Cahier du GERAD G-2009-74

Z. Coulibaly · D. Orban

An ℓ_1 Elastic Interior-Point Methods for Mathematical Programs with Complementarity Constraints

November 13, 2009

Abstract. We propose an interior-point algorithm based on an elastic formulation of the ℓ_1-penalty merit function for mathematical programs with complementarity constraints. The method generalizes that of Gould, Orban, and Toint (2003) and naturally converges to a strongly stationary point or delivers a certificate of degeneracy without recourse to second-order intermediate solutions. Remarkably, the method allows for a unified treatment of both general, unstructured, and structured degenerate problems, such as problems with complementarity constraints, with no changes to accommodate one class or the other. Numerical results on a standard test set illustrate the efficiency and robustness of the approach.

Contents

1. Introduction

We consider the solution of mathematical programs with complementarity constraints of the form

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\ f(x) \quad \text{subject to}\ c_{\mathcal{E}}(x)=0,\ c_{\mathcal{I}}(x)\geq 0,\ \min\{x_1,x_2\}=0, \tag{MPCC}$$

where $f:\mathbb{R}^n\to\mathbb{R}$, $c_{\mathcal{E}}:\mathbb{R}^n\to\mathbb{R}^{n_{\mathcal{E}}}$ and $c_{\mathcal{I}}:\mathbb{R}^n\to\mathbb{R}^{n_{\mathcal{I}}}$ are twice continuously differentiable, $\mathcal{E}$ and $\mathcal{I}$ are index sets with $n_{\mathcal{E}}$ and $n_{\mathcal{I}}$ elements respectively, and where $x\in\mathbb{R}^n$ is partitioned into $x=(x_0,x_1,x_2)$ with $x_1,x_2\in\mathbb{R}^p$ and $x_0\in\mathbb{R}^{n-2p}$. The last set of constraints in (MPCC), which is understood componentwise, characterizes a mathematical program with complementarity constraints, or MPCC for short. We implicitly assume that there are no other complementarity constraints in the general equality constraints of (MPCC). In practice, complementarity constraints may occur in the more general form $\min\{F_1(x),F_2(x)\}=0$ but it is easy to see that after adding slack variables $s_1=F_1(x)$ and $s_2=F_2(x)$, we recover a problem of the form (MPCC). A difficulty is the lack of differentiability of the complementarity constraints which are often recast as equivalent smooth inequalities and/or equalities, such as

$$x_1\geq 0,\quad x_2\geq 0,\quad \text{and}\quad X_1x_2=0, \tag{1.1}$$

Z. Coulibaly: GERAD and Department of Mathematics and Industrial Engineering, École Polytechnique, Montréal, QC, Canada. E-mail: zoumana.coulibaly@gerad.ca

D. Orban: GERAD and Department of Mathematics and Industrial Engineering, École Polytechnique, Montréal, QC, Canada. E-mail: dominique.orban@gerad.ca

Mathematics Subject Classification (2000): 90C06, 90C20, 90C30, 90C51, 90C53, 90C55, 65F10, 65F50

or

$$x_1 \geq 0, \quad x_2 \geq 0, \quad \text{and} \quad X_1 x_2 \leq 0, \tag{1.2}$$

where we denoted X_1 the $p \times p$ diagonal matrix $\text{diag}(x_1)$. However, it is easy to see that both (1.1) and (1.2) fail to satisfy the Mangasarian and Fromowitz Constraint Qualification, or MFCQ, and therefore, their feasible set admits no strict interior, which precludes usage of efficient interior-point methods. The same happens with dot-product formulations of the form $x_1^T x_2 = 0$ or $x_1^T x_2 \leq 0$. This form of *degeneracy* of MPCCs is nonetheless structured in the sense that their local solutions may still be characterized by the existence of Lagrange multipliers.

A common way to circumvent this lack of qualification is to *regularize* the problem, i.e., enlarge the feasible set by introducing parameters, transforming, e.g., $X_1 x_2 \leq 0$ into $X_1 x_2 \leq t$ as in (Ralph and Wright, 2004; Scholtes, 2001), or (1.2) into $x_1 \geq -\delta_1$, $x_2 \geq -\delta_2$ and $X_1 x_2 \leq \theta$ for some positive user-controlled regularization parameters δ_1, δ_2 and θ as in (DeMiguel et al., 2005). The goal of the regularization is to have the ability to use powerful methods for regular nonlinear programs, such as SQP methods or interior-point methods. However, with it comes the disadvantage that general-purpose numerical implementations must possess an *MPCC mode* in which complementarity constraints are explicitly declared or detected and processed appropriately.

We propose using the smooth elastic ℓ_1-penalized reformulations described for general nonlinear programs by Gould et al. (2003). The result is a smooth problem parametrized by a penalty parameter $\nu > 0$ with inequality constraints and bounds only. The attractive feature of this approach is that all feasible points of the latter problem satisfy the MFCQ. The idea is then to combine a primal-dual log-barrier approach with the ℓ_1 penalty and iteratively solve unconstrained problems globalized by a trust-region mechanism. A major advantage over other regularization schemes is that the role of the regularization parameters is played by actual variables of the problem controlled by the optimization method—most often, a variation of Newton's method. The inner problems are solved using an appropriately preconditioned trust-region method. Following Conn et al. (2000), gradients are measured using the corresponding dual norm and this allows for early truncation in the subproblems and promotes fast convergence. An advantage over the ℓ_∞-elastic method of Anitescu et al. (2007) is that there is no need to compute an approximate second-order point at each iteration.

In the next sections, we study the elastic problem corresponding to (MPCC) when the complementarity constraints are reformulated as (1.1) or (1.2). As in all exact penalty methods, the penalty parameter remains bounded provided (MPCC) satisfies an appropriate constraint qualification—in this case, the MPCC-MFCQ. Global convergence occurs in one of several forms. In the first, the penalty parameter is updated finitely many times and all limit points of the sequence of iterates are *strongly stationary*—the strongest form in the hierarchy of stationarity concepts appropriate for (MPCC). In the second, the problem is found to be locally infeasible, the penalty parameter diverges and all limit points are stationary for the infeasibility measure. In the third and last form, the penalty parameter diverges yet there are feasible limit points. In this case, upon adequately updating the penalty parameter, the algorithm delivers a certificate of failure of the MPCC-MFCQ. This certifies that, as MPCCs go, the one being solved is degenerate.

As a result, the ℓ_1-elastic algorithm of this paper may be seen as a general-purpose method for degenerate programming, where *degeneracy* is understood as a lack of a sufficient constraint qualification. Whether this degeneracy is *structured* as in the case of (MPCC) or *unstructured*, as in the case of a general nonlinear program with local solutions where the MFCQ fails to hold, the method presented here, when it converges to a feasible point, is able to either identify a local solution or a feasible point violating the relevant basic constraint qualification condition. In our experience, the latter situation still produces a solution to the problem in practice—only one for which no Lagrange multipliers exist.

The rest of this paper is organized as follows. We recall constraint qualification and stationarity concepts relevant to MPCCs in §1.3. In §2, we detail our S-ℓ_1-QP elastic approach and associated regularity properties and first-order optimality conditions. We also establish a correspondence between strongly stationary points of an MPCC and KKT points of the associated elastic problem. In §3, we apply a primal-dual interior-point method to the elastic problem and formulate an elastic algorithm. Global convergence properties are given in §4. Numerical results on a test set from the MacMPEC collection (Leyffer, 2004) are presented and summarized in §5. Finally, we draw some conclusions and perspectives in §6 and give the complete numerical results in Appendix A.

1.1. Related Research

The literature on numerical methods for MPCCs is rather extensive and we cannot possibly cite all pertinent references here. We will cite references most relevant to our approach and refer the interested reader to references therein.

Anitescu (2005) uses an ℓ_∞ penalty function to formulate an elastic problem subsequently treated by an SQP method. Benson et al. (2006) and Leyffer et al. (2006) use penalty approach coupled with an interior-point method. More specifically, the latter push the complementarity constraints into the objective by adding a term of the form $\nu\|X_1 x_2\|_1$, leaving only the bounds $x_1 \geq 0$ and $x_2 \geq 0$ in the constraints. Their approach is related to the one in the present paper although it is specific to MPCCs and does not directly generalize to unstructured degenerate problems. They illustrate the important need for adaptive strategies when updating the penalty parameter.

As already mentioned, Ralph and Wright (2004) and Scholtes (2001) relax $X_1 x_2 \leq 0$ to $X_1 x_2 \leq t$ where $t > 0$ is a user-controlled parameter. Raghunathan and Biegler (2005) take on a similar approach, adapting an interior-point solver to suit this relaxation in order to adress the fact that the strict interior of the relaxed feasible set is empty in the limit. DeMiguel et al. (2005) propose a two-sided relaxation and use an interior-point method. Global and local convergence results are derived under a second order optimality condition.

1.2. Notation

Throughout the paper, we use the following standard notation for diagonal matrices. For any vector x, the capital letter X denotes the matrix $\operatorname{diag}(x)$. Whenever $x \in \mathbb{R}^n$ is decomposed as (x_0, x_1, x_2), we denote x_{1i} the ith component of x_1 and x_{2i} the ith component of x_2. We denote $e_\mathcal{E}$ and $e_\mathcal{I}$ the vectors of ones of $\mathbb{R}^{n_\mathcal{E}}$ and $\mathbb{R}^{n_\mathcal{I}}$, respectively. Unless otherwise noted, $\|\cdot\|$ is used to denote the Euclidean norm. For any $a \in \mathbb{R}$, we denote $a^- = \min\{0, a\}$ and $a^+ = \max\{0, a\}$ the negative and positive parts of a, respectively. Similarly if $a \in \mathbb{R}^n$, the ith component of the vector a^- is defined as $a_i^- \equiv \max\{0, -a_i\}$. In particular, $\|a^-\|_1 = \sum_i \max\{0, -a_i\}$.

1.3. Assumptions and Basic Results

Throughout the paper, our main assumption is that f, $c_\mathcal{E}$ and $c_\mathcal{I}$ are twice-continuously differentiable.

Consider a given generic nonlinear program

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \ \ f(x) \quad \text{subject to } \ c_\mathcal{E}(x) = 0, \ c_\mathcal{I}(x) \geq 0, \tag{NLP}$$

which may or may not include constraints of the form (1.1) or (1.2) and let x be feasible for (NLP). For the inequality constraints of (MPCC) or those of (NLP), let $\mathcal{A}(x)$ be the set of active indices at x, i.e.,

$$\mathcal{A}(x) = \{i \in \mathcal{I} \mid c_i(x) = 0\}.$$

The Lagrangian associated to (NLP) is

$$(x, \lambda_\mathcal{E}, \lambda_\mathcal{I}) \mapsto f(x) - \lambda_\mathcal{E}^T c_\mathcal{E}(x) - \lambda_\mathcal{I}^T c_\mathcal{I}(x),$$

where $\lambda_\mathcal{E} \in \mathbb{R}^{n_\mathcal{E}}$ and $\lambda_\mathcal{I} \in \mathbb{R}^{n_\mathcal{I}}$ are vectors of Lagrange multipliers. The first-order necessary conditions for optimality, or Karush-Kuhn-Tucker (KKT) conditions, at $x \in \mathbb{R}^n$ are that there exists Lagrange multipliers such that

$$\nabla f(x) - J_\mathcal{E}(x)^T\lambda_\mathcal{E} - J_\mathcal{I}(x)^T\lambda_\mathcal{I} = 0, \quad c_\mathcal{E}(x) = 0, \quad C_\mathcal{I}(x)\lambda_\mathcal{I} = 0, \quad \text{and} \quad (c_\mathcal{I}(x), \lambda_\mathcal{I}) \geq 0.$$

Existence of Lagrange multipliers is only guaranteed under a *constraint qualification* condition. The *Mangasarian and Fromovitz Constraint Qualification* condition holds at x if and only if the vectors $\nabla c_i(x)$ for $i \in \mathcal{E}$ are linearly independent and if there exists $d \neq 0$ such that $\nabla c_i(x)^T d = 0$ for all $i \in \mathcal{E}$ and $\nabla c_i(x)^T d > 0$ for all $i \in \mathcal{A}(x)$. For short, we refer to the latter condition as the MFCQ. It is well known that the MFCQ is satisfied at a solution x of (NLP) if and only if the set of optimal Lagrange multipliers associated to x is nonempty and compact (Gauvin, 1977). A stronger condition is the *strict MFCQ* which requires that

$$\{\nabla c_i(x) \mid i \in \mathcal{A}(x),\ \lambda_i > 0\} \cup \{\nabla c_i(x) \mid i \in \mathcal{E}\}$$

be a set of linearly independent vectors, and that there exist $d \neq 0$ such that $\nabla c_i(x)^T d = 0$ for all $i \in \mathcal{E}$ and

$$\begin{aligned} \nabla c_i(x)^T d > 0 &\quad \text{for all } i \in \mathcal{A}(x),\ \lambda_i = 0, \\ \nabla c_i(x)^T d = 0 &\quad \text{for all } i \in \mathcal{A}(x),\ \lambda_i > 0. \end{aligned}$$

It is possible to show that the SMFCQ is a necessary and sufficient condition for the existence and uniqueness of Lagrange multipliers associated to a solution x (Kyparisis, 1985). Finally, a much stronger condition is the *Linear Independence Constraint Qualification* condition, or LICQ, which requires that

$$\{\nabla c_i(x) \mid i \in \mathcal{A}(x)\} \cup \{\nabla c_i(x) \mid i \in \mathcal{E}\}$$

be a set of linearly independent vectors. The LICQ is sufficient for existence and uniqueness of Lagrange multipliers, but is stronger than the SMFCQ.

We do not assume that any constraint qualification holds for (MPCC). In fact, it is easy to verify that the usual MFCQ fails to hold at any point satisfying (1.1) or (1.2). As the following results affirm, it is a set of weaker conditions that are relevant to MPCCs. We do not assume either that (MPCC) is feasible—we wish to be able to detect local infeasibility and identify a stationary point of the infeasibility measure.

The MPCC Lagrangian (Scheel and Scholtes, 2000) associated to (MPCC) is defined as

$$L(x, \alpha, \lambda, z) = \alpha f(x) - \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x) - \lambda_{\mathcal{I}}^T c_{\mathcal{I}}(x) - z_1^T x_1 - z_2^T x_2, \tag{1.3}$$

where $\alpha \geq 0$, $\lambda_{\mathcal{E}} \in \mathbb{R}^{n_{\mathcal{E}}}$, $\lambda_{\mathcal{I}} \in \mathbb{R}^{n_{\mathcal{I}}}_+$, $z_1 \in \mathbb{R}^p_+$ and $z_2 \in \mathbb{R}^p_+$ are Fritz-John multipliers. Whenever $\alpha > 0$ and upon dividing all other multipliers by α, the latter are referred to as Lagrange multipliers. To any feasible x of (MPCC), we associate the active sets

$$\mathcal{A}_1(x) = \{i = 1, \dots, p \mid x_{1i} = 0\}, \quad \text{and} \quad \mathcal{A}_2(x) = \{i = 1, \dots, p \mid x_{2i} = 0\}. \tag{1.4}$$

For brevity and when the context is sufficiently clear, we will simply write $\mathcal{A}_1$ and $\mathcal{A}_2$ instead of $\mathcal{A}_1(x)$ and $\mathcal{A}_2(x)$. Note that by construction, $\mathcal{A}_1 \cup \mathcal{A}_2 = \{1, \dots, p\}$. The set $\mathcal{A}_1 \cap \mathcal{A}_2$ is the set of *biactive* or *corner* variables while its complement is the set of *branch* variables. We let $n_{\mathcal{C}} = n_{\mathcal{E}} + n_{\mathcal{I}} + 2p$ be the total number of constraints.

We now recall the most useful qualification concepts for (MPCC). We say that the MPCC-MFCQ, MPCC-SFMCQ or MPCC-LICQ holds at a feasible point x for (MPCC) if and only if the usual MFCQ, SFMCQ or LICQ holds at x for the set of constraints

$$c_{\mathcal{E}}(x) = 0, \quad c_{\mathcal{I}}(x) \geq 0, \quad x_1 \geq 0, \quad x_2 \geq 0.$$

Note that the relevant active set in those conditions is $\mathcal{A}(x) \cup \mathcal{A}_1 \cup \mathcal{A}_2$. We will say that (MPCC) is *degenerate* if it possesses a solution at which the MPCC-MFCQ is violated. A feature of the algorithm proposed in the next sections will be the ability to converge to such degenerate solutions while at the same time delivering a certificate of failure of the MPCC-MFCQ.

The first result states the form of the Fritz-John conditions for (MPCC), which are weaker than the KKT conditions but hold regardless of whether or not a constraint qualification condition is satisfied.

Lemma 1.1 (**Scheel and Scholtes, 2000**). *Let x be a solution of (MPCC). Then there exist non vanishing Fritz-John multipliers $(\alpha, \lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}, z_1, z_2)$ such that*

$$\alpha \nabla f(x) - J_{\mathcal{E}}(x)^T \lambda_{\mathcal{E}} - J_{\mathcal{I}}(x)^T \lambda_{\mathcal{I}} - \begin{bmatrix} 0 \\ z_1 \\ z_2 \end{bmatrix} = 0, \tag{1.5a}$$
$$X_1 z_1 = 0, \tag{1.5b}$$
$$X_2 z_2 = 0, \tag{1.5c}$$
$$C_{\mathcal{I}}(x) \lambda_{\mathcal{I}} = 0, \tag{1.5d}$$
$$X_1 x_2 = 0, \tag{1.5e}$$
$$z_{1i} z_{2i} \geq 0, \quad i \in \mathcal{A}_1 \cap \mathcal{A}_2 \tag{1.5f}$$
$$c_{\mathcal{E}}(x) = 0, \tag{1.5g}$$
$$(x_1, x_2) \geq 0, \tag{1.5h}$$
$$(c_{\mathcal{I}}(x), \lambda_{\mathcal{I}}) \geq 0. \tag{1.5i}$$

In Lemma 1.1, note that the left-hand side of (1.5a) is the gradient of the Lagrangian (1.3) with respect to x, (1.5b)–(1.5d) are the complementarity conditions, (1.5e) together with (1.5h) are equivalent to the complementarity constraint of (MPCC), and that (1.5g) and (1.5i) impose feasibility and non negativity of the Lagrange multipliers associated to inequality constraints. The unusual condition is (1.5f) which imposes that the multipliers associated to *corner* variables have the same sign.

In (1.5), if $\alpha > 0$, we say that x is Clarke-stationary, or C-stationary for short. The next result says that when a constraint qualification is satisfied, then $\alpha > 0$ and we can at least hope for a C-stationary point.

Theorem 1.1 (**Scheel and Scholtes, 2000; Ralph and Wright, 2004**). *Let x be a solution of (MPCC) where the MPCC-MFCQ is satisfied. Then there exists a compact set $\Lambda \subseteq \mathbb{R}^{nc}$ of Lagrange multipliers such that for all $(\lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}, z_1, z_2) \in \Lambda$, (1.5) are satisfied with $\alpha = 1$.*

Under stronger constraint qualifications, not only is there a unique choice of Lagrange multipliers, but the stationarity conditions also become more demanding.

Theorem 1.2 (**Scheel and Scholtes, 2000**). *Let x be a solution of (MPCC) where the MPCC-SMFCQ is satisfied. Then there exists a unique vector of Lagrange multipliers $(\lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}, z_1, z_2)$ such that (1.5) are satisfied with $\alpha = 1$ and with (1.5f) replaced by*

$$(z_{1i}, z_{2i}) \geq 0, \quad \text{for all } i \in \mathcal{A}_1 \cap \mathcal{A}_2. \tag{1.6}$$

If x satisfies the conclusions of Theorem 1.2, it is said to be *strongly stationary*. The multipliers associated to the corner variables now must be nonnegative. In essence, this implies that multipliers associated to branch variables are unsigned and that the constraints on those variables are, for all practical purposes, acting as equality constraints. Of course, it may happen that x is strongly stationary even if the MPCC-SMFCQ fails to hold at x. Recall also that the MPCC-LICQ implies the MPCC-SMFCQ and therefore that Theorem 1.2 also holds under this stronger assumption.

We now introduce the following qualification condition which will be useful in proving that the algorithm of §2 may converge to feasible points violating the MPCC-MFCQ.

Definition 1.1 (MPCC-BCQ). *Let x be feasible for (MPCC). We say that the basic constraint qualification holds for (MPCC) if and only if the only vector $(\lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}, z_1, z_2) \in \mathbb{R}^{nc}$ to satisfy $\lambda_i \geq 0$ for $i \in \mathcal{A}(x)$, $z_{1i} \geq 0$ for $i \in \mathcal{A}_1(x)$, $z_{2i} \geq 0$ for $i \in \mathcal{A}_2(x)$, and*

$$J_{\mathcal{E}}(x)^T \lambda_{\mathcal{E}} + \sum_{i \in \mathcal{A}(x)} \lambda_i \nabla c_i(x) + \begin{bmatrix} 0 \\ z_1 \\ z_2 \end{bmatrix} = 0, \tag{1.7}$$

is $(\lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}, z_1, z_2) = (0, 0, 0, 0)$.

Note that in conjunction with (1.5d) and (1.5i), setting $\alpha = 0$ in (1.5a) yields (1.7). The following well known lemma from Motzkin will be instrumental for the production of a degeneracy certificate. See e.g., (Mangasarian, 1994) for this and other alternative theorems.

Lemma 1.2 (Motzkin's Alternative Theorem). *Let A and C be given matrices, with A being nonvacuous. Then either*

1. *$Ad > 0$, $Cd = 0$ has a solution d, or*
2. *$A^T\lambda_{\text{A}} + C^T\lambda_{\text{C}} = 0$ has a solution $(\lambda_{\text{A}}, \lambda_{\text{C}})$ such that $\lambda_{\text{A}} \geq 0$ with $\lambda_{\text{A}} \neq 0$,*

but never both.

We now establish the following consequence of Lemma 1.2 connecting the MPCC-BCQ and the MPCC-MFCQ.

Lemma 1.3. *Let x be feasible for (MPCC). Then the MPCC-BCQ is satisfied at x if and only if the MPCC-MFCQ is satisfied at x.*

Proof. Define $J_{\mathcal{A}}(x)$ to be the $|\mathcal{A}(x)| \times n$ matrix whose rows are those of $J_{\mathcal{I}}(x)$ with index in $\mathcal{A}(x)$, i.e., the vectors $\nabla c_i(x)$ for $i \in \mathcal{A}(x)$. Similarly, define $\lambda_{\mathcal{A}}$ to be the sub-vector of $\lambda_{\mathcal{I}}$ corresponding to indices in $\mathcal{A}(x)$. Let E, A and C be the $((|\mathcal{A}_1| + |\mathcal{A}_2|) \times n)$, $((|\mathcal{A}(x)| + |\mathcal{A}_1| + |\mathcal{A}_2|) \times n)$ and $(n_{\mathcal{E}} \times n)$ matrices

$$E \equiv \begin{bmatrix} 0 & E_{\mathcal{A}_1} & 0 \\ 0 & 0 & E_{\mathcal{A}_2} \end{bmatrix}, \quad A \equiv \begin{bmatrix} J_{\mathcal{A}}(x) \\ E \end{bmatrix} \quad \text{and} \quad C \equiv J_{\mathcal{E}}(x),$$

where the rows of $E_{\mathcal{A}_1}$ are those of the $(p \times p)$ identity matrix with indices in $\mathcal{A}_1$, and where we define $E_{\mathcal{A}_2}$ similarly. In Lemma 1.2, identify $\lambda_{\text{A}} \equiv (\lambda_{\mathcal{A}}, z_1, z_2)$ and $\lambda_{\text{C}} \equiv \lambda_{\mathcal{E}}$. With this notation, (1.7) may be rewritten $A^T\lambda_{\text{A}} + C^T\lambda_{\text{C}} = 0$ while the MPCC-MFCQ reads $Ad > 0$ and $Cd = 0$ with the additional requirement that C be full row rank. The conclusion follows from a straightforward application of Lemma 1.2. □

2. The S-ℓ_1-QP Elastic Approach

We follow the approach suggested by Gould et al. (2003) and apply an ℓ_1-penalty approach to any smooth reformulation of (MPCC). Since the resulting penalty problem is nondifferentiable on the boundary of the feasible set, it is in turn reformulated in terms of *elastic variables*. For the purpose of illustration, we concentrate on the formulation (1.1) in the rest of this paper, but keep in mind that all results will equally apply, with appropriate modifications, to any other smooth reformulation of the complementarity constraints.

The ℓ_1-penalty problem associated to

$$\begin{array}{lll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) & \\ \text{subject to} & c_{\mathcal{E}}(x) = 0, & c_{\mathcal{I}}(x) \geq 0, \\ & X_1 x_2 = 0, & (x_1, x_2) \geq 0, \end{array} \tag{2.1}$$

is

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \; f(x) + \nu\|c_{\mathcal{E}}(x)\|_1 + \nu\|c_{\mathcal{I}}^-(x)\|_1 + \nu\|X_1 x_2\|_1 + \nu\|x_1^-\|_1 + \nu\|x_2^-\|_1, \tag{2.2}$$

where $\nu > 0$ is a penalty parameter.

There are several ways to cast (2.2) as an elastic problem (Gould et al., 2003). Our analysis concentrates on the following:

$$\begin{array}{lll} \underset{x \in \mathbb{R}^n,\, s \in \mathbb{R}^{n_C}}{\text{minimize}} & \phi^{\text{P}}(x, s; \nu) & \\ \text{subject to} & c_{\mathcal{E}}(x) + s_{\mathcal{E}} \geq 0, & s_{\mathcal{E}} \geq 0, \\ & c_{\mathcal{I}}(x) + s_{\mathcal{I}} \geq 0, & s_{\mathcal{I}} \geq 0, \\ & x_1 + s_1 \geq 0, & s_1 \geq 0, \\ & x_2 + s_2 \geq 0, & s_2 \geq 0, \\ & X_1 x_2 + s_3 \geq 0, & s_3 \geq 0, \end{array} \tag{2.3}$$

where

$$\phi^{\mathrm{P}}(x,s;\nu) = f(x) + \nu\sum_{i\in\mathcal{E}}(c_i(x)+2s_i) + \nu\sum_{i\in\mathcal{I}} s_i + \nu\sum_{i=1}^{p}(x_{1i}x_{2i} + 2s_{3i} + s_{1i} + s_{2i}), \tag{2.4}$$

and where $s = (s_{\mathcal{E}}, s_{\mathcal{I}}, s_1, s_2, s_3)$ is the vector of *elastic variables*. Again we keep in mind that appropriate reformulations of all of our results will continue to hold for other elastic forms of (2.2). We refer to (2.3) as the *elastic problem*. Note that in the latter problem, the constraints $x_1 \geq 0$ and $x_2 \geq 0$ are treated as any other inequality constraint and $X_1x_2 = 0$ is treated as any other equality constraint. In other words, the formulation (2.3) does not treat complementarity constraints in any special manner and because of this, the algorithm developed in the next sections is suited to any sort of degenerate problem, whether arising from the reformulation of an MPCC or not. We come back to this point in §6.

The attractive feature of (2.3) is that the standard MFCQ is satisfied at *all* feasible point (x,s) (Gould et al., 2003, Theorem 2.2). Therefore, and as is apparent from inspection of the constraints, the relative interior of the feasible set is nonempty and primal-dual interior-point methods thus seem like natural candidates for approximately solving the elastic problem.

The first-order optimality conditions of (2.3) may be written

$$\nabla f(x) - J_{\mathcal{E}}(x)^T(y_{\mathcal{E}} - \nu e_{\mathcal{E}}) - J_{\mathcal{I}}(x)^T y_{\mathcal{I}} - \begin{bmatrix} 0 \\ X_2y_3 - \nu x_2 + y_1 \\ X_1y_3 - \nu x_1 + y_2 \end{bmatrix} = 0, \tag{2.5a}$$

$$\begin{bmatrix} \nu e_{\mathcal{E}} - (y_{\mathcal{E}} - \nu e_{\mathcal{E}}) - u_{\mathcal{E}} \\ \nu e_{\mathcal{I}} - y_{\mathcal{I}} - u_{\mathcal{I}} \end{bmatrix} = 0, \tag{2.5b}$$

$$\begin{bmatrix} \nu e - y_1 - u_1 \\ \nu e - y_2 - u_2 \\ \nu e - (y_3 - \nu e) - u_3 \end{bmatrix} = 0, \tag{2.5c}$$

$$(C_{\mathcal{E}}(x) + S_{\mathcal{E}})y_{\mathcal{E}} = 0, \tag{2.5d}$$

$$(C_{\mathcal{I}}(x) + S_{\mathcal{I}})y_{\mathcal{I}} = 0, \tag{2.5e}$$

$$(X_1 + S_1)y_1 = 0, \tag{2.5f}$$

$$(X_2 + S_2)y_2 = 0, \tag{2.5g}$$

$$(X_1X_2 + S_3)y_3 = 0, \tag{2.5h}$$

$$Su = 0, \tag{2.5i}$$

$$(c_{\mathcal{E}}(x) + s_{\mathcal{E}}, c_{\mathcal{I}}(x) + s_{\mathcal{I}}, x_1 + s_1, x_2 + s_2, X_1x_2 + s_3, s) \geq 0, \tag{2.5j}$$

$$(y,u) \geq 0, \tag{2.5k}$$

where $u = (u_{\mathcal{E}}, u_{\mathcal{I}}, u_1, u_2, u_3)$ is the vector of Lagrange multipliers associated to the bound constraints on the elastic variables and $y = (y_{\mathcal{E}}, y_{\mathcal{I}}, y_1, y_2, y_3)$ is the vector of Lagrange multipliers associated to the other inequalities. In order to avoid any confusion, multipliers for (2.3) will be denoted y and u while multipliers for (MPCC) will remain λ and z.

Note that (2.5a)–(2.5c) are the gradient with respect to x and s of the Lagrangian

$$\begin{aligned}\mathcal{L}(x,s,y,u;\nu) = \phi^{\mathrm{P}}(x,s;\nu) &- y_{\mathcal{E}}^T(c_{\mathcal{E}}(x) + s_{\mathcal{E}}) - y_{\mathcal{I}}^T(c_{\mathcal{I}}(x) + s_{\mathcal{I}}) \\ &- y_1^T(x_1 + s_1) - y_2^T(x_2 + s_2) - y_3^T(X_1x_2 + s_3) - u^Ts.\end{aligned}$$

The following properties are derived from (Gould et al., 2003) and generalized to the case of complementarity constraints.

Theorem 2.1. *If (x,s,y,u) is a KKT point of (2.3) for some $\nu > 0$ and if x is feasible for (MPCC) then $s = 0$, and x is strongly stationary for (MPCC) with multipliers*

$$(\lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}, z_1, z_2) = (y_{\mathcal{E}} - \nu e_{\mathcal{E}}, y_{\mathcal{I}}, X_2(y_3 - \nu e) + y_1, X_1(y_3 - \nu e) + y_2). \tag{2.6}$$

Proof. The definition of $\lambda_{\mathcal{E}}$ and $\lambda_{\mathcal{I}}$ is covered in (Gould et al., 2003, Theorem 2.3). We restrict our attention to z_1 and z_2.

To show that $s = 0$, we distinguish two cases. Assume first that $x_{ki} = 0$ for some $k \in \{1, 2\}$. If $s_{ki} > 0$, we deduce from (2.5i) that $u_{ki} = 0$, and with (2.5c), we get $y_{ki} = \nu > 0$. But (2.5f)-(2.5h) yield $s_{ki} = 0$, which is a contradiction. Thus $s_{ki} = 0$. If on the other hand $x_{ki} > 0$, (2.5f)-(2.5h) imply $y_{ki} = 0$. This combines with (2.5c) to give $u_{ki} = \nu > 0$. But then (2.5i) implies that $s_{ki} = 0$. Consequently $s_1 = s_2 = 0$. We show similarly that $s_3 = 0$ by noticing that $X_1 x_2 = 0$.

By feasibility of x, we have $X_2 x_1 = X_2 x_1 = 0$, and therefore the definition of z given in (2.6) implies (1.5a). Since $s = 0$, we deduce from (2.5f)-(2.5h) that $X_1 y_1 = X_2 y_2 = X_1 X_2 y_3 = 0$. Hence, $X_1 z_1 = X_1 y_1 - \nu X_1 x_2 + X_1 X_2 y_3 = 0$ and $X_2 z_2 = X_2 y_2 - \nu X_2 x_1 + X_1 X_2 y_3 = 0$, so that (1.5b)-(1.5c) are satisfied. From (2.5j), we deduce (1.5h). Finally, for all bi-actives indices $i \in \mathcal{A}_1 \cap \mathcal{A}_2$, (2.6) yields $(z_{1i}, z_{2i}) = (y_{1i}, y_{2i})$, and from (2.5k), we deduce that $(z_{1i}, z_{2i}) \geq 0$ for all $i \in \mathcal{A}_1 \cap \mathcal{A}_2$. Therefore x is strongly stationary for (MPCC). □

Theorem 2.2. *Assume x is strongly stationary for (MPCC) and associated to finite values (λ, z) such that (1.5) is satisfied with $\alpha = 1$ and with (1.5f) replaced by (1.6). Set*

$$\bar{\nu} = \max \left[\|\lambda\|_\infty, \max_{i \in \mathcal{A}_1 \setminus \mathcal{A}_2} \left\{ -\frac{z_{1i}}{x_{2i}} \;\middle|\; z_{1i} < 0 \right\}, \max_{i \in \mathcal{A}_2 \setminus \mathcal{A}_1} \left\{ -\frac{z_{2i}}{x_{1i}} \;\middle|\; z_{2i} < 0 \right\} \right]. \tag{2.7}$$

Then $(x, 0)$ is a KKT point for (2.3) for all $\nu \geq \bar{\nu}$ with multipliers $y_{\mathcal{E}} = \lambda_{\mathcal{E}} + \nu e_{\mathcal{E}}$, $y_{\mathcal{I}} = \lambda_{\mathcal{I}}$,

$$y_1 \equiv z_1 + \nu x_2, \quad y_2 \equiv z_2 + \nu x_1, \quad y_{3i} \equiv \begin{cases} (z_{1i} + \nu x_{2i})/x_{2i}, & i \in \mathcal{A}_1 \setminus \mathcal{A}_2, \\ (z_{2i} + \nu x_{1i})/x_{1i}, & i \in \mathcal{A}_2 \setminus \mathcal{A}_1, \\ 0, & i \in \mathcal{A}_1 \cap \mathcal{A}_2, \end{cases} \tag{2.8}$$

and

$$u_1 \equiv \nu e - y_1, \quad u_2 \equiv \nu e - y_2, \quad u_3 \equiv 2\nu e - y_3. \tag{2.9}$$

Proof. The proof was established for the general constraints in (Gould et al., 2003, Theorem 2.5) and states that ν must be larger than $\|\lambda\|_\infty$. Here, we restrict ourselves to the complementarity constraints. Since x is strongly stationary, we have $(z_{1i}, z_{2i}) \geq 0$ for all $i \in \mathcal{A}_1 \cap \mathcal{A}_2$. The given multipliers satisfy $(y, u) \geq 0$ for $\nu \geq \bar{\nu}$.

We now show that $(x, 0, y, u)$ is a KKT point for all penalty parameter $\nu \geq \bar{\nu}$. By construction, $X_2 y_3 - \nu x_2 + y_1 = z_1$ and $X_1 y_3 - \nu x_1 + y_2 = z_2$ so that (2.5a) is satisfied. Similarly, (2.5c) is verified. From $s = 0$ we deduce (2.5i) and the feasibility of x along with (2.8) implies (2.5f)–(2.5h). Conditions (2.5k) follow immediately from definitions (2.8) and (2.9). Finally with $s = 0$, (1.5h) yields (2.5j). Hence $(x, 0,\ y,\ u)$ is a KKT point for (2.3) for all penalty parameter $\nu \geq \bar{\nu}$. □

In Theorem 2.2, one among other assumptions that would guarantee finiteness of the multipliers (y, z) is the MPCC-MFCQ. However, there might exist a set of finite multipliers in other circumstances.

3. Primal-Dual Interior-Point Framework

In this section, we apply a primal-dual interior-point method to (2.3) and formulate an algorithm which parallels that of Gould et al. (2003).

The logarithmic barrier problem associated to (2.3) is to

$$\underset{x \in \mathbb{R}^n,\, s \in \mathbb{R}^{n_C}}{\text{minimize}} \quad \phi^{\mathrm{B}}(x, s; \nu, \mu), \tag{3.1}$$

where

$$\begin{aligned} \phi^{\mathrm{B}}(x, s; \nu, \mu) = & \phi^{\mathrm{P}}(x, s; \nu) - \mu \sum_{i \in \mathcal{C}} \log(c_i(x) + s_i) - \mu \sum_{i \in \mathcal{C}} \log(s_i) - \mu \sum_{i=1}^{p} \log(x_{1i} + s_{1i}) - \mu \sum_{i=1}^{p} \log(s_{1i}) \\ & - \mu \sum_{i=1}^{p} \log(x_{2i} + s_{2i}) - \mu \sum_{i=1}^{p} \log(s_{2i}) - \mu \sum_{i=1}^{p} \log(x_{1i} x_{2i} + s_{3i}) - \mu \sum_{i=1}^{p} \log(s_{3i}), \end{aligned}$$

where $\mu > 0$ is the barrier parameter and where $\mathcal{C} = \mathcal{E} \cup \mathcal{I}$. The primal-dual system consists in an equivalent rewriting of the first-order optimality conditions for (3.1). It may be viewed as a perturbation of (2.5) in which the right-hand side of the complementarity conditions (2.5d)–(2.5i) are changed to μe, where e is the vector of ones of appropriate size. In this case, the vectors y and u are the primal-dual estimates of the optimal Lagrange multipliers.

Our prototype algorithm, Algorithm 3.1, consists in an outer and an inner iteration. The role of the inner iteration is to seek an approximate solution to the primal-dual system for fixed values of $\nu > 0$ and $\mu > 0$. Once this is achieved, the outer iteration updates the penalty and barrier parameters.

Algorithm 3.1 Prototype Algorithm—Outer Iteration

Step 0. Let the forcing functions $\epsilon^{\mathrm{D}}(\cdot)$, $\epsilon^{\mathrm{C}}(\cdot)$ and $\epsilon^{\mathrm{U}}(\cdot)$ be given, and let $\kappa_\nu > 0$. Choose $x^0 \in \mathbb{R}^n$, $s^0 \in \mathbb{R}^{nc}_+$ such that $c(x^0) + s^0 > 0$, $x_1^0 + s_1^0 > 0$, $x_2^0 + s_2^0 > 0$ and $X_1^0 x_2^0 + s_3^0 > 0$, initial dual estimates $y^0, u^0 \in \mathbb{R}^{nc}_+$, and penalty and barrier parameters ν^0 and $\mu^0 > 0$, and set $k = 0$.

Step 1. Choose a suitable preconditioning norm $\|.\|_{[P^{k+1}]}$ and find a new primal-dual iterate $w^{k+1} = (x^{k+1}, s^{k+1}, y^{k+1}, u^{k+1})$ satisfying

$$\left\| \nabla_{xs} \mathcal{L}(x^{k+1}, s^{k+1}, y^{k+1}, u^{k+1}; \nu^k) \right\|_{[P^{k+1}]} \le \epsilon^{\mathrm{D}}(\mu^k) \tag{3.2a}$$

$$\left\| \begin{bmatrix} (C_{\mathcal{E}}(x^{k+1}) + S_{\mathcal{E}}^{k+1}) y_{\mathcal{E}}^{k+1} - \mu^k e_{\mathcal{E}} \\ (C_{\mathcal{I}}(x^{k+1}) + S_{\mathcal{I}}^{k+1}) y_{\mathcal{I}}^{k+1} - \mu^k e_{\mathcal{I}} \\ (X_1^{k+1} + S_1^{k+1}) y_1^{k+1} - \mu^k e_p \\ (X_2^{k+1} + S_2^{k+1}) y_2^{k+1} - \mu^k e_p \\ (X_1^{k+1} X_2^{k+1} + S_3^{k+1}) y_3^{k+1} - \mu^k e_p \end{bmatrix} \right\| \le \epsilon^{\mathrm{C}}(\mu^k) \tag{3.2b}$$

$$\left\| S^{k+1} u^{k+1} - \mu^k e \right\| \le \epsilon^{\mathrm{U}}(\mu^k) \tag{3.2c}$$

$$\left(c_{\mathcal{E}}(x^{k+1}) + s_{\mathcal{E}}^{k+1}, c_{\mathcal{I}}(x^{k+1}) + s_{\mathcal{I}}^{k+1}, x_1^{k+1} + s_1^{k+1}, x_2^{k+1} + s_2^{k+1}, X_1^{k+1} x_2^{k+1} + s_3^{k+1}, s^{k+1} \right) > 0 \tag{3.2d}$$

$$\text{and} \quad \left(\nu^k [e + e_{\mathcal{E}}^0] + \kappa_\nu e, \nu^k [e + e_{\mathcal{E}}^0] + \kappa_\nu e \right) \ge \left(y^{k+1}, u^{k+1} \right) > 0 \tag{3.2e}$$

by (for example) approximately solving (3.1).

Step 2. Select a new barrier parameter, $\mu^{k+1} \in (0, \mu^k]$ such that $\lim \mu^k = 0$. If necessary, adjust the penalty parameter, ν^k. Increment k by one, and return to Step 1.

In Algorithm 3.1, *forcing* functions $\mu \mapsto \epsilon(\mu) \ge 0$ are functions defined for $\mu \ge 0$ such that $\epsilon(\mu) = 0$ if and only if $\mu = 0$. The preconditioning norm is defined by $\|r\|_{[P]}^2 \equiv r^T d$, where the vector d solves

$$K(w)d \equiv \begin{bmatrix} P + J(x)^T \Theta(w) J(x) & J(x)^T \Theta(w) \\ \Theta(w) J(x) & \Theta(w) + U S^{-1} \end{bmatrix} \begin{bmatrix} d_x \\ d_s \end{bmatrix} = \begin{bmatrix} r_x \\ r_s \end{bmatrix} \equiv r. \tag{3.3}$$

in which J is the Jacobian matrix of the constraints of (2.3) and

$$\Theta(w) \equiv \begin{bmatrix} Y_1(X_1 + S_1)^{-1} & & \\ & Y_2(X_2 + S_2)^{-1} & \\ & & Y_3(X_1 X_2 + S_3)^{-1} \end{bmatrix}.$$

The matrix P in (3.3) is a suitable preconditioning approximation of the Hessian of the Lagrangian of (2.3), $P \approx \nabla_{xx} \mathcal{L}$. The preconditioner P is chosen such that the matrix K of (3.3) is positive definite (Gould et al., 2003, §3.5).

In principle, any globally convergent algorithm ensuring satisfaction of (3.2a)–(3.2e) in finitely many iterations may be used in the inner iteration. In our implementation, described in §5, we elected to choose the trust-region method of Conn et al. (2000). That the latter is guaranteed to converge and meet the inner iteration requirements is established in Gould et al. (2003).

4. Global Convergence

We now state convergence properties of the sequences generated by Algorithm 3.1. Our results extend those of Gould et al. (2003). We assume that the penalty parameter is updated iteratively according to:

$$\nu^{k+1} = \begin{cases} \max\{\tau_1\nu^k,\ \nu^k+\tau_2\} & \text{if } \min\{\|c_{\mathcal{I}}(x)^-\|,\ \|x_1^-\|,\ \|x_2^-\|\} > \eta_1^k \\ & \text{or } \min\{\|c_{\mathcal{E}}(x)\|,\ \|X_1x_2\|\} > \eta_2^k, \\ \nu^k & \text{otherwise,} \end{cases} \tag{4.1}$$

for some sequences $\{\eta_1^k\}$ and $\{\eta_2^k\}$, converging to zero, and given constants $\tau_1 > 1,\ \tau_2 > 0$. We also assume that the following assumptions are true:

Assumption 4.1. *the functions f, $c_{\mathcal{E}}$ and $c_{\mathcal{I}}$ are twice-continuously differentiable over an open set that contains all iterates encountered. Over this set, the derivatives $\nabla f(x)$, $\nabla_{xx}f(x)$, $\nabla c_i(x)$ and $\nabla_{xx}c_i(x)$ remain uniformly bounded, for all $i \in \mathcal{C}$.*

Assumption 4.2. *Each preconditioning matrix P^k is bounded from above in norm and is such that the smallest eigenvalue of the matrix K of system (3.3) is uniformly positive for all iterates k.*

Assumption 4.3. *The forcing functions ϵ^C, ϵ^D and ϵ^U satisfy*

$$\epsilon^C(\mu) \le \kappa_c\mu, \quad \epsilon^U(\mu) \le \kappa_c\mu, \quad \text{and} \quad \epsilon^D(\mu) \le \kappa_d\mu^{\gamma^k+\frac{1}{2}}, \tag{4.2}$$

for some preset constants $\kappa_c \in (0,1)$ and $\kappa_d > 0$, and a sequence $\{\gamma^k\} > 0$.

Under Assumption 4.3, (Gould et al., 2003, Lemma 4.3) shows that Algorithm 3.1 with update strategy (4.1) generates iterates $w^{k+1} = (x^{k+1}, s^{k+1}, y^{k+1}, u^{k+1})$ such that

$$\|v\| \le \kappa(\nu^k + \kappa_\nu)(\mu^k)^{\gamma^k}, \tag{4.3}$$

for all v satisfying

$$\|v\|_{[P^{k+1}]} \le \epsilon^D(\mu^k).$$

On the other hand, by using the definition of the preconditioning norm $\|\cdot\|_{[P^k]}$, it follows that

$$\|r\|^2_{[P^k]} = r^Td^k \le \|r\|_2\|d^k\|_2 = \|(K^k)^{-1}r\|_2\|r\|_2 \le \frac{1}{\lambda_{\min}(K^k)}\|r\|_2^2 \le \frac{1}{\lambda^*_{\min}}\|r\|_2^2, \tag{4.4}$$

where the definition of $\lambda^*_{\min} > 0$ follows from Assumption 4.2.

Our first convergence result concerns that case where Algorithm 3.1 generates a bounded penalty parameter sequence $\{\nu^k\}$.

Theorem 4.1. *Let $\{v^k\} = \{(x^k, s^k, y^k, u^k)\}$ be an infinite sequence of primal-dual variables generated by Algorithm (3.1). Assume that the penalty parameter $\{\nu^k\}$ reaches its final value ν^* in a finite number of iterations. Then*

1. *the sequences $\{s^k\}$, $\{y^k\}$, and $\{u^k\}$ are bounded,*
2. *if (x^*, s^*, y^*, u^*) is a limit point of $\{v^k\}$, then $s^* = 0$, and x^* is strongly stationary for (MPCC).*

Proof. As in the previous proofs, we need only consider the complementarity constraints. Let also $x_3 \equiv X_1x_2$ to simplify. By assumption, there exists an integer $N > 0$, such that for all $k \ge N$, $\nu^k = \nu^*$. As a consequence, (4.1) implies that for all $k \ge N$,

$$x_1^k \ge 0, \quad x_2^k \ge 0, \quad \text{and} \quad \lim_{k\to\infty} X_1^kx_2^k = 0. \tag{4.5}$$

To show that $\{s^k\}$ is bounded, assume by contradiction that $s_{li}^k \to \infty$ for some $(l,\ i) \in \{1,2,3\} \times \{1,\dots,p\}$. It follows from the properties of forcing functions ϵ^D, ϵ^C and ϵ^U, the fact that $\mu^k \downarrow 0$, and

(3.2c) that $u_{li}^k \to 0$. Therefore, (3.2a) and (4.3) imply that $\{y_{li}^k\}$ is bounded, and from (3.2b), we have $x_{li}^k \to -\infty$, which contradicts (4.5). Thus $\{s^k\}$ is bounded.

Suppose now that the subsequence $\{w^k\}_{\mathcal{K}} \subset \{w^k\}$ is such that $\lim_{k\in\mathcal{K}} w^k = w^* = (x^*, s^*, y^*, u^*)$. We show that w^* is a KKT point for the elastic problem (2.3) and that x^* is feasible for (MPCC). The result will then follow from Theorem 2.1.

By using (4.3), the stopping condition (3.2a) and the fact that $\lim_{k\in\mathcal{K}} \mu^k = 0$,

$$\begin{aligned}\|\nabla_{xs}\mathcal{L}(x^*, s^*, y^*, u^*; \nu^*)\| &= \lim_{k\in\mathcal{K}} \left\|\nabla_{xs}\mathcal{L}(x^{k+1}, s^{k+1}, y^{k+1}, u^{k+1}; \nu^k)\right\| \\ &\le \lim_{k\in\mathcal{K}} \kappa(\nu^k + \kappa_\nu)(\mu^k)^{\gamma^k} = 0. \end{aligned} \tag{4.6}$$

This implies that w^* satisfies (2.5b)–(2.5c). From (3.2b), we obtain $\lim_{k\in\mathcal{K}}(X_l^k + S_l^k)y_l^k = (X_l^* + S_l^*)y_l^* = 0$, for $l = 1, 2, 3$ and (3.2c) implies $S^*u^* = 0$, which yield (2.5f)–(2.5i) for w^*. By using (3.2d), we have $\lim_{k\in\mathcal{K}} (x_l^{k+1} + s_l^{k+1},\ s_l^{k+1}) = (x_l^* + s_l^*, s_l^*) \ge 0$, $l = 1, 2, 3$. From (3.2e), we obtain $\lim_{k\in\mathcal{K}} (y^{k+1},\ u^{k+1}) = (y^*, u^*) \ge 0$. Finally, (4.5) shows that x^* is feasible for (MPCC), which concludes the proof. □

Assume now that the sequence of penalty parameters generated by Algorithm 3.1 is unbounded.

Theorem 4.2. *Under assumptions (4.1)–(4.3), let $w_P^k = (x^k, s^k)$ and $w_D^k = (y^k, u^k)$ be sequences generated by Algorithm 3.1. Assume the penalty parameter ν^k is updated infinitely many times at iterations $k \in \mathcal{K} \subseteq \mathbb{N}$. Then*

1. *the subsequence $\{w_D^k\}_{\mathcal{K}} = \{(y^k, u^k)\}_{\mathcal{K}}$ is unbounded,*
2. *every limit point w_P^* of $\{w_P^k\}$ is stationary for the ℓ_1 infeasibility measure.*

Proof. Along $\mathcal{K}$, the update strategy (4.1) implies $\nu^{k+1} \ge \nu^k + \tau_2$, $\tau_2 > 0$ so that $\{\nu^k\}_{\mathcal{K}} \to \infty$. Since $\{\nu^k\}$ is nondecreasing, it follows that $\{\nu^k\} \to \infty$. Assume $\{(y^k, u^k)\}_{\mathcal{K}}$ is bounded and let (y^*, u^*) be one of its limit points. Reducing to a further subsequence if necessary, we may assume that $\{y^k\}_{\mathcal{K}} \to y^*$ and $\{u^k\}_{\mathcal{K}} \to u^*$, so that for $k \in \mathcal{K}$ sufficiently large, $\|y^k\| \le 2\|y^*\|$ and $\|u^k\| \le 2\|u^*\|$. From the latter bounds, the part of (4.6) corresponding to (2.5c), and the inverse triangle inequality, we deduce:

$$(\sqrt{6p} - \kappa(\mu^{k-1})^{\gamma^{k-1}})\nu^{k-1} \le \|y^k\| + \|u^k\| + \kappa\kappa_\nu(\mu^{k-1})^{\gamma^{k-1}} \le 2(\|y^*\| + \|u^*\|) + \kappa\kappa_\nu(\mu^{k-1})^{\gamma^{k-1}},$$

for all sufficiently large $k \in \mathcal{K}$. For $k \to \infty$, the latter condition implies that $\{\nu^{k-1}\}$ is bounded, which is a contradiction. Therefore $\{w_D^k\}_{\mathcal{K}}$ is unbounded.

Assume now that $\{w_P^k\}_{\mathcal{K}} \to w_P^*$ and define

$$\bar{y}^{k+1} = \frac{y^{k+1}}{\nu^k}, \quad \bar{u}^{k+1} = \frac{u^{k+1}}{\nu^k} \quad \text{and} \quad \bar{\mu}^k = \frac{\mu^k}{\nu^k}. \tag{4.7}$$

The stopping criteria of Algorithm (3.1), scaled by ν^k, read

$$\left\|(\nu^k)^{-1}\nabla_{xs}\mathcal{L}(x^{k+1}, s^{k+1}, \bar{y}^{k+1}, \bar{u}^{k+1}; \nu^k)\right\| \le \kappa_p(\mu^k)^{\gamma^k} \tag{4.8a}$$

$$\left\|\begin{bmatrix}(X_1^{k+1} + S_1^{k+1})\bar{y}_1^{k+1} - \bar{\mu}^k e_p \\ (X_2^{k+1} + S_2^{k+1})\bar{y}_2^{k+1} - \bar{\mu}^k e_p \\ (X_1^{k+1}X_2^{k+1} + S_3^{k+1})\bar{y}_3^{k+1} - \bar{\mu}^k e_p\end{bmatrix}\right\| \le (\nu^k)^{-1}\epsilon^{\mathrm{C}}(\mu^k) \tag{4.8b}$$

$$\left\|S^{k+1}\bar{u}^{k+1} - \bar{\mu}^k e\right\| \le (\nu^k)^{-1}\epsilon^{\mathrm{U}}(\mu^k) \tag{4.8c}$$

$$\left(x_1^{k+1} + s_1^{k+1}, x_2^{k+1} + s_2^{k+1}, X_1^{k+1}x_2^{k+1} + s_3^{k+1}, s^{k+1}\right) > 0 \tag{4.8d}$$

$$\text{and} \quad \left([e + e_{\mathcal{E}}^0] + \kappa_\nu e, [e + e_{\mathcal{E}}^0] + \kappa_\nu e\right) \ge \left(\bar{y}^{k+1}, \bar{u}^{k+1}\right) > 0 \tag{4.8e}$$

where $\kappa_p \equiv (\nu^0)^{-1}\kappa(1 + \kappa_\nu)$. From (4.8e) we deduce that the sequence $\{(\bar{y}^{k+1},\ \bar{u}^{k+1})\}_{\mathcal{K}}$ is bounded. Without loss of generality, $\{(\bar{y}^{k+1}, \bar{u}^{k+1})\}_{\mathcal{K}} \to (y^*, u^*)$. By using facts that $\mu^k \downarrow 0$, $\nu^k \to \infty$ and $(\nu^k)^{-1}\max\{\mu^k, \epsilon^{\mathrm{C}}(\mu^k), \epsilon^{\mathrm{U}}(\mu^k)\} \le (\nu^0)^{-1}\max\{\epsilon^{\mathrm{C}}(\mu^k), \epsilon^{\mathrm{U}}(\mu^k)\} \to 0$, and taking the limit in (4.8), we see that the limit point (x^*, s^*, y^*, u^*) satisfies the KKT first order optimality condition for the ℓ_1 infeasibility measure for (MPCC). □

In order to strengthen the previous result, we now define an additional condition under which the penalty parameter is updated:

$$\nu^{k+1} = \max\{\tau_1\nu^k, \nu^k + \tau_2\} \quad \text{if} \quad \|y^{k+1} - \nu^k e^0_{\mathcal{E},3}\| > \gamma\nu^k, \tag{4.9}$$

where $0 < \gamma < 1$ is a given constant. This condition is used in addition to (4.1). Here, $y^k = (y^k_{\mathcal{E}}, y^k_{\mathcal{I}}, y^k_1, y^k_2, y^k_3)$ denotes the multipliers of *all* constraints and $e^0_{\mathcal{E},3}$ denotes a vector of $\mathbb{R}^{n_c+3p}$ with unit components in those places corresponding to $y_{\mathcal{E}}$ and to y_3 and zero components everywhere else. The shift of ν^k thus only applies to components associated to general equality constraints and the formulation of the complementarity condition as an equality constraint. In this sense, $X_1x_2 = 0$ is treated as any other equality constraint.

Now assume that the penalty parameter ν^k is unbounded due to the updating rule (4.9). The next result states that if a feasible point of (MPCC) is nevertheless approached, it must be a point where the MPCC-MFCQ fails to hold.

Theorem 4.3. *Assume that Assumptions (4.1)–(4.3) are satisfied and let $\{(x^k, s^k)\}$ and $\{(y^k, u^k)\}$ be sequences generated by Algorithm 3.1. Assume the penalty parameter ν^k is updated infinitely many times at iterations $k \in \mathcal{K} \subseteq \mathbb{N}$ because of (4.9). Then*

1. *the subsequence $\{(y^k, u^k)\}_{\mathcal{K}}$ is unbounded,*
2. *if (x^*, s^*) is a limit point of $\{(x^k, s^k)\}$, then x^* is a feasible point of (MPCC) where MPCC-MFCQ fails to hold.*

Proof. From Theorem 4.2, $\{(y^k, u^k)\}$ is unbounded. From (4.9), since $\{\nu^k\}$ is unbounded, we have for all $k \in \mathcal{K}$,

$$\|y^{k+1} - \nu^k e^0_{\mathcal{E},3}\| > \gamma\nu^k \to +\infty.$$

Since the unboundedness of ν^k is not due to failure of the updating rule (4.1), x^* is feasible for (MPCC). Define now $\alpha^k = \|y^{k+1} - \nu^k e^0_{\mathcal{E},3}\|_\infty$ and the scaled sequences

$$\bar{y}^{k+1}_{\mathcal{E}} = \frac{y^{k+1}_{\mathcal{E}} - \nu^k e_{\mathcal{E}}}{\alpha^k}, \quad \bar{y}^{k+1}_{\mathcal{I}} = \frac{y^{k+1}_{\mathcal{I}}}{\alpha^k}, \quad \bar{u}^{k+1} = \frac{u^{k+1}}{\alpha^k}, \quad \text{and} \quad \bar{\nu}^{k+1} = \frac{\nu^k}{\alpha^k}.$$

Similarly, let $\bar{y}^{k+1}_i = y^{k+1}_i/\alpha^k$, $(i = 1, 2)$, and

$$\bar{y}^{k+1}_3 = \frac{y^{k+1}_3 - \nu^k e}{\alpha^k}, \quad \bar{z}^{k+1}_1 = X^{k+1}_2\bar{y}^{k+1}_3 + \bar{y}^{k+1}_1, \quad \text{and} \quad \bar{z}^{k+1}_2 = X^{k+1}_1\bar{y}^{k+1}_3 + \bar{y}^{k+1}_2.$$

It is easy to see that those sequences remain bounded and, without loss of generality, we may thus assume that $\{(\bar{y}^{k+1}, \bar{u}^{k+1}, \bar{z}^{k+1}, \bar{\nu}^{k+1})\}_{\mathcal{K}} \to (\bar{y}, \bar{u}, \bar{z}, \bar{\nu})$. Moreover, by construction, $\|\bar{y}^{k+1}\|_\infty = 1$ for all k. The stopping criterion (3.2a), scaled by α^k, becomes

$$\left\| \frac{1}{\alpha^k}\nabla f(x^{k+1}) - J_{\mathcal{E}}(x^{k+1})^T\bar{y}^{k+1}_{\mathcal{E}} - J_{\mathcal{I}}(x^{k+1})^T\bar{y}^{k+1}_{\mathcal{I}} - \begin{bmatrix} 0 \\ \bar{z}^{k+1}_1 \\ \bar{z}^{k+1}_2 \end{bmatrix} \right\| \le \frac{\kappa(\nu^k + \kappa_\nu)}{\alpha^k}(\mu^k)^{\gamma^k}.$$

Upon taking limits and using (3.2b), it follows that

$$J_{\mathcal{E}}(x^*)\bar{y}_{\mathcal{E}} + J_{\mathcal{A}}(x^*)\bar{y}_{\mathcal{A}} + \begin{bmatrix} 0 \\ \bar{z}_1 \\ \bar{z}_2 \end{bmatrix} = 0.$$

We have shown that (1.7) is satisifed for nonzero multipliers. Because of Lemma 1.3, x^* cannot satisfy the MPCC-MFCQ. □

5. Implementation and Numerical Results

Our implementation is realized in the Python programming language as part of the NLPy toolkit and programming environment of Orban (2009). The modeling of (MPCC) is done in the AMPL modeling language (Fourer et al., 2002). An interface between AMPL and Python by way of the AMPL Solver Library (Gay, 1997) lets users interact with AMPL models from within Python.

Preconditioning systems (3.3) are reformulated as the equivalent but potentially much sparser

$$\begin{bmatrix} P & J(x)^T \\ J(x) & -\Theta^{-1} - U^{-1}S \end{bmatrix} \begin{bmatrix} d_x \\ \xi \end{bmatrix} = \begin{bmatrix} r_x \\ -U^{-1}Sr_s \end{bmatrix}, \tag{5.1}$$

from which d_s may be recovered via

$$d_s = U^{-1}S(r_s - \xi).$$

The coefficient matrix of the latter system becomes, however, indefinite. In our implementation, it is factorized via the Harwell Subroutine Library (2007) subroutine MA57 of Duff (2004). Initially, P is chosen as a band matrix of semi-bandwidth 5 extracted from the Hessian of the Lagrangian $\nabla_{xx}\mathcal{L}(w)$. The values on the diagonal of P are iteratively increased until the coefficient matrix of (3.3) is positive definite, or, equivalent, until the coefficient matrix of (5.1) has precisely $n_{\mathcal{C}}$ negative eigenvalues.

The implementation defines three different penalty parameters. The first, $\nu_{\mathcal{E}}$, is assigned to the terms containing the general equality constraints in (2.4). The second, ν_s, is assigned to the terms containing the elastic variables for general inequality constraints. The third, ν_t, is assigned to the terms containing the elastic variables for bound constraints. This choice is appropriate since it permits to penalize separately each type of constraints and take constraint scaling into consideration. The penalty parameters are initialized by way of one of the three following strategies:

Strategy 1: $\nu^0 = \max\{1, \|\nabla f(x^0)\|_\infty\}$,
Strategy 2: $\nu^0_{s,t} = \|u^0 + y^0\|_\infty$, and $\nu^0_{\mathcal{E}} = \frac{1}{2}\|u^0 + y^0\|_\infty$,
Strategy 3: $\nu^0_{s,t} = \|u^0 + y^0\|_\infty$ and $\nu^0_{\mathcal{E}} = \max\left\{\|\nabla c_i(x^0)\| \mid i \in \mathcal{E}\right\}$.

In these formulae, x^0 is the initial iterate—by default x^0 is as specified in the original model, or zero if no initial guess is specified,—and u^0, y^0 are the Lagrange multipliers associated to x^0. The latter are initialized to vectors of ones unless a different initial guess is specified. The first strategy is inspired by the initialization in SNOPT (Gill et al., 2005). The other two come directly from (2.5b)–(2.5c).

The barrier parameter is initialized and updated using the simple rule $\mu^0 = 5$ and $\mu^{k+1} = \frac{1}{5}\mu^k$. The penalty parameter is updated according to (4.1) and (4.9) where we selected $\tau_1 = 2$, $\tau_2 = 1$, $\eta_1^k = \eta_2^k = \min\{0.2, \max\{10(\mu^k)^{0.4}, 10^{-6}\}\}$, and $\gamma = 0.999$. The power 0.4 of μ^k in the definition of η_1^k and η_2^k is meant to account for constraints that may not be strictly complementary at the solution. Indeed in this case, the distance between an exact solution $w(\mu^k)$ and w^* is proportional to $\sqrt{\mu^k}$ (Wright and Orban, 2002, Corollary 19). The definition of η_1^k and η_2^k thus avoids over-penalization due to such constraints. We imposed a maximum of 18 outer iterations with 800 inner iterations in each. The forcing functions are chosen as

$$\epsilon^{\mathrm{D}}(\mu) = 1.1\min\{\mu, \mu^{1.0001}\}, \quad \epsilon^{\mathrm{C}}(\mu) = \epsilon^{\mathrm{U}}(\mu) = 1.1\mu.$$

Optimality is declared attained as soon as the ℓ_∞-norm of the residual of (1.5) with $\alpha = 1$ and (1.5f) replaced with (1.6) falls below 10^{-6}.

A failure is declared when the maximum number of iterations is reached or when μ attains its smallest allowed value of 10^{-14} without the optimality conditions being satisfied.

We tested Algorithm 3.1 on problems from the MacMPEC collection of Leyffer (2004). The following problems were eliminated because we were not able to solve them in a reasonable time: *bem-milanc30-s*, *incid-set2-32*, *pack-comp1-32*, *pack-comp1c-32*, *pack-comp1p-32*, *pack-comp2-32*, *pack-comp2c-32*, *pack-comp2p-32*, *qpec-200-1*, *qpec-200-2*, *qpec-200-3*, *qpec-200-4*, *siouxfls*. Removing those problems left 129 problems in the test set.

It appears that Strategy 2 gives the best robustness for both the equality and inequality formulations of (MPCC): 85.27% of the test problems were solved successfully for (1.1) and 82.17% for (1.2). For

Strategy 1, these rates are 68.21% and 66.67%, respectively. Strategy 3 yields 73.09% and 65.12%, respectively.

To compare the performance of Algorithm 3.1 on the formulations (1.1) and (1.2), we use the performances profiles of Dolan and Moré (2002). Figures 5.1, 5.2 and 5.3 give performance profiles in terms of CPU time, number of iterations, and number of function evaluations, respectively. In each plot, the solid line corresponds to Strategy 1, the dashed line to Strategy 2 and the dotted line to Strategy 3.

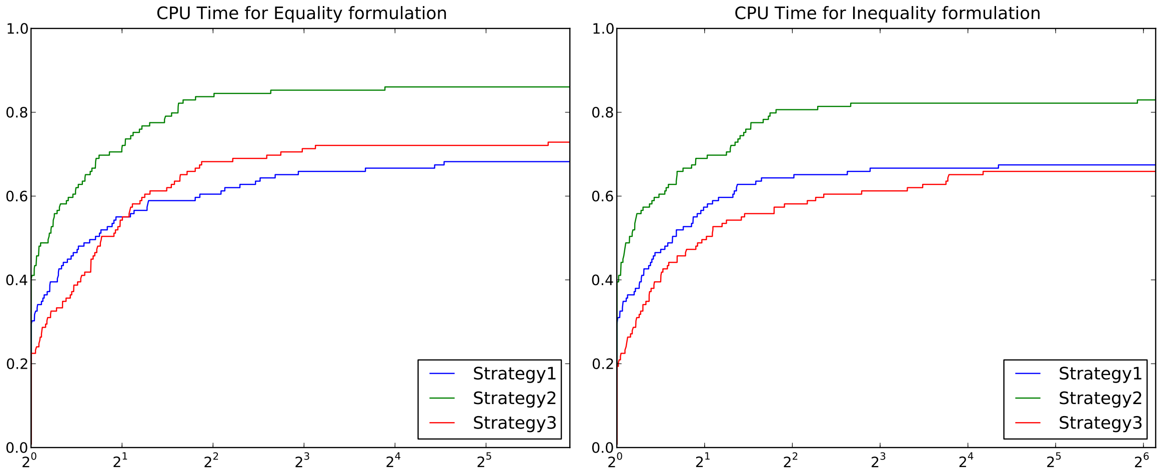

Fig. 5.1. Performance profiles for (1.1) (left) and (1.2) (right) in terms of CPU time.

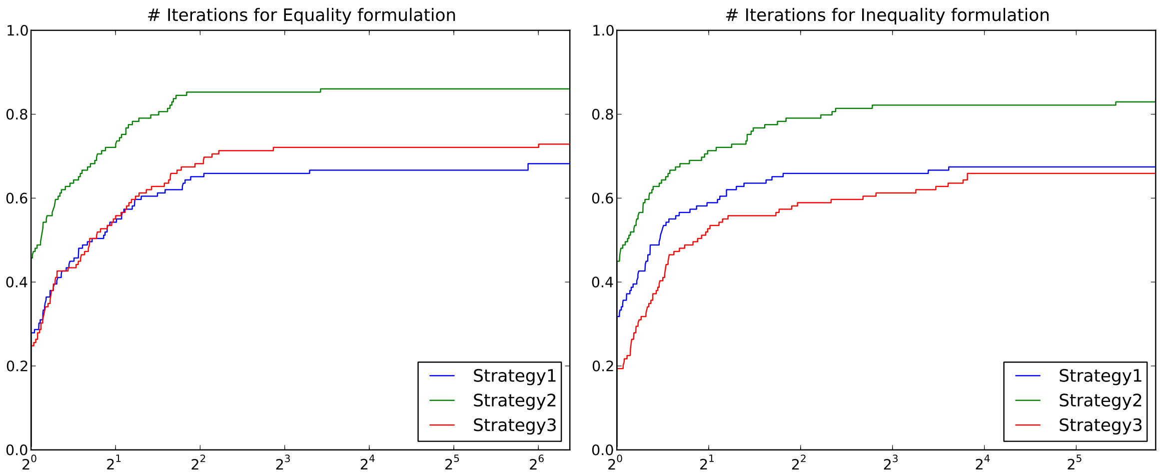

Fig. 5.2. Performance profiles for (1.1) (left) and (1.2) (right) in terms of number of inner iterations.

Globally, for each strategy, the profile in the leftmost part of Figures 5.1, 5.2 and 5.3 is above that in the rightmost part. This seems to indicate that all strategies perform better with the equality-constrained formulation (1.1) than the inequality-constrained formulation (1.2). We report complete numerical results in AppendixA.

For the purpose of comparing our approach with a closely-related one, we ran the method of DeMiguel et al. (2005) on the same set of 129 test problems. The results appear in Fig. 5.4. The relaxation algorithm of DeMiguel et al. (2005) finds a stationarity point on 86 problems among the 129, reaching a convergence rate of 66.67%. We stress however that the CPU time profile should be

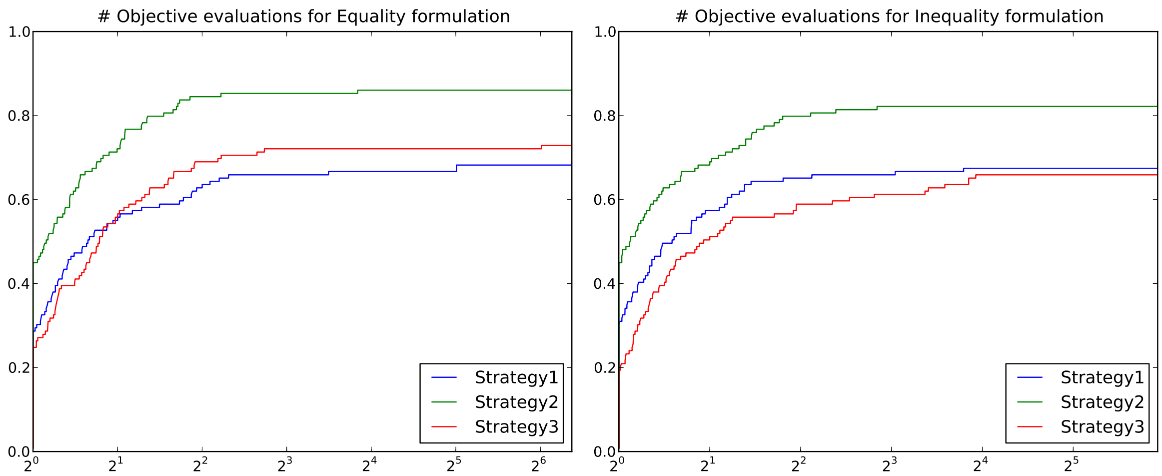

Fig. 5.3. Performance profiles for (1.1) (left) and (1.2) (right) in terms of number of objective evaluations.

taken with a grain of salt because the Matlab implementation of DeMiguel et al. (2005) uses only dense linear algebra.

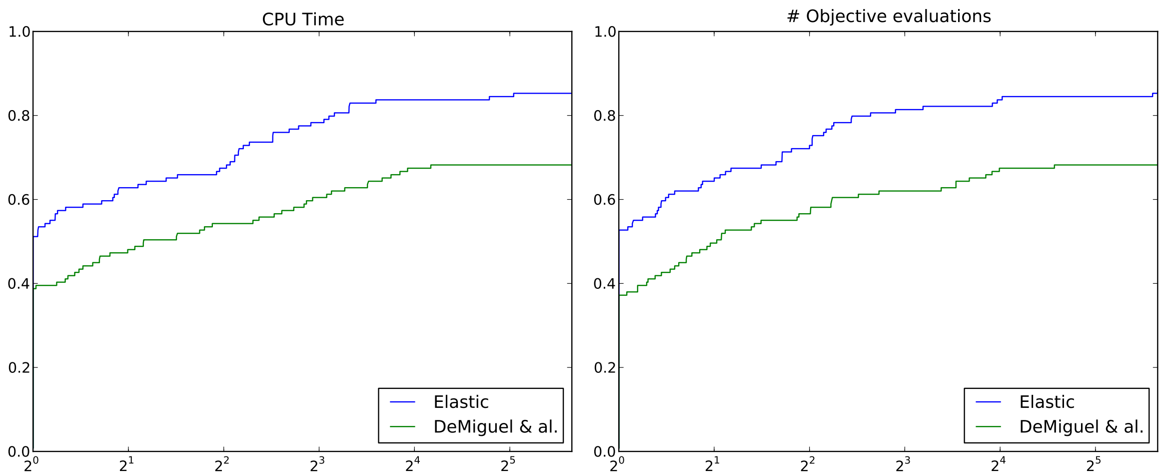

Fig. 5.4. Performances profiles comparing Algorithm 3.1 on (1.1) with initialization Strategy 2 (solid line) with the relaxation of DeMiguel et al. (2005) (dashed line) in terms of CPU time (left) and number of function evaluations (right).

6. Discussion

The method presented in this paper is a general method for nonlinear programming which has attractive features when the problem is degenerate. In particular, MPCCs and general nonlinear programs are treated in the same way. There is therefore no such thing as an “MPEC-mode” in our implementation. From a theoretical point of view, it is important to mention that we do not need to identify iterates satisfying a second-order optimality condition to prove global convergence to strongly stationary points. It is very encouraging to notice that this preliminary implementation is able to solve over 85% of the problems in the MacMPEC collection of Leyffer (2004). Increasing numerical robustness by developping robust initializations of penalty and barrier parameters will be investigated in the future. Dynamic updates of the penalty parameter will deserve special attention, since they can significantly increase robustness, as reported by Leyffer et al. (2006). They consist in occasional decreases in the penalty parameters under certain circumstances. Finally, the elastic approach presented here can equally be formulated as an "implicit elastic" strategy in the spirit of (Gould et al., 2003).

We defer the local convergence analysis of Algorithm 3.1 for (MPCC) to future research because the elastic problem (2.3) does not satisfy the LICQ even if (MPCC) satisfies the MPCC-LICQ. Consider for example the simple problem

$$\begin{array}{ll} \underset{x_1,x_2}{\text{minimize}} & f(x_1,x_2) = x_1 + x_2, \\ \text{subject to} & \min\{x_1,x_2\} = 0, \end{array} \tag{6.1}$$

which has the solution $x^* = (0,0)$. The MPCC-LICQ is obviously satisfied at x^* but the elastic problem associated to (6.1) is

$$\begin{array}{lll} \underset{x_1,x_2,s}{\text{minimize}} & x_1 + x_2 + \nu(s_1 + s_2 + x_1x_2 + 2s_3), & \\ \text{subject to} & x_1 + s_1 \geq 0, & s_1 \geq 0, \\ & x_2 + s_2 \geq 0, & s_2 \geq 0, \\ & x_1x_2 + s_3 \geq 0, & s_3 \geq 0. \end{array} \tag{6.2}$$

All constraints of (6.2) are active at $(x^*, s^*) = (0,0,0,0)$ but it is easy to see that their Jacobian matrix at (x^*, s^*) cannot have full row rank.

It can be shown that if MPCC-LICQ holds at x^* for (MPCC) and if $\mathcal{A}_1 \cap \mathcal{A}_2(x^*) = \emptyset$, the LICQ holds at $(x^*, s^*) = (0,0)$ for (2.3). Clearly, this assumption is unreasonable. Instead, we believe that fast local convergence can be shown to take place as follows. Assuming the penalty parameter remains finite, Theorem 4.1 applies. Upon denoting ν^* the final value of the penalty parameter and x^* a limit point of $\{x^k\}$, Algorithm 3.1 identifies x^* by solving (2.3) with $\nu = \nu^*$. This latter problem satisfies the MFCQ without any assumption on (MPCC). It is not difficult to show that if (MPCC) satisfies the following standard strict complementarity assumption

$$\lambda_i^* \neq 0 \text{ for all } i \in \mathcal{A}(x^*), \text{and } (z_{1i}^*, z_{2i}^*) > 0 \text{ for all } i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*),$$

then (2.3) satisfies the usual strict complementarity assumption at $(x^*, 0)$. Similarly, a standard second-order sufficiency assumption on (MPCC) translates into the usual strong second-order sufficiency on (2.3). We believe that it is then possible to utilize the results of Wright and Orban (2002) to show that the rate of convergence of $\{w^k\}$ to w^* is the same as the rate of convergence of $\{\mu^k\}$ to zero. An advantage of this approach over other local convergence analyses, such as those of DeMiguel et al. (2005) and Leyffer et al. (2006), is that it dispenses with the assumption that (MPCC) satisfies a constraint qualification at x^*. Because such analysis would apply not only to the case of MPCCs but also to the general nonlinear programs tackled in Gould et al. (2003) or the mathematical programs with vanishing constraints of Curatolo and Orban (2009), we defer it to a subsequent report.

A. Detailed Results

We provide below the detailed results of Algorithm 3.1 with Strategy 2 on all problems from our test set. For convenience, we separate the results in three tables. Table A.1 lists all problems solved to optimality and for which our method finds a final objective value (nearly) identical to that reported in Leyffer (2004). Table A.2 lists all problems solved to optimality and for which our method finds a final objective value different from that reported by Leyffer (2004). Table A.3 lists all problems on which our method failed to identify a stationary point within the limits imposed. The columns in each table report, from left to right, the problem name, the number of variables, the CPU time spent looking for a solution, the "official" optimal objective value as reported in Leyffer (2004), the final objective value found by our implementation, the final dual feasibility residual, the largest final elastic value, the final infeasibility measure, the number of inner iterations, the largest Lagrange multiplier for (MPCC) in absolute value, the final complementarity measure and the largest final penalty parameter.

Table A.1: Problems solved to optimality for formulation (1.1).

Name	#vars	time	f_L^*	$f(x^*)$	$\|\nabla L\|_\infty$	$\|(s,t)\|_\infty$	infeas.	# iter	$\|y\|_\infty$	$\max\{s_i t_i\}$	$\nu_{\max}$
bard1	11	0.62	17	17	1.3e−07	5.7e−09	3.3e−09	48	1.8e+02	5.2e−07	180
bard3	8	8.23	-12.6787	-12.6787	2.6e−07	4.5e−11	1.0e−11	1629	3.1e+02	4.1e−09	246
bard3m	10	1.30	-12.6787	-12.6787	7.4e−08	1.7e−08	2.5e−07	167	4.8e+03	6.5e−07	4800
bilevel1m	10	4.17	-60	-10	4.0e−09	1.1e−09	2.0e−12	411	6.4e+02	1.0e−07	642
bilevel2	32	10.94	-6600	-6600	1.1e−07	1.0e−12	1.9e−13	324	4.2e+03	8.2e−10	4320
bilevel2m	20	5.51	-6600	-6.6e+03	3.7e−09	3.0e−12	4.6e−13	282	1.3e+03	8.2e−10	1440
bilevel3	12	9.6	-12.6787	-12.6787	8.3e−07	6.8e−10	2.1e−10	1180	1.6e+02	2.0e−08	160
desilva	8	0.90	-1	-1.	3.6e−07	1.1e−09	4.5e−11	150	1.8e+02	1.0e−07	184
df1	3	0.21	0	5.08e−09	9.9e−09	2.1e−09	5.1e−09	46	1.8e+02	1.5e−07	180
ex911	13	1.86	-13	-1.3e+01	5.2e−07	2.3e−10	5.1e−11	132	9.9e+02	2.0e−08	640
ex912	10	1.81	-6.25	-6.25	4.0e−09	1.1e−09	4.7e−11	83	7.3e+02	1.0e−07	640
ex914	10	1.01	-37	-3.7e+01	2.1e−08	5.2e−11	9.1e−11	106	4.8e+03	4.1e−09	1e+4
ex915	13	0.32	-1	-1.0e−00	6.6e−08	1.6e−09	2.5e−09	30	1.8e+02	1.0e−07	180
ex918	14	0.34	-3.25	-3.25e+00	5.5e−07	1.4e−09	4.4-08	37	4.8e+03	7.5e−07	4800
ex919	12	2.86	3.11111	3.11e+00	1.7e−09	4.3e−10	7.3e−10	394	1.4e+04	1.0e−07	4e+4
ex9110	14	0.33	-3.25	-3.25e+00	5.6e−07	1.4e−09	4.4e−08	37	4.8e+03	7.4e−07	4800
ex924	8	0.36	.5	5.0e−01	5.6e−08	4.6e−11	6.2e−11	31	1.8e+02	4.1e−09	180
ex925	8	0.91	9	9.0e+00	3.2e−09	2.3e−10	9.7e−12	90	7.3e+02	2.0e−08	640
ex926	16	0.51	-1	-1.0e−00	7.7e−08	1.1e−09	2.0e−09	21	1.8e+02	1.0e−07	180
ex928	6	0.21	1.5	1.5e+00	1.1e−07	2.3e−10	3.1e−10	34	1.9e+02	2.0e−08	188
ex929	9	0.62	2	2.0e+00	2.1e−07	2.4e−10	2.1e−10	75	1.8e+02	2.1e−08	180
flp2	6	0.44	0	3.31e−13	4.8e−07	1.2e−08	7.6e−09	47	1.2e+02	5.1e−07	80
flp4-1	110	25.8	0	1.58120e−06	7.6e−09	1.7e−08	2.4e−08	94	6.0e+01	5.1e−07	60
flp4-2	170	99.05	0	1.33969e−06	1.3e−08	1.7e−08	2.4e−08	117	6.0e+01	5.1e−07	60
flp4-3	210	168.4	0	1.76035e−06	3.9e−08	1.7e−08	2.4e−08	146	6.0e+01	5.1e−07	60
flp4-4	300	1424.96	0	1.09878e−07	1.8e−07	6.8e−10	9.8e−10	155	6.0e+01	2.1e−08	60
gauvin	5	0.20	20	2.0e+01	1.1e−08	1.2e−09	5.7e−10	41	1.8e+02	1.0e−07	180
gnash10	17	2.12	-230.823	-230.823	1.4e−07	5.7e−09	1.2e−09	203	1.8e+02	5.1e−07	180
gnash11	17	1.42	-129.912	-129.912	1.4e−08	2.3e−10	3.6e−11	111	1.8e+02	2.1e−08	180
gnash12	17	1.27	-36.9331	-3.69331e+01	2.4e−08	2.3e−10	2.3e−11	57	1.8e+02	2.1e−08	180
gnash13	17	1.53	-7.06178	-7.06178e+00	1.1e−08	4.6e−11	7.1e−12	93	1.8e+02	4.1e−09	180
gnash14	17	1.78	-.179046	-1.79046e−01	1.1e−08	2.3e−10	5.8e−11	108	1.8e+02	2.0e−08	180
gnash16	17	3.11	-241.442	-2.41442e+02	6.6e−09	4.6e−11	4.4e−11	174	4.8e+02	4.1e−09	480
gnash17	17	1.79	-90.7491	-9.07491e+01	4.0e−12	1.1e−09	1.0e−09	146	4.8e+02	1.0e−07	480
gnashm10	13	1.52	-230.823	-2.30823e+02	1.62e−09	1.71e−08	1.24e−08	135	6.0e+01	5.12e−07	200
gnashm11	13	0.36	-129.912	-1.29912e+02	3.39e−08	6.83e−10	3.420e−10	40	6.0e+01	2.05e−08	124
gnashm12	13	0.33	-36.9331	-3.69331e+01	2.177e−08	6.83e−10	1.662e−10	37	6.0e+01	2.05e−08	60
gnashm13	13	0.81	-7.06178	-7.06178e+00	1.74e−08	6.83e−10	6.84e−11	47	6.0e+01	2.05e−08	60
gnashm14	13	0.66	-.179046	-1.79046e−01	1.38e−12	3.41e−09	6.28e−11	48	6.0e+01	1.02e−07	60
incid-set1-8	166	38.16	3.816e−17	2.36e−05	2.8e−07	5.4e−08	1.9e−08	214	2.0e+01	5.1e−07	20
incid-set1c-8	166	32.50	3.816e−17	2.36e−05	7.5e−07	5.4e−08	2.1e−08	220	2.0e+01	5.2e−07	20
incid-set2-8	166	802.52	4.518e−3	4.518e−03	1.0e−08	9.38e−11	5.6e−11	5030	2.0e+01	8.2e−10	20
incid-set2c-8	166	284.24	5.471e−3	5.478e−03	2.9e−10	1.17e−08	6.6e−09	1841	2.0e+01	1.0e−07	20
jr1	3	0.04	.5	5.0e−01	1.0e−07	5.7e−08	6.1e−08	11	2.0e+01	5.1e−07	20
jr2	3	0.04	.5	5.0e−01	5.6e−08	4.7e−08	1.1e−07	12	2.0e+01	5.1e−07	20
kth1	2	0.04	0	9.7e−07	1.7e−08	5.1e−08	2.4e−13	12	1.9e+01	4.9e−07	20
kth2	3	0.05	0	-3.51e−09	8.3e−07	5.2e−08	3.5e−09	14	2.0e+01	5.1e−07	20
kth3	3 2	0.047993	.5	.5	8.4e−09	2.04e−08	4.6e−08	11	3.3e+01	5.1e−07	33
liswet1-050	202	6.95	1.399e−02	1.4e−02	7.2e−07	3.7e−09	5.5e−09	47	6.0e+01	1.0e−07	60
liswet1-100	402	25.91	1.373e−02	1.37e−02	3.4e−08	7.9e−10	1.1e−09	59	60	2.0e−08	60
liswet1-200	802	112.29	1.701e−02	1.70e−02	5.2e−07	4.96e−09	5.4e−09	74	60	1.0e−07	60
nash1	8	0.15	7.8861e−30	2.097e−13	2.3e−07	1.7e−08	2.3e−08	27	6.0e+01	5.1e−07	60
outrata32	5	5.39	3.4494	3.4494	6.2e−09	5.8e−09	8.1e−09	764	1.8e+02	5.1e−07	181
outrata33	5	3.25	4.60425	4.60425	2.2e−11	1.2e−09	1.8e−09	528	1.8e+02	1.0e−07	181
outrata34	5	3.60	6.59268	6.59268	1.0e−09	5.7e−09	9.4e−09	586	1.8e+02	5.1e−07	181
pack-comp1-8	156	39.11	.6	.6	1.63e−07	2.5e−08	1.9e−08	287	2.2e+01	1.0e−07	22
pack-comp1c-8	156	36.50	.6	.6	3.95e−07	2.6e−08	2.0e−08	268	2.2e+01	1.1e−07	22
pack-comp2c-8	156	115.93	.673458	6.73e−01	2.7e−11	2.0e−09	1.6e−09	758	4.8e+03	2.0e−08	4800
pack-rig1-4	17	0.21	7.19e−01	7.19e−01	6.3e−07	5.1e−08	3.4e−08	19	2.2e+01	5.3e−07	22
pack-rig1-8	123	8.47	.787932	7.87934e−01	3.2e−07	1.07e−08	2.0e−08	83	2.2e+01	1.0e−07	22
pack-rig1c-4	17	0.21	7.21e−01	7.21e−01	2.4e−07	5.1e−08	3.1e−08	18	2.2e+01	5.2e−07	22
pack-rig1c-8	123	2.85	.7883	.7883	1.3e−08	1.1e−08	2.0e−08	31	2.2e+01	1.0e−07	22
pack-rig1p-4	32	0.94	6.0e−01	6.0e−01	1.1e−08	5.1e−08	8.9e−08	48	2.2e+01	5.2e−07	22
pack-rig2-4	17	0.21	6.95e−01	6.95e−01	1.6e−08	1.02e−08	1.4e−08	19	2.2e+01	1.0e−07	22
pack-rig2-8	123	17.75	.780404	.780406	7.78e−09	1.14e−09	2.0e−09	163	1.4e+04	1.0e−07	1e+4
pack-rig2c-4	17	0.28	7.12e−01	7.12e−01	6.6e−07	5.12e−08	8.7e−08	25	2.2e+01	5.1e−07	22
pack-rig2c-8	123	42.20	.799306	.799306	3.0e−09	4.9e−11	7.9e−11	459	1.4e+04	4.1e−09	1e+4
pack-rig2p-4	17	1.67	6.0e−01	6.0e−01	1.1e−08	1.0e−08	2.4e−08	68	2.2e+01	1.4e−07	22
portfl-i-1	87	1.99	1.502e−5	3.16e−05	6.2e−07	5.1e−08	9.7e−08	28	2.0e+01	5.6e−07	20
portfl-i-2	87	2.00	1.526e−05	3.15e−05	8.1e−07	5.1e−08	9.5e−08	27	2.0e+01	5.7e−07	20
portfl-i-3	87	1.97	6.265e−6	1.44e−05	8.3e−07	1.0e−08	1.8e−08	26	2.0e+01	3.0e−07	20
portfl-i-4	87	2.05	2.177e−6	5.67e−06	7.1e−08	1.0e−08	1.8e−08	29	2.0e+01	1.1e−07	20
portfl-i-6	87	2.09	2.361e−6	2.08e−05	3.1e−07	5.1e−08	9.0e−08	27	2.0e+01	5.7e−07	20
qpec-100-3	110	37.04	-5.48287	-5.48287	5.9e−07	7.0e−10	1.4e−09	50	6.0e+01	2.1e−08	60
qpec1	80	0.348947	80	8.0e+01	7.4e−07	5.8e−09	7.1e−09	17	1.8e+02	5.1e−07	180
ralph2	4	0.06	0	-6.11e−08	3.5e−07	6.2e−09	3.e−08	19	2.0e+01	1.4e−07	20
scale1	2	0.18	1	1.0e−00	9.4e−13	3.2e−13	4.4e−12	42	5.4e+02	1.6e−10	540
scale2	2	0.05	1	1.0e−00	3.5e−08	3.2e−08	9.6e−08	10	2.0e+01	5.1e−07	20
scale3	2	0.16	1	1.0e−00	7.5e−15	3.2e−13	4.4e−12	46	5.4e+02	1.6e−10	540
scale5	2	0.89	100	1.0e+02	2.8e−11	5.5e−12	1.3e−11	105	4.9e+03	2.0e−08	4860
scholtes1	4	1.05	2	2.0e+00	5.6e−11	1.1e−09	1.7e−09	314	1.8e+02	1.0e−07	180
scholtes2	4	0.18	15	1.5e+01	8.6e−09	3.9e−09	9.3e−10	51	5.6e+01	9.5e−08	60
scholtes3	2	0.07	.5	.5	1.8e−07	1.1e−08	1.1e−09	19	2.0e+01	1.0e−07	20
scholtes5	6	0.049	1	1.00e+00	7.5e−09	2.7e−08	2.0e−08	10	3.3e+01	5.1e−07	33
sl1	11	0.40	.0001	1.00e−04	4.4e−07	6.8e−10	1.7e−09	52	6.3e+01	3.1e−08	63
stackelberg1	3	0.31	-3266.67	-3266.67	1.3e−12	1.1e−09	1.2e−11	82	6.4e+02	1.0e−07	640
tap-09	122	141.59	109.143	109.131	6.6e−07	3.4e−07	1.9e−08	763	9.2e+00	5.1e−07	8
tap-15	293	1471.92	184.295	184.295	4.3e−07	6.8e−08	3.7e−09	1482	8.7e+00	1.0e−07	8

Table A.2: Problems solved to optimality for formulation (1.1) with a different objective value.

Name	#vars	time	f_L^*	$f(x^*)$	$\|\nabla L\|_\infty$	$\|(s,t)\|_\infty$	infeas.	# iter	$\|y\|_\infty$	$\max\{s_i t_i\}$	$\nu_{\max}$
bard1m	9	0.37	17	2	1.5e−10	3.7e−09	5.3e−09	68	4.8e+03	1.0e−07	4800
bard2	16	1.06	-6598	6163	6.2e−08	2.5e−11	2.9e−11	75	1.4e+04	4.1e−09	1e+4
bard2m	20	2.20	-6598	-6600	6.1e−07	4.1e−11	3.0e−11	133	1.4e+04	4.1e−09	1e+4

Table A.2: Problems solved to optimality for formulation (1.1) with a different objective value (continued).

Name	#vars	time	f_L^*	$f(x^*)$	$\|\nabla L\|_\infty$	$\|(s,t)\|_\infty$	infeas.	# iter	$\|y\|_\infty$	$\max\{s_i t_i\}$	$\nu_{\max}$
bilevel1m	10	4.17	-60	-10	4.0e−09	1.1e−09	2.0e−12	411	6.4e+02	1.0e−07	642
bilin	14	0.6	5.6	1.46e+01	7.9e−10	6.2e−10	4.8e−10	84	1.8e+02	7.1e−08	180
dempe	4	11.12	31.25	28.25	1.0e−08	1.0e−08	2.8e−08	4035	2.0e+01	1.0e−07	20
ex913	23	9.76	-29.2	-6.e+00	2.1e−08	1.0e−10	1.6e−10	397	1.8e+02	4.1e−09	180
ex916	14	3.98	-15	-1.10e+01	8.2e−09	9.1e−12	4.9e−13	201	7.6e+02	8.2e−10	640
ex917	17	1.55	-6	-2.6e+01	1.2e−08	1.4e−08	1.0e−08	101	1.8e+02	5.1e−07	180
ex921	10	1.63	-1.25	1.7e+01	6.5e−08	2.3e−10	4.9e−11	124	1.0e+03	2.0e−08	640
ex927	10	1.56	25	1.70e+01	6.5e−08	2.3e−10	4.9e−11	124	1.0e+03	2.0e−08	640
gnash15	13	2.70	-354.699	-3.05671e+02	1.5e−08	7.0e−09	1.4e−08	153	4.8e+02	5.1e−07	480
gnash18	13	0.47	-25.6982	-7.06e+00	2.6e−08	3.2e−10	2.8e−10	49	4.8e+02	2.0e−08	480
gnash19	13	0.53	-6.11671	-1.79e−01	3.7e−08	1.0e−09	8.0e−10	51	1.6e+02	2.1e−08	160
gnashm16	9	1.25	-241.44	-2.23e+02	6.52e−08	1.34e−08	9.837e−09	107	1.280e+03	5.87e−07	1280
gnashm17	9	0.69	-90.75	-7.94e+01	3.73e−08	8.16e−09	5.998e−09	52	1.6e+02	1.02e−07	160
gnashm18	9	0.66	-25.7	-1.77e+01	2.78e−08	1.39e−09	1.35e−09	48	1.6e+02	2.194e−08	160
gnashm19	9	1.01	-6.11671	-2.45	2.908e−09	1.42e−09	1.169e−09	54	1.600e+02	2.048e−08	160
hs044-i	26	0.70	15.6117	6.29e−06	6.2e−08	6.8e−10	4.5e−10	57	6.0e+01	2.062e−08	60
monteiro	216	1028.99	37.53	-3.82e+02	3.0e−07	3.4e−14	2.5e−14	4010	5.2e+05	2.4e−09	3e+7
outrata31	9	17.93	3.2077	2.60	5.3e−08	1.2e−09	9.1e−10	1900	1.8e+02	1.0e−07	181
pack-comp1p-8	107	3858.52	0.6	-2.18e+06	1.0e+11	7.9	1.0e+01	6013	8.0e+17	8.8e+5	2e+9
pack-comp2-8	107	724.31	.673117	6.42e−01	1.5e+06	1.2e−14	1.3e−03	2610	1.0e+05	1.3e−12	1782
qpec-100-1	205	26.76	.0990028	2.41e−01	4.0e−08	3.6e−09	9.7e−09	40	6.0e+01	1.0e−07	60
qpec-100-2	210	40.92	-6.26049	-6.43	7.9e−07	3.8e−09	6.3e−09	51	6.0e+01	1.05e−07	60
qpec-100-4	220	35.17	-3.60073	-3.91	5.8e−09	3.7e−09	7.3e−09	36	6.0e+01	1.0e−07	60

Table A.3: Failures for formulation (1.1).

Name	#vars	time	f_L^*	$f(x^*)$	$\|\nabla L\|_\infty$	$\|(s,t)\|_\infty$	infeas.	# iter	$\|y\|_\infty$	$\max\{s_i t_i\}$	$\nu_{\max}$
bar-truss-3	35	298.44	10166.6	2920	2.2e+10	1.6e+02	1.6e+02	4783	1.6e+10	3.1e+9	1e+9
bilevel1	16	55.03	-60	-3.33	4.2e+08	3.3e+00	3.3e+00	2128	4.2e+08	2.9e+7	1e+9
design-cent1	15	147.39	1.806	3.62e−05	1.4e+04	1.5e−14	1.0e−14	10180	1.8e+02	1.3e−12	180
ex922	10	1.92	55.23	9.997e+01	6.5e−10	1.5e−14	2.033e−05	115	1.3e+3	1.3e−12	1280
ex923	16	48.92	-55	4.0e+00	1.7e+09	2.0e+00	2.0e+00	1161	4.6e+08	7.4e+6	1e+9
gnashm15	13	59.46	-354.699	-3.79e+02	7.63e+06	1.83e−08	4.83e+01	5834	3.36e+08	1.28e−05	1e+09
hakonsen	11	161.24	24.3668	1.99e+01	1.3e+08	1.4e−01	1.4e−01	12475	5.6e+7	6.4e+6	1e+8
monteiroB	216	4066.1	827.859	1.99e−05	2.1e+11	3.8e+0	3.8e+0	1604	3.4e+8	9.9e+3	1e+9
pack-rig1p-8	156	3582.47	.787932	-2.36e+06	1.6e+13	1.3e+03	1.4e+03	5599	2.1e+16	7.5e+8	3e+9
pack-rig2p-8	156	3634.72	.780404	-2.31e+06	1.2e+13	5.6e+02	1.2e+03	6115	2.5e+08	1.1e+10	3e+9
qpec2	30	0.47	45	4.47e+01	7.3e−11	4.8e−15	5.4e−05	22	5.40e+02	1.3e−12	540
ralph1	3	0.10	0	-1.67e−02	1.39e−13	4.3e−14	2.8e−04	25	6.0e+01	1.3e−12	180
ralphmod	204	35486.4	683.033	-6.7e+02	1.5e+10	1.01e−01	1.6	1293	1.4e+07	3.03e+03	2e+09
scale4	2	0.07	1	4.99e−07	2.5e−14	6.6e−14	1.0e−04	18	2.0e+01	1.3e−12	20
scholtes4	4	11.29	-3.07336e−7	-9.99e−02	1.5e+05	2.5e−03	2.5e−03	3156	3.0e+06	7.5e+0 3	4860
taxmcp	19	54.15	.818705	1.0	4.7e+08	2.2e−15	2.0e+00	2463	4.3e+08	4.1e−09	3e+9
water-FL	169	3229.62	929.169	-8.42e+07	2.2e+13	1.4e+03	1.4e+03	4072	1.2e+13	1.3e+11	3e+9
water-net	52	880.47	3411.92	1.61e+02	7.0e+15	9.7e+00	9.7e+00	5792	3.4e+08	2.0e+4	1e+9

As mentioned by Benson et al. (2006) and DeMiguel et al. (2005), some problems in Table (A.3) are ill-posed in the sense that they do not admit a strongly stationary point. It is the case with *ex9.2.2*, *qpec2*, *ralph1* and *scholtes4*.

According to DeMiguel et al. (2005) the *pack* problems have an empty strictly feasible region, *ralphmod* is unbounded, and *design-cent-3* is infeasible. The ill-posedness of these MPCCs may explain why our method fails. We may also observe here that Algorithm 3.1 can solve the problem *tap-15* although DeMiguel et al. (2005) claim that it does not have a strongly stationary point.

References

M. Anitescu. On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 15(4):1203–1236, 2005.

M. Anitescu, P. Tseng, and S. J. Wright. Elastic-mode algorithms for mathematical programs with equilibrium constraints: global convergence and stationarity properties. *Mathematical Programming*, 110:337–371, 2007.

H. Y. Benson, A. Sen, D. F. Shanno, and R. J. Vanderbei. Interior-point algorithms, penalty methods and equilibrium problems. *Computational Optimization and Applications*, 34(2):155–182, 2006.

A. R. Conn, N. I. M. Gould, D. Orban, and Ph. L. Toint. A primal-dual trust-region algorithm for non-convex nonlinear programming. *Mathematical Programming*, 87(2):215–249, 2000.

P.-R. Curatolo and D. Orban. An elastic penalty method for mathematical programs with vanishing constraints. Technical Report Cahier du GERAD G-2009-xx, GERAD, Montréal, QC, Canada, 2009. In preparation.

V. DeMiguel, M. P. Friedlander, F. J. Nogales, and S. Scholtes. A two-sided relaxation scheme for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 16(2):587–609, 2005.

E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2002.

I. S. Duff. MA57—a code for the solution of sparse Symmetric Definite and Indefinite Systems. *Transactions of the ACM on Mathematical Software*, 30(2):118–144, 2004.

R. Fourer, D. M. Gay, and B. W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Press / Brooks/Cole Publishing Company, second edition, 2002.

J. Gauvin. A necessary and sufficient regularity condition to have bounded multipliers in non convex programming. *Mathematical Programming*, 12:136–138, 1977.

D. M. Gay. Hooking your solver to Ampl. Technical Report 97-4-06, Lucent Technologies Bell Labs Innovations, Murray Hill, NJ, 1997. www.ampl.com/REFS/HOOKING.

P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM Review*, 47(1):99–131, 2005.

N. I. M. Gould, D. Orban, and Ph. L. Toint. An interior-point ℓ_1-penalty method for nonlinear optimization. Technical Report RAL-TR-2003-022, Rutherford Appleton Laboratory, Chilton, Oxfordshire, England, 2003.

Harwell Subroutine Library. *A collection of Fortran codes for large-scale scientific computation.* AERE Harwell Laboratory, Harwell, Oxfordshire, England, 2007. www.cse.clrc.ac.uk/nag/hsl.

J. Kyparisis. On uniqueness of Kuhn-Tucker multipliers in nonlinear programming. *Mathematical Programming*, 32:242–246, 1985.

S. Leyffer. MacMPEC: AMPL collection of MPECs. www.mcs.anl.gov/∼leyffer/MacMPEC, 2004.

S. Leyffer, G. López-Calva, and J. Nocedal. Interior methods for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 17(1):52–77, 2006.

O. L. Mangasarian. *Nonlinear Programming*. Number 10 in Classics in Applied Mathematics. SIAM, Philadelphia, PA, 2nd edition, 1994.

D. Orban. NLPy—a large-scale optimization toolkit in Python. Technical Report Cahier du GERAD G-2009-xx, GERAD, Montréal, QC, Canada, 2009. In preparation.

A. U. Raghunathan and L. T. Biegler. An interior point method for mathematical programs with complementarity constraints (mpccs). *SIAM Journal on Optimization*, 15(3):720–750, 2005.

D. Ralph and S. J. Wright. Some properties of regularization and penalization schemes for MPECs. *Optimization Methods and Software*, 19(5):527–556, 2004.

H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000.

S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936, 2001.

S. J. Wright and D. Orban. Local Convergence of the Newton/Log-Barrier Method for Degenerate Problems. *Mathematics of Operations Research*, 27(3):585–613, 2002.