MIXED INTEGER SECOND ORDER CONE PROGRAMMING

SARAH DREWES* AND STEFAN ULBRICH †

Abstract. This paper deals with solving strategies for mixed integer second order cone problems. We present different lift-and-project based linear and convex quadratic cut generation techniques for mixed 0-1 second-order cone problems and present a new convergent outer approximation based approach to solve mixed integer SOCPs. The latter is an extension of outer approximation based approaches for continuously differentiable problems to subdifferentiable second order cone constraint functions.
We give numerical results for some application problems, where the cuts are applied in the context of a nonlinear branch-and-cut method and the branch-and-bound based outer approximation algorithm. The different approaches are compared to each other.

Key words. Mixed Integer Nonlinear Programming, Second Orde Cone Programming, Outer Approximation, Cuts

AMS(MOS) subject classifications. 90C11

1. Introduction. Mixed Integer Second Order Cone Programs (MISOCP) can be formulated as

$$\begin{array}{rrcll} \min & c^T x & & & \\ s.t. & Ax & = & b & \\ & x & \succeq & 0 & \\ & x_j & \in & [l_j, u_j] & (j \in J), \\ & x_j & \in & \mathbb{Z} & (j \in J), \end{array} \tag{1.1}$$

where $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$, $l_j, u_j \in \mathbb{R}$ and $x \succeq 0$ denotes that $x \in \mathbb{R}^n$ consists of *noc* part vectors $x_i \in \mathbb{R}^{k_i}$ lying in second order cones defined by

$$\mathcal{K}_i = \{x_i = (x_{i0}, x_{i1}^T)^T \in \mathbb{R} \times \mathbb{R}^{k_i - 1} : \|x_{i1}\|_2 \leq x_{i0}\}.$$

Mixed integer second order cone problems have various applications in finance or engineering, for example turbine balancing problems, cardinality-constrained portfolio optimization (cf. Bertsimas and Shioda in [12]) or the problem of finding a minimum length connection network also known as the Euclidian Steiner Tree Problem (ESTP) (cf. Fampa, Maculan in [11]).

Available convex MINLP solvers like `BONMIN` [19] by Bonami et al. or `FilMINT` [22] by Abhishek et al. are not applicable for (1.1), since the occurring second order cone constraints are not continuously differentiable.

Branch-and-cut methods for convex mixed 0-1 problems had been discussed

*Research Group Nonlinear Optimization, Department of Mathematics, Technische Universität Darmstadt, Germany.

†Research Group Nonlinear Optimization, Department of Mathematics, Technische Universität Darmstadt, Germany.

by Stubbs and Mehrotra in [1] and [6]. In [3] Çezik and Iyengar discuss cuts for general self-dual conic programming problems and investigate their applications on the maxcut and the traveling salesman problem. Atamtürk and Narayanan present in [8] integer rounding cuts for conic mixed-integer programming by investigating polyhedral decompositions of the second order cone conditions. There is also an article [7] dealing with outer approximation techniques for MISOCPs by Vielma et al. which is based on Ben-Tal and Nemirovskii's polyhedral outer approximation of second order cone constraints [9].

In this paper we present lift-and-project based linear and quadratic cuts for mixed 0-1 problems by extending results from [1] by Stubbs, Mehrotra and [3] by Çezik, Iyengar. Furthermore, a hybrid branch&bound based outer approximation approach for MISOCPs is developed. Thereby linear outer approximations based on subgradients satisfying the Karush Kuhn Tucker (KKT) optimality conditions of the occurring SOCP problems enable us to extend the convergence result for continuously differentiable constraints to subdifferentiable second order cone constraints. In numerical experiments the latter algorithm is compared to a nonlinear branch-and-bound approach and the impact of the cutting techniques is investigated in the context of both algorithms.

2. Lift-and-Project Cuts for Mixed 0-1 SOCPs. The cuts presented in this section are based on lift-and-project based relaxations that will be introduced in Section 2.1. Cuts based on similar relaxation hierarchies have previously been developed for mixed 0-1 linear programming problems, see for example [10] by Balas et al..

2.1. Relaxations. In [1], Stubbs and Mehrotra generalize the lift-and-project relaxations described in [10] to the case of mixed 0-1 convex programming. We describe these relaxations with respect to second order cone constraints. Throughout the rest of this section we consider mixed-0-1 second order cone problems of the form (1.1), where $l_j = 0, u_j = 1$, for all $j \in J$. We define the following sets associated with (1.1): The binary feasible set $C^0 := \{x \in \mathbb{R}^n : Ax = b, x \succeq 0, x_k \in \{0,1\}, k \in J\}$, its continuous relaxation $C := \{x \in \mathbb{R}^n : Ax = b, x \succeq 0, x_k \in [0,1], k \in J\}$ and $C^j := \{x \in \mathbb{R}^n : x \in C, x_j \in \{0,1\}\} \quad (j \in J)$.
In the binary case it is possible to generate a hierarchy of relaxations that is based on the continuous relaxation C and finally describes $\mathrm{conv}(C^0)$, the convex hull of C^0. For a lifting procedure that yields a description of $\mathrm{conv}(C^j)$, we introduce further variables $u^0 \in \mathbb{R}^n, u^1 \in \mathbb{R}^n, \lambda^0 \in \mathbb{R}, \lambda^1 \in \mathbb{R}$

and define the set

$$M_j(C) = \left\{ (x, u^0, u^1, \lambda^0, \lambda^1) : \begin{array}{l} \lambda^0 u^0 + \lambda^1 u^1 = x \\ \lambda^0 + \lambda^1 = 1, \lambda^0, \lambda^1 \geq 0 \\ Au^0 = b \\ Au^1 = b \\ u^0 \succeq 0 \\ u^1 \succeq 0 \\ (u^0)_k \in [0,1] \quad (k \in J, k \neq j) \\ (u^1)_k \in [0,1] \quad (k \in J, k \neq j) \\ (u^0)_j = 0, (u^1)_j = 1 \end{array} \right\}.$$

To eliminate the nonconvex bilinear equality constraint we use substitution $v^0 := \lambda^0 u^0$ and $v^1 := \lambda^1 u^1$ and get

$$\tilde{M}_j(C) = \left\{ (x, v^0, v^1, \lambda^0, \lambda^1) : \begin{array}{l} v^0 + v^1 = x \\ \lambda^0 + \lambda^1 = 1, \lambda^0, \lambda^1 \geq 0 \\ Av^0 - \lambda^0 b = 0 \\ Av^1 - \lambda^1 b = 0 \\ v^0 \succeq 0 \\ v^1 \succeq 0 \\ (v^0)_k \in [0, \lambda^0] \quad (k \in J, k \neq j) \\ (v^1)_k \in [0, \lambda^1] \quad (k \in J, k \neq j) \\ (v^0)_j = 0, (v^1)_j = \lambda^1 \end{array} \right\}. \tag{2.1}$$

Note that if $\lambda^i > 0$ $(i = 0, 1)$ $u^i \succeq 0 \Leftrightarrow \lambda^i u^i \succeq 0$, as well as $Au^i = b \Leftrightarrow \lambda^i A u^i = \lambda^i b$ hold and thus the conic and linear conditions remain invariant under the above transformation. In the case of $\lambda^i = 0$ $(i = 0, 1)$, the bilinear term $\lambda^i u^i$ vanishes as well as v^i vanishes due to $v^i_k \in [0, \lambda^i]$, for $k \neq j$ and $v^i_j = \lambda^i$. Thus, the projections of $M_j(C)$ and $\tilde{M}_j(C)$ on x are equivalent. We denote this projection by

$$P_j(C) := \{x : (x, v^0, v^1, \lambda^0, \lambda^1) \in \tilde{M}_j(C)\}. \tag{2.2}$$

Applying this lifting procedure for an entire subset of indices $B \subseteq J$, $B := \{i_1, \dots i_p\}$ yields

$$\tilde{M}_B(C) := \left\{ \begin{pmatrix} x \\ \begin{pmatrix} v^{0j} \\ v^{1j} \\ \lambda^{0j} \\ \lambda^{1j} \end{pmatrix} \\ j \in \{1, \dots p\} \end{pmatrix} : \begin{array}{l} v^{0j} + v^{1j} = x \\ \lambda^{0j} + \lambda^{1j} = 1, \lambda^{0j}, \lambda^{1j} \geq 0 \\ Av^{0j} - \lambda^{0j} b = 0 \\ Av^{1j} - \lambda^{1j} b = 0 \\ v^{0j} \succeq 0 \\ v^{1j} \succeq 0 \\ v^{1j}_{i_k} = v^{1k}_{i_j} \quad j < k \in \{1, \dots p\} \\ (v^{0j})_k \in [0, \lambda^{0j}] \quad (k \in J \setminus i_j) \\ (v^{1j})_k \in [0, \lambda^{1j}] \quad (k \in J \setminus i_j) \\ v^{0j}_{i_j} = 0, v^{1j}_{i_j} = \lambda^{1j} \end{array} \right\}. \tag{2.3}$$

Here we used the symmetry condition $v^{1j}_{i_k} = v^{1k}_{i_j}$ for all $k, j \in \{1, \dots p\}$from Theorem 6 in [1]. We denote the projection of $\tilde{M}_B(C)$ by

$$P_B(C) := \{x : (x, (v^{0j}, v^{1j}, \lambda^{0j}, \lambda^{1j})\ j \in \{1, \dots p\}) \in \tilde{M}_B(C)\}. \qquad (2.4)$$

The sets $P_B(C)$ are convex sets with $C^0 \subseteq P_B(C) \subseteq C$. Due to Theorem 7 in [1]

$$V_B - x_B x_B^T \succeq_{sd} 0 \qquad (2.5)$$

is another valid inequality for $P_B(C) \cap C^0$.We use this inequality to get a further tightening of the set $\tilde{M}_B(C)$:

$$\tilde{M}^+_B(C) := \{ \quad (x, (v^{0j}, v^{1j}, \lambda^{0j}, \lambda^{1j})\ j \in \{1, \dots p\}) \in \tilde{M}_B(C) : \\ V_B - x_B x_B^T \succeq_{sd} 0\}. \qquad (2.6)$$

Its projection on x will be denoted by

$$P^+_B(C) := \{x : (x, (v^{0j}, v^{1j}, \lambda^{0j}, \lambda^{1j})\ j \in \{1, \dots p\}) \in \tilde{M}^+_B(C)\}. \qquad (2.7)$$

The sequential applications of these lift-and-project procedures that generate the sets $P_j(C)$ in (2.2), $P_B(C)$ in (2.4) and $P^+_B(C)$ in (2.7), define a hierarchy of relaxations of C^0 containing conv(C^0), for which the following connections are cited from [1] and [3].

THEOREM 2.1. *Let $B \subseteq J$, $j \in J$ and $|J| = l$, then*

1. $P_j(C) = \text{conv}(C^j)$,
2. $P^+_B(C) \subseteq P_B(C) \subseteq \cap_{j \in B}\text{conv}(C^j)$,
3. $C^0 \subseteq P^+_B(C)$,
4. $P_{i_l}(P_{i_{l-1}}(\cdots P_{i_1})) = \text{conv}(C^0)$.
5. $(P_J)^l(C) = (P^+_J)^l(C) = \text{conv}(C^0)$,
   if $(P_J)^0(C) = (P^+_J)^0(C) = C$ *and* $(P_J)^k(C) = P_J((P_J)^{k-1}(C))$, $(P^+_J)^k(C) = P^+_J((P^+_J)^{k-1}(C))$, *for* $k = 1, \dots l$.

Proof: Part 1 and 2 follow by construction, 3 follows from (2.5). Part 4 and 5 follow from Theorem 1 and 6 in [1]. □

Note that the relaxations $P_B(C)$ and $P^+_B(C)$ are described by $\mathcal{O}(n|B|)$ variables and $\mathcal{O}(|B|)$ m-dimensional conic constraints. Thus, the number of variables and constraints grow linearly with $|B|$.

2.2. Cut Generation using Subgradients. Stubbs and Mehrotra showed in [1] that cuts for mixed 0-1 convex programming problems can be generated using the following theroem.

THEOREM 2.2. *Let $B \subseteq J$, $\bar{x} \notin P_B(C)$ and $\hat{x}$ be the optimal solution of the minimum distance problem* $\min_{x \in P_B(C)} f(x) := \|x - \bar{x}\|$. *Then there exists a subgradient ξ of $f(x)$ at $\hat{x}$, such that $\xi^T(x - \hat{x}) \geq 0$ is a valid linear inequality for every $x \in P_B(C)$ that cuts off $\bar{x}$.*

Proof. This result was shown by Stubbs and Mehrotra in [1], Theorem 3. □

If we choose the Euclidian norm as objective function, $f(x) := \|x-\bar{x}\|_2$, the minimum distance problem is a second order cone problem and we can use Theorem 2.2 to get a valid cut for (1.1).

PROPOSITION 2.1. *Let $B \subseteq J$, $\bar{x} \notin P_B(C)$ and $\hat{x}$ be the optimal solution of the minimum distance problem* $\min_{x \in P_B(C)} f(x) := \|x - \bar{x}\|_2$. *Then*

$$(\hat{x} - \bar{x})^T x \geq \hat{x}^T(\hat{x} - \bar{x}) \tag{2.8}$$

is a valid linear inequality for $x \in P_B(C)$ that cuts off $\bar{x}$.

Proof. Follows from Theorem 2.2, since f is differentiable on $P_B(C)$ with $\nabla f(\hat{x}) = \frac{1}{\|\hat{x}-\bar{x}\|_2}(\hat{x} - \bar{x})$. □

Note that the linear inequality (2.8) from Proposition 2.1 is obtained by solving a single SOCP.

2.3. Cut Generation by Application of Duality. In this section results of Çezik and Iyengar presented for conic programming in [3] are investigated and extended.

To derive valid cuts for (1.1) we first state conditions that define valid inequalities for the lifted set $\tilde{M}_B^+(C)$. Later we will show, how valid linear and quadratic cuts in the variable x can be deduced from that.

For the next results, we introduce some additional notation. At first we introduce the inner product of two matrices A, B in $\mathbb{R}^{m,n}$ by $A \bullet B = \sum_{i=1}^m \sum_{j=1}^n A_{ij}B_{ij}$. Furthermore, an upper index k of a vector v or a matrix M v^k or M^k is used to give a name to that vector or matrix and lower indices v_k or $M_{k,j}$ denote the k-th component of a vector v or the (k,j)-th element of a matrix M.

THEOREM 2.3. *Suppose* $int(\text{conv}(C^0)) \neq \emptyset$. *Fix $B \subseteq J$, $B = \{i_1, \dots i_p\}$. Let $V_B^1 = [v_{i_k}^{1j}]_{j,k=1,\dots p}$. Then*

$$Q \bullet V_B^1 + \alpha^T x \geq \beta, \quad Q = Q^T = (q^1, \dots q^p) \in \mathbb{R}^{p,p} \tag{2.9}$$

is valid for all $(x, (v^{0k}, v^{1k}, \lambda^{0k}, \lambda^{1k}) \quad k \in \{1, \dots p\}) \in \tilde{M}_B^+(C)$ if and only if there exist $y^{1,k} \in \mathbb{R}^n, y^2 \in \mathbb{R}^p, y^3 \in \mathbb{R}^p, y^4 \in \mathbb{R}^p, y^{5,k} \in \mathbb{R}^m, y^{6,k} \in \mathbb{R}^m, y^{7,k} \in \mathbb{R}^{\frac{p(p-1)}{2}}$, $s_x \succeq 0$, $s_{v^{0k}}, s_{v^{1k}} \succeq 0$, $s_{\lambda^{0k}}, s_{\lambda^{1k}} \geq 0$, $s_{h_{j1}^{0k}} \geq 0$, $(s_{h_{j2}^{0k}}, s_{h_{j3}^{0k}})^T \succeq 0$, $s_{h_{j1}^{1k}} \geq 0$, $(s_{h_{j2}^{1k}}, s_{h_{j3}^{1k}})^T \succeq 0$ for $j = 1, \dots p$, $j \neq k$, $k \in \{1, \dots p\}$ and symmetric $S^6 \in \mathbb{R}^{p+1,p+1}$, $S^6 \succeq_{sd} 0$ satisfying

$$-\sum_{k=1}^p y^{1,k} + (e_{i_1}^n, \dots e_{i_p}^n, 0^n)(S_{p+1,\cdot}^6)^T + (e_{i_1^n}, \dots e_{i_p^n}, 0^n)S_{\cdot,p+1}^6 + s_x = \alpha, \tag{2.10}$$

$$I^n y^{1,k} + y_k^3 e^n_{i_k} + A^T y^{5,k} - (s_{h^{0k}_{j1}} e^n_{i_j} + s_{h^{0k}_{j3}} e^n_{i_j})_{j=1,\dots p, j\neq k} + s_{v^{0k}} = 0, \quad (2.11)$$

$$\begin{aligned}
&\text{for } j = 1, \dots k-1 \\
&y^{1k}_{i_j} + A^T_{i_j} y^{6,k} - y^{7,j}_{k-j} + S^6_{j,k} - s_{h^{1k}_{j1}} - s_{h^{1k}_{j3}} + (s_{v^{1k}})_{i_j} &= q^k_j \\
&\text{for } j = k \\
&y^{1k}_{i_k} + A^T_{i_k} y^{6,k} + y^4_k + S^6_{k,k} + (s_{v^{1k}})_{i_k} &= q^k_k \\
&\text{for } j = k+1, \dots p \\
&y^{1k}_{i_j} + A^T_{i_j} y^{6,k} + y^{7,k}_{j-k} + S^6_{j,k} - s_{h^{1k}_{j1}} - s_{h^{1k}_{j3}} + (s_{v^{1k}})_{i_j} &= q^k_j \\
&\text{for } j = p+1, \dots n: \\
&y^{1k}_{i_j} + A^T_{i_j} y^{6,k} + (s_{v^{1k}})_{i_j} &= 0
\end{aligned} \quad (2.12)$$

$$y^2_k - b^T y^{5,k} - \sum_{j=1, j\neq k}^{p} s_{h^{0k}_{j2}} + s_{\lambda^{0k}} = 0, \quad (2.13)$$

$$y^2_k - y^4_k - b^T y^{6,k} - \sum_{j=1, j\neq k}^{p} s_{h^{1k}_{j2}} + s_{\lambda^{1k}} = 0, \quad (2.14)$$

$$\sum_{k=1}^{p} y^2_k - S^6_{p+1,p+1} - \beta = 0, \quad (2.15)$$

where 0^n is the zero column vector in $\mathbb{R}^n$, I^n is the identity matrix in $\mathbb{R}^{n,n}$ and $e^n_{i_j}$ is the i_j-th unit vector in $\mathbb{R}^n$.

Proof. We investigate the problem

$$\begin{aligned}
\min \quad & Q \bullet V^1_B + \alpha^T x \\
s.t. \quad & (x, (v^{0k}, v^{1k}, \lambda^{0k}, \lambda^{1k})\ k = 1 \dots p) \in \tilde{M}^+_B(C)
\end{aligned} \quad (2.16)$$

that has linear constraints, conic constraints and boundary constraints of the form $v \in [0, \lambda]$. We introduce nonnegative auxiliary variables to rewrite this boundary constraints as linear constraints and gain thus a standard conic programming problem. The dual feasibility conditions of this problem comply with conditions (2.10)-(2.14) and condition (2.15) sets the dual objective value to β.

Due to assumption $int(\text{conv}(C^0)) \neq \emptyset$, we can conclude that $int(\tilde{M}^+_B(C)) \neq \emptyset$. Thus, the feasible set of the primal problem has non-empty interior. We can conclude immediately that every dual feasible point with objective value β, that is a point satisfying (2.10)-(2.15), provides a lower bound on the primal objective – compare [13].

For the other direction, assume (2.9) holds and thus the primal objective value is bounded below by β. Then we can deduce that the dual problem is solvable. Now we can show that the dual objective value is unbounded below. From here we can deduce with continuity of the objective and convexity of the feasible set that for every β between $-\infty$ and the smallest primal objective value, we can find a dual feasible point with objective value β, that is a point satisfying (2.10)-(2.15). A detailed proof is given in Drewes' [23] $\square$

Remark: Apart from the restriction to SOCP and some technicalities, the last theorem equates to Theorem 2 in [3] by Çezik and Iyengar. One important difference is that we did not assume the relaxed binary conditions to be present in our problem formulation $Ax = b, x \succeq 0$. Indeed, the implication $int(C^0) \neq \emptyset \Rightarrow \tilde{M}_B^+(C) \neq \emptyset$ holds only under that technically important assumption compare Lemma 2.1.7 in [23] for details.

Due to Theorem 2.3, conditions (2.10)-(2.15) and the semidefinite and second order cone conditions define the valid inequality (2.9) in the variables (x, V_B^1) for the lifted set $\tilde{M}_B^+(C)$. The same statement is true for the lifted set $\tilde{M}_B(C)$ when conditions (2.10)-(2.15) are satisfied with $S^6 = 0$.

PROPOSITION 2.2. *Suppose* $int(\text{conv}(C^0)) \neq \emptyset$. *Fix* $B \subseteq J$, $B = \{i_1, \ldots i_p\}$. *Let* $V_B^1 = [v_{i_k}^{1j}]_{j,k=1,\ldots p}$. *Then*

$$Q \bullet V_B^1 + \alpha^T x \geq \beta, \quad Q = Q^T = (q^1, \ldots q^p) \in \mathbb{R}^{p \times p}$$

is valid for all $(x, (v^{0k}, v^{1k}, \lambda^{0k}, \lambda^{1k}) \quad \forall k \in \{1, \ldots p\}) \in \tilde{M}_B(C)$ *if and only if there exist* $y^{1,k} \in \mathbb{R}^n, y^2 \in \mathbb{R}^p, y^3 \in \mathbb{R}^p, y^4 \in \mathbb{R}^p, y^{5,k} \in \mathbb{R}^m, y^{6,k} \in \mathbb{R}^m, y^{7,k} \in \mathbb{R}^{\frac{p(p-1)}{2}}$, $s_x \succeq 0$, $s_{v^{0k}}, s_{v^{1k}} \succeq 0$, $s_{\lambda^{0k}}, s_{\lambda^{1k}} \geq 0$, $s_{h_{j1}^{0k}} \geq 0$, $(s_{h_{j2}^{0k}}, s_{h_{j3}^{0k}})^T \succeq 0$, $s_{h_{j1}^{1k}} \geq 0$, $(s_{h_{j2}^{1k}}, s_{h_{j3}^{1k}})^T \succeq 0$ *for* $j = 1, \ldots p$, $j \neq k$ *for all* $k \in \{1, \ldots p\}$ *and* $S^6 \in \mathbb{R}^{p+1,p+1}$, $S^6 = 0$ *satisfying conditions (2.10)-(2.15).*

Proof. The proof is analogous to the proof of Theorem 2.3 and $\tilde{M}_B(C)$ instead of $\tilde{M}_B^+(C)$. $\square$

In the following we apply Theorem 2.3 and Proposition 2.2 to generate valid cuts for (1.1).

LEMMA 2.1 (Linear and quadratic cut generation). *Let* $int(\text{conv}(C^0)) \neq \emptyset$ *and* $B \subseteq J$.

1) *The inequality* $\alpha^T x \geq \beta$ *is valid for* $P_B(C)$ *if there exist* $(Q = 0, \alpha, \beta)$ *that satisfy conditions (2.10)-(2.15) with* $S^6 = 0$.
2) *The convex quadratic inequality* $x_B^T Q x_B + \alpha^T x \geq \beta$ *is valid for* $P_B^+(C)$, *if* (Q, α, β) *with* $-Q \succeq_{sd} 0$ *satisfy conditions (2.10)-(2.15).*

Proof:

1) Follows straightforward from Proposition 2.2.
2) From $V_B^1 - x_B x_B^T \succeq_{sd} 0$ and $-Q \succeq_{sd} 0$ follows that $-V_B^1 \bullet Q + x_B x_B^T \bullet Q \geq 0$ (cf. [14], Lemma 1.2.3), which is equivalent to $x_B^T Q x_B \geq V_B^1 \bullet Q$. Now, part 2 follows from Theorem 2.3. $\square$

The last lemma is analogous to Lemma 4 from [3], whereas part 1 of the

lemma here is formulated based on Proposition 2.2 instead of Theorem 2.3. For this reason the cut defining conditions (2.10)-(2.15) with $S^6 = 0$ are linear equality conditions and second order cone constraints in the variables y and s. Since α also appears only linearly in (2.10)-(2.15), generating linear cuts can be done by solving a second order cone problem.
To generate deep cuts with respect to a fractional relaxed solution $\bar{x}$ we solve the problem

$$\begin{array}{ll} \min & \alpha^T \bar{x} - \beta \\ s.t. & (Q = 0, \alpha, \beta) \text{ satisfy conditions (2.10)-(2.15) with } S^6 = 0, \\ & \|\alpha\|_2 \leq 1. \end{array} \tag{2.17}$$

If $\bar{x} \notin P_B(C)$, the optimal solution of (2.17) provides a valid linear cut $\alpha^T x - \beta \geq 0$ that is violated by $\bar{x}$.
To generate quadratic cuts we solve the problem

$$\begin{array}{ll} \min & \bar{x}_B^T Q \bar{x}_B + \alpha^T \bar{x} - \beta \\ s.t. & (Q, \alpha, \beta), \text{ satisfy conditions (2.10)-(2.15)} \\ & -Q \succeq_{sd} 0, \\ & \|\alpha\|_2 \leq 1. \end{array} \tag{2.18}$$

Since the columns of Q as well as α and β appear linearly in (2.10)-(2.15), the quadratic cut generating problem (2.18) is a conic program with semidefinite and second order cone constraints. The optimal solution provides a valid cut $x_B^T Q x_B + \alpha^T x - \beta \geq 0$ violated by $\bar{x}$, if $\bar{x} \notin P_B^+(C)$.

Next, we consider diagonal matrices $Q = diag(q_{11}, \dots q_{pp})$, with $q_{ii} \in \mathbb{R}$, $q_{ii} \leq 0$, $(i = 1, \dots p)$. With this choice, we can show that the condition

$$Q \bullet V_B^1 \leq x_B^T Q x_B \tag{2.19}$$

holds for $(x, (v^{0k}, v^{1k}, \lambda^{0k}, \lambda^{1k}) \quad k \in \{1, \dots p\}) \in \tilde{M}_B(C)$.

LEMMA 2.2 (Diagonal quadratic cut generation). *Let $int(\text{conv}(C^0)) \neq \emptyset$ and $B \subseteq J$. The convex quadratic inequality $x_B^T Q x_B + \alpha^T x \geq \beta$ is valid for $P_B(C)$, if (Q, α, β) with $Q = diag(q_{11}, \dots q_{pp})$, $q_{ii} \leq 0$ satisfy conditions (2.10)-(2.15) with $S^6 = 0$.*

Proof. Condition (2.19) is equivalent to

$$\begin{array}{ll} & v_B^{11,T} q_1 + \cdots + v_B^{1p,T} q_p \leq (x_{i_1} x_B)^T q_1 + \cdots + (x_{i_p} x_B)^T q_p \\ \Leftrightarrow & v_{i_1}^{11} q_{11} + \cdots + v_{i_p}^{1p} q_{pp} \leq x_{i_1}^2 q_{11} + \cdots + x_{i_p}^2 q_{pp}. \end{array} \tag{2.20}$$

Since the quadratic terms are positive and $q_{ii} \leq 0 \quad \forall i$, inequality (2.20) is true if $v_{i_k}^{1k} \geq x_{i_k}^2 \quad \forall i = 1, \dots k$. Since $x_{i_k} = v_{i_k}^{1k} + v_{i_k}^{0k}$ and $(v^{0k})_{i_k} = 0$ induce $x_{i_k} = v_{i_k}^{1k}$, the inequality follows from $x_{i_k} \in [0,1]$ $\forall k = 1, \dots p$. □

Therefore, we only have to modify conditions (2.10)-(2.15) with $S^6 = 0$ for

diagonal matrices Q and add the nonnegativity conditions $-q_{ii} \geq 0$ to get cut defining linear and second order cone conditions.
The optimal solution of

$$\begin{array}{ll} \min & \bar{x}_B Q \bar{x}_B + \alpha^T \bar{x} - \beta \\ s.t. & (Q, \alpha, \beta) \text{ satisfy } (2.10) - (2.15) \text{ with } S^6 = 0, \\ & Q_{ij} = 0, i \neq j, \forall i, j = 1, \dots p, \\ & Q_{ii} \leq 0, \forall i = 1, \dots p, \\ & \|\alpha\|_2 \leq 1. \end{array} \tag{2.21}$$

provides the valid quadratic inequality $x_B Q x_B + \alpha^T x - \beta \geq 0$ that is violated by $\bar{x}$, if $\bar{x} \notin P_B(C)$.

3. Branch&Bound based Outer Approximation.

We develop a branch&bound based outer approximation approach as proposed by Bonami et al. in [5] on the basis of Fletcher, Leyffer's [4] and Quesada, Grossmann's [2]. The idea is to iteratively compute integer feasible solutions of a (sub)gradient based linear outer approximation of (1.1) and to tighten this outer approximation by solving nonlinear continuous problems. We introduce the following notations. The objective function gradient c consists of noc part vectors $c_i = (c_{i0}, c_{i1}^T)^T \in \mathbb{R}^{k_i}$, the matrix $A = (A_1, \dots A_{noc})$ consits of noc part matrices $A_i \in \mathbb{R}^{m,k_i}$, and the matrix $I_J = ((I_J)_1, \dots (I_J)_{noc})$ maps x to the integer variables, where $(I_J)_i \in \mathbb{R}^{|J|,k_i}$ is the block of columns of I_J belonging to the i-th cone of dimension k_i.

3.1. Nonlinear Subproblems.

For a given integer configuration x_J^k, we define the nonlinear (SOCP) subproblem

$$\begin{array}{lll} \min & c^T x & \\ s.t. & Ax & = b, \\ & x & \succeq 0, \\ & x_J & = x_J^k. \end{array} \qquad (NLP(x_J^k))$$

We make the following assumptions:

- **A1** The set $\{x : Ax = b, x_J \in [l, u]\}$ is bounded.
- **A2** Every nonlinear subproblem $F(x_J^k)$ or $NLP(x_J^k)$ that is obtained from (1.1) by fixing the integer variables x_J has nonempty interior (Slater constraint qualification).

These assumptions comply with assumptions A1 and A3 made by Fletcher and Leyffer in [4] with the difference that any constraint qualification suffices in their case and we do not assume the constraint functions to be differentiable. Due to that, our convergence analysis requires a constraint qualification that guarantees primal-dual optimality.

Remark: **A2** might be expected as a very strong assumption, since it is violated as soon as a leading cone variable x_{i0} is fixed to zero In that case, all variables belonging to that cone are eliminated in our implementation and the Slater condition may hold now for the reduced problem. Otherwise the algorithm uses another technique to ensure convergence – compare the remark at the end of section 3.4.

3.2. Subgradient Based Linear Outer Approximations. Assume $g : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex and subdifferentiable function on $\mathbb{R}^n$. Then due to the convexity of g, the inequality $g(x) \geq g(\bar{x}) + \xi^T(x - \bar{x})$ holds for all $\bar{x}, x \in \mathbb{R}^n$ and every subgradient $\xi \in \partial g(\bar{x})$ – see for example [15]. Thus, we yield a linear outer approximation of the region $\{x : g(x) \leq 0\}$ applying constraints of the form

$$g(\bar{x}) + \xi^T(x - \bar{x}) \leq 0. \tag{3.1}$$

In the case of (1.1), the feasible region is described by constraints $g_i(x) := -x_{i0} + \|x_{i1}\| \leq 0,\ i = 1, \ldots noc$, where $g_i(x)$ is differentiable on $\mathbb{R}^n \setminus \{x : \|x_{i1}\| = 0\}$ with $\nabla g_i(x_i) = (-1, \frac{x_{i1}^T}{\|x_{i1}\|})$ and subdifferentiable if $\|\bar{x}_{i1}\| = 0$.

LEMMA 3.1. *The convex function $g_i(x_i) := -x_{i0} + \|x_{i1}\|$ is subdifferential in $x_i = (x_{i0}, x_{i1}^T)^T = (a, 0^T)^T$, $a \in \mathbb{R}$, with $\partial g_i((a, 0^T)^T) = \{\xi = (\xi_0, \xi_1^T)^T, \xi_0 \in \mathbb{R}, \xi_1 \in \mathbb{R}^{k_i - 1} : \xi_0 = -1, \|\xi_1\| \leq 1\}$.*

Proof. Follows from the subgradient inequality in $(a, 0^T)^T$. □

The following technical lemma will be used in the subsequent proofs.

LEMMA 3.2. *Assume $\mathcal{K}$ is the second order cone of dimension k and $x = (x_0, x_1^T)^T \in \mathcal{K}, s = (s_0, s_1^T)^T \in \mathcal{K}$ satisfy the condition $x^T s = 0$, then*

1. *$x \in int(\mathcal{K}) \ \Rightarrow \ s = (0, \ldots 0)^T$,*
2. *$x \in bd(\mathcal{K}) \setminus \{0\} \ \Rightarrow \ s \in bd(\mathcal{K})$ and $\exists \gamma \geq 0 : \ s = \gamma(x_0, -x_1^T)$.*

Proof. 1.: Assume $\|x_1\| > 0$ and $s_0 > 0$. Due to $x_0 > \|\bar{x}_1\|$ it holds that $s^T x = s_0 x_0 + s_1^T x_1 > s_0\|x_1\| + s_1^T x_1 \geq s_0\|x_1\| - \|s_1\|\|x_1\|$. Then $x^T s = 0$ can only be true, if $s_0\|x_1\| - \|s_1\|\|x_1\| < 0 \Leftrightarrow s_0 < \|s_1\|$ which contradicts $s \in \mathcal{K}$. Thus, $s_0 = 0 \Rightarrow s = (0, \ldots 0)^T$. If $\|x_1\| = 0$, then $s_0 = 0$ follows directly from $x_0 > 0$.
2.: Due to $\bar{x}_0 = \|\bar{x}_1\|$, we have $s^T x = 0 \Leftrightarrow -s_1^T x_1 = s_0\|x_1\|$. Since $s_0 \geq \|s_1\| \geq 0$ we have $-s_1^T x_1 = s_0\|x_1\| \geq \|x_1\|\|s_1\|$. Cauchy -Schwarz's inequality yields $-s_1^T x_1 = \|x_1\|\|s_1\|$ inducing both $s_1 = -\gamma x_1,\ \gamma \in \mathbb{R}$ and $s_0 = \|s_1\|$. It follows that $-x_1^T s_1 = \gamma x_1^T x_1 \geq 0$. Together with $s_0 = \|s_1\|$ and $\|x_1\| = x_0$ we get that there exists $\gamma \geq 0$, such that $s_1 = (\| - \gamma x_1\|, -\gamma x_1^T)^T = \gamma(x_0, -x_1^T)^T$. □

Using the definitions

$$I_0(\bar{x}) := \{i : \bar{x}_i = (0, \ldots 0)^T\},$$
$$I_a(\bar{x}) := \{i : g_i(\bar{x}) = 0, \bar{x}_i \neq (0, \ldots 0)^T\}$$

we show now, how to choose an appropriate element of the subdifferential $\partial g_i(\bar{x})$ for solutions $\bar{x}$ of $NLP(x_J^k)$.

LEMMA 3.3. *Assume **A1** and **A2**. Let $(\bar{x}, \bar{s}, \bar{y})$ be the primal-dual solution of $NLP(x_J^k)$. Then there exist Lagrange multipliers $\bar{\mu} = -\bar{y}$ and $\bar{\lambda}_i \geq 0$ $(i \in I_0 \cup I_a)$ that solve the KKT conditions in $\bar{x}$ with subgradients*

$$\bar{\xi}_i = \begin{pmatrix} -1 \\ -\frac{\bar{s}_{i1}}{\bar{s}_{i0}} \end{pmatrix}, \text{ if } \bar{s}_{i0} > 0, \ \bar{\xi}_i = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \text{ if } \bar{s}_{i0} = 0 \qquad (i \in I_0(\bar{x})).$$

Proof. A1 and A2 guarantee the existence of such a solution $(\bar{x}, \bar{s}, \bar{y})$ satisfying the primal dual optimality system

$$c_i - (A_i^T, (I_J)_i^T)\bar{y} = \bar{s}_i, \ i = 1, \dots noc, \tag{3.2}$$

$$A\bar{x} = b, \ I_J\bar{x} = x_J^k, \tag{3.3}$$

$$\bar{x}_{i0} \geq \|\bar{x}_{i1}\|, \ \bar{s}_{i0} \geq \|\bar{s}_{i1}\|, \ i = 1, \dots noc, \tag{3.4}$$

$$\bar{s}_i^T\bar{x}_i = 0, \ i = 1, \dots noc. \tag{3.5}$$

Since $NLP(x_J^k)$ is convex and due to A2, there also exist Lagrange multipliers $\mu \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^{noc}$, such that $\bar{x}$ satisfies the KKT-conditions

$$\begin{array}{rll} c_i + (A_i^T, (I_J)_i^T)\mu + \lambda_i\xi_i & = 0, & i \in I_0(\bar{x}), \\ c_i + (A_i^T, (I_J)_i^T)\mu + \lambda_i\nabla g_i(\bar{x}_i) & = 0, & i \in I_a(\bar{x}), \\ c_i + (A_i^T, (I_J)_i^T)\mu & = 0, & i \notin I_0(\bar{x}) \cup I_a(\bar{x}), \end{array} \tag{3.6}$$

We now compare both optimality systems to each other.
First, we consider $i \notin I_0 \cup I_a$. Since $\bar{x}_i \in int(\mathcal{K}_i)$, Lemma 3.2, part 1 induces $\bar{s}_i = (0, \dots 0)^T$. Conditions (3.2) for $i \notin I_0 \cup I_a$ are thus equal to $c_i - (A_i^T, (I_J)_i^T)\bar{y} = 0$ and thus $\bar{\mu} = -\bar{y}$ satisfies the KKT-condition (3.6) for $i \notin I_0 \cup I_A$.
Next we consider $i \in I_a(\bar{x})$, where $x_i \in bd(\mathcal{K}) \setminus \{0\}$. Lemma 3.2, part 2 yields

$$\bar{s}_i = \begin{pmatrix} \|-\gamma\bar{x}_{i1}\| \\ -\gamma\bar{x}_{i1} \end{pmatrix} = \gamma \begin{pmatrix} \bar{x}_{i0} \\ -\bar{x}_{i1} \end{pmatrix} \tag{3.7}$$

for $i \in I_a(\bar{x})$. Inserting $\nabla g_i(\bar{x}) = (-1, \frac{\bar{x}_{i1}^T}{\|\bar{x}_{i1}\|})^T$ for $i \in I_a$ into (3.6) yields the existence of $\lambda_i \geq 0$ such that

$$c_i + (A_i^T, (I_J)_i^T)\mu = \lambda_i \begin{pmatrix} 1 \\ -\frac{\bar{x}_{i1}}{\|\bar{x}_{i1}\|} \end{pmatrix}, \ i \in I_a(\bar{x}). \tag{3.8}$$

Insertion of (3.7) into (3.2) and comparison with (3.8) yields the existence of $\gamma \geq 0$ such that $\bar{\mu} = -\bar{y}$ and $\bar{\lambda}_i = \gamma\bar{x}_{i0} = \gamma\|\bar{x}_{i1}\| \geq 0$ satisfy the KKT-conditions (3.6) for $i \in I_a(\bar{x})$.
For $i \in I_0(\bar{x})$, condition (3.6) is satisfied by $\mu \in \mathbb{R}^m$, $\lambda_i \geq 0$ and subgradients ξ_i of the form $\xi_i = (-1, v^T)^T$, $\|v\| \leq 1$. Since $\bar{\mu} = -\bar{y}$ satisfies (3.6) for $i \notin I_0$, we look for a suitable v and $\lambda_i \geq 0$ satisfying

$c_i - (A_i^T, (I_J)_i^T)\bar{y} = \lambda_i(1, -v^T)^T$ for $i \in I_0(\bar{x})$. Comparing the last condition with (3.2) yields that if $\|\bar{s}_{i1}\| > 0$, then $\lambda_i = \bar{s}_{i0}, -v = \frac{\bar{s}_{i1}}{\bar{s}_{i0}}$ satisfy condition (3.6) for $i \in I_0(\bar{x})$. Since $\bar{s}_{i0} \geq \|\bar{s}_{i1}\|$ we obviously have $\lambda_i \geq 0$ and $\|v\| = \|\frac{\bar{s}_{i1}}{\bar{s}_{i0}}\| = \frac{1}{\bar{s}_{i0}}\|\bar{s}_{i1}\| \leq 1$. If $\|\bar{s}_{i1}\| = 0$, the required condition (3.6) is satisfied by $\lambda_i = \bar{s}_{i0}, -v = (0, \ldots 0)^T$. □

3.3. Infeasibility in Nonlinear Problems. If the nonlinear program $NLP(x_J^k)$ is infeasible for x_J^k, the algorithm solves a feasibility problem of the form

$$\begin{array}{rcl} \min\ u & & \\ s.t.\quad Ax & = & b, \\ -x_{i0} + \|x_{i1}\| & \leq & u,\ i = 1, \ldots noc, \\ u & \geq & 0, \\ x_J & = & x_J^k. \end{array} \qquad F(x_J^k)$$

It has the property that the optimal solution $(\bar{x}, \bar{u})$ minimizes the maximal violation of the conic constraints. One necessity for convergence of the outer approximation approach is the following. If $NLP(x_J^k)$ is not feasible, then the solution of the feasibility problem $F(x_J^k)$ must tighten the outer approximation such that the current integer assignment x_J^k is no longer feasible for the linear outer approximation. For this purpose, we must identify the subgradients at the solution of $F(x_J^k)$, that satisfy the KKT conditions.
We define the index sets of active constraints in a solution $(\bar{x}, \bar{u})$ of $F(x_J^k)$,

$$\begin{array}{rcl} I_F := I_F(\bar{x}) & := & \{i \in \{1, \ldots noc\} : -\bar{x}_{i0} + \|\bar{x}_{i1}\| = \bar{u}\}, \\ I_{F0} := I_{F0}(\bar{x}) & := & \{i \in I_F : \|\bar{x}_{i1}\| = 0\}, \\ I_{F1} := I_{F1}(\bar{x}) & := & \{i \in I_F : \|\bar{x}_{i1}\| \neq 0\}. \end{array} \tag{3.9}$$

LEMMA 3.4. *Assume **A1** and **A2** hold. Let $(\bar{x}, \bar{u})$ solve $F(x_J^k)$ with $\bar{u} > 0$ and let $(\bar{s}, \bar{y})$ be the solution of its dual program. Then there exist Lagrange multipliers $\bar{\mu} = -\bar{y}$ and $\bar{\lambda}_i \geq 0$ $(i \in I_F)$ that solve the KKT conditions in $(\bar{x}, \bar{u})$ with subgradients*

$$\bar{\xi}_i = \begin{pmatrix} -1 \\ -\frac{\bar{s}_{i1}}{\bar{s}_{i0}} \end{pmatrix}, \text{ if } \bar{s}_{i0} > 0,\ \bar{\xi}_i = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \text{ if } \bar{s}_{i0} = 0 \tag{3.10}$$

for $i \in I_{F0}(\bar{x})$.
Proof: Since $F(x_J^k)$ has interior points, there exist Lagrange multipliers $\mu \in \mathbb{R}^m$, $\lambda \geq 0$, such that optimal solution $(\bar{x}, \bar{u})$ of $F(x_J^k)$ satisfies the KKT-conditions

$$A_i^T \mu_A + (I_J)_i^T \mu_J = 0,\ i \notin I_F, \tag{3.11}$$

$$\nabla g_i(\bar{x}_i)\lambda_{g_i} + A_i^T \mu_A + (I_J)_i^T \mu_J = 0,\ i \in I_{F1}, \tag{3.12}$$

$$\xi_i \lambda_{g_i} + A_i^T \mu_A + (I_J)_i^T \mu_J = 0,\ i \in I_{F0}, \tag{3.13}$$

$$\sum_{i \in I_F} (\lambda_g)_i = 1, \tag{3.14}$$

with $\xi_i \in \partial g_i(\bar{x}_i)$ plus the feasibility conditions, where we already used the complementary conditions for $\bar{u} > 0$ and the inactive constraints. Due to the nonempty interior of $F(x_J^k)$, $(\bar{x}, \bar{u})$ satisfies also the primal-dual optimality system

$$Ax = b,$$
$$u \geq 0,$$
$$-A_i^T y_A - (I_J^T)_i y_J = s_i,\ i = 1, \ldots noc, \tag{3.15}$$
$$x_{i0} + u \geq \|\bar{x}_{i1}\|,\ \sum_{i=1}^{noc} s_{i0} = 1, \tag{3.16}$$
$$s_{i0} \geq \|\bar{s}_{i1}\|,\ i = 1, \ldots noc, \tag{3.17}$$
$$s_{i0}(x_{i0} + u) + s_{i1}^T x_{i1} = 0,\ i = 1, \ldots noc, \tag{3.18}$$

where we again used complementarity for $\bar{u} > 0$.
First we investigate $i \notin I_F$, where $\bar{x}_{i0} + \bar{u} > \|\bar{x}_{i1}\|$ inducing $s_i = (0, \ldots 0)^T$ (cf. Lemma 3.2, part 1). Thus, the KKT conditions (3.11) are satisfied by $\mu_A = -y_A$ and $\mu_J = -y_J$.
Next, we consider $i \in I_{F1}$ for which by definition $\bar{x}_{i0} + \bar{u} = \|\bar{x}_{i1}\| > 0$ holds. Applying Lemma 3.2, part 2 yields there exists $\gamma \geq 0$ with $s_{i1} = -\gamma \bar{x}_{i1}$. Insertion into (3.15) yields

$$-A_i^T y_A - (I_J)_i y_J + \gamma \|\bar{x}_{i1}\| \begin{pmatrix} -1 \\ \frac{\bar{x}_{i1}}{\|\bar{x}_{i1}\|} \end{pmatrix} = 0,\ i \in I_{F1}.$$

Since $\nabla g_i(\bar{x}_i) = (-1, \frac{\bar{x}_{i1}^T}{\|\bar{x}_{i1}\|})^T$, we obtain that the KKT-condition (3.12) is satisfied by $\mu_A = -y_A$, $\mu_J = -y_J$ and $\lambda_i = s_{i0} = \gamma \|\bar{x}_{i1}\| \geq 0$.
Finally, we investigate $i \in I_{F0}$, where $\bar{x}_{i0} + \bar{u} = \|\bar{x}_{i1}\| = 0$. Since $\mu_A = -y_A$, $\mu_J = -y_J$ satisfy the KKT-conditions for $i \notin I_{F0}$, we are going to derive a subgradient ξ_i that satisfies (3.13) with that choice. In analogy to Lemma 3.3 from subsection 3.1 we derive that $\xi_{i1} = \frac{-s_{i1}}{s_{i0}}$, if $s_{i0} > 0$ and $\xi_{i1} = 0$ otherwise, are suitable together with $\lambda_i = s_{i0} \geq 0$. Due to $\lambda_i = s_{i0}$ for all $i \in I_F$, (3.16) yields, that the last KKT condition (3.14) is satisfied by this choice, too. □

Every subgradient ξ of $g_i(\bar{x}) - \bar{u}$ with respect to $\bar{x}$ provides a subgradient $(\xi^T, -1)^T$ of $g_i(\bar{x}) - \bar{u}$ with respect to $(\bar{x}, \bar{u})$ and thus an inequality $g_i(\bar{x}) + \xi^T(x - \bar{x}) \leq 0$ that is valid for the feasible region of (1.1). The next lemma states that the subgradients (3.10) of Lemma 3.4 together with the gradients of the differentiable functions g_i in the solution of $F(x_J^k)$ provide inequalities that separate the last integer solution.

LEMMA 3.5. *Assume **A1** and **A2** hold. If $NLP(x_J^k)$ is infeasible and thus $(\bar{x}, \bar{u})$ solve $F(x_J^k)$ with positive optimal value $\bar{u} > 0$, then every x satisfying the linear equalities $Ax = b$ with $x_J = x_J^k$, is infeasible in the* *constraints*

$$\begin{array}{rcl} -x_{i0} + \frac{\bar{x}_{i1}^T}{\|\bar{x}_{i1}\|} x_{i1} & \leq & 0, \ i \in I_{F1}(\bar{x}), \\ -x_{i0} - \frac{\bar{s}_{i1}^T}{\bar{s}_{i0}} x_{i1} & \leq & 0, \ i \in I_{F0}, \bar{s}_{i0} \neq 0, \\ -x_{i0} & \leq & 0, \ i \in I_{F0}, \bar{s}_{i0} = 0, \end{array} \tag{3.19}$$

where I_{F1} and I_{F0} are defined by (3.9) and $(\bar{s}, \bar{y})$ is the solution of the dual program of $F(x_J^k)$. **Proof**: The proof is done in analogy to Lemma 1 in [4]. Due to assumption **A1** and **A2**, the optimal solution of $F(x_J^k)$ is attained. We further know from Lemma 3.4, that there exist $\lambda_{g_i} \geq 0$, with $\sum_{i \in I_F} \lambda_{g_i} = 1$, μ_A and μ_J satisfying the KKT conditions

$$\sum_{i \in I_{F1}} \nabla g_i(\bar{x}) \lambda_{g_i} + \sum_{i \in I_{F0}} \xi_i^n \lambda_{g_i} + A^T \mu_A + I_J^T \mu_J = 0 \tag{3.20}$$

in $\bar{x}$ with subgradients (3.10). To show the result of the lemma, we assume now that x, with $x_J = x_J^k$, satisfies conditions (3.19) which are equivalent to

$$\begin{array}{r} g_i(\bar{x}) + \nabla g_i(\bar{x})^T (x - \bar{x}) \leq 0, \ i \in I_{F1}(\bar{x}), \\ g_i(\bar{x}) + \xi_i^{n,T} (x - \bar{x}) \leq 0, \ i \in I_{F0}(\bar{x}). \end{array}$$

We multiply the inequalities by $(\lambda_g)_i \geq 0$ and add all inequalities. Since $g_i(\bar{x}) = \bar{u}$ for $i \in I_F$ and $\sum_{i \in I_F} \lambda_{g_i} = 1$ we get

$$\sum_{i \in I_{F1}} (\lambda_{g_i} \bar{u} + \lambda_{g_i} \nabla g_i(\bar{x})^T (x - \bar{x})) + \sum_{i \in I_{F0}} (\lambda_{g_i} \bar{u} + \lambda_{g_i} \xi_i^{n,T} (x - \bar{x})) \leq 0$$

$$\Leftrightarrow \qquad \bar{u} + \left(\sum_{i \in I_{F1}} \lambda_{g_i} \nabla g_i(\bar{x}) + \sum_{i \in I_{F0}} (\lambda_{g_i} \xi_i^n) \right)^T (x - \bar{x}) \leq 0.$$

Insertion of (3.20) yields

$$\begin{array}{rcl} & \bar{u} + (-A^T \mu_A - I_J^T \mu_J)^T (x - \bar{x}) & \leq 0 \\ \Leftrightarrow^{Ax = A\bar{x} = b} & \bar{u} - \mu_J^T (x_J - \bar{x}_J) & \leq 0 \\ \Leftrightarrow^{x_J = x_J^k = \bar{x}_J} & \bar{u} & \leq 0. \end{array}$$

This is a contradiction to the assumption $\bar{u} > 0$.□

Thus, the solution $\bar{x}$ of $F(x_J^k)$ produces new constraints (3.19) that strengthen the outer approximation such that the integer solution x_J^k is no longer feasible. If $NLP(x_J^k)$ is infeasible, the active set $I_F(\bar{x})$ is not empty and thus, at least one constraint (3.19) can be added.
Let $T \subset \mathbb{R}^n$ contain solutions of nonlinear subproblems $NLP(x_J^k)$ and

$S \subset \mathbb{R}^n$ contains solutions of feasibility problems $F(x_J^k)$. Using the subgradients from Lemma 3.5 and 3.4 we build the linear outer approximation problem

$$\begin{array}{rrcl} & \min\ c^T x & & \\ s.t. & Ax & = & b \\ & c^T x & < & c^T\bar{x},\ \bar{x} \in T, \\ & -\|\bar{x}_{i1}\| x_{i0} + \bar{x}_{i1}^T x_{i1} & \le & 0,\ i \in I_a(\bar{x}),\ \bar{x} \in T, \\ & -\|\bar{x}_{i1}\| x_{i0} + \bar{x}_{i1}^T x_{i1} & \le & 0,\ i \in I_{F1}(\bar{x})\ \bar{x} \in S, \\ & -x_{i0} & \le & 0,\ i \in I_0(\bar{x}),\ \bar{s}_{i0} = 0,\ \bar{x} \in T, \\ & -x_{i0} - \frac{1}{\bar{s}_{i0}} \bar{s}_{i1}^T x_{i1} & \le & 0,\ i \in I_0(\bar{x}),\ \bar{s}_{i0} > 0,\ \bar{x} \in T, \\ & -x_{i0} - \frac{\bar{s}_{i1}^T}{\bar{s}_{i0}} x_{i1} & \le & 0,\ i \in I_{F0}(\bar{x}), \bar{s}_{i0} \neq 0,\ \bar{x} \in S, \\ & -x_{i0} & \le & 0,\ i \in I_{F0}(\bar{x}), \bar{s}_{i0} = 0,\ \bar{x} \in S, \\ & x_j & \in & [l_j, u_j],\ (j \in J) \\ & x_j & \in & \mathbb{Z},\ (j \in J). \end{array} \qquad \text{(OA(T,S))}$$

3.4. The Algorithm. We define nodes N^k consisting of lower and upper bounds on the integer variables that can be interpreted as branch&bound nodes for (1.1) as well as $OA(T,S)$. Let $(MISOC^k)$ denote the mixed integer SOCP defined by the bounds of N^k and $OA^k(T,S)$ its MILP outer approximation with continuous relaxation $(\widetilde{MISOC^k})$ and $\widetilde{OA^k}(T,S)$. The following hybrid algorithm integrates branch&bound and the outer approximation approach as proposed by Bonami et al. in [5] for general differentiable MINLPs.

Algorithm 1 HYBRID OA/B-A-B FOR (1.1)

Input: Problem (1.1)
Output: Optimal solution x^* or indication of infeasibility
Initialization: $CUB := \infty$, solve $(\widetilde{MISOC})$ with solution x^0,
 if $((\widetilde{MISOC})$ infeasible) STOP, problem infeasible
 else set $S = \emptyset$, $T = \{x^0\}$ and solve MILP $OA(T)$.
1. **if** $(OA(T)$ infeasible) STOP, problem infeasible
 else solution $x^{(1)}$ found:
  if $(NLP(x_J^{(1)})$ feasible)
   compute solution $\bar{x}$ of $NLP(x^{(1)})$, $T := T \cup \{\bar{x}\}$,
   if $(c^T\bar{x} < CUB)$ $CUB = c^T\bar{x}$, $x^* = \bar{x}$ **endif**.
  else compute solution $\bar{x}$ of $F(x_J^{(1)})$, $S := S \cup \{\bar{x}\}$.
 $Nodes := \{N^0 = (lb^0 = l, ub^0 = u)\}$, $ll := 0$, $L := 10$, $i := 0$
2. **while** $Nodes \neq \emptyset$ **do** select N^k from $Nodes$, $Nodes := Nodes \setminus N^k$
 2a. **if** $(ll = 0 \mod L)$ solve $\widetilde{MISOC^k}$
  if $(\widetilde{MISOC^k}$ feasible): solution $\bar{x}$, $T := T \cup \{\bar{x}\}$
   if $(\bar{x}_J$ integer):
    if $(c^T\bar{x} < CUB)$ $CUB = c^T\bar{x}$, $x^* = \bar{x}$

go to 2.
else go to 2.
2b. solve $\widetilde{OA^k}(T,S)$ with solution x^k
while ($\widetilde{OA^k}(T,S)$ feasible) & (x^k_J integer) & ($c^T x^k < CUB$)
if ($NLP(x^k_J)$ is feasible with solution $\bar{x}$) $T := T \cup \{\bar{x}\}$
if ($c^T \bar{x} < CUB$) $CUB = c^T \bar{x}$, $x^* = \bar{x}$
else solve $F(x^k_J)$ with solution $\bar{x}$, $S := S \cup \{\bar{x}\}$
compute solution x^k of updated $\widetilde{OA^k(T,S)}$
2c. **if** ($c^T x^k < CUB$) branch on variable $x^k_j \notin \mathbb{Z}$,
create $N^{i+1} = N^k$, with $ub^{i+1}_j = \lfloor x^k_j \rfloor$,
create $N^{i+2} = N^k$, with $lb^{i+2}_j = \lceil x^k_j \rceil$,
set $i = i+2$, $ll = ll+1$.

Note that if $L = 1$, then step 2 performs a nonlinear branch&bound search. If $L = \infty$ Algorithm 1 resembles a branch&bound based outer approximation algorithm. Convergence of the outer approximation approach in case of continuously differentiable constraint functions was shown in [4], Theorem 2. Convergence of Algorithm 1 is stated in the next theorem.

THEOREM 3.1. *Assume **A1** and **A2**. Then the outer approximation algorithm terminates in a finite number of steps at an optimal solution of (1.1) or with the indication, that it is infeasible.*

Proof. We show that no integer assignment x^k_J is generated twice by showing that $x_J = x^k_J$ is infeasible in the linearized constraints created in the solutions of $NLP(x^k_J)$ or $F(x^k_J)$. The finiteness follows then from the boundedness of the feasible set. A1 and A2 guarantee the solvability, presence of KKT conditions and primal-dual optimality of the nonlinear subproblems $NLP(x^k_J)$ and $F(x^k_J)$. Lemma 3.5 yields thus the result for $F(x^k_J)$.
It remains to consider the case, when $NLP(x^k_J)$ is feasible with solution $\bar{x}$. Assume $\tilde{x}$, with $\tilde{x}_J = \bar{x}_J$ is the optimal solution of $OA(\widetilde{T \cup \{\bar{x}\}}, S)$. Then

$$\begin{aligned}
c^T_{\bar{J}}\tilde{x}_{\bar{J}} + c^T_J \bar{x}_J < c^T_{\bar{J}}\bar{x}_{\bar{J}} + c^T_J \bar{x}_J, \Leftrightarrow\ c^T_{\bar{J}}\tilde{x}_{\bar{J}} &< c^T_{\bar{J}}\bar{x}_{\bar{J}} && (3.21)\\
(\nabla g_i(\bar{x}))^T_{\bar{J}}(\tilde{x}_{\bar{J}} - \bar{x}_{\bar{J}}) &\le 0,\ i \in I_a(\bar{x}), && (3.22)\\
(\bar{\xi}_i)^T_{\bar{J}}(\tilde{x}_{\bar{J}} - \bar{x}_{\bar{J}}) &\le 0,\ i \in I_0(\bar{x}), && (3.23)\\
A_{\bar{J}}(\tilde{x}_{\bar{J}} - \bar{x}_{\bar{J}}) &= 0, && (3.24)
\end{aligned}$$

must hold with $\bar{\xi}_i$ from Lemma 3.3. Due to A2 we know that there exist $\mu \in \mathbb{R}^m$ and $\lambda \in \mathbb{R}^{|I_0 \cup I_a|}_+$ satisfying the KKT conditions (3.6) of $NLP(x^k_J)$ in $\bar{x}$, that is

$$\begin{array}{lll}
-c_i &= A^T_i \mu + \lambda_i \xi_i, & i \in I_0(\bar{x}),\\
-c_i &= A^T_i \mu + \lambda_i \nabla g_i(\bar{x}), & i \in I_a(\bar{x}),\\
-c_i &= A^T_i \mu, & i \notin I_0(\bar{x}) \cup I_a(\bar{x})
\end{array} \qquad (3.25)$$

with the subgradients $\bar{\xi}_i$ chosen from Lemma 3.3. Farkas' Lemma (cf. [16]) states that (3.25) is equivalent to the fact that that as long as $(\tilde{x} - \bar{x})$ satisfies (3.22) - (3.24), then $c_{\bar{J}}^T(\tilde{x}_{\bar{J}} - \bar{x}_{\bar{J}}) \geq 0 \Leftrightarrow c_{\bar{J}}^T\tilde{x}_{\bar{J}} \geq c_{\bar{J}}^T\bar{x}_{\bar{J}}$ must hold, which is a contradiction to (3.21). □

Version without Slater condition. Assume N^k is a node such that A2 is violated by $\widetilde{NLP}(x_J^k)$ and assume x_J^k is feasible for the updated outer approximation $\widetilde{OA}^k(T \cup \{\bar{x}\}, S)$.Then the inner while-loop in step 2b becomes infinite and Algorithm 1 does not converge. In the implementation we detect, whenever this situation occurs by checking, if an integer assignment is generated twice. In that case, the outer approximation approach is not working for the node N^k and we solve the SOCP relaxation $(\widetilde{MISOC^k})$ instead. If that problem is not infeasible and has no integer feasible solution, we branch on the solution of this SOCP relaxation to explore the subtree of N^k. For details of this strategy see Section 4.5 in [23].

4. Numerical results. We implemented a pure branch&bound algorithm ('B&B'), a classical branch&cut approach ('B&C') as well as the outer approximation approach Algorithm 1 ('B&B-OA'). Thereby each presented cutting technique was applied seperately. The suffix behind the name of the solver specifies the applied cutting technique, where 'Linear' solves cut generating problem (2.17), 'SOC Quad' solves cut generating problem (2.21), 'SDP Quad' solves cut generating problem (2.18) and 'Subgrad' solves the minimum distance problem from Proposition 2.1. The SOCP problems are solved with our own implementation of an infeasible primal-dual interior point approach (cf. [23], Chapter 1), the linear programs are solved with CPLEX 10.0.1 and the cut SDPs are solved using Sedumi [18].

First, we report our results for mixed 0-1 formulations of nine different ESTP test problems ($n = 58/114, m = 41/79, noc = 40/78, |J| = 9/18$) from Beasley's website [17]. Each ESTP problem was tested in combination with the depth first search and the best bound first node selection strategy and three different branching rules (most fractional branching, combined fractional branching and pseudocost branching). The resulting 54 test instances were tested with nonlinear branch&bound and branch&cut where we applied five cutting loops in the root node. We tested Algorithm 1 on these instances – without cuts and with one cut generation in every occuring SOCP relaxation.

For each algorithm we display the number of solved SOCP nodes and LP nodes needed to solve all test instances, the percentage to which this number is reduced by the specified cut (see 'Node Recution to') and the minimal reduction that was achieved for at least one problem instance (see 'Minimal Reduction to'). Furthermore we show the number of test instances reduced by the applied cutting technique. As displayed in Table 1, in combination with branch&cut, lift-and-project cuts reduce the number of

Solver	Nodes (SOCP)	Node Reduction to (%)	Minimal Reduction to (%)	Reduced problems (%)
B&B	12979	-	-	-
B&C Linear	8414	64.83	6.22	85.19
B&C SOC Quad	8414	64.83	6.22	85.19
B&C SDP Quad	11741	90.46	41.79	11.11
B&C Subgrad	11349	87.44	41.53	68.52

Table 1
B&C for ESTP Problems

Solver	Nodes (SOCP/LP)	Node Reduction to (%)	Minimal Reduction to (%)	Reduced problems (%)
B&B-OA	3927 / 15455	-	-	
B&B-OA Linear	3956 / 15484	100.30	71.43	9.26
B&B-OA SOC Quad	3956 / 15484	100.30	71.43	9.26
B&B-OA SDP Quad	3615 / 14156	91.69	52.15	50.00
B&B-OA Subgrad	3757 /13748	90.32	55.93	62.96

Table 2
B&B-OA for ESTP Problems

solved nodes down to between 64.83% and 90.46 % for all instances and down to 6.22% for single test instances. Thereby, the linear and quadratic cuts based on SOCP problems reduce the search trees of most of the problems and lead to the best reductions. Although the SDP based quadratic cuts have the tightest underlying relaxation, these cuts do not achieve the best reductions, which is different, when cuts are generated in every node of the search tree. In that case SDP based cuts achieve the best minimal reductions. Due to the high computational costs of this approach we do not discuss it further at this point. Table2 shows that in the context of Algorithm 1 reductions of the search trees are achieved by the subgradient based and SDP based quadratic cuts and also for single instances with the SOCP based linear and quadratic dual cuts, which lead to a small increase of the total number of nodes with respect to all ESTP test instances. Since the gut generating problems are high-dimensional SOCP problems the ob-

	B&B/ B&C	B&B-OA
SOCP- Nodes	391/391	54
LP- Nodes	-	780
Time in sec. Wallclock (CPU)	964 (275)	196 (55)

TABLE 3
Balancing Problem

served recuctions of solved nodes do not lead necessarily to a decrease of the running time.
The algorithms were also applied to several engineering problems arising in the area of turbine balancing. Table 3 reports the results achieved by the different algorithms for such a problem ($n = 212, m = 145, noc = 153, |J| = 56$). For these kind of problems, application of cuts only in the root node does not lead to any reduction, whereas applying one cut in every node achieves reductions, but that becomes very expensive. A comparison of the branch&cut approach and Algorithm 1 on the basis of Tables 1 to 3 shows, that the latter algorithm solves remarkable fewer SOCP problems. We observed for almost all test instances that the branch&bound based outer approximation approach is preferable regarding running times, since the LP problems stay moderately in size since only linearizations of active constraints are added. Thus also the balancing problems are solved in moderate running times.

5. Summary. We presented different cutting techniques based on lift-and-project relaxation of the feasible region of mixed 0-1 SOCPs as well as a convergent branch&bound based outer approximation approach using subgradient based linearizations. We presented numerical results for some application problems. The impact of the different cutting techniques in a classical branch and cut framework and the outer approximation algorithm was investigated. A comparison of the algorithms showed that the outer approximation approach solves alomst all problems in signficantly shorter running time.

REFERENCES

[1] ROBERT A. STUBBS AND SANJAY MEHROTRA, *A branch-and-cut method for 0-1 mixed convex programming* in Mathematical Programming, 1999, 86: pp.515-532

[2] I. QUESADA AND I.E. GROSMANN, *An LP/NLP based Branch and Bound Algorithm for Convex MINLP Optimization Problems*, in Computers and Chemical Engineering, 1992, 16:(10,11) pp. 937-947

[3] M.T. ÇEZIK AND G. IYENGAR, *Cuts for Mixed 0-1 Conic Programming*, in Mathematical Programming, Ser. A, 2005, 104: pp. 179-200

[4] Roger Fletcher and Sven Leyffer, *Solving Mixed Integer Nonlinear Programs by Outer Approximation*, in Mathematical Programming, 1994, 66: pp. 327-349.

[5] P. Bonami and L.T. Biegler and A.R.Conn and G. Cornuejols and I.E. Grossmann and C.D.Laird and J. Lee and A.Lodi and F.Margot and N.Sawaya and A. Wchter , *An Algorithmic Framework for Convex Mixed Integer Nonlinear Programs*, IBM Research Division, New York, 2005

[6] Robert A. Stubbs and Sanjay Mehrotra, *Generating Convex Polynomial Inequalities for Mixed 0-1 Programs*, Journal of global optimization, 2002, 24: pp. 311-332

[7] Juan Pablo Vielma and Shabbir Ahmed and George L. Nemhauser, *A Lifted Linear Programming Branch-and-Bound Algorithm for Mixed Integer Conic Quadratic Programs*, INFORMS Journal on Computing, 2008,20(3): pp. 438-450

[8] Alper Atamtürk and Vishnu Narayanan, *Cuts for Conic Mixed-Integer Programming*, Mathematical Programming, Ser. A, DOI 10.1007/s10107-008-0239-4, 2007

[9] Aharon Ben-Tal and Arkadi Nemirovski, *On Polyhedral Approximations of the Second-Order Cone*, in Mathematics of Operations Research, 2001, 26(2):pp. 193–205

[10] Egon Balas and Sebastián Ceria and Gérard Cornuéjols, *A lift-and-project cutting plane algorithm for mixed 0-1 programs*, in Mathematical Programming, 1993, 58: pp. 295-324

[11] Marcia Fampa and Nelson Maculan, *A new relaxation in conic form for the Euclidian Steiner Tree Problem in* $\mathbb{R}^n$, in RAIRO Operations Research, 2001,35: pp. 383-394

[12] Dimitris Bertsimas and Romy Shioda, *Algorithm for cardinality-constrained quadratic optimization*, in Computational Optimization and Applications, 2007,91: pp. 239-269

[13] Yurii Nesterov and Arkadii Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, 2001

[14] Christoph Helmberg, *Semidefinite Programming for Combinatorial Optimization*, Konrad-Zuse-Zentrum fr Informationstechnik, 2000, Berlin, Habilitationsschrift

[15] R. Tyrrell Rockafellar, *Convex Analysis*, Princeton University Press, 1970

[16] Carl Geiger and Christian Kanzow, *Theorie und Numerik restringierter Optimierungsaufgaben*, Springer Verlag Berlin Heidelberg New York, 2002

[17] John E. Beasley, *OR Library: Collection of test data for Euclidean Steiner Tree Problems*, howpublished = `http://people.brunel.ac.uk/ mastjjb/jeb/orlib/esteininfo.html`,

[18] Jos F. Sturm, *SeDuMi*, `http://sedumi.ie.lehigh.edu/`

[19] Pietro Belotti and Pierre Bonami and John J. Forrest and Lazlo Ladanyi and Carl Laird and Jon Lee and Francois Margot and Andreas Wächter, *BonMin*, `http://www.coin-or.org/Bonmin/`

[20] Roger Fletcher and Sven Leyffer, *User Manual of filterSQP*,`http://www.mcs.anl.gov/ leyffer/papers/SQP_manual.pdf`

[21] Carl Laird and Andreas Wächter, *IPOPT*, `https://projects.coin-or.org/Ipopt`

[22] Kumar Abhishek and Sven Leyffer and Jeffrey T. Linderoth, *FilMINT: An Outer Approximation-Based Solver for Nonlinear Mixed Integer Programs*, Argonne National Laboratory, Mathematics and Computer Science Division,2008

[23] Sarah Drewes, *Mixed Integer Second Order Cone Programming*, PhD Thesis, submitted April, 2009