A Characterization of Approximability for Biased CSPs

Euiwoong Lee∗
University of Michigan
Ann Arbor, United States
euiwoong@umich.edu

Suprovat Ghoshal∗
University of Michigan
Ann Arbor, United States
suprovat@umich.edu

ABSTRACT
A $\mu$-biased Max-CSP instance with predicate $\psi : \{0, 1\}^r \rightarrow \{0, 1\}$ is an instance of Constraint Satisfaction Problem (CSP) where the objective is to find a labeling of relative weight at most $\mu$ which satisfies the maximum fraction of constraints. Biased CSPs are versatile and express several well-studied problems such as Densest-$k$-Sub(Hyper)graph and SmallSetExpansion.

In this work, we explore the role played by the bias parameter $\mu$ on the approximability of biased CSPs. We show that the approximability of such CSPs can be characterized (up to loss of factors in arity $r$) using the bias-approximation curve of Densest-$k$-SubHypergraph (D&S). In particular, this gives a tight characterization of predicates which admit approximation guarantees that are independent of the bias parameter $\mu$.

Motivated by the above, we give new approximation and hardness results for D&S. In particular, assuming the Small Set Expansion Hypothesis (SSEH), we show that D&S with arity $r$ and $k = \mu r$ is NP-hard to approximate to a factor of $\Omega((2^r/2r)^{r-1}\log(1/\mu))$ for every $r \geq 2$ and $\mu < 2^{-r}$. We also give a $O(\mu r^{-1}\log(1/\mu))$-approximation algorithm for the same setting. Our upper and lower bounds are tight up to constant factors, when the arity $r$ is a constant, and in particular, imply the first tight approximation bounds for the Densest-$k$-Subgraph problem in the linear bias regime. Furthermore, using the above characterization, our results also imply matching algorithms and hardness for every biased CSP of constant arity.

CSC CONCEPTS
• Theory of computation → Problems, reductions and completeness: Approximation algorithms analysis.

KEYWORDS
Constraint Satisfaction Problems, Inapproximability, Densest-$k$-SubHypergraph

ACM Reference Format:
Euiwoong Lee and Suprovat Ghoshal. 2022. A Characterization of Approximability for Biased CSPs. In Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC ’22), June 20–24, 2022, Rome, Italy.

1 INTRODUCTION
Constraint Satisfaction Problems (CSPs) are a class of extensively studied combinatorial optimization problems in theoretical computer science. Typically, an instantiation of an $r$-CSP $\Psi(V, E, [R])$, $(\Pi_e)_{e \in E}$ is characterized by an underlying $r$-ary hypergraph $\Psi = (V, E)$ with label set $[R]$, and a constraint $\Pi_e \subset [R]^r$ for every edge $e \in E$. The objective is to find a labeling $\sigma : V \rightarrow [R]$ that satisfies the maximum fraction of constraints — here, the labeling $\sigma$ satisfies a hyperedge $e = (v_1, \ldots, v_r)$ if $(\sigma(v_1), \sigma(v_2), \ldots, \sigma(v_r)) \in \Pi_e$. The expressive power of CSPs is evident from the long list of fundamental and well-studied combinatorial optimization problems that can be expressed as a CSP: Max-Cut [15,31], Coloring [17, 37], Unique Games [18, 21] are all examples of CSPs, each of which has been studied extensively by themselves (see [25] for a comprehensive overview). The tight interplay between CSPs and Probabilistically Checkable Proofs (PCPs) has led to a line of works spanning decades resulting in substantial progress in the theory of approximation of CSPs, eventually leading to landmark results such as tight upper and lower bounds for every CSP assuming the Unique Games Conjecture [32].

A well-studied variant of CSPs are CSPs with cardinality constraints i.e., CSPs where there are global constraints on the relative weight of vertices that can be assigned a particular label. Perhaps one of the simplest instantiations of a CSP with a cardinality constraint is the Densest-$k$-Subgraph (D&S) problem. Here, given an undirected graph $G = (V, E)$ and a parameter $k \in \mathbb{N}$, the objective is to find a subset of $k$-vertices such that the number edges induced by the subset is maximized. It is easy to see that this is an instantiation of a Boolean CSP of arity 2 with the underlying graph as $G$, the edge constraints being AND, and the global cardinality constraint is that exactly $k$ vertices of the CSP can be assigned the label 1. Furthermore, this is also a relaxation of the $k$-CLIQUE problem, and consequently, it is not surprising that there has been many works which study its approximability [5, 6, 12, 26].

While the unconstrained version of this problem – i.e., MAX-AND with no negations – is trivially polynomial time solvable, the additional simple cardinality constraint makes the problem significantly harder. In particular, Raghavendra and Steurer [33] showed that assuming the Small Set Expansion Hypothesis (SSEH), D&S is NP-hard to approximate to any constant factor. Furthermore, Manurangsi [26] showed that assuming the Exponential Time Hypothesis (ETH), there is no polynomial time algorithm which can approximate D&S up to an almost polynomial ratio. A similar phenomena was also observed by Austrin and Stankovic [4] for the setting of cardinality constrained Max-Cut as well. Furthermore, the nature of how the approximability of the CSP is affected is also predicated...
dependent. For instance, while in the case of DkS, the cardinality constraint makes it constant factor inapproximable for any constant, in the case of MAX-CUT, there exists a 0.858-approximation factor for any \( k = \mu n \) with \( \mu \in (0, 1) \). Hence we are motivated to ask the following question:

> Can we characterize predicates which admit approximation factors which are independent of the bias parameter \( \mu \)?

In this work, we focus on understanding the above phenomena at a more fine grained level. In particular, we aim to explicitly quantify the role of the bias parameter \( \mu := k/n \) in the approximability of a Boolean CSP with a cardinality constraint. Formally, for any \( \mu \in (0, 1) \), the \( \mu \)-biased instance of a Boolean CSP \( \Psi(V, E) \) with \( r \)-ary predicate \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \) (denoted by \( \Psi(\mu)(V, E) \)) is an instantiation where the objective is to find a labeling \( \sigma : V \rightarrow \{0, 1\} \) of relative weight at most \( \mu \) which satisfies the maximum fraction of constraint. Furthermore, let \( \alpha_{\leq \mu}(\psi) \) be referred to as the bias-approximation curve – denote the optimal approximation factor efficiently \(^4\) achievable for \( \mu \)-biased vertex weighted instances on predicate \( \psi \). Given this setup, it is natural to ask the following:

> Can we give matching upper and lower bounds for \( \alpha_{\leq \mu}(\psi) \) for every constant bias \( \mu \) and predicate \( \psi \)?

The above, despite being a natural question, has only been studied for very specific instantiations of \( \psi \) such as \( \text{Max-k-Vertex Cover} \) [26], \( \text{MAX-CUT} \) [4]. Furthermore, tight lower bounds are known for even fewer settings such as the almost satisfiable regimes of \( \text{SmallSetExpansion} \), \( \text{Max-Bisection} \), \( \text{BalancedSeparator} \) [35], and as such, a finer understanding of \( \alpha_{\leq \mu}(\cdot) \) is absent even for natural problems such as DkS.

### 1.1 Our Main Results

In this work, we make substantial progress towards answering the above questions. In order to formally state our results, we need to introduce some additional notation. Given a predicate \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \), let \( \psi^{-1}(1) \) be the set of accepting strings for predicate \( \psi \). Let \( M_\psi \) denote the set of minimal elements of \( \psi^{-1}(1) \) under the ordering imposed by the containment relationship\(^2\). We will think of instances of MAX-CSPs as being vertex weighted, and an instance of \( \mu \)-biased CSP with predicate \( \psi \) is one where the objective is to find a global labeling of the vertices with relative weight \(^3\) at most \( \mu \) which satisfies the maximum fraction of constraints. Furthermore, we say that a predicate \( \psi \) is bias dependent if \( \inf_{\mu \in (0, 1/2)} \alpha_{\leq \mu}(\psi) = 0 \). For any \( i \geq 2 \), we use DkSH\(_i\) to denote the DENSEST-\( k \)-SUBGRAPH problem on hypergraphs of arity \( i \). Finally, for any \( i \in \mathbb{Z}_{\geq 2} \), we use \( a^{\text{upper}}_\nu(DkSH_i) \) to denote the bias approximation curve of uniformly weighted DkSH\(_i\) instance.

Our first result is the following theorem which completely characterizes predicates which are bias dependent.

**Theorem 1.1.** The following holds assuming SSEH. A predicate \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \) is bias independent if and only if \( M_\psi \subseteq S_{\leq 1} \).

\(^4\)Here, we say that an \( \alpha \)-factor approximation is efficiently achievable if the problem of finding an \( \alpha \)-approximate solution to the biased MAX-CSP problem is in \( P \).

\(^2\)Here the containment relationship refers to the containment relationship induced by interpreting the Boolean strings as indicators of subsets.

\(^3\)Given a labeling \( \sigma : V \rightarrow \{0, 1\} \), its relative weight with respect to vertex weight function \( w : V \rightarrow \{0, 1\} \) is defined as \( w(\sigma) := \sum_{i \sigma(i) = 1} w(i)/w(V) \), where \( w(V) \) denotes the total vertex weight.

where \( S_{\leq 1} \) is the set of \( r \)-length strings of Hamming weight at most 1.

As a useful exercise, we instantiate the above theorem for \( \psi := \text{NEQ} \) (i.e., Biased MAX-CUT) and \( \psi := \text{AND} \) (i.e., DkS). Note that the NEQ\(^{-1}(1) = \{(0, 1), (1, 0)\} \subset S_{\leq 1} \), where as AND\(^{-1}(1) = \{(1, 1)\} \notin S_{\leq 1} \), which using Theorem 1.1 implies that the former admits an independent approximation factor, whereas the latter would be bias dependent. Our next theorem gives an unconditional tight characterization (up to factors of \( r \)) of the bias-approximation curve of a predicate in terms of DkSH.

**Theorem 1.2.** For every integer \( r \geq 2 \) and for any \( \mu \in (0, 1/2) \), the following holds for any predicate \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \):

\[
\alpha_{\leq \mu}(\psi) \asymp_r \min_{\beta \in M_\psi} \mu^\beta \cdot (\text{DkSH}_i)^{(\beta)}
\]

where \( \| \cdot \|_0 \) denotes the Hamming weight of a string, and \( \asymp_r \) is used to denote that the two sides are equal up to multiplicative factors depending on \( r \).

**Hardness and Approximation for DkSH.** Theorem 1.2 directly implies that we can reduce the task of understanding the bias-approximation curve of general Boolean predicates to that of DkSH, up to loss of multiplicative factors dependent on \( r \). We introduce an additional notation: let \( \delta^{(r)}_{\mu}(H) \) denotes the optimal value of DkSH\(_r\) with bias \( \mu \) on hypergraph \( H \) of arity \( r \).

Our first result here is the following theorem which gives the first bias dependent hardness for DkSH.

**Theorem 1.3.** The following holds assuming SSEH for every \( r \geq 2 \) and \( \mu < 2^{-r} \). Given a hypergraph \( H = (V, E) \) of arity \( r \), it is \( \text{NP-hard} \) to distinguish between the following two cases:

**YES Case:** \( \delta^{(r)}_{\mu}(H) \geq \frac{\mu}{r^2 \log(1/\mu)} \)

**NO Case:** \( \delta^{(r)}_{\mu}(H) \leq \mu^2 \).

The above theorem implies that DkSH\(_r\) with bias parameter \( \mu \) is hard to approximate up to factor a \( \Theta(r^2) \cdot \mu^{-2} \log(1/\mu) \). We complement the above hardness result with the following theorem which gives bias dependent approximation for DkSH.

**Theorem 1.4.** The following holds for any \( \mu \in (0, 1) \) and \( r \geq 2 \). There exists a randomized polynomial time algorithm which on input a hypergraph \( H = (V, E) \) of arity \( r \) returns a set \( S \subset V \) such that \( |S| = \mu n \) and \( |E_H[S]| \geq C_\mu^{-1} \log(1/\mu) \cdot \delta^{(r)}_{\mu}(H) \).

The upper and lower bounds on the optimal approximation factor from the above theorems are tight up to factor \( O(r^2) \), and are therefore tight up to multiplicative constants for constant \( r \). In particular, for the setting of \( r = 2 \) i.e. \( \text{DENSEST-k-SUBGRAPH} \), the above imply the tight approximation bound of \( \Theta(\mu \log(1/\mu)) \). Finally, Theorems 1.2, 1.3 and 1.4 together imply the following corollary which gives tight bias dependent approximation bounds for every constant \( r \).

**Corollary 1.5.** The following holds for any predicate \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \) assuming SSEH.

\[
\alpha_{\leq \mu}(\psi) \asymp_r \min_{\beta \in M_\psi} \mu^\beta (\beta)_{\leq 1} \log(1/\mu).
\]
Remark 1.6. The above results also generalize readily to the setting where the variables are allowed to be negated by applying the above results (Theorems 1.1 and 1.2) to each of the $2^k$ predicates obtained by applying the $2^k$ negation patterns to the literals.

Remark 1.7. We point out that in our setting, we allow algorithms to output solutions with relative weight slightly larger than $\mu$, say $\mu(1+\eta)$, where $\eta$ is a constant. This additional multiplicative slack is indeed necessary as Theorem 1.2 does not hold in the case where algorithms are constrained to output a solution of relative weight at most $\mu$. This is mainly due to the observation that in general, weighted instances of $k$-SH can be much harder than unweighted instances (even for the same $k$). Allowing a constant multiplicative slack in the relative weight enables us to bypass this technical difficulty. Furthermore, we note that our upper bound for the bias approximation curve (i.e., the hardness) also holds for algorithms which are allowed this multiplicative slack – for details we refer the readers to the full version of our paper [14].

Application to $\text{Max-k-CSP}$. Extending our techniques from Theorem 1.3, we also prove the following new approximation bound for $\text{Max-k-CSP}$s in the large alphabet regime.

Theorem 1.8. The following holds assuming the Unique Games Conjecture, for every $k \geq 2$ and $R \geq 2^k$. Given a $\text{Max-k-CSP}$ instance $\Psi(V, E, [R], \{\Pi_e\}_{e \in E})$, it is NP-hard to distinguish between the following two cases:

**YES Case**: $\text{Opt}(\Psi) \geq \frac{C_1}{k^2 \log(R)}$

**NO Case**: $\text{Opt}(\Psi) \leq \frac{C_2}{R^{k^2}}$

where $C_1, C_2 > 0$ are absolute positive constants independent of $R$ and $k$.

The above implies that $\text{Max-k-CSP}$s on label sets $[R]$ are Unique Games hard to approximate up to a factor of $\Omega(k^2 R^{-(k-1) \log(R)})$, this improves on the previous lower bound $\Omega(k^2 R^{-(k-1) \log(R)})$ implicit in the work of Khot and Saket [22]. Furthermore, since [28] gave a $O(R^{-(k-1) \log(R)})$-approximation algorithm for $\text{Max-k-CSP}$, Theorem 1.8 is tight up to factor of $O(k^2)$, and in particular is tight for all constant $k$.

1.2 Related Works

$\text{CSPs with Global Constraints}$. There have been several works which study specific instances of CSPs with global constraints. Of particular interest is the $\text{Max-Bisection}$ problem which is $\text{MAX-CUT}$ with a global bisection constraint. The question of whether $\text{Max-Bisection}$ is strictly harder than $\text{Max-CUT}$ has been a tantalizingly open question that has been studied by several works [13][35][39][36], the current best known approximation factor being that of 0.8776 by Austen, Benabbas and Georgiou [3]. Another problem in this framework is the $\text{SmallSetExpansion}$ problem, which is well studied due to its connection to the SSEH [33] and its consequences. In particular, Raghavendra, Steurer and Tetali [34] gave an algorithm, which on input a graph having a set of volume at most $\delta$ with expansion $\epsilon$, outputs a set of volume at most $O(\delta^{1/2} \log(1/\delta))$ – this was later shown to be tight by Raghavendra, Steurer and Tulsiani [35] assuming SSEH. There have been several works which also give frameworks for approximating general CSPs with global constraints. Guruswami and Sinop [16] gave Lasserre hierarchy based algorithms for the setting when underlying label extended graph has low threshold rank. Raghavendra and Tan [36] also propose a Lasserre hierarchy based framework for general settings. More recently, [1] also study such CSPs using $\text{Sticky Brownian Motion}$ based rounding algorithms.

$\text{DkS and DkSH}$. There is a long line of works which study the complexity of approximating $\text{DkS}$. Feige [10] showed that $\text{DkS}$ is APX-hard assuming the hardness of refuting random 3-SAT formulas. Subsequently Khot [20] also established APX-hardness assuming no sub-exponential time algorithms exist for SAT. Stronger inapproximability results are known under alternative hypotheses. The SSEH of Raghavendra and Steurer [33] immediately implies constant factor inapproximability of $\text{DkS}$ where as Manurangsi [26] showed almost polynomial ratio ETH based hardness. There are also results which establish running time lower bounds under alternative hypotheses [7], [29]. On the algorithmic front, Feige and Seldser [12] give an $n^{1/k}$-approximation algorithm for $\text{DkS}$. For $k$-independent bounds, Feige, Kortsarz and Peleg [11] gave a $O(n^{1/3})$ approximation algorithm, which was later improved to a $O(n^{1/4+\epsilon})$ by [5]. In comparison, there have been relatively fewer works which study the hypergraph variant i.e. $\text{DkSH}$. Assuming the existence of certain one way functions, Applebaum [2] showed that $\text{DkSH}$ is hard to approximate for hypergraphs on $n$-vertices up to a factor of $n^e$, for some constant $e > 0$. The results of [27] also implies that assuming SSEH, $\text{DkSH}$ is inapproximable for any constant factor with large enough arity.

$\text{Max-CSPs on large alphabets}$. There is a vast literature which study CSPs on non-boolean alphabets. For the setting of arity 2, Kindler, Kolla and Trevisan [23] gave a $\Omega((1/R) \log R)$ approximation algorithm, which matches the Unique Games based hardness from [21]. For the setting of larger arities, Makarychev and Makarychev [24] gave a $O(k^{-1} \cdot R^{k^2})$ approximation algorithm, which was then shown to be tight for the setting of $k \geq R$ by the work of Chan [8]. For the setting of small arities i.e., $k \leq R$, a $\Omega(k^3 R^{-k-1}) \log R)$-UGC based hardness can be derived using the results of [22]. In particular, combining their result with a known $R^{-k+1}$-integrality gap for linear programs implies the lower bound.

2 OVERVIEW: BIAS INDEPENDENCE CHARACTERIZATION

In this section, we briefly describe the challenges towards establishing our results and the techniques used to address them.

2.1 Characterization Of Bias Independence Via Minimal Sets

Our first step is to understand what makes the optimal approximation factor for a predicate bias dependent. For the purpose of...
exposition, we shall just focus on the behavior of predicates in the range \( \mu \in (0, 1/2) \). As a warm up, we will first restrict our attention to symmetric Boolean predicates i.e., predicates whose set of accepting strings is permutation invariant. In particular, one can always express a symmetric predicate \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \) as
\[
\psi := \psi_1 \lor \psi_2 \lor \ldots \lor \psi_t
\]
for some \( t \leq r \). However, we can still establish an approximate one-to-one correspondence that \( \mu \)-approximation algorithm for \( \text{DKSH}_r \) implies a \( \alpha \mu \)-approximation algorithm for \( \mu \)-biased CSPs with predicate \( \psi \).

The above arguments combined together roughly establishes the following:
\[
\alpha \mu (\text{DKSH}_r) \leq_{r} \alpha \mu (\psi) \leq_{r} \alpha \mu (\text{DKSH}_1),
\]
where \( \leq \) hides multiplicative factors in \( r \). In particular, (2) completely characterizes the bias dependence of predicate \( (\psi_1)_{i \in [r]} \).

Handling General Symmetric Predicates. Now recall that a symmetric predicate \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \) can be always expressed as \( \psi = \vee_{i \in [r]} \psi_i \). Clearly, the approximability of \( \psi \) is determined by the choice of \( \psi_i \). In particular, it is natural to suggest that \( \psi \) is as easy to approximate as the easiest predicate i.e., \( \psi_i \) with the smallest \( i \) which we denote by \( i^* \). It turns out that this is indeed the right characterization as we can establish that
\[
\alpha \mu (\psi) \approx_{r} \max_{i \in \{1, \ldots, r\}} \alpha \mu (\text{DKSH}_i),
\]
where \( \approx \) implies that the LHS is within multiplicative factors of \( r \) of the RHS. While the hardness of \( \psi \) using \( \text{DKSH}_r \) again follows by introducing dummy vertices with infinite weights, establishing the converse, i.e., \( \psi \) is as easy as \( \text{DKSH}_r \) requires more work due to the following issue. Given a \( \mu \)-biased CSP \( \Psi(\mu)(V, E) \) on predicate \( \psi \), it might be the case that all labelings which assign strings of weight \( r^* \) to a significant fraction of edges only satisfy a small fraction of constraints in comparison to the optimal \( \mu \)-biased labeling. Since our previous argument relied on the existence of good labelings which assign strings of weight \( r^* \) to a large fraction of constraints, it cannot be used to argue good approximation for such instances as is. We address this issue by ruling out the existence of such instances. In particular, we can show given any labeling \( \sigma \), by sub-sampling we can construct another labeling \( \sigma' \) which satisfies a significant fraction of edges in comparison to \( \sigma \), while assigning them strings of weight exactly \( r^* \). This in turn allows us to conclude that for any \( \mu \)-biased CSP on predicate \( \psi \), there exists a \( \mu \)-biased labeling that satisfies at least \( 2^{-r' \cdot \delta} \) fraction of the optimal edge using strings from \( \psi_i \rightarrow \{0, 1\} \) - this observation can then be combined with the previous arguments to establish the converse direction.

Handling General Predicates using Minimal Sets. Now we relax our setting to that of general Boolean predicates \( \psi : \{0, 1\}^r \rightarrow \{0, 1\} \). Note that if \( \psi \) is not symmetric, then it is no longer guaranteed to admit a decomposition of the form (1), and therefore it is not clear if one can still characterize the bias approximation curve of \( \psi \) using that of \( \text{DKSH} \). In order to motivate our characterization here, consider the following notion of partial ordering among predicates. For a pair of Boolean predicates \( \psi, \psi' : \{0, 1\}^r \rightarrow \{0, 1\} \), we say \( \psi \geq \psi' \) if \( \psi_i^{-1}(1) \supset= \psi'_{i}(1) \), and for

To be precise, since we want the bias parameter for the resulting instance to be still constant (for constant \( r \)), in the actual reduction we assign the dummy vertices a weight of \( |V| \).

In the actual reduction, the biases of the labeling can differ up to a multiplicative factor of \( r \), but we ignore this issue here to keep the presentation simple.
every accepting string $\beta \in \psi^{-1}(1)$, there exists $\tilde{\beta} \in \psi^{-1}(1) \cap \tilde{\psi}^{-1}(1)$ such that $\text{supp}(\beta) \supseteq \text{supp}(\tilde{\beta})$. For any such $\psi, \tilde{\psi}$-pair it is not too difficult to show that $\alpha_{\leq \mu}(\tilde{\psi}) \geq r \alpha_{\leq \mu}(\psi)$.

To see this, consider a $\mu$-biased instance $\Psi_{(\mu)}(V,E)$ on predicate $\psi$, and let $\sigma : V \to \{0,1\}$ be the $\mu$-biased labeling which achieves the optimal value, say $\gamma$. Now given $\sigma$, consider the following sub-sampling process to construct $\sigma' : V \to \{0,1\}$. For every $i \in V$ we do the following independently: if $\sigma(i) = 1$, sample $\sigma'(i)$ uniformly from $\{0,1\}$ otherwise, we set it to 0. Now fix a constraint $e$ satisfied by $\sigma$ and let $\tilde{\beta}_e \in \psi^{-1}(1) \cap \tilde{\psi}^{-1}(1)$ such that $\text{supp}(\sigma(e)) \supseteq \text{supp}(\tilde{\beta}_e)$. Now observe that the sub-sampling always ensures that the resulting labeling $\sigma'$ is at most $\mu$ biased. Furthermore,

$$\begin{align*}
E_{\sigma \sim E} & \Pr_\epsilon \left[ \tilde{\psi}(\sigma'(e)) = 1 \right] \\
& \geq \Pr_{\sigma \sim E} \left[ \sigma(e) = 1 \right] \Pr_{\sigma \sim \sigma'} \left[ \sigma'(e) = \tilde{\beta}_e \right] \\
& \geq 2^{-r} \gamma,
\end{align*}$$

i.e., the labeling $\sigma'$ satisfies at $2^{-r} \gamma$-fraction of constraints $e$ in $\Psi$ by assigning strings from $\tilde{\psi}^{-1}(1)$. Furthermore, since $\psi^{-1}(1) \supseteq \tilde{\psi}^{-1}(1)$, if a labeling satisfies at least $\gamma$-fraction of edges in $\Psi$ with respect to predicate $\tilde{\psi}$, it also satisfies at least $\gamma$-fraction of edges with respect to predicate $\psi$. Hence, any $\alpha_{\leq \mu}(\tilde{\psi})$-approximation algorithm for $\tilde{\psi}$ is also a $\Omega(2^{-r} \alpha_{\leq \mu}(\psi))$-approximation algorithm for $\psi$. In summary, this establishes that whenever $\tilde{\psi} \geq \psi$ we have $\alpha_{\leq \mu}(\tilde{\psi}) \leq r \alpha_{\leq \mu}(\psi)$.

The above partial ordering and its properties immediately imply that a predicate $\phi$ will be at least as easy as the set of minimal elements dominated by it. A reduction based argument will also show that it is as hard as its minimal elements. Furthermore, a straightforward argument also shows that the minimal elements $\tilde{\psi}$ dominated by $\psi$ are predicates for which the accepting set is a singleton set i.e. they satisfy $|\tilde{\psi}^{-1}(1)| = 1$. It turns out that for such predicates, the arguments used in the setting of symmetric predicates generalize readily. These observations taken together imply the following characterization. Given a predicate $\psi$, we have

$$\alpha_{\leq \mu}(\psi) \times_r \min_{\beta \in M_\psi} \alpha_{\leq \mu}(\text{DKSH}_{H|\beta||\mu}) \leq \alpha_{\leq \mu}(\psi), \quad (4)$$

where $M_\psi$ is the set of minimal elements of $\psi^{-1}(1)$. In particular, the above immediately reduces our task to characterizing the bias-approximation curve of DKSH, which is a key technical contribution of this work and is discussed in details in Section 2.2.

About Weighted vs. Unweighted settings. We conclude the first part of the overview by discussing some additional complications that arise while handling vertex weights. While we ignore the difference between weighted and non-weighted settings in the above discussion, the precise statements of our results (Theorem 1.2 in particular) actually relate the bias approximation curve of weighted biased CSP problems to that of unweighted DKSH. While this does not affect the arguments used to upper bound the bias approximation curve, it presents several subtle challenges in the direction of the lower bound where we show that algorithms for unweighted DKSH can be used as black-boxes for solving weighted biased CSP instance. In particular, to achieve matching upper and lower bounds, we allow the vertex weights of the CSP to be polynomially large (e.g., recall that the reduction from DKSH to biased CSPs sets the weights of the dummy vertices to infinite). However, for such weighted instances, techniques for reducing to the unweighted setting don’t apply as is. This is to be expected since biased CSPs with arbitrary vertex weights can be strictly harder than unweighted biased CSPs, even under the same bias constraint. This issue is addressed by allowing a multiplicative slack in the bias of the labeling i.e., where we allow algorithms to output solutions with relative weight at most $\mu(1+\eta)$ for some constant $\eta$ – this multiplicative slack is crucially used in relating the approximation curve of weighted DKSH to that of unweighted DKSH. Note that this relaxation does not change the lower bound on the approximation curve since our reduction from DKSH to biased CSP holds as is even in this setting. For more details on this point, we refer the readers to Sections 5, 6 and 7 of the full version of the paper [14].

2.2 Overview: Hardness Of DKSH

While hardness of approximating DKSH for general arities is relatively less explored, there has been substantial work on lower bounds for DKSH. The strongest known results here are the constant factor inapproximability by Raghavendra and Steurer [33] assuming SSEH, and the almost polynomial ratio hardness by Manurangsi [26] assuming ETH. However the techniques from the above works don’t apply to our setting since here we seek to quantify the approximation curve as an explicit function of the bias parameter $\mu = k/n$. Instead, our approach towards establishing Theorem 1.3 would be to treat DKSH as instances of MAX-AND subject to a global cardinality constraint. Hence, as is standard, our reduction will use the framework of composing a dictatorship test with an appropriate outer verifier. This is a well studied approach that has been used successfully to show often optimal inapproximability bounds for MAX-CSPs (see [19] for an overview of such reductions).

Informally for a bias $\mu$, the $(c(\mu), s(\mu))$ dictatorship test (for AND predicate) in our setting is a distribution $D$ over tuples of element i.e., hyperedges $(x_1, \ldots, x_\ell)$ drawn from some product probability space $\Omega^\ell$. The distribution naturally defines a weighted hypergraph $H = (\Omega^\ell, D, w)$ on the set of vertices $\Omega^\ell$ and every Boolean function $f : \Omega^\ell \to \{0,1\}$ induces a subset $S_f$ of $\Omega^\ell$, and therefore can be associated with weight $w(S_f)$. In addition, we seek the following properties from the distribution:

- **Completeness.** If $f : \Omega^\ell \to \{0,1\}$ is a dictator of weight $\mu$, then the set $S_f$ induces at least $c(\mu)$-fraction of hyperedges in $H$.

- **Soundness.** If $f : \Omega^\ell \to \{0,1\}$ is a function of weight $\mu$ such that $S_f$ induces at least $s(\mu)$-fraction of hyperedges in $H$, then $f$ has at least one influential coordinate.

Here the notion of dictators and influential coordinates are the natural analogues of their counterparts in a long code test. The above is typically the key component in dictatorship test reductions, where it is well understood that a $(c(s), s(\mu))$-dictatorship test for a predicate $\psi$ almost immediately leads to a $(s(\mu))$-hardness of approximation (for the unconstrained MAX-CSP) by composing it with a suitable outer verifier such as LABEL COVER or UNIQUE GAMES. Therefore, the obvious first challenge here is to design a family of bias dependent dictatorship tests with the right completeness soundness.
A somewhat loose restatement of our completeness and soundness properties from above would be that we require a distribution over a space $\Omega^r$ with the property that all sets which induce large number of edges, or equivalently, sets which are noise stable, must have influential coordinates. Objects with this property, including the noisy hypercube and its variants, have been studied extensively in the context of hardness of approximation, and have been instrumental in showing optimal hardness of several problems such as MaxCut, UniqueGames [21], SmallSetExpansion [35]. Due to the additional bias constraint, this motivates us to study the $\mu$-biased $(1-\rho)$-noisy hypercube $\mathcal{H}_{\mu,\rho}$. We describe the dictatorship test as a distribution over hyperedges for $\mathcal{H}_{\mu,\rho}$ in Figure 1 below.

**Hyperedge Distribution on $\mathcal{H}_{\mu,\rho}$**

**Distribution.**

1. Sample $x \sim \{0,1\}^t_{\mu}$.  
2. Sample independent $\rho$-correlated copies $x_1, \ldots, x_r \sim x$.  
3. Output hyperedge $(x_1, \ldots, x_r)$.

**Figure 1: Biased Noisy Hypercube Test**

Here $\{0,1\}^t_{\mu}$ denotes the probability space where each bit $i \in [t]$ is independently sampled from the Bernoulli distribution with bias $\mu$. Furthermore, fixing a $x \in \{0,1\}^t_{\mu}$, a $\rho$-correlated copy $x'$ of $x$ is sampled by setting each bit $x'(i)$ to $x(i)$ with probability $\rho$ and re-sampling $x'(i)$ with probability $1-\rho$ independently. The completeness and soundness properties of the above test can be derived using standard Fourier analytic techniques. A useful first observation is that the fraction of hyperedges induced by a set indicated by a function $f : \{0,1\}^t \rightarrow \{0,1\}$ can be expressed as

$$w(E[S_f]) = E_{x_1,\ldots,x_r} \left[ \prod_{i \in [r]} f(x_i) \right].$$

For the completeness analysis of the test, we see that when $f$ is the dictator function $f(x) = x(1)$, then

$$E_{x_1,\ldots,x_r} \left[ \prod_{i \in [r]} f(x_i) \right] \geq \Pr_x \left[ x(1) = 1 \right] \Pr_{(x_i)_{x(1)=1}} \left[ \forall i \in [r], x_i(1) = x(1) \right] \geq \mu \rho^r$$

and $w(S_f) = E_x \left[ x(1) \right] = \mu$ i.e., $f$ satisfies the weight constraint. On the other hand, for analyzing the soundness guarantee, fix a function $f$ having no influential coordinates and $w(S_f) = \mu$. Then using standard arguments involving the Invariance Principle and Gaussian Stability bounds (for e.g., see Theorem 2.10 [22]), it can be shown that

$$E_{x_1,\ldots,x_r} \left[ \prod_{i \in [r]} f(x_i) \right] \leq 2\rho^r \tag{5}$$

when $\rho := 1/\sqrt{\tau} \log(1/\mu)$. Combining the above arguments, we get a test with completeness-soundness ratio of $\mu^{-1} (\log(1/\mu))^{r/2}$, which is worse by a factor of $(\log(1/\mu))^{r/2-1}$ from the intended ratio. The issue here is that the completeness value of the test has a $\rho^r$ multiplier due to the independent $\rho$-correlated re-sampling. In particular, conditioned on $x(1) = 1$, the completeness pays an extra multiplier of $\rho^r$ since for every $i \in [r]$, $x_i(1)$ can be chosen to resampled independently with probability $1-\rho$.

To fix the above issue, we allow the noise pattern of variables $x_1, \ldots, x_r$ to be correlated instead of being fully independent. Formally, observe that we can reinterpret the original $\rho$-correlated sampling along a coordinate $j \in [t]$ in the following way.

(i) Sample $\theta_1(j), \theta_2(j), \ldots, \theta_r(j) \sim \{0,1\}_\rho$.
(ii) For every $i \in [r]$, do the following: if $\theta(j) = 1$, set $x_i(j) = x(j)$ otherwise sample $x_i(j) \sim \{0,1\}_\rho$ independently.

The above results in a distribution where each pair of $x_i, x_j$ variables are $\rho^2$-correlated. Now, the crucial observation here is that, as is the case with noise stability type arguments, the soundness analysis for the above test distribution just relies on the second moment structure of the test distribution, and in particular, just uses the fact that the $(x_i, x_j)$ variables are pairwise $\rho^2$-correlated. This is due to the folklore observation that for any distribution on $x_1, \ldots, x_r$ that is pairwise $\rho^2$-correlated, using techniques from [30] one can show that the following holds:

$$E_{x_1,\ldots,x_r} \left[ \prod_{i \in [r]} f(x_i) \right] \approx \prod_{i \in [r]} E_{x_i} \left[ f(x_i) \right].$$

Furthermore, it is well known (for e.g. [22]) that this weaker condition on the test distribution can be realized with more correlated noise patterns. In particular, we can consider the following alternative distribution:

W.p. $\rho^2$, set $\theta_1(j), \ldots, \theta_r(j)$ to 1, otherwise set $\theta_1(j), \ldots, \theta_r(j)$ to 0 In other words, for any $j \in [t]$, with probability $\rho^2$, we set all $x_i(j)$ variables to $x(j)$, otherwise we resample all variables independently. It is easy to see that the above again results in a distribution where each $(x_i, x_j)$ variable pair is $\rho^2$-correlated. However, note that under the new distribution, the probability of realizing a all-ones assignment along any coordinate $j \in [t]$ is at least

$$\Pr \left[ x(j) = 1 \right] \cdot \Pr \left[ \theta(j) = 1 \right] = \rho^2 \mu,$$

which improves on $\mu \rho^r$ from the previous test distribution – this is the key consequence of our choice of distribution that is responsible for deriving the improved completeness-soundness ratio. We
conclude our discussion by giving a brief sketch of the completeness and soundness analysis of the test with respect to the new distribution. Call the new distribution over \( r \)-tuples \( \mathcal{D}' \). For the completeness, for a dictator function \( f(x) = x(1) \), we proceed as before and get that

\[
E_{(x_1,\ldots,x_r) \sim \mathcal{D}'} \left( \prod_{i \in [r]} f(x_i) \right) \\
\geq \Pr_{x \sim \{0,1\}^r} \left( x(1) = 1 \right) \Pr_{(x_i)_{i=1}^r | x(1) = 1} \left( \forall i \in [r], x_i(1) = x(1) \right) \\
\geq \mu \Pr_{\theta_1,\ldots,\theta_r \sim \mathcal{D}'} \left( \forall i \in [r], \theta_i(1) = 1 \right) \\
= \mu \rho^r.
\]

For the soundness direction, we employ the noise stability analysis from Khot and Saket [22]. Consider a function \( f : \{0,1\}^t \to \{0,1\} \) having no influential coordinates satisfying \( E_\mu [f(x)] = \mu \). The first step is to observe that since under the test distribution, the variables \( x_1,\ldots, x_r \) are pairwise \( \rho^2 \)-correlated, using a multidimensional version of Borell’s Isoperimetric Inequality [30], one can show that

\[
E_{x_1,\ldots,x_r} \left( \prod_{i \in [r]} f(x_i) \right) \leq \Gamma^{(r)}_{\rho^2}(\mu)
\]

where \( \Gamma^{(r)}_{\rho^2}(\mu) \) is the iterated \( r \)-ary Gaussian stability of a halfspace with volume \( \mu \). Furthermore, when \( \rho^2 \leq O(1/(r^2 \log(1/\mu))) \), [22] shows that

\[
\Gamma^{(r)}_{\rho^2}(\mu) \leq \mu^r
\]

which concludes the soundness analysis.

### 2.4 Choice of Outer Verifier

Given the above dictatorship test, we now proceed to discuss the composition step. Typically, for CSPs without global constraints, a dictatorship test for a predicate can be plugged in almost immediately into UNIQUE GAMES (or often even LABEL COVER), and result in hardness matching the completeness soundness ratio. However, that technique fails to work for CSPs with global cardinality constraints since if the outer verifier does not provide local bias control – we elaborate on this issue below.

Informally, the composition step with the above dictatorship test will introduce a long code table \( f_\omega : \Omega^t \to \{0,1\} \) for every vertex \( v \) of the outer verifier CSP (say UNIQUE GAMES), denoted by \( \Psi \). Then, it embeds the dictatorship test with the outer verifier in such a way that the overall reduction can be thought of as the following two step process.

- Sample a vertex \( \omega \in \Psi \). Let \( g_\omega \) denote the weighted average of the long codes of the neighbors of \( \omega \) in \( \Psi \).
- Test \( g_\omega \) on the distribution \((x_1,\ldots,x_t)\) i.e. accept if and only if

\[
g_\omega(x_1) = \cdots = g_\omega(x_t) = 1.
\]

The completeness and soundness guarantees of the test can then be used to argue the completeness and soundness properties of the full reduction. For the completeness direction, if the outer verifier \( \Psi \) admits a labeling \( \sigma \) which satisfies most edges, then one can show that the dictator assignment \( f_\omega = x_\sigma(\omega) \) induces a large fraction of edges. This is because since most constraints in \( \Psi \) are consistent with the labeling \( \sigma \), this translates to the effect that even the averaged function \( g_\omega \) still behaves like \( f_\omega \) which is a dictator, and hence, the test accepts with probability close to the completeness of the distribution. On the other hand, if the optimal value of \( \Psi \) is small, then for any fixed labeling, most edges will be inconsistent, and therefore most averaged functions will not have influential coordinates. Now suppose in addition, we could guarantee that for most averaged function \( g_\omega \), we have \( E[g_\omega] \approx \mu \), we can use the soundness guarantee of the distribution to argue that the test accepts with probability at most the soundness value.

However, note that the composition step as is can only guarantee that the “expected” bias of a long code \( \{f_\omega\} \) for a randomly chosen vertex \( v \) is \( \mu \), and as such this does not imply the aforementioned “concentration of biases” property (i.e., \( E[g_\omega] \approx \mu \) for most \( v \)’s) that is needed to argue soundness. In particular, in the context of the reduction, failure to ensure the concentration property can allow the adversary to devise cheating assignments where it can set a subset of long codes to be the constant all ones functions and the remaining to be all zeros. This in turn can cause the overall soundness analysis of the reduction to fail since the soundness guarantee of the dictatorship test only holds for functions whose average value is close to the global average \( \mu \). Therefore, in order to ensure that even under such assignments, the biases of the averaged long codes concentrate around the global bias \( \mu \), the averaging operator \( \{f_\omega\} \mapsto \{g_\omega\} \) should have good mixing properties, or equivalently, have large spectral gap. This requires the use of non-standard outer verifiers such as the Quasirandom PCP [20] or SMALLSET-EXPANSION [33, 35]. In particular, we shall use the SSE based framework introduced in [35] which we describe informally below.

The key component in the framework introduced in [35] is a family of noise operators \( \{M_\omega\} \), referred to as Noise Operators with Leakage. Formally, given a string \( z \in \{\bot, \top\}^t \) the corresponding noise operator \( M_\omega \) on the space \( \Omega^t \) is defined as follows. For any \( \omega \in \Omega^t \), one can sample \( \omega' \sim M_\omega(\omega) \) using the following process. For every \( i \in [t] \), do the following independently: if \( z(i) = \top \), set \( \omega'(i) = \omega(i) \) otherwise re-sample \( \omega'(i) \sim \Omega \). In particular, for a random draw of \( (\bot, \top)^t \), the noise operator \( M_\omega \) behaves like the standard noise operator \( T_{\omega} \) on the space \( L_2(\Omega) \). By incorporating the above noise operator in the averaging step, one can guarantee that the spectral gap of the averaging operator is at least \( 1 - \beta \), thus guaranteeing the aforementioned concentration on the biases. We describe our overall reduction in Figure 2.

Let \( G = (V,E) \) be a SMALLSET-EXPANSION instance.

**Input.** Long Codes \( \{f_\omega\}_{\omega \in V} \) with \( f_\omega : \Omega^t \to \{0,1\} \), where \( \Omega = \{0,1\}_t \times \{\bot, \top\}_t \).

**Test:**

1. Sample \( A \in V^t \) and \( B_1,\ldots, B_t \sim V^t \).
2. Sample \( (x_1, z_1),\ldots,(x_t, z_t) \sim (\mathcal{D}^*)^\otimes t \) such that marginally \( (x_i, z_i) \sim \Omega^t \).
3. For every \( i \in [r] \), sample \( (B_{i1}, x_{i1}, z_{i1}) \sim M_\omega(B_i, x_i) \).
4. Sample random permutations \( \pi_1,\ldots, \pi_t \sim S_t \).
5. Output hyperedge \( \{\pi_1(B'_{i1}, x'_{i1}, z_{i1}),\ldots, \pi_t(B'_{i1}, x'_{i1}, z_{i1})\} \).

Figure 2: Reduction from SSE (Informal)
The analysis of the above reduction combines the arguments from Section 2.3 with techniques from [35]. A key difference is that since our test distribution works in the almost uncorrelated regime, we need a more careful completeness analysis i.e., the reduction from [35] works in the setting \( \rho \to 1 \) whereas our tests are based in the setting \( \rho \to 0 \). Another additional component in our proof is that we strengthen \( k \)-ary noise stability estimate for functions over the \([R]\)-ary hypercube (i.e., (5)) with small low degree influences to the setting of arbitrary probability spaces.

2.5 Approximation Algorithm For DkSH

We conclude this section with a brief discussion on the approximation algorithm for DkSH from Theorem 1.4. Our approximation algorithm for DkSH is similar in spirit to the \( O(R^{-(k-1)} \log R) \)-approximation algorithm for Max-k-CSP’s from [28]. Following their approach, the overall algorithm consists of two components:

- A \( O(\mu \log(1/\mu)) \)-approximation algorithm for DkS.
- A reduction from DkS, to DkSH which shows that

\[
\alpha(\mu)(\text{DkSH}_n) \geq \mu^{-2} \alpha(\mu)(\text{DkS}).
\]

For the first point, we use a \( O(\mu \log(1/\mu)) \)-approximation algorithm that can be obtained by using a reduction ([9]) from DkS to Max-2-CSP on label set \( \mu \) and then combining it with the \( O(\log(R/R')) \)-approximation algorithm for Max-2-CSP from [23]. The second point is based on the reduction from Max-k-CSP to Max-2-CSP in [28] – in our setting, this translates to reducing a DkS instance \( H \) into the DkS instance \( G \) by adding the clique-expansion of every edge in \( H \) to \( G \) with appropriate choices of vertex and edge weights. Combining these ideas immediately yields the \( O(\mu^{-3} \log(1/\mu)) \)-approximation algorithm for DkSH as claimed in Theorem 1.4.

3 CONCLUSION

In this work, we initiate a systematic study of how the bias parameter and predicate structure affects the approximability of CSPs with global cardinality constraints. While our work addresses some key questions, it also presents several interesting questions that are fundamental towards developing a well-rounded understanding of CSPs with global constraints. We list some key such questions below:

1. **Optimal \( r \)-dependent bias approximation curve**: Throughout the present work, our upper and lower bounds are matching up to multiplicative factors of the arity \( r \), and hence an immediate next step would be to derive upper and lower bounds that are matching even in terms of their dependency on \( r \). As an intermediate but still important step, the question of optimal \( r \)-dependent approximation bounds for DkS remains open.

2. **Approximability in sub-constant bias regime**: While this work addresses the approximability of Biased CSPs in the regime where the bias \( \mu \) is constant, one can also ask if matching bounds can be derived in the regime where the bias is strictly sub-constant. For e.g., can we derive matching bounds when \( \mu = n^{-\epsilon} \) where \( \epsilon \in (0,1) \) is a constant?  

3. **Matching bounds under weaker hypotheses**: Our explicit upper bounds on the bias approximation curve (for e.g., Theorem 1.3, Corollary 1.5) are all based on SSEH, and hence it would be interesting to understand if quantitatively similar bounds can also be shown be established under weaker hypotheses such as UGC?

**ACKNOWLEDGMENTS**

EL is partially supported by a gift from Google.

**REFERENCES**

[1] Sepehr Abbasi-Zadeh, Nikhil Bansal, Guru Guruganesh, Aleksandar Nikolov, Roy Schwartz, and Mohit Singh. 2020. Sticky brownian rounding and its applications to constraint satisfaction problems. In Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 843–873. https://doi.org/10.1137/1.9781611975994.52

[2] Benny Applebaum. 2013. Pseudorandom generators with low stretch and low locality from random local one-way functions. SIAM J. Comput. 42, 5 (2013), 2008–2037. https://doi.org/10.1137/120884857

[3] Per Austrin, Siavosh Benabbas, and Konstantinos Georgiou. 2016. Better balance by being biased: A 0.8776-approximation for max bisection. ACM Transactions on Algorithms (TALG) 13, 1 (2016), 1–27. https://doi.org/10.1145/2907022

[4] Per Austrin and Aleksa Stankovic. 2019. Global Cardinality Constraints Make Approximating Some Max-2-CSPs Harder. 145 (2019), 24:1–24:17. https://doi.org/10.1139/LIPIcs-APPROX-RANDOM.2019.24

[5] Aditya Bhaskara, Moses Charikar, Eden Chlamtac, Uriel Feige, and Aravindan Vijayaraghavan. 2010. Detecting high log-densities: An \( O(n/4) \) approximation for densest \( k \)-subgraph. In Proceedings of the forty-second ACM symposium on Theory of computing, 201–210. https://doi.org/10.1145/1866689.1866719

[6] Aditya Bhaskara, Moses Charikar, Venkatesan Guruswami, Aravindan Vijayaraghavan, and Yuan Zhou. 2012. Polynomial integrality gaps for strong sdp relaxations of densest \( k \)-subgraph. In Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms. SIAM, 388–405. https://doi.org/10.1137/1.9781611973089.34

[7] Parinya Chalermsook, Marek Cygan, Guy Kortsarz, Bundit Laekhanukit, Pasin Manurangsi, Danupon Nanongkai, and Luca Trevisan. 2020. From Gap-Exponential Time Hypothesis to Fixed Parameter Tractable Inapproximability: Clique, Dominating Set, and More. SIAM J. Comput. 49, 4 (2020), 772–810. https://doi.org/10.1137/18M1166869

[8] Siu On Chan. 2016. Approximation resistance from pairwise-independent subgroups. Journal of the ACM (JACM) 63, 3 (2016), 1–52. https://doi.org/10.1145/2873054

[9] Moses Charikar, MohammadTaghi Hajiaghayi, and Howard J. Karloff. 2011. Improved Approximation Algorithms for Label Cover Problems. Algorithmica 61, 1 (2011), 190–206. https://doi.org/10.1007/s00453-010-9464-3

[10] Uriel Feige. 2002. Relations between average case complexity and approximation complexity. In Proceedings of the thirty-fourth annual ACM symposium on Theory of computing, 534–543. https://doi.org/10.1145/509907.509985

[11] Uriel Feige, David Peleg, and Guy Kortsarz. 2001. The dense \( k \)-subgraph problem. Algorithmica 29, 3 (2001), 410–421. https://doi.org/10.1007/978-0-387-21809-8_19

[12] U Feige and M Seltzer. 1997. On the densest \( k \)-subgraph problems. In** Annals of Mathematics**. 1997. Improved approximation algorithms for MAX \( k \)-CUT and MAX BIEDGE连接**. Algorithmica 18, 1 (1997), 67–81. https://doi.org/10.1007/BF0253688

[13] Suprovat Ghoshal and Euiwoong Lee. 2022. A Characterization of Approximability for Biased CSPs. arXiv preprint arXiv:2201.04617 (2022).

[14] Michel X. Goemans and David P. Williamson. 1994. 0.879-approximation algorithms for MAX CUT and MAX SAT. In Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, 23-25 May 1994, Montreal, Quebec, Canada, Frank Thomson Leighton and Michael T. Goodrich (Eds.). ACM, 422–431. https://doi.org/10.1145/195089.195216

[15] Venkatesan Guruswami and Ali Kemal Sinop. 2011. Lasserre Hierarchy, Higher Eigenvalues, and Approximation Schemes for Graph Partitioning and Quadratic Integer Programming with PSD Objectives. In IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011, Rafail Ostrovsky (Ed.). IEEE Computer Society, 482–491. https://doi.org/10.1109/FOCS.2011.36

[16] David Karger, Rajeev Motwani, and Madhu Sudan. 1998. Approximate graph coloring by semidefinite programming. Journal of the ACM (JACM) 45, 2 (1998), 246–265. https://doi.org/10.1145/274878.274791

[17] Subhash Khot. 2002. On the power of unique 2-prover 1-round games. In Proceedings of the 44th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada, 767–775. https://doi.org/10.1145/509907.510017

[18] Subhash Khot. 2005. Guest column: inapproximability results via Long Code based PCPs. SIGACT News 36, 2 (2005), 25–42. https://doi.org/10.1145/1067309.1067318

[19] Subhash Khot. 2006. Ruling out PTAS for graph min-bisection, dense \( k \)-subgraph, and bipartite clique. SIAM J. Comput. 36, 4 (2006), 1025–1071. https://doi.org/10.1002/sim.2873
1117/S0097539705447037
[21] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O’Donnell. 2007. Optimal
inapproximability results for MAX-CUT and other 2-variable CSPs. SIAM J.
Comput. 37, 1 (2007), 319–357. https://doi.org/10.1137/S0097539705447372
[22] Subhash Khot and Rishi Saket. 2015. Approximating CSPs Using LP Relaxation. In
Automata, Languages, and Programming - 42nd International Colloquium, ICALP
2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I (Lecture Notes in Computer
Science, Vol. 9134). Springer, 822–833. https://doi.org/10.1007/978-3-662-47672-
7_67
[23] Guy Kindler, Alexandra Kolla, and Luca Trevisan. 2016. Approximation of non-
boolean 2CSP. (2016), 1705–1714. https://doi.org/10.1137/1.9781611974331.ch117
[24] Konstantin Makarychev and Yury Makarychev. 2012. Approximation algorithm
for non-boolean max k-csp. In Approximation, Randomization, and Combinatorial
Optimization. Algorithms and Techniques. Springer, 254–265. https://doi.org/10.
1007/978-3-642-32512-0_22
[25] Konstantin Makarychev and Yury Makarychev. 2017. Approximation algorithms
for CSPs. In Dagstuhl Follow-Ups, Vol. 7. Schloss Dagstuhl-Leibniz-Zentrum fuer
Informatik. https://doi.org/10.4230/LIPIcs.DOI-VOl7.15091.11
[26] Pasin Manurangsi. 2017. Almost-polynomial ratio ETH-hardness of approximat-
ing densest k-subgraph. In Proceedings of the 49th Annual ACM SIGACT Sympo-
sium on Theory of Computing, 954–961. https://doi.org/10.1145/3055399.3055412
[27] Pasin Manurangsi. 2018. Inapproximability of Maximum Bichromatic Problems,
Minimum k-Cut and Densest k-At-Least-k-Subgraph from the Small Set Expansion
Hypothesis. Algorithms 11, 1 (2018), 10. https://doi.org/10.3390/a11010010
[28] Pasin Manurangsi, Prerum Nakkarin, and Luca Trevisan. 2016. Near-Optimal
UGC-hardness of Approximating Max k-CSP. Approximation, Randomization,
and Combinatorial Optimization. Algorithms and Techniques (2016). https://doi.
org/10.4230/LIPIcs.APPROX-RANDOM.2016.15
[29] Pasin Manurangsi, Aviad Rubinstein, and Tselil Schramm. 2021. The Strongish
Planted Clique Hypothesis and Its Consequences. In 12th Innovations in Theoreti-
cal Computer Science Conference, ITCS 2021, January 6-8, 2021, Virtual Conference
(LIPIcs, Vol. 185), James R. Lee (Ed.). 10:1–10:21. https://doi.org/10.4230/LIPIcs.
ITCS.2021.10
[30] Elchanan Mossel. 2010. Gaussian bounds for noise correlation of functions.
Geometric and Functional Analysis 19, 4 (2010), 1713–1756.
[31] Ryan O’Donnell and Yi Wu. 2008. An optimal SDP algorithm for Max-Cut,
and equally optimal Long Code tests. In Proceedings of the forty-th annual ACM
symposium on Theory of computing. 335–344. https://doi.org/10.1145/1374376.
1374425
[32] Prasad Raghavendra. 2008. Optimal algorithms and inapproximability results for
every CSP?. In STOC. 245–254. https://doi.org/10.1145/1374376.1374414
[33] Prasad Raghavendra and David Steurer. 2010. Graph expansion and the unique
games conjecture. In Proceedings of the forty-second ACM symposium on Theory
of computing. ACM, 755–764. https://doi.org/10.1145/1806689.1806792
[34] Prasad Raghavendra, David Steurer, and Prasad Tetali. 2010. Approximations
for the isoperimetric and spectral profile of graphs and related parameters. In
Proceedings of the forty-second ACM symposium on Theory of computing. ACM,
631–640. https://doi.org/10.1145/1806689.1806776
[35] Prasad Raghavendra, David Steurer, and Madhur Tulsiani. 2012. Reductions
between expansion problems. In 2012 IEEE 27th Conference on Computational
Complexity. IEEE, 64–73. https://doi.org/10.1109/CCC.2012.43
[36] Prasad Raghavendra and Ning Tan. 2012. Approximating CSPs with global
 cardinality constraints using SDP hierarchies. In Proceedings of the twenty-third
 annual ACM-SIAM symposium on Discrete Algorithms. SIAM, 373–387. https:
//doi.org/10.1137/1.9781611973099.33
[37] Arv Wigderson. 1982. A new approximate graph coloring algorithm. In Proceed-
ing of the fourteenth annual ACM symposium on Theory of computing. 325–329.
https://doi.org/10.1145/800070.802207
[38] Yinyu Ye. 2001. A 699-approximation algorithm for Max-Bisection. Mathematical
programming 90, 1 (2001), 101–111. https://doi.org/10.1007/PL00011415
[39] Uri Zwick. 2002. Computer assisted proof of optimal approximability results. In
STOC. 496–505. http://dl.acm.org/citation.cfm?id=545381.545448

997