ON THE CONNECTION OF THE SHERALI-ADAMS CLOSURE AND BORDER BASES

SEBASTIAN POKUTTA AND ANDREAS S. SCHULZ

Abstract. The Sherali-Adams lift-and-project hierarchy is a fundamental construct in integer programming, which provides successively tighter linear programming relaxations of the integer hull of a polytope. We initiate a new approach to understanding the Sherali-Adams procedure by relating it to methods from computational algebraic geometry. Our two main results are the equivalence of the Sherali-Adams procedure to the computation of a border basis, and a refinement of the Sherali-Adams procedure that arises from this new connection. We present a modified version of the border basis algorithm to generate a hierarchy of linear programming relaxations that are stronger than those of Sherali and Adams, and over which one can still optimize in polynomial time (for a fixed number of rounds in the hierarchy). In contrast to the well-known Gröbner bases approach to integer programming, our procedure does not create primal solutions, but constitutes a novel approach of using computer-algebraic methods to produce dual bounds.

1. Introduction

In integer programming, there are several well-known hierarchies of linear or semi-definite programs that provide successively tighter relaxations of a 0/1-polytope. At one end of the spectrum of these relaxations is the original relaxation $P = \{x \in [0,1]^n \mid Ax = b\}$; the integer hull $P_I := \mathrm{conv}(P \cap \mathbb{Z}^n)$ is at the other end. The list of such hierarchies includes the Gomory-Chvátal, Sherali-Adams, lift-and-project, Lovász-Schrijver, and Lasserre hierarchies; see [8, 24] for details. Here, we consider the Sherali-Adams hierarchy [30], which has recently received increasing attention: In the context of propositional proof systems, it was shown that the level of the Sherali-Adams hierarchy required to prove the pigeon hole principle or the least number principle is $\Omega(n)$ [29]. Moreover, all first-order sentences that fail in all finite structures, but have an infinite model require at least a poly-logarithmic level of the Sherali-Adams hierarchy to be refuted [11]. In terms of integrality gaps, with n denoting the dimension of the polytope, it was shown that for every $\epsilon > 0$ there is a $\delta > 0$ such that the optimum over the relaxation at level n^δ of the Sherali-Adams hierarchy is a factor of $2 - \epsilon$ away from the true optimum for both the vertex-cover and the max-cut problem [6]. Moreover, for the complete graph K_{2d+1}, the complete description of the matching polytope is only obtained at level $2d-1$ of the Sherali-Adams hierarchy [26]. Still, the Sherali-Adams hierarchy is an important construct to systematically "convexify" the set of 0/1-points contained in a linear system, and it is known to be stronger than the lift-and-project hierarchy (the Lovász-Schrijver hierarchy without semidefinite cuts). In this paper, we look at the Sherali-Adams hierarchy from a computational algebra lens and uncover an interesting connection between the Sherali-Adams procedure and the border basis algorithm.

Originally, border bases were introduced as a generalization of Gröbner bases to address numerical instabilities in the computation of polynomial ideals. Empirical examples of the numerical advantages of border bases are, for instance, given in [1, 16, 23]. Border bases have also been successfully used for solving zero-dimensional systems of polynomial equations; see, e.g., [2, 27, 28]. At that time, the notion of border basis was not fully established, and the following articles provided the first concise framework: [17, 18, 19, 22]. Zero-dimensional systems of polynomial equations include systems of polynomials equations with solutions in 0/1. Solutions to those systems were, for example, read of the eigenvalue spectrum of the associated endomorphism matrices. Also, the *Nullstellensatz* and a variant of the border basis algorithm have been used as a proof system to establish infeasibility of combinatorial problems. For example, in [14, Section 2.3] and [13, 12], infeasibility of 3-colorability of graphs (and other combinatorial problems) is certified using this approach. We will take a completely different point

Date: November 2010/Draft.

of view and ask whether we can derive sensible relinearizations and projections of the computed bases in order to approximate the integer hull of polytopes $P \subseteq [0,1]^n$.

The preference of border bases over Gröbner bases partly arises from the iterative generation of linear syzygies, the difference polynomials of two generators that have been multiplied by exactly one variable each, inherent in the border basis algorithm, which allows for successively approximating the 0/1 solution set of these systems; we will show how this successive approximation of a border basis can be turned into a notion of hierarchy and that this new hierarchy refines the Sherali-Adams hierarchy. This new hierarchy gives still rise to linear relaxations, and for any fixed $d \in \{1,\dots,n\}$, with n denoting the dimension of the polytope, we can optimize over the corresponding closure in polynomial time.

Our paper is organized as follows. In Section 2, we review the Sherali-Adams hierarchy and give a brief introduction into the theory of border bases. In Section 3, we establish the first connection between this theory and the reformulation-linearization technique of Sherali and Adams. In Section 4, we present the new hierarchy. Section 5 contains our concluding remarks.

2. Preliminaries

2.1. **The Sherali-Adams hierarchy.** We briefly recall the Sherali-Adams hierarchy as defined in [30]. Let P be a rational polytope contained in the n-dimensional 0/1-cube. We may assume, without loss of generality, that P is given as $P = \{x \in [0,1]^n : Ax = b\}$ with $A \in \mathbb{Z}^{m\times n}$ and $b \in \mathbb{Z}^m$. (If there were inequalities different from $0 \le x \le e$, one could express their slack in binary encoding by using at most a polynomial number of additional variables; e denotes the all-one vector.) We will refer to the equations given by the rows of the system $Ax = b$ as $a_i x = b_i$. We are interested in the integral hull $P_I := \operatorname{conv}(P \cap \mathbb{Z}^n)$ of P. Sherali and Adams defined a series of ever tighter relaxations by transforming the linear system $0 \le x \le e$, $Ax = b$ into a system of polynomials of maximum degree d, for $1 \le d \le n$, which is subsequently projected onto the space of x-variables. For convenience, let $[n] := \{1,\dots,n\}$ and consider $J_1, J_2 \subseteq [n]$. If $J_1 \cap J_2 = \emptyset$ and $|J_1 \cup J_2| = d$ we say that the pair (J_1,J_2) is of *order* d. For each pair (J_1,J_2) of order d we define

$$F_d(J_1,J_2) := \Big(\prod_{j\in J_1} x_j\Big)\Big(\prod_{j\in J_2}(1-x_j)\Big),$$

with the understanding that the product over the empty set is equal to one. Note that the polynomial $F_d(J_1,J_2)$ has degree d, and $F_d(J_1,J_2) \ge 0$ for $x \in [0,1]^n$. While the original Sherali-Adams procedure, which is formulated for systems of linear inequalities, multiplies each inequality with $F_d(J_1,J_2)$ for all pairs (J_1,J_2) of order d, the following lemma implies that for systems of linear equations it suffices to consider pairs (J_1,J_2) of order at most d such that $J_2 = \emptyset$.

Lemma 2.1. [30, p. 422] *Let (J_1,J_2) be of order d. Then*

$$F_d(J_1,J_2) = \sum_{p=|J_1|}^{d} \sum_{\substack{J\subseteq J_2\\ |J|=p-|J_1|}} (-1)^{|J|} F_p(J_1 \cup J, \emptyset).$$

One can therefore state the Sherali-Adams procedure as follows:

Algorithm 2.2. (*Sherali-Adams procedure for equality systems*)

Input: $P = \{x \in [0,1]^n : Ax = b\}$, $d \in [n]$.

Output: *The polytope* $\mathrm{AS}^{[d]}(P) \subseteq [0,1]^n$.

(1) *Multiply each equation $a_i x = b_i$ with $F_p(J,\emptyset)$ for all $(J,\emptyset)$ of order p with $0 \le p \le d$. We obtain a system of polynomial equations.*

(2) *Substitute any occurrence of x_j^2 by x_j for all $j \in [n]$.*

(3) *Add all polynomial inequalities of the form $F_d(J_1,J_2) \ge 0$ where (J_1,J_2) is a pair of order d.*

(4) *Linearize the polynomial system by substituting w_J for all monomials $\prod_{j\in J} x_j$ with $|J| \ge 2$. Let M^d be the resulting linear system.*

(5) *Set* $\mathrm{AS}^{[d]}(P) := \mathrm{proj}_X M^d$ *where* $X := \{x_1, \dots, x_n\}$.

A few remarks are in order. The substitution in Step 2 is motivated by the fact that any 0/1-solution x satisfies $x_j^2 - x_j = 0$. This substitution step is the one that actually tightens the linear relaxation. If one would skip Step 2 one would end up with the initial polytope P. With $AS^{[0]} = P$, one has the following hierarchy of relaxations.

Theorem 2.3. [30, Theorem 1, Theorem 3] *Let* $P \subseteq [0,1]^n$ *be a polytope. Then*

$$P = \mathrm{AS}^{[0]}(P) \supseteq \mathrm{AS}^{[1]}(P) \supseteq \cdots \supseteq \mathrm{AS}^{[n]}(P) = P_I.$$

Alternatively, the result also follows from Balas' sequential convexification [3] and the fact that $\mathrm{AS}^{[1]}(P)$ is contained in the lift-and-project closure N_0 of P (see, e.g., [8]). It is also possible to iterate the Sherali-Adams operator, by defining $\mathrm{AS}^{(i+1)}(P) := \mathrm{AS}^{[1]}(\mathrm{AS}^{(i)}(P))$. It still holds that $\mathrm{AS}^{(n)}(P) = P_I$ because $\mathrm{AS}^{[1]}(P)$ is contained in the lift-and-project closure. However, the iterated Sherali-Adams operator is weaker in general.

Lemma 2.4. [24, 25] *Let* $P \subseteq [0,1]^n$ *be a polytope and* $d \in [n]$. *Then*

$$\mathrm{AS}^{[d]}(P) \subseteq \mathrm{AS}^{(d)}(P).$$

The iterated version of the Sherali-Adams operator is in fact identical with the N operator of Lovász and Schrijver [25, 24]. In this sense, the Sherali-Adams hierarchy refines the Lovász-Schrijver N hierarchy. Let us finally note that it is not necessary to compute the explicit projection in Step 5 of Algorithm 2.2 as one can optimize over M^d directly. In particular, for fixed d, one can optimize in polynomial time over the relaxation on the d-th level of the Sherali-Adams hierarchy.

2.2. **Border bases.** We will now give a very brief introduction into the theory of border bases and computational commutative algebra. A more detailed introduction to computational commutative algebra can be found in [10, 21]. For an extensive coverage of border bases we refer to [19, 17, 18, 22]. Our exposition here follows that of [18]. Border bases represent a convenient way to characterize the solutions of a system of polynomial equations and can be considered as a generalization of the well-known Gröbner bases. In fact, we will later see that, under suitable assumptions, any border basis contains a reduced Gröbner basis.

We consider the polynomial ring $K[X]$ over the field K with indeterminates $X = \{x_1, \dots, x_n\}$. For convenience, we define $x^a := \prod_{j \in [n]} x_j^{a_j}$ for $a \in \mathbb{N}^n$ and let $\mathbb{T}^n := \{x^a \mid a \in \mathbb{N}^n\}$ be the *monoid of terms*. For any $d \in \mathbb{N}$, let $\mathbb{T}^n_{\leq d} := \{x^a \in \mathbb{T}^n \mid ||a||_1 \leq d\}$ be the set of monomials of total degree at most d. Furthermore, we fix the following total order σ on the monomials over X. Let x^a, x^b be two monomials, then we say that $x^a <_\sigma x^b$ if either $||a||_1 < ||b||_1$ or $||a||_1 = ||b||_1$ and $\min_{j \in \mathrm{supp}(a)} j < \min_{j \in \mathrm{supp}(b)} j$ where $\mathrm{supp}(m) := \{j \in [n] \mid m_j \neq 0\}$ is the *support of* $m \in \mathbb{N}^n$. All computations of border bases and Gröbner bases are done with respect to this ordering σ. For a polynomial $p \in K[X]$ with $p = \sum_{i=1}^l a_i x^{m_i}$ we define the *support of* p to be $\mathrm{supp}(p) := \{x^{m_i} \mid i \in [l], a_i \neq 0\}$, and, for a set of polynomials $P \subseteq K[X]$, we define the *support of* P to be $\mathrm{supp}(P) := \bigcup_{p \in P} \mathrm{supp}(p)$. The *leading term* $\mathrm{LT}(p)$ of a polynomial p is $\mathrm{LT}(p) := t$ with $t \in \mathrm{supp}(p)$ such that for all $t' \in \mathrm{supp}(P) \setminus \{t\}$ we have $t >_\sigma t'$; the *leading coefficient* $\mathrm{LC}(p)$ of p is the associated coefficient belonging to $\mathrm{LT}(p)$. We define the *degree of a polynomial* $p \in K[X]$ as $\deg(p) := \max_{x^m \in \mathrm{supp}(p)} ||m||_1$. For a set $\mathcal{M} \subseteq \mathbb{T}^n$ we define $[[\mathcal{M}]] := \{tm \mid t \in \mathbb{T}^n, m \in \mathcal{M}\}$, the monomial ideal generated by $\mathcal{M}$.

Definition 2.5. Let $\mathcal{O}$ be a finite subset of $\mathbb{T}^n$. If for all $t \in \mathcal{O}$ and $t' \in \mathbb{T}^n$ such that $t' \mid t$ we have $t' \in \mathcal{O}$, i.e., $\mathcal{O}$ is closed under division, then we call $\mathcal{O}$ an *order ideal*. Furthermore, the *border* $\partial\mathcal{O}$ of a non-empty order ideal $\mathcal{O}$ is the set of terms $\partial\mathcal{O} := \{x_j t \notin \mathcal{O} \mid j \in [n], t \in \mathcal{O}\}$; we set $\partial\emptyset := \{1\}$ for the empty order ideal.

An illustration of the definition is provided in Figure 2.1. The blue elements are in the order ideal and the grey elements constitute the border. The light-grey element is a border element but it is not the leading term of any Gröbner basis element as we will see later. In the following we will frequently switch between considering polynomials $P = \{p_1, \dots, p_s\}$ and the associated vector space whose coordinates are indexed by the monomials in the support of P. We will denote this vector space by $\langle P \rangle_K$. If A, B, and C are vector spaces, recall that $A = B \oplus C$, if $A = B + C$ and $B \cap C = \{0\}$. Let $\mathcal{O}$ be an order ideal, then the $\mathcal{O}$-border basis is a special set of polynomials:

Definition 2.6. Let $\mathcal{O} = \{t_1, \dots, t_\mu\}$ be an order ideal with border $\partial\mathcal{O} = \{b_1, \dots, b_\nu\}$. Further let $I \subseteq K[X]$ be an ideal and $\mathcal{G} = \{g_1, \dots, g_\nu\} \subseteq I$ be a finite set of polynomials.

(1) The set $\mathcal{G}$ is an *$\mathcal{O}$-border prebasis* if $g_j = b_j - \sum_{i=1}^{\mu} \alpha_{ij} t_i$ with $\alpha_{ij} \in K$ for all $j \in [\nu]$.

(2) An $\mathcal{O}$-border prebasis $\mathcal{G}$ is an *$\mathcal{O}$-border basis of I*, if $\langle \mathcal{G} \rangle = I$, i.e., $\mathcal{G}$ generates I and $K[X] = I \oplus \langle \mathcal{O} \rangle_K$ as vector spaces.

(3) If there exists an $\mathcal{O}$-border basis of I then we say that *$\mathcal{O}$ supports a border basis of I.*

Note that the condition $\langle \mathcal{G} \rangle = I$ is already implied by $\mathcal{G} \subseteq I$ and $K[X] = I \oplus \langle \mathcal{O} \rangle_K$, as was shown in [19]. Moreover, for any given order ideal $\mathcal{O}$ and ideal I the $\mathcal{O}$-border basis of I is unique as b_j has a unique representation in $K[X] = I \oplus \langle \mathcal{O} \rangle_K$ for all $j \in [\nu]$. An ideal I is zero-dimensional if $K[X]/I$ is finite dimensional vector space. We will now formulate the border basis algorithm for the computing a border basis of a zero-dimensional ideal with respect to an order ideal $\mathcal{O}$ that is induced by the term ordering σ.

Definition 2.7. Let W be a finite set of polynomials. We define the *neighborhood extension of W* to be

$$W^+ := W \cup W x_1 \cup \dots \cup W x_n.$$

Definition 2.8. Let $L \subseteq \mathbb{T}^n$ and let F be a finite set of polynomials such that $\operatorname{supp}(F) \subseteq L$. We inductively define the following sets of polynomials:

$$F_0 := F \quad \text{ and } \quad F_{k+1} := F_k^+ \cap L \text{ for } k \geq 0.$$

The union $F_L := \bigcup_{k \geq 0} F_k$ of the ascending chain $F_0 \subseteq F_1 \subseteq \dots$ is called the *L-stable span*.

In the following, we will explain how the L-stable span can be computed explicitly for $L = \mathbb{T}^n_{\leq d}$. We will need a modified version of Gaussian elimination:

Lemma 2.9. [18, Lemma 12] *Let $V = \{v_1, \dots, v_r\} \subseteq K[X] \setminus \{0\}$ be a finite set of polynomials such that* $\mathrm{LT}(v_i) \neq \mathrm{LT}(v_j)$ *whenever $i, j \in [r]$ with $i \neq j$ and* $\mathrm{LC}(v_i) = 1$ *for all $i \in [r]$. Further let $G = \{g_1, \dots, g_s\}$ be a finite set of polynomials. Then Algorithm 2.10 computes a finite set of polynomials $W \subseteq K[X]$ with* $\mathrm{LC}(w) = 1$ *for all $w \in W$,* $\mathrm{LT}(u_1) \neq \mathrm{LT}(u_2)$ *for any distinct $u_1, u_2 \in V \cup W$, and $\langle V \cup W \rangle_K = \langle V \cup G \rangle_K$.*

We will use a version of Gaussian elimination for polynomials which is identical to Gaussian elimination on the coefficient matrix with columns being indexed by the monomials and the rows being indexed by the polynomials. An exact formulation of the algorithm can be found in Appendix A.

Algorithm 2.10. (*Gaussian elimination for polynomials* - `GaussEl`)
Input: *V, G as in Lemma 2.9.*
Output: *$W \subseteq K[X]$ as in Lemma 2.9.*

We sometimes apply Gaussian elimination to a set V of polynomials that do not satisfy the assumptions of Lemma 2.9. In this case, it is implicitly assumed that V itself is replaced by $\mathtt{GaussEl}(\emptyset, V)$. Also, observe that the Gaussian elimination described above computes fully reduced elements, i.e., the associated matrix of coefficients is maximally interreduced in the linear algebra sense. We can now compute the L-stable span using Algorithm 2.10:

Lemma 2.11. [18, Proposition 13] *Let $F := \{f_1, \dots, f_r\} \subseteq K[X]$ be a finite set of polynomials and $L := \mathbb{T}^n_{\leq d}$ such that* $\operatorname{supp}(F) \subseteq L$ *and* $d := \max_{i \in [r]} \deg(f_i)$*. Then Algorithm 2.12 computes a vector space basis V of F_L with pairwise different leading terms.*

Algorithm 2.12. (*L-stable span computation* - LStabSpan)
Input: *F, L as in Lemma 2.11.*
Output: *V as in Lemma 2.11.*

```
(1) $V := \texttt{GaussEl}(\emptyset, F)$.
(2) $W' := \texttt{GaussEl}(V, V^+ \setminus V) \setminus V$.
(3) $W := \{w \in W' \mid \deg(w) \leq d\}$.
(4) If $|W| > 0$ set $V := V \cup W$ and go to step (2).
```

For the sake of simplicity, we also let $\texttt{LStabSpan}(F, d) := \texttt{LStabSpan}(F, \mathbb{T}^n_{\leq d})$. The last ingredient that we need in order to formulate the border basis algorithm is the final reduction algorithm. This algorithm applies linear algebra to interreduce the elements so that they only have support in the leading term and $\mathcal{O}$, as required by Definition 2.6.

Lemma 2.13. [18, Proposition 17] *Let $F = \{f_1, \ldots, f_s\}$ be a system of generators of a zero-dimensional ideal I. Let $\mathcal{L}$ be an order ideal and $L = \langle \mathcal{L} \rangle_K$ be the associated vector space. If V is a vector space basis of F_L with pairwise different leading terms and $\mathcal{O} := \mathcal{L} \setminus \mathrm{LT}(V)$ such that $L = F_L \oplus \langle \mathcal{O} \rangle_K$ and $\partial \mathcal{O} \subseteq \mathcal{L}$, then Algorithm 2.14 computes the $\mathcal{O}$-border basis $\mathcal{G} = \{g_1, \ldots, g_\nu\}$ of I.*

Algorithm 2.14. (*Final Reduction Algorithm* - FinalRed)
Input: *V, $\mathcal{O}$ as in Lemma 2.13.*
Output: *$\mathcal{G}$ as in Lemma 2.13.*

```
(1) Let $V_R := \emptyset$.
(2) If $V = \emptyset$ return $\emptyset$ and stop.
(3) Let $v \in V$ such that $v$ has minimal leading term. Set $V := V \setminus \{v\}$.
(4) Let $H := \mathrm{supp}(v) \setminus (\mathrm{LT}(v) \cup \mathcal{O})$.
(5) If $H = \emptyset$ then append $v/\mathrm{LC}(v)$ to $V_R$ and go to step (2).
(6) For each $h \in H$ choose $w_h \in V_R$ and $c_h \in K$ such that $\mathrm{LT}(w_h) = h$ and $h \notin \mathrm{supp}(v - c_h w_h)$.
(7) Set $v := v - \sum_{h \in H} c_h w_h$, append $v/\mathrm{LC}(v)$ to $V_R$, and go to step (2).
(8) Return $\mathcal{G} := \{v \in V_R : \mathrm{LT}(v) \in \partial \mathcal{O}\}$.
```

We can now formulate the border basis algorithm.

Proposition 2.15. [18, Proposition 18] *Let $F = \{f_1, \ldots, f_s\} \subseteq K[X]$ be a finite set of polynomials that generate a zero-dimensional ideal $I = \langle F \rangle_K$. Then Algorithm 2.16 computes the $\mathcal{O}$-border basis $\mathcal{G}$ of I.*

Algorithm 2.16. (*Border basis algorithm* - BBasis)
Input: *F as in Proposition 2.15.*
Output: *$\mathcal{G}$ as in Proposition 2.15.*

```
(1) Let $d := \max_{f \in F}\{\deg(f)\}$ and put $L := \mathbb{T}^n_{\leq d}$.
(2) $V = \{v_1, \ldots, v_r\} := \texttt{LStabSpan}(F, L)$.
(3) Let $\mathcal{O} := \mathbb{T}^n_{\leq d} \setminus \{\mathrm{LT}(v_1), \ldots, \mathrm{LT}(v_r)\}$.
(4) If $\partial \mathcal{O} \not\subseteq L$ then set $d := d + 1$ and put $L := \mathbb{T}^n_{\leq d}$ and go to step (2).
(5) Set $\mathcal{G} := \texttt{FinalRed}(V, \mathcal{O})$.
```

In our specified setting, where the border bases are derived from a degree-compatible term ordering σ, the border basis of a finite set of polynomials F that generates a zero-dimensional ideal $\langle F \rangle$ contains a reduced Gröbner basis of the ideal $\langle F \rangle$: If $\mathcal{G}$ is the $\mathcal{O}$-border basis of $\langle F \rangle$, then $\tilde{\mathcal{G}} := \{g \in \mathcal{G} \mid \forall t \in \mathbb{T}^n$ with $t | \mathrm{LT}(g)$ we have $t \in \mathcal{O}\}$ is a reduced (σ-)Gröbner basis [18]. In the next section, we will consider finite systems of polynomials of the form $F = \{Ax - b, x_j^2 - x_j \mid j \in [n]\}$ with $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$. Even in this restricted case, the border basis and the Gröbner basis need not coincide:

Example 2.17. Consider the system $F := \{x_1 + x_2 - 1, x_1^2 - x_1, x_2^2 - x_2\}$ with σ being the degree-lexicographic ordering. The border basis $\mathcal{G}$ of F is given by $\mathcal{G} := \{x_1 + x_2 - 1, x_2^2 - x_2, x_1 x_2\}$ with

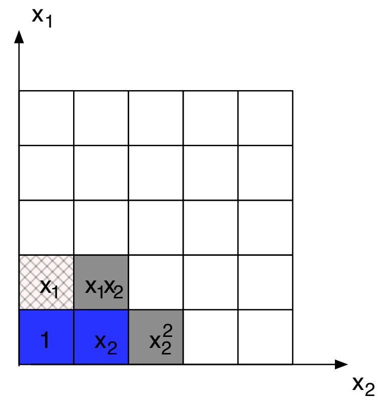

Figure 2.1. Order ideal and border

border $\{x_1, x_1x_2, x_2^2\}$. As $\mathrm{LT}(x_1+x_2-1) = x_1 | x_1x_2 = \mathrm{LT}(x_1x_2)$ and $x_1 \in \mathrm{LT}(\mathcal{G})$ and, therefore, $x_1 \notin \mathcal{O}$, it follows that $\mathcal{G}$ is not a (reduced) Gröbner basis. The example is depicted in Figure 2.1. The light-grey element is indeed in the border but is not a leading term of any element in the Gröbner basis, basically because $x_1 | x_1x_2$.

We will finish this section with an observation that in the case of F being of the form $\{Ax-b, x_j^2-x_j \mid j \in [n]\}$, the border basis algorithm stops for some $d \in [n]$. Note that this is not true in general for generic systems of generators F of a zero-dimensional ideal.

Proposition 2.18. *Let $F = \{Ax-b, x_j^2-x_j \mid j \in [n]\}$ with $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$ be a finite set of polynomials. Then the border bases algorithms stops for some $d \in [n]$.*

Proof. Let F be as above. Suppose that for $d \in [n-1]$ and $\{v_1, \dots, v_r\} := \mathtt{LStabSpan}(F, d)$ we have $\partial(\mathbb{T}^n_{\leq d} \setminus \{\mathrm{LT}(v_1), \dots, \mathrm{LT}(v_r)\}) \not\subseteq \mathbb{T}^n_{\leq d}$. We will show that $\partial(\mathbb{T}^n_{\leq n} \setminus \{\mathrm{LT}(v_1), \dots, \mathrm{LT}(v_r)\}) \subseteq \mathbb{T}^n_{\leq n}$ with $\{v_1, \dots, v_r\} := \mathtt{LStabSpan}(F, n)$ and then the assertion follows. Let $\{v_1, \dots, v_r\} = \mathtt{LStabSpan}(F, n)$. We claim that for every monomial $x^w \in \mathbb{T}^n_{\leq n}$ with $\deg(x^w) = n$ there exists $i \in [r]$ such that $\mathrm{LT}(v_i) = x^w$. Suppose that there exists $j \in [n]$ such that $w_j > 1$. Write $w = (w-2e_j) + 2e_j$. Then $x^{2e_j} = \mathrm{LT}(x_j^2 - x_j)$ and as $x_j^2 - x \in F_j$ we have that $x^{w-2e_j}x^{2e_j} = \mathrm{LT}(x^{w-2e_j}(x_j^2-x_j)) = x^{w-2e_j}\mathrm{LT}(x_j^2-x_j)$. Note that $x^{w-2e_j}(x_j^2-x_j)$ is contained in $\mathtt{LStabSpan}(F, n)$ by construction. Now we consider the case that $w = e$. We have to distinguish two cases. If $\mathrm{defect}(A) = n$, then the border bases algorithm stops after one round as $\{x_j^2 - x_j \mid j \in [n]\}$ constitutes a border bases[1]. Now consider the case that $\mathrm{defect}(A) < n$. then there exists $j \in [n]$ and $i \in [r]$ such that $x_j = \mathrm{LT}(v_i)$. We can write $w = (e-e_j)+e_j$ and argue as above. Thus every monomial $x^w \in \mathbb{T}^n_{\leq n}$ with $\deg(x^w) = n$ occurs as a leading term. Therefore $\max\{\deg(m) \mid m \in \mathbb{T}^n_{\leq n} \setminus \{\mathrm{LT}(v_1), \dots, \mathrm{LT}(v_r)\}\} \leq n-1$ and so $\max\{\deg(m) \mid m \in \partial(\mathbb{T}^n_{\leq d} \setminus \{\mathrm{LT}(v_1), \dots, \mathrm{LT}(v_r)\})\} \leq n$ and the proof is completed. □

3. A first connection between the Sherali-Adams procedure and the border basis algorithm

We will now establish the connection between the Sherali-Adams procedure and the border basis algorithm. We consider a polytope $P := \{x \in [0,1]^n \mid Ax = b\} \subseteq [0,1]^n$ with $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$ to which we will apply the Sherali-Adams procedure. Simultaneously, we will consider the polynomial equality system $F := \{Ax-b, x_j^2-x_j \mid j \in [n]\}$ to which we apply the border basis algorithm. We will now slightly rewrite the border basis algorithm (Algorithm 2.16) to reflect different hierarchies for $d \in [n]$ similar to the hierarchies implied by the Sherali-Adams operator $\mathrm{AS}^{[d]}$ for $d \in [n]$. The new algorithm performs exactly the same operations except for stopping after a predefined number d of iterations and skipping the final reduction step. To motivate the last part, let us have another

[1] Here, $\mathrm{defect}(A) := n - \mathrm{rank}(A)$, i.e., the column defect.

look at the border basis algorithm. In the main part of the algorithm, we compute an L-stable span of sufficiently high degree. In the second part, the final reduction algorithm removes all polynomials whose leading terms are not contained in the border as those polynomials can be regenerated by the border bases in the ring theoretic setting. Contrary to this, in the K-vector space setting (the basis of linear programming) we cannot multiply two polynomials with each other and therefore we need to keep these polynomials. In fact, we would otherwise lose crucial information as shown in Examples B.5 and B.3. We therefore consider the L-stable span generation only.

Algorithm 3.1. (*d-stable span computation* - `dStabSpan`)

```
Input: A finite set of polynomials $F = \{f_1, \dots, f_s\}$ and $d \in [n]$.
Output: A finite set of polynomials $V$.
  (1) $V := \texttt{GaussEl}(\emptyset, F)$. Set $m := 1$.
  (2) If $m < d$ then
      (a) $W' := \texttt{GaussEl}(V, V^+ \setminus V) \setminus V$.
      (b) $W := \{w \in W' \mid \deg(w) \leq d\}$. Set $V := V \cup W$.
      (c) Set $m := m + 1$ and go to step (2).
```

We also slightly adapt the border basis algorithm to perform a predefined number of iterations:

Algorithm 3.2. (*d-border basis algorithm* - `dBBasis`)

```
Input: A finite set of polynomials $F = \{f_1, \dots, f_s\}$ that generate a zero-dimensional ideal and $d \in [n]$.
Output: A finite set of polynomials $V$.
  (1) $V = \{v_1, \dots, v_r\} := \texttt{dStabSpan}(F, d)$.
```

We can now link the border basis algorithm and the Sherali-Adams procedure. For a finite system of polynomial equations F, we define the *projection* $\mathscr{R}(F)$ to be the polytope obtained by relinearization of F and subsequent projection onto the space of monomials of degree ≤ 1 together with boundary constraints $0 \leq x_j \leq 1$ for all $j \in [n]$. This well-known *reformulation-linearization-technique* (*RLT*) (cf. [30, 24]) is the conceptual basis of the Sherali-Adams hierarchy, where this corresponds to steps (3), (4), and (5) of Algorithm 2.2. Note first that Gaussian elimination does not affect the projection onto the polytope $\mathscr{R}(F)$:

Lemma 3.3. *Let $L, M \subseteq K[X]$ be two finite sets of polynomials. Then $\langle \texttt{GaussEl}(L, M)\rangle_K = \langle L \cup M\rangle_K$.*

If we consider a finite set of polynomials L, we have that $L^+ = L \cup \bigcup_{j\in[n]} x_j L$ or, equivalently as K-vector spaces,

$$\langle L^+\rangle_K = \langle L \cup \bigcup_{j\in[n]} x_j L\rangle_K = \langle L\rangle_K + \sum_{j\in[n]} \langle x_j L\rangle_K.$$

As the maps $f \mapsto x_j f$ from $K[X]$ to $K[X]$ are K-vector space homomorphisms, we have $\langle L\rangle_K + \sum_{j\in[n]}\langle x_j L\rangle_K = \langle L\rangle_K + \sum_{j\in[n]} x_j \langle L\rangle_K =: \langle L\rangle_K^+$ and thus $\langle L^+\rangle_K = \langle L\rangle_K^+$. Together with Lemma 3.3 we can conclude:

Lemma 3.4. *Let $L \subseteq K[X]$ be a finite sets of polynomials. Then $\langle \texttt{GaussEl}(L, L^+\setminus L)\rangle_K = \langle \texttt{GaussEl}(\emptyset, L)\rangle_K^+$.*

Proof. Observe that $\langle \texttt{GaussEl}(L, L^+ \setminus L)\rangle_K = \langle L^+\rangle_K$ by Lemma 3.3. With the discussion from above we have $\langle L^+\rangle_K = \langle L\rangle_K^+$ and again applying Lemma 3.3 yields $\langle L\rangle_K^+ = \langle \texttt{GaussEl}(\emptyset, L)\rangle_K^+$. □

If $L, M \subseteq K[X]$ are two finite sets of polynomials and $\mathscr{R}$ is the linearization-projection map from above, then clearly $\mathscr{R}(L) = \mathscr{R}(M)$ if $\langle L\rangle_K = \langle M\rangle_K$ as K-vector spaces:

Lemma 3.5. *Let $L, M \subseteq K[X]$ be two finite sets of polynomials. Then $\mathscr{R}(\texttt{GaussEl}(L, M)) = \mathscr{R}(L \cup M)$.*

Lemmas 3.3, 3.4, and 3.5 lead to our first main result: The truncated border basis algorithm `dStabSpan` and the Sherali-Adams procedure generate the same linear relaxation. For a set of polynomials $M \subseteq K[X]$ let $M^{\leq d} := \{f \in M \mid \deg(f) \leq d\}$. Note that $\langle M\rangle^{\leq d} \neq \langle M^{\leq d}\rangle$ as degree truncation is not a vector space homomorphism.

Theorem 3.6. *Let $P = \{x \in [0,1]^n \mid Ax = b\}$ be a rational polytope, and let $F = \{Ax-b, x_j^2-x_j \mid j \in [n]\}$ be the associated set of polynomials. Then* $\mathrm{AS}^{[d]}(P) = \mathscr{R}(\texttt{dStabSpan}(F, d+1))$.

Proof. Note that the degree truncation only affects polynomials that arise from combinations with one of the polynomials $x_j^2 - x_j$, by a degree argument. All these polynomials only perform substitutions $x_j^2 \mapsto x_j$, written in matrix form. Moreover, truncation only occurs after the last extension $+$. By Lemma 3.4, $\langle \texttt{dStabSpan}(F, d+1)\rangle_K = \langle \texttt{GaussEl}(\emptyset, ((F^+)^{\cdots})^+)^{\le d+1}\rangle_K$ where the neighborhood extension $+$ is performed $d+1$ times. With Lemma 3.3, we have $\langle \texttt{GaussEl}(\emptyset, ((F^+)^{\cdots})^+)^{\le d+1}\rangle_K = \langle(((F^+)^{\cdots})^+)^{\le d+1}\rangle_K$. Finally, observe that $(((F^+)^{\cdots})^+)^{\le d+1}$ coincides with the unprojected Sherali-Adams system $M^d(P)$. Thus,

$$\langle M^d(P)\rangle_K = \langle(((F^+)^{\cdots})^+)^{\le d+1}\rangle_K = \langle \texttt{GaussEl}(\emptyset, ((F^+)^{\cdots})^+)^{\le d+1}\rangle_K = \langle \texttt{dStabSpan}(F, d+1)^{\le d+1}\rangle_K.$$

Lemma 3.5 implies that $\mathrm{AS}^{[d]}(P) = \mathscr{R}(M^d(P)) = \mathscr{R}(\texttt{dStabSpan}(F, d+1))$. □

4. A new hierarchy of relaxations refining the Sherali-Adams hierarchy

Using the border basis framework, we will now derive a new hierarchy of relaxations that refines the Sherali-Adams hierarchy by exploiting Gaussian elimination. For this, we need a version of the border basis algorithm that uses the original LStabSpan procedure, but is limited to the computational universe $\mathbb{T}^n_{\le d}$, for given $d \in [n+1]$.

Algorithm 4.1. (*Border basis algorithm* for *(fixed) degree d* - BBasis)
Input: *A finite set of polynomials $F = \{f_1, \ldots, f_s\}$ generating a zero-dimensional ideal and $d \in [n+1]$.*
Output: *A finite set of polynomials V.*

```
(1) $V = \{v_1, \ldots, v_r\} :=$ LStabSpan($F, d$).
```

Clearly, if we choose $d \in [n]$ as obtained at the end of Algorithm 2.16, then for any given finite set of polynomials $F = \{f_1, \ldots, f_s\}$ that generate a zero-dimensional ideal the outputs of Algorithm 4.1 for the input F, d followed by final reduction, and Algorithm 2.16 coincide. Abusing notation slightly, we denote both algorithms by BBasis as the distinction will be clear from the context. The reason that Algorithm 4.1 yields stronger relaxations than Algorithm 3.1 (and hence, by Theorem 3.6, than Sherali-Adams) is that, for given $d \in [n+1]$, Algorithm 3.1 stops after d rounds. In contrast, Algorithm 4.1 may perform additional neighborhood extensions, temporarily creating additional polynomials that exceed the maximum allowed degree d, which are subsequently reduced by Gaussian elimination. The reason why this is possible is that degree truncation is not a vector space homomorphism. We start with the following observation:

Lemma 4.2. *Let $L, M \subseteq K[X]$ be two finite sets of polynomials. Then* $\mathrm{LT}(\texttt{GaussEl}(L, M)) \supseteq \mathrm{LT}(L \cup M)$.

Proof. Let $p \in L \cup M$ with leading term m. Without loss of generality suppose that p is chosen first among all polynomials in $L \cup M$ with leading term m by the GaussEl procedure. If $q \in L \cup M$ with $\mathrm{LT}(q) = m$, then GaussEl replaces q with $q - \frac{\mathrm{LC}(q)}{\mathrm{LC(p)}}p$ and therefore $\mathrm{LT}(q - \frac{\mathrm{LC}(q)}{\mathrm{LC(p)}}p) < \mathrm{LT}(q)$. The first occurrence p though is replaced with $\frac{1}{\mathrm{LC}(p)}p$ and thus has leading term m. Therefore $m \in \mathrm{LT}(\texttt{GaussEl}(L, M))$. □

It is worthwhile to observe that the replacement $q - \frac{\mathrm{LC}(q)}{\mathrm{LC(p)}}p$ might have a leading term that is potentially not contained in $\mathrm{LT}(L \cup M)$. This can indeed occur as we will see later in Remark B.2 and exactly this fact is the basis for our refinement of the Sherali-Adams hierarchy. We will now show that LStabSpan(F, d) stops after a small number of rounds.

Theorem 4.3. *Let $F = \{Ax - b, x_j^2 - x_j \mid j \in [n]\}$ with $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$. Further let $d \in [n+1]$. Then* LStabSpan(F, d) *terminates after at most $d + \tau$ executions of Step (2) where $\tau = \binom{d+\mathrm{defect}(A)+1}{\mathrm{defect}(A)+1}$ if* $\mathrm{defect}(A) > 0$*, and after at most d iterations if* $\mathrm{defect}(A) = 0$.

Proof. We use the index i to refer to the sets V, W and W' in round i. In the first round, $\texttt{LStabSpan}(F, L)$ computes a reduced representation V_1 of F such that any two polynomials $f_1, f_2 \in F$ have pairwise different leading terms. Let $\mathscr{L}_i = \mathrm{LT}(V_i)$ denote the set of leading terms at the beginning of round i, and let $ax - b$ be an arbitrary linear equality in V_1. Then $\mathrm{LT}(ax - b) = x_j$ for some $j \in [n]$. Let $\mathscr{M} = \{x_j \mid j \in [n]\} \setminus \mathscr{L}_1$ be the set of variables which do not occur as leading terms in the initial reduced system V_1. In round i, we compute $\widetilde{W_i} := V_i^+ \setminus V_i$. Thus, if $x^m = \mathrm{LT}(v)$ with $v \in V_i$, we obtain $x^{m+e_j} = \mathrm{LT}(x_j v) = x_j \mathrm{LT}(v)$, and so we have $\mathrm{LT}(\widetilde{W_i}) = \{x_j t \mid j \in [n], t \in \mathscr{L}_i\} = \bigcup_{j \in [n]} x_j \mathscr{L}_i$. If we now consider $W_i' := \texttt{GaussEl}(V_i, V_i^+ \setminus V_i) \setminus V_i$ then it is easy to see that $\mathrm{LT}(\widetilde{W_i}) \subseteq \mathrm{LT}(W_i')$ by Lemma 4.2. We can therefore conclude

$$\mathscr{L}_{i+1} = \mathrm{LT}(V_i) \cup \mathrm{LT}(W_i') \supseteq \mathrm{LT}(V_i) \cup \mathrm{LT}(\widetilde{W_i}) = \mathrm{LT}(V_i) \cup \bigcup_{j \in [n]} x_j \mathscr{L}_i = \mathscr{L}_i \cup \bigcup_{j \in [n]} x_j \mathscr{L}_i.$$

This recursive relation can be unfolded to

$$\mathscr{L}_i \supseteq \bigcup_{||m||_1 < i} x^m \mathscr{L}_1.$$

Now let $x^w \in \mathbb{T}^n_{\leq d}$ with $x^w \notin [[\mathscr{M}]]$. We claim that $x^w \in \mathscr{L}_{||w||_1}$. If $||w||_1 = 1$, then $x^w \in \mathscr{L}_1$ by definition of $\mathscr{M}$. Now suppose that there exists $j \in [n]$ such that $w_j = \ell > 1$. Write $w = (w - 2e_j) + 2e_j$. Then $x^{2e_j} = \mathrm{LT}(x_j^2 - x_j)$ and thus $x^{2e_j} \in \mathscr{L}_1$ and therefore $x^{w-2e_j} x^{2e_j} \in \mathscr{L}_{||w||_1 - 1} \subseteq \mathscr{L}_{||w||_1}$. Suppose that $w \leq e$. As $x^w \notin [[\mathscr{M}]]$ there exists $j \in [n]$ such that $x_j \notin \mathscr{M}$ and thus we can consider $w = (w - e_j) + e_j$. With the same argumentation as before, $x^j \in \mathscr{L}_1$ and $x^{(w-e_j)} x^{e_j} \in \mathscr{L}_{||w||_1}$.

We can thus conclude that after d iterations any leading term $x^w \in \mathbb{T}^n_{\leq d}$ with $x^w \notin [[\mathscr{M}]]$ has been generated. If $\mathrm{defect}(A) = 0$ then $\mathscr{M} = \emptyset$ and thus all potential leading terms of degree $\leq d$ have been generated and therefore the algorithm

Suppose now that $\mathrm{defect}(A) > 0$, i.e., $\mathscr{M} \neq \emptyset$ and that there are still leading terms $x^w \in [[\mathscr{M}]]$ with $\deg(x^w) \leq d$ missing. Then any further iteration leads to at least one new leading term; otherwise the procedure stops. Note that $|\mathscr{M}| = \mathrm{defect}(A)$ and thus there are at most $\tau := \binom{d+\mathrm{defect}(A)+1}{\mathrm{defect}(A)+1}$ leading terms missing. Therefore, after at most τ additional iterations the procedure stops. □

We will now show that $\texttt{LStabSpan}(F, d) \supseteq \texttt{dStabSpan}(F, d)$, which establishes that our new procedure leads to stronger relaxations than Sherali-Adams.

Lemma 4.4. *Let $F = \{Ax - b, x_j^2 - x_j \mid j \in [n]\}$ with $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$, and let $d \in [n+1]$. Then* $\texttt{LStabSpan}(F, d) \supseteq \texttt{dStabSpan}(F, d)$.

Proof. It suffices to observe that in Algorithm $\texttt{dStabSpan}$ whenever $W_m = \emptyset$ for some round m, then $W_l = \emptyset$ for all $l > m$. Put differently, one could terminate Algorithm $\texttt{dStabSpan}$ at the first occurrence of $W_m = \emptyset$, without changing the output of the algorithm. The claim follows. □

The essential difference between $\texttt{dStabSpan}$ and $\texttt{LStabSpan}$ lies in the intermediate Gauss eliminations. These steps remove duplicate leading terms and thus permit a monomial that is not a leading term to become a leading term, as we have seen in Lemma 4.2. Of course, if $\mathrm{defect}(A) = 0$, then $\texttt{LStabSpan}(F, d) = \texttt{dStabSpan}(F, d)$:

Lemma 4.5. *Let $F = \{Ax - b, x_j^2 - x_j \mid j \in [n]\}$ with $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$ such that* $\mathrm{defect}(A) = 0$*, and let $d \in [n+1]$. Then* $\texttt{LStabSpan}(F, d) = \texttt{dStabSpan}(F, d)$.

Proof. By Lemma 4.4, we have $\texttt{LStabSpan}(F, d) \supseteq \texttt{dStabSpan}(F, d)$. As $\mathrm{defect}(A) = 0$, Theorem 4.3 implies that $\texttt{LStabSpan}(F, d)$ performs at most d rounds. We conclude that $\texttt{LStabSpan}(F, d) \subseteq \texttt{dStabSpan}(F, d)$ as both procedures are identical except for the latter performing exactly d rounds and thus generating at least the polynomials contained in $\texttt{LStabSpan}(F, d)$. This completes the proof. □

In Section B we include examples which show that $\texttt{LStabSpan}(F,d) = \texttt{dStabSpan}(F,d)$ does not hold in general. We will now define a new hierarchy of relaxations that is derived from the border basis algorithm of the polynomial system $F = \{Ax - b, x_j^2 - x_j \mid j \in [n]\}$ associated with the polytope $P = \{x \in [0,1]^n \mid Ax = b\}$ with $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$.

Definition 4.6. Let $P = \{x \in [0,1]^n \mid Ax = b\}$ be a rational polytope and let $F = \{Ax - b, x_j^2 - x_j \mid j \in [n]\}$ be the associated set of polynomials. The *d-th border basis closure of P* is defined as $\mathrm{BC}^{[d]}(P) := \mathscr{R}(\texttt{LStabSpan}(F, d+1))$.

Lemma 4.4 implies that the border basis hierarchy refines the Sherali-Adams hierarchy:

Theorem 4.7. *Let $P = \{x \in [0,1]^n \mid Ax = b\}$ be a 0/1 polytope and $d \in [n]$. Then $\mathrm{BC}^{[d]}(P) \subseteq \mathrm{AS}^{[d]}(P)$.*

This additional strength arises from the Gaussian elimination step which permits additional leading terms (and hence polynomial equations) to be generated (see also Theorem 4.3). In fact, we generate all possible (linearly generated) syzygies that are contained in our computational universe. The consideration of syzygies is the crucial point in the classical Gröbner bases algorithm and finds its natural resemblance in the $\texttt{LStabSpan}$ procedure in our setting. This refinement is a genuine refinement as shown in Example B.1.

Observe that $\mathrm{BC}^{[d]}(P)$ is still a linear closure in contrast to, e.g., the Lasserre hierarchy, which also refines the Sherali-Adams hierarchy. We will show now that one can still optimize in polynomial time over $\mathrm{BC}^{[d]}(P)$, provided d is fixed. (The same was known to be true for $\mathrm{AS}^{[d]}(P)$.)

Theorem 4.8. *Let $P = \{x \in [0,1]^n \mid Ax = b\}$ with $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$, and let $c \in \mathbb{Q}^n$. Furthermore, let $d \in [n]$ be fixed. Then one can optimize c over $\mathrm{BC}^{[d]}(P)$ in time polynomial in n and the input size of A, b, and c.*

Proof. Let $F = \{Ax - b, x_j^2 - x_j \mid j \in [n]\}$. Then F has $m + n$ equations. Moreover, there are at most $\binom{n+d+2}{d+2}$ monomials of degree $\leq d+1$ that can occur as leading terms of the polynomials in $\texttt{LStabSpan}(F, d+1)$. As the new variables introduced by relinearization correspond to monomials of degree at least two, the resulting linear system has at most $\mathrm{O}(n^{d+1})$ variables.

It remains to show that the size of the largest coefficient of the resulting system $\texttt{LStabSpan}(F, d+1)$ is polynomial in the size of A, b, and c. But this is clear as the coefficients are only subject to changes due to Gaussian elimination steps. The result follows, as we can optimize over the relinearized system (without having to explicitly perform the projection). □

5. Concluding Remarks

It is interesting to note that neither border bases nor Gröbner bases can be used in reduced form to obtain the integral hull of a polytope through projection (cf. Examples B.5 and B.3). Here, the final reduction algorithm destroys important information contained in the polynomial systems that in turn is not accessible in the relinearization-projection step. Hence, the final reduction algorithm must not be used when aiming for the integral hull through relinearization. Skipping the final reduction algorithm, the border basis algorithm leads to a stronger version of the Sherali-Adams procedure.

Border bases have also been used in coding theory [5]. The connection established in this paper may give rise to possible applications in crypto analysis, where border bases have been used to solve sparse quadratic systems of equalities. Such systems naturally arise from crypto systems (such as AES, BES, HFE, DES, CTC variants, etc.) when rewriting the S-boxes as polynomial equations. Our results provide the missing link between these algebraic methods and the Sherali-Adams procedure. It follows from our work that the celebrated XL, XSL, MutantXL attacks, which are based on relinearization methods, are essentially equivalent to the reformulation-linearization-technique of Sherali and Adams. In fact, it turns out that the XL algorithm (see, e.g., [20, 9]) in its classical form is identical to a level d Sherali-Adams closure. Interpreting the XL algorithm as the Sherali-Adams procedure in a cutting-plane framework opens up a wide variety of techniques from cutting plane theory that could be applied in order to attack ciphers.

References

[1] J. Abbott, C. Fassino, and M.-L. Torrente. Stable border bases for ideals of points. *Journal of Symbolic Computation*, 43:883–894, 2008.

[2] W. Auzinger and H.J. Stetter. An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations. In *Proceedings of the International Conference on Numerical Mathematics*, pages 11–30. National University of Singapore, May 31-June 4, 1988, Birkhäuser, 1988.

[3] E. Balas. Disjunctive programming: Properties of the convex hull of feasible points. *Discrete Applied Mathematics*, 89:3–44, 1998.

[4] D. Bertsimas, G. Perakis, and S. Tayur. A New Algebraic Geometry Alogorithm for Integer Programming. *Management Science*, pages 999–1008, 2000.

[5] M. Borges-Quintana, M.A. Borges-Trenard, and E. Martínez-Moro. An Application of Möller's Algorithm to Coding Theory. In M. Sala, T. Mora, L. Perret, S. Sakata, and C. Traverso, editors, *Gröbner Bases, Coding, and Cryptography*, pages 379–384. Springer, 2009.

[6] M. Charikar, K. Makarychev, and Y. Makarychev. Integrality gaps for Sherali-Adams relaxations. *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 283–292, 2009.

[7] CoCoATeam. CoCoA: a system for doing Computations in Commutative Algebra. Available at `http://cocoa.dima.unige.it`.

[8] G. Cornuéjols. Valid inequalities for mixed integer linear programs. *Mathematical Programming*, 112:3–44, 2008.

[9] N. Courtois, A. Klimov, J. Patarin, and A. Shamir. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. *Lecture Notes in Computer Science*, 1807:392–407, 2000.

[10] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2007.

[11] S. Dantchev. Rank complexity gap for Lovász-Schrijver and Sherali-Adams proof systems. *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 311–317, 2007.

[12] J.A. De Loera, J. Lee, P.N. Malkin, and S. Margulies. Hilbert's Nullstellensatz and an algorithm for proving combinatorial infeasibility. In *Proceedings of the twenty-first international symposium on Symbolic and algebraic computation - ISSAC '08*, 2008.

[13] J.A. De Loera, J. Lee, S. Margulies, and S. Onn. Expressing Combinatorial Problems by Systems of Polynomial Equations and Hilbert's Nullstellensatz. *Combinatorics, Probability and Computing*, 18(4):551–582, 2009.

[14] J.A. De Loera, P.N. Malkin, and P.A. Parrilo. Computation with polynomial equations and inequalities arising in combinatorial optimization. *preprint*, 2009.

[15] E. Gawrilow and M. Joswig. polymake: a framework for analyzing convex polytopes. In Gil Kalai and Günter M. Ziegler, editors, *Polytopes — Combinatorics and Computation*, pages 43–74. Birkhäuser, 2000.

[16] D. Heldt, M. Kreuzer, S. Pokutta, and H. Poulisse. Approximate computation of zero-dimensional polynomial ideals. *Journal of Symbolic Computation*, 44(11):1566–1591, 2009.

[17] A. Kehrein and M. Kreuzer. Characterizations of border bases. *Journal of Pure and Applied Algebra*, 196:251–270, 2005.

[18] A. Kehrein and M. Kreuzer. Computing border bases. *Journal of Pure and Applied Algebra*, 205:279–295, 2006.

[19] A. Kehrein, M. Kreuzer, and L. Robbiano. An algebraist's view on border bases. In *Solving Polynomial Equations: Foundations, Algorithms, and Applications*, pages 169–202. Springer, 2005.

[20] M. Kreuzer. Algebraic attacks galore! *Preprint*, 2009.

[21] M. Kreuzer and L. Robbiano. *Computational Commutative Algebra 1*. Springer, 2000.

[22] M. Kreuzer and L. Robbiano. *Computational Commutative Algebra 2*. Springer, 2005.

[23] M. Kreuzer and L. Robbiano. Deformations of border bases. *Collectanea Mathematica*, 59:275–297, 2008.

[24] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research*, 28:470–496, 2003.

[25] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1:166–190, 1991.

[26] C. Mathieu and A. Sinclair. Sherali-Adams relaxations of the matching polytope. *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 293–302, 2009.

[27] H.M. Möller. Systems of algebraic equations solved by means of endomorphisms. *Lecture Notes in Computer Science*, 673:43–56, 1993.

[28] B. Mourrain. A new criterion for normal form algorithms. *Lecture Notes in Computer Science*, 1719:430–443, 1999.

[29] M. Rhodes. Rank lower bounds for the Sherali-Adams operator. *Lecture Notes in Computer Science*, 4497:648–659, 2007.

[30] H.D. Sherali and W.P. Adams. A hierarchy of relaxations between the continous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3:411–430, 1990.

Appendix A. Gaussian Elimination for polynomials

Algorithm A.1. (*Gaussian elimination for polynomials* - `GaussEl`)

```
Input: $V$, $G$ as in Lemma 2.9.
Output: $W \subseteq K[X]$ as in Lemma 2.9.
  (1) Let $H := G$ and $\eta := 0$.
  (2) If $H = \emptyset$ then return $W := \{v_{r+1}, \dots, v_{r+\eta}\}$ and stop.
  (3) Choose $f \in H$ and remove it from $H$. Let $i := 1$.
  (4) If $f = 0$ or $i > r + \eta$ then go to step (7).
  (5) If $\mathrm{LT}(v_i) \in \mathtt{Support}(f)$ then replace $f$ with $f - \mathtt{Coeff}(f, \mathrm{LT}(v_i)) \cdot v_i$. Set $i := 1$; go to step (4).
  (6) Else, set $i := i + 1$. Go to step (4).
  (7) If $f \neq 0$ then put $\eta := \eta + 1$ and let $v_{r+\eta} := f / \mathrm{LC}(f)$. Go to step (2).
```

Here `Support`(f) *denotes the (monomial) support of the polynomial* f *and* `Coeff`(f, m) *the coefficient of the monomial* m *in* f.

Appendix B. Examples

We present a few examples highlighting the differences of `LStabSpan` and `dStabSpan`. In particular, we present an example showing that the border basis rank can be strictly smaller than the Sherali-Adams rank without generating polynomials of degree higher than d. Therefore the border basis hierarchy is a genuine refinement of the Sherali-Adams hierarchy. This additional strength arises from the Gaussian elimination step which permits additional leading terms (and hence polynomial equations) to be generated (see also Theorem 4.3 and Remark B.2). In fact we generate all possible (linearly generated) syzygies that are contained in our computational universe. The consideration of syzygies is the crucial point in the classical Gröbner basis algorithm and finds its natural resemblance in the `LStabSpan` procedure in our setting. We will further show that a border basis or Gröbner basis cannot be used (in reduced form) in a projection framework to obtain the integral hull of 0/1 polytopes.

All computations in this section were performed using `CoCoA 4.7.5` ([7]) and `polymake 2.9.7` ([15]).

Example B.1. Consider the polytope $P \subseteq [0,1]^4$

$$P := \{x \in [0,1]^4 \mid x_1 + 2x_2 + 3x_3 + 5x_4 = 4\}$$

with $P_I = \{(1,0,1,0)\}$. The first Sherali-Adams closure $\mathrm{AS}^{[1]}(P)$ is given by the projection of the following 9 polynomials:

$$x_1x_2 + 3/2x_1x_3 + 5/2x_1x_4 + 3x_2 + 9/2x_3 + 15/2x_4 - 6$$

$$x_1x_3 - 2x_2x_3 + 5/3x_1x_4 - 10/3x_2x_4 + 10/3x_2 + 3x_3 + 5x_4 - 4$$

$$x_2x_3 - 5/12x_1x_4 + 5/6x_2x_4 + 5/4x_3x_4 - 5/6x_2 - x_3 - 5/4x_4 + 1$$

$$x_1x_4 + 2x_2x_4 + 3x_3x_4 + x_4$$

and $x_1^2 - x_1, x_4^2 - x_4, x_2^2 - x_2, x_3^2 - x_3, x_1 + 2x_2 + 3x_3 + 5x_4 - 4$, which are the initial polynomials.

We obtain the polytope $M^1(P) = \mathrm{conv}(\{p_1, \dots, p_4\}) \subseteq [0,1]^{10}$ with

$$\begin{aligned}
p_1 &= (0, \frac{5}{9}, \frac{2}{3}, 1, 0, 0, 0, \frac{1}{9}, \frac{1}{3}, 1), \\
p_2 &= (0, \frac{1}{27}, \frac{8}{9}, \frac{2}{3}, 0, 0, 0, \frac{7}{27}, 0, 1), \\
p_3 &= (0, 1, 0, 1, 0, 0, 0, 0, 1, 0), \\
p_4 &= (0, \frac{8}{9}, \frac{2}{3}, 0, 0, 0, 0, \frac{4}{9}, 0, 0).
\end{aligned}$$

with the index set of the vector given by the ordered tuple

$$(x_4, x_3, x_2, x_1, x_3x_4, x_2x_4, x_1x_4, x_2x_3, x_1x_3, x_1x_2).$$

Therefore we have $\mathrm{AS}^{[1]}(P) \neq P_I$. The LStabSpan procedure does not stop after 2 rounds though and generates additional relations

$$x_2x_4 + 3/4x_3x_4 - 1/12x_2 - 1/10x_3 + 1/12x_4 + 1/10,$$

$$x_3x_4 - 1/45x_2 + 2/45x_3 + 1/9x_4 - 2/45,$$

$$x_2 - 1/2x_3 + 5/2x_4 + 1/2,$$

$$x_3 - 35/59x_4 - 1,$$

$$x_4.$$

The projection of this extended system which is the first border basis closure satisfies $\mathrm{BC}^{[1]}(P) = P_I = \{(0,1,0,1)\}$, i.e., $x_3 = x_1 = 1$ and $x_4 = x_2 = 0$ is the only possible solution.

In view of Example B.1 the following observation captures a crucial property of the border basis closure:

Remark B.2. Let F be a finite set of polynomials generating a zero-dimensional ideal and let $d \in [n]$ be arbitrary. Furthermore let $\mathscr{M} \subseteq \{x_j \mid j \in [n]\}$ be the set of variables which do not occur as leading terms of polynomials generated by the Sherali-Adams closure. The additional polynomials generated by the border basis closure have leading terms in $[[\mathscr{M}]]$ and we potentially discover polynomials of low (or even linear) degree. The linear equations obtained can be used as cutting planes without any further computation. We will now try to illustrate this point by revisiting Example B.1. The system in Example B.1 is:

$$F = \{x_1 + 2x_2 + 3x_3 + 5x_4 - 4\} \cup \{x_j^2 - x_j \mid j \in [4]\}.$$

We will analyze the iterations within the LStabSpan$(F, 2)$ procedure. Before the first iteration we have polynomials with leading terms $\{x_1, x_1^2, x_2^2, x_3^2, x_4^2\}$. After the first iteration we obtained the additional leading terms $\{x_1x_2, x_1x_3, x_2x_3, x_1x_4\}$. Note that by now all polynomials in F have been multiplied by one variable in $\{x_j \mid j \in [4]\}$. Something essential happens in the next iteration: We again extend these polynomials and the resulting polynomials would have degree 3. Now the effect of the Gaussian elimination becomes clear: Suppose we have two polynomials p_1, p_2 with leading terms $\mathrm{LT}(p_1) = x_1x_2$ and $\mathrm{LT}(p_2) = x_2x_3$. After having multiplied p_1, p_2 with an additional variables these polynomials have degree 3. For example we obtain x_3p_1 and x_1p_2 in this iteration. Obviously both polynomials have degree 3 and if we would not perform any further operations p_1, p_2 would be removed as their degree exceeds two. But since $\mathrm{LT}(x_3p_1) = x_3\mathrm{LT}(p_1) = x_1x_2x_3 = x_1\mathrm{LT}(p_2) = \mathrm{LT}(x_1p_2)$, the GaussEl procedure replaces x_3p_1 by $x_3p_1 - x_1p_2$ (or vice versa, depending on which is processed first). Note that $\mathrm{LT}(x_3p_1 - x_1p_2) < \mathrm{LT}(x_3p_1)$ and thus we might actually generate new polynomials with smaller leading terms (we sketched this behavior in Figure B.1). This is exactly what happens in our example and what gives the border basis closure additional strength: After the second iteration we obtain new leading terms $\{x_2x_4, x_3x_4\}$ and after the third iteration we obtain polynomials with leading terms $\{x_2, x_3, x_4\}$ which correspond to the polynomials $\{x_2 - 1/2x_3 + 5/2x_4 + 1/2, x_3 - 35/59x_4 - 1, x_4\}$.

As these polynomials are linear equations these can by applied as cuts immediately.

Direct Gröbner bases closures. We will now show that if $\mathscr{T}$ is the Gröbner basis operator, i.e., $\mathscr{T}(F) = \mathscr{G}$ where $\mathscr{G}$ is the Gröbner bases of F then we cannot expect in general that $\mathscr{T}(P) = P_I$.

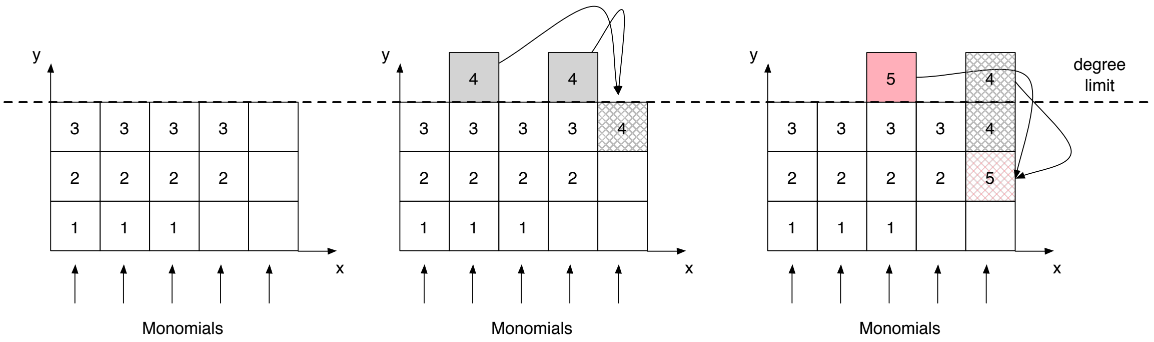

FIGURE B.1. GaussEl generating additional leading terms

Example B.3. Let $P = \{x \in [0,1]^4 \mid x_1 + 2x_2 + x_3 + x_4 = 3\} \subseteq [0,1]^4$. If $\mathscr{G}$ is the Gröbner basis of the associated system F, then $\mathscr{R}(\mathscr{G})$ has the following vertices

$$\begin{aligned} p_1 &= (0,1,1,0), \\ p_2 &= (0,1,0,1), \\ p_3 &= (1,1,0,0), \\ p_4 &= (0,\frac{1}{2},1,1), \\ p_5 &= (1,0,1,1), \end{aligned}$$

and thus is not integral.

This is surprising insofar as the Gröbner basis $\mathscr{G}$ of F contains a lot of information about P. For example $\mathscr{G} = \{1\}$ if and only if $P \cap \mathbb{Z}^n = \emptyset$. Thus we can use the Gröbner basis to decide infeasibility. If we compute the Gröbner basis with respect to a different ordering $\sigma = \texttt{Lex}$ (the lexicographic ordering), it is well-known that the Gröbner basis of F can be used to construct feasible solutions in time linear in the number of elements in $\mathscr{G}$ or to enumerate the set $P \cap \mathbb{Z}^n$ (for a nice summary of these properties see [4]). But even in the case of a Gröbner basis $\mathscr{G}$ with respect to the ordering Lex we cannot expect that $\mathscr{R}(\mathscr{T}(F)) = P_I$ holds as the following example shows:

Example B.4. Let P be as in Example B.3. If $\mathscr{G}$ is the Gröbner basis of the associated system F computed with respect to the term ordering $\sigma = \texttt{Lex}$, then $\mathscr{R}(\mathscr{G})$ has the following vertices

$$\begin{aligned} p_1 &= (1,1,0,0), \\ p_2 &= (1,0,1,1), \\ p_3 &= (1,\frac{1}{2},1,0), \\ p_4 &= (1,\frac{1}{2},0,1), \\ p_5 &= (0,\frac{1}{2},1,1), \\ p_6 &= (0,1,0,1), \\ p_7 &= (0,1,1,0) \end{aligned}$$

and thus $\mathscr{R}(\mathscr{G})$ is not integral. What is interesting to note is that the vertex set depends on the term ordering. Whereas in Example B.3 $\mathscr{R}(\mathscr{G})$ has 5 vertices, it has 7 vertices here.

Direct border bases closures. We will now show that the same problem occurs when the closure is derived from a border basis to which the final reduction algorithm has been applied. Whereas the information obtained after the final reduction algorithm is still sufficient in the ring theoretic setting where multiplication of variables is allowed and the full border basis information is still available, we lose information in the linear setting that we consider here. The missing information due to the reduction cannot be reconstructed when multiplication of variables is not permitted. This observation explains why it is crucial to skip the final reduction algorithm.

Example B.5. Let $P = \{x \in [0,1]^5 \mid x_1 + x_2 + 3x_3 + x_4 + 2x_5 = 5\} \subseteq [0,1]^5$. If $\mathscr{G}$ is the border basis of the associated system F, then $\mathscr{R}(\mathscr{G})$ has the following vertices

$$\begin{aligned}
p_1 &= (1,1,1,0,0),\\
p_2 &= (1,1,0,1,1),\\
p_3 &= (1,0,1,1,0),\\
p_4 &= (0,0,1,0,1),\\
p_5 &= (\frac{1}{3},\frac{1}{3},\frac{2}{3},1,\frac{2}{3}),\\
p_6 &= (1,\frac{1}{3},\frac{2}{3},\frac{1}{3},\frac{2}{3}),\\
p_7 &= (0,1,1,1,0),\\
p_8 &= (\frac{1}{3},1,\frac{2}{3},\frac{1}{3},\frac{2}{3}),
\end{aligned}$$

and thus $\mathscr{R}(\mathscr{G})$ is not integral. Thus, although the border bases contains all the necessary information about the integral hull, we cannot expect that the relinearization and projection is equal to the integral hull. In contrast to this, the border basis closure rank of the system is 3, i.e., $\mathrm{BC}^{[3]}(P) = P_I$.

Massachusetts Institute of Technology, USA
E-mail address: pokutta@mit.edu

Massachusetts Institute of Technology, USA
E-mail address: schulz@mit.edu