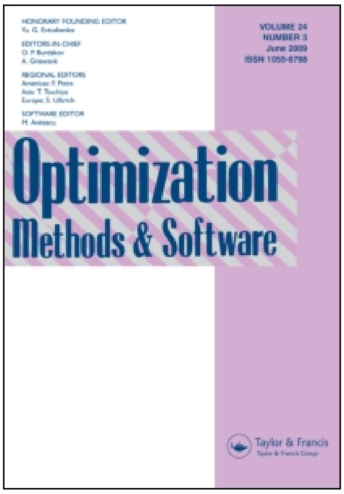

Optimization Methods and Software
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713645924

A p-cone sequential relaxation procedure for 0-1 integer programs

Samuel Burer [a]; Jieqiu Chen [a]
[a] Department of Management Sciences, University of Iowa, Iowa City, IA, USA

Online Publication Date: 01 August 2009

To cite this Article Burer, Samuel and Chen, Jieqiu(2009)'A p-cone sequential relaxation procedure for 0-1 integer programs',Optimization Methods and Software,24:4,523 — 548

To link to this Article: DOI: 10.1080/10556780903057341

URL: http://dx.doi.org/10.1080/10556780903057341

Optimization Methods & Software
Vol. 24, Nos. 4–5, August–October 2009, 523–548

A p-cone sequential relaxation procedure for 0-1 integer programs

Samuel Burer* and Jieqiu Chen

Department of Management Sciences, University of Iowa, Iowa City, IA 52242-1994, USA

(*Received 11 February 2008; final version received 13 May 2009*)

Several authors have introduced sequential relaxation techniques – based on linear and/or semi-definite programming – to generate the convex hull of 0-1 integer points in a polytope in at most n steps. In this paper, we introduce a sequential relaxation technique, which is based on p-order cone programming ($1 \leq p \leq \infty$). We prove that our technique generates the convex hull of 0-1 solutions asymptotically. In addition, we show that our method generalizes and subsumes several existing methods. For example, when $p = \infty$, our method corresponds to the well-known procedure of Lovász and Schrijver based on linear programming. Although the p-order cone programs in general sacrifice some strength compared to the analogous linear and semi-definite programs, we show that for $p = 2$ they enjoy a better theoretical iteration complexity. Computational considerations of our technique are discussed.

Keywords: global optimization; 0-1 integer programming; second-order cone programming; cone programming; convex relaxation; successive convex relaxations

1. Introduction

Consider solving the 0-1 integer program

$$\begin{aligned} \min \quad & c^{\mathrm{T}} x \\ \text{s.t.} \quad & a_i^{\mathrm{T}} x \leq b_i \quad \forall\, i = 1, \ldots, m \\ & x \in \{0, 1\}^n. \end{aligned} \tag{1}$$

Beyond the basic linear programming (LP) relaxation P of the feasible set of Equation (1), many authors have considered general techniques for achieving tighter relaxations [6,8,12,17,18,20,21,25,26]. One recurring theme is to lift the feasible set of Equation (1) into a higher dimensional space, where a convex relaxation is constructed, and then to project this relaxation back into the original space of variables, thus obtaining a relaxation P^1, which is possibly tighter than P. The choice of lifting and relaxation determines the strength of P^1. In all previous works, LP and semi-definite programming (SDP) relaxations have been used in the

*Corresponding author. Email: samuel-burer@uiowa.edu

ISSN 1055-6788 print/ISSN 1029-4937 online

DOI: 10.1080/10556780903057341
http://www.informaworld.com

higher dimensional space. Kim and Kojima [15] suggest a modification of the SDPs that only enforces positive semi-defintеness on 2×2 principal submatrices, which can be modelled with second-order cone programming (SOCP) [14,16].

In [6,17,18,21], the idea of sequential relaxation is also introduced. Stated simply, the idea is to repeat the lift-and-project procedure on the tighter relaxation P^1, thus obtaining an even tighter relaxation P^2. If P^k denotes the kth relaxation obtained inductively, a fundamental question is whether $\{P^k\}$ converges to P^{01}, the convex hull of the feasible set of Equation (1). In each study referenced above, the answer is positive; in fact, $P^k = P^{01}$ for all $k \geq n$. Computationally, one can typically optimize over P^k in polynomial time as long as k is constant. Kojima and Tunçel [17,18] apply their techniques to a much broader class of quadratically constrained problems and show asymptotic convergence to the convex hull of solutions.

Sherali and Adams [26], Lasserre [20], and Parrilo [25] do not explicitly employ the idea of sequential relaxation. Rather, they lift to ever higher dimensions before projecting a technique that is analogous to sequential relaxation. Here, too, the authors show that lifting to a finite dimension (dependent in some manner on n) achieves P^{01}. Similar to the work of Kojima and Tunçel [17,18], these authors' techniques can be applied to more general problem classes, but one may have to lift 'infinitely' to achieve the convex hull of solutions.

Although these lift-and-project procedures are very powerful theoretically, they present significant computational challenges, even after a single iteration. One must deal with more variables in the higher dimensional space as well as additional constraints introduced by lifting. For example, after one iteration of the LP-based procedure of [6], the resulting LP contains $2n$ variables and $2m + 1$ constraints assuming that the constraints $a_i^{\mathrm{T}} x \leq b_i$ already imply the bounds $0 \leq x_j \leq 1$. The procedure of Lovász and Schrijver [21] requires even more variables and constraints. After one iteration, their LP-based procedure has $\mathcal{O}(n^2)$ variables and $\mathcal{O}(nm)$ linear constraints, and their SDP-based procedure contains an additional semi-definite constraint on an order $n \times n$ matrix. In both of these particular cases, computational progress has been achieved by exploiting the structure of constraints generated by these procedures [5,10].

The purpose of this paper is to explore p-order cone programming (POCP) relaxations, which include SOCP relaxations when $p = 2$. Our interest in POCP arises from the fact that POCP is becoming a well-understood tool in convex optimization [3,11,19,29]. Moreover, there are currently several high quality implementations for SOCP [22,27,28], and in fact POCP can be formulated exactly via SOCP [2,7,19]. Our hope is that, by introducing POCP relaxations, we might discover new lift-and-project procedures that have their own theoretical and computational advantages.

In this paper, we introduce a family of lift-and-project procedures parameterized by $p \in [1, \infty]$ and prove that each asymptotically yields P^{01}, the convex hull of 0-1 solutions (Theorem 4.1). A feature of this family of procedures is the ability to lift and project with respect to different subsets of variables at different iterations. Although we do not achieve finite convergence in general so that our procedure is weaker than existing methods in this sense, we do observe theoretical advantages. In particular, the theoretical iteration complexity of solving the POCP relaxations via interior-point methods is minimized when $p = 2$ at which the iteration complexity is an order of magnitude less than solving existing LP and SDP relaxations (Corollary 3.6). In addition, our family of procedures unifies existing approaches. For example, when $p = \infty$, we recover the LP-based procedure of Lovász and Schrijver [21].

In Section 2, we describe the basic p-order cone lift-and-project procedure and give an alternate derivation of it. This is just one iteration of the entire sequential relaxation approach, which is described in Section 4. We then compare and contrast our procedure with three existing approaches: [6,17,18,21]. In particular, we point out that our method includes the LP-based lift-and-project procedure of Lovász and Schrijver [21] and the relaxation of Balas *et al.* [6] as special cases.

In Section 3, we study fundamental properties of the p-order cone procedure. In particular, we examine duality properties and two types of monotonicity. One monotonicity property establishes that the strength of the procedure increases with p, so that the strength is maximized at $p = \infty$. The other monotonicity property says that after applying the p-cone procedure to two compact and convex sets $Q \subseteq P$, the obtained set for Q is contained in that for P. We also study the iteration complexity of solving the resultant p-order cone relaxation via interior-point methods, where it is shown that the lowest iteration complexity is obtained for $p = 2$. Following these results is the main technical result of the paper (Theorem 3.15): the p-order cone procedure, when applied to a generic compact, convex set P, cuts off all fractional extreme points of P. Theorem 3.15 is motivated and proven in three steps. We first show the result holds when P is a polytope, which is the easiest case. Then we establish the result when P is a ball, which finally allows us to prove the theorem for general P.

In Section 4, we describe the sequential relaxation approach based on repetition of the p-order cone procedure and prove that it generates the convex hull of 0-1 solutions asymptotically (Theorem 4.1). An explicit example is provided to show that, in general, the iterated procedure may indeed require an infinite number of repetitions to converge.

In Section 5, we consider computational issues associated with the p-cone sequential relaxation procedure – particularly with optimizing over the first-iteration relaxation. We compare the SOCP relaxation ($p = 2$, lowest theoretical complexity) to the LP relaxation of Lovász–Schrijver ($p = \infty$, tightest relaxation). We observe that, even though the SOCP relaxation enjoys a much lower theoretical iteration complexity, the LPs solve more quickly and produce better bounds.

Finally, in Section 6, we conclude with final remarks.

2. Relaxation procedure and comparisons

In this section, we first introduce notation and terminology and then formally describe our p-cone lift-and-project procedure, comparing and contrasting it with the methods of Lovász and Schrijver [21], Kojima and Tunçel [17,18], and Balas *et al.* [6].

2.1 *Notation and terminology*

$\Re^n$ refers to n-dimensional Euclidean space, and $\Re^{n\times n}$ is the set of real, $n \times n$ matrices. We let $e_i \in \Re^n$ represent the ith unit coordinate vector and $e \in \Re^n$ represent the vector of all ones. We denote by $[n]$ the set $\{1, 2, \ldots, n\}$. For $\mathcal{J} \subseteq [n]$ an index set, $x_{\mathcal{J}} \in \Re^{|\mathcal{J}|}$ is defined as the vector composed of entries of x that are indexed by $\mathcal{J}$. Similarly, given a matrix $A \in \Re^{n\times n}$, $A_{\mathcal{J}\cdot}$ represents the $|\mathcal{J}| \times n$ matrix composed of the rows of A indexed by $\mathcal{J}$. $\mathrm{Diag}(x)$ denotes the diagonal matrix with diagonal x, and $A \succeq 0$ means that the matrix A is symmetric positive semidefinite; its dimension will be clear from context. Finally, given a set S defined over variables (x, y), $\mathrm{proj}_x(S)$ denotes the projection of S onto the coordinates in x.

For $p \in [1, \infty]$, the usual p-norm on $\Re^n$ is defined as

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

We also define, when $p = \infty$, $\|x\|_\infty := \max_{i=1}^{n} |x_i|$. Associated with $p \in [1, \infty]$ is $q \in [1, \infty]$ such that $p^{-1} + q^{-1} = 1$. Both the p-norm and q-norm give rise to closed, convex cones in $\Re^{1+n}$:

$$\mathcal{K}_p := \{(x_0, x) \in \Re^{1+n} : x_0 \geq \|x\|_p\}$$

$$\mathcal{K}_q := \{(y_0, y) \in \Re^{1+n} : y_0 \geq \|y\|_q\}.$$

It is well known that $\mathcal{K}_q$ is dual to $\mathcal{K}_p$, i.e.

$$\mathcal{K}_q = \{(y_0, y) \in \Re^{1+n} : y_0 x_0 + y^{\mathrm{T}} x \geq 0 \quad \forall\, (x_0, x) \in \mathcal{K}_p\},$$

which is written as $K_q = K_p^*$; see for example, [9, p. 51]. Important special cases occur when $p = 2$ or $p = \infty$. When $p = 2$, $K_p = K_q$, i.e., K_p and K_q are self-dual. When $p = \infty$, $q = 1$, and both K_p and K_q are polyhedral cones.

2.2 *Relaxation procedure*

For the purposes of generality, particularly, with regards to Section 4, we consider a slightly different form of the feasible set of the integer program (1), the only difference being that the linear constraints are indexed by an arbitrary set $\mathcal{I}$ (possibly infinite) as follows:

$$F := \{x \in \{0, 1\}^n : a_i^{\mathrm{T}} x \leq b_i \quad \forall\, i \in \mathcal{I}\}.$$

This semi-infinite representation for F, as opposed to a finite one, does not affect the theoretical exposition of the p-cone procedure, though it may pose computational issues. Associated with F is its continuous convex relaxation

$$P := \{x \in \Re^n : a_i^{\mathrm{T}} x \leq b_i \quad \forall\, i \in \mathcal{I}\},$$

obtained by relaxing the integrality requirements on x. We assume P is contained in $[0, 1]^n$, for example, by including bounds on x via explicit constraints $a_i^{\mathrm{T}} x \leq b_i$. So P is compact convex.

We wish to generate a compact convex relaxation of F, which is tighter than P. Unless stated otherwise, we assume throughout this section the fixed choice of $p \in [1, \infty]$ and $\emptyset \neq \mathcal{J} \subseteq [n]$. We will denote the proposed convex relaxation as $N_{(p,\mathcal{J})}(P)$, or more often simply as $N(P)$. Defining

$$P^{01} := \mathrm{conv}(F),$$

our goal is to produce $N(P)$ such that $P^{01} \subseteq N(P) \subseteq P$.

Our first step is to lift F into a higher dimensional space. We will make use of the following simple key geometric proposition.

PROPOSITION 2.1 *Define* $r := \sqrt[p]{|\mathcal{J}|}/2$ *and* $d := e/2 \in \Re^{|\mathcal{J}|}$. *Then* $x_{\mathcal{J}} \in \{0, 1\}^{|\mathcal{J}|}$ *implies* $\|x_{\mathcal{J}} - d\|_p \leq r$.

It is important to keep in mind that r depends on p and $|\mathcal{J}|$ and that d depends on $|\mathcal{J}|$. This proposition establishes the existence of a family of p-balls circumscribing the integer points $\{0, 1\}^{|\mathcal{J}|}$. In fact, $p' \geq p$ implies that the p'-ball is contained in the p-ball (see Proposition 3.9), with $p = \infty$ corresponding to the convex hull $[0, 1]^{|\mathcal{J}|}$ of the integer points.

Using Proposition 2.1, F can be rewritten redundantly as

$$F = \{x \in \Re^n : x = x^2,\ a_i^{\mathrm{T}} x \leq b_i \quad \forall\, i \in \mathcal{I},\ \|x_{\mathcal{J}} - d\|_p \leq r\}.$$

We note that

$$\left.\begin{array}{r} a_i^{\mathrm{T}} x \leq b_i \\ \|x_{\mathcal{J}} - d\|_p \leq r \end{array}\right\} \Longrightarrow \|(b_i - a_i^{\mathrm{T}} x)(x_{\mathcal{J}} - d)\|_p \leq r(b_i - a_i^{\mathrm{T}} x), \tag{2}$$

which in turn implies

$$F = \{x \in \Re^n : x = x^2,\ \|b_i x_{\mathcal{J}} - x_{\mathcal{J}} x^{\mathrm{T}} a_i - (b_i - a_i^{\mathrm{T}} x) d\|_p \leq r(b_i - a_i^{\mathrm{T}} x) \quad \forall\, i \in \mathcal{I}\}$$

since $b_i - a_i^{\mathrm{T}}x$ is kept non-negative. Next, introducing an $n \times n$ matrix variable X satisfying $X = xx^{\mathrm{T}}$ and defining

$$\hat{F} := \left\{ (x, X) \in \Re^n \times \Re^{n \times n} : \begin{array}{c} X = xx^{\mathrm{T}} \quad \mathrm{diag}(X) = x \\ \|b_i x_{\mathcal{J}} - X_{\mathcal{J}\cdot} a_i - (b_i - a_i^{\mathrm{T}}x)d\|_p \le r(b_i - a_i^{\mathrm{T}}x) \quad \forall\, i \in \mathcal{I} \end{array} \right\}$$

we see that $F = \mathrm{proj}_x(\hat{F})$, i.e., $\hat{F}$ is the lifted version of F. In addition, dropping the non-convex constraint $X = xx^{\mathrm{T}}$ from $\hat{F}$, we obtain a convex relaxation of $\hat{F}$:

$$\hat{P} := \left\{ (x, X) \in \Re^n \times \Re^{n \times n} : \begin{array}{c} \mathrm{diag}(X) = x \\ \|b_i x_{\mathcal{J}} - X_{\mathcal{J}\cdot} a_i - (b_i - a_i^{\mathrm{T}}x)d\|_p \le r(b_i - a_i^{\mathrm{T}}x) \quad \forall\, i \in \mathcal{I} \end{array} \right\}.$$

Finally, we define $N(P)$ as the projection of $\hat{P}$:

$$N(P) := \mathrm{proj}_x(\hat{P}).$$

The desired property of $N(P)$ is immediate.

Proposition 2.2 $P^{01} \subseteq N(P) \subseteq P$.

Proof $P^{01} \subseteq N(P)$ by construction. Moreover, the definition of $\hat{P}$ implies that every $x \in N(P)$ satisfies $r(b_i - a_i^{\mathrm{T}}x) \ge 0$ for all $i \in \mathcal{I}$. Since $r > 0$, this implies $x \in P$. So $N(P) \subseteq P$. ■

Using the assumption that the constraints $a_i^{\mathrm{T}}x \le b_i$ include the bounds $0 \le x_j \le 1$ explicitly and the fact that $N(P) \subseteq P$ is bounded, one can see from the definition of $\hat{P}$ that $\mathrm{proj}_{(x, X_{\mathcal{J}\cdot})}(\hat{P})$ is bounded and hence compact convex. Since the rows $X_{j\cdot}$ for $j \notin \mathcal{J}$ are not constrained in $\hat{P}$ except for the entries X_{jj}, it is also clear that

$$N(P) = \mathrm{proj}_x(\mathrm{proj}_{(x, X_{\mathcal{J}\cdot})}(\hat{P})).$$

Since the projection of compact convex sets are compact convex, we conclude that $N(P)$ is compact convex.

Just like P, $N(P)$ has its own semi-infinite outer description, which can be the basis of lift-and-project applied to $N(P)$ itself. This will be the main idea behind the iterated procedure of Section 4.

As an example, consider the feasible set

$$F = \left\{ x \in \{0, 1\}^2 : \begin{array}{c} -x_1 \le 0 \\ -x_2 \le 0 \\ x_1 + 2x_2 \le 2.5 \\ 3x_1 + x_2 \le 2.5 \end{array} \right\} = \{(0, 0), (0, 1)\} \tag{3}$$

so that

$$P^{01} = \{(0, x_2) \in \Re^2 : 0 \le x_2 \le 1\}.$$

Let $\mathcal{J} = \{1, 2\}$. In Figure 1, we illustrate the p-cone procedure by depicting the four sets

$$P \supseteq N_{(1,\mathcal{J})}(P) \supseteq N_{(2,\mathcal{J})}(P) \supseteq N_{(\infty,\mathcal{J})}(P) = P^{01}$$

containing P^{01}. Figure 1 was drawn by determining a collection of points on the boundary of each set via a collection of linesearch procedures. For each of the four sets, a single linesearch started at (0, 0) and moved into the first quadrant at an angle $\theta \in [0, \pi/2]$. The point on the boundary was

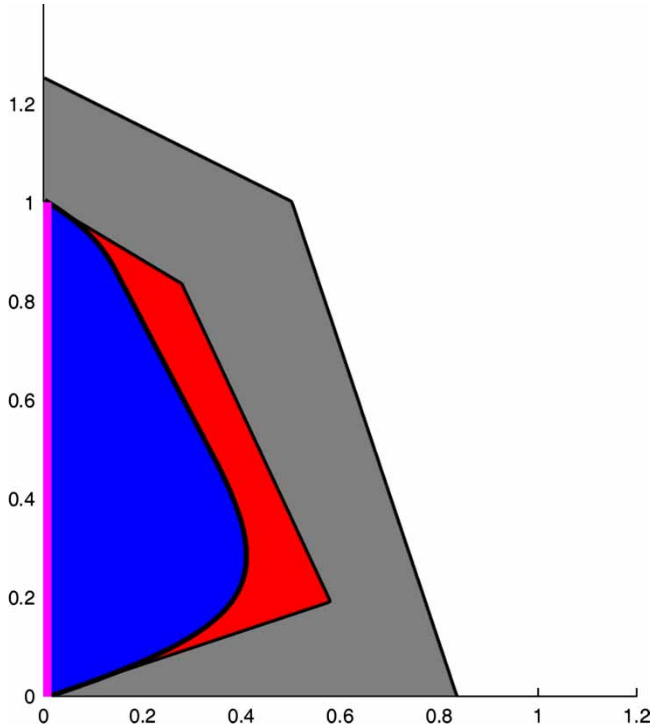

Figure 1. The four sets $P \supseteq N_{(1,\mathcal{J})}(P) \supseteq N_{(2,\mathcal{J})}(P) \supseteq N_{(\infty,\mathcal{J})}(P)$ relative to the example feasible set F in Equation (3), where $\mathcal{J} = \{1, 2\}$. Note that $N_{(\infty,\mathcal{J})}(P) = P^{01}$ in this example. Available in colour online.

precisely the point where the linesearch left the set. The linesearch was repeated for a sufficiently fine grid on $[0, \pi/2]$ for each of the four sets.

Recall that P is the continuous LP relaxation of F; in the figure, it is the largest set. The next largest is $N_{(1,\mathcal{J})}$, a polyhedral set since $p = 1$. $N_{(2,\mathcal{J})}$ is the projection of a second-order cone set and hence has a curved boundary. Finally, the depicted line segment between $(0, 0)$ and $(0, 1)$ is $N_{(\infty,\mathcal{J})}$, which equals P^{01} in this example.

2.3 *A different derivation*

In the derivation of $N(P)$, we have relied on the implication (2), which can be thought of as replacing two constraints by their product respecting non-negativity. We now show that one can obtain an alternate representation of the right-hand side of Equation (2) via an alternate representation of $\|x_{\mathcal{J}} - d\|_p \le r$. Multiplication of constraints is still the key idea.

Consider $\mathcal{K}_p$ and $\mathcal{K}_q$ as described in Section 2.1. Because $\mathcal{K}_q = \mathcal{K}_p^*$, it holds that

$$\begin{aligned} \|x_{\mathcal{J}} - d\|_p \le r &\iff (r, x_{\mathcal{J}} - d) \in \mathcal{K}_p \\ &\iff vr + u^{\mathrm{T}}(x_{\mathcal{J}} - d) \ge 0 \quad \forall\, (v, u) \in \mathcal{K}_q. \end{aligned} \tag{4}$$

PROPOSITION 2.3 *For a given $x \in \Re^n$, the right-hand side of Equation* (2) *holds if and only if*

$$(b_i - a_i^{\mathrm{T}} x)(vr + u^{\mathrm{T}}(x_{\mathcal{J}} - d)) \ge 0 \quad \forall\, (v, u) \in \mathcal{K}_q. \tag{5}$$

Proof ($\Rightarrow$): The right-hand side of Equation (2) implies $b_i - a_i^{\mathrm{T}} x \ge 0$. If $b_i - a_i^{\mathrm{T}} x = 0$, then clearly Equation (5) holds. On the other hand, if $b_i - a_i^{\mathrm{T}} x > 0$, then dividing the right-hand side of Equation (2) by $b_i - a_i^{\mathrm{T}}x$ shows $\|x_{\mathcal{J}} - d\|_q \le r$, which in turn implies $vr + u^{\mathrm{T}}(x_{\mathcal{J}} - d) \ge 0$ for all $(v, u) \in \mathcal{K}_q$ by Equation (4). Now multiplying with $b_i - a_i^{\mathrm{T}}x$ implies Equation (5).

($\Leftarrow$): If $b_i - a_i^{\mathrm{T}}x = 0$, then the right-hand side of Equation (2) holds trivially. On the other hand, if $b_i - a_i^{\mathrm{T}}x \neq 0$, then for any non-zero u the two inequalities

$$(b_i - a_i^{\mathrm{T}}x)(\|u\|_q r + u^{\mathrm{T}}(x_{\mathcal{J}} - d)) \ge 0$$
$$(b_i - a_i^{\mathrm{T}}x)(\|u\|_q r - u^{\mathrm{T}}(x_{\mathcal{J}} - d)) \ge 0$$

together imply $b_i - a_i^{\mathrm{T}}x > 0$. As a consequence, $vr + u^{\mathrm{T}}(x_{\mathcal{J}} - d) \ge 0$ for all $(v, u) \in \mathcal{K}_q$, which means $\|x_{\mathcal{J}} - d\|_q \le r$ by Equation (4), which in turn implies the right-hand side of Equation (2). ■

An immediate consequence of Proposition 2.3 is that $\hat{P}$ defined in the derivation of $N(P)$ can be equivalently expressed using the semi-infinite collection of inequalities

$$(b_i - a_i^{\mathrm{T}}x)(vr - u^{\mathrm{T}}d) + b_i u^{\mathrm{T}}x_{\mathcal{J}} - u^{\mathrm{T}}X_{\mathcal{J}\cdot}a_i \ge 0 \quad \forall\, (i, (v, u)) \in \mathcal{I} \times \mathcal{K}_q,$$

thus providing an equivalent definition of $N(P)$.

2.4 *Comparison with existing approaches*

In the derivation of $N(P)$, we did not use the full strength of the relationship $X = xx^{\mathrm{T}}$ in the relaxation $\hat{P}$. In particular, we could have also imposed in $\hat{P}$ the following two convex conditions, which are implied by $X = xx^{\mathrm{T}}$:

$$X = X^{\mathrm{T}} \quad \text{and} \quad \begin{pmatrix} 1 & x^{\mathrm{T}} \\ x & X \end{pmatrix} \succeq 0. \tag{6}$$

If we had imposed these conditions, then $N(P)$ would be even tighter. However, we have purposely not imposed them because they are not necessary for the theoretical convergence of the iterated procedure in Section 4. In practice, one would certainly want to impose as many constraints that can be handled efficiently. In particular, imposing symmetry $X = X^{\mathrm{T}}$ can be useful to eliminate variables.

We mention conditions (6) here because they facilitate comparison with existing lift-and-project methods in the following subsections.

2.4.1 *Lovász–Schrijver*

The LP-based approach of Lovász and Schrijver [21] is derived like the p-cone procedure except that the following lifted and relaxed sets serve in the place of our $\hat{F}$ and $\hat{P}$ as follows:

$$\hat{F}_{\mathrm{LS}} = \left\{ (x, X) \in \Re^n \times \Re^{n\times n} : \begin{array}{c} X = xx^{\mathrm{T}} \quad \mathrm{diag}(X) = x \\ (b_i - a_i^{\mathrm{T}}x)x_k \ge 0 \quad \forall\, (i, k) \in \mathcal{I} \times [n] \\ (b_i - a_i^{\mathrm{T}}x)(1 - x_k) \ge 0 \quad \forall\, (i, k) \in \mathcal{I} \times [n] \end{array} \right\}$$
$$\hat{P}_{\mathrm{LS}} = \left\{ (x, X) \in \Re^n \times \Re^{n\times n} : \begin{array}{c} X = X^{\mathrm{T}} \quad \mathrm{diag}(X) = x \\ b_i x_k - X_{k\cdot}a_i \ge 0 \quad \forall\, (i, k) \in \mathcal{I} \times [n] \\ (b_i - a_i^{\mathrm{T}}x) - (b_i x_k - X_{k\cdot}a_i) \ge 0 \quad \forall\, (i, k) \in \mathcal{I} \times [n] \end{array} \right\}.$$

We next establish relations between $\hat{F}_{\mathrm{LS}}$ and $\hat{F}$, $\hat{P}_{\mathrm{LS}}$, and $\hat{P}$.

PROPOSITION 2.4 *Let $p = \infty$ and $\mathcal{J} = [n]$. If the p-cone lift-and-project procedure also enforces the symmetry condition of Equation* (6)*, then $\hat{F} = \hat{F}_{\mathrm{LS}}$ and $\hat{P} = \hat{P}_{\mathrm{LS}}$.*

Proof Note that $r = 1/2$ with $p = \infty$ and $\mathcal{J} = [n]$. It suffices to show that the conditions

$$\|b_i x - xx^{\mathrm{T}} a_i - (b_i - a_i^{\mathrm{T}} x) d\|_p \le r(b_i - a_i^{\mathrm{T}} x) \quad \forall\, i \in \mathcal{I}$$

of $\hat{F}$ are equivalent to the conditions

$$\begin{aligned} (b_i - a_i^{\mathrm{T}} x) x_k &\ge 0 \quad \forall\, (i,k) \in \mathcal{I} \times [n] \\ (b_i - a_i^{\mathrm{T}} x)(1 - x_k) &\ge 0 \quad \forall\, (i,k) \in \mathcal{I} \times [n] \end{aligned} \tag{7}$$

of $\hat{F}_{\mathrm{LS}}$. By Proposition 2.3, the conditions of $\hat{F}$ can be replaced by

$$(b_i - a_i^{\mathrm{T}} x)(vr + u^{\mathrm{T}}(x - d)) \ge 0 \quad \forall\, (i, (v, u)) \in \mathcal{I} \times \mathcal{K}_1.$$

Since $\mathcal{K}_1$ is finitely generated by $\{(1, \pm e_1), \ldots, (1, \pm e_n)\}$, we obtain

$$\begin{aligned} (b_i - a_i^{\mathrm{T}} x)(r + e_k^{\mathrm{T}}(x - d)) &\ge 0 \quad \forall\, (i,k) \in \mathcal{I} \times [n] \\ (b_i - a_i^{\mathrm{T}} x)(r - e_k^{\mathrm{T}}(x - d)) &\ge 0 \quad \forall\, (i,k) \in \mathcal{I} \times [n], \end{aligned}$$

which reduce to Equation (7), as desired. ■

The next theorem follows directly.

THEOREM 2.5 *Let $p = \infty$ and $\mathcal{J} = [n]$, and suppose the p-cone lift-and-project procedure enforces the symmetry condition of Equation* (6)*. Then the p-cone procedure reduces to the LP-based Lovász–Schrijver lift-and-project procedure.*

Lovász and Schrijver [21] also proposed an SDP-based procedure that enforces the semi-definiteness condition of Equation (6) in $\hat{P}_{\mathrm{LS}}$. Just as the p-cone procedure with symmetry replicates the LP-based Lovász–Schrijver procedure, the p-cone procedure with Equation (6) replicates the SDP-based Lovász–Schrijver procedure.

THEOREM 2.6 *Let $p = \infty$ and $\mathcal{J} = [n]$, and suppose the p-cone lift-and-project procedure enforces the symmetry and semi-definiteness conditions of Equation* (6)*. Then the p-cone procedure reduces to the SDP-based Lovász–Schrijver lift-and-project procedure.*

2.4.2 *Kojima–Tunçel*

Kojima and Tunçel [17,18] present their method as a direct extension of the approach of Lovász and Schrijver [21] to general quadratic optimization problems. Their approach essentially reduces to that of Lovász–Schrijver in the case of 0-1 integer programming – with one important difference, which we explain next. This difference, in particular, has relevance to our discussion and proofs in Section 4.

As discussed in the previous subsection, the Lovász–Schrijver approach is based on lifting with respect to the collection of constraints

$$\begin{aligned}(b_i - a_i^{\mathrm{T}} x)x_k &\geq 0 \quad \forall\,(i,k) \in \mathcal{I} \times [n] \\ (b_i - a_i^{\mathrm{T}} x)(1 - x_k) &\geq 0 \quad \forall\,(i,k) \in \mathcal{I} \times [n].\end{aligned}$$

In contrast, Kojima and Tunçel [17] (see Section 6, p. 767) lift with respect to the larger, extended collection

$$\begin{aligned}(b_i - a_i^{\mathrm{T}} x)x_k &\geq 0 \quad \forall\,(i,k) \in \mathcal{I} \times [n] \\ (b_i - a_i^{\mathrm{T}} x)(1 - x_k) &\geq 0 \quad \forall\,(i,k) \in \mathcal{I} \times [n] \\ (b_i - a_i^{\mathrm{T}} x)(b_h - a_h^{\mathrm{T}} x) &\geq 0 \quad \forall\,(i,h) \in \mathcal{I} \times \mathcal{I} \\ x_j x_k &\geq 0 \quad \forall\,(j,k) \in [n] \times [n] \\ x_j(1 - x_k) &\geq 0 \quad \forall\,(j,k) \in [n] \times [n] \\ (1 - x_j)(1 - x_k) &\geq 0 \quad \forall\,(j,k) \in [n] \times [n],\end{aligned}$$

which reduces to lifting with respect to

$$(b_i - a_i^{\mathrm{T}} x)(b_h - a_h^{\mathrm{T}} x) \geq 0 \quad \forall\,(i,h) \in \mathcal{I} \times \mathcal{I},$$

since P implies the constraints $0 \leq x_k \leq 1$ by assumption. For more insight on this point, refer to Section 3.3 for a discussion on the monotonicity properties of lift-and-project procedures.

This broader lifting guarantees the Kojima–Tunçel approach is at least as strong as the Lovász–Schrijver approach and at least as strong as our approach for $p = \infty$ and $\mathcal{J} = [n]$.

We remark that Lovász and Schrijver [21] did consider the broader lifting of Kojima and Tunçel [17,18] but chose not to focus on it for algorithmic reasons. This point is explained in detail by Kojima and Tunçel [17,18].

2.4.3 *Balas–Ceria–Cornuéjols*

The approach of Balas *et al.* [6] can also be viewed as a special case of our approach. The authors choose a single index j and then apply the lift-and-project procedure outlined above in Section 2.2, replacing $\hat{F}$ and $\hat{P}$ by the following:

$$\hat{F}_{\mathrm{BCC}} = \left\{ (x,X) \in \Re^n \times \Re^{n\times n} : \begin{array}{c} X = xx^{\mathrm{T}} \quad \mathrm{diag}(X) = x \\ (b_i - a_i^{\mathrm{T}} x)x_j \geq 0 \quad \forall\, i \in \mathcal{I} \\ (b_i - a_i^{\mathrm{T}} x)(1 - x_j) \geq 0 \quad \forall\, i \in \mathcal{I} \end{array} \right\}$$

$$\hat{P}_{\mathrm{BCC}} = \left\{ (x,X) \in \Re^n \times \Re^{n\times n} : \begin{array}{c} \mathrm{diag}(X) = x \\ b_i x_j - X_{j\cdot} a_i \geq 0 \quad \forall\, i \in \mathcal{I} \\ (b_i - a_i^{\mathrm{T}} x) - (b_i x_j - X_{j\cdot} a_i) \geq 0 \quad \forall\, i \in \mathcal{I} \end{array} \right\}.$$

We point out two important details. First, $\hat{P}_{\mathrm{BCC}}$ does not enforce the symmetry condition $X = X^{\mathrm{T}}$ of Equation (6). Second, because all rows $X_{k\cdot}$ for $k \neq j$ are unconstrained except for the equation $X_{kk} = x_k$, $\hat{P}_{\mathrm{BCC}}$ may be reduced to

$$\hat{P}_{\mathrm{BCC}} = \left\{ (x,y) \in \Re^n \times \Re^n : \begin{array}{c} y_j = x_j \\ b_i x_j - a_i^{\mathrm{T}} y \geq 0 \quad \forall\, i \in \mathcal{I} \\ (b_i - a_i^{\mathrm{T}} x) - (b_i x_j - a_i^{\mathrm{T}} y) \geq 0 \quad \forall\, i \in \mathcal{I} \end{array} \right\}$$

without affecting the projection onto x.

We claim that the Balas–Ceria–Cornuéjols approach is a special case of our method with $\mathcal{J} = \{j\}$ and arbitrary p.

PROPOSITION 2.7 *Let $p \in [1, \infty]$ and $\mathcal{J} = \{j\}$ for some fixed index j. It holds that $\hat{F}_{\rm BCC} = \hat{F}$ and $\hat{P}_{\rm BCC} = \hat{P}$.*

Proof It suffices to show that the conditions

$$\|b_i x_j - x_j x^{\rm T} a_i - (b_i - a_i^{\rm T} x) d\|_p \le r(b_i - a_i^{\rm T} x) \quad \forall\, i \in \mathcal{I} \tag{8}$$

of $\hat{F}$ are equivalent to the conditions

$$(b_i - a_i^{\rm T} x) x_j \ge 0 \quad \forall i \in \mathcal{I}$$
$$(b_i - a_i^{\rm T} x)(1 - x_j) \ge 0 \quad \forall i \in \mathcal{I}$$

of $\hat{F}_{\rm BCC}$. Noting that $r = 1/2$ and $d = 1/2$ and that the p-norm is applied to a scalar in this case, Equation (8) can be rewritten as

$$\left|(b_i - a_i^{\rm T} x)\left(x_j - \frac{1}{2}\right)\right| \le \frac{1}{2}(b_i - a_i^{\rm T} x) \quad \forall\, i \in \mathcal{I},$$

which is clearly equivalent to the conditions of $\hat{F}_{\rm BCC}$. ■

We thus have the following theorem.

THEOREM 2.8 *Suppose $p \in [1, \infty]$ and $\mathcal{J} = \{j\}$ for some fixed index j. Then the p-cone lift-and-project procedure reduces to the LP-based Balas–Ceria–Cornuéjols lift-and-project procedure.*

3. Duality, complexity, monotonicity, and fractional extreme points

In this section, we examine fundamental properties of the p-cone lift-and-project procedure outlined in Section 2.2. The first main result, proved in Section 3.2, establishes the theoretical iteration complexity of optimizing over $N_{(p,\mathcal{J})}(P)$. The second main result, proved below in Section 3.4.3, is that $N(P)$ contains no extreme points of P having fractional entries in $\mathcal{J}$, a result which will prove critical in Section 4.

Unless stated otherwise, we assume throughout this section that the pair $(p, \mathcal{J})$ is fixed, and we use the short notation $N(P)$ in place of $N_{(p,\mathcal{J})}(P)$. We also assume throughout that $\mathcal{J} = \{1, \ldots, |\mathcal{J}|\}$, i.e. $\mathcal{J}$ specifies the first $|\mathcal{J}|$ variables in x; this is for notational simplicity only.

3.1 *Duality*

Consider the relaxation $\min\{c^{\rm T} x : x \in N(P)\}$ of the 0-1 integer program $\min\{c^{\rm T} x : x \in F\}$. Its explicit p-cone representation is

$$\begin{aligned} \min \quad & c^{\rm T} x \\ \text{s.t.} \quad & \operatorname{diag}(X) = x \\ & (r(b_i - a_i^{\rm T} x), b_i x_{\mathcal{J}} - X_{\mathcal{J}\cdot} a_i - (b_i - a_i^{\rm T} x) d) \in \mathcal{K}_p \quad \forall\, i \in \mathcal{I}, \end{aligned} \tag{9}$$

where $(x, X) \in \Re^n \times \Re^{n\times n}$. The associated dual is

$$\begin{aligned}
\max \quad & \sum_{i\in\mathcal{I}} b_i(d^{\mathrm{T}}u^i - rv_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{I}} \left((d^{\mathrm{T}}u^i - rv_i)a_i + b_i \begin{pmatrix} u^i \\ 0 \end{pmatrix} \right) + \lambda = c \\
& \sum_{i\in\mathcal{I}} \begin{pmatrix} u^i \\ 0 \end{pmatrix} a_i^{\mathrm{T}} + \mathrm{Diag}(\lambda) = 0 \\
& (v_i, u^i) \in \mathcal{K}_q \quad \forall\, i \in \mathcal{I},
\end{aligned} \tag{10}$$

where the dual variables are $\lambda \in \Re^n$ and $(v_i, u^i) \in \Re^{1+|\mathcal{J}|}$ for all $i \in \mathcal{I}$. To illustrate the dual derivation without going deep into the details, we prove weak duality between Equations (9) and (10) in the following proposition.

PROPOSITION 3.1 *Suppose $|\mathcal{I}| < \infty$, i.e. P is a polytope. The dual of the p-cone relaxation* (9) *is the q-cone optimization* (10). *In particular, weak duality holds. If, in addition, both Equations* (9) *and* (10) *have interior feasible solutions, then strong duality holds.*

Proof We prove weak duality to illustrate the dual nature of Equations (9) and (10). The strong duality result is standard [9]. Let (x, X) be feasible for Equation (9) and let $(\lambda, (v_i, u^i))$ be feasible for Equation (10). Also, let $s_i := b_i - a_i^{\mathrm{T}}x$. Then

$$\begin{aligned}
c^{\mathrm{T}}x - \sum_{i\in\mathcal{I}} b_i(d^{\mathrm{T}}u^i - rv_i) &= \left(\sum_{i\in\mathcal{I}} \left((d^{\mathrm{T}}u^i - rv_i)a_i + b_i \begin{pmatrix} u^i \\ 0 \end{pmatrix} \right) + \lambda \right)^{\mathrm{T}} x - \sum_{i\in\mathcal{I}} b_i(d^{\mathrm{T}}u^i - rv_i) \\
&= \sum_{i\in\mathcal{I}} (rv_i - d^{\mathrm{T}}u^i)s_i + \sum_{i\in\mathcal{I}} b_i\, x^{\mathrm{T}} \begin{pmatrix} u^i \\ 0 \end{pmatrix} + \lambda^{\mathrm{T}}x \\
&= \sum_{i\in\mathcal{I}} (rs_iv_i + [b_ix_{\mathcal{J}} - s_id]^{\mathrm{T}}u^i) + \lambda^{\mathrm{T}}x \\
&= \sum_{i\in\mathcal{I}} (rs_iv_i + [b_ix_{\mathcal{J}} - s_id]^{\mathrm{T}}u^i) - \left(\sum_{i\in\mathcal{I}} \begin{pmatrix} u^i \\ 0 \end{pmatrix} a_i^{\mathrm{T}} \right) \bullet X \\
&= \sum_{i\in\mathcal{I}} (rs_iv_i + [b_ix_{\mathcal{J}} - s_id]^{\mathrm{T}}u^i) - \left(\sum_{i\in\mathcal{I}} u^i a_i^{\mathrm{T}} \right) \bullet X_{\mathcal{J}\cdot} \\
&= \sum_{i\in\mathcal{I}} (rs_iv_i + [b_ix_{\mathcal{J}} - s_id - X_{\mathcal{J}\cdot}a_i]^{\mathrm{T}}u^i) \\
&\geq \sum_{i\in\mathcal{I}} 0 = 0.
\end{aligned}$$

■

Related to the primal and dual problems (9) and (10), we consider the following question and derive a duality result: given $\bar{x} \in P \subseteq [0, 1]^n$, is $\bar{x} \in N(P)$? To answer this question, we must determine whether or not the set

$$\left\{X \in \Re^{n\times n} : \begin{array}{c} \mathrm{diag}(X) = \bar{x} \\ (r\bar{s}_i, b_i\bar{x}_{\mathcal{J}} - X_{\mathcal{J}\cdot}a_i - \bar{s}_i d) \in \mathcal{K}_p \quad \forall\, i \in \mathcal{I} \end{array}\right\} \tag{11}$$

is empty, where $\bar{s}_i := b_i - a_i^{\mathrm{T}}\bar{x}$. This question is in turn related to the following set by Proposition 3.2 below:

$$\left\{(\lambda, (v_i, u^i)) \in \Re^n \times \mathcal{K}_q^{|\mathcal{I}|} : \begin{array}{c} \sum_{i\in\mathcal{I}} \begin{pmatrix} u^i \\ 0 \end{pmatrix} a_i^{\mathrm{T}} + \mathrm{Diag}(\lambda) = 0 \\ \bar{x}^{\mathrm{T}}\lambda + \sum_{i\in\mathcal{I}}(r\bar{s}_i\, v_i + (b_i\bar{x}_{\mathcal{J}} - \bar{s}_i d)^{\mathrm{T}} u^i) < 0 \end{array}\right\}. \tag{12}$$

PROPOSITION 3.2 *Suppose $|\mathcal{I}| < \infty$, i.e. P is a polytope. Let $\bar{x} \in P$, and define $\bar{s}_i := b_i - a_i^{\mathrm{T}}\bar{x}$ for all $i \in \mathcal{I}$. Then Equation* (11) *is empty, i.e. $\bar{x} \notin N(P)$, if and only if Equation* (12) *is non-empty.*

Proof We first argue that, when feasible, the set

$$\left\{(\lambda, (v_i, u^i)) : \sum_{i\in\mathcal{I}} \begin{pmatrix} u^i \\ 0 \end{pmatrix} a_i^{\mathrm{T}} + \mathrm{Diag}(\lambda) = 0\right\}$$

has non-empty interior with respect to the cones $\mathcal{K}_q \ni (v_i, u^i)$. This follows because we may arbitrarily increase each v_i without affecting the matrix equation. The proposition is now a straightforward application of the conic version of Farkas' lemma [4]. ■

Sets of the form (12) will be crucial for our analysis in the remainder of Section 3 and in Section 4. For ease of reference and in order to facilitate the derivation of various results, we now establish definitions, notations, and basic results related to Equation (12). We still assume fixed $(p, \mathcal{J})$.

Let $\hat{A} \in \Re^{\hat{m}\times n}$, $\hat{b} \in \Re^{\hat{m}}$, and $\hat{x} \in \Re^n$ be given, where $\hat{m} < \infty$. Also define the polyhedron

$$P(\hat{A}, \hat{b}) := \{x \in \Re^n : \hat{A}x \le \hat{b}\}.$$

PROPOSITION 3.3 *With the above definitions, it holds that $\hat{x} \notin N(P(\hat{A}, \hat{b}))$ if and only if*

$$\mathcal{C}(\hat{A}, \hat{b}, \hat{x}) := \left\{(\lambda, (v_i, u^i)) \in \Re^n \times \mathcal{K}_q^n : \begin{array}{c} \begin{pmatrix} U \\ 0 \end{pmatrix} \hat{A} + \mathrm{Diag}(\lambda) = 0 \\ \hat{x}^{\mathrm{T}}\lambda + \hat{x}_{\mathcal{J}}^{\mathrm{T}} U\hat{b} + (\hat{b} - \hat{A}\hat{x})^{\mathrm{T}}(r\, v - U^{\mathrm{T}} d) < 0 \end{array}\right\}$$

is non-empty, where $U = (u^1, \ldots, u^n) \in \Re^{|\mathcal{J}|\times n}$ and $v = (v_1, \ldots, v_n)^{\mathrm{T}} \in \Re^n$.

Proof This is just a restatement of Proposition 3.2 for the generic polyhedron $P(\hat{A}, \hat{b})$. ■

In addition, we define a *proposed canonical solution* associated with $\mathcal{C}(\hat{A}, \hat{b}, \hat{x})$ when $\hat{A}$ is square and invertible as follows:

$$U(\hat{A}) := (u^1(\hat{A}), \ldots, u^n(\hat{A})) := [\hat{A}^{-1}]_{\mathcal{J}\cdot}, \tag{13a}$$

$$v(\hat{A}) := (\|u^1(\hat{A})\|_q, \ldots, \|u^n(\hat{A})\|_q)^{\mathrm{T}}, \tag{13b}$$

$$\lambda := \begin{pmatrix} -e^{\mathrm{T}} & 0 \end{pmatrix}^{\mathrm{T}}, \tag{13c}$$

where e is the all-ones vector of length $|\mathcal{J}|$.

PROPOSITION 3.4 *If $\hat{A}$ is square and invertible, then the proposed canonical solution* $(U, v, \lambda) := (U(\hat{A}), v(\hat{A}), \lambda)$ *given by Equation* (13) *is feasible for* $\mathcal{C}(\hat{A}, \hat{b}, \hat{x})$ *if and only if*

$$-e^{\mathrm{T}}\hat{x}_{\mathcal{J}} + \hat{x}_{\mathcal{J}}^{\mathrm{T}} U\hat{b} + (\hat{b} - \hat{A}\hat{x})^{\mathrm{T}}(r\,v - U^{\mathrm{T}}d) < 0.$$

Proof By construction, (U, v, λ) satisfies all of the conditions for membership in $\mathcal{C}(\hat{A}, \hat{b}, \hat{x})$ except possibly the strict inequality. ■

3.2 *Iteration complexity*

For the discussion in this subsection, we assume that $|\mathcal{I}| < \infty$, i.e. P is a polytope.

The general interior-point methodology of Nesterov and Nemirovskii [23] can be used to derive iteration complexity results for solving the p-cone relaxation (9) and/or its dual (10). Stated with respect to Equation (9), the key result is as follows:

THEOREM 3.5 [23] *Suppose that Equation* (9) *is interior feasible with finite optimal value* v^*. *Given a polynomial-time self-concordant barrier for* $\mathcal{K}_p$ *with barrier parameter* θ_p*, an interior feasible solution* (x^0, X^0)*, and a tolerance* $\varepsilon > 0$*, there exists an algorithm ('short-step primal-only interior-point algorithm'), which delivers a solution* (x^*, X^*) *satisfying* $c^{\mathrm{T}}x^* - v^* < \varepsilon$ *within* $\mathcal{O}(\sqrt{\theta_p|\mathcal{I}|}\log(\varepsilon^{-1}(c^{\mathrm{T}}x^0 - v^*)))$ *iterations, each of which takes polynomial time.*

As is evident from the theorem, the key ingredient determining the iteration complexity of the interior-point algorithm is the barrier parameter θ_p. It is well known that there exists a self-concordant barrier for $\mathcal{K}_2$ with $\theta_2 = 2 = \mathcal{O}(1)$, and when $p \neq 2$, Andersen *et al*. [3] show the existence of a self-concordant barrier for $\mathcal{K}_p$ such that $\theta_p = 4|\mathcal{J}| = \mathcal{O}(|\mathcal{J}|)$. This implies the following corollary.

COROLLARY 3.6 *With respect to Theorem* 3.5*, Equation* (9) *can be solved in* $\mathcal{O}(\sqrt{|\mathcal{I}|})$ *iterations when* $p = 2$ *and* $\mathcal{O}(\sqrt{|\mathcal{J}||\mathcal{I}|})$ *iterations otherwise.*

It is interesting that the iteration complexity does not depend on $|\mathcal{J}|$ when $p = 2$.

This corollary illustrates that, among all relaxations as p varies in $[1, \infty]$, the second-order cone relaxation has the lowest overall theoretical iteration complexity. In addition, when $p = q = 2$, one can also apply the stronger algorithmic framework of Nesterov and Todd [24] for homogeneous self-dual cones to obtain long-step primal–dual path-following algorithms, which have high quality practical implementations.

From a theoretical point of view, then, one may be interested only in the relaxations when $p = 2$ (lowest iteration complexity) and $p = \infty$ (strongest relaxation and same iteration complexity as all other $p \neq 2$). Of course, what happens in practice may differ from theory, as we will see in Section 5.

To close this subsection, we remark that, for $p = 1$ and $p = \infty$, the iteration complexity given by Corollary 3.6 matches the iteration complexity obtained if one first formulates (9) as its standard LP representation and then applies an LP interior-point method to that representation.

3.3 *Two types of monotonicity*

Monotonicity is a relatively simple, yet important, property of the p-cone procedure outlined in Section 2.2. In fact, there are two types of monotonicity though both are derived from the same principle. The first involves the effect of the p-cone procedure on P and its subsets for fixed $(p, \mathcal{J})$, while the second involves the effect on P under different values of p.

The monotonicity properties that we wish to prove for $N_{(p,\mathcal{J})}(P)$ stem directly from the derivation of the p-cone procedure and particularly from the fact that F, $\hat{F}$, and $\hat{P}$ are defined with respect to the inequalities

$$(b_i - a_i^{\mathrm{T}} x)(r - \|x_{\mathcal{J}} - d\|_p) \geq 0 \quad \forall\, i \in \mathcal{I}; \tag{14}$$

see also Equation (2). It is evident that any strengthening of these inequalities in the representations of F, $\hat{F}$, and $\hat{P}$ can yield a corresponding strengthening of $N_{(p,\mathcal{J})}(P)$ around P^{01}. This is the key observation for the following two monotonicity properties.

The first monotonicity property involves strengthening the portion $b_i - a_i^{\mathrm{T}} x$ of Equation (14).

PROPOSITION 3.7 *Let $p \in [1, \infty]$ and $\emptyset \neq \mathcal{J} \subseteq [n]$ be fixed. Suppose Q is a convex set such that $F \subseteq Q \subseteq P$. Then $P^{01} \subseteq N_{(p,\mathcal{J})}(Q) \subseteq N_{(p,\mathcal{J})}(P) \subseteq P$.*

Proof The inclusions $P^{01} \subseteq N(Q)$ and $N(P) \subseteq P$ are derived directly from Proposition 2.2. To prove $N(Q) \subseteq N(P)$, we simply note that, with respect to $N(Q)$, the sets F, $\hat{F}$, and $\hat{P}$ are based on the inequalities

$$(g_\ell - f_\ell^{\mathrm{T}} x)(r - \|x_{\mathcal{J}} - d\|_p) \geq 0 \quad \forall\, \ell \in \mathcal{L},$$

where $Q = \{x \in \Re^n : f_\ell^{\mathrm{T}} x \leq g_\ell \quad \forall\ \ell \in \mathcal{L}\}$. Since $Q \subseteq P$, these inequalities are clearly a strengthening of Equation (14), and so $N(Q) \subseteq N(P)$. ■

The second monotonicity property involves strengthening the portion $r - \|x_{\mathcal{J}} - d\|_p$ of Equation (14) and requires the following lemma.

LEMMA 3.8 *Let $p' \geq 1$, and suppose $v \in \Re^s$ satisfies $\|v\|_{p'}^{p'} \leq s$. Then $\|v\|_p^p \leq s$ for all $p \in [1, p']$.*

Proof Without loss of generality, we replace v by its component-wise absolute value, i.e., we assume $v_j \geq 0$ for all j.

As a function of p (v fixed), $f(p) := \|v\|_p^p = \sum_{j=1}^s v_j^p$ is convex, and so its maximum over $[1, p']$ occurs at 1 or p'. So to prove the lemma it suffices to show $f(1) \leq s$, i.e.

$$\sum_{j=1}^s v_j \leq s \iff \left(\sum_{j=1}^s v_j\right)^{p'} \leq s^{p'}.$$

Next, $g(a) := a^{p'}$ is convex for non-negative a since $p' \geq 1$. In particular,

$$\begin{aligned}
\left(\sum_{j=1}^s v_j\right)^{p'} &= g\left(\sum_{j=1}^s v_j\right) = g\left(\sum_{j=1}^s \frac{1}{s} \cdot s v_j\right) \\
&\leq \sum_{j=1}^s \frac{1}{s}\, g(s v_j) = \sum_{j=1}^s s^{p'-1} v_j^{p'} = s^{p'-1} \sum_{j=1}^s v_j^{p'} \\
&= s^{p'-1} \|v\|_{p'}^{p'} \leq s^{p'},
\end{aligned}$$

as desired. ■

PROPOSITION 3.9 *Let $\emptyset \neq \mathcal{J} \subseteq [n]$ and $1 \le p \le p' \le \infty$ be given. Define $r := \sqrt[p]{|\mathcal{J}|}/2$, $r' := \sqrt[p']{|\mathcal{J}|}/2$, and $d = e/2 \in \Re^{|\mathcal{J}|}$ in accordance with Proposition 2.1. If $x \in \Re^n$ satisfies $\|x_{\mathcal{J}} - d\|_{p'} \le r'$, then $\|x_{\mathcal{J}} - d\|_p \le r$. As a consequence, $N_{(p',\mathcal{J})}(P) \subseteq N_{(p,\mathcal{J})}(P)$.*

Proof Regarding the first statement of the proposition, we can rearrange the desired implication as

$$\|2x_{\mathcal{J}} - e\|_{p'}^{p'} \le |\mathcal{J}| \Longrightarrow \|2x_{\mathcal{J}} - e\|_p^p \le |\mathcal{J}|,$$

which holds because of Lemma 3.8. Next, the inclusion $N_{(p',\mathcal{J})}(P) \subseteq N_{(p,\mathcal{J})}(P)$ follows because Equation (14) is strengthened when p and r are replaced by p' and r'. ■

3.4 *Elimination of fractional extreme points*

We introduce the following definition.

DEFINITION 3.10 *Let $\emptyset \neq \mathcal{J} \subseteq [n]$ be given. For any $\bar{x} \in \Re^n$, we say that $\bar{x}$ is $\mathcal{J}$-fractional if the subvector $\bar{x}_{\mathcal{J}}$ is contained in $[0, 1]^{|\mathcal{J}|}$ and has one or more fractional entries.*

In this subsection, we prove that $N(P)$ contains no $\mathcal{J}$-fractional points, which are extreme in P. Said differently, we show that $N(P)$ cuts off all $\mathcal{J}$-fractional extreme points of P.

We start with the case that $|\mathcal{I}| < \infty$, i.e. P is a polytope. Then we extend the ideas to balls. Finally, we use the analysis with balls to show our main result that, no matter the geometric structure of P, all $\mathcal{J}$-fractional points of P are cut off by $N(P)$.

3.4.1 *Polytopes*

When $|\mathcal{I}| < \infty$, P is a polytope, and since P is bounded in $[0, 1]^n$, we know that $|\mathcal{I}| > n$ and that P has extreme points. We prove the following proposition.

PROPOSITION 3.11 *Suppose $|\mathcal{I}| < \infty$, i.e. P is a polytope, and $\bar{x}$ is a $\mathcal{J}$-fractional extreme point of P. Then $\bar{x} \notin N(P)$.*

We give two proofs since we believe that both are instructive. The first is a direct proof that the set (11) is empty, which implies $\bar{x} \notin N(P)$; see the discussion in Section 3.1. The second proof follows the approach of Propositions 3.3 and 3.4.

Proof We must show Equation (11) is empty, where $\bar{s}_i := b_i - a_i^{\mathrm{T}}\bar{x}$. Because $\bar{x}$ is an extreme point of P, there exists $\mathcal{T} \subseteq \mathcal{I}$ of size n such that $\bar{s}_i = 0$ for all $i \in \mathcal{T}$ and the set $\{a_i : i \in \mathcal{T}\}$ is linearly independent. Hence, any X in Equation (11) must satisfy, for all $i \in \mathcal{T}$,

$$\begin{aligned}(0, b_i\bar{x}_{\mathcal{J}} - X_{\mathcal{J}\cdot}a_i) \in \mathcal{K}_p &\Longleftrightarrow X_{\mathcal{J}\cdot}a_i = b_i\bar{x}_{\mathcal{J}} \\ &\Longleftrightarrow X_{\mathcal{J}\cdot}a_i = (a_i^{\mathrm{T}}\bar{x})\bar{x}_{\mathcal{J}} \\ &\Longleftrightarrow (X_{\mathcal{J}\cdot} - \bar{x}_{\mathcal{J}}\bar{x}^{\mathrm{T}})a_i = 0.\end{aligned}$$

By the linear independence of $\{a_i : i \in \mathcal{T}\}$, it follows that X must satisfy $X_{\mathcal{J}\cdot} = \bar{x}_{\mathcal{J}}\bar{x}^{\mathrm{T}}$ with $\mathrm{diag}(X) = \bar{x}$. However, since $\bar{x}$ is $\mathcal{J}$-fractional, these conditions are inconsistent. So Equation (11) is empty. ■

Proof This proof assumes without loss of generality that $\mathcal{J} = \{1, \ldots, |\mathcal{J}|\}$ in line with Equation (12). Let $\mathcal{T}$ be defined as in the previous proof; we assume for simplicity that $\mathcal{T} = [n]$.

We first apply Propositions 3.3 and 3.4 with $(\hat{A}, \hat{b}, \hat{x}) := (A_{\mathcal{T}\cdot}, b_{\mathcal{T}}, \bar{x})$. Consider the proposed canonical solution $(U, v, \lambda) := (U(\hat{A}), v(\hat{A}), \lambda)$ given by Equation (13). Note that $\hat{b} - \hat{A}\hat{x} = 0$ and $U\hat{b} = \hat{x}_{\mathcal{J}}$. By Proposition 3.4, (U, v, λ) is feasible for $\mathcal{C}(\hat{A}, \hat{b}, \hat{x})$ since

$$\begin{aligned} -e^{\mathrm{T}}\hat{x}_{\mathcal{J}} + \hat{x}_{\mathcal{J}}^{\mathrm{T}}U\hat{b} + (\hat{b} - \hat{A}\hat{x})^{\mathrm{T}}(r\,v - U^{\mathrm{T}}d) &= -e^{\mathrm{T}}\hat{x}_{\mathcal{J}} + \hat{x}_{\mathcal{J}}^{\mathrm{T}}U\hat{b} + 0^{\mathrm{T}}(r\,v - U^{\mathrm{T}}d) \\ &= -e^{\mathrm{T}}\hat{x}_{\mathcal{J}} + \hat{x}_{\mathcal{J}}^{\mathrm{T}}\hat{x}_{\mathcal{J}} = -e^{\mathrm{T}}\bar{x}_{\mathcal{J}} + \bar{x}_{\mathcal{J}}^{\mathrm{T}}\bar{x}_{\mathcal{J}} < 0, \end{aligned}$$

where the inequality follows because $\bar{x}$ is $\mathcal{J}$-fractional. So, by Proposition 3.3, $\bar{x} \notin N(P(\hat{A}, \hat{b}))$.

Since $P(\hat{A}, \hat{b}) \supseteq P$, it holds by the monotonicity property of Proposition 3.7 that $N(P(\hat{A}, \hat{b})) \supseteq N(P)$. Hence, $\bar{x} \notin N(P)$ as desired. ■

3.4.2 *Balls*

In the previous subsection, we showed that, if P is a polytope, then $N(P)$ cuts off all $\mathcal{J}$-fractional extreme points from P. The proof used the fact that every extreme point in a polytope corresponds to n linearly independent active constraints. For general P, however, extreme points do not necessarily correspond to n active constraints. For example, if P is a ball in the interior of $[0, 1]^n$, then all extreme points of P have exactly one active constraint in the semi-infinite LP representation of P. In this subsection, we study balls to establish that $N(P)$ does in fact cut off all $\mathcal{J}$-fractional extreme points in this case as well. To avoid notational confusion with the P defined in Section 2.2, however, we will use B to denote the ball under investigation.

Let B be a ball centred at $h \in \Re^n$ with radius $R > 0$, i.e.

$$\begin{aligned} B &:= \{x : \|x - h\|_2 \le R\} \\ &= \{x : w^{\mathrm{T}}(x - h) \le R \quad \forall w \text{ s.t. } \|w\|_2 = 1\}. \end{aligned} \tag{15}$$

In keeping with the development of the p-cone procedure, we could just as well assume that B is the intersection of a ball and $[0, 1]^n$, but this is actually not necessary for the result that we present. Moreover, the analysis is a bit simpler without the assumption. The result is as follows.

PROPOSITION 3.12 *Suppose B is a ball given by Equation* (15) *for some centre $h \in \Re^n$ and radius $R > 0$. Suppose $\bar{x}$ is a $\mathcal{J}$-fractional extreme point of B. Then $\bar{x} \notin N(B)$.*

The proof of Proposition 3.12, although related to the proof of Proposition 3.11 for polytopes, is technically different. The fundamental difference is that, for balls, we have only one active constraint at $\bar{x}$, whereas for polytopes, we have n linearly independent ones. Nevertheless, the idea of the proof below is to carefully select n linearly independent constraints, which are *nearly* active at $\bar{x}$. By analysing those constraints, we see that they have the effect of cutting off $\bar{x}$, yielding a proof similar in spirit to that of Proposition 3.11.

Proof Since $\bar{x}$ is extreme, there exists some $\bar{w}$ with $\|\bar{w}\|_2 = 1$ such that $\bar{w}^{\mathrm{T}}(\bar{x} - h) = R$. In fact, $\bar{w} = R^{-1}(\bar{x} - h)$ since $\|\bar{x} - h\|_2 = R$. Related to $\bar{w}$, we define two additional vectors $\eta, \beta \in \Re^n$. First, let η be any vector having all non-zero entries such that $\eta^{\mathrm{T}}\bar{w} \neq 0$. For example, η could be taken as a small perturbation of $\bar{w}$ itself. Second, define $\beta := \eta^{-1}$, whose components are the inverses of the components of η.

For small $\theta > 0$, define the following collection of n vectors, each of which is a unit-length perturbation of $\bar{w}$:

$$\hat{a}_i := \ell_i^{-1}((1-\theta)\bar{w} + \theta\beta_i\, e_i) \quad \forall\, i = 1, \ldots, n,$$

where

$$\ell_i := \|(1-\theta)\bar{w} + \theta\beta_i\, e_i\|_2.$$

Note that both $\hat{a}_i$ and ℓ_i depend on θ even though our notation does not reflect this dependence. Note also that $\ell_i > 0$ for $\theta \approx 0$ and so $\hat{a}_i$ is well defined. Using the notation of Propositions 3.3 and 3.4, we define

$$\hat{A} := (\hat{a}_1, \ldots, \hat{a}_n)^{\mathrm{T}}$$
$$\hat{b} := Re + \hat{A}h$$

and consider the polyhedron

$$P(\hat{A}, \hat{b}) = \{x : \hat{A}x \leq \hat{b}\} = \{x : \hat{a}_i^{\mathrm{T}}(x - h) \leq R \quad \forall\, i = 1, \ldots, n\},$$

which contains B since its defining inequalities are a subset of those defining B. Our proof strategy will be to verify $\bar{x} \notin N(P(\hat{A}, \hat{b}))$ for $\theta \approx 0$ via Proposition 3.3 by showing that $\mathcal{C}(\hat{A}, \hat{b}, \bar{x})$ is non-empty. Since $N(B) \subseteq N(P(\hat{A}, \hat{b}))$ due to monotonicity (see Proposition 3.7), this will imply $\bar{x} \notin N(B)$ as desired.

To show $\mathcal{C}(\hat{A}, \hat{b}, \bar{x}) \neq \emptyset$, we now consider Proposition 3.4 and, in particular, show that the proposed canonical solution $(U, v, \lambda) := (U(\hat{A}), v(\hat{A}), \lambda)$ given by Equation (13) satisfies

$$-e^{\mathrm{T}}\bar{x}_{\mathcal{J}} + \bar{x}_{\mathcal{J}}^{\mathrm{T}} U\hat{b} + (\hat{b} - \hat{A}\bar{x})^{\mathrm{T}}(r\,v - U^{\mathrm{T}}d) < 0. \tag{16}$$

The proposed canonical solution (U, v, λ) is well defined because

$$\hat{A} = \mathrm{Diag}(\ell^{-1} \circ \beta)((1-\theta)\eta\bar{w}^{\mathrm{T}} + \theta I)$$

is invertible. In particular, its inverse via the Sherman–Morrison–Woodbury formula is

$$\hat{A}^{-1} = \theta^{-1}\left[I - \left(\frac{1-\theta}{\theta + (1-\theta)\bar{w}^{\mathrm{T}}\eta}\right)\eta\bar{w}^{\mathrm{T}}\right]\mathrm{Diag}(\ell \circ \eta) \tag{17}$$

Note that the denominator $\theta + (1-\theta)\bar{w}^{\mathrm{T}}\eta$ is non-zero for sufficiently small θ since $\bar{w}^{\mathrm{T}}\eta \neq 0$ by construction. So we now investigate the left-hand side of Equation (16) and show that its limit as $\theta \to 0_+$ exists and is less than 0, which suffices to establish that $\mathcal{C}(\hat{A}, \hat{b}, \bar{x}) \neq \emptyset$ for $\theta \approx 0$. Recall that $\hat{A}$ and $\hat{b}$ depend on θ, and hence so do U and v.

We first show that $\hat{A}^{-1}e$ equals $\bar{w} + \mathcal{O}(\theta)$. In other words, as $\theta \to 0_+$, $\hat{A}^{-1}e$ approaches $\bar{w}$ such that the error $\hat{A}^{-1}e - \bar{w}$ goes to 0 at least as fast as θ itself. A Taylor series expansion of ℓ about $\theta = 0$ shows that

$$\ell = (1-\theta)e + \theta\,\beta \circ \bar{w} + \mathcal{O}(\theta^2),$$

and so

$$\ell \circ \eta = (1-\theta)\eta + \theta\bar{w} + \mathcal{O}(\theta^2)$$

since $\beta = \eta^{-1}$. In addition, using $\bar{w}^{\mathrm{T}}\bar{w} = 1$, we have

$$\bar{w}^{\mathrm{T}}(\ell \circ \eta) = (1-\theta)\bar{w}^{\mathrm{T}}\eta + \theta + \mathcal{O}(\theta^2).$$

Therefore, from Equation (17) and the preceding equations,

$$\begin{aligned}
\hat{A}^{-1}e &= \theta^{-1}\left[I - \left(\frac{1-\theta}{\theta + (1-\theta)\bar{w}^{\mathrm{T}}\eta}\right)\eta\bar{w}^{\mathrm{T}}\right](\ell \circ \eta) \\
&= \theta^{-1}\left[\ell \circ \eta - \left(\frac{1-\theta}{\theta + (1-\theta)\bar{w}^{\mathrm{T}}\eta}\right)\eta \cdot \bar{w}^{\mathrm{T}}(\ell \circ \eta)\right] \\
&= \theta^{-1}\left[\ell \circ \eta - \left(\frac{1-\theta}{\theta + (1-\theta)\bar{w}^{\mathrm{T}}\eta}\right)(\theta + (1-\theta)\bar{w}^{\mathrm{T}}\eta + \mathcal{O}(\theta^2))\eta\right] \\
&= \theta^{-1}[\ell \circ \eta - (1-\theta)(1 + \mathcal{O}(\theta^2))\eta] \\
&= \theta^{-1}[\ell \circ \eta - (1-\theta)\eta + \mathcal{O}(\theta^2)] \\
&= \theta^{-1}[(1-\theta)\eta + \theta\bar{w} - (1-\theta)\eta + \mathcal{O}(\theta^2)] \\
&= \bar{w} + \mathcal{O}(\theta),
\end{aligned}$$

where the fourth equality follows from

$$\frac{\mathcal{O}(\theta)^2}{\theta + (1-\theta)\bar{w}^{\mathrm{T}}\eta} = \mathcal{O}(\theta^2),$$

which holds since $\bar{w}^{\mathrm{T}}\eta \neq 0$.

With $\hat{A}^{-1}e = \bar{w} + \mathcal{O}(\theta)$ now established, it follows that

$$\begin{aligned}
\hat{A}^{-1}\hat{b} &= \hat{A}^{-1}(Re + \hat{A}h) = R\hat{A}^{-1}e + h \\
&= R\bar{w} + h + \mathcal{O}(\theta), \\
\hat{A}^{-1}(\hat{b} - \hat{A}\bar{x}) &= \hat{A}^{-1}(Re + \hat{A}(h - \bar{x})) = R\bar{w} + h - \bar{x} + \mathcal{O}(\theta) \\
&= R \cdot R^{-1}(\bar{x} - h) + h - \bar{x} + \mathcal{O}(\theta) \\
&= \mathcal{O}(\theta),
\end{aligned}$$

where we have used the fact that $\bar{w} = R^{-1}(\bar{x} - h)$. From Equation (13), we have $U\hat{b} = R\bar{w}_{\mathcal{J}} + h_{\mathcal{J}} + \mathcal{O}(\theta)$, $U(\hat{b} - \hat{A}\bar{x}) = \mathcal{O}(\theta)$, and $v^{\mathrm{T}}(\hat{b} - \hat{A}\bar{x}) = \mathcal{O}(\theta)$. Thus, the left-hand side of Equation (16) is

$$\begin{aligned}
-e^{\mathrm{T}}\bar{x}_{\mathcal{J}} + \bar{x}_{\mathcal{J}}^{\mathrm{T}}(R\bar{w}_{\mathcal{J}} + h_{\mathcal{J}} + \mathcal{O}(\theta)) + \mathcal{O}(\theta) &= -e^{\mathrm{T}}\bar{x}_{\mathcal{J}} + \bar{x}_{\mathcal{J}}^{\mathrm{T}}(R \cdot R^{-1}(\bar{x}_{\mathcal{J}} - h_{\mathcal{J}}) + h_{\mathcal{J}}) + \mathcal{O}(\theta) \\
&= -e^{\mathrm{T}}\bar{x}_{\mathcal{J}} + \bar{x}_{\mathcal{J}}^{\mathrm{T}}\bar{x}_{\mathcal{J}} + \mathcal{O}(\theta).
\end{aligned}$$

As desired, the limit as $\theta \to 0_+$ of the left-hand side of Equation (16) equals $-e^{\mathrm{T}}\bar{x}_{\mathcal{J}} + \bar{x}_{\mathcal{J}}^{\mathrm{T}}\bar{x}_{\mathcal{J}}$, which is negative since $\bar{x}$ is $\mathcal{J}$-fractional. ■

We will actually use a feature of the above proof again for the proof of Theorem 4.1 in Section 4. So we record this result for easy reference.

COROLLARY 3.13 *Let B and $\bar{x}$ be as in Proposition 3.12. Then there exists a polyhedron $P(\hat{A}, \hat{b}) := \{x \in \Re^n : \hat{A}x \leq \hat{b}\}$, for some $(\hat{A}, \hat{b})$, such that $P(\hat{A}, \hat{b}) \supseteq B$ and $\bar{x} \notin N(P(\hat{A}, \hat{b})) \supseteq N(B)$.*

Proof The desired polyhedron is $P(\hat{A}, \hat{b})$, depending on small $\theta > 0$, in the proof of Proposition 3.12. ■

3.4.3 *The general case*

We now show that $N(P)$ cuts off all $\mathcal{J}$-fractional extreme points of P. The basic idea is that, given a $\mathcal{J}$-fractional extreme point $\bar{x} \in P$, there exists a ball $B \supseteq P$ such that $\bar{x}$ is also a $\mathcal{J}$-fractional extreme point of B. Thus, by Proposition 3.12 and the monotonicity property of Proposition 3.7, $\bar{x} \notin N(B) \supseteq N(P)$.

We first establish the existence of the ball B just described. A similar result has been used in Kojima and Tunçel [17,18].

PROPOSITION 3.14 *Let $\bar{x}$ be a $\mathcal{J}$-fractional extreme point of P. Then there exists a ball B such that $P \subseteq B$ and $\bar{x}$ is a $\mathcal{J}$-fractional extreme point of B.*

Proof This proposition is just a simple application of standard convex analysis. Recall that P is compact convex. Hence, there exists a hyperplane $H := \{x : \alpha^{\mathrm{T}} x = \beta\}$ supporting P at $\bar{x}$, i.e. $\bar{x} \in H$ and $P \setminus \{\bar{x}\} \subseteq H_{++} := \{x : \alpha^{\mathrm{T}} x > \beta\}$. We also define $H_{+} := \{x : \alpha^{\mathrm{T}} x \geq \beta\}$.

Next, given $\gamma > 0$, we define a ball $B(\gamma)$ dependent on $\bar{x}$ and α:

$$B(\gamma) := \{x : \|x - (\bar{x} + \gamma\alpha)\|_2 \leq \gamma\|\alpha\|_2\}.$$

It is easy to check that $B(\gamma) \subseteq H_{+}$ and that $\bar{x}$ is an extreme point of $B(\gamma)$. Furthermore, for every $x \in H_{++}$, there exists sufficiently large γ such that $x \in B(\gamma)$. Hence, because $P \setminus \{\bar{x}\} \subseteq H_{++}$ is bounded, there exists sufficiently large γ such that $P \setminus \{\bar{x}\} \subseteq B(\gamma)$, and so $P \subseteq B(\gamma)$. For any such large γ, we can take $B := B(\gamma)$ to achieve the proposition. ∎

The proof of the ensuing theorem then follows from Proposition 3.14.

THEOREM 3.15 *Let $\bar{x}$ be a $\mathcal{J}$-fractional extreme point of P. Then $\bar{x} \notin N(P)$.*

Proof Let B be the ball of Proposition 3.14. Then, by Proposition 3.12, $\bar{x} \notin N(B)$. Since $N(B) \supseteq N(P)$ by monotonicity of Proposition 3.7, $\bar{x} \notin N(P)$. ∎

4. Iterated procedure and convergence

So far we have discussed how the p-cone procedure produces $N_{(p,\mathcal{J})}(P)$ from P for a given $(p, \mathcal{J})$. Because $N(P)$ is compact convex with its own semi-infinite outer description which may or may not be known explicitly, we may conceptually apply the p-cone procedure – perhaps for a different choice of $(p, \mathcal{J})$ – to $N(P)$ itself. In fact, we may repeat the p-cone procedure *ad infinitum*. A key question is whether the resultant sequence of compact convex sets converges to P^{01}.

More formally, let $\{(p_k, \mathcal{J}^k)\}_{k\geq 1}$ be a sequence of choices $p_k \in [1, \infty]$ and $\emptyset \neq \mathcal{J}^k \subseteq [n]$, and define $N^1(P) := N_{(p_1,\mathcal{J}^1)}(P)$ and $N^k(P) := N_{(p_k,\mathcal{J}^k)}(N^{k-1}(P))$ for all $k > 1$. We then ask whether $\lim_{k\to\infty} N^k(P)$ equals P^{01}.

Lovász and Schrijver [21], Kojima and Tunçel [17,18], and Balas *et al.* [6] have all considered the same question for their own procedures. In particular, one may interpret Lovász and Schrijver [21] as taking $p_k = \infty$ and $\mathcal{J}^k = [n]$ for all k; they show finite convergence after n iterations, i.e. $N^n(P) = P^{01}$. Recall that the method of Kojima and Tunçel [17,18], when applied to 0-1 programs, essentially reduces to that of Lovász and Schrijver [21]; so they take the same p_k and $\mathcal{J}^k$. Finally, one may interpret Balas *et al.* [6] as choosing p_k arbitrarily and $\mathcal{J}^k$ a single element in $[n]$. The authors prove that, if $\mathcal{J}^1 \cup \cdots \cup \mathcal{J}^n = [n]$, then $N^n(P) = P^{01}$.

We show in Theorem 4.1 below that the iterated p-cone procedure converges asymptotically for arbitrary $\{p_k\}_{k=1}^{\infty}$ as long as each index $j \in [n]$ appears infinitely often in the sequence $\{\mathcal{J}^k\}_{k=1}^{\infty}$. Before stating and proving the theorem, we discuss a few items.

First, we have mentioned that the approach of Kojima and Tunçel [17,18] obtains asymptotic convergence in general. It is reasonable to ask if their approach or proof techniques subsume ours and hence prove convergence of our procedure. However, this is not the case since their asymptotic analysis uses *all* valid 'rank-2' quadratic inequalities, i.e. valid inequalities $(b_i - a_i^{\mathrm{T}}x)(b_h - a_h^{\mathrm{T}}x) \geq 0$ obtained by multiplying *any* pair of valid linear inequalities $b_i - a_i^{\mathrm{T}}x \geq 0$ and $b_h - a_h^{\mathrm{T}}x \geq 0$ for P. In contrast, our approach and analysis require only a partial subset of such inequalities, which are obtained by multiplying valid linear inequalities of the p-cone constraint $\|x_{\mathcal{J}} - d\|_p \leq r$ with the inequalities $b_i - a_i^{\mathrm{T}}x \geq 0$ defining P; see Section 2.4.2. In this sense, one can think of our approach as proving asymptotic convergence under weaker conditions than those used by Kojima and Tunçel [17,18] although the sets we consider are less general than those studied by Kojima and Tunçel.

Second, after the proof of Theorem 4.1, we provide an example where an infinite number of iterations is required to converge. So our p-cone successive relaxation procedure does not have finite convergence in general. Of course, for specific sequences $\{(p_k, \mathcal{J}^k)\}_{k=1}^{\infty}$, it may be possible to prove finiteness as with Lovász and Schrijver [21] and Balas *et al.* [6].

Third, we suspect that obtaining a rate of asymptotic convergence is difficult. This perceived difficulty stems from the methodology used to prove Theorem 3.15, which establishes that $\mathcal{J}$-fractional extreme points are cut off by the p-cone procedure. To establish a rate of convergence, it seems necessary to establish how 'deep' these cuts are with respect to P^{01}, but the methodology of Theorem 3.15 uses the existence of cuts with no quantitative knowledge of their strength.

We are now ready to state and prove the theorem.

THEOREM 4.1 *Let* $\{(p_k, \mathcal{J}^k)\}_{k\geq 1}$ *be a sequence of choices* $p_k \in [1,\infty]$ *and* $\emptyset \neq \mathcal{J}^k \subseteq [n]$*, which give rise to compact, convex sets* $N^k(P) \supseteq P^{01}$ *via the definitions* $N^1(P) := N_{(p_1,\mathcal{J}^1)}(P)$ *and* $N^k(P) := N_{(p_k,\mathcal{J}^k)}(N^{k-1}(P))$ *for all* $k > 1$*. Then* $N^k(P) \supseteq N^{k+1}(P)$ *so that* $\lim_{k\to\infty} N^k(P)$ *exists and equals* $\cap_{k\geq 1} N^k(P)$*. In addition, if* $\cup_{k\geq \bar{k}} \mathcal{J}^k = [n]$ *for all* $\bar{k}$*, then* $\lim_{k\to\infty} N^k(P) = P^{01}$*.*

Proof Since each $N^k(P)$ is compact and convex and contained in $N^{k-1}(P)$, $\lim_{k\to\infty} N^k(P)$ exists and equals $Z := \cap_{k=1}^{\infty} N^k(P)$. This proves the first part of the theorem.

To prove the second part, we first claim that every extreme point of Z is integral. Suppose for contradiction that $\bar{z}$ is a fractional extreme point of Z, and let j be any index where $\bar{z}_j$ is fractional. Next, let $\mathcal{S} := \{\mathcal{J} \subseteq [n] : j \in \mathcal{J}\}$. Theorem 3.15 implies $\bar{z} \notin N_{(1,\mathcal{J})}(Z)$ for all $\mathcal{J} \in \mathcal{S}$.

A continuity argument, which we prove two paragraphs below, implies that, for each $\mathcal{J} \in \mathcal{S}$, there exists $k_{\mathcal{J}}$ large enough so that $\bar{z} \notin N_{(1,\mathcal{J})}(N^{k-1}(P))$ for all $k \geq k_{\mathcal{J}}$. Define $\hat{k} := \max\{k_{\mathcal{J}} : \mathcal{J} \in \mathcal{S}\}$. In particular, consider $k \geq \hat{k}$ such that $j \in \mathcal{J}^k$. Since $\mathcal{J}^k \in \mathcal{S}$, it holds that $\bar{z} \notin N_{(1,\mathcal{J}^k)}(N^{k-1}(P))$. By the monotonicity property of Proposition 3.9, it also holds that $\bar{z} \notin N_{(1,\mathcal{J}^k)}(N^{k-1}(P)) \supseteq N_{(p_k,\mathcal{J}^k)}(N^{k-1}(P)) = N^k(P)$, which contradicts the statement $\bar{z} \in Z$. Hence, we conclude that every extreme point of Z is integral.

Since $P^{01} \subseteq Z$ by construction, it thus follows that $Z = P^{01}$.

Now we prove the continuity argument from above, i.e. for each $\mathcal{J} \in \mathcal{S}$, we prove the existence of $k_{\mathcal{J}}$ large enough so that $\bar{z} \notin N_{(1,\mathcal{J})}(N^{k-1}(P))$ for all $k \geq k_{\mathcal{J}}$. So let $\mathcal{J} \in \mathcal{S}$ be fixed. Since $\bar{z}$ is a $\mathcal{J}$-fractional extreme point of Z, by Proposition 3.14 there exists a ball $B \supseteq Z$ such that $\bar{z}$ is a $\mathcal{J}$-fractional extreme point of B. Furthermore, by Corollary 3.13, there exists a polyhedron

$$P(\hat{A}, \hat{b}) := \{x \in \Re^n : \hat{A}x \leq \hat{b}\},$$

for some $(\hat{A}, \hat{b})$, containing B such that $\bar{z} \notin N_{(1,\mathcal{J})}(P(\hat{A}, \hat{b}))$. By Proposition 3.3, this is equivalent to $\mathcal{C}(\hat{A}, \hat{b}, \bar{z}) \neq \emptyset$. For small $\varepsilon > 0$, this in turn implies $\mathcal{C}(\hat{A}, \hat{b} + \varepsilon e, \bar{z}) \neq \emptyset$ or, equivalently, $\bar{z} \notin N_{(1,\mathcal{J})}(P(\hat{A}, \hat{b} + \varepsilon e))$.

Since $N^{k-1}(P)$ converges to $Z \subseteq P(\hat{A}, \hat{b})$, for any $\varepsilon > 0$, there exists k_ε large enough so that $N^{k-1}(P) \subseteq P(\hat{A}, \hat{b} + \varepsilon e)$ for all $k \geq k_\varepsilon$. Then for small ε and large k, it follows that $\bar{z} \notin N_{(1,\mathcal{J})}(N^{k-1}(P))$ by the previous paragraph and monotonicity property. ■

We now give an example, which shows that the p-cone successive relaxation procedure may require an infinite number of iterations to converge.

To construct the example, we first analyse the behaviour of a particular class of two-dimensional polytopes under a single iteration of the procedure with $p < \infty$ and $\mathcal{J} = [2]$. For any $a > 1$, consider the following two-dimensional polytope:

$$P(a) := \left\{ (x_1, x_2) \geq 0 : \begin{array}{c} ax_1 + x_2 \leq a \\ x_1 + ax_2 \leq a \end{array} \right\}.$$

Note that $P(a)$ is symmetric about the line $x_1 = x_2$ and has the four vertices $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(a/(a+1), a/(a+1))$, the last of which is greater than $(1/2, 1/2)$ but smaller than $(1, 1)$. In particular, $P(a)$ is contained in $[0, 1]^2$, and its integer convex hull is

$$P^{01} = \{(x_1, x_2) \geq 0 : \ x_1 + x_2 \leq 1\} = P(1).$$

PROPOSITION 4.2 *Let $a > 1$ be given. For any $p < \infty$ and $\mathcal{J} = [2]$, there exists $a' \in (1, a)$ such that*

$$P(a') \subseteq N_{(p,\mathcal{J})}(P(a)).$$

Proof It suffices to show the existence of $1/2 < \delta < 1$ such that $(\delta, \delta) \in N(P(a))$ since then

$$\operatorname{conv}\{(0, 0), (0, 1), (1, 0), (\delta, \delta)\} = P\left(\frac{\delta}{1-\delta}\right)$$

is contained in $N(P(a))$. If such δ exists, we may take $a' = \delta/(1-\delta)$, which necessarily satisfies $a' \in (1, a)$.

By the discussion prior to Proposition 3.2, $(\delta, \delta) \in N(P(a))$ if and only if the set (11) with $x_1 = x_2 = \delta$ is non-empty. In our case, after eliminating $X_{11} = X_{22} = \delta$, the relevant set is

$$\hat{P}_\delta := \left\{ (0, 0) \leq (X_{12}, X_{21}) \leq (\delta, \delta) : \begin{array}{l} \left\| \begin{pmatrix} (a-1)\delta - aX_{21} - \frac{1}{2}\beta \\ X_{12} + \frac{1}{2}\beta \end{pmatrix} \right\|_p \leq r\beta \\ \left\| \begin{pmatrix} (a-1)\delta - aX_{12} - \frac{1}{2}\beta \\ X_{21} + \frac{1}{2}\beta \end{pmatrix} \right\|_p \leq r\beta \end{array} \right\},$$

where $r = \sqrt[p]{2}/2$ and $\beta := a - \delta(a+1)$. It then suffices to show that the set

$$\begin{aligned} \hat{P}'_\delta &:= \hat{P}_\delta \cap \{(X_{12}, X_{21}) : X_{12} = X_{21}\} \\ &= \left\{ 0 \leq y \leq \delta : \left\| \begin{pmatrix} (a-1)\delta - ay - \frac{1}{2}\beta \\ y + \frac{1}{2}\beta \end{pmatrix} \right\|_p \leq r\beta \right\} \end{aligned}$$

is non-empty.

We next consider the closely related set

$$\hat{P}'_{1/2} := \left\{ 0 \le y \le \frac{1}{2} : \left\| \begin{pmatrix} (a-1)/4 - ay \\ (a-1)/4 + y \end{pmatrix} \right\|_p \le \frac{r}{2}(a-1) \right\},$$

which is obtained by substituting $\delta \leftarrow 1/2$ in the definition of $\hat{P}'_\delta$. Within this convex set, note that $y = 0$ is feasible and makes the cone constraint active. Since $a > 1$ and $p < \infty$, it is then easy to see that increasing y to a small positive number satisfies the cone constraint strictly. In other words, $\hat{P}'_{1/2}$ has non-empty interior.

By continuity, it then follows that for δ sufficiently close to $1/2$, $\hat{P}'_\delta$ is non-empty, as desired. ■

With Proposition 4.2 in hand, we can construct the example having infinite convergence. For any $a > 1$, consider the sequence $N^k(P(a))$ from Theorem 4.1 based on any choice $\{(p_k, \mathcal{J}^k)\}_{k\ge 0}$ satisfying $p_k < \infty$ and $\mathcal{J}^k = [2]$ for all k. By Proposition 4.2, the monotonicity property of Proposition 3.7, and induction there exists $a_k > 1$ such that $P(a_k) \subseteq N^k(P(a))$ for all k. Hence, $N^k(P(a)) \ne P^{01}$ for all k, which ensures convergence only in the limit.

Note that this example also shows infinite convergence even if symmetry $X = X^{\mathrm{T}}$ is enforced in the p-cone procedure since symmetry is enforced in the proof of Proposition 4.2.

5. Computational considerations

For fixed $\mathcal{J}$, a single application of the p-cone lift-and-project procedure gives rise to a family of relaxations of P^{01} parameterized by p. From the monotonicity property of Proposition 3.9, we know that the larger p is, the tighter the corresponding relaxation. So $p = \infty$ is the tightest. On the other hand, we have shown in Corollary 3.6 that optimizing over the $p = 2$ relaxation induces the lowest theoretical iteration complexity – in fact, an order of magnitude less than all other p, which themselves share the same iteration complexity. Thus, one may be particularly interested in the cases $p = 2$ (SOCP) and $p = \infty$ (LP).

In this section, we computationally test these cases using state-of-the-art SOCP and LP software. We had hypothesized that the lower iteration complexity combined with the high quality of modern SOCP software would make solving $p = 2$ quicker than solving $p = \infty$ – perhaps much quicker so as to justify the loss in relaxation quality. However, as described next, we have observed that solving $p = \infty$ is faster with better bounds.

We conduct experiments on two sets of problems. The first set includes eight instances of the maximum stable set problem from the Center for Discrete Mathematics and Theoretical Computer Science [13]. For a graph G with vertex set $[n]$ and edge set $E \subseteq [n] \times [n]$, the (unweighted) maximum stable set problem is

$$\alpha := \max\{e^{\mathrm{T}}x \mid x_i + x_j \le 1,\ (i, j) \in E,\ x \in \{0, 1\}^n\}.$$

Table 1 contains a basic description of the instances.

We set $\mathcal{J} = [n]$ and solve both the $p = 2$ and $p = \infty$ relaxations. We also enforce the symmetry condition $X = X^{\mathrm{T}}$ of Equation (6) so as to eliminate about half of the variables in X. The SOCPs were solved using MOSEK 5.0, and the LPs were solved using both CPLEX 9.0 and MOSEK 5.0. Pre-solving was turned off for all solvers, and computations were performed under the Linux operating system with a single 2.8 GHz AMD Opteron processor and 4 GB of RAM.

Regarding the solution of the LPs, we used CPLEX to solve the dual form (10) using the dual simplex method, which gave better results than, for example, solving Equation (9) with the

Table 1. Description of the stable set instances.

Name	n	$\|E\|$
johnson8-2-4	28	168
MANN-a9	45	72
hamming6-2	64	192
keller4	171	5100
brock200_1	200	5066
san200_07_1	200	5970
sanr200_07	200	6032
c-fat200-1	200	18366

dual simplex method. On the other hand, MOSEK's LP solver optimizes Equations (9) and (10) simultaneously using a primal–dual interior-point method.

Table 2 compares α, the bounds, and the solution times (in seconds). The values for those cells containing '$*$' were unavailable due to the solvers running out of memory.

Table 2 clearly shows the overall superiority of the LP relaxation as solved by MOSEK for the maximum stable set problems. The $p = \infty$ relaxation can be solved faster and the bounds are better as well. Still, it is worth noting that the SOCP relaxations solve more quickly than the LP relaxations via the dual simplex method.

The second set of test problems includes seven mixed-integer programming problems from MIPLIB 2003 [1]. A description of the problems is shown in Table 3. Specifically, *rows* and *columns* give the number of constraints and variables of those problems. There are two types of variables, binary and continuous, whose numbers are listed under the last two columns of Table 3. The column *non-zero* indicates the number of non-zero entries in the constraints. We note that the problems contain no general integer variables. These problems are among some of the smallest problems of MIPLIB 2003, but not all of them are considered easy to solve. For example, *markshare1* is small in size but is classified as 'hard' in MIPLIB 2003.

Just as with the stable set instances, we set $\mathcal{J}$ equal to the index set of binary variables and solve the $p = \infty$ and $p = 2$ relaxations using CPLEX 9.0 and MOSEK 5.0. The computing environment is similar, the only difference being that we turn on pre-solve for both CPLEX and MOSEK to avoid out-of-memory failures. The bounds of the relaxations and the times in seconds are reported in Table 4 along with the optimal and LP-relaxation values. Note that these are minimization problems.

Table 2. The bounds and times (in seconds) for solving the $p = \infty$ and $p = 2$ relaxations of the stable set instances from Table 1.

					Times		
			Bounds		$p = \infty$		
Name	LP value	α	$p = \infty$	$p = 2$	CPLEX	MOSEK	$p = 2$
johnson8-2-4	14	4	9.33	12.2	9.00e−2	1.30e−1	2.40e−1
MANN-a9	22.5	16	18	20.5	1.76e+0	2.40e−1	1.15e+0
hamming6-2	32	32	32	32	2.39e+1	1.91e+0	1.00e+1
keller4	85.5	11	57	80.9	2.14e+3	1.44e+3	4.67e+3
brock200_1	100	21	66.7	95.1	3.36e+4	8.09e+3	1.18e+4
san200_07_1	100	30	66.7	95.1	1.00e+5	9.00e+2	3.05e+4
sanr200_07	100	18	66.7	95	5.76e+4	6.93e+3	1.21e+4
c-fat200-1	100	12	$*$	$*$	$*$	$*$	$*$

Note. Each LP is solved using two methods: the dual simplex method (CPLEX) and the primal–dual interior-point method (MOSEK). An asterisk ($*$) indicates that the corresponding solver ran out of memory. A time limit of 100,000 s is enforced for each run.

Table 3. Description of selected instances from MIPLIB 2003.

Name	Rows	Columns	Non-zero	Binary	Continuous
markshare1	6	62	312	50	12
pk1	45	86	915	55	31
pp08a	136	240	480	64	176
pp08aCUTS	246	240	839	64	176
modglob	291	422	968	98	324
danoint	664	521	3232	56	465
qiu	1192	840	3432	48	792

Table 4. The bounds and times (in seconds) for solving $p = \infty$ and $p = 2$ relaxations of selected instances from MIPLIB 2003.

			Bounds			Times		
			$p = \infty$			$p = \infty$		
Name	LP value	IP value	CPLEX	MOSEK	$p = 2$	CPLEX	MOSEK	$p = 2$
markshare1	0	1	0	0	0	9.12e+0	1.45e+0	4.02e+0
pk1	1.47e−9	1.10e+1	0	0	3.43e−3	1.58e+2	9.10e+0	2.50e+1
pp08a	2.75e+3	7.35e+3	6.40e+3	6.40e+3	3.11e+3	2.38e+2	3.21e+1	3.91e+1
pp08aCUTS	5.48e+3	7.35e+3	6.77e+3	6.77e+3	5.61e+3	7.19e+2	8.13e+1	1.49e+2
modglob	2.04e+7	2.07e+7	–	2.06e+7	2.04e+7	1.00e+5	3.10e+3	4.28e+3
danoint	6.26e+1	6.57e+1	–	6.28e+1	6.27e+1	1.00e+5	4.74e+2	4.53e+4
qiu	−9.32e+2	−1.33e+2	−3.08e+2	−3.08e+2	−8.37e+2	3.95e+4	1.06e+2	2.19e+2

Note. The relaxations of $p = \infty$ were solved using both CPLEX and MOSEK. A time limit of 100,000 s was enforcedand and '–' means the solver did not find the optimal value when time limit was reached.

The conclusion, we draw from Table 4, which is similar to that of Table 2: overall, the $p = \infty$ relaxation outperforms the $p = 2$ relaxation in terms of both bounds and CPU times. For *pk1*, the bound for $p = 2$ is better than that for $p = \infty$, and the reason is that MOSEK only obtained a 'near optimal' solution for *pk1*.

Although these computational results do not align with the theoretical result of Corollary 3.6 that solving the $p = 2$ relaxation has a lower iteration complexity, we are hopeful that improvements in SOCP software may make the p-cone procedure more competitive in the future. These results are a testament to the high quality of current LP software.

6. Conclusions

In this paper, we have introduced lift-and-project procedures for 0-1 integer programming based on POCP. From a theoretical point of view, our approach generalizes and unifies several existing methods, which have been based on linear and SDP. Asymptotic convergence of the repeated application of our procedure has also been established, and for $p = 2$, when applying one iteration of the p-cone procedure, our method enjoys a theoretical iteration complexity, which is an order of magnitude faster than existing lift-and-project techniques. From the computational point of view, solving the SOCP corresponding to $p = 2$ is not competitive with solving the LP for $p = \infty$. Overall, we believe that the p-cone procedure makes a solid theoretical contribution to the literature on lift-and-project procedures, with possible computational improvements in the future as SOCP solvers become more efficient.

We conclude with a final observation. We have mentioned in the Introduction that Kim and Kojima [15] have derived SOCP relaxations from SDP lift-and-project relaxations by replacing semi-definiteness with just the semi-definiteness of 2×2 principal submatrices, which can be modelled by SOCP. In addition, Kim and Kojima [14] and Kim *et al*. [16] use SDP lift-and-project relaxations to generate valid convex quadratic constraints, which are SOCP-representable and then enforced in place of semi-definiteness. When $p = 2$, the approach in this paper is different. In particular, our method is not derived from semi-definiteness. In fact, irrespective of p, semi-definiteness can be applied in our procedure to further enhance its strength, and so the above SOCP ideas can also be applied to our procedure as well for any p.

Acknowledgements

The authors would like to thank two anonymous referees for many helpful suggestions that have improved this paper immensely. Both authors supported in part by NSF Grant CCF-0545514.

References

[1] T. Achterberg, T. Koch, and A. Martin, *MIPLIB 2003*, Oper. Res. Lett. 34(4) (2006), pp. 1–12. Available at http://www.zib.de/Publications/abstracts/ZR-05-28/. See http://miplib.zib.de.
[2] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Math. Program. 95(1, Ser. B) (2003), pp. 3–51. ISMP 2000, Part 3 (Atlanta, GA).
[3] E.D. Andersen, C. Roos, and T. Terlaky, *Notes on duality in second order and p-order cone optimization*, Optimization 51(4) (2002), pp. 627–643.
[4] E.J. Anderson and P. Nash, *Linear programming in infinite-dimensional spaces*. *Wiley-interscience series in discrete mathematics and optimization*, *Theory and Applications*, A Wiley-Interscience Publication, John Wiley & Sons Ltd., Chichester, 1987.
[5] E. Balas and M. Perregaard, *A precise correspondence between lift-and-project cuts, simple disjunctive cuts, and mixed integer Gomory cuts for 0-1 programming*, Math. Program. 94(2–3, Ser. B) (2003), pp. 221–245. The Aussois 2000 Workshop in Combinatorial Optimization.
[6] E. Balas, S. Ceria, and G. Cornuéjols, *A lift-and-project cutting plane algorithm for mixed 0-1 programs*, Math. Program. 58 (1993), pp. 295–324.
[7] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization*, MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. ISBN 0-89871-491-5. Analysis, algorithms, and engineering applications.
[8] D. Bienstock and M. Zuckerberg, *Subset algebra lift operators for 0-1 integer programming*, SIAM J. Optim. 15(1) (2004), pp. 63–95 (electronic).
[9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004, ISBN 0-521-83378-7.
[10] S. Burer and D. Vandenbussche, *Solving lift-and-project relaxations of binary integer programs*, SIAM J. Optim. 16(3) (2006), pp. 493–512.
[11] F. Glineur and T. Terlaky, *Conic formulation for lp-norm optimization*, J. Optim. Theory Appl. 122(2) (2004), pp. 285–307.
[12] R.E. Gomory, *An algorithm for integer solutions to linear programs*, in *Recent Advances in Mathematical Programming*, R. Graves and P. Wolfe, eds., McGraw-Hill, New York, 1963, pp. 269–302.
[13] D. Johnson and M. Trick, *Cliques, Coloring, and Satisfiability: Second DIMACS Implementation Challenge*, American Mathematical Society, Providence, RI, 1996.
[14] S. Kim and M. Kojima, *Second order cone programming relaxation of nonconvex quadratic optimization problems*, Optim. Methods Softw. 15(3–4) (2001), pp. 201–224.
[15] S. Kim and M. Kojima, *Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations*, Comput. Optim. Appl. 26(2) (2003), pp. 143–154.
[16] S. Kim, M. Kojima, and M. Yamashita, *Second order cone programming relaxation of a positive semidefinite constraint*, Optim. Methods Softw. 18(5) (2003), pp. 535–541.
[17] M. Kojima and L. Tunçel, *Cones of matrices and successive convex relaxations of nonconvex sets*, SIAM J. Optim. 10(3) (2000), pp. 750–778.
[18] M. Kojima and L. Tunçel, *Discretization and localization in successive convex relaxation methods for nonconvex quadratic optimization*, Math. Program. 89(1, Ser. A) (2000), pp. 79–111. ISSN 0025-5610.
[19] P. Krokhmal and P. Soberanis, *Risk optimization with p-order conic constraints: a linear programming approach*, Working paper, University of Iowa, Iowa City, IA, 2008.
[20] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. 11(3) (2001), pp. 796–817.
[21] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and 0-1 optimization*, SIAM J. Optim. 1 (1991), pp. 166–190.
[22] MOSEK, Inc. *The MOSEK optimization tools manual 5.0*, 2007.

[23] Y.E. Nesterov and A.S. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, Society for Industrial and Applied Mathematics, Philadelphia, 1994.
[24] Y.E. Nesterov and M.J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Math. Oper. Res. 22(1) (1997), pp. 1–42.
[25] P.A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program. 96(2, Ser. B) (2003), pp. 293–320. Algebraic and geometric methods in discrete optimization.
[26] H.D. Sherali and W.P. Adams, *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems*, SIAM J. Discrete Math. 3(3) (1990), pp. 411–430.
[27] J.F. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optim. Methods Softw. 11/12(1–4) (1999), pp. 625–653. Available at http://sedumi.mcmaster.ca/.
[28] R.H. Tütüncü, K.C. Toh, and M.J. Todd, *SDPT3: a Matlab software package for semidefinite-quadratic-linear programming, version 3.0*, August 2001. Available at http://www.math.nus.edu.sg/∼mattohkc/sdpt3.html.
[29] G. Xue and Y. Ye, *An efficient algorithm for minimizing a sum of p-norms*, SIAM J. Optim. 10(2) (2000), pp. 551–579.