Convex Hulls for Non-Convex Mixed-Integer Quadratic Programs with Bounded Variables

Laura Galli[*] Adam N. Letchford[†]

December 2019

Abstract

We consider non-convex mixed-integer quadratic programs in which all variables are explicitly bounded. Many exact methods for such problems use additional variables, representing products of pairs of original variables. We study the convex hull of feasible solutions in this extended space. Some other approaches use bit representation to convert bounded integer variables into binary variables. We study the convex hulls associated with some of these formulations as well. We also present extensive computational results.

Keywords: mixed-integer nonlinear programming, global optimisation, polyhedral combinatorics, cutting planes

1 Introduction

A wide range of problems in Operational Research, Statistics and Finance can be formulated as *mixed-integer quadratic programs* (MIQPs); that is, optimisation problems with a mixture of continuous and integer-constrained variables, linear constraints, and a quadratic objective function. When the objective function is convex, one can use a generic algorithm for convex mixed-integer nonlinear programs, such as those surveyed in [5, 13, 19]. The non-convex case, however, still presents a formidable challenge; see, e.g., [9–11, 24–26].

In this paper, we focus on non-convex MIQPs in which all variables are explicitly bounded. A classic approach to such problems, due to McCormick [20], introduces additional variables that represent squares or products of original variables, and uses these to construct a linear programming relaxation of the MIQP. (One can also construct semidefinite programming relaxations; see, e.g., [9, 11].)

[*]Department of Computer Science, University of Pisa, Largo B. Pontecorvo 3, 56124 Pisa, Italy. E-mail: laura.galli@unipi.it

[†]Department of Management Science, Lancaster University, Lancaster LA1 4YW, United Kingdom. E-mail: a.n.letchford@lancaster.ac.uk

Another classic approach, due to Watters [27], is to use *bit representation* to convert bounded integer variables into binary variables. This enables one to reformulate the MIQP as a mixed 0-1 quadratic (or sometimes even linear) program, which may be easier to solve. (In fact, the reformulation can be done in several different ways; see, e.g., [3, 15, 18, 27].)

In this paper, we consider the convex hull of feasible solutions for the approach of McCormick and for two of the approaches based on bit representation. We explore conditions under which the convex hulls are polyhedral, and present results concerned with their dimension. We then consider various known and new families of valid inequalities, and examine conditions under which they define faces of maximal dimension. Finally, we give extensive computational results, on several different kinds of test problem, to gain insight into the strengths and weaknesses of the three approaches.

The paper has the following structure. Section 2 reviews the literature. Sections 3 and 4 are concerned with the McCormick and bit-representation approaches, respectively. Section 5 contains the computational results, and Section 6 contains some concluding remarks.

Throughout the paper, we use the acronyms "LP", "BP" and "QP" for linear, bilinear and quadratic programs, respectively. We also use the prefixes "I" and "MI" for integer and mixed-integer, respectively. So, for example, "MIBP" means "mixed-integer bilinear program". We also assume that MIQPs are written in the following form:

$$\min \left\{ x^T Q x + c \cdot x : \; Ax \le b, \; x \le u, \; x \in \mathbb{R}^n_+, \; x_i \in \mathbb{Z} \; (i \in I) \right\},$$

where $Q \in \mathbb{Q}^{n \times n}$, $c \in \mathbb{Q}^n$, $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$, $u \in \mathbb{Z}^n_+$ and $I \subseteq \{1, \ldots, n\}$. We assume w.l.o.g. that Q is symmetric.

We also use the following notation throughout. We let N denote $\{1, \ldots, n\}$. For all $i \in I$, we let r_i denote $\lfloor \log_2 u_i \rfloor$. We write "conv" for convex hull and "cl" for closure. Finally, we sometimes refer to the following polytopes:

$$\begin{aligned} P &= \left\{ x \in \mathbb{R}^n_+ : \; Ax \le b, \; x \le u \right\} \\ P_I &= \operatorname{conv} \left\{ x \in P : \; x_i \in \mathbb{Z} \; (i \in I) \right\}. \end{aligned}$$

Note that P is the feasible region of the continuous relaxation of the MIQP, and P_I is the so-called *integral hull* of P.

2 Literature Review

In this section, we review the relevant literature, in chronological order. Due to space restrictions, we will be rather selective.

In 1967, Watters [27] showed how to convert bounded integer variables into 0-1 variables. For a given $i \in I$, replace x_i with

$$\sum_{s=0}^{r_i} 2^s \, \tilde{x}_{is},$$

where the $\tilde{x}_{is}$ are new binary variables. If $u_i + 1$ is not a power of two, also add the constraint:

$$\sum_{s=0}^{r_i} 2^s \, \tilde{x}_{is} \le u_i. \tag{1}$$

In 1974, Glover & Woolsey [16] proposed to convert 0-1 QPs into 0-1 LPs, as follows. Replace each quadratic term, say $x_i x_j$, with a new binary variable, say X_{ij}, and add the constraints

$$X_{ij} \ge 0,\ X_{ij} \le x_i,\ X_{ij} \le x_j,\ X_{ij} \ge x_i + x_j - 1. \tag{2}$$

In 1975, Glover [15] showed that a bounded MIBP can be converted into a mixed 0-1 LP having $O\big(\sum_{i \in I} r_i\big)$ variables and constraints, provided that no bilinear term is the product of two continuous variables.

In 1976, McCormick [20] proposed to construct LP relaxations of bounded MIQPs as follows. For each pair $i, j \in N$ with $i < j$ and $Q_{ij} \neq 0$, replace the quadratic term $x_i x_j$ with a new continuous variable, say X_{ij}, and add the linear inequalities

$$X_{ij} \ge 0,\ X_{ij} \le u_j x_i,\ X_{ij} \le u_i x_j,\ X_{ij} \ge u_j x_i + u_i x_j - u_i u_j. \tag{3}$$

Similarly, for each $i \in N$ with $Q_{ii} \neq 0$, replace x_i^2 with X_{ii}, and add

$$X_{ii} \ge 0,\ X_{ii} \le u_i x_i,\ X_{ii} \ge 2u_i x_i - u_i^2. \tag{4}$$

Note that this is a relaxation rather than a reformulation.

In 1986, Adams & Sherali [1] showed that one can strengthen the LP relaxation of the Glover-Woolsey formulation as follows. Take any linear constraint from the original 0-1 QP, say $\alpha \cdot x \le \beta$, and any $i \in N$, and note that the quadratic inequalities $(\alpha \cdot x)x_i \le \beta x_i$ and $(\alpha \cdot x)(1 - x_i) \le \beta(1 - x_i)$ are valid. Linearising them yields $2n$ inequalities of the form:

$$\textstyle\sum_{j \neq i} \alpha_j X_{ij} \le (\beta - \alpha_i) x_i \tag{5}$$

$$\textstyle\sum_{j \neq i} \alpha_j (x_j - X_{ij}) \le \beta(1 - x_i). \tag{6}$$

This approach is now called the *Reformulation-Linearization Technique* (RLT), and the inequalities (5), (6) are called *RLT* inequalities. The RLT has been generalised and applied to non-convex MIQP and other problems [25].

In 1989, Padberg [21] studied the following polytope, which he called the *Boolean quadric polytope*:

$$\text{conv}\left\{ (x, X) \in \{0,1\}^{n + \binom{n}{2}} :\ X_{ij} = x_i x_j \, (1 \le i < j \le n) \right\}.$$

He showed that the constraints (2) define facets, and derived several other families of facet-defining inequalities, such as the following *triangle* inequalities:

$$x_i + x_j + x_k \le X_{ij} + X_{ik} + X_{jk} + 1 \quad (\{i, j, k\} \subseteq N) \tag{7}$$

$$X_{ij} + X_{ik} \le x_i + X_{jk} \quad \big(i \in N,\ \{j, k\} \subseteq N \setminus \{i\}\big). \tag{8}$$

We will let BQP_n denote the Boolean quadric polytope of order n.

In 1993, Ramana [22] studied the following unbounded convex set

$$\mathrm{conv}\left\{(x,X) \in \mathbb{R}^{n+\binom{n+1}{2}} : \; X_{ij} = x_i x_j \; (1 \le i \le j \le n)\right\}.$$

He showed that this set is described by the following linear inequalities:

$$(2s)v \cdot x - \sum_{i \in N} v_i^2 X_{ii} - 2 \sum_{1 \le i < j \le n} v_i v_j X_{ij} \; \le \; s^2 \qquad (v \in \mathbb{R}^n, s \in \mathbb{R}). \tag{9}$$

The validity of (9) follows from the fact that $(v \cdot x - s)^2 \ge 0$. We will follow [8] in calling (9) *positive semidefinite* (psd) inequalities.

In 1993, Boros & Hammer [6] discovered the following family of valid inequalities for BQP_n, which includes all those of Padberg:

$$\sum_{i \in N} v_i(2s+1-v_i)x_i - 2 \sum_{1 \le i < j \le n} v_i v_j X_{ij} \; \le \; s(s+1) \qquad (v \in \mathbb{Z}^n, s \in \mathbb{Z}). \tag{10}$$

The validity of (10) follows from the fact that $(v \cdot x - s)(v \cdot x - s - 1) \ge 0$ when v and s are integral, and the fact that $x_i = x_i^2$ when x_i is binary.

In 1997, Harjunkoski *et al.* [18] gave a new approach for bounded IBPs. First, each integer variable x_i is replaced with binary variables $\tilde{x}_{is}$, as usual. Then, for all pairs $i, j \in N$, and for $s = 0, \ldots, r_i$, an additional continuous variable, say y_{isj}, is defined, which represents the product $\tilde{x}_{is} x_j$. They then replace all terms of the form $x_i x_j$ with $\sum_{s=0}^{r_i} 2^s y_{isj}$. Finally, they add the following linear inequalities for all pairs i, j and for $s = 0, \ldots, r_i$:

$$y_{isj} \ge 0, \; y_{isj} \le u_j \tilde{x}_{is}, \; y_{isj} \le x_j, \; y_{isj} \ge u_j \tilde{x}_{is} + x_j - u_j. \tag{11}$$

The result is again a mixed 0-1 LP.

In 1998, Yajima & Fujie [28] studied the following convex set:

$$\mathrm{conv}\left\{(x,X) \in [0,1]^{n+\binom{n+1}{2}} : \; X_{ij} = x_i x_j \; (1 \le i \le j \le n)\right\}.$$

We will follow [8] in calling it QPB_n (short for "quadratic programming with box constraints"). It was shown in [8] that any inequality valid for BQP_n is valid also for QPB_n. See also [2].

In 2008, Billionnet *et al.* [3] rediscovered the approach in [18], in the context of bounded IQPs. They also used the RLT to derive cutting planes for the resulting mixed 0-1 LPs. In addition, they noted that the following equations are valid for all $\{i, j\} \subseteq N$:

$$\sum_{s=0}^{r_i} 2^s \, y_{isj} \; = \; \sum_{s=0}^{r_j} 2^s \, y_{jsi}. \tag{12}$$

In 2012, Billionnet *et al.* [4] used bit representation to show that, under certain conditions, non-convex bounded MIQPs can be converted to convex mixed 0-1 QPs.

In 2013, Gupte *et al.* [17] adapted the method in [18] to the case of bounded MIBPs in which each bilinear term is the product of an integer variable and a continuous variable. They also derived some cutting planes, as follows. For a given $i \in I$, let S_i^0 and S_i^1 be the sets of bits that take the value zero or one, respectively, in the bit representation of u_i. For any $s \in S_i^0$, if we let $C(s)$ denote $\{t \in S_i^1 : t > s\}$, then the cover inequality

$$\sum_{t \in C(s) \cup \{s\}} \tilde{x}_{it} \leq |C(s)| \tag{13}$$

is valid. Following the RLT, we can multiply each such inequality by either x_j or $u_j - x_j$, for any $j \in N \setminus I$, to obtain the inequalities

$$\textstyle\sum_{t \in C(s) \cup \{s\}} y_{itj} \leq |C(s)| x_j \tag{14}$$

$$u_j \textstyle\sum_{t \in C(s) \cup \{s\}} \tilde{x}_{it} - \sum_{t \in C(s) \cup \{s\}} y_{itj} \leq |C(s)|(u_j - x_j). \tag{15}$$

In 2014, Burer & Letchford [10] studied the following unbounded convex set:

$$\text{cl conv} \left\{ (x, X) \in \mathbb{R}^{n + \binom{n+1}{2}} : \, X_{ij} = x_i x_j \, (1 \leq i \leq j \leq n), \; x_i \in \mathbb{Z} \, (i \in I) \right\}.$$

They pointed out that, when $v_i = 0$ for all $i \in N \setminus I$, the psd inequalities (9) can be strengthened to:

$$(2s+1) v \cdot x - \sum_{i \in I} v_i^2 X_{ii} - 2 \sum_{\{i,j\} \subseteq I} v_i v_j X_{ij} \; \leq \; s(s+1). \tag{16}$$

Note that these reduce to (10) when all variables are binary.

Finally, we mention that there are several other papers concerned with strengthening McCormick and/or RLT relaxations of bounded MIQPs (e.g., [7, 9, 11, 14, 24]). We omit details for the sake of brevity.

3 Convex Hulls in the McCormick Space

In this section, we consider convex sets associated with the McCormick approach. For ease of exposition, we consider the continuous case in Subsection 3.1 and the mixed-integer case in Subsection 3.2. Throughout this section, for notational convenience, we identify X_{ij} and X_{ji} for $1 \leq i < j \leq n$. We also assume for simplicity that all possible X variables are present; that is, that the variable X_{ij} has been defined for $1 \leq i \leq j \leq n$. (When this is not the case, some of our valid inequalities may not be applicable.)

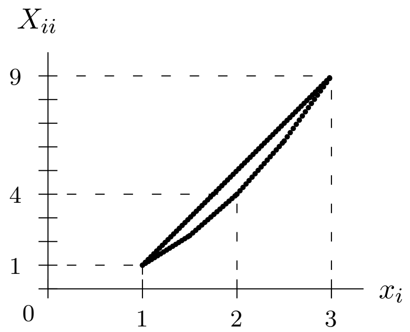

Figure 1: Projection of $C(P)$ when $\ell_i = 1$ and $u_i = 3$.

3.1 The continuous case

The continuous case arises when $I = \emptyset$. In this case, it helps to use the following notation. Given a point $x^* \in P$, we let $M(x^*)$ denote the corresponding point in $\mathbb{R}^{n+\binom{n+1}{2}}$; that is, the pair (x^*, X^*), where $X^*_{ij} = x^*_i x^*_j$ for $1 \leq i \leq j \leq n$. We are interested in the following convex hull:

$$C(P) = \text{conv}\Big\{ M(x) \in \mathbb{R}^{n+\binom{n+1}{2}} : \ x \in P \Big\}.$$

We remark that the case in which P is the unit hypercube was studied in [2, 8, 28], and the case in which P is a simplex was studied in [2].

The following lemma can be proved in the same way as Lemma 2 in [8]:

Lemma 1 *A pair* (x, X) *is an extreme point of* $C(P)$ *if and only if* $x \in P$ *and* $(x, X) = M(x)$.

The fact that P is a polytope implies that $C(P)$ is a closed and bounded convex set. It is however not polyhedral, except for one trivial case.

Lemma 2 $C(P)$ *is polyhedral if and only if* P *consists of a single point.*

Proof. If P is a point, then $C(P)$ is a point as well. If P is not a point, then there exists at least one variable, say x_i, that can take more than one value. Suppose that x_i can take values in the interval $[\ell_i, u_i]$, and let P' be the projection of $C(P)$ into a two-dimensional subspace, having x_i and X_{ii} as axes. Figure 1 shows P' for the case $\ell_i = 1$, $u_i = 3$. As one can see, the projection is not polyhedral, since the lower convex envelope is a segment of the parabola defined by the convex quadratic inequality $X_{ii} \geq x_i^2$. $\square$

The next thing to settle is the dimension of $C(P)$. Suppose that P has dimension k. Then its affine hull can be described by $n - k$ linearly independent equations. Without loss of generality, assume that these equations take the form

$$x_r = \sum_{i=1}^{k} \alpha_{ir} x_i + \beta_r \qquad (r = k+1, \ldots, n). \tag{17}$$

These equations are clearly also valid for $C(P)$. Moreover, using the RLT, we can construct $n(n-k)$ additional valid equations for $C(P)$. It suffices to multiply each of the equations (17) by each x variable in turn, and linearise, to obtain:

$$X_{jr} = \sum_{i=1}^{k} \alpha_{ir} X_{ij} + \beta_r x_j \qquad (j = 1, \ldots, n;\ r = k+1, \ldots, n). \tag{18}$$

It turns out that the equations (17) and (18) give a complete description of the affine hull of $C(P)$. However, when $k < n-1$, the equations turn out to be linearly dependent, so there is some redundancy. This is a consequence of the following theorem and its proof.

Theorem 1 *If $\dim(P) = k$, then $\dim\big(C(P)\big) = k + \binom{k+1}{2}$.*

Proof. Suppose $\dim(P) = k$. Let x^* be a point in the relative interior of P, and let ϵ be a small positive quantity. For any $i \in \{1, \ldots, k\}$, we can construct another point in P, which we call x^i, by taking x^*, increasing x_i by ϵ, and adjusting $x_{k+1}, \ldots, x_n$ according to (17). Similarly, for any $i \in \{1, \ldots, k\}$, we can construct another point in P, which we call $\tilde{x}^i$, by decreasing x_i by ϵ, and adjusting $x_{k+1}, \ldots, x_n$ accordingly. Moreover, for any pair $\{i, j\}$ with $1 \le i < j \le n$, we can construct yet another point in P, which we call x^{ij}, by increasing both x_i and x_j by ϵ, and adjusting $x_{k+1}, \ldots, x_n$. This gives $1 + 2k + \binom{k}{2} = 1 + k + \binom{k+1}{2}$ points in P. The corresponding points in $C(P)$ are $M\big(x^*\big)$, $M\big(x^1\big)$ and so on. These points are easily shown to be affinely independent. So $\dim\big(C(P)\big) \ge k + \binom{k+1}{2}$.

Now, consider the subset of (18) obtained by deleting the equations with $j > r$. If we order the remaining equations by index r first and index j second, we find that each equation contains exactly one variable that does not appear in the previous equations in the ordered list. The corresponding matrix of left-hand side coefficients is easily shown to have full row rank, for example, by performing Gaussian elimination. So the equations are linearly independent, and there are $k(n-k) + \binom{n-k}{2}$ of them. Moreover, the equations (17) do not involve any X variables, and they are therefore also linearly independent. This makes $(k+1)(n-k) + \binom{n-k}{2}$ linearly independent equations in total. So $\dim\big(C(P)\big) \le n + \binom{n}{2} - (k+1)(n-k) - \binom{n-k}{2} = k + \binom{k+1}{2}$. □

The RLT can also be used to derive valid inequalities for $C(P)$. Suppose that the inequalities $\alpha \cdot x \le \beta$ and $\gamma \cdot x \le \delta$ define distinct facets of P.

Then the quadratic inequality $(\beta - \alpha \cdot x)(\delta - \gamma \cdot x) \geq 0$ is valid for $C(P)$. Linearising yields:

$$(\gamma\beta + \delta\alpha) \cdot x - \sum_{i,j \in N} \alpha_i \gamma_j X_{ij} \leq \beta\delta. \tag{19}$$

It turns out that these RLT inequalities are facet-defining.

Proposition 1 *Each RLT inequality (19) defines a distinct facet of $C(P)$.*

Proof. Let F_1 and F_2 be the facets of P defined by the inequalities $\alpha \cdot x \leq \beta$ and $\gamma \cdot x \leq \delta$, and let G be the face of $C(P)$ defined by the corresponding RLT inequality. From the way in which the RLT inequality is derived, together with Lemma 1, it follows that an extreme point of $C(P)$ lies in G if and only if it is of the form $M(x^*)$ for some $x^* \in F_1 \cup F_2$. This immediately proves that each RLT inequality defines a distinct face of $C(P)$. The proof that they define facets is similar to that of Theorem 1. □

Corollary 1 *The McCormick inequalities (3) define facets of $C(P)$ for all pairs $\{i, j\} \subset N$ such that the bounds $0 \leq x_i \leq u_i$ and $0 \leq x_j \leq u_j$ define facets of P. Also, the McCormick inequality $X_{ii} \leq u_i x_i$ defines a facet of $C(P)$ for all $i \in N$ such that the bounds $0 \leq x_i \leq u_i$ define facets of P.*

Proof. Each of the inequalities (3) can be derived as an RLT inequality from two distinct bounds. For example, the inequality $X_{ij} \geq u_j x_i + u_i x_j - u_i u_j$ can be derived from the bounds $x_i \leq u_i$ and $x_j \leq u_j$, via the quadratic inequality $(u_i - x_i)(u_j - x_j) \geq 0$. Similarly, the inequality $X_{ii} \leq u_i x_i$ can be derived from the bounds $x_i \geq 0$ and $x_i \leq u_i$, via the quadratic inequality $x_i(u_i - x_i) \geq 0$. □

The following lemma deals with the remaining inequalities in (4).

Lemma 3 *For any $i \in N$, the McCormick inequalities $X_{ii} \geq 0$ and $X_{ii} \geq 2u_i x_i - u_i^2$ do not define maximal faces of $C(P)$.*

Proof. Any point in $C(P)$ with $X_{ii} = 0$ also has $x_i = 0$, and therefore also satisfies the inequality $X_{ii} \leq u_i x_i$ at equality. Similarly, any point in $C(P)$ with $X_{ii} = 2u_i x_i - u_i^2$ also has $x_i = u_i$, and therefore also satisfies the inequality $X_{ii} \leq u_i x_i$ at equality. □

Finally, we observe that Ramana's psd inequalities (9) are also valid for $C(P)$. The following result can be proved in the same way as Theorem 2 in [8].

Proposition 2 *The psd inequality (9) defines a maximal face of $C(P)$ if and only if there exists a point x in the relative interior of P with $v \cdot x = s$. In that case, the face has dimension $(k-1) + \binom{k}{2} = dim\big(C(P)\big) - (k+1)$, where $k = dim(P)$.*

We remark that Proposition 2 implies Lemma 3, since the inequalities mentioned in the lemma can be viewed as trivial psd inequalities.

An interesting question is whether the RLT and psd inequalities give a complete description of $C(P)$ when $n = 2$. (It is shown in [2] that they do not give a complete description when $n = 3$ and P is a cube.)

3.2 The mixed-integer case

Now we consider the mixed-integer case, in which I may be non-empty. We will find it helpful to use the following notation:

$$\begin{aligned} S(P,I) &= \big\{x \in P : \ x_i \in \mathbb{Z} \ (i \in I)\big\} \\ C(P,I) &= \text{conv}\left\{M(x) \in \mathbb{R}^{n+\binom{n+1}{2}} : \ x \in S(P,I)\right\}. \end{aligned}$$

(Here, "S" stands for "solutions" and "C" stands for "convex hull".)

The following lemma is a straightforward analogue of Lemma 2.

Lemma 4 *$C(P,I)$ is polyhedral if and only if $|S(P,I)|$ is finite.*

Lemma 4 implies of course that $C(P,I)$ is polyhedral in the pure integer case (i.e., when $I = N$). We remark that, if $|S(P,I)|$ is finite in the mixed-integer case, then there must exist rationals r_{ij} and p_i such that:

$$x_i = \sum_{j \in I} r_{ij} x_j + p_i \qquad (i \in N \setminus I, \ x \in P).$$

That is, $C(P,I)$ can be polyhedral in the mixed-integer setting only if the continuous variables effectively play the role of slack variables. This implies of course that $\dim(P) \leq |I|$ in this case.

Interestingly, no-one seems to have worked on computing the affine hull of $C(P,I)$ for general P. This is surprising, since a characterisation of the affine hull would appear to be essential if one wished to derive necessary and/or sufficient conditions for valid inequalities to define facets (or maximal faces) of $C(P,I)$.

A first observation is that, if $\alpha \cdot x = \beta$ is an implicit equation for P_I, then it is also an implicit equation for $C(P,I)$. Moreover, given any such equation, we can use the RLT to derive further implicit equations for $C(P,I)$, as in the previous subsection. The proof of Theorem 1 then yields the following lemma:

Lemma 5 *If $\dim\big(P_I\big) = k$, then $\dim\big(C(P,I)\big) \leq k + \binom{k+1}{2}$.*

Unfortunately, computing the dimension of P_I is $\mathcal{NP}$-hard in general (by an easy reduction from the problem of testing feasibility of an ILP). Not only that, but the dimension of $C(P,I)$ can be smaller than suggested by Lemma 5. As a trivial example, if x_i is binary, then we have the additional implicit equation $x_i = X_{ii}$. This is a special case of the following result.

Lemma 6 *Suppose there exists a vector $v \in \mathbb{Z}^n$ and scalars $s_1, s_2 \in \mathbb{Z}$ such that:*

- $v_i = 0$ *for all* $i \in N \setminus I$;
- $v^T x \in \{s_1, s_2\}$ *for all* $x \in S(P, I)$.

Then the equation

$$(s_1 + s_2) v \cdot x + \sum_{i \in N} v_i^2 X_{ii} + 2 \sum_{1 \le i < j \le n} v_i v_j X_{ij} + s_1 s_2 = 0$$

is satisfied by all points in $C(P, I)$.

Proof. The quadratic equation $(v^T x - s_1)(v^T x - s_2) = 0$ is satisfied by all $x \in S(P, I)$. Expanding and linearising this equation yields the result. □

Perhaps surprisingly, there may exist additional implicit equations that cannot be found by the methods used so far. For example, suppose that $n = 2$ and $S(P, I)$ contains the following six points: $(0, 1)$, $(1, 0)$, $(1, 1)$, $(2, 1)$, $(1, 2)$ and $(3, 1)$. One can verify that these points satisfy the quadratic equation $(x_1 - 1)(x_2 - 1) = 0$. As a result, the equation $x_1 + x_2 - X_{12} = 1$ is valid for $C(P, I)$. This is a special case of the following result.

Proposition 3 *Suppose there exists vectors $v^1, v^2 \in \mathbb{Z}^n$ and scalars $s_1, s_2 \in \mathbb{Z}$ such that all $x \in S(P, I)$ satisfy $\left(v^1 \cdot x = s_1\right) \vee \left(v^2 \cdot x = s_2\right)$. Then the equation*

$$\left(s_2 v^1 + s_1 v^2\right) \cdot x + \sum_{i \in I} v_i^1 v_i^2 X_{ii} + \sum_{1 \le i < j \le n} \left(v_i^1 v_j^2 + v_i^2 v_j^1\right) X_{ij} + s_1 s_2 = 0$$

is satisfied by all points in $C(P, I)$.

Proof. All points in $S(P, I)$ satisfy the quadratic equation $\left(v^1 \cdot x - s_1\right)\left(v^2 \cdot x - s_2\right) = 0$. Expanding and linearising this equation yields the result. □

An important open question is whether the equations mentioned in this subsection completely describe the affine hull of $C(P, I)$. Another open question is whether the conditions of Lemma 6 and Proposition 3 can be checked in polynomial time. We suspect that, unfortunately, these tasks are again $\mathcal{NP}$-hard.

Next, we make some remarks about the RLT inequalities. Note that, if one wishes to generate RLT inequalities that are as strong as possible, one should multiply pairs of inequalities that define distinct facets of P_I, rather than of P. This may not be easy in practice, however, either because one does not know many facets of P_I, or because the number of known facets is huge. Moreover, even if one does use facets of P_I, the resulting RLT inequality is not guaranteed to define a facet of $C(P, I)$ in general. This is shown in the following example.

Example 1 *Suppose that* $n = |I| = 2$*, and let both* P *and* P_I *be defined by the inequalities* $0 \le x_1 \le 2$*,* $0 \le x_2 \le 2$ *and* $x_1 + x_2 \le 3$*. One can check (either by hand or with the help of a computer) that* $dim\big(P_I\big) = 2$ *and* $dim\big(C(P,I)\big) = 5$*. One can also check that the inequalities* $x_1 \le 2$ *and* $x_2 \le 2$ *define distinct facets of* P_I*. Thus, the quadratic inequality* $(2 - x_1)(2 - x_2) \ge 0$ *is satisfied by all points in* $S(P,I)$*. Note however that this quadratic inequality is satisfied at equality by only four points in* $S(P,I)$*:* $(0,2)$*,* $(1,2)$*,* $(2,0)$ *and* $(2,1)$*. So the corresponding RLT inequality,* $2x_1 + 2x_2 - X_{12} \le 4$*, is satisfied at equality by only four extreme points of* $C(P,I)$*. It therefore cannot define a facet of* $C(P,I)$*.* □

In general, finding a necessary and sufficient condition for an RLT inequality to define a facet looks like a complex task. The same applies to the psd inequalities (9), the strengthened psd inequalities (16), and the inequalities mentioned in [14, 24].

We remark that, for each $i \in I$, the McCormick inequalities $X_{ii} \ge 0$ and $X_{ii} \ge 2u_i x_i - u_i^2$ are dominated by the inequalities $X_{ii} \ge x_i$ and $X_{ii} \ge (2u_i - 1)x_i - u_i(u_i - 1)$, which are special cases of (16). We do not have conditions for these inequalities to define facets either.

We close this section with one last result.

Proposition 4 *Suppose the inequality* $\alpha \cdot x + \beta \cdot X \le \gamma$ *is valid for* QPB_n*. Then the "stretched" inequality*

$$\sum_{i \in N} \frac{\alpha_i}{u_i} x_i + \sum_{1 \le i \le j \le n} \frac{\beta_{ij}}{u_i u_j} X_{ij} \le \gamma$$

is valid for $C(P,I)$*.*

Proof. By definition, $C(P,I)$ is contained in the following convex set:

$$\text{conv}\left\{(x,X) \in \mathbb{R}_+^{n+\binom{n+1}{2}} : x \le u, \; X_{ij} = x_i x_j \, (1 \le i \le j \le n)\right\}.$$

This latter set can be obtained by stretching QPB_n appropriately. (It suffices to multiply x_i by u_i and X_{ij} by $u_i u_j$). □

Although this result may appear trivial, it can yield facet-defining inequalities that are different from any mentioned so far. Suppose, for example, that $n = 3$, $I = N$ and $P = [0,2]^3$, and consider Padberg's triangle inequality (7). From the result of Burer & Letchford [8] mentioned in Section 2, the triangle inequality is valid for QPB_n as well as BQP_n. Applying Proposition 4 and multiplying the resulting inequality by 4 then yields:

$$2x_1 + 2x_2 + 2x_3 \le X_{ij} + X_{ik} + X_{jk} + 4.$$

One can check that this inequality defines a facet of $C(P,I)$, but is not an RLT, psd or strengthened psd inequality.

4 Convex Hulls from Bit Representation

In this section, we study some convex hulls that arise when one uses bit representation. We will see that bit representation allows one to obtain stronger LP relaxations. Of course, this comes at the cost of using more variables and constraints. Moreover, as we will see, the convex hulls become even more complex.

The rest of this section consists of three subsections, concerned with the approaches of Watters [27], Harjunkowski *et al.* [18] and Glover & Woolsey [16], respectively. Throughout, we let L denote $\sum_{i \in I}(r_i + 1)$ and $\bar{n}$ denote $|N \setminus I|$.

4.1 Polytopes from the Watters approach

Recall that Watters approach can be used to convert a bounded non-convex MIQP into a non-convex mixed 0-1 QP. In this subsection, we look at the convex hull of solutions to this mixed 0-1 QP, which we denote by $\tilde{P}$. Note that $\tilde{P}$ is a polytope, and it "lives" in $\mathbb{R}^{L+\bar{n}}$. Although $\tilde{P}$ is not our primary object of interest, results on $\tilde{P}$ will prove to be useful in the following two subsections.

As usual, we start with some remarks about the dimension. Clearly, any implicit equation for P_I can be converted into an implicit equation for $\tilde{P}$. Thus, if $\dim(P_I) = n - k$, then $\dim(\tilde{P}) \leq L + \bar{n} - k$. Unfortunately, there may be additional implicit equations, as the following example shows.

Example 2 *Suppose that $n = |I| = 2$, $u_1 = u_2 = 2$, and the system $Ax \leq b$ consists of the constraints $x_1 \leq 2x_2$ and $x_2 \leq 2x_1$. One can check that $S(P, I)$ contains five feasible solutions: $(0,0)$, $(1,1)$, $(1,2)$, $(2,1)$ and $(2,2)$. So P_I is full-dimensional. Applying bit representation yields the following five extreme points of $\tilde{P}$:*

- $\tilde{x}_{10} = \tilde{x}_{11} = \tilde{x}_{20} = \tilde{x}_{21} = 0$;
- $\tilde{x}_{10} = \tilde{x}_{20} = 1$, $\tilde{x}_{11} = \tilde{x}_{21} = 0$;
- $\tilde{x}_{10} = \tilde{x}_{21} = 1$, $\tilde{x}_{11} = \tilde{x}_{20} = 0$;
- $\tilde{x}_{11} = \tilde{x}_{20} = 1$, $\tilde{x}_{10} = \tilde{x}_{21} = 0$;
- $\tilde{x}_{11} = \tilde{x}_{21} = 1$, $\tilde{x}_{10} = \tilde{x}_{20} = 0$.

One can check that these five solutions satisfy the equation $\tilde{x}_{10} + \tilde{x}_{11} = \tilde{x}_{20} + \tilde{x}_{21}$. Therefore, $\tilde{P}$ is not full-dimensional. □

In general, determining $\dim(\tilde{P})$ looks difficult, even if one is given an explicit description of the affine hull of P. On the positive side, there are some obvious valid inequalities for $\tilde{P}$:

- The bounds $0 \leq \tilde{x}_{is} \leq 1$ for $i \in I$ and $s = 0, \ldots, r_i$ are valid, and so are the bounds $0 \leq x_i \leq u_i$ for $i \in N \setminus I$.
- The cover inequalities (13) are valid.
- If $\alpha \cdot x \leq \beta$ is any valid inequality for P_I, then the inequality

$$\sum_{i \in I} \alpha_i \sum_{s=0}^{r_i} 2^s \tilde{x}_{is} + \sum_{i \in N \setminus I} \alpha_i x_i \leq \beta.$$

  is valid for $\tilde{P}$.

Unfortunately, there may exist facet-defining inequalities that cannot be obtained in any of the above ways:

Example 3 *Suppose that $n = |I| = 2$, $u_1 = u_2 = 2$, and the system $Ax \leq b$ consists of the single inequality $x_1 + x_2 \leq 3$. Since it is impossible for both variables to take the value 2 simultaneously, the inequality $\tilde{x}_{11} + \tilde{x}_{21} \leq 1$ is valid for $\tilde{P}$. One can check that this inequality defines a facet of $\tilde{P}$, yet does not belong to the any of the families of inequalities mentioned above.* □

In light of the above results, we believe that the polytope $\tilde{P}$ itself warrants further study.

4.2 Hulls from the Harjunkowski approach

We now move on to the approach of Harjunkowski *et al.* As mentioned in Section 2, the approach was originally designed for IBPs, and then extended to IQPs and MIBPs in [3] and [17]. It can also be extended to the case of bounded MIQPs. The procedure is as follows:

1. For $i \in I$, replace x_i with $\sum_{s=0}^{r_i} 2^s \tilde{x}_{is}$ in the objective and constraints.
2. For $i \in I$, $s = 0, \ldots, r_i$ and $j \in N$, introduce a continuous variable y_{isj}, representing $\tilde{x}_{is} x_j$. (We permit $j = i$ here, unlike in the bilinear case.)
3. For each pair $i, j \in N \setminus I$ with $i < j$, introduce a continuous variable X_{ij}, representing $x_i x_j$.
4. Use the y and X variables to linearise the objective function.
5. For $i \in I$, $s = 0, \ldots, r_i$ and $j \in N \setminus I$, add the Harjunkowski constraints (11).
6. For $i, j \in I$, not necessarily distinct, and $s = 0, \ldots, r_i$, add the constraints

$$y_{isj} \geq 0,\ y_{isj} \leq u_j \tilde{x}_{is},\ y_{isj} \leq \sum_{t=0}^{r_j} 2^t \tilde{x}_{jt},\ y_{isj} \geq u_j \tilde{x}_{is} + \sum_{t=0}^{r_j} 2^t \tilde{x}_{jt} - u_j. \tag{20}$$

7. For $i, j \in N \setminus I$ with $i < j$, add the McCormick constraints (3).

8. For $i \in N \setminus I$, add the McCormick constraints (4).

The result of this procedure is a mixed 0-1 LP with $O\big(nL+\bar{n}^2\big)$ variables and $O\big(m+nL+\bar{n}^2\big)$ constraints. Note that, if $Q_{ij} = 0$ for all $i, j \in N \setminus I$, the mixed 0-1 LP is a *reformulation* of the original MIQP; otherwise, it is just a *relaxation*. (In fact, in the former case, the X variables and McCormick constraints can be discarded.)

Given any $x \in S(P, I)$, we let $H(x)$ denote the corresponding feasible solution to the mixed 0-1 LP. (The "H" stands for "Harjunkowski".) Note that $H(x)$ is actually a quadruple $(\tilde{x}, \bar{x}, y, X)$, where $\tilde{x} \in \{0,1\}^L$, $\bar{x} \in \mathbb{R}^{\bar{n}}_+$, $y \in \mathbb{R}^{nL}_+$ and $X \in \mathbb{R}^{\binom{\bar{n}+1}{2}}_+$. We then define

$$C_H(P, I) = \text{conv}\Big\{H(x) : \; x \in S(P, I)\Big\}.$$

The following lemma is an analogue of Lemma 4:

Lemma 7 *$C_H(P, I)$ is polyhedral if and only if $|S(P, I)|$ is finite.*

Now, observe that any valid linear equation or inequality for $C(P, I)$ can be converted into a valid equation or inequality for $C_H(P, I)$ using the following linear identities:

$$\begin{array}{ll} x_i = \sum_{s=0}^{r_i} 2^s \tilde{x}_{is} & (i \in I) \\ X_{ij} = \sum_{s=0}^{r_i} 2^s y_{isj} & (i \in I, j \in N). \end{array}$$

In other words, $C(P, I)$ is the *projection* of $C_H(P, I)$ into the space of the x and X variables. This implies that, given any convex relaxation that uses "McCormick" variables, we can find a convex relaxation that uses "Harjunkowski" variables, and is at least as strong.

The following two subsections present some equations and inequalities that are *not* implied by those for $C(P, I)$.

4.2.1 Additional equations

Observe that, for $\{i, j\} \subseteq I$, the equation (12) is an implicit equation for $C_H(P, I)$ that is not implied by those for $C(P, I)$. Some more are presented in the following propositions.

Proposition 5 *Consider an implicit equation for $\tilde{P}$, written in the form:*

$$\alpha \cdot \tilde{x} + \sum_{i \in N \setminus I} \beta_i \cdot x_i = \gamma, \tag{21}$$

and suppose that it is not implied by implicit equations for P_I. (That is, it is not the case that $\alpha_{is} = 2\alpha_{i,s-1}$ for $i \in I$ and $s = 1, \ldots, r_i$.) Then the *following are implicit equations for $C_H(P, I)$ that are not implied by implicit equations for $C(P, I)$:*

$$\begin{array}{lr} \sum_{i \in I} \sum_{s=0}^{r_i} \alpha_{is} y_{isj} + \sum_{i \in N \setminus I} \beta_i \sum_{s=0}^{r_j} 2^s y_{jsi} = \gamma \sum_{s=0}^{r_j} 2^s \tilde{x}_{js} & (j \in I) \\ \sum_{i \in I} \sum_{s=0}^{r_i} \alpha_{is} y_{isj} + \sum_{i \in N \setminus I} \beta_i X_{ij} = \gamma x_j & (j \in N \setminus I). \end{array}$$

Proof. To show that the equations are valid, it suffices to multiply (21) by each variable x_j in turn, and then linearise the result. The fact that they are not implied by equations for $C(P, I)$ comes from the stated condition on α. □

Proposition 6 *Consider an implicit equation for P_I that involves only continuous variables, written in the form:*

$$\sum_{i \in N \setminus I} \beta_i \cdot x_i = \gamma. \tag{22}$$

The following are implicit equations for $C_H(P, I)$ that are not implied by implicit equations for $C(P, I)$:

$$\sum_{i \in N \setminus I} \beta_i y_{jti} = \gamma \tilde{x}_{jt} \qquad (j \in I,\ t = 0, \ldots, r_j).$$

Proof. To show that the equations are valid, it suffices to multiply (22) by each variable $\tilde{x}_{jt}$ in turn, and then linearise the result. The fact that they are not implied by implicit equations for $C(P, I)$ comes from the fact that they involve only one $\tilde{x}$ variable each. □

In fact, $C_H(P, I)$ can have even more implicit equations, in addition to those already mentioned. For example, when $u_i = 2^k$ for some non-negative integer k, every feasible solution with $\tilde{x}_{ik} = 1$ also has $y_{iki} = 2^k$, which yields the implicit equation $y_{iki} = 2^k \tilde{x}_{ik}$.

4.2.2 Additional inequalities

One can also derive valid inequalities for $C_H(P, I)$ that are not implied by valid inequalities for $C(P, I)$. One way to do this is to use the RLT, as follows. Let $\alpha x \leq \beta$ be any valid inequality for P_I, and let $\gamma \tilde{x} + \delta \bar{x} \leq \epsilon$ be any valid inequality for $\tilde{P}$ that is not implied by valid inequalities for P_I. Form the quadratic inequality $(\beta - \alpha x)(\epsilon - \gamma \tilde{x} - \delta \bar{x}) \geq 0$, and linearise it to obtain a valid inequality for $C_H(P, I)$. Note that this procedure yields the constraints (11), (14), (15) and (20) as special cases. Another interesting special case is obtained by taking a bound $x_j \geq 0$ or $x_j \leq u_j$ as our valid inequality for P_I, where $j \in I$, and taking a cover inequality (13) as our valid inequality for $\tilde{P}$. The resulting valid inequalities for $C_H(P, I)$ are:

$$\sum_{t \in C(s) \cup \{s\}} y_{itj} \leq |C(s)| \sum_{t=0}^{r_j} 2^t \tilde{x}_{jt} \tag{23}$$

$$u_j \sum_{t \in C(s) \cup \{s\}} \tilde{x}_{it} - \sum_{t \in C(s) \cup \{s\}} y_{itj} \leq |C(s)| \left(u_j - \sum_{t=0}^{r_j} 2^t \tilde{x}_{jt} \right). \tag{24}$$

The following two propositions present some additional valid inequalities for $C_H(P, I)$, which are not derived via the RLT.

Proposition 7 *When $i = j$, we can strengthen the constraints (20) as follows:*

- *Replace the first inequality in (20) with*

$$y_{isi} \geq 2^s \, \tilde{x}_{is}. \tag{25}$$

- *Replace the second inequality in (20) with*

$$y_{isi} \leq \lambda^1_{is} \, \tilde{x}_{is}, \tag{26}$$

  where λ^1_{is} is the largest value that x_i can take when $\tilde{x}_{is} = 1$.

- *Replace the fourth inequality in (20) with*

$$y_{isi} \geq \sum_{s=0}^{r_i} 2^s \tilde{x}_{is} + \lambda^0_{is}(\tilde{x}_{is} - 1), \tag{27}$$

  where λ^0_{is} is the largest value that x_i can take when $\tilde{x}_{is} = 0$.

Proof. If $\tilde{x}_{is} = 0$, then $y_{isi} = 0$. If $\tilde{x}_{is} = 1$, then both x_i and y_{isi} must lie between 2^s and λ^1_{is}. Either way, the inequalities (25) and (26) are satisfied. As for the inequality (27), its right-hand side is equivalent to x_i when $\tilde{x}_{is} = 1$, and to $x_i - \lambda^0_{is}$ when $\tilde{x}_{is} = 0$. In either case, the inequality is satisfied. □

Proposition 8 *When $i = j$, the inequalities (24) can be strengthened to:*

$$\tilde{\lambda}_{is} \sum_{t \in \{s\} \cup C(s)} \tilde{x}_{it} - \sum_{t \in \{s\} \cup C(s)} y_{iti} \leq |C(s)| \left(\tilde{\lambda}_{is} - \sum_{t=0}^{r_j} 2^t \tilde{x}_{it} \right), \tag{28}$$

where $\tilde{\lambda}_{is}$ is the largest value that x_i can take when at least two of the bits in $\{s\} \cup C(s)$ must take the value zero.

Proof. Observe that x cannot take a value larger than $\tilde{\lambda}_{is}$ if the inequality (13) has a positive slack. This implies the following quadratic inequality:

$$\left(|C(s)| - \sum_{t \in C(s) \cup \{s\}} \tilde{x}_{it}\right) \left(\tilde{\lambda}_{is} - x\right) \geq 0.$$

(To see this, note that the first quantity is always non-negative, and the second term can only be negative if the first is zero.) Expanding the quadratic inequality and linearising yields (28). □

We close this subsection with a remark. Using the software package `PORTA` [12], we have found that, even when $n = |I| = 1$, the polytope $C_H(P, I)$ is remarkably complicated. For example, when $u = 31$, we obtained over 15,000 facets.

4.3 Hulls from the Glover–Woolsey approach

To complete the picture, we now extend the approach of Glover & Woolsey [16] to bounded MIQPs. The procedure is as follows:

1. Perform steps 1, 3, 7 and 8 of the procedure mentioned in the previous subsection.

2. For all $i \in I$ such that $u_i + 1$ is not a power of two, add the constraint (1).

3. For $i, j \in I$ with $i < j$, $s = 0, \ldots, r_i$ and $t = 0, \ldots, r_j$, introduce a binary variable $\tilde{X}_{isjt}$, representing $\tilde{x}_{is}\, \tilde{x}_{jt}$.

4. For $i \in I$ and $0 \leq s < t \leq r_i$, introduce a binary variable $\tilde{X}_{isit}$, representing $\tilde{x}_{is}\, \tilde{x}_{it}$.

5. For $i \in I$, $s = 0, \ldots, r_i$ and $j \in N \setminus I$, introduce a continuous variable y_{isj}, representing $\tilde{x}_{is}\, x_j$.

6. Use the $\tilde{X}$, y and X variables to linearise the objective function.

7. For $i, j \in I$ with $i < j$, $s = 0, \ldots, r_i$ and $t = 0, \ldots, r_j$, add the Glover–Woolsey constraints:

$$\tilde{X}_{isjt} \leq \tilde{x}_{is},\ \tilde{X}_{ijst} \leq \tilde{x}_{jt},\ \tilde{x}_{is} + \tilde{x}_{jt} - \tilde{X}_{isjt} \leq 1. \tag{29}$$

8. For $i \in I$ and $0 \leq s < t \leq r_i$, add the Glover–Woolsey constraints

$$\tilde{X}_{isit} \leq \tilde{x}_{is},\ \tilde{X}_{isit} \leq \tilde{x}_{it},\ \tilde{x}_{is} + \tilde{x}_{it} - \tilde{X}_{isit} \leq 1. \tag{30}$$

9. For $i \in I$, $j \in N \backslash I$ and $s = 0, \ldots, r_i$, add the Harjunkowski constraints (11).

The resulting mixed 0-1 LP has $\mathcal{O}\big((L+\bar{n})^2\big)$ variables and $\mathcal{O}\big((L+\bar{n})^2+m\big)$ constraints. As before, if $Q_{ij}=0$ for all $i,j \in N \setminus I$, the mixed 0-1 LP is a reformulation of the MIQP; otherwise, it is a relaxation.

Given any $x \in S(P,I)$, we let $G(x)$ denote the corresponding feasible solution to the mixed 0-1 LP. (The "G" stands for "Glover".) Note that $G(x)$ is actually a quintuple $(\tilde{x}, \bar{x}, y, \tilde{X}, \bar{X})$, where $\tilde{x} \in \{0,1\}^L$, $\bar{x} \in \mathbb{R}^{\bar{n}}_+$, $y \in \mathbb{R}^{\bar{n}L}_+$, $\tilde{X} \in \{0,1\}^{\binom{L}{2}}$ and $\bar{X} \in \mathbb{R}_+^{\binom{\bar{n}+1}{2}}$. We then define

$$C_G(P,I) = \text{conv}\Big\{G(x): \; x \in S(P,I)\Big\}.$$

As usual, $C_G(P,I)$ is polyhedral if and only if $|S(P,I)|$ is finite.

Now, observe that any valid linear equation or inequality for $C_H(P,I)$ can be converted into a valid equation or inequality for $C_G(P,I)$ using the following linear identities:

$$y_{isj} = \sum_{t=0}^{r_j} 2^t \tilde{X}_{isjt} \qquad (i,j \in I;\; s=0,\ldots,r_i).$$

In other words, $C_H(P,I)$ is a projection of $C_G(P,I)$, just as $C(P,I)$ is a projection of $C_H(P,I)$. This implies that, given any convex relaxation that uses "Harjunkowski" variables, we can find a convex relaxation that uses "Glover–Woolsey" variables, that is at least as strong.

The following proposition presents some implicit equations for $C_G(P,I)$ that are not implied by those for $C_H(P,I)$.

Proposition 9 *Let (21) be an implicit equation for $\tilde{P}$ that is not implied by those for P_I. Then, for $j \in I$ and $t = 0,\ldots,r_j$, the equation*

$$\sum_{s \neq t} \alpha_{js} \tilde{X}_{jsjt} + \sum_{i \in I \setminus \{j\}} \sum_{s=0}^{r_i} \alpha_{is} \tilde{X}_{isjt} + \sum_{i \in N \setminus I} \beta_i y_{jti} = (\gamma - \alpha_{jt}) \tilde{x}_{jt}$$

is an implicit equation for $C_G(P,I)$ that is not implied by those for $C_H(P,I)$.

Proof. To show that the equation is valid, multiply (21) by $\tilde{x}_{jt}$ and linearise. The fact that it is not implied by implicit equations for $C_H(P,I)$ comes from the fact that the left-hand side coefficient for $\tilde{X}_{isjt}$ is not always twice that of $\tilde{X}_{i,s-1,j,t}$. □

In fact, $C_G(P,I)$ can have even more implicit equations. For example, if $i \in I$, $u_i > 1$ and $u_i < 2^{r_i} + 2^k$ for some $k \in \{0,\ldots,r_i-1\}$, then it is impossible for $\tilde{x}_{ir}$ and $\tilde{x}_{ik}$ to take the value 1 simultaneously. As a result, we have the implicit equation $\tilde{X}_{ikir} = 0$.

One can also derive valid inequalities for $C_G(P,I)$ that are not implied by valid inequalities for $C_H(P,I)$. As before, one way to do this is to use the RLT. Let $\alpha^k\tilde{x} + \beta^k\bar{x} \leq \gamma_k$, for $k = 1, 2$, be valid inequalities for $\tilde{P}$ that are not implied by valid inequalities for P_I. Form the quadratic inequality $\left(\gamma_1 - \alpha^1\tilde{x} - \beta^1\bar{x}\right)\left(\gamma_2 - \alpha^2\tilde{x} - \beta^2\bar{x}\right) \geq 0$, and linearise it to obtain a valid inequality for $C_G(P, I)$. Note that this procedure yields the constraints (29) and (30) as special cases. Other interesting special cases are obtained by multiplying a cover inequality (13) by an $\tilde{x}$ variable or its complement, or multiplying two cover inequalities together.

One can also adapt Ramana's psd inequalities (9) to the case of $C_G(P, I)$. Specifically, take vectors $v \in \mathbb{R}^L$ and $v' \in \mathbb{R}^{\bar{n}}$, and a scalar $s \in \mathbb{R}$, and form the quadratic inequality $(v \cdot \tilde{x} + v' \cdot \bar{x} - s)^2 \geq 0$. Expanding and linearising yields the psd inequality. Provided that v_{is} is not equal to $2v_{i,s-1}$ for $i \in I$ and $s = 1, \ldots, r_i$, the psd inequality is not implied by the psd inequalities for $C(P, I)$.

In a similar way, one can adapt Burer and Letchford's inequalities (16) to the case of $C_G(P, I)$. Specifically, take a vector $v \in \mathbb{Z}^L$ and a scalar $s \in \mathbb{Z}$, and form the quadratic inequality $(v \cdot \tilde{x} - s)(v \cdot \tilde{x} - s - 1) \geq 0$. Then expand and linearise. As before, provided that v_{is} is not equal to $2v_{i,s-1}$ for all i and s, the resulting inequality is not equivalent to an inequality for $C(P, I)$.

We close this section with a remark. When $n = |I| = 1$ and $u_1 = 2^k - 1$ for some positive integer k, $C_G(P, I)$ is affinely congruent to the Boolean quadric polytope of order k. This implies that optimising an arbitrary linear function over $C_G(P, I)$ is strongly $\mathcal{NP}$-hard even when $n = |I| = 1$. This makes it very unlikely that one could find a complete linear description of $C_G(P, I)$ even in this very special case.

5 Computational Experiments

In this section, we present the results of some computational experiments. The experiments were conducted on a 2.299 GHz AMD Opteron 6376 with 16Gb RAM, under a 64 bit Linux operating system (Ubuntu 18.04). We used `CPLEX` (v. 12.9) to solve all LPs and mixed 0-1 LPs.

5.1 Test instances

For simplicity and brevity, we consider "box-constrained" IQPs of the form:

$$\min\left\{x^T Q x + c \cdot x :\ x \in \{0, \ldots, u\}^n\right\},$$

where n is set to 25 and u ranges from 1 to 7.

For the objective function, we consider three different cases: convex, concave and indefinite. To generate the convex instances, we pick a random point $x^* \in (0, u)^n$ and a random matrix $M \in \mathbb{Z}^{n \times n}$ with entries in $\{-10, \ldots, 10\}$. We then minimise $||M(x - x^*)||_2^2$. This corresponds to setting Q to $M^T M$ and c to $-2Qx^*$. The concave instances are similar, except that we replace Q with $-Q$ and c with $-c$. The indefinite instances are obtained by picking two random points x^1, x^2 and two random matrices M^1, M^2, and minimising $||M^1(x - x^1)||_2^2 - ||M^2(x - x^2)||_2^2$.

For each of the three cases, and for $u \in \{1, \ldots, 7\}$, we created five random instances. This makes $3 \times 7 \times 5 = 105$ instances in total.

5.2 LP relaxations

For each of the 105 instances, we solved five LP relaxations:

1. The McCormick relaxation, containing only constraints (3) and (4).
2. The Harjunkowski-type relaxation, containing only constraints (20).
3. The Glover–Woolsey-type relaxation, containing constraints (29) and (30) and, if $u + 1$ is not a power of 2, the constraint (1) for all $i \in I$.
4. An "enhanced" Harjunkowski-type relaxation, in which (a) Propositions 7 and 8 are applied, (b) the RLT constraints (23), (24) are added (when $u + 1$ is not a power of two), and (c) the equations mentioned at the end of Subsection 4.2.1 are added (when u is a power of two).
5. An "enhanced" Glover–Woolsey-type relaxation, in which we add (a) the RLT constraints obtained by multiplying cover inequalities by either bounds on $\tilde{x}$ variables or other cover inequalities (when $u + 1$ is not a power of two), and (b) some $\tilde{X}$ variables are fixed to zero (for certain values of u; see Subsection 4.3).

We remark that the McCormick, Harjunkowski and Glover–Woolsey relaxations have $O(n^2)$, $O(nL)$ and $O(L^2)$ variables and constraints, respectively. The same is true of the enhanced versions.

Table 1 shows the gap between the lower bound and the optimum, expressed as a percentage of the optimum, for each relaxation and combination of parameters. Here, "conv", "conc" and "indef" stand for convex, concave and indefinite, respectively. Each figure is the average over five random instances. All results are to 2 d.p. (We do not report computing times, because they were negligible, always much less than one second.)

The results are rather surprising. In the first place, most of the integrality gaps are huge. Second, the concave instances have the smallest integrality gaps, and the convex ones have the largest. This is the opposite of what we expected. (A possible explanation is that the optimal solution of a concave instance will always be an extreme point of the box.) Third, the McCormick bound is better than the Harjunkowski bound, despite having fewer variables and constraints. On the other hand, the "enhanced" Harjunkowski bound dominates the McCormick bound. Finally, we see no difference between the bounds from the two "enhanced" relaxations.

Case	u	Original			Enhanced	
		Mc	Ha	GW	Ha	GW
conv	1	580.16	580.16	324.95	324.95	324.95
	2	464.09	576.91	804.10	255.09	255.09
	3	453.71	453.71	339.78	339.78	339.78
	4	449.90	599.88	1139.28	247.96	247.96
	5	447.10	538.28	726.75	309.39	309.39
	6	446.23	488.24	499.38	333.91	333.91
	7	445.96	445.96	359.36	359.36	359.36
conc	1	20.25	20.25	20.25	20.25	20.25
	2	20.24	45.52	207.18	20.24	20.24
	3	20.20	20.20	20.20	20.20	20.20
	4	20.20	53.87	332.17	20.20	20.20
	5	20.23	41.60	162.99	20.23	20.23
	6	20.22	30.51	73.13	20.22	20.22
	7	20.23	20.23	20.23	20.23	20.23
indef	1	48.30	48.30	45.70	45.70	45.70
	2	48.01	76.31	208.64	45.43	45.43
	3	48.07	48.07	46.63	46.63	46.63
	4	48.07	86.53	315.45	45.48	45.48
	5	48.07	71.64	171.96	46.31	46.31
	6	48.13	58.98	94.11	46.69	46.69
	7	48.12	48.12	47.01	47.01	47.01

Table 1: Average percentage integrality gap.

		Original		Enhanced	
Case	u	Ha	GW	Ha	GW
conc	1	0.50	0.20	0.23	0.20
	2	8.25	14.29	0.41	0.66
	3	8.80	4.77	1.05	4.70
	4	45.73	149.06	4.02	7.03
	5	26.21	47.46	1.78	7.68
	6	23.29	76.25	3.94	11.24
	7	17.19	22.73	2.23	22.54
indef	1	1.14	0.94	0.58	0.96
	2	49.06	35.32	3.19	4.64
	3	76.13	13.63	7.26	14.06
	4	3094.44	3397.36	119.20	125.05
	5	1171.46	205.49	79.35	102.17
	6	1284.09	2157.26	142.68	170.61
	7	294.79	66.83	126.53	67.35

Table 2: Average branch-and-bound time.

Observe that, for the Harjunkowski and Glover–Woolsey relaxations, the instances with $u \in \{1, 3, 7\}$ tend to have smaller gaps than the others. This is presumably related to the fact that the inequalities $\sum_{s=0}^{r_i} \tilde{x}_{is} \leq u$ are unnecessary when $u + 1$ is a power of 2. This effect disappears for the enhanced versions of those relaxations.

On balance, the enhanced Harjunkowski relaxation looks most promising here (since it yields good bounds, but has fewer variables and constraints than the enhanced Glover–Woolsey relaxation.)

5.3 Branch-and-bound

We then took four of the above relaxations, declared the $\tilde{x}$ variables to be binary, and fed the resulting mixed 0-1 LPs into a branch-and-bound solver. Surprisingly, the convex instances turned out to be the hardest by far, with many instances taking over an hour to solve. Table 2 shows the average branch-and-bound time, in seconds, for the concave and indefinite instances. As before, each figure is the average over five random instances.

It is clear that the "enhanced" formulations take significantly less time to solve, on average, than the original formulations. This shows that the effort spent in tightening the relaxations pays off. It is also clear that the indefinite instances are harder to solve than the concave instances. This is probably due to the greater integrality gaps mentioned above. Finally, as in the previous subsection, we see that instances with $u \in \{1, 3, 7\}$ tend to be easier than the others, but only when the formulations are not enhanced.

Overall, the enhanced Harjunkowski formulation is a clear winner. This is probably due to the fact that it has fewer variables and constraints than the enhanced Glover–Woolsey formulation, but yields the same LP bound.

6 Conclusion

Although there are many papers on valid inequalities for specific mixed-integer quadratic problems, this is the first time that results have been given for MIQPs in general. Moreover, we have derived results for several different formulation approaches, including ones that use bit representation. The computational results suggest that, among the different possible approaches, the most promising is the one based on adding cuts to the formulation of Harjunkowski *et al.* [18]. However, further experimentation would be needed to confirm this.

There are several other interesting possible topics for future research. One is to solve the enhanced Harjunkowski *et al.* formulation with a *spatial branch-and-bound* algorithm (see [20, 23, 26]). The idea here is that, after branching on the domain of a variable, one can tighten some of the constraints (11), (14), (15) and (20) accordingly. Potentially, this could reduce the number of branch-and-bound nodes. Some of our cutting planes can be tightened in a similar way.

A second topic worth a look is that of exploiting *sparsity* in the quadratic cost matrix Q. Indeed, if $Q_{ij} = 0$, then we can in principle omit the corresponding variables (i.e., X_{ij}, y_{isj} or $\tilde{X}_{isjt}$, depending on the formulation). Provided that one can figure out how to modify the constraints accordingly, this may lead to a much smaller formulation, which may enable instances to be solved more quickly. However, omitting variables corresponds to projecting the convex hulls into a subspace, and this may not be easy.

Acknowledgement: The authors gratefully acknowledge the partial financial support from the Italian Ministry of Education, University and Research (MIUR), under the project "Nonlinear and combinatorial aspects of complex networks" (grant `PRIN 2015B5F27W`).

References

[1] W.P. Adams & H.D. Sherali (1986) A tight linearization and an algorithm for zero-one quadratic programming problems. *Mgmt. Sci.*, 32, 1274–1290.

[2] K.M. Anstreicher & S. Burer (2010) Computable representations for convex hulls of low-dimensional quadratic forms. *Math. Program.*, 124, 33–43.

[3] A. Billionnet, S. Elloumi & A. Lambert (2008) Linear reformulations of integer quadratic programs. In L.T. Hoai An, P. Bouvry & P.D. Tao (eds.) *Modelling, Computation and Optimization in Information Systems and Management Sciences*, pp. 43–51. Heidelberg: Springer.

[4] A. Billionnet, S. Elloumi & A. Lambert (2012) Extending the QCR method to general mixed-integer programs. *Math. Program.*, 131, 381–401.

[5] P. Bonami, M. Kilinç & J. Linderoth (2012) Algorithms and software for convex mixed integer nonlinear programs. In: J. Lee & S. Leyffer (eds.), *Mixed Integer Nonlinear Programming*, pp. 1–40. New York: Springer.

[6] E. Boros & P.L. Hammer (1993) Cut-polytopes, Boolean quadric polytopes and nonnegative quadratic pseudo-Boolean functions. *Math. Oper. Res.*, 18, 245–253.

[7] C. Buchheim & E. Traversi (2015) On the separation of split inequalities for non-convex quadratic integer programming. *Discr. Optim.*, 15, 1–14.

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

[9] S. Burer & A.N. Letchford (2012) Non-convex mixed-integer nonlinear programming: a survey. *Surveys in Oper. Res. & Mgmt. Sci.*, 17, 97–106.

[10] S. Burer & A.N. Letchford (2014) Unbounded convex sets for non-convex mixed-integer quadratic programming. *Math. Program.*, 143, 231–256.

[11] S. Burer & A. Saxena (2012) The MILP road to MIQCP. In J. Lee & S. Leyffer (eds.), *Mixed Integer Nonlinear Programming*, pp. 373–405. New York: Springer.

[12] T. Christof & A. Loebl: `PORTA` (polyhedron representation transformation algorithm). Software package, available for download at `http://www.iwr.uni-heidelberg.de/groups/comopt/software`

[13] C. D'Ambrosio & A. Lodi (2013) Mixed integer nonlinear programming tools: an updated practical overview. *Ann. Oper. Res.*, 204, 301–320.

[14] L. Galli, K. Kaparis & A.N. Letchford (2011) Gap inequalities for non-convex mixed-integer quadratic programs. *Oper. Res. Lett.*, 39, 297–300.

[15] F. Glover (1975) Improved linear integer programming formulations of nonlinear integer problems. *Mngt. Sci.*, 22, 455–460.

[16] F. Glover & E. Woolsey (1974) Converting the 0-1 polynomial program to a 0-1 linear program. *Oper. Res.*, 22, 180–182.

[17] A. Gupte, S. Ahmed, M.S. Cheon & S. Dey (2013) Solving mixed integer bilinear problems using MILP formulations. *SIAM J. Optim.*, 23, 721–744.

[18] I. Harjunkoski, R. Pörn, T. Westerlund & H. Skrivars (1997) Different strategies for solving bilinear integer non-linear programming problems with convex transformations. *Comput. Chem. Eng.*, 21, 5487–5492.

[19] J. Kronqvist, D.E. Bernal, A. Lundell & I.E. Grossmann (2019) A review and comparison of solvers for convex MINLP. *Optim. & Engin.*, 20, 397–455.

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

[21] M.W. Padberg (1989) The boolean quadric polytope: some characteristics, facets and relatives. *Math. Program.*, 45, 139–172.

[22] M. Ramana (1993) *An Algorithmic Analysis of Multiquadratic and Semidefinite Programming Problems.* Ph.D. thesis, Johns Hopkins University, Baltimore, MD.

[23] H.S. Ryoo & N.V. Sahinidis (1996) A branch-and-reduce approach to global optimization. *J. Glob. Optim.*, 8, 107–138.

[24] A. Saxena, P. Bonami & J. Lee (2010) Convex relaxations of non-convex mixed integer quadratically constrained programs: extended formulations. *Math. Program.*, 124, 383–411.

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

[26] M. Tawarmalani & N.V. Sahinidis (2003) *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming.* Dordrecht: Kluwer.

[27] L.J. Watters (1967) Reduction of integer polynomial programming problems to zero-one linear programming problems. *Oper. Res.*, 15, 1171–1174.

[28] Y. Yajima & T. Fujie (1998) A polyhedral approach for nonconvex quadratic programming problems with box constraints. *J. Glob. Optim.*, 13, 151–170.