Hypercontractive inequalities via SOS, with an application to Vertex-Cover

Manuel Kauers* Ryan O'Donnell† Li-Yang Tan‡ Yuan Zhou†§

December 24, 2012

Abstract

Our main result is a formulation and proof of the reverse hypercontractive inequality in the sum-of-squares (SOS) proof system. As a consequence, we show that for any constant $\gamma > 0$, the $O(1/\gamma)$-round SOS/Lasserre SDP hierarchy certifies the statement "$\mathsf{Min\text{-}Vertex\text{-}Cover}(G^n_\gamma) \geq (1 - o_n(1))|V|$", where $G^n_\gamma = (V, E)$ is the "Frankl–Rödl graph" with $V = \{0, 1\}^n$ and $(x, y) \in E$ whenever $\Delta(x, y) = (1 - \gamma)n$. This is despite the fact that k rounds of various LP and SDP hierarchies fail to certify the statement "$\mathsf{Min\text{-}Vertex\text{-}Cover}(G^n_\gamma) \geq (\frac{1}{2} + \epsilon)|V|$" once $\gamma = \gamma(k, \epsilon) > 0$ is small enough. Finally, we also give an SOS proof of (a generalization of) the sharp $(2, q)$-hypercontractive inequality for any even integer q.

1 Introduction

Hypercontractive inequalities play an important role in analysis of Boolean functions. They are concerned with the *noise operator* T_ρ which acts on functions $f : \{-1, 1\}^n \to \mathbb{R}$ via $T_\rho f(x) = \mathbf{E}[f(\boldsymbol{y})]$, where $\boldsymbol{y}$ is a "ρ-correlated copy" of x. Equivalently, $T_\rho f = \sum_{S\subseteq[n]} \rho^{|S|} \widehat{f}(S)\chi_S$, where the numbers $\widehat{f}(S)$ are the Fourier coefficients of f. The standard hypercontractivity inequality was first proved by Bonami [Bon70] and the reverse hypercontractivity inequality was first proved by Borell [Bor82]. We state both, recalling the notation $\|f\|_p = \mathbf{E}_{\boldsymbol{x}\sim\{-1,1\}^n}[|f(\boldsymbol{x})|^p]^{1/p}$.

Hypercontractive Inequality. *Let $f : \{-1, 1\}^n \to \mathbb{R}$, let $1 \leq p \leq q \leq \infty$, and let $0 \leq \rho \leq \sqrt{(p - 1)/(q - 1)}$. Then $\|T_\rho f\|_q \leq \|f\|_p$.*

Reverse Hypercontractive Inequality. *Let $f : \{-1, 1\}^n \to \mathbb{R}^{\geq 0}$, let $-\infty \leq q \leq p \leq 1$, and let $0 \leq \rho \leq \sqrt{(1 - p)/(1 - q)}$. Then $\|T_\rho f\|_q \geq \|f\|_p$.*

The hypercontractive inequality is almost always used with either $p = 2$ or $q = 2$. The $(2, 4)$-hypercontractivity inequality — i.e., the case $q = 4$, $p = 2$, $\rho = 1/\sqrt{3}$ — is a particularly useful case, as is the following easy corollary:

Theorem 1.1. *For $k \in \mathbb{N}$, let $\mathcal{P}^{\leq k}$ be the projection operator which maps $f : \{-1, 1\}^n \to \mathbb{R}$ to its low-degree part $\mathcal{P}^{\leq k} f = \sum_{|S|\leq k} \widehat{f}(S)\chi_S$. Then the $2 \to 4$ operator norm of $\mathcal{P}^{\leq k}$ is at most $3^{k/2}$. I.e., $\|\mathcal{P}^{\leq k} f\|_4 \leq 3^{k/2}\|f\|_2$.*

Theorem 1.1 is known to have a proof which is noticeably simpler than that of the general hypercontractive inequality [MOO10]. Theorem 1.1 can be used to prove, e.g., the KKL Theorem [KKL88], the sharp small-set expansion statement for the 1/3-noisy hypercube, and the Invariance Principle of [MOO10]. More generally, the hypercontractivity inequality has the following corollary:

Theorem 1.2. *For any $q \geq 2$ and $f : \{-1, 1\}^n \to \mathbb{R}$ we have $\|\mathcal{P}^{\leq k} f\|_q \leq (q - 1)^{k/2}\|f\|_2$.*

This corollary is often use to control the behavior of low-degree polynomials of random bits.

Reverse hypercontractivity is perhaps most often used to show that if $A, B \subseteq \{-1, 1\}^n$ are large sets and $(\boldsymbol{x}, \boldsymbol{y})$ is a ρ-correlated pair of random strings then there is a substantial chance that $\boldsymbol{x} \in A$ and $\boldsymbol{y} \in B$. This was first deduced in [MOR+06] by deriving the following consequence of reverse hypercontractivity:

*Research Institute for Symbolic Computation, Johannes Kepler Universität. Supported by FWF grant Y464-N18.

†Department of Computer Science, Carnegie Mellon University. Supported by NSF grants CCF-0747250 and CCF-1116594, a Sloan fellowship, and a grant from the MSR–CMU Center for Computational Thinking.

‡Department of Computer Science, Columbia University. Research done while visiting CMU.

§Also supported by a grant from the Simons Foundation (Award Number 252545).

Theorem 1.3. *Let $f, g : \{-1,1\}^n \to \mathbb{R}^{\geq 0}$, let $0 \leq q \leq 1$, and let $0 \leq \rho \leq 1 - q$. Then $\mathbf{E}[f(\boldsymbol{x})g(\boldsymbol{y})] \geq \|f\|_q \|g\|_q$ when $(\boldsymbol{x}, \boldsymbol{y})$ is a pair of ρ-correlated random strings.*

The reverse hypercontractive inequality has been used, e.g., in problems related to approximability and hardness of approximation [FKO07, She09, BHM12], and problems in quantitative social choice [MOO10, Mos12a, MOS12b, Kel12, MR12].

1.1 Sum-of-squares proofs of hypercontractive inequalities

The present work is concerned with proving hypercontractive inequalities via "sums of squares" (SOS); i.e., in the Positivstellensatz proof system introduced by Grigoriev and Vorobjov [GV01]. A recent work of Barak et al. [BBH+12] showed that the Khot–Vishnoi [KV05] SDP integrality gap instances for Unique-Games are actually well-solved by the "4-round Lasserre SDP hierarchy"; equivalently, the "degree-8 SOS hierarchy". This is despite the fact that they are strong gap instances for superconstantly many rounds of other weaker SDP hierarchies such as Lovász–Schrijver$^+$ and Sherali–Adams$^+$ [RS09, KS09]. The key to analyzing the optimum value of the Khot–Vishnoi instances is the hypercontractive inequality, and perhaps the key technical component of the Barak et al. result is showing that Theorem 1.1 has a degree-4 "SOS proof". That is, if we treat the each $f(x)$ as a formal "indeterminate", then $9^k\|f\|_2^4 - \|\mathcal{P}^{\leq k} f\|_4^4$ is a degree-4 polynomial in the 2^n indeterminates, and Barak et al. showed that it is a sum of squared polynomials (hence always nonnegative).

The connection between SOS proofs and SDP relaxations for optimization problems was made independently by Lasserre [Las00] and Parrilo [Par00]. Roughly speaking, if a system of n-variate polynomial inequalities can be refuted within the degree-d SOS proof system of Grigoriev and Vorobjov [GV01], then this refutation can also be found efficiently by solving a semidefinite program of size $n^{O(d)}$. (For more details, see e.g. [OZ13].) The associated "degree-d SOS hierarchy" for approximating optimization problems is known to be at least as strong as the Lovász–Schrijver$^+$ and Sherali–Adams$^+$ SDP hierarchies, and the [BBH+12] result shows that it can be noticeably stronger for the notorious Unique Games problem.

Later, [OZ13] showed that the degree-4 SOS hierarchy correctly analyzes the value of the [DKSV06] instances of Balanced-Separator, which are known to be superconstant-factor integrality gap instances for superconstantly many rounds of the "LH SDP hierarchy" [RS09]. It was also shown in [OZ13] that the degree-$O(1)$ SOS hierarchy certifies the value of the [KV05] instances of Max-Cut to within factor .952, whereas superconstantly many rounds of the Sherali–Adams$^+$ hierarchy are still off by a factor of .878 [RS09, KS09]. The key to the former result was an SOS proof of the KKL Theorem (relying on [BBH+12]'s SOS proof of Theorem 1.1); the key to the latter was an SOS proof of an Invariance Principle variant, which in turned needed an SOS proof Theorem 1.2. The work [OZ13] was unable to actually obtain Theorem 1.2 with an SOS proof, but instead obtained a weaker version which sufficed for their purposes.

Still, the full power of the SOS hierarchy is far from well-understood. Analyzing what can and cannot be proved with low-degree SOS proofs is evidently very important; for example, it's consistent with our current knowledge that the degree-4 SOS hierarchy refutes the Unique-Games Conjecture, gives a 1.01-approximation for Uniform Sparsest-Cut, a .94-approximation for Max-Cut, and a 1.4-approximation for Vertex-Cover.

In particular, hypercontractive inequalities have played a key role in many of the sophisticated SDP integrality gap instances. Thus it is natural to ask: Can a sharp version of the hypercontractive inequality be proved in the SOS proof system? Can any version of the reverse hypercontractive inequality be proved? As we will see, the latter question is particularly relevant for the known SDP integrality instances of the Vertex-Cover problem, a basic optimization task not yet analyzed via SOS.

1.2 Our results

The main result in this paper is an SOS proof of the reverse hypercontractivity Theorem 1.3 for all q equal to the reciprocal of an even integer. As an application of this, we show that the $O(1)$-degree SOS hierarchy correctly analyzes the Frankl–Rödl SDP integrality gap instances for Vertex-Cover for a large range of parameters. Finally, we also give an SOS proof of the sharp $(2, q)$-hypercontractive inequality for all even integers q; in fact, a version with relaxed moment conditions. We find it interesting to see that the two powerful hypercontractive inequalities admit proofs as "elementary" as sum-of-squares proofs. On the other hand, to obtain these proofs we had to use somewhat elaborate methods, including computer algebra techniques.

The hypercontractive inequality for even integer norms. As mentioned, Barak et al. [BBH+12] gave an SOS proof of Theorem 1.1, that $\|\mathcal{P}^{\leq k} f\|_4^4 \leq 9^k \|f\|_2^4$. Although there is a very easy proof of this theorem "in ZFC" [MOO10], that proof uses the Cauchy–Schwarz inequality, whose square-roots do not obviously translate into SOS statements. The SOS proof in [BBH+12] gets around this by proving the *generalized* statement $\mathbf{E}[(\mathcal{P}^{\leq k} f)^2 (\mathcal{P}^{\leq k'} g)^2] \leq 3^{k+k'} \mathbf{E}[f^2]\,\mathbf{E}[g^2]$, allowing them to replace Cauchy–Schwarz with $XY \leq \frac{1}{2}X^2 + \frac{1}{2}Y^2$. In [OZ13] this SOS proof was very slightly generalized to cover the $(2,4)$-hypercontractive inequality, $\mathbf{E}[(T_\rho f)^2 (T_\rho g)^2] \leq \mathbf{E}[f^2]\,\mathbf{E}[g^2]$ for $\rho = 1/\sqrt{3}$. That work also gave an SOS proof of a weakened version of Theorem 1.2 for all even integers q, namely $\|\mathcal{P}^{\leq k} f\|_q^q \leq q^{O(qk/2)} \|f\|_2^q$. (Attention is restricted to even integers q because the $(2,q)$-hypercontractive inequality cannot even be stated as a polynomial inequality otherwise.)

In Section 3 we prove the full $(2,q)$-hypercontractive inequality for all even integers q. Our strategy is as follows. First, we give a simple proof ("in ZFC") of $(2,q)$-hypercontractivity for all even integers q; our proof works not just for random ± 1 bits but for any random variables satisfying fairly liberal moment bounds. Indeed, we are not aware of any previous work showing that such moment bounds are sufficient for hypercontractivity. However this proof relies on the well-known fact that the hypercontractivity inequality tensorizes [KS88], which in turn uses the triangle inequality for the $(q/2)$-norm, an inequality that cannot even be stated in SOS. For our SOS extension of this result we move to a $(q/2)$-function version of the statement as in [BBH+12]; this requires a little more work and a very mild use of computer algebra at the end to prove a binomial sum identity. Our final theorem is as follows:

Theorem 1.4. *(Informal.) Let $s \in \mathbb{N}^+$ and write $q = 2s$. Let $0 \leq \rho \leq \frac{1}{\sqrt{q-1}}$. Let $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ be a sequence of independent real random variables, with each $\boldsymbol{x}_i$ satisfying*

$$\mathbf{E}[\boldsymbol{x}_i^{2j-1}] = 0, \quad \mathbf{E}[\boldsymbol{x}_i^{2j}] \leq (2s-1)^j \frac{\binom{s}{j}}{\binom{2s}{2j}} \quad \text{for all integers } 1 \leq j \leq s;$$

further assume that $\mathbf{E}[\boldsymbol{x}_i^2] = 1$ for each i. (Rademachers and standard Gaussians qualify.) Then for functions $f_1, \ldots, f_s : \{-1,1\}^n \to \mathbb{R}$ there is an SOS proof of

$$\mathbf{E}\left[\prod_{i=1}^{s} (T_\rho f_i(\boldsymbol{x}))^2\right] \leq \prod_{i=1}^{s} \mathbf{E}[f_i(\boldsymbol{x})^2].$$

As corollaries we have SOS proofs of $\|T_\rho f\|_q^q \leq \|f\|_2^q$ and $\|\mathcal{P}^{\leq k} f\|_q^q \leq (q-1)^{qk/2} \|f\|_2^q$.

The reverse hypercontractive inequality. Giving an SOS proof of this theorem proved to be significantly more difficult; it is our main result and the source of our application to Vertex-Cover integrality gaps. The theorem cannot even be stated in the SOS proof system directly since the p-"norms" are not polynomials in the values $f(x)$ for $p < 1$. We turn to the 2-function version from [MOR+06], Theorem 1.3; if $q = \frac{1}{2k}$ for some $k \in \mathbb{N}^+$ and if we replace f and g by f^{2k} and g^{2k} then we get a polynomial statement (and we can even drop the hypothesis that f and g are nonnegative). The resulting theorem is:

Theorem 1.5. *Let $k \in \mathbb{N}^+$ and let $0 \leq \rho \leq 1 - \frac{1}{2k}$. Then for functions $f, g : \{-1,1\}^n \to \mathbb{R}$ there is an SOS proof of*

$$\mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y}) \\ \rho\text{-corr'd}}} [f(\boldsymbol{x})^{2k} g(\boldsymbol{y})^{2k}] \geq \mathbf{E}[f]^{2k}\,\mathbf{E}[g]^{2k}.$$

We prove this result in Section 4. An induction on n easily reduces the problem to the $n = 1$ case; for each k, this is an inequality in four real indeterminates. Then by homogeneity we can further reduce to an inequality in just two indeterminates. Nevertheless, giving an SOS-proof of this "two-point inequality" for all k seems to be surprisingly tricky. As an example of the problem we need to solve (the $k = 3$ case), the reader is invited to try the following puzzle:

"Show that

$$\tfrac{11}{24}(1+a)^6(1+b)^6 + \tfrac{11}{24}(1-a)^6(1-b)^6 + \tfrac{1}{24}(1+a)^6(1-b)^6 + \tfrac{1}{24}(1-a)^6(1+b)^6 - 1$$

is a sum of squared polynomials in a and b."

Our solution is presented in Section 4.1. Our high level approach is to employ a change of variables which reduces the task to proving a sequences of one-variable real inequalities. This is helpful because every nonnegative univariate polynomial is SOS; hence we can use any mathematical technique to verify the one-variable inequalities. We establish the one-variable inequalities using techniques from computer algebra. Peculiarly, this approach only works for the specific choice $\rho = 1 - \frac{1}{2k}$; however the proof for general $0 \leq \rho \leq 1 - \frac{1}{2k}$ can be deduced since the two-point inequality is linear in ρ.

1.2.1 Application to integrality gap instances for minimum vertex cover

Along the lines of [BBH+12, OZ13], our SOS proof of the reverse hypercontractive inequality also has application to integrality gap instances; specifically, for the Minimum Vertex-Cover problem. Let us briefly recall the relevant work on Vertex-Cover, and then state our result.

Previous work on approximating Vertex-Cover. We are interested in the problem of finding a C-approximate minimum vertex cover in a graph; i.e., one whose cardinality is at most C times that of the minimum vertex cover. For $C = 2$ the problem is easily solved in linear time [GJ79, Gavril 1974]; for $C = 1.36$ the problem is known to be NP-hard [DS05]. But for $C = 1.5$, say, we know neither polynomial-time solvability, nor NP-hardness. Based on the nearly 30-year lack of progress on the algorithms side, it is reasonable to suspect that there is no efficient $(2-\epsilon)$-approximation algorithm. Indeed, this is known to be true assuming the Unique-Games Conjecture [KR08]. However there is reasonable doubt about the Unique-Games Conjecture [ABS10] and one may seek alternative evidence of hardness. One fairly strong form of evidence is showing that various linear programming-based and/or semidefinite programming-based generic polynomial-time optimization algorithms fail to even $(2-\epsilon)$-certify minimum vertex covers. Here we are using the following definition:

Definition 1.6. We say that an algorithm *C-certifies* the minimum vertex cover in a graph if, on input a graph G, it outputs a number α guaranteed to satisfy $\alpha \le \mathsf{Min\text{-}Vertex\text{-}Cover}(G) \le C\alpha$. (This is a strictly easier problem than actually *finding* C-approximate minimum vertex covers.)

There is a long line of work of this type [KG98, Cha02, ABL02, ABLT06, Tou06, FO06, STT07a, STT07b, GMT08, GM08, Sch08, CMM09, Tul09, GMPT10, GM10, BCGM11]; see Georgiou's thesis [Geo10] for a recent survey. Briefly, it is known that the following algorithms fail to provide $(2-\epsilon)$-certifications for the minimum vertex cover: $N^{\Omega(1)}$ levels of the Sherali–Adams LP hierarchy [CMM09] (where N is the number of vertices); $\Omega(\sqrt{\frac{\log N}{\log\log N}})$ levels of the Lovász–Schrijver$^+$ SDP hierarchy; and, 6 levels of the Sherali–Adams$^+$ SDP hierarchy [BCGM11]. Further, [BCGM11] conjectures (based on numerical evidence) that their 6-level result can be extended to any constant number of levels. Since the Sherali–Adams$^+$ hierarchy is stronger than the Sherali–Adams and Lovász–Schrijver$^+$ hierarchies, this conjecture would subsume the other two mentioned results, at least with regards to ruling out polynomial-time (constant-level) algorithms.

The *Frankl–Rödl graphs*, defined as follows, are used to show failure of the Sherali–Adams$^+$ hierarchy to $(2-\epsilon)$-certify the minimum vertex cover. As far as we are aware, these are the only known "hard instances" even for the basic SDP relaxation (for factor $(2-\epsilon)$).

Definition 1.7. Let $n \in \mathbb{N}$ and let $0 \le \gamma \le 1$ be such that $(1-\gamma)n$ is an even integer. The *Frankl–Rödl* graph G^n_γ is the undirected graph with vertex set $\{-1,1\}^n$ and edge set $\{(x,y) : \Delta(x,y) = (1-\gamma)n\}$, where $\Delta(\cdot,\cdot)$ denotes Hamming distance.

The following theorem is essentially due to Frankl and Rödl [FR87], though a few details of the proof are only worked out in [GMPT10].

Theorem 1.8. *There exists $\gamma_0 > 0$ such that for all $\gamma \le \gamma_0$ it holds that $\mathsf{Min\text{-}Vertex\text{-}Cover}(G^n_\gamma) \ge (1 - n(1-\gamma^2/64)^n)N$, where $N = 2^n$ is the number of vertices in the graph. In particular, $\mathsf{Min\text{-}Vertex\text{-}Cover}(G^n_\gamma) \ge (1 - o_N(1))N$ for $.1\sqrt{\frac{\log n}{n}} \le \gamma \le \gamma_0$.*

On the other hand, Benabbas, Chan, Georgiou, and Magen [BCGM11] showed that for all $\epsilon > 0$ there exists $\gamma > 0$ such that the level-6 Sherali–Adams$^+$ algorithm can only certify that $\mathsf{Min\text{-}Vertex\text{-}Cover}(G^n_\gamma) \ge (1/2+\epsilon)N$ (for n sufficiently large). They further conjectured that this remains true for level-k Sherali–Adams$^+$ assuming $\gamma = \gamma(\epsilon, k)$ is small enough. Specifically, they suggest fixing $\gamma = \Theta(\sqrt{\frac{\log n}{n}})$.

Our result. As an application of our SOS proof for the reverse hypercontractive inequality, we are able to show that the SOS/Lasserre hierarchy *is* an effective certification algorithm for the Frankl–Rödl graphs for any constant $\gamma > 0$. Specifically, we show that the $O(1/\gamma)$-degree SOS hierarchy provides a $(1+o_n(1))$-certification for the minimum vertex cover in the Frankl–Rödl graph G^n_γ. In fact, our analysis works even for any $\gamma > \omega(\frac{1}{\log n})$; whether it continues to hold for γ as small as $\Theta(\sqrt{\frac{\log n}{n}}) = \Theta(\sqrt{\frac{\log\log N}{\log N}})$ (the parameter setting suggested in [GMPT10, BCGM11]) remains open.

To accomplish our goal, we need to show an SOS proof for the Frankl–Rödl Theorem. A key ingredient in Frankl and Rödl's original proof is the vertex isoperimetric inequality on $\{-1,1\}^n$, due to Harper. The standard proof of this inequality involves a "shifting" argument which we do not see how to carry out with SOS. However, it is known that inequalities of this type can also be proved using the reverse hypercontractive inequality studied in this paper. In particular, Benabbas, Hatami, and Magen [BHM12] have very recently proven a "density" variation of the Frankl–Rödl Theorem using the reverse hypercontractive inequality. We obtain the SOS proof for the Frankl–Rödl Theorem by combining our SOS proof for the reverse hypercontractive inequality and an SOS version of the Benabbas–Hatami–Magen proof; see Section 5.

2 Preliminaries

The SOS proof system. We describe the SOS (Positivstellensatz) proof system of Grigoriev and Vorobjov [GV01] using the notation from the work [OZ13]; for more details, please see that paper.

Definition 2.1. Let $X = (X_1, \dots, X_n)$ be indeterminates, let $q_1, \dots, q_m, r_1, \dots, r_{m'} \in \mathbb{R}[X]$, and let

$$A = \{q_1 \geq 0, \dots, q_m \geq 0\} \cup \{r_1 = 0, \dots, r_{m'} = 0\}.$$

Given $p \in \mathbb{R}[X]$ we say that A *SOS-proves* $p \geq 0$ *with degree* k, written

$$A \quad \vdash_k \quad p \geq 0,$$

whenever

$$\exists v_1, \dots, v_{m'} \text{ and SOS } u_0, u_1, \dots, u_m \text{ such that}$$
$$p = u_0 + \sum_{i=1}^{m} u_i q_i + \sum_{j=1}^{m'} v_j r_j, \qquad \text{with } \deg(u_0), \deg(u_i q_i), \deg(v_j r_j) \leq k \ \forall i \in [m], j \in [m'].$$

Here we use the abbreviation "$w \in \mathbb{R}[X]$ is SOS" to mean $w = s_1^2 + \cdots + s_t^2$ for some $s_i \in \mathbb{R}[X]$. We say that A has a *degree-k SOS refutation* if

$$A \quad \vdash_k \quad -1 \geq 0.$$

Finally, when $A = \emptyset$ we will sometimes use the shorthand

$$\vdash_k \quad p \geq 0,$$

which simply means that p is SOS and $\deg(p) \leq k$.

Analysis of boolean functions. Let us recall some standard notation from the field. We write $\boldsymbol{x} \sim \{-1,1\}^n$ to denote that the string $\boldsymbol{x}$ is drawn uniformly at random from $\{-1,1\}^n$. Given $f : \{-1,1\}^n \to \mathbb{R}$ we sometimes use abbreviations like $\mathbf{E}[f]$ for $\mathbf{E}_{\boldsymbol{x} \sim \{-1,1\}^n}[f(\boldsymbol{x})]$. For $f, g : \{-1,1\}^n \to \mathbb{R}$ we also write $\langle f, g \rangle = \mathbf{E}[fg] = \mathbf{E}_{\boldsymbol{x} \sim \{-1,1\}^n}[f(\boldsymbol{x})g(\boldsymbol{x})]$. For $-1 \leq \rho \leq 1$ we say that $(\boldsymbol{x}, \boldsymbol{y}) \sim \{-1,1\}^n \times \{-1,1\}^n$ is a pair of *ρ-correlated random strings* if the pairs $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ are independent for $i \in [n]$ and satisfy $\mathbf{E}[\boldsymbol{x}_i] = \mathbf{E}[\boldsymbol{y}_i] = 0$ and $\mathbf{E}[\boldsymbol{x}_i \boldsymbol{y}_i] = \rho$. The operator T_ρ acts on functions $f : \{-1,1\}^n \to \mathbb{R}$ via $T_\rho f(x) = \mathbf{E}[f(\boldsymbol{y}) \mid \boldsymbol{x} = x]$, where $(\boldsymbol{x}, \boldsymbol{y})$ is a pair of ρ-correlated random strings.

Simple SOS facts and lemmas. We will use the following facts and lemmas in our SOS proofs. The first one, in particular, we use throughout without comment.

Lemma 2.2.

$$\begin{aligned} &\textit{If} \quad A \quad \vdash_k \quad p \geq 0, \quad A' \quad \vdash_{k'} \quad p' \geq 0, \\ &\textit{then} \quad A \cup A' \quad \vdash_{\max(k,k')} \quad p + p' \geq 0. \end{aligned}$$

The following fact is a well-known consequence of the Fundamental Theorem of Algebra.

Fact 2.3. *A univariate polynomial $p(x)$ is SOS if it is nonnegative. In other words, we have*

$$\vdash_{\deg(p)} p(x) \geq 0,$$

when $p(x) \geq 0$ for all $x \in \mathbb{R}$.

It is also well known that for homogeneous polynomials, one can reduce the number of variables by 1 by "dehomogenizing" the polynomial, getting an SOS representation (if there is one), and rehomogenizing it to get an SOS representation of the original polynomial. Applying this trick to Fact 2.3, we get:

Fact 2.4. *A homogeneous bivariate polynomial $p(x, y)$ is SOS if it is nonnegative.*

Here are some additional lemmas:

Lemma 2.5. *Let $c \geq 0$ be a constant and X an indeterminate. Then for any $k \in \mathbb{N}^+$,*

$$X \geq c \quad \vdash_k \quad X^k \geq c^k.$$

Proof. This follows because

$$X^k - c^k = (X - c + c)^k - c^k = \sum_{i=1}^{k} \tbinom{k}{i} c^{k-i}(X-c)^i$$

and each power $(X - c)^i$ is either a square or $(X - c)$ times a square. □

Lemma 2.6. *For any $k \in \mathbb{N}^+$ we have*

$$\vdash_{2k} \quad \left(\tfrac{X+Y}{2}\right)^{2k} \leq \tfrac{X^{2k}+Y^{2k}}{2}.$$

Proof. Since $\frac{X^{2k}+Y^{2k}}{2} - \left(\frac{X+Y}{2}\right)^{2k}$ is a degree-$2k$ homogeneous polynomial, the claim follows from Fact 2.4: the inequality is indeed true by convexity of $t \mapsto t^k$. □

3 The hypercontractive inequality in SOS

As a warmup, we give a simple proof ("in ZFC") of the $(2, q)$-hypercontractive inequality $\|T_\rho f\|_q \leq \|f\|_q$ for all even integers q, which implies Theorem 1.2 for all even integers q. As mentioned, we do this under a significantly weakened moment condition:

"s-Moment Conditions." *For a real random variable $\boldsymbol{x}_i$, the condition is that $\mathbf{E}[\boldsymbol{x}_i^2] = 1$ and*

$$\mathbf{E}[\boldsymbol{x}_i^{2j-1}] = 0, \quad \mathbf{E}[\boldsymbol{x}_i^{2j}] \leq (2s-1)^j \frac{\binom{s}{j}}{\binom{2s}{2j}} \quad \text{for all integers } 1 \leq j \leq s.$$

Our proof will show that these moment conditions are sharp; none of them can be relaxed.

Remark 3.1. By converting to factorials and expanding, one verifies that

$$(2s-1)^j \frac{\binom{s}{j}}{\binom{2s}{2j}} = (2j-1)!! \cdot \prod_{i=1}^{j-1} \frac{2s-1}{2s-(2i+1)}.$$

It follows that for each fixed $j \in \mathbb{N}^+$, the quantity decreases as a function of s (for $s \geq j$) to the limit $(2j-1)!!$, which is the $(2j)$th moment of a standard Gaussian. This shows that a standard Gaussian and a uniformly random ± 1 bit both satisfy all of the above moment conditions.

Theorem 3.2. *Let $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ be a sequence independent real random variables satisfying the s-Moment Conditions. Let $f : \{-1, 1\}^n \to \mathbb{R}$, $s \in \mathbb{N}^+$, and $0 \leq \rho \leq \sqrt{1/(2s-1)}$. Then $\|T_\rho f(\boldsymbol{x})\|_{2s} \leq \|f(\boldsymbol{x})\|_2$.*

Proof. It is well-known that the hypercontractive inequality tensorizes [KS88] and so it suffices to treat the case $n = 1$. By homogeneity we may also assume $\mathbf{E}[f] = 1$; we thus write $f(\boldsymbol{x}_1) = 1 + \epsilon \boldsymbol{x}_1$ for some $\epsilon \in \mathbb{R}$. Raising both sides of the inequality to the $(2s)^{\text{th}}$ power and using the odd moment conditions ($\mathbf{E}[\boldsymbol{x}_1^{2j-1}] = 0$ for all integers $1 \leq j \leq s$), we have

$$\|T_\rho f(\boldsymbol{x}_1)\|_{2s}^{2s} = \sum_{j=0}^{s} \binom{2s}{2j} \rho^{2j} \epsilon^{2j}\, \mathbf{E}[x_1^{2j}] \tag{1}$$

$$\|f(\boldsymbol{x}_1)\|_2^{2s} = \sum_{j=0}^{s} \binom{s}{j} \epsilon^{2j}. \tag{2}$$

By the even moment conditions $\mathbf{E}[\boldsymbol{x}_1^{2j}] \leq (2s-1)^j \binom{s}{j} / \binom{2s}{2j}$, each summand in (1) is at most the corresponding term in (2) and the proof is complete. □

By considering $\epsilon \to 0$ in (1) and (2) it is easy to see for each $j = 1, 2, \ldots, s$ in turn that the associated s-moment condition cannot be further relaxed.

Our SOS extension of this result requires the following lemma:

Lemma 3.3. *Let v be an even positive integer and let $G_1, \ldots, G_v, H_1, \ldots, H_v$ be indeterminates. Then*

$$\vdash_{2v} \quad \prod_{i=1}^{v} G_i H_i \le \frac{1}{\binom{v}{v/2}} \sum_{\substack{T \subset [v] \\ |T| = v/2}} \prod_{i \in T} G_i^2 \prod_{i \in [v] \setminus T} H_i^2.$$

Proof. The non-SOS proof would be to just apply the AM-GM inequality. For the SOS proof we first trivially write

$$\prod_{i \in [v]} G_i H_i = \frac{1}{\binom{v}{v/2}} \sum_{\substack{T \subseteq V \\ |T| = v/2}} \left(\prod_{i \in T} G_i \prod_{i \in [v] \setminus T} H_i \right) \left(\prod_{i \in [v] \setminus T} G_i \prod_{i \in T} H_i \right).$$

We then apply the fact that $\vdash_2 XY \le \frac{1}{2} X^2 + \frac{1}{2} Y^2$ to each summand to deduce

$$\begin{aligned} \vdash_{2v} \quad \prod_{i=1}^{v} G_i H_i &\le \frac{1}{2\binom{v}{v/2}} \sum_{\substack{T \subseteq [v] \\ |T| = v/2}} \prod_{i \in T} G_i^2 \prod_{i \in [t] \setminus T} H_i^2 + \frac{1}{2\binom{v}{v/2}} \sum_{\substack{T \subseteq [v] \\ |T| = v/2}} \prod_{i \in [v] \setminus T} G_i^2 \prod_{i \in T} H_i^2 \\ &= \frac{1}{\binom{v}{v/2}} \sum_{\substack{T \subset [v] \\ |T| = v/2}} \prod_{i \in T} G_i^2 \prod_{i \in [v] \setminus T} H_i^2. \end{aligned}$$

□

We are now ready to state and prove the full version of Theorem 1.4.

Theorem 3.4. *Fix $s \in \mathbb{N}^+$ and write $q = 2s$. Let $0 \le \rho \le \frac{1}{\sqrt{q-1}}$. Let $n \in \mathbb{N}$ and for each $1 \le i \le s$ and each $S \subseteq [n]$, introduce an indeterminate $\widehat{f}_i(S)$. For each $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we write*

$$f_i(x) = \sum_{S \subseteq [n]} \widehat{f}_i(S) \prod_{j \in S} x_i, \quad T_\rho f_i(x) = \sum_{S \subseteq [n]} \rho^{|S|} \widehat{f}_i(S) \prod_{j \in S} x_i.$$

Let $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ be a sequence independent real random variables satisfying the s-Moment Conditions. Then

$$\vdash_q \quad \mathbf{E}\left[\prod_{i=1}^{s} (T_\rho f_i(\boldsymbol{x}))^2 \right] \le \prod_{i=1}^{s} \mathbf{E}[f_i(\boldsymbol{x})^2]. \tag{3}$$

Proof. We prove (3) by induction on n. The base case, $n = 0$, is trivial. For general $n \ge 1$, we can decompose each $f_i(x)$ as

$$f_i(x_1, \ldots, x_n) = x_n g_i(x_1, \ldots, x_{n-1}) + h_i(x_1, \ldots, x_{n-1}).$$

Formally, this means introducing the shorthand $h_i(x_1, \ldots, x_{n-1}) = \sum_{S \not\ni n} \widehat{f}_i(S) \prod_{j \in S} x_i$, and similarly for g_i. We also introduce the notation $\boldsymbol{F}_i = f_i(\boldsymbol{x})$, $\widetilde{\boldsymbol{F}}_i = T_\rho f_i(\boldsymbol{x})$ for each i, and similarly $\boldsymbol{G}_i, \widetilde{\boldsymbol{G}}_i, \boldsymbol{H}_i, \widetilde{\boldsymbol{H}}_i$. Note that these latter four do not depend on $\boldsymbol{x}_n$. By definition we have $\widetilde{\boldsymbol{F}}_i = \rho \boldsymbol{x}_n \widetilde{\boldsymbol{G}}_i + \widetilde{\boldsymbol{H}}_i$.

Using the fact that $\boldsymbol{x}_n$ is independent of all $\boldsymbol{G}_i$, $\boldsymbol{H}_i$ and has zero odd moments, the left-hand side of (3) can be written as follows:

$$\begin{aligned} &\mathbf{E}\left[\prod_{i=1}^{s} \left(\rho^2 \boldsymbol{x}_n^2 \widetilde{\boldsymbol{G}}_i^2 + 2\rho \boldsymbol{x}_n \widetilde{\boldsymbol{G}}_i \widetilde{\boldsymbol{H}}_i + \widetilde{\boldsymbol{H}}_i^2 \right) \right] \\ &= \sum_{\substack{\text{partitions} \\ (U,V,W) \text{ of } [s]}} \rho^{2|U|+|V|} 2^{|V|} \, \mathbf{E}\left[\boldsymbol{x}_n^{2|U|+|V|} \right] \mathbf{E}\left[\prod_{i \in U} \widetilde{\boldsymbol{G}}_i^2 \prod_{i \in V} \widetilde{\boldsymbol{G}}_i \widetilde{\boldsymbol{H}}_i \prod_{i \in W} \widetilde{\boldsymbol{H}}_i^2 \right] \\ &= \sum_{u=0}^{s} \sum_{\substack{v=0 \\ v \text{ even}}}^{s-u} \rho^{2u+v} 2^v \, \mathbf{E}[\boldsymbol{x}_n^{2u+v}] \sum_{\substack{(U,V,W) \\ |U|=u \\ |V|=v}} \mathbf{E}\left[\prod_{i \in U} \widetilde{\boldsymbol{G}}_i^2 \prod_{i \in W} \widetilde{\boldsymbol{H}}_i^2 \prod_{i \in V} \widetilde{\boldsymbol{G}}_i \widetilde{\boldsymbol{H}}_i \right]. \end{aligned} \tag{4}$$

We apply Lemma 3.3 to each $\prod_{i\in V}\widetilde{\boldsymbol{G}}_i\widetilde{\boldsymbol{H}}_i$ (notice that each is multiplied against an SOS polynomial) to obtain

$$\vdash_q \quad (4) \le \sum_{u=0}^{s}\sum_{\substack{v=0\\v\text{ even}}}^{s-u}\frac{\rho^{2u+v}2^v}{\binom{v}{v/2}}\mathbf{E}[\boldsymbol{x}_n^{2u+v}]\sum_{\substack{(U,V,W)\\|U|=u\\|V|=v}}\sum_{\substack{T\subseteq V\\|T|=v/2}}\mathbf{E}\left[\prod_{i\in U\cup T}\widetilde{\boldsymbol{G}}_i^2\prod_{i\in W\cup(V\setminus T)}\widetilde{\boldsymbol{H}}_i^2\right]$$

$$\le \sum_{u=0}^{s}\sum_{\substack{v=0\\v\text{ even}}}^{s-u}\frac{2^v}{\binom{v}{v/2}}\frac{\binom{s}{u+v/2}}{\binom{2s}{2u+v}}\sum_{\substack{(U,V,W)\\|U|=u\\|V|=v}}\sum_{\substack{T\subseteq V\\|T|=v/2}}\mathbf{E}\left[\prod_{i\in U\cup T}\widetilde{\boldsymbol{G}}_i^2\prod_{i\in W\cup(V\setminus T)}\widetilde{\boldsymbol{H}}_i^2\right]$$

$$\le \sum_{u=0}^{s}\sum_{\substack{v=0\\v\text{ even}}}^{s-u}\frac{2^v}{\binom{v}{v/2}}\frac{\binom{s}{u+v/2}}{\binom{2s}{2u+v}}\sum_{\substack{(U,V,W)\\|U|=u\\|V|=v}}\sum_{\substack{T\subseteq V\\|T|=v/2}}\prod_{i\in U\cup T}\mathbf{E}[\boldsymbol{G}_i^2]\prod_{i\in W\cup(V\setminus T)}\mathbf{E}[\boldsymbol{H}_i^2], \tag{5}$$

where the second inequality uses the s-Moments Condition and the bound on ρ, and the third inequality uses the induction hypothesis. (Again, note that each inequality is multiplied against an SOS polynomial.) It is easy to check that $\mathbf{E}[\boldsymbol{F}_i^2] = \mathbf{E}[\boldsymbol{G}_i^2] + \mathbf{E}[\boldsymbol{H}_i^2]$ and so the right-hand side of (3) is simply

$$\sum_{R\subseteq[s]}\prod_{i\in R}\mathbf{E}[\boldsymbol{G}_i^2]\prod_{i\in[s]\setminus R}\mathbf{E}[\boldsymbol{H}_i^2].$$

Thus to complete the inductive proof, it suffices to show that for each $R \subseteq [s]$, the coefficient on $\prod_{i\in R}\mathbf{E}[\boldsymbol{G}_i^2]\prod_{i\in[s]\setminus R}\mathbf{E}[\boldsymbol{H}_i^2]$ in (5) is equal to 1. By symmetry, and taking the sum over v first in (5), it suffices to check that for each $r = |R| = |U \cup T| \in \{0, 1, \ldots, s\}$ we have

$$\sum_{v'=0}^{r}\frac{2^{2v'}}{\binom{2v'}{v'}}\frac{\binom{s}{r}}{\binom{2s}{2r}}\binom{r}{v'}\binom{s-r}{v'} = 1 \quad\Leftrightarrow\quad \sum_{v'=0}^{r}\frac{2^{2v'}}{\binom{2v'}{v'}}\binom{r}{v'}\binom{s-r}{v'} = \frac{\binom{2s}{2r}}{\binom{s}{r}}.$$

This identity can be proved computationally using Zeilberger's algorithm [Zei90, PWZ97]. Denoting the sum on the left side of the left equation by $T(r)$, the algorithm delivers the recurrence equation $T(r+1) - T(r) = 0$, which together with the initial value $T(0) = 1$ yields that $T(r) = 1$ for all r. $\square$

4 The reverse hypercontractive inequality in SOS

This section is devoted to providing a proof Theorem 1.5, the reverse hypercontractivity in the SOS proof system. More precisely:

Theorem 4.1. *Let $k, n \in \mathbb{N}^+$, let $0 \le \rho \le 1 - \frac{1}{2k}$, and let $f(x)$, $g(x)$ be indeterminates for each $x \in \{-1,1\}^n$. Then*

$$\vdash_{4k} \quad \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y})\\ \rho\text{-corr'd}}}[f(\boldsymbol{x})^{2k}g(\boldsymbol{y})^{2k}] \ge \mathbf{E}[f]^{2k}\,\mathbf{E}[g]^{2k}.$$

For each fixed k, we prove Theorem 4.1 by induction on n. The $n = 1$ base case of the induction is the following 4-variable inequality:

Theorem 4.2. *Let $k \in \mathbb{N}^+$ and let $0 \le \rho \le 1 - \frac{1}{2k}$. Let F_0, F_1, G_0, G_1 be real indeterminates. Then*

$$\vdash_{4k} \quad (\tfrac14 + \tfrac14\rho)\Big(F_0^{2k}G_0^{2k} + F_1^{2k}G_1^{2k}\Big) + (\tfrac14 - \tfrac14\rho)\Big(F_0^{2k}G_1^{2k} + F_1^{2k}G_0^{2k}\Big) \ge \Big(\tfrac{F_0+F_1}{2}\Big)^{2k}\Big(\tfrac{G_0+G_1}{2}\Big)^{2k}.$$

Proving this base case will be the key challenge; for now, we give the induction which proves Theorem 4.1.

Proof of Theorem 4.1. Let $n > 1$. Given indeterminates $f(x)$, $g(x)$ for $x \in \{-1,1\}^n$, let $f_0(x)$ be shorthand for $f(x_1, \ldots, x_{n-1}, 1)$, let $f_1(x)$ be shorthand for $f(x_1, \ldots, x_{n-1}, -1)$, and similarly define shorthands g_0, g_1. Now

$$\begin{aligned}\mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y})\\ \rho\text{-corr'd}}}[f(\boldsymbol{x})^{2k}g(\boldsymbol{y})^{2k}] &= (\tfrac14 + \tfrac14\rho)\,\mathbf{E}[f_0(\boldsymbol{x})^{2k}g_0(\boldsymbol{y})^{2k}]\\ &\quad + (\tfrac14 + \tfrac14\rho)\,\mathbf{E}[f_1(\boldsymbol{x})^{2k}g_1(\boldsymbol{y})^{2k}]\\ &\quad + (\tfrac14 - \tfrac14\rho)\,\mathbf{E}[f_0(\boldsymbol{x})^{2k}g_1(\boldsymbol{y})^{2k}]\\ &\quad + (\tfrac14 - \tfrac14\rho)\,\mathbf{E}[f_1(\boldsymbol{x})^{2k}g_0(\boldsymbol{y})^{2k}].\end{aligned}$$

By four applications of induction, we deduce

$$\vdash_{4k} \quad \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y})\\ \rho\text{-corr'd}}}[f(\boldsymbol{x})^{2k} g(\boldsymbol{y})^{2k}] \geq (\tfrac{1}{4} + \tfrac{1}{4}\rho)\,\mathbf{E}[f_0(\boldsymbol{x})]^{2k}\,\mathbf{E}[g_0(\boldsymbol{y})]^{2k} + (\tfrac{1}{4} + \tfrac{1}{4}\rho)\,\mathbf{E}[f_1(\boldsymbol{x})]^{2k}\,\mathbf{E}[g_1(\boldsymbol{y})]^{2k} + (\tfrac{1}{4} - \tfrac{1}{4}\rho)\,\mathbf{E}[f_0(\boldsymbol{x})]^{2k}\,\mathbf{E}[g_1(\boldsymbol{y})]^{2k} + (\tfrac{1}{4} - \tfrac{1}{4}\rho)\,\mathbf{E}[f_1(\boldsymbol{x})]^{2k}\,\mathbf{E}[g_0(\boldsymbol{y})]^{2k}.$$

Now applying the $n = 1$ base case of the induction (Theorem 4.2) to the right-hand side of the above we conclude that

$$\vdash_{4k} \quad \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y})\\ \rho\text{-corr'd}}}[f(\boldsymbol{x})^{2k} g(\boldsymbol{y})^{2k}] \geq \left(\tfrac{\mathbf{E}[f_0(\boldsymbol{x})]+\mathbf{E}[f_1(\boldsymbol{x})]}{2}\right)^{2k} \left(\tfrac{\mathbf{E}[g_0(\boldsymbol{y})]+\mathbf{E}[g_1(\boldsymbol{y})]}{2}\right)^{2k} = \mathbf{E}[f]^{2k}\,\mathbf{E}[g]^{2k}. \qquad \square$$

Our remaining task is to prove the 4-variable base case, Theorem 4.2. Let us make a few simplifications. First, we claim it suffices to prove it in the case $\rho = \rho^* = 1 - \frac{1}{2k}$. To see this, note that

$$(\tfrac{1}{4} + \tfrac{1}{4}\rho)\Big(F_0^{2k}G_0^{2k} + F_1^{2k}G_1^{2k}\Big) + (\tfrac{1}{4} - \tfrac{1}{4}\rho)\Big(F_0^{2k}G_1^{2k} + F_1^{2k}G_0^{2k}\Big) - \Big(\tfrac{F_0+F_1}{2}\Big)^{2k}\Big(\tfrac{G_0+G_1}{2}\Big)^{2k}$$

is linear in ρ. Thus if we can show it is SOS for both $\rho = 0$ and $\rho = \rho^*$, it follows easily that it is SOS for all $0 < \rho < \rho^*$. And for $\rho = 0$ the task is easy:

$$\vdash_{4k} \quad \tfrac{1}{4}\Big(F_0^{2k}G_0^{2k} + F_1^{2k}G_1^{2k}\Big) + \tfrac{1}{4}\Big(F_0^{2k}G_1^{2k} + F_1^{2k}G_0^{2k}\Big) = \Big(\tfrac{F_0^{2k}+F_1^{2k}}{2}\Big)\Big(\tfrac{G_0^{2k}+G_1^{2k}}{2}\Big) \geq \big(\tfrac{F_0+F_1}{2}\big)^{2k}\big(\tfrac{G_0+G_1}{2}\big)^{2k}$$

by Lemma 2.6. Next, for clarity we make a change of variables; our task becomes showing that for real indeterminates μ, ν, α, β,

$$\vdash_{4k} \quad (\tfrac{1}{4} + \tfrac{1}{4}\rho^*)\Big((\mu+\alpha)^{2k}(\nu+\beta)^{2k} + (\mu-\alpha)^{2k}(\nu-\beta)^{2k}\Big) + (\tfrac{1}{4} - \tfrac{1}{4}\rho^*)\Big((\mu+\alpha)^{2k}(\nu-\beta)^{2k} + (\mu-\alpha)^{2k}(\nu+\beta)^{2k}\Big) - \mu^{2k}\nu^{2k} \geq 0. \tag{6}$$

Finally, by homogeneity we can reduce the above to proving the following "two-point inequality":

Two-Point Inequality. *Let $k \in \mathbb{N}^+$ and let $\rho^* = 1 - \frac{1}{2k}$. Then*

$$\vdash_{4k} \quad P_k(a,b) := (\tfrac{1}{4} + \tfrac{1}{4}\rho^*)\Big((1+a)^{2k}(1+b)^{2k} + (1-a)^{2k}(1-b)^{2k}\Big) + (\tfrac{1}{4} - \tfrac{1}{4}\rho^*)\Big((1+a)^{2k}(1-b)^{2k} + (1-a)^{2k}(1+b)^{2k}\Big) - 1 \geq 0.$$

Proof that (6) follows from the Two-Point Inequality. Suppose we show that $P_k(a,b)$ is equal to a sum of squares, say $\sum_{i=1}^m R_i(a,b)^2$ where each $R_i(a,b)$ is a bivariate polynomial. Viewing this as an SOS identity in a only, we deduce that $\deg_a(R_i) \leq k$ for each i; similarly, $\deg_b(R_i) \leq k$ for each i. Then

$$\sum_{i=1}^m (\mu^k \nu^k R_i(\tfrac{\alpha}{\mu}, \tfrac{\beta}{\nu}))^2 = \mu^{2k}\nu^{2k} \sum_{i=1}^m R_i(\tfrac{\alpha}{\mu}, \tfrac{\beta}{\nu})^2 = \mathrm{LHS}(6),$$

and in the summation each expression $\mu^k\nu^k R_i(\frac{\alpha}{\mu}, \frac{\beta}{\nu})$ is a polynomial in μ, ν, α, β. $\square$

It remains to establish the Two-Point Inequality via an SOS proof.

Remark 4.3. We remind the reader that there is of course a "ZFC" proof of the Two-Point Inequality, since it follows as a special case of the reverse hypercontractive inequality.

4.1 The Two-Point Inequality in SOS

This section is devoted to proving the Two-Point Inequality; i.e., showing $P_k(a,b)$ is SOS. The crucial idea turns out to be rewriting it under the following substitutions:

$$r = a - b, \qquad s = a + b, \qquad t = ab.$$

We may then express

$$P_k(a,b) = -1 + \left(\tfrac{1}{4}+\tfrac{1}{4}\rho^*\right)\left((1+t+s)^{2k}+(1+t-s)^{2k}\right) + \left(\tfrac{1}{4}-\tfrac{1}{4}\rho^*\right)\left((1-t+r)^{2k}+(1-t-r)^{2k}\right)$$
$$= -1 + \left(\tfrac{1}{2}+\tfrac{1}{2}\rho^*\right)\sum_{i=0}^{k}\tbinom{2k}{2i}(1+t)^{2k-2i}s^{2j} + \left(\tfrac{1}{2}-\tfrac{1}{2}\rho^*\right)\sum_{j=0}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}r^{2j},$$

where we used the identity

$$\tfrac{1}{2}\left((c+d)^{2k}+(c-d)^{2k}\right) = \sum_{i=0}^{k}\tbinom{2k}{2i}c^{2k-2i}d^{2i}.$$

Next we use $r^2 = s^2 - 4t$ to eliminate r, obtaining

$$P_k(a,b) = -1 + \left(\tfrac{1}{2}+\tfrac{1}{2}\rho^*\right)\sum_{i=0}^{k}\tbinom{2k}{2i}(1+t)^{2k-2i}s^{2i} + \left(\tfrac{1}{2}-\tfrac{1}{2}\rho^*\right)\sum_{j=0}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}(s^2-4t)^j.$$

Now we expand $(s^2-4t)^j$ in the latter sum so that we can write it as an even polynomial in s. We get

$$\sum_{j=0}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}(s^2-4t)^j = \sum_{j=0}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}\sum_{i=0}^{j}\tbinom{j}{i}s^{2i}(-4t)^{j-i}$$
$$= \sum_{i=0}^{k}s^{2i}\sum_{j=i}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}\tbinom{j}{i}(-4t)^{j-i}.$$

Thus we have

$$P_k(a,b) = -1 + \sum_{i=0}^{k}\left(\left(\tfrac{1}{2}+\tfrac{1}{2}\rho^*\right)\tbinom{2k}{2i}(1+t)^{2k-2i} + \left(\tfrac{1}{2}-\tfrac{1}{2}\rho^*\right)\sum_{j=i}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}\tbinom{j}{i}(-4t)^{j-i}\right)s^{2i}$$
$$= Q_{k,0}(t) + Q_{k,1}(t)s^2 + Q_{k,2}(t)s^4 + \cdots + Q_{k,k}(t)s^{2k}, \tag{7}$$

where

$$Q_{k,0}(t) = -1 + \left(\tfrac{1}{2}+\tfrac{1}{2}\rho^*\right)(1+t)^{2k} + \left(\tfrac{1}{2}-\tfrac{1}{2}\rho^*\right)\sum_{j=0}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}(-4t)^j,$$

$$Q_{k,i}(t) = \left(\tfrac{1}{2}+\tfrac{1}{2}\rho^*\right)\tbinom{2k}{2i}(1+t)^{2k-2i} + \left(\tfrac{1}{2}-\tfrac{1}{2}\rho^*\right)\sum_{j=i}^{k}\tbinom{2k}{2j}(1-t)^{2k-2j}\tbinom{j}{i}(-4t)^{j-i}, \qquad i = 1\ldots k.$$

Suppose we could show that $Q_{k,0}(t)$ and also $Q_{k,1}(t),\ldots,Q_{k,k}(t)$ are nonnegative for all $t \in \mathbb{R}$. Then by Fact 2.3 they are also SOS, and hence $P_k(a,b)$ is SOS in light of (7). This would complete the proof of the Two-Point Inequality.

In fact that is precisely what we show below, using some computer algebra assistance. We remark, though, that is not a priori clear that this strategy should work; i.e., that $Q_{k,0}(t),\ldots,Q_{k,k}(t)$ should be nonnegative. It does not follow from the truth of the Two-Point Inequality. To see this, observe that whereas the Two-Point Inequality is known to hold for any $0 \le \rho \le \rho^*$, it is *not* true that $Q_{k,0}(t) \ge 0$ for all $0 \le \rho \le \rho^*$. In fact, for $k = 1$ we have

$$Q_{1,0}(t) = t^2 - (2-4\rho^*)t \tag{8}$$

which is nonnegative for all t *only* for the specific choice $\rho^* = 1 - \frac{1}{2k} = \frac{1}{2}$.

Nevertheless, we now complete the proof of the Two-Point Inequality by showing that $Q_{k,0}(t),\ldots,Q_{k,k}(t)$ are all nonnegative.

Proposition 4.4. *For each $k \in \mathbb{N}^+$ (with $\rho^* = 1 - \frac{1}{2k}$), the polynomial $Q_{k,0}(t)$ is nonnegative.*

Proof. For $k = 1$ we have $Q_{1,0}(t) = t^2$ (as noted in (8)); henceforth we may assume $k \ge 2$. For $t < 0$ we substitute $a = \sqrt{-t}$, $b = -\sqrt{-t}$ into (7); since $s = a + b = 0$ we get $P_k(\sqrt{-t}, -\sqrt{-t}) = Q_{k,0}(t)$. By Remark 4.3 we have $P_k(\sqrt{-t}, -\sqrt{-t}) \ge 0$ and hence $Q_{k,0}(t) \ge 0$ for all $t < 0$.

For $t \geq 0$ we first rewrite

$$Q_{k,0}(t) = -1 + (1+t)^{2k} \left(1 - \frac{1}{4k} + \frac{1}{4k} \sum_{j=0}^{k} \binom{2k}{2j} \frac{(1-t)^{2k-2j}}{(1+t)^{2k}} (-4t)^j \right).$$

Denoting the sum in this expression by $S_k(t)$, Zeilberger's algorithm [Zei90, PWZ97] finds the recurrence equation

$$(t+1)^2 S_{k+2}(t) - 2(t^2 - 6t + 1) S_{k+1}(t) + (t+1)^2 S_k(t) = 0,$$

valid for all $k \geq 0$. Since the coefficients in this recurrence do not depend on k but only on t, the recurrence can be solved in closed form. Together with the initial values $S(0) = 1$ and $S(1) = \frac{t^2-6t+1}{(t+1)^2}$, it follows that $S(t) = \cos(4k \arctan(\sqrt{t}))$. (Not every computer algebra system may deliver the solution in this form; however, for the correctness of the proof it is sufficient to check that $\cos(4k \arctan(\sqrt{t}))$ is indeed a solution of the recurrence. This is easy to verify.) Hence,

$$\begin{aligned} Q_{k,0}(t) &= -1 + (1+t)^{2k} \left(1 - \tfrac{1}{4k} + \tfrac{1}{4k} \cos(4k \arctan(\sqrt{t})) \right) \\ &\geq -1 + (1 + 2kt) \left(1 - \tfrac{1}{4k} + \tfrac{1}{4k} \cos(4k \arctan(\sqrt{t})) \right), \end{aligned}$$

using the fact that the parenthesized expression is clearly nonnegative. We now split into two cases.

Case 1: $t \geq \frac{1}{2k(2k-1)}$. In this case we simply use that $\cos(4k \arctan(\sqrt{t})) \geq -1$ to obtain

$$Q_{k,0}(t) \geq -1 + (1 + 2kt) \left(1 - \tfrac{1}{2k}\right) = -\tfrac{1}{2k} + (2k-1)t,$$

which is indeed nonnegative when $t \geq \frac{1}{2k(2k-1)}$.

Case 2: $0 \leq t \leq \frac{1}{2k(2k-1)}$. In this case we use the following estimates:

$$\arctan(\sqrt{t}) \leq \sqrt{t} \quad \forall t \geq 0, \qquad \cos(x) \geq \kappa(x) := 1 - \tfrac{1}{2}x^2 + \tfrac{1}{24}x^4 - \tfrac{1}{720}x^6 \quad \forall x \in \mathbb{R}.$$

Note that $4k \arctan(\sqrt{t}) \leq 4k\sqrt{t} \leq 4k\sqrt{\frac{1}{2k(2k-1)}}$, and the latter quantity is at most $\sqrt{2}$ for all $k \geq 2$. Since $\cos(x)$ is decreasing for $x \in [0, \sqrt{2}]$ we have

$$\begin{aligned} &\cos(4k \arctan(\sqrt{t})) \geq \cos(4k\sqrt{t}) \geq \kappa(4k\sqrt{t}) \\ &\qquad \Rightarrow \quad Q_{k,0}(t) \geq -1 + (1 + 2kt) \left(1 - \tfrac{1}{4k} + \tfrac{1}{4k} \kappa(4k\sqrt{t}) \right) \\ &\qquad\qquad = -1 + (1 + 2kt) \left(1 - \tfrac{1}{4k} + \tfrac{1}{4k} \left(1 - 8k^2 t + \tfrac{32}{3} k^4 t^2 - \tfrac{256}{45} k^6 t^3\right)\right) \\ &\qquad\qquad = q(t) t^2, \qquad \text{where } q(t) = -\tfrac{128}{45} k^6 t^2 - \left(\tfrac{64}{45} k - \tfrac{16}{3}\right) k^4 t + \left(\tfrac{8}{3} k - 4\right) k^2. \end{aligned}$$

It remains to show that $q(t) \geq 0$ for $0 \leq t \leq \frac{1}{2k(2k-1)}$. Since $q(t)$ is a quadratic polynomial with negative leading coefficient, we only need to check that $q(0),\ q(\frac{1}{2k(2k-1)}) \geq 0$. We have $q(0) = \left(\frac{8}{3}k - 4\right) k^2$, which is clearly nonnegative for $k \geq 2$. Finally, one may check that

$$q\left(\tfrac{1}{2k(2k-1)}\right) = \tfrac{4k^2}{45(2k-1)^2} \left(19 + 136(k-2) \left((k - \tfrac{13}{34})^2 + \tfrac{103}{1156}\right)\right),$$

which is evidently nonnegative for $k \geq 2$. ☐

Proposition 4.5. *For all $1 \leq i \leq k \in \mathbb{N}^+$, the polynomial $Q_{k,i}(t)$ is nonnegative (with $\rho^* = 1 - \frac{1}{2k}$).*

Proof. In fact, we will prove the stronger claim that each $Q_{k,i}(t)$ is nonnegative even when ρ^* is set to 0. I.e., we will show that

$$\widetilde{Q}_{k,i}(t) := \tfrac{1}{2}\binom{2k}{2i}(1+t)^{2k-2i} + \tfrac{1}{2} \sum_{j=i}^{k} \binom{2k}{2j} (1-t)^{2k-2j} \binom{j}{i} (-4t)^{j-i}$$

is nonnegative. To see that this is indeed stronger, simply note that $Q_{k,i}(t)$ and $\widetilde{Q}_{k,i}(t)$ are convex combinations of the same two main quantities, but $\widetilde{Q}_{k,i}(t)$ has less of its "weight" on the first quantity $\binom{2k}{2i}(1+t)^{2k-2i}$, which is clearly nonnegative. We will furthermore show that even $\widetilde{Q}_{k,0}(t) \geq 0$.

One may check that the following recurrence holds for all integers $0 \le i \le k$:

$$(1+i)(1+k)\widetilde{Q}_{k+2,i+1}(t) = (1+i)(2+k)(1+t)^2\widetilde{Q}_{k+1,i+1}(t) + (2+k)(2+2k-i)\widetilde{Q}_{k+1,i}(t).$$

(This was found by guessing the form of a polynomial recurrence and then solving via computer.) Therefore we only need to prove $\widetilde{Q}_{k,i}(t) \ge 0$ for the cases that $k = i$ and $i = 0$; the nonnegativity of $\widetilde{Q}_{k,i}(t)$ for general k and i then follows by induction. For $k = i$ we have $\widetilde{Q}_{k,k}(t) = 1 \ge 0$. For $i = 0$ the proof of nonnegativity is similar to, but easier than, that of Proposition 4.4. For $t < 0$ it's obvious from its definition that $\widetilde{Q}_{k,0}(t)$ is nonnegative. For $t \ge 0$, the proof of Proposition 4.4 gives

$$\widetilde{Q}_{k,0}(t) = \tfrac{1}{2}(1+t)^{2k}\left(1+\cos(4k\arctan(\sqrt{t}))\right) \ge 0. \qquad \square$$

5 Vertex-Cover and the Frankl-Rödl Theorem in SOS

It will be slightly more convenient to work with the maximum independent set (Max-IS) problem, rather than the minimum vertex cover problem; there is an equivalence between the problems because the complement of a vertex cover is an independent set, and vice versa. We will consider the fractional size of independent sets.

Definition 5.1. Given a graph $G = (V, E)$ we define

$$\textsf{Max-IS}(G) = \max\{|S|/|V| : S \subseteq V \text{ is an independent set}\} \in [0,1],$$

where S is said to be an independent set if $E \cap (S \times S) = \emptyset$.

We repeat the Frankl–Rödl Theorem with this notation:

Theorem 5.2. *There exists $\gamma_0 > 0$ such that for all $\gamma \le \gamma_0$ it holds that $\textsf{Max-IS}(G^n_\gamma) \le n(1-\gamma^2/64)^n$. In particular, $\textsf{Max-IS}(G^n_\gamma) \le o_n(1)$ for $.1\sqrt{\frac{\log n}{n}} \le \gamma \le \gamma_0$.*

The main goal of this section is to give a low-degree proof of "$\textsf{Max-IS}(G^n_\gamma) \le o_n(1)$" (for all constant $\gamma > 0$) in the SOS proof system. We state our theorem as follows.

Theorem 5.3. *Let $n \in \mathbb{N}^+$ and let $\frac{1}{\log n} \le \gamma \le \frac{1}{2}$ such that $(1-\gamma)n$ is an even integer. Given the Frankl–Rödl graph $G^n_\gamma = (V, E)$, for each $x \in V = \{-1,1\}^n$ let $f(x)$ be an indeterminate. Then there is a degree-$O(1/\gamma)$ SOS refutation of the system expressing the statement that $\textsf{Max-IS}(G) \ge O(n^{-\gamma/10})$; i.e.,*

$$\{f(x)^2 = f(x)\ \forall x \in V,\quad f(x)f(y) = 0\ \forall (x,y) \in E,\quad \tfrac{1}{|V|}\sum_{x \in V} f(x) \ge Cn^{-\gamma/10}\} \vdash_{O(1/\gamma)} \quad -1 \ge 0$$

for a universal constant C.

This theorem implies that the $N^{O(1/\gamma)}$-time SOS/Lasserre SDP hierarchy algorithm certifies that $\textsf{Max-IS}(G^n_\gamma) \le O(n^{-\gamma/10})$; in other words, that $\textsf{Min-Vertex-Cover}(G^n_\gamma) \ge (1 - O(n^{-\gamma/10}))N = (1 - O(n^{-\gamma/10}))2^n$. Note that our bound is nontrivial only for $\gamma \gg \frac{1}{\log n}$.

We prove Theorem 5.3 by "SOS-izing" the Benabbas–Hatami–Magen Fourier-theoretic proof [BHM12] of the following "density" version of the Frankl–Rödl Theorem:

Theorem 5.4. *([BHM12]) Fix $0 < \gamma < 1/2$ and $0 < \alpha \le 1$. In the graph $G^n_\gamma = (V, E)$, if $S \subseteq V$ has $|S|/2^n \ge \alpha$ then*

$$\Pr_{(\boldsymbol{x},\boldsymbol{y})\sim E}[\boldsymbol{x} \in S, \boldsymbol{y} \in S] \ge 2(\alpha/2)^{1/\gamma} - o_n(1).$$

Here the $o_n(1)$ goes to 0 rather slowly in n, which means that the Benabbas–Hatami–Magen proof only recovers the Frankl–Rödl Theorem for $\gamma > \omega(\frac{1}{\log n})$.

5.1 The Benabbas–Hatami–Magen argument in SOS

Benabbas, Hatami, and Magen [BHM12] introduce the following operator:

Definition 5.5. For integer $0 \le d \le n$ the operator S_d is defined on functions $f : \{-1,1\}^n \to \mathbb{R}$ by $S_d f(x) = \mathbf{E}_{\boldsymbol{y}}[f(\boldsymbol{y})]$, where $\boldsymbol{y}$ is a chosen uniformly at random subject to $\Delta(x, \boldsymbol{y}) = d$.

The key technical contribution of [BHM12] is showing how to pass between the S_d operators (which are relevant for Frankl–Rödl analysis) and the T_ρ operators (for which we have reverse hypercontractivity). Intuitively, the operators S_d and $T_{1-2d/n}$ should be similar (at least if d/n is bounded away from 0 and 1). However there is one caveat: "parity" issues with S_d. For example, if $f : \{-1,1\}^n \to \{0,1\}$ is the indicator of the strings of even Hamming weight, then

$$\langle f, S_d f\rangle = \begin{cases} 0 & \text{if } d \text{ is odd,} \\ \frac{1}{2} & \text{if } d \text{ is even;} \end{cases} \qquad \text{but, } \langle f, T_{1-2d/n} f\rangle \approx \tfrac{1}{4} \text{ for } d \text{ odd or even.}$$

Benabbas, Hatami, and Magen evade this parity issue by considering the operator $\frac{1}{2}S_d + \frac{1}{2}S_{d+1}$.

Definition 5.6. For integer $0 \le d < n$, we define the operator $S'_d = \frac{1}{2}S_d + \frac{1}{2}S_{d+1}$.

The crucial theorem in [BHM12]'s work is the following:

Theorem 5.7. *(Follows from Lemma 3.4 in [BHM12].) Let $f : \{-1,1\}^n \to \mathbb{R}$. Let $d = n - c$ for some integer $e^2\sqrt{n} \le c \le n/2$. Then for $\rho = 1 - 2d/n$,*

$$\langle f, S'_d f\rangle - \langle f, T_\rho f\rangle = \sum_{U \subseteq [n]} \widehat{f}(U)^2 \cdot \delta(U),$$

where each real number $\delta(U)$ satisfies

$$|\delta(U)| \le O(\max\{n^{-1/5}, \tfrac{n}{c^2}\log^2(\tfrac{c^2}{n})\}).$$

Given this Theorem 5.7, Benabbas, Hatami, and Magen are able to deduce their main Theorem 5.4 from the reverse hypercontractivity result Theorem 1.3 without too much trouble. We now show that this deduction can also be carried out in the SOS proof system. Specifically, we give here the proof of our Theorem 5.3, relying on the SOS proof of hypercontractivity (Theorem 4.1) from Section 4.

Proof of Theorem 5.3. Write $d = (1-\gamma)n$ (where $\frac{1}{\log n} \le \gamma \le \frac{1}{2}$) and write $\rho' = 1 - 2d/n = -(1-2\gamma)$. For $i = 0, 1$ let us denote

$$f_i(x) = \begin{cases} f(x) & \text{if } x\text{'s Hamming weight equals } i \bmod 2, \\ 0 & \text{else.} \end{cases}$$

We have

$$\{f(x)f(y) = 0\ \forall \Delta(x,y) = d\} \quad \vdash_2 \quad \langle f_0, S'_d f_0\rangle + \langle f_1, S'_d f_1\rangle = 0$$

because if x's Hamming weight has the same parity as y's then their distance can only be d (an even integer) not $d+1$ (an odd one). Using Theorem 5.7 it follows that

$$\begin{aligned}\{f(x)f(y) = 0\ \forall \Delta(x,y) = d\} \quad &\vdash_2 \\ &\langle f_0, T_{\rho'} f_0\rangle + \langle f_1, T_{\rho'} f_1\rangle \le \delta \sum_U \widehat{f}_0(U)^2 + \delta \sum_U \widehat{f}_1(U)^2 = \delta(\mathbf{E}[f_0^2] + \mathbf{E}[f_1^2]) = \delta\, \mathbf{E}[f^2].\end{aligned}$$

where

$$\delta = O(\max\{n^{-1/5}, \tfrac{1}{\gamma^2 n}\log^2(\gamma^2 n)\}) = O(n^{-1/5})$$

(with the second bound using $\gamma \ge \frac{1}{\log n}$.) We now write $g_i(x)$ to denote $f_i(-x)$ and also $\rho = -\rho' = 1-2\gamma$; then

$$\langle f_i, T_{\rho'} f_i\rangle = \langle f_i, T_\rho g_i\rangle = \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y}) \\ \rho\text{-corr'd}}}[f_i(\boldsymbol{x})g_i(\boldsymbol{y})]$$

so we conclude

$$\{f(x)f(y) = 0\ \forall \Delta(x,y) = d\} \quad \vdash_2 \quad \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y}) \\ \rho\text{-corr'd}}}[f_0(\boldsymbol{x})g_0(\boldsymbol{y})] + \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y}) \\ \rho\text{-corr'd}}}[f_1(\boldsymbol{x})g_1(\boldsymbol{y})] \le \delta\, \mathbf{E}[f^2].$$

Next, let $k = \lceil \frac{1}{4\gamma} \rceil$. We have $f(x)^2 = f(x) \vdash_{2k} f(x)^{2k} = f(x)$, from which we may easily deduce

$$\begin{aligned}\{f(x)^2 = f(x)\ \forall x, \quad f(x)f(y) = 0\ \forall \Delta(x,y) = d\}& \\ \vdash_{2k} \quad \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y}) \\ \rho\text{-corr'd}}}[f_0(\boldsymbol{x})^{2k} g_0(\boldsymbol{y})^{2k}] &+ \mathop{\mathbf{E}}_{\substack{(\boldsymbol{x},\boldsymbol{y}) \\ \rho\text{-corr'd}}}[f_1(\boldsymbol{x})^{2k} g_1(\boldsymbol{y})^{2k}] \le \delta\, \mathbf{E}[f].\end{aligned}$$

Since $\rho = 1 - 2\gamma \leq 1 - \frac{1}{2k}$ we may apply our reverse hypercontractivity result Theorem 4.1 (in Section 4) to deduce

$$\{f(x)^2 = f(x)\ \forall x, \quad f(x)f(y) = 0\ \forall \Delta(x,y) = d\} \quad \vdash_{2k} \quad \mathbf{E}[f_0]^{2k}\,\mathbf{E}[g_0]^{2k} + \mathbf{E}[f_1]^{2k}\,\mathbf{E}[g_1]^{2k} \leq \delta\,\mathbf{E}[f].$$

We're now almost done. First, $\mathbf{E}[f_i] = \mathbf{E}[g_i]$ formally for each $i = 0, 1$. Second, for simplicity we use the bound

$$\{f(x)^2 = f(x)\ \forall x\} \quad \vdash_2 \quad \delta\,\mathbf{E}[f] = \delta\,\mathbf{E}[2f - f^2] = \delta\,\mathbf{E}[1 - (1-f)^2] \leq \delta.$$

Thus we have

$$\begin{aligned} \{f(x)^2 = f(x)\ \forall x, \quad f(x)f(y) = 0\ \forall \Delta(x,y) = d\} \quad \vdash_{2k} \quad & \delta \geq \mathbf{E}[f_0]^{4k} + \mathbf{E}[f_1]^{4k} \\ & \geq 2\left(\frac{\mathbf{E}[f_0] + \mathbf{E}[f_1]}{2}\right)^{4k} \\ & = 2(\mathbf{E}[f]/2)^{4k} \\ \Rightarrow \quad & 2^{4k-1}\delta \geq \mathbf{E}[f]^{4k}, \end{aligned}$$

where the second inequality is Lemma 2.6. Finally, from Lemma 2.5 we may deduce

$$\mathbf{E}[f] \geq Cn^{-\gamma/10} \quad \vdash_{4k} \quad \mathbf{E}[f]^{4k} \geq C^{4k} n^{-4k\gamma/10} \geq C^{4k} n^{-\gamma/5} \geq 2^{4k}\delta$$

for C sufficiently large. By combining the previous two statements we can get

$$\{f(x)^2 = f(x)\ \forall x, \quad f(x)f(y) = 0\ \forall \Delta(x,y) = d, \quad \mathbf{E}[f] \geq Cn^{-\gamma/10}\} \quad \vdash_{4k} \quad -1 \geq 0,$$

as required. $\square$

6 Conclusions

We describe here a few questions left open by our work. Regarding reverse hypercontractivity, it seems we may not have given the Book proof of the SOS Two-Point Inequality. We would be happy to see a more elegant "human proof", but even more interesting would be a computer algebra technique that could automatically prove SOS-ness, symbolically for all k.

An additional open question regarding the Frankl–Rödl Theorem is whether the Benabbas–Hatami–Magen proof can be improved to work for γ as small as $\sqrt{\frac{\log n}{n}}$. However even if this is possible, the resulting SOS proof would (presumably) be of degree $\Omega(\sqrt{\frac{n}{\log n}}) = \Omega(\sqrt{\frac{\log N}{\log\log N}})$, slightly superconstant. Could there be an $O(1)$-degree SOS of the Frankl–Rödl Theorem? An interesting toy version of this question is the following: The vertex isoperimetric inequality for the hypercube immediately implies that if $A, B \subseteq \{-1,1\}^n$ satisfy $\mathrm{dist}(A,B) \geq \sqrt{n \log n}$ then $\frac{|A|}{2^n}\frac{|B|}{2^n} = o_n(1)$. Does this have an $O(1)$-degree SOS proof?

Acknowledgments

The authors would like to thank Siavosh Benabbas and Hamed Hatami for the advance copy of [BHM12]. Proposition 4.5 was independently proven by Fedor Nazarov; we are very grateful to him for sharing his proof with us. Thanks also to Toni Pitassi and Doron Zeilberger for helpful discussions.

References

[ABL02] Sanjeev Arora, Béla Bollobás, and László Lovász. Proving integrality gaps without knowing the linear program. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 313–322, 2002. 1.2.1

[ABLT06] Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis. Proving integrality gaps without knowing the linear program. *Theory of Computing*, 2(1):19–51, 2006. 1.2.1

[ABS10] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for Unique Games and related problems. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science*, pages 563–572, 2010. 1.2.1

[BBH+12] Boaz Barak, Fernando Brandão, Aram Harrow, Jonathan Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing*, pages 307–326, 2012. 1.1, 1.2, 1.2.1

[BCGM11] Siavosh Benabbas, Siu On Chan, Konstantinos Georgiou, and Avner Magen. Tight gaps for Vertex Cover in the Sherali–Adams SDP hierarchy. In *Proceedings of the 32nd Annual IARCS Conference on Foundations of Software Technology and Theoretical Computer Science*, pages 41–54, 2011. 1.2.1, 1.2.1, 1.2.1

[BHM12] Siavosh Benabbas, Hamed Hatami, and Avner Magen. An isoperimetric inequality for the Hamming cube with applications for integrality gaps in degree-bounded graphs. Unpublished, 2012. 1, 1.2.1, 5, 5.4, 5.1, 5.1, 5.1, 5.7, 6

[Bon70] Aline Bonami. Étude des coefficients Fourier des fonctions de $L^p(G)$. *Annales de l'Institute Fourier*, 20(2):335–402, 1970. 1

[Bor82] Christer Borell. Positivity improving operators and hypercontractivity. *Mathematische Zeitschrift*, 180:225–234, 1982. 1

[Cha02] Moses Charikar. On semidefinite programming relaxations for graph coloring and vertex cover. In *Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 616–620, 2002. 1.2.1

[CMM09] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Integrality gaps for Sherali–Adams relaxations. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 283–292. ACM, 2009. 1.2.1

[DKSV06] Nikhil Devanur, Subhash Khot, Rishi Saket, and Nisheeth Vishnoi. Integrality gaps for Sparsest Cut and Minimum Linear Arrangement problems. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 537–546, 2006. 1.1

[DS05] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1):439–485, 2005. 1.2.1

[FKO07] Uriel Feige, Guy Kindler, and Ryan O'Donnell. Understanding parallel repetition requires understanding foams. In *Proceedings of the 22nd Annual IEEE Conference on Computational Complexity*, pages 179–192, 2007. 1

[FO06] Uriel Feige and Eran Ofek. Random 3CNF formulas elude the Lovász theta function. Technical Report cs/0603084, arXiv, 2006. 1.2.1

[FR87] Péter Frankl and Vojtěch Rödl. Forbidden intersections. *Transactions of the American Mathematical Society*, 300(1):259–286, 1987. 1.2.1

[Geo10] Konstantinos Georgiou. *Integrality gaps for strong linear programming and semidefinite programming relaxations*. PhD thesis, University of Toronto, 2010. 1.2.1

[GJ79] Michael Garey and David Johnson. *Computers and Intractability: a guide to the theory of NP-completeness*. W. H. Freeman and Company, 1979. 1.2.1

[GM08] Konstantinos Georgiou and Avner Magen. Expansion fools the Sherali–Adams system: compromising local and global arguments. Technical Report CSRG–587, University of Toronto, November 2008. 1.2.1

[GM10] Konstantinos Georgiou and Avner Magen. Tight integrality gap for Sherali–Adams SDPs for Vertex Cover. Manuscript, 2010. 1.2.1

[GMPT10] Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis. Integrality gaps of $2 - o(1)$ for Vertex Cover SDPs in the Lovász–Schrijver hierarchy. *SIAM Journal on Computing*, 39(8):3553–3570, 2010. 1.2.1, 1.2.1, 1.2.1

[GMT08] Konstantinos Georgiou, Avner Magen, and Iannis Tourlakis. Vertex Cover resists SDPs tightened by local hypermetric inequalities. In *Proceedings of the 12th Annual Conference on Integer Programming and Combinatorial Optimization*, IPCO'08, pages 140–153, Berlin, Heidelberg, 2008. Springer-Verlag. 1.2.1

[GV01] Dima Grigoriev and Nicolai Vorobjov. Complexity of Null- and Positivstellensatz proofs. *Annals of Pure and Applied Logic*, 113(1):153–160, 2001. 1.1, 2

[Kel12] Nathan Keller. A tight quantitative version of Arrow's impossibility theorem. *Journal of the European Mathematical Society (JEMS)*, 14(5):1331–1355, 2012. 1

[KG98] Jon Kleinberg and Michel Goemans. The Lovász theta function and a semidefinite programming relaxation of vertex cover. *SIAM Journal on Discrete Mathematics*, 11(2):196–204, 1998. 1.2.1

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on Boolean functions. In *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science*, pages 68–80, 1988. 1

[KR08] Subhash Khot and Oded Regev. Vertex Cover might be hard to approximate to within $2-\epsilon$. *Journal of Computer and System Sciences*, 74(3):335–349, 2008. 1.2.1

[KS88] Wieslaw Krakowiak and Jerzy Szulga. Hypercontraction principle and random multilinear forms. *Probability Theory and Related Fields*, 77(3):325–342, 1988. 1.2, 3

[KS09] Subhash Khot and Rishi Saket. SDP integrality gaps with local ℓ_1-embeddability. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 565–574, 2009. 1.1

[KV05] Subhash Khot and Nisheeth Vishnoi. The Unique Games Conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 53–62, 2005. 1.1

[Las00] Jean Lasserre. Optimisation globale et théorie des moments. *Comptes Rendus de l'Académie des Sciences*, 331(11):929–934, 2000. 1.1

[MOO10] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Annals of Mathematics*, 171(1), 2010. 1, 1, 1.2

[MOR+06] Elchanan Mossel, Ryan O'Donnell, Oded Regev, Jeffrey Steif, and Benjamin Sudakov. Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami–Beckner inequality. *Israel Journal of Mathematics*, 154:299–336, 2006. 1, 1.2

[Mos12a] Elchanan Mossel. A quantitative Arrow theorem. *Probability Theory and Related Fields*, 154(1-2):49–88, 2012. 1

[MOS12b] Elchanan Mossel, Krzysztof Oleszkiewicz, and Arnab Sen. On reverse hypercontractivity. Technical Report 1108.1210, arXiv, 2012. 1

[MR12] Elcahan Mossel and Miklós Rácz. A quantitative Gibbard–Satterthwaite theorem without neutrality. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 1041–1060, 2012. 1

[OZ13] Ryan O'Donnell and Yuan Zhou. Approximability and proof complexity. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2013. 1.1, 1.2, 1.2.1, 2

[Par00] Pablo Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization.* PhD thesis, California Institute of Technology, 2000. 1.1

[PWZ97] Marko Petkovšek, Herbert Wilf, and Doron Zeilberger. $A = B$. AK Peters, Ltd., 1997. 3, 4.1

[RS09] Prasad Raghavendra and David Steurer. Integrality gaps for strong SDP relaxations of Unique Games. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 575–585, 2009. 1.1

[Sch08] Grant Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 593–602, 2008. 1.2.1

[She09] Jonah Sherman. Breaking the multicommodity flow barrier for $O(\sqrt{\log n})$-approximations to Sparsest Cut. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 363–372, 2009. 1

[STT07a] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. A linear round lower bound for Lovász–Schrijver SDP relaxations of Vertex Cover. In *Proceedings of the 22nd Annual IEEE Conference on Computational Complexity*, pages 205–216, 2007. 1.2.1

[STT07b] 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*, pages 302–310. ACM, 2007. 1.2.1

[Tou06] Iannis Tourlakis. New lower bounds for Vertex Cover in the lovász–Schrijver hierarchy. In *Proceedings of the 21st Annual IEEE Conference on Computational Complexity*, pages 170–182. IEEE Computer Society, 2006. 1.2.1

[Tul09] Madhur Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 303–312, 2009. 1.2.1

[Zei90] Doron Zeilberger. A fast algorithm for proving terminating hypergeometric identities. *Discrete Mathematics*, 80(2):207–211, 1990. 3, 4.1

A Solution to the puzzle

$$\begin{aligned} a^6b^6 + 15\,(a+b)^2\,(1+ba)^4 + 10\,(a+b)^4\,(1+ba)^2 + 5\left(a^3b+b^2+a^2+ab^3\right)^2 \\ +\ 35a^4b^4 + \left(a^3+b^3\right)^2 + \tfrac{17}{3}\left(a^2b+ab^2\right)^2 + 35a^2b^2 + \tfrac{148}{3}a^2b^4 + \tfrac{148}{3}a^4b^2. \end{aligned}$$