Clique Is Hard on Average for Unary Sherali-Adams

Susanna F. De Rezende
Lund University

Aaron Potechin
University of Chicago

Kilian Risse
EPFL

***Abstract*—We prove that unary Sherali-Adams requires proofs of size $n^{\Omega(d)}$ to rule out the existence of an $n^{\Theta(1)}$-clique in Erdős-Rényi random graphs whose maximum clique is of size $d \leq 2\log n$. This lower bound is tight up to the multiplicative constant in the exponent. We obtain this result by introducing a technique inspired by pseudo-calibration which may be of independent interest. The technique involves defining a measure on monomials that precisely captures the contribution of a monomial to a refutation. This measure intuitively captures progress and should have further applications in proof complexity.**

***Index Terms*—Proof Complexity, Clique, Unary Sherali Adams**

I. Introduction

The problem of identifying a maximum clique in a given graph, that is, finding a fully connected subgraph of maximum size, is one of the fundamental problems of theoretical computer science. It is one of the first combinatorial problems proven **NP**-hard by Karp in the 1970s [1] and was even mentioned in Cook's paper [2] introducing the theory of **NP**-complete problems. This problem is also notoriously hard to approximate: unless $\mathbf{P} = \mathbf{NP}$, the size of the maximum clique cannot be approximated within a factor of $n^{1-\varepsilon}$ [3], [4].

The k-clique problem, determining whether there is a clique of size k in a given n-vertex graph, can be solved by iterating over all subsets of vertices of size k and checking whether one of them is a clique. Somewhat surprisingly, this naïve algorithm, which runs in time $O(n^k)$, is believed to be essentially tight: the constant in the exponent can be slightly improved by a clever use of matrix multiplication [5] but unless the class of fixed parameter tractable problems collapses to $\mathbf{W}[1]$, there must be some dependence on k in the exponent [6], and under the exponential time hypothesis [7] this dependency must be linear [8].

While the clique problem is quite well understood in the worst case and under standard hardness assumptions, much less is known in the average-case setting. For example, it is not known whether there are algorithms running in time $n^{o(k)}$ that, given an Erdős-Rényi graph with edge probability just below the threshold of containing a k-clique, can determine that the graph does not contain a k-clique. Even if we only require the algorithm to refute the existence of a clique of size $n^{\varepsilon} \gg k$, this problem is still conjectured to be hard.

Such average-case questions seem difficult to relate to worst-case hardness assumptions such as $\mathbf{P} \neq \mathbf{NP}$. Therefore, instead of studying this average-case question in the Turing model under standard hardness assumptions, we study these questions for limited models of computation but prove *unconditional* lower bounds. This approach has turned out to be quite fruitful and several results of this form have emerged over the past few decades. For Boolean circuits, Rossman [9], [10] proved two remarkable results: he showed that monotone circuits, i.e., circuits consisting of $\vee$ and $\wedge$ gates only, as well as circuits of constant depth require size $\Omega(n^{k/4})$ to refute the existence of a k-clique in the average-case setting.

Instead of studying circuits, we approach this problem from the lens of proof complexity. Very broadly, proof complexity studies certificates of unsatisfiability of propositional formulas. As we cannot argue about certificates of unsatisfiability in general we consider certificates of a certain form, or in terms of proof complexity, refutations in a given proof system. For instance, if we prove that any certificate in a proof system P that witnesses that a given n-vertex graph contains no k-clique requires length $n^{\Omega(k)}$ on average, then we immediately obtain average-case $n^{\Omega(k)}$ running time lower bounds for any algorithm whose trace can be interpreted as a proof in the system P. It is often the case that state-of-the-art algorithms can be captured by seemingly simple proof systems, as was shown to be the case for clique algorithms [11].

It is often the case that weak proof systems are sensitive to the precise encoding of principles. The k-clique formula is no exception: it is somewhat straightforward to prove almost optimal $n^{\Omega(k)}$ resolution size lower bounds for the less usual binary encoding of the k-clique formula [12] and these lower bounds can even be extended to an $n^{\Omega(k)}$ lower bound for the Res(s) proof system for constant s [13]. For the more natural unary encoding, not much is known. There are essentially optimal $n^{\Omega(k)}$ average-case size lower bounds for regular resolution [11], [14] and tree-like resolution [15], [16]. For resolution, there are two average-case lower bounds that hold in different regimes: for $n^{5/6} \ll k \leq n/3$, Beame et al. [17] proved an average-case $\exp(n^{\Omega(1)})$ size lower bound and for $k \leq n^{1/3}$, Pang [14] proved a $2^{k^{1-o(1)}}$ lower bound. It is a long standing open problem, mentioned, e.g., in [15], to prove an unconditional $n^{\Omega(k)}$ resolution size lower bound for the

Supported by the Approximability and Proof Complexity project funded by the Knut and Alice Wallenberg Foundation. Part of this work was carried out while all authors were associated with KTH Royal Institute of Technology. Other parts were carried out while taking part in the semester program *Lower Bounds in Computational Complexity* in the fall of 2018 and the semster programs *Meta-Complexity* and *Satisfiability: Extended Reunion* in the spring of 2023 at the Simons Institute for the Theory of Computing at UC Berkeley.

Susanna F. de Rezende received funding from ELLIIT, from Knut and Alice Wallenberg grants KAW 2018.0371 and KAW 2021.0307, and from the Swedish Research Council grant 2021-05104. Aaron Potechin is in part supported by NSF grant CCF-2008920. Kilian Risse is supported by Swiss National Science Foundation project 200021-184656 "Randomness in Problem Instances and Randomized Algorithms".

unary encoding—even in the worst case. If we wish to extend these results to stronger proof systems that can reason about different formulations of the same problem, our lower bound techniques should also be oblivious to the precise encoding of problems. As we explain later on, this is one of the strengths of our proof strategy.

Little is known about the average-case hardness of the k-clique formula in the semi-algebraic setting. There are optimal degree lower bounds for $k \leq n^{1/2-\varepsilon}$ for the Sum-of-Squares proof system [18]–[20], but there are no non-trivial lower bounds on size. For Nullstellensatz, however, if restricted to not use dual variables, then size lower bounds follow by a simple syntactic argument [21]. Prior to our work no other size lower bounds were known for algebraic or semi-algebraic proof systems.

A. Our Result

In this work, we obtain the first size-lower bound on the clique formula for a semi-algebraic proof system. We show that the unary Sherali-Adams proof system, which is incomparable to resolution [22], requires size $n^{\Omega(D)}$ to refute the k-clique formula on random graphs whose maximum clique is of size $D \leq 2\log n$. Our result even applies in the approximate setting, where the formula states that the graph contains a clique of size $n^{1/100}$ but with high probability there is no clique of size D.

Theorem I.1 (Informal)**.** *For all integers $n \in \mathbb{N}^+$ and $D \leq 2\log n$, if $G \sim \mathcal{G}(n, n^{-2/D})$ is an Erdős-Rényi random graph, then it holds asymptotically almost surely that unary Sherali-Adams requires size at least $n^{\Omega(D)}$ to refute the claim that G contains a clique of size k, for any $k \leq n^{1/67}$.*

We note that our result also holds for the SubCubeSum proof system [23]. In fact our proof strategy gives a lower bound on the sum of coefficients of a Sherali-Adams refutation, ignoring Boolean axioms.

Let us stress that the size lower bound holds regardless of the degree of the refutation. This is a somewhat unique feature of our technique—all other lower bound strategies for Sherali-Adams and Sum-of-Squares are tailored to proving degree lower bounds, which, if strong enough, imply size lower bounds by the size-degree relation [24]. Since the clique formula has refutations of degree D we cannot expect to obtain size lower bounds through this connection for $D \leq \sqrt{n}$. We therefore introduce a new technique, inspired by pseudo-calibration [19], that is more refined—for any monomial m, of arbitrary degree, we determine a lower bound on the size of the smallest unary Sherali-Adams proof of m.

B. Organization

The rest of this paper is organized as follows. In Section II we introduce some basic terminology to then outline our proof strategy in Section III where we also attempt to convey some intuition. With the motivation at hand from Section III we then go on to define the central combinatorial concept of a *core* of a graph in Section IV and a notion of pseudorandomness in Section V. We proceed in Section VI to prove the main theorem for any graph satisfying our notion of pseudorandomness, albeit omitting the proof of one of the main lemmas. Finally, in Section VII we conclude with some open problems. We refer to the full-length version of this paper for the missing proofs, including that Erdős-Rényi random graphs satisfy the pseudorandom properties we define.

II. Preliminaries

Natural logarithms (base e) are denoted by $\ln$, whereas base 2 logarithms are denoted by $\log$. For integers $n \geq 1$ we introduce the shorthand $[n] = \{1, 2, \ldots, n\}$ and sometimes identify singletons $\{u\}$ with the element u. Let $\binom{S}{\ell}$ denote the set of subsets of S of size ℓ and, for a given a random variable X and an event P, we denote by $\mathbb{1}_P(X)$ the indicator random variable that is 1 if P holds and 0 otherwise.

Instead of working with an Erdős-Rényi random graph, we work in the block model [11], [17] as defined below. From [11], [17] we know that a lower bound on the block model implies a lower bound for Erdős-Rényi random graphs. Before introducing the block model, we need to set up some terminology and notation.

For the remainder of this paper G always denotes a k-partite graph with partitions $V_1, \ldots, V_k$ of size n each. We call a partition V_i a *block* and, for $S \subseteq [k]$, denote by V_S the vertices in blocks in S, that is, $V_S = \bigcup_{i \in S} V_i$. For disjoint sets $W_1, \ldots, W_s$ we let a *tuple* $t = (w_1, \ldots, w_s)$ be a sequence of vertices satisfying $w_i \in W_i$ for all $i \in [s]$. All tuples we consider are defined with respect to the partition $V_1, \ldots, V_k$, though, at times, may only be defined over a subset of the blocks, that is, not all tuples are of size k. For a tuple $t = (v_1, \ldots, v_k)$ and a set $S \subseteq [k]$ we denote the projection of t onto S by $t_S = (v_i \mid i \in S)$. An *s-tuple* is a tuple of size s and sometimes it is convenient for us to think of a tuple as a set of vertices. We take the liberty to interchangeably identify a tuple as a sequence as well as a set and hope that this causes no confusion.

A set Q of tuples is a *rectangle* if it can be written as the Cartesian product of sets $U_i \subseteq V_i$ (possibly empty) for $i \in [k]$. In other words, $Q = \bigtimes_{i \in [k]} U_i$ or, equivalently, there is a set $S \subseteq [k]$ such that Q contains all tuples $t = (u_1, \ldots, u_s)$ satisfying $u_i \in U_i$ for $i \in S$. Rectangles, unless explicitly stated, consist of k-tuples only, that is, if $Q = \bigtimes_{i \in [k]} U_i$, then we usually assume that all U_i are non-empty. Given a rectangle Q and a set $S \subseteq [k]$ we let Q_S be the projection of Q onto the blocks in S: if $Q = \bigtimes_{i \in [k]} U_i$, then $Q_S = \bigtimes_{i \in S} U_i$ and, in particular, we have $Q_i = U_i$ for $i \in [k]$.

While G always denotes a large graph, the graphs H and F denote small graphs: throughout the paper H and F are graphs on k labeled vertices. Usually these graphs have a small vertex cover and graphs denoted by F furthermore have many isolated vertices. For a graph H we denote the minimum vertex cover by $\mathrm{vc}(H)$ and sometimes refer to H as a *pattern graph*, whereas F is usually a *core graph* (see Section IV). We denote by $\mathcal{H}$ the set of graphs on k labeled vertices and for a parameter $i \in \mathbb{N}^+$ let $\mathcal{H}_i \subseteq \mathcal{H}$ be the family of graphs with

a minimum vertex cover of size at most i, that is, all graphs $H \in \mathcal{H}_i$ satisfy $\mathrm{vc}(H) \leq i$.

Given k blocks $V_1, \ldots, V_k$ of size n and a real number $0 \leq p \leq 1$ we denote by $\mathcal{G}(n,k,p)$ the distribution over graphs on the vertex set $V_{[k]}$ defined by sampling each edge $e = \{u,v\}$ independently with probability p if u and v are in distinct blocks. Edges within the same block are never included and hence $\mathcal{G}(n,k,p)$ is a distribution over k-partite graphs.

A. Clique Formula

Below we present a polynomial encoding of the k-clique formula. As our lower bound strategy is quite agnostic to the precise encoding we could equally well define the formula as a translation of a CNF. For the sake of exposition we choose to work with the following encoding.

Given a k-partite graph G with blocks $V_1, \ldots, V_k$ of size n we define the *k-clique formula over G* as follows. The formula is defined over $2kn$ variables: each vertex $v \in V_{[k]}$ is associated with two variables x_v and $\bar{x}_v$. All variables are Boolean and thus for each variable y (where y is either x_v or $\bar{x}_v$ for some $v \in V_{[k]}$) we introduce the Boolean axiom $y(1-y)$. Through the negation axioms $1 - \bar{x}_v - x_v$ we ensure that the variables associated with a single vertex v are the negation of each other. For each block V_i we introduce the *block axiom* $\sum_{v \in V_i} x_v - 1$ stating that precisely one vertex from each block is chosen and for each pair of vertices $\{u,v\} \notin G$ in distinct blocks we introduce the *edge axiom* $x_u x_v$ that ensures that non-neighbors are not simultaneously selected. Let us remark that we could add edge axioms for pairs of vertices in the same block but for ease of exposition we choose not to include them.

It should be evident that this formula is satisfiable if and only if there is a k-tuple t such that the vertex induced subgraph $G[t]$ is a clique.

B. Unary Sherali-Adams

Let $\mathcal{P} = \{p_1 = 0, \ldots, p_m = 0\}$ be a polynomial system of equations over Boolean variables $x_1, \ldots, x_n$ and their twin variables $\bar{x}_1, \ldots, \bar{x}_n$. If we assume that $\mathcal{P}$ contains all the necessary Boolean axioms as well as the negation axioms, then a Sherali-Adams refutation of $\mathcal{P}$ is a sequence of polynomials $(g_1, \ldots, g_m, f_0)$ such that f_0 is of the form

$$f_0 = \sum_{\substack{A,B \subseteq [n] \\ \alpha_{A,B} \geq 0}} \alpha_{A,B} \prod_{i \in A} x_i \prod_{i \in B} \bar{x}_i \tag{1}$$

and it holds that

$$\sum_{j \in [m]} g_j p_j + f_0 = -1 \ . \tag{2}$$

The *size* of a refutation is the number of monomials on the left hand side of Equation (2) when all polynomials are expanded out (without any cancellations) as a sum of monomials. The *coefficient size* of a Sherali-Adams refutation is the sum of the magnitudes of the coefficients of all monomials occurring in the proof (again, without any cancellations).

Unary Sherali-Adams is a subsystem of Sherali-Adams where all coefficients of monomials are either $+1$ or -1 and the right-hand-side of Equation (2) is any negative integer

$$\sum_{j \in [m]} g_j p_j + f_0 = -M \ , \tag{3}$$

where f_0 is again a non-negative sum of monomials (sometimes also called a conical junta).

Proposition II.1. *If Sherali-Adams requires coefficient size s to refute $\mathcal{P}$, then unary Sherali-Adams requires size at least s to refute $\mathcal{P}$.*

Proof. We can transform any unary Sherali-Adams refutation of size s, summing to an integer $-M$, to a Sherali-Adams refutation of coefficient size at most s by dividing the left hand side by $M \geq 1$. □

We may define the clique formula, as introduced in Section II-A, over any graph $G = (V,E)$ on kn vertices by partitioning the vertices into k sets $V = V_1 \dot\cup \cdots \dot\cup V_k$ of equal size and defining the clique formula with respect to that partition. It may be more natural to define the k-clique formula for such G with $V_1 = \cdots = V_k$. The following proposition, essentially due to Beame et al. [17], states that the Sherali-Adams coefficient size required to refute the latter is lower bounded by the coefficient size required to refute the clique formula defined with respect to a k-partition.

Proposition II.2 ([17]). *For $k, n \in \mathbb{N}^+$ and any graph G on kn vertices, the minimum Sherali-Adams coefficient size to refute the k-clique formula over G is bounded from below by the coefficient size required to refute the k-clique formula defined with respect to any equal sized k-partition of G.*

This proposition was proven in [17] for resolution size via a restriction argument, and it is straightforward to see that the same proof holds for Sherali-Adams coefficient size.

C. Some Auxiliary Lemmas

Lemma II.3. *There are at most $2^{c \log k + b(c-(b+1)/2)} \leq 2^{c(b+\log k)}$ graphs H over k vertices with a vertex cover of size b and $\left|V\big(E(H)\big)\right| \leq c$.*

Proof. We first choose the b vertices from the k vertices that form the vertex cover. Then, from the remaining $k-b$ vertices, we choose $c-b$ vertices that may be incident to an edge. We can add edges that are incident to the vertex cover and the other $c-b$ vertices and thus get that there are at most

$$\binom{k}{b}\binom{k-b}{c-b} 2^{\binom{b}{2}} 2^{b(c-b)} \leq 2^{c \log k + b(c-(b+1)/2)} \tag{4}$$

many such graphs. □

Recall that a maximal matching of H is a matching that cannot be extended in H.

Proposition II.4. *Any maximal matching in a graph H is of size at least $\lceil \mathrm{vc}(H)/2 \rceil$.*

Proof. Since M is maximal, all edges of H are incident to $V(M)$. Thus the set $V(M)$ is a vertex cover of H. □

III. Main Theorem and Proof Overview

The main result in this paper is a tight, up to constants in the exponent, size lower bound for unary Sherali-Adams for k-clique formulas over Erdős-Rényi random graphs, which we state formally next.

Theorem III.1. *For all integers $n \in \mathbb{N}^+$, $D \leq 2\log n$ and $k \leq n^{1/66}$, if $G \sim \mathcal{G}(n, k, n^{-2/D})$, then it holds asymptotically almost surely that unary Sherali-Adams requires size at least $n^{\Omega(D)}$ to refute the k-clique formula over G.*

Note that Theorem I.1 follows directly from Theorem III.1 along with Proposition II.2.

In the rest of this section we outline our proof strategy. We intend to come up with a so-called *pseudo-measure* which lower bounds the size of a unary Sherali-Adams refutation. In fact it proves something slightly stronger: the existence of a pseudo-measure implies a lower bound on the sum of the magnitude of the coefficients of a (general) Sherali-Adams refutation. Before we get ahead of ourselves let us define what a pseudo-measure is. A similar notion has previously appeared in [25] for the Nullstellensatz proof system over the reals.

Definition III.2 (pseudo-measure)**.** Let $\delta > 0$ and $\mathcal{P}$ be a set of polynomials over the polynomial ring $\mathbb{R}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$. A linear function $\mu \colon \mathbb{R}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n] \to \mathbb{R}$, mapping polynomials to reals, is a *δ-pseudo-measure for $\mathcal{P}$* if for all monomials m and all polynomials $p \in \mathcal{P}$ it holds that

1) $|\mu(m \cdot p)| \leq \delta$, and
2) $\mu(m) \geq -\delta$.

We have the following simple proposition.

Proposition III.3. *If μ is a δ-pseudo-measure for $\mathcal{P}$, then any Sherali-Adams refutation of $\mathcal{P}$ requires coefficient size at least $\mu(1)/\delta$. In particular, this implies that unary Sherali-Adams requires size at least $\mu(1)/\delta$ to refute $\mathcal{P}$.*

Proof. Suppose we have a δ-pseudo measure μ for $\mathcal{P}$ and a Sherali-Adams refutation

$$\sum_{p \in \mathcal{P}} g_p \cdot p + f_0 = -1 \tag{5}$$

of $\mathcal{P}$. Apply μ to the refutation. Observe that the left-hand side has to sum to $-\mu(1)$. For any $p \in \mathcal{P}$ and any monomial m occurring in g_p with coefficient c_m, it holds that $|\mu(m \cdot p)| \leq c_m \cdot \delta$ and similarly for a monomial $m \in f_0$ occurring with a positive coefficient c_m it holds that $\mu(m) \geq -c_m \cdot \delta$. Thus Sherali-Adams requires coefficient size at least $\mu(1)/\delta$, as claimed. The size lower bound for unary Sherali-Adams follows by virtue of Proposition II.1. □

A. Our Pseudo-Measure

In what follows we define our pseudo-measure μ for the k-clique formula. We may think of μ as a progress measure: it assigns to each monomial a real value which can be thought of as the contribution of this monomial towards the refutation of the k-clique formula. Thus, intuitively, we would like to associate each monomial with the fraction of potentially satisfying assignments that it rules out. In order to define this a bit more formally, let us introduce the set of potentially satisfying assignments.

We say that an assignment α is *potentially satisfying* for the k-clique formula if there is a graph G such that the k-clique formula defined over G is satisfied by α. This set of assignments can be easily characterized: if we associate each k-tuple t with the assignment ρ_t that sets all variables x_u to 1 if $u \in t$ and to 0 otherwise, then the set of potentially satisfying assignments of the k-clique formula is

$$\{\rho_t \mid t \in V_1 \times V_2 \times \cdots \times V_k\} \ . \tag{6}$$

We say that a monomial m *rules out* an assignment ρ if $\rho(m) = 1$. As there is a one-to-one correspondence between potentially satisfying assignments and tuples, it is convenient to think of the tuples that a monomial rules out. We thus associate each monomial m with the set

$$Q(m) = \{t \mid \rho_t(m) = 1\} \tag{7}$$

of ruled out k-tuples. Note that $Q(1)$ is the set of all tuples, that is, $Q(1) = V_1 \times V_2 \times \cdots \times V_k$ and the set $Q(x_u x_v)$ associated with an edge axiom $x_u x_v$ consists of all k-tuples that contain the vertices u and v.

More generally, it is not too hard to see that the set of ruled out tuples of a monomial is a rectangle and that for each rectangle Q there is at least one monomial m such that Q is the set of tuples ruled out by m. We thus often discuss rectangles and it is implicitly understood that if a statement holds for all rectangles, then it also holds for all monomials. Finally, observe that if a monomial m satisfies $m = m_1 \cdot m_2$, then $Q(m) \subseteq Q(m_1)$.

For intuition we will now discuss two naïve, and fatally flawed, attempts to define a pseudo-measure. For our first attempt, we simply associate each monomial with the fraction of ruled out tuples, that is we map a monomial m to

$$\frac{|Q(m)|}{|Q(1)|} = n^{-k} \cdot |Q(m)| \ . \tag{8}$$

This measure is clearly non-negative and hence satisfies Property 2 of Definition III.2 for any $\delta > 0$. Furthermore, again for any $\delta > 0$, it satisfies Property 1 of Definition III.2 for the Boolean axioms, the negation axioms as well as the block axioms. Only the edge axioms cause trouble: the rectangle $Q(x_u x_v)$ associated with the edge axiom $x_u x_v$ is a n^{-2} fraction of all tuples. As such, this pseudo-measure may only gives us an n^2 lower bound—not quite what we are after.

We may remedy this by not associating a monomial m with *all* tuples in $Q(m)$ but rather only with a subset of $Q(m)$ that depends on the graph G. One very naïve attempt would be to associate m with the number of k-cliques that it rules out, that is, we may associate a monomial m with the normalized measure

$$n^{-k} \sum_{t \in Q(m)} 2^{\binom{k}{2}} \mathbb{1}_{\{t \text{ is a clique}\}}(G) \ . \tag{9}$$

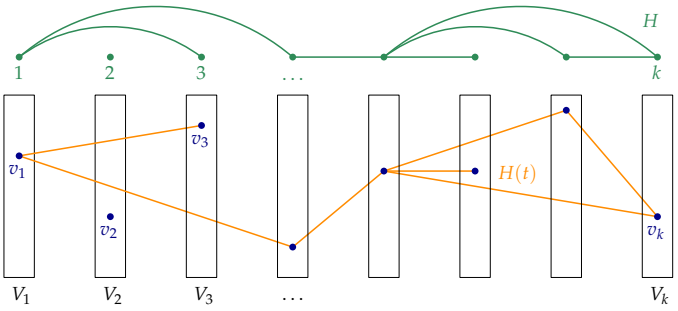

Figure 1. A pattern graph H mapped onto a tuple $t = (v_1, \dots, v_k)$

This definition, at least *in expectation* over $G \sim \mathcal{G}(n, k, 1/2)$, satisfies all properties of a pseudo-measure: the monomial 1 is mapped to 1, the axioms are all mapped to 0 and the measure is non-negative.

The obvious problem is that all graphs we consider do *not* contain a k-clique and hence everything (including the monomial 1) is mapped to 0. Following the lead of Barak et al. [19] we expand Equation (9) in the Fourier basis and truncate the resulting expression. By a careful choice of our truncation we can with some significant effort prove that this measure satisfies all required properties of Definition III.2. In order to state the precise definition of μ we need some notation.

For each potential edge e we have a character $\chi_e(G)$ defined by

$$\chi_e(G) = \begin{cases} \frac{1-p}{p} & \text{if } e \in E(G) \\ -1 & \text{otherwise;} \end{cases} \tag{10}$$

and for a set of potential edges E we let $\chi_E(G) = \prod_{e \in E} \chi_e(G)$. Note that for $p = 1/2$ this is the usual ± 1 Fourier basis. First time readers are advised to keep this case in mind for the remainder of the article.

To concisely state our pseudo-measure we need some further notation. We consider sums of tuples and want to treat edge sets that are equal up to the mapping onto a k-tuple as the same. More precisely, if we have two k-tuples $t = (v_1, \dots, v_k), t' = (v'_1, \dots, v'_k)$ and edge sets $E \subseteq \binom{t}{2}$ and $E' \subseteq \binom{t'}{2}$ such that $\{v_i, v_j\} \in E$ if and only if $\{v'_i, v'_j\} \in E'$, then we want to identify E and E' as the same edge set. To this end we consider *pattern graphs* H (similar to the *shape graphs* in the terminology of [19]) over the vertex set $[k]$. For a tuple $t = (v_1, \dots, v_k)$ and a graph H over $[k]$ we let $H(t)$ be the edge set that contains the edge $\{v_i, v_j\}$ if and only if the edge $\{i, j\}$ is present in H. See Figure 1 for an illustration. With this notation at hand we define our pseudo-measure as

$$\mu(m) = \mu_d\big(Q(m)\big) = n^{-k} \sum_{t \in Q(m)} \sum_{\substack{H \\ \mathrm{vc}(H) \le d}} \chi_{H(t)}(G) \ , \tag{11}$$

where the second sum is over all graphs H over $[k]$ vertices with vertex cover at most d, and $d = \eta D$ is a small constant $\eta > 0$ times the maximum clique size of G. It is convenient for us to work with the above (non-standard) basis as it allows us to easily cancel characters in case an edge is missing.

Observe that Boolean axioms, the negation axioms and the block axioms multiplied by an arbitrary monomial are all mapped to 0 by μ. Hence it remains to prove that the measure μ maps the constant 1 monomial to a large value, that μ is small on subrectangles of edge axioms, i.e., any edge axiom multiplied by a monomial is mapped to a small value, and that all monomials are mapped to an approximately non-negative value.

By inspecting the second moment of $\mu(1)$ it is not too hard to see that there is quite a bit of freedom on how to choose the truncation in the definition of μ while maintaining the property that $\mu(1) = 1 \pm n^{-\Omega(1)}$ asymptotically almost surely. However, ensuring that the edge axioms are associated with small measure is more delicate. Here we heavily rely on our choice to truncate according to the minimum vertex cover. More specifically we rely on two crucial properties of graphs H satisfying $\mathrm{vc}(H) = d$: firstly, we use the fact that such graphs contain a matching of size $\lceil d/2 \rceil$ (see Proposition II.4) and, secondly, that the family of these graphs satisfies a monotonicity property which leads to a useful partition of this family. For more details about this partition we refer to Section IV. Let us mention that it is conceivable that one could increase the bound on k for which our results hold by truncating according to the size of the maximum matching. As we do not know how to define the above mentioned partition with respect to maximum matching we truncate according to the minimum vertex cover.

In the following sections we try to present some intuition as to why μ_d is a pseudo-measure, that is, why it satisfies Definition III.2. In Section III-B we verify that G sampled from $\mathcal{G}(n, k, 1/2)$ asymptotically almost surely satisfies $\mu(1) = \mu_d(\bigtimes_{i \in [k]} V_i) = 1 \pm n^{-\Omega(1)}$. As mentioned, this follows by a straightforward second moment argument. In Section III-C we outline why any subrectangle Q of an edge axiom satisfies $|\mu_d(Q)| \le n^{-\Omega(d)}$. This proof motivates the definitions in Sections IV and V. Finally, in Section III-D, we provide some high-level overview of how to prove that any rectangle Q is mapped to an approximately non-negative value, that is, it holds that $\mu_d(Q) \ge -n^{-\Omega(d)}$. This is the most technically challenging part of the paper.

B. Expected Behavior of Our Pseudo-Measure

The measure $\mu_d(Q)$ of any rectangle Q satisfies

$$\mathbb{E}_G[\mu_d(Q)] = n^{-k} \sum_{t \in Q} \sum_{H \in \mathcal{H}_d} \mathbb{E}_G[\chi_{H(t)}(G)] \tag{12}$$

$$= n^{-k} \sum_{t \in Q} \mathbb{E}_G[\chi_{\emptyset(t)}(G)] \tag{13}$$

$$= n^{-k} |Q| \ . \tag{14}$$

In particular, as $Q(1) = \bigtimes_{i \in [k]} V_i$, it holds that $\mathbb{E}_G[\mu(1)] = 1$. In what follows we show that, for $p = 1/2$, the measure is somewhat concentrated around the expected value. The concentration, though, is far from enough to perform a union bound over all rectangles to argue that the measure behaves as expected on all rectangles simultaneously.

We show that the measure concentrates by an application of Chebyshev's inequality. To this end we analyze the second moment: for $p = 1/2$ we have

$$\mathbb{E}_G[\mu_d^2(Q)] = n^{-2k} \sum_{H \in \mathcal{H}_d} \sum_{t,t' \in Q} \mathbb{E}_G[\chi_{H(t)}(G)\chi_{H(t')}(G)] \quad (15)$$

$$= n^{-2k} \sum_{H \in \mathcal{H}_d} \sum_{\substack{t,t' \in Q: \\ t_{V(E(H))} = t'_{V(E(H))}}} \mathbb{E}_G[\chi_{H(t)}(G)\chi_{H(t')}(G)] \quad (16)$$

$$= n^{-2k} \sum_{H \in \mathcal{H}_d} |\{(t,t') \colon t_{V(E(H))} = t'_{V(E(H))}\}| \quad (17)$$

$$= n^{-2k} \sum_{H \in \mathcal{H}_d} |Q_{V(E(H))}| \cdot |Q_{[k] \setminus V(E(H))}|^2 \quad (18)$$

$$\leq n^{-k}|Q|\Big(1 + \sum_{\substack{H \in \mathcal{H}_d \\ H \neq \emptyset}} n^{-|V(E(H))|}\Big) \ . \quad (19)$$

A careful application of Lemma II.3 allows us to bound the number of pattern graphs H we sum over in (19) to conclude that $\mathbb{E}[\mu_d^2(Q)] = |Q|n^{-k}\big(1 \pm n^{-\Omega(1)}\big)$, as long as k and d are small. By virtue of Chebyshev's inequality we then conclude that $\mu(1) = 1 \pm n^{-\Omega(1)}$ asymptotically almost surely.

A natural attempt to prove that the measure is mostly non-negative is to analyze higher moments in the hope that these are closely concentrated around the (positive) expected value. The fundamental difficulty in analyzing the pseudo-measure μ_d is that we have to analyze exponentially many rectangles simultaneously. Since there is such a large number of rectangles, for each input graph G, there will be some rectangles where the value of μ_d differs considerably from the expected value.

For example, the measure on a rectangle Q with only a few vertices Q_i in some block V_i heavily depends on the behavior of the edges incident to the vertices in Q_i. Hence, if Q_i is small enough, we expect large deviations from the expected value. A slightly simplified, though more concrete, example of this phenomenon goes as follows: let $v_1 \in V_1$ and $v_2 \in V_2$, let Q be the rectangle that consists of all tuples that contain v_1 as well as v_2, and let H be the graph with the single edge $\{1,2\}$. In this setting the sum $\sum_{t \in Q} \chi_{H(t)}(G)$ heavily depends on whether the edge $\{v_1, v_2\}$ is present in G: if the edge is present, then the sum is equal to $n^{k-2}\frac{1-p}{p}$ and, if the edge is not present, then it is equal to $-n^{k-2}$. This indicates that on some rectangles the measure heavily depends on a few edges and we can thus not hope to naïvely prove concentration of the measure over all rectangles.

This slightly simplified example can be generalized to show that for a fixed H there is always a small number of rectangles where the value contributed by H is much larger than expected. Part of the technical challenge of the proof is to identify these bad rectangles and to handle them separately.

C. Edge Axioms Should Have Small Measure

We now explain the main ideas for bounding the magnitude of the measure of edge axioms. Recall that all other axioms are mapped to 0 by μ and we are thus just left to show that the value of the edge axioms is closely concentrated around 0.

For every pair of vertices $\{u,v\} \notin E(G)$ in distinct blocks we have an edge axiom $p_{uv} = x_u x_v$ stating that at least one of x_u and x_v are set to 0. Let Q be a subrectangle of $Q(p_{uv})$. Note that for every such rectangle Q there is a monomial m such that $Q = Q(m \cdot p_{uv})$ and hence these are the correct rectangles to consider if we want to prove Property 1 of Definition III.2. In other words, if we manage to show for all such Q that $|\mu_d(Q)| \leq n^{-\Omega(d)}$, then it follows that for all monomials m it holds that $|\mu_d(m \cdot p_{uv})| \leq n^{-\Omega(d)}$, as wanted.

We first show that for a fixed pair of vertices $\{u,v\} \notin E(G)$, with good probability, all such subrectangles Q have small absolute measure. By a union bound over all missing edges we then conclude that all subrectangles Q of an edge axiom satisfy $|\mu_d(Q)| \leq n^{-\Omega(d)}$. Let us fix an edge $\{u,v\} \notin E(G)$.

If Q is empty, then there is nothing to prove as $\mu_d(Q)$ is trivially 0. Hence we may assume that Q is non-empty, that is, Q has at least one vertex per block and hence each tuple in Q contains both u and v. Let $i \neq j \in [k]$ such that $u \in V_i$ and $v \in V_j$. For $e = \{i,j\}$ we may write

$$\mu_d(Q) = n^{-k} \sum_{t \in Q} \sum_{H \in \mathcal{H}_d} \chi_{H(t)}(G) \quad (20)$$

$$= n^{-k} \sum_{t \in Q} \Big(\sum_{\substack{H \in \mathcal{H}_d \\ e \notin H}} \chi_{H(t)}(G) + \sum_{\substack{H \in \mathcal{H}_d \\ e \in H}} \chi_{H(t)}(G) \Big) \quad (21)$$

$$= n^{-k} \sum_{t \in Q} \sum_{\substack{H \colon \mathrm{vc}(H) = d, \\ \mathrm{vc}(H \cup \{e\}) = d+1}} \chi_{H(t)}(G) \ , \quad (22)$$

where the last equality follows from the fact that every tuple $t \in Q$ contains u and v and thus, if $e \notin H$, then $\chi_{H(t)}(G) = -\chi_{H(t)}(G) \cdot \chi_{\{u,v\}}(G) = -\chi_{(H \cup \{e\})(t)}(G)$ as $\{u,v\} \notin E(G)$.

The naïve approach to bounding $|\mu_d(Q)|$ is to try to bound the magnitude of $\sum_{t \in Q} \chi_{H(t)}(G)$ for each H separately and to then multiply this bound by the number of graphs H we sum over. Recall from Lemma II.3 that there are about 2^{dk} graphs with a minimum vertex cover of size d.

Unfortunately the magnitude of $\sum_{t \in Q} \chi_{H(t)}(G)$ may simply be too large: it can be of magnitude $n^{-O(d)}|Q|$. In particular for large Q this bound is insufficient

$$|\mu_d(Q)| \leq n^{-k} \sum_{\substack{H \colon \mathrm{vc}(H) = d, \\ \mathrm{vc}(H \cup \{e\}) = d+1}} |Q| n^{-O(d)} \quad (23)$$

$$\leq 2^{dk} n^{-O(d)} = n^{\Omega(dk/\log n - d)} \ , \quad (24)$$

as k is much larger than both d and $\log n$. Instead of bounding each H separately, we bundle some graphs H together and then proceed to bound the magnitude of the sum over each such bundle. More precisely, we have families of graphs, indexed by graphs F with at most $3d$ non-isolated vertices, of the form

$$\mathcal{H}(F, E_F^\star) = \{H \mid E(H) = E(F) \cup E, \text{where } E \subseteq E_F^\star\} \ , \quad (25)$$

that partition the set of graphs H satisfying $\mathrm{vc}(H) = d$ and $\mathrm{vc}(H \cup \{e\}) = d+1$. Using these families we can bound the

magnitude of $\mu_d(Q)$ by

$$|\mu_d(Q)| = n^{-k}\Big|\sum_{t\in Q}\sum_{\substack{H:\,\mathrm{vc}(H)=d,\\ \mathrm{vc}(H\cup\{e\})=d+1}}\chi_{H(t)}(G)\Big| \tag{26}$$

$$\le n^{-k}\sum_{F}\Big|\sum_{t\in Q}\sum_{H\in\mathcal{H}(F,E_F^\star)}\chi_{H(t)}(G)\Big| \tag{27}$$

$$= n^{-k}\sum_{F}\Big|\sum_{t\in Q}\chi_{F(t)}(G)\sum_{E\subseteq E_F^\star}\chi_{E(t)}(G)\Big| \ . \tag{28}$$

Observe that the innermost sum is, up to normalization, the indicator function of whether the edge set $E_F^\star(t)$ is present in G. In fact the innermost sum, with the appropriate definition of $E_F^\star$, is simply a statement about the common neighborhood sizes of different subsets of t in G. We will need to argue that for random graphs, with high probability, all such sets behave as expected and the innermost sums are therefore bounded.

Furthermore, since each graph F has at most $3d$ with incident edges, there are fewer such graphs: according to Lemma II.3 at most $2^{3d(d+\log k)}$. Since $k \le n^{1/66}$ and $d \le 2\eta\log n$, for some small constant η, it holds that there are at most $2^{d(d+\log k)} \lesssim n^{d/50}$ many such graphs F. Thus, an upper bound of $n^{k-\Omega(d)}$ on the absolute value of two innermost sums in Equation (28) can now be used to obtain the claimed bound $|\mu_d(Q)| \le n^{-\Omega(d)}$. This completes the proof sketch for bounding the measure on edge axioms.

In Section IV we formally define these *core* graphs F and the families $\mathcal{H}(F, E_F^\star)$. In Section V we introduce the pseudo-randomness property of graphs we rely on in order to bound the two innermost sums in Equation (28). In Section VI-A we formally prove that the measure on subrectangles of axioms is bounded in absolute value. The verification that random graphs indeed satisfy our notion of pseudorandomness can be found in the full version.

D. Rectangles Should Be Approximately Non-Negative

To show that all rectangles Q have essentially non-negative measure, the main idea is to decompose Q into a collection $\mathcal{Q}$ of rectangles satisfying the following properties.

1) The collection $\mathcal{Q}$ is small, that is, $|\mathcal{Q}| \le n^{O(d)}$.
2) Each rectangle $Q \in \mathcal{Q}$ is either
   a) small: $|Q| \le n^{(1-\varepsilon)k}$ and hence $|\mu_d(Q)|$ is negligible,
   b) a subrectangle of an axiom; $|\mu_d(Q)|$ is bounded, or
   c) all common neighborhoods in Q are of expected size and therefore $\mu_d(Q) \approx |Q|/n^k > 0$.

In other words, $\mathcal{Q}$ contains some rectangles that have negligible measure and a collection of larger rectangles on which the measure behaves as expected. As the latter rectangles have strictly positive measure we may conclude that our pseudo-measure is essentially non-negative on all rectangles.

We bound the measure on small rectangles by summing the maximum possible magnitude of any character appearing in the definition of our pseudo measure.

Lemma III.4. *Any rectangle Q satisfies $|\mu_d(Q)| \le O\big(|Q|n^{-k}k^d p^{-dk}\big)$.*

Proof. We bound $\mu_d(Q)$ by counting the number of pattern graphs H we sum over multiplied by the maximum magnitude of each such character. We have that

$$|\mu_d(Q)| \le n^{-k}\sum_{i=0}^{d}\sum_{j=i}^{ik}\Big|\sum_{t\in Q}\sum_{\substack{H:\\ \mathrm{vc}(H)=i\\ |E(H)|=j}}\chi_{H(t)}(G)\Big| \tag{29}$$

$$\le |Q|n^{-k}\sum_{i=0}^{d}\binom{k}{i}\sum_{j=0}^{ik}\binom{ik}{j}\Big(\frac{1-p}{p}\Big)^j \tag{30}$$

$$= |Q|n^{-k}\sum_{i=0}^{d}\binom{k}{i}\frac{1}{p^{ik}} \le O(|Q|n^{-k}k^d p^{-dk}) \ , \tag{31}$$

as claimed. □

We implement the above proof outline in Section VI-B. Proving that our pseudo-measure concentrates around a positive value on rectangles as described in Item 2c is the most delicate part of our proof. In fact, above proof outline is somewhat inaccurate in that the value the pseudo-expectation concentrates around is not simply $|Q|/n^k$ but further depends on the number of small blocks in the rectangle Q. We refer to Definition VI.6 for the precise definition of these rectangles and to Lemma VI.7 for the claimed concentration inequality. We do not include the proof of Lemma VI.7 in this extended abstract—it can be found in the full-length version of this paper.

IV. Cores

In this section we introduce the notion of a core of a pattern graph, which will be used extensively throughout the rest of the paper. Our notion of a core seems to be loosely connected to the notion of a *vertex cover kernel* as used in parameterized complexity (see, e.g., the survey by Fellows et al. [26]).

A. Cores and Boundaries

Recall that when bounding the measure of subrectangles of axioms A_e, we were left with sums over graphs H such that $\mathrm{vc}(H) = d$ and $\mathrm{vc}(H\cup\{e\}) = d+1$ (see Equation (22)). Such graphs motivate the following definition of sets of graphs in the *boundary* of an edge.

Definition IV.1 (boundary)**.** Let $i \in \mathbb{N}$, H be a graph and $e \in \binom{V(H)}{2}$ be an edge. The graph H is in the (i,e)-*boundary*, denoted by $\mathcal{H}_i(e)$, if and only if $\mathrm{vc}(H) = i$ and $\mathrm{vc}(H\cup\{e\}) = i+1$. Furthermore, we say that H is in the e-*boundary* if and only if H is in an (i,e)-boundary for some $i \in \mathbb{N}$.

As mentioned in the proof sketch bounding the edge axioms, we cannot bound each H in the e-boundary separately (there are too many pattern graphs H) so we partition such graphs according to *cores* as explained below.

Definition IV.2 (core)**.** A vertex induced subgraph F of H is a *core* if any minimum vertex cover of F is also a vertex cover of H.

The notions of cores and (i,e)-boundaries interact nicely in the following sense.

Proposition IV.3. *A core of a graph H is in the (i,e)-boundary if and only if H is.*

Proof. Let F be a core of H. We first argue that if a core F of the graph H is in the (i,e)-boundary, then so is H. Indeed, by definition it holds that $\mathrm{vc}(F) = \mathrm{vc}(H) = i$. Moreover, F being in the (i,e)-boundary implies that the minimum vertex cover of $F \cup \{e\}$ has size $i+1$, and therefore the minimum vertex cover of $H \cup \{e\}$ must also be $i+1$ since F is a subgraph of H.

It remains to argue that if H is in the (i,e)-boundary, then so is the core F. By definition of core, $\mathrm{vc}(F) = \mathrm{vc}(H) = i$. Suppose, for the sake of contradiction, that F is not in the (i,e)-boundary and thus $\mathrm{vc}(F \cup \{e\}) = i$. Let W be a minimum-sized vertex cover of $F \cup \{e\}$. Since $|W| = i$, it holds that W is also a minimum-sized vertex cover of F and thus, by definition of core, W is also a vertex cover of H. But this contradicts the assumption that H is in the (i,e)-boundary since W also covers the edge e and hence is a vertex cover of size i of $H \cup \{e\}$. □

Recall that $\mathcal{H}$ is the set of graphs on k labeled vertices. We consider a map core from $\mathcal{H}$ to small cores that satisfies certain properties as described in the lemma below.

Lemma IV.4. *There is a map* core *that maps graphs $H \in \mathcal{H}$ to a core of H with the following properties. For every graph F in the image of* core *we have that $|V(E(F))| \leq 3 \cdot \mathrm{vc}(F)$ and that there exists an edge set $E_F^\star \subseteq V\big(E(F)\big) \times \big([k] \setminus V\big(E(F)\big)\big)$ such that $\mathrm{core}(H) = F$ if and only if $E(H) = E(F) \cup E$ for $E \subseteq E_F^\star$.*

From now on we only consider the cores given by the map core as in Lemma IV.4. With a slight abuse of nomenclature we say that $\mathrm{core}(H)$ is *the core* of H. Note that for a graph F in the image of core we have that $\mathrm{core}^{-1}(F) = \mathcal{H}(F, E_F^\star) = \{H \mid E(H) = E(F) \cup E, \text{ for } E \subseteq E_F^\star\}$, as introduced in Section III-C.

We refer to the full paper for the proof of Lemma IV.4. In the following we sketch the construction of a core for intuition without proving that it satisfies the stated properties. Given a graph H with lexicographic minimum vertex cover W we let U_1 be the lex first maximal set of vertices with a matching from U_1 to W that covers all vertices in U_1. Similarly we let U_2 be the lex first maximal set of vertices in $H \setminus U_1$ with a matching from U_2 to W covering all vertices in U_2 to define $\mathrm{core}(H) = H[D \cup U_1 \cup U_2]$. An illustration can be found in Figure 2.

V. Well-Behaved Graphs

In this section, we define the notion of *well-behaved* graphs, which is based on two combinatorial properties of graphs related to common neighborhoods of small tuples, and two analytic properties that bound certain character sums. In the following sections we prove that our measure satisfies the required conditions to obtain our unary Sherali-Adams lower bound for any well-behaved graph.

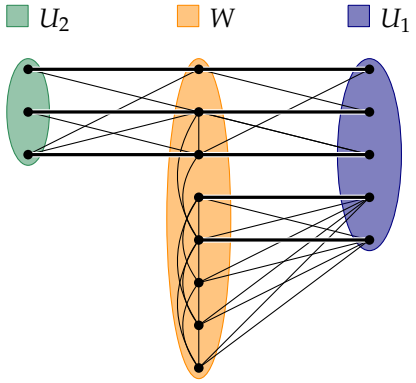

Figure 2. A candidate core with edges in M_1 and M_2 highlighted

Let us start by introducing the concepts needed to define well-behaved graphs. We say a rectangle Q is *s-small* if $|Q_i| \leq s$ for all $i \in [k]$ and, given a set $A \subseteq [k]$, a rectangle Q is said to be *(s,A)-large* if $|Q_i| > s$ for all $i \in A$. For any set $\mathcal{D}$ we say that a function $f : \mathcal{D} \to \mathbb{R}^+$ is *r-bounded* if $f(x) \leq r$ for all $x \in \mathcal{D}$.

We require some terminology and notation from graph theory. The neighborhood of a vertex $v \in V$ in a graph $G = (V,E)$ is $N(v) = N_G(v) = \{u \mid \{u,v\} \in E\}$ and the neighborhood of a set of vertices $U \subseteq V$ is $N(U) = N_G(U) = \{v \notin U \mid \exists u \in U : \{u,v\} \in E\}$. For a set $W \subseteq V$ the neighborhood of a vertex v in W is $N(v,W) = N(v) \cap W$ and similarly for a set U we let the neighborhood of U in W be $N(U,W) = N(U) \cap W$. The common neighborhood of U is $N^\cap(U) = \bigcap_{u \in U} N(u)$ and the common neighborhood of U in W is $N^\cap(U,W) = N^\cap(U) \cap W$. This notation is naturally extended to a tuple t by considering t as a set of vertices.

The next two definitions are purely combinatorial. They are similar to definitions that have appeared in previous papers on k-clique [11], [15], [17]. Recall that throughout the paper graphs denoted by G are k-partite with partitions $V_1, \ldots, V_k$ of size n each.

Definition V.1 (bounded common neighborhoods)**.** A graph G has *(β,p)-bounded common neighborhoods from $Q = \bigtimes_{i \in A} Q_i$ to $R \subseteq V(G)$* if it holds that for all $B \subseteq A$ and all $t \in Q_B$

$$|N^\cap(t,R)| \in (1 \pm \beta)p^{|t|}|R| \ .$$

A graph G has *(β,p,d)-bounded common neighborhoods in every block* if for all $A \subseteq [k]$ of size at most d and all $i \in [k] \setminus A$, G has (β,p)-bounded common neighborhoods from V_A to V_i.

While it turns out that random graphs do have bounded common neighborhoods, the graph induced by a rectangle may certainly have tuples with ill-behaved common neighborhoods: we may for example have an isolated vertex in a rectangle. The following definition roughly states that, while there may be tuples with ill-behaved neighborhoods in a rectangle, there is a large sub-rectangle which has bounded common neighborhoods.

Definition V.2 (bounded error sets)**.** A graph G has (s, w, β, p, d)-*bounded error sets* if for all rectangles $Q = \bigtimes_{i\in[k]} Q_i$ satisfying $|Q_i| \geq s$ or $|Q_i| = 0$ it holds that there exists a small set of vertices $W \subseteq V(G)$, $|W| \leq w$, such that for all $S \subseteq [k]$ of size at most d it holds that all tuples $t \in \bigtimes_{i\in S}(Q_i \setminus W)$ satisfy

$$\big|N^{\cap}(t, Q_j \setminus W)\big| \in (1 \pm \beta)p^{|t|}\big|Q_j \setminus W\big|$$

for all $j \in [k] \setminus S$. We refer to W as the *error set of* Q.

Recall from the edge axiom proof sketch in Section III-C that we require bounds of the form $n^{k-\Omega(\mathrm{vc}(F))}$ on the absolute value of certain character sums. It turns out that, in order to prove that monomials are mapped to an essentially non-negative value, we need tighter (depending on $|Q|$) as well as "localized" versions of these bounds. For conciseness we introduce the following terminology.

Definition V.3 (bounded character sums)**.** Let $s \in \mathbb{N}^+$, $B \subseteq [k]$, $Q_B = \bigtimes_{i\in B} Q_i$ and F be a core graph. A graph G has s-*bounded character sums over* Q_B *for* F if it holds that

$$\Big|\sum_{t\in Q_B}\sum_{H\in\mathcal{H}(F,E_F^\star[B])} \chi_{H[B](t)}(G)\Big| \leq s \ .$$

We are now ready to state the pseudorandomness property of graphs that allows us to prove average-case unary Sherali-Adams lower bounds for the k-clique formula. As Properties 3 and 4 are somewhat difficult to parse we give an informal description upfront.

Property 3 states that all character sums over the families $\mathcal{H}(F, E_F^\star)$ are of bounded magnitude if the rectangle considered has large minimum block size. Smaller rectangles are unfortunately not as well-behaved. However, for certain rectangles, we can guarantee something similar: Property 4 states that if the common neighborhood of small tuples in a rectangle are bounded, then the mentioned character sums can still be bounded.

First time readers may, for now, choose to skip the formal definition of Property 4. It might be more insightful to first read Section VI and return Property 4 once it is used.

Definition V.4 (well-behaved graph)**.** We say that a k-partite graph G with partitions of size n is D-*well-behaved* if, for $p = n^{-2/D}$, the following properties hold:

1) G has $(1/k, p, D/4)$-bounded common neighborhoods in every block.
2) There exists a constant $C \in \mathbb{R}^+$ such that G has $(2s, s, 1/k, p, \ell)$-bounded error sets for all $\ell \leq D/4$ and $s \geq Ck^4\ell \ln n/p^{2\ell}$.
3) For any core F with $\mathrm{vc}(F) \leq D/4$ and any $\big(n/2, V\big(E(F)\big)\big)$-large rectangle Q it holds that G has s-bounded character sums over $Q_{[n]}$ for F, where
$$s = 6 \cdot p^{-|E(F)|} \cdot n^{k-\lambda\,\mathrm{vc}(F)/4} \ ,$$
for any $\lambda < 1 - \log(k)/\log(n)$.
4) For any $\Lambda \geq 20\,k \log n$, any core F with $\mathrm{vc}(F) \leq D/4$, any $B \subseteq [k]$ and any (4Λ)-small rectangle Q that is also (Λ, B)-large the following holds for $A = V\big(E(F)\big) \cap B$. If G has $(3/k, p)$-bounded common neighborhoods from Q_A to Q_i, for every $i \in B \setminus A$, then G has s-bounded character sums over Q_B for F, where
$$s = O\big(p^{-|E(F[B])|} \cdot (\Lambda/10\,k\log n)^{-\,\mathrm{vc}(F[B])/4} \cdot |Q_B|\big) \ .$$

In what follows we often state that a graph G is D-well-behaved in which case it is implicitly understood that G is k-partite with partitions of size n. In the full-length version of this paper, we prove that a graph G, sampled from the distribution $\mathcal{G}(n, k, n^{-2/D})$, is asymptotically almost surely D-well-behaved, as stated next.

Theorem V.5. *If n is a large enough integer, $k \in \mathbb{N}^+$ and $D \in \mathbb{R}^+$ satisfy $4 \leq D \leq 2\log n$ and $k \leq n^{1/5}$, then $G \sim \mathcal{G}(n, k, n^{-2/D})$ is asymptotically almost surely D-well-behaved.*

VI. Clique Is Hard on Well-Behaved Graphs

In this section we prove that our measure μ_d is an $n^{-\Omega(D)}$-pseudo-measure for the k-clique formula, if the formula is defined over a D-well-behaved graph G.

Theorem VI.1. *There are constants $\eta, c > 0$ and $D_0 \in \mathbb{N}$ such that the following holds for large enough $n \in \mathbb{N}$ and all D satisfying $D_0 < D \leq 2\log n$. If $D \leq k \leq n^{1/66}$, $d = \eta D$ and G is a D-well-behaved k-partite graph with n vertices per block, then the measure μ_d is an n^{-cD}-pseudo-measure for the k-clique formula over G and, furthermore, satisfies $\mu_d(1) \geq 1 - n^{-\Omega(1)}$.*

From Theorem V.5 and Theorem VI.1 along with Proposition III.3 we obtain Theorem III.1.

In order to prove that the measure μ_d satisfies the properties of a pseudo-measure as listed in Definition III.2, we show that μ_d maps any axiom multiplied by a monomial to approximately 0 and that all monomials are associated with an essentially non-negative value. Finally, we argue that $\mu_d(1) \geq 1 - n^{-\Omega(1)}$.

For all axioms except the edge axioms, it is easy to see that μ_d maps the axiom times any monomial not only to approximately 0 but to precisely 0. Indeed, by definition of μ_d, for any monomial m and variable x, we have $\mu_d\big(m(1 - \bar{x} - x)\big) = 0$ and $\mu_d\big(m(x^2 - x)\big) = 0$. Similarly we may observe that $\mu_d\big(m(\sum_{v\in V_i} x_v - 1)\big) = 0$ by linearity of μ over tuples. With regards to bounding the measure on axioms we are left to prove that for any edge axioms $p_{uv} = x_u x_v$ and any monomial m it holds $|\mu_d(mp_{uv})| \leq n^{-cD}$. The following lemma states a slightly more precise bound which we use when proving Lemma VI.3.

Lemma VI.2. *Let G be a D-well-behaved graph, let n be a large enough integer and let $d = \eta D \leq 2\eta \log n$ for some constant $\eta > 0$. It holds that all edge axioms p_{uv} and all rectangles $Q \subseteq Q(p_{uv})$ satisfy $|\mu_d(Q)| \leq O\big(\big(n^{\lambda/4 - 12\eta}/(2k)^3\big)^{-d}\big)$ for any $\lambda < 1 - \log(k)/\log(n)$.*

Note that by choosing $\lambda = 1/2$, and considering $k \leq n^{1/66}$ and $\eta > 0$ small enough, Lemma VI.2 implies that any

subrectangle of an edge axiom satisfies $|\mu_d(Q)| \leq n^{-cD}$ for some small enough constant c. We postpone the proof of Lemma VI.2 to Section VI-A.

In addition to the bound on the magnitude of the measure on the axioms we also need to prove that the measure is essentially non-negative. We state this formally below and defer the proof to Section VI-B.

Lemma VI.3. *There are constants $\eta, c > 0$ such that if G is a D-well-behaved graph, n is large enough, $d = \eta D \leq 2\eta \log n$ and $D \leq k \leq n^{1/66}$, then any rectangle Q satisfies $\mu_d(Q) \geq -n^{-cD}$.*

In Section III-B we argued that, with high probability, $\mu_d(1)$ is approximately 1 if G is a random graph. In what follows we show that this holds for any D-well-behaved graph.

Lemma VI.4. *There are constants $\eta, c > 0$ such that for n large enough, $k \leq n^{1/20}$, $D \leq 2\log n$ and $d = \eta D$ it holds that if G is a D-well-behaved graph, then $\mu_d(1) \geq 1 - n^{-c}$.*

Proof. This is a direct consequence of the definition of a D-well behaved graph and Lemma IV.4. Recall the map core from Lemma IV.4 and the families

$$\mathcal{H}(F, E_F^\star) = \{H \mid E(H) = E(F) \cup E, \text{where } E \subseteq E_F^\star\}\ , \quad (32)$$

defined for core graphs $F \in \mathrm{img}(\mathrm{core})$ which satisfy $\mathrm{vc}(F) \leq d$. From Property 3 of Definition V.4, choosing $\lambda = 4/5$, it follows that for every $F \, \mathrm{img}(\mathrm{core})$ we have

$$n^{-k}\Big|\sum_{t\in Q(1)}\sum_{H\in\mathcal{H}(F,E_F^\star)}\chi_{H(t)}(G)\Big| \leq n^{-\mathrm{vc}(F)/6}\ , \quad (33)$$

where we use the bound $p^{-|E(F)|} \leq p^{-3d\,\mathrm{vc}(F)} \leq n^{6\eta\,\mathrm{vc}(F)}$ and the fact that η is a small enough constant. As the families defined in Equation (32) partition the set of graphs H of vertex cover at most d, using the bound from Equation (33), it holds that

$$\begin{aligned}
\mu_d(1) &= 1 + n^{-k}\sum_{\substack{H\in\mathcal{H}_d\\ H\neq\emptyset}}\sum_{t\in Q(1)}\chi_{H(t)}(G) && (34)\\
&\geq 1 - n^{-k}\sum_{i=1}^{d}\sum_{\substack{F\in\mathrm{img(core)}\\ \mathrm{vc}(F)=i}}\Big|\sum_{t\in Q(1)}\sum_{H\in\mathcal{H}(F,E_F^\star)}\chi_{H(t)}(G)\Big| && (35)\\
&\geq 1 - \sum_{i=1}^{d} 2^{3i(d+\log k)} n^{-i/6} && (36)\\
&\geq 1 - n^{-c}\ , && (37)
\end{aligned}$$

for some constant $c > 0$. In above inequalities we relied on Lemma II.3 for an upper bound on the number of cores in $\mathrm{img}(\mathrm{core})$ with vertex cover of size i, on the fact that $d \leq 2\eta \log n$, that η is a small enough constant and that $k \leq n^{1/20}$. This completes the proof of Lemma VI.4. □

This completes the proof of Theorem VI.1 modulo Lemma VI.2 and Lemma VI.3, which we prove in Section VI-A and Section VI-B, respectively.

A. Axioms Have Small Measure

In this section, we show that, with high probability, any subrectangle of an edge axiom has small measure in absolute value. We rely on the following technical lemma.

Lemma VI.5. *If G is a D-well-behaved graph, then for any core graph F and any rectangle Q, we have that*

$$\Big|\sum_{t\in Q}\sum_{H\in\mathcal{H}(F,E_F^\star)}\chi_{H(t)}(G)\Big| \leq 6\cdot 2^{|A|}p^{-|E(F)|}n^{k-\lambda\,\mathrm{vc}(F)/4}\ ,$$

where $A = V\big(E(F)\big)$ and $\lambda < 1 - \log(k)/\log(n)$.

Proof. Let F be a core graph, let $A = V\big(E(F)\big)$ and $s = 6\cdot p^{-|E(F)|}\cdot n^{k-\lambda\,\mathrm{vc}(F)/4}$. By Property 3 of Definition V.4 we have that if Q is $(n/2, A)$-large (i.e., if Q satisfies $|Q_i| \geq n/2$ for all $i \in A$), then $\big|\sum_{t\in Q}\sum_{H\in\mathcal{H}(F,E_F^\star)}\chi_{H(t)}(G)\big| \leq s$.

Given any rectangle Q (not necessarily $(n/2, A)$-large), let $T \subseteq A$ be the set of blocks of Q such that $|Q_i| < n/2$. By a simple inclusion-exclusion argument, we have that

$$Q = \sum_{S\subseteq T}(-1)^{|S|}\Big(\bigtimes_{i\in S}(V_i \setminus Q_i)\Big)\times\Big(\bigtimes_{i\in T\setminus S}V_i\Big)\times\Big(\bigtimes_{i\in[k]\setminus T}Q_i\Big)\ . \quad (38)$$

For $S \subseteq T$, denote by Q^S the rectangle $\big(\bigtimes_{i\in S}(V_i \setminus Q_i)\big) \times \big(\bigtimes_{i\in T\setminus S}V_i\big) \times \big(\bigtimes_{i\in[k]\setminus T}Q_i\big)$. Note that Q^S is $(n/2, A)$-large and therefore, by Property 3 of Definition V.4, G has s-bounded character sums over Q^S for F. This implies that

$$\begin{aligned}
&\Big|\sum_{t\in Q}\sum_{H\in\mathcal{H}(F,E_F^\star)}\chi_{H(t)}(G)\Big|\\
&\qquad\leq \sum_{S\subseteq T}\Big|\sum_{t\in Q^S}\sum_{H\in\mathcal{H}(F,E_F^\star)}\chi_{H(t)}(G)\Big| \leq 2^{|A|}\cdot s\ ,
\end{aligned} \quad (39)$$

as claimed. □

We are now ready to prove Lemma VI.2.

Proof of Lemma VI.2. Fix an edge $\{u, v\} \notin E(G)$, let $i, j \in [k]$ such that $u \in V_i$ and $v \in V_j$, consider the edge axiom $p_{uv} = x_u x_v$ and let $Q \subseteq Q(p_{uv})$ be an arbitrary subrectangle of this edge axiom. Recall from Section III-C that every tuple $t \in Q$ contains the vertices u and v and thus, for $e = \{i, j\}$, we may write

$$\begin{aligned}
\mu_d(Q) &= n^{-k}\sum_{t\in Q}\sum_{H\in\mathcal{H}_d}\chi_{H(t)}(G) && (40)\\
&= n^{-k}\sum_{t\in Q}\sum_{\mathcal{H}_d(e)}\chi_{H(t)}(G)\ , && (41)
\end{aligned}$$

where $\mathcal{H}_d(e)$, as defined in Section IV, denotes the set of graphs in the (d, e)-boundary. Let the map core be as guaranteed by Lemma IV.4. Recall that, according to Proposition IV.3, the graph $\mathrm{core}(H)$ is in $\mathcal{H}_d(e)$ if and only if H is. Hence the sets

$$\{\mathrm{core}^{-1}(F) = \mathcal{H}(F, E_F^\star) \mid F \in \mathcal{H}_d(e) \wedge F \in \mathrm{img}(\mathrm{core})\} \quad (42)$$

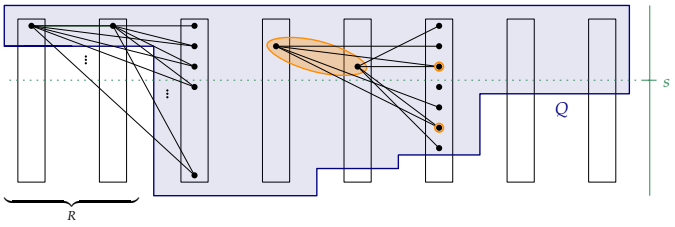

Figure 3. The rectangle Q is a good rectangles as the vertices in R have all vertices as neighbors, the blocks outside R are large and small tuples on these blocks have common neighborhoods of expected size

partition the (d,e)-boundary $\mathcal{H}_d(e)$ and we may thus write

$$|\mu_d(Q)| = \left|n^{-k} \sum_{t\in Q} \sum_{H\in\mathcal{H}_d(e)} \chi_{H(t)}(G)\right| \tag{43}$$

$$= \left|n^{-k} \sum_{\substack{F\in\mathcal{H}_d(e)\\ F\in\mathrm{img(core)}}} \sum_{t\in Q} \sum_{H\in\mathcal{H}(F,E_F^\star)} \chi_{H(t)}(G)\right| \tag{44}$$

$$\leq n^{-k} \sum_{\substack{F\in\mathcal{H}_d(e)\\ F\in\mathrm{img(core)}}} \left|\sum_{t\in Q} \sum_{H\in\mathcal{H}(F,E_F^\star)} \chi_{H(t)}(G)\right| . \tag{45}$$

By Lemma VI.5 each inner part can be bounded by $6\cdot 2^{|A|}\cdot p^{-|E(F)|}\cdot n^{k-\lambda\,\mathrm{vc}(F)/4}$. Note that $\mathrm{vc}(F)=d$ and, according to Lemma IV.4, it holds that $|A|\leq 3d$ and $p^{-|E(F)|}\leq p^{-3d^2} = n^{6\eta d}$, using the assumption that $d\leq 2\eta\log n$. Hence

$$|\mu_d(Q)| \leq \sum_{\substack{F\in\mathcal{H}_d(e)\\ F\in\mathrm{img(core)}}} 6\cdot 2^{3d}\cdot n^{-d(\lambda/4-6\eta)} \tag{46}$$

$$\leq 2^{3d(d+\log k)}\cdot 6\cdot 2^{3d}\cdot n^{-d(\lambda/4-6\eta)} \tag{47}$$

$$\leq 6\cdot\left(n^{\lambda/4-12\eta}/(2k)^3\right)^{-d} , \tag{48}$$

where we used Lemma II.3 to bound the number of core graphs and the assumption $d\leq 2\eta\log n$. This concludes the proof of Lemma VI.2. □

B. All Rectangles Are Approximately Non-Negative

Before defining good rectangles formally, let us give an informal description. A good rectangle Q consists of two parts. The first part is very small: on a few blocks the rectangle Q only consists of single vertices. Each vertex in this small part is adjacent to *all* other vertices in Q. Equivalently, on this small part we have a clique and the remaining vertices in Q are in the common neighborhood of this clique.

On the other blocks, where Q does not consist of a single vertex, we require that these blocks are large, of size at least $s=\mathrm{poly}(n)$. In addition we also require that all common neighborhoods are bounded on this large part. An illustration of a good rectangle can be found in Figure 3. The formal definition follows.

Definition VI.6 (good rectangle)**.** Let G be a k-partite graph and let $s,\beta,p,d\in\mathbb{R}^+$ and $R\subseteq[k]$. A rectangle $Q=\bigtimes_{i\in[k]} Q_i$ is (s,β,p,d,R)-*good for* G if it satisfies the following properties.

1) If $i\in R$, then $Q_i=\{v_i\}$; otherwise $|Q_i|\geq s$.
2) For all $i\in R$ it holds that $N(v_i)\supseteq\bigcup_{j\neq i} Q_j$.
3) For all $S\subseteq[k]\setminus R$ of size at most d and for all $i\notin R\cup S$, G has (β,p)-bounded common neighborhoods from Q_S to Q_i.

On good rectangles the measure is tightly concentrated around the expected value. In the full-length version of the present paper we prove the following concentration bound.

Lemma VI.7. *For constants $\varepsilon>0$ and $\eta<1/25$, for $n,k,d\in\mathbb{N}$ and $p=n^{-2/D}\leq 1/2$ satisfying $d\leq\eta D$ and $D\leq k\leq n$ the following holds. If $s\geq k^{13}n^{48\eta+\varepsilon}\log n$ and G is a D-well-behaved graph, then any $(s,1/k,p,d,R)$-good rectangle Q for G with $|R|=\ell<d$ satisfies*

$$\mu_d(Q) = p^{-\ell(k-(\ell+1)/2)}\,|Q|\,n^{-k}\big(1\pm O(n^{-\varepsilon/8})\big) .$$

In the remainder of this section we prove Lemma VI.3, assuming Lemma VI.7. As outlined in Section III-D, we intend to decompose any rectangle Q into a small family $\mathcal{Q}$ of rectangles such that each rectangle in $\mathcal{Q}$ either contains few tuples, is a subrectangle of an edge axiom or is a good rectangle. The following lemma summarizes our claim.

Lemma VI.8. *Let G be a D-well-behaved graph, let $p=n^{-2/D}$, $d\leq D/4$ and $s\geq Ck^4 d\ln n/p^{2d}$ for some large enough constant C. Then any rectangle Q_0 can be partitioned into a set of rectangles $\mathcal{Q}$ of size $|\mathcal{Q}|\leq 2kn(2s)^d$ such that each $Q\in\mathcal{Q}$ satisfies that either*

1) *Q is small: $|Q|\leq O\big((n\cdot p^d)^{k-d}\big)$,*
2) *Q is a subrectangle of an edge axiom, or*
3) *Q is $(s,1/k,p,d,R)$-good for G, where $R\subseteq[k]$ satisfies $|R|<d$.*

Before proving Lemma VI.8, let us show how Lemma VI.3 follows. The idea of the proof is to apply Lemma VI.8 to a given rectangle Q_0 to obtain a collection $\mathcal{Q}$ of rectangles. It holds that $\mu_d(Q_0)=\sum_{Q\in\mathcal{Q}}\mu_d(Q)$. By Lemma III.4 there is a $\delta>0$ such that all small rectangles $Q\in\mathcal{Q}$ satisfy $|\mu_d(Q)|\leq n^{-\delta D}$ and similarly by Lemma VI.2 the same holds for $Q\in\mathcal{Q}$ that are a subrectangle of an edge axiom. Further, by our choice of parameters, the size of $\mathcal{Q}$ is small—we may think of it as $n^{\delta D/2}$. We can thus lower bound

$$\mu_d(Q_0)=\sum_{Q\in\mathcal{Q}}\mu_d(Q)\geq -n^{-\delta D/2}+\sum_{\substack{Q\in\mathcal{Q}\\ Q\text{ is good}}}\mu_d(Q) . \tag{49}$$

Lemma VI.7 states that the remaining good rectangles in above sum have strictly positive value. Thus $\mu_d(Q_0)\geq -n^{-\delta D/2}$ as claimed. In what follows we verify that this indeed holds for our choice of parameters.

Proof of Lemma VI.3. Let Q_0 be any rectangle. Our goal is to show that $\mu_d(Q_0)\geq -n^{-cD}$, for a sufficiently small constant c. Let $D\leq k\leq n^{1/66}$ be as in the statement of the lemma and choose $\lambda=1-\varepsilon-\log(k)/\log(n)$ for sufficiently small constants $\varepsilon>0$ and $\eta>0$ such that for $s=k^{13}n^{48\eta+\varepsilon}\log n$ it holds that $s\leq n^{\lambda/4-12\eta-\varepsilon}/k^3$. Let $d=\eta D\leq 2\eta\log n$ and $p=n^{-2/D}$. Note that for our choice of parameters it holds

that $s = \omega(k^4 n^{4\eta} \log^2 n)$, hence $s = \omega(k^4 d \ln n / p^{2d})$, and we may thus apply Lemma VI.8 with $d = \eta D$ to the rectangle Q_0 to obtain a family $\mathcal{Q}$ of size at most $|\mathcal{Q}| \leq 2kn(2s)^d$.

By Lemma VI.2, any subrectangle of an axiom has measure bounded by $O\big(\big(n^{\lambda/4-12\eta}/(2k)^3\big)^{-d}\big)$. Moreover, according to Lemma III.4 each small rectangle $Q \in \mathcal{Q}$ has measure of magnitude at most

$$|\mu_d(Q)| \leq O\big(|Q| n^{-k} k^d p^{-dk}\big) = O(n^{-d/2}) \ , \qquad (50)$$

which is even smaller than the bound on axioms. Since $s \leq n^{\lambda/4-12\eta-\varepsilon}/k^3$, we conclude that the measure of all small rectangles and all subrectangles of axioms in $\mathcal{Q}$ add up, in magnitude, to at most $|\mathcal{Q}| \cdot O\big(\big(n^{\lambda/4-12\eta}/(2k)^3\big)^{-d}\big) \leq n^{-cD}$, for a small enough constant c.

Hence the measure of Q_0 is mostly on the good rectangles of $\mathcal{Q}$ and on these rectangles we know that it is closely concentrated around a strictly positive value. Indeed, we can apply Lemma VI.7 to any rectangle Q which is $(s, 1/k, p, d, R)$-good for G to conclude that

$$\mu_d(Q) = p^{-\ell(k-(\ell+1)/2)} |Q| \, n^{-k} \big(1 \pm O(n^{-\varepsilon/8})\big) > 0 \ . \quad \square$$

Let us proceed to prove Lemma VI.8.

Proof of Lemma VI.8. Let us describe a recursive decomposition procedure that can be applied to any rectangle $Q = \bigtimes_{i \in [k]} Q_i$.

If either Q is small, a subrectangle of an axiom or $(s, 1/k, p, d, R)$-good for some $R \subseteq [k]$, then return Q. Otherwise decompose in the following recursive fashion.

1) If there is a singleton $Q_i = \{v_i\}$ such that $N(v_i) \not\supseteq \bigcup_{j \neq i} Q_j$, then we decompose Q into $|Q \setminus N(v_i)| + 1$ many rectangles as follows. Denote by $u_1, u_2, \ldots, u_m$ the vertices in Q that are not a neighbor of v_i and assume that they are in blocks $j_1, j_2, \ldots, j_m$. For $\nu = 1, \ldots, m$ we remove all tuples that contain the vertex u_ν: let $R^0 = Q$ so we can write
$$\begin{aligned} Q^\nu &= \{u_\nu\} \times \bigtimes_{j \neq j_\nu} R_j^{\nu-1} \text{ and} \\ R^\nu &= \big(R_{j_\nu}^{\nu-1} \setminus u_\nu\big) \times \bigtimes_{j \neq j_\nu} R_j^{\nu-1} \ . \end{aligned} \qquad (51)$$
Note that the rectangles $Q^1, \ldots, Q^m, R^m$ partition Q. Add the Q^ν to the partition as these are subrectangles of edge axioms and recursively decompose R^m.
2) If there is a block $i \in [k]$ of size $1 < |Q_i| \leq 2s$, then split Q into the $|Q_i|$ rectangles
$$\big\{\{v_i\} \times \bigtimes_{j \neq i} Q_j : v_i \in Q_i\big\} \qquad (52)$$
and recursively decompose each of these rectangles.
3) Let A be the set of blocks of size greater than $2s$. Because G is D-well-behaved, by Property 2 of Definition V.4, it holds that G has $(2s, s, 1/k, p, d)$-bounded error sets. In particular Q_A has an error set $U = \{u_1, \ldots, u_m\}$ of size at most s. Decompose Q into $Q^1, \ldots, Q^m$ and R^m as in Case 1. By definition the rectangle R^m is $(s, 1/k, p, d, [k] \setminus A)$-good and we may thus add it to the partition. Recursively decompose the rectangles $Q^1, \ldots, Q^m$.

This completes the description of the decomposition procedure. We need to argue that the decomposition $\mathcal{Q}$ created by above procedure is not too large, that is, of size $|\mathcal{Q}| \leq 2kn \cdot (2s)^d$. Let us start with a few observations.

Because G is D-well-behaved it holds that G has $(1/k, p, D/4)$-bounded common neighborhoods in every block (see Property 1 of Definition V.4). Let Q be a rectangle with d blocks of size 1 and with the remaining vertices contained in the common neighborhood of these singletons. All such rectangles Q are small. Thus the decomposition procedure does not need to decompose such rectangles Q any further.

Whenever we decompose a rectangle in Cases 2 and 3 all rectangles that we need to recursively decompose have one more singleton. Because we can stop decomposing after identifying d singletons and in Cases 2 and 3 we create at most $2s$ many rectangles that require further decomposition we end up with at most $2(2s)^d$ many rectangles. We ignored the rectangles from Case 1 so far. But each rectangle that requires further decomposition from Cases 2 and 3 results in at most another kn many rectangles from Case 1. Thus the size of the family of rectangles is bounded by $2kn \cdot (2s)^d$. $\square$

VII. Concluding Remarks

For $k \leq n^{1/100}$ we prove an essentially tight average-case $n^{\Omega(D)}$ size lower bound on unary Sherali-Adams refutations of the k-clique formula for Erdős-Rényi random graphs with maximum clique of size D. In fact, we obtain a lower bound on the sum of the magnitude of the coefficients appearing in a (general) Sherali-Adams refutation. The obvious problem left open is to prove an $n^{\Omega(D)}$ *monomial* size lower bound on Sherali-Adams refutations of the clique formula.

One possible avenue to prove such a monomial size lower bound is to argue that any Sherali-Adams proof of the clique formula can be converted into a proof of the same monomial size but with small coefficients. In fact, a slightly weaker statement would suffice: recall that our lower bound only counts the size of the coefficients of generalized monomials as well as of monomials multiplied by edge axioms. As such we would just need to be able to convert a general Sherali-Adams refutation into a refutation with low coefficients for such monomials.

In contrast to previous lower bounds for clique, our proof strategy is not purely combinatorial. It might be fruitful to obtain an explicit combinatorial description of μ_d—we believe this could potentially be used to prove average-case clique lower bounds for other proof systems, including resolution.

A strength of our lower bound approach is that it is quite oblivious to the encoding: one can introduce all possible extension variables depending on a *single* block and the lower bound argument still goes through. This is because the only property we require of a monomial m is that the set of tuples $Q_m = \{t \mid \rho_t(m) \neq 0\}$ whose associated assignment ρ_t sets m to non-zero is a rectangle. By extending ρ in the natural

manner to extension variables it is easy to see that Q_m is still a rectangle.

Our lower bound strategy seems to fail quite spectacularly once the edge probability is increased well beyond $1/2$. More precisely, once $D = \omega(\log n)$, we fail to counter exponential in d^2 factors that arise from encoding the core graphs: as long as $D = O(\log n)$ we can counter these with s^{-d} terms, where s is the minimum block size of a good rectangle. As s is clearly bounded by the block size n, this approach fails once $D = \omega(\log n)$. We leave it as an open problem to extend our result to the dense setting.

We rely on rather unorthodox pseudorandomness properties of the underlying graph. It is natural to wonder whether these properties follow from a previously studied notion of pseudorandomness. Furthermore, it is wide open whether our lower bound can be made explicit. In particular, we have not investigated whether graphs that satisfy our pseudorandomness property can be constructed deterministically.

Another application of our pseudo-measure μ_d is in communication complexity. Suppose we consider the k-player number-in-hand model, where player i obtains a single node u_i from block V_i. The goal of the k players is to find an edge missing in the induced subgraph by the tuple $(u_1, \ldots, u_k)$. Consider the leaves of such a communication protocol. Note that each leaf ℓ is associated with a subrectangle Q_ℓ of an edge axiom. As the family of these associated rectangles Q_ℓ partition the whole space, but $|\mu_d(Q_\ell)| \leq n^{-\Omega(D)}$, there must be at least $n^{\Omega(D)}$ leaves.

Finally, we have not investigated whether our technique can be used to obtain lower bounds for other proof systems. For example, is it possible that with similar ideas one could obtain tree-like cutting planes lower bounds with bounded coefficients? Possibly even with unbounded coefficients? The communication complexity view of the problem suggests that this may be a viable approach.

ACKNOWLEDGEMENTS

The authors are grateful to Albert Atserias, Per Austrin, Johan Håstad, Jakob Nordström, Pavel Pudlák, Dmitry Sokolov, Joseph Swernofsky, and Neil Thapen for helpful discussions and feedback. In particular we would like to thank Albert Atserias who observed that cores seem to be related to kernels.

REFERENCES

[1] R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations*, ser. The IBM Research Symposia Series. Springer, 1972, pp. 85–103.

[2] S. A. Cook, "The complexity of theorem-proving procedures," in *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing (STOC '71)*, May 1971, pp. 151–158.

[3] J. Håstad, "Clique is hard to approximate within $n^{1-\epsilon}$," *Acta Mathematica*, vol. 182, pp. 105–142, 1999, preliminary version in *FOCS '96*.

[4] D. Zuckerman, "Linear degree extractors and the inapproximability of max clique and chromatic number," *Theory of Computing*, vol. 3, no. 6, pp. 103–128, Aug. 2007, preliminary version in *STOC '06*.

[5] J. Nešetřil and S. Poljak, "On the complexity of the subgraph problem," *Commentationes Mathematicae Universitatis Carolinae*, vol. 026, no. 2, pp. 415–419, 1985.

[6] R. Downey and M. R. Fellows, "Fixed-parameter tractability and completeness II: Completeness for W[1]," *Theoretical Computer Science A*, vol. 141, no. 1–2, pp. 109–131, Apr. 1995.

[7] R. Impagliazzo and R. Paturi, "On the complexity of k-SAT," *Journal of Computer and System Sciences*, vol. 62, no. 2, pp. 367–375, Mar. 2001, preliminary version in *CCC '99*.

[8] J. Chen, X. Huang, I. A. Kanj, and G. Xia, "Linear FPT reductions and computational lower bounds," in *Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC '04)*, Jun. 2004, pp. 212–221.

[9] B. Rossman, "On the constant-depth complexity of k-clique," in *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, C. Dwork, Ed. ACM, 2008, pp. 721–730. [Online]. Available: https://doi.org/10.1145/1374376.1374480

[10] ——, "The monotone complexity of k-clique on random graphs," in *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*. IEEE Computer Society, 2010, pp. 193–201. [Online]. Available: https://doi.org/10.1109/FOCS.2010.26

[11] A. Atserias, I. Bonacina, S. F. de Rezende, M. Lauria, J. Nordström, and A. A. Razborov, "Clique is hard on average for regular resolution," *J. ACM*, vol. 68, no. 4, pp. 23:1–23:26, 2021. [Online]. Available: https://doi.org/10.1145/3449352

[12] M. Lauria, P. Pudlák, V. Rödl, and N. Thapen, "The complexity of proving that a graph is Ramsey," *Combinatorica*, vol. 37, no. 2, pp. 253–268, Apr. 2017, preliminary version in *ICALP '13*.

[13] S. S. Dantchev, N. Galesi, A. Ghani, and B. Martin, "Proof complexity and the binary encoding of combinatorial principles," *CoRR*, vol. abs/2008.02138, 2020. [Online]. Available: https://arxiv.org/abs/2008.02138

[14] S. Pang, "Large clique is hard on average for resolution," in *Computer Science - Theory and Applications - 16th International Computer Science Symposium in Russia, CSR 2021, Sochi, Russia, June 28 - July 2, 2021, Proceedings*, ser. Lecture Notes in Computer Science, R. Santhanam and D. Musatov, Eds., vol. 12730. Springer, 2021, pp. 361–380. [Online]. Available: https://doi.org/10.1007/978-3-030-79416-3_22

[15] O. Beyersdorff, N. Galesi, M. Lauria, and A. A. Razborov, "Parameterized bounded-depth Frege is not optimal," *ACM Transactions on Computation Theory*, vol. 4, no. 3, pp. 7:1–7:16, Sep. 2012, preliminary version in *ICALP '11*.

[16] M. Lauria, "Cliques enumeration and tree-like resolution proofs," *Inf. Process. Lett.*, vol. 135, pp. 62–67, 2018. [Online]. Available: https://doi.org/10.1016/j.ipl.2018.03.001

[17] P. Beame, R. Impagliazzo, and A. Sabharwal, "The resolution complexity of independent sets and vertex covers in random graphs," *Comput. Complex.*, vol. 16, no. 3, pp. 245–297, 2007. [Online]. Available: https://doi.org/10.1007/s00037-007-0230-0

[18] R. Meka, A. Potechin, and A. Wigderson, "Sum-of-squares lower bounds for planted clique," in *Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC '15)*, Jun. 2015, pp. 87–96.

[19] B. Barak, S. Hopkins, J. Kelner, P. K. Kothari, A. Moitra, and A. Potechin, "A nearly tight sum-of-squares lower bound for the planted clique problem," *SIAM Journal on Computing*, vol. 48, no. 2, pp. 687–735, 2019. [Online]. Available: https://doi.org/10.1137/17M1138236

[20] S. Pang, "SOS lower bound for exact planted clique," in *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, ser. LIPIcs, V. Kabanets, Ed., vol. 200. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, pp. 26:1–26:63. [Online]. Available: https://doi.org/10.4230/LIPIcs.CCC.2021.26

[21] S. Margulies, "Computer algebra, combinatorics, and complexity: Hilbert's Nullstellensatz and NP-complete problems," Ph.D. dissertation, University of California, Davis, 2008.

[22] M. Göös, A. Hollender, S. Jain, G. Maystre, W. Pires, R. Robere, and R. Tao, "Separations in proof complexity and TFNP," in *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*. IEEE, 2022, pp. 1150–1161. [Online]. Available: https://doi.org/10.1109/FOCS54457.2022.00111

[23] Y. Filmus, M. Mahajan, G. Sood, and M. Vinyals, "MaxSAT resolution and subcube sums," *ACM Transactions on Computational Logic*, vol. 24, no. 1, pp. 1–27, Jan. 2023.

[24] A. Atserias and T. Hakoniemi, "Size-Degree Trade-Offs for Sums-of-Squares and Positivstellensatz Proofs," in *34th Computational Complexity Conference (CCC 2019)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), A. Shpilka, Ed., vol. 137. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019, pp. 24:1–24:20. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2019/10846

[25] A. Potechin and A. Zhang, "Bounds on the total coefficient size of nullstellensatz proofs of the pigeonhole principle and the ordering principle," 2022.

[26] M. R. Fellows, L. Jaffke, A. I. Király, F. A. Rosamond, and M. Weller, *What Is Known About Vertex Cover Kernelization?* Cham: Springer International Publishing, 2018, pp. 330–356. [Online]. Available: https://doi.org/10.1007/978-3-319-98355-4_19