J Glob Optim (2012) 53:29–51
DOI 10.1007/s10898-010-9644-3

On linear programs with linear complementarity constraints

Jing Hu · John E. Mitchell · Jong-Shi Pang · Bin Yu

Received: 5 December 2010 / Accepted: 23 December 2010 / Published online: 8 January 2011

Abstract The paper is a manifestation of the fundamental importance of the linear program with linear complementarity constraints (LPCC) in disjunctive and hierarchical programming as well as in some novel paradigms of mathematical programming. In addition to providing a unified framework for bilevel and inverse linear optimization, nonconvex piecewise linear programming, indefinite quadratic programs, quantile minimization, and ℓ_0 minimization, the LPCC provides a gateway to a mathematical program with equilibrium constraints, which itself is an important class of constrained optimization problems that has broad applications.

It is our great pleasure to dedicate this work to Professor Richard W. Cottle on the occasion of his 75th birthday in 2009. Professor Cottle is the father of the linear complementarity problem (LCP) [16]. The linear program with linear complementarity constraints (LPCC) treated in this paper is a natural extension of the LCP; our hope is that the LPCC will one day become as fundamental as the LCP, thereby continuing Professor Cottle's legacy, bringing it to new heights, and extending its breadth.
The work of the author John E. Mitchell was supported by the National Science Foundation under grant DMS-0715446 and by the Air Force Office of Sponsored Research under grant FA9550-08-1-0081.
The work of the author Jong-Shi Pang was supported by the Office of Naval Research under grant no. N00014-06-1-0014, by the Air Force Office of Sponsored Research under grant FA9550-08-1-0061, and by the National Science Foundation under grant CMMI-0969600.

J. Hu
Market Analytics, Inc., 500 Davis Street, Suite 1001, Evanston, IL 60201, USA

J. E. Mitchell
Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180-1590, USA
e-mail: mitchj@rpi.edu

J.-S. Pang (✉)
Department of Industrial and Enterprise Systems Engineering, University of Illinois,
Urbana, IL 61801, USA
e-mail: jspang@illinois.edu

B. Yu
Department of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute,
Troy, NY 12180-1590, USA
e-mail: yub@rpi.edu

We describe several approaches for the global resolution of the LPCC, including a logical Benders approach that can be applied to problems that may be infeasible or unbounded.

Keywords Linear programs with linear complementarity constraints · Inverse programming · Hierarchical programming · Piecewise linear programming · Quantile minimization · Cross-validated support vector regression

1 Introduction

A mathematical program with complementarity constraints (MPCC) is a constrained optimization problem subject to certain complementarity conditions on pairs of variables. The latter conditions classify the MPCC as a nonconvex, disjunctive program. A linear program with complementarity constraints (LPCC) is a special case of the MPCC in which the objective function and all constraints are linear, except for the complementarity conditions. Complementarity constraints are very natural in describing certain logical relations. An early occurrence of these constraints is in piecewise linear optimization, wherein the complementarity condition expresses the simple fact that a linear segment of the function should not be invoked until its immediate predecessor is fully utilized. This condition is not needed in the minimization of a convex piecewise linear function, but cannot be removed in a nonconvex minimization problem.

With the goal of establishing the MPCC as a fundamental class of disjunctive programs of practical significance, the present paper documents a number of novel optimization models in which complementarity occurs naturally in the algebraic and/or logical description of the model objectives and/or constraints in Sects. 3, 4, and 5. Such models include hierarchical, inverse, quantile, and ℓ_0 optimization, as well as optimization problems with equilibrium constraints. In turn, with its linear structures, the LPCC occupies a central niche in these nonconvex problems, playing the same role as a linear program does in the domain of convex programming. Thus, the LPCC provides an important gateway to a large class of nonlinear disjunctive programs; as such, it is imperative that efficient algorithms be developed to facilitate the global resolution of the LPCC. An effort along this line is described in [28] and two algorithms are discussed in Sect. 6. Methods for improving relaxations of LPCCs are given in Sect. 7 and computational results are described in Sect. 8.

2 Problem formulation

Since this paper focuses on the LPCC and the emphasis is on the complementarity constraints, we restrict the presentation of the models to linear ones. We begin by giving a general formulation of the LPCC in the form suggested by Scheel and Scholtes [48]. Given vectors and matrices: $c \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, $e \in \Re^m$, $b \in \mathbb{R}^k$, $A \in \mathbb{R}^{k\times n}$, $B \in \mathbb{R}^{k\times m}$, and $C \in \mathbb{R}^{k\times m}$, the LPCC is to find a triple $(x, y, w) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ in order to globally

$$\begin{array}{cl} \underset{(x,y,w)}{\text{minimize}} & c^T x + d^T y + e^T w \\ \text{subject to} & Ax + By + Cw \ \geq \ b \\ \text{and} & 0 \leq y \perp w \geq 0, \end{array} \tag{1}$$

where the $\perp$ notation denotes the perpendicularity between two vectors. Thus, without the orthogonality condition: $y \perp w$, the LPCC is a linear program (LP). With this condition, the

LPCC is equivalent to 2^m LPs, each called a *piece* of the problem and defined by a subset $\mathcal{I}$ of $\{1, \cdots, m\}$:

$$\begin{array}{ll} \underset{(x,y,w)}{\text{minimize}} & c^T x + d^T y + e^T w \\ \text{subject to} & Ax + By + Cw \geq b \\ & y_i = 0 \leq w_i, \quad i \in \mathcal{I} \\ \text{and} & y_i \geq 0 = w_i, \quad i \notin \mathcal{I}. \end{array}$$

The global resolution of the LPCC means the generation of a certificate showing that the problem is in one of its 3 possible states: (a) it is infeasible, (b) it is feasible but unbounded below, or (c) it attains a finite optimal solution. Needless to say, linear equations (in addition to linear inequalities as stated above) connecting the variables (x, y, w) are allowed in the constraints of the LPCC; for convenience of presentation, such equality constraints are omitted.

A frequently occurred special case of (1) is the following:

$$\begin{array}{ll} \underset{(x,y)}{\text{minimize}} & c^T x + d^T y \\ \text{subject to} & Ax + By \geq b \\ \text{and} & 0 \leq y \perp q + Nx + My \geq 0, \end{array} \tag{2}$$

in which $N \in \mathbb{R}^{m \times n}$ and $M \in \mathbb{R}^{m \times m}$. Generalizing the standard linear complementarity problem (LCP): $0 \leq y \perp q + Nx + My \geq 0$, affine variational constraints also lead to the problem (1). In particular, consider the problem:

$$\begin{array}{ll} \underset{(x,y)}{\text{minimize}} & c^T x + d^T y \\ \text{subject to} & Ax + By \geq b \\ & y \in K \quad \text{and} \quad (y' - y)^T (q + Nx + My) \geq 0, \ \forall\, y' \in K, \end{array} \tag{3}$$

where $K \triangleq \{y : Ey \leq h\}$ is a given polyhedron, with $E \in \mathbb{R}^{\ell \times m}$ and $h \in \mathbb{R}^{\ell}$. By letting $\lambda \in \mathbb{R}^{\ell}$ be the multipliers of the inequalities defining K, the problem (3) has the equivalent formulation:

$$\begin{array}{ll} \underset{(x,y,\lambda)}{\text{minimize}} & c^T x + d^T y \\ \text{subject to} & Ax + By \geq b \\ & 0 = q + Nx + My + E^T \lambda \\ & 0 \leq \lambda \perp h - Ey \geq 0. \end{array} \tag{4}$$

In turn, the affine variational inequality of finding $y \in K$ such that $(y' - y)^T (q + Nx + My) \geq 0$ for all $y' \in K$ provides a unified formulation for convex quadratic programs and a host of equilibrium problems with affine structures [19].

Since the LPCC is a generalization of the linear complementarity problem, it is NP-hard. We also show explicitly in Sect. 5.1 how an integer program can be reduced to an LPCC. Analogously to integer programming, there does not appear to be a simple way to characterize *a priori* instances of LPCC that are hard computationally. Of course, the number of complementarities is important, as to a lesser degree are the dimensions of the other variables. Looking at formulation (2), the structure of B and of M may play a role. For example, if $B = 0$ and M is copositive plus then it follows from results for LCPs that for any given x a pivoting algorithm can be used to determine whether a feasible y exists.

Global optimization methods for solving linear complementarity problems can be found in [25,41]. Collections of global optimization test problems include [20,21].

3 Complementarity constraints enforcing KKT conditions

Beginning in this section, we present various applications of the LPCC (1) and its special cases. These applications show that the complementarity constraints often arise in practical modeling. In this section, we consider applications where the complementarity conditions are used to model KKT optimality conditions that must be satisfied by some of the variables. The KKT conditions can either be those of a subproblem or of the problem itself. The applications in this section demonstrate multiple modeling paradigms, from heirarchical optimization to inverse optimization to data fitting and even to quadratic programming.

In subsequent sections, we look at applications where the complementarity conditions are used to model nonconvex piecewise linear functions in Sect. 4, and other applications are discussed in Sect. 5.

3.1 Hierarchical optimization

In a bilevel optimization problem, feasible solutions are constrained to correspond to optimal solutions to a lower level problem. If the lower level problem is convex and satisfies a constraint qualification then it can be replaced by its KKT optimality conditions [18]. Hence such problems naturally lead to MPCCs. If the upper level problem is linear and if lower level problem is a linear program or a convex quadratic program then the problem can be reformulated as an LPCC.

A hierarchical optimization problem may have more than one lower level problem, and these lower level problems may have subproblems of their own. In order to be able to formulate the problem as an LPCC, we restrict attention to hierarchical problems with a single layer of subproblems. In particular, we consider the following problem:

$$\begin{array}{lll} \underset{(x,y)}{\text{minimize}} & c^T x + \sum_{i=1}^{r} {h^i}^T y^i & \\ \text{subject to} & Ax + \sum_{i=1}^{r} B^i y^i \ \geq \ b & \\ \text{and} & y^i \ \in \ \underset{v^i}{\text{argmin}} & {d^i}^T v^i + \frac{1}{2} (v^i)^T Q^i v^i \\ & \text{subject to} & C^i v^i \ \geq \ g^i - Fx - \sum_{j \neq i, j=1}^{r} G^j y^j, \end{array} \tag{5}$$

where $x \in \Re^n$, $y^i, v^i \in \Re^{p_i}$, $b \in \Re^m$, $g^i \in \Re^{q_i}$, each Q^i is symmetric and positive semidefinite, and c, d^i, h^i, A, B^i, C^i, F, G^i and Q^i are all dimensioned appropriately. This problem arises in Stackelberg games, where there is a single leader with decision variables x and there are r followers with decision variables y^i, and each follower is optimizing its own subproblem.

Since each subproblem is convex with linear constraints, an optimal solution must satisfy the KKT conditions. If a subproblem is infeasible or unbounded then formally the argmin of the subproblem is empty. Hence, (5) can be reformulated as the following equivalent LPCC:

$$
\begin{array}{lll}
\underset{(x,y)}{\text{minimize}} & c^T x + \sum_{i=1}^{r} h^{i\,T} y^i & \\
\text{subject to} & Ax + \sum_{i=1}^{r} B^i y^i \ \geq \ b & \\
& Fx + C^i y^i + \sum_{j \neq i, j=1}^{r} G^j y^j \ - \ w^i \ = \ g^i \ \text{ for } i = 1, \ldots, r & \\
& d^i + Q^i y^i - (C^i)^T \lambda^i \ = \ 0 \ \text{ for } i = 1, \ldots, r & \\
\text{and} & 0 \ \leq \ w^i \ \perp \ \lambda^i \ \geq \ 0 \ \text{ for } i = 1, \ldots, r &
\end{array}
\tag{6}
$$

where $w^i, \lambda^i \in \Re^{q_i}$. If we assume that the dual feasible region of the ith subproblem given by

$$\{(v^i, \pi^i) \in \Re^{p_i+q_i} \ : \ d^i + Qv^i - (C^i)^T\pi^i = 0, \ \pi^i \geq 0\}$$

is nonempty for each i, then the subproblem is either infeasible or achieves its minimum at a KKT point.

Surveys of bilevel optimization problems include [15,17], and hierarchical optimization problems are surveyed in [3]. The next two sections give examples of the LPCC formulation of bilevel optimization problems.

3.2 Inverse convex quadratic programming

Inverse convex quadratic programming pertains to the inversion of the inputs to a convex quadratic program (QP) so that a secondary objective function is optimized; when the latter is linear, then we obtain an LPCC. Inverse optimization problems are surveyed in [2]. Inverse quadratic programs are investigated in [57,58]. Inverse conic programs are considered in [29] and inverse linear complementarity problems in [49].

To illustrate, consider a standard convex quadratic program:

$$
\begin{array}{ll}
\underset{x \in \mathbb{R}^n}{\text{minimize}} & c^T x + \tfrac{1}{2} x^T Q x \\
\text{subject to} & Ax \ \leq \ b,
\end{array}
\tag{7}
$$

where Q is a symmetric positive semidefinite matrix. Solving this program for a given tuple (Q, A, b, c) is the forward problem. An inverse problem is as follows. Given a target triple $(\bar{b}, \bar{c}, \bar{x})$, which could represent historical data or empirical observations, and a given pair of matrices (Q, A) that identifies the forward optimization model, we want to construct a pair (b, c) and an optimal solution x of the forward QP so that (b, c, x) is least deviated from $(\bar{b}, \bar{c}, \bar{x})$. Using a polyhedral (say, the ℓ_1 or ℓ_∞) norm $\| \bullet \|$ to measure the deviation, we obtain the bilevel optimization formulation for this inverse program:

$$
\begin{array}{llll}
\underset{(x,b,c)}{\text{minimize}} & \| (x, b, c) - (\bar{x}, \bar{b}, \bar{c}) \| & & \\
\text{subject to} & (b, c) \ \in \ \mathcal{F} & \text{(a polyhedron)} & \\
\text{and} & x \ \in \ \underset{x'}{\text{argmin}} & c^T x' + \tfrac{1}{2} (x')^T Q x' & \\
& \qquad \text{subject to} & Ax' \ \leq \ b. &
\end{array}
\tag{8}
$$

A variant of this inverse problem is the following: given the constraint matrix A, the positive semidefinite matrix Q, a positive scalar $\varepsilon > 0$, and the triple $(\bar{x}, \bar{b}, \bar{c})$, we want to find a triple (x, b, c) so that the pair (b, c) is least deviated from $(\bar{b}, \bar{c})$ and that $\bar{x}$ is at a distance of at most ε from an optimal solution of the convex QP (Q, A, b, c).

Writing out the Karush–Kuhn–Tucker (KKT) conditions of the inner-level QP in (8), we arrive at the following LPEC formulation of the above inverse quadratic program:

$$\begin{array}{ll} \underset{(x,b,c)}{\text{minimize}} & \| (x,b,c) - (\bar{x}, \bar{b}, \bar{c}) \| \\ \text{subject to} & (b,c) \in \mathcal{F} \\ \text{and} & c + Qx + A^T\lambda = 0 \\ & 0 \le b - Ax \perp \lambda \ge 0. \end{array} \tag{9}$$

A noteworthy point about the above inverse problem is that the pair of matrices (Q, A) is fixed. If they are part of the inversion process, then we obtain a *nonlinear* program with complementarity constraints instead.

The notion of inverting an optimization problem provides an illustration of the process of model selection in the presence of historical data and/or empirical observations. Similar inversions arise in many related contexts pertaining to the parameter identification in a forward optimization or equilibrium problem, for the purpose of optimizing a prescribed performance function. Such an inverse process is very common in the field of partial differential equations wherein the forward process is defined by these equations. When the forward process is a continuous optimization problem with inequality constraints, the inverse optimization problem is an instance of a bilevel program, which leads to an MPCC when the forward (i.e., the inner) optimization problem is formulated in terms of its KKT conditions, and to an LPCC in particular cases.

3.3 Cross-validated support vector regression

The support vector machine (SVM) is a well-known statistical learning method for data mining [54]. Mathematically, the SVM is formulated as a convex quadratic program with 2 hyper-parameters—the regularization constant C and the tube width ε, which are typically selected by cross validation based on the mean square error (MSE) or mean absolute deviation (MAD) measured on certain out-of-sample data. Traditionally, such a selection is done in an ad hoc manner. Several recent papers in machine learning [12–14] have suggested embedding the SVM in a bilevel optimization framework for the choice of (C, ε) via the minimization of an outer-level out-of-sample error. In what follows, we present a bilevel programming formulation for a cross-validated support vector regression problem with (C, ε) as the design variables. Further discussion of this problem can be found in the references [9,28,33,34].

Suppose that the regression data are described by the ℓ points $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_\ell, y_\ell)\}$ in the Euclidean space $\mathbb{R}^{n+1}$ for some positive integers ℓ and n. We partition these points into N mutually disjoint subsets, Ω_t for $t = 1, \ldots, N$, such that $\bigcup_{t=1}^N \Omega_t = \{1, \ldots, \ell\}$. Let $\overline{\Omega}_t \equiv \{1, \ldots, \ell\} \setminus \Omega_t$ be the subset of the data other than those in groups Ω_t. Our goal is to fit a hyperplane $y = x^T w + b$ to the given data points based on their partitioning. This is accomplished by solving a bilevel model selection problem, which is to find the parameters (C, ε) and $(\mathbf{w}^t, b_t)$ for $t = 1, \ldots, N$ in order to

$$\begin{array}{ll} \underset{C,\varepsilon,\mathbf{w}^t,b_t}{\text{minimize}} & \displaystyle\frac{1}{N}\sum_{t=1}^{N}\frac{1}{|\,\Omega_t\,|}\sum_{i\in\Omega_t} |\,\mathbf{x}_i^T\mathbf{w}^t + b_t - y_i\,| \\ \text{subject to} & C, \varepsilon \ge 0 \\ \text{and for} & t = 1, \ldots, N, \end{array} \tag{10}$$

$$(\mathbf{w}^t, b_t) \in \underset{\mathbf{w},b}{\text{argmin}} \left\{ C \sum_{j\in\overline{\Omega}_t} \max\left(|\,\mathbf{x}_j^T\mathbf{w} + b - y_j\,| - \varepsilon, 0 \right) + \frac{1}{2}\|\,\mathbf{w}\,\|_2^2 \right\}, \tag{11}$$

where the argmin in the last constraint denotes the set of optimal solutions to the convex optimization problem (11) in the variable $\mathbf{w}$ for given hyper-parameters (C, ε). Note that the inner problem is strictly convex in w, so each w^t will be the same. Clearly, the inner problem (11) is equivalent to the convex quadratic program:

$$\begin{array}{ll} \underset{\mathbf{w},b,e_j}{\text{minimize}} & C \sum\limits_{j \in \overline{\Omega}_t} e_j + \frac{1}{2} \parallel \mathbf{w} \parallel_2^2 \\ \text{subject to} & \left\{ \begin{array}{lcl} e_j & \geq & \mathbf{x}_j^T \mathbf{w} + b - y_j - \varepsilon \\ e_j & \geq & -\mathbf{x}_j^T \mathbf{w} - b + y_j - \varepsilon \\ e_j & \geq & 0 \end{array} \right\} \quad j \in \overline{\Omega}_t. \end{array}$$

Thus the overall bilevel cross-validated support vector regression is an instance of an LPCC after we write out the KKT conditions of the above QP. The complete LPCC formulation is as follows:

$$\begin{array}{lll} \underset{C,\varepsilon,\mathbf{w}^t,b_t,e^t,\eta^{t\pm}}{\text{minimize}} & \frac{1}{N} \sum\limits_{t=1}^{N} \frac{1}{\mid \Omega_t \mid} \sum\limits_{i \in \Omega_t} \mid \mathbf{x}_i^T \mathbf{w}^t + b_t - y_i \mid & \\ \text{subject to} & C, \varepsilon \geq 0 & \\ & 0 \leq \eta_j^{t+} \perp e_j^t - \mathbf{x}_j^T \mathbf{w}^t - b_t + y_j + \varepsilon \geq 0 & \forall\, j \in \overline{\Omega}_t, \text{ for } t = 1, \ldots, N \\ & 0 \leq \eta_j^{t-} \perp e_j^t + \mathbf{x}_j^T \mathbf{w}^t + b_t - y_j + \varepsilon \geq 0 & \forall\, j \in \overline{\Omega}_t, \text{ for } t = 1, \ldots, N \\ & 0 \leq e_j^t \perp C - \eta_j^{t+} - \eta_j^{t-} \geq 0 & \forall\, j \in \overline{\Omega}_t, \text{ for } t = 1, \ldots, N \\ & \sum\limits_{j \in \overline{\Omega}_t} (\eta_j^{t-} - \eta_j^{t+}) = 0 & \text{for } t = 1, \ldots, N \\ & \mathbf{w}^t = \sum\limits_{j \in \overline{\Omega}_t} (\eta_j^{t-} - \eta_j^{t+}) \mathbf{x}_j & \text{for } t = 1, \ldots, N. \end{array} \tag{12}$$

The last constraint can be used to substitute for $\mathbf{w}^t$ elsewhere in the problem.

The cross-validated support vector regression approach could be embedded in another level of cross-validation. In particular, the cross-validation support vector regression model could be applied to a subset of the data, and the resulting model tested on the remaining data. This might improve the generalizability of the resulting model.

3.4 Indefinite quadratic programs

In this application, the KKT conditions are imposed on the problem itself rather than on a subproblem, and the objective function value of a KKT point is given by a linear function. Consider the QP (7) where the matrix Q is symmetric and indefinite. We assume that the program is feasible but not necessarily bounded. The problem of deciding by a finite algorithm whether (7) has a finite optimal solution or is unbounded was not fully resolved until the recent paper [27] in which an LPCC is introduced whose global resolution provides the answer to this decision problem. Previous approaches were based on the assumption that the QP is known to have a finite optimal solution. Indeed, an early result of Giannessi and Tomasin [22] states under the solvability assumption, the QP (7) is equivalent to the LPCC of minimizing a certain linear objective function (involving the constraint multipliers) over the set of KKT conditions of the QP. This equivalence breaks down for an unbounded QP.

In what follows, we present the equivalent LPCC formulation for the QP (7) assuming only its feasibility. We refer the reader to the cited reference for the derivation details of this LPCC. Since (7) is obviously equivalent to

$$\begin{array}{ll} \underset{x^{\pm} \in \mathbb{R}^{2n}}{\text{minimize}} & \frac{1}{2}\,(\,x^{+} - x^{-}\,)^{T}\, Q(\,x^{+} - x^{-}\,) + c^{T}\,(\,x^{+} - x^{-}\,) \\ \text{subject to} & A(\,x^{+} - x^{-}\,) \,\leq\, b \\ \text{and} & x^{\pm} \,\geq\, 0, \end{array}$$

we may assume, to simplify the notation, that the recession cone $\mathcal{D} \equiv \{d \in \mathbb{R}^n : Ad \leq 0\}$ of the feasible set is contained in the nonnegative orthant $\mathbb{R}^n_+$. It is then shown in [27] that the QP (7) is unbounded below if and only if the LPCC below has a feasible solution with a negative objective value:

$$\begin{array}{ll} \underset{(x,d,\xi,\lambda,\mu,t,s) \in \mathbb{R}^{2n+3m+2}}{\text{minimize}} & -t \\ \text{subject to} & 0 \,=\, c + Qx + A^T \xi + t\,\mathbf{1}_n \\ & 0 \,=\, Qd + A^T \lambda - A^T \mu + s\,\mathbf{1}_n \\ & 0 \,\leq\, \xi \,\perp\, b - Ax \,\geq\, 0 \\ & 0 \,\leq\, \mu \,\perp\, b - Ax \,\geq\, 0 \\ & 0 \,\leq\, \lambda \,\perp\, -Ad \,\geq\, 0 \\ & 0 \,\leq\, \xi \,\perp\, -Ad \,\geq\, 0 \\ & 0 \,\leq\, \mu \,\perp\, -Ad \,\geq\, 0 \\ & 0 \,\leq\, s, \quad \mathbf{1}_n^T d \,\geq\, 1. \end{array} \tag{13}$$

If Q is copositive on $\mathcal{D}$, then the QP (7) is unbounded below if and only if the following somewhat simplified LPCC:

$$\begin{array}{ll} \underset{(x,d,\xi,\lambda,t) \in \mathbb{R}^{2(n+m)+1}}{\text{minimize}} & -t \\ \text{subject to} & 0 \,=\, c + Qx + A^T \xi + t\,\mathbf{1}_n \\ & 0 \,=\, Qd + A^T \lambda \\ & 0 \,\leq\, \xi \,\perp\, b - Ax \,\geq\, 0 \\ & 0 \,\leq\, \lambda \,\perp\, -Ad \,\geq\, 0 \\ & 0 \,\leq\, \xi \,\perp\, -Ad \,\geq\, 0 \\ & 1 \,\leq\, \mathbf{1}_n^T d \end{array} \tag{14}$$

has a feasible solution with a negative objective value. Detailed investigation of how to solve (13) or (14) has yet to be undertaken. Some preliminary computational results with unbounded problems can be found in [27]. These results exploit constraints that require that the second order optimality conditions be satisfied at a solution. Theoretically, this class of constraints enables the solution of certain classes of nonconvex quadratic constraints to be solved in polynomial time—see the cited reference for details.

4 Complementarity constraints enforcing piecewise linearity

A piecewise linear function can be modeled using disjunctive constraints, even if the function is nonconvex. In Sect. 4.1, we derive an LPCC formulation for a problem where the piecewise linear function is given explicitly. Piecewise linear functions occur implicitly in quantile minimization, and LPCC formulations of such problems are discussed in Sect. 4.2.

4.1 Piecewise linear programming

The classical problem of a separable, piecewise linear program can be written:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \sum_{i=1}^{n} f_i(x_i) \\ \text{subject to} & Ax \geq b, \end{array} \tag{15}$$

where each $f_i(x_i)$ is a (possibly nonconvex) piecewise linear function given as:

$$f_i(x_i) \triangleq \begin{cases} \alpha_{i,1} + \beta_{i,1}\, x_i & \text{if } -\infty < x_i \leq \gamma_{i,1} \\ \alpha_{i,2} + \beta_{i,2}\, x_i & \text{if } \gamma_{i,1} \leq x_i \leq \gamma_{i,2} \\ \vdots & \vdots \\ \alpha_{i,p} + \beta_{i,p}\, x_i & \text{if } \gamma_{i,p-1} \leq x_i \leq \gamma_{i,p} \\ \alpha_{i,p+1} + \beta_{i,p+1}\, x_i & \text{if } \gamma_{i,p} \leq x_i < \infty, \end{cases}$$

for some constants $\alpha_{i,j}$, $\beta_{i,j}$, and $\gamma_{i,j}$ with $\gamma_{i,1} < \cdots < \gamma_{i,k}$ and $\alpha_{i,j} + \beta_{i,j}\gamma_{i,j} = \alpha_{i,j+1} + \beta_{i,j+1}\gamma_{i,j}$ for all $j = 1, \ldots, p$. While the latter equations ensure the continuity of f_i at the breakpoints $\gamma_{i,j}$, there is no guarantee that f_i is a convex function. To formulate (15) as an LPCC, let $y_{i,j}$ denote the portion of x_i in the interval $[\gamma_{i,j-1}, \gamma_{i,j}]$, where $\gamma_{i,0} = -\infty$ and $\gamma_{i,p+1} \triangleq \infty$. The variables $y_{i,j}$ satisfy the following conditions:

$$0 \leq \widehat{\gamma}_{i,j} - y_{i,j} \perp y_{i,j+1} \geq 0, \quad \forall\, j = 1, \ldots, p, \tag{16}$$

where

$$\widehat{\gamma}_{i,j} \equiv \begin{cases} \gamma_{i,1} & \text{if } j = 1 \\ \gamma_{i,j} - \gamma_{i,j-1} & \text{if } j = 2, \ldots, p. \end{cases}$$

In terms of the auxiliary variables $y_{i,j}$, we can write

$$x_i = \sum_{j=1}^{p+1} y_{i,j}, \quad \text{and} \quad f_i(x_i) = \alpha_{i,1} + \beta_{i,1}\, y_{i,1} + \sum_{j=2}^{p+1} \beta_{i,j}\, y_{i,j} \tag{17}$$

Substituting the expression of x_i into the constraint $Ax \geq b$, we obtain the following LPCC formulation of (15):

$$\begin{array}{ll} \underset{y_{i,j}}{\text{minimize}} & \sum_{i=1}^{n} \left[\alpha_{i,1} + \beta_{i,1}\, y_{i,1} + \sum_{j=2}^{p+1} \beta_{i,j}\, y_{i,j} \right] \\ \text{subject to} & \sum_{i=1}^{n} a_{\ell i} \sum_{j=1}^{p+1} y_{i,j} \geq b_\ell, \quad \ell = 1, \ldots, k \\ \text{and} & \text{the complementarity conditions(16).} \end{array} \tag{18}$$

Note that the complementarity constraints (16) cannot be dropped from the above formulation if the functions f_i are not convex. A simple 1-dimensional counter-example is given by: maximize $|t|$ subject to $t \in [-1, 1]$. The LPCC formulation of this problem is:

$$\begin{array}{ll} \text{maximize} & -t_1 + t_2 \\ \text{subject to} & t_1 \geq -1,\ t_2 \leq 1 \\ \text{and} & 0 \leq -t_1 \perp t_2 \geq 0. \end{array}$$

Without the complementarity constraint, the optimal solution is $(t_1, t_2) = (-1, 1)$, yielding $t = t_1 + t_2 = 0$ that is not optimal for the original absolute-value maximization problem.

It is useful to note how the complementarity constraints arise from this problem versus the previous hierarchical optimization problem. Previously, these constraints were needed to describe the optimality of an inner-level quadratic program, whereas here they are needed to express a logical relation between the linear segments of a piecewise linear function. Another noteworthy remark about the LPCC (18) is that (18) is equivalent to p^n LPs, which compares well with the exponential 2^{np} LP pieces in this LPCC with np complementarities. Finally, we mention that the complementarity representation (16) and (17) of a piecewise linear function allows the latter to appear in the constraints of an optimization problem; thus piecewise linear constraints can be modeled as linear complementarity constraints.

Nonconvex piecewise linear optimization has a long history. The common approach for treating this problem is to formulate it as a mixed integer program using special ordered sets. The Ph.D. thesis [30] and the subsequent references [31,32,55,56] study this problem and its extensions extensively and investigate branch and cut algorithms that are based on valid inequalities for special-ordered sets of type 2. An advantage of this contemporary approach is that the 0-1 variables are handled implicity by special branching rules. The LPCC formulation offers an alternative approach to these existing approaches; the detailed investigation of this complementarity formulation is regrettably beyond the scope of this paper.

4.2 Quantile minimization

Quantiles are fundamental statistical quantities. They have recently been used in risk analysis to assess probabilities of investment losses and as criteria for portfolio management. In what follows, we first give an LPCC formulation for a general quantile minimization problem and then examine the global minimization of the value-at risk (VaR) associated with a portfolio of risky assets using a scenario approach.

In order statistics, we are given m linear functions $b_i - a_i^T x$ and wish to choose x to minimize the kth largest. We assume x is constrained to lie in a polyhedron P (possibly $\mathbb{R}^n$). Here, b_i a scalar and a_i an n-vector. This problem can arise, for example, in chance constrained programming [42]. In particular, consider the problem

$$\min_{\alpha,x} \{\alpha \ : \ P_\xi[\alpha \geq f(x,\xi)] \geq 1-\gamma, \ x \in P\} \tag{19}$$

where $0 < \gamma < 1$, ξ is a random parameter and P is a polyhedron. Assume the uncertainty can be represented by m equally likely scenarios and if $f(x,\xi)$ is a linear function of x in each of these scenarios. If γ is not an integer multiple of $1/m$ then the problem can be represented as minimizing the kth largest of these m linear functions, with $k = \lceil \gamma m \rceil$.

Minimizing the maximum function can be formulated as a linear program. For other choices of k this is a nonconvex problem that can be expressed as the following LPCC:

$$\begin{array}{ll} \underset{\alpha,\beta,x,s}{\text{minimize}} & \alpha \\ \text{subject to} & \alpha + \beta_i \geq b_i - a_i^T x \quad i = 1,\ldots,m \\ & 0 \leq \beta \perp s \geq 0 \\ & \mathbf{1}^T s = m - k + 1 \\ & 0 \leq s \leq \mathbf{1}, \quad x \in P, \end{array} \tag{20}$$

where $x \in \mathbb{R}^n$, s and β are m-vectors, and α is a scalar. If $\beta_i > 0$ then α is smaller than the function value at the current x. The complementarity condition allows no more than $k-1$ components of β to be strictly positive. Given x, the values of β_i are nonnegative for the $k-1$ largest values of $f_i(x)$ and are equal to zero for the smallest $m-k+1$ functions, so α gives the value of the kth largest function for this x. Minimizing α leads to the minimum of the kth largest function.

We now consider the situation where the uncertainty in the chance constrained program (19) is represented by m scenarios where the scenarios are no longer assumed to be equally likely. The LPCC (20) can be generalized to handle this situation. Let p_i be the probability of the ith scenario, where $f(x, \xi_i) = b_i - a_i^T x$. The resulting formulation is

$$\begin{array}{ll} \underset{\alpha,\beta,x,s}{\text{minimize}} & \alpha \\ \text{subject to} & \alpha + \beta_i \geq b_i - a_i^T x \quad i = 1, \ldots, m \\ & 0 \leq \beta \perp s \geq 0 \\ & p^T s \geq 1 - \gamma \\ & 0 \leq s \leq \mathbf{1}, \quad x \in P. \end{array} \tag{21}$$

Any feasible solution to this formulation satisfies

$$\begin{aligned} P_\xi[\alpha \geq f(x, \xi)] &= \sum_{i:\alpha \geq b_i - a_i^T x} p_i \geq \sum_{i:\beta_i = 0} p_i \geq \sum_{i:s_i > 0} p_i \\ &\geq \sum_i p_i s_i = p^T s \geq 1 - \gamma, \end{aligned}$$

so it is feasible in (19). It can be shown similarly that a feasible solution to (19) gives a feasible solution to (21).

One application of quantile minimization arises in financial optimization. In a nutshell, the VaR minimization problem in portfolio selection is to choose, for a prescribed confidence level $\zeta \in (0, 1)$ of risk, an investment portfolio that is characterized by a deterministic vector $x \in \mathbb{R}^n$, where n is the number of financial instruments, so as to minimize the VaR of the portfolio subject to various restrictions on x; in turn, the VaR is the threshold of loss so that the probability of loss not exceeding this value is at least the given confidence level. To formulate this optimization problem mathematically, let r denote an n-dimensional random vector whose components represent the random losses of some financial instruments. Let $X \subseteq \mathbb{R}^n$ be a polyhedron representing the set of feasible investments. Adopting a scenario approach, let $\{r^1, \ldots, r^k\}$ be the finite set of scenario values of r, and $\{p^1, \ldots, p^k\}$ be the associated probabilities of the respective scenarios. As shown in [40], the portfolio selection problem of minimizing the VaR can be stated as the following LPCC:

$$\begin{array}{lllll} \underset{m,x,\tau,w,\lambda}{\text{minimize}} & m & & & \\ \text{subject to} & 0 \leq \tau_i \perp & \frac{p_i}{1-\zeta} - \lambda_i \geq 0 & & i = 1, \ldots, k \\ & 0 \leq \lambda_i \perp & w_i \triangleq m + \tau_i - x^T r^i \geq 0 & & i = 1, \ldots, k \\ \text{and} & x \in X \text{ and } & 1 = \sum_{i=1}^{k} \lambda_i. & & \end{array} \tag{22}$$

The formulation (22) constrains $0 \leq \lambda_i \leq \frac{p_i}{1-\zeta}$. On any piece of the LPCC, the bounded variables λ only appear in one constraint, namely $\sum_{i=1}^n \lambda_i = 1$. Thus, in any basic feasible solution at most one component λ_i will be basic; all other components will be nonbasic at either their upper or lower bounds. If $\lambda_i = 0$ then $\tau_i = 0$; if $\lambda_i = \frac{p_i}{1-\zeta}$ then $w_i = 0$; if $0 < \lambda_i < \frac{p_i}{1-\zeta}$ then both $\tau_i = 0$ and $w_i = 0$. Thus, τ and s are complementary. The possible pieces of the LPCC can be enumerated by considering the various cases for λ. The number of pieces where every component of λ is at a bound is no larger than 2^k. The number of pieces of the LPCC formulation of VaR where exactly one component of λ is not at a bound is no larger than $k2^{k-1}$: one component is basic and each of the remaining $k-1$ variables is at one of its bounds. Hence, the number of pieces of the LPCC formulation of VaR is no larger than $(k+2)2^{k-1}$.

Note that the complementarity restrictions impose additional limits on the number of components of τ and of s that can be positive. For example, if each $p_i = 1/k$ then exactly $r := \lceil k(1-\zeta)\rceil$ components of λ are positive, so enforcing complementarity requires fixing r components of s to zero and $k-r$ (if $k(1-\zeta)$ is integer) or $k-r+1$ components of τ to zero. Hence the number of pieces is only $\binom{k}{r}$ if $k(1-\zeta)$ is integer and $r\binom{k}{r}$ otherwise.

The VaR can also be modeled as a chance-constrained program [1,37]. The corresponding formulation (21) can be related to (22) through the change of variables $p_i(1-s_i) = (1-\zeta)\lambda_i$. It should also be noted that formulation (22) follows from the result that Value-at-Risk can be expressed as the optimal solution to a bilevel program where the inner problem is to minimize the Conditional Value-at-Risk (CVaR) [44,45].

5 Other applications

An LPCC is a representation of a problem with linear constraints together with some disjunctions. Any problem that is linear together with some either/or constraints can be naturally formulated as an LPCC. These either/or constraints can arise from combinatorial restrictions, and we consider a fundamental example of such a problems in Sect. 5.1. Linear programs with complementarity constraints can arise as approximations of more general problems. They can also arise in testing optimality of more general problems, as we discuss in Sect. 5.2.

Disjunctive constraints can be modeled as integer programs if a “Big-M” is introduced. This parameter depends on approximations of upper bounds on constraints or variables. The advantage of the LPCC approach to disjunctive constraints is that it does not require knowledge of bounds on constraints, so it is not necessary to introduce a “big-M” into the model.

5.1 Binary integer programming

In this subsection, we formally show that the general LPCC is NP-hard. One simple example of a combinatorial disjunction is in binary integer programming: a variable is either equal to one or to zero. A generic binary integer program can be written

$$\begin{array}{ll} \underset{x\in\{0,1\}^n}{\text{minimize}} & c^T x \\ \text{subject to} & Ax \ge b. \end{array}$$

This is equivalent to the LPCC

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & Ax \ge b \\ \text{and} & 0 \le x \perp \mathbf{1} - x \ge 0, \end{array}$$

where $\mathbf{1}$ denotes the vector of ones. Binary integer programming is NP-hard, so this reduction shows that the general LPCC is NP-hard. Of course, in practice, one would usually choose to solve the integer program directly rather than transform it into an LPCC.

5.2 B-stationarity of MPCCs

Consider an extension of the LPCC in which the objective function is nonlinear:

$$\begin{array}{ll} \underset{(x,y,w)}{\text{minimize}} & f(x,y,w) \\ \text{subject to} & Ax + By + Cw \geq b \\ \text{and} & 0 \leq y \perp w \geq 0. \end{array} \tag{23}$$

The concept of a B-stationary optimal point for an MPCC was formulated in the text [38] and named in the survey paper [48]. It was investigated further in [39]. A B-stationary point $(\bar{x}, \bar{y}, \bar{w})$ is one that is an optimal solution to the problem obtained by linearizing the original problem at the point $(\bar{x}, \bar{y}, \bar{w})$.

Let I denote the indices of the constraints $Ax + By + Cw \geq b$ that are active at the point $(\bar{x}, \bar{y}, \bar{w})$. The point is B-stationary if and only if the optimal value of the following LPCC is equal to zero:

$$\begin{array}{ll} \underset{(d_x,d_y,d_w)}{\text{minimize}} & \nabla_x f(\bar{x},\bar{y},\bar{w})^T d_x + \nabla_y f(\bar{x},\bar{y},\bar{w})^T d_y + \nabla_w f(\bar{x},\bar{y},\bar{w})^T d_w \\ \text{subject to} & A_i d_x + B_i d_y + C_i d_w \geq 0 \quad \text{for } i \in I \\ \text{and} & 0 \leq d_{y_j} \perp d_{w_j} \geq 0 \;\; \forall j \text{ with } \bar{y}_j = \bar{w}_j = 0 \\ & d_{y_j} \geq 0 \;\; \forall j \quad \text{with } \bar{y}_j = 0 \text{ and } \bar{w}_j > 0 \\ & d_{w_j} \geq 0 \;\; \forall j \quad \text{with } \bar{y}_j > 0 \text{ and } \bar{w}_j = 0 \end{array} \tag{24}$$

where A_i, B_I and C_i denote the ith rows of A, B, and C, respectively. The directions d_x, d_y, and d_w have the same dimensions as the original variables. This problem determines whether a certain type of improving direction exists. This is a homogeneous LPCC since it is concerned with the existence of a direction, so it either has optimal value zero or it has unbounded optimal value.

6 Algorithms

In this section we discuss two algorithms for finding a global optimum of an LPCC. The first approach can be used when simple bounds are available for all components of y and w. The second algorithm is a logical Benders decomposition approach [28] which can also determine if an LPCC is unbounded or infeasible.

6.1 Branch-and-cut

When upper bounds $y \leq \bar{y}$ and $w \leq \bar{w}$ are known, problem (1) is equivalent to the following integer programming problem:

$$\begin{array}{ll} \underset{(x,y,w,z)}{\text{minimize}} & c^T x + d^T y + e^T w \\ \text{subject to} & Ax + By + Cw \geq b \\ \text{and} & 0 \leq y_i \leq z_i \bar{y}_i \quad \text{for } i = 1, \ldots, m \\ & 0 \leq w_i \leq (1 - z_i)\bar{w}_i \quad \text{for } i = 1, \ldots, m \\ & z \text{ binary} \end{array} \tag{25}$$

where $z \in \mathbb{B}^m$. This formulation can then be solved with any standard integer programming solver. Relaxations of this formulation can be made tighter using the techniques discussed below in Sect. 7. The bounds $\bar{y}$ and $\bar{w}$ should be chosen as small as possible in order to make LP relaxations tighter.

If the complementary pair of variables y_i and w_i are both basic and strictly positive in a basic feasible solution to an LP relaxation of (1) then the rows of the simplex tableau corresponding to the basic variables y_i and w_i can be used to generate a valid linear constraint that is violated by the current point. This derivation is called a *simple cut* by Audet et al. [4], and is based on intersection cuts for 0-1 programming [5], and has also been investigated by Balas and Perregaard [8].

A branch-and-cut approach can be used even if explicit bounds on the variables are not available. Disjunctive cuts (including simple cuts) can be used to tighten the relaxation, as discussed in Sect. 7. The complementarities can be branched on. The relaxation at each node is a linear program, obtained by relaxing all unfixed complementarities. We have conducted preliminary computational testing of such an algorithm using a strong branching scheme, with very encouraging results, and some results are included in Sect. 8.

6.2 Logical benders decomposition

We developed a logical Benders decomposition approach [28] to handle the situation when simple upper bounds are not available on all the variables y and w. The existence of this algorithm makes it possible to formulate and solve various classes of problems as LPCCs; these problem classes could not be solved directly using integer programming because of the lack of bounds on at least some of the variables. Classical Benders decomposition can be used to solve (25) when bounds on the variables are available. In this subsection, we show the relationship between the classical Benders cutting planes and the cuts derived in the logical Benders decomposition approach. This relationship can be exploited when bounds are available for a proper subset of the y and w variables.

In the logical Benders decomposition approach, we first introduce a conceptually large scalar parameter Θ and construct the integer program:

$$\begin{array}{lrcll} \underset{(x,y,w,z)}{\text{minimize}} & c^T x + d^T y + e^T w & & & \\ \text{subject to} & Ax + By + Cw & \geq & b & \\ \text{and} & 0 \leq y_i & \leq & z_i \Theta & \text{for } i = 1, \ldots, m \\ & 0 \leq w_i & \leq & (1 - z_i)\Theta & \text{for } i = 1, \ldots, m \\ & z \text{ binary} & & & \end{array} \tag{26}$$

The LPCC is equivalent to the limiting case of this integer program as $\Theta \to \infty$. For a fixed value of $z = \bar{z}$, (26) is a linear program. The limiting dual linear program is

$$\begin{array}{ll} \underset{(\lambda,u,v)}{\text{maximize}} & b^T \lambda \\ \text{subject to} & A^T \lambda = c \\ & B^T \lambda - v \leq d \\ & C^T \lambda - u \leq e \\ & \bar{z}^T u + (\mathbf{1} - \bar{z})^T v = 0 \\ & (\lambda, u, v) \geq 0 \end{array} \tag{27}$$

where $(\lambda, u, v) \in \mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}^m$. We define $\varphi(\bar{z})$ to the value of this limiting dual LP. Let

$$\Xi \equiv \{(\lambda, u, v) \in \mathbb{R}^k_+ \times \mathbb{R}^m_+ \times \mathbb{R}^m_+ : A^T \lambda = c,\ B^T \lambda - v \leq d,\ C^T \lambda - u \leq e\}. \tag{28}$$

If $\Xi = \emptyset$ then the LPCC is either infeasible or unbounded.

The algorithm initializes with a master problem where every binary vector $z \in \mathbb{B}^m$ feasible. This corresponds to considering every possible assignment of the complementarities. As satisfiability constraints on z are added, various complementarity assignments are ruled out, either because the assignment is infeasible, or because the assignment cannot give a value better than a known feasible solution. A feasible $\bar{z}$ is chosen and subproblem (27) is solved with this $\bar{z}$. The solution to the subproblem provides information about the master problem and allows the generation of constraints to shrink the feasible region of the master problem. There are several cases:

- If (27) has a finite optimal value $\varphi(\bar{z})$ then we obtain a feasible solution to the LPCC. Let $(\bar{\lambda}, \bar{u}, \bar{v})$ be the optimal solution. Any z for which $(\bar{\lambda}, \bar{u}, \bar{v})$ is feasible must have value at least $\varphi(\bar{z})$, so such a z cannot be better than $\bar{z}$ and need not be considered further. A valid constraint can be added to the master problem so that such z are cut off. This constraint is called a *point cut*.
- If (27) is unbounded then the choice of $\bar{z}$ is infeasible in the original LPCC, so a *ray cut* is added to the master problem to ensure that subsequent choices of z do not allow the same ray.
- If (27) is infeasible then either $\bar{z}$ leads to an unbounded solution to the original LPCC, or the choice of $\bar{z}$ is infeasible. In the latter case, a ray exists in the homogeneous version of (27) and so a *ray cut* is added to the master problem. In the former case, the homogeneous version has optimal value 0 and this leads to confirmation that the corresponding primal piece is feasible and hence the original LPCC is unbounded.

The algorithm is summarized below:

1. **Initialize** the Master Problem with all binary z feasible.
2. Find a **feasible** $\bar{z}$ for the Master Problem.
3. Solve the subproblem (27).
    - (27) finite: Add a point cut to Master Problem.
    - (27) unbounded: Add a ray cut to the Master Problem.
    - (27) infeasible: Either LPCC is unbounded so **STOP**, or add a ray cut to the Master Problem.
4. If the Master Problem is infeasible, **STOP** with determination of the solution of LPCC: either it is infeasible, or the best feasible solution found is optimal.
5. **Return** to Step 2.

Given the optimal dual solution $(\bar{\lambda}, \bar{u}, \bar{v})$ to (27), the point cut has the form

$$\sum_{i:\bar{v}_i>0} (1 - z_i) + \sum_{i:\bar{u}_i>0} z_i \geq 1. \tag{29}$$

If we abuse notation and allow $(\bar{\lambda}, \bar{u}, \bar{v})$ to represent the ray when $\bar{z}$ is infeasible then the ray cut is also of the form (29). These point and ray cuts should be sparsified if possible, in order to strengthen them: the fewer terms that appear in the constraint, the larger the number of binary vectors that violate it. From the structure of (29), it is clear that the Master Problem is a satisfiability problem.

Benders decomposition can also be used to solve problem (25) when bounds are available on y and w. The approach constructs a subgradient approximation to the function $\varphi(z)$.

The extended real valued function $\varphi(.)$ is convex provided $\Xi \neq \emptyset$. For fixed $\bar{z}$, the dual to the LP relaxation of (25) can be written

$$\begin{array}{ll} \underset{(\lambda,u,v)}{\text{maximize}} & b^T\lambda - \sum_{i=1}^{m} \bar{y}_i(1-\bar{z}_i)v_i - \sum_{i=1}^{m} \bar{w}_i\bar{z}_i u_i \\ \text{subject to} & (\lambda, u, v) \in \Xi \end{array} \tag{30}$$

If $(\bar{\lambda}, \bar{u}, \bar{v})$ solves (30) for $z = \bar{z}$ then $\xi \in \mathbb{R}^m$ with $\xi_i = \bar{y}_i\bar{v}_i - \bar{w}_i\bar{u}_i$ is a subgradient of $\varphi(z)$ at $\bar{z}$. The corresponding subgradient inequality is

$$\begin{aligned} \varphi(z) &\geq b^T\bar{\lambda} - \sum_{i=1}^{m} \bar{y}_i(1-z_i)\bar{v}_i - \sum_{i=1}^{m} \bar{w}_i z_i \bar{u}_i \\ &= \varphi(\bar{z}) + \sum_{i=1}^{m}(z_i - \bar{z}_i)(\bar{y}_i\bar{v}_i - \bar{w}_i\bar{u}_i) \end{aligned} \tag{31}$$

Inequality (31) is the standard inequality used in Benders decomposition, specialized to problem (25). Note that the validity of this inequality only requires that $(\bar{\lambda}, \bar{u}, \bar{v})$ be feasible in (30). Thus, lower bounding inequalities on $\varphi(z)$ can be created from any feasible solution to (30). Note that this subgradient inequality is related to the point cut (29). In particular, we have the following theorem:

Theorem 1 *Let* $(\bar{\lambda}, \bar{u}, \bar{v})$ *solve* (30) *for* $z = \bar{z} \in \mathbb{B}^m$. *Any* $z \in \mathbb{B}^m$ *with* $\varphi(z) < \varphi(\bar{z})$ *must satisfy*

$$\sum_{i:\bar{v}_i>0,\bar{z}_i=1} (1-z_i) + \sum_{i:\bar{u}_i>0,\bar{z}_i=0} z_i \geq 1 \tag{32}$$

Proof The proof is by contraposition. Assume (32) is violated by z, so if $\bar{v}_i > 0$ then $z_i \geq \bar{z}_i$ and if $\bar{u}_i > 0$ then $z_i \leq \bar{z}_i$. The point $(\bar{\lambda}, \bar{u}, \bar{v})$ is feasible in (30) for any z, so

$$\begin{aligned} \varphi(z) &\geq b^T\bar{\lambda} - \sum_{i=1}^{m} \bar{y}_i(1-z_i)\bar{v}_i - \sum_{i=1}^{m} \bar{w}_i z_i \bar{u}_i \\ &\geq b^T\bar{\lambda} - \sum_{i=1}^{m} \bar{y}_i(1-\bar{z}_i)\bar{v}_i - \sum_{i=1}^{m} \bar{w}_i \bar{z}_i \bar{u}_i \\ &= \varphi(\bar{z}) \end{aligned}$$

giving the desired result. □

A similar result can be derived when (30) is unbounded:

Theorem 2 *Assume* $\varphi(\bar{z}) = +\infty$. *Let* $(\bar{\lambda}, \bar{u}, \bar{v})$ *be an optimal ray for* (30) *for* $z = \bar{z} \in \mathbb{B}^m$. *Any* $z \in \mathbb{B}^m$ *with finite optimal value must satisfy*

$$\sum_{i:\bar{v}_i>0,\bar{z}_i=1} (1-z_i) + \sum_{i:\bar{u}_i>0,\bar{z}_i=0} z_i \geq 1 \tag{33}$$

Proof The proof is by contraposition. Assume (33) is violated by z, so if $\bar{v}_i > 0$ then $z_i \geq \bar{z}_i$ and if $\bar{u}_i > 0$ then $z_i \leq \bar{z}_i$. The triple $(\bar{\lambda}, \bar{u}, \bar{v})$ is a ray in (30) for any z, and

$$b^T\bar{\lambda} - \sum_{i=1}^{m} \bar{y}_i(1-z_i)\bar{v}_i - \sum_{i=1}^{m} \bar{w}_i z_i \bar{u}_i \geq b^T\bar{\lambda} - \sum_{i=1}^{m} \bar{y}_i(1-\bar{z}_i)\bar{v}_i - \sum_{i=1}^{m} \bar{w}_i \bar{z}_i \bar{u}_i > 0.$$

It follows that $\varphi(z) = +\infty$. □

Thus, the standard subgradient cut implies the corresponding point or ray cut when bounds are available on all components of y and w. If bounds are available on only some of the components, a dual problem can be constructed that is a combination of (27) and (30). If the optimal dual solution only has positive components of u_i and v_j for the bounded variables then the subgradient cut (31) can be used; otherwise the point cut (29) can be added to the master problem.

7 Tightening the relaxation

The linear programming relaxation of the LPCC can be tightened by using disjunctive cuts or lift-and-project (see Balas et al. [6,7], Lovasz and Schrijver [36], and Sherali and Adams [50]). A tightened relaxation helps directly in an integer programming formulation of an LPCC. It also helps in the logical Benders decomposition approach when trying to sparsify the cuts (29), because the sparsification approach requires relaxing some of the complementarity restrictions.

Lift-and-project approaches involve forming products of variables and then linearizing the products. For example, the following quadratic inequalities are valid in the LPCC:

$$\sum_{j=1}^{n} A_{ij} x_j y_q + \sum_{l=1}^{m} B_{ik} y_l y_q + \sum_{p\neq q, p=1}^{m} C_{ip} w_p y_q \geq b_i y_q \quad \text{for } i = 1, \ldots, k, \; q = 1, \ldots, m$$

$$\sum_{j=1}^{n} A_{ij} x_j w_q + \sum_{l\neq q, l=1}^{m} B_{ik} y_l w_q + \sum_{p=1}^{m} C_{ip} w_p w_q \geq b_i w_q \quad \text{for } i = 1, \ldots, k, \; q = 1, \ldots, m,$$

as are the equalities $y_i w_i = 0$ for $i = 1, \ldots, m$. These quadratic constraints can be linearized by introducing matrices of variables. To simplify notation, we let

$$\gamma := \begin{pmatrix} x \\ y \\ w \end{pmatrix} \in \mathbb{R}^{n+2m}. \tag{34}$$

Let

$$M := \begin{bmatrix} X & \zeta^T & \Lambda^T \\ \zeta & Y & \Upsilon^T \\ \Lambda & \Upsilon & W \end{bmatrix} := \begin{bmatrix} x \\ y \\ w \end{bmatrix} \begin{bmatrix} x^T & y^T & w^T \end{bmatrix} = \gamma\gamma^T. \tag{35}$$

The diagonal entries of Υ must be zero in any feasible solution to the LPCC. The quadratic terms in the inequalities can be replaced by appropriate entries of the matrices defined in (35). The nonlinear equality (35) can itself be relaxed. For example, we could construct valid linear inequalities by using quadratic inequalities of the form

$$\left(\alpha + \sum_{i=1}^{n+2m} \beta_i \gamma_i \right)^2 \geq 0 \tag{36}$$

for constants $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}^{n+2m}$, and then replacing quadratic terms by the appropriate entries in M to give the valid linear constraint

$$\alpha^2 + 2\alpha \sum_{i=1}^{n+2m} \beta_i \gamma_i + \sum_{i=1}^{n+2m} \beta_i^2 M_{ii} + 2 \sum_{i=1}^{n+2m-1} \sum_{j=i+1}^{n+2m} \beta_i \beta_j M_{ij} \geq 0. \tag{37}$$

The vector β can be chosen to be sparse. It can also be chosen as a cutting plane based on eigenvectors of M, which could then be sparsified if desired. Approximating a semidefiniteness constraint using linear constraints is discussed in, for example, [35,43,51]. Any inequality derived in this way is implied by the positive semidefiniteness of the matrix

$$\hat{M} := \begin{bmatrix} 1 & \gamma^T \\ \gamma & M \end{bmatrix}, \tag{38}$$

as shown in the following standard lemma.

Lemma 3 *The set of γ and M which satisfy (37) for every $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}^{n+2m}$ is equal to*

$$\left\{ (\gamma, M) : \begin{bmatrix} 1 & \gamma^T \\ \gamma & M \end{bmatrix} \succeq 0 \right\} = \left\{ (\gamma, M) : M - \gamma\gamma^T \succeq 0 \right\}.$$

Proof Constraint (37) can be written as

$$\begin{bmatrix} \alpha & \beta^T \end{bmatrix} \begin{bmatrix} 1 & \gamma^T \\ \gamma & M \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \geq 0.$$

showing the first part of the lemma. The Schur complement result states that the matrix

$$\begin{bmatrix} 1 & \gamma^T \\ \gamma & M \end{bmatrix}$$

is positive semidefinite if and only if $M - \gamma\gamma^T \succeq 0$ [24]. □

In this way, a semidefinite programming relaxation of an LPCC can be constructed, with many potential linear constraints to tighten the relaxation. In certain applications, ideas of matrix completion [23] can be exploited, where positive semidefiniteness of $\hat{M}$ can be guaranteed by ensuring an appropriate submatrix of $\hat{M}$ is positive semidefinite. From a practical point of view, it may be useful to work with just a submatrix of $\hat{M}$ even when matrix completion is not available.

The semidefiniteness constraint only enforces one side of the matrix equality $M = \gamma\gamma^T$. Recently, the reference [46,47] has discussed methods for generating disjunctive cuts to enforce the nonconvex constraint $\gamma\gamma^T - M \succeq 0$. Disjunctive cuts can also be enforced as an algorithm proceeds, and used to cut off points that are not in the convex hull of the set of feasible solutions. See [7] for more details.

8 Computational experience

The logical Benders decomposition method of Hu et al. [28] has been used to find global optimal solutions of general LPCCs effectively. It has also been used to identify infeasible LPCCs and unbounded LPCCs. For more details, see the cited reference and also the doctoral thesis of Hu [26]. This method has been used to solve nonconvex global optimal solutions [27].

Vandenbussche and Nemhauser [52,53] developed an integer programming approach to solving the LPCC formulation of nonconvex quadratic programming problems with box constraints. Burer and Vandenbussche [10,11] used a semidefinite programming approach within a branch-and-bound algorithm to solve the LPCC formulation of box-constrained nonconvex quadratic programs, with impressive results.

Table 1 Computational experience with cross-validated support vector regression problems of dimension 5

Problem		Values		Branch-and-Cut			CPLEX	
Train	Test	RelaxLP	LPCCmin	LP	Nodes	Time	Nodes	Time
10	10	3.2191	8.9582	594	87	0.11	1,487	0.58
10	10	5.4535	15.2389	640	102	0.14	1,348	0.41
10	10	5.7514	10.4474	1,138	150	0.22	2,978	0.94
20	20	12.1247	15.7522	8,556	754	4.3	270,796	211.2
20	20	10.3882	16.3427	5,746	492	3.4	29,901	23.8
20	20	12.4111	18.3917	1,680	158	1.0	9,102	8.6
30	30	19.5499	23.6281	8,430	451	11.1	99,893	169.1
30	30	18.4676	24.2991	11,030	656	13.3	64,938	84.6
30	30	15.7274	18.6622	3,048	208	3.6	34,328	54.3
40	40	31.3081	39.3843	24,886	1,184	51.8	539,128	1407.9
40	40	34.9672	42.6682	39,080	2,312	75.7	719,346	2145.5
40	40	30.0506	34.0263	6,478	373	14.1	81,957	205.1
50	50	33.5479	36.7149	14,004	729	51.9	172,908	981.0
50	50	37.3623	39.3964	19,508	1,132	63.1	422,080	1699.7
50	50	41.7643	50.9952	90,258	5,109	288.9	–	–
60	60	37.6553	50.3336	136,708	6,569	707.4	–	–
60	60	49.1863	51.5881	26,772	1,296	143.9	348,278	2522.7
60	60	40.8572	47.6830	59,118	2,439	347.4	–	–
70	70	47.4771	55.1966	62,670	2,118	542.1	–	–
70	70	45.6660	56.0318	133,786	6,793	1178.3	–	–
70	70	46.7967	54.3473	45,546	2,097	400.0	–	–
80	80	64.3488	69.2778	41,852	1,724	550.0	–	–
80	80	69.2562	74.0414	46,624	1,571	586.5	–	–
80	80	55.5947	59.9417	60,956	1,741	866.4	–	–
90	90	75.5819	77.4303	43,332	1,626	644.0	–	–
90	90	64.6352	76.0012	196,554	7,870	2962.5	–	–
90	90	74.2927	78.9140	68,278	2,259	1052.3	–	–
100	100	85.2575	87.154	48,526	1,432	1085.2	–	–
100	100	81.3640	84.1875	49,998	1,534	1069.9	–	–
100	100	69.3205	75.2687	302,730	11,001	5672.2	–	—

Recently, we have experimented with a specialized branching algorithm for solving the cross-validated support vector regression problem described in Sect. 3.3. Branching is performed to fix the complementarities. Multiple linear programming relaxations of the problem are solved at each node of a tree, with branching decisions based on the solutions of these linear programs. The algorithm hence has similarities to strong branching rules in algorithms for integer programming. Preliminary computational results with this algorithm are contained in Table 1.

In the notation of Sect. 3.3, the test problems have $N = 1$, the number of test points is equal to $|\Omega|$, the number of training points is $|\overline{\Omega}|$, and the dimension of the vectors $\mathbf{w}$ and $\mathbf{x}_i$ is equal to five. The table gives the optimal value of the LP relaxation, the optimal value of the

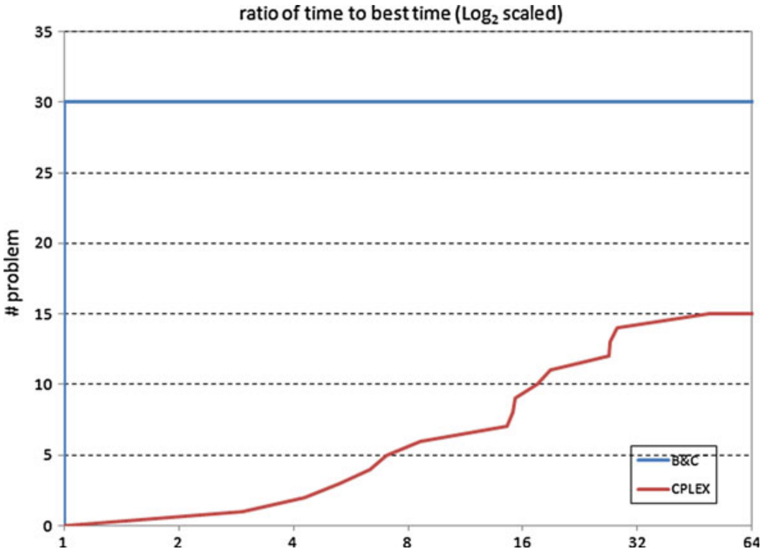

Fig. 1 Performance profile of runtimes on the 30 cross-validated support vector regression problems. The *left* axis gives the number of problems solved, and the *plot* indicates the number of problems solved within a given ratio of the best time for that problem. Ratios for the branch-and-cut code and for CPLEX are given

LPCC formulation, the number of linear programming subproblems solved, the size of the branch and bound tree, and the runtime for the branch-and-cut approach, and the size of the branch-and-bound tree and the runtime for CPLEX. An upper bound of 18000 seconds was imposed on the runtime. The runtimes are in seconds on an AMD Phenom II X4 955 4 core CPU @3.2GHZ with 4 gb of memory, using a 64 bit windows operating system, and running on a single core. The tolerance for optimality is 10^{-6} and for complementarity is 10^{-5}. The CPLEX result is from solving an MIP formulation of the LPCC, using the CPLEX 11.0 callable library with indicator constraints, with the default CPLEX settings. The branch-and-cut algorithm used CPLEX 11.0 to solve the LP subproblems. A performance profile of these results is contained in Fig. 1. Note that the scale of the time axis in the figure is logarithmic. It is clear that our algorithm dramatically outperforms a default application of CPLEX for these problems.

9 Concluding remarks

Disjunctive constraints arise in many settings. For example, they arise when constructing optimality conditions for nonlinear programs; when these nonlinear programs appear as constraints in an optimization problem, a natural formulation for the complete problem is often a mathematical program with complementarity constraints, of which the LPCC is the simplest instance. An LPCC is an NP-hard nonconvex nonlinear program so finding a global optimum is non-trivial. The importance and breadth of applications of LPCCs make further research on methods for determining global optimality imperative. Multiple alternative approaches are possible, and several have been investigated. Determining efficient methods (or combination of methods) for particular applications is an intriguing challenge that requires sustained investigation. By presenting a host of realistic contexts where the LPCC arises, we hope to have established the fundamental significance of this class of global optimization problems and have motivated its further research.

References

1. Ahmed, S., Shapiro, A.: Solving chance-constrained stochastic programs via sampling and integer programming. In: Chen, Z.L., Raghavan, S. (eds.) TutORials in Operations Research, chapter 12, pp. 261–269. INFORMS, London (2008)
2. Ahuja, R.K., Orlin, J.B.: Inverse optimization. Oper. Res. **49**(5), 771–783 (2001)
3. Anandalingam, G., Friesz, T.L.: Hierarchical optimization: an introduction. Ann. Oper. Res. **34**(1), 1–11 (1992)
4. Audet, C., Savard, G., Zghal, W.: New branch-and-cut algorithm for bilevel linear programming. J. Optim. Theory Appl. **38**(2), 353–370 (2007)
5. Balas, E.: Intersection cuts — a new type of cutting planes for integer programming. Oper. Res. **19**, 19–39 (1971)
6. Balas, E., Ceria, S., Cornuéjols, G.: A lift-and-project cutting plane algorithm for mixed 0-1 programs. Math. Program. **58**, 295–324 (1993)
7. Balas, E., Perregaard, M.: Lift-and-project for mixed 0-1 programming: recent progress. Dis. Appl. Math. **123**(1–3), 129–154 (2002)
8. Balas, E., Perregaard, M.: A precise correspondence between lift-and-project cuts, simple disjunctive cuts, and mixed integer Gomory cuts for 0-1 programming. Math. Program. **94**(2-3), 221–245 (2003)
9. Bennett, K.P., Ji, X., Hu, J., Kunapuli, G., Pang, J.-S.: Model selection via bilevel programming. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN'06) Vancouver, pp. 1922–1929. B.C. Canada, July 16–21 (2006)
10. Burer, S., Vandenbussche, D.: A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations. Math. Program. **113**(2), 259–282 (2008)
11. Burer, S., Vandenbussche, D.: Globally solving box-constrained nonconvex quadratic programs with semidefinite-based finite branch-and-bound. Comput. Optim. Appl. **43**(2), 181–195 (2009)
12. Chang, M.W., Lin, C.J.: Leave-one-out bounds for support vector regression model selection. Neural Comput. **17**, 1188–1222 (2005)
13. Chapelle, O., Vapnik, V., Bousquet, O., Mukherjee, S.: Choosing multiple parameters with support vector machines. Mach. Learn. **46**, 131–159 (2002)
14. Chung, K.W., Kao, W.C., Sun, C.L., Wang, L.L., Lin, C.J.: Radius margin bounds for support vector machines with the RBF kernel. Neural Comput. **15**, 2643–2681 (2003)
15. Colson, B., Marcotte, P., Savard, G.: An overview of bilevel optimization. Ann. Oper. Res. **153**(1), 235–256 (2007)
16. Cottle, R.W., Pang, J.S., Stone, R.E.: The linear complementarity problem. SIAM Classics in Applied Mathematics 60, Philadelphia (2009) [Originally published by Academic Press, Boston (1992)]
17. Dempe, S.: Foundations of Bilevel Programming. Kluwer, Dordrecht (2002)
18. Dempe, S., Dutta, J.: Is bilevel programming a special case of a mathematical program with complementarity constraints? Math. Program (2010) (online first)
19. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems: vols. I and II. Springer, New York (2003)
20. Floudas, C.A., Pardalos, P.M.: A Collection of Test Problems for Constrained Global Optimization Algorithms, vol. 455 of Lecture Notes in Computer Science. Springer, New York (1990)
21. Floudas, C.A., Pardalos, P.M., Adjiman, C., Esposito, W.R., Gümüs, Z.H., Harding, S.T., Klepeis, J.L., Meyer, C.A., Schweiger, C.A.: Handbook of Test Problems in Local and Global Optimization, vol. 33 of Nonconvex Optimization and its Applications. Kluwer, Dordrecht (1999)
22. Giannessi, F., Tomasin, E.: Nonconvex quadratic programs, linear complementarity problems, and integer linear programs. In: Conti, R., Ruberti, A. (eds.) Fifth Conference on Optimization Techniques (Rome 1973), Part I, vol. 3 of Lecture Notes in Computer Science, pp. 437–449. Springer, Berlin (1973)
23. Grone, B., Johnson, C.R., Marques de Sa, E., Wolkowicz, H.: Positive definite completions of partial Hermitian matrices. Lin. Alg. Appl. **58**, 109–124 (1984)
24. Horn, R.A., Johnson, C.: Matrix Analysis. Cambridge University Press, Cambridge (1985)
25. Horst, R., Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization, 2nd edn, Vol. 48 of Nonconvex Optimization and its Applications. Kluwer, Dordrecht (2000)
26. Hu, J.: On linear programs with linear complementarity constraints. PhD thesis, Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, August (2009)

27. Hu, J., Mitchell, J.E., Pang, J.S.: An LPCC approach to nonconvex quadratic programs. Technical report, Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180. Accepted for publication in Mathematical Programming. May (2008)
28. Hu, J., Mitchell, J.E., Pang, J.S., Bennett, K.P., Kunapuli, G.: On the global solution of linear programs with linear complementarity constraints. SIAM J. Optim. **19**(1), 445–471 (2008)
29. Iyengar, G., Kang, W.: Inverse conic programming with applications. Oper. Res. Lett. **33**(3), 319–330 (2005)
30. Keha, A.B.: A polyhedral study of nonconvex piecewise linear optimization. PhD thesis, Industrial Engineering, Georgia Institute of Technology, Atlanta, GA, August (2003)
31. Keha, A.B., de Farias, I.R. Jr., Nemhauser, G.L.: Models for representing piecewise linear cost functions. Oper. Res. Lett. **32**(1), 44–48 (2004)
32. Keha, A.B., de Farias, I.R. Jr., Nemhauser, G.L.: A branch-and-cut algorithm without binary variables for nonconvex piecewise linear optimization. Oper. Res. **54**(5), 847–858 (2006)
33. Kunapuli, G., Hu, J., Bennett, K.P., Pang, J.S.: Bilevel model selection for support vector machines. In: Data Mining and Mathematical Programming, vol 45 of CRM Proceedings and Lecture Notes, pp. 129–158. Centre de Recherches Mathématiques (2008)
34. Kunapuli, G., Hu, J., Bennett, K.P., Pang, J.S.: Classification model selection via bilevel programming. Optim. Methods Softw. **23**(4), 475–489 (2008)
35. Krishnan, K., Mitchell, J.E.: A unifying framework for several cutting plane methods for semidefinite programming. Optim. Methods Softw. **21**(1), 57–74 (2006)
36. Lovász, L., Schrijver, A.: Cones of matrices and set-functions and 0-1 optimization. SIAM J. Optim. **1**(2), 166–190 (1991)
37. Luedtke, J.: Integer programming approaches for some non-convex and stochastic optimization problems. PhD thesis, Industrial and Systems Engineering, Georgia Institute of Technology (2007)
38. Luo, Z.-Q., Pang, J.-S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
39. Pang, J.-S., Fukushima, M.: Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. Comput. Optim. Appl. **13**(1–3), 111–136 (1999)
40. Pang, J.-S., Leyffer, S.: On the global minimization of the Value-at-Risk. Optim. Methods Softw. **19**(5), 611–631 (2004)
41. Pardalos, P.M.: The linear complementarity problem. In: Gomez, S., Hennart, J.P. (eds.) Advances in Optimization and Numerical Analysis, pp. 39–49 (1994)
42. Prékopa, A.: Probabilistic programming. In: Ruszczynski, A., Shapiro, A. (eds.) Stochastic Programming, chapter 5, Handbooks in Operations Research and Management Science, vol. 10., Elsevier, Amsterdam (2003)
43. Qualizza, A., Belotti, P., Margot, F.: Linear programming relaxations of quadratically constrained quadratic programs. IMA Volume Series Springer, accepted (2010)
44. Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. J. Risk **2**, 21–41 (2000)
45. Rockafellar, R.T., Uryasev, S.: Conditional value-at-risk for general loss distribution. J. Bank. Finance **26**, 1443–1471 (2002)
46. Saxena, A., Bonami, P., Lee, J.: Convex relaxations of non-convex mixed integer quadratically constrained programs: extended formulations. Math. Program. **124**(1–2), 383–411 (2010)
47. Saxena, A., Bonami, P., Lee, J.: Convex relaxations of non-convex mixed integer quadratically constrained programs: projected formulations. Technical Report RC24695, IBM. Accepted for publication in Mathematical Programming August (2008)
48. Scheel, H., Scholtes, S.: Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. Math. Oper. Res. **25**(1), 1–22 (2000)
49. Scholtes, S.: Active set methods for inverse linear complementarity problems. Technical report, Judge Institute of Management Science, Cambridge University, Cambridge CB2 1AG, November (1999)
50. Sherali, H.D., Adams, W.P.: A hierarchy of relaxation between the continuous and convex hull representations for zero-one programming problems. SIAM J. Dis. Math. **3**, 411–430 (1990)
51. Sherali, H.D., Fraticelli, B.M.P.: Enhancing RLT relaxations via a new class of semidefinite cuts. J. Global Optim. **22**, 233–261 (2002)
52. Vandenbussche, D., Nemhauser, G.L.: A branch-and-cut algorithm for nonconvex quadratic programs with box constraints. Math. Program. **102**(3), 559–575 (2005)
53. Vandenbussche, D., Nemhauser, G.L.: A polyhedral study of nonconvex quadratic programs with box constraints. Math. Program. **102**(3), 531–557 (2005)
54. Vapnik, V.: Statistical Learning Theory. Wiley-Interscience, New York (1998)

55. Vielma, J.P., Ahmed, S., Nemhauser, G.L.: Mixed-integer models for nonseparable piecewise linear optimization: Unifying framework and extensions. Oper. Res. **58**(2), 303–315 (2010)
56. Vielma, J.P., Keha, A.B., Nemhauser, G.L.: Nonconvex, lower semicontinuous piecewise linear optimization. Dis. Optim. **5**(2), 467–488 (2008)
57. Xiao, X., Zhang, L., Zhang, J.: A smoothing Newton method for a type of inverse semi-definite quadratic programming problem. J. Comput. Appl. Math. **223**(1), 485–498 (2009)
58. Zhang, J., Zhang, L.: An augmented Lagrangian method for a type of inverse quadratic programming problems. Appl. Math. Optim. **61**(1), 57–83 (2010)