Strong formulations for convex functions over nonconvex sets

Daniel Bienstock and Alexander Michalka
Columbia University
New York

November, 2011

version Wed.Oct.10.215343.2012

1 Introduction

In this paper we derive strong linear inequalities for systems representing convex quadratics over nonconvex sets, and we present, in several cases, convex hull characterizations by polynomially separable linear inequalities in the original space. A class of examples we consider is of the form

$$\{ (x, q) \in \mathbb{R}^d \times \mathbb{R} \ : \ q \geq Q(x), \quad x \in \mathbb{R}^d - \text{int}(P) \},$$

where $Q(x) : \mathbb{R}^d \to \mathbb{R}$ is a positive definite quadratic function, $P \subset \mathbb{R}^d$ is full-dimensional and convex and "int" denotes interior. Particular cases we consider are those where P is a polyhdedron or an ellipsoid. We similarly characterize sets of the form

$$\{(x, w, z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \ : \ z \geq F(x), \quad w \leq G(x) \}$$

where both F and G are positive definite quadratics.

Preliminaries. Several important classes of optimization problems include nonlinearities in the objective or constraints. Often this results in nonconvexities and a current research thrust addresses the computation of global bounds and exact solution techniques for such problems. The field is not new; one of the earliest results is the characterization of the convex hull of a box-constrained bilinear form $x_1 x_2$ [21], [2]. Recently, some interesting new results in this direction have been obtained [20]. [9] contains a survey. Also see [10], [5], [26], [27].

A frequently used approach has been to borrow ideas from the field of mixed-integer programming, even when no binary variables are present. The concept of *lifting* arose in (linear) mixed-integer programming [22]. It has also been extended to the continuous setting [11], [15], [19]. Lifting techniques are compelling in that when applicable they provide a computationally practicable way to stengthen valid inequalities. An interesting use of this idea appears in [25], which approximates, using lifted linear inequalities, SDP relaxations of quadratically constrained sets. [7] lifts "tangent" inequalities to approximate multilinear functions.

Our main approach also makes use of lifting. Our contributions are that in each case we characterize the set of nondominated valid linear inequalities for the appropriate region and that we show that these are lifted linear inequalities, which furthermore are efficiently separable, and in the original space of variables. In some cases we obtain closed-form expressions for the lifting coefficients.

Our results focus on quadratics. A great deal of attention has, in fact, been recently focused on problems involving quadratics, and a number of deep results have followed, which provide alternative (but related) methodologies for addressing the problems we consider. See, e.g. [3], [8], [4]. A frequently-applied technique is the Reformulation-Linearization method (RLT) and semidefinite programming extensions. See e.g. [28], [29].

This paper is organized as follows. The polyhedral case is considered in Section 2.3; Section 2.4 addresses the ellipsoidal case. Sections 3 and 3.1 present results for indefinite quadratics. Sections 2, 2.1 and 2.2 introduce some of our general ideas.

2 The positive-definite case

We consider sets of the form

$$S \doteq \{ (x,q) \in \mathbb{R}^d \times \mathbb{R} \,:\, q \ge Q(x), \quad x \in \mathbb{R}^d - \mathrm{int}(P) \}, \tag{1}$$

where $Q(x) : \mathbb{R}^d \to \mathbb{R}$ is a *positive-definite* quadratic function, and each connected component of $P \subset \mathbb{R}^d$ is a homeomorph of either a half-plane or a ball. Thus, each connected component of P is a closed set with nonempty interior.

Since $Q(x)$ is positive definite, we may assume without loss of generality that $Q(x) = \|x\|^2$ (achieved via a linear transformation). For any $y \in \mathbb{R}^d$, the linearization inequality

$$q \;\ge\; 2y^T(x-y) + \|y\|^2 \;=\; 2y^T x - \|y\|^2 \tag{2}$$

is valid for all $(x,q) \in \mathbb{R}^d \times \mathbb{R}$. We seek ways of making this inequality stronger.

Definition 2.1 *Given $\mu \in \mathbb{R}^d$ and $R \ge 0$, we write $\mathcal{B}(\mu, R) = \{ x \in \mathbb{R}^d : \|x-\mu\| \le R \}$.*

2.1 Geometric characterization

Let $x \in \mathbb{R}^d$. Then $x \in \mathbb{R}^d - \mathrm{int}(P)$ if and only if

$$\|x-\mu\|^2 \ge \rho, \qquad \text{for each ball } \mathcal{B}(\mu, \sqrt{\rho}) \subseteq P. \tag{3}$$

In terms of our set S, we can rewrite (3) as

$$q \ge 2\mu^T x - \|\mu\|^2 + \rho, \qquad \text{for each ball } \mathcal{B}(\mu, \sqrt{\rho}) \subseteq P. \tag{4}$$

On the other hand, suppose

$$\delta q \;-\; 2\beta^T x \;\ge\; \beta_0 \tag{5}$$

is valid for S. Since $\mathbb{R}^d - P$ contains points with arbitrarily large norm it follows $\delta \ge 0$. Suppose that $\delta > 0$: then without loss of generality $\delta = 1$. Further, given $x \in \mathbb{R}^d$, (5) is satisfied by (x,q) with $q \ge \|x\|^2$ if and only if it is satisfied by $(x, \|x\|^2)$, and so if and only if we have

$$\|x-\beta\|^2 \ge \;\; \|\beta\|^2 + \beta_0. \tag{6}$$

Since (5) is valid for S, we have that (6) holds for each $x \in \mathbb{R}^d - \mathrm{int}(P)$. Assuming further that (5) is not trivial, that is to say, it is violated by some $(z, \|z\|^2)$ with $z \in \mathrm{int}(P)$, we must therefore have that $\|\beta\|^2 + \beta_0 > 0$ and $\mathcal{B}(\beta, \sqrt{\|\beta\|^2 + \beta_0}) \subseteq P$, i.e. statement (6) is an example of (3). Below we discuss several ways of sharpening these observations.

2.2 Lifted first-order cuts

Let $y \in \partial P$. Then we can always find a ball $\mathcal{B}(\mu, \sqrt{\rho}) \subseteq P$ such that $\|\mu - y\|^2 = \rho$, possibly by setting $\mu = y$ and $\rho = 0$.

Definition 2.2 *Given $y \in \partial P$, we say P is locally flat at y if there is a ball $\mathcal{B}(\mu, \sqrt{\rho}) \subseteq P$ with $\|\mu - y\|^2 = \rho$ and $\rho > 0$.*

Suppose P is locally flat at y and let $\mathcal{B}(\mu, \sqrt{\rho})$ be as in the definition. Let $a^T x \ge a_0$ be a supporting hyperplane for $\mathcal{B}(\mu, \sqrt{\rho})$ at y, i.e. $a^T y = a_0$ and $a^T x \ge a_0$ for all $x \in \mathcal{B}(\mu, \sqrt{\rho})$. We claim that

$$q \;\ge\; 2y^T x - \|y\|^2 + 2\alpha(a^T x - a_0) \tag{7}$$

is valid for S if $\alpha \ge 0$ is small enough. To see this, note that since $a^T x \ge a_0$ supports $\mathcal{B}(\mu, \sqrt{\rho})$ at y, it follows that $\mu - y = \bar{\alpha} a$ for positive $\bar{\alpha}$, i.e.,

$$\mathcal{B}(y + \bar{\alpha} a\,,\, \sqrt{\bar{\alpha}^2 \|a\|^2}) \;=\; \mathcal{B}(\mu, \sqrt{\rho}). \tag{8}$$

Now, assume $\alpha \le \bar{\alpha}$. Then $(v, \|v\|^2)$ violates (7) iff

$$\begin{aligned}
\|v\|^2 &< 2y^T v - \|y\|^2 + 2\alpha(a^T v - a_0) && (9)\\
&= 2(y+\alpha a)^T v - \|y+\alpha a\|^2 + \alpha^2\|a\|^2 + 2\alpha(y^T a - a_0) && (10)\\
&= 2(y+\alpha a)^T v - \|y+\alpha a\|^2 + \alpha^2\|a\|^2, \quad \text{that is,} && (11)\\
v &\in \mathcal{B}(y+\alpha a\,,\, \sqrt{\alpha^2\|a\|^2}) \subseteq \mathcal{B}(\mu, \sqrt{\rho}) && (12)
\end{aligned}$$

since $\alpha \le \bar{\alpha}$. In other words, for small enough, but positive α, (7) is valid for S.

In fact, the above derivation implies a stronger statement: since $a^T x \ge a_0$ supports $\mathcal{B}(y+\alpha a\,,\, \sqrt{\alpha^2\|a\|^2})$ at y, for any $\alpha > 0$, it follows (7) is valid for S iff $\mathcal{B}(y+\alpha a\,,\, \sqrt{\alpha^2\|a\|^2}) \subseteq P$. Define

$$\hat{\alpha} = \hat{\alpha}(P, y) \doteq \sup\{\, \alpha : \text{(7) is valid}\,\} \qquad (13)$$

If there exists $v \notin P$ such that $a^T v > a_0$ then the assumptions on P imply that $\hat{\alpha} < +\infty$ and the 'sup' is a 'max'. If on the other hand $a^T v \le a_0$ for all $v \notin P$ then $\hat{\alpha} = +\infty$ (and, of course, $a^T x \le a_0$ is valid for S). In the former case, we call

$$q \ge 2y^T x - \|y\|^2 + 2\hat{\alpha}(a^T x - a_0) \qquad (14)$$

a *lifted first-order inequality.*

Theorem 2.3 *Any linear inequality*

$$\delta q - \beta^T x \ge \beta_0 \qquad (15)$$

valid for S either has $\delta = 0$ (in which case the inequality is valid for $\mathbb{R}^d - P$), or $\delta > 0$ and (15) is dominated by a lifted first-order inequality or by a linearization inequality (2).

Proof. Consider a valid inequality (15). As above we either have $\delta = 0$, in which case we are done, or without loss of generality $\delta = 1$, and by increasing β_0 if necessary we have that (15) is tight at some point $(y, \|y\|^2) \in \mathbb{R}^d \times \mathbb{R}$.
Write

$$\beta^T x + \beta_0 = 2y^T x - \|y\|^2 + 2\gamma^T x + \gamma_0, \qquad (16)$$

for appropriate γ and γ_0. Suppose first that $y \in \text{int}(\mathbb{R}^d - P)$. Then $(\gamma, \gamma_0) = (0, 0)$, or else (15) would not be valid in a neighborhood of y. Thus, (15) is a linearization inequality.
Suppose next that $y \in \partial P$, and that (15) is not a linearization inequality, i.e. $(\gamma, \gamma_0) \ne (0, 0)$. We can write (15) as

$$\begin{aligned}
q &\ge 2y^T x - \|y\|^2 + 2\gamma^T x + \gamma_0\\
&= 2(y+\gamma)^T x - \|y+\gamma\|^2 - 2\gamma^T y - \|\gamma\|^2 + \gamma_0. && (17)
\end{aligned}$$

Since (15) is not a linearization inequality, and is tight at $(y, \|y\|^2)$ there exist points $(v, \|v\|^2)$ (with v near y) which do not satisfy it. Necessarily, any such v must not lie in $\mathbb{R}^d - P$ (since (15) is valid for S). Using (17) this happens iff

$$\begin{aligned}
\|v\|^2 &< 2(y+\gamma)^T v - \|y+\gamma\|^2 - 2\gamma^T y - \|\gamma\|^2 + \gamma_0, \quad \text{that is,} && (18)\\
v &\in \text{int}\left(\mathcal{B}\left(y+\gamma\,,\, \sqrt{-2\gamma^T y - \|\gamma\|^2 + \gamma_0}\right) \right). && (19)
\end{aligned}$$

In other words, the set of points that violate (15) is the interior of some ball $\mathcal{B}$ with positive radius, which necessarily must be contained in P. Since $(y, \|y\|^2)$ satisfies (15) with inequality, y is in the boundary of $\mathcal{B}$. Thus, P is locally flat at y; writing $a^T x = a_0$ to denote the hyperplane orthogonal to γ through y, we have that (15) is dominated by the resulting lifted first-order inequality. ■

2.3 The polyhedral case

Here we will discuss an efficient separation procedure for lifted first-order inequalities in the case that P is a polyhedron. Further properties of these inequalities are discussed in [23].

Suppose that $P = \{ x \in \mathbb{R}^d \,:\, a_i^T x \geq b_i,\ 1 \leq i \leq m\}$ is a full-dimensional polyhedron, where each inequality is facet-defining and the representation of P is minimal. For $1 \leq i \leq m$ let $H_i \doteq \{ x \in \mathbb{R}^d \,:\, a_i^T x = b_i\}$. For $i \neq j$ let $H_{\{i,j\}} \doteq \{ x \in \mathbb{R}^d \,:\, a_i^T x = b_i,\ a_j^T x = b_j\}$. Assuming $H_{\{i,j\}} \neq \emptyset$ (i.e. H_i and H_j are not parallel) $H_{\{i,j\}}$ is $(d-2)-$dimensional; in that case we denote by ω_{ij} the unique unit norm vector orthogonal to both H_{ij} and a_i (unique up to reversal).

Consider a fixed pair of indices $i \neq j$ with $H_{\{i,j\}} \neq \emptyset$, and let $\mu \in \text{int}(P)$. Let Ω_{ij} be the 2-dimensional hyperplane through μ generated by a_i and ω_{ij}. By construction, therefore, Ω_{ij} is orthogonal to $H_{\{i,j\}}$ and is thus the orthogonal complement to $H_{\{i,j\}}$ through μ. It follows that $\Omega_{ij} = \Omega_{ji}$ and that this hyperplane contains the orthogonal projection of μ onto H_i (which we denote by $\pi_i(\mu)$ and the orthogonal projection of μ onto H_j ($\pi_j(\mu)$, respectively). Further, $\Omega_{ij} \cap H_{\{i,j\}}$ consists of a single point $k_{\{i,j\}}(\mu)$ satisfying

$$\begin{aligned} \|\mu - k_{\{i,j\}}(\mu)\|^2 &= \|\mu - \pi_i(\mu)\|^2 + \|\pi_i(\mu) - k_{\{i,j\}}(\mu)\|^2 \\ &= \|\mu - \pi_j(\mu)\|^2 + \|\pi_j(\mu) - k_{\{i,j\}}(\mu)\|^2. \end{aligned} \tag{20}$$

Now we return to the question of separating lifted first-order inequalities. Note that P is locally flat at a point y if and only if y is in the relative interior of one of the facets. Suppose that y is in the relative interior of the i^{th} facet. Denoting, for $j \neq i$,

$$P^{i,j} \doteq \{x \in \mathcal{R}^d \,:\, a_i^T x \geq b_i,\ a_j^T x \geq b_j\}, \quad \text{and} \tag{21}$$

we clearly have (see (13))

$$\hat{\alpha} = \min_{j \neq i} \hat{\alpha}(P^{i,j}, y).$$

We will argue that for $j \neq i$, $\hat{\alpha}(P^{i,j}, y)$ is an **affine** function of y, i.e.

$$\hat{\alpha}(P^{i,j}, y) = p_{ij} y + q_{ij} \tag{22}$$

for appropriate constants p_{ij} and q_{ij}.

Assume first that $H_{\{i,j\}} = \emptyset$, i.e. H_i and H_j are parallel and thus without loss of generality $a_i = -a_j$ and $b_i < -b_j$. But as per (for example) equation (12) the lifting coefficient at y is proportional to the largest radius of a ball that can be inscribed in the region delimited by H_i and H_j, i.e. $\{x \in \mathcal{R}^d \,:\, b_i \leq a_i^T x \leq -b_j\}$. This largest radius equals exactly half the distance between H_i and H_j, and is therefore independent of y, i.e. it is trivially an affine function of y.

Thus we assume that $H_{\{i,j\}} \neq \emptyset$. Then

$$y = \pi_i(\mu) \text{ and } \hat{y} = \pi_j(\mu), \tag{23}$$

$$y - k_{\{i,j\}}(\mu) \text{ is parallel to } \omega_{ij} \text{ and } \hat{y} - k_{\{i,j\}}(\mu) \text{ is parallel to } \omega_{ji}, \tag{24}$$

$$\|\mu - y\|^2 = \|\mu - \hat{y}\|^2 = \rho, \quad \text{and by (20)}, \tag{25}$$

$$\|y - k_{\{i,j\}}(\mu)\| = \|\hat{y} - k_{\{i,j\}}(\mu)\|, \quad \text{and} \tag{26}$$

$$\|\mu - y\| = \tan\phi \, \|y - k_{\{i,j\}}(\mu)\|, \tag{27}$$

where 2ϕ is the angle formed by ω_{ij} and ω_{ji}. By the preceding discussion, $\rho = (\, \hat{\alpha}(P^{i,j}, y)\|a_i\|\,)^2$; using (25) and (27) we will complete the argument that $\hat{\alpha}(P^{i,j}, y)$ is an affine function of y.

Let $h^g_{\{i,j\}}$ $(1 \leq g \leq d-2)$ be a basis for $\{ x \in \mathbb{R}^d \,:\, a_i^T x = a_j^T x = 0\}$. Then a_i, together with ω_{ij} and the $h^g_{\{i,j\}}$ form a basis for $\mathbb{R}^d$. Let

- O_i be the projection of the origin onto H_i – hence O_i is a multiple of a_i,
- N_i be the projection of O_i onto $H_{\{i,j\}}$.

We have

$$y \;=\; O_i \,+\, (N_i - O_i) \,+\, (k_{\{i,j\}}(\mu) - N_i) \,+\, (y - k_{\{i,j\}}(\mu)), \tag{28}$$

and thus, since $N_i - O_i$ and $y - k_{\{i,j\}}(\mu)$ are parallel to ω_{ij}, and $k_{\{i,j\}}(\mu) - N_i$ and O_i are orthogonal to ω_{ij},

$$\omega_{ij}^T y \;=\; \omega_{ij}^T (N_i - O_i) + \omega_{ij}^T (y - k_{\{i,j\}}(\mu)) \;=\; \omega_{ij}^T (N_i - O_i) + \|\omega_{ij}\| \|y - k_{\{i,j\}}(\mu)\|, \tag{29}$$

or

$$\|y - k_{\{i,j\}}(\mu)\| \;=\; \|\omega_{ij}\|^{-1} \omega_{ij}^T \,(y - N_i + O_i). \tag{30}$$

Consequently,

$$\hat{\alpha}(P^{i,j}, y) \;=\; \frac{\rho}{\|a_i\|} \;=\; \frac{\tan\phi}{\|a_i\|} \,\|y - k_{\{i,j\}}(\mu)\| \tag{31}$$

$$=\; \frac{\tan\phi}{\|a_i\|} \,\|\omega_{ij}\|^{-1} \omega_{ij}^T \,(y - N_i + O_i), \tag{32}$$

which is affine in y, as desired.

Now let $x^* \in \mathbb{R}^d$. The problem of finding the strongest possible lifted first-order inequality at x^* chosen from among those obtained by starting from a point on face i, can thus be written as follows:

$$\begin{aligned}
\min \quad & -2y^T x^* \,+\, \|y\|^2 \,-\, 2\alpha(a^T x^* - a_0) && (33)\\
s.t. \quad & y \in P && (34)\\
& a_i^T y = b_i && (35)\\
& 0 \le \alpha \le p_{ij} y + q_{ij} \quad \forall\, j \neq i. && (36)
\end{aligned}$$

[Here, (36) is valid because for $y \in H_{\{i,j\}}$ expression (32) yields $\hat{\alpha} = 0$, since ω_{ij} is orthogonal to both a_i and $H_{\{i,j\}}$.] This is a linearly constrained, convex quadratic program with $d+1$ variables and $2m-1$ constraints. By solving this problem for each choice of $1 \le i \le m$ we obtain the strongest inequality overall.

2.3.1 The Disjunctive Approach

For $1 \le i \le m$ let $\bar{P}^i = \{\, x \in \mathbb{R}^d \,:\, a_i^T x \le b_i \}$; thus $\mathbb{R}^d - P = \bigcup_i \bar{P}^i$. Further, for $1 \le i \le m$ write:

$$\bar{Q}^i = \{\, (x, q) \in \mathbb{R}^d \times \mathbb{R} \,:\, a_i^T x \le b_i,\; q \ge \|x\|^2 \}.$$

Thus, $(x^*, q^*) \in \mathrm{conv}(S)$ if and only if (x^*, q^*) can be written as a convex combination of points in the sets $\bar{Q}^i$. This is the approach pioneered in Ceria and Soares [14] (also see [30]). The resulting separation problem is carried out by solving a second-order cone program with m conic constraints and md variables, and then using second-order cone duality in order to obtain a linear inequality (details in [23]).

Thus, the derivation we presented above amounts to a possibly simpler alternative to the Ceria-Soares approach, which also makes explicit the geometric nature of the resulting cuts.

2.4 The ellipsoidal case

In this section we will discuss an efficient separation procedure for lifted first-order inequalities in the case that P is a convex ellipsoid with nonempty interior. Write

$$P \;=\; \{x \,\in \mathbb{R}^d \,:\, x^T A x \,-\, 2c^T x \,+\, b \,\le\, 0\}$$

for appropriate $A \succ 0$, c and b. Suppose we are given a point $\bar{x} \in \mathrm{int}(P)$. The problem of finding the strongest inequality at $\bar{x}$ is:

$$\begin{aligned}
\min_{\mu,\rho} \quad & \|\mu\|^2 - \rho - 2\bar{x}^T \mu && (37)\\
\text{Subject to:} \quad & \{x : \|x - \mu\|^2 \le \rho\} \subseteq P && (38)
\end{aligned}$$

Constraint (38) forces the ball of excluded points to be contained in the ellipsoid P. The S-Lemma [31], [24], [6], tells us that (μ,ρ) is feasible for (38) if and only if there is some nonnegative $\theta = \theta(\mu,\rho)$ such that

$$\|x-\mu\|^2 - \rho - \theta(x^TAx - 2b^Tx + c) \geq 0 \quad \forall x \in \mathbb{R}^d.$$

This is equivalent to saying that there is $\theta \geq 0$ with

$$\min_x\{\|x-\mu\|^2 - \rho - \theta(x^TAx - 2b^Tx + c)\} \;\geq\; 0,$$

or equivalently

$$\min_x\{x^T(I-\theta A)x - 2(\mu-\theta b)^Tx + (\|\mu\|^2 - \rho - \theta c)\} \;\geq\; 0. \tag{39}$$

Clearly, we must have $\theta \leq \frac{1}{\lambda_{max}}$ for this to hold. Now consider an optimal pair $(\hat\mu,\hat\rho)$ for problem (37)-(38), and the corresponding value $\hat\theta$. We will show next that $\hat\theta = \frac{1}{\lambda_{max}}$.

Aiming for a contradiction, assume $\hat\theta < \frac{1}{\lambda_{max}}$. Then $(I-\hat\theta A)$ is invertible, and the optimal solution to the minimization problem (39) is given by

$$x^* = (I-\hat\theta A)^{-1}(\hat\mu - \hat\theta b).$$

Substituting this expression in (39), we obtain

$$-\left[\hat\mu^T(I-\hat\theta A)^{-1}\hat\mu - 2\hat\theta\hat\mu^T(I-\hat\theta A)^{-1}b + \hat\theta^2 b^T(I-\hat\theta A)^{-1}b\right] + \hat\mu^T\hat\mu - \hat\rho - \hat\theta c \;\geq\; 0.$$

Thus (via another application of the S-Lemma) problem (37)-(38) can be rewritten as:

$$\begin{aligned} \min_{\mu,\rho} \quad & \|\mu\|^2 - \rho - 2\bar{x}^T\mu \\ \text{Subject to:} \quad & -\left[\mu^T(I-\hat\theta A)^{-1}\mu - 2\hat\theta\mu^T(I-\hat\theta A)^{-1}b + \hat\theta^2 b^T(I-\hat\theta A)^{-1}b\right] + \|\mu\|^2 - \rho - \hat\theta c \geq 0 \end{aligned}$$

This is a convex QCQP. Notice the term $\|\mu\|^2 - \rho$ which appears both in the objective and the constraint. From this we can see that the constraint will hold with equality at the optimal (μ,ρ), so we can substitute into the objective to get the unconstrained separation problem:

$$\min_\mu \quad \hat\theta c + \mu^T(I-\hat\theta A)^{-1}\mu - 2\hat\theta b^T(I-\hat\theta A)^{-1}\mu + \hat\theta^2 b^T(I-\hat\theta A)^{-1}b - 2\bar{x}^T\mu.$$

This is a convex QP; using KKT conditions we get that its optimal solution is given by

$$\hat\mu = \hat\theta b + (I-\hat\theta A)\bar{x},$$

and plugging this into the ojective gives a value of

$$-\|\bar{x}\|^2 + \hat\theta(\bar{x}^TA\bar{x} - 2b^T\bar{x} + c).$$

Since $\bar{x} \in \text{int}(P)$ we have $\bar{x}^TA\bar{x} - 2b^T\bar{x} + c < 0$, so this objective value is decreasing linearly in $\hat\theta$. Since our objective in problem (37)-(38) is to minimize, the optimal $\hat\theta$ will be as large as possible: $\frac{1}{\lambda_{max}}$, as desired.

[Note that we can then determine the optimal squared radius $\hat\rho$ by:

$$\|\hat\mu\|^2 - \hat\rho - 2\bar{x}^T\hat\mu = -\bar{x}^2 + \hat\theta(\bar{x}^TA\bar{x} - 2b^T\bar{x} + c).$$

This again shows that any $\hat\theta < \frac{1}{\lambda_{max}}$ is not optimal - we always get a better cut by slightly increasing $\hat\theta$.]

Assuming now that $\hat\theta = \lambda_{max}^{-1}$, the following approach is almost identical to the above. Write the separation problem as:

$$\min_{\mu,\rho} \quad \|\mu\|^2 - \rho - 2\bar{x}^T\mu \tag{40}$$

$$\text{Subject to:} \quad \min_x\{x^T(I-\hat\theta A)x - 2(\mu-\hat\theta b)^Tx + (\|\mu\|^2 - \rho - \hat\theta c)\} \;\geq\; 0 \tag{41}$$

or equivalently, pulling out a few terms in the constraint which don't depend on x:

$$\min_{\mu,\rho} \quad \|\mu\|^2 - \rho - 2\bar{x}^T\mu \tag{42}$$

$$\text{Subject to:} \quad \|\mu\|^2 - \rho + \min_x\{x^T(I-\hat{\theta}A)x - 2(\mu-\hat{\theta}b)^T x + \hat{\theta}c\} \;\geq\; 0. \tag{43}$$

Clearly the constraint will hold with equality, so we can transform the constrained problem into an unconstrained one:

$$\min_{\mu} \left[-2\bar{x}^T\mu - \min_x\{x^T(I-\hat{\theta}A)x - 2(\mu-\hat{\theta}b)^T x - \hat{\theta}c\} \right].$$

The optimal μ must be such that the optimal value of the inner minimization problem (the one over x) is finite. That is, for any $\delta \in \mathbb{R}^d$,

$$(I-\hat{\theta}A)\delta = 0 \quad \text{implies} \quad (\mu-\hat{\theta}b)^T\delta = 0.$$

Using the Farkas Lemma, this is equivalent to μ being of the form

$$\mu = \hat{\theta}b + (I-\hat{\theta}A)\pi \qquad \text{for some } \pi \in \mathbb{R}^d$$

Then the optimal solution to the inner minimization is any x satisfying

$$(I-\hat{\theta}A)x = \mu - \hat{\theta}b = (I-\hat{\theta}A)\pi.$$

Clearly π is a minimizer, and the resulting optimal value is

$$-\pi^T(I-\hat{\theta}A)\pi - \hat{\theta}c$$

We can then rewrite the separation problem again as:

$$\min_{\pi} \left[-2\bar{x}^T(\hat{\theta}b + (I-\hat{\theta}A)\pi) + \pi^T(I-\hat{\theta}A)\pi + \hat{\theta}c \right]$$

This is an unconstrained convex QP, its optimal solution is

$$\hat{\pi} = \bar{x},$$

which means the optimal center $\hat{\mu}$ is

$$\hat{\mu} = \hat{\theta}b + (I-\hat{\theta}A)\bar{x}$$

and the optimal squared radius $\hat{\rho}$ is

$$\hat{\rho} = \|\hat{\mu}\|^2 - 2\bar{x}^T\hat{\mu} + \|\bar{x}\|^2 - \hat{\theta}(\bar{x}^TA\bar{x} - 2b^T\bar{x} + c).$$

3 Indefinite Quadratics

The general case of a set $\{\,(x,q) \in \mathbb{R}^d \times \mathbb{R} \;:\; q \geq Q(x), \quad x \in \mathbb{R}^d - \text{int}(P)\,\}$, where $Q(x)$ is a semidefinite quadratic can be approached in much the same way as that employed above, but with some important differences.

We first consider the case where P is a polyhedron. Let $P = \{(x,w) \in \mathbb{R}^{d+1} \;:\; a_i^T x - w \leq b_i,\ 1 \leq i \leq m\}$ (here, w is a scalar). Consider a set of the form

$$S \;\doteq\; \left\{ (x,w,q) \in \mathbb{R}^{d+2} \,:\, q \geq \|x\|^2,\ (x,w) \in \mathbb{R}^{d+1} - P \right\}. \tag{44}$$

Many examples can be brought into this form, or similar, by an appropriate affine transformation. Consider a point (x^*,w^*) in the relative interior of the i^{th} facet of P. We seek a lifted first-order inequality of the form

$$(2x^* - \alpha a_i)^T x + \alpha w + \alpha b_i - \|x^*\|^2 \leq q,$$

for appropriate $\alpha \geq 0$. If we are lifting to the j^{th} facet, then we must have $v_{ij} = \alpha b_i - \|x^*\|^2$, where

$$v_{ij} \doteq \min \|x\|^2 - (2x^* - \alpha a_i)^T x - \alpha w \tag{45}$$

$$\text{s.t.} \quad a_j^T x - w = b_j. \tag{46}$$

To solve this optimization problem, consider its Lagrangian:

$$\mathcal{L}(x, w, \nu) = \|x\|^2 - (2x^* - \alpha a_i)^T x - \alpha w - \nu(a_j^T x - w - b_j)$$

Taking the gradient in x and setting it to 0:

$$\begin{aligned} \nabla_x \mathcal{L} = 0 \quad &\Leftrightarrow \quad 2x - 2x^* + \alpha a_i - \nu a_j = 0 \\ &\Leftrightarrow \quad x = x^* - \frac{\alpha}{2} a_i + \frac{\nu}{2} a_j \end{aligned}$$

Now doing the same for w:

$$\begin{aligned} \nabla_w \mathcal{L} = 0 \quad &\Leftrightarrow \quad -\alpha + \nu = 0 \\ &\Leftrightarrow \quad \nu = \alpha \end{aligned}$$

Combining these two gives

$$x = x^* - \frac{\alpha}{2} a_i + \frac{\alpha}{2} a_j$$

then using the constraint $a_j^T x - w = b_j$ gives

$$w = a_j^T x^* - b_j - \frac{\alpha}{2} a_j^T a_i + \frac{\alpha}{2} a_j^T a_j$$

Next we expand out the objective value using the expressions we have derived for x and w, and set the result equal to $\alpha b_i - \|x^*\|^2$. Omitting the intermediate algebra, the result is the quadratic equation

$$\alpha(a_i^T x^* - b_i - (a_j^T x^* - b_j)) - \frac{1}{4}\alpha^2(a_i^T a_i - 2a_i^T a_j + a_j^T a_j) = 0$$

One root of this equation is $\alpha = 0$. The other root is

$$\hat{\alpha} \doteq \frac{4(a_i^T x^* - b_i - (a_j^T x^* - b_j))}{a_i^T a_i - 2a_i^T a_j + a_j^T a_j}. \tag{47}$$

Since $a_i^T x^* - w^* = b_i$, and $a_j^T x^* - w^* \leq b_j$, we have

$$a_i^T x^* - b_i - (a_j^T x^* - b_j) > 0$$

so $\hat{\alpha} > 0$ (the denominator is a squared distance between some two vectors so it is non-negative). Moreover, the expression for $\hat{\alpha}$ is an affine function of x^*. Thus, as in Section 2.3, the computation of a maximally violated lifted first-order inequality is a convex optimization problem.

In this case there is an additional detail of interest: note that the points $(x, w, \|x\|^2)$ cut-off by the inequality are precisely those such that

$$(2x^* - \hat{\alpha} a_i)^T x + \hat{\alpha} w + \hat{\alpha} b_i - \|x^*\|^2 > \|x\|^2. \tag{48}$$

This condition defines the interior of a *paraboloid*; this is the proper generalization of condition (3) in the indefinite case.

3.1 Tightening a general quadratic expression

Consider a set of the form

$$\Pi \doteq \{ (x,w,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \ : \ z \geq x^T Q x + q^T x, \quad w \leq x^T A x \} \tag{49}$$

where both A and Q are symmetric positive definite $d \times d$ matrices. We will show below that this system can be characterized through a family of polynomially separable linear inequalities in x, w and z; we develop a streamlined construction in Section 3.1.1. An application is described later.

Let $Q = LL^T$, where L is lower triangular and invertible, and let $V \Lambda V^T$ be the spectral decomposition of $L^{-1} A L^{-T}$. Writing $p = V^T L^T x$, and so $x = L^{-T} V p$, we therefore have:

$$\begin{aligned} x^T Q x &= p^T V^T L^{-1} L L^T L^{-T} V p \ = \ p^T p, \ \text{ and} \\ x^T A x &= p^T V^T L^{-1} A L^{-T} V p \ = \ p^T \Lambda p. \end{aligned}$$

Thus, without loss of generality Π is described by the system

$$z \geq \|x\|^2 + q^T x \tag{50}$$

$$w \leq x^T \Lambda x, \tag{51}$$

where $\Lambda \succ 0$ is diagonal. Define

$$P \doteq \{(x,w) \in \mathbb{R}^d \times \mathbb{R} \, : \, x^T \Lambda x - w \leq 0\}.$$

This is a paraboloid in (x,w)-space whose interior is the set of points in in (x,w)-space that are cut-off by (51). Write

$$\lambda_{max} \doteq \max_i \lambda_i, \tag{52}$$

and, given $\mu \in \mathbb{R}^d$ and $\nu \in \mathbb{R}$,

$$M(\mu,\nu) \ \doteq \ \{(x,w) \ \in \ \mathbb{R}^d \times \mathbb{R} \ : \ \lambda_{max}\|x-\mu\|^2 + (\nu - w) \ \leq \ 0\}.$$

Then it is seen that

$$x \in \mathbb{R}^d - \text{int}(P) \ \text{ iff } \ x \in \mathbb{R}^d - \text{int}(M(\mu,\nu)), \text{ for all } \mu,\nu \text{ such that } M(\mu,\nu) \subseteq P. \tag{53}$$

Using this characterization together with (50) we have that for each pair $(\mu,\nu) \in \mathbb{R}^d \times \mathbb{R}$ with $M(\mu,\nu) \subseteq P$ the following inequality is valid for the set (50)-(51):

$$\lambda_{max}\|\mu\|^2 - \lambda_{max}(2\mu + q)^T x + (\nu - w) + \lambda_{max} z \geq 0, \tag{54}$$

which precisely cuts-off $\text{int}(M(\mu,\nu))$ in the sense that given $(\hat{x},\hat{w}) \in \text{int}(M(\mu,\nu))$ if $\hat{z} \leq \|\hat{x}\|^2 + q^T\hat{x}$ then $(\hat{x},\hat{w},\hat{z})$ violates (54).

By definition, for $(\mu,\nu) \in \mathbb{R}^d \times \mathbb{R}$ we have $M(\mu,\nu) - P \neq \emptyset$ iff there exists $(\hat{x},\hat{w}) \in M(\mu,\nu)$ with $\hat{w} < \hat{x}^T \Lambda \hat{x}$, and therefore iff there exists $\hat{x}$ such that $(\hat{x}, \hat{x}^T \Lambda \hat{x}) \in \text{int}(M(\mu,\nu))$. Consequently,

$$M(\mu,\nu) \subseteq P \quad \text{iff} \quad \nu + \min_x \{ \, \lambda_{max} \, \|x - \mu\|^2 - x^T \Lambda x \, \} \geq 0.$$

Therefore, the maximum violation of an inequality (54) at a point $(\bar{x},\bar{w}) \in \text{int}(P)$ is obtained by solving the problem

$$\max_{\mu,\nu} \qquad \bar{w} - \nu - \lambda_{max}\|\mu\|^2 + 2\lambda_{max}\bar{x}^T\mu + 2\lambda_{max} q^T \bar{x} \tag{55}$$

$$\text{subject to:} \qquad \nu + \min_x \{ \, \lambda_{max} \, \|x - \mu\|^2 - x^T \Lambda x \, \} \geq 0. \tag{56}$$

We will show that this problem can be solved in polynomial time; in fact we will provide an explicit expression for an optimal solution. Let (μ^*, ν^*) be optimal. Clearly (56) will hold with equality, otherwise we could just decrease ν^* and obtain a better solution. Further, if x^* is the minimizer in the constraint, we have

$$\mu^*_i = 0 \quad \text{for all } i \text{ with } \lambda_i = \lambda_{max} \tag{57}$$

and

$$x_i^* = \frac{\lambda_{max}\mu_i^*}{\lambda_{max} - \lambda_i} \quad \text{for all } i \text{ with } \lambda_i < \lambda_{max}.$$

Thus,

$$\nu^* + \lambda_{max} \sum_{\{i:\lambda_i<\lambda_{max}\}} \left(\frac{\mu_i^* \lambda_{max}}{\lambda_{max} - \lambda_i} - \mu_i^* \right)^2 - \sum_{\{i:\lambda_i<\lambda_{max}\}} \lambda_i \frac{\lambda_{max}^2 (\mu_i^*)^2}{(\lambda_{max} - \lambda_i)^2} = 0, \quad \text{or} \tag{58}$$

$$\nu^* + \lambda_{max} \sum_{\{i:\lambda_i<\lambda_{max}\}} \frac{(\mu_i^*)^2 \lambda_i^2}{(\lambda_{max} - \lambda_i)^2} - \sum_{\{i:\lambda_i<\lambda_{max}\}} \frac{\lambda_i \lambda_{max}^2 (\mu_i^*)^2}{(\lambda_{max} - \lambda_i)^2} = 0, \quad \text{and therefore} \tag{59}$$

$$\nu^* + \lambda_{max} \sum_{\{i:\lambda_i<\lambda_{max}\}} \frac{\mu_i^2 \lambda_i^2 - \lambda_{max} \lambda_i \mu_i^2}{(\lambda_{max} - \lambda_i)^2} = 0, \quad \text{thus,} \tag{60}$$

$$\nu^* = \lambda_{max} \sum_{\{i:\lambda_i<\lambda_{max}\}} \frac{\mu_i^2 \lambda_i}{\lambda_{max} - \lambda_i}. \tag{61}$$

Now we can rewrite the separation problem (55)-(56):

$$\max_{\mu} \quad -\lambda_{max}\|\mu\|^2 + 2\lambda_{max}\bar{x}^T\mu + 2\lambda_{max} q^T \bar{x} + \bar{w} - \lambda_{max} \sum_{\{i:\lambda_i<\lambda_{max}\}} \frac{\lambda_i {\mu_i}^2}{\lambda_{max} - \lambda_i} \tag{62}$$

$$\text{subject to:} \quad \mu_i = 0 \text{ for all } i \text{ with } \lambda_i = \lambda_{max}; \tag{63}$$

dividing the objective by λ_{max} and ignoring constant terms we get

$$\max_{\mu} - \left(\sum_{\{i:\lambda_i<\lambda_{max}\}} \left(1 + \frac{\lambda_i}{\lambda_{max} - \lambda_i} \right) \mu_i^2 \right) + 2 \sum_{\{i:\lambda_i<\lambda_{max}\}} \mu_i \bar{x}_i. \tag{64}$$

Note that the coefficient of μ_i^2 is $-\lambda_{max}/(\lambda_{max} - \lambda_i) < 0$, and thus the quadratic maximized in (64) is negative definite. Setting its gradient to zero, we obtain that the optimal solution is

$$\mu_i^* = \frac{\lambda_{max} - \lambda_i}{\lambda_{max}} \bar{x}_i. \tag{65}$$

which, together with (57) implies, by substituting into (61)

$$\nu^* = \frac{1}{\lambda_{max}} \sum_{\lambda_i<\lambda_{max}} \lambda_i(\lambda_{max} - \lambda_i)\bar{x}_i^2. \tag{66}$$

We thus obtain an explicit solution to the separation problem (55)-(56), and therefore, using the geometrical statement (53) a characterization of (50)-(51) by polynomially separable linear inequalities.

We now prove a domination result for the cuts (54) similar to that in Theorem 2.3. We will show that these are lifted inequalities, and that they dominate all valid inequalities. Given $\bar{x} \in \mathbb{R}^d$, and scalar $\alpha \geq 0$, the inequality

$$\left(\begin{bmatrix} 2\bar{x} + q \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -2\Lambda\bar{x} \\ 1 \end{bmatrix} \right)^T \begin{bmatrix} x - \bar{x} \\ w - \bar{x}^T\Lambda\bar{x} \end{bmatrix} + \|\bar{x}\|^2 + q^T\bar{x} \leq z \tag{67}$$

is termed a *lifted* inequality at $(\bar{x}, \bar{x}^T\Lambda\bar{x})$ with lifting coefficient α. We will also equivalently rewrite (67) as

$$z \geq \|\bar{x}\|^2 + q^T\bar{x} + (2\bar{x} + q)^T(x - \bar{x}) + \alpha(w - \bar{x}^T\Lambda\bar{x} - 2\bar{x}^T\Lambda(x - \bar{x})); \tag{68}$$

thus (67) strengthens the valid inequality $z \geq \|\bar{x}\|^2 + q^T\bar{x} + (2\bar{x} + q)^T(x - \bar{x})$ in the (infeasible) region where $w > x^T\Lambda x$.

Theorem 3.1 *Given $\bar{x} \in \mathbb{R}^d$ the lifted inequality (67) is valid iff $\alpha \leq \lambda_{max}^{-1}$, and when $\alpha = \lambda_{max}^{-1}$ it coincides with the strongest inequality (54) at $(\bar{x}, \bar{x}^T \Lambda \bar{x})$.*

Proof. Suppose without loss of generality that $\lambda_{max} = \lambda_1$. Write $e_1 = (1, 0, \ldots, 0)^T \in \mathbb{R}^d$, and define, for $\delta \in \mathbb{R}$,

$$\bar{x}(\delta) \doteq \bar{x} + \delta e_1, \tag{69}$$

$$\bar{w}(\delta) \doteq \bar{x}^T(\delta)\, \Lambda\, \bar{x}(d) \;=\; \bar{x}^T \Lambda \bar{x} + 2\delta \bar{x}_1 \lambda_{max} + \delta^2 \lambda_{max}, \quad \text{and} \tag{70}$$

$$\bar{z}(\delta) \doteq \|\bar{x}(\delta)\|^2 + q^T \bar{x}(\delta) \;=\; \|\bar{x}\|^2 + q^T \bar{x} + 2\delta \bar{x}_1 + \delta^2 + \delta q_1. \tag{71}$$

Then by construction $(\bar{x}(\delta), \bar{w}(\delta), \bar{z}(\delta))$ is feasible (it is contained in Π, defined by (50)-(51)). Evaluating (68) at $(\bar{x}(\delta), \bar{w}(\delta))$, on the other hand, we obtain

$$z \;\geq\; \|\bar{x}\|^2 + q^T \bar{x} + 2\delta \bar{x}_1 + \delta q_1 + \alpha \delta^2 \lambda_{max} \;>\; \bar{z}(\delta), \quad \text{if } \alpha > \lambda_{max}^{-1}. \tag{72}$$

Thus (67) is *not* valid if $\alpha > \lambda_{max}^{-1}$. Next, by definition if (67) is *in*valid for a certain value $\hat{\alpha} > 0$, then there is a triple (x, w, z) in Π for which the right-hand side of (68) exceeds z. This can only happen if the last term in the right-hand side of (68) is positive; i.e.

$$w - \bar{x}^T \Lambda \bar{x} \;-\; 2\bar{x}^T \Lambda (x - \bar{x}) > 0.$$

But in that case (67) will be invalid for any lifting coefficient $\alpha \geq \hat{\alpha}$. Thus, in order to complete the proof of the theorem, it suffices to show that when $\alpha = \lambda_{max}^{-1}$ inequality (67) coincides with the most violated inequality (54) at $(\bar{x}, \bar{x}^T \Lambda \bar{x})$. To do so, first write the lifted cut at $(\bar{x}, \bar{x}^T \Lambda \bar{x})$ and using coefficient $1/\lambda_{max}$ as:

$$(2\bar{x} - 2\lambda_{max}^{-1} \Lambda \bar{x} + q)^T x + \lambda_{max}^{-1} w - z \leq -\lambda_{max}^{-1} \bar{x}^T \Lambda \bar{x} + \|x\|^2$$

Now we just show that the coefficients of our separating cut (54) match the coefficients of the lifted cut. The coefficients of x match if and only if

$$2\bar{x} - \frac{2}{\lambda_{max}} \Lambda \bar{x} = 2\mu \quad \Leftrightarrow \quad \mu = \bar{x} - \frac{1}{\lambda_{max}} \Lambda \bar{x}$$

$$\Leftrightarrow \quad \mu_i = \left(\frac{\lambda_{max} - \lambda_i}{\lambda_{max}} \right) \bar{x}_i \quad \forall i$$

The constant terms (right-hand-sides of the two cuts) match if and only if

$$\begin{aligned}
\nu &= \bar{x}^T \left(-\Lambda + \lambda_{max} I\right) \bar{x} - \lambda_{max} \|\mu\|^2 \\
&= \sum_{i=1}^{n} (\lambda_{max} - \lambda_i) \bar{x}_i^2 - \lambda_{max} \sum_{i=1}^{n} \left(\frac{\lambda_{max} - \lambda_i}{\lambda_{max}} t_i \right)^2 \\
&= \sum_{i=1}^{n} \bar{x}_i^2 \left((\lambda_{max} - \lambda_i) - \frac{(\lambda_{max} - \lambda_i)^2}{\lambda_{max}} \right) \\
&= \frac{1}{\lambda_{max}} \sum_{i=1}^{n} \bar{x}_i^2 (\lambda_i (\lambda_{max} - \lambda_i));
\end{aligned}$$

this matches expressions (65), (66) for the optimal μ and ν we get when computing the strongest cut (54) at $(\bar{x}, \bar{x}^T \Lambda \bar{x})$. ■

Theorem 3.2 *Any inequality*

$$c^T x + \gamma w - \tau z \leq d \tag{73}$$

valid for Π with $\gamma \neq 0$ and $\tau \neq 0$ is dominated by a lifted inequality.

Proof. Clearly $\gamma > 0$ and $\tau > 0$; without loss of generality $\tau = 1$; for convenience we restate the inequality as

$$c^T x + \gamma w - z \leq d. \tag{74}$$

Without loss of generality, assume (74) is not dominated by another valid inequality. Since (74) is valid, we have

$$0 \leq \|x\|^2 + (q-c)^T x - \gamma x^T \Lambda x + d, \quad \forall\ x \in \mathbb{R}^d, \tag{75}$$

and in consequence $I - \gamma\Lambda \succeq 0$. Thus the expression in the right-hand side (75) attains its minimum at some point $\bar{x} \in \mathbb{R}^d$. Writing

$$\bar{z} = \|\bar{x}\|^2 + q^T \bar{x} \quad \text{and} \quad \bar{w} = \bar{x}^T \Lambda \bar{x},$$

and since by assumption (74) is not dominated, we therefore have that (74) is tight at $(\bar{x}, \bar{w}, \bar{z})$.
The set of points $(x, w, \|x\|^2 + q^T x)$ violating (74) is

$$\begin{aligned} \left\{(x,w) : c^T x + \gamma w - (\|x\|^2 + q^T x) > d\right\} &= \left\{(x,w) : (c-q)^T x - \|x\|^2 + \gamma w > d\right\} \\ &= \left\{(x,w) : \|x\|^2 + (q-c)^T x + d < \gamma w\right\} \\ &= \left\{(x,w) : \frac{1}{\gamma}\|x\|^2 + \frac{1}{\gamma}(q-c)^T x + \frac{1}{\gamma} d < w\right\} \end{aligned}$$

which is the interior of a paraboloid in the (x, w) space. Since (74) is tight at $(\bar{x}, \bar{w})$, we know that $(\bar{x}, \bar{w})$ must be on the boundary of this paraboloid, and we have

$$\frac{1}{\gamma}\|\bar{x}\|^2 + \frac{1}{\gamma}(q-c)^T \bar{x} + \frac{1}{\gamma} d = \bar{w} = \bar{x}^T \Lambda \bar{x}.$$

Given these facts, we can determine what the vector c must be. We want to show that (74) must have the form of the "lifted first-order cut"[1]:

$$\left(\begin{bmatrix} 2\bar{x} + q \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -2\Lambda\bar{x} \\ 1 \end{bmatrix} \right)^T \begin{bmatrix} x - \bar{x} \\ w - \bar{w} \end{bmatrix} + \bar{z} \leq z \tag{76}$$

where $\alpha \geq 0$ is the lifting coefficient. Note that (76) is equivalent to:

$$(2\bar{x} + q - 2\alpha\Lambda\bar{x})^T x + \alpha w - z \leq -\alpha\bar{w} + \|\bar{x}\|^2. \tag{77}$$

We will show that $c = 2\bar{x} + q - 2\gamma\Lambda\bar{x}$, that is to say, inequality (74) is a lifted inequality with lifting coefficient $\alpha = \gamma$. Suppose $c \neq 2\bar{x} + q - 2\gamma\Lambda\bar{x}$. This means that the system

$$\alpha = \gamma, \quad 2\alpha\Lambda\bar{x} = 2\bar{x} + q - c$$

is infeasible, and by the Farkas Lemma, there exists a vector (π, ρ) with

$$\pi^T(2\bar{x} + q - c) + \rho\gamma < -1 \quad \text{and} \quad \pi^T(2\Lambda\bar{x}) + \rho = 0.$$

At $(\bar{x}, \bar{w})$, the inequality

$$(2\Lambda\bar{x})(x - \bar{x}) + \bar{w} \leq w$$

supports P, where as before $P = \{\, (x,w) \in \mathbb{R}^d \times \mathbb{R} \,:\, x^T \Lambda x - w \leq 0\}$. Thus the opposite inequality

$$(2\Lambda\bar{x}, -1) \cdot (x, w) \ \geq\ \bar{w}$$

gives a sufficient condition to guarantee $(x, w) \notin \text{int}(P)$. Now consider the point $(\bar{x} + \epsilon\pi, \bar{w} - \epsilon\rho)$ where ϵ is a scalar. For this point to be feasible, it is therefore enough that:

$$\begin{aligned} (2\Lambda\bar{x})^T(\bar{x} + \epsilon\pi) - (\bar{w} - \epsilon\rho) \geq \bar{w} \quad &\Leftrightarrow \quad 2\bar{w} - \bar{w} + 2\epsilon\pi^T\Lambda\bar{x} + \epsilon\rho \geq \bar{w} \\ &\Leftrightarrow \quad \epsilon(2\pi^T\Lambda\bar{x} + \rho) \geq 0 \\ &\Leftrightarrow \quad \epsilon \cdot 0 \geq 0 \end{aligned}$$

[1] Borrowing terminology from Section 2.2

which holds for all $\epsilon \in \mathbb{R}$. Then, since the inequality (74) is valid, it must hold for all points $(\bar{x}+\epsilon\pi, \bar{w}-\epsilon\rho)$. This requires in particular that for all $\epsilon > 0$, we must have:

$$\begin{aligned}
& c^T(\bar{x}+\epsilon\pi) + \gamma(\bar{w}-\epsilon\rho) - \|\bar{x}+\epsilon\pi\|^2 - q^T(\bar{x}+\epsilon\pi) \ \leq \ d \\
\Leftrightarrow \quad & c^T\bar{x} + \gamma\bar{w} + \epsilon c^T\pi - \gamma\epsilon\rho - \bar{x}^T\bar{x} - 2\epsilon\pi^T\bar{x} - \epsilon\pi^T\pi - q^T\bar{x} - \epsilon q^T\pi \ \leq \ d \\
\Leftrightarrow \quad & \epsilon c^T\pi - \gamma\epsilon\rho - 2\epsilon\pi^T\bar{x} - \epsilon^2\|\pi\|^2 - \epsilon q^T\pi \ \leq \ 0 \\
\Leftrightarrow \quad & \epsilon(c^T\pi - \gamma\rho - 2\pi^T\bar{x} - q^T\pi) - \epsilon^2\|\pi\|^2 \ \leq \ 0 \\
\Leftrightarrow \quad & \epsilon\|\pi\|^2 + \pi^T(2\bar{x}+q-c) + \gamma\rho \ \geq \ 0.
\end{aligned}$$

However, since $\pi^T(2\bar{x}+q-c) + \gamma\rho < -1$, this fails to hold for small ϵ, a contradiction. ■

As a summary of the above we have:

Theorem 3.3 *Let A, Q be positive definite $d \times d$ matrices. Any nondominated valid inequality for the set $\{\,(x,w,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \ : \ z \geq x^TQx + q^Tx, \quad w \leq x^TAx\,\}$ is a lifted inequality, and given a point $(\bar{x}, \bar{w})$ in the interior of $\{(x,w) \in \mathbb{R}^d \times \mathbb{R} \, : \, x^T\Lambda x - w \leq 0\}$ we can compute a strongest lifted inequality at $(\bar{x}, \bar{w})$ in polynomial time.* ■

3.1.1 No-spectrum implementation

The construction above requires the computation of the spectrum of the $d \times d$ matrix A found in the initial description of the set Π (eq. (49). This step was needed in order to derive the various relationships obtained above; but might prove expensive if d is large. Here we describe an equivalent construction that avoids the computation of eigenvalues, other than the largest. We assume, therefore, that we have a system of the form:

$$\begin{aligned}
& z \geq x^Tx + q^Tx \\
& w \leq x^TAx
\end{aligned}$$

where A is positive-definite. Suppose we have a point $(\bar{x}, \bar{w}, \bar{z})$, with $\bar{x}^TA\bar{x} < \bar{w}$ which we want to separate. We have shown that valid tight cuts must be of the form

$$(2\mu + q, \alpha, -1) \cdot (x, w, z) \leq \alpha\nu + \|\mu\|^2$$

where $\alpha \geq 0$. Rearranging terms, we write the constraint as

$$z \geq (2\mu + q)^Tx + \alpha w - \alpha\nu - \|\mu\|^2$$

A point $(x, w, x^Tx + q^Tx)$ violates such a cut if and only if

$$\begin{aligned}
x^Tx + q^Tx < 2\mu^Tx + q^Tx + \alpha w - \alpha\nu - \|\mu\|^2 \quad & \Leftrightarrow \quad x^Tx - 2\mu^Tx + \mu^T\mu - \alpha w + \alpha\nu < 0 \\
& \Leftrightarrow \quad \beta\|x - \mu\|^2 - w + \nu < 0
\end{aligned}$$

where in the last line we define $\beta = \alpha^{-1}$. The excluded region is a paraboloid in the (x, w) space. For the cut to be valid, we need this paraboloid to be contained in the infeasible region. That is, we need

$$\big\{(x,w) | \beta\|x-\mu\|^2 - w + \nu < 0\big\} \subseteq \big\{(x,w) | x^TAx - w < 0\big\}$$

By the S-Lemma, this is equivalent to the existence of some $\theta \geq 0$ with

$$\beta\|x-\mu\|^2 - w + \nu - \theta(x^TAx - w) \geq 0 \quad \forall (x,w) \in \mathbb{R}^{d+1}$$

Clearly we must have $\theta = 1$, or else we could fix x and send w to $\pm\infty$. So the validity of the separating cut is equivalent to

$$\nu + \beta\|\mu\|^2 \geq x^T(A - \beta I)x + 2\beta\mu^Tx \quad \forall x \in \mathbb{R}^d$$

or

$$\nu + \beta\|\mu\|^2 \geq \max_x \big\{x^T(A - \beta I)x + 2\beta\mu^Tx\big\}$$

The objective of the separation problem is

$$\max_{\mu,\nu,\alpha}: \quad (2\mu+q)^T\bar{x} + \alpha\bar{w} - \alpha\nu - \|\mu\|^2$$

or, using the definition $\beta = \alpha^{-1}$:

$$\max_{\mu,\nu,\beta}: \quad (2\mu+q)^T\bar{x} + \frac{1}{\beta}\bar{w} - \frac{1}{\beta}\nu - \|\mu\|^2$$

Adding in the validity constraint for our separating cut, the separation problem is:

$$\begin{array}{ll} \text{maximize:} & (2\mu+q)^T\bar{x} + \frac{1}{\beta}\bar{w} - \frac{1}{\beta}\nu - \|\mu\|^2 \\ \text{subject to:} & \nu + \beta\|\mu\|^2 \geq \max_x \left\{x^T(A-\beta I)x + 2\beta\mu^T x\right\} \end{array}$$

Clearly the constraint will hold with equality at the optimum; we can move it into the objective to get the equivalent unconstrained problem:

$$\max_{\mu,\beta}: \quad (2\mu+q)^T\bar{x} + \frac{1}{\beta}\bar{w} - \frac{1}{\beta}\max_x \left\{x^T(A-\beta I)x + 2\beta\mu^T x\right\}$$

In the optimal solution, the value of the inner maximization must be finite. This implies that we must have $\beta \geq \lambda_{max}(A)$ and $\beta\mu = (A-\beta I)\pi$ for some π.

Suppose first that we have fixed $\beta > \lambda_{max}(A)$, so $(A-\beta I)$ is negative-definite and invertible. The optimal x for the inner maximization is given by

$$-\beta(A-\beta I)^{-1}\mu$$

and results in an optimal value of

$$-\beta^2\mu^T(A-\beta I)^{-1}\mu$$

We can then rewrite the separation problem as:

$$\begin{array}{rl} \max_{\mu}: & (2\mu+q)^T\bar{x} + \frac{1}{\beta}\bar{w} + \frac{1}{\beta}\left(\beta^2\mu^T(A-\beta I)^{-1}\mu\right) \\ \Leftrightarrow \quad \max_{\mu}: & \beta\mu^T(A-\beta I)^{-1}\mu + 2\bar{x}^T\mu + q^T\bar{x} + \frac{1}{\beta}\bar{w} \end{array}$$

This is a convex QP whose optimal solution is

$$\mu^* = \frac{-1}{\beta}(A-\beta I)\bar{x}$$

The resulting objective value is

$$\begin{aligned} & \frac{1}{\beta}\bar{x}^T(A-\beta I)\bar{x} - \frac{2}{\beta}\bar{x}^T(A-\beta I)\bar{x} + q^T\bar{x} + \frac{1}{\beta}\bar{w} \\ = \; & \frac{-1}{\beta}\bar{x}^T A\bar{x} + \bar{x}^T\bar{x} + q^T\bar{x} + \frac{1}{\beta}\bar{w} \\ = \; & \bar{x}^T\bar{x} + q^T\bar{x} + \frac{1}{\beta}(\bar{w} - \bar{x}^T A\bar{x}) \end{aligned}$$

which is decreasing in β, since $\bar{w} > \bar{x}^T A\bar{x}$. So we want to have β at its lower bound of $\lambda_{max}(A)$.

Define $\lambda_{max} = \lambda_{max}(A)$. We can restate the separation problem as:

$$\max_{\mu}: \quad (2\mu+q)^T\bar{x} + \frac{1}{\lambda_{max}}\bar{w} - \frac{1}{\lambda_{max}}\max_x \left\{x^T(A-\lambda_{max}I)x + 2\lambda_{max}\mu^T x\right\}$$

The optimal solution for the inner minimization is any x satisfying

$$(A-\lambda_{max}I)x = -\lambda_{max}\mu$$

Since we had the condition $\lambda_{max}\mu = (A - \lambda_{max}I)\pi$, we have that $-\pi$ is a maximizer. The resulting maximum value is

$$-\pi^T(A - \lambda_{max}I)\pi$$

and the separation problem becomes:

$$\max_{\pi} : \quad \frac{1}{\lambda_{max}}\pi^T(A - \lambda_{max}I)\pi + \frac{2}{\lambda_{max}}\bar{x}^T(A - \lambda_{max}I)\pi + q^T\bar{x} + \frac{1}{\lambda_{max}}\bar{w}$$

Again, the separation problem is a convex QP. Its optimal solution is any π satisfying

$$\frac{1}{\lambda_{max}}(A - \lambda_{max}I)\pi = \frac{-1}{\lambda_{max}}(A - \lambda_{max}I)\bar{x}$$

so setting

$$\pi = -\bar{x}$$

gives a maximizer. The resulting optimal μ is

$$\mu^* = \frac{1}{\lambda_{max}}(A - \lambda_{max}I)\pi = \frac{1}{\lambda_{max}}(\lambda_{max}I - A)\bar{x}$$

Using this and the constraint from the first formulation of the separation problem (which we know will hold with equality) we can get the optimal ν:

$$\begin{aligned}
\nu^* &= -\lambda_{max}\|\mu^*\|^2 + \max_x \left\{ x^T(A - \lambda_{max}I)x + 2\lambda_{max}{\mu^*}^T x \right\} \\
&= -\lambda_{max}\|\mu^*\|^2 + \bar{x}^T(A - \lambda_{max}I)\bar{x} + 2\lambda_{max}{\mu^*}^T\bar{x} \\
&= -\lambda_{max}\|\mu^*\|^2 + \bar{x}^T(A - \lambda_{max}I)\bar{x} - 2\bar{x}^T(A - \lambda_{max}I)\bar{x} \\
&= -\lambda_{max}\|\mu^*\|^2 + \bar{x}^T(\lambda_{max}I - A)\bar{x}
\end{aligned}$$

The reader may verify that these expressions for μ^* and ν^* coincide with (65) and (66) (resp.) when A is diagonal.

3.1.2 Application

Consider an optimization problem with an objective function of the form

$$\min \quad x^T M x \, + \, v^T x \, + \, c, \tag{78}$$

or a constraint of the form

$$x^T M x \, + \, v^T x \, + \, c \geq \; 0, \tag{79}$$

where $M \in \mathbb{R}^d \times \mathbb{R}^d$ is symmetric. By using the spectral decomposition of M to change coordinates, and if necessary adding and subtracting terms of the form x_i^2, and finally scaling, without loss of generality we obtain an expression of the form

$$\sum_{i=1}^{d} x_i^2 \; - \; \sum_{i=1}^{d} \lambda_i x_i^2 \; + \; v^T x \; + \; c, \quad \text{where } \lambda_i > 0 \text{ for all } i.$$

In case of an optimization problem with objective (78), we can lift to an equivalent system of the form

$$\min\{\, z - w \, + \, c \, : \, \text{s.t. } z \geq \sum_{i=1}^{d} x_i^2 \, + \, v^T x, \quad w \leq \sum_{i=1}^{d} \lambda_i x_i^2 \},$$

whose constraint set is exactly of the form (50)-(51).

3.1.3 Example

Consider the bilinear form

$$f(x) \doteq 2(x_1x_2 \,+\, x_1x_3 \,+ x_2x_3)$$

over the unit cube $[0,1]^3$. Writing, for $1 \leq i < j \leq 3$, $f_{ij} = x_ix_j$, the McCormick relaxation for f_{ij} amounts to:

$$f_{ij} \,\geq\, x_i + x_j - 1, \quad f_{ij} \leq \min\{x_i, x_j\}.$$

At $\bar{x} = (1/2, 1/2, 1/2)^T$, the lower bound on $f(\bar{x})$ produced by the McCormick relaxation is zero (for more complex examples see [18]). We show next how our procedures may be used to generate a formulation that proves a positive lower bound on $f(\bar{x})$. We stress that what we have here is an adhoc construction – we plan to return to this topic in a future work.
We have $f(x) = U(x) - L(x)$, where

$$\begin{aligned} U(x) &\doteq (x_1 + x_2)^2 + (x_1 + x_3)^2 + (x_2 + x_3)^2, \\ L(x) &\doteq 2(x_1^2 + x_2^2 + x_3^2). \end{aligned} \tag{80, 81}$$

Now we apply the techniques from Section 3.1. We have $U(x) = x^TQx$ and $L(x) = x^TAx$, where

$$Q = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \qquad A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The Cholesky decomposition of Q is

$$Q = LL^T = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & \sqrt{3/2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{6} & \sqrt{4/3} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} & 1/\sqrt{6} \\ 0 & 0 & \sqrt{4/3} \end{bmatrix}.$$

Let $V\Lambda V^T$ be the eigendecomposition of $L^{-1}AL^{-T} = 2L^{-1}L^{-T} = 2(L^TL)^{-1}$:

$$V = \begin{bmatrix} \sqrt{3}/6 & 1/2 & 2/\sqrt{6} \\ 1/6 & -\sqrt{3}/2 & \sqrt{2}/3 \\ -2\sqrt{2}/3 & 0 & 1/3 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}.$$

The transformation we use is $p = V^TL^Tx$, or $x = L^{-T}Vp$. Note:

$$L^{-T}V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{6} & -\sqrt{3}/6 \\ 0 & 2/\sqrt{6} & -\sqrt{3}/6 \\ 0 & 0 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} \sqrt{3}/6 & 1/2 & 2/\sqrt{6} \\ 1/6 & -\sqrt{3}/2 & \sqrt{2}/3 \\ -2\sqrt{2}/3 & 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & \sqrt{2}/2 & \sqrt{3}/6 \\ 1/\sqrt{6} & -\sqrt{2}/2 & \sqrt{3}/6 \\ -2/\sqrt{6} & 0 & \sqrt{3}/6 \end{bmatrix}.$$

Thus, we have

$$x \in [0,1]^3 \Leftrightarrow \begin{bmatrix} -L^{-T}V \\ L^{-T}V \end{bmatrix} p \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Let $\breve{H}$ be the image of $[0,1]^3$ under the mapping. It can be seen that for any x we have

$$p_3(x) = \frac{2}{\sqrt{3}}\,(x_1 + x_2 + x_3)$$

and thus our point of interest, $\bar{x}$, is mapped to $\bar{p} = (0, 0, \sqrt{3})^T$.

Further, in p-space, $f(x)$ is represented as

$$F(p) \doteq (p_1^2 + p_2^2 + p_3^2) - (2p_1^2 + 2p_2^2 + \frac{1}{2}p_3^2).$$

Consider the "paraboloid cut"

$$p_1^2 + p_2^2 + (p_3 - 2\alpha\sqrt{3})^2 + \epsilon \quad \geq \quad 2p_1^2 + 2p_2^2 + \frac{1}{2}p_3^2. \tag{82}$$

For $\alpha = \epsilon = 1/10$, a calculation shows that (82) is valid for all $p \in \breve{H}$ with $p_3 \geq \sqrt{3}$ (or, informally, it is valid for all $x \in [0,1]^3$ with $\sum_i x_i \geq 3/2$.) In the region of validity, we therefore have

$$F(p) \geq 4\alpha\sqrt{3}p_3 - 12\alpha^2 - \epsilon = \frac{2}{5}\sqrt{3}p_3 - \frac{11}{50}.$$

In other words, for $x \in [0,1]^3$ with $\sum_i x_i \geq 3/2$,

$$f(x) \geq \frac{4}{5}(x_1 + x_2 + x_3) - \frac{11}{50}.$$

Consider now the paraboloid cut (82) with $\alpha = 1/2$ and $\epsilon = 2$. A calculation shows that in that case (82) is valid for all $p \in \breve{H}$ with $p_3 \leq \sqrt{3}$. Where it is valid we get

$$F(p) \geq 4\alpha\sqrt{3}p_3 - 12\alpha^2 - \epsilon = 2\sqrt{3}p_3 - 5,$$

and thus, for $x \in [0,1]^3$ with $\sum_i x_i \leq 3/2$,

$$f(x) \geq 4(x_1 + x_2 + x_3) - 5.$$

We now have a disjunction between two polyhedra:

$$\Theta \doteq \left\{ (x,f) \ : \ x \in [0,1]^3, \ \sum_j x_j \geq 3/2, \ f \geq \frac{4}{5}(x_1 + x_2 + x_3) - \frac{11}{50} \right\}, \quad \text{and}$$

$$\Pi \doteq \left\{ (x,f) \ : \ x \in [0,1]^3, \ \sum_j x_j \leq 3/2, \ f \geq 4(x_1 + x_2 + x_3) - 5 \right\}.$$

Thus, solving the linear program

$$\begin{aligned} & \min \ f \\ \text{s.t.} \quad & (x,f) \in conv(\Theta \cup \Pi) \\ & x = \bar{x} \end{aligned}$$

yields a valid lower bound on $f(\bar{x})$. The value of this LP is (slightly greater than) .3114275. Using LP duality, one also obtains the valid cut

$$f(x) \geq 1.245714(x_1 + x_2 + x_3) - 1.5571435$$

which likewise implies $f(\bar{x}) \geq .3114275$.

As the example makes clear, issues of numerical precision are of paramount importance in this context. We plan to return to these questions in a future work.

References

[1] F. Alizadeh and D. Goldfarb, Second-Order Cone Programming, *Mathematical Programming* **95** (2001), 3 – 51.

[2] F. Al-Khayyal, and J. Falk, Jointly constrained biconvex programming, *Math. Oper. Res.* **8** (1983), 273-286.

[3] K.M. Anstreicher, Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming, *J. Global Optimization* **43** (2009), 471–484.

[4] K.M. Anstreicher and S. Burer, Computable representations for convex hulls of low-dimensional quadratic forms, *Mathematical Programming (Series B)* **124** (2010), 33–43.

[5] X. Bao, N.V. Sahinidis, and M. Tawarmalani, Multiterm polyhedral relaxations for nonconvex, quadratically constrained quadratic programs, *Optimization Methods and Software* **24** (2009), 485–504.

[6] A. Ben-Tal and A. Nemirovsky, em Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications (2001) MPS-SIAM Series on Optimization, SIAM, Philadelphia, PA.

[7] P. Belotti, A.J. Miller and M. Namazifar, Valid Inequalities and Convex Hulls for Multilinear Functions, *Electronic Notes in Discrete Mathematics* **36** (2010), 805 - 812.

[8] S. Burer and A. N. Letchford, On non-convex quadratic programming with box constraints, *SIAM Journal on Optimization*, **20** (2009), 1073–1089.

[9] S. Burer and A. N. Letchford, Non-Convex Mixed-Integer Nonlinear Programming: A Survey, *Optimization Online*, February 2012.

[10] S. Cafieri, J. Lee, and L. Liberti, On convex relaxations of quadrilinear terms, *Journal of Global Optimization*, **47** (2010), 661-685.

[11] D. Bienstock, Computational study of a family of mixed-integer quadratic programming problems, *Math. Programming* **74** (1996), 121– 140.

[12] D. Bienstock, Eigenvalue techniques for proving bounds for convex objective, nonconvex programs, *Proc. IPCO 2010*.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004).

[14] S. Ceria and J. Soares, Convex programming for disjunctive convex optimization, *Mathematical Programming* **86** (1999), 595 - 614.

[15] I.R. de Farias JR., E.L. Johnson and G.L. Nemhauser, Facets of the Complementarity Knapsack Polytope, *Mathematics of Operations Research* **27** (2002), 210 –226.

[16] D. Goldfarb and G. Iyengar, Robust Portfolio Selection Problems, *Mathematics of Operations Research*, **28** (2002), 1 – 38.

[17] G.H. Golub, Some modified matrix eigenvalue problems, *SIAM Review* **15** (1973), 318 – 334.

[18] M. Kilinc, J. Linderoth and J. Luedtke, Effective Separation of Disjunctive Cuts for Convex Mixed Integer Nonlinear Programs, *Optimization Online* (2010).

[19] A.B. Keha, I.R. de Farias JR. and G.L. Nemhauser, A Branch-and-Cut Algorithm without Binary Variables for Nonconvex Piecewise Linear Optimization," *Operations Research* **54** (2006), 847 – 858.

[20] J. Luedtke, M. Namazifar and J. Linderoth, Some Results on the Strength of Relaxations of Multilinear Functions, Optimization Online, August 2010.

[21] G.P. McCormick, Computability of global solutions to factorable nonconvex programs: Part IConvex underestimating problems. *Math. Program.* **10** (1976), 147 – 175.

[22] G.L. Nemhauser and L.A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York (1988).

[23] A. Michalka, Ph.D. Dissertation, Columbia University (in preparation).

[24] I. Pólik and T. Terlaky, A survey of the S-lemma, *SIAM Review* **49** (2007), 371 – 418.

[25] A. Qualizza, P. Belotti, F. Margot, Linear Programming Relaxations of Quadratically Constrained Quadratic Programs, manuscript, 2011.

[26] A. Saxena, P. Bonami and J. Lee, Convex relaxations of non-convex mixed integer quadratically constrained programs: Extended formulations. *Mathematical Programming B* **124** (2010) 383 – 411.

[27] A. Saxena, P. Bonami and J. Lee, Convex relaxations of non-convex mixed integer quadratically constrained programs: Projected formulations. To appear, *Mathematical Programming*.

[28] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, Kluwer, Dordrecht (1998).

[29] Sherali, H. D. and Adams, W. P. A, Reformulation-Linearization Technique (RLT) for Semi-Infinite and Convex Programs under Mixed 0-1 and General Discrete Restrictions, *Discrete Applied Mathematics*, **157**, (2009) 1319–1333.

[30] R.A. Stubbs and S. Mehrotra, A branch-and-cut method for 0-1 mixed convex programming, *Mathematical Programming* **86** (1999), 515 - 532.

[31] V. A. Yakubovich, S-procedure in nonlinear control theory, Vestnik Leningrad University, **1** (1971) 62–777.