DOI: 10.1137/S1052623400377591

SIAM J. OPTIM.
Vol. 12, No. 2, pp. 303–321

ON THE RESOLUTION OF THE GENERALIZED NONLINEAR COMPLEMENTARITY PROBLEM*

ROBERTO ANDREANI†, ANA FRIEDLANDER‡, AND SANDRA A. SANTOS‡

Abstract. Minimization of a differentiable function subject to box constraints is proposed as a strategy to solve the generalized nonlinear complementarity problem (GNCP) defined on a polyhedral cone. It is not necessary to calculate projections that complicate and sometimes even disable the implementation of algorithms for solving these kinds of problems. Theoretical results that relate stationary points of the function that is minimized to the solutions of the GNCP are presented. Perturbations of the GNCP are also considered, and results are obtained related to the resolution of GNCPs with very general assumptions on the data. These theoretical results show that local methods for box-constrained optimization applied to the associated problem are efficient tools for solving the GNCP. Numerical experiments are presented that encourage the use of this approach.

Key words. box-constrained optimization, complementarity

AMS subject classifications. 65H10, 90C33, 90C30

PII. S1052623400377591

1. Introduction.

The generalized nonlinear complementarity problem (GNCP) is to find $x \in \mathbb{R}^m$ such that

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

where F and G are continuous functions from $\mathbb{R}^m$ to $\mathbb{R}^n$, $\mathcal{K}$ is a nonempty closed convex cone in $\mathbb{R}^n$, and $\mathcal{K}^\circ$ denotes the polar cone of $\mathcal{K}$.

We consider the case $n = m$, $F, G \in C^1$, and $\mathcal{K}$ a polyhedral cone in R^n; that is, given $A \in \mathbb{R}^{q\times n}$ and $B \in \mathbb{R}^{s\times n}$, we have

$$\mathcal{K} = \{v \in \mathbb{R}^n \,|\, Av \geq 0,\ Bv = 0\}$$

and

$$\mathcal{K}^\circ = \{u \in \mathbb{R}^n \,|\, u = A^T\lambda_1 + B^T\lambda_2,\ \lambda_1 \geq 0\}.$$

This problem has many interesting applications, and its solution using special techniques has been considered extensively in the literature. See [16, 17, 24] among others. If $\mathcal{K} = \mathbb{R}^m_+ \equiv \{x \in \mathbb{R}^m \,|\, x \geq 0\}$, $G(x) = x - F(x)$, and $F : \mathbb{R}^m \to \mathbb{R}^m$, the GNCP$(F, G, \mathcal{K})$ reduces to the so-called *implicit complementarity problem* [20, 21]. In particular, if $G(x) = x$, the GNCP reduces to the *nonlinear complementarity problem*, denoted by NCP.

Our approach in this paper is to formulate the GNCP as an equivalent bound-constrained smooth optimization problem. Differentiable bound-constrained minimization is a well-developed area of practical optimization, and many methods and reliable software are available for large-scale problems. See, for example, [7, 8, 12, 26].

*Received by the editors September 5, 2000; accepted for publication (in revised form) May 11, 2001; published electronically November 7, 2001.
http://www.siam.org/journals/siopt/12-2/37759.html

†Department of Computer Science and Statistics, University of the State of São Paulo (UNESP), CP 136, CEP 15054-000, São José do Rio Preto SP, Brazil (andreani@nimitz.dcce.ibilce.unesp.br). This author was supported by FAPESP and CNPq.

‡Department of Applied Mathematics, IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas SP, Brazil (friedlan@ime.unicamp.br, sandra@ime.unicamp.br). The second and third authors were supported by FAPESP, CNPq, FINEP, and FAEP-UNICAMP.

This motivated the authors to find equivalences between variational and complementarity problems and smooth box-constrained minimization problems (see [1, 13, 14, 15]).

We prove here that the GNCP is equivalent to a bound-constrained optimization problem in the sense that a global minimizer with zero objective function value is a solution of the GNCP. We also establish conditions for proving that stationary points of the minimization problems are global minimizers and, consequently, solutions of the GNCP. The GNCP (or GCP in other references) is a problem related to the *variational inequality problem* (VIP). The VIP and other related problems were reformulated by many authors as different minimization problems and systems of equations. See [18, 22, 24, 25]. The reformulations of related problems as bound-constrained problems in [1, 14, 15] that use the same approach as the one presented here cannot be extended to obtain a merit function with the properties of the reformulation proposed in this paper. As pointed out by one of the referees, the GNCP can be reformulated as a *mixed complementarity problem* (MCP). In [4], Andreani and Martínez prove results for the MCP based on their work on the bounded VIP [5]. These results applied to the GNCP lead to sufficient conditions on the functions F and G stronger than the ones obtained in this paper. The sufficient conditions given in this paper on the functions F and G that guarantee that stationary points of the merit function solve the GNCP cannot be obtained from any of the previous results.

The objective functions of the minimization problems have a very simple structure that consists of a sum of terms that are polynomials in the original problem data plus an additional term of the type $(x^T z)^p$, with $p > 1$. This term plays a fundamental role in the proof of the equivalence results, and $p = 2$ is especially interesting for linear programming and linear complementarity problems, because in these cases the objective function to be minimized is just a polynomial of fourth degree. It is important to remark that no penalty parameters are needed in these problem formulations, which we call the *quartic approach*. In [1, 13, 14, 15] some very simple counterexamples show that when $p = 1$ the existence of stationary points that are not global minimizers is possible. For the complementarity problem, [1, Theorem 2.4] shows that if F' is positive definite, the merit function with $p = 1$ is such that its stationary points are solutions of the original problem.

These merit functions preserve all the derivatives of the functions that define the GNCP. Consequently, the global and local convergence properties depend on the algorithm used for box-constrained minimization. This is a very important feature, since it makes viable the use of algorithms that need high-order derivatives or their approximations, such as the tensor methods of [23]. Any efficient algorithm for smooth box-constrained minimization can be used, in particular, algorithms that do not rest upon matrix factorizations at all, allowing us to deal with large-scale problems.

Complementarity and related problems have also been solved using algorithms based on the projection equation. See [10] and references therein. These methods are very efficient; however, their behavior is strongly dependent on the monotonocity of the function that defines the problem. Failure of this condition results in divergence of the sequences generated by these algorithms. Unlike the formulations in [22, 25], the computation of the objective function of the equivalent minimization problem considered here is straightforward, and projections on convex sets are not necessary to compute either the objective function or the derivatives. Therefore, special algorithms for dealing with nonsmoothness do not need to be devised. In [24], to obtain the fundamental equivalence result for a cone that is not necessarily polyhedral, the

authors assume the same conditions on the problem as we do here. However, even for polyhedral cones, the implementation of the algorithm proposed there requires projections that, in general, are very expensive to compute.

Using the same merit function of [17], a stronger result is obtained in [16] where the GNCP is reformulated as a system of semismooth equations, and an unconstrained differentiable formulation is given if $\mathcal{K}$ is the positive orthant. The conditions established to ensure that a stationary point x_* of the unconstrained minimization problem is a solution of the GNCP are essentially that the Jacobian of F at x_* (denoted by $F'(x_*)$) is invertible and that $D(G'(x_*)[F'(x_*)]^{-1})_{RR}D$ is an S_0-matrix, where D is a convenient nonsingular diagonal matrix and R is the set of indices for which (1) does not hold at x_*. ($B \in \mathbb{R}^{n\times n}$ is an S_0-*matrix* if there exists $v \in \mathbb{R}^n$ such that $v \geq 0$, $v \neq 0$, and $Bv \geq 0$.) A trust-region method is proposed in [16] for solving the GNCP based on these reformulations. This algorithm was implemented by the authors and tested for some problems.

In [17] an unconstrained minimization reformulation of the GNCP is considered such that the merit function is differentiable when $\mathcal{K} = \mathbb{R}^n_+$. The sufficient conditions for a stationary point x_* of the merit function to be a global minimizer are that $F'(x_*)$ is nonsingular and the product $G'(x_*)[F'(x_*)]^{-1}$ is a P_0-matrix. ($B \in \mathbb{R}^{n\times n}$ is a P_0-*matrix* if its principal minors are all nonnegative.) The authors suggest the use of a first-order method for minimizing the merit function due to the fact that it is once but not twice continuously differentiable.

The case of a general cone $\mathcal{K}$ was considered in [24], using an unconstrained reformulation for the GNCP. It is proved there that x_* is a solution of the GNCP if $F'(x_*)$ is nonsingular and $G'(x_*)[F'(x_*)]^{-1}$ is positive definite. The evaluation of the corresponding objective function is rather complicated and requires projections that in general are not easy to compute.

Here we require, essentially, the same conditions as in [24] to guarantee that stationary points of the minimizing problems are solutions of the GNCP. These assumptions cannot be relaxed for a general cone $\mathcal{K}$ as we show with an example in section 3. If $\mathcal{K} = \mathbb{R}^n_+$, we require a weaker condition on matrix $G'(x_*)F'(x_*)^{-1}$. If F and G are affine functions with $\mathcal{K}$ polyhedral, the conditions are that $G'F'^{-1}$ is positive semidefinite in the null space of B and the GNCP is feasible. Finally, an even weaker condition is needed if F and G are affine and $\mathcal{K} = \mathbb{R}^n_+$.

If $\mathcal{K}$ is a general cone and it is not possible to ensure that $G'F'^{-1}$ is positive definite at a stationary point of the merit function, a sequence of perturbed problems can be constructed for which the strict monotonicity property holds and such that the sequence of solutions of these perturbed problems converges to a solution of the original one. The results related to this construction are valid for a general cone and may be applied also to the results in [24].

The paper is organized as follows: In section 2 we associate with (1) a box-constrained minimization problem, and we prove that assuming a local strict monotonicity condition, stationary points of this problem are solutions of (1). In section 3 we consider perturbations of the original problem that allow us to deal with monotone (not necessarily strict) functions. Numerical experiments are presented in section 4. Finally, conclusions and lines for future research are discussed in section 5.

Notation. We denote by $\langle\cdot,\cdot\rangle$ the Euclidean inner product on $\mathbb{R}^n$ and by $\|\cdot\|$ the norm induced by this inner product and its corresponding matricial norm. If B is a real $n \times n$ matrix, $B \geq 0$ ($B > 0$) means that B is positive semidefinite (positive definite).

2. Equivalence results. The following minimization problem with simple bounds is associated with the GNCP$(F, G, \mathcal{K})$ defined in (1):

$$\begin{aligned} \min \quad & f(x, z, \lambda) \\ \text{subject to} \quad & \begin{cases} z^1 \geq 0, \\ \lambda^1 \geq 0, \end{cases} \end{aligned} \tag{2}$$

where

$$f(x, z, \lambda) = \|RF(x) - z\|^2 + \|G(x) - R^T\lambda\|^2 + \rho\langle z^1, \lambda^1\rangle^2$$

and

$$R = \begin{pmatrix} A \\ B \end{pmatrix}, \quad z = \begin{pmatrix} z^1 \\ 0 \end{pmatrix} \in \mathbb{R}^q \times \mathbb{R}^s, \quad \lambda = \begin{pmatrix} \lambda^1 \\ \lambda^2 \end{pmatrix} \in \mathbb{R}^q \times \mathbb{R}^s.$$

The next theorem states that solving problem GNCP$(F, G, \mathcal{K})$ is equivalent to finding the global minimizer of the optimization problem (2).

THEOREM 1. *If (x_*, z_*, λ_*) is a global minimizer of problem* (2) *with $f(x_*, z_*, \lambda_*) = 0$, then x_* is a solution of the* GNCP$(F, G, \mathcal{K})$. *Conversely, if x_* is a solution of the* GNCP$(F, G, \mathcal{K})$, *then there exist z_*, λ_* such that (x_*, z_*, λ_*) is a global minimizer of* (2) *with $f(x_*, z_*, \lambda_*) = 0$.*

Proof. If $f(x_*, z_*, \lambda_*) = 0$, then

$$\begin{gathered} AF(x_*) = z^1 \geq 0, \quad BF(x_*) = 0, \quad \text{implying that} \quad F(x_*) \in \mathcal{K}, \\ G(x_*) = A^T\lambda_*^1 + B^T\lambda_*^2, \quad \text{with} \quad \lambda_*^1 \geq 0, \quad \text{so} \quad G(x_*) \in \mathcal{K}^\circ, \end{gathered}$$

and

$$\langle F(x_*), G(x_*)\rangle = \langle F(x_*), R^T\lambda_*\rangle = \langle z_*, \lambda_*\rangle = \langle z_*^1, \lambda_*^1\rangle = 0.$$

Conversely, if x_* is a solution of the GNCP$(F, G, \mathcal{K})$ then there exists $\lambda_* = (\lambda_*^1, \lambda_*^2)$ with $\lambda_*^1 \geq 0$ such that $G(x_*) = A^T\lambda_*^1 + B^T\lambda_*^2$, $z_*^1 = AF(x_*) \geq 0$, and $0 = F(x_*)^T G(x_*) = F(x_*)^T(A^T\lambda_*^1 + B^T\lambda_*^2) = (z_*^1)^T\lambda_*^1 + (BF(x_*))^T\lambda_*^2 = (z_*^1)^T\lambda_*^1$.

Therefore, calling $z_* = (z_*^1, 0)^T$, we have that $f(x_*, z_*, \lambda_*) = 0$. ∎

Global minimizers are very hard to find, especially in large-scale problems. Most efficient large-scale algorithms for box-constrained optimization are guaranteed to converge only to stationary points of the problem. Therefore, it is desirable to relate stationary points of (2) to solutions of the GNCP.

THEOREM 2. *Let $F(x), G(x) \in C^1$. If (x_*, z_*, λ_*) is a stationary point of* (2) *and $G'(x_*)[F'(x_*)]^{-1}$ is positive definite in the null space of B, then x_* is a solution of the* GNCP$(F, G, \mathcal{K})$.

Proof. Let

$$\begin{aligned} H_* &= G'(x_*)[F'(x_*)]^{-1}, \\ w_* &= AF(x_*) - z_*^1, \\ u_* &= BF(x_*), \\ v_* &= G(x_*) - R^T\lambda_*, \\ \theta_* &= \langle z_*^1, \lambda_*^1\rangle. \end{aligned}$$

If (x_*, z_*, λ_*) is a stationary point of (2), then there exist $\mu \in \mathbb{R}^p_+$ and $\gamma \in \mathbb{R}^s_+$ such that

$$2G'(x_*)^T v_* + 2F'(x_*)^T(A^T w_* + B^T u_*) = 0, \tag{3}$$

$$-2Av_* + 2\rho\theta_* z^1_* - \mu = 0, \tag{4}$$

$$Bv_* = 0, \tag{5}$$

$$-2w_* + 2\rho\theta_*\lambda^1_* - \gamma = 0, \tag{6}$$

$$\langle \mu, \lambda^1_* \rangle = 0, \quad \langle \gamma, z^1_* \rangle = 0, \tag{7}$$

$$\lambda^1_* \geq 0, \quad \mu \geq 0, \quad \gamma \geq 0, \quad z^1_* \geq 0. \tag{8}$$

By (3) we have

$$H^T_* v_* + A^T w_* + B^T u_* = 0. \tag{9}$$

Now, by (4), (6), and (7), we obtain

$$4\langle Av_*, w_* \rangle = 4\rho^2\theta^2_* + \langle \mu, \gamma \rangle, \tag{10}$$

and (5), (9), and (10) imply that

$$\langle v_*, H^T_* v_* \rangle + \langle Av_*, w_* \rangle = \langle v_*, H^T_* v_* \rangle + \rho^2\theta^3_* + \frac{\langle \mu, \gamma \rangle}{4} = 0. \tag{11}$$

Therefore, by (5) and the fact that $\langle v_*, H^T_* v_* \rangle > 0$ in the null space of B, (11) implies

$$\theta_* = 0, \quad \langle v_*, H^T_* v_* \rangle = 0. \tag{12}$$

Since H^T_* is positive definite in the null space of B, by (12), necessarily,

$$v_* = 0. \tag{13}$$

Thus, by (12) and (6),

$$2w_* = -\gamma. \tag{14}$$

If a_i denotes the ith row of matrix A, using (13) and replacing w_* and v_* in (9), we get

$$A^T w_* + B^T u_* = \sum_{i=1}^{q} a_i(\langle a_i, F(x_*) \rangle - (z^1_*)_i) + B^T BF(x_*) = 0. \tag{15}$$

Let

$$\mathcal{I} = \{i \in \{1, \ldots, q\} \,|\, (z^1_*)_i = 0\};$$

then, if $i \notin \mathcal{I}$, we have that $(z^1_*)_i > 0$. But, by (7), we also have $\gamma_i = 0$. So, by (14),

$$(w_*)_i = \langle a_i, F(x_*) \rangle - (z^1_*)_i = 0 \quad \forall\, i \notin \mathcal{I}. \tag{16}$$

Now, by (15) and (16)

$$A^T w_* + B^T u_* = \sum_{i \in \mathcal{I}} a_i \langle a_i, F(x_*) \rangle + B^T BF(x_*) = 0. \tag{17}$$

Premultiplying (17) by $F(x_*)^T$, we obtain

$$\sum_{i\in\mathcal{I}} \langle a_i, F(x_*)\rangle^2 + \|BF(x_*)\|^2 = 0, \tag{18}$$

and by (18)

$$u_* = BF(x_*) = 0, \quad (w_*)_i = \langle a_i, F(x_*)\rangle = 0 \quad \forall\, i \in \mathcal{I}. \tag{19}$$

Finally, (12), (13), (16), and (19) imply that $f(x_*, z_*, \lambda_*) = 0$. ∎

In the following theorem we show that the hypothesis of Theorem 2 can be relaxed if the functions F and G are affine.

THEOREM 3. *Let $F(x), G(x)$ be affine, $G'F'^{-1}$ positive semidefinite in the null space of B and* GNCP$(F, G, \mathcal{K})$ *feasible. If (x_*, z_*, λ_*) is a stationary point of* (2), *then x_* is a solution of* GNCP$(F, G, \mathcal{K})$.

Proof. As in Theorem 2, we obtain (3)–(12). Since $\theta_* = 0$, the optimality conditions read as

$$2G'^T v_* + 2F'^T(A^T w_* + B^T u_*) = 0, \tag{20}$$

$$-2Av_* - \mu = 0, \tag{21}$$

$$Bv_* = 0, \tag{22}$$

$$-2w_* - \gamma = 0, \tag{23}$$

$$\langle \mu, \lambda_*^1\rangle = 0, \quad \langle \gamma, z_*^1\rangle = 0, \tag{24}$$

$$\lambda_*^1 \geq 0, \quad \mu \geq 0, \quad \gamma \geq 0, \quad z_*^1 \geq 0. \tag{25}$$

Relations (20)–(25) are the necessary and sufficient conditions for a global minimizer of the following convex quadratic minimization problem:

$$\begin{aligned} &\min \quad f(x, z, \lambda) = \|RF(x) - z\|^2 + \|G(x) - R^T\lambda\|^2 \\ &\text{subject to} \quad \begin{cases} z^1 \geq 0, \\ \lambda^1 \geq 0. \end{cases} \end{aligned} \tag{26}$$

Since, by hypothesis, the GNCP$(F, G, \mathcal{K})$ is feasible, it turns out that (x_*, z_*, λ_*) is a global solution of (26) with objective function value zero, and as $\theta_* = 0$, we get $f(x_*, z_*, \lambda_*) = 0$. ∎

The hypotheses of Theorem 2 can also be weakened if $\mathcal{K} = \mathbb{R}^n_+$, as we show in the following theorem.

DEFINITION 1. *A matrix $B \in \mathbb{R}^{n\times n}$ is* column-sufficient *if for $v \in \mathbb{R}^n$, $v_i(Bv)_i \leq 0\ \forall\, i$ implies $v_i(Bv)_i = 0\ \forall\, i$. A matrix B is called* row-sufficient *if B^T is column-sufficient.*

DEFINITION 2. *A matrix $B \in \mathbb{R}^{n\times n}$ is called an S-*matrix *if there exists $v \in \mathbb{R}^n$ such that $v \geq 0$ and $Bv > 0$.*

THEOREM 4. *Let $\mathcal{K} = \mathbb{R}^n_+$ and $F(x), G(x) \in C^1$. If (x_*, z_*, λ_*) is a stationary point of* (2) *and $G'(x_*)[F'(x_*)]^{-1}$ is a row-sufficient S-matrix, then x_* is a solution of the GNCP$(F, G, \mathcal{K})$.*

Proof. In this case the optimization problem is

$$\begin{aligned} &\min \quad \|F(x) - z\|^2 + \|G(x) - \lambda\|^2 + \rho\langle z, \lambda\rangle^2 \\ &\text{subject to} \quad \begin{cases} z \geq 0, \\ \lambda \geq 0. \end{cases} \end{aligned} \tag{27}$$

Defining

$$\begin{aligned} w_* &= F(x_*) - z_*, \\ v_* &= G(x_*) - \lambda_*, \\ H_* &= G'(x_*)[F'(x_*)]^{-1}, \end{aligned}$$

and

$$\theta_* = z_*^T \lambda_*,$$

the optimality conditions read as

$$2G'(x_*)^T v_* + 2F'(x_*)^T w_* = 0, \tag{28}$$

$$-2v_* + 2\rho\theta_* z_* - \mu = 0, \tag{29}$$

$$-2w_* + 2\rho\theta_* \lambda_* - \gamma = 0, \tag{30}$$

$$\langle \mu, \lambda_* \rangle = 0, \quad \langle \gamma, z_* \rangle = 0, \tag{31}$$

$$\lambda_* \geq 0, \quad \mu \geq 0, \quad \gamma \geq 0, \quad z_* \geq 0. \tag{32}$$

By (29) and (30),

$$4(w_*)_i(v_*)_i = 4\rho^2\theta_*^2(\lambda_*)_i(z_*)_i + \mu_i\gamma_i \tag{33}$$

for $i \in \{1, \ldots, n\}$. We can write (28) as

$$H_*^T v_* + w_* = 0. \tag{34}$$

Therefore, by (33) and (34),

$$4(v_*)_i(H_*^T v_*)_i + 4\rho^2\theta_*^2(\lambda_*)_i(z_*)_i + \mu_i\gamma_i = 0 \tag{35}$$

for $i \in \{1, \ldots, n\}$. Since H_* is row-sufficient, (35) implies that

$$(v_*)_i(H_*^T v_*)i = 0 \quad \text{for } i \in \{1, \ldots, n\} \quad \text{and} \quad \theta_* = 0. \tag{36}$$

Using (29), (30), (33), (34), and (36), we have that

$$H_*^T v_* = -w_* = \frac{\gamma}{2} \geq 0 \tag{37}$$

and

$$v_* = -\frac{\mu}{2} \leq 0. \tag{38}$$

If $v_* \neq 0$, (37) and (38) contradict the fact that H_* is an S-matrix (see [11]), and therefore

$$v_* = 0, \quad w_* = 0. \tag{39}$$

Finally, by (36) and (39), $f(x_*, z_*, \lambda_*) = 0$. ∎

If F and G are affine functions and $\mathcal{K}$ is the positive orthant, then the following result holds.

THEOREM 5. *Let $\mathcal{K} = \mathbb{R}^n_+$, $F(x)$, $G(x)$ be affine such that $G'F'^{-1}$ is a row-sufficient matrix. If* GNCP$(F, G, \mathcal{K})$ *is feasible and (x_*, z_*, λ_*) is a stationary point of* (2), *then x_* is a solution of* GNCP$(F, G, \mathcal{K})$.

Proof. As in Theorem 4, we obtain (28) and (36). The rest of the proof mimics that of Theorem 3. ∎

Remark. The results of Theorems 4 and 5 are valid with the following hypothesis: There exists a partition of $I = \{1, \ldots, n\}$, $I = [I_0, I_1]$, where

$$\widetilde{F}^T = (F^T_{i\in I_0}, G^T_{i\in I_1}) \quad \text{and} \quad \widetilde{G}^T = (G^T_{i\in I_0}, F^T_{i\in I_1})$$

such that $\widetilde{G}'(x_*)[\widetilde{F}'(x_*)]^{-1}$ is a row-sufficient S-matrix or just row-sufficient if F and G are affine.

3. Perturbed problems. The finite *variational inequality problem* $\mathrm{VIP}(\widehat{F}, \Omega)$, where $\widehat{F} : \mathbb{R}^n \to \mathbb{R}^n$ and $\Omega \subseteq \mathbb{R}^n$ is a closed convex set, is to find $x \in \Omega$ such that $\langle \widehat{F}(x), y - x\rangle \geq 0 \ \forall\, y \in \Omega$.

In [1], for $\Omega = \{x \in \mathbb{R}^n \,|\, g(x) \leq 0, Bx = c, x \geq 0\}$, where $g = (g_1, \ldots, g_m)^T$, $g_i \in C^1(\mathbb{R}^n)$ is convex $\forall\, i = 1, \ldots, m$, $B \in \mathbb{R}^{q\times n}$, and $c \in \mathbb{R}^q$, the authors reformulated the $\mathrm{VIP}(\widehat{F}, \Omega)$ as an equivalent box-constrained smooth optimization problem. The properties of the merit function proposed there are similar to the one considered in section 2 of this paper for the GNCP.

We relate now the $\mathrm{GNCP}(F, G, \mathcal{K})$ with the $\mathrm{VIP}(G\circ F^{-1}, \mathcal{K})$ whenever F^{-1} exists.

LEMMA 6. *If F^{-1} exists, then x_* is a solution of the* $\mathrm{GNCP}(F, G, \mathcal{K})$ *if and only if $F(x_*)$ is a solution of the* $\mathrm{VIP}(G\circ F^{-1}, \mathcal{K})$.

Proof. If x_* is a solution of $\mathrm{GNCP}(F, G, \mathcal{K})$, then

$$F(x_*) \in \mathcal{K}, \quad G(x_*) \in \mathcal{K}^\circ, \quad \langle F(x_*), G(x_*)\rangle = 0. \tag{40}$$

Since F^{-1} exists,

$$\langle G(x_*), F(x_*)\rangle = \langle G \circ F^{-1}(F(x_*)), F(x_*)\rangle = 0 \tag{41}$$

and, as $G(x_*) \in \mathcal{K}^\circ$,

$$\langle G(x_*), y\rangle \geq 0 \quad \forall\, y \in \mathcal{K}. \tag{42}$$

By (40)–(42), $F(x_*) \in \mathcal{K}$ and

$$\langle G \circ F^{-1}(F(x_*)), y - F(x_*)\rangle \geq 0 \quad \forall\, y \in \mathcal{K}. \tag{43}$$

This implies that $F(x_*)$ is a solution of $\mathrm{VIP}(G\circ F^{-1}, \mathcal{K})$.

Conversely, if $F(x_*)$ is a solution of $\mathrm{VIP}(G\circ F^{-1}, \mathcal{K})$, then

$$F(x_*) \in \mathcal{K}. \tag{44}$$

So, for $0 \leq \varepsilon \leq 1$,

$$(1 + \varepsilon)F(x_*) \in \mathcal{K} \quad \text{and} \quad (1 - \varepsilon)F(x_*) \in \mathcal{K} \tag{45}$$

and, since (43) holds for any $y \in \mathcal{K}$, we obtain

$$\langle G \circ F^{-1}(F(x_*)), F(x_*)\rangle = \langle G(x_*), F(x_*)\rangle = 0. \tag{46}$$

By (43) and (46), $\langle G(x_*), y\rangle \geq 0 \ \forall\, y \in \mathcal{K}$, so $G(x_*) \in \mathcal{K}^\circ$. Then, by (44) and (46), x_* is a solution of $\mathrm{GNCP}(F, G, \mathcal{K})$. ∎

If F and G are affine functions we can guarantee that, if $G \circ F^{-1}$ is (not necessarily strictly) monotone, stationary points of the merit function are solutions of the GNCP.

In general, we can have stationary points of the associated problem that are not solutions of the original problem. Consider, for instance, the following example.

Example. Let $F : \mathbb{R} \to \mathbb{R}$ be defined as $F(x) = x$, $G(x) = -1$ if $x \leq 1$ and $G(x) = (x-1)^2 - 1$ if $x \geq 1$, and $\mathcal{K} = \mathbb{R}_+$. Observe that $G \circ F^{-1}$ is monotone and convex. The GNCP$(F, G, \mathcal{K})$ has the unique solution $x_* = 2$. The merit function is given in this case by

$$f(x, v, \lambda) = \begin{cases} (-1-v)^2 + (x-\lambda)^2 + (\lambda v)^2 & \text{if } x \leq 1, \\ ((x-1)^2 - 1 - v)^2 + (x-\lambda)^2 + (\lambda v)^2 & \text{if } x \geq 1, \end{cases}$$

and $(0, 0, 0)^T$ is a stationary point of reformulation (2) that corresponds to this problem.

In [1, Theorem 3.2] the authors proved that if $\widehat{F}$ is (not necessarily strictly) monotone, the sequence of solutions of the perturbed problems $\widehat{F} + \varepsilon_k I$, where I is the identity matrix, converges to the unique solution of minimum norm of the VIP$(\widehat{F}, \mathcal{K})$.

In a similar way, given a sequence of strictly positive ε_k such that $\varepsilon_k \downarrow 0$, we can associate with the GNCP$(F, G, \mathcal{K})$ a family of perturbed problems, as follows. For all $k \in \mathbb{N}$ and $x \in \mathbb{R}^n$ we define

$$G_k(x) = G(x) + \varepsilon_k F(x).$$

In the following theorems we relate the solutions of the perturbed problems to the solution of GNCP$(F, G, \mathcal{K})$, where $\mathcal{K}$ is not necessarily a polyhedral cone. Thus, these results may be used with the formulation proposed in [24].

THEOREM 7. *If* GNCP$(F, G_k, \mathcal{K})$ *admits a solution* x_k $\forall\, k \in \mathbb{N}$ *and the sequence of solutions* $\{x_k\}$ *is bounded, then every limit point of* $\{x_k\}$ *is a solution of the* GNCP$(F, G, \mathcal{K})$.

Proof. Since $\{x_k\}$ is bounded, it admits a convergent subsequence. Let K_1 be an infinite subset of $\mathbb{N}$, and x_* be such that

$$\lim_{k \in K_1} x_k = x_*.$$

If x_k is a solution of GNCP$(F, G_k, \mathcal{K})$, then

$$F(x_k) \in \mathcal{K}, \quad G(x_k) + \varepsilon_k F(x_k) \in \mathcal{K}^\circ, \quad \langle F(x_k), G(x_k) + \varepsilon_k F(x_k) \rangle = 0. \tag{47}$$

By the continuity of F and G and the closedness of $\mathcal{K}$, $\lim_{k \in K_1} F(x_k) = F(x_*) \in \mathcal{K}$, $\lim_{k \in K_1} G(x_k) = G(x_*) \in \mathcal{K}^\circ$, and $F(x_*)^T G(x_*) = 0$. $\square$

Remark. In Theorem 7 there is no assumption of monotonicity on either the original problem or the perturbed ones.

The result of [1, Theorem 3.2] is used next to characterize x_* in the set of solutions of GNCP$(F, G, \mathcal{K})$, denoted by SOL(GNCP$(F, G, \mathcal{K})$). Also, SOL(VIP) denotes the set of solutions of a VIP.

THEOREM 8. *Assume that* $G \circ F^{-1}$ *is monotone and that the set of solutions of the* GNCP$(F, G, \mathcal{K})$ *is not empty. Then the sequence* $\{x_k\}$ *of solutions of the GNCP*$(F, G_k, \mathcal{K})$ *converges to a solution* x_* *of the* GNCP$(F, G, \mathcal{K})$ *that is the unique solution of the problem*

$$\min \|F(x)\| \text{ subject to } x \in \mathrm{SOL}(\mathrm{GNCP}(F, G, \mathcal{K})). \tag{48}$$

Proof. If x_k is a solution of GNCP$(F, G_k, \mathcal{K})$, then, by Lemma 6, $F(x_k)$ is a solution of the VIP$(G_k \circ F^{-1}, \mathcal{K})$.

Since $G \circ F^{-1}(x)$ is monotone and

$$G_k(x) \circ F^{-1}(x) = (G + \varepsilon_k F) \circ F^{-1}(x) = G \circ F^{-1}(x) + \varepsilon_k x, \tag{49}$$

we have that $G_k \circ F^1(x)$ is strictly monotone. As F is an homeomorphism, [1, Theorem 3.2] implies that $\lim_{k\to\infty} F(x_k) = F(x_*)$, where $F(x_*)$ is the unique minimum norm solution of VIP$(G \circ F^{-1}, \mathcal{K})$ and solves the problem

$$\min \|F(x)\| \text{ subject to } F(x) \in \text{SOL(VIP)}.$$

Then, by Lemma 6, x_* is a solution of GNCP$(F, G, \mathcal{K})$ and is the unique solution of

$$\min \|F(x)\| \text{ subject to } x \in \text{SOL(GNCP}(F, G, \mathcal{K})). \qquad \square$$

The results obtained in this section allow us to solve GNCPs such that $G \circ F^{-1}$ is monotone using the approach developed in section 2 for the perturbed problems.

4. Computational experiments. Our set of experiments contains four families: randomly generated problems in the positive orthant, implicit complementarity problems from Outrata and Zowe [19], problems with general cones in $\mathbb{R}^n$, and problems in three-dimensional cones with control of generated faces.

For the first family of problems, functions F and G are affine and both cones are the positive orthant. Although quite simple, these problems contain essential elements to start the investigation. By varying dimensions and features of the matrices that define F and G, we have produced an extensive set of tests for which the theoretical hypothesis of equivalence might hold or not.

In the second family our main objective was to solve problems already addressed in the literature. We also extended the family of implicit complementarity problems proposed in [19] to variable dimension, producing large-scale tests. For such problems, however, the cones are the positive orthant as well.

General polyhedral cones were treated in the third and fourth families of problems. In the third one, functions F and G are affine and the matrices A and B that define the cones are generated to accomplish well defined problems, but without any specific control. In the fourth family, we produced three-dimensional tests, so that geometrical features of the cone, like control of edges and number of faces, were exploited to a great extent.

The equivalent minimization problems (2), with simple bounded variables, were solved using `BOX-QUACAN`, software developed by our research group at the State University of Campinas. It is based on the trust-region approach for solving large-scale bound-constrained minimization and uses the infinity norm to define the trust-region, so that the quadratic subproblems also have simple bounded variables. The subproblems are solved by combining conjugate gradients with projected gradients and a mild active set strategy (see [6, 12] or [9, p. 459]).

The code was developed in Fortran 77 double precision (Microsoft PowerStation) and run on a Pentium 64MB RAM. The stopping criteria used are tolerance for the objective function value $\varepsilon_f = 10^{-10}$ and tolerance for the norm of the continuous projected gradient $\varepsilon_g = 10^{-6}$. We set $\rho = 1$ for all the tests.

4.1. Randomly generated problems in the positive orthant. In our first set of experiments we considered the problem of finding $x \in \mathbb{R}^n$ such that $Mx+c \geq 0$, $Px + d \geq 0$, and $(Mx + c)^T(Px + d) = 0$, where matrices $M, P \in \mathbb{R}^{n\times n}$ and vectors $c, d \in \mathbb{R}^n$ are given.

The problems were randomly generated to exploit specific features of matrices M and P in a total of fourteen families as follows: M and P may be identical (families 1 to 6) or not (families 7 to 14); M and P may be symmetric (families 1 to 3 and 7 to 10) or not (4–6, 11–14); and matrices M and P may be regular (1, 2, 4, 5, 7, 8, 11, and 12) or singular (3, 6, 9, 10, 13, and 14). For each family, four values for the dimension n were used (5, 50, 500, and 5000). For each dimension, three problems were solved, with different seeds. For details on the generation, see [2].

Whenever M or P is invertible, the theoretical hypotheses of the equivalence results of section 2 can be verified by analyzing properties of matrices PM^{-1} or MP^{-1}. There were some problems, from families 8, 12, and 13, that converged to local nonglobal minimizers of (2), with merit function value greater than 10^{-1}. For problems from the first, second, fourth, and fifth sets, the theoretical hypotheses hold, representing 28.5% of the total number of tests. For families 1, 2, 4, 5, and 7, the algorithm computed the same solution that was generated for assembling the problem data. For families 3, 6, 10, and 14, since both matrices M and P are singular, the theoretical hypotheses fail, representing 28.5% of tests. For these tests, however, the global solution of (2) was always obtained. There is no guarantee that the theoretical hypotheses are valid for the test problems of sets 7, 8, 9, 11, 12, and 13, which represent 43% of tests. In fact, in 18 out of the 60 problems of these last six sets, at least one of the values $u^TPM^{-1}u$ or $v^TMP^{-1}v$, where $u = Mx + c - z$ and $v = Px + d - \lambda$, was negative. In the total of 168 problems solved, the hypotheses fail for 66 (39%), but only 16 converged to local solutions of (2), which correspond to 24% of the candidates for failure, and to 9.5% of the total of tests.

4.2. Implicit complementarity problems from Outrata and Zowe. In the second set of experiments we solved implicit complementarity problems (see [19]) of the following form:

Find $y \in \mathbb{R}^n$ such that

$$y - m(y) \geq 0, \quad F(y) \geq 0, \quad \text{and} \quad \langle F(y), y - m(y)\rangle = 0,$$

where $m_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, n$,

$$F(y) = Ay + b = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} y + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \tag{50}$$

and $m(y) = \varphi(Ay + b)$, with $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ twice continuously differentiable.

As in [19, Examples 4.3 and 4.4], the following choices for function φ defined our test problems:

$$\text{POZ1:} \quad \varphi_i(\lambda) = -0.5 - \lambda_i, \quad i = 1, 2, 3, 4, \quad \text{and}$$

$$\text{POZ2:} \quad \varphi_i(\lambda) = -1.5\lambda_i + 0.25\lambda_i^2, \quad i = 1, 2, 3, 4.$$

For each problem, three starting vectors were used, namely,

(a) $(0.0, 0.0, 0.0, 0.0)^T$,
(b) $(-0.5, -0.5, -0.5, -0.5)^T$,
(c) $(-1.0, -1.0, -1.0, -1.0)^T$.

In [19], Newtonian strategies were adopted to solve problems POZ1 and POZ2. In the *first approach*, the iterative scheme to compute fixed points of an operator S was

$$y_{k+1} = y_k - (E - V^k)^{-1}(y_k - S(y_k)),$$

where $V^k \in \partial S(y_k)$. In the *second approach*, a Newton variant scheme was applied to the semismooth operator

$$H(y) := \min\{y - m(y), F(y)\} = 0,$$

where min denotes the componentwise minimum of the two vectors in brackets.

Problems POZ1 and POZ2 were also solved in [16], with a trust-region approach for solving the GNCP$(F, G, \mathbb{R}^n_+)$ using the merit function $\Phi : \mathbb{R}^n \to R$ defined by

$$\Phi(x) := \frac{1}{2}\sum_{i=1}^{n} \phi(F_i(x), G_i(x))^2.$$

The function $\phi(a, b) = \sqrt{a^2 + b^2} - a - b$ is the Fischer–Burmeister one, with the property $\phi(a, b) = 0 \Leftrightarrow a \geq 0,\ b \geq 0,\ ab = 0$.

In Tables 1 and 2 we present, for comparative purposes, numerical results of [19] and [16] for problems POZ1 and POZ2, respectively. Our results are reported in Table 3, where the notation `INNER`, `MVP`, `OUTER`, and `FE` is used to indicate the number of iterations and matrix-vector products performed by the inner (quadratic) solver, and the number of iterations and functional evaluations performed by the outer (trust-region) algorithm. We also included the final value of our merit function $f(x, z, \lambda)$, together with the norm of the projected gradient $\|g_p\|$ at the final approximation.

Table 1
Previous results: Problem 1 (POZ1: $n = 4$).

OZ95			JFQS98		
Start	First approach `ITER`	Second approach `ITER`	`ITER`	`FE`	Φ
(a)	2	14	5	17	7.65D−18
(b)	2	41	4	16	9.71D−15
(c)	V^2 singular	56	5	11	3.43D−24

Table 2
Previous results: Problem 2 (POZ2: $n = 4$).

OZ95			JFQS98		
Start	First approach `ITER`	Second approach `ITER`	`ITER`	`FE`	Φ
(a)	3	15	5	17	1.05D−18
(b)	V^2 singular	15	4	16	4.89D−15
(c)	V^2 singular	No convergence	5	11	7.05D−22

The results of our approach compared quite well with [16] and were, by far, superior to the results of [19]. For problem POZ1, starting points (a) and (b) provide similar results in terms of computational effort, although point (b) generates a solution with slightly better quality. For this problem, starting with point (c), on the other hand, requires twice as many inner iterations and matrix-vector products as starting

TABLE 3
Results using our approach ($n = 4$).

Problem	Start	OUTER	FE	INNER	MVP	$f(x, z, \lambda)$	$\|g_p\|$
POZ1	(a)	4	5	24	30	2.31D−10	8.61D−06
	(b)	4	5	22	39	1.55D−14	7.03D−08
	(c)	4	5	45	68	6.63D−11	7.77D−06
POZ2	(a)	5	6	48	74	4.25D−12	2.33D−06
	(b)	6	8	104	171	1.15D−14	8.25D−08
	(c)	3	4	31	60	9.43D−11	2.25D−05

TABLE 4
Additional tests with larger dimensions.

Problem	Start	OUTER	FE	INNER	MVP	$f(x, z, \lambda)$	$\|g_p\|$
POZ1	(a)	7	10	125	236	1.26D−11	5.48D−06
$n = 40$	(b)	6	8	102	313	3.16D−13	4.52D−07
	(c)	5	7	84	176	1.11D−10	6.62D−06
POZ1	(a)	8	12	146	205	2.42D−12	8.69D−07
$n = 400$	(b)	7	10	126	201	5.66D−12	1.59D−06
	(c)	6	8	94	206	1.44D−11	2.32D−06
POZ1	(a)	9	14	143	311	1.59D−12	7.79D−07
$n = 4000$	(b)	8	12	123	377	7.43D−12	2.26D−06
	(c)	7	9	99	289	1.96D−11	2.91D−06
POZ2	(a)	7	11	127	248	4.36D−12	2.45D−06
$n = 40$	(b)	6	9	116	201	1.89D−11	2.56D−06
	(c)	6	8	104	176	6.90D−13	7.44D−07
POZ2	(a)	9	14	143	227	6.64D−13	5.35D−07
$n = 400$	(b)	7	11	135	367	1.75D−11	2.74D−06
	(c)	7	10	120	203	6.30D−13	4.93D−07
POZ2	(a)	10	15	157	394	2.98D−12	9.12D−07
$n = 4000$	(b)	9	14	161	385	7.84D−11	5.18D−06
	(c)	8	12	161	309	1.21D−12	4.94D−07

TABLE 5
Average results of our approach.

Problem	n	OUTER	FE	INNER	MVP
POZ1	4	4.0	5.0	30.3	45.7
	40	6.0	8.3	103.7	241.7
	400	7.0	10.0	122.0	204.0
	4000	8.0	11.7	121.7	325.7
POZ2	4	4.7	6.0	61.0	101.7
	40	6.3	9.3	115.7	208.3
	400	7.7	11.7	132.7	265.7
	4000	9.0	13.7	159.7	362.7

with (a) or (b). For problem POZ2, the starting point that generated the highest cost was (b).

To assess the reliability of our approach, we enlarged the dimension n of problems POZ1 and POZ2, allowing $n = 40$, $n = 400$, and $n = 4000$. Matrix $A \in \mathbb{R}^{n\times n}$ and vector $b \in \mathbb{R}^n$ are the natural extensions of (50), as are the starting vectors (a), (b), and (c). Results are presented in Table 4, where one can see that the computational effort grows very slowly as n increases. The greatest difference happens between $n = 4$ and $n = 40$, but from 40 to 400 and from 400 to 4000 the cost does not grow as much as in the first case. Such differences in the increasing factors can be better appreciated by the average values shown in Table 5.

4.3. Problems with general polyhedral cones in $\mathbb{R}^n$. In this third set of experiments we address the problem of finding $x \in \mathbb{R}^n$ such that $Mx + c \in \mathcal{K}$, $Px + d \in \mathcal{K}^\circ$, and $(Mx + c)^T(Px + d) = 0$, where the sets $\mathcal{K}$, $\mathcal{K}^\circ$ are defined by

$$\mathcal{K} = \{v \in \mathbb{R}^n \mid Av \geq 0,\ Bv = 0\},$$
$$\mathcal{K}^\circ = \{u \in \mathbb{R}^n \mid u = A^T\lambda_1 + B^T\lambda_2,\ \lambda_1 \geq 0\},$$

with $A \in \mathbb{R}^{q\times n}$, $B \in \mathbb{R}^{s\times n}$ given. Matrices $M, P \in \mathbb{R}^{n\times n}$ and vectors $c, d \in \mathbb{R}^n$ are also given.

The problems were randomly generated quite similarly to our first set of experiments. For details, see [2]. According to the features of matrices M and P, we divided the set of tests into three families: (1) $M = P$, indefinite and nonsymmetric; (2) $M = P$, indefinite and symmetric; (3) $M \neq P$, indefinite, nonsymmetric, and singular. For families (1) and (2) the theoretical hypotheses of the equivalence results hold since $PM^{-1} = I$.

For each family, six sets for the dimensions (n, q, s) were considered: $(10, 5, 1)$, $(10, 10, 1)$, $(10, 15, 1)$, $(100, 50, 5)$, $(100, 100, 5)$, and $(100, 150, 5)$. For each set of dimensions, three problems were generated, with different seeds. The arithmetic means of the results are reported in Tables 6 and 7, where we present the number of iterations (`INNER`) and matrix-vector products (`MVP`) performed by the inner (quadratic) solver, and the number of iterations (`OUTER`) and functional evaluations (`FE`) performed by the outer (trust-region) algorithm.

Table 6
Average results: Problems with $n = 10$, $s = 1$.

q	Family	INNER	MVP	OUTER	FE
5	1	136.7	170.3	9.0	10.0
10		184.0	257.0	11.3	12.3
15		309.0	436.7	18.0	19.0
5	2	168.0	213.0	11.3	12.3
10		168.7	232.3	11.8	12.8
15		208.3	282.3	12.7	13.7
5	3	208.7	253.3	10.0	11.0
10		278.7	371.7	13.0	14.0
15		485.7	640.7	19.7	20.7

Table 7
Average results: Problems with $n = 100$, $s = 5$.

q	Family	INNER	MVP	OUTER	FE
50	1	1021.0	1373.0	35.7	36.7
100		2199.3	2971.3	72.0	73.0
150		3946.3	5103.0	113.3	114.0
50	2	1064.7	1421.0	37.0	38.0
100		2167.7	2833.0	67.3	68.3
150		4291.3	5720.3	124.3	125.3
50	3	7397.7	7922.0	101.0	102.0
100		160724.0	166259.0	1856.0	1857.0
150		102189.0	112886.0	957.3	963.0

We denote the figures of Tables 6 and 7 by T_{ij}^k, where $k \in \{1, 2, 3\}$ represents each family, $i \in \{1, 2, 3\}$ corresponds to rows with $q = 5, 10, 15$ (Table 6), $i \in \{4, 5, 6\}$ corresponds to rows with $q = 50, 100, 150$ (Table 7), and $j \in \{1, 2, 3, 4\}$ is the corre-

sponding column with the values `INNER`, `MVP`, `OUTER`, and `FE`. Based on these values, we define cost measures to guide our analysis.

Concerning the effort spent by the algorithm, there are two aspects we would like to address: How is such effort related to the problem dimension, and how is it related to the problem features? Considering each dimension separately, we started by defining two cost measures, per inner iteration (`MVP`/`INNER`) and global (`INNER`/`OUTER`), as follows:

$$me_1(i) = \frac{1}{3}\sum_k \frac{T_{i2}^k}{T_{i1}^k} \quad \text{and} \quad me_2(i) = \frac{1}{3}\sum_k \frac{T_{i1}^k}{T_{i3}^k}$$

for $i = 1, 2, 3, 4, 5, 6$.

For a better understanding of the average values represented by these two measures, we also computed the minimum and maximum values:

$$m_1(i) = \min_k \frac{T_{i2}^k}{T_{i1}^k}, \quad M_1(i) = \max_k \frac{T_{i2}^k}{T_{i1}^k}, \quad m_2(i) = \min_k \frac{T_{i1}^k}{T_{i3}^k}, \quad \text{and} \quad M_2(i) = \max_k \frac{T_{i1}^k}{T_{i3}^k}.$$

Results are reported in Table 8, where the triples contain

$$(m_1(i), me_1(i), M_1(i)) \quad \text{and} \quad (m_2(i), me_2(i), M_2(i))$$

for $i = 1, \ldots, 6$.

TABLE 8
Measures of effort per problem dimension.

Dimension (q)	(m_1, me_1, M_1)	(m_2, me_2, M_2)
5	(1.22, 1.24, 1.26)	(14.69, 16.54, 20.03)
10	(1.34, 1.36, 1.38)	(14.79, 17.50, 21.34)
15	(1.33, 1.37, 1.42)	(16.49, 19.28, 24.11)
50	(1.09, 1.25, 1.34)	(28.24, 41.04, 66.12)
100	(1.06, 1.24, 1.35)	(30.49, 48.74, 83.55)
150	(1.09, 1.24, 1.34)	(34.24, 54.99, 95.91)

With the aim of analyzing results according to the family of generated problems, we define two additional measures for each one of sets 1 to 3. The weights $\ln(n+2q+s)$ and $\sqrt{\ln(n+2q+s)}$ were introduced to filter dependence of dimension and somehow make uniform the computed values:

$$me_3(k) = \frac{1}{6}\left(\sum_{i=1}^{3} \frac{T_{i2}^k}{\ln(11+10i)T_{i1}^k} + \sum_{i=4}^{6} \frac{T_{i2}^k}{\ln(100i-195)T_{i1}^k}\right)$$

and

$$me_4(k) = \frac{1}{6}\left(\sum_{i=1}^{3} \frac{T_{i2}^k}{\sqrt{\ln(11+10i)}T_{i1}^k} + \sum_{i=4}^{6} \frac{T_{i2}^k}{\sqrt{\ln(100i-195)}T_{i1}^k}\right)$$

for $k = 1, 2, 3$. We stress that the values $11+10i$, $i = 1, 2, 3$, and $100i-195$, $i = 4, 5, 6$ are, respectively, the dimensions 21, 31, 41 and 205, 305, 405 used in the tests. Results are shown in Table 9, where we also include minimum (m_3, m_4) and maximum values (M_3, M_4).

TABLE 9
Measures of effort per problem family.

Family	(m_3, me_3, M_3)	(m_4, me_4, M_4)
1	(0.50, 0.54, 0.58)	(23.08, 37.16, 54.70)
2	(0.51, 0.54, 0.56)	(23.19, 37.02, 53.86)
3	(0.43, 0.48, 0.54)	(31.34, 81.58, 151.38)

Observing Table 8, one can see that the effort of the inner solver is always inferior to 1.5 matrix-vector products per iteration. Moreover, it is slightly larger for smaller problems (dimensions $n + 2q + s \in \{21, 31, 41\}$) than for larger ones ($n + 2q + s \in \{205, 305, 405\}$), although the dispersion between minimum and maximum values grows with increasing q. This last comment also applies to the global effort measure me_2, that grows as q increases, together with the length of intervals $[m_2, M_2]$. Although dimension differs by a factor of ten for the two sets of problems, figures of (m_2, me_2, M_2) are about twice as large when the two sets are compared.

Concerning Table 9, the main conclusions are that symmetry of matrices M and P does not seem to interfere in the performance of our approach, since families 1 and 2 produced quite similar results for both triples (m_3, me_3, M_3) and (m_4, me_4, M_4). The singularity of matrices M and P, on the other hand, showed significant effects, especially as far as the global performance is concerned.

This set of experiments consists of 54 tests. For the 27 problems of smaller dimension, the final objective function value was always inferior to 10^{-5}. Considering the 27 large ones, for 8 problems of the third family the final objective function values were greater than 10^{-2}, indicating convergence to a local nonglobal solution. This amounts to 55.6% success among problems for which the theoretical condition of equivalence does not hold. We stress, however, that whenever the hypothesis is valid, a global solution was reached.

4.4. Problems in three-dimensional cones with control of generated faces. In the fourth set of experiments we addressed the problem of finding $x \in \mathcal{K} = \{v \in \mathbb{R}^n \mid Av \geq 0\}$ such that $Tx + c \in \mathcal{K}^\circ = \{v \in \mathbb{R}^n \mid A^T\lambda = v, \lambda \geq 0\}$ and $x^T(Tx + c) = 0$. We generated the polyhedral cones $\mathcal{K}$ with q faces, such that their edges were the lines

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\left(\frac{2\pi}{q}k\right) \\ r\sin\left(\frac{2\pi}{q}k\right) \\ 1 \end{pmatrix} t, \quad t \in \mathbb{R}, \quad k = 1, \ldots, q.$$

Therefore, $\mathcal{K}$ was defined by computing the rows of matrix A as the normal vectors to the support planes of the cone faces. In other words, the vector that defines the ith row of matrix A ($i = 1, \ldots, q$) is given by the cross-product

$$\begin{pmatrix} \cos\left(\frac{2\pi}{q}(i-1)\right) \\ \sin\left(\frac{2\pi}{q}(i-1)\right) \\ \frac{1}{r} \end{pmatrix} \times \begin{pmatrix} r\cos\left(\frac{2\pi}{q}i\right) \\ r\sin\left(\frac{2\pi}{q}i\right) \\ 1 \end{pmatrix} = \begin{pmatrix} \sin\left(\frac{2\pi}{q}i\right)\left(\cos\frac{2\pi}{q} - 1\right) - \cos\left(\frac{2\pi}{q}i\right)\sin\frac{2\pi}{q} \\ \cos\left(\frac{2\pi}{q}i\right)\left(1 - \cos\frac{2\pi}{q}\right) - \sin\left(\frac{2\pi}{q}i\right)\sin\frac{2\pi}{q} \\ r\sin\frac{2\pi}{q} \end{pmatrix}.$$

The problems were generated as follows. Given the values of the radius r and of the dimension q (number of faces of cone $\mathcal{K}$), we built matrix A and created two

types of solutions x_*, at the boundary and in the interior of $\mathcal{K}$, respectively. Next we randomly generated matrix T, keeping it symmetric, and produced four families of problems, namely, (1) T indefinite, (2) T positive definite, (3) T positive semidefinite, and (4) T negative semidefinite. For more details, see [2].

The tests were produced by varying $r \in \{0.1, 1, 10\}$, $q \in \{3, 4, 5, 6, 9, 12\}$, the four families of matrices T, and the two kinds of generated solution x_*, which amounted to 144 problems. Three distinct seeds were chosen to generate problems for each selection of r, q, T, and x_*.

To analyze the robustness of the proposed approach, since half of the generated problems do not satisfy the hypothesis of the equivalence result (families 1 and 4, with matrices T indefinite and negative semidefinite, respectively), we observed that for the 72 problems with x_* generated at the boundary of the cone, 29 out of the 72×3 tests stopped at local nonglobal solutions. This corresponds to success for 86.6% of the total and 73.2% of the candidates for failure. For problems with x_* generated in the interior of the cone, six problems converged to local nonglobal solutions, in a total of 72×3 problems. In this case, the measures of success are 97.2% of the total and 94.4% of the problems without theoretical guarantee of convergence. Summing up the two blocks of tests, there were 35 failures, representing success in 91.9% of total and 83.8% of the universe of problems that do not satisfy the hypothesis of equivalence result.

There are some salient features that emerge from the results. First, the computational cost of the inner solver grows with the problem dimension, reaching its maximum for $q = 9$ and $q = 5$ if x_* is generated at the boundary and in the interior of $\mathcal{K}$, respectively.

It is also evident that the degree of difficulty of the generated problems grows as the radius r decreases: $r = 10$ produces the easiest problems whereas $r = 0.1$ generates the most difficult ones. Recall that in this set of experiments our problem is to find $x \in \mathcal{K} = \{v \in \mathbb{R}^n \mid Av \geq 0\}$ such that $Tx + q \in \mathcal{K}^\circ = \{v \in \mathbb{R}^n \mid A^T\lambda = v, \lambda \geq 0\}$, so the requirements for $\mathcal{K}$ and $\mathcal{K}^\circ$ are different.

Grouping problems according to the features of matrix T, there are 36 problems for each family (6 dimensions q, 3 values for r, and 2 types of generated x_*). We have computed the ratios $\texttt{INNER}/n_t$ and $\texttt{OUTER}/n_t$, where $n_t = n + 2q$ is the dimension of problem (2), and calculated average values, presented in Table 10, together with minimum and maximum values.

TABLE 10
Measures of effort per problem features.

Family	$\texttt{INNER}/n_t$			$\texttt{OUTER}/n_t$		
	Minimum	Average	Maximum	Minimum	Average	Maximum
1	6.3	15.6	51.1	0.3	0.6	1.4
2	1.6	12.2	33.4	0.2	0.6	1.2
3	3.9	13.6	35.8	0.3	0.6	1.1
4	1.7	16.4	153.3	0.1	0.6	2.1

Observing the figures of Table 10, we see that solving problems from families 1 and 4 (T indefinite and negative semidefinite, respectively) demands more effort than solving those from families 2 and 3 (T positive definite and positive semidefinite, respectively). The largest dispersion, that is, the largest interval (minimum, maximum), occurs for the fourth family, because of an outlier. Removing this discrepant value, the triples become $(1.7, 14.5, 46.2)$ and $(0.1, 0.7, 1.2)$, with dispersions similar to those of the first family.

5. Conclusions. We proposed a smooth box-constrained minimization reformulation of the GNCP$(F, G, \mathcal{K})$, assuming that $\mathcal{K}$ is a polyhedral cone. Any efficient minimization algorithm for solving this kind of problems may be used. The study of perturbed problems gives information about the solutions of a GNCP$(F, G, \mathcal{K})$ for a general cone $\mathcal{K}$ with very mild assumptions on the problem data.

Computational experiments are presented which encourage the use of our approach. Four groups of problems were addressed: randomly generated problems in the positive orthant, implicit complementarity problems from Outrata and Zowe, problems with general cones in $\mathbb{R}^n$, and problems in three-dimensional cones with control of generated faces.

The numerical results showed that the solution of the GNCP using (2) was found in the majority of the tests, even without accomplishment of theoretical hypothesis, meaning that the behavior of the method does not depend strongly on the sufficient conditions that guarantee the equivalence. Quantifying this robustness, considering only the universe of problems without theoretical support for convergence, for the first set of experiments the amount of failure was 24%. In the third and fourth sets, local nonglobal solutions were reached in 44% and 16% of the tests, respectively. No doubt, in the absence of theoretical support, the convergence to global solutions is more frequent for problems of smaller dimensions. The second set of problems, included for comparative purposes, formed by implicit complementarity problems, contained large-scale experiments (dimension up to $3 \times 4000 = 12000$) for which our approach had a very good performance. The third set of experiments revealed that general polyhedral cones might produce quite difficult problems, especially as the dimension increases. The fourth group of tests was created to investigate geometrical features of the cone $\mathcal{K}$. Besides noticing that, for the generated three-dimensional problems, thinner cones need more effort than wider ones, we observed that the increasing number of edges and faces did not substantially augment the amount of effort needed to solve the problems. As a natural extension of this work, we would like to investigate the possibility of approximating a general cone by a polyhedral one. This leads us to look for further connections between theory and practice concerning geometrical and algebraic properties of general cones and their relationship with GNCP defined in these sets. We are also interested in studying the behavior of our approach applied to problems with nonlinear functions F and G and polyhedral cones.

An important question that arises concerns whether limit points of the sequences generated by the minimization algorithm exist. The boundedness of the level sets of the merit function is a sufficient condition for the existence of these limit points, and results in this sense are given in [3, 17]. This matter deserves future research.

REFERENCES

[1] R. Andreani, A. Friedlander, and J. M. Martínez, *On the solution of finite-dimensional variational inequalities, using smooth optimization with simple bounds*, J. Optim. Theory Appl., 94 (1997), pp. 635–657.

[2] R. Andreani, A. Friedlander, and S. A. Santos, *Solving Generalized Nonlinear Complementarity Problems: Numerical Experiments on Polyhedral Cones*, Technical Report, IMECC, State University of Campinas, Campinas, Brazil, 2001; also available online from http://www.dcce.ibilce.unesp.br/∼andreani.

[3] R. Andreani and J. M. Martínez, *On the reformulation of nonlinear complementarity problems using the Fischer-Burmeister function*, Appl. Math. Lett., 12 (1999), pp. 7–12.

[4] R. Andreani and J. M. Martínez, *Reformulation of variational inequalities on a simplex and compactification of complementarity problems*, SIAM J. Optim., 10 (2000), pp. 878–895.

[5] R. Andreani and J. M. Martínez, *On the solution of bounded and unbounded mixed complementarity problems*, Optimization, 50 (2001), pp. 265–278.

[6] R. H. Bielschowsky, A. Friedlander, F. A. M. Gomes, J. M. Martínez, and M. Raydan, *An adaptive algorithm for bound constrained quadratic minimization*, Investigación Operativa, 7 (1997), pp. 67–102.

[7] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, *A limited memory algorithm for bound constrained optimization*, SIAM J. Sci. Comput., 16 (1995), pp. 1190–1208.

[8] A. R. Conn, N. I. M. Gould, and Ph. L. Toint, *Global convergence of a class of trust region algorithms for optimization with simple bounds*, SIAM J. Numer. Anal., 25 (1988), pp. 433–460; see also SIAM J. Numer. Anal., 26 (1989), pp. 764–767.

[9] A. R. Conn, N. I. M. Gould, and Ph. L. Toint, *Trust-Region Methods*, MPS/SIAM Ser. on Optim. 1, SIAM, Philadelphia, 2000.

[10] M. C. Ferris and C. Kanzow, *Complementarity and related problems*, in Handbook of Applied Optimization, P. M. Pardalos and M. G. C. Resende, eds., to appear.

[11] M. Fiedler and V. Pták, *Some generalizations of positive definiteness and monotonicity*, Numer. Math., 9 (1966), pp. 163–172.

[12] A. Friedlander, J. M. Martínez, and S. A. Santos, *A new trust region algorithm for bound constrained minimization*, Appl. Math. Optim., 30 (1994), pp. 235–266.

[13] A. Friedlander, J. M. Martínez, and S. A. Santos, *On the resolution of linearly constrained convex minimization problems*, SIAM J. Optim., 4 (1994), pp. 331–339.

[14] A. Friedlander, J. M. Martínez, and S. A. Santos, *Solution of linear complementarity problems using minimization with simple bounds*, J. Global Optim., 6 (1995), pp. 1–15.

[15] A. Friedlander, J. M. Martínez, and S. A. Santos, *A new strategy for solving variational inequalities on bounded polytopes*, Numer. Funct. Anal. Optim., 6 (1995), pp. 653–668.

[16] H. Jiang, M. Fukushima, L. Qi, and D. Sun, *A trust region method for solving generalized complementarity problems*, SIAM J. Optim., 8 (1998), pp. 140–157.

[17] C. Kanzow and M. Fukushima, *Equivalence of the generalized complementarity problem to differentiable unconstrained minimization*, J. Optim. Theory Appl., 90 (1996), pp. 581–603.

[18] O. L. Mangasarian and M. V. Solodov, *Nonlinear complementarity problem as unconstrained and constrained minimization*, Math. Program., 62 (1993), pp. 277–297.

[19] J. V. Outrata and J. Zowe, *A Newton method for a class of quasi-variational inequalities*, Comput. Optim. Appl., 4 (1995), pp. 5–21.

[20] J.-S. Pang, *The implicit complementarity problem*, in Nonlinear Programming 4, O. Mangasarian, M. Robinson, and R. Meyer, eds., Academic Press, New York, 1981, pp. 487–518.

[21] J.-S. Pang, *Complementarity problem*, in Handbook of Global Optimization, R. Horst and P. Pardalos, eds., Kluwer Academic Publishers, Boston, 1995, pp. 271–338.

[22] J. M. Peng, *Equivalence of variational inequality problems to unconstrained optimization*, Math. Programming, 78 (1997), pp. 347–355.

[23] R. B. Schnabel and T.-T. Chow, *Tensor methods for unconstrained optimization using second derivatives*, SIAM J. Optim., 1 (1991), pp. 293–315.

[24] P. Tseng, N. Yamashita, and M. Fukushima, *Equivalence of complementarity problems to differentiable minimization: A unified approach*, SIAM J. Optim., 6 (1996), pp. 446–460.

[25] N. Yamashita, K. Taji, and M. Fukushima, *Unconstrained optimization reformulations of variational inequality problems*, J. Optim. Theor. Appl., 92 (1997), pp. 439–456.

[26] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, *L-BFGS-B Fortran Subroutines for Large-Scale Bound Constrained Optimization*, Technical report, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, 1995.