J Glob Optim (2008) 40:505–527
DOI 10.1007/s10898-006-9109-x

ORIGINAL PAPER

Box-constrained minimization reformulations of complementarity problems in second-order cones

R. Andreani · A. Friedlander · M. P. Mello · S. A. Santos

Received: 16 October 2006 / Accepted: 17 October 2006 / Published online: 21 February 2007

Abstract Reformulations of a generalization of a second-order cone complementarity problem (GSOCCP) as optimization problems are introduced, which preserve differentiability. Equivalence results are proved in the sense that the global minimizers of the reformulations with zero objective value are solutions to the GSOCCP and vice versa. Since the optimization problems involved include only simple constraints, a whole range of minimization algorithms may be used to solve the equivalent problems. Taking into account that optimization algorithms usually seek stationary points, a theoretical result is established that ensures equivalence between stationary points of the reformulation and solutions to the GSOCCP. Numerical experiments are presented that illustrate the advantages and disadvantages of the reformulations.

Keywords Complementarity problems · Minimization algorithms · Reformulations

Mathematical Subject Classifications 90C33 · 90C30

Given $F, G : \mathbb{R}^n \to \mathbb{R}^n$, we consider the following generalized second-order cone complementarity problem GSOCCP$(F, G, \mathcal{K})$ of finding $x \in \mathbb{R}^n$ such that

$$G(x) \in \mathcal{K}, F(x) \in \mathcal{K}^{\circ}, F(x)^T G(x) = 0, \tag{1}$$

Supported by FAPESP (01/04597-4), CNPq, PRONEX-Optimization, FAEPEX-Unicamp.

R. Andreani (✉) · A. Friedlander · M. P. Mello · S. A. Santos
Department of Applied Mathematics, IMECC-UNICAMP, University of Campinas,
CP 6065, 13081-970 Campinas, SP, Brazil
e-mail: andreani@ime.unicamp.br

A. Friedlander
e-mail: friedlan@ime.unicamp.br

M. P. Mello
e-mail: margarid@ime.unicamp.br

S. A. Santos
e-mail: sandra@ime.unicamp.br

where $\mathcal{K}$ is the convex cone

$$\mathcal{K} = \left\{ x \in \mathbb{R}^n \mid x_1^2 \geq \sum_{j=2}^{n} a_j^2 x_j^2, x_1 \geq 0 \right\}$$

and its dual cone $\mathcal{K}^\circ$ is defined by

$$\mathcal{K}^\circ = \{x \in \mathbb{R}^n \mid \forall y \in \mathcal{K}, \langle x, y \rangle \geq 0\}.$$

The case usually considered in the literature is to have $a_i = 1$, for $i = 1, \ldots, n$, that is, $\mathcal{K}$ is a Lorentz cone. In some articles results are obtained for the case where $\mathcal{K}$ is the Cartesian product of Lorentz cones.

If we let, without loss of generality, $A = \mathrm{diag}(1, -a_2^2, -a_3^2, \ldots, -a_p^2, 0, \ldots, 0)$ (where $a_i \neq 0$ for $2 \leq i \leq p$), $\bar{A} = \mathrm{diag}(1, -1/a_2^2, \ldots, -1/a_p^2, 0, \ldots, 0)$ and $M = \mathrm{diag}(m_1, \ldots, m_n)$, where $m_i = 0$ for $1 \leq i \leq p$ and $m_i = 1$ for $i > p$, the convex cones considered may be expressed in matrix form as

$$\mathcal{K} = \left\{ x \in \mathbb{R}^n \mid \frac{1}{2} x^T A x \geq 0, x_1 \geq 0 \right\}$$

and

$$\mathcal{K}^\circ = \left\{ x \in \mathbb{R}^n \mid \frac{1}{2} x^T \bar{A} x \geq 0, x_1 \geq 0, Mx = 0 \right\}.$$

Note that $A\bar{A} = \mathrm{diag}(\mathbf{1}_p, 0, \ldots, 0)$ and $AM = MA = \bar{A}M = M\bar{A} = 0$, where $\mathbf{1}_p = (1, \ldots, 1) \in \mathbb{R}^p$. Of course, by a convenient scaling of x, we may assume $a_i = 1$, for $i = 2, \ldots, p$. We do adopt this assumption henceforth to simplify notation even further. Still the cone considered here is more general, because p may be strictly less than n. This implies, in particular, that $\mathcal{K}^\circ$ may be different from $\mathcal{K}$.

The GSOCCP($F, G, \mathcal{K}$) is a difficult problem because, although $\mathcal{K}$ is a convex set, there isn't a really "nice" way of defining it. In the widely used definition $\mathcal{K} = \{x \in \mathbb{R}^n \mid g_N(x) = -x_1 + \|(x_2, \ldots, x_p)\| \leq 0\}$, the function g_N in the constraint set is convex but it is not differentiable at the origin. In the equivalent definition adopted herein, namely $\mathcal{K} = \{x \in \mathbb{R}^n \mid x_1 \geq 0,\ \ g_D(x) = -x_1^2 + \|(x_2, \ldots, x_p)\|^2 \leq 0\}$, the two functions that appear in the constraint set are smooth, but g_D is not convex. Furthermore, at the origin we have $\nabla g_D(0) = 0$, and thus the origin is not a regular point. Thus the task of reformulating this problem via nonlinear programming is particularly challenging.

One important special case of GSOCCP is the Karush–Kuhn–Tucker (KKT) optimality conditions for the second-order cone program (SOCP), that consists in a generalization of the linear programming problem where the positivity of the variables is substituted by the requirement that the variables belong to a cone $\mathcal{K}$. On the other hand, if we allow $\mathcal{K}$ in (1) to be the Cartesian product of second-order cones, then the nonlinear complementarity problem (NCP) is also a special case of GSOCCP. These problems and their numerous applications are extensively discussed in [1,11,18].

We are interested in reformulations that preserve the smoothness properties of F and G in (1). In [3] the authors analyzed the case where $\mathcal{K}$ is a polyhedral cone. In [8,9], smooth merit functions, based on the well known Fischer-Burmeister NCP-function [12], are presented for the second-order cone complementarity problem (SOCCP): finding orthogonal $F(x)$ and $G(x)$ belonging to $\mathcal{K}$, a self-dual cone. Notice that this precludes the case $p < n$. In particular, [8,9] focus on the case where $\mathcal{K}$ is the Cartesian product of Lorentz cones. The results obtained in these papers are extensions of the results for the NCP in [10,14]. Nice properties of the FB-function are extended to the SOCCP via the *Jordan product*, see [8]. The authors prove that if ∇F and $-\nabla G$ are column monotone, every stationary point of the merit function they propose solves the SOCCP. In [9] similar results are obtained for a modification of the merit function defined in [8], where the authors extend the results in [20]. The approach suggested in [8] was implemented and compared with ours.

Hayashi et al. [16] focus on the special (SOCCP) where $G(x) = x$ and $\mathcal{K}$ is the Cartesian product of Lorentz cones. They also employ results from Jordan algebra to construct a nondifferentiable merit function. An algorithm is constructed that converges quadratically to a solution of SOCCP under certain assumptions. The algorithm works with a sequence of approximations to the original merit function, combining regularization and smoothing strategies. This involves the introduction of two sequences of parameters that are adjusted throughout the algorithm. The algorithm is tested on several problems, some randomly generated and one nonlinear SOCCP. We solve a subset of these problems in Sect. 4.4 using the approach proposed herein.

In this paper we construct two reformulations of the GSOCCP, using ten and five extra variables per cone, respectively. Both are nonlinear minimization problems with box constraints. The merit functions inherit the same degree of differentiability as the original data. Any efficient bound constrained minimization algorithm for large scale problems can be used to solve the reformulated problem. Our choice of code for the numerical tests was just a matter of convenience. The strong points of our second reformulation are the easiness of implementation, smoothness preservation and good discrimination capacity, i.e., stationary points which are not solutions have high objective function values (in comparison with the threshold adopted) and thus do not lead to “false positives.”

The paper is organized as follows. In Sect. 1 we introduce the first reformulation and prove a global equivalence result. In Sect. 2 we discuss some computational aspects and propose a second alternative reformulation. In Sect. 3 we give conditions under which a stationary point of a problem originating from the first reformulation provides a solution of the original GSOCCP. Section 4 presents numerical experiments which include small problems, ana application of SOCP in grasping force optimization, suggested in [18], with data adapted from [15], and ramdonly generated problems with larger dimension, as in [16]. To conclude, Sect. 5 is dedicated to final remarks.

1 An equivalent reformulation

If x^* is a solution of the generalized cone complementarity problem (1), then $G(x^*)$ and $F(x^*)$ solve (2) and (3) below, respectively.

$$\begin{array}{ll} \min & \langle F(x^*), x\rangle, \\ \text{s.t.} & x \in \mathcal{K}, \end{array} \tag{2}$$

and

$$\begin{array}{ll} \min & \langle G(x^*), x\rangle, \\ \text{s.t.} & x \in \mathcal{K}^\circ. \end{array} \tag{3}$$

An approach that has worked before, e.g. in [3], was to formulate a merit function that embodies the KKT conditions of either (2) or (3). This role is played by the functions f and g defined below. Consequently, in addition to the original set of variables, these merit functions also depend on an extra set, that includes the Lagrange multipliers associated with (2) and (3), respectively. The ultimate objective is to show that GSOCCP is equivalent to the problem of minimizing the merit function, in the sense that, if GSOCCP has a solution then the merit function attains its lower bound of zero and vice-versa. This involves, of course, being able to ascertain the existence of Lagrange multipliers at the optimal solution to (2), or (3). The difficulty here is that the origin does not satisfy any type of constraint qualification, see [4,5], since $\nabla g_D(0) = 0$. We make up for this lack of regularity by uniting the two functions in a convex combination. Notice, however, that in the special case $G(x) = x$, matters could be simplified by first verifying whether $F(0) \in \mathcal{K}^\circ$, in which case $x^* = 0$ solves the generalized cone complementarity problem (1). If not, then the origin is not the solution, and f alone could be used as a merit function.

Let

$$f(x,\lambda,\mu,z,y) = \|F(x) - \lambda AG(x) - \mu e_1\|^2 + \left(\frac{1}{2}G(x)^T AG(x) - z\right)^2 + (G_1(x) - y)^2 + (\lambda z)^2 + (\mu y)^2 \tag{4}$$

and

$$g(x,\xi,\nu,w,s,\zeta) = \|G(x) - \xi \bar{A}F(x) - \nu e_1 - M\zeta\|^2 + \left(\frac{1}{2}F(x)^T \bar{A}F(x) - w\right)^2 + (F_1(x) - s)^2 + \|MF(x)\|^2 + (\xi w)^2 + (\nu s)^2, \tag{5}$$

where $x, \zeta \in \mathbb{R}^n$, $\lambda, \mu, z, y, \xi, \nu, w, s \in \mathbb{R}$ and $e_1 = (1, 0, \ldots, 0)^T \in \mathbb{R}^n$. Setting $v_f = (\lambda, \mu, z, y)$ and $v_g = (\xi, \nu, w, s)$, we define $\phi(x, v_f, v_g, \zeta, r) = rf(x, v_f) + (1-r)\, g(x, v_g, \zeta)$. The optimization problem

$$\begin{aligned} \min\ & \phi(x, v_f, v_g, \zeta, r), \\ \text{s.t.}\ & 1 \geq r \geq 0, \\ & v_f,\ v_g \geq 0. \end{aligned} \tag{6}$$

is a reformulation of the GSOCCP$(F, G, \mathcal{K})$ in the sense spelled out in the following theorem.

Theorem 1 *If $(x^*, v_f^*, v_g^*, \zeta^*, r^*)$ is a global minimizer of (6) with objective value zero then x^* is a solution to GSOCCP$(F, G, \mathcal{K})$.*
Conversely, if x^* is a solution to GSOCCP$(F, G, \mathcal{K})$, then there exist

$$v_f^*, v_g^* \geq 0, \quad 1 \geq r^* \geq 0 \quad \text{and} \quad \zeta^*$$

such that $(x^*, v_f^*, v_g^*, \zeta^*, r^*)$ is a global minimizer of (6) with objective value zero.

Proof First of all, notice that, for fixed x, v_f, v_g, ζ, the objective function value, for feasible values of r, is the convex combination of $f(x, v_f)$ and $g(x, v_g, \zeta)$, that is, a number on the line segment

$$\left[\min(f(x, v_f), g(x, v_g, \zeta)), \max(f(x, v_f), g(x, v_g, \zeta))\right].$$

But the minimum value on this line segment is achieved at its left end. Therefore, minimizing the convex combination of $f(x, v_f)$ and $g(x, v_g, \zeta)$ is equivalent to minimizing the minimum of $f(x, v_f)$ and $g(x, v_g, \zeta)$.

Now, since f and g are nonnegative, the objective value is zero only if

$$f(x^*, v_f^*) = 0 \quad \text{or} \quad g(x^*, v_g^*, \zeta^*) = 0.$$

First consider the case $f(x^*, v_f^*) = 0$. Clearly $G(x^*) \in \mathcal{K}$. Moreover,

$$\begin{aligned} F(x^*)^T \bar{A} F(x^*) &= (\lambda^* AG(x^*) + \mu^* e_1)^T \bar{A} (\lambda^* AG(x^*) + \mu^* e_1) \\ &= \lambda^{*2} G(x^*)^T A\bar{A}AG(x^*) + 2\lambda^*\mu^* G(x^*)^T A\bar{A}e_1 + \mu^{*2} e_1^T \bar{A} e_1 \\ &= \lambda^{*2} G(x^*)^T AG(x^*) + 2\lambda^*\mu^* G_1(x^*) + \mu^{*2} \\ &= 2\lambda^{*2} z^* + 2\lambda^*\mu^* y^* + \mu^{*2} \geq 0. \end{aligned}$$

Also,

$$F_1(x^*) = \lambda^* G_1(x^*) + \mu^* = \lambda^* y^* + \mu^* \geq 0$$

and, for $i > p$, $F_i(x^*) = \lambda^* a_{ii} G_i(x^*) = 0$.

Therefore $F(x^*) \in \mathcal{K}^\circ$. It remains to show that the complementarity condition is satisfied:

$$\begin{aligned} F(x^*)^T G(x^*) &= (\lambda^* AG(x^*) + \mu^* e_1)^T G(x^*) \\ &= \lambda^* G(x^*)^T AG(x^*) + \mu^* G_1(x^*) \\ &= 2\lambda^* z^* + \mu^* y^* = 0. \end{aligned}$$

Now, assuming $g(x^*, v_g^*, \zeta^*) = 0$, it follows easily that $F(x^*) \in \mathcal{K}^\circ$. In order to verify that $G(x^*) \in \mathcal{K}$, we calculate

$$\begin{aligned} G(x^*)^T AG(x^*) &= \left(\xi^* \bar{A} F(x^*) + \nu^* e_1 + M\zeta^*\right)^T A \left(\xi^* \bar{A} F(x^*) + \nu^* e_1 + M\zeta^*\right) \\ &= \xi^{*2} F(x^*)^T \bar{A} A \bar{A} F(x^*) + \nu^{*2} e_1^T A e_1 + \zeta^{*T} MAM\zeta^* \\ &\quad + 2\xi^*\nu^* F(x^*)^T \bar{A} A e_1 + 2\xi^* F(x^*)^T \bar{A} AM\zeta^* + 2\nu^* e_1^T AM\zeta^* \\ &= \xi^{*2} F(x^*)^T \bar{A} F(x^*) + \nu^{*2} + 0 + 2\xi^*\nu^* F_1(x^*) + 0 + 0 \\ &= 2\xi^{*2} w^* + \nu^{*2} + 2\xi^*\nu^* s^* = \nu^{*2} \geq 0. \end{aligned}$$

The nonnegativity of $G(x^*)$'s first component follows analogously

$$G_1(x^*) = \xi^* F_1(x^*) + \nu^* = \xi^* s^* + \nu^* \geq 0,$$

completing the proof that $G(x^*) \in \mathcal{K}$.

Checking the complementarity condition:

$$\begin{aligned} F(x^*)^T G(x^*) &= F(x^*)^T \left(\xi^* \bar{A} F(x^*) + \nu^* e_1 + M\zeta^*\right) \\ &= \xi^* F(x^*)^T \bar{A} F(x^*) + \nu^* F_1(x^*) + F(x^*)^T M\zeta^* \\ &= 2\xi^* w^* + \nu^* s^* + 0 = 0. \end{aligned}$$

Conversely, suppose x^* is a solution to the GSOCCP$(F, G, \mathcal{K})$. The possible cases are $G(x^*) = 0$ and $G(x^*) \neq 0$. They are treated separately:

(a) Assume $G(x^*) = 0$. Let $w^* = 1/2\ F(x^*)^T \bar{A} F(x^*)$, $s^* = F_1(x^*)$, and set all remaining variables to zero. Using the fact that $F(x^*) \in \mathcal{K}^\circ$, we conclude that $w^*, s^* \geq 0$ and $g(x^*, v_g^*, \zeta^*) = 0$. Thus the objective value will be $r^* f(x^*, v_f^*) + (1 - r^*) g(x^*, v_g^*, \zeta^*) = g(x^*, v_g^*, \zeta^*) = 0$, a global minimum, since f and g are nonnegative.

(b) Suppose $G(x^*) \neq 0$ and consider the optimization problem (2), rewritten below for convenience.

$$\begin{aligned} \text{Minimize } & F(x^*)^T u, \\ \text{s.t. } & \frac{1}{2} u^T A u \geq 0, \\ & u_1 \geq 0. \end{aligned} \tag{7}$$

Since $F(x^*) \in \mathcal{K}^\circ$, the objective function is nonnegative for a feasible u. Given that x^* is a solution to the GSOCCP$(F, G, \mathcal{K})$, $u^* = G(x^*)$ satisfies the constraints and is an optimal solution to (7). Notice that $u_1^* > 0$, thus there is at most one active constraint at u^*. If $(u^*)^T A u^* > 0$ then both constraints are superfluous and the gradient of the objective must be zero at u^*, that is, $F(x^*) = 0$. If $(u^*)^T A u^* = 0$, then the gradient of the unique active constraint is $Au^* = (u_1^*, \ldots)^T \neq 0$, forming, thus, a linearly independent set, implying that constraint qualifications hold at u^*. In both cases there exist Lagrange multipliers $\lambda^* \geq 0$ and $\mu^* = 0$ (since $u_1^* > 0$) such that

$$\begin{aligned} F(x^*) - \lambda^* A G(x^*) - \mu^* e_1 &= 0, \\ \frac{\lambda^*}{2} G(x^*)^T A G(x^*) &= 0, \\ G_1(x^*) \mu^* &= 0. \end{aligned}$$

Let $z^* = G(x^*)^T A G(x^*)/2$, $y^* = G_1(x^*)$, $r^* = 1$ and set all remaining variables to zero. Taking into account that $G(x^*) \in \mathcal{K}$, it follows that $z^* \geq 0$, $y^* \geq 0$ and $f(x^*, v_f^*) = 0$. Since the function g assumes nonnegative values only, $r^* f(x^*, v_f^*) + (1 - r^*) g(x^*, v_g^*, \zeta^*) = f(x^*, v_f^*) = 0$ is the global minimum of (6).

Therefore, given a solution to the GSOCCP$(F, G, \mathcal{K})$, we are able to construct an optimal solution to (6) with objective value zero. □

Notice that Theorem 1 is easily generalized for the case where $\mathcal{K}$ is a Cartesian product of cones. In this case ϕ would be replaced by a sum of like terms, one for each cone.

2 A simple instance and an alternative formulation

The reformulation (6) was tested for the following instance of GSOCCP$(F, G, \mathcal{K})$:

$$F(x) = \begin{pmatrix} x_1 + 1 \\ x_2 + 2 \end{pmatrix}, \quad G(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad A = \bar{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8}$$

Figure 1 below gives a graphical representation of this instance. The lighter cone identifies the region $\{x \mid F(x) \in \mathcal{K}^\circ\}$, the region $\{x \mid G(x) \in \mathcal{K}\}$ corresponds to the intermediate shade of gray and their overlapping is indicated by the darkest shade. The circle with center $(-1/2, -1)$ and radius $\sqrt{5}/2$ is the loci of points satisfying the orthogonality condition $F(x)^T G(x) = 0$. Clearly, the unique solution to this instance is the vertex of the darkest cone: $(1/2, -1/2)$.

The code `easy` was used for this and all subsequent nonlinear optimizations. This is a Fortran double-precision code for solving nonlinear programming problems, based

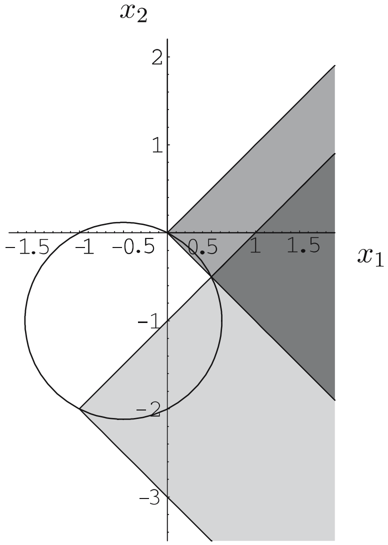


Fig. 1 Simple instance

on augmented Lagrangian [17], trust region [13] and projected gradients combined with a mild active set strategy [6]. It is available at http://www.ime.unicamp.br/∼martinez. The trust-region augmented Lagrangian implemented in `easy` formulates a quadratic model of the augmented Lagrangian, which, in the case of the reformulations considered herein, coincides with the objective function since there are no explicit constraints (only simple bounds), and updates the model and the trust-region at each outer iteration. Roughly speaking, the quadratic model uses a numerical approximation of the Hessian at the current point. This quadratic model is optimized by the quadratic solver, which combines conjugate gradient with a mild active set strategy. The computational effort is expressed by the number of iterations of the trust-region algorithm for simple-bounded minimization, functional evaluations performed, iterations of the inner quadratic solver and matrix-vector products (MVP) computed, denoted by `ITBOX`, `FE`, `ITQUA` and `MVP`, respectively. This choice of code was purely a matter of convenience, it should be stressed that any other code for bound constrained nonlinear optimization could be used.

Table 1 contains the outcomes obtained running `easy` on reformulation (6) of problem (8) with 200 initial points $(x^0, v_f^0, v_g^0, \zeta^0, r^0)$, with x^0 ramdomly generated in the box $\left[-10, 10\right] \times \left[-10, 10\right]$ and remaining variables set to 0.5. The stopping criterion was norm of projected gradient less than 10^{-8}, achieved for 82% of the tests. The remaining 18% stopped with too small a step (less than 10^{-8}), meaning that possibly the end point is close to a local minimizer. The column with header ϕ^* contains the final objective function value of (6).

The three types of end points obtained were I(0.5, −0.5), II(−1.0021, −1.9958), and III(0.0000, −0.0387), with ϕ_{I}^*, ϕ_{II}^* and ϕ_{III}^* of order 10^{-20}, 10^{-6} and 10^{-7}, resp. The unique solution (point I) was obtained in 59% of the tests. The remaining ended in points close to (−1, −2) (17% to point II) or to (0,0) (25% to point III). It is worth mentioning that of the 13% of tests that stopped with a too small step, 4.5% ended at the solution, whereas the remaining 13.5% ended at point III.

Table 1 Results for instance (8)—reformulation (6)

	ITBOX	FE	ITQUA	MVP	ϕ^*
Minimum	7	9	35	56	0
Average	284.2	409.5	2366.4	3193.2	
Maximum	1020	1470	24658	29207	2E−6

Despite the very small objective function values reached at points II and III, they are not feasible for GSOCCP$(F, G, \mathcal{K})$. At these points we have approximately orthogonality and membership in exactly one (but not both) of the cones. In fact, they are far from the feasible set (the darkest cone in Fig. 1).

Consider, for instance, point II obtained in 17% of the tests. We observed that, for this point, the corresponding final value $r^* = 0$ was always achieved, and so the objective value was equal to $g(-1.0021, -1.9958) \approx 10^{-6}$. The second and third terms in the expression (5) of $g(-1.0021, -1.9958)$ enforce the membership of $F(x)$ in the dual cone. Orthogonality is enforced by the terms $(\xi w)^2$ and $(\nu s)^2$. But membership of $G(x)$ in $\mathcal{K}$ is enforced only indirectly by the first term, $\|G(x) - \xi \bar{A}F(x) - M\zeta\|^2$. We can drive this term to zero (this was achieved by lowering the parameter used as stopping criterion in `easy`) and still have $G(x)$ outside $\mathcal{K}$. For point III, since the corresponding value $r^* = 1$ was obtained and $f(0.0000, -0.0387) \approx 10^{-7}$, the previous reasoning holds with g and G replaced by f and F, respectively.

Given that we are dealing with a difficult problem, with possibly many local optima, it is not so bad to end in a point that is not a solution, but it is very bad to end at a point that looks as if it is a solution, because the objective value is very small. After all, algorithms for nonlinear problems will seldom arrive at the exact solution. In this case, solutions with objective values close to zero may seem to be close to the global solution. It would be preferable to have higher objective function values associated with solutions II and III.

This phenomenon motivated the following equivalent reformulation. Given

$$\begin{aligned}\Xi(x, \lambda, z, y, w, s) = \frac{1}{2}\Bigg(&\|\lambda F(x) - (1-\lambda)AG(x)\|^2 + (\lambda w)^2 + ((1-\lambda)z)^2 \\ &+ \left(\frac{1}{2}G(x)^T AG(x) - z\right)^2 + (G_1(x) - y)^2 \\ &+ \left(\frac{1}{2}F(x)^T \bar{A}F(x) - w\right)^2 + (F_1(x) - s)^2 + \|MF(x)\|^2 \Bigg),\end{aligned}$$

define

$$\begin{aligned} &\min \ \Xi(x, \lambda, z, y, w, s), \\ &\text{s.t.} \quad 1 \geq \lambda \geq 0, \\ &\qquad\quad z, y, w, s \geq 0. \end{aligned} \tag{9}$$

By removing the multiplication factors r and $1 - r$, membership in both cones is directly enforced in the objective function. Next we state the equivalence between GSOCCP$(F, G, \mathcal{K})$ and problem (9). Extension to Cartesian product of cones is straightforward.

Theorem 2 *If* $(x^*, \lambda^*, z^*, y^*, w^*, s^*)$ *is a global minimizer of* (9) *with objective value zero then* x^* *is a solution to* $GSOCCP(F, G, \mathcal{K})$. *Conversely, if* x^* *is a solution to* $GSOCCP(F, G, \mathcal{K})$*, then there exist*

$$z^*, y^*, w^*, s^* \geq 0, \quad \text{and} \quad 1 \geq \lambda^* \geq 0$$

such that $(x^*, \lambda^*, z^*, y^*, w^*, s^*)$ *is a global minimizer of* (9) *with objective value zero.*

Proof The proof is analogous to the one of Theorem 1. Only the converse statement merits a few comments. First we recall that in the proof of Theorem 1 we showed that, if $G(x^*) \neq 0$ at a solution x^*, then there exists $\epsilon^* \geq 0$ such that $F(x^*) = \epsilon^* AG(x^*)$. Adapting that reasoning to the case $F(x^*) \neq 0$, it can be shown that there exist $\xi^* \geq 0$ and ζ^* such that $G(x^*) = \xi^* \bar{A} F(x^*) + M\zeta^*$. Multiplying the last equality by A we obtain $AG(x^*) = \xi^* \text{diag}(\mathbf{1}_p, 0, \ldots, 0) F(x^*)$, which is the same as $AG(x^*) = \xi^* F(x^*)$, if $F(x^*) \in \mathcal{K}^\circ$. Thus, at a solution x^* to GSOCCP$(F, G, \mathcal{K})$, the two cases may be combined in the equality $aF(x^*) = bAG(x^*)$, where both numbers a and b are nonnegative, and at least one may be assumed to be nonzero. Dividing both sides by the positive number $a + b$ we obtain $\lambda^* = a/(a + b) \in [0, 1]$ that will give part of the solution to (9). The assignment of the other variables follows the scheme in the proof of Theorem 1. □

Some of the extra variables in reformulation (6) were interpreted as Lagrange multipliers of auxiliary optimization problems (2) and (3). In the interest of brevity, Theorem 2 was proved using an adequate adaptation of the proof of Theorem 1. There is, however, a purely algebraic motivation behind the λ variable in the first term of the second reformulation (9), which could have been used instead in the proof of the converse part of Theorem 2, see the Appendix.

3 Conditions on stationary points

From a practical point of view, once a solution of a reformulation is attained, one can easily check whether its x-part is (close to) a solution to the original problem (1). Now, most algorithms for nonlinear programming, when successful, end in stationary points. Hence it is interesting, from the theoretical point of view, to establish conditions under which a stationary point of the reformulation will have objective function zero, and thus be a global solution thereof, containing a solution of (1). The next theorem tackles this problem, regarding the first reformulation.

Theorem 3 *Let* $(x^*, v_f^*, v_g^*, \zeta^*, r^*)$ *be a stationary point of (6), and define*

$$H_g = \nabla F(x^*)^{-1} \nabla G(x^*) - \xi^* \bar{A} \quad \text{and} \quad H_f = \nabla G(x^*)^{-1} \nabla F(x^*) - \lambda^* A.$$

(a) If $r^* = 1$ *and* H_f *is positive definite then* x^* *is a solution to GSOCCP.*
(b) If $r^* = 0$*, and* H_g *is positive definite then* x^* *is a solution to GSOCCP.*
(c) If $0 < r^* < 1$ *then* $\phi(x^*, v_f^*, v_g^*, \zeta^*, r)$ *is constant for* $0 \leq r \leq 1$ *and*
  (i) if H_g *or* H_f *are positive definite then* x^* *is a solution to GSOCCP.*
  or
  (ii) $(\nabla_x f(x^*, v_f^*), 0, \ldots, 0)$ *is a descent direction for* ϕ *from* $(x^*, v_f^*, v_g^*, \zeta^*, 1)$ *and* $(\nabla_x g(x^*, v_g^*, \zeta^*), 0, \ldots, 0)$ *is a descent direction for* ϕ *from* $(x^*, v_f^*, v_g^*, \zeta^*, 0)$.

Proof Let

$$\begin{aligned}\ell_1 &= F(x^*) - \lambda^* AG(x^*) - \mu^* e_1,\\ \ell_2 &= \tfrac{1}{2}G(x^*)^T AG(x^*) - z^*,\\ \ell_3 &= G_1(x^*) - y^*,\\ \ell_4 &= G(x^*) - \xi^* \bar{A}F(x^*) - \nu^* e_1 - M\zeta^*,\\ \ell_5 &= \tfrac{1}{2}F(x^*)^T \bar{A}F(x^*) - w^*,\\ \ell_6 &= F_1(x^*) - s^*,\\ \ell_7 &= MF(x^*).\end{aligned}$$

The first order necessary optimality conditions (KKT) of (6) are

$$f(x^*, v_f^*) - g(x^*, v_g^*, \zeta^*) - \gamma_1 + \gamma_2 = 0, \tag{10}$$

$$\gamma_1 r = 0, \ \ \gamma_2(1-r) = 0, \ \ \gamma_1 \geq 0, \ \ \gamma_2 \geq 0, \ \ 0 \leq r \leq 1, \tag{11}$$

$$r\nabla_x f(x^*, v_f^*) + (1-r)\nabla_x g(x^*, v_g^*, \zeta^*) = 0, \tag{12}$$

$$-\ell_1^T AG(x^*) + (\lambda^* z^*)z^* - \theta_1 = 0, \quad \lambda^*\theta_1 = 0, \quad \lambda^* \geq 0, \quad \theta_1 \geq 0, \tag{13}$$

$$-\ell_1^T e_1 + (\mu^* y^*)y^* - \theta_3 = 0, \quad \mu^*\theta_3 = 0, \quad \mu^* \geq 0, \quad \theta_3 \geq 0, \tag{14}$$

$$-\ell_2 + (\lambda^* z^*)\lambda^* - \theta_2 = 0, \quad z^*\theta_2 = 0, \quad z^* \geq 0, \quad \theta_2 \geq 0, \tag{15}$$

$$-\ell_3 + (\mu^* y^*)\mu^* - \theta_4 = 0, \quad y^*\theta_4 = 0, \quad y^* \geq 0, \quad \theta_4 \geq 0, \tag{16}$$

$$-\ell_4^T \bar{A}F(x^*) + (\xi^* w^*)w^* - \pi_1 = 0, \quad \xi^*\pi_1 = 0, \quad \xi^* \geq 0, \quad \pi_1 \geq 0, \tag{17}$$

$$-\ell_5 + (\xi^* w)\xi^* - \pi_2 = 0, \quad w^*\pi_2 = 0, \quad w^* \geq 0, \quad \pi_2 \geq 0, \tag{18}$$

$$-\ell_4^T e_1 + (\nu^* s^*)s^* - \pi_3 = 0, \quad \nu^*\pi_3 = 0, \quad \nu^* \geq 0, \quad \pi_3 \geq 0, \tag{19}$$

$$-\ell_6 + (\nu^* s^*)\nu^* - \pi_4 = 0, \quad s^*\pi_4 = 0, \quad s^* \geq 0, \quad \pi_4 \geq 0, \tag{20}$$

$$-M\ell_4 = 0. \tag{21}$$

(a) If $r^* = 1$, by (12),

$$\begin{aligned}\frac{1}{2}\nabla G(x^*)^{-1}\nabla_x f(x^*, v_f^*) &= \left[\nabla G(x^*)^{-1}\nabla F(x^*) - \lambda^* A\right]\ell_1 + \ell_2 AG(x^*) + \ell_3 e_1\\ &= H_f \ell_1 + \left[AG(x^*)\right]\ell_2 + \ell_3 e_1 = 0.\end{aligned} \tag{22}$$

From (13) and (15) we get

$$\ell_2\left[AG(x^*)\right]^T \ell_1 = (\lambda^* z^*)^3 + \theta_1\theta_2, \tag{23}$$

and (14) and (16) imply

$$\ell_3 e_1^T \ell_1 = (\mu^* y^*)^3 + \theta_3\theta_4. \tag{24}$$

Premultiplying (22) by ℓ_1 we obtain

$$0 = \ell_1^T H_f^T \ell_1 + \ell_1^T \left[AG(x^*)\right] \ell_2 + \ell_1^T B \ell_3 = \ell_1^T H_f^T \ell_1 \; + \underbrace{(\lambda^* z^*)^3 + \theta_1 \theta_2}_{\geq 0} + \underbrace{((\mu^*)^T y^*)^3 + \theta_3 \theta_4}_{\geq 0}. \tag{25}$$

By (25) and the assumption that H_f is positive definite,

$$\ell_1 = 0, \quad \lambda^* z^* = 0, \quad (\mu^*)^T y^* = 0. \tag{26}$$

Therefore, using (22) and (26), $G_1(x^*)\ell_2 + \ell_3 = 0$ and $G_i(x^*)\ell_2 = 0$, for $i = 2, \ldots, p$. If, for some $k = 2, \ldots, p$, $G_k(x^*) \neq 0$, then $\ell_2 = 0$, and necessarily also $\ell_3 = 0$, implying that $\phi(x^*, v_f^*, v_g^*, \zeta^*, r^*) = 0$. If $G_i(x^*) = 0$ for $i = 2, \ldots, p$, then

(i) If $z^* > 0$, (26) and (15) imply $\theta_2 = 0$, $\ell_2 = -\theta_2 = 0$ and, like before, $\ell_3 = 0$ and $\phi(x^*, v_f^*, v_g^*, \zeta^*, r^*) = 0$.

(ii) If $z^* = 0$, by (15) and the definition of ℓ_2 we obtain

$$0 \geq -\theta_2 = \ell_2 = \frac{1}{2} G_1(x^*)^2 \geq 0 \quad \Longrightarrow \quad \ell_2 = 0, \tag{27}$$

and again $\ell_3 = 0$ and $\phi(x^*, v_f^*, v_g^*, \zeta^*, r^*) = 0$.

(b) If $r^* = 0$ by (11)

$$\frac{1}{2} \nabla F(x^*)^{-1} \nabla_x g(x^*, v_g^*, \zeta^*) = \left[\nabla F(x^*)^{-1} \nabla G(x^*) - \xi^* \bar{A}\right] \ell_4 + \ell_5 \bar{A} F(x^*) + \ell_6 e_1 + M \ell_7$$
$$= H_g \ell_4 + \ell_5 \bar{A} F(x^*) + \ell_6 e_1 + M \ell_7 = 0. \tag{28}$$

By (17)–(20)

$$\ell_5 \left[\bar{A} F(x^*)\right]^T \ell_4 = (\xi^* w^*)^3 + \pi_1 \pi_2, \tag{29}$$

$$\ell_6 e_1^T \ell_4 = (\nu^* s^*)^3 + \pi_3^T \pi_4. \tag{30}$$

Premultiplying (28) by ℓ_4 and using (29), (30) and (21)

$$\ell_4^T \left[H_g\right]^T \ell_4 + \underbrace{(\xi^* w^*)^3 + \pi_1 \pi_2 + (\nu^* s^*)^3 + \pi_3^T \pi_4}_{\geq 0} = 0. \tag{31}$$

Therefore,

$$\ell_4 = 0, \quad \xi^* w^* = 0 \quad \text{and} \quad (\nu^*)^T s^* = 0. \tag{32}$$

Now, by (28) we get

$$\ell_5 \bar{A} F(x^*) + \ell_6 e_1 + M \ell_7 = 0. \tag{33}$$

The following equalities are a consequence of (28) and (33):

$$\begin{aligned} F_1(x^*)\ell_5 + \ell_6 &= 0, \\ -F_i(x^*)\ell_5 &= 0, \quad i = 2, \ldots, p, \\ F_i(x^*) &= 0, \quad i = p+1, \ldots, n. \end{aligned} \tag{34}$$

These equalities then imply that $\ell_7 = 0$.

If, for some $k = 2,\dots,p$, $F_k(x^*) \neq 0$, then $\ell_5 = 0$, and it follows that $\ell_6 = 0$ and $\phi(x^*, v_f^*, v_g^*,$
$\zeta^*, r^*) = 0$.
If $F_i(x^*) = 0$ for $i = 2,\dots,p$, then
(i) If $w^* > 0$, by (32) and (19) we have that $\pi_2 = 0$, $\ell_5 = -\pi_2 = 0$.
By (34), $\ell_6 = 0$, and it follows that $\phi(x^*, v_f^*, v_g^*, \zeta^*, r^*) = 0$.
(ii) If $w^* = 0$, by the definition of ℓ_6 and (19), we have

$$0 \geq -\pi_2 = \ell_5 = \frac{1}{2}F_1(x^*)^2 \geq 0 \quad \Longrightarrow \quad \ell_5 = 0. \tag{35}$$

Again by (34), $\ell_6 = 0$ and $\phi(x^*, v_f^*, v_g^*, \zeta^*, r^*) = 0$.

(c) If $0 < r^* < 1$ by (10) and (11) then $f(x^*, v_f^*) = g(x^*, v_g^*, \zeta^*)$. By (12)

$$\nabla_x f(x^*, v_f^*) = \frac{(r-1)}{r}\nabla_x g(x^*, v_g^*, \zeta^*). \tag{36}$$

If $\nabla_x f(x^*, v_f^*) \neq 0$, taking $r^* = 0$ or $r^* = 1$, we have that $(\nabla_x f(x^*, v_f^*), 0, \dots, 0)$ and $(\nabla_x g(x^*, v_g^*, \zeta^*),$
$0, \dots, 0)$ are descent directions for ϕ at $(x^*, v_f^*, v_g^*, \zeta^*, 1)$ and $(x^*, v_f^*, v_g^*, \zeta^*, 0)$, respectively. If $\nabla_x f(x^*, v_f^*) = 0$ and H_g or H_f are positive definite the proof is just as in items (a) or (b).

□

While the description adopted for cone $\mathcal{K} = \{x \in \mathbb{R}^n \mid x^T Ax/2 \geq 0, x_1 \geq 0\}$ has the virtue of smoothness, it is nevertheless a nonconvex way of describing a convex set. On the other hand, reformulation (6) is suggested by KKT conditions on auxiliary optimization problems like (7), whose constraint set contains precisely the offending nonconvex constraint $u^T Au/2 \geq 0$. The conditions imposed in Theorem 3 are direct consequences of these choices.

4 Numerical experiments

All numerical experiments were carried out by solving the appropriate nonlinear optimization problem in `easy`. Unfortunately, though not surprisingly, we could not find in the literature classes of actual problems that require the general framework considered here. Not one instance with F and G both nonlinear was encountered and degenerate cones are not heard of. We did test our approach with the few available examples, as detailed below.

In Sects. 4.1, 4.2 and 4.3, for the sake of comparison, we tested our reformulation against the merit function Ψ_{BF} formulated in [8]. In this case the corresponding nonlinear optimization problem is unconstrained. Two hundred trials of all examples were run, with same initial setup (original variables randomly set in the interval $[-10, 10]$, others set to 0.5). The notation employed in Table 1 is maintained.

Experiments comparing reformulations (6) (objective function ϕ) and (9) (objective function Ξ) are reported in Sects. 4.1 and 4.2. The second reformulation proved more robust, which led us to adopt it in the remaining tests.

In Sect. 4.3 we describe a larger example (27 original x variables), with a more realistic flavor, coming from a problem in robotics. We compare the more successful

Table 2 Results for instance (8)—objective Ξ

	ITBOX	FE	ITQUA	MVP	Ξ^*
Minimum	6	7	26	34	1E−32
Average	11.3	15.4	70.2	102.8	
Maximum	18	27	161	225	1E−17

reformulation (9) using merit function Ξ with Chen and Tseng's reformulation using Ψ_{BF}. In this case the latter's proved less efficient.

In Sect. 4.4 we tackle larger problems, randomly generated, and another nonlinear small (five variables originally) one. The problems are described in [16], where they are solved by an algorithm specifically tailored for the class of second-order cone complementarity problems (SOCCP).

4.1 Affine bidimensional case

The first instance of GSOCCP($F, G, \mathcal{K}$) tested was (8). Results are given in previous Table 1 and in Table 2 below. Runs using reformulation (9) converged to point I in 100% of the tests, and all the tests stopped with norm of the projected gradient less than 10^{-8} for this formulation. The results using function Ψ_{BF} are reported in Table 3. Convergence to point I occurred in all trials using Ψ_{BF} as well.

An additional comment deserves to be made concerning reformulation (6) and the sufficient condition provided by Theorem 3. The Jacobians of functions F and G given in (8) are the identity matrix. Analyzing matrices H_f and H_g at the end points, one can see that, for $r^* = 1$, H_f is never positive definite, whereas for $r^* = 0$, H_g is positive definite for 83 runs out of the 200 initial points. Therefore, the sufficient conditions of Theorem 3 hold for 42% of these tests. It is worth mentioning that there were 34 runs (17%) for which convergence to point I was obtained and $r^* = 0$, without positive definiteness of matrix H_g.

4.2 Affine functions of Peng and Yuan

The second instance of GSOCCP($F, G, \mathcal{K}$) tested was taken from [19, Problem 3]:

$$F(x) = \begin{pmatrix} 15x_1 - 5x_2 - x_3 + 4x_4 - 5x_5 \\ 5x_2 + x_5 \\ -x_1 - 3x_2 + 8x_3 + 2x_4 - 3x_5 \\ 2x_1 - 4x_2 + 2x_3 + 9x_4 - 4x_5 \\ -5x_2 + 10x_5 - 1 \end{pmatrix}, \quad G(x) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \tag{37}$$

with $A = \bar{A} = \mathrm{diag}\,(1, -1, -1, -1, -1)$. Three distinct end points were obtained, namely

$$\begin{aligned} \text{(unique solution) I } &(0.049185, -0.0030997, 0.0096024, 0.0031883, 0.048033), \\ \text{II } &(0.048919, -0.0031088, 0.0096054, 0.0032299, 0.048093), \\ \text{and III } &(0.020424, -0.016125, 0.022867, 0.021211, 0.0879190). \end{aligned}$$

Tables 4 and 5 summarize the results of 200 runs. For reformulation (6), 96% of tests ended with norm of projected gradient less than 10^{-8}, whereas 4% stopped with too small a step (infinity norm smaller than 10^{-8}). In terms of quality of results, 93.5% of

Table 3 Results for instance (8)—objective Ψ_{BF}

	ITBOX	FE	ITQUA	MVP	Ψ_{BF}
Minimum	5	6	6	12	0
Average	7.1	8.1	8.0	15.1	
Maximum	8	9	11	19	5E−17

Table 4 Results for instance (37)—reformulation (6)

	ITBOX	FE	ITQUA	MVP	ϕ^*
Minimum	31	44	238	483	1E−25
Average	220.3	319.8	2591.4	3504.3	
Maximum	348	538	5645	7358	1E−8

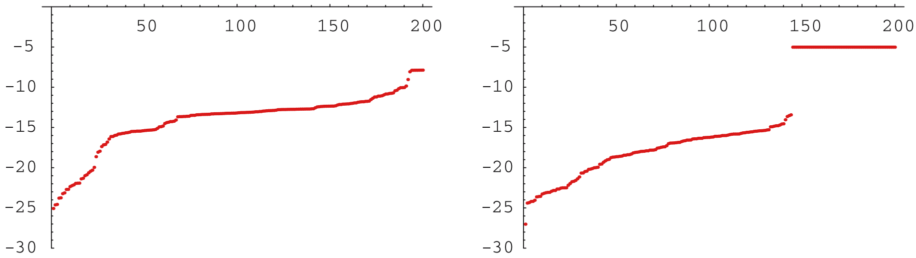


Fig. 2 Test problem $\times \log_{10}(\phi^*)$ (left) and test problem $\times \log_{10}(\Xi^*)$ (right) for the 200 tests of instance (37)

tests converged to the unique solution, with $\phi^*_{\rm I} < 10^{-11}$, although only 56.5% reached $\phi^*_{\rm I} < 10^{-13}$, the same threshold obtained with reformulation (9). Point II was reached for 2.5% of tests ($\phi^*_{\rm II} \approx 10^{-10}$) and point III for the remaining 4%, with $\phi^*_{\rm III} \approx 10^{-8}$.

As far as reformulation (9) is concerned, 100% of tests stopped with norm of projected gradient less than 10^{-8}, 72% ended at point I ($\Xi^*_{\rm I} < 10^{-13}$) and 28% at point III ($\Xi^*_{\rm III} > 10^{-7}$). Sorting the 200 final objective function values in ascending order and ploting the resulting vector produced the graphs in Fig. 2, where it can be seen that reformulation (9) produces a better discrimination between the solutions obtained.

In terms of success rates, reformulation (9) was much superior to reformulation (6) in instance (8) (100% as opposed to 59%), but not so in instance (37) (72–93.5%). Nevertheless, its reasonable success rate allied to its power to discriminate the desired stationary points suggest a more robust character, further evidenced in the experiments carried out in [2]. Thus, reformulation (6) will not be considered in the remaining tests.

All trials with merit function Ψ_{BF} converged to Point I. Table 6 contains the relevant data. The only drawback was the increased coding effort for the objective function and gradient, since these functions are defined by parts.

For the two small examples (originally in two and five variables) described in this and the previous section, better results were obtained using the merit function Ψ_{BF}. This is not the case for the next problem.

Table 5 Results for instance (37)—reformulation (9)

	ITBOX	FE	ITQUA	MVP	Ξ^*
Minimum	21	32	189	244	1E−27
Average	69.6	96.6	667.3	907.5	
Maximum	262	389	2823	3904	1E−5

Table 6 Results for instance (37)—objective Ψ_{BF}

	ITBOX	FE	ITQUA	MVP	Ψ^*_{BF}
Minimum	12	13	30	42	1.57E−22
Average	18.3	21.3	49.6	77.4	
Maximum	32	46	94	196	4E−17

4.3 Grasping force optimization

Grasp analysis is pivotal to the study of robotic systems with multi-fingered hands. Han et al. [15] formulate optimization problems modeling several aspects of grasp analysis (force closure, force feasibility and force optimization). The consideration of nonlinear friction models leads to the introduction of second-order cone constraints in these problems. Previously these were dealt with by means of linearization, which simplified matters algorithmically speaking, but at a cost. If the linearization turns out to be too loose a relaxation of the true nonlinear constraint, the solution obtained may violate the latter. Once this problem is detected, the usual remedy is to refine the approximation, increasing, perhaps prohibitively, the computational effort. The concentrated research in SOCP of the last decade opened various alternative ways of foregoing the linearization, taking the original nonlinear constraints directly into account. In the grasp force optimization arena, we may cite the positive definiteness formulation pioneered by Buss et al. [7] and the linear matrix inequality (LMI) framework adopted by Han et al. [15]. While the former work included the development of specific programs for solving the problems, the latter was able to apply existing interior point software for a special class of convex optimization problems with LMI constraints.

The grasp force optimization problem considered herein concerns a 3-D object grasped by a mechanical hand, with several contact points between its fingers and the object. The objective is to minimize some function of the contact forces subject to restrictions which represent equilibrium conditions (external forces should be balanced), admissibility (properties of the mechanism should be taken into account), bounds on joint efforts and friction constraints. When the latter are modeled as second-order cone constraints, the optimization problem falls into the SOCP category. Typically the constraints contain linear equalities and inequalities, as well as cone constraints, that is, the contact force vector must satisfy linear constraints and lie in the Cartesian product of cones (uni-, three- or four-dimensional ones, depending on whether the contact is, respectively, frictionless, point contact with friction, or a soft finger contact with elliptic approximation). The objective function embodies the criterion used to select the "most desirable" force vector, amongst the feasible ones. Thus it may consist of a simple min/max criterion (one will find the "gentlest grasp" able to effectively hold the object), minimum weighted norm criterion, the maxdet criterion of [15] (which induces robust choices of grasp forces with respect to the friction cone constraints), etc.

Since the numerical experiments have so far indicated that the equivalent reformulation (9) presented in Sect. 1 was more promising, we elected to solve the grasping force problem only with this variant. Furthermore, the six equality constraints of our example were taken into account by means of allowing for a degenerate cone, which afforded greater efficiency, both in the problem formulation as well as in the numerical performance. For the reformulation using merit function Ψ_{BF}, however, these equations correspond to yet another cone, since degenerate cones are not allowed.

The grasping force optimization problem falls into the following framework:

$$\begin{aligned} &\min f^T x,\\ &\text{s.t. } \|A_i x + b_i\| \le c_i^T x + d_i, \quad i = 1, \dots, \ell, \end{aligned} \tag{38}$$

where f and x belong to $\mathbb{R}^n$, $A_i \in \mathbb{R}^{(n_i-1)\times n}$, $b_i \in \mathbb{R}^{n_i-1}$, $c_i \in \mathbb{R}^n$ and $d_i \in \mathbb{R}$. Notice that linear equality and inequality constraints can be put in the format considered above, by letting the right-hand-side or the left-hand-side be zero, respectively. Its dual problem is

$$\begin{aligned} &\min_{u,w} \ \textstyle\sum_{i=1}^{\ell}(b_i^T u_i + d_i w_i),\\ &\text{s.t.} \quad \textstyle\sum_{i=1}^{\ell}(A_i^T u_i + c_i w_i) = f,\\ &\qquad \|u_i\| \le w_i, u_i \in \mathbb{R}^{(n_i-1)}, \ w_i \in \mathbb{R}, \ i = 1, \dots, \ell. \end{aligned} \tag{39}$$

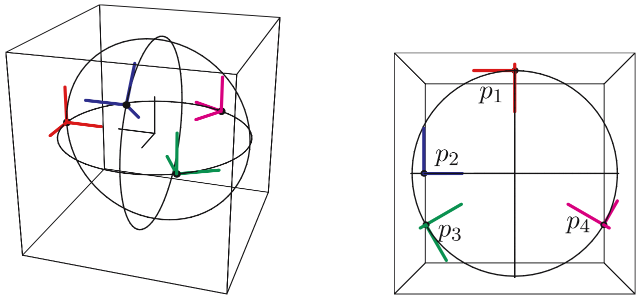
Letting $\mathcal{K}^{n_i} = \{(u_i, w_i) \in \mathbb{R}^{n_i-1} \times \mathbb{R} \mid w_i \ge \|u_i\|\}$ and $\varpi = (u_1, w_1, \dots, u_\ell, w_\ell)$, the second set of constraints in (39) becomes $\varpi \in \mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_\ell}$ and the dual problem can be stated as

$$\begin{aligned} &\min g^T \varpi,\\ &\text{s.t. } B\varpi = f,\\ &\qquad \varpi \in \mathcal{K}, \end{aligned} \tag{40}$$

where $g \in R^L, B \in R^{n\times L}, f \in \mathbb{R}^n, L = n_1 + \cdots + n_\ell$ and $\operatorname{rank}(B) = n$. Let $\hat{\varpi}$ be such that $B\hat{\varpi} = f$ and $Z \in \mathbb{R}^{L\times(L-n)}$ such that $\operatorname{Range}(Z) = \operatorname{Kernel}(B)$. Define $F(\xi) = \widehat{\varpi} + Z\xi_1$ and $G(\xi) = g + B^T\xi_2$, where $\xi = (\xi_1, \xi_2) \in \mathbb{R}^{L-n} \times \mathbb{R}^n$. Then duality theory [1] implies that sufficient optimality conditions for the primal-dual pair (38)–(40) constitute the GSOCCP$(F, G, \mathcal{K})$. In fact, the latter complementarity problem generalizes linear programming's complementary slackness conditions.

Gentlest grasp force. The grasping force problem's setup consists of a unit sphere that is in contact with four mechanical fingers, as shown in Fig. 3, where the sphere's center coincides with the reference coordinate system's origin. It is adapted from [15]. A point contact with friction was assumed at all fingertips, with common friction coefficient $\mu = 0.4$.

The problem is to find contact forces $x_i \in \mathbb{R}^3$, for $i = 1, \dots, 4$, expressed in local contact frames, such that the maximum of their normal components is minimized, while obeying Coulomb's friction law, maintaining static equilibrium and satisfying upper and lower bounds:



Contact points coordinates

$p_1^T = (1, 0, 0)$ $\quad$ $p_2^T = (0, \cos(\pi/5), \sin(\pi/5))$

$p_3^T = (\cos(2\pi/3), \sin(2\pi/3), 0)$ $\quad$ $p_3^T = (\cos(4\pi/3), \sin(4\pi/3), 0)$

Fig. 3 Setup of grasping force problem

$$
\begin{aligned}
&\min_{x,t}\ t \\
&\text{s.t.} \quad (x_i)_3 \le t, i = 1, \ldots, 4 && t \ge \text{maximum of normal components} \\
&\qquad \|((x_i)_1, (x_i)_2)\| \le \mu (x_i)_3, i = 1, \ldots, 4 && \text{friction cone constraints} \qquad (41) \\
&\qquad Gx + h_{\text{ext}} = 0 && \text{static equilibrium} \\
&\qquad -10 \le (x_i)_j \le 10, i = 1, \ldots, 4, j = 1, 2, 3, && \text{bounds on forces,}
\end{aligned}
$$

where $G \in \mathbb{R}^{6\times 12}$ is the grasp map, that transforms applied finger forces expressed in local contact frames to resultant object wrenches. The resultant generalized contact force Gx must balance the external load $h_{\text{ext}}^T = (2.1, -0.2, -4.3, 0.4, -1.5, 0.6)$ experienced by the object. Cone constraints allow us to remove most of the upper/lower bounds, since they imply $(x_i)_3 \ge 0$ and, on the other hand, $(x_i)_3 \le 10$ plus the cone constraints imply $|(x_i)_{1,2}| \le \mu 10 < 10$.

The dual problem of (41) has 26 variables: $\varpi = (w_1, \ldots, w_4, u_{11}, u_{12}, w_5, u_{21}, u_{22}, w_6, u_{31}, u_{32}, w_7, u_{41}, u_{42}, u_{43}, u_{44}, u_{45}, u_{46}, u_{47}, u_{48}, w_8, w_9, w_{10}, w_{11}, w_{12})$. The first four variables in ϖ are the dual variables associated with the inequalities involving t in (41), and belong to $\mathcal{K}^1$ (the set of nonnegative reals). Then we have (u_1, w_5), (u_2, w_6) and $(u_3, w_7) \in \mathcal{K}^3$, variables associated with the first three friction constraints. The cone containing (u_4, w_8) is degenerate—$\{(u_4, w_8) \mid w_8 \ge \|(u_{41}, u_{42})\|\}$—and encompasses the last friction constraint as well as the equilibrium equality constraints. The last four variables are associated with the (remaining) upper bound constraints. Adapting reformulation (9) to this case, we end up with a box constrained problem on 62 variables. Numerical tests used same choice of initialization as previous ones. Success, that is, objective function values less than or equal 10^{-15}, was achieved in 194 out of 200 trials. Table 7 summarizes the performance and the plot of Fig. 4 shows the base 10 logarithm of the final objective function values in ascending order. The six values corresponding to the trials that did not converge ($\Xi^* = 1.28$) to the global solution

[1] We follow here the convention adopted in [18], where the last variable, instead of the first, is greater than or equal to the norm of the remaining ones.

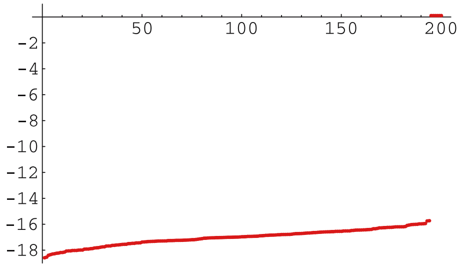


Fig. 4 Test problem $\times \log_{10}(\Xi^*)$ (grasp problem)

Table 7 Grasp problem—objective Ξ

	ITBOX	FE	ITQUA	MVP	Ξ^*
Minimum	20	28	769	989	2.58E−19
Average	36.6	50.3	2069.3	2462.6	
Maximum	81	107	5950	6445	1.28

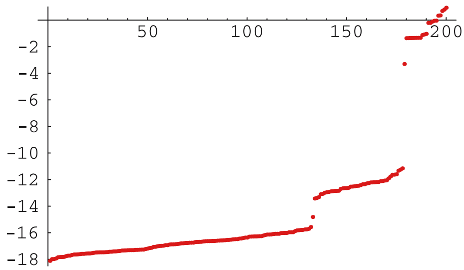


Fig. 5 Test problem $\times \log_{10}(\Psi^*_{\text{BF}})$ (grasp problem)

are almost imperceptible, right above the horizontal axis. It should be noted, however, that even in these cases `easy` ended in a stationary point, that is, the stopping criterion was norm of projected gradient smaller than preset tolerance (10^{-8}).

The optimization problem using Ψ_{BF} has 27 variables. There were 22 failures out of 200 trials, almost a fourfold increase comparing to the reformulation using Ξ. The failures correspond to the dots in the upper right corner of the plot in Fig. 5. Furthermore, it should be noted that `easy` stopped unfavorably in 68 trials, that is, it stopped due to lack of progress and not because the norm of the projected gradient was small. The acknowledged successful trials correspond to the first 132 points in the plot, forming an almost horizontal curve, before the first jump. Since `easy` was designed for box-constrained optimization, large artificial boxes were defined in the coding of this unconstrained problem. The failures are not related to these artificial boundaries. In all cases the boundaries were never touched during the execution of the algorithm. In the absence of active bounds, the quadratic solver reduces to a truncated Newton method.

Concerning the comparison between the two reformulations, embodied by the functions Ξ and Ψ_{BF}, we point out that though there were fewer outer iterations (ITBOX) and function evaluations (FE), see Tables 7 and 8, the computational effort of the inner quadratic solver (ITQUA) and on the MVP was higher when optimizing Ψ_{BF}. It is also interesting to notice the much larger range of variability in the statistics of Table 8. The ratios max/min for the four markers on effort were 4.05, 3.82, 7.74, 6.52

Table 8 Grasp problem—objective Ψ_{BF}

	ITBOX	FE	ITQUA	MVP	Ψ^*_{BF}
Minimum	9	19	325	436	7.65E−19
Average	25.7	42.3	3327.7	3461.4	
Maximum	116	155	19620	20229	8.5636

for Table 7 and 12.90, 8.16, 60.40, 46.40 for Table 8, suggesting that the first approach was more robust for this particular example.

The ratio ITQUA/ITBOX is an estimate of the number of iterations of the quadratic solver. Theoretically, this inner algorithm should converge in at most n iterations, where n is the dimension of the quadratic problem. However, it is well known that the number of iterations may be much larger for ill conditioned quadratics. Calculating these ratios for the average figures in Tables 7 and 8 we obtain 56.5 and 129.5, respectively. The first figure compares favorably with the dimension of the problem (62), whereas the second is more than 10 times the corresponding dimension (27). This may well be related to the fact that the function Ψ_{BF} does not have continuous second derivatives.

Comparing the graphs in Figs. 4 and 5, we see that the first reformulation provides a better discrimination of the results. Finally, we should mention that our intention in considering the grasping problem was just to include a test problem with physical appeal, in addition to the randomly generated or toy problems presented in the other sections.

4.4 Test problems of Hayashi et al.

In their recent paper on SOCCP, Hayashi et al. [16] test their algorithm on classes of monotone problems. We employ our approach in two of them. The first is a set of problems randomly generated according to the recipe given in [16] and the second is a nonlinear SOCCP. In these problems $G(x)$ is the identity function and $F(x)$ is affine ($= Mx + q$) in the first set and nonlinear in the second. The cone $\mathcal{K}$ is the Lorentz cone in $\mathbb{R}^n$ and the product of two Lorentz cones in $\mathbb{R}^3 \times \mathbb{R}^2$, respectively. Hayashi et al.'s optimality threshold is adopted, that is, a successful outcome is obtained when the optimal objective function value Ξ^* is smaller than 10^{-16}.

The matrix M and vector q are generated so that the problem is feasible, M is rank-deficient positive semidefinite and there exists $\bar{x} \in \operatorname{int}\mathcal{K}$ such that $M\bar{x} + q \in \operatorname{int}\mathcal{K}$. Tables 9 and 10 summarize the results of the two subsets of tests involving these problems. In the first subset, n was kept constant and equal to 100, and the rank of M varies from 10 to 99. Each instance was solved 100 times with different initial points, randomly generated. The rate of success doesn't seem to depend on the rank, varying from 87% to 95%. In the second subset the dimension n assumes values 100, …, 1,000, and the rank of M is randomly chosen in the interval $[0.9n, n-1]$. For each value of n we solved 10 instances, using 10 randomly generated initial points for each of them. We observed higher rates of success in the smaller dimensions. The discrimination property was maintained, the number of successes is practically invariant if the threshold is increased to 10^{-5}, so there is a clear distinction between stationary points that correspond to global solutions and other endpoints. The nonlinear reformulation (9) has $n + 5$ variables. The convenience of randomly generated problems quickly

Table 9 Results for linear SOCCP's with various degrees of rank deficiency

Rank	10	20	30	40	50	60	70	80	90	99
% success	87	93	94	95	95	91	90	94	93	89

Table 10 Results for linear SOCCP's with various problem sizes

Dimension	100	200	300	400	500	600	700	800	900	1000
% success	89	74	68	56	56	53	40	45	37	31

leads to a wealth of data, awkward to report and not necessarily very enlightening. Thus the more concise character of Tables 9 and 10, which could easily expand on to many more tables, were we to provide the same level of detail as in the tables of previous sections.

The nonlinear function used in the second set of tests is given by

$$F(x) = \begin{pmatrix} 24(2x_1 - x_2)^3 + \exp(x_1 - x_3) - 4x_4 + x_5 \\ -12(2x_1 - x_2)^3 + 3(3x_2 + 5x_3)/\sqrt{1 + (3x_2 + 5x_3)^2} - 6x_4 - 7x_5 \\ -\exp(x_1 - x_3) + 5(3x_2 + 5x_3)/\sqrt{1 + (3x_2 + 5x_3)^2} - 3x_4 + 5x_5 \\ 4x_1 + 6x_2 + 3x_3 - 1 \\ -x_1 + 7x_2 - 5x_3 + 2 \end{pmatrix}$$

The algorithm converged either to $x^* = (0.23240, -0.073079, 0.22061, 0.53390, -0.53390)^T$ (the true solution), in 76% of the 200 trials, or to the stationary point $x = (0.16415, -0.073443, 0.26353, 0.53517, -0.25708)^T$, with optimal objective function value of the order of 10^{-4}, in the remaining 24%. The version of (9) corresponding to this problem has 15 variables, the original five, and five extra ones for each cone.

5 Conclusions

We proved that the GSOCCP can be reformulated as a box-constrained minimization problem, preserving the smoothness of the original data. Furthermore, the implementation of the merit function is a straightforward task, with complexity closely related to that of the functions of the original problem. We obtained sufficient conditions under which a stationary point of the first reformulated problem is a global solution and thus provides a solution of the GSOCCP.

The second reformulation introduced shares the first one's format (box-constrained minimization), but not the theoretical results concerning stationary points, see [2]. Nevertheless, its performance in the numerical experiments suggests a very desirable feature: the elimination of "false positives," those points that have very small objective function values and satisfy all kinds of stringent stopping criteria but yet are far from actual solutions. We observed a better discrimination between true and false solutions.

The merit function Ψ_{BF} of Chen and Tseng's was also implemented and compared with ours, using the same nonlinear code, the software `easy`. This reformulation was more successful in the small problems (two and five variables) but less so in the larger

(27 variables) one tested. The poor performance of `easy` in the optimization of $\Psi_{\rm BF}$ in the larger problems of Sect. 4.3 indicates the inherent difficulty of this problem. Despite the nice theoretical result concerning the optimality of its stationary points, the topology of the objective function $\Psi_{\rm BF}$ proves to be a tall order for the code. It seems that the algorithm is unable to travel far along the "narrow valleys" of the objective function, zigzagging instead along its walls, finally stopping due to lack of progress. This is a general feature of reformulations of complementarity and related problems. We do not mean to suggest that there is a particular reformulation dramatically more efficient than the others. The point is to stress the possibility of using reformulations as alternatives to solving the GSOCCP directly, and the appeal of our approach is its straightforward use. Indeed, the implementation of $\Psi_{\rm BF}$ (and that of its gradient) consists more of a challenge.

Higher dimensional tests were done using the recipe for randomly generated SOCCPs in [16]. The problems considered had $G(x) = x$ and $F(x) = Mx + q$. The order of the symmetric positive semidefinite M varied from 100 to 1,000 and it was constructed so that its rank was deficient. The rate of success doesn't seem to depend on the rank but does deteriorate with the increase in dimension.

Apparently the reformulations considered herein result in "difficult" objective functions, in the sense that they seem to exhibit narrow valleys, as mentioned above, known obstacles for optimization algorithms. Nevertheless, the lack of good theoretical results for the stationary points of Ξ was somehow compensated by the good behavior of its second derivative. This may be very relevant from a practical perspective, as evidenced in the numerical tests.

It is to be expected that algorithms tailored to the task should present a better performance at solving SOCCPs than a general purpose nonlinear programming code applied to a reformulation. It must be stressed, however, that our reformulations encompass more general classes of problems (general nonlinear function G and degenerated cone $\mathcal{K}$), are easy to implement, do not require the tuning of additional parameters, preserve the smoothness of the original data and provide a good discrimination between stationary points that are global solutions and the rest.

Acknowledgements We thank J.M. Martínez for the fruitful discussions and helpful suggestions. We also thank the anonymous referees for their constructive criticisms.

Appendix

The proof of the converse part of Theorem 2 depends on the existence of an appropriately valued λ, obtained below in a purely algebraic manner, without recourse to auxiliary optimization problems as done in the text.

Theorem 4 *If $y \in \mathcal{K}$, $w \in \mathcal{K}^\circ$ and $y^T w = 0$ then there exists $\lambda \in [0, 1]$ such that*

$$\lambda w = (1 - \lambda) A y. \tag{42}$$

Furthermore, if w (resp., y) is in the relative interior of $\mathcal{K}^\circ$ (resp., $\mathcal{K}$) then $\lambda = 0$ (resp., $\lambda = 1$).

Proof We denote by $y_{i:j}$ the subvector of y containing components $y_i, \ldots, y_j$. With this notation, $\mathcal{K} = \{x \in \mathbb{R}^n \mid x_1^2 \geq \|x_{2:p}\|^2, x_1 \geq 0\}$ and $\mathcal{K}^\circ = \{x \in \mathbb{R}^n \mid x_1^2 \geq \|x_{2:p}\|^2, x_1 \geq 0, x_{p+1:n} = 0\}$.

First assume that w belongs to the relative interior of $\mathcal{K}^\circ$, that is, $w^T A w > 0, w_1 > 0$, $Mw = 0$. Thus $w_1^2 > \|w_{2:p}\|^2 \geq 0$ and $w^T y = 0$ implies

$$y_1 = -\sum_{i=2}^{p} \frac{w_i}{w_1} y_i = -\frac{1}{w_1} \langle w_{2:p}, y_{2:p} \rangle. \tag{43}$$

Therefore, since $y \in \mathcal{K}$, we have that

$$\frac{1}{w_1^2} \langle w_{2:p}, y_{2:p} \rangle^2 \geq \|y_{2:p}\|^2. \tag{44}$$

But

$$\frac{1}{w_1^2} \langle w_{2:p}, y_{2:p} \rangle^2 = \frac{1}{w_1^2} \|w_{2:p}\|^2 \|y_{2:p}\|^2 \cos^2 \theta. \tag{45}$$

Substituting (45) in (44), we obtain

$$\frac{1}{w_1^2} \|w_{2:p}\|^2 \|y_{2:p}\|^2 \cos^2 \theta \geq \|y_{2:p}\|^2. \tag{46}$$

Now if $w_{2:p} = 0$ then $y_{2:p} = 0$, and (43) implies $y_{1:p} = 0$. If $w_{2:p} \neq 0$ and $y_{2:p} \neq 0$, from (46) we obtain

$$\cos^2 \theta \geq \frac{w_1^2}{\|w_{2:p}\|^2}, \tag{47}$$

an impossibility, since w in the relative interior of $\mathcal{K}^\circ$ implies that the latter fraction is greater than 1. Thus we conclude that $y_{1:p} = 0$ and (42) is true with $\lambda = 0$.

Now assume that w does not belong to the relative interior of $\mathcal{K}^\circ$. If $w = 0$, then (42) is true with $\lambda = 1$. Consider the case $w \neq 0$. The facts that $w \neq 0$ and is not in the relative interior imply $w_1^2 = \|w_{2:p}\|^2$, $w_1 > 0$ and $w_{2:p} \neq 0$. Thus, orthogonality between w and y implies (43), as before. If $y_{2:p} = 0$, then $y_1 = 0$ and (42) is true with $\lambda = 0$. On the other hand, if $y_{2:p} \neq 0$, we may obtain (47), as above. This time, this implies $\cos\theta = \pm 1$. Now

$$0 \leq y_1 = -\frac{1}{w_1} \|w_{2:p}\| \; \|y_{2:p}\| \cos\theta \tag{48}$$

from which follows that $\cos\theta = -1$. Therefore there exists $\alpha > 0$ such that

$$y_{2:p} = -\alpha w_{2:p} \tag{49}$$

and

$$y_1 = -\frac{1}{w_1} (-\alpha) \|w_{2:p}\|^2 = \alpha w_1. \tag{50}$$

Rewriting (49) and (50) we obtain

$$w = \frac{1}{\alpha} A y. \tag{51}$$

Letting $\lambda = (1 + 1/\alpha)^{-1}$ we arrive at expression (42).

The case where y is in the relative interior of $\mathcal{K}$ is similar to the first case considered, exchanging the roles of w and y. We have $y^T Ay > 0$, $y_1 > 0$ and orthogonality implying $w_1 = -(1/y_1)\langle w_{2:p}, y_{2:p}\rangle$, from which we may conclude that $w_{1:p} = 0$. Since $Mw = 0$, relation (42) is satisfied with $\lambda = 1$. □

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. **95**, 3–51 (2003)
2. Andreani, R., Friedlander, A., Mello, M.P., Santos, S.A.: Generalized second-order complementarity problems: theory and numerical experiments. Technical Report, IMECC, Unicamp, January (2005) Available at http://www.ime.unicamp.br/ andreani
3. Andreani, R., Friedlander, A., Santos, S.A.: On the resolution of generalized nonlinear complementarity problems. SIAM J. Optimiz. **12**, 303–321 (2001)
4. Andreani, R., Martínez, J.M., Schuverdt, M.L.: On the relation between constant positive linear dependence condition and quasinormality constraint qualification. J. Optimiz. Theory Appl. **125**, 473–485 (2005)
5. Bertsekas, D.P.: Nonlinear Programming, 2nd edn. Athena Scientific, Belmont, Massachusetts (1999)
6. Bielchowsky, R.H., Friedlander, A., Gomes, F.A.M., Martínez, J.M., Raydan, M.: An adaptive algorithm for bound constrained quadratic minimization. Invest. Operat. **7**, 67–102 (1997)
7. Buss, M., Hashimoto, H., Moore, J.B.: Dextrous hand grasping force optimization. IEEE Trans. Robot. Automat. **12**, 406–418 (1996)
8. Chen, J.-S., Tseng, P.: An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. Technical Report, July (2004)
9. Chen, J.-S.: A new merit function and its related properties for the second-order cone complementarity problem. Technical Report, October (2004).
10. Facchinei, F., Soares, J.: A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optimiz. **7**, 225–247 (1997)
11. Ferris, M.C., Pang, J.-S.: Engineering and economic applications for nonlinear complementarity problems. SIAM Rev. **39**, 669–713 (1997)
12. Fischer, A.: A special Newton-type optimization methods. Optimization **24**, 269–284 (1992)
13. Friedlander, A., Martínez, J.M., Santos, S.A.: A new trust region algorithm for bound constrained minimization. Appl. Math. Optimiz. **30**, 235–266 (1994)
14. Geiger, C., Kanzow, C.: On the resolution of monotone complementarity problems. Comput. Optimiz. Appl. **5**, 155–173 (1996)
15. Han, L., Trinkle, J.C., Li, Z.X.: Grasp analysis as linear matrix inequality problems. IEEE Trans. Robot. Automat. **16**(06), 663–674 (2000)
16. Hayashi, S., Yamashita, N., Fukushima, M.: A combined smoothing and regularization method for monotone second-order cone complementarity Problems. SIAM J. Optimiz. **15**, 593–615 (2005)
17. Krejić, N., Martínez, J.M., Mello, M.P., Pilotta, E.A.: Validation of an augmented Lagrangian algorithm with a Gauss–Newton Hessian approximation using a set of hard-spheres problems. Comput. Optimiz. Appl. **16**, 247–263 (2000)
18. Lobo, M., Vandenberghe, L., Boyd, S., Lebret, H.: Applications of second-order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
19. Peng , J.-M., Yuan, Y.-X.: Unconstrained methods for generalized complementarity problems. J. Comput. Math. **15**, 253–264 (1997)
20. Yamada, K., Yamashita, N., Fukushima, M.: A new derivative-free descent method for the nonlinear complementarity problem. In: Pillo, G.D., Giannessi, F. (eds.) Nonlinear Optimization and Related Topics, pp. 463–489. Kluwer Academic Publishers, Netherlands (2000)