On the Hardest Problem Formulations
for the 0/1 Lasserre Hierarchy

Adam Kurpisz[(✉)], Samuli Leppänen, and Monaldo Mastrolilli

IDSIA, 6928 Manno, Switzerland
{adam,samuli,monaldo}@idsia.ch

Abstract. The Lasserre/Sum-of-Squares (SoS) hierarchy is a systematic procedure for constructing a sequence of increasingly tight semidefinite relaxations. It is known that the hierarchy converges to the 0/1 polytope in n levels and captures the convex relaxations used in the best available approximation algorithms for a wide variety of optimization problems.

In this paper we characterize the set of 0/1 integer linear problems and unconstrained 0/1 polynomial optimization problems that can still have an integrality gap at level $n - 1$. These problems are the hardest for the Lasserre hierarchy in this sense.

1 Introduction

The *Sum of Squares* (SoS) proof system introduced by Grigoriev and Vorobjov [20] is a proof system based on the *Positivstellensatz*. Shor [37], Nesterov [30], Parrilo [33] and Lasserre [24] show that it can be efficiently automatized using semidefinite programming (SDP) such that any n-variable degree-d proof can be found in time $n^{O(d)}$. The SDP, often called the Lasserre/SoS[1] hierarchy, is the dual of the SoS proof system, meaning that the Lasserre hierarchy value at "level d/2" of an optimization problem is equal to the best provable bound using a degree-d SoS proof (see the monograph by Laurent [26]). For a brief history of the different formulations from [20], [24], [33] and the relations between them and results in real algebraic geometry we refer the reader to [32].

The Lasserre hierarchy can be seen as a systematic procedure to strengthen a relaxation of an optimization problem by constructing a sequence of increasingly tight SDP relaxations. The tightness of the relaxation is parametrized by its *level* or *round*, which corresponds to the degree of the proof in the proof system. Moreover, it captures the convex relaxations used in the best available approximation algorithms for a wide variety of optimization problems. For example, the first round of the hierarchy for the INDEPENDENT SET problem implies the Lovász θ-function [28] and for the MAX CUT problem it gives the Goemans-Williamson

---

This work replaces and improves an early version of the paper titled "The Lasserre hierarchy in almost diagonal form" appeared in arXiv.

[1] For brevity, we will interchange Lasserre hierarchy with SoS hierarchy since they are essentially the same in our context.

relaxation [15]. The ARV relaxation of the SPARSEST CUT [2] problem is no stronger than the relaxation given in the third round of the Lasserre hierarchy, and the subexponential time algorithm for UNIQUE GAMES [1] is implied by a sublinear number of rounds [5,21]. More recently, it has been shown that $O(1)$ levels of the Lasserre hierarchy is equivalent in power to any polynomial size SDP extended formulation in approximating maximum constraint satisfaction problems [27]. Other approximation guarantees that arise from the first $O(1)$ levels of the Lasserre (or weaker) hierarchy can be found in [5,6,9,10,12,13,21,29,34]. For a more detailed overview on the use of hierarchies in approximation algorithms, see the surveys [11,25,26].

The limitations of the Lasserre hierarchy have also been studied. Most of the known lower bounds for the hierarchy originated in the works of Grigoriev [17,18] (also independently rediscovered later by Schoenebeck [36]). In [18] it is shown that random 3XOR or 3SAT instances cannot be solved by even $\Omega(n)$ rounds of SoS hierarchy. Lower bounds, such as those of [7,38] rely on [18,36] plus gadget reductions. For a different technique to obtain lower bounds, see the recent paper [4].

A particular weakness of the hierarchy revolves around the fact that it has hard time reasoning about terms of the form $x_1 + ... + x_n$ using the fact that all x_i's are 0/1. Grigoriev [17] showed that $\lfloor n/2 \rfloor$ levels of Lasserre are needed to prove that the polytope $\{x \in [0,1]^n | \sum_{i=1}^n x_i = \lfloor n/2 \rfloor + 1/2\}$ contains no integer point. A simplified proof can be found in [19].

In [8] Cheung considered a simple instance of the MIN KNAPSACK problem, i.e. the minimization of $\sum_{i=1}^n x_i$ for 0/1 variables such that $\sum_{i=1}^n x_i \geq \delta(n)$, for some $\delta(n) < 1$ that depends on n. Cheung proved that the Lasserre hierarchy requires n levels to converge to the integral polytope. This is shown by providing a feasible solution at level $n-1$ of value $\frac{n}{n+1}$, whereas the smallest integral solution has value 1. This gives an integrality gap[2] of $1 + \frac{1}{n}$ that vanishes with n.

We emphasize that the main interest in the work of Cheung revolves around understanding how fast the Lasserre hierarchy converges to the integral polytope and not how fast the integrality gap reduces, therefore not ruling out the possibility that the integrality gap might decrease slowly with the number of levels. This is conceptually an important difference. For the MAX KNAPSACK (or MIN KNAPSACK) problem the presence of an integrality gap at some "large" level $t(n)$, that depends on n, is promptly implied by $P \neq NP$, whereas the existence of a "large" integrality gap at some "large" level $t(n)$ is not immediately clear (since both MAX KNAPSACK and MIN KNAPSACK problems admit an FPTAS). With this regard, note that Cheung's result also implies that for the MAX KNAPSACK the Lasserre hierarchy requires n levels to converge to the integral polytope. However, in [23] it is shown that only $O(1/\varepsilon)$ levels are needed to obtain an integrality gap of $1 - \varepsilon$, for any arbitrarily small constant $\varepsilon > 0$. It is also worth pointing out that currently the Cheung knapsack result [8] is the

---

[2] The *integrality gap* is defined to be the measure of the quality of the relaxation described by the ratio between the optimal integral value and the relaxed optimal value. If this ratio is different from 1 we will say that "there is an integrality gap".

only known integrality gap result for Lasserre/Sum-of-Squares hierarchy at level $n - 1$.

Our results. With n variables, the n-th level of the Lasserre hierarchy is sufficient to obtain the 0/1 polytope, where the only feasible solutions are convex combinations of feasible integral solutions [24]. This can be proved by using the *canonical lifting lemma* (see Laurent [25]), where the feasibility of a solution to the Lasserre relaxation at level n reduces to showing that a certain diagonal matrix is positive semidefinite (PSD).

The main challenge in analyzing integrality gap instances at level smaller than n is showing that a candidate solution satisfies the positive semidefinite constraints. In this paper, we first show that the feasibility of a solution to the Lasserre relaxation at level $n - 1$ reduces to showing that a matrix differing from a diagonal matrix by a rank one matrix (almost diagonal form) is PSD. We analyze the eigenvalues of the almost diagonal matrices and obtain compact necessary and sufficient conditions for the existence of an integrality gap of the Lasserre relaxation at level $n - 1$. This result can be seen as the opposite of [16] where they consider the case when the first order Lasserre relaxation is exact.

Interestingly, for 0/1 integer linear programs the existence of a gap at level $n - 1$ implies that the problem formulation contains only constraints of the form we call *Single Vertex Cutting* (SVC). An SVC constraint only excludes one vertex of the $\{0, 1\}^n$ hypercube. It can thus be seen as the most generic non-trivial form of constraint, since the feasible set of any integer linear program can be modeled using only constraints of this form.

This characterization allows us to show that n levels of Lasserre are needed to prove that a polytope defined by (exponentially many) SVC constraints contains no integer point. No other example of this kind was known at level n (the previously known example in [17] requires $\lfloor n/2 \rfloor$ levels).

One problem where SVC constraints can arise naturally is the KNAPSACK problem. By applying the computed conditions, we improve the Cheung [8] MIN KNAPSACK integrality gap of the Lasserre relaxation at level $n - 1$ from $1 + 1/n$ to any arbitrary large number. This shows a substantial difference between the MIN KNAPSACK and the MAX KNAPSACK when we take into consideration the integrality gap size of the Lasserre relaxation.

Furthermore, we show that a similar result holds beyond the class of integer linear programs. More precisely, we show that any unconstrained 0/1 polynomial optimization problem exhibiting an integrality gap at level $n - 1$ of the Lasserre relaxation has necessarily an objective function given by a polynomial of degree n. This rules out the existence of any integrality gap at level $n - 1$ for any k-ary boolean constraint satisfaction problem with $k < n$. Finally, we provide an example of an unconstrained 0/1 polynomial optimization problem with an integrality gap at level $n - 1$ of the Lasserre hierarchy, and discuss why the problem can be seen as a constraint satisfaction version of an SVC constraint. Our result complements the recent paper [14] where it is shown that the Lasserre relaxation does not have any gap at level $\lceil \frac{n}{2} \rceil$ when optimizing n-variate 0/1 polynomials of degree 2.

2   The Lasserre Hierarchy

In this section we provide a definition of the Lasserre hierarchy [24]. For the applications that we have in mind, we restrict our discussion to optimization problems with 0/1-variables and linear constraints. More precisely, we consider the following general optimization problem $\mathbb{P}$: Given a multilinear polynomial $f : \{0,1\}^n \to \mathbb{R}$

$$\mathbb{P} : \quad \min\{f(x)|x \in \{0,1\}^n, g_\ell(x) \geq 0 \text{ for } \ell \in [m]\} \tag{1}$$

where $\{g_\ell(x) : \ell \in [m]\}$ are linear functions of x.

Many basic optimization problems are special cases of $\mathbb{P}$. For example, any k-ary boolean constraint satisfaction problem, such as MAX CUT, is captured by (1) where a degree k function $f(x)$ counts the number of satisfied constraints, and no linear constraints $g_\ell(x) \geq 0$ are present. Also any 0/1 integer linear program is a special case of (1), where $f(x)$ is a linear function.

Lasserre [24] proposed a hierarchy of SDP relaxations for increasing δ,

$$\min\{L(f)|L : \mathbb{R}[X]_{2\delta} \to \mathbb{R}, L(1) = 1, \text{ and } L(u^2), L(u^2 g_\ell) \geq 0, \forall \text{ polynomial } u\} \tag{2}$$

where $L : \mathbb{R}[X]_{2\delta} \to \mathbb{R}$ is a linear map with $\mathbb{R}[X]_{2\delta}$ denoting the ring $\mathbb{R}[X]$ restricted to polynomials of degree at most 2δ.[3] In particular for 0/1 problems L vanishes on the truncated ideal generated by $x_i^2 - x_i$. Note that (2) is a relaxation since one can take L to be the evaluation map $f \to f(x^*)$ for any optimal solution x^*.

Relaxation (2) can be equivalently formulated in terms of *moment matrices* [24]. In the context of this paper, this matrix point of view is more convenient to use and it is described below. In our notation we mainly follow the survey of Laurent [25] (see also [35]).

Variables and Moment Matrix. Throughout this paper, vectors are written as columns. Let N denote the set $\{1, \ldots, n\}$. The collection of all subsets of N is denoted by $\mathcal{P}(N)$. For any integer $t \geq 0$, let $\mathcal{P}_t(N)$ denote the collection of subsets of N having cardinality at most t. Let $y \in \mathbb{R}^{\mathcal{P}(N)}$. For any nonnegative integer $t \leq n$, let $M_t(y)$ denote the matrix with (I, J)-entry $y_{I \cup J}$ for all $I, J \in \mathcal{P}_t(N)$. Matrix $M_t(y)$ is termed in the following as the *t-moment matrix* of y. For a linear function $g(x) = \sum_{i=1}^n g_i \cdot x_i + g_0$, we define $g * y$ as a vector, often called *shift operator*, where the I-th entry is $(g * y)_I = \sum_{i=1}^n g_i y_{I \cup \{i\}} + g_0 y_I$. Let f denote the vector of coefficients of polynomial $f(x)$ (where f_I is the coefficient of monomial $\Pi_{i \in I} x_i$ in $f(x)$).

Definition 1. *The Lasserre relaxation of problem* (1) *at the t-th level, denoted as* $\text{LAS}_t(\mathbb{P})$*, is the following*

---

[3] In [3], $L(p)$ is written $\tilde{\mathbb{E}}[p]$ and called the "pseudo-expectation" of p.

$$\text{LAS}_t(\mathbb{P}): \quad \min \left\{ \sum_{I \subseteq N} f_I y_I \,|\, y \in \mathbb{R}^{\mathcal{P}_{2t+2d}(N)} \text{ and } y \in \mathbb{M} \right\} \tag{3}$$

where $\mathbb{M}$ is the set of vectors $y \in \mathbb{R}^{\mathcal{P}_{2t+2d}(N)}$ that satisfy the following PSD conditions

$$y_\varnothing = 1 \tag{4}$$

$$M_{t+d}(y) \succeq 0 \tag{5}$$

$$M_t(g_\ell * y) \succeq 0 \qquad \ell \in [m] \tag{6}$$

where $d = 0$ if $m = 0$ (no linear constraints) otherwise $d = 1$.

We will use the following known facts (see e.g. [25,35]). Consider any vector $w \in \mathbb{R}^{\mathcal{P}(N)}$ (vector w is intended to be either the vector $y \in \mathbb{R}^{\mathcal{P}(N)}$ of variables or the shifted vector $g * y$ for any $g \in \mathbb{R}^{\mathcal{P}(N)}$). For any $I \in \mathcal{P}(N)$, variables $\{w_I^N : I \subseteq N\}$ are defined as follows:

$$w_I^N := \sum_{H \subseteq N \setminus I} (-1)^{|H|} w_{H \cup I}$$

Note that $w_I = \sum_{I \subseteq J} w_J^N$ (by using inclusion-exclusion principle, see [35]). The latter with $y_\emptyset = 1$ implies that $\sum_{J \subseteq N} y_J^N = 1$, and that the objective function can be rewritten as follows:

$$\sum_{I \subseteq N} f_I y_I = \sum_{I \subseteq N} f(x_I) y_I^N$$

where $f(x_I)$ denotes the value of $f(x)$ when $x_i = 1$ for $i \in I$ and $x_i = 0$ for $i \notin I$.

Congruent transformations are known not to change the sign of the eigenvalues (see e.g. [22]). It follows that in studying the positive-semidefiniteness of matrices we can focus on congruent matrices without loss of generality. Let $D_t(w)$ denote the diagonal matrix in $\mathbb{R}^{\mathcal{P}_t(N) \times \mathcal{P}_t(N)}$ with (I, I)-entry equal to w_I^N for all $I \in \mathcal{P}_t(N)$.

Lemma 1. [25] Matrix $M_n(w)$ is congruent to the diagonal matrix $D_n(w)$.

By Lemma 1, $M_n(y) \succeq 0$ implies that the variables in $\{y_I^N : I \subseteq N\}$ can be interpreted as a probability distribution (see [25,35]), where y_I^N is the probability that the variables with index in I are set to one and the remaining to zero.

Lemma 2. [25] For any polynomial g of degree at most one, $y \in \mathbb{R}^{\mathcal{P}(N)}$ and $z = g * y$ we have $z_I^N = g(x_I) \cdot y_I^N$ where $g(x_I) = \sum_{i \in I} a_i + b$.

Note that, by using Lemma 1 and Lemma 2, it can be easily shown the well known fact that at level n any solution can be written as a convex combination of feasible integral solutions. The latter implies that any integrality gap vanishes at level n.

3   The $(n-1)$-Moment Matrix

In the following we show that $M_{n-1}(w)$ is congruent to the diagonal matrix $D_{n-1}(w)$ perturbed by a rank one matrix, and analyze its eigenvalues. For ease of notation, we will use D to denote $D_{n-1}(w)$ throughout this section.

Lemma 3. *Matrix $M_{n-1}(w)$ is congruent to the matrix $D + w_N^N \cdot vv^\top$, where v is a $|\mathcal{P}_{n-1}(N)|$-dimensional vector with $v_I = (-1)^{n+1-|I|}$ for any $I \in \mathcal{P}_{n-1}(N)$.*

3.1   Positive Semidefiniteness of $M_{n-1}(y)$

In this section we derive the necessary and sufficient conditions for $M_{n-1}(w) \succeq 0$. From Lemma 3 we have that $M_{n-1}(y) \succeq 0 \Leftrightarrow D + w_N^N vv^\top \succeq 0$, where $vv^\top$ is a rank one matrix with entries ± 1.

Lemma 4. *If $w_N^N \neq 0$ then, for any $I \subseteq N$, $\lambda = w_I^N$ is an eigenvalue of the matrix $D + w_N^N vv^\top$ if and only if there is another $J \neq I$ with $w_I^N = w_J^N$, $J \subseteq N$; The remaining eigenvalues are the solutions λ of the following equation*

$$\sum_{N \neq I \subseteq N} \frac{1}{\lambda - w_I^N} = \frac{1}{w_N^N} \tag{7}$$

Proof. Consider the zeroes λ of the characteristic polynomial of $D + w_N^N vv^\top$:

$$\det(\lambda I - (D + w_N^N vv^\top)) = \det(D_\lambda - w_N^N vv^\top) = 0 \tag{8}$$

where $D_\lambda = \lambda I - D$. Applying Cauchy's formula for the determinant of a rank-one perturbation [22, p. 26] we can write this as

$$\det(D_\lambda) - w_N^N v^\top \operatorname{adj}(D_\lambda) v = 0 \tag{9}$$

Consider a solution λ to (9). Exactly one of the following three cases must hold:

1. D_λ is nonsingular, meaning that $\lambda \neq w_I^N$ for all $N \neq I \subseteq N$. Then $\operatorname{adj}(D_\lambda) = (\det D_\lambda) D_\lambda^{-1}$ and the above becomes

$$\det(D_\lambda)(1 - w_N^N v^\top D_\lambda^{-1} v) = 0$$

   which simplifies to (7).
2. D_λ is singular, and $\lambda = w_I^N$ for exactly one $N \neq I \subseteq N$. Then $\operatorname{adj}(D_\lambda) = \alpha e_I e_I^\top$ for some nonzero α [22, p. 22-23], where $(e_I)_J = 1$ if $I = J$ and $(e_I)_J = 0$ otherwise. Now (9) simplifies to

$$w_N^N v^\top (\alpha e_I e_I^\top) v = 0$$

   which can only hold if $w_N^N = 0$. Hence such λ cannot be a solution to (8).
3. D_λ is singular and there are more than one $N \neq I \subseteq N$ such that $\lambda = w_I^N$. Then $\operatorname{adj}(D_\lambda) = 0$ [22, p. 22] and λ is a solution to (8).

$\square$

Lemma 5. *Matrix $D + w_N^N vv^\top$ is positive-semidefinite if and only if either $w_I^N \geq 0$ for all $I \subseteq N$, or the following holds*

$$w_K^N < 0, \qquad \text{for exactly one } K \subseteq N, \tag{10}$$

$$w_J^N > 0, \qquad \text{for all } K \neq J \subseteq N, \tag{11}$$

$$\sum_{I \subseteq N} \frac{1}{w_I^N} \leq 0 \tag{12}$$

Proof. If $w_I^N \geq 0$ for all $I \subseteq N$ then $D + w_N^N vv^\top \succeq 0$ since it is the sum of two PSD matrices. Otherwise, there exists $I \subseteq N$ with $w_I^N < 0$ and we distinguish between the following complementary cases.

If there are two different sets $N \neq I, J \subseteq N$ such that $w_I^N = w_J^N < 0$, then by Lemma 4 the matrix $D + w_N^N vv^\top$ has a negative eigenvalue. Therefore we may assume that all the negative entries of D are different from each other. Then by Lemma 4, any potentially negative eigenvalue is given by (7). With this in mind, let $f(\lambda) = \sum_{N \neq I \subseteq N} \frac{1}{\lambda - w_I^N}$ and study the points λ where $f(\lambda)$ intersects the line given by $\frac{1}{w_N^N}$.

There are three cases:

1. For two sets $N \neq I, J \subseteq N$ we have $w_I^N < w_J^N \leq 0$. Then since the function $f(\lambda)$ has vertical asymptotes at the points w_I^N and w_J^N, there must be a point $\lambda < 0$ such that $f(\lambda) = \frac{1}{w_N^N}$ regardless of the value of w_N^N (see Figure 1 (i)).
2. For exactly one $N \neq I \subseteq N$ we have $w_I^N \leq 0$ and $w_N^N < 0$. Then $f(\lambda)$ has one vertical asymptote in $(-\infty, 0]$ and thus the line $\frac{1}{w_N^N}$ crosses the graph of $f(\lambda)$ at least in one $\lambda < 0$ (see Figure 1 (ii)).
3. For exactly one $I \subseteq N$ we have $w_I^N < 0$ and the rest are strictly positive. Then we note that there can be at most one $\lambda < 0$ such that $f(\lambda) = \frac{1}{w_N^N}$. Inspecting the form of the graph shows that there is no intersection in the negative half-plane if and only if $f(0) \geq \frac{1}{w_N^N}$ (see Figure 1 (iii) and (iv) for the case $I = N$). □

4   Integrality Gaps of Lasserre Hierarchy at Level $n - 1$

In this section we characterize the set of problems $\mathbb{P}$ of the form (1) that can have an integrality gap at level $n-1$ of the Lasserre relaxation. In particular, we prove that in order to exhibit an integrality gap, a constrained problem can only have constraints each of which rule out only one point of the $\{0, 1\}^n$ hypercube. We fully characterize what this means in the case where the constraints are linear. We also discuss two examples of problems with such constraints, and in particular, we exhibit a simple instance of the MIN KNAPSACK problem that has an unbounded integrality gap. Finally, we show that if $\mathbb{P}$ is an unconstrained problem that has an integrality gap at level $n - 1$, then the objective function of $\mathbb{P}$ must be a polynomial of degree n.
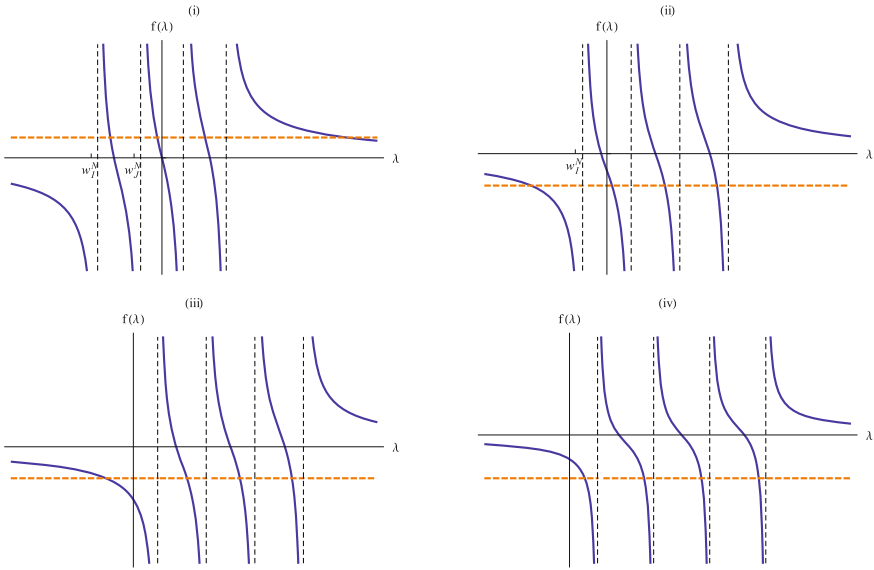
Fig. 1. A conceptual plot of different relevant arrangements of the graph of $f(\lambda)$ and the graph of $\frac{1}{w_N^N}$ (dotted lines)

4.1   Problems with Linear Constraints

In this subsection we focus on 0/1-integer linear programs $\mathbb{P}$ of the form (1). We will assume, w.l.o.g., that if constraint $g(x) \geq 0$ is satisfied by all integral points then it is redundant and no one of these redundant constraints is present.

Theorem 1. *Let $\mathbb{P}$ be a 0/1-integer linear program of the form* (1). *The Lasserre relaxation* $\text{LAS}_{n-1}(\mathbb{P})$ *has an integrality gap if and only if there exists a solution* $\{y_I^N | I \subseteq N\}$ *that satisfies the following conditions:*

$$y_I^N > 0 \qquad \text{for all } I \subseteq N, \tag{13}$$

$$\sum_{I \subseteq N} y_I^N = 1 \tag{14}$$

$$g_\ell(x_{K_\ell})y_{K_\ell}^N < 0 \qquad \text{for exactly one } K_\ell \subseteq N \text{ for each } \ell \in [m], \tag{15}$$

$$g_\ell(x_J)y_J^N > 0 \qquad \text{for all } \ell \in [m], \text{ for all } K_\ell \neq J \subseteq N, \tag{16}$$

$$\sum_{I \subseteq N} \frac{1}{g_\ell(x_I)y_I^N} \leq 0 \qquad \text{for all } \ell \in [m], \tag{17}$$

$$\sum_{I \subseteq N} y_I^N f(x_I) < f(x_{I^*}) \tag{18}$$

where $f(x_{I^*})$ is a minimal integral feasible solution.

Definition 2. *We call $g(x) \geq 0$ a* Single Vertex Cutting *(SVC) constraint if there exists only one $I \subseteq N$ such that $g(x_I) < 0$ and for every other $I \neq J \subseteq N$ it holds $g(x_J) > 0$.*

Corollary 1. *Let $f(x_{I^*})$ denote the integral optimum of (1). If there is an integrality gap, i.e., $y \in \mathrm{LAS}_{n-1}(\mathbb{P})$ such that $\sum_{I \subseteq N} y_I^N f(x_I) < f(x_{I^*})$, then the constraints in (1) are SVC.*

We are considering only problems with linear constraints over $\{0, 1\}^n$, so it is straightforward to characterize the SVC constraints.

Lemma 6. *Let $g(x) = \sum_{i=1}^{n} a_i x_i - b \geq 0$ be a linear SVC constraint. Then $b \neq 0$ and $a_i \neq 0$ for all i, and if P is the set of indices such that $a_i < 0 \Leftrightarrow i \in P$, then $\sum_{i \in P} a_i < b$, but $\sum_{i \in Q} a_i > b$ for all $P \neq Q \subseteq N$.*

4.2   Example Problems with SVC Constraints at Level $n - 1$

As proved in Corollary 1, SVC constraints are in some sense the most difficult constraints to handle for the Lasserre hierarchy. Each such constraint excludes only one point of the $\{0, 1\}^n$ hypercube, and thus the feasible set of any integer linear program can be modeled using only these constraints. It follows that if modeled in this way, any integer linear program can potentially have an integrality gap at level $n - 1$ of the Lasserre hierarchy. In this section we give two examples of problems where the Lasserre hierarchy does not converge to the integer polytope even at level $n - 1$.

Unbounded Integrality Gap for the Min Knapsack. One problem where the SVC constraint naturally arises is the KNAPSACK problem. We show that the minimization version of the problem has an unbounded integrality gap at level $n - 1$ of the Lasserre hierarchy. Indeed, consider the following simple instance of the MIN-KNAPSACK:

$$(GapKnap) \min\{\textstyle\sum_{i=1}^{n} x_i \mid \sum_{i=1}^{n} x_i \geq 1/P, x_i \in \{0, 1\} \text{ for } i \in [n]\} \qquad (19)$$

Notice that the optimal *integral* value of $(GapKnap)$ is one. The optimal value of the linear programming relaxation of $(GapKnap)$ is $1/P$, so the integrality gap of the LP is P and can be arbitrarily large.

By using Theorem 1 we prove the following dichotomy-type result. If we allow a "large" P (exponential in the number of variables n), then the Lasserre hierarchy is of no help to limit the unbounded integrality gap of $(GapKnap)$, even at level $(n - 1)$. This analysis is tight since $\mathrm{LAS}_n(GapKnap)$ admits an optimal integral solution with n variables. We also show that the requirement that P is exponential in n is necessary for having a "large" gap at level $(n - 1)$.

Corollary 2. *(Integrality Gap Bounds for* MIN-KNAPSACK*) The integrality gap of $\mathrm{LAS}_{n-1}(GapKnap)$ is k, for any $k \geq 2$ if and only if $P = \Theta(k) \cdot 2^{2n}$.*

Remark 1. We observe that the instance (19) can be easily ruled out by requiring that each coefficient of any variable must be not larger than the constant term in the knapsack constraint. However, even with this pruning step, the integrality gap can be made unbounded up to the last but two levels of the Lasserre hierarchy: add an additional variable x_{n+1} only in the constraint (not in the objective function) and increase the constant term to $1 + 1/P$. Any solution for $\text{LAS}_{n-1}(GapKnap)$ can be easily turned into a feasible solution for the augmented instance by setting the new variables $y'_I = y_{I \setminus \{n+1\}}$ for any $I \in \mathcal{P}_{2t+2}([n+1])$ and observing that any principal submatrix of the new moment matrices has either determinant equal to zero or it is a principal submatrix in the moment matrix of the reduced problem.

Undetected Empty Integer Hull. As discussed at the beginning of this section, any integer linear problem can be modeled using SVC constraints. Formulating the problem in this "pathological" way can potentially hinder the convergence of the Lasserre hierarchy. We demonstrate this by showing an extreme example, where the Lasserre hierarchy cannot detect that the integer hull is empty even at level $n - 1$.

Consider the feasible set given by (exponentially many) inequalities of the form

$$\sum_{i \in P}(1 - x_i) + \sum_{i \in N \setminus P} x_i \geq b \tag{20}$$

for each $P \subseteq N$. Clearly, any integral assignment I such that $x_i = 1$ if $i \in I$ and $x_i = 0$ otherwise, cannot satisfy all of the inequalities when b is positive. However, there exists an assignment of the variables y_I^N that satisfies the conditions of Theorem 1, and is hence a feasible solution to the Lasserre relaxation of the polytope described above at level $n - 1$, as shown below.

Consider a symmetric solution $y_I^N = \frac{1}{2^n}$ for every $I \subseteq N$ and some constraint of the form (20) corresponding to a given set $P \subseteq N$. Now the variables $z_I^N = g(x_I)y_I^N$ satisfy (15) and (16), and we need to check that it is possible to satisfy (17):

$$\sum_{I \subseteq N} \frac{1}{z_I^N} = \frac{1}{2^n} \sum_{I \subseteq N} \frac{1}{|P \setminus I| + |I \setminus P| - b} \leq 0 \Leftrightarrow \sum_{\varnothing \neq I \subseteq N} \frac{1}{|I| - b} \leq \frac{1}{b}$$

When $0 < b < \frac{1}{2}$, the above is implied by $\sum_{\varnothing \neq I \subseteq N} 2 \leq \frac{1}{b}$, so choosing $b = \frac{1}{2^{n+1}}$ makes (17) satisfied.

4.3   Unconstrained Problems at Level $n - 1$

Let $f : \{0,1\}^n \to \mathbb{R}$ be an objective function of a polynomial minimization problem normalized such that $\min_{x \in \{0,1\}^n} f(x) = 0$ and $\max_{x \in \{0,1\}^n} f(x) = 1$. We start with the conditions that an unconstrained polynomial optimization problem has to satisfy in order do admit a gap at level $n - 1$.

Theorem 2. *Let* $\mathbb{P}$ *denote an unconstrained polynomial optimization problem of the form* (1). *The Lasserre relaxation* $\mathrm{LAS}_{n-1}(\mathbb{P})$ *has an integrality gap if and only if there exists a solution* $\{y_I^N | I \subseteq N\}$ *that satisfies* (18) *and the following conditions:*

$$\sum_{I \subseteq N} y_I^N = 1 \tag{21}$$

$$y_K^N < 0 \qquad \textit{for exactly one } K \subseteq N, \tag{22}$$

$$y_J^N > 0 \qquad \textit{for all } K \neq J \subseteq N, \tag{23}$$

$$\sum_{I \subseteq N} \frac{1}{y_I^N} \leq 0 \tag{24}$$

We note that f can always be represented as a multivariate polynomial of degree at most n. The main result of this section is Theorem 3.

Theorem 3. *If* f *is a function such that* f *has an integrality gap at level* $n-1$, *then* f *is a multivariate polynomial of degree* n.

Proof. We will use some elementary Fourier analysis of boolean functions (see e.g. [31, Ch.1]). To follow an established convention, we switch from studying the function $f : \{0,1\}^n \to \mathbb{R}$ to $h : \{-1,1\}^n \to \mathbb{R}$ via the bijective transform $f(x) = h(1 - 2x)$. Observe that f is of degree t if and only if h is of degree t, and for any $S \subseteq N$ we have $f(x_S) = h(w_S)$, where $w_i = -1$ if $i \in S$ and $w_i = 1$ otherwise.

Assume as before that for some $I_1 \subseteq N$, $h(w_{I_1}) = 1$ and $0 \leq h(w_I) \leq 1$. We assume that $|I_1|$ is even and let $I_2 \subseteq N$ be some fixed set such that $|I_2|$ is odd (the case where $|I_1|$ is odd is symmetric). We assume that h has an integrality gap, so by Lemma 7 (see below) necessarily $\sum_{I \subseteq N} h(w_I) < 2$, which we rewrite in a more convenient form (using $h(w_{I_1}) = 1$)

$$h(w_{I_2}) < 1 - \sum_{I_1 \neq I \neq I_2} h(w_I) \tag{25}$$

Assume now that h has a degree smaller than n, or in other words, its Fourier coefficient $\hat{h}(N)$ is 0:

$$\hat{h}(N) = 2^{-n} \sum_{S \subseteq N} h(w_S)(-1)^{|S|} = 0$$

Removing the normalizing constant and reordering the sum the above implies (using the assumptions on the parity of $|I_1|, |I_2|$)

$$\sum_{\substack{S \neq I_1 \\ |S| \text{ even}}} h(w_S) - \sum_{\substack{S \neq I_2 \\ |S| \text{ odd}}} h(w_S) = -1 + h(w_{I_2}) < - \sum_{I_1 \neq I \neq I_2} h(w_I)$$

by (25). Moving all the h terms to the left hand side yields

$$2 \sum_{\substack{S \neq I_1 \\ |S| \text{ even}}} h(w_S) < 0$$

which contradicts the assumption that $h(w) \geq 0$. $\square$

Lemma 7. *Let $f(x_{I_1}) = 1$ and $0 \leq f(x) \leq 1$ for every $x \in \{0,1\}^n$. If f is such that $\sum_{I \subseteq N} f(x_I) \geq 2$ then there is no gap at level $n-1$.*

We point out that there exists a function of degree n that exhibits an integrality gap at level $n-1$. Consider the function given by

$$f(x) = 1 - \sum_{\varnothing \neq I \subseteq N} (-1)^{|I|} \prod_{i \in I} x_i$$

This function has the value 1 when all the variables are 0, and 0 elsewhere. It is a straightforward application of Theorem 2 to show that $f(x)$ exhibits an integrality gap at level $n-1$. We remark that $f(x)$ can be seen as a constraint satisfaction version of an SVC constraint.

References

1. Arora, S., Barak, B., Steurer, D.: Subexponential algorithms for unique games and related problems. In: FOCS, pp. 563–572 (2010)
2. Arora, S., Rao, S., Vazirani, U.V.: Expander flows, geometric embeddings and graph partitioning. Journal of the ACM **56**(2) (2009)
3. Barak, B., Brandão, F.G.S.L., Harrow, A.W., Kelner, J.A., Steurer, D., Zhou, Y.: Hypercontractivity, sum-of-squares proofs, and their applications. In: STOC, pp. 307–326 (2012)
4. Barak, B., Chan, S.O., Kothari, P.: Sum of squares lower bounds from pairwise independence. In: STOC (2015)
5. Barak, B., Raghavendra, P., Steurer, D.: Rounding semidefinite programming hierarchies via global correlation. In: FOCS, pp. 472–481 (2011)
6. Bateni, M., Charikar, M., Guruswami, V.: Maxmin allocation via degree lower-bounded arborescences. In: STOC, pp. 543–552 (2009)
7. Bhaskara, A., Charikar, M., Vijayaraghavan, A., Guruswami, V., Zhou, Y.: Polynomial integrality gaps for strong sdp relaxations of densest k-subgraph. In: SODA, pp. 388–405 (2012)
8. Cheung, K.K.H.: Computation of the Lasserre ranks of some polytopes. Mathematics of Operations Research **32**(1), 88–94 (2007)
9. Chlamtac, E.: Approximation algorithms using hierarchies of semidefinite programming relaxations. In: FOCS, pp. 691–701 (2007)
10. Chlamtac, E., Singh, G.: Improved approximation guarantees through higher levels of SDP hierarchies. In: Goel, A., Jansen, K., Rolim, J.D.P., Rubinfeld, R. (eds.) APPROX and RANDOM 2008. LNCS, vol. 5171, pp. 49–62. Springer, Heidelberg (2008)
11. Chlamtac, E., Tulsiani, M.: Convex relaxations and integrality gaps. In: Anjos, M.F., Lasserre, J.B. (eds.) Handbook on semidefinite, conic and polynomial optimization. International Series in Operations Research & Management Science, vol. 166, pp. 139–169. Springer, Heidelberg (2012)

12. Cygan, M., Grandoni, F., Mastrolilli, M.: How to sell hyperedges: the hypermatching assignment problem. In: SODA, pp. 342–351 (2013)
13. de la Vega, W.F., Kenyon-Mathieu, C.: Linear programming relaxations of maxcut. In: SODA, pp. 53–61 (2007)
14. Fawzi, H., Saunderson, J., Parrilo, P.: Sparse sum-of-squares certificates on finite abelian groups (2015). CoRR, abs/1503.01207
15. Goemans, M.X., Williamson, D.P.: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. Journal of the ACM **42**(6), 1115–1145 (1995)
16. Gouveia, J., Parrilo, P.A., Thomas, R.R.: Theta bodies for polynomial ideals. SIAM Journal on Optimization **20**(4), 2097–2118 (2010)
17. Grigoriev, D.: Complexity of positivstellensatz proofs for the knapsack. Computational Complexity **10**(2), 139–154 (2001)
18. Grigoriev, D.: Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. Theoretical Computer Science **259**(1–2), 613–622 (2001)
19. Grigoriev, D., Hirsch, E.A., Pasechnik, D.V.: Complexity of semi-algebraic proofs. In: Alt, H., Ferreira, A. (eds.) STACS 2002. LNCS, vol. 2285, p. 419. Springer, Heidelberg (2002)
20. Grigoriev, D., Vorobjov, N.: Complexity of null-and positivstellensatz proofs. Annals of Pure and Applied Logic **113**(1–3), 153–160 (2001)
21. Guruswami, V., Sinop, A.K.: Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with psd objectives. In: FOCS, pp. 482–491 (2011)
22. Horn, R.A., Johnson, C.R.: Matrix analysis. Cambridge University Press (2013)
23. Karlin, A.R., Mathieu, C., Nguyen, C.T.: Integrality gaps of linear and semi-definite programming relaxations for knapsack. In: Günlük, O., Woeginger, G.J. (eds.) IPCO 2011. LNCS, vol. 6655, pp. 301–314. Springer, Heidelberg (2011)
24. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM Journal on Optimization **11**(3), 796–817 (2001)
25. Laurent, M.: A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0–1 programming. Mathematics of Operations Research **28**(3), 470–496 (2003)
26. Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. Emerging Applications of Algebraic Geometry **149**, 157–270 (2009)
27. Lee, J.R., Raghavendra, P., Steurer, D.: Lower bounds on the size of semidefinite programming relaxations. In: STOC (to appear, 2015)
28. Lovász, L.: On the shannon capacity of a graph. IEEE Transactions on Information Theory **25**, 1–7 (1979)
29. Magen, A., Moharrami, M.: Robust algorithms for on minor-free graphs based on the Sherali-Adams hierarchy. In: APPROX-RANDOM, pp. 258–271 (2009)
30. Nesterov, Y.: Global quadratic optimization via conic relaxation, pp. 363–384. Kluwer Academic Publishers (2000)
31. O'Donnell, R.: Analysis of Boolean Functions. Cambridge University Press (2014)
32. O'Donnell, R., Zhou, Y.: Approximability and proof complexity. In: SODA, pp. 1537–1556 (2013)
33. Parrilo, P.: Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. PhD thesis, California Institute of Technology (2000)

34. Raghavendra, P., Tan, N.: Approximating csps with global cardinality constraints using sdp hierarchies. In: SODA, pp. 373–387 (2012)
35. Rothvoß, T.: The lasserre hierarchy in approximation algorithms. Lecture Notes for the MAPSP 2013 - Tutorial, June 2013
36. Schoenebeck, G.: Linear level Lasserre lower bounds for certain k-csps. In: FOCS, pp. 593–602 (2008)
37. Shor, N.: Class of global minimum bounds of polynomial functions. Cybernetics **23**(6), 731–734 (1987)
38. Tulsiani, M.: Csp gaps and reductions in the Lasserre hierarchy. In: STOC, pp. 303–312 (2009)