $A_{i^*(A,x)} \in S$, from (4) and from union bound, we have

$$\Pr\left[\exists p \in [\ell], \tilde{B}_{i^*(A,x)}^p \notin S \mid E\right] \leq 2\varepsilon_V \ell + \eta \ell.$$

From the above two inequalities, we have

$$c_3 \leq 1/2 + \varepsilon_T + 2\varepsilon_V \ell + \eta \ell + \Pr\left[e \notin T'_0 \mid E \wedge \left(\forall p \in [\ell], \tilde{B}_{i^*(A,x)}^p \in S\right)\right].$$

Conditioned on the above event, $e \notin T'_0$ implies that there exists $p \in [\ell]$ and some $i \neq i^*(A,x)$ in this (j -th) block such that $\tilde{B}_i^p \in S$, and $x_i \neq \perp$ or $i \in D$. We have bounded an almost identical probability before in the case $W(A,x,j) = \emptyset$ when we bound c_2 . Similary, here we have an upper bound of $O(\varepsilon_T/\beta) + O(\eta \ell/\beta) + O(\varepsilon_V \ell/\beta)$ on this probability. Hence,

$$c_3 \leq 1/2 + O(\varepsilon_T/\beta) + O(\eta \ell/\beta) + O(\varepsilon_V \ell/\beta)$$

By combining our bounds on c_1, c_2, c_3 with (2), we immediately arrive at the desired bound:

$$E_H(T'_0) \geq 1/2 - O(\varepsilon_T/\beta) - O(\eta \ell/\beta) - O(\varepsilon_V \ell/\beta) - 2^{-\Omega(k)}.$$

□

4.5 Soundness

Let us consider any set T such that $\mu_H(T) \leq 1/2$. We would like to give an upper bound on $E_H(T)$. From how we define hyperedges, we can assume w.l.o.g. that $(A,x) \in T$ if and only if $\pi(A',x) \in T$ for every $A' \in M_x(A)$ and $\pi \in \Pi_{R,k}$. We call such T $\Pi_{R,k}$ -invariant.

Let $f : V^R \times \Omega^R \rightarrow \{0,1\}$ denote the indicator function for T , i.e., $f(A,x) = 1$ if and only if $(A,x) \in T$. Note that $\mathbb{E}_{A \sim V^R, x \sim \Omega^R} f(A,x) = \mu_H(T) \leq 1/2$. Following notation from [RST12], we write $f_A(x)$ as a shorthand for $f(A,x)$. In addition, for each $A \in V^R$, we will write $\tilde{B} \sim \Gamma(A)$ as a shorthand for \tilde{B} generated randomly by sampling $\tilde{A} \sim T_V(A)$, $B \sim G^{\otimes R}(\tilde{A})$ and $\tilde{B} \sim T_V(B)$ respectively. Let us restate Raghavendra et al.'s [RST12] lemma regarding the variance of $\mathbb{E}_x f_A(x)$ in a more convenient formulation below.

Lemma 13 ([RST12, Lemma 6.6]⁹) *For every $A \in V^R$, let $\mu_A \triangleq \mathbb{E}_{x \sim \Omega^R} f_A(x)$. We have*

$$\mathbb{E}_{A \sim V^R} \left(\mathbb{E}_{\tilde{B} \sim \Gamma(A)} \mu_{\tilde{B}} - \mu_H(T) \right)^2 \leq \beta.$$

To see how the above lemma helps us decode an UG assignment, observe that, if our test accepts on $f_{\tilde{B}^1}, \dots, f_{\tilde{B}^\ell}, x, D$, then it also accepts on any subset of the functions (with the same x, D); hence, to apply Theorem 11, it suffices that t of the functions have means ≤ 0.99 . We will choose ℓ to be large compared to t . Using above lemma and a standard tail bound, we can argue that Theorem 11 is applicable for almost all tuples $\tilde{B}^1, \dots, \tilde{B}^\ell$, as stated below.

Lemma 14 *For any positive integer $t \leq 0.01\ell$,*

$$\Pr_{A \sim V^R, \tilde{B}^1, \dots, \tilde{B}^\ell \sim \Gamma(A)} [|\{i \in [\ell] \mid \mu_{\tilde{B}^i} \leq 0.99\}| \geq t] \geq 1 - 10\beta - 2^{-\ell/100}.$$

⁹Lemma 6.6 in [RST12] involves symmetrizing f 's, but we do not need it here since T is $\Pi_{R,k}$ -invariant.

Proof. First, note that, since $\mu_H(T) \leq 1/2$, we can use Chebychev's inequality and Lemma 6.6 to arrive at the following bound, which is analogous to Lemma 6.7 in [RST12]:

$$\Pr_{A \sim V^R} \left[\mathbb{E}_{\tilde{B} \sim \Gamma(A)} \mu_{\tilde{B}} \geq 0.9 \right] \leq 10\beta. \quad (10)$$

Let us call $A \in V^R$ such that $\mathbb{E}_{\tilde{B} \sim \Gamma(A)} \mu_{\tilde{B}} \geq 0.9$ *bad* and the rest of $A \in V^R$ *good*.

For any good $A \in V^R$, Markov's inequality implies that $\Pr_{\tilde{B} \sim \Gamma(A)} [\mu_{\tilde{B}} > 0.99] \leq 0.9/0.99 < 0.95$. As a result, an application of Chernoff bound gives the following inequality.

$$\Pr_{\tilde{B}^1, \dots, \tilde{B}^\ell \sim \Gamma(A)} [|\{i \in [\ell] \mid \mu_{\tilde{B}^i} \leq 0.99\}| < t \mid A \text{ is good}] \leq 2^{-\ell/100}. \quad (11)$$

Finally, observe that (10) and (11) immediately yields the desired bound. \square

4.5.1 Decoding an Unique Games Assignment

With Lemma 14 ready, we can now decode an UG assignment via a similar technique from [BK09].

Lemma 15 *For any $\varepsilon_T, \gamma, \beta > 0$, let $t = t(\varepsilon_T, \gamma, \beta)$, $\kappa = \kappa(\varepsilon_T, \gamma, \beta)$ and $d = d(\varepsilon_T, \gamma, \beta)$ be as in Theorem 11. For any integer $\ell \geq 100t$, if there exists $T \subseteq V_H$ of such that $\mu_H(T) \leq 1/2$ and $E_H(T) \geq 2\gamma + 10\beta + 2^{-\ell/100}$, then there exists $F : V^R \rightarrow [R]$ such that*

$$\Pr_{A \sim V^R, \tilde{B} \sim \Gamma(A), \pi_A, \pi_B \sim \Pi_{R,k}} [\pi_A^{-1}(F(\pi_A(\tilde{A}))) = \pi_B^{-1}(F(\pi_B(\tilde{B})))] \geq \frac{\gamma \kappa^2}{4d^2 \ell^2}.$$

Proof. The decoding procedure is as follows. For each $A \in V^R$, we construct a set of candidate labels $\text{Cand}[A] \triangleq \{j \in [R] \mid \text{infl}_j^d(f_A) \geq \kappa\}$. We generate F randomly by, with probability $1/2$, setting $F(A)$ to be a random element of $\text{Cand}[A]$ and, with probability $1/2$, sampling $\tilde{B} \sim \Gamma(A)$ and setting $F(A)$ to be a random element from $\text{Cand}[\tilde{B}]$. Note that, if the candidate set is empty, then we simply pick an arbitrary assignment.

From our assumption that T is $\Pi_{R,k}$ -invariant, it follows that, for every $A \in V^R, \pi \in \Pi_{R,k}$ and $j \in [R]$, $\Pr_F[\pi^{-1}(F(\pi(A))) = j] = \Pr_F[F(A) = j]$. In other words, we have

$$\Pr_{F, A \sim V^R, \tilde{B} \sim \Gamma(A), \pi_A, \pi_B \sim \Pi_{R,k}} [\pi_A^{-1}(F(\pi_A(\tilde{A}))) = \pi_B^{-1}(F(\pi_B(\tilde{B})))] = \Pr_{F, A \sim V^R, \tilde{B} \sim \Gamma(A)} [F(\tilde{A}) = F(\tilde{B})]. \quad (12)$$

Next, note that, from how our reduction is defined, $E_H(T)$ can be written as

$$E_H(T) = \Pr_{A \sim V^R, \tilde{B}^1, \dots, \tilde{B}^\ell \sim \Gamma(A), x \sim \Omega^R, D \sim S_{\varepsilon_T}(R)} \left[\bigwedge_{i=1}^{\ell} f_{\tilde{B}^i}(C_D(x)) \equiv 1 \right].$$

From $E_H(T) \geq 2\gamma + 10\beta + 2^{-\ell/100}$ and from Lemma 14, we can conclude that

$$\Pr_{A, \tilde{B}^1, \dots, \tilde{B}^\ell, x, D} \left[\left(\bigwedge_{i=1}^{\ell} f_{\tilde{B}^i}(C_D(x)) \equiv 1 \right) \wedge \left(|\{i \in [\ell] \mid \mu_{\tilde{B}^i} \leq 0.99\}| \geq t \right) \right] \geq 2\gamma.$$

From Markov's inequality, we have

$$\gamma \leq \Pr_{A, \tilde{B}^1, \dots, \tilde{B}^\ell} \left[\Pr_{x, D} \left[\left(\bigwedge_{i=1}^{\ell} f_{\tilde{B}^i}(C_D(x)) \equiv 1 \right) \wedge \left(|\{i \in [\ell] \mid \mu_{\tilde{B}^i} \leq 0.99\}| \geq t \right) \right] \geq \gamma \right]$$

$$= \Pr_{A, \tilde{B}^1, \dots, \tilde{B}^\ell} \left[\left(\Pr_{x, D} \left[\bigwedge_{i=1}^{\ell} f_{\tilde{B}^i}(C_D(x)) \equiv 1 \right] \geq \gamma \right) \wedge \left(|\{i \in [\ell] \mid \mu_{\tilde{B}^i} \leq 0.99\}| \geq t \right) \right].$$

A tuple $(A, \tilde{B}^1, \dots, \tilde{B}^\ell)$ is said to be *good* if $\Pr_{x \sim \Omega^R, D \sim S_{\varepsilon_T}(R)} \left[\bigwedge_{i=1}^{\ell} f_{\tilde{B}^i}(C_D(x)) \equiv 1 \right] \geq \gamma$ and $|\{i \in [\ell] \mid \mu_{\tilde{B}^i} \leq 0.99\}| \geq t$. For such tuple, Theorem 11 implies that there exist $i_1 \neq i_2 \in [\ell], j \in [R]$ s.t. $\text{infl}_j^d(f_{\tilde{B}^{i_1}}), \text{infl}_j^d(f_{\tilde{B}^{i_2}}) \geq \kappa$. This means that $\text{Cand}(\tilde{B}^{i_1}) \cap \text{Cand}(\tilde{B}^{i_2}) \neq \emptyset$.

Hence, if we sample a tuple $(A, \tilde{B}^1, \dots, \tilde{B}^\ell)$ at random, and then sample two different \tilde{B}, \tilde{B}' randomly from $\tilde{B}^1, \dots, \tilde{B}^\ell$, then the tuple is good with probability at least γ and, with probability $1/\ell^2$, we have $\tilde{B} = \tilde{B}^{i_1}, \tilde{B}' = \tilde{B}^{i_2}$. This gives the following bound:

$$\Pr_{A, \tilde{B}, \tilde{B}'} [\text{Cand}(\tilde{B}) \cap \text{Cand}(\tilde{B}') \neq \emptyset] \geq \frac{\gamma}{\ell^2}.$$

Now, observe that \tilde{B} and \tilde{B}' above are distributed in the same way as if we pick both of them independently with respect to $\Gamma(A)$. Recall that, with probability $1/2$, $F(A)$ is a random element of $\text{Cand}(\tilde{B})$ where $\tilde{B} \sim \Gamma(A)$ and, with probability $1/2$, $F(\tilde{B}')$ is a random element of $\text{Cand}(\tilde{B}')$. Moreover, since the sum of degree d -influence is at most d [MOO10, Proposition 3.8], the candidate sets are of sizes at most d/κ . As a result, the above bound yields

$$\Pr_{A \sim V^R, \tilde{B}' \sim \Gamma(A)} [F(A) = F(\tilde{B}')] \geq \frac{\gamma \kappa^2}{4d^2 \ell^2},$$

which, together with (12), concludes the proof of the lemma. \square

4.5.2 Decoding a Small Non-Expanding Set

To relate our decoded UG assignment back to a small non-expanding set in G , we use the following lemma of [RST12], which roughly states that, with the right parameters, the soundness case of SSEH implies that only small fraction of constraints in the UG can be satisfied.

Lemma 16 ([RST12, Lemma 6.11]) *If there exists $F : V^R \rightarrow [R]$ such that*

$$\Pr_{A \sim V^R, \tilde{B} \sim \Gamma(A), \pi_A, \pi_B \sim \Pi_{R,k}} [\pi_A^{-1}(F(\pi_A(\tilde{A}))) = \pi_B^{-1}(F(\pi_B(\tilde{B})))] \geq \zeta,$$

then there exists a set $S \subseteq V$ with $\frac{|S|}{|V|} \in \left[\frac{\zeta}{16R}, \frac{3k}{\varepsilon_V R} \right]$ with $\Phi(S) \leq 1 - \frac{\zeta}{16k}$.

By combining the above lemma with Lemma 15, we immediately arrive at the following:

Lemma 17 *For any $\varepsilon_T, \gamma, \beta > 0$, let $t = t(\varepsilon_T, \gamma, \beta)$, $\kappa = \kappa(\varepsilon_T, \gamma, \beta)$ and $d = d(\varepsilon_T, \gamma, \beta)$ be as in Theorem 11. For any integer $\ell \geq 100t$ and any $\varepsilon_V > 0$, if there exists $T \subseteq V_H$ with $\mu_H(T) \leq 1/2$ such that $E_H(T) \geq 2\gamma + 10\beta + 2^{-\ell/100}$, then there exists a set $S \subseteq V$ with $\frac{|S|}{|V|} \in \left[\frac{\zeta}{16R}, \frac{3k}{\varepsilon_V R} \right]$ with $\Phi(S) \leq 1 - \frac{\zeta}{16k}$ where $\zeta = \frac{\gamma \kappa^2}{4d^2 \ell^2}$.*

4.6 Putting Things Together

We can now deduce inapproximability of MUCHB by simply picking appropriate parameters.

Proof of Lemma 4. The parameters are chosen as follows:

- Let $\beta = \varepsilon/30, \gamma = \varepsilon/6$, and $k = \Omega(\log(1/\varepsilon))$ so that the term $2^{-\Omega(k)}$ in Lemma 12 is $\leq \varepsilon/4$.
- Let $\varepsilon_T = O(\beta\varepsilon)$ so that the error term $O(\varepsilon_T/\beta)$ in Lemma 12 is at most $\varepsilon/4$.
- Let $t = t(\varepsilon_T, \gamma, \beta), \kappa = \kappa(\varepsilon_T, \gamma, \beta)$ and $d = d(\varepsilon_T, \gamma, \beta)$ be as in Theorem 11.
- Let $\zeta = \frac{\gamma \kappa^2}{4d^2 \ell^2}$ be as in Lemma 17 and let $\ell = \max\{100t, 1000 \log(1/\varepsilon)\}$.
- Let $\varepsilon_V = O(\varepsilon\beta/\ell)$ where so that the error term $O(\varepsilon_V \ell/\beta)$ in Lemma 12 is at most $\varepsilon/4$.
- Let $\eta = \min\{\frac{\zeta}{32k}, O(\varepsilon\beta/\ell)\}$ so that the error term $O(\eta\ell/\beta)$ in Lemma 12 is at most $\varepsilon/4$.

- Let $M = \max\{\frac{16k}{\beta\zeta}, \frac{3\beta}{\varepsilon_V}\}$.
- Finally, let $R = \frac{k}{\beta\delta}$ where $\delta = \delta(\eta, M)$ is the parameter from the SSEH (Conjecture 8).

Let $G = (V, E, w)$ be an instance of SSEH(η, δ, M) and let $H = (V_H, G_H)$ be the hypergraph resulted from our reduction. If there exists $S \subseteq V$ of size $\delta|V|$ of expansion at most η , Lemma 12 implies that there is a bisection (T_0, T_1) of V_H such that $E_H(T_0), E_H(T_1) \geq 1/2 - \varepsilon$.

As for the soundness, Lemma 17 with our choice of parameters implies that, if there exists a set $T \subseteq V_H$ with $\mu(T) \leq 1/2$ and $E_H(T_0) \geq \varepsilon$, there exists $S \subseteq V$ with $|S| \in [\frac{\delta|V|}{M}, \delta|V|M]$ whose expansion is less than $1 - \eta$. The contrapositive of this yields the soundness property. \square

5 Conclusion

In this work, we prove essentially tight inapproximability of MEB, MBB, Minimum k -Cut and DALkS based on SSEH. Our results, especially for the biclique problems, demonstrate further the applications of the hypothesis and particularly the RST technique [RST12] in proving hardness of graph problems that involve some form of expansion. Given that the technique has been employed for only a handful of problems [RST12, LRV13], an obvious but intriguing research direction is to try to utilize the technique to other problems. One plausible candidate problem to this end is the 2-Catalog Segmentation Problem [KPR04] since a natural candidate reduction for this problem fails due to a similar counterexample as in Section 4.2.

Another interesting question is to derandomize graph product used in the gap amplification step for biclique problems. For Maximum Clique, this step has been derandomized before [AFWZ95, Zuc07]; in particular, Zuckerman [Zuc07] derandomized Håstad's result [Hås01] to achieve $n^{1-\varepsilon}$ ratio NP-hardness for approximating Maximum Clique. Without going into too much detail, we would like to note that Zuckerman's result is based on a construction of dispersers with certain parameters; properties of dispersers then imply soundness of the reduction whereas completeness is trivial from the construction since Håstad's PCP has perfect completeness. Unfortunately, our PCP does not have perfect completeness and, in order to use Zuckerman's approach, additional properties are required in order to argue about completeness of the reduction.

Acknowledgement

I am grateful to Prasad Raghavendra for providing his insights on the Small Set Expansion problem and techniques developed in [RS10, RST12] and Luca Trevisan for lending his expertise in PCPs with small free bits. I would also like to thank Haris Angelidakis for useful discussions on Minimum k -Cut and Daniel Reichman for inspiring me to work on the two biclique problems. Finally, I thank ICALP 2017 anonymous reviewers for their useful comments and, more specifically, for pointing me to [BGH⁺16].

References

- [AAM⁺11] Noga Alon, Sanjeev Arora, Rajsekar Manokaran, Dana Moshkovitz, and Omri Weinstein. Inapproximability of densest k -subgraph from average case hardness. Unpublished Manuscript, 2011.
- [AC09] Reid Andersen and Kumar Chellapilla. Finding dense subgraphs with size bounds. In *WAW*, pages 25–37, 2009.
- [AFWZ95] Noga Alon, Uriel Feige, Avi Wigderson, and David Zuckerman. Derandomized graph products. *Computational Complexity*, 5(1):60–75, 1995.
- [AHI02] Yuichi Asahiro, Refael Hassin, and Kazuo Iwama. Complexity of finding dense subgraphs. *Discrete Appl. Math.*, 121(1–3):15 – 26, 2002.
- [ALM⁺98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, May 1998.

- [AMS11] Christoph Ambühl, Monaldo Mastrolilli, and Ola Svensson. Inapproximability results for maximum edge biclique, minimum linear arrangement, and sparsest cut. *SIAM J. Comput.*, 40(2):567–596, April 2011.
- [And07] Reid Andersen. Finding large and small dense subgraphs. *CoRR*, abs/cs/0702032, 2007.
- [AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, January 1998.
- [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 *STOC*, pages 201–210, 2010.
- [BCV⁺12] Aditya Bhaskara, Moses Charikar, Aravindan Vijayaraghavan, Venkatesan Guruswami, and Yuan Zhou. Polynomial integrality gaps for strong SDP relaxations of densest k -subgraph. In *SODA*, pages 388–405, 2012.
- [BGH⁺16] Amey Bhangale, Rajiv Gandhi, Mohammad Taghi Hajiaghayi, Rohit Khandekar, and Guy Kortsarz. Bicovering: Covering edges with two small subsets of vertices. In *ICALP*, pages 6:1–6:12, 2016.
- [BK09] Nikhil Bansal and Subhash Khot. Optimal long code test with one free bit. In *FOCS*, pages 453–462, 2009.
- [BKRW17] Mark Braverman, Young Kun Ko, Aviad Rubinstein, and Omri Weinstein. ETH hardness for densest- k -subgraph with perfect completeness. In *SODA*, pages 1326–1341, 2017.
- [Blu91] Avrim Blum. Algorithms for approximate graph coloring. Technical report, Cambridge, MA, USA, 1991.
- [BS92] Piotr Berman and Georg Schnitger. On the complexity of approximating the independent set problem. *Inf. Comput.*, 96(1):77–94, 1992.
- [CMMV17] Eden Chlamták, Pasin Manurangsi, Dana Moshkovitz, and Aravindan Vijayaraghavan. Approximation algorithms for label cover and the log-density threshold. In *SODA*, pages 900–919, 2017.
- [Din16] Irit Dinur. Mildly exponential reduction from gap 3SAT to polynomial-gap label-cover. *ECCC*, 23:128, 2016.
- [DS14] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *STOC*, pages 624–633, 2014.
- [Fei02] Uriel Feige. Relations between average case complexity and approximation complexity. In *STOC*, pages 534–543, 2002.
- [FK04] Uriel Feige and Shimon Kogan. Hardness of approximation of the balanced complete bipartite subgraph problem. Technical report, Weizmann Institute of Science, Israel, 2004.
- [FKP01] Uriel Feige, Guy Kortsarz, and David Peleg. The dense k -subgraph problem. *Algorithmica*, 29(3):410–421, 2001.
- [FL01] Uriel Feige and Michael Langberg. Approximation algorithms for maximization problems arising in graph partitioning. *J. Algorithms*, 41(2):174–211, November 2001.
- [FS97] Uriel Feige and Michael Seltser. On the densest k -subgraph problem. Technical report, Weizmann Institute of Science, Rehovot, Israel, 1997.
- [GH94] Olivier Goldschmidt and Dorit S Hochbaum. A polynomial algorithm for the k -cut problem for fixed k . *Mathematics of Operations Research*, 19(1):24–37, 1994.
- [GJ79] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1979.

- [GL09] Doron Goldstein and Michael Langberg. The dense k subgraph problem. *CoRR*, abs/0912.5327, 2009.
- [Hås96] Johan Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. In *FOCS*, pages 627–636, 1996.
- [Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, July 2001.
- [IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, December 2001.
- [Jer92] Mark Jerrum. Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4), 1992.
- [Joh87] David S. Johnson. The NP-completeness column: An ongoing guide. *J. Algorithms*, 8(5):438–448, September 1987.
- [Kho01] Subhash Khot. Improved inapproximability results for MaxClique, chromatic number and approximate graph coloring. In *FOCS*, pages 600–609, 2001.
- [Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *STOC*, pages 767–775, 2002.
- [Kho06] Subhash Khot. Ruling out PTAS for graph min-bisection, dense k -subgraph, and bipartite clique. *SIAM J. Comput.*, 36(4), 2006.
- [KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O’Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007.
- [KP93] Guy Kortsarz and David Peleg. On choosing a dense subgraph (extended abstract). In *FOCS*, pages 692–701, 1993.
- [KP06] Subhash Khot and Ashok Kumar Ponnuswami. Better inapproximability results for MaxClique, chromatic number and Min-3Lin-Deletion. In *ICALP*, pages 226–237, 2006.
- [KPR04] Jon Kleinberg, Christos Papadimitriou, and Prabhakar Raghavan. Segmentation problems. *J. ACM*, 51(2):263–280, March 2004.
- [KR08] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2 - \varepsilon$. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008.
- [KS09] Samir Khuller and Barna Saha. On finding dense subgraphs. In *ICALP*, pages 597–608, 2009.
- [Kuč95] Luděk Kučera. Expected complexity of graph partitioning problems. *Discrete Appl. Math.*, 57(2-3):193–212, February 1995.
- [LRV13] Anand Louis, Prasad Raghavendra, and Santosh Vempala. The complexity of approximating vertex expansion. In *FOCS*, pages 360–369, 2013.
- [Man15] Pasin Manurangsi. On approximating projection games. Master’s thesis, Massachusetts Institute of Technology, Cambridge, MA, January 2015.
- [Man17a] Pasin Manurangsi. Almost-polynomial ratio ETH-hardness of approximating densest k -subgraph. In *STOC*, 2017. To appear.
- [Man17b] Pasin Manurangsi. Inapproximability of maximum edge biclique, maximum balanced biclique and minimum k -cut from the small set expansion hypothesis. In *ICALP*, 2017. To appear.
- [MOO10] Elchanan Mossel, Ryan O’Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. Math.*, pages 295–341, 2010.
- [Mos15] Dana Moshkovitz. The projection games conjecture and the NP-hardness of ln n -approximating set-cover. *Theory of Computing*, 11:221–235, 2015.
- [MR17] Pasin Manurangsi and Prasad Raghavendra. A birthday repetition theorem and complexity of approximating dense CSPs. In *ICALP*, 2017. To appear.

- [NR01] Joseph (Seffi) Naor and Yuval Rabani. Tree packing and approximating k -cuts. In *SODA*, pages 26–27, 2001.
- [O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.
- [Pee03] René Peeters. The maximum edge biclique problem is NP-complete. *Discrete Appl. Math.*, 131(3):651–654, September 2003.
- [RS02] R. Ravi and Amitabh Sinha. Approximating k -cuts via network strength. In *SODA*, pages 621–622, 2002.
- [RS10] Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *STOC*, pages 755–764, 2010.
- [RST12] Prasad Raghavendra, David Steurer, and Madhur Tulsiani. Reductions between expansion problems. In *CCC*, pages 64–73, 2012.
- [SV95] Huzur Saran and Vijay V. Vazirani. Finding k cuts within twice the optimal. *SIAM J. Comput.*, 24(1):101–108, February 1995.
- [Sve13] Ola Svensson. Hardness of vertex deletion and project scheduling. *Theory of Computing*, 9(24):759–781, 2013.
- [SW98] Anand Srivastav and Katja Wolf. Finding dense subgraphs with semidefinite programming. In *APPROX 1998*, pages 181–191, 1998.
- [XYC11] Mingyu Xiao, Leizhen Cai, and Andrew Chi-Chih Yao. Tight approximation ratio of a general greedy splitting algorithm for the minimum k -way cut problem. *Algorithmica*, 59(4):510–520, 2011.
- [ZNI01] Liang Zhao, Hiroshi Nagamochi, and Toshihide Ibaraki. Approximating the minimum k -way cut in a graph via minimum 3-way cuts. *J. Comb. Optim.*, 5(4):397–410, 2001.
- [Zuc07] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3(6):103–128, 2007.

A Reduction from MUCHB to Biclique Problems

Proof of Lemma 9 from Lemma 4. The reduction from MUCHB to the two biclique problems are simple. Given a hypergraph $H = (V_H, E_H)$, we create a bipartite graph $G = (L, R, E)$ by letting $L = R = E_H$ and creating an edge (e_1, e_2) for $e_1 \in L, e_2 \in R$ iff $e_1 \cap e_2 = \emptyset$.

Completeness. If there is a bisection (T_0, T_1) of V_H such that $|E_H(T_0)|, |E_H(T_1)| \geq (1/2 - \varepsilon)|E_H|$, then $E_H(T_0) \subseteq L$ and $E_H(T_1) \subseteq R$ induce a complete bipartite subgraph with at least $(1/2 - \varepsilon)n$ vertices on each side in G .

Soundness. Suppose that, for every $T \subseteq V_H$ of size at most $|V_H|/2$, we have $|E_H(T)| \leq \varepsilon|V_H|$. We will show that G does not contain $K_{\varepsilon n+1, \varepsilon n+1}$ as a subgraph. Suppose for the sake of contradiction that G contains $K_{\varepsilon n+1, \varepsilon n+1}$ as a subgraph. Consider one such subgraph; let $e_1, \dots, e_{\varepsilon n+1} \in L$ and $e'_1, \dots, e'_{\varepsilon n+1} \in R$ be the vertices of the subgraph. From how the edges are defined, $T_0 \triangleq e_1 \cup \dots \cup e_{\varepsilon n+1}$ and $T_1 \triangleq e'_1 \cup \dots \cup e'_{\varepsilon n+1}$ are disjoint. At least one of the two sets is of size at most $|V_H|/2$, assume without loss of generality that $|T_0| \leq |V_H|/2$. This is a contradiction since T_0 contains at least $\varepsilon n + 1$ hyperedges $e_1, \dots, e_{\varepsilon n+1}$.

By picking $\varepsilon = \delta/2$ and $n > 2/\varepsilon$, we arrive at the desired hardness result. \square

B Gap Amplification via Randomized Graph Product

In this section, we provide a full proof of the gap amplification step for biclique problems, thereby proving Theorem 3. The argument provided below is almost the same as the randomized graph product based analysis

of Khot [Kho06, Appendix D], which is in turn based on the analysis of Berman and Schnitger [BS92], except that we modify the construction slightly so that the reduction time is polynomial¹⁰.

Specifically, we will prove the following statement, which immediately implies Theorem 3.

Lemma 18 *Assuming SSEH and $NP \neq BPP$, for every $\varepsilon > 0$, no polynomial time algorithm can, given a bipartite graph $G' = (L', R', E')$ with $|L'| = |R'| = N$, distinguish between the following two cases:*

- (Completeness) G' contains $K_{N^{1-\varepsilon}, N^{1-\varepsilon}}$ as a subgraph.
- (Soundness) G' does not contain $K_{N^\varepsilon, N^\varepsilon}$ as a subgraph.

Proof. Let $\varepsilon > 0$ be any constant. Assume without loss of generality that $\varepsilon < 1$. Consider any bipartite graph $G = (L, R, E)$ with $|L| = |R| = n$. Let $\delta = 2^{-4/\varepsilon}$, $k = \log n$ and $N = (1/\delta)^k$. We construct a graph $G' = (L', R', E')$ where $|L'| = |R'| = N$ as follows. For $i \in [N]$, pick random elements $U^i \sim L^k$ and $V^i \sim R^k$, and add them to L' and R' respectively. Finally, for every $U \in L'$ and $V \in R'$, there is an edge between U and V in G' if and only if there is an edge in the original graph G between U_{j_1} and V_{j_2} for every $j_1, j_2 \in [k]$.

Completeness. Suppose that the original graph G contains $K_{(1/2-\delta)n, (1/2-\delta)n}$ as a subgraph. Let one such subgraph be (S, T) where $S \subseteq L$ and $T \subseteq R$. Observe that $(L' \cap S^k, R' \cap T^k)$ induces a biclique in G' . For each $i \in [N]$, note that $U_i \in S^k$ with probability $(1/2 - \delta)^k \geq (1/4)^k = N^{-\varepsilon/2}$ independent of each other. As a result, from Chernoff bound, $|L' \cap S^k| \geq N^{1-\varepsilon}$ with high probability. Similarly, we also have $|R' \cap T^k| \geq N^{1-\varepsilon}$ with high probability. Thus, G' contains $K_{N^{1-\varepsilon}, N^{1-\varepsilon}}$ as a subgraph with high probability.

Soundness. Suppose that G does not contain $K_{\delta n, \delta n}$ as a subgraph. We will show that, with high probability, G' does not contain $K_{N^\varepsilon, N^\varepsilon}$ as a subgraph. To do so, we will first prove the following proposition.

Proposition 19 *For any set A , let $\mathcal{P}(A)$ denote the power set of A . Moreover, let $\mathcal{F} : \mathcal{P}(L^k \cup R^k) \rightarrow \mathcal{P}(L \cup R)$ be the “flattening” operation defined by $\mathcal{F}(A) \triangleq \bigcup_{U \in A} \{U_i \mid i \in [k]\}$. Then, with high probability, we have $|\mathcal{F}(S')| \geq \delta n$ for every subset $S' \subseteq L'$ of size N^ε and $|\mathcal{F}(T')| \geq \delta n$ for every subset $T' \subseteq R'$ of size N^ε .*

Proof of Proposition 19. Let us consider the probability that there exists a set $S' \subseteq L'$ of size N^ε such that $|\mathcal{F}(S')| < \delta n$. This can be bounded as follows.

$$\begin{aligned} \Pr[\exists S' \subseteq L', |S'| = N^\varepsilon \wedge \mathcal{F}(S') < \delta n] &= \Pr[\exists S \subseteq L, |S| < \delta n \wedge |S^k \cap L'| \geq N^\varepsilon] \\ &\leq \sum_{S \subseteq L, |S| < \delta n} \Pr[|S^k \cap L'| \geq N^\varepsilon] \\ &= \sum_{S \subseteq L, |S| < \delta n} \Pr \left[\sum_{i \in [N]} \mathbb{1}[U_i \in S^k] \geq N^\varepsilon \right]. \end{aligned}$$

Observe that, for each $i \in [N]$, $\mathbb{1}[U_i \in S^k]$ is simply an independent Bernoulli random variable with mean $(|S|/n)^k < \delta^k = 1/N$. Hence, by Chernoff bound, we have

$$\Pr[\exists S' \subseteq L', |S'| = N^\varepsilon \wedge \mathcal{F}(S') < \delta n] \leq \sum_{S \subseteq L, |S| < \delta n} 2^{-\Omega(N^\varepsilon)} \leq \sum_{S \subseteq L, |S| < \delta n} 2^{-\Omega(n^2)} = 2^{-\Omega(n^2)}$$

as desired.

Analogously, we also have $|\mathcal{F}(T')| \geq \delta n$ for every subset $T' \subseteq R'$ of size N^ε with high probability, thereby concluding the proof of Proposition 19. \square

With Proposition 19 ready, let us proceed with our soundness proof. Suppose that the event in Proposition 19 occurs. Consider any subset $S' \subseteq L'$ of size N^ε and any subset $T' \subseteq R'$ of size N^ε . Since $|\mathcal{F}(S')| \geq \delta n$, $|\mathcal{F}(T')| \geq \delta n$ and G does not contain $K_{\delta n, \delta n}$ as a subgraph, there exists $u \in \mathcal{F}(S')$ and $v \in \mathcal{F}(T')$ such that $(u, v) \notin E$. From the definition of G' , this implies that S' and T' do not induce a biclique in G' . As a

¹⁰Using Khot’s result directly would result in the construction time being quasi-polynomial.

result, G' does not contain $K_{N^\varepsilon, N^\varepsilon}$ as a subgraph. From this and from Proposition 19, G' does not contain $K_{N^\varepsilon, N^\varepsilon}$ as a subgraph with high probability, concluding our soundness argument.

Since Lemma 9 asserts that distinguishing between the two cases above are NP-hard (assuming SSEH) and the above reduction takes polynomial time, we can conclude that, assuming SSEH and $\text{NP} \neq \text{BPP}$, no polynomial time algorithm can distinguish the two cases stated in the lemma. \square

C Comparison Between SSEH and Strong UGC

In this section, we briefly discuss the similarities and differences between the classical Unique Games Conjecture [Kho02], the Small Set Expansion Hypothesis [RS10] and the Strong Unique Games Conjecture [BK09]. Let us start by stating the Unique Games Conjecture, proposed by Khot in his influential work [Kho02]:

Conjecture 20 (Unique Games Conjecture (UGC) [Kho02]) *For every $\varepsilon, \eta > 0$, there exists $R = R(\varepsilon, \eta)$ such that, given an UG instance $(\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W}), [R], \{\pi_e\}_{e \in \mathcal{E}})$ such that \mathcal{G} is regular, it is NP-hard to distinguish between the following two cases:*

- (Completeness) There exists an assignment $F : \mathcal{V} \rightarrow [R]$ such that $\text{val}_U(F) \geq 1 - \varepsilon$.
- (Soundness) For every assignment $F : \mathcal{V} \rightarrow [R]$, $\text{val}_U(F) \leq \eta$.

In other words, Khot's UGC states that it is NP-hard to distinguish between an UG instance which is almost satisfiable from one in which only small fraction of edges can be satisfied. While SSEH as stated in Conjecture 2 is not directly a statement about an UG instance, it has a strong connection with the UGC. Raghavendra and Steurer [RS10], in the same work in which they proposed the conjecture, observes that SSEH is implied by a variant of UGC in which the soundness is strengthened so that the constraint graph is also required to be a small-set expander (i.e. every small set has near perfect edge expansion). In a subsequent work, Raghavendra, Steurer and Tulsiani [RST12] showed that the two conjectures are in fact equivalent. More formally, the following variant of UGC is equivalent to SSEH:

Conjecture 21 (UGC with Small-Set Expansion (UGC with SSE) [RS10]) *For every $\varepsilon, \eta > 0$, there exist $\delta = \delta(\varepsilon) > 0$ and $R = R(\varepsilon, \eta)$ such that, given an UG instance $(\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W}), [R], \{\pi_e\}_{e \in \mathcal{E}})$ such that \mathcal{G} is regular, it is NP-hard to distinguish between the following two cases:*

- (Completeness) There exists an assignment $F : \mathcal{V} \rightarrow [R]$ such that $\text{val}_U(F) \geq 1 - \varepsilon$.
- (Soundness) For every assignment $F : \mathcal{V} \rightarrow [R]$, $\text{val}_U(F) \leq \eta$. Moreover, \mathcal{G} satisfies $\Phi(S) \geq 1 - \varepsilon$ for every $S \subseteq \mathcal{V}$ of size δn .

While our result is based on SSEH (which is equivalent to UGC with SSE), Bhangale et al. [BGH⁺16] relies on another strengthened version of the UGC, which requires the following additional properties:

- There is not only an assignment that satisfies almost all constraints, but also a partial assignment to almost the whole graph such that every constraint between two assigned vertices is satisfied.
- The graph in the soundness case has to satisfy the following vertex expansion property: for every not too small subset of \mathcal{V} , its neighborhood spans almost the whole graph.

More formally, the conjecture can be stated as follows.

Conjecture 22 (Strong UGC (SUGC) [BK09]) *For every $\varepsilon, \eta, \delta > 0$, there exists $R = R(\varepsilon, \eta, \delta)$ such that, given an UG instance $(\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W}), [R], \{\pi_e\}_{e \in \mathcal{E}})$ such that \mathcal{G} is regular, it is NP-hard to distinguish between the following two cases:*

- (Completeness) There exists a subset $S \subseteq \mathcal{V}$ of size at least $(1 - \varepsilon)|\mathcal{V}|$ and a partial assignment $F : S \rightarrow [R]$ such that every edge inside S is satisfied.
- (Soundness) For every assignment $F : \mathcal{V} \rightarrow [R]$, $\text{val}_U(F) \leq \eta$. Moreover, \mathcal{G} satisfies $|\Gamma(S)| \geq (1 - \delta)|\mathcal{V}|$ for every $S \subseteq \mathcal{V}$ of size δn where $\Gamma(S)$ denote the set of all neighbors of S .

The conjecture was first formulated by Bansal and Khot [BK09]. We note here that the name “Strong UGC” was not given by Bansal and Khot, but was coined by Bhangale et al. [BGH⁺16]. In fact, the name “Strong UGC” was used earlier by Khot and Regev [KR08] to denote a different variant of UGC, in which

the completeness is strengthened to be the same as in Conjecture 22 but the soundness does not include the vertex expansion property. Interestingly, this variant of UGC is equivalent to the original version of the conjecture [KR08]. Moreover, as pointed out in [BK09], it is not hard to see that the soundness property of SUGC can also be achieved by simply adding a complete graph with negligible weight to the constraint graph. In other words, both the completeness and soundness properties of SUGC can be achieved separately. However, it is not known whether SUGC is implied by UGC.

To the best of our knowledge, it is not known if one of Conjecture 21 and Conjecture 22 implies the other. In particular, while the soundness cases of both conjectures require certain expansion properties of the graphs, Conjecture 21 deals with edge expansion whereas Conjecture 22 deals with vertex expansion; even though these notations are closely related, they do not imply each other. Moreover, as pointed out earlier, the completeness property of SUGC is stronger than that of UGC with SSE; we are not aware of any reduction from SSE to UG that achieves this while maintaining the same soundness as in Conjecture 21.

Finally, we note that both soundness and completeness properties of SUGC are crucial for Bhangale et al.'s reduction [BGH⁺16]. Hence, it is unlikely that their technique applies to SSEH. Similarly, our reduction relies crucially on edge expansion properties of the graph and, thus, is unlikely to be applicable to SUGC.