SIAM J. OPTIM.
Vol. 30, No. 1, pp. 798–822




# HIGH DEGREE SUM OF SQUARES PROOFS, BIENSTOCK–ZUCKERBERG HIERARCHY, AND CHVÁTAL–GOMORY CUTS*

MONALDO MASTROLILLI†

*A mia mamma, che esiste per mancanza.*

**Abstract.** Chvátal–Gomory cuts (CG-cuts) and the Bienstock–Zuckerberg hierarchy capture useful linear programs that the standard bounded degree sum-of-squares (SoS) hierarchy fails to capture. In this paper we present a novel polynomial time SoS hierarchy for 0/1 problems with a custom subspace of high degree polynomials (not the standard subspace of low degree polynomials). We show that the new SoS hierarchy recovers the Bienstock–Zuckerberg hierarchy. Our result implies a linear program that reproduces the Bienstock–Zuckerberg hierarchy as a polynomial-sized, efficiently constructible extended formulation that satisfies all constant pitch inequalities. The construction is also very simple, and it is fully defined by giving the supporting polynomials. Moreover, for a class of polytopes (e.g., set cover and packing problems), the resulting SoS hierarchy optimizes in polynomial time over the polytope resulting from any constant rounds of CG-cuts, up to an arbitrarily small error in the solution value. Arguably, this is the first example where different basis functions can be useful in *asymmetric* situations to obtain a hierarchy of relaxations.

**Key words.** sum-of-squares hierarchy, Chvátal–Gomory cuts, Lasserre hierarchy, lift and project methods

**AMS subject classifications.** 90C10, 90C27

**DOI.** 10.1137/17M1150712

## 1. Introduction.

The Lasserre/sum-of-squares (SoS) hierarchy [17, 25, 27, 31] is a systematic procedure for constructing a sequence of increasingly tight semidefinite relaxations. The SoS hierarchy is parameterized by its *level* (or *degree*) $d$, such that the formulation gets tighter as $d$ increases, and a solution of accuracy $\varepsilon > 0$ can be found by solving a semidefinite program of size $(mn\log(1/\varepsilon))^{O(d)}$, where $n$ is the number of variables and $m$ the number of constraints in the original problem. In this paper we consider 0/1 problems. In this setting, 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 (see, e.g., [3, 6, 18, 19] and the references therein).

In a recent paper Kurpisz, Leppänen, and Mastrolilli [15] characterize the set of 0/1 integer linear problems that still have an (arbitrarily large) integrality gap at level $n-1$. These problems are the "hardest" for the SoS hierarchy in this sense. In another paper, the same authors [16] consider a problem that is solvable in $O(n\log n)$ time and proved that the integrality gap of the SoS hierarchy is unbounded at level $\Omega(\sqrt{n})$ even after incorporating the objective function as a constraint (a classical trick that sometimes helps to improve the quality of the relaxation). All these SoS-hard instances are covering problems.

Chvátal–Gomory (CG) rounding is a popular cut generating procedure that is often used in practice (see, e.g., [7] and section 6 for a short introduction). There

*Received by the editors October 5, 2017; accepted for publication (in revised form) January 21, 2020; published electronically March 3, 2020. Preliminary version appeared in *Proceedings of IPCO'17*, 2017.
https://doi.org/10.1137/17M1150712
**Funding:** This work was supported by Swiss National Science Foundation project 200020-169022, "Lift and Project Methods for Machine Scheduling Through Theory and Experiments."
†IDSIA, Lugano, Switzerland (monaldo@idsia.ch).

are several prominent examples of CG-cuts in polyhedral combinatorics, including the odd-cycle inequalities of the stable set polytope, the blossom inequalities of the matching polytope, the simple Möbius ladder inequalities of the acyclic subdigraph polytope, and the simple comb inequalities of the symmetric traveling salesman polytope, to name a few. CG-cuts capture useful and efficient linear programs that the bounded degree SoS hierarchy fails to capture. Indeed, the SoS-hard instances studied in [15] are the "easiest" for CG-cuts, in the sense that they are captured within the *first* CG closure. It is worth noting that it is NP-hard [21] to optimize a linear function over the first CG closure, an interesting contrast to lift-and-project hierarchies where one can optimize in polynomial time for any constant number of levels.[1]

Interestingly, Bienstock and Zuckerberg [5] proved that, in the case of set cover, one can separate over all CG-cuts to an arbitrary fixed precision in polynomial time. The result in [5] is based on another result [4] by the same authors, namely on a (positive semidefinite) lift-and-project operator (which we denote BZ herein) that is quite different from the previously proposed operators. This lift-and-project operator generates different variables for different relaxations. They showed that this flexibility can be very useful in attacking relaxations of some set cover problems.

These three methods, SoS, CG, and BZ, are to some extent incomparable, roughly meaning that there are instances where one succeeds while the other fails (see, to name a few, [2] for a comparison between SoS and BZ and [15] for "easy" cases for CG-cuts that are "hard" for SoS, and finally note that clique constraints are "easy" for SoS but "hard" for CG-cuts [28]).

One can think of the standard Lasserre/SoS hierarchy at level $d$ as optimizing an objective function over linear functionals that sends $n$-variate polynomials of degree at most $d$ (over $\mathbb{R}$) to real numbers. The restriction to polynomials of degree $d$ is the standard way (as suggested in [17, 27] and used in most of the applications) to bound the complexity, implying a semidefinite program of size $n^{O(d)}$. However, this is not strictly necessary for getting a polynomial time algorithm and it can be easily extended by considering more general subspaces having a "small" (i.e., polynomially bounded) set of basis functions (see, e.g., Chapter 3 in [6] and [9, 10]). This is a less explored direction and it will play a key role in this paper. Indeed, the more general view of the SoS approach has been used so far to exploit very symmetric situations (see, e.g., [9, 10, 29]). For symmetric cases the use of different basis functions has been proved to be very useful.

To the best of the author's knowledge, in this paper we give the first example where different basis functions can be useful in *asymmetric* situations to obtain a hierarchy of relaxations. More precisely, we focus on 0/1 problems and show how to reframe the BZ hierarchy [4] as an augmented version of the SoS hierarchy that uses high degree polynomials (in section 4 we consider the set cover problem, which is the main known application of the BZ approach, and in section 5 we sketch the general framework that is based on the set cover case). The resulting high degree SoS approach retains in one single unifying SoS framework the best from the standard bounded degree SoS hierarchy, incorporates the BZ approach, and allows us to get, in polynomial time for any fixed $t \in \mathbb{N}$ and $\varepsilon > 0$, a solution that satisfies the $t$th CG closure and that is at most $\varepsilon$-times worse than the optimal solution value for both

[1]It has often been claimed in recent papers that one can optimize over degree-$d$ SoS via the ellipsoid algorithm in $n^{O(d)}$ time. In a recent work, O'Donnell [26] observed that this often repeated claim is far from true. However, this issue does not apply to most of the results published so far and to the applications of this paper. See also [24] for recent news.

set cover and packing problems (BZ guarantees this only for set cover problems). Moreover, the proposed framework is very simple and, assuming a basic knowledge in SoS machinery (see section 2), it is fully defined by giving the supporting polynomials. This is in contrast to the BZ hierarchy that requires an elaborate description [4, 33]. Finally, as observed in [1] (see Propositions 25 and 26 in [1]), the performances of the BZ hierarchy depend on the presence of redundant constraints.[2] The proposed approach removes these unwanted features.

We emphasize that one can also generalize the Sherali–Adams hierarchy/proof system in the same manner to obtain the covering results. We will give a detailed description of this in the following. So the formulation that we are going to describe for the set cover problem is actually an explicit linear program (see section 4.2) that reproduces the BZ hierarchy as a polynomial-sized, efficiently constructible extended formulation that satisfies all constant pitch inequalities.

*Paper structure.* In order to make this article as self-contained as possible and accessible to nonexpert readers, in section 2 we give a basic introduction to SoS proofs/relaxations. However, we provide an introduction from a more general point of view, namely in terms of a generic subspace of polynomials. This is the "non-standard" flavor that will be advocated in this paper.

In section 3 we consider a family of elementary CG-cuts that are "hard" for the standard Lasserre/$d$-SoS relaxation. More precisely, for every $L$, we show that there exists $\varepsilon > 0$ such that the set $\{x \in [0,1]^n : \sum_{i=1}^n x_i \geq L + \varepsilon\}$ has Lasserre rank at least $n - L$. On the other side, this can be easily fixed by using a different basis of high degree polynomials.

Our main application is given in section 4, where we show that the SoS framework equipped with a suitably chosen polynomial-size spanning set of high degree polynomials produces a relaxation, actually a compact linear program, for set cover problems for which all valid inequalities of a given, fixed pitch hold (Theorem 4.2). The general BZ approach is discussed in section 5.

In section 6, we give the packing analogue of Theorem 4.2. In this case the standard SoS hierarchy is sufficient. Moreover, we show that the optimal value of maximizing a linear function over the $d$th CG closure of a packing polytope (an NP-hard problem in general) can be approximated, to arbitrary precision and in polynomial time, by using the standard SoS hierarchy.

Final remarks and future directions are given in section 7. The full version of the paper can be found in [23].

**2. Sum-of-squares proofs and relaxations.** In this section we give a brief introduction to SoS proofs/relaxations. We refer to the monograph [19] for an excellent in-depth overview. We emphasize that there is no mathematical innovation in this section; all the details herein are basically known. However, instead of the "standard" SoS description in terms of bounded degree monomials, we provide a definition as a function of a generic subspace of polynomials. This is used in the remainder of the paper.

We will use the following notation. Let $\mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n]$ be the ring of polynomials over the reals in $n$ variables. Let $\mathbb{R}[x]_d$ denote the subspace of $\mathbb{R}[x]$ of degree at most $d \in \mathbb{N}$. If $\mathcal{S} = \{s_1, \ldots, s_k\}$ is a set of polynomials in $\mathbb{R}[x]$, then the *span* of $\mathcal{S}$, denoted $\langle \mathcal{S} \rangle$, is the set of all linear combinations of the polynomials in $\mathcal{S}$, i.e., $\langle \mathcal{S} \rangle := \{\sum_{i=1}^k c_i \cdot s_i : c_i \in \mathbb{R}\}$, and $\mathcal{S}$ is called the *spanning set* of $\langle \mathcal{S} \rangle$.

[2] I thank Levent Tunçel for pointing out his work to me [1].

The set $\mathcal{F}$ of feasible solutions of an optimization problem is usually described by a finite number of polynomial equations and/or inequalities. This is formalized by the following definition. Let $\mathcal{F} \subset \mathbb{R}^n$ be defined as

$$\mathcal{F} = \{x \in \mathbb{R}^n : f_i(x) = 0 \ \forall i \in [\ell], g_j(x) \geq 0 \ \forall j \in [m]\}, \tag{2.1}$$

where for each $i \in [\ell]$ and $j \in [m]$, $f_i(x), g_j(x) \in \mathbb{R}[x]$ and where $[\ell]$ denote $\{1, 2, \ldots, \ell\}$. Here, $\mathcal{F}$ is called a *basic closed semialgebraic set.* For the sake of brevity, throughout this document, while referring to a semialgebraic set, we implicitly assume a *basic closed* semialgebraic set.

One could write many other constraints that are equally valid on the set $\mathcal{F}$. For example, we are able to produce further polynomials vanishing on the set $\mathcal{F}$ by considering linear combinations of $f_i(x)$ with polynomial coefficients. The set of all polynomials generated this way is a polynomial ideal.

Definition 2.1. *The* ideal *generated by a finite set* $\{f_1, \ldots, f_\ell\}$ *of polynomials in* $\mathbb{R}[x]$ *is defined as*

$$\mathbf{I}(f_1, \ldots, f_\ell) := \left\{ \sum_{i=1}^{\ell} t_i \cdot f_i \ : \ t_1, \ldots, t_\ell \in \mathbb{R}[x] \right\}.$$

A polynomial $p \in \mathbb{R}[x]$ is an SoS if it can be written as the SoS of some other polynomials. If these last polynomials belong to a subspace $\langle \mathcal{S} \rangle \subseteq \mathbb{R}[x]$, for a given spanning set $\mathcal{S} \subseteq \mathbb{R}[x]$, then we say that $p$ is $\mathcal{S}$-SoS.

Definition 2.2. *For* $\mathcal{S} \subseteq \mathbb{R}[x]$*, a polynomial* $p \in \mathbb{R}[x]$ *is* $\mathcal{S}$-SoS *if* $p \in \Sigma_{\mathcal{S}}$ *where*

$$\Sigma_{\mathcal{S}} := \left\{ p \in \mathbb{R}[x] : p = \sum_{i=1}^{r} q_i^2 \text{ for some } r \in \mathbb{N} \text{ and } q_1, \ldots, q_r \in \langle \mathcal{S} \rangle \right\}.$$

As for the vanishing polynomials on $\mathcal{F}$, we are able to produce further valid inequalities for set $\mathcal{F}$ by multiplying $g_j(x)$ against SoS polynomials, or by taking conic combinations of valid constraints. This gives the notion of quadratic module.

Definition 2.3. *For* $\mathcal{S} \subseteq \mathbb{R}[x]$*, the* $\mathcal{S}$-quadratic module *generated by a finite set* $\{g_1, \ldots, g_m\}$ *of polynomials in* $\mathbb{R}[x]$ *is defined as*

$$\mathbf{qmodule}_{\mathcal{S}}(g_1, \ldots, g_m) := \left\{ s_0 + \sum_{i=1}^{m} s_i \cdot g_i : s_0, s_1, \ldots, s_m \in \Sigma_{\mathcal{S}} \right\}.$$

Certifying that a polynomial $p \in \mathbb{R}[x]$ is nonnegative over a semialgebraic set $\mathcal{F}$ is an important problem in optimization, as certificates of nonnegativity can often be leveraged into optimization algorithms. For example, let $p := p' - \gamma$, where $p' \in \mathbb{R}[x]$ and $\gamma$ is a real number. If we can certify that $p$ is nonnegative over $\mathcal{F}$, then the infimum of $p'$ is not smaller than $\gamma$. We will elaborate more on this in section 2.1.

Definition 2.4. *For* $\mathcal{S} \subseteq \mathbb{R}[x]$ *and* $p(x) \in \mathbb{R}[x]$*, an* $\mathcal{S}$-SoS *certificate of nonnegativity of* $p(x)$ *over* $\mathcal{F}$ *(see* (2.1)*) is given by a polynomial identity of the form*

$$p(x) = f(x) + g(x) \tag{2.2}$$

*for some* $f(x) \in \mathbf{I}(f_1, \ldots, f_\ell)$ *and* $g(x) \in \mathbf{qmodule}_{\mathcal{S}}(g_1, \ldots, g_m)$*.*

Notice that for all $x \in \mathcal{F}$ the right-hand side of (2.2) is manifestly nonnegative, thereby certifying that $p(x) \geq 0$ over $\mathcal{F}$.

In the following, whenever $\mathcal{S} = \mathbb{R}[x]$, we drop $\mathcal{S}$ from the notation. So $\Sigma$, SoS and $\mathbf{qmodule}(g_1, \ldots, g_m)$ denote $\Sigma_{\mathbb{R}[x]}$, $\mathbb{R}[x]$-SoS and $\mathbf{qmodule}_{\mathbb{R}[x]}(g_1, \ldots, g_m)$, respectively.

A natural question arises: Can all valid constraints be generated this way? Unless further assumptions are made, the answer is negative (see, e.g., [6]). However, for the applications of this paper, we are interested in the case in which $\mathcal{F}$ is the set of feasible solutions of a 0/1 integer linear program, with $n$ variables and $m$ linear constraints:

$$\mathcal{F}_{01} := \left\{x \in \mathbb{R}^n : x_i^2 - x_i = 0 \ \forall i \in [n],\ g_j(x) \geq 0 \ \forall j \in [m]\right\}, \tag{2.3}$$

where $x_i^2 - x_i = 0$ encodes $x_i \in \{0, 1\}$ and each constraint $g_j(x) \geq 0$ is linear. Under this assumption the answer of the above question is positive, as shown in the following. (Actually, the linearity of the constraints is not necessary for this purpose.) We review this derivation from a slightly different perspective, by highlighting several aspects that will play a role in our proofs.

We start with some preliminaries. The set of polynomials in $\mathbb{R}[x]$ that vanish on the Boolean hypercube $\mathbb{Z}_2^n$ is the ideal

$$\mathbf{I}_{01} := \mathbf{I}\left(x_1^2 - x_1, \ldots, x_n^2 - x_n\right).$$

Definition 2.5. *Let* $\mathbf{I}$ *be an ideal, and let* $f, g \in \mathbb{R}[x]$. *We say that* $f$ *and* $g$ *are* congruent modulo $\mathbf{I}$, *written* $f \equiv g \pmod{\mathbf{I}}$, *if* $f - g \in \mathbf{I}$.

From the above definition, an $\mathcal{S}$-SoS certificate of nonnegativity of $p(x)$ over $\mathcal{F}_{01}$ is given by a polynomial congruence of the form

$$p(x) \equiv g(x) \pmod{\mathbf{I}_{01}} \tag{2.4}$$

for some $g(x) \in \mathbf{qmodule}_{\mathcal{S}}(g_1, \ldots, g_m)$. For the sake of brevity, whenever we use "$\equiv$" we assume that the congruence is modulo $\mathbf{I}_{01}$ (unless differently specified).

Let us introduce an indicator multilinear polynomial that will play an important role throughout this paper. For $I \subseteq Z \subseteq [n]$, the *Kronecker delta* polynomial is defined as

$$\delta_I^Z := \prod_{i \in I} x_i \prod_{j \in Z \setminus I} (1 - x_j). \tag{2.5}$$

If $Z = \emptyset$ we assume that $\delta_I^Z = 1$. Let $x_I^Z$ denote the 0/1 (partial) assignment with $x_i = 1$ for $i \in I$, and $x_j = 0$ for $j \in Z \setminus I$. Notice that $\delta_I^Z$ is an indicator polynomial that is 1 when its variables get assigned values according to $x_I^Z$. Moreover, the following identities hold:

$$\sum_{I \subseteq Z} \delta_I^Z = 1, \tag{2.6}$$

$$\left(\delta_I^Z\right)^2 \equiv \delta_I^Z, \tag{2.7}$$

$$\delta_I^Z \delta_J^Z \equiv 0 \text{ for } I, J \subseteq Z \text{ with } I \neq J. \tag{2.8}$$

By using (2.7) and (2.8) we have (for $Z \subseteq [n]$ and $W \subseteq 2^Z$)

$$\left(\sum_{I \in W} \delta_I^Z\right)^2 \equiv \sum_{I \in W} \delta_I^Z. \tag{2.9}$$

For any given $p(x) \in \mathbb{R}[x]$, let us use $p(x_I^Z)$ to denote $p(x)$ after the partial assignment defined by $x_I^Z$: for example, if $p(x) = p_0 + \sum_{i=1}^n p_i \cdot x_i$, then $p(x_I^Z) = p_0 + \sum_{i\in I} p_i + \sum_{[n]\setminus Z} p_i \cdot x_i$. Then the following holds:

$$\delta_I^Z p(x) \equiv \delta_I^Z p\left(x_I^Z\right). \tag{2.10}$$

These basic facts will be used several times.

SoS *proofs over the Boolean hypercube.* For any given polynomial $p(x) \in \mathbb{R}[x]$ that is nonnegative over $\mathcal{F}_{01}$, we are interested in certifying this property by exhibiting an SoS certificate. With this aim, partition the Boolean hypercube into two sets $N^+ := \{I \subseteq [n] : p(x_I^{[n]}) \geq 0\}$ and $N^- := \{I \subseteq [n] : p(x_I^{[n]}) < 0\}$. If $p(x)$ is nonnegative over $\mathcal{F}_{01}$, then for each $I \in N^-$ there exists a constraint that is violated on $x_I^{[n]}$, i.e., there is a mapping $h : N^- \to [m]$ such that $g_{h(I)}(x_I^{[n]}) < 0$. To ease the notation, we drop the exponent "$[n]$" from $x_I^{[n]}$ and $\delta_I^{[n]}$. Then,

$$\begin{aligned} p(x) &= \overbrace{\left(\sum_{I\subseteq[n]} \delta_I\right)}^{=1 \text{ by } (2.6)} p(x) \overset{\text{by } (2.10)}{\equiv} \sum_{I\in N^+} \delta_I p(x_I) + \sum_{I\in N^-} \delta_I \frac{p(x_I)}{g_{h(I)}(x_I)} g_{h(I)}(x_I) \\ &\overset{\text{by (2.9) and (2.10)}}{\equiv} \underbrace{\left(\sum_{I\in N^+} \delta_I \sqrt{p(x_I)}\right)^2}_{s_0} + \sum_{I\in N^-} \underbrace{\left(\delta_I \sqrt{\frac{p(x_I)}{g_{h(I)}(x_I)}}\right)^2}_{s_{h(I)}} g_{h(I)}(x). \end{aligned} \tag{2.11}$$

It follows that any nonnegative polynomial over $\mathcal{F}_{01}$ admits an $\mathcal{S}$-SoS certificate where $\mathcal{S}$ is the set $\{\delta_I : I \subseteq [n]\}$ of Kronecker delta multilinear polynomials. The quotient ring $\mathbb{R}[x]/\mathbf{I}_{01}$ is the set of equivalence classes for congruence modulo $\mathbf{I}_{01}$. Polynomials from the quotient ring $\mathbb{R}[x]/\mathbf{I}_{01}$ are in bijection with square-free (also known as multilinear) polynomials in $\mathbb{R}[x]$. We will use $\mathbb{R}[x]/\mathbf{I}_{01}$ to denote the subspace of multilinear polynomials. The aforementioned Kronecker delta polynomials form a basis for the subspace of multilinear polynomials $\mathbb{R}[x]/\mathbf{I}_{01}$. The next proposition summarizes the above.

PROPOSITION 2.6. *Let $\langle \mathcal{S} \rangle = \mathbb{R}[x]/\mathbf{I}_{01}$. If $p(x) \in \mathbb{R}[x]$ is nonnegative over $\mathcal{F}_{01}$, then it admits an $\mathcal{S}$-SoS certificate of the form*

$$p(x) \equiv g(x) \pmod{\mathbf{I}_{01}} \tag{2.12}$$

*for some $g(x) \in \mathbf{qmodule}_{\mathcal{S}}(g_1, \ldots, g_m)$.*

The existence of an $\mathcal{S}$-SoS certificate can be decided by solving a semidefinite programming (SDP) feasibility problem whose matrix dimension is bounded by $O(|\mathcal{S}|)$. We refer to [6, 9] and [23] for details and an example.

If $\langle \mathcal{S} \rangle = \mathbb{R}[x]/\mathbf{I}_{01}$, then the SDP has exponential size. The "standard," namely the "most used" way to bound the complexity, is to restrict the spanning set $\mathcal{S}$ of $\mathcal{S}$-SoS certificates to be the standard monomial basis of constant degree $d = O(1)$. This bounds the degrees of the polynomials in $\mathcal{S}$-SoS certificates to be a constant, and a nonnegativity certificate is computed by solving a semidefinite program of size $n^{O(d)}$. Clearly this restriction imposes severe limitations on the kind of proofs that can be obtained. This type of approach was first proposed by Shor [31], and the idea was taken further by Parrilo [27] and Lasserre [17].

However, this modus operandi with bounded degree monomials can be extended to other subspaces $\langle \mathcal{S} \rangle$ having "small" spanning sets $\mathcal{S}$, i.e., with $|\mathcal{S}| = n^{O(d)}$ for some $d = O(1)$. This is a less explored direction and it will play a key role in this paper.

**2.1. 0/1 optimization and SoS relaxations.** As already remarked, a number $\gamma$ is a global lower bound of a polynomial $p(x)$ over $\mathcal{F}_{01}$ if and only if $p(x) - \gamma$ is nonnegative over $\mathcal{F}_{01}$. For 0/1 problems, without loss of generality (w.l.o.g.), we can assume that $p(x)$ is in multilinear form and therefore $p(x) \in \mathbb{R}[x]_n$. For $\langle \mathcal{S} \rangle \subseteq \mathbb{R}[x]/\mathbf{I}_{01}$, a relaxation of the above optimization problem is obtained by computing the $\sup \gamma$ such that $p(x) - \gamma$ has an $\mathcal{S}$-SoS certificate of nonnegativity:

$$\sup_{\gamma} \{\gamma : p(x) - \gamma \in \mathcal{C}_{\mathcal{S}}\}, \tag{2.13}$$

where

$$\mathcal{C}_{\mathcal{S}} := \{q + r : q \in \mathbf{qmodule}_{\mathcal{S}}(g_1, \ldots, g_m), r \in \mathbf{I}_{01} \cap \mathbb{R}[n]_{2n}\} \tag{2.14}$$

is a set of $\mathcal{S}$-SoS certificates. Note that (2.13) is indeed an approximation, since it could be that $p(x) - \gamma$ is nonnegative for some $\gamma$, but the set $\mathcal{S}$ is "too small" so that an $\mathcal{S}$-SoS certificate does not exist. However, enlarging $\mathcal{S}$ increases the number of possible certificates and thus tightens the approximation. For 0/1 semialgebraic sets and multilinear $p(x)$, we can always reduce to the case where the polynomials of SoS certificates have degree at most $2n$, since for $\langle \mathcal{S} \rangle = \mathbb{R}[x]/\mathbf{I}_{01}$, the relaxation (2.13) is actually exact, as shown in (2.11). This explains why we can restrict $r \in \mathbf{I}_{01} \cap \mathbb{R}[n]_{2n}$ in (2.14).

**2.2. Duality and the Lasserre/SoS hierarchy.** The linear space of all real polynomials of $n$ variables and degree at most $d$ is isomorphic to the Euclidean space $\mathbb{R}^{\binom{n+d}{d}}$. Indeed, a simple combinatorial argument shows that any degree-$d$ polynomial $p(x)$ can have at most $\binom{n+d}{d}$ monomials, which we can order in some arbitrary way (ordered basis). Then, we can put the coefficients in a column vector $p$, in the selected order, and thus obtain a bijective mapping to $\mathbb{R}^{\binom{n+d}{d}}$. We will say that $p$ is the *column vector representation* of $p(x)$ in the (ordered) *standard monomial basis*. Then, for $\mathcal{S} \subseteq \mathbb{R}[x]/\mathbf{I}_{01}$, set $\mathcal{C}_{\mathcal{S}}$ (see (2.14)) is (isomorphic to) a subset of $\mathbb{R}^{\binom{n+2n}{2n}}$, and it can be shown to form a *cone* in the sense of convex geometry.

We emphasize that in the above arguments we have chosen the standard monomials as a basis for the column vector representation of polynomials. It is clear that other bases are possible. Actually, for our main application we will use a different basis. More generally, any linear space $V$ (of polynomials) is isomorphic to the space of column vectors of a certain dimension: choose an ordered basis $b^\top = (b_1, \ldots, b_k)$ for the linear space $V$ (of polynomials); the column vector representation of $p(x) \in V$ is a vector $p \in \mathbb{R}^k$ such that $p(x) = b^\top p$.

*Dual program.* Recall that in linear algebra, a *linear functional* $y$ is a linear map from a linear space $V$ to its field $\mathbb{F}$ of scalars. A linear functional $y$ is a *linear* function:

$$y(\alpha \cdot v + \beta \cdot w) = \alpha \cdot y(v) + \beta \cdot y(w) \quad \forall v, w \in V, \ \forall \alpha, \beta \in \mathbb{F}. \tag{2.15}$$

In $\mathbb{R}^k$, for $k \in \mathbb{N}$, linear functionals are represented as vectors and their action on vectors is given by the inner product: letting $y, z \in \mathbb{R}^k$, the evaluation of $y$ at $z$ is denoted by the inner product $\langle z, y \rangle$, that is, $\langle z, y \rangle = y(z)$. Let $\mathcal{C}$ be a set in $\mathbb{R}^k$ equipped with an inner product $\langle z, y \rangle = y(z)$. The *dual* cone[3] of $\mathcal{C}$ is defined by

[3]Recall, in finite dimension, *topological* and *algebraic* duals are the same.

$$\mathcal{C}^* = \{y \in \mathbb{R}^k : y(z) \geq 0 \ \forall z \in \mathcal{C}\}. \tag{2.16}$$

In other words, the dual cone is the set of linear functionals that are nonnegative on the primal cone. Consider a standard conic program over a cone $\mathcal{C}$ and its dual:

$$\textbf{Primal} : \sup_z \{\langle c, z\rangle : p - Az \in \mathcal{C}\}; \quad \textbf{Dual} : \inf_y \{\langle p, y\rangle : A^\top y = c; y \in \mathcal{C}^*\}. \tag{2.17}$$

To find the dual program of (2.13) as a conic optimization problem, choose an (ordered) *basis* for the polynomials in $\mathcal{C}_\mathcal{S}$ (we will say a little bit more about this later). The dimension of this basis defines the dimension of the linear functionals $y$: there is one entry in $y$ for each polynomial in the basis. Set $p$ in (2.17) to be the column vector representation of polynomial $p(x)$ in (2.13) according to the chosen (ordered) basis. Consider representing the variable $\gamma$ as the constant term of a polynomial $z(x)$. Let $z$ be the column vector representation of $z(x)$ and maximize its inner product with a suitably chosen vector $c$ so that $\langle c, z\rangle = \gamma$.

For the standard monomial basis choose $c = (1, 0, \ldots, 0)^\top$ and the matrix $A$ such that $A_{0,0} = 1$ and $A_{i,j} = 0$ elsewhere. So under this choice, we get as the dual

$$\inf_y \{\langle p, y\rangle : y_0 = 1; \ y \in \mathcal{C}^*_\mathcal{S}\}. \tag{2.18}$$

The dual cone $\mathcal{C}^*_\mathcal{S}$ of $\mathcal{C}_\mathcal{S}$ turns out to have some nice properties, as explained below. For any given polynomial $p(x) \in \mathbb{R}[x]$, we will use $y[p(x)]$ to denote $y(p)$ (or $\langle p, y\rangle$), where $p$ is the column vector representation of $p(x)$ according to the chosen (ordered) basis, and $y$ is a linear functional. With respect to any chosen vector basis for the polynomials from $\mathcal{C}_\mathcal{S}$, the elements of the dual space $\mathcal{C}^*_\mathcal{S}$ define linear functionals $y[\cdot]$ (sometimes called *pseudoexpectation* functionals and denoted with $\tilde{\mathbb{E}}[\cdot]$) on polynomials that satisfy

(1) (Normalization) $y[1] = 1$;
(2) (Linearity) $y[\alpha \cdot p(x) + \beta \cdot q(x)] = \alpha \cdot y[p(x)] + \beta \cdot y[q(x)]$ for all $p(x), q(x) \in \mathcal{C}_\mathcal{S}$ and $\alpha, \beta \in \mathbb{R}$;
(3) (Positivity) $y\big[q(x)^2\big] \geq 0$ for all $q(x) \in \langle\mathcal{S}\rangle$;
(4) (Positivity) $y\big[q(x)^2 \cdot g_i(x)\big] \geq 0$ for all $q(x) \in \langle\mathcal{S}\rangle$, for all $i \in [m]$;
(5) (Multilinearity) $y\big[t(x) \cdot (x_i^2 - x_i)\big] = 0$ for all $t(x) \in \mathbb{R}[x]$, for all $i \in [n]$.

Condition (1) says that the constant polynomial 1 is mapped to 1. Note that in (2.18), $y_0 = 1$ comes directly from (1) (in the standard monomial basis we have $y[1] = \big\langle (1, 0, \ldots, 0)^\top, y\big\rangle = y_0$).

Condition (2) follows from the linearity of linear functionals (see (2.15)). Note that assigning arbitrary values to the entries of the linear functional $y$ guarantees linearity. Indeed, the entries of $y$ are linearly independent because they correspond to the "linearization" of the polynomials that form a basis for $\mathcal{C}_\mathcal{S}$, which are linearly independent. This is the only place where we need linear independence. Alternatively, we can choose a spanning set of polynomials for $\mathcal{C}_\mathcal{S}$ and impose the linearity condition (2).

Conditions (3), (4), and (5) follow from the definition of the dual cone (see (2.16)) of $\mathcal{C}_\mathcal{S}$ (see (2.14)). Note that the multilinearity condition (5) can be easily enforced by restricting to multilinear polynomials: any given polynomial $p(x)$ will be replaced by its *multilinear form*, denoted $\overline{p(x)}$, i.e., the normal form after polynomial division by the Gröbner basis $\{x_i^2 - x_i : i \in [n]\}$.[4] So in conditions (3) and (4), we replace $q(x)^2$

[4]The multilinear form of $p(x)$ is obtained by replacing every occurrence in $p(x)$ of "$x_i^k$" with "$x_i$" whenever $i \in [n]$ and $k \geq 2$; for example, $x_1 \cdot x_2 + 2 \cdot x_2$ is the multilinear form of $x_1^3 \cdot x_2 + 2 \cdot x_2^2$.

and $q(x)^2 \cdot g_i(x)$ with their multilinear forms $\overline{q(x)^2}$ and $\overline{q(x)^2 \cdot g_i(x)}$, respectively. From now on, we will restrict to the subspace of multilinear polynomials $\mathbb{R}[x]/\mathbf{I}_{01}$. This allows us to enforce the multilinearity condition (5).

By the above arguments, we can restrict to the polynomials from $\mathcal{C}_\mathcal{S}$ that are multilinear. These polynomials are spanned by the set of multilinear polynomials

$$T := \{\overline{x_i \cdot p \cdot q} : p, q \in \mathcal{S}, i \in [n] \cup \{0\}\},$$

where $x_0 := 1$, and recall that we are considering linear constraints $g_i(x) \geq 0$ for $i \in [m]$. So, the vector $y$ has one entry for each polynomial that belongs to a chosen basis for the span $\langle T \rangle$ of $T$. By assuming this, we can reformulate (2.18), with respect to a chosen basis for $\langle T \rangle$, as

$$\inf\ y[p] \tag{2.19}$$
$$s.t.\ y[1] = 1; \tag{2.20}$$
$$y\left[\overline{q(x)^2}\right] \geq 0 \quad \forall q(x) \in \langle \mathcal{S} \rangle; \tag{2.21}$$
$$y\left[\overline{q(x)^2 \cdot g_i(x)}\right] \geq 0 \quad \forall q(x) \in \langle \mathcal{S} \rangle, \forall i \in [m]. \tag{2.22}$$

The program (2.19)–(2.22) is actually a *semidefinite program* whose matrix dimension is bounded by $O(|\mathcal{S}|)$, which we call $\mathcal{S}$-SoS *relaxation*. To see this, let $b^\top = (b_1, \ldots, b_k)$ be an (ordered) basis for $\langle \mathcal{S} \rangle$ for some $k \leq |\mathcal{S}|$. Then, consider any polynomial $q(x) \in \langle \mathcal{S} \rangle$, and let $q$ be its column vector representation according to the (ordered) basis $b^\top$, i.e., $q(x) = b^\top q$. Then $q(x)^2 = \langle qq^\top, bb^\top \rangle$. Let $M(y)$ be a $|b| \times |b|$ square matrix indexed by the pairs $(b_i, b_j) \in b \times b$ such that the $(b_i, b_j)$th entry of $M(y)$ is equal to $y[\overline{b_i b_j}]$. Recall that $y[q(x)^2]$ is equal to $\langle y, p \rangle$, where $p$ is the column vector representation of $q(x)^2$. By simple inspection note that $y[\overline{q(x)^2}] = \langle qq^\top, M(y) \rangle$. It follows that condition (2.21) is equivalent to imposing $\langle qq^\top, M(y) \rangle \geq 0$ for all $q$, which is equivalent to requiring $M(y)$ to be positive semidefinite. A similar argument holds for condition (2.22).

*Standard and generalized* SoS *relaxations.* When $\mathcal{S}$ is the standard (multilinear) monomial basis of degree $\leq d$, then the $\mathcal{S}$-SoS relaxation (2.19)–(2.22) is the (standard) *Lasserre/*SoS*-hierarchy* parameterized by the degree $d \in \mathbb{N}$, in short, denoted by $d$-SoS. $\mathcal{S}$-SoS generalizes $d$-SoS relaxations by working with a generic set $\mathcal{S}$ of polynomials. In this case, the aforementioned matrix $M(y)$ is the so-called (truncated) moment matrix.

Note that in standard SoS, set $T$ forms a basis for $\langle T \rangle$, and it is the set of all (multilinear) monomials of degree at most $2d + 1$. The variables in $d$-SoS are the entries of the linear functionals $y$, which correspond to the "linearization" of the polynomials from $T$.

*Standard and generalized Sherali–Adams relaxations.* If $\mathcal{S}$ is again the standard monomial basis of degree $\leq d$ *and* we further relax (2.21) and (2.22) by

$$y\left[\overline{q(x)}\right] \geq 0 \quad \forall q(x) \in \mathcal{S}, \tag{2.23}$$
$$y\left[\overline{q(x) \cdot g_i(x)}\right] \geq 0 \quad \forall q(x) \in \mathcal{S}, \forall i \in [m], \tag{2.24}$$

then we obtain the so-called Sherali–Adams hierarchy of relaxations, denoted $d$-SA and defined by (2.19), (2.20), (2.23), (2.24). This is again parameterized by $d$, but it is a *linear program* (this follows from (2.23), (2.24), and the definition of linear functionals where their action on vectors is given by the dot product) of size $O(|\mathcal{S}|) = n^{O(d)}$. Note that both hierarchies, $d$-SoS and $d$-SA, have the same spanning set $\mathcal{S}$ of monomials, which are *nonnegative* over the Boolean hypercube.

In the definition of $d$-SA we restrict to working with polynomials from $\overline{q(x) \cdot g_i(x)}$ for $q(x) \in \mathcal{S}$, and $i \in [m]$. Let

$$T_{SA} := \{\overline{x_i \cdot p} : p \in \mathcal{S}, i \in [n] \cup \{0\}\}.$$

When $\mathcal{S}$ is the standard monomial basis then $T_{SA}$ is a basis for $\langle T_{SA} \rangle$, and it is the set of all (multilinear) monomials of degree at most $d+1$. It follows that the variables in $d$-SA are the entries of the linear functionals $y$, which correspond to the "linearization" of the polynomials from $T_{SA}$.

We generalize $d$-SA relaxations to work with a generic set $\mathcal{S}$ of polynomials (that is nonnegative over the Boolean hypercube) and obtain $\mathcal{S}$-SA. The relaxation $\mathcal{S}$-SA is a linear program with $O(|\mathcal{S}|)$ linear constraints, which correspond to (2.20), (2.23), and (2.24).

We conclude our overview on SoS relaxations by pointing out the following fact.

PROPOSITION 2.7. *If $p(x)$ admits an $\mathcal{S}$-SoS certificate of nonnegativity over $\mathcal{F}_{01}$, then $y[p(x)] \geq 0$ holds for the corresponding $\mathcal{S}$-SoS$(\mathcal{F}_{01})$ relaxation* (2.20)–(2.22).

*Proof.* By assumption, for some $f(x) \in \mathbf{I}_{01}$ and $g(x) \in \mathbf{qmodule}_{\mathcal{S}}(g_1, \ldots, g_m)$, we have $p(x) = f(x) + g(x)$. Then, $y[p(x)] = y[f(x)] + y[g(x)] = 0 + y[s_0] + \sum_{i=1}^{m} y[s_i \cdot g_i]$ for some $s_0, s_1, \ldots, s_m \in \Sigma_{\mathcal{S}}$. By (2.21) and (2.22), each addend of the sum is nonnegative and we have $y[p(x)] \geq 0$. ∎

By Proposition 2.7, if $p(x) := \sum_i a_i x_i - a_0 \geq 0$ is a valid linear inequality for all $x \in \mathcal{F}_{01}$ that admits an $\mathcal{S}$-SoS certificate, then $y[p(x)] = \sum_i a_i y[x_i] - a_0 \geq 0$. Note that $\{y[x_1], \ldots, y[x_n]\}$ is the solution $y$ of (2.20)–(2.22) projected to the original space of the variables. So, the (projected) solution of the $\mathcal{S}$-SoS relaxation (2.20)–(2.22) satisfies $p(x) \geq 0$. This implies the following informal "recipe" that we will follow in the remainder of the paper. (Similar arguments hold for $\mathcal{S}$-SA.)

**Recipe**. Assume that we are looking for a "small" relaxation for $\mathcal{F}_{01}$ that satisfies a potentially "large" set of linear constraints $Ax \geq b$ that are valid for all $x \in \mathcal{F}_{01}$. With this aim, search for a "small" spanning set $\mathcal{S} \subseteq \mathbb{R}[x]$ (if one exists) such that $Ax - b$ admits an $\mathcal{S}$-SoS certificate. If we succeed, then the corresponding $\mathcal{S}$-SoS relaxation (2.20)–(2.22) satisfies our goal.

## 3. A simple Chvátal–Gomory cut that is hard for *d*-SoS.

For illustrative purposes, in this section we consider a simple example where the standard Lasserre/$d$-SoS relaxation provably fails for "large" $d$. However, this can be easily fixed by using $\mathcal{S}$-SoS with a "small" spanning set $\mathcal{S}$ of high degree polynomials.

The example is motivated by the following situation. Consider the rational polyhedra $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ with $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$. Inequalities of the form $(\lambda^\top A)x \geq \lceil \lambda^\top b \rceil$, with $\lambda \in \mathbb{R}^m_+$, $\lambda^\top A \in \mathbb{Z}^n$, and $\lambda^\top b \notin \mathbb{Z}$, are commonly referred to as Chvátal–Gomory cuts or CG-cuts (further information on CG-cuts is provided in section 6). It is a natural question to study how many levels (or degree $d$) of the "standard" SoS hierarchy, i.e., $d$-SoS, are necessary to strengthen $(\lambda^\top A)x \geq \lambda^\top b$ to get $(\lambda^\top A)x \geq \lceil \lambda^\top b \rceil$. With this aim, consider the following semialgebraic set:

$$\mathcal{F}_{01} = \left\{ x \in \mathbb{R}^n : x_i^2 - x_i = 0 \ \forall i \in [n], \sum_{i=1}^n x_i \geq b \right\}, \tag{3.1}$$

where $b \in \mathbb{Q}_+$ is intended to be a positive fractional number. Obviously, any feasible integral solution satisfies $\sum_{i=1}^n x_i \geq \lceil b \rceil$, and this is promptly captured by the first CG closure.

The following Theorem 3.1 (the proof can be found in [23]) shows that regardless of whether $b$ is "small," i.e., $b = O(1)$, or "large," i.e., $b = \Omega(n)$, $d$-SoS$(\mathcal{F}_{01})$ fails to enforce the simple CG-cut when $d = o(n)$.

THEOREM 3.1. *Let $\mathcal{F}_{01}$ be defined as in* (3.1)*, with $P$ sufficiently large (that depends on $n$), $L \in \left\{0, 1, \ldots, \lceil \frac{n}{2} \rceil - 1\right\}$, and $b := L + 1/P$. Then, the $d$-*SoS$(\mathcal{F}_{01})$*-relaxation requires $d \geq n - L$ for enforcing $\sum_{i=1}^n x_i \geq \lceil b \rceil$.*

We remark that Grigoriev, Hirsch, and Pasechnik gave in [11] a very interesting and influential result that is related to our Theorem 3.1 but is *significatively different* in terms of both lower bounds and techniques. We defer the interested reader to section 3.1 for a discussion on this point and for a more precise meaning of "significatively different."

The result in Theorem 3.1 is disappointing for at least two reasons: the considered CG-cut looks pathetically trivial, and the proof that $d$-SoS$(\mathcal{F}_{01})$ fails for small $d$ is relatively complicated (see [23]).

On the other side, it would be sufficient to have in the "bag" $\langle \mathcal{S} \rangle$ the set of symmetric polynomials, i.e., polynomials which do not change under permutations of the variables, to promptly enforce this CG-cut within $\mathcal{S}$-SoS$(\mathcal{F}_{01})$. The proof is basically the same as the one given in (2.11):

$$\sum_{i=1}^n x_i - \lceil b \rceil = \overbrace{\left( \sum_{i=0}^n \sum_{I \subseteq [n]:|I|=i} \delta_I \right)}^{=1} \left( \sum_{i=1}^n x_i - \lceil b \rceil \right) \overset{(2.10)}{\equiv} \sum_{i=0}^n \overbrace{\left( \sum_{I \subseteq [n]:|I|=i} \delta_I \right)}^{\text{symmetric}} (i - \lceil b \rceil)$$

$$\overset{(2.9)}{\equiv} \underbrace{\sum_{i=\lceil b \rceil}^n \left( \overbrace{\sum_{I \subseteq [n]:|I|=i} \delta_I}^{\text{symmetric}} \sqrt{i - \lceil b \rceil} \right)^2}_{s_0(x)} + \underbrace{\sum_{i=0}^{\lceil b \rceil - 1} \left( \overbrace{\sum_{I \subseteq [n]:|I|=i} \delta_I}^{\text{symmetric}} \sqrt{\frac{i - \lceil b \rceil}{i - b}} \right)^2}_{s_1(x)} \underbrace{\left( \sum_{i=1}^n x_i - b \right)}_{g_1(x)}.$$

Note that $s_0(x)$ and $s_1(x)$ are sums of squares of *symmetric* polynomials. It is a well-known fact that every symmetric polynomial can be written uniquely as a polynomial in the $n + 1$ elementary symmetric polynomials (see, e.g., [32]). Therefore, it is sufficient to define $\mathcal{S}$ as the set of elementary symmetric polynomials to guarantee that $\sum_{i=1}^n x_i - \lceil b \rceil$ admits an $\mathcal{S}$-SoS certificate. We refer to [9, 10, 29] for other more interesting symmetric situations.

We emphasize that in this paper we show how to handle some *asymmetric* situations by exploiting the problem structure, which is our main result.

**3.1. On a related result by Grigoriev, Hirsch, and Pasechnik.** Grigoriev, Hirsch, and Pasechnik (see Theorem 8.1 in [11]) gave a result related to Theorem 3.1 but also significatively different, as explained in this section. In [11], the *symmetric knapsack* is defined as follows:

$$\mathcal{F}'_{01} = \left\{ x \in \mathbb{R}^n : x_i^2 - x_i = 0 \ \forall i \in [n], \sum_{i=1}^{n} x_i = b \right\}. \tag{3.2}$$

Note that $\mathcal{F}'_{01}$ is a more constrained version of the set $\mathcal{F}_{01}$ defined in (3.1).

The Positivstellensatz calculus [11] is a proof system for languages consisting of unsolvable systems of polynomial equations. Note that (3.2) is unsolvable when $b$ is a nonintegral value. A degree $d$ infeasibility certificate consists of a set of degree $d$ polynomials, say $\{h_1 \dots, h_l\}$, and a derivation of $\sum_j h_j^2 = -1$ from $\mathcal{F}'_{01}$. Let $\delta$ denote the step function defined as follows:

$$\delta(x) = \begin{cases} 2 & \text{if } x \notin [0, n]; \\ 2k + 4 & \text{if } x \in [k, k+1] \cup [n-k-1, n-k] \text{ for all integers } 0 \le k < n/2. \end{cases}$$

In [11] the following result is proved.

THEOREM 3.2 (see [11]). *Any Positivstellensatz calculus refutation of the symmetric knapsack problem $\mathcal{F}'_{01}$ (see (3.2)) has degree* $\min\{\delta(b), \lceil (n-1)/2 \rceil + 1\}$.

Notice that any Positivstellensatz calculus lower bound for the more constrained set $\mathcal{F}'_{01}$ gives an SoS lower bound for the set $\mathcal{F}_{01}$ defined in (3.1). However, for $b < n/2$, the bounds given by Theorem 3.2 (see [11]), when applied to set $\mathcal{F}_{01}$, are weaker and also considerably weaker than the ones provided by our Theorem 3.1. For example, for any given constant $k$ and $b \in (k, k+1)$, the degree lower bound in Theorem 3.2 is $2k + 4 = O(1)$, whereas by Theorem 3.1 the degree lower bound is $n - k$.

Regarding the technique, Theorem 3.1 is proved by building on a result given in [13]. The latter has been shown to be very powerful in several other situations (see [13, 14] for more examples).

Finally, we observe that the study of the number of levels necessary to strengthen *inequalities*, as in Theorem 3.1, is useful for analyzing the SoS ability to strengthen convex combinations of valid covering inequalities, as explained at the beginning of section 3. Analyzing *equalities*, like in (3.2), is less appropriate for these purposes.

**4. SoS derivation of pitch inequalities for set cover.** In this section we consider set cover problems. For a given $m \times n$ matrix $A$ with 0/1 entries, the feasible region $\mathcal{F}_A$ for the set cover problem is defined by

$$\mathcal{F}_A = \{x \in \mathbb{R}^n : x_i^2 - x_i = 0 \ \forall i \in [n], Ax \ge e\}, \tag{4.1}$$

where $e$ is the vector of 1s. We focus on the concept of *pitch* introduced in [4, 33].

DEFINITION 4.1. *For any given inequality $a^\top x - a_0 \ge 0$, with indices ordered so that $0 < a_1 \le a_2 \le \cdots \le a_h$ and $a_j = 0$ for $j > h$, its* pitch $\pi(a, a_0)$ *is the minimum integer such that $\sum_{i=1}^{\pi(a,a_0)} a_i - a_0 \ge 0$.*

We start emphasizing that valid inequalities for $\mathcal{F}_A$ of pitch at most $\pi$ are "hard" to enforce within "standard" hierarchies of relaxations, and this happens already with the first nontrivial pitch value, namely $\pi = 2$ as shown by the following example.

*Example* 4.1. Consider a set cover instance defined by a full-circulant constraint matrix $FC$ as follows:

$$\mathcal{F}_{FC} = \left\{ x \in \mathbb{R}^n : x_i^2 - x_i = 0 \ \forall i \in [n], \sum_{j \in [n] \setminus \{i\}} x_j \ge 1 \ \forall i \in [n] \right\}. \tag{4.2}$$

Observe that $\sum_{j=1}^{n} x_j \geq 2$ is a pitch 2 valid inequality for the feasible region of this set cover instance. However, to enforce this inequality we need $n-3$ levels for a lifting operator stronger than the Sherali–Adams hierarchy [4] and require at least $d = \Omega(\log^{1-\varepsilon} n)$ [13], with $\varepsilon > 0$ arbitrarily small, for the standard $d$-SoS hierarchy (conjectured to be $n/4$ in [4]).

This instance will be used in the following to exemplify our approach (see Examples 4.2 and 4.3).

Conversely, we show that there is an $\mathcal{S}_A(\pi)$-SoS relaxation, where $\mathcal{S}_A(\pi)$ is a set of high degree polynomials of polynomial size, that satisfies all valid inequalities of constant pitch $\pi = O(1)$.

THEOREM 4.2. *Consider a set cover problem given by a matrix $A$, and let $\pi = O(1)$ be a fixed nonnegative integer. There is a polynomial-size $\mathcal{S}_A(\pi)$-SoS relaxation that satisfies all valid inequalities for $\mathcal{F}_A$ of pitch at most $\pi$.*

Note that the $\mathcal{S}_A(\pi)$-SoS relaxation of Theorem 4.2 is completely determined by defining the set $\mathcal{S}_A(\pi)$ (see section 2.2 for a discussion on the size and on the set of variables that appear in a generic $\mathcal{S}_A(\pi)$-SoS relaxation). A closer look will reveal (see section 4.2) that the $\mathcal{S}_A(\pi)$-SoS relaxation is actually a linear program corresponding to the generalized Sherali–Adams relaxation $\mathcal{S}_A(\pi)$-SA (see section 2.2).

*Preliminaries.* Given a vector $a \in \mathbb{R}^n$, the support of $a$, denoted supp(a), is the set $\{i \in [n] : a_i \neq 0\}$. Let $A_i \subseteq \{1, \ldots, n\}$ be the support of the $i$th row of $A$. By overloading notation, we also use $A_i$ to denote the corresponding set of variables $\{x_j : j \in A_i\}$. We assume that $A$ is *minimal*, i.e., there is no $i \neq j$ such that $A_i \subseteq A_j$.

For any given $T, F \subseteq [n]$ with $T \cap F = \emptyset$, let $\mathcal{F}_{A_{(T,F)}}$ denote the subregion of $\mathcal{F}_A$ with $x_i = 1$ for all $i \in T$, and $x_j = 0$ for all $j \in F$. Let $A_{(T,F)}$ be the matrix that is obtained from $A$ by removing all the rows where $x_i$ appears for $i \in T$ (these constraints are satisfied when $x_i = 1$ for all $i \in T$) and setting to zero the $j$th column for all $j \in F$. We will assume that $A_{(T,F)}$ is minimal by removing the dominated rows. Therefore, $\mathcal{F}_{A_{(T,F)}} = \{x \in \{0,1\}^n : A_{(T,F)}x \geq e, x_i = 1\ \forall i \in T, x_j = 0\ \forall j \in F\}$ and $\mathcal{F}_{A_{(T,F)}} \subseteq \mathcal{F}_A$.

For the sake of simplicity, we add the nonnegative constraints $x_i \geq 0$ for $i \in [n]$ to the set of valid constraints that define the semialgebraic set (4.1). This is not strictly necessary, since $x_i = x_i^2$ and therefore $x_i \geq 0$, but it will simplify the exposition.

### 4.1. Proof of Theorem 4.2.

Let $a^\top x - a_0 \geq 0$ be a valid inequality over $\mathcal{F}_A$ of pitch $\pi(a, a_0) \leq \pi$, with $a \geq 0$. First, we show an SoS certificate of nonnegativity for $a^\top x - a_0$. Then, we collect the polynomials we used in the SoS certificate and put them in the "bag" $\mathcal{S}_A(\pi)$. So, the set of polynomials $\mathcal{S}_A(\pi)$ of Theorem 4.2 will be completely defined at the end of this proof, and its definition will naturally follow from the given SoS certificate.

For the time being, it is sufficient to say that $\mathcal{S}_A(\pi)$ is a set of polynomials of size $(mn)^{O(1)}$ for any fixed $\pi = O(1)$. In $\mathcal{S}_A(\pi)$ every polynomial has the following form: $\sum_{J \in W} \delta_J^V$ for some $V \subseteq [n]$ and $W \subseteq 2^V$. In short, we will say that set $\mathcal{S}_A(\pi)$ is *delta-structured* to denote this structure.

By (2.9), note that $\sum_{J \in W} \delta_J^V \equiv (\sum_{J \in W} \delta_J^V)^2$, and therefore $q(x) \equiv q(x)^2$ for all $q(x) \in \mathcal{S}_A(\pi)$. Moreover, every polynomial in $\mathcal{S}_A(\pi)$ is nonnegative over the Boolean hypercube. In the remainder a certificate of nonnegativity will be congruent (mod $\mathbf{I}_{01}$) to the following form:

(4.3)

$$\sum_i q_i(x) \overbrace{\left(\lambda_i^\top (Ax-e) + \gamma_i^\top x + \mu_i\right)}^{\text{conical combination of constraints}} \qquad \text{for some } q_i(x) \in \mathcal{S}_A(\pi), \lambda_i, \gamma_i, \mu_i \geq 0.$$

By the above properties, this certificate can be immediately transformed into an $\mathcal{S}_A(\pi)$-SoS certificate.

The proof of Theorem 4.2 will be by induction on the pitch value $\pi$. The base of the induction $\pi = 0$ is trivial: in this case we must have $a_0 \leq 0$, and $\mathcal{S}_A(0) = \{1\}$ is sufficient to prove that $-a_0 \geq 0$. Note that $\mathcal{S}_A(0) = \{1\}$ is independent on the matrix $A$ and it is delta-structured (recall if $V = \emptyset$, then $\delta_J^V = 1$).

By the induction hypothesis, for any given $0 \leq p \leq \pi - 1$ and any given constraint matrix $A'$, we assume that any valid pitch-$p$ inequality for $\mathcal{F}_{A'}$ admits an $\mathcal{S}_{A'}(p)$-SoS certificate where $\mathcal{S}_{A'}$ is delta-structured. We will prove that the induction hypothesis also holds for pitch $\pi$ (induction step).

We proceed "backward," as in (2.11). We start multiplying $a^\top x - a_0$ by $\sum_{I\subseteq V} \delta_I^V$ for a suitably chosen set $V \subseteq [n]$ that will be specified soon. Recall that $\sum_{I\subseteq V} \delta_I^V = 1$ (see (2.6)). Let $(a^\top x - a_0)_{(T,F)}$ denote $(a^\top x - a_0)$ after setting $x_i = 1$ for $i \in T$ and $x_j = 0$ for $j \in F$. By (2.10), note that $\delta_J^V(a^\top x - a_0) \equiv \delta_J^V(a^\top x - a_0)_{(J,V\setminus J)}$. Let $\delta_{\geq\pi}^V := \sum_{I\subseteq V, |I|\geq\pi} \delta_I^V$ (zero if $|V| < \pi$). It follows that

(4.4)

$$\begin{aligned} a^\top x - a_0 &= \overbrace{\left(\sum_{I\subseteq V} \delta_I^V\right)}^{=1} (a^\top x - a_0) \\ &\equiv \underbrace{\delta_\emptyset^V (a^\top x - a_0)_{(\emptyset,V)}}_{\textsc{First}} + \underbrace{\left(\sum_{J\subseteq V, 0<|J|<\pi} \delta_J^V (a^\top x - a_0)_{(J,V\setminus J)}\right)}_{\textsc{Second}} + \underbrace{\left(\delta_{\geq\pi}^V\right)\left(a^\top x - a_0\right)}_{\textsc{Third}}. \end{aligned}$$

Therefore, showing an SoS certificate for $a^\top x - a_0$ boils down to providing an SoS certificate for each of the three summands, First, Second, and Third, in (4.4). Before doing this we need to specify the set $V \subseteq [n]$.

*How to choose $V$.* Set $V$ is chosen according to the following Lemma 4.3 (see [4, 33]), which gives a structural property of valid inequalities for set cover. The statement of Lemma 4.3 is slightly different from Proposition 4.22 in [33] (or Theorem 6.3 in [4]). The main difference is given by property (4.7) (see Lemma 4.3). This property is not explicitly given in [4, 33], but it can be easily derived by their construction as explained in the proof that follows.

Lemma 4.3 (see [4, 33]). *Suppose $a^\top x - a_0 \geq 0$ is a valid inequality for $\mathcal{F}_A$ with $a \geq 0$. Then there is a subset $C = C(a, a_0)$ of the rows of $A$ with $|C| \leq \pi(a, a_0)$ such that*

$$A_i \subseteq \text{supp}(a) \quad \forall i \in C, \tag{4.5}$$

$$(a^\top x - a_0)_{(\emptyset,V)} \geq 0 \text{ is valid for } \mathcal{F}_C, \tag{4.6}$$

$$\mathcal{F}_{A_{(\emptyset,V)}} \neq \emptyset, \tag{4.7}$$

*where $V := \bigcup_{i,j\in C, i\neq j} A_i \cap A_j$ is the set of variables occurring in more than one row of $C$, and $\mathcal{F}_C := \{x \in [0,1]^n : (\sum_{j\in A_i} x_j - 1)_{(\emptyset,V)} \geq 0\ \forall i \in C\}$.*

*Proof.* The proof is by induction on $\pi = \pi(a, a_0)$. If $\pi = 0$, then $|C| = 0$, and it follows that $\mathcal{F}_C = \{x : x \in [0,1]^n\}$ and $V = \emptyset$. A pitch zero inequality must have $a_0 \leq 0$. So, since $a \geq 0$, $a^\top x - a_0 \geq 0$ is indeed valid for $\mathcal{F}_C$ and for $\mathcal{F}_{A_{(\emptyset,\emptyset)}} = \mathcal{F}_A$ $(\neq \emptyset)$.

Now, assume that the claim holds for all valid inequalities of pitch $p$ with $0 \leq p \leq \pi - 1$ and $\pi \geq 1$. Consider a valid inequality $a^\top x - a_0 \geq 0$ of pitch $\pi$. Note that there must be some $v \in [m]$ such that $A_v \subseteq \mathrm{supp}(a)$ or, otherwise, we could set $x_j = 0$ for all $j \in \mathrm{supp}(a)$, and $x_j = 1$ everywhere else, and thereby satisfy every constraint and nevertheless have $a^\top x = 0$ (so contradicting the hypothesis that $a^\top x - a_0 \geq 0$ is a valid inequality of pitch $\pi \geq 1$). Choose $A_v \subseteq \mathrm{supp}(a)$. Note that we are assuming, w.l.o.g., that $A$ is minimal, so there is no $A_i$, with $i \in [m]$ and $i \neq v$, that is a proper subset of $A_v$. Let $v(1) \in A_v$ be the index of the minimum coefficient $a_j : j \in A_v$, where $a_j$ is the coefficient of variable $x_j$ in the valid inequality $a^\top x - a_0 \geq 0$.

We first set to zero all the variables from $V_v$, where $V_v$ are all the variables from all $A_i$, with $i \neq v$, that appear in $A_v - \{v(1)\}$, i.e., $V_v := (A_v - \{v(1)\}) \bigcap (\cup_{i \neq v} A_i)$. Consider $\mathcal{F}_{A_{(\emptyset,V_v)}}$ and note that $\mathcal{F}_{A_{(\emptyset,V_v)}} \neq \emptyset$ because by assumption no $A_j \subset A_v$ and therefore $(a^\top x - a_0)_{(\emptyset,V_v)} \geq 0$ is a valid inequality for $\mathcal{F}_{A_{(\emptyset,V_v)}}$. Set $x_{v(1)} = 1$ in $(a^\top x - a_0)_{(\emptyset,V_v)} \geq 0$ to get $(a^\top x - a_0)_{(\{v(1)\},V_v)} \geq 0$, which is a valid inequality for $\mathcal{F}_{A_{(\emptyset,V_v)}}$. Note that the pitch $p$ of $(a^\top x - a_0)_{(\{v(1)\},V_v)}$ is such that $p \leq \pi - 1$ and therefore, by the induction hypothesis, it satisfies the properties of the claim when we consider $(a^\top x - a_0)_{(\{v(1)\},V_v)} \geq 0$ as a valid inequality for $\mathcal{F}_{A_{(\emptyset,V_v)}}$. Let $a'$ be the vector that is obtained from $a$ by setting to zero all the coefficients from $V_v \cup \{v\}$ and let $a'_0 := a_0 - a_{v(1)}$, so $(a^\top x - a_0)_{(\{v(1)\},V_v)} = a'^\top x - a'_0$. By the induction hypothesis there must be a subset $C'$ of the rows from $A' := A_{(\emptyset,V_v)}$ such that $|C'| \leq p$ and

$$A'_i \subseteq \mathrm{supp}(a') \quad \forall i \in C', \tag{4.8}$$

$$\left(a'^\top x - a'_0\right)_{(\emptyset,V')} \geq 0 \text{ is valid for } \mathcal{F}_{C'}, \tag{4.9}$$

$$\mathcal{F}_{A'_{(\emptyset,V')}} \neq \emptyset, \tag{4.10}$$

where $V'$ is the set of variables occurring in more than one row from $C'$ and $\mathcal{F}_{C'} = \{x \in [0,1]^n : (\sum_{j \in A'_i} x_j - 1)_{(\emptyset,V')} \geq 0, i \in C'\}$.

Define $C := \{v\} \cup C'$. Therefore condition (4.5) is satisfied by construction. Moreover, note that in $\mathcal{F}_C$ (as defined in the statement of Lemma 4.3) all the constraints are disjoint, and basic feasible solutions are integral (if needed, we refer to [33] for more details). Suppose that we are given an arbitrary $\tilde{x} \in \{0,1\}^n$ that satisfies $\mathcal{F}_C$. Consider that we must have $\tilde{x}_j = 1$ for some $j \in A_v$ such that $a_j \geq a_{v(1)}$. If we define $x'$ to be the same as $\tilde{x}$ but with $x'_j = 0$, then $x'$ still satisfies $\mathcal{F}_{C'}$. Thus by induction $a'^\top x' \geq a_0 - a_{v(1)}$, which implies that $a^\top \tilde{x} = a'^\top x' + a_j \tilde{x}_j = a'^\top x' + a_j \geq a_0 - a_{v(1)} + a_j \geq a_0$. This proves property (4.6).

To prove property (4.7) we show that we can set to zero all the overlapping variables from the rows in $C$, namely the variables from $V$, and still get a nonempty set of integral solutions, i.e., $\mathcal{F}_{A_{(\emptyset,V)}} \neq \emptyset$. Indeed, by the induction hypothesis we have that $\mathcal{F}_{A'_{(\emptyset,V')}} \neq \emptyset$, where $A'_{(\emptyset,V')} = A_{(\emptyset,V_v \cup V')}$. Therefore $\mathcal{F}_{A_{(\emptyset,V)}} \neq \emptyset$ because $V \subseteq V_v \cup V'$. ∎

First SoS *certificate.* Consider the First summand in (4.4). By Lemma 4.3, we have that $(a^\top x - a_0)_{(\emptyset,V)} \geq 0$ is valid for $\mathcal{F}_C$ (see (4.6)). Note that the linear constraints that define the feasible region $\mathcal{F}_C$ are just a subset of the linear constraints from $\mathcal{F}_A$ after setting to zero all the variables from $V$. It follows that $(a^\top x - a_0)_{(\emptyset,V)} = (\lambda^\top(Ax-e)+\gamma^\top x+\mu)_{(\emptyset,V)}$ for some $\lambda,\gamma,\mu \geq 0$. Then,

$$\delta_\emptyset^V \left(a^\top x - a_0\right)_{(\emptyset,V)} = \delta_\emptyset^V \left(\lambda^\top(Ax-e)+\gamma^\top x+\mu\right)_{(\emptyset,V)} \overset{(2.10)}{\equiv} \delta_\emptyset^V \left(\lambda^\top(Ax-e)+\gamma^\top x+\mu\right).$$

The latter has the form given by (4.3), and it yields an SoS certificate. In order to obtain such a certificate it is sufficient to include in $\mathcal{S}_A(\pi)$ the multilinear polynomial $\delta_\emptyset^V$. With this aim, by using Lemma 4.3, let $\mathcal{C}(\pi) := \{C : C \subseteq [m] \wedge |C| \leq \pi\}$ and $V_C := \bigcup_{i,j\in C, i\neq j} A_i \cap A_j$ be the set of variables occurring in more than one row with index from $C \in \mathcal{C}(\pi)$; add to $\mathcal{S}_A(\pi)$ all $\delta_\emptyset^{V_C}$ with $C \in \mathcal{C}(\pi)$. For any given constant pitch $\pi$, there are polynomially many such $\delta_\emptyset^{V_C}$, and one of them is equal to $\delta_\emptyset^V$ by Lemma 4.3.

SECOND SoS *certificate.* Consider the SECOND summand in (4.4). By property (4.7) we know that by setting to zero all the variables from $V$ we obtain a nonempty subset of feasible integral solutions. It follows that by setting $x_j = 1$, for all $j \in J$, and $x_h = 0$, for all $h \in V \setminus J$, we obtain a nonempty subset of feasible integral solutions, i.e., $\mathcal{F}_{A_{(J,V\setminus J)}} \neq \emptyset$ and $(a^\top x - a_0)_{(J,V\setminus J)} \geq 0$ is a valid inequality for the solutions in $\mathcal{F}_{A_{(J,V\setminus J)}}$ (since $a^\top x - a_0 \geq 0$ is by assumption a valid inequality for any feasible integral solution). Moreover the pitch $p$ of $(a^\top x - a_0)_{(J,V\setminus J)} \geq 0$ is strictly smaller than $\pi$, $0 \leq p \leq \pi - |J|$. By the induction hypothesis, it follows that $(a^\top x - a_0)_{(J,V\setminus J)}$ has an $\mathcal{S}_{A_{(J,V\setminus J)}}(p)$-SoS certificate, which means that there is a $q(x) \in \mathcal{S}_{A_{(J,V\setminus J)}}(p)$ such that

$$\left(a^\top x - a_0\right)_{(J,V\setminus J)} \equiv q(x)\left(\lambda_J^\top(Ax-e)+\gamma_J^\top x+\mu_J\right)_{(J,V\setminus J)}$$

for some $\lambda_J,\gamma_J,\mu_J \geq 0$. The claim follows by observing that

$$\delta_J^V q(x)\left(\lambda_J^\top(Ax-e)+\gamma_J^\top x+\mu_J\right)_{(J,V\setminus J)} \overset{(2.10)}{\equiv} \delta_J^V q(x)\left(\lambda_J^\top(Ax-e)+\gamma_J^\top x+\mu_J\right).$$

Again, the latter has the form given by (4.3). Note that $\delta_J^V q(x)$ is delta-structured. We define the set $\mathcal{S}_A(\pi)$ so that it includes $p(x) := \delta_J^{V_C} \cdot q(x)$ for all $q(x) \in \mathcal{S}_{A_{(J,V_C\setminus J)}}(\pi - |J|)$ and for all $J \subseteq V_C, 0 < |J| < \pi$, and $C \in \mathcal{C}(\pi)$.

THIRD SoS *certificate.* Finally, consider the THIRD summand from (4.4). Recall (see Definition 4.1) that $0 < a_1 \leq a_2 \leq \cdots \leq a_h$ and $a_j = 0$ for $j > h$ for some $h \in [n]$, so $\mathrm{supp}(a) = \{1,\ldots,h\}$. By (4.5), $V \subseteq \mathrm{supp}(a)$. If $|V| < \pi$, then $\delta_{\geq\pi}^V$ is the null polynomial and we are done. Otherwise, let $a_i' := a_i$ for $i \in [\pi]$, $a_i' := a_\pi$ for $i = [h]\setminus[\pi]$ and $a_i' := 0$ for $i \in \mathrm{supp}(a)\setminus V$. It follows that

$$\begin{aligned}
\delta_{\geq\pi}^V\left(\sum_{i=1}^h a_i x_i - a_0\right) &= \delta_{\geq\pi}^V\left(\sum_{i\in V} a_i x_i - a_0 + \sum_{i\in supp(a)\setminus V} a_i x_i\right)\\
&= \delta_{\geq\pi}^V\left(\sum_{i\in V} a_i' x_i - a_0 + \sum_{i\in supp(a)} (a_i - a_i') x_i\right)\\
&\overset{(2.10)}{\equiv} \sum_{I\subseteq V\cap[\pi]} \sum_{k=\pi-|I|}^{|V|} \left(\overbrace{\sum_{\substack{J\subseteq V\setminus[\pi]\\ |J|=k}} \delta_{I\cup J}^V}^{p_{I,k}(x)} \left(\overbrace{\sum_{i\in I} a_i' + k a_\pi - a_0}^{\geq 0} + \sum_{i\in supp(a)} (a_i - a_i') x_i\right)\right).
\end{aligned}$$

The latter has again the form given by (4.3), and it yields an SoS certificate. We define the set $\mathcal{S}_A(\pi)$ so that it includes the polynomials $p_{I,k}(x)$ that are used in the above formula. Note that each $p_{I,k}(x)$ is a symmetric polynomial with respect to the variables indexed by set $V \setminus [\pi]$; therefore it admits a succinct representation by the mean of elementary symmetric polynomials.

**4.1.1. Set $\boldsymbol{\mathcal{S}_A(\pi)}$.** We summarize the definition of $\mathcal{S}_A(\pi)$. Let

$$\mathcal{C}(\pi) := \{C : C \subseteq [m] \wedge |C| \leq \pi\}, \tag{4.11}$$

$$V_C := \bigcup_{i,j \in C, i \neq j} A_i \cap A_j. \tag{4.12}$$

Set $\mathcal{S}_A(\pi)$ includes the following polynomials:

$$\left\{\delta_\emptyset^{V_C} : C \in \mathcal{C}(\pi)\right\} \qquad \text{(First)},$$

$$\left\{\delta_J^{V_C} \cdot q(x) : C \in \mathcal{C}(\pi), J \subseteq V_C \text{ with } 0 < |J| < \pi, q(x) \in \mathcal{S}_{A_{(J, V_C \setminus J)}}(\pi - |J|)\right\} \qquad \text{(Second)},$$

$$\left\{\sum_{\substack{J \subseteq V_C \setminus [\pi] \\ |J| = k}} \delta_{I \cup J}^V : C \in \mathcal{C}(\pi), I \subseteq V_C \text{ with } |I| \leq \pi, k = \pi - |I|, \dots, |V_C|\right\} \qquad \text{(Third)}.$$

Note that when $\pi \in \{0, 1\}$ then $\mathcal{S}_A(\pi) = \{1\}$. By a simple counting argument, we have $|\mathcal{S}_A(\pi)| = (mn)^{O(1)}$ for any fixed $\pi = O(1)$.

*Example* 4.2 (pitch 2 certificate). Consider the set cover instance defined by (4.2), in Example 4.1, namely by a full-circulant constraint matrix FC. The entries of the $i$th row of matrix $FC$ are all equal to 1 but the $i$th entry, which is zero. Let $FC_i := [n] \setminus \{i\}$ denote the support of the $i$th row of matrix $FC$. Let $g_i(x) := \sum_{j \in FC_i} x_j - 1 \geq 0$ be the $i$th constraint corresponding to row $FC_i$.

As already observed, $\sum_{j \in [n]} x_j \geq 2$ is a pitch 2 valid inequality for the feasible region of this set cover instance, and this inequality is "hard" to enforce by "standard" hierarchies like Lasserre/$d$-SoS and $d$-SA (Sherali–Adams).

We start considering the spanning set $\mathcal{S}_{FC}(2)$ (defined in section 4.1.1). According to the definition of set $\mathcal{C}(2)$ (see (4.11)), note that $\{1, 2\} \in \mathcal{C}(2)$; then (see (4.12)), $V_{\{1,2\}} = FC_1 \cap FC_2 = \{3, \dots, n\}$. For short let $V := V_{\{1,2\}}$. The following set $\mathcal{P}$ of polynomials is a subset of $\mathcal{S}_{FC}(2)$:

$$\mathcal{P} := \overbrace{\left\{\delta_\emptyset^V\right\}}^{\mathcal{P}_0} \cup \overbrace{\left\{\delta_{\{i\}}^V : i \in V\right\}}^{\mathcal{P}_1} \cup \overbrace{\left\{\sum_{\substack{I \subseteq V: \\ |I| = k}} \delta_I^V : k = 2, \dots, n\right\}}^{\mathcal{P}_2} \subseteq \mathcal{S}_{FC}(2). \tag{4.13}$$

In addition to those listed above, note that in set $\mathcal{S}_{FC}(2)$ there are also other polynomials. These polynomials are all the same under a permutation of the variables and they play a similar role due to the symmetry of the example. By using the above polynomials we obtain a proof of nonnegativity as follows:

(4.14)

$$\sum_{j\in[n]} x_j - 2 = \overbrace{\left(\delta^V_\emptyset + \sum_{i\in V}\delta^V_{\{i\}} + \sum_{k=2}^{n}\left(\sum_{\substack{I\subseteq V:\\|I|=k}}\delta^V_I\right)\right)}^{=1}\left(\sum_{j\in[n]} x_j - 2\right)$$

$$\equiv \underbrace{\delta^V_\emptyset\,(x_1+x_2-2)}_{\text{First}} + \underbrace{\left(\sum_{i\in V}\delta^V_{\{i\}}\,(x_1+x_2-1)\right)}_{\text{Second}} + \underbrace{\sum_{k=2}^{n}\left(\sum_{\substack{I\subseteq V:\\|I|=k}}\delta^V_I\right)(x_1+x_2+k-2)}_{\text{Third}}$$

$$\equiv \underbrace{\delta^V_\emptyset\,(g_1(x)+g_2(x))}_{\text{First}} + \underbrace{\left(\sum_{i\in V}\delta^V_{\{i\}} g_i(x)\right)}_{\text{Second}} + \underbrace{\sum_{k=2}^{n}\left(\sum_{\substack{I\subseteq V:\\|I|=k}}\delta^V_I\right)(x_1+x_2+(k-2))}_{\text{Third}}.$$

The latter has the form given by (4.3), and it yields an SoS certificate (and it is a Sherali–Adams certificate) for the considered pitch 2 inequality.

**4.2. An explicit compact LP formulation.** For any fixed $\pi = O(1)$, in the proof of Theorem 4.2 we have shown that every valid inequality $a^\top x - a_0 \geq 0$ of pitch at most $\pi$ admits a certificate of nonnegativity that is congruent (mod $\mathbf{I}_{01}$) to (4.3). By reformulating this result in an equivalent way, we have shown that $a^\top x - a_0$ belongs to the following cone of polynomials:

(4.15)

$$\mathcal{C}_{\mathcal{S}_A(\pi)} = \left\{\overline{\sum_i q_i(x)\left(\lambda_i^\top(Ax-e)+\gamma_i^\top x+\mu_i\right)} : q_i(x)\in\mathcal{S}_A(\pi), \lambda_i,\gamma_i,\mu_i\geq 0\right\}.$$

The dual cone $\mathcal{C}^*_{\mathcal{S}_A(\pi)}$ is the set of linear functionals $y[\cdot]$ that are nonnegative on the primal cone satisfying (see the discussion in section 2.2, constraints (2.20), (2.23), (2.24))

$$y[1] = 1; \tag{4.16}$$

$$y\left[\overline{q(x)}\right] \geq 0 \quad \forall q(x)\in\mathcal{S}_A(\pi); \tag{4.17}$$

$$y\left[\overline{q(x)\cdot g_i(x)}\right] \geq 0 \quad \forall q(x)\in\mathcal{S}_A(\pi), \forall i\in[m+n]; \tag{4.18}$$

where $g_i(x)\geq 0$, for $i\in[m+n]$, denotes a constraint from $Ax\geq e$, or $x_j\geq 0$ for $j\in[n]$.

As already discussed in section 2.2, the linear functional inequalities (4.16), (4.17), and (4.18) yield a *linear program* of size $O(|\mathcal{S}_A(\pi)|)$. It is actually a *hierarchy* of linear programs parameterized by the pitch $\pi$. This relaxation can be seen as a generalized Sherali–Adams relaxation, where the standard monomial basis of degree $\leq d$ has been replaced with the set $\mathcal{S}_A(\pi)$ of high degree polynomials.

*Example* 4.3 (pitch 2 LP). We provide an explicit LP for the set cover instance considered in Examples 4.1 and 4.2. With this aim, we can compute either an ordered basis for the cone of polynomials (4.15) or, alternatively, an ordered spanning set and impose the linearity conditions (see the discussion in section 2.2 and condition (2)). Here, we follow the second option.

Let $T := \{\overline{x_i \cdot p} : p \in \mathcal{S}_{FC}(2), i \in [n] \cup \{0\}\}$, and note that $T$ is a spanning set for (4.15). The dimension of $T$ is equal to the dimension of the linear functionals $y$: there is one entry in $y$ for each polynomial in $T$. So vector $y$ is indexed by the polynomials in $T$. Consider set $\mathcal{P} \subseteq \mathcal{S}_{FC}(2)$ of polynomials (see (4.13)).

- *Variables.* The LP variables are the entries of vector $y$. In particular there are the following variables: $y[\overline{q(x)x_j}]$ for $q(x) \in \mathcal{P} = \mathcal{P}_0 \cup \mathcal{P}_1 \cup \mathcal{P}_2$ and $j \in [n] \cup \{0\}$ (recall $x_0 := 1$).
- *Constraints.* By (4.17), (4.18) we have the following linear constraints in the LP formulation:

(4.19)
$$y\left[\overline{q(x)x_j}\right] \geq 0 \quad \forall q(x) \in \mathcal{P}_2, j = 0, 1, 2;$$

(4.20)
$$y\left[\overline{\delta_\emptyset^V \cdot g_i(x)}\right] = y\left[\delta_\emptyset^V x_i\right] - y\left[\delta_\emptyset^V\right] \geq 0 \quad \forall i = 1, 2;$$

(4.21)
$$y\left[\overline{\delta_{\{i\}}^V \cdot g_i(x)}\right] = y\left[\delta_{\{i\}}^V x_1\right] + y\left[\delta_{\{i\}}^V x_2\right] - y\left[\delta_{\{i\}}^V\right] \geq 0 \quad \forall i \in V = \{3, \ldots, n\}.$$

The following valid inequality can be obtained by a conical combination of (4.19)–(4.21):

$$\begin{aligned} y\left[\delta_\emptyset^V x_1\right] + y\left[\delta_\emptyset^V x_2\right] - 2y\left[\delta_\emptyset^V\right] + \sum_{i \in V} \left( y\left[\delta_{\{i\}}^V x_1\right] + y\left[\delta_{\{i\}}^V x_2\right] - y\left[\delta_{\{i\}}^V\right] \right) \\ + \sum_{q(x) \in \mathcal{P}_2} (y[q(x)x_1] + y[q(x)x_2] + (k-2)y[q(x)]) \geq 0. \end{aligned} \tag{4.22}$$

Note that $\sum_{q(x) \in \mathcal{P}} q(x) = 1$, and therefore by the linearity conditions (see the discussion in section 2.2 and condition (2)) the following is part of the set of the LP constraints (for $j = 1, 2$):

$$y\left[\delta_\emptyset^V x_j\right] + \sum_{i \in V} y\left[\delta_{\{i\}}^V x_j\right] + \sum_{q(x) \in \mathcal{P}_2} y[q(x)x_j] = y[x_j].$$

Analogously, note $\sum_{k=1}^n k(\sum_{\substack{I \subseteq V: \\ |I|=k}} \delta_I^V) = \sum_{i \in V} x_i$, which by linearity gives the following constraint that holds for the linear functional $y$ (and that is part of the LP formulation):

$$\sum_{k=1}^n k \cdot y \left[ \sum_{\substack{I \subseteq V: \\ |I|=k}} \delta_I^V \right] = \sum_{i \in V} y[x_i].$$

Then, (4.22) and the linearity conditions imply the pitch 2 inequality

$$\sum_{i \in [n]} y[x_i] - 2 \geq 0.$$

**5. The Bienstock–Zuckerberg hierarchy.** The BZ hierarchy [4, 33] generalizes the approach for set cover. The full description requires several layers of details and here we sketch only the main points. We refer to the original manuscripts for a more precise and comprehensive description.

Any nontrivial constraint can be rewritten in the set cover form, $\sum_{i\in I} a_i x_i + \sum_{j\in J} a_j(1-x_j) \geq b$, with all the coefficients $a, b$ nonnegative. Then the BZ hierarchy uses the standard concept of minimal covers[5] (see, e.g., [7]): a *minimal cover* is an inclusion-minimal set $C \subseteq \text{supp}(a)$ such that $\sum_{j\notin C} a_j < b$ and therefore $\sum_{j\in C} x'_j \geq 1$ is a valid inequality (where $x'_j = x_j$ if $j \in I$ or $x'_j = 1 - x_j$ else). In general, the number of minimal covers can be exponential so the idea in BZ is to generate only the "$k$-small" ones, which are added to the original relaxation. Here with "$k$-small" we mean all the valid minimal covers with all the variables from $I$ (or $J$) but at most $k$, or at most $k$ from $I$ (or $J$). These minimal covers can be enumerated in polynomial time for any fixed $k$. Then the set cover approach is applied to the set cover problem given by the $k$-small minimal covers. If the minimal covers are polynomially bounded, this allows us to generate the pitch bounded valid inequalities as for set cover (see the application below). Roughly speaking, the "power" of the BZ approach is given by the presence of the $k$-small minimal covers; if this set is empty, then the hierarchy is not stronger than a variant of the Sherali–Adams hierarchy (see [2]).

The BZ approach can be reframed into the SoS framework by choosing the appropriate spanning polynomials. We omit the complete mapping because this would require the full description of BZ, which that is quite lengthy. Moreover, the most important application of BZ currently known is given by the set cover problem, which has been widely explained in previous sections. By way of example, we show in [23] that we do not need to explicitly add the $k$-small minimal covers, since they can be implied by adding the "right" polynomials. By using the explained ideas, it should be easy to fill in the missing details.

**6. Chvátal–Gomory cuts.** Consider a rational polyhedra $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ with $A \in \mathbb{Z}^{m\times n}$ and $b \in \mathbb{Z}^m$. Inequalities of the form $(\lambda^\top A)x \geq \lceil \lambda^\top b \rceil$, with $\lambda \in \mathbb{R}^m_+$, $\lambda^\top A \in \mathbb{Z}^n$, and $\lambda^\top b \notin \mathbb{Z}$, are commonly referred to as CG-cuts; see, e.g., [7]. CG-cuts are valid for the integer hull $P^*$ of $P$.

The following rational polyhedron is commonly referred to as the *first CG closure*:

$$P^{(1)} := \left\{x \in \mathbb{R}^n : (\lambda^\top A)x \geq \lceil \lambda^\top b \rceil, \lambda \in [0,1]^m, \lambda^\top A \in \mathbb{Z}^n\right\}. \tag{6.1}$$

In particular $P^{(1)}$ is a stronger relaxation of $P^*$ than $P$, i.e., $P^* \subseteq P^{(1)} \subseteq P$. We can iterate the closure process to obtain the CG closure of $P^{(1)}$. We denote by $P^{(2)}$ this second CG closure. Iteratively, we define the $t$th CG closure $P^{(t)}$ of $P$ to be the CG closure of $P^{(t-1)}$, for $t \geq 2$ integer. An inequality that is valid for $P^{(t)}$ but not for $P^{(t-1)}$ is said to have *CG-rank* $t$.

Eisenbrand and Schulz [8] proved that for any polytope $P$ contained in the unit cube $[0,1]^n$, one can choose $t = O(n^2 \log n)$ and obtain the integer hull $P^{(t)} = P^*$. Rothvoß and Sanitá [30] proved that there is a polytope contained in the unit cube whose CG-rank has order $n^2$, thus showing that the above bound is tight, up to a logarithmic factor.

The CG-cuts that are valid for $P^{(1)}$ and that can be derived by using coefficients in $\lambda$ of value 0 or 1/2 only are called $\{0, 1/2\}$-cuts. In [21] it is shown that the separation problem for $\{0, 1/2\}$-cuts remains strongly NP-hard, even when all integer

[5] More precisely, in [4, 33] a closely related concept that is called *obstruction* is used.

variables are binary, $P = \{x \in \mathbb{R}^n_+ : Ax \leq e\}$ with $A \in \{0,1\}^{m\times n}$, and $e$ denote the all-one vector with $m$ entries. As pointed out in [21], the latter hardness proof can easily be adapted to set partitioning and set cover problems. This result implies that it is NP-hard to optimize a linear function over the first closure $P^{(1)}$.

For min set cover problems, Bienstock and Zuckerberg [5] obtained the following result. For an arbitrary fixed precision $\varepsilon > 0$ and a fixed $t \in \mathbb{N}$, choose $\pi$ such that $\left(\frac{\pi+1}{\pi}\right)^t \leq 1+\varepsilon$. For any given set cover instance, let *opt* denote the optimal integral value and let $opt^{(t)}$ ($\leq opt$) denote the optimal value over the $t$th closure $P^{(t)}$. Bienstock and Zuckerberg [5] considered the optimal solution $x^*_\pi$ of value $opt_\pi$ ($\leq opt$) of a relaxation $R_\pi$ that satisfies all pitch-$\pi$ valid inequalities for the integer hull. Then, either ($opt \geq$) $opt_\pi \geq opt^{(t)}$, implying therefore that $R_\pi$ is a better relaxation than the $t$th closure $P^{(t)}$, or ($opt \geq$) $opt^{(t)} \geq opt_\pi$. In the latter case, they proved that $x^*_\pi$ can be rounded to satisfy all the CG-cuts of rank $t$. Moreover, the value of the rounded solution is at most $1+\varepsilon$ times larger than $opt_\pi$. This implies that $(1+\varepsilon)opt_\pi \geq opt^{(t)}$ and therefore $opt_\pi \geq (1-\varepsilon)opt^{(t)}$. This gives a polynomial time approximation scheme (PTAS) for approximating $opt^{(t)}$, i.e., for the minimization of set cover objective functions over $P^{(t)}$. It follows that the generalized SoS (or Sherali–Adams) relaxation with high degree polynomials described in this paper yields also to a PTAS for approximating set cover objective functions over $P^{(t)}$.

In the next section we present a more general result for packing problems, meaning that the coefficients of the nonnegative matrix $A$ are no longer restricted to be 0/1, or bounded (see [23]), as for the set cover case. It remains an interesting open question to extend the results for the set cover problem to general covering problems, namely covering problems with general nonnegative matrices $A$.

**6.1. Approximating fixed-rank CG closure for packing problems.** In this section we consider packing problems and show that $d$-SoS yields a PTAS for approximating over the $t$th CG closure $P^{(t)}$, for any fixed $t$. It follows that the SoS approach can be used for approximating to any arbitrary precision, over any constant CG closure, for both packing and set cover problems (BZ guarantees this only for set cover problems).

Consider any given $m \times n$ nonnegative matrix $A$ and a vector $b \in \mathbb{R}^m_+$. Let $\mathcal{F}_{A,b}$ be the feasible region for the 0-1 packing problem defined by $A$ and $b$:

$$\mathcal{F}_{A,b} = \{x \in \mathbb{R}^n : x_i^2 - x_i = 0\ \forall i \in [n], Ax \leq b\}.$$

We extend the definition of pitch also for packing inequalities as follows.

DEFINITION 6.1. *For any given packing inequality $a_0 - a^\top x \geq 0$, with $a_0, a \geq 0$ and indices ordered so that $0 < a_1 \leq a_2 \leq \cdots \leq a_h$ and $a_j = 0$ for $j > h$, its* pitch *$\pi(a, a_0)$ is the* maximum *integer such that $a_0 - \sum_{i=1}^{\pi(a,a_0)} a_i \geq 0$.*

For example, the classical clique inequality $\sum_{i\in C} x_i \leq 1$, where $C$ is a clique, has pitch equal to one.

The following result for packing problems can be seen as the dual of Theorem 4.2 for set cover. It can be derived by using the so-called decomposition theorem due to Karlin, Mathieu, and Nguyen [12]. Here we give a direct simple proof that follows the approach used throughout this paper.

LEMMA 6.2. *Consider any packing problem instance given by a matrix $A \in \mathbb{R}^{m\times n}_+$ and a vector $b \in \mathbb{R}^m_+$. Let $\pi = O(1)$ be a fixed positive integer. Then, $(\pi+1)$-SoS satisfies all valid inequalities for $\mathcal{F}_{A,b}$ of pitch at most $\pi$.*

*Proof.* Suppose $a_0 - a^\top x \geq 0$ is a valid inequality for $\mathcal{F}_{A,b}$ of pitch $\pi$ with $a_0, a \geq 0$. The claim follows from Proposition 2.7 by showing that $a_0 - a^\top x$ admits a $(\pi+1)$-SoS certificate.

Let $S := \text{supp}(a)$ and $x_I := \prod_{i\in I} x_i$ for $I \subseteq [n]$. By (2.10) (choose $Z = I$), for any given $I \subseteq S$ we have $x_I(a_0 - a^\top x) \equiv x_I(a_0 - \sum_{i\in I} a_i - \sum_{i\notin I} a_i x_i) \pmod{\mathbf{I}_{01}}$.

Let $F := \{I : I \subseteq S, (a_0 - \sum_{i\in I} a_i) < 0\}$ and $T := \{J : J \subseteq S, J \notin F\}$ (and therefore if we set to 1 all the variables $x_i$ with $i \in I$, for $I \in F$, then the assumed valid inequality $a_0 - a^\top x \geq 0$ is violated). Let $V := \{x \in \mathbb{R}^n : x_I = 0\ \forall I \in F,\ x_k^2 - x_k = 0\ \forall k \in [n]\}$ and note that any feasible integral solution belongs to $V$.

For any given $\delta_J^S$, let $\bar{\delta}_J^S$ denote the "truncated" version of $\delta_J^S$ obtained from $\delta_J^S$ by zeroing all the monomials $x_I$ with $I \in F$ (observe that $\bar{\delta}_J^S = 0$ for $J \in F$). Clearly, $\deg(\bar{\delta}_J^S) \leq \pi$, since $a_0 - a^\top x \geq 0$ has pitch at most $\pi$. Note that $\sum_{I\subseteq S} \bar{\delta}_I^S = \sum_{I\in T} \bar{\delta}_I^S = 1$, $(\bar{\delta}_I^S)^2 \equiv \bar{\delta}_I^S \pmod{\mathbf{I}(V)}$, and $\bar{\delta}_I^S(a_0 - a^\top x) \equiv \bar{\delta}_I^S(a_0 - \sum_{i\in I} a_i) \pmod{\mathbf{I}(V)}$. These can be derived by *multilinearizing* and by *zeroing* all the monomials from $\mathbf{I}(V)$ that are on the left- and right-hand sides of (2.6), (2.7), and (2.10), respectively.

If follows that

$$\bar{\delta}_I^S \left(a_0 - a^\top x\right) = \bar{\delta}_I^S \left(a_0 - \sum_{i\in I} a_i\right) + h_I(x) \quad \text{for some } h_I(x) \in \mathbf{I}(V). \tag{6.2}$$

As said before, the term $\bar{\delta}_I^S(a_0 - \sum_{i\in I} a_i)$, which is on the right-hand side of (6.2), is obtained from the left-hand side of (6.2) by multilinearizing, so replacing each occurrence of $x_i^2$ with $x_i$, and by zeroing all the monomials from $\mathbf{I}(V)$ that appear on the left-hand side. Note that these latter monomials have degree at most $\pi + 1$, since they derive from multiplying a degree $\pi$ polynomial $\bar{\delta}_I^S$ with a linear function. Therefore $\deg(h_I(x)) \leq \pi + 1$. Then

$$a_0 - a^\top x = \left(a_0 - a^\top x\right) \overbrace{\left(\sum_{I\in T} \bar{\delta}_I^S\right)}^{=1} \overset{(6.2)}{=} \sum_{I\in T} \overbrace{\left(a_0 - \sum_{i\in I} a_i\right)}^{\geq 0} \bar{\delta}_I^S + f(x) \tag{6.3}$$

for $f(x) = \sum_{I\in T} h_I(x) \in \mathbf{I}(V)$ with $\deg(f) \leq \pi + 1$.

From the above equivalence we see that $a_0 - a^\top x$ can be written $\pmod{\mathbf{I}(V)}$ as a conical combination of polynomials from $\{\bar{\delta}_I^S : I \in T\}$ of degree at most $\pi$. The claim follows by transforming the above congruence $\pmod{\mathbf{I}(V)}$ (6.3) into a congruence $\pmod{\mathbf{I}_{01}}$, while still using bounded degree polynomials.

Since $f(x) \in \mathbf{I}(V)$, by looking at the definition of $\mathbf{I}(V)$ note that every monomial in $f(x)$ belongs to $\mathbf{I}(V)$ as well. Then $f(x) = \sum_{I\in U} f_I \cdot x_I$ for some $U \subseteq 2^{[n]}$ such that, for all $I \in U$, we have $x_I \in \mathbf{I}(V)$ and $f_I \in \mathbb{R}$ (and, as already observed, $\deg(x_I) \leq \pi + 1$).

If $f_I \geq 0$, then $f_I \cdot x_I \equiv f_I \cdot (x_I)^2 \pmod{\mathbf{I}_{01}}$; otherwise (i.e., $f_I < 0$), since $x_I \in \mathbf{I}(V)$. Then, for some $\lambda, \gamma \geq 0$, there is a valid constraint from a conical combination of valid constraints $c_I(x) := \left(\lambda^\top(b - Ax) + \gamma^\top x\right) \geq 0$ that is violated by setting $x_i = 1$ for $i \in I$, i.e., $c(x_I) < 0$. Therefore (recall $x_I \cdot c_I(x_I) \equiv x_I \cdot c_I(x)$ $\pmod{\mathbf{I}_{01}}$)

$$f_I \cdot x_I \equiv \left(\sqrt{\frac{f_I}{c_I(x_I)}}\, x_I\right)^2 c_I(x) \pmod{\mathbf{I}_{01}}.$$

It follows that $a_0 - a^\top x$ admits a $(\pi+1)$-SoS certificate:

$$a_0 - a^\top x \equiv s_0 + \sum_{i=1}^{m} s_i g_i \pmod{\mathbf{I}_{01}} \quad \text{for some } s_i \in \Sigma_{\pi+1}, \tag{6.4}$$

where $g_i \geq 0$, for $i \in [m]$, denotes the $i$th constraint from $b - Ax \geq 0$ and $\Sigma_{\pi+1} := \{\sum_i q_i^2 : q_i \in \mathbb{R}[x]_{\pi+1}\}$. ∎

Let $P = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1 \ \forall i \in [n], Ax \leq b\}$ denote the linear relaxation of $\mathcal{F}_{A,b}$. For $t \in \mathbb{N}$, recall that $P^*$ and $P^{(t)}$ denote the integer hull and the $t$th CG closure, respectively, of the starting linear program $P$, and $opt^{(t)}(c) := \max\{c^\top x : x \in P^{(t)}\}$. Without loss of generality, we will assume that $c \in \mathbb{R}^n_+$ (otherwise it is always optimal to set $x_i = 0$ whenever $c_i \leq 0$). Let $Sol(d)$ denote the set of feasible solutions for $d$-SoS projected to the original space of the variables. Let $opt_d(c) := \max\{c^\top x : x \in Sol(d)\}$.

The following result shows that fixed-rank CG closures of packing problems can be approximated to any arbitrary precision, and in polynomial time, by using the standard SoS hierarchy.

THEOREM 6.3. *For any fixed $t \in \mathbb{N}$ and $\varepsilon > 0$, there is an integer $d = d(t, \varepsilon)$ such that $opt_d(c) \leq (1+\varepsilon) opt^{(t)}(c)$ for all $c \in \mathbb{R}^n_+$.*

*Proof.* For any fixed $t \in \mathbb{N}$ and $\varepsilon > 0$, choose $d \in \mathbb{N}$ such that $(d/(d-1))^t \leq 1+\varepsilon$. Let $opt_d$ (or $opt^{(t)}$) denote $opt_d(c)$ (or $opt^{(t)}(c)$), for short.

If $opt_d \leq opt^{(t)}$, then we are done. Otherwise ($opt_d > opt^{(t)}$), let $x^{(t)} := \phi_t \cdot x^*$ where $\phi_t := (\frac{d-1}{d})^t$. It follows that $opt_d = c^\top x^* \leq (1+\varepsilon) c^\top x^{(t)}$. We show that $x^{(t)}$ is feasible for the rank-$t$ CG closure. This imples that $c^\top x^{(t)} \leq opt^{(t)}$, and the claim follows since $opt_d \leq (1+\varepsilon) c^\top x^{(t)} \leq (1+\varepsilon) opt^{(t)}$.

The proof is by induction on $t$. As a base of induction note that when $t = 0$ then clearly $x^{(0)} \in P = P^{(0)}$.

Assume now, by the induction hypothesis, that $x^{(t-1)} \in P^{(t-1)}$ for any rank equal to $(t-1)$ with $t \geq 1$. We need to show that it is valid also for rank $t$. If the pitch of a generic rank-$t$ valid inequality for $P^{(t)}$ is at most $d-1$, then by Lemma 6.2 it follows that any feasible solution $x \in Sol(d)$ (and therefore $x^{(t)}$) satisfies this inequality. Otherwise, consider a generic rank-$t$ valid inequality $\lfloor a_0 \rfloor - a^\top x \geq 0$ of pitch larger than $d-1$, where $a_0 - a^\top x \geq 0$ is any valid inequality from the closure $P^{(t-1)}$. By the induction hypothesis note that $a_0 - a^\top x^{(t-1)} \geq 0$. Since the pitch is higher than $d-1$ then $a_0 > d-1$ (vector $a$ can be assumed, w.l.o.g., to be nonnegative and integral) and therefore $\frac{a_0}{\lfloor a_0 \rfloor} \leq \frac{d}{d-1}$, and by multiplying the solution $x^{(t-1)} \in P^{(t-1)}$ by $(d-1)/d$ we obtain a feasible solution for the rank-t CG closure. ∎

**7. Conclusions and future directions.** A breakthrough result [20] of Lee, Raghavendra, and Steurer shows that the standard SoS is *optimal* for constraint satisfaction problems among all semidefinite programs of comparable size. In [15] and [16], the standard SoS is shown to be *pessimal* for simple problems, meaning that it requires exponential size to get any bounded approximation.

The standard SoS has been defined with respect to the standard monomial basis, which looks like a "natural" choice, but in fact it turns out to be an *arbitrary* choice. This way can be "good" or "bad" depending on the problem at hand.

In this paper, we have shown a first example of SoS equipped with a different basis, which is useful in *asymmetric* situations. The proposed approach overcomes some provable limitations of the standard SoS.

A very challenging open question is to understand what is the "right" basis for the problem that we want to address. Roughly speaking, can we transform the recipe in section 2 into an effective algorithm? Any progress in this direction would be of considerable interest.

## REFERENCES


[1] Y. H. Au and L. Tunçel, *A comprehensive analysis of polyhedral lift-and-project methods*, SIAM J. Discrete Math., 30 (2016), pp. 411–451, https://doi.org/10.1137/130950173.

[2] Y. H. Au and L. Tunçel, *Elementary polytopes with high lift-and-project ranks for strong positive semidefinite operators*, Discrete Optim., 27 (2018), pp. 103–129, https://doi.org/10.1016/j.disopt.2017.10.001.

[3] N. Bansal, *LP/SDP Hierarchies Reading Group*, http://www.win.tue.nl/~nikhil/hierarchies/.

[4] D. Bienstock and M. Zuckerberg, *Subset algebra lift operators for* 0-1 *integer programming*, SIAM J. Optim., 15 (2004), pp. 63–95.

[5] D. Bienstock and M. Zuckerberg, *Approximate fixed-rank closures of covering problems*, Math. Program., 105 (2006), pp. 9–27, https://doi.org/10.1007/s10107-005-0598-z.

[6] G. Blekherman, P. Parrilo, and R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Ser. Optim. 13, SIAM, Philadelphia, 2013, https://books.google.it/books?id=O7hxRDy9vrIC.

[7] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer Programming*, Grad. Texts in Math. 271, Springer, New York, 2014, https://books.google.it/books?id=antqBQAAQBAJ.

[8] F. Eisenbrand and A. S. Schulz, *Bounds on the Chvátal rank of polytopes in the* 0/1*-cube*, Combinatorica, 23 (2003), pp. 245–261, https://doi.org/10.1007/s00493-003-0020-5.

[9] H. Fawzi, J. Saunderson, and P. A. Parrilo, *Equivariant semidefinite lifts and sum-of-squares hierarchies*, SIAM J. Optim., 25 (2015), pp. 2212–2243, https://doi.org/10.1137/140966265.

[10] K. Gatermann and P. A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Algebra, 192 (2004), pp. 95–128, https://doi.org/10.1016/j.jpaa.2003.12.011.

[11] D. Grigoriev, E. A. Hirsch, and D. V. Pasechnik, *Complexity of semialgebraic proofs*, Moscow Math. J., 2 (2002), pp. 647–679.

[12] A. R. Karlin, C. Mathieu, and C. T. Nguyen, *Integrality gaps of linear and semi-definite programming relaxations for knapsack*, in Proceedings of the 15th International Conference on Integer Programming and Combinatoral Optimization, 2011, pp. 301–314.

[13] A. Kurpisz, S. Leppänen, and M. Mastrolilli, *Sum-of-squares hierarchy lower bounds for symmetric formulations*, in Proceedings of the 18th International Conference on Integer Programming and Combinatorial Optimization, 2016, pp. 362–374, https://doi.org/10.1007/978-3-319-33461-5_30.

[14] A. Kurpisz, S. Leppänen, and M. Mastrolilli, *Tight sum-of-squares lower bounds for binary polynomial optimization problems*, in Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming, 2016, pp. 78:1–78:14, https://doi.org/10.4230/LIPIcs.ICALP.2016.78.

[15] A. Kurpisz, S. Leppänen, and M. Mastrolilli, *On the hardest problem formulations for the* 0/1 *Lasserre hierarchy*, Math. Oper. Res., 42 (2017), pp. 135–143, https://doi.org/10.1287/moor.2016.0797.

[16] A. Kurpisz, S. Leppänen, and M. Mastrolilli, *An unbounded sum-of-squares hierarchy integrality gap for a polynomially solvable problem*, Math. Program., 166 (2017), pp. 1–17, https://doi.org/10.1007/s10107-016-1102-7.

[17] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817.

[18] M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for* 0-1 *programming*, Math. Oper. Res., 28 (2003), pp. 470–496.

[19] M. Laurent, *Sums of Squares, Moment Matrices and Optimization Over Polynomials*, Springer, New York, 2009, pp. 157–270, https://doi.org/10.1007/978-0-387-09686-5_7.

[20] J. R. Lee, P. Raghavendra, and D. Steurer, *Lower bounds on the size of semidefinite programming relaxations*, in Proceedings of the 47th Annual ACM on Symposium on Theory of Computing, 2015, pp. 567–576, https://doi.org/10.1145/2746539.2746599.

[21] A. N. Letchford, S. Pokutta, and A. S. Schulz, *On the membership problem for the* $\{0, 1/2\}$*-closure*, Oper. Res. Lett., 39 (2011), pp. 301–304, https://doi.org/10.1016/j.orl.2011.07.003.

[22] M. Mastrolilli, *High degree Sum-of-Squares proofs, Bienstock-Zuckerberg hierarchy and CG cuts*, in Proceedings of the 19th International Conference on Integer Programming and Combinatorial Optimization, 2017, pp. 405–416.

[23] M. Mastrolilli, *High degree sum of squares proofs, Bienstock-Zuckerberg hierarchy and Chvátal-Gomory cuts*, CoRR, https://arxiv.org/abs/1709.07966, 2017.

[24] M. Mastrolilli, *The complexity of the ideal membership problem for constrained problems over the Boolean domain*, in Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms, San Diego, CA, 2019, pp. 456–475, https://doi.org/10.1137/1.9781611975482.29.

[25] Y. Nesterov, *Squared Functional Systems and Optimization Problems*, Springer, New York, 2000, pp. 405–440, https://doi.org/10.1007/978-1-4757-3216-0_17.

[26] R. O'Donnell, *SOS is not obviously automatizable, even approximately*, in Proceedings of the 8th Innovations in Theoretical Computer Science Conference, Berkeley, CA, 2017, pp. 59:1–59:10, https://doi.org/10.4230/LIPIcs.ITCS.2017.59.

[27] P. A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., 96 (2003), pp. 293–320.

[28] P. Pudlák, *Lower bounds for resolution and cutting plane proofs and monotone computations*, J. Symb. Log., 62 (1997), pp. 981–998, https://doi.org/10.2307/2275583.

[29] A. Raymond, J. Saunderson, M. Singh, and R. R. Thomas, *Symmetric sums of squares over k-subset hypercubes*, Math. Program., 167 (2018), pp. 315–354, https://doi.org/10.1007/s10107-017-1127-6.

[30] T. Rothvoß and L. Sanitá, 0/1 *polytopes with quadratic Chvátal rank*, in Proceedings of the 16th International Conference on Integer Programming and Combinatorial Optimization, 2013, pp. 349–361, https://doi.org/10.1007/978-3-642-36694-9_30.

[31] N. Shor, *Class of global minimum bounds of polynomial functions*, Cybernet. Syst. Anal., 23 (1987), pp. 731–734.

[32] B. Sturmfels, *Algorithms in Invariant Theory*, 2nd ed., Springer, New York, 2008.

[33] M. Zuckerberg, *A Set Theoretic Approach to Lifting Procedures for* 0-1 *Integer Programming*, Ph.D. thesis, Columbia University, 2004.