Optimal Sherali-Adams Gaps from Pairwise Independence

Konstantinos Georgiou[1⋆], Avner Magen[1⋆], and Madhur Tulsiani[2⋆⋆]

[1] Department of Computer Science, University of Toronto
[2] Computer Science Division, University of California, Berkeley
{cgeorg,avner}@cs.toronto.edu, {madhurt}@cs.berkeley.edu

Abstract. This work considers the problem of approximating fixed predicate constraint satisfaction problems (MAX k-CSP(P)). We show that if the set of assignments accepted by P contains the support of a balanced pairwise independent distribution over the domain of the inputs, then such a problem on n variables cannot be approximated better than the trivial (random) approximation, even using $\Omega(n)$ levels of the Sherali-Adams LP hierarchy.
It was recently shown [3] that under the Unique Game Conjecture, CSPs with predicates with this condition cannot be approximated better than the trivial approximation. Our results can be viewed as an unconditional analogue of this result in the restricted computational model defined by the Sherali-Adams hierarchy. We also introduce a new generalization of techniques to define consistent "local distributions" over partial assignments to variables in the problem, which is often the crux of proving lower bounds for such hierarchies.

1 Introduction

A constraint satisfaction problem (CSP) consists of a set of constraints that seek a universal solution. In the maximization version (MAX-CSP) one tries to maximize the number of constraints that can be simultaneously satisfied. The most standard family of CSPs arise from Boolean predicates P with bounded support k. In their generality, the predicates are defined over an alphabet $\{0, 1, \ldots, q-1\} = [q]$ and they can be thought as functions $P : [q]^k \to \{0, 1\}$. A constraint is defined by the predicate P applied to a k-tuple of literals $(x_1 + b_1 \bmod q, \ldots, x_k + b_k \bmod q)$, where $b_i \in [q]$, and is said to be satisfied by some assignment on $(x_1, \ldots, x_k)$ if the predicate evaluates to 1. Given some predicate P, an instance of the MAX k-CSP(P) problem is a collection of constraints as above and the objective is to maximize the number of constraints that can be satisfied simultaneously. As a special case, we can obtain all well studied MAX-CSP problems, e.g. MAX k-SAT, MAX-CUT etc. When the predicate to be used in different constraints is not fixed we simply refer to the problem as MAX k-CSP.

⋆ Funded in part by NSERC.
⋆⋆ Supported by the NSF grants CCF-0515231 and CCF-0729137 and by US-Israel BSF grant 2006060.

The MAX k-CSP problem is NP-hard for $k \geq 2$, and a lot of effort has been devoted in determining the true inapproximability of the problem. In general, the inapproximability of the MAX k-CSP depends on the size of alphabet over which literals are valued. For the case of Boolean variables, Samorodnitsky and Trevisan [19] proved that the problem is hard to approximate better than a factor of $2^{2\sqrt{k}}/2^k$, which was improved to $2^{\sqrt{2k}}/2^k$ by Engebresten and Holmerin [9]. Later Samorodnitsky and Trevisan [20] showed that it is Unique-Games-hard to approximate the same problem with factor better than $2^{\lceil \log k+1 \rceil}/2^k$. For the more general case of q-ary variables (MAX k-CSP$_q$), Guruswami and Raghavendra [13] showed a hardness ratio of q^2k/q^k when q is a prime.

In a very general result which captures all the above ones, Austrin and Mossel [3] showed that if $P : [q]^k \to \{0,1\}$ is a predicate such that the set of accepted inputs $P^{-1}(1)$ contains the support of a balanced pairwise independent distribution μ on $[q]^k$, then MAX k-CSP(P) is UG-hard to approximate better than a factor of $|P^{-1}(1)|/q^k$. Considering that a random assignment satisfies $|P^{-1}(1)|/q^k$ fraction of all the constraints, this is the strongest result one can get for a predicate P. Using appropriate choices for the predicate P, this then implies hardness ratios of $kq^2(1+o(1))/q^k$ for general $q \geq 2$, $q(q-1)k/q^k$ when q is a prime power, and $(k + O(k^{0.525})/2^k$ for $q = 2$.

We study the inapproximability of such a predicate P (which we call promising) in the hierarchy of linear programs defined by Sherali and Adams. In particular, we show an unconditional analogue of the result of Austrin and Mossel in this hierarchy.

Hierarchies of Linear and Semidefinite Programs A standard approach in approximating NP-hard problems, and therefore MAX k-CSP, is to formulate the problem as a 0-1 integer program and then relax the integrality condition to get a linear (or semidefinite) program which can be solved efficiently. The quality of such an approach is intimately related to the *integrality gap* of the relaxation, namely, the ratio between the optimum of the relaxation and that of the integer program.

Several methods (or procedures) were developed in order to obtain tightenings of relaxations in a systematic manner. These procedures give a sequence or a *hierarchy* of increasingly tighter relaxations of the starting program. The commonly studied ones include the hierarchies defined by Lovász-Schrijver [16], Sherali-Adams [24], and Lasserre [14] (see [15] for a comparison). Stronger relaxations in the sequence are referred to as higher *levels* of the hierarchy. It is known for all these hierarchies that for a starting program with n variables, the program at level n has integrality gap 1, and that it is possible to optimize over the program at the rth level in time $n^{O(r)}$.

Many known linear (semidefinite) programs can be captured by constant many levels of the Sherali-Adams (Lasserre) hierarchy. Fernández de la Vega and Kenyon-Mathieu [11] have provided a PTAS for Max Cut in dense graphs using Sherali-Adams. In [17] it is shown how to get a Sherali-Adams based PTAS for Vertex-Cover and Max-Independent-Set in minor-free graphs, while recently

Mathieu and Sinclair [18] showed that the integrality gap for the matching polytope is asymptotically $1+1/r$, and Bateni, Charikar and Guruswami [4] that the integrality gap for a natural LP formulation of the MaxMin allocation problem has integrality gap at most $n^{1/r}$, both after r many Sherali-Adams tightenings. Chlamtac [7] and Chlamtac and Singh [8] gave an approximation algorithm for Max-Independent-Set in hypergraphs based on the Lasserre hierarchy, with the performance depending on the number of levels.

Lower bounds in these hierarchies amount to showing that the integrality gap remains large even after many levels of the hierarchy. Integrality gaps for $\Omega(n)$ levels can be seen as unconditional lower bounds (as they rule out even exponential time algorithms obtained by the hierarchy) in a restricted (but still fairly interesting) model of computation. Considerable effort was invested in proving lower bounds (see $[2, 26, 25, 23, 5, 10, 1, 22, 12, 11]$). For CSPs in particular, strong lower bounds ($\Omega(n)$ levels) were proved recently for the Lasserre hierarchy (which is the strongest) by [21] and [27], who showed a factor 2 integrality gap for MAX k-XOR and factor $2^k/2k$ integrality gap for MAX k-CSP respectively.

Our Result and Techniques Both the results in the Lasserre hierarchy (and previous analogues in the Lovász-Schrijver hierarchy) seemed to be heavily relying on the structure of the predicate for which the integrality gap was proven, as being some system of linear equations. It was not clear if the techniques could be extended using only the fact that the predicate is promising (which is a much weaker condition). In this paper, we try to explore this issue, proving $\Omega(n)$ level gaps for the (admittedly weaker) Sherali-Adams hierarchy.

Theorem 1. *Let $P : [q]^k \to \{0,1\}$ be predicate such that $P^{-1}(1)$ contains the support of a balanced pairwise independent distribution μ. Then for every constant $\zeta > 0$, there exist $c = c(q, k, \zeta)$ such that for large enough n, the integrality gap of* MAX k-CSP*(P) for the tightening obtained by cn levels of the Sherali-Adams hierarchy applied to the standard LP*[3] *is at least* $\dfrac{q^k}{|P^{-1}(1)|} - \zeta$.

We note that $\Omega(n^\delta)$-level gaps for these predicates can also be deduced via reductions from the recent result of [6] who obtained $\Omega(n^\delta)$-level gaps for Unique Games, where $\delta \to 0$ as $\zeta \to 0$.

A first step in achieving our result is to reduce the problem of a level-t gap to a question about family of distributions over assignments associated with sets of variables of size at most t. These distributions should be (a) supported only on satisfying (partial) assignments and (b) should be consistent among themselves, in the sense that for $S_1 \subseteq S_2$ which are subsets of variables, the distributions over S_1 and S_2 should be equal on S_1. The second requirement guarantees that the obtained solution is indeed feasible, while the first implies that the solution achieves objective value that corresponds to satisfying *all* the constraints of the instance.

[3] See the resulting LP in section 2.3.

The second step is to come up with these distributions! We explain why the simple method of picking a uniform distribution (or a reweighting of it according to the pairwise independent distribution that is supported by P) over the satisfying assignments cannot work. Instead we introduce the notion of "advice sets". These are sets on which it is "safe" to define such simple distributions. The actual distribution for a set S we use is then the one induced on S by a simple distribution defined on the advice-set of S. Getting such advice sets heavily relies on notions of expansion of the constraints graph. In doing so, we use the fact that random instances have inherently good expansion properties. At the same time, such instances are highly unsatisfiable, ensuring that the resulting integrality gap is large.

Arguing that it is indeed "safe" to use simple distributions over the advice sets relies on the fact that the predicate P in question is promising, namely $P^{-1}(1)$ contains the support of a balanced pairwise independent distribution. We find it interesting and somewhat curious that the condition of pairwise independence comes up in this context for a reason very different than in the case of UG-hardness. Here, it represents the limit to which the expansion properties of a random CSP instance can be pushed to define such distributions.

2 Preliminaries and Notation

2.1 Constraint Satisfaction Problems

For an instance Φ of $\mathsf{MAX\ k\text{-}CSP}_q$, we denote the variables by $\{x_1, \ldots, x_n\}$, their domain $\{0, \ldots, q-1\}$ by $[q]$ and the constraints by $C_1, \ldots, C_m$. Each constraint is a function of the form $C_i : [q]^{T_i} \to \{0, 1\}$ depending only on the values of the variables in the ordered tuple T_i with $|T_i| \leq k$.

For a set of variables $S \subseteq [n]$, we denote by $[q]^S$ the set of all mappings from the set S to $[q]$. In context of variables, these mappings can be understood as partial assignments to a given subset of variables. For $\alpha \in [q]^S$, we denote its projection to $S' \subseteq S$ as $\alpha(S')$. Also, for $\alpha_1 \in [q]^{S_1}, \alpha_2 \in [q]^{S_2}$ such that $S_1 \cap S_2 = \emptyset$, we denote by $\alpha_1 \circ \alpha_2$ the assignment over $S_1 \cup S_2$ defined by α_1 and α_2.

We shall prove results for constraint satisfaction problems where every constraint is specified by the same Boolean predicate $P : [q]^k \to \{0, 1\}$. We denote the set of assignments which the predicate evaluates to 1 by $P^{-1}(1)$. A CSP instance for such a problem is a collection of constraints of the form of P applied to k-tuples of *literals*. For a variable x with domain $[q]$, we take a literal to be $(x + a) \bmod q$ for any $a \in [q]$. More formally,

Definition 1. *For a given $P : [q]^k \to \{0, 1\}$, an instance Φ of $\mathsf{MAX\ k\text{-}CSP}_q(P)$ is a set of constraints $C_1, \ldots, C_m$ where each constraint C_i is over a k-tuple of variables $T_i = \{x_{i_1}, \ldots, x_{i_k}\}$ and is of the form $P(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$ for some $a_{i_1}, \ldots, a_{i_k} \in [q]$. We denote the maximum number of constraints that can be simultaneously satisfied by $\mathsf{OPT}(\Phi)$.*

2.2 Expanding CSP Instances

For an instance Φ of MAX k-CSP$_q$, define its constraint graph G_Φ, as the following bipartite graph from L to R. The left hand side L consists of a vertex for each constraint C_i. The right hand side R consists of a vertex for every variable x_j. There is an edge between a constraint-vertex i and a variable-vertex j, whenever variable x_j appears in constraint C_i. When it is clear from the context, we will abbreviate G_Φ by G.

For $C_i \in L$ we denote by $\Gamma(C_i) \subseteq R$ the neighbors $\Gamma(C_i)$ of C_i in R. For a set of constraints $\mathcal{C} \subseteq L$, $\Gamma(\mathcal{C})$ denotes $\cup_{c_i \in \mathcal{C}} \Gamma(C_i)$. For $S \subseteq R$, we call a constraint $C_i \in L$, S-dominated if $\Gamma(C_i) \subseteq S$. We denote by $G|_{-S}$ the bipartite subgraph of G that we get after removing S and all S-dominated constraints. Finally, we also denote by $\mathcal{C}(S)$ the set of all S-dominated constraints.

Our result relies on set of constraints that are well expanding. We make this notion formal below.

Definition 2. *Consider a bipartite graph $G = (V, E)$ with partition L, R. The boundary expansion of $X \subset L$ is the value $|\partial X|/|X|$, where $\partial X = \{u \in R : |\Gamma(u) \cap X| = 1\}$. G is (r, e) boundary expanding if the boundary expansion for all subsets of L of size at most r is at least e.*

2.3 The Sherali-Adams Hierarchy

Below we present a relaxation for the MAX k-CSP$_q$ problem as it is obtained by applying a level-t Sherali-Adams tightening of the standard LP formulation of some instance Φ of MAX k-CSP$_q$. A well known fact states that the level-n Sherali-Adams tightening provides a perfect formulation, i.e. the integrality gap is 1 (see [24] or [15] for a proof).

The intuition behind the level-t Sherali-Adams tightening is the following. Note that an integer solution to the problem can be given by a single mapping $\alpha_0 \in [q]^{[n]}$, which is an assignment to all the variables. Using this, we can define 0/1 variables $X_{(S,\alpha)}$ for each $S \subseteq [n]$ such that $|S| \leq t$ and $\alpha \in [q]^S$. The intended solution is $X_{(S,\alpha)} = 1$ if $\alpha_0(S) = \alpha$ and 0 otherwise. We introduce $X_{(\emptyset,\emptyset)}$ which is intended to be 1. By relaxing the integrality constraint on the variables, we obtain the level-t Sherali-Adams LP tightening.

Level-t (for $t \geq k$) Sherali-Adams LP tightening for a MAX k-CSP$_q$ instance Φ

$$\text{maximize} \quad \sum_{i=1}^{m} \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) \cdot X_{(T_i,\alpha)}$$

$$\text{subject to} \quad \sum_{j \in [q]} X_{(S \cup \{i\}, \alpha \circ j)} = X_{(S,\alpha)} \quad \forall S \text{ s.t. } |S| < t,\ \forall i \notin S, \alpha \in [q]^S$$

$$X_{(S,\alpha)} \geq 0 \qquad \forall S \text{ s.t. } |S| \leq t,\ \forall \alpha \in [q]^S$$

$$X_{(\emptyset,\emptyset)} = 1$$

For an LP formulation of $\mathsf{MAX\ k\text{-}CSP}_q$, and for a given instance Φ of the problem, we denote by $\mathsf{FRAC}(\Phi)$ the LP (fractional) optimum, and by $\mathsf{OPT}(\Phi)$ the integral optimum. For the particular instance Φ, the integrality gap is then defined as $\mathsf{FRAC}(\Phi)/\mathsf{OPT}(\Phi)$. The integrality gap of the LP formulation is the supremum of integrality gaps over all instances.

Next we give a sufficient condition for the existence of a solution to the level-t Sherali-Adams LP tightening for a $\mathsf{MAX\ k\text{-}CSP}_q$ instance Φ.

Lemma 1. *Consider a family of distributions* $\{\mathcal{D}(S)\}_{S\subseteq[n]:|S|\leq t}$*, where each* $\mathcal{D}(S)$ *is defined over* $[q]^S$*. If for every* $S \subseteq T \subseteq [n]$ *with* $|T| \leq t$*, the distributions* $\mathcal{D}(S), \mathcal{D}(T)$ *are equal on* S*, then*

$$X_{(S,\alpha)} = \Pr_{\mathcal{D}(S)}[\alpha]$$

satisfy the above level-t Sherali-Adams tightening.

Proof. Consider some $S \subseteq [n]$, $|S| < t$, and some $i \notin S$. Note that the distributions $\mathcal{D}(S), \mathcal{D}(S \cup \{i\})$ are equal on S, and therefore we have

$$\begin{aligned}
\sum_{j\in[q]} X_{(S\cup\{i\},\alpha\circ j)} &= \sum_{j\in[q]} \Pr_{\beta\sim\mathcal{D}(S\cup\{i\})}[\beta = \alpha \circ j] \\
&= \sum_{j\in[q]} \Pr_{\beta\sim\mathcal{D}(S\cup\{i\})}[(\beta(i) = j) \wedge (\beta(S) = \alpha)] \\
&= \Pr_{\beta\sim\mathcal{D}(S\cup\{i\})}[\beta(S) = \alpha] \\
&= \Pr_{\beta'\sim\mathcal{D}(S)}[\beta' = \alpha] \\
&= X_{(S,\alpha)}.
\end{aligned}$$

The same argument also shows that if $S = \emptyset$, then $X_{(\emptyset,\emptyset)} = 1$. Finally, it is clear that all linear variables are assigned non negative values completing the lemma.

2.4 Pairwise Independence and Approximation Resistant Predicates

We say that a distribution μ over variables $x_1, \ldots, x_k$, is a balanced pairwise independent distribution over $[q]^k$, if we have

$$\forall j \in [q].\forall i.\ \Pr_\mu[x_i = j] = \frac{1}{q} \quad \text{and} \quad \forall j_1, j_2 \in [q].\forall i_1 \neq i_2.\ \Pr_\mu[(x_{i_1} = j_1)\wedge(x_{i_2} = j_2)] = \frac{1}{q^2}.$$

A predicate P is called approximation resistant if it is hard to approximate the $\mathsf{MAX\ k\text{-}CSP}_q(P)$ problem better than using a random assignment. Assuming the Unique Games Conjecture, Austrin and Mossel [3] show that a predicate is approximation resistant if it is possible to define a balanced pairwise independent distribution μ such that P is always 1 on the support of μ.

Definition 3. *A predicate* $P : [q]^k \to \{0, 1\}$ *is called* promising*, if there exist a distribution supported over a subset of* $P^{-1}(1)$ *that is pairwise independent and balanced. If* μ *is such a distribution we say that* P *is promising supported by* μ*.*

3 Towards Defining Consistent Distributions

To construct valid solutions for the Sherali-Adams LP tightening, we need to define distributions over every set S of bounded size as is required by Lemma 1. Since we will deal with promising predicates supported by some distribution μ, in order to satisfy consistency between distributions we will heavily rely on the fact that μ is a balanced pairwise independent distribution.

Consider for simplicity that μ is uniform over $P^{-1}(1)$ (the intuition for the general case is not significantly different). It is instructive to think of $q = 2$ and the predicate P being k-XOR, $k \geq 3$. Observe that the uniform distribution over $P^{-1}(1)$ is pairwise independent and balanced. A first attempt would be to define for every S, the distribution $\mathcal{D}(S)$ as the uniform distribution over all consistent assignments of S. We argue that such distributions are in general problematic. This follows from the fact that satisfying assignments are not always extendible. Indeed, consider two constraints $C_{i_1}, C_{i_2} \in L$ that share a common variable $j \in R$. Set $S_2 = T_{i_1} \cup T_{i_2}$, and $S_1 = S_2 \setminus \{j\}$. Assuming that the support of no other constraint is contained in S_2, we get that distribution $\mathcal{D}(S_1)$ maps any variable in S_1 to $\{0, 1\}$ with probability 1/2 independently, but some of these assignments are not even extendible to S_2 meaning that $\mathcal{D}(S_2)$ will assign them with probability zero.

Thus, to define $\mathcal{D}(S)$, we cannot simply sample assignments satisfying all constraints in $\mathcal{C}(S)$ with probabilities given by μ. In fact the above example shows that any attempt to blindly assign a set S with a distribution that is supported on all satisfying assignments for S is bound to fail. At the same time it seems hard to reason about a distribution that uses a totally different concept. To overcome this obstacle, we take a two step approach:

1. For a set S we define a superset $\overline{S}$ such that $\overline{S}$ is "global enough" to contain sufficient information, while it also is "local enough" so that $\mathcal{C}(\overline{S})$ is not too large. We require the property of such sets that if we remove $\overline{S}$ and $\mathcal{C}(\overline{S})$, then the remaining graph $G|_{-\overline{S}}$ still has good expansion. We deal with this in Section 3.1.
2. The distribution $\mathcal{D}(S)$ is going to be the uniform distribution over satisfying assignments in $\overline{S}$. In the case that μ is not uniform over $P^{-1}(1)$, we give a natural generalization to the above uniformity. We show how to define distributions, which we denote by $\mathcal{P}_\mu(S)$, such that for $S_1 \subseteq S_2$, the distributions are guaranteed to be consistent if $G|_{-S_1}$ *has good expansion*. This appears in Section 3.2.

We then combine the two techniques and define $\mathcal{D}(S)$ according to $\mathcal{P}_\mu(\overline{S})$. This is done in section 4.

3.1 Finding Advice-Sets

We now give an algorithm below to obtain a superset $\overline{S}$ for a given set S, which we call the advice-set of S. It is inspired by the "expansion correction" procedure in [5].

Algorithm Advice

The input is an (r, e_1) boundary expanding bipartite graph $G = (L, R, E)$, some $e_2 \in (0, e_1)$, and some $S \subseteq R$, $|S| < (e_1 - e_2)r$, with some order $S = \{x_1, \ldots, x_t\}$.

Initially set $\overline{S} \leftarrow \emptyset$ and $\xi \leftarrow r$
For $j = 1, \ldots, |S|$ **do**
  $M_j \leftarrow \emptyset$
  $\overline{S} \leftarrow \overline{S} \cup \{x_j\}$
  If $G|_{-\overline{S}}$ is not (ξ, e_2) boundary expanding **then**
    Find a maximal $M_j \subset L$ in $G|_{-\overline{S}}$, such that $|M_j| \leq \xi$ in $G|_{-\overline{S}}$ and $|\partial M_j| \leq e_2 |M_j|$
    $\overline{S} \leftarrow \overline{S} \cup \partial M_j$
    $\xi \leftarrow \xi - |M_j|$
Return $\overline{S}$

Theorem 2. *Algorithm Advice, with internal parameters e_1, e_2, r, returns $\overline{S} \subseteq R$ such that (a) $G|_{-\overline{S}}$ is (ξ_S, e_2) boundary expanding, (b) $\xi_S \geq r - \frac{|S|}{e_1 - e_2}$, and (c) $|\overline{S}| \leq \frac{e_1 |S|}{e_1 - e_2}$.*

Proof. Suppose that the loop terminates with $\xi = \xi_S$. Then $\sum_{j=1}^{t} |M_j| = r - \xi_S$. Since G is (r, e_1) boundary expanding, the set $M = \cup_{j=1}^{t} M_j$ has initially at least $e_1(r - \xi_S)$ boundary neighbors in G. During the execution of the while loop, each set M_j has at most $e_2 |M_j|$ boundary neighbors in $G|_{-\overline{S}}$. Therefore, at the end of the procedure M has at most $e_2(r - \xi_S)$ boundary neighbors in $G|_{-S}$. It follows that $|S| + e_2(r - \xi_S) \geq e_1(r - \xi_S)$, which implies (b).

From the bound size of S we know that $\xi_S > 0$. In particular, ξ remains positive throughout the execution of the while loop. Next we identify a loop invariant: $G|_{-\overline{S}}$ is (ξ, e_2) boundary expanding.

Indeed, note that the input graph G is (ξ, e_1) boundary expanding. At step j consider the set $\overline{S} \cup \{x_j\}$, and suppose that $G_{-(\overline{S} \cup \{x_j\})}$ is not (ξ, e_2) boundary expanding. We find maximal M_j, $|M_j| \leq \xi$, such that $|\partial M_j| \leq e_2 |M_j|$. We claim that $G_{-(\overline{S} \cup \{x_j\} \cup \partial M_j)}$ is $(\xi - |M_j|, e_2)$ boundary expanding (recall that since ξ remains positive, $|M_j| < \xi$). Now consider the contrary. Then, there must be $M' \subset L$ such that $|M'| \leq \xi - |M_j|$ and such that $|\partial M'| \leq e_2 |M'|$. Consider then $M_j \cup M'$ and note that $|M_j \cup M'| \leq \xi$. More importantly $|\partial(M_j \cup M')| \leq e_2 |M_j \cup M'|$, and therefore we contradict the maximality of M_j; (a) follows.

Finally note that $\overline{S}$ consists of S union the boundary neighbors of all M_j. From the arguments above, the number of those neighbors does not exceed $e_2(r - \xi_S)$ and hence $|\overline{S}| \leq |S| + e_2(r - \xi_S) \leq |S| + \frac{e_2 |S|}{e_1 - e_2} = \frac{e_1 |S|}{e_1 - e_2}$, which proves (c).

3.2 Defining the Distributions $\mathcal{P}_\mu(S)$

We now define for every set S, a distribution $\mathcal{P}_\mu(S)$ such that for any $\alpha \in [q]^S$, $\Pr_{\mathcal{P}_\mu(S)}[\alpha] > 0$ only if α satisfies all the constraints in $\mathcal{C}(S)$. For a constraint C_i

with set of inputs T_i, defined as $C_i(x_{i_1}, \ldots, x_{i_k}) \equiv P(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$, let $\mu_i : [q]^{T_i} \to [0, 1]$ denote the distribution

$$\mu_i(x_{i_1}, \ldots, x_{i_k}) = \mu(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$$

so that the support of μ_i is contained in $C_i^{-1}(1)$. We then define the distribution $\mathcal{P}_\mu(S)$ by picking each assignment $\alpha \in [q]^S$ with probability proportional to $\prod_{C_i \in \mathcal{C}(S)} \mu_i(\alpha(T_i))$. Formally,

$$\Pr_{\mathcal{P}_\mu(S)}[\alpha] \;=\; \frac{1}{Z_S} \cdot \prod_{C_i \in \mathcal{C}(S)} \mu_i(\alpha(T_i)) \tag{1}$$

where $\alpha(T_i)$ is the restriction of α to T_i and Z_S is a normalization factor given by

$$Z_S \;=\; \sum_{\alpha \in [q]^S} \prod_{C_i \in \mathcal{C}(S)} \mu_i(\alpha(T_i)).$$

To understand the distribution, it is easier to think of the special case when μ is just the uniform distribution on $P^{-1}(1)$ (like in the case of MAX k-XOR). Then $\mathcal{P}_\mu(S)$ is simply the uniform distribution on assignments satisfying all the constraints in $\mathcal{C}(S)$. When μ is not uniform, then the probabilities are weighted by the product of the values $\mu_i(\alpha(T_i))$ for all the constraints [4]. However, we still have the property that if $\Pr_{\mathcal{P}_\mu(S)}[\alpha] > 0$, then α satisfies all the constraints in $\mathcal{C}(S)$.

In order for the distribution $\mathcal{P}_\mu(S)$ to be well defined, we need to ensure that $Z_S > 0$. The following lemma shows how to calculate Z_S if G is sufficiently expanding, and simultaneously proves that if $S_1 \subseteq S_2$, and if $G|_{-S_1}$ is sufficiently expanding, then $\mathcal{P}_\mu(S_1)$ is consistent with $\mathcal{P}_\mu(S_2)$ over S_1.

Lemma 2. *Let Φ be a* MAX k-CSP*(P) instance as above and $S_1 \subseteq S_2$ be two sets of variables such that both G and $G|_{-S_1}$ are $(r, k-2-\delta)$ boundary expanding for some $\delta \in (0, 1)$ and $|\mathcal{C}(S_2)| \le r$. Then $Z_{S_2} = q^{|S_2|}/q^{k|\mathcal{C}(S_2)|}$, and for any $\alpha_1 \in [q]^{S_1}$*

$$\sum_{\substack{\alpha_2 \in [q]^{S_2} \\ \alpha_2(S_1) = \alpha_1}} \Pr_{\mathcal{P}_\mu(S_2)}[\alpha_2] = \Pr_{\mathcal{P}_\mu(S_1)}[\alpha_1].$$

Proof. Let $\mathcal{C} = \mathcal{C}(S_2) \backslash \mathcal{C}(S_1)$ be given by the set of t many constraints $C_{i_1}, \ldots, C_{i_t}$ with each C_{i_j} being on the set of variables T_{i_j}. Some of these variables may be fixed by α_1. Also, any α_2 consistent with α_1 can be written as $\alpha_1 \circ \alpha$ for some $\alpha \in [q]^{S_2 \backslash S_1}$. Below, we express these probabilities in terms the product of μ on the constraints in $\mathcal{C}(S_2) \setminus \mathcal{C}(S_1)$.

Note that the equations below are still correct even if we haven't shown $Z_{S_2} > 0$ (in that case both sides are 0). In fact, replacing S_1 by $\emptyset$ in the same

[4] Note however that $\mathcal{P}_\mu(S)$ is not a product distribution because different constraints in $\mathcal{C}(S)$ may share variables.

calculation will give the value of Z_{S_2}.

$$\begin{aligned}
Z_{S_2} \cdot \sum_{\substack{\alpha_2 \in [q]^{S_2} \\ \alpha_2(S_1)=\alpha_1}} \Pr_{\mathcal{P}_\mu(S_2)}[\alpha_2] &= \sum_{\alpha \in [q]^{S_2 \setminus S_1}} \prod_{C_i \in \mathcal{C}(S_2)} \mu_i((\alpha_1 \circ \alpha)(T_i)) \\
&= \left(\prod_{C_i \in \mathcal{C}(S_1)} \mu_i(\alpha_1(T_i)) \right) \sum_{\alpha \in [q]^{S_2 \setminus S_1}} \prod_{j=1}^{t} \mu_{i_j}((\alpha_1 \circ \alpha)(T_{i_j})) \\
&= \left(Z_{S_1} \cdot \Pr_{\mathcal{P}_\mu(S_1)}[\alpha_1] \right) \sum_{\alpha \in [q]^{S_2 \setminus S_1}} \prod_{j=1}^{t} \mu_{i_j}((\alpha_1 \circ \alpha)(T_{i_j})) \\
&= \left(Z_{S_1} \cdot \Pr_{\mathcal{P}_\mu(S_1)}[\alpha_1] \right) \cdot q^{|S_2 \setminus S_1|} \mathop{\mathbb{E}}_{\alpha \in [q]^{S_2 \setminus S_1}} \left[\prod_{j=1}^{t} \mu_{i_j}((\alpha_1 \circ \alpha)(T_{i_j})) \right]
\end{aligned}$$

The following claim, whose proof can be found in the Appendix, lets us calculate this expectation conveniently using the expansion of $G|_{-S_1}$.

Claim. Let $\mathcal{C}$ be as above. Then there exists an ordering $C_{i'_1}, \ldots, C_{i'_t}$ of constraints in $\mathcal{C}$ and a partition of $S_2 \setminus S_1$ into sets of variables $F_1, \ldots, F_t$ such that for all j, $F_j \subseteq T_{i'_j}$, $|F_j| \geq k-2$, and

$$\forall j \;\; F_j \cap \left(\cup_{l>j} T_{i'_l} \right) = \emptyset.$$

Using this decomposition, the expectation above can be split as

$$\mathop{\mathbb{E}}_{\alpha \in [q]^{S_2 \setminus S_1}} \left[\prod_{j=1}^{t} \mu_{i_j}(\alpha_1 \circ \alpha(T_{i_j})) \right] = \mathop{\mathbb{E}}_{\beta_t \in [q]^{F_t}} \left[\mu_{i'_t} \cdots \mathop{\mathbb{E}}_{\beta_2 \in [q]^{F_2}} \left[\mu_{i'_2} \mathop{\mathbb{E}}_{\beta_1 \in [q]^{F_1}} \left[\mu_{i'_1} \right] \right] \cdots \right]$$

where the input to each $\mu_{i'_j}$ depends on α_1 and $\beta_j, \ldots, \beta_t$ but not on $\beta_1, \ldots, \beta_{j-1}$.

We now reduce the expression from right to left. Since F_1 contains at least $k-2$ variables and $\mu_{i'_1}$ is a balanced pairwise independent distribution,

$$\mathop{\mathbb{E}}_{\beta_1 \in [q]^{F_1}} \left[\mu_{i'_1} \right] = \frac{1}{q^{|F_1|}} \cdot \Pr_\mu[(\alpha_1 \circ \beta_2 \ldots \circ \beta_t)(T_{i'_1} \setminus F_1)] = \frac{1}{q^k}$$

irrespective of the values assigned by $\alpha_1 \circ \beta_2 \circ \ldots \circ \beta_t$ to the remaining (at most 2) variables in $T_{i'_1} \setminus F_1$. Continuing in this fashion from right to left, we get that

$$\mathop{\mathbb{E}}_{\alpha \in [q]^{S_2 \setminus S_1}} \left[\prod_{j=1}^{t} \mu_{i_j}((\alpha_1 \circ \alpha)(T_{i_j})) \right] = \left(\frac{1}{q^k} \right)^t = \left(\frac{1}{q^k} \right)^{|\mathcal{C}(S_2) \setminus \mathcal{C}(S_1)|}$$

Hence, we get that

$$Z_{S_2} \cdot \sum_{\substack{\alpha_2 \in [q]^{S_2} \\ \alpha_2(S_1)=\alpha_1}} \Pr_{\mathcal{P}_\mu(S_2)}[\alpha_2] = \left(Z_{S_1} \cdot \frac{q^{|S_2 \setminus S_1|}}{q^{k|\mathcal{C}(S_2) \setminus \mathcal{C}(S_1)|}} \right) \Pr_{\mathcal{P}_\mu(S_1)}[\alpha_1]. \quad (2)$$

Summing over all $\alpha_1 \in [q]^{S_1}$ on both sides gives

$$Z_{S_2} \;=\; Z_{S_1} \cdot \frac{q^{|S_2 \setminus S_1|}}{q^{k|\mathcal{C}(S_2)\setminus\mathcal{C}(S_1)|}}.$$

Since we know that G is $(r, k-2-\delta)$ boundary expanding, we can replace S_1 by $\emptyset$ in the above equation to obtain $Z_{S_2} = q^{|S_2|}/q^{k|\mathcal{C}(S_2)|}$ as claimed. Also note that since $\mathcal{C}(S_1) \subseteq \mathcal{C}(S_2)$, $Z_{S_2} > 0$ implies $Z_{S_1} > 0$. Hence, using equation (2) we get

$$\sum_{\substack{\alpha_2 \in [q]^{S_2} \\ \alpha_2(S_1)=\alpha_1}} \Pr_{\mathcal{P}_\mu(S_2)}[\alpha_2] \;=\; \Pr_{\mathcal{P}_\mu(S_1)}[\alpha_1]$$

which proves the lemma.

4 Constructing the Integrality Gap

We now show how to construct integrality gaps using the ideas in the previous section. For a given promising predicate P, our integrality gap instance will be random instance Φ of the $\mathsf{MAX\ k\text{-}CSP}_q(P)$ problem. To generate a random instance with m constraints, for every constraint C_i, we randomly select a k-tuple of distinct variables $T_i = \{x_{i_1}, \ldots, x_{i_k}\}$ and $a_{i_1}, \ldots, a_{i_k} \in [q]$, and put $C_i \equiv P(x_{i_1} + a_{i_1}, \ldots, x_{i_k} + a_{i_k})$. It is well known and used in various works on integrality gaps and proof complexity (e.g. [5], [1], [22] and [21]), that random instances of CSPs are both highly unsatisfiable and highly expanding. We capture the properties we need in the lemma below (for a proof see e.g. [27]).

Lemma 3. *Let $\epsilon, \delta > 0$ and a predicate $P : [q]^k \to \{0,1\}$ be given. Then there exist $\gamma = O(q^k \log q/\epsilon^2)$, $\eta = \Omega((1/\gamma)^{10/\delta})$ and $N \in \mathbb{N}$, such that if $n \geq N$ and Φ is a random instance of* $\mathsf{MAX\ k\text{-}CSP}$*$(P)$ with $m = \gamma n$ constraints, then with probability $1 - o(1)$*

1. $\mathsf{OPT}(\Phi) \leq \frac{|P^{-1}(1)|}{q^k}(1+\epsilon) \cdot m$.
2. *For any set $\mathcal{C}$ of constraints with $|\mathcal{C}| \leq \eta n$, we have $|\partial(\mathcal{C})| \geq (k-2-\delta)|\mathcal{C}|$.*

Let Φ be an instance of $\mathsf{MAX\ k\text{-}CSP}_q$ on n variables for which G_Φ is $(\eta n, k-2-\delta)$ boundary expanding for some $\delta < 1/2$, as in Lemma 3. For such a Φ, we now define the distributions $\mathcal{D}(S)$.

For a set S of size at most $t = \eta\delta n/4k$, let $\overline{S}$ be subset of variables output by the algorithm Advice when run with input S and parameters $r = \eta n, e_1 = (k-2-\delta), e_2 = (k-2-2\delta)$ on the graph G_Φ. Theorem 2 shows that

$$|\overline{S}| \;\leq\; (k-2-\delta)|S|/\delta \;\leq\; \eta n/4.$$

We then use (1) to define the distribution $\mathcal{D}(S)$ for sets S of size at most $\delta\eta n/4k$ as

$$\Pr_{\mathcal{D}(S)}[\alpha] \quad = \quad \sum_{\substack{\beta \in [q]^{\overline{S}} \\ \beta(S)=\alpha}} \Pr_{\mathcal{P}_\mu(\overline{S})}[\beta].$$

Using the properties of the distributions $\mathcal{P}_\mu(\overline{S})$, we can now prove that the distributions $\mathcal{D}(S)$ are consistent.

Claim. Let the distributions $\mathcal{D}(S)$ be defined as above. Then for any two sets $S_1 \subseteq S_2 \subseteq [n]$ with $|S_2| \leq t = \eta\delta n/4k$, the distributions $\mathcal{D}(S_1), \mathcal{D}(S_2)$ are equal on S_1.

Proof. The distributions $\mathcal{D}(S_1), \mathcal{D}(S_2)$ are defined according to $\mathcal{P}_\mu(\overline{S}_1)$ and $\mathcal{P}_\mu(\overline{S}_2)$ respectively. To prove the claim, we show that $\mathcal{P}_\mu(\overline{S}_1)$ and $\mathcal{P}_\mu(\overline{S}_2)$ are equal to the distribution $\mathcal{P}_\mu(\overline{S}_1 \cup \overline{S}_2)$ on $\overline{S}_1, \overline{S}_2$ respectively (note that it need not be the case that $\overline{S}_1 \subseteq \overline{S}_2$).

Let $S_3 = \overline{S}_1 \cup \overline{S}_2$. Since $|\overline{S}_1|, |\overline{S}_2| \leq \eta n/4$, we have $|S_3| \leq \eta n/2$ and hence $|\mathcal{C}(S_3)| \leq \eta n/2$. Also, by Theorem 2, we know that both $G|_{-\overline{S}_1}$ and $G|_{-\overline{S}_2}$ are $(2\eta n/3, k-2-2\delta)$ boundary expanding. Thus, using Lemma 2 for the pairs $(\overline{S}_1, S_3)$ and $(\overline{S}_2, S_3)$, we get that

$$\begin{aligned} \Pr_{\mathcal{D}(S_1)}[\alpha_1] &= \sum_{\substack{\beta_1 \in [q]^{\overline{S}_1} \\ \beta_1(S_1)=\alpha_1}} \Pr_{\mathcal{P}_\mu(\overline{S}_1)}[\beta_1] \\ &= \sum_{\substack{\beta_3 \in [q]^{S_3} \\ \beta_3(S_1)=\alpha_1}} \Pr_{\mathcal{P}_\mu(S_3)}[\beta_3] \\ &= \sum_{\substack{\beta_2 \in [q]^{\overline{S}_2} \\ \beta_2(S_1)=\alpha_1}} \Pr_{\mathcal{P}_\mu(\overline{S}_2)}[\beta_2] \\ &= \sum_{\substack{\alpha_2 \in [q]^{S_2} \\ \alpha_2(S_1)=\alpha_1}} \Pr_{\mathcal{D}(S_2)}[\alpha_2] \end{aligned}$$

which shows that $\mathcal{D}(S_1)$ and $\mathcal{D}(S_2)$ are equal on S_1.

It is now easy to prove the main result.

Theorem 3. *Let $P : [q]^k \to \{0,1\}$ be a promising predicate. Then for every constant $\zeta > 0$, there exist $c = c(q,k,\zeta)$, such that for large enough n, the integrality gap of* MAX k-CSP*(P) for the tightening obtained by cn levels of the Sherali-Adams hierarchy is at least* $\dfrac{q^k}{|P^{-1}(1)|} - \zeta$.

Proof. We take $\epsilon = \zeta/q^k, \delta = 1/4$ and consider a random instance Φ of MAX k-CSP(P) with $m = \gamma n$ as given by Lemma 3. Thus, $\mathsf{OPT}(\Phi) \leq \frac{|P^{-1}(1)|}{q^k}(1+\epsilon)\cdot m$.

On the other hand, by Claim 4 we can define distributions $\mathcal{D}(S)$ over every set of at most $\delta\eta n/4k$ variables such that for $S_1 \subseteq S_2$, $\mathcal{D}(S_1)$ and $\mathcal{D}(S_2)$ are consistent over S_1. By Lemma 1 this gives a feasible solution to the LP obtained by $\delta\eta n/4k$ levels. Also, by definition of $\mathcal{D}(S)$, we have that $\Pr_{\mathcal{D}(S)}[\alpha] > 0$ only if α satisfies all constraints in $\mathcal{C}(S)$. Hence, the value of $\mathsf{FRAC}(\Phi)$ is given by

$$\sum_{i=1}^m \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) X_{(T_i,\alpha)} = \sum_{i=1}^m \sum_{\alpha \in [q]^{T_i}} C_i(\alpha) \Pr_{\mathcal{D}(T_i)}[\alpha] = \sum_{i=1}^m \sum_{\alpha \in [q]^{T_i}} \Pr_{\mathcal{D}(T_i)}[\alpha] = m.$$

Thus, the integrality gap after $\delta\eta n/4k$ levels is at least

$$\frac{\mathsf{FRAC}(\Phi)}{\mathsf{OPT}(\Phi)} = \frac{q^k}{|P^{-1}(1)|(1+\epsilon)} \geq \frac{q^k}{|P^{-1}(1)|} - \zeta.$$

References

1. Mikhail Alekhnovich, Sanjeev Arora, and Iannis Tourlakis. Towards strong nonapproximability results in the Lovász-Schrijver hierarchy. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*. ACM, 2005.
2. Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis. Proving integrality gaps without knowing the linear program. *Theory of Computing*, 2(2):19–51, 2006.
3. Per Austrin and Elchanan Mossel. Approximation resistant predicates from pairwise independence. In *IEEE Conference on Computational Complexity*, pages 249–258. IEEE Computer Society, 2008.
4. Mohammad Hossein Bateni, Moses Charikar, and Venkatesan Guruswami. MaxMin allocation via degree lower-bounded arborescences. In *STOC'2009*. ACM Press, 2009.
5. Joshua Buresh-Oppenheim, Nicola Galesi, Shlomo Hoory, Avner Magen, and Toniann Pitassi. Rank bounds and integrality gaps for cutting planes procedures. *Theory of Computing*, 2(4):65–90, 2006.
6. Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Integrality gaps for Sherali-Adams relaxations. In *STOC'2009*. ACM Press, 2009.
7. Eden Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. In *FOCS: IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 691–701, 2007.
8. Eden Chlamtac and Gyanit Singh. Improved approximation guarantees through higher levels of SDP hierarchies. In *Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX) 2008*, 2008.
9. Lars Engebretsen and Jonas Holmerin. More efficient queries in PCPs for NP and improved approximation hardness of maximum CSP. In *STACS 2005, 22nd Annual Symposium on Theoretical Aspects of Computer Science, Stuttgart, Germany, February 24-26, 2005, Proceedings*. Springer, 2005.
10. Uriel Feige and Robert Krauthgamer. The probable value of the Lovász-Schrijver relaxations for maximum independent set. *SICOMP: SIAM Journal on Computing*, 32(2):345–370, 2003.
11. Wenceslas Fernández de la Vega and Claire Kenyon-Mathieu. Linear programming relaxations of maxcut. In *SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 53–61, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics.
12. Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis. Integrality gaps of $2 - o(1)$ for Vertex Cover SDPs in the Lovász-Schrijver hierarchy. In *Proceedings of the 47th IEEE Symposium on Foundations of Computer Science*, pages 702–712, 2007.
13. Venkatesan Guruswami and Prasad Raghavendra. Constraint satisfaction over a non-boolean domain: Approximation algorithms and unique-games hardness. In *11th International Workshop, APPROX 2008, Boston, MA, USA, August 25-27, 2008. Proceedings*. Springer, 2008.

14. Jean B. Lasserre. An explicit exact SDP relaxation for nonlinear 0-1 programs. In *Integer programming and combinatorial optimization (Utrecht, 2001)*, volume 2081 of *Lecture Notes in Comput. Sci.*, pages 293–303. Springer, Berlin, 2001.
15. Monique Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Math. Oper. Res.*, 28(3):470–496, 2003.
16. László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, May 1991.
17. Avner Magen and Mohammad Moharrami. Sherali-Adams based polynomial approximation schemes for NP-hard problems on planar and minor-free graphs. Manuscript, 2008.
18. Claire Mathieu and Alistair Sinclair. Sherali-Adams relaxations of the matching polytope. In *STOC'2009*. ACM Press, 2009.
19. Alex Samorodnitsky and Luca Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing, STOC'2000 (Portland, Oregon, May 21-23, 2000)*, pages 191–199. ACM Press, 2000.
20. Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and PCPs. In *stoc06*, pages 11–20, 2006.
21. Grant Schoenebeck. Linear level lasserre lower bounds for certain k-CSPs. In *FOCS*, pages 593–602. IEEE Computer Society, 2008.
22. Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. A linear round lower bound for Lovász-Schrijver SDP relaxations of vertex cover. In *IEEE Conference on Computational Complexity*, pages 205–216. IEEE Computer Society, 2007.
23. Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. Tight integrality gaps for Lovász-Schrijver LP relaxations of vertex cover and max cut. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 11-13, 2007*. ACM, 2007.
24. Hanif D. Sherali and Warren P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.*, 3(3):411–430, 1990.
25. Iannis Tourlakis. Towards optimal integrality gaps for hypergraph vertex cover in the Lovász-Schrijver hierarchy. In *8th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2005*, volume 3624, pages 233–244. Springer, 2005.
26. Iannis Tourlakis. New lower bounds for vertex cover in the Lovász-Schrijver hierarchy. In *Proceedings of the 21st IEEE Conference on Computational Complexity*, pages 170–182. IEEE Computer Society, 2006.
27. Madhur Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *STOC'2009*. ACM Press, 2009.

Appendix

Proof. (of Claim 3.2) We build the sets F_j inductively using the fact that $G|_{-S_1}$ is $(r, k-2-\delta)$ boundary expanding.

Start with the set of constraints $\mathcal{C}_1 = \mathcal{C}$. Since $|\mathcal{C}_1| = |\mathcal{C}(S_2) \setminus \mathcal{C}(S_1)| \leq r$, this gives that $|\partial(\mathcal{C}_1) \setminus S_1| \geq (k-2-\delta)|\mathcal{C}_1|$. Hence, there exists $C_{i_j} \in \mathcal{C}_1$ such that $|T_{i_j} \cap (\partial(\mathcal{C}_1) \setminus S_1)| \geq k-2$. Let $T_{i_j} \cap (\partial(\mathcal{C}_1) \setminus S_1) = F_1$ and $i'_1 = i_j$. We then take $\mathcal{C}_2 = \mathcal{C}_1 \setminus \{C_{i'_1}\}$ and continue in the same way.

Since at every step, we have $F_j \subseteq \partial(\mathcal{C}_j) \setminus S_1$, and for all $l > j$ $\mathcal{C}_l \subseteq \mathcal{C}_j$, F_j shares no variables with $\Gamma(\mathcal{C}_l)$ for $l > j$. Hence, we get $F_j \cap \left(\cup_{l>j} T_{i'_l}\right) = \emptyset$ as claimed.