



Solving asymmetric variational inequalities via convex optimization

Michele Aghassi^a, Dimitris Bertsimas^{b,*}, Georgia Perakis^b

^aOperations Research Center, Massachusetts Institute of Technology, E40-131, Cambridge, MA 02139, USA

^bSloan School of Management and Operations Research Center, Massachusetts Institute of Technology, E53-363, Cambridge, MA 02139, USA

Received 20 May 2005; accepted 13 September 2005

Available online 6 December 2005

Abstract

Using duality, we reformulate the asymmetric variational inequality (VI) problem over a conic region as an optimization problem. We give sufficient conditions for the convexity of this reformulation. We thereby identify a class of VIs that includes monotone affine VIs over polyhedra, which may be solved by commercial optimization solvers.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Variational inequalities; Convex optimization; Duality

1. Introduction

The variational inequality (VI) problem has engaged members of the optimization, mathematics, transportation science, engineering, and economics communities. Given a set $K \subseteq \mathbb{R}^n$ and a mapping $\mathbf{F} : K \rightarrow \mathbb{R}^n$, the VI problem, denoted $VI(K, \mathbf{F})$, is to find an $\mathbf{x}^* \in K$ such that

$$\mathbf{F}(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) \geq 0 \quad \forall \mathbf{x} \in K. \quad (1)$$

VIs, first introduced by Stampacchia and his collaborators [18,19,23,27,28], subsume many other well-studied mathematical problems, including the solution

of systems of equations, complementarity problems, and a class of fixed point problems. In addition, for any optimization problem over a closed, convex feasible region, the first-order optimality conditions comprise a VI. Accordingly, the VI problem also generalizes convex optimization.

For a complete discussion and history of the VI problem and associated solution methods, we refer the interested reader to the recent survey text by Facchinei and Pang [12] and the monograph by Patriksson [24]. The survey article by Harker and Pang [16] and the Ph.D. thesis of Hammond [15], as well as the references therein, also provide insightful reviews of the VI problem and associated algorithms.

One class of techniques for solving the VI problem exploits the fact that the Karush–Kuhn–Tucker (KKT) conditions of a VI comprise a mixed complementarity

* Corresponding author.

E-mail addresses: maghassi@alum.mit.edu (M.Aghassi), dbertsim@mit.edu (D. Bertsimas), georgiap@mit.edu (G. Perakis).

problem (MiCP), involving both equations and non-negativity constraints. One may solve these KKT conditions using interior point methods.

In contrast, instead of reformulating the VI problem as a system of equations and inequalities, other contributors to the VI literature have suggested reformulating the VI as an equivalent optimization problem. Most simply, it is well known (see, e.g., [12, Theorem 1.3.1]) that if the Jacobian of \mathbf{F} , denoted by $J\mathbf{F}(\mathbf{x})$, is symmetric $\forall \mathbf{x} \in K$, then there exists a function $f : K \rightarrow \mathbb{R}$ such that

$$\nabla f(\mathbf{x}) = \mathbf{F}(\mathbf{x}) \quad \forall \mathbf{x} \in K.$$

If, in addition, K is closed and convex, then \mathbf{x}^* solves $VI(K, \mathbf{F})$ iff it is a stationary point of

$$\min_{\mathbf{x} \in K} f(\mathbf{x}). \quad (2)$$

When \mathbf{F} is furthermore monotone over K , or equivalently, when its Jacobian matrix is positive semidefinite over K , the resulting optimization problem (2) is convex, and \mathbf{x}^* solves $VI(K, \mathbf{F})$ iff it is an optimal solution of (2). Depending on the exact form of this convex program, one may apply any of a number of commercial solvers, including ILOG CPLEX, ILOG Solver, Xpress-MP, and LINDO. Moreover, when \mathbf{F} is affine and K is polyhedral, problem (2) is a linearly constrained quadratic program (LCQP).

For a VI with possibly asymmetric $J\mathbf{F}(\mathbf{x})$, the concept of merit functions provides for the reformulation as an equivalent optimization problem. For $X \supseteq K$, where X is a closed set, a merit function for $VI(K, \mathbf{F})$ is defined to be a non-negative $\theta : X \rightarrow \mathbb{R}$ such that \mathbf{x}^* solves $VI(K, \mathbf{F})$ iff $\mathbf{x}^* \in X$ and $\theta(\mathbf{x}^*) = 0$. Thus, $VI(K, \mathbf{F})$ is equivalent to $\min_{\mathbf{x} \in X} \theta(\mathbf{x})$. The earliest and most intuitive merit function, proposed by Zuhovickii et al. [34] in the context of game-theoretic equilibrium computation, is the classical primal gap function, given by

$$\theta_{\text{gap}}(\mathbf{x}) \triangleq \sup_{\mathbf{y} \in K} \mathbf{F}(\mathbf{x})'(\mathbf{x} - \mathbf{y}). \quad (3)$$

This gap function and other more sophisticated merit functions give rise to specialized iterative descent methods for their minimization and, as a consequence, for the solution of the VI problem.

1.1. Complexity of VI algorithms and VI solver availability

While the favored reformulations and associated algorithms for the VI problem have proved to be practically useful, they are not without drawbacks. For instance, the merit function approach to solving VIs involves an iterative descent method in which, at each iteration, evaluation of the merit function and its gradient requires the computation of a projection onto K . Under certain conditions, usually involving continuity and some version of monotonicity, iterative descent methods for merit function minimization yield limit points that are stationary points of the corresponding merit-function optimization problem (see, e.g., [12, Section 10.6] for an overview). However, even when these requirements are met, the iterative descent algorithms may not converge in finite, let alone polynomial time.

In fact, the topic of complexity is not as often discussed in the VI literature as in other areas of optimization. Most algorithms for VIs are guaranteed to yield asymptotic convergence, with the exceptions of ellipsoid methods [21,22,25], and interior point algorithms (see, e.g., [17,29–33]). Under certain conditions, ellipsoid and interior point methods for VIs ensure polynomial-time convergence. While ellipsoid algorithms provide important insights into the complexity of the VI problem, they are not applied in practice. Moreover, although interior point methods can efficiently solve interesting classes of the VI problem, commercial-grade software packages for these algorithms are not available to the same extent as are commercial products for solving general convex QPs and other convex optimization problems.

Accordingly, to solve instances of the VI problem, one would ideally like to harness the power of these industrial-strength convex optimization software packages. To the best of our knowledge, in the absence of symmetry of $J\mathbf{F}(\mathbf{x})$, the reformulation of a VI as a single-level, convex optimization problem, which is many-times continuously differentiable, is not well understood.

As already noted, the VI literature includes reformulations for problem instances in which \mathbf{F} has a symmetric Jacobian over K . While the merit-function technique further provides an optimization-based approach for the asymmetric case, merit functions are

usually at most once continuously differentiable (see, e.g., the discussion in [12, Section 10.6], [13, Section 4.2]). More importantly, for almost all merit functions, evaluation of the function itself requires solution of an optimization problem, generally involving a maximization over K . Thus, the merit function approach requires the solution of a bilevel optimization problem, a form that commercial optimization solvers cannot handle.

Finally, let us return to the case of a VI problem satisfying the symmetry and monotonicity requirements ensuring the existence of an objective function f , as in (2), inducing an equivalent convex optimization problem. Even in this case, identification of such an objective function f requires the evaluation, in closed form, of the indefinite line integral

$$\begin{aligned} f(\mathbf{x}) &= \oint \mathbf{F}(\mathbf{x})' d\mathbf{x} \\ &= \int_0^1 \mathbf{F}(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0))' (\mathbf{x} - \mathbf{x}^0) dt, \end{aligned}$$

where \mathbf{x}^0 is any vector in K . This integration is easy if \mathbf{F} is affine. However, if \mathbf{F} is nonlinear, and especially if n , the dimension of the problem, is large, analytically evaluating such an integral may be impractical. Ideally, one would like to develop a reformulation of the VI problem that does not require such preprocessing in order to generate a convex optimization problem for input into a commercial solver.

1.2. Contributions of the paper

In this paper, we study instances of $VI(K, \mathbf{F})$ in which K is a polyhedral or more general conic region. Our contributions are as follows:

1. Using duality, we reformulate any such VI as a single-level, and many-times continuously differentiable optimization problem, even if the associated cost function has an asymmetric Jacobian matrix. (See Section 2.1.) Our method thereby avoids the difficulties discussed in Section 1.1.
2. We give sufficient conditions for the convexity of this reformulation. We thereby identify a class of VIs, of which monotone affine (and possibly asymmetric) VIs over polyhedra are a special case, which may be solved using widely available and

commercial-grade convex optimization software. (See Section 2.3.)

3. In addition, we note that the VI problem may be viewed as a special instance of a robust constraint, and that robust optimization therefore subsumes the VI problem. (See Section 2.2.)

1.3. Notation

We use the following notation conventions throughout the paper. Boldface letters denote vectors and matrices, upper case letters signify matrices, while lower case denotes vectors. We use subscripts to denote elements of a vector or matrix and superscripts to denote one entire vector or matrix in a sequence. For a square, but not necessarily symmetric matrix \mathbf{A} , $\mathbf{A} \succeq 0$ denotes that \mathbf{A} is positive semidefinite.

2. Applying duality to VIs

In this section, we apply duality to the VI problem in order to reformulate it as an equivalent, single-level optimization problem. We initially focus our attention on the class of VIs in which K is polyhedral, and we then extend our results to more general conic regions. We make no assumptions on the existence of solutions to the VI problems we consider.

2.1. Reformulation of the VI as a single-level optimization problem

For K polyhedral, without loss of generality, we may restrict our attention to standard-form instances,

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\} \neq \emptyset, \tag{4}$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Rewriting K in standard form may require parallel changes in \mathbf{F} . For instance, recasting non-positive variables as non-negative ones requires changes of sign in \mathbf{F} , and replacement of free variables by the difference of non-negative variables requires augmenting the dimension of \mathbf{F} .

We now state and prove a constraint equivalence.

Theorem 1. *Suppose that K is the non-empty polyhedron given by (4). Then, \mathbf{x}^* solves $VI(K, \mathbf{F})$ iff $\exists \lambda^* \in$*

\mathbb{R}^m such that $(\mathbf{x}, \lambda) = (\mathbf{x}^*, \lambda^*)$ satisfies

$$\mathbf{F}(\mathbf{x})' \mathbf{x} = \mathbf{b}' \lambda,$$

$$\mathbf{A} \mathbf{x} = \mathbf{b},$$

$$\mathbf{x} \geq 0,$$

$$\mathbf{A}' \lambda \leq \mathbf{F}(\mathbf{x}). \tag{5}$$

Proof. As is well known in the VI literature (see, e.g., [12, Section 1.2]), by the definition of the VI problem, $\mathbf{x}^* \in K$ satisfies (1) iff the following relation holds.

$$\begin{aligned} \mathbf{F}(\mathbf{x}^*)' \mathbf{x}^* = \min_{\mathbf{x}} \quad & \mathbf{F}(\mathbf{x}^*)' \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0. \end{aligned} \tag{6}$$

That is, \mathbf{x}^* must itself optimize the linear optimization problem (LP) (6) it induces. This well-known observation is credited to Eaves [10], who originally noted this equivalence in the context of the complementarity problem.

In LP (6), \mathbf{x}^* is treated as data and \mathbf{x} is the vector of decision variables. Since \mathbf{x}^* in this way parameterizes LP (6), we refer to this LP as $LP(\mathbf{x}^*)$. Its dual, to which we refer as $DLP(\mathbf{x}^*)$, is

$$\begin{aligned} \max_{\lambda} \quad & \mathbf{b}' \lambda \\ \text{s.t.} \quad & \mathbf{A}' \lambda \leq \mathbf{F}(\mathbf{x}^*). \end{aligned} \tag{7}$$

Suppose that \mathbf{x}^* solves $VI(K, \mathbf{F})$. Then $LP(\mathbf{x}^*)$ has bounded optimal value, given by $\mathbf{F}(\mathbf{x}^*)' \mathbf{x}^*$. By LP strong duality, its dual, $DLP(\mathbf{x}^*)$, is also feasible with bounded optimal value equal to that of $LP(\mathbf{x}^*)$. Let λ^* denote an optimal solution of $DLP(\mathbf{x}^*)$. Then $(\mathbf{x}^*, \lambda^*)$ satisfies (5).

For the reverse direction, suppose that $(\mathbf{x}^*, \lambda^*)$ satisfies system (5). Then, \mathbf{x}^* and λ^* are primal and dual feasible for $LP(\mathbf{x}^*)$ and $DLP(\mathbf{x}^*)$, respectively. Since $\mathbf{F}(\mathbf{x}^*)' \mathbf{x}^* = \mathbf{b}' \lambda^*$, by LP weak duality, \mathbf{x}^* must be optimal for $LP(\mathbf{x}^*)$. Therefore, \mathbf{x}^* solves $VI(K, \mathbf{F})$. \square

Remark. Let us note two points relating to Theorem 1. First, our use of duality to reformulate the VI as a single-level optimization problem is inspired by the robust optimization literature’s use of duality (see, e.g., [8]) to reformulate a robust counterpart of an optimization problem under uncertainty as a nominal problem, i.e., one involving no uncertainty. We further discuss

the connection between the VI problem and robust optimization in Section 2.2.

Second, since the KKT conditions of any LP are necessary and sufficient for its optimality, \mathbf{x}^* solves $VI(K, \mathbf{F})$, with K polyhedral, iff there exist KKT multipliers for which \mathbf{x}^* and these multipliers satisfy the KKT system corresponding to $LP(\mathbf{x}^*)$. This equivalence is well known in the VI literature (see, e.g., [12, Proposition 1.2.1]). Accordingly, to prove Theorem 1, rather than using strong duality, we could have used the KKT conditions of $LP(\mathbf{x}^*)$. For such an alternate, KKT-based proof, see Aghassi [1].

Theorem 1 implies the following equivalence of the VI with an optimization problem.

Corollary 1. *Suppose that K is the non-empty polyhedron given by (4). \mathbf{x}^* solves $VI(K, \mathbf{F})$ iff the following mathematical optimization problem (MP) has optimal value zero and $\exists \lambda^* \in \mathbb{R}^m$ such that $(\mathbf{x}^*, \lambda^*)$ is an optimal solution*

$$\begin{aligned} \min_{\mathbf{x}, \lambda} \quad & \mathbf{F}(\mathbf{x})' \mathbf{x} - \mathbf{b}' \lambda \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \\ & \mathbf{A}' \lambda \leq \mathbf{F}(\mathbf{x}). \end{aligned} \tag{8}$$

Proof. To begin, note that $\forall (\mathbf{x}, \lambda)$ feasible for MP (8), \mathbf{x} is a feasible solution of $LP(\mathbf{x})$, and λ is a feasible solution of $DLP(\mathbf{x})$. In addition, the objective function of MP (8) represents the duality gap of this primal-dual pair of solutions. Consequently, if MP (8) is feasible, by LP weak duality, its objective value is always non-negative. From this observation and the equivalence proved in Theorem 1, the result immediately follows. \square

Remark. It is worth pausing to note that, if $VI(K, \mathbf{F})$ has no solution, then MP (8) either is infeasible or has a strictly positive optimal value. Consider any $\mathbf{x}^* \in K$ that does *not* solve $VI(K, \mathbf{F})$. While we make no assumptions on the boundedness of K , because $K \neq \emptyset$, there are only two possibilities for $LP(\mathbf{x}^*)$. If this LP is unbounded, then there does not exist a $\lambda \in \mathbb{R}^m$ such that (\mathbf{x}^*, λ) is feasible for MP (8). Otherwise, $LP(\mathbf{x}^*)$ must be bounded but have optimal value strictly less than $\mathbf{F}(\mathbf{x}^*)' \mathbf{x}^*$. In this case, $\forall \lambda \in \mathbb{R}^m$ such that (\mathbf{x}^*, λ) is feasible for MP (8), the corresponding objective value is, by weak duality, strictly positive.

Corollary 1 establishes that any VI with polyhedral K can *always* be reformulated as a single-level optimization problem, even under asymmetry of the Jacobian matrix of the associated $\mathbf{F}(\mathbf{x})$. Thus, any such problem is always equivalent to an optimization problem of a format compatible with commercial optimization software, which does not accept constraint or objective functions that themselves involve optimization problems. Furthermore, the objective and constraint functions of our reformulation (8) are, in contrast to most merit functions in the VI literature, continuously differentiable as many times as $\mathbf{F}(\mathbf{x})$ is. For the affine VI, Corollary 1 implies that this problem can always be reformulated as a linearly constrained quadratic program (LCQP).

The equivalence established in Corollary 1 generalizes the well-known equivalence (see, e.g., [12, Section 1.5.3]) between the complementarity problem, i.e., identify \mathbf{x} such that $\mathbf{x} \geq 0$, $\mathbf{F}(\mathbf{x}) \geq 0$, and $\mathbf{F}(\mathbf{x})' \mathbf{x} = 0$, and the optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{F}(\mathbf{x})' \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x} \geq 0, \quad \mathbf{F}(\mathbf{x}) \geq 0. \end{aligned} \tag{9}$$

Note that this complementarity problem is the special case of $VI(K, \mathbf{F})$, in which $K = \mathbb{R}_+^n$, i.e., the non-negative orthant. Indeed, MP (8) reduces to MP (9) when $\mathbf{A} = 0$ and $\mathbf{b} = 0$.

We have thus far considered only polyhedral K . In fact, we may extend our VI reformulation result to the more general case, in which K is defined in terms of an arbitrary cone $C \subseteq \mathbb{R}^n$. In this setting, C may be a cone other than the non-negative orthant, in terms of which all polyhedra are defined. Without loss of generality, consider

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq_C 0\} \neq \emptyset. \tag{10}$$

We use the convention that, for any cone C , $\mathbf{x} \geq_C 0$ denotes that $\mathbf{x} \in C$, and $\mathbf{x} \succ_C 0$ denotes that $\mathbf{x} \in \text{int}(C)$, the interior of C . Let C^* denote the cone dual to C , i.e.,

$$C^* = \{\boldsymbol{\eta} \in \mathbb{R}^n \mid \boldsymbol{\eta}' \mathbf{x} \geq 0, \forall \mathbf{x} \in C\}.$$

We omit the proof of the following theorem, since it is analogous to that of Corollary 1. For a review of conic duality, we refer the interested reader to Ben-Tal and Nemirovski [5].

Theorem 2. Consider $VI(K, \mathbf{F})$, where K is given by (10). Suppose the following MP has optimal value zero and $\exists \boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ is an optimal solution

$$\begin{aligned} \min_{\mathbf{x}, \boldsymbol{\lambda}} \quad & \mathbf{F}(\mathbf{x})' \mathbf{x} - \mathbf{b}' \boldsymbol{\lambda} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq_C 0 \\ & \mathbf{A}' \boldsymbol{\lambda} \leq_{C^*} \mathbf{F}(\mathbf{x}). \end{aligned} \tag{11}$$

Then \mathbf{x}^* solves $VI(K, \mathbf{F})$.

Conversely, suppose that \mathbf{x}^* solves $VI(K, \mathbf{F})$, that $\exists \mathbf{x} \in K$ such that $\mathbf{x} \succ_C 0$, and that $\exists \boldsymbol{\lambda} \in \mathbb{R}^m$ such that $\mathbf{A}' \boldsymbol{\lambda} \prec_{C^*} \mathbf{F}(\mathbf{x}^*)$. Then, MP (11) has optimal value zero, and $\exists \boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ is an optimal solution.

Let $\mathbb{M}^{m,n} \subseteq \mathbb{R}^{mn}$ denote the space of $m \times n$ matrices, $\mathbb{S}^n \subseteq \mathbb{R}^{n^2}$ denote the space of symmetric $n \times n$ matrices, and \mathbb{S}_+^n denote the self-dual cone of symmetric, positive semidefinite, $n \times n$ matrices. In addition, let us define the inner product of $\mathbf{X} \in \mathbb{M}^{m,n}$ with $\mathbf{Y} \in \mathbb{M}^{m,n}$ as

$$\mathbf{X} \bullet \mathbf{Y} = \sum_{i=1}^m \sum_{j=1}^n X_{ij} Y_{ij},$$

where X_{ij} and Y_{ij} are the (i, j) th elements of the matrices \mathbf{X} and \mathbf{Y} , respectively. Setting $C = \mathbb{S}_+^n$ in Theorem 2, we obtain the following corollary, specific to VIs over subsets of \mathbb{S}_+^n . Note that, in terms of the notation conventions set forth in Section 1.3 and in this section, for $\mathbf{X} \in \mathbb{M}^{n,n}$, $\mathbf{X} \geq_{\mathbb{S}_+^n}$ is a stronger statement than $\mathbf{X} \geq 0$, since the former requires that \mathbf{X} is symmetric, while the latter does not.

Corollary 2. Consider an arbitrary $\mathbf{F} : \mathbb{M}^{n,n} \rightarrow \mathbb{M}^{n,n}$ and $VI(K, \mathbf{F})$, where

$$K = \left\{ \mathbf{X} \in \mathbb{M}^{n,n} \mid \mathbf{A}^i \bullet \mathbf{X} = b_i, \right. \\ \left. i = 1, \dots, m; \mathbf{X} \geq_{\mathbb{S}_+^n} 0 \right\}, \tag{12}$$

$\mathbf{A}^i \in \mathbb{M}^{n,n}$, and $b_i \in \mathbb{R}$, $i \in \{1, \dots, m\}$. Suppose the following MP has optimal value zero and $\exists \boldsymbol{\lambda}^* \in \mathbb{R}^m$

such that $(\mathbf{X}^*, \lambda^*)$ is an optimal solution.

$$\begin{aligned}
 & \min_{\substack{\mathbf{X} \in \mathbb{M}^{n,n} \\ \lambda \in \mathbb{R}^m}} \mathbf{F}(\mathbf{X}) \bullet \mathbf{X} - \mathbf{b}'\lambda \\
 & \text{s.t. } \mathbf{A}^i \bullet \mathbf{X} = b_i, \quad i = 1, \dots, m \\
 & \quad \mathbf{X} \succeq_{\mathbb{S}_+^n} \mathbf{0} \\
 & \quad \sum_{i=1}^m \lambda_i \mathbf{A}^i \preceq_{\mathbb{S}_+^n} \mathbf{F}(\mathbf{X}). \tag{13}
 \end{aligned}$$

Then \mathbf{X}^* solves $VI(K, \mathbf{F})$.

Conversely, suppose that \mathbf{X}^* solves $VI(K, \mathbf{F})$, that $\exists \mathbf{X} \in K$ such that $\mathbf{X} \succ_{\mathbb{S}_+^n} \mathbf{0}$, and that $\exists \lambda \in \mathbb{R}^m$ such that $\sum_{i=1}^m \lambda_i \mathbf{A}^i \prec_{\mathbb{S}_+^n} \mathbf{F}(\mathbf{X}^*)$. Then, MP (13) has optimal value zero, and $\exists \lambda^* \in \mathbb{R}^m$ such that $(\mathbf{X}^*, \lambda^*)$ is an optimal solution.

Remark. Note that when

$$\mathbf{F}(\mathbf{X}) = \sum_{i=1}^n \sum_{j=1}^n X_{ij} \mathbf{G}^{ij} + \mathbf{H},$$

where $\mathbf{H} \in \mathbb{M}^{n,n}$, and $\mathbf{G}^{ij} \in \mathbb{M}^{n,n}$, for $i, j \in \{1, \dots, n\}$, the objective function of MP (13) is quadratic in \mathbf{X} and λ .

2.2. Robust optimization generalizes the VI problem

Before discussing the convexity of our reformulations of the VI, we digress to note an interesting connection between the VI problem and robust optimization. Recall that robust optimization (see, e.g., [3,4,7,8,11,26]), comprises a deterministic, worst-case approach to data uncertainty. In particular, suppose a vector $\tilde{\phi}$ of data parameters defining optimization problem

$$\begin{aligned}
 & \min_{\mathbf{x}} f(\mathbf{x}) \\
 & \text{s.t. } \mathbf{x} \in X(\tilde{\phi}), \tag{14}
 \end{aligned}$$

where \mathbf{x} is the vector of decision variables, is subject to uncertainty. In the robust optimization framework, we assume we know only an uncertainty set U of possible values that $\tilde{\phi}$ may realize. The robust counterpart of MP (14) is given by

$$\begin{aligned}
 & \min_{\mathbf{x}} f(\mathbf{x}) \\
 & \text{s.t. } \mathbf{x} \in X(\tilde{\phi}) \quad \forall \tilde{\phi} \in U.
 \end{aligned}$$

Table 1

Size of the original VI formulation versus that of our single-level optimization reformulation

	Original VI	Optimization reformulation
Variables	n	$m + n$
Constraints	$(m + n) + K $	$m + 2n$

Although $VI(K, \mathbf{F})$ involves no data uncertainty whatsoever, for arbitrary K and \mathbf{F} , it is in fact a special instance of a robust constraint. Namely, VI (1) is a robust constraint in which it is as if \mathbf{x} is subject to uncertainty and known only to belong to K . Accordingly, the robust optimization framework encompasses the VI problem.

2.3. Convexity of reformulations

Having reformulated the VI as a single-level optimization problem, we next use this reformulation to analyze the complexity of the original VI. In particular, converting the VI into a single-level optimization problem facilitates certification of polynomial-time complexity. Indeed, one need not bother proving the polynomial-time convergence of an algorithm specifically tailored to the VI. Rather, if our proposed optimization problem reformulation is solvable in polynomial time (e.g., if it is convex and satisfies other complexity conditions), then the complexity of the original VI is polynomial time.

To understand the justification of this result for the case of polyhedral K , note that if the standard-form representation of K involves $m + n$ constraints and n variables, then the size of reformulation (8) of $VI(K, \mathbf{F})$ is polynomial in m and n . The number of variables and constraints in this reformulation are given in Table 1. While our reformulation is in a space of higher dimension than the original VI formulation, this increase in size is modest and furthermore allows the VI to be expressed in terms of finitely many constraints (polynomially many with respect to m and n), rather than infinitely many constraints (when $|K|$, the cardinality of K , is infinite). Moreover, one can easily reduce the dimensionality of the reformulation by manipulating equality constraints in order to eliminate redundant variables, as is a standard technique in optimization (see, e.g., [9]).

In the following theorem, we give a sufficient condition for the convexity of reformulations (5) and (8).

Theorem 3. *Suppose that K is the non-empty polyhedron given by (4). If $F_j(\mathbf{x})$ is a concave function over K , $\forall j \in \{1, \dots, n\}$, and $\mathbf{F}(\mathbf{x})'\mathbf{x}$ is a convex function over K , then system (5) defines a convex feasible region and MP (8) defines a convex optimization problem.*

Proof. The result for MP (8) follows immediately from Corollary 1. If we replace the constraint $\mathbf{F}(\mathbf{x})'\mathbf{x} = \mathbf{b}'\lambda$ in system (5) with $\mathbf{F}(\mathbf{x})'\mathbf{x} \leq \mathbf{b}'\lambda$, we obtain a resulting system that is equivalent to (5). The reason is that, as explained in the proof of Corollary 1, $\mathbf{F}(\mathbf{x})'\mathbf{x} - \mathbf{b}'\lambda$ is the duality gap of primal-dual pair (6) and (7) and therefore cannot be strictly negative. Accordingly, the convexity of the solution set of system (5) follows immediately from Theorem 1. \square

Remark. It is worth noting that if K is not, without reformulation, a subset of the non-negative orthant, then the concavity requirements on $F_j(\mathbf{x})$ in Theorem 3 may be directly stated as follows. If, $\forall \mathbf{x} \in K$, $x_j \geq 0$, then the sufficient condition asks that $F_j(\mathbf{x})$ be concave, as before. Alternatively, if, $\forall \mathbf{x} \in K$, $x_j \leq 0$, then the sufficient condition asks that $F_j(\mathbf{x})$ be convex. Otherwise, the sufficient condition asks that $F_j(\mathbf{x})$ be affine.

Returning to our discussion of the implications and significance of Theorem 3, if $VI(K, \mathbf{F})$ satisfies the theorem's conditions, it is equivalent to a convex optimization problem, given by MP (8), and can therefore be solved by commercial convex optimization software. For the affine case, the conditions of Theorem 3 simplify considerably. In particular, when $\mathbf{F}(\mathbf{x})$ is affine, the following equivalence holds. $\mathbf{F}_j(\mathbf{x})$ is a concave function over K , $\forall j \in \{1, \dots, n\}$, and $\mathbf{F}(\mathbf{x})'\mathbf{x}$ is a convex function over K iff \mathbf{F} is monotone over K . The reason is that monotonicity holds iff $J\mathbf{F}(\mathbf{x}) \geq 0$ (see, e.g., [12, Proposition 2.3.2]). Accordingly, any affine monotone VI is polynomially solvable using a commercial QP solver applied to a convex LCQP. In contrast, when $\mathbf{F}(\mathbf{x})$ is not affine, the aforementioned equivalence fails. In particular, if $\mathbf{F}_j(\mathbf{x})$ is a concave function over K , $\forall j \in \{1, \dots, n\}$, but $\mathbf{F}(\mathbf{x})$ is not affine, then monotonicity on K is necessary, but not sufficient for $\mathbf{F}(\mathbf{x})'\mathbf{x}$ to be a convex function

over K . For example, for $n = 1$, $K = [0, \infty)$, and $F(x) = -e^{-x} + 1$, $F(x)$ is concave and monotone over K , but $F(x) \cdot x$ is concave for $x \geq 2$.

We summarize the above discussion of the affine case in the following corollary.

Corollary 3. *Suppose that K is the non-empty polyhedron given by (4), and that $\mathbf{F}(\mathbf{x}) = \mathbf{G}\mathbf{x} + \mathbf{h}$, with $\mathbf{G} \geq 0$, but not necessarily symmetric. Then, system (5) is a set of quadratic constraints defining a convex feasible region and MP (8) is a convex LCQP.*

Similarly, for the setting of $VI(K, \mathbf{F})$ with $K \subseteq \mathbb{S}_+^n$, we have the following analogous result.

Corollary 4. *Suppose that K is given by (12), and that*

$$\mathbf{F}(\mathbf{X}) = \sum_{i=1}^n \sum_{j=1}^n X_{ij} \mathbf{G}^{ij} + \mathbf{H},$$

where $\mathbf{H} \in \mathbb{M}^{n,n}$, and $\mathbf{G}^{ij} \in \mathbb{M}^{n,n}$, for $i, j \in \{1, \dots, n\}$. Suppose that $\mathcal{G} \geq 0$, where

$$\mathcal{G} = \left[\text{vec}(\mathbf{G}^{11}), \text{vec}(\mathbf{G}^{12}), \dots, \text{vec}(\mathbf{G}^{1n}), \dots, \text{vec}(\mathbf{G}^{n1}), \dots, \text{vec}(\mathbf{G}^{nn}) \right],$$

and $\text{vec}(\mathbf{G}^{ij})$ denotes the column vector obtained by stacking the row vectors of the matrix \mathbf{G}^{ij} one on top of another. Then, MP (13) is convex, with a quadratic objective.

Proof. The result follows, since $\mathbf{F}(\mathbf{X}) = \text{vec}(\mathbf{X})'\mathcal{G} \text{vec}(\mathbf{X}) + \text{vec}(\mathbf{H})$. \square

Our VI reformulation sheds additional light on VI complexity analysis, but, more importantly, it extends the applicability of commercial, single-level optimization solvers to a larger class of VIs. In contrast, the VI literature has focused on developing specialized iterative algorithms for these problems. Since commercial optimization packages for single-level problems are available, supported, and refined through practice to a greater extent than are solvers specific to VIs, we believe that our reformulation has practical applicability.

2.4. Reformulation of an MPEC as a single-level optimization problem

Let $SOL(K, \mathbf{F})$ denote the solution set of $VI(K, \mathbf{F})$. Using this notation, the MPEC is the following optimization problem whose constraints include a parameterized VI:

$$\begin{aligned} \min_{\mathbf{u} \in \mathbb{R}^{n_1}, \mathbf{x} \in \mathbb{R}^{n_2}} \quad & g(\mathbf{u}, \mathbf{x}) \\ \text{s.t.} \quad & (\mathbf{u}, \mathbf{x}) \in S \\ & \mathbf{x} \in SOL(K(\mathbf{u}), \mathbf{F}(\mathbf{u}; \cdot)). \end{aligned} \tag{15}$$

$K(\mathbf{u})$ denotes a feasible region parameterized by the vector \mathbf{u} of so-called upper-level decision variables, and $\mathbf{F}(\mathbf{u}; \cdot)$ denotes a function, which is also parameterized by \mathbf{u} and whose values are elements in \mathbb{R}^{n_2} . Analogously, \mathbf{x} is called the vector of lower-level decision variables. In this section, we extend the VI reformulation results from Section 2.1 to MPECs.

In general, the exact value of \mathbf{u} may determine not only the coefficients in the constraints defining $K(\mathbf{u})$ but also the number of constraints. Let us focus on instances in which \mathbf{u} affects only the coefficients in the constraints, i.e., cases in which the following condition holds.

Condition 1. $\forall \mathbf{u} \in \mathbb{R}^{n_1}$ such that $\exists \mathbf{x} \in \mathbb{R}^{n_2}$ with $(\mathbf{u}, \mathbf{x}) \in S$, $\exists m < \infty$ for which $K(\mathbf{u})$ is a nonempty polyhedron given in standard form by

$$K(\mathbf{u}) = \{\mathbf{x} \in \mathbb{R}^{n_2} \mid [\mathbf{A}(\mathbf{u})]\mathbf{x} = \mathbf{b}(\mathbf{u}), \mathbf{x} \geq 0\},$$

with $[\mathbf{A}(\mathbf{u})] \in \mathbb{R}^{m \times n_2}$ and $\mathbf{b}(\mathbf{u}) \in \mathbb{R}^m$.

Theorem 1 yields the following corollary.

Corollary 5. Consider MPEC (15) satisfying Condition 1. Then, $(\mathbf{u}^*, \mathbf{x}^*)$ is an optimal solution of MPEC (15) iff $\exists \boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that $(\mathbf{u}, \mathbf{x}, \boldsymbol{\lambda}) = (\mathbf{u}^*, \mathbf{x}^*, \boldsymbol{\lambda}^*)$ is an optimal solution of the following problem:

$$\begin{aligned} \min_{\mathbf{u}, \mathbf{x}, \boldsymbol{\lambda}} \quad & g(\mathbf{u}, \mathbf{x}) \\ \text{s.t.} \quad & (\mathbf{u}, \mathbf{x}) \in S \\ & \mathbf{F}(\mathbf{u}; \mathbf{x})' \mathbf{x} = \mathbf{b}(\mathbf{u})' \boldsymbol{\lambda} \\ & [\mathbf{A}(\mathbf{u})]\mathbf{x} = \mathbf{b}(\mathbf{u}) \\ & \mathbf{x} \geq 0 \\ & [\mathbf{A}(\mathbf{u})]' \boldsymbol{\lambda} \leq \mathbf{F}(\mathbf{u}; \mathbf{x}). \end{aligned} \tag{16}$$

The fact that the MPEC with polyhedral $K(\mathbf{u})$ may be converted into a single-level optimization problem is well known in the literature (see, e.g., [20]). The classical single-level MPEC reformulation is KKT-based, in that it dictates replacement of the VI constraint with its KKT conditions. In contrast, the derivation method used in Corollary 5 and in the proof of Theorem 1 is based on LP duality. Especially when $K(\mathbf{u})$ is not, without augmentation of the lower-level space of variables, in standard form, the duality-based reformulation may be appealing. In particular, although the two approaches are closely related, unlike its KKT-based analog, the duality-based reformulation does not give rise to complementarity constraints in the resulting single-level optimization equivalent of the MPEC. In some settings, this lack of complementarity constraints may facilitate identification of convexity of the MPEC reformulation and may provide greater flexibility for further manipulation of this reformulation.

In general, certifying the convexity of MPEC reformulation (16) is a bit more involved than doing the same for VI reformulation (5). Specifically, if $K(\mathbf{u})$ and $F(\mathbf{u}; \cdot)$ truly depend on \mathbf{u} , in general, the reformulation will be non-convex. For example, if the dependence on \mathbf{u} is linear, then reformulation (16) will contain terms involving products of the upper- and lower-level decision variables. Otherwise, in the simple case in which $K(\mathbf{u})$ and $F(\mathbf{u}; \cdot)$ are constant with respect to all feasible \mathbf{u} , reformulation (16) is convex if S is convex, if the upper-level objective g is convex, and if, $\forall \mathbf{u}$ feasible for the MPEC, $K(\mathbf{u})$ and $\mathbf{F}(\mathbf{u}; \cdot)$ satisfy the conditions of Theorem 3.

Lastly, we note that, since the MPEC is an optimization problem with a parameterized VI constraint, robust optimization subsumes the MPEC, just as it encompasses the VI.

2.5. Examples admitting reformulation as convex programs

Returning to the topic of VIs, in this section, we illustrate our VI reformulation results with examples of $VI(K, \mathbf{F})$ satisfying the conditions of Theorem 3. Recall that these conditions are sufficient for the convexity of the VI reformulations given in Section 2.1. As noted in that section, if $\mathbf{F}(\mathbf{x})$ is affine, then $VI(K, \mathbf{F})$ is always equivalent to an LCQP, even if $J\mathbf{F}(\mathbf{x})$ is

asymmetric. Moreover, this LCQP is convex as long as $\mathbf{F}(\mathbf{x})$ is monotone. Since it is obvious that there exist such instances involving monotone $\mathbf{F}(\mathbf{x})$ with asymmetric $J\mathbf{F}(\mathbf{x})$, we give examples only of instances with nonlinear $\mathbf{F}(\mathbf{x})$.

Example 1. Let $K \subseteq \mathbb{R}_+^n$ be a bounded polyhedron. Let $\mathbf{F}(\mathbf{x}) = \mathbf{H}(\mathbf{x}) + \mathbf{G}\mathbf{x} + \mathbf{h}$, where \mathbf{G} is asymmetric, $H_j(\mathbf{x}) = \ln x_j$, $j = 1, \dots, n$, $\mathbf{G} \succeq 0$. Because \mathbf{F} is continuous and K is compact and convex, this VI problem is guaranteed to possess a solution (see, e.g., [12, Corollary 2.2.5]).

Example 2. Consider the same setting as Example 1, but with $H_j(\mathbf{x}) = \sqrt{x_j}$, $j = 1, \dots, n$.

Example 3. Finally, consider any $K \subseteq \mathbb{R}_+^2 \setminus \{(0, 0)\}$ such that $(\frac{1}{3}, \frac{1}{3}) \in K$. Let

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} \ln(2x_1 + x_2) \\ \ln(x_1 + 2x_2) \end{pmatrix}.$$

Clearly, $(\frac{1}{3}, \frac{1}{3})$ is a solution of this VI. By inspection, it is not obvious that this example satisfies the sufficient conditions of Theorem 3. However, one may symbolically compute, e.g., using MAPLE, the eigenvalues of the Hessians of $F_1(\mathbf{x})$, $F_2(\mathbf{x})$, and $\mathbf{F}(\mathbf{x})' \mathbf{x}$. In doing so, it becomes clear by inspection that, because $K \subseteq \mathbb{R}_+^2$, the signs of these eigenvalues over K imply the concavity of $F_j(\mathbf{x})$, $j \in \{1, 2\}$, and the convexity of $\mathbf{F}(\mathbf{x})' \mathbf{x}$ over K (see, e.g., [6, Proposition B.4]).

2.6. Connection with the classical gap function

In this section, we consider a connection between our approach and the merit-function approach, the latter of which induces bilevel programming reformulations of VIs. Recall the classical primal gap function,

$$\theta_{\text{gap}}(\mathbf{x}) \triangleq \sup_{\mathbf{y} \in K} \mathbf{F}(\mathbf{x})'(\mathbf{x} - \mathbf{y}).$$

Despite its simplicity, the VI literature on merit functions has deprecated the use of this gap function in practice, because it may not be differentiable. Until 1989, the VI community regarded as an open question the issue of whether there exists a continuously differentiable optimization problem that equivalently reformulates $VI(K, \mathbf{F})$. In that year Auchmuty [2], and

independently and soon thereafter Fukushima [14], proposed the regularized gap function, which altered the classical gap function in a way that guaranteed continuous differentiability.

Interestingly, for K given by (4), our reformulation of $VI(K, \mathbf{F})$ as equivalent optimization problem (8) is in fact equivalent to $\min_{\mathbf{x} \in K} \theta_{\text{gap}}(\mathbf{x})$. To see why, let $K_D(\mathbf{x})$ denote the feasible region of $DLP(\mathbf{x})$, as given by (7). Indeed, since K is closed,

$$\begin{aligned} \theta_{\text{gap}}(\mathbf{x}) &= \mathbf{F}(\mathbf{x})' \mathbf{x} - \min_{\mathbf{y} \in K} \mathbf{F}(\mathbf{x})' \mathbf{y} \\ &= \mathbf{F}(\mathbf{x})' \mathbf{x} - \max_{\lambda \in K_D(\mathbf{x})} \mathbf{b}' \lambda \\ &= \min_{\lambda \in K_D(\mathbf{x})} [\mathbf{F}(\mathbf{x})' \mathbf{x} - \mathbf{b}' \lambda]. \end{aligned}$$

Recall that the objective and constraint functions of reformulation (8) are continuously differentiable as many times as $\mathbf{F}(\mathbf{x})$ is. Accordingly, the classical primal gap function does in fact induce a continuously differentiable merit-type function, but it is one that is defined over a modestly higher-dimensional space.

Acknowledgments

We would like to thank Rob Freund for enlightening discussions of this material. We are also grateful to the reviewers of this paper for their insightful comments. This research was partially supported by the PECASE Award DMI-9984339 from the National Science Foundation, by a National Science Foundation Graduate Research Fellowship and by the Singapore-MIT Alliance.

References

- [1] M. Aghassi, Robust optimization, game theory, and variational inequalities, Ph.D. Dissertation, Operations Research Center, MIT, Cambridge, 2005.
- [2] G. Auchmuty, Variational principles for variational inequalities, *Numer. Funct. Anal. Optim.* 10 (1989) 863–874.
- [3] A. Ben-Tal, A. Nemirovski, Robust solutions of uncertain linear programs, *Oper. Res. Lett.* 25 (1999) 1–13.
- [4] A. Ben-Tal, A. Nemirovski, Robust solutions of linear programming problems contaminated with uncertain data, *Math. Programming Ser. A* 88 (2000) 411–424.
- [5] A. Ben-Tal, A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS/SIAM Series on Optimization, SIAM, Philadelphia, 2001.

- [6] D. Bertsekas, *Nonlinear Programming*, second ed., Athena Scientific, Belmont, 1995.
- [7] D. Bertsimas, D. Pachamanova, M. Sim, Robust linear optimization under general norms, *Oper. Res. Lett.* 32 (2004) 510–516.
- [8] D. Bertsimas, M. Sim, The price of robustness, *Oper. Res.* 52 (2004) 35–53.
- [9] D. Bertsimas, J. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont, 1997.
- [10] B.C. Eaves, On the basic theorem of complementarity, *Math. Programming* 1 (1971) 68–75.
- [11] L. El Ghaoui, H. Lebret, Robust solutions to least-squares problems with uncertain data, *SIAM J. Matrix Anal. Appl.* 18 (1997) 1035–1064.
- [12] F. Facchinei, J. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, 2003.
- [13] M. Ferris, C. Kanzow, Complementarity and related problems: a survey, in: P. Pardalos, M. Resende (Eds.), *Handbook of Applied Optimization*, Oxford UP, New York, 2002, pp. 514–530.
- [14] M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, *Math. Programming* 53 (1992) 99–110.
- [15] J. Hammond, Solving asymmetric variational inequality problems and systems of equations with generalized nonlinear programming algorithms, Ph.D. Dissertation, Department of Mathematics, MIT, Cambridge, 1984.
- [16] P. Harker, J. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications, *Math. Programming, Ser. B* 48 (1990) 161–220.
- [17] P. Harker, B. Xiao, A polynomial-time algorithm for affine variational inequalities, *Appl. Math. Lett.* 4 (1991) 31–34.
- [18] P. Hartman, G. Stampacchia, On some nonlinear elliptic differential functional equations, *Acta Math.* 115 (1966) 271–310.
- [19] J. Lions, G. Stampacchia, Variational inequalities, *Commun. Pure Appl. Math.* 20 (1967) 493–519.
- [20] Z. Luo, J. Pang, D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge UP, New York, 1996.
- [21] H. Lüthi, On the solution of variational inequalities by the ellipsoid method, *Math. Oper. Res.* 10 (1985) 515–522.
- [22] T. Magnanti, G. Perakis, A unifying geometric solution framework and complexity analysis for variational inequalities, *Math. Programming* 71 (1995) 327–351.
- [23] O. Mancino, G. Stampacchia, Convex programming and variational inequalities, *J. Optim. Theory Appl.* 9 (1972) 3–23.
- [24] M. Patriksson, *Nonlinear programming and variational inequality problems: a unified approach*, *Applied Optimization*, vol. 23, Kluwer, Dordrecht, 1998.
- [25] G. Perakis, Geometric, interior point, and classical methods for solving finite dimensional variational inequality problems, Ph.D. Dissertation, Department of Applied Mathematics, Brown University, Providence, RI, 1992.
- [26] A.L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, *Oper. Res.* 21 (1973) 1154–1157.
- [27] G. Stampacchia, Formes bilineaires coercives sur les ensembles convexes, *Comptes Rendus Academie Sci. Paris* 258 (1964) 4413–4416.
- [28] G. Stampacchia, Variational inequalities, Theory and Applications of Monotone Operators, in: *Proceedings of the NATO Advanced Study Institute, Venice, 1968*, pp. 101–192.
- [29] J. Sun, G. Zhao, Global linear and local quadratic convergence of a long-step adaptive-mode interior point method for some monotone variational inequality problems, *SIAM J. Optim.* 8 (1998) 123–139.
- [30] J. Sun, G. Zhao, Quadratic convergence of a long-step interior-point method for non-linear monotone variational inequality problems, *J. Optim. Theory Appl.* 97 (1998) 471–491.
- [31] P. Tseng, Global linear convergence of a path-following algorithm for some monotone variational inequality problems, *J. Optim. Theory Appl.* 75 (1992) 265–279.
- [32] J. Wu, Interior-point algorithms for some monotone variational inequality problems, Publication 879, Centre de Recherche sur les Transports, Université de Montréal, 1993.
- [33] J. Wu, Long-step primal path-following algorithm for monotone variational inequality problems, *J. Optim. Theory Appl.* 99 (1998) 509–531.
- [34] S. Zuhovickii, R. Poljak, M. Primak, Two methods of search for equilibrium points of n -person concave games, *Soviet Math. Dokl.* 10 (1969) 279–282.