On the Power of Symmetric LP and SDP Relaxations

James R. Lee
Department of Computer Science and Engineering
University of Washington
Email: jrl@cs.washington.edu

Prasad Raghavendra
Computer Science Division
U. C. Berkeley
Email: prasad@cs.berkeley.edu

David Steurer
Department of Computer Science
Cornell University
Email: dsteurer@cs.cornell.edu

Ning Tan
Computer Science Division
U. C. Berkeley
Email: ntan@eecs.berkeley.edu

***Abstract*—We study the computational power of general symmetric relaxations for combinatorial optimization problems, both in the linear programming (LP) and semidefinite programming (SDP) case. We show new connections to explicit LP and SDP relaxations, like those obtained from standard hierarchies.**

Concretely, for $k < n/4$, we show that k-rounds of *sum-of-squares / Lasserre* relaxations of size $k\binom{n}{k}$ achieve best-possible approximation guarantees for Max CSPs among all symmetric SDP relaxations of size at most $\binom{n}{k}$. This result gives the first lower bounds for symmetric SDP relaxations of Max CSPs, and indicates that the sum-of-squares method provides the "right" SDP relaxation for this class of problems.

Moreover, for $k < n/4$, we show the existence of symmetric LP relaxations of size $O(n^{2k})$ for the traveling salesman problem that achieve per instance best-possible approximation guarantees among all symmetric LP relaxations of size roughly $\binom{n}{k}$.

***Keywords*-semidefinite programs, extended formulations, linear programs, constraint satisfaction problems, travelling salesman problem.**

I. Introduction

The best known (approximation) algorithms for a vast range of combinatorial optimization problems are based on (polynomial-size) symmetric LP or SDP relaxations. This work studies the computational power of such relaxations and compares it to the power of explicit relaxation, e.g., obtained from hierarchies [1]–[4]. The motivation for this comparison is two fold: On the one hand, we can deduce new lower bounds for general symmetric relaxations (from known lower bounds for hierarchies). On the other hand, our comparison identifies the best symmetric relaxations of a certain size. These relaxations are therefore a promising basis for new approximation results.

A groundbreaking work of Yannakakis [5] initiated the study of general LP formulations and showed exponential lower bounds on the size of symmetric LP formulations for traveling salesman and maximum matching. This work also provided a framework for proving lower bounds on general LP formulations (based on the notion of *nonnegative rank* of matrices). Recent breakthroughs [6], [7] extended Yannakakis's lower bounds to the non-symmetric case using techniques from communication complexity.

There has been some progress to extend these lower bounds on LP formulations to the approximation setting [8]–[10], but so far only for clique[1] and Max CSPs. In the SDP setting, no lower bounds are known for explicit problems (neither exact nor approximate). This work gives the first lower bounds for general symmetric SDP relaxations.

The ultimate goal of this line of research is to identify the "right" LP and SDP relaxations (not necessarily symmetric) for classes of optimization problems. We conjecture that Sherali–Adams and sum-of-squares relaxations of polynomial size indeed achieve the best possible approximation guarantees among all polynomial-size LP and SDP relaxations for many problems. Some of our proof techniques are tailored toward the symmetric case (especially the group-theoretic arguments). However, our basic framework also works in the non-symmetric case and could therefore form the basis of a proof for the non-symmetric case, in the same way that Yannakakis's framework was instrumental in the lower bound results for general LP formulations.

Symmetric SDP Relaxations for Max CSPs: Semidefinite programming marries linear programming and spectral methods. Prominent examples like max cut and sparsest cut show that semidefinite relaxations can achieve approximation guarantees that are not (known to be) achievable by linear relaxations or spectral methods on their own [11], [12].

The Unique Games Conjecture [13] predicts that a particularly simple SDP relaxation achieves best-possible approximation guarantees for every Max CSP [14]. It's an outstanding open question whether more complicated SDP relaxation can refute this conjecture (by providing better approximations than the basic SDP relaxation). Indeed, recent works show that polynomial-size SDP relaxations based on

[1]In the case of clique, the LP relaxations considered in the lower bounds do not subsume all LP relaxations for clique that appear in the literature.

the sum-of-squares method / Lasserre hierarchy provide better approximations on families of instances for which many other methods fail [15].

Analogous to Yannakakis's characterization of general LP relaxations, there exists a characterization of general SDP relaxations (in terms of the notion of *positive-semidefinite rank* of matrices) [6], [16], but no explicit lower bounds are known. We provide an alternative characterization in terms of sums-of-squares of linear subspaces, inspired by the viewpoint developed in previous work [10]. This characterization allows us to compare the power of general symmetric SDP relaxation and the power of low-degree sum-of-squares relaxations [3], [4] for Max CSPs.

Theorem I.1. *For every Max CSP, $k < n/4$, degree-k sum-of-squares relaxations achieve the best-possible approximation guarantees among all symmetric SDP relaxations of size at most $\binom{n}{k}$.*

(This result also holds if k is a function of n, up to exponential-size relaxations.)

Moreover, we exhibit an augmented degree k sum-of-squares relaxation that achieves the best approximation guarantees among all symmetric SDP relaxations on an instance-by-instance basis. Specifically, we show the following:

Theorem I.2. *For every Max CSP* Max-Π*, $k < n/4$, there exists an augmented degree-k sum-of-squares relaxation of size n^{k+10} that on every instance $\mathfrak{I}$ of the* Max-Π*, achieves the best-possible approximation guarantees among all symmetric SDP relaxations of size at most $\binom{n}{k}$.*

It is interesting that the guarantee of optimality holds on every instance, and therefore would apply even when one is interested in special classes of instances such as planar instances.

Combined with known lower bounds for sum-of-squares relaxations [17]–[19], this result implies the first explicit lower bounds for general symmetric SDP relaxations of natural optimization problems. (A recent work shows that random 0/1 polytopes require exponential-size SDP relaxations [20], but these polytopes do not correspond to natural combinatorial optimization problems.) A concrete implication is that for every positive constant $\varepsilon > 0$, symmetric SDP relaxation require exponential size to achieve approximation ratio $7/8 + \varepsilon$ for Max 3-Sat.

Symmetric LP Relaxations for Traveling Salesman: Recent years have seen a lot of progress on the approximability of constraint satisfaction problems (e.g., in the context of the Unique Games Conjecture). It is a very interesting question whether these results could lead to new insights about other notorious combinatorial optimization problems, e.g., traveling salesman.

Previous work showed that symmetric LP relaxations for Max CSPs are exactly as powerful as Sherali–Adams relaxations [10]. Here, we show an analogous result for traveling salesman.

Theorem I.3. *For every $k \in \mathbb{N}$, $k < n/4$, there exists an symmetric LP relaxation $\mathcal{L}$ for* traveling salesman *on n sites with $O(n^{2k})$ constraints that can be generated in time $O(n^{4k+3})$ such that the following holds – For every instance $\mathfrak{I}$ and every symmetric LP relaxation $\mathcal{L}'$ of size at most $\binom{n}{k}$ we have*

$$\mathcal{L}'(\mathfrak{I}) \leqslant \mathcal{L}(\mathfrak{I}) \leqslant \mathrm{opt}(\mathfrak{I}),$$

i.e., $\mathcal{L}$ is a better approximation to opt *than $\mathcal{L}'$.*

Related Work: In an independent effort, Fawzi, Saunderson, and Parillo [21] show relate lower bounds for symmetric semidefinite programs. The notion of symmetry used in [21] is stronger than the one we use (see Section III-A for more details), but many of our results are incomparable.

II. Preliminaries

Constraint Satisfaction Problem: Constraint Satisfaction Problems (CSPs) are a broad class of discrete optimization problems that include Max Cut and Max 3-Sat. The main focus of this work is CSPs over a boolean domain; the same ideas can be generalized to CSPs over general finite domains.

Fix some $k \in \mathbb{N}$. A *k-ary predicate* is a mapping $P : \{-1,1\}^k \to \{0,1\}$. For a given $n \in \mathbb{N}$ and a subset $S \subseteq [n]$ with $|S| = k$, we use the notation $P^S : \{-1,1\}^n \to \{0,1\}$ to denote the mapping

$$P^S(x_1, x_2, \ldots, x_n) = P(x_S),$$

where $x_S \in \{-1,1\}^k$ denotes the projection of $x \in \{-1,1\}^n$ to the coordinates in S.

Let Π be a collection of k-ary predicates. We will often refer to such a collection as a *k-ary CSP*. An instance of $\mathfrak{I}$ of Max-Π consists of n boolean variables $x_1, x_2, \ldots, x_n$, m predicates $P_1, P_2, ..., P_m \in \Pi$, and m subsets $S_1, S_2, \ldots, S_m \subseteq [n]$. The constraints of the CSP are naturally of the form $P_i^{S_i}(x) = 1$. The associated optimization problem is to find an assignment $x \in \{-1,1\}^n$ that satisfies as many constraints as possible, i.e. that maximizes

$$\mathrm{val}_{\mathfrak{I}}(x) = \frac{1}{m}\sum_{i=1}^{m} P_i^{S_i}(x).$$

Given a CSP instance $\mathfrak{I}$, we denote its optimal value by $\mathrm{opt}_{\mathfrak{I}} = \max_{x\in\{-1,1\}^n} \mathrm{val}_{\mathfrak{I}}(x)$ Finally, we will use Max-Π_n to denote the set of Max-Π instances on n variables.

Positive Semi-definite Matrices: We will use the notation $\mathcal{S}_k^+$ to denote the cone of $k\times k$ symmetric, positive semidefinite (PSD) matrices with real entries. We equip $\mathcal{S}_k^+$ with the Frobenius inner product $\langle U, V\rangle = \mathrm{Tr}(U^T V) = \sum_{i=1}^k \sum_{j=1}^k U_{ij} V_{ij}$.

One may naturally identify $\mathcal{S}_k^+$ with a subset of $\mathbb{R}^{k(k+1)/2}$ so that the inner product of two PSD matrices is equal to the inner product of the corresponding vectors. We will use these two representations interchangeably when the context is clear.

We will now define the notion of an "SDP relaxation" for a CSP. Let Π be a k-ary CSP and let $n \in \mathbb{N}$. An SDP relaxation for $\textsc{Max}\text{-}\Pi_n$ consists of two objects: A linearization and a spectrahedron. Fix a number $R \in \mathbb{N}$ called the *size of the relaxation.*

Linearization: A linearization associates to each assignment $x \in \{-1,1\}^n$ an element $\tilde{x} \in \mathcal{S}_R^+$ and to each instance $\mathfrak{I}$ a vector $\tilde{\mathfrak{I}} \in \mathbb{R}^{R(R+1)/2}$ satisfying the property that $\mathrm{val}_{\mathfrak{I}}(x) = \langle \tilde{\mathfrak{I}}, \tilde{x} \rangle$.

Spectrahedron: A spectrahedron $\mathcal{S}$ is the intersection of the PSD cone with an affine linear subspace, i.e.

$$\mathcal{S} = \{y \in \mathbb{R}^{R(R+1)/2} \mid Ay = b, y \in \mathcal{S}_R^+\},$$

where A is an $\frac{R(R+1)}{2} \times \frac{R(R+1)}{2}$ matrix and $b \in \mathbb{R}^{R(R+1)/2}$. To be a valid relaxation, $\mathcal{S}$ must contain all the integral points, i.e. $\{\tilde{x} : x \in \{-1,1\}^n\} \subseteq \mathcal{S}$.

The SDP thus associated with a $\textsc{Max}\text{-}\Pi_n$ instance $\mathfrak{I}$ is given by

$$\begin{aligned} \text{maximize} \quad & \langle \tilde{\mathfrak{I}}, y \rangle \\ \text{subject to} \quad & Ay = b \\ & y \in \mathcal{S}_R^+ . \end{aligned}$$

It is worth noting that the spectrahedron is independent of the instance $\mathfrak{I}$ (note also that one has a possibly different spectrahedron for every input size n). The instance itself only enters the relaxation through the objective function.

Although we refer to R as the size of the SDP relaxation, note that it has $R(R+1)/2$ variables and equality constraints. Finally, we say that an SDP relaxation is a (c,s)*-approximation* for $\textsc{Max}\text{-}\Pi_n$ if, for every instance $\mathfrak{I}$, the following implication holds true:

$$\mathrm{opt}(\mathfrak{I}) \leqslant s \implies \max_{y \in \mathcal{S}} \langle \tilde{\mathfrak{I}}, y \rangle \leqslant c.$$

Sum of Squares Hierarchy: We now briefly review the *Sum of squares (SoS)* SDP hierarchy for binary CSPs; for an in-depth discussion, one may consult the monograph [22]. For a detailed description of the present perspective, we refer readers to [15].

A solution to the the d-round SoS hierarchy consists of vectors $v_{S,\alpha}$ for all sets of variables $S \subseteq [n]$ with $|S| \leqslant d$ and assignments $\alpha \in \{-1,1\}^S$. The constraints are described as follows. For every subset S such that $|S| \leqslant d$, there should exist a probability distribution μ_S on $\{-1,1\}^S$. Furthermore, these distributions should be consistent in the sense that for any two subsets S and T with $|S|, |T| \leqslant d$, the marginal distributions of μ_S and μ_T on $S \cap T$ should be identical. One then requires that for any subsets $S, T \subseteq [n]$ with $|S \cup T| \leqslant d$ and any assignments $\alpha \in \{-1,1\}^S, \beta \in \{-1,1\}^T$, we have

$$\langle v_{S,\alpha}, v_{T,\beta} \rangle = \mathbb{P}_{\mu_{S \cup T}}\{X_S = \alpha, X_T = \beta\},$$

where X denotes a random variable distributed according to $\mu_{S \cup T}$ and X_S and X_T denote the projections of X to the coordinates in S and T, respectively.

Alternatively, one can think of the SoS SDP as optimizing an objective function over "local expectation" functionals. Consider a map $\tilde{\mathbb{E}}$ that sends n-variate polynomials of degree at most d (over $\mathbb{R}$) to real numbers. We say that $\tilde{\mathbb{E}}$ is a *level-d pseudoexpectation functional* (d-p.e.f) if it satisfies the following properties:

- **Linearity.** For every pair of n-variate real polynomials P and Q with $\deg(P), \deg(Q) \leqslant d$, and every pair of numbers $a, b \in \mathbb{R}$, we have
$$\tilde{\mathbb{E}}(aP + bQ) = a\tilde{\mathbb{E}}(P) + b\tilde{\mathbb{E}}(Q).$$
- **Positivity.** For every polynomial P with $\deg(P) \leqslant d/2$, we have $\tilde{\mathbb{E}}(P^2) \geqslant 0$
- **Normalization.** $\tilde{\mathbb{E}}(1) = 1$.

For CSPs over the boolean cube $\{-1,1\}^n$, we may assume the following additional constraints on the functionals; this is because $x_i^2 = 1$ for $x_i \in \{-1,1\}$.

- **Folding.** For every monomial $x_\alpha = \prod_{i=1}^n (x_i)^{\alpha_i}$ of degree at most d, we have $\tilde{\mathbb{E}}[x_\alpha] = \tilde{\mathbb{E}}[x_{\alpha \bmod 2}]$, where $x_{\alpha \bmod 2} = \prod_i (x_i)^{\alpha_i \bmod 2}$.

Consider now a k-ary CSP instance $\mathfrak{I}$. We may naturally consider the functional $\mathrm{val}_{\mathfrak{I}} : \{-1,1\}^n \to [0,1]$ as a multilinear polynomial of degree at most k by expressing it in the Fourier basis: $\mathrm{val}_{\mathfrak{I}} = \sum_{S \subseteq [n]: |S| \leqslant k} a_S \chi_S$, where $\chi_S(x) = \prod_{i=1}^n x_i$. By abuse of notation, we can consider $\mathrm{val}_{\mathfrak{I}}$ also as a multilinear polynomial over $\mathbb{R}^n$.

We can now express the *degree-d SoS value of the the instance* $\mathfrak{I}$ by

$$\mathrm{SoS}_d(\mathfrak{I}) = \max \left\{ \tilde{\mathbb{E}}\left[\mathrm{val}_{\mathfrak{I}}\right] : \tilde{\mathbb{E}} \text{ is } d\text{-p.e.f} \right\}.$$

III. Symmetric SDPs

A. Symmetry

Let S_n denote the set of permutations on n objects. Clearly S_n acts on $\mathbb{R}^n$ by permutation of the coordinates. We call a subset $S \subseteq \mathbb{R}^n$ *symmetric* if it is invariant under the action of S_n. In [5], an extended formulation of an n-dimensional convex polytope $\mathcal{P} \subseteq \mathbb{R}^n$ is a convex polytope $Q \subseteq \mathbb{R}^{n+n'}$ such that $\mathcal{P}$ is the projection of Q to the first n coordinates. Suppose $\mathcal{P}$ is symmetric. One says that the *extended formulation is symmetric* if, for every $\sigma \in S_n$, there is a $\sigma' \in S_{n'}$ such that the permutation $(\sigma, \sigma') \in S_{n+n'}$ preserves Q, i.e. $Q = (\sigma, \sigma')Q$.

A direct analog of this definition is unsuitable for SDPs. Consider again the natural identification of $\mathcal{S}_R^+$ with a subset of $\mathbb{R}^{R(R+1)/2}$. If $\sigma \in S_{R(R+1)/2}$ and $Y \in \mathbb{R}^{R(R+1)/2}$ is PSD, it is not necessarily the case that σY is PSD. It is more natural to define the action of S_R on $\mathbb{R}^{R(R+1)/2}$ as that which permutes the rows and columns simultaneously. Thus if $\sigma \in S_R$ and $Y = (Y_{ij}) \in \mathbb{R}^{R(R+1)/2}$, we define $\sigma \cdot Y = (Y_{\sigma(i)\sigma(j)})_{ij} \in \mathbb{R}^{R(R+1)/2}$. It is manifestly clear that $\mathcal{S}_R^+ \subseteq \mathbb{R}^{R(R+1)/2}$ is invariant under this action. If one thinks about an SDP as a vector program, this corresponds naturally to permuting the underlying variables. It leads to the following notion of symmetry.

Definition III.1. An SDP relaxation of size R for Max-Π_n is *symmetric* if, for any $\sigma \in S_n$, there is a $\tilde{\sigma} \in S_R$, such that for every $x \in \{-1, 1\}^n$,

$$\widetilde{\sigma(x)} = \tilde{\sigma} \cdot \tilde{x},$$

where $\tilde{x}$ is the linearization of x and $\widetilde{\sigma(x)}$ is the linearization of $\sigma(x)$.

Remark III.2. Fawzi et al. [21] use a more general notion of symmetry wherein for any $\sigma \in S_n$ there exists an invertible matrix $\rho(\sigma)$ such that $\widetilde{\sigma(x)} = \rho(\sigma)\tilde{x}\rho(\sigma)^T$. In our setup, the matrices $\rho(\sigma)$ are restricted to being permutation matrices.

B. Function families

We now present a necessary condition for there to exist a good SDP relaxation for Max-Π_n in terms of families of functions on the discrete cube. This is analogous to the characterization for LPs given in [10], and follows closely the semidefinite generalization of Yannakakis' factorization theorem presented in [6]. In what follows, $\|\cdot\|$ denotes the Euclidean norm.

Theorem III.3. *Consider some boolean CSP Π_n. Suppose that for some $c \geqslant s \geqslant 0$, there exists an SDP relaxation of size R that (c, s)-approximates* Max-Π_n. *Then there exists a family of functions $f_1, f_2, \ldots, f_R : \{-1, 1\}^n \to \mathbb{R}^R$, such that for each instance $\Im$ with $\mathrm{opt}(\Im) \leqslant s$, there are numbers $\{\lambda_{i,j} : 1 \leqslant i, j \leqslant R\} \subseteq \mathbb{R}$ and $\eta \geqslant 0$ satisfying: For all $x \in \{-1, 1\}^n$,*

$$c - val_{\Im}(x) = \sum_{i=1}^{R} \left\| \sum_{j=1}^{R} \lambda_{i,j} f_j(x) \right\|^2 + \eta\,.$$

Furthermore, if the SDP relaxation is symmetric, then the family $\{f_i : 1 \leqslant i \leqslant R\}$ is invariant under permutation of inputs, i.e. for all $\sigma \in S_R$,

$$\{\sigma f_i : 1 \leqslant i \leqslant R\} = \{f_i : 1 \leqslant i \leqslant R\}\,.$$

Proof:

Let $\mathcal{S}$ be the spectrahedron associated with an SDP relaxation of size R that (c, s)-approximates Max-Π_n and write

$$\mathcal{S} = \{y \mid Ay = b, y \in \mathcal{S}_R^+\}\,.$$

Suppose that $\mathrm{opt}(\Im) \leqslant s$. Since the SDP relaxation (c, s)-approximates Max-Π_n, we have $\langle \tilde{\Im}, y\rangle \leqslant c$ for all $y \in \mathcal{S}$.

In particular, this implies that $c - \langle \tilde{\Im}, y\rangle \geqslant 0$ is valid for all $y \in \mathcal{S}$. Therefore, by the strong separation theorem (and the fact that the SDP cone is self-dual), there exists a PSD matrix $\Lambda \in \mathcal{S}_R^+$, a vector $\beta \in \mathbb{R}^{R(R+1)/2}$ and a number $\eta \geqslant 0$ such that for all $y \in \mathcal{S}$,

$$c - \langle \tilde{\Im}, y\rangle = \langle \Lambda, y\rangle + \langle \beta, Ay - b\rangle + \eta\,.$$

Specializing to $y = \tilde{x}$ for $x \in \{-1, 1\}^n$, we have

$$c - \mathrm{val}_{\Im}(x) = c - \langle \tilde{\Im}, \tilde{x}\rangle = \langle \Lambda, \tilde{x}\rangle + \langle \beta, A\tilde{x} - b\rangle + \eta$$

As $\tilde{x} \in \mathcal{S}$ for $x \in \{-1, 1\}^n$ (by the definition of a valid relaxation), we will have $A\tilde{x} - b = 0$, which implies that

$$c - \mathrm{val}_{\Im}(x) = \langle \Lambda, \tilde{x}\rangle + \eta \qquad \forall x \in \{-1, 1\}^n\,.$$

Write $\Lambda = \sum_{i=1}^R \lambda_i \lambda_i^T$ for a set of vectors $\{\lambda_i\} \subseteq \mathbb{R}^R$. For each $x \in \{-1, 1\}^n$, let $\tilde{x} = L_x L_x^T$ be a Cholesky decomposition of $\tilde{x}$, and define the functions $\{f_i\}$ so that $f_1(x), f_2(x), \ldots, f_R(x)$ are the rows of L_x. In this case, we have

$$\begin{aligned} &c - \mathrm{val}_{\Im}(x) \\ &\quad= \left\langle \sum_i \lambda_i \lambda_i^T, \tilde{x} \right\rangle + \eta \\ &\quad= \left\langle \sum_i \lambda_i \lambda_i^T, \sum_i f_i(x) f_i(x)^T \right\rangle + \eta \\ &\quad= \sum_{i=1}^{R} \left\| \sum_{j=1}^{R} \lambda_{i,j} f_j(x) \right\|^2 + \eta\,. \end{aligned}$$

Suppose now that the SDP relaxation is symmetric. By definition, for each permutation $\sigma \in S_n$, there exists a permutation $\tilde{\sigma} \in S_R$ such that

$$\widetilde{\sigma(x)} = \tilde{\sigma} \cdot \tilde{x}$$

for all $x \in \{-1, 1\}^n$.

Note that $f_i(x)$ is the i^{th} row of the Cholesky decomposition of $\tilde{x}$. From the above condition, it is clear that the i^{th} row of the Cholesky decomposition of $\widetilde{\sigma(x)}$ is the $\tilde{\sigma}(i)^{th}$ row of $\tilde{x}$. Hence we will have

$$f_i(\sigma(x)) = f_{\tilde{\sigma}(i)}(x)$$

for all $x \in \{-1, 1\}^n$ and therefore the function family $\{f_i : 1 \leqslant i \leqslant R\}$ is invariant under the action of S_R, as desired. ■

C. Instance optimal symmetric SDPs

We now present an augmented version of the SoS hierarchy and show that the approximation it achieves on *every* Max-Π_n instance is at least as good as *any* symmetric SDP of roughly the same size. Our starting point is a structural lemma on symmetric families of functions.

Definition III.4. A function $f : \{-1, 1\}^n \to \mathbb{R}$ is a *k-near-junta* if $f(x_1, x_2, \ldots, x_n)$ depends on at most k variables and the value $\sum_{i=1}^n x_i$. In other words, there is a subset $S \subseteq [n]$ with $|S| = k$ such that if x and x' have $\sum_{i=1}^n x_i = \sum_{i=1}^n x'_i$ and differ only on coordinates outside S, then $f(x) = f(x')$.

The proof of the following lemma is very similar to an analogous claim in the work of Yannakakis [5].

Lemma III.5 ([10], Lemma 4.3)**.** *Let $\mathcal{F}$ be a finite family of functions of the form $f : \{-1, 1\}^n \to \mathbb{R}$ such that $|\mathcal{F}| \leqslant \binom{n}{k}$ for some $k < n/4$. If $\mathcal{F}$ is invariant under the action of S_n, then each $f \in \mathcal{F}$ is a k-near-junta.*

Recall that the d-round SoS hierarchy corresponds to a normalized pseudo-expectation functional over low degree polynomials. Specifically, the pseudoexpectation functional $\tilde{\mathbb{E}}$ is a linear functional that maps polynomials of degree at most d to $\mathbb{R}$ and satisfies linearity and positivity constraints. Note that this functional can be represented by a table containing the pseudo-expectations of every monomial of degree at most d, and the positivity constraint is equivalent to the quadratic form $P \mapsto \tilde{\mathbb{E}}P^2$ being positive semidefinite.

In the augmented SoS hierarchy, we require a pseudo-expectation functional on a slightly larger class of polynomials than low-degree polynomials. Fix a positive integer d. Consider the vector space of polynomials of the form

$$P = \sum_{0\leqslant i\leqslant 2n} P_i(x)\left(\sum_j x_j\right)^i, \tag{III.1}$$

where each $P_i(x)$ is a polynomial of degree at most d. Note that the dimension of this vector space is at most $2n$ times the dimension of the vector space of degree d polynomials.

In the augmented SoS SDP we will maximize the objective function over pseudo-expectation functionals on this vector space of polynomials. Similar to SoS hierarchy, we require the pseudo-expectation functional $\tilde{\mathbb{E}}$ to satisfy the following properties:

– **Linearity**
  $\tilde{\mathbb{E}}(P+Q) = \tilde{\mathbb{E}}P + \tilde{\mathbb{E}}Q$ for every polynomial P and Q of the form III.1. This is slightly more subtle compared to the usual SoS, since assigning an arbitrary table of values of $\tilde{\mathbb{E}}m(x)(\sum_i x_i)^k$ for every monomial $m(x)$ of degree at most d and $k \leqslant n$ no longer guarantees linearity, as they're not linearly independent. However we can specify a basis of the space spanned by these polynomials and let the SDP output the pseudo-moments of the basis.
  Compared to SoS, the size of this SDP is at most $2n$ times bigger, as the number of polynomials in the basis is at most $2n$ times bigger.
– **Positivity**
  We want that for $P = \sum_{0\leqslant i\leqslant n} P_i(x)(\sum_j x_j)^i$, $\deg(P_i) \leqslant d/2$, $\tilde{\mathbb{E}}P^2 \geqslant 0$. Once we specify the basis, this is equivalent to the quadratic form being semidefinite.
– **Normalization**
  $\tilde{\mathbb{E}}1 = 1$

Finally as the CSP is over the boolean cube $\{-1,1\}^n$, the following additional constraints on the functional arise from the fact that $x_i^2 = 1$.

– **Folding** For every monomial $x_\alpha = \prod_i (x_i)^{\alpha_i}$ of degree $\leqslant d$, we will have $\tilde{\mathbb{E}}[x_\alpha (\sum_i x_i)^j] = \tilde{\mathbb{E}}[x_{\alpha \bmod 2} (\sum_i x_i)^j]$ for all $j \in \{0,\ldots,2n\}$ wherein $x_{\alpha \bmod 2} = \prod_i (x_i)^{\alpha_i \bmod 2}$.

Now we prove that this augmented SoS is instance-wise optimal. Recall that an SDP relaxation is said to give an α-approximation on an instance $\Im \in$ MaxCSP if the value of the SDP relaxation is at most $\alpha \cdot \mathrm{opt}(\Im)$.

Theorem III.6. *Given an instance $\Im$ of Max-Π_n, suppose $2d$-rounds of the augmented SoS hierarchy do not give an α-approximation, then no symmetric SDP of size $\binom{n}{d}$ can achieve an α-approximation on $\Im$.*

Proof: We prove the result by contradiction. Suppose there exists a symmetric SDP that achieves an α-approximation on $\Im$. By Theorem III.3, there exists a family of vector valued functions $\{f_i\}$ such that

$$\alpha \cdot \mathrm{opt}(\Im) - \mathrm{val}_\Im(x) = \sum_i \left\| \sum_j \lambda_{i,j} f_j(x)) \right\|^2 + \eta\,, \tag{III.2}$$

for some $\eta > 0$ and real numbers $\lambda_{i,j}$. Note that by Lemma III.5, each f_i is d near-junta.

Let $f_{j,k}$ be the k-th coordinate of f_j, it is easy to see that $f_{j,k}$ is also d near-junta. Therefore,

$$f_{j,k} = \sum_{l=0}^{n-1}\left(\sum_t x_t\right)^l P_{j,k,l}\,, \tag{III.3}$$

for some polynomials $P_{j,k,l}$ with degree at most d. Here we are also using the fact that $\sum_t x_t$ takes at most $n+1$ different values.

Let $\tilde{E}$ denote the pseudo-expectation functional obtained by solving the $2d$-round augmented SoS hierarchy on the instance $\Im$. Clearly, $\tilde{E}$ can be evaluated on the LHS of (III.2) since $\mathrm{val}_\Im$ is a low-degree polynomial. By (III.3), the pseudoexpectation can also be evaluated on the RHS of (III.2).

On evaluating $\tilde{E}$ on the RHS of (III.2),

$$\begin{aligned} &\tilde{\mathbb{E}} \sum_i \|\sum_j \lambda_{i,j} f_j\|^2 + \eta \\ &= \sum_{i,k} \tilde{\mathbb{E}}\left(\sum_j \lambda_{i,j} f_{j,k}\right)^2 + \eta \\ &= \sum_{i,k} \tilde{\mathbb{E}}\left(\sum_j \lambda_{i,j,l} P_{j,k,l}(\sum_t x_t)^l\right)^2 + \eta \\ &\geqslant 0 \end{aligned}$$

However, on the LHS we will have

$$\tilde{\mathbb{E}}(\alpha \cdot \mathrm{opt}(\Im) - \mathrm{val}_\Im) = \alpha \cdot \mathrm{opt}(\Im) - \mathrm{SoS}(\Im) < 0\,,$$

a contradiction. ∎

Note that one can also modify the Sherali-Adam Hierarchy in the same manner to obtain instance optimal LP for CSPs.

We remark that this augmented SoS SDP is not stronger than the usual SoS SDP in terms of general approximation guarantee (that is, the worst case approximation ratio over all possible instances), as we will show in the next section. However, it is possible that this SDP performs better than SoS on some specific instances.

D. Sum of Squares SDPs

In this section we prove that the Sum of Squares SDP achieves the best possible approximation amongst symmetric SDPs of similar size (not per instance-wise). Specifically we show that the approximation guarantee of SoS SDP on instances with n variables is at least as good as the approximation guarantee of *any* symmetric SDP of similar size on $2n$ variables.

Lemma III.7. *Suppose that the conditions of Theorem III.3 hold for $N = 2n$, then there exists a family of k-juntas $\{g_i\}$ on n variables of size at most $2^k n^k$, such that for every instance $\Im$ on n variables, there exists $\lambda_{i,j}$, such that,*

$$c' - val_{\Im} = \sum_i \|(\sum_j \lambda_{i,j} g_j)\|^2$$

Proof: Given a MaxCSP instance $\Im$, we construct another instance $\Im'$ of size $2n$ by adding n extra dummy variables, while keeping the constraints the same on first n variables. There are no constraints amongst the dummy variables. Since the conditions of Theorem III.3 hold, we have

$$c' - \mathrm{val}_{\Im'}(y) = \sum_i \|(\sum_j \lambda_{i,j} f_j(y))\|^2$$

for every $y \in \{-1, 1\}^{2n}$.

In particular, we have

$$c' - \mathrm{val}_{\Im'}(x, -x) = \sum_i \|(\sum_j \lambda_{i,j} f_j(x, -x))\|^2$$

Define $g_j(x) = f_j(x, -x)$, since f_j is k near-junta, g_j is k-junta.

It's easy to see that $\mathrm{val}_{\Im'}(x, -x) = \mathrm{val}_{\Im}(x)$, hence we have

$$\begin{aligned} & c' - \mathrm{val}_{\Im}(x) \\ =& c' - \mathrm{val}_{\Im'}(x, -x) \\ =& \sum_i \|(\sum_j \lambda_{i,j} f_j(x, -x))\|^2 \\ =& \sum_i \|(\sum_j \lambda_{i,j} g_j)\|^2 \end{aligned}$$

■

Now we prove the main theorem of this section.

Theorem III.8. *Given MaxCSP Π, suppose that $2k$-rounds SoS relaxation cannot achieve (c, s)-approximation on instances with n variables, then no symmetric SDP of size $\binom{N}{k}$ achieves (c, s)-approximation on instances with N variables, with $N = 2n$.*

Proof: We prove it by contradiction. Suppose there exists an SDP relaxation that achieves (c, s)-approximation on instances with N variables, by Lemma III.7, there exists a family of k-juntas g_i such that for every $\Im$ on n variables,

$$c' - \mathrm{val}_{\Im}(x) = \sum_i \|(\sum_j \lambda_{i,j} g_j(x))\|^2$$

In particular, the equation holds for the instance $\Im_0$ where SoS fails to achieve (c, s)-approximation.

Let $\tilde{\mathbb{E}}$ be the pseudo-expectation functional defined by the SoS solution on $\Im_0$, by linearity of $\tilde{\mathbb{E}}$, we have

$$\begin{aligned} \tilde{\mathbb{E}} P &= \tilde{\mathbb{E}} c' - \tilde{\mathbb{E}} \mathrm{val}_{\Im} \\ &= c' - \mathrm{SoS}(\Im) \\ &\leqslant c - \mathrm{SoS}(\Im) \\ &< 0 \end{aligned}$$

However on the other hand, by positivity of $\tilde{\mathbb{E}}$ we have,

$$\tilde{\mathbb{E}} P = \sum_i \tilde{\mathbb{E}} \|(\sum_j \lambda_{i,j} g_j)\|^2 \geqslant 0 ,$$

a contradiction. ■

IV. Instance Optimal Symmetric LPs for TSP

In this section, we will present an *instance-optimal* linear program for Travelling Salesman problem. We begin by recalling the setup of lower bounds for general linear programs. The setup is similar to that of semidefinite programs for MaxCSP outlined in Section II.

An LP relaxation of size R for Travelling Salesman on n sites consists of the following.

Linearization: A linearization that associates to each tour $\sigma \in S_n$ an element $\tilde{\sigma} \in R^m$ and to each instance $\Im$ a vector $\tilde{\Im} \in \mathbb{R}^m$ such that $\mathrm{val}_{\Im}(\sigma) = \langle \tilde{\Im}, \tilde{\sigma} \rangle$ for every instance $\Im$ and every tour $\sigma \in S_n$.

Polyhedron: A convex polytope $P \subseteq \mathbb{R}^m$ described by R linear inequalities, such that $\tilde{\sigma} \in P$ for each tour $\sigma \in S_n$.

Given an instance $\Im$ of Travelling Salesman, the LP relaxation $\mathcal{L}$ outputs the value $\mathcal{L}(\Im) = \min_{y \in P} \langle \tilde{\Im}, y \rangle$. Finally, the LP relaxation is said to be symmetric if for every permutation $\pi \in S_n$ of the sites, there exists a corresponding symmetry $\tilde{\pi}$ of the polytope P such that $\widetilde{\pi(\sigma)} = \tilde{\pi} \cdot \tilde{\sigma}$.

In this section we show that for every constant k, there exists a symmetric LP of size $O(n^k)$, such that on every instance of the Travelling Salesman problem, its approximation ratio α is no worse than that of every symmetric LP of size $\binom{n}{k}$. Specifically, we will prove Theorem I.3 (restated here for convenience).

Theorem IV.1. *(Theorem I.3 restated) For every $k \in \mathbb{N}$, $k < n/4$, there exists an symmetric LP relaxation $\mathcal{L}$ for* traveling salesman *on n sites with $O(n^{2k})$ constraints that can be generated in time $O(n^{4k+3})$ such that the following holds – For every instance $\Im$ and every symmetric LP relaxation $\mathcal{L}'$ of size at most $\binom{n}{k}$ we have*

$$\mathcal{L}'(\Im) \leqslant \mathcal{L}(\Im) \leqslant \mathrm{opt}(\Im) ,$$

i.e., $\mathcal{L}$ is a better approximation to opt *than $\mathcal{L}'$.*

To prove this result, we will need the following tailored version of Theorem 1 in [8] and Theorem 2.2 of [10]

Theorem IV.2. *For every symmetric LP $\mathcal{L}$ for* Travelling Salesman *of size s, there exists a corresponding family of s functions $\mathcal{F} = \{f_i : S_n \mapsto \mathbb{R}^{\geqslant 0}\}$ with the following properties:*

- *The family $\mathcal{F}$ is invariant under permutations of sites, i.e., for every permutation π and $f \in \mathcal{F}$ we have $f \circ \pi \in \mathcal{F}$.*
- *For every instance $\mathfrak{I}$ of* Travelling Salesman *with n vertices, we have*

$$\mathcal{L}(\mathfrak{I}) = \text{ largest } c \text{ such that } \mathrm{val}_{\mathfrak{I}} - c \in Cone(\mathcal{F}),$$

where $Cone(\mathcal{F})$ is the cone generated by the family of functions $\mathcal{F}$.

Remark IV.3. For any family of non-negative functions $\mathcal{F} = \{f : S_n \to \mathbb{R}^{\geqslant 0}\}$, there is a natural linear program $\mathcal{L}$ for Travelling Salesman associated with it

$$\begin{aligned} &\text{Maximize } c \\ &\text{Subject to } \mathrm{val}_{\mathfrak{I}} - c \in Cone(\mathcal{F}) \end{aligned}$$

The optimum $\mathcal{L}(\mathfrak{I})$ always satisfies $\mathcal{L}(\mathfrak{I}) \leqslant \mathrm{opt}(\mathfrak{I})$. This follows from the fact that the function $\mathrm{val}_{\mathfrak{I}} - \mathcal{L}(\mathfrak{I})$ by virtue of being in the $Cone(\mathcal{F})$ is a non-negative function on tours. Finally, as stated above, the linear program $\mathcal{L}$ is not tractable since $Cone(\mathcal{F}) \in \mathbb{R}^{n!}$.

In order to prove Theorem I.3, we will construct a symmetric linear program $\mathcal{L}$ of size $O(n^{2k})$ with an associated family of functions $\mathcal{F}$ such that:

1) For every family of functions $\mathcal{F}' = \{f : S_n \mapsto \mathbb{R}^{\geqslant 0}\}$ that is invariant under permutations of sites and $|\mathcal{F}'| \leqslant \binom{n}{k}$,
$$Cone(\mathcal{F}') \subseteq Cone(\mathcal{F})$$
2) The linear program $\mathcal{L}$ can be generated in time $O(n^{4k+3})$.

By property 1 above, for any other symmetric linear program $\mathcal{L}'$ of size $\binom{n}{k}$, the corresponding family of functions $\mathcal{F}'$ will satisfy $Cone(\mathcal{F}') \subseteq Cone(\mathcal{F})$. From Theorem IV.2, this implies that we will have $\mathcal{L}(\mathfrak{I}) \geqslant \mathcal{L}'(\mathfrak{I})$ for every instance. Moreover, since $\mathcal{L}$ can be generated efficiently, the above construction will suffice to prove Theorem I.3.

First, we begin by describing the family of $O(n^{2k})$ non-negative functions $\mathcal{F}$.

Definition IV.4. Let S and T denote ordered tuples of at most k vertices such that $|S| = |T|$. Let $I_{S,T,\mathrm{odd}} : S_n \mapsto \{0, 1\}$ be the indicator function such that $I_{S,T,\mathrm{odd}}(\sigma) = 1$ if and only if $\sigma(S) = T$ and σ is an odd permutation. Similarly let $I_{S,T,\mathrm{even}}$ be the indicator function such that $I_{S,T,\mathrm{even}}(\sigma) = 1$ if and only if $\sigma(S) = T$ and σ is an even permutation.

Define $\mathcal{F}$ to be the family of all these indicator functions, i.e.,

$$\mathcal{F} = \{I_{S,T,\mathrm{odd}}, I_{S,T,\mathrm{even}} | S, T \subset [n], |S| = |T| \leqslant k\}$$

Now we will show that the family of functions $\mathcal{F}$ satisfies property 1. Formally, we will prove the following:

Lemma IV.5. *If $\mathcal{F}'$ is a family of non-negative functions that is invariant under permutation of vertices and satisfies $|\mathcal{F}'| \leqslant \binom{n}{k}$ then*

$$Cone(\mathcal{F}') \subseteq Cone(\mathcal{F})$$

In order to prove Lemma IV.5, we will first show an analogue to Lemma III.5. Roughly speaking, if a family of functions on S_n is invariant under permutations, then each function only depends on few locations of the tour, and possibly the parity of the tour.

Claim IV.6. *Suppose a family of $\binom{n}{k}$ functions $\{f_i : S_n \mapsto \mathbb{R}^{\geqslant 0}\}$ is invariant under permutation of its inputs, then for each f_i, there exists a set of indices J_i such that f_i only depends on the positions of J_i and the parity of the input permutation.*

Proof: Let $\mathrm{Orb}(f)$ denote the orbit of a function f under permutation of its inputs. Therefore we have $|\mathrm{Orb}(f_i)| \leqslant \binom{n}{k}$. Hence for each f_i the automorphism group that preserves f_i is large. Now we appeal to the following claim from the work of Yannakakis [5],

Claim IV.7. *([5], Claim 2) Let H be a group of permutations whose index in S_n is at most $\binom{n}{k}$ for some $k < n/4$. Then there exists a set J of size at most k such that H contains all even permutations that fix the elements of J.*

By Claim IV.7, the automophism group $\mathrm{Aut}(f_i)$ contains all even permutations that fix a subset of coordinates J_i with $|J_i| \leqslant k$. Therefore the function only depends on the positions of the indices in J and the parity of the permutation. ■

Proof of Lemma IV.5: By Claim IV.6, each function in $\{f_i\}$ only depends on the positions of at most k indices and the parity of the permutation, therefore the function can be written as non-negative combination of the indicator functions in $\{g_i\}$, which implies $\mathrm{cone}(\{f_i\}) \subseteq \mathrm{cone}(\{g_i\})$. □

The natural linear program for Travelling Salesman associated with the family of functions $\mathcal{F}$ is given by,

$$\begin{aligned} &\text{Maximize } c \\ &\text{Subject to } \mathrm{val}_{\mathfrak{I}} - c \in Cone(\mathcal{F}) \end{aligned} \tag{IV.1}$$

Note that the condition $\mathrm{val}_{\mathfrak{I}} - c \in Cone(\mathcal{F})$ translates in to the following set of linear constraints on a set of variables $\{\lambda_f\}_{f \in \mathcal{F}}$,

$$\begin{aligned} \mathrm{val}_{\mathfrak{I}}(\sigma) - c &= \sum_{f \in \mathcal{F}} \lambda_f f(\sigma) && \forall \sigma \in S_n \\ \lambda_f &\geqslant 0 && \forall f \in \mathcal{F} \end{aligned}$$

The above linear program has $O(n^{2k})$ variables and $O(n^{2k})$ inequalities as desired, but has $n!$ equalities. Now, we will show how to find a succinct representation of the linear program with $O(n^{2k})$ variables and $O(n^{2k})$ constraints. To this end, let us write $\mathrm{val}_{\mathfrak{I}}(\sigma)$ as a sum of the indicator functions of pairwise events. Define $J_{(i,a,b)}$ to be the indicator function of the event that the i-th edge of the tour is (a,b). That is

$$J_{(i,a,b)} = 1[\sigma(i) = a \wedge \sigma(i+1 \mod n) = b]$$

Hence we have

$$\mathrm{val}_{\mathfrak{I}}(\sigma) = \sum_{i=1}^{n} \sum_{a,b} D(a,b) J_{(i,a,b)} \,,$$

where $D(a,b)$ is the cost of traversing the edge (a,b). Rewriting the linear program IV.1 in terms of these functions, we have

$$\begin{aligned} &\text{maximize } c \\ &\text{subject to } \sum_{i=1}^{n} \sum_{a,b} D(a,b) J_{(i,a,b)} - c - \sum_{f \in \mathcal{F}} \lambda_f f = 0 \qquad \text{(IV.2)} \\ &\qquad \lambda_f \geqslant 0 \qquad \text{(IV.3)} \end{aligned}$$

We will rewrite the above linear program in an alternate basis, so as to reduce $n!$ equations to n^{2k} equations. To achieve this, we begin by making the following observation:

Observation IV.8. *The inner product $\langle f, f' \rangle$ between any pair of functions $f, f' \in \mathcal{F}$ can be computed in time $O(n^3)$.*

Proof: Consequence of the simple combinatorial structure of the functions in $\mathcal{F}$. In particular, there are explicit formulae for the inner products of the indicator functions. ■

Let V be the $|\mathcal{F}| \times |\mathcal{F}|$ matrix with entries indexed by $f, f' \in \mathcal{F}$, given by $V_{ff'} = \langle f, f' \rangle$. By the observation, we can compute the matrix V in time $O(n^{4k+3})$. Given the matrix V, for every vector $\Lambda = (\lambda_f)_{f \in \mathcal{F}}$ we will have,

$$\begin{aligned} \sum_{f \in \mathcal{F}} \lambda_f f = 0 &\iff \langle \sum_{f \in \mathcal{F}} \lambda_f f, \sum_{f \in \mathcal{F}} \lambda_f f \rangle = 0 \\ &\iff \Lambda^T V \Lambda = 0 \\ &\iff V\Lambda = 0 \end{aligned}$$

Note that while $\sum_{f \in \mathcal{F}} \lambda_f f = 0$ is a system of $n!$ equations in the variables λ_f, $V\Lambda = 0$ is an equivalent system of $O(n^{2k})$ equations.

Note that the indicator functions $J_{i,a,b}$ appearing in the equalities IV.2 can also be written as,

$$J_{(i,a,b)} = I_{(i,i+1),(a,b),odd} + I_{(i,i+1),(a,b),even} \,.$$

Further, we can write the constant function 1 as $1 = I_{odd} + I_{even}$ where $I_{odd}, I_{even} \in \mathcal{F}$.

Let w denote a vector indexed by the family $\mathcal{F}$ given by

$$w_f = \begin{cases} D(a,b) - \lambda_f & \text{if } f = I_{(i,i+1 \mod n),(a,b),odd} \\ D(a,b) - \lambda_f & \text{if } f = I_{(i,i+1 \mod n),(a,b),even} \\ c - \lambda_f & \text{if } f \in \{I_{odd}, I_{even}\} \\ -\lambda_f & \text{otherwise} \end{cases}$$

Note that each entry of w is a linear function of the variables $(\lambda_f)_{f \in \mathcal{F}}$ and c (the distances $D(a,b)$ are constants depending on the instance $\mathfrak{I}$). We can rewrite (IV.2) as,

$$\sum_{f \in \mathcal{F}} w_f f = 0 \,,$$

or equivalently,

$$Vw = 0$$

Therefore, we can rewrite the linear program $\mathcal{L}$ in the following form with $O(n^{2k})$ variables namely $\{\lambda_f, w_f\}_{f \in \mathcal{F}} \cup \{c\}$ and $O(n^{2k})$ constraints.

$$\begin{aligned} &\text{maximize } c \\ &\text{subject to } Vw = 0 \\ &\qquad w_f = \begin{cases} D(a,b) - \lambda_f & \text{if } f = I_{(i,i+1 \mod n),(a,b),odd} \\ D(a,b) - \lambda_f & \text{if } f = I_{(i,i+1 \mod n),(a,b),even} \\ c - \lambda_f & \text{if } f \in \{I_{odd}, I_{even}\} \\ -\lambda_f & \text{otherwise} \end{cases} \\ &\qquad \lambda_f \geqslant 0 \qquad \forall f \in \mathcal{F} \end{aligned}$$

Acknowledgment

James R. Lee's research is supported by NSF grant CCF-1217256. Prasad Raghavendra and Ning Tan acknowledge support from an NSF Career Award and an Alfred P. Sloan Research Fellowship. David Steurer's research is supported by an NSF Career Award and an Alfred P. Sloan Fellowship.

References

[1] L. Lovász and A. Schrijver, "Cones of matrices and set-functions and 0-1 optimization," *SIAM Journal on Optimization*, vol. 1, no. 2, pp. 166–190, 1991. 1

[2] H. D. Sherali and W. P. Adams, "A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems," *SIAM Journal on Discrete Mathematics*, vol. 3, no. 3, pp. 411–430, 1990. 1

[3] P. A. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," Ph.D. dissertation, 2000. 1, 2

[4] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001. 1, 2

[5] M. Yannakakis, "Expressing combinatorial optimization problems by linear programs," *Journal of Computer and System Sciences*, vol. 43, no. 3, pp. 441–466, 1991. 1, 3, 4, 7

[6] S. Fiorini, S. Massar, S. Pokutta, H. R. Tiwary, and R. de Wolf, "Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds," in *Proceedings of the 44th symposium on Theory of Computing*. ACM, 2012, pp. 95–106. 1, 2, 4

[7] T. Rothvoß, "The matching polytope has exponential extension complexity," *arXiv preprint arXiv:1311.2369*, 2013. 1

[8] G. Braun, S. Fiorini, S. Pokutta, and D. Steurer, "Approximation limits of linear programs (beyond hierarchies)," in *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*. IEEE, 2012, pp. 480–489. 1, 6

[9] M. Braverman and A. Moitra, "An information complexity approach to extended formulations," in *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing*. ACM, 2013, pp. 161–170. 1

[10] S. O. Chan, J. R. Lee, P. Raghavendra, and D. Steurer, "Approximate constraint satisfaction requires large lp relaxations," *arXiv preprint arXiv:1309.0563*, 2013. 1, 2, 4, 6

[11] M. X. Goemans and D. P. Williamson, ". 879-approximation algorithms for max cut and max 2sat," in *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*. ACM, 1994, pp. 422–431. 1

[12] S. Arora, S. Rao, and U. Vazirani, "Expander flows, geometric embeddings and graph partitioning," *Journal of the ACM (JACM)*, vol. 56, no. 2, p. 5, 2009. 1

[13] S. Khot, "On the power of unique 2-prover 1-round games," in *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*. ACM, 2002, pp. 767–775. 1

[14] P. Raghavendra, "Optimal algorithms and inapproximability results for every csp?" in *Proceedings of the 40th annual ACM symposium on Theory of computing*. ACM, 2008, pp. 245–254. 1

[15] B. Barak, F. G. Brandão, A. W. Harrow, J. Kelner, D. Steurer, and Y. Zhou, "Hypercontractivity, sum-of-squares proofs, and their applications," in *Proceedings of the 44th symposium on Theory of Computing*. ACM, 2012, pp. 307–326. 2, 3

[16] J. Gouveia, P. A. Parrilo, and R. R. Thomas, "Lifts of convex sets and cone factorizations," *Mathematics of Operations Research*, vol. 38, no. 2, pp. 248–264, 2013. 2

[17] D. Grigoriev, "Linear lower bound on degrees of positivstellensatz calculus proofs for the parity," *Theoretical Computer Science*, vol. 259, no. 1, pp. 613–622, 2001. 2

[18] G. Schoenebeck, "Linear level lasserre lower bounds for certain k-csps," in *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*. IEEE, 2008, pp. 593–602. 2

[19] M. Tulsiani, "Csp gaps and reductions in the lasserre hierarchy," in *Proceedings of the 41st annual ACM symposium on Theory of computing*. ACM, 2009, pp. 303–312. 2

[20] J. Briët, D. Dadush, and S. Pokutta, "On the existence of 0/1 polytopes with high semidefinite extension complexity," in *Algorithms–ESA 2013*. Springer, 2013, pp. 217–228. 2

[21] H. Fawzi, J. Saunderson, and P. A. Parrilo, "Equivariant semidefinite lifts and sum-of-squares hierarchies," *Arxiv*, vol. arXiv:1312.6662, 2013. 2, 4

[22] M. Laurent, "Sums of squares, moment matrices and optimization over polynomials," in *Emerging applications of algebraic geometry*, ser. IMA Vol. Math. Appl. Springer, New York, 2009, vol. 149, pp. 157–270. [Online]. Available: http://dx.doi.org.offcampus.lib.washington.edu/10.1007/978-0-387-09686-5_7 3