Interior Methods for Mathematical Programs with Complementarity Constraints*

Sven Leyffer[†], Gabriel López-Calva[‡] and Jorge Nocedal[§]

December 17, 2004

Abstract

This paper studies theoretical and practical properties of interior-penalty methods for mathematical programs with complementarity constraints. A framework for implementing these methods is presented, and the need for adaptive penalty update strategies is motivated with examples. The algorithm is shown to be globally convergent to strongly stationary points, under standard assumptions. These results are then extended to an interior-relaxation approach. Superlinear convergence to strongly stationary points is also established. Two strategies for updating the penalty parameter are proposed, and their efficiency and robustness are studied on an extensive collection of test problems.

Keywords: MPEC, MPCC, nonlinear programming, interior-point methods, exact penalty, equilibrium constraints, complementarity constraints.
AMS-MSC2000: 90C30, 90C33, 90C51, 49M37, 65K10.

1 Introduction

In this paper we study the numerical solution of mathematical programs with complementarity constraints (MPCCs) of the form

$$
\begin{aligned}
\text{minimize} \quad & f(x) && \text{(1.1a)}\\
\text{subject to} \quad & c_i(x) = 0, \quad i \in \mathcal{E} && \text{(1.1b)}\\
& c_i(x) \geq 0, \quad i \in \mathcal{I} && \text{(1.1c)}\\
& 0 \leq x_1 \perp x_2 \geq 0. && \text{(1.1d)}
\end{aligned}
$$

The variables have been divided as $x = (x_0, x_1, x_2)$, with $x_0 \in \mathbb{R}^n$, $x_1, x_2 \in \mathbb{R}^p$. The complementarity condition (1.1d) stands for

$$
x_1 \geq 0,\ x_2 \geq 0, \ \text{ and either } \ x_{1i} = 0 \ \text{ or } \ x_{2i} = 0, \ \text{ for } \ i = 1, \ldots, p, \tag{1.2}
$$

where x_{1i}, x_{2i} are the ith components of vectors x_1 and x_2, respectively.

Complementarity (1.2) represents a logical condition (a disjunction) and must be expressed in analytic form if we wish to solve the MPCC using nonlinear programming methods. A popular

*Preprint ANL/MCS-P1211-1204
[†]Mathematics and Computer Science Division, Argonne National Laboratory, leyffer@mcs.anl.gov
[‡]IEMS Department, Northwestern University, g-lopez-calva@northwestern.edu
[§]ECE Department, Northwestern University, nocedal@ece.northwestern.edu

reformulation of the MPCC is

$$\begin{aligned} \text{minimize} \quad & f(x) && \text{(1.3a)} \\ \text{subject to} \quad & c_i(x) = 0, \quad i \in \mathcal{E} && \text{(1.3b)} \\ & c_i(x) \geq 0, \quad i \in \mathcal{I} && \text{(1.3c)} \\ & x_1 \geq 0, \quad x_2 \geq 0 && \text{(1.3d)} \\ & x_{1i} x_{2i} \leq 0 \;\; i = 1, \ldots, p. && \text{(1.3e)} \end{aligned}$$

This formulation preserves the solution set of the MPCC but is not totally adequate because it violates the Mangasarian-Fromowitz constraint qualification (MFCQ) at any feasible point. This lack of regularity can create problems when applying *classical* nonlinear programming algorithms. Sequential quadratic programming (SQP) methods can give rise to inconsistent constraint linearizations. Interior methods exhibit inefficiencies caused by the conflicting goals of enforcing complementarity while keeping the variables x_1, x_2 away from their bounds.

Modern nonlinear programming algorithms include, however, regularization techniques and other safeguards to deal with degeneracy, and one cannot rule out the possibility that they can cope with the difficulties created by the formulation (1.3) without having to exploit the special structure of MPCCs. If this level of robustness could be attained (and this is a laudable goal) there might be no need to develop algorithms specifically for MPCCs.

Numerical experiments by Fletcher and Leyffer [10] suggest that this goal is almost achieved by modern active-set SQP methods. In [10], FILTERSQP [9] was used to solve the problems in the MacMPEC collection [16], which contains over a hundred MPCCs, and fast convergence was almost always observed. The reason for this practical success is that, even though the formulation (1.3) fails to satisfy MFCQ, it is locally equivalent to a nonlinear program that satisfies MFCQ, and a robust SQP solver is able to identify the right set of active constraints in the well-behaved program and converge to the solution. Failures, however, are still possible for the SQP approach. Fletcher et al. [11] give several examples that illustrate ways in which an SQP method may fail to converge.

Interior methods are less successful when applied directly to the nonlinear programming formulation (1.3). Fletcher and Leyffer [10] tested LOQO [24] and KNITRO [4] (both codes as of 2002) and observed that they were slower and less reliable than SQP methods. This result contrasts starkly with the experience in nonlinear programming, where interior methods compete well with SQP methods. These studies have stimulated considerable interest in developing interior methods for MPCCs that guarantee both global convergence and efficient practical performance. The approaches are in two broad categories.

The first category comprises relaxation approaches, where (1.3) is replaced by a family of problems in which (1.3e) is changed to

$$x_{1i} x_{2i} \leq \theta, \quad i = 1, ..., p, \tag{1.4}$$

and the relaxation parameter $\theta > 0$ is driven to zero. This type of approach has been studied from a theoretical perspective by Scholtes [23] and Ralph and Wright [21]. Interior methods based on the relaxation (1.4) have been proposed by Liu and Sun [17] and Raghunathan and Biegler [20]. In both studies, the parameter θ is proportional to the barrier parameter μ and is updated only at the end of each barrier problem. Raghunathan and Biegler focus on local analysis and report very good numerical results on the MacMPEC collection. Liu and Sun analyze global convergence of their algorithm and report limited numerical results. Numerical difficulties may arise when the relaxation parameter gets small, since the interior of the regularized problem shrinks toward the empty set.

DeMiguel et al. [7] address this problem by proposing a different relaxation scheme where, in addition to (1.4), the nonnegativity bounds on the variables are relaxed to

$$x_{1i} \geq -\delta, \quad x_{2i} \geq -\delta. \tag{1.5}$$

Under fairly general assumptions, their algorithm drives either θ or δ, but not both, to zero. This provides the resulting family of problems with a strict interior, even when the appropriate relaxation parameters are approaching zero, which is a practical advantage over the previous relaxation approach. The drawback is that the algorithm has to correctly identify the parameters that must be driven to zero, a requirement that can be difficult to meet in some cases.

The second category involves a regularization technique based on an exact-penalty reformulation of the MPCC. Here, (1.3e) is moved to the objective function in the form of an ℓ_1-penalty term

$$f(x) + \pi x_1^T x_2, \tag{1.6}$$

where $\pi > 0$ is a penalty parameter. If π is chosen large enough, the solution of the MPCC can be recast as the minimization of a single penalty function. The appropriate value of π is, however, unknown in advance and must be estimated during the course of the minimization.

This approach was first proposed by Anitescu [1] in the context of active set SQP methods. It has been adopted as a heuristic to solve MPCCs with interior methods in LOQO by Benson et al. [3], who present very good numerical results on the MacMPEC set. A more general class of exact penalty functions was analyzed by Hu and Ralph [15], who derive global convergence results for a sequence of penalty problems that are solved exactly. Anitescu [2] derives similar global results in the context of inexact solves.

In this paper, we focus on the penalization approach, because we view it as a general tool for handling degenerate nonlinear programs. Our goal is to study global and local convergence properties of interior-penalty methods for MPCCs and to propose efficient and robust practical implementations.

We start in Section 2 by presenting the interior-penalty framework; some examples motivate the need for proper updating strategies for the penalty parameter. Section 3 shows that the proposed interior-penalty method converges globally to strongly stationary points, under standard assumptions. These results are then extended to the interior-relaxation approaches considered in [17] and [20]. In Section 4 we show that, near a solution that satisfies some standard regularity properties, the penalty parameter is not updated, and the iterates converge superlinearly to the solution. Section 5 presents two practical implementations of the interior-penalty method with different updating strategies for the penalty parameter. Our numerical experiments, reported in the same section, favor a dynamic strategy that assesses the magnitude of the penalty parameter at every iteration. We close the paper with some concluding remarks.

2 An Interior-Penalty Method for MPCCs

To circumvent the difficulties caused by the complementarity constraints, we replace (1.3) by the ℓ_1-penalty problem

$$\begin{array}{ll} \text{minimize} & f(x) + \pi x_1^T x_2 \\ \text{subject to} & c_i(x) = 0, \quad i \in \mathcal{E} \\ & c_i(x) \geq 0, \quad i \in \mathcal{I} \\ & x_1 \geq 0, \ x_2 \geq 0, \end{array} \tag{2.1}$$

where $\pi > 0$ is a penalty parameter. In principle, the ℓ_1-penalty term should have the form $\sum_i \max\{0, x_{1i}x_{2i}\}$, but we can write it as $x_1^T x_2$ if we enforce the nonnegativity of x_1, x_2. This exact penalty reformulation of MPCCs has been studied in [1, 2, 3, 15, 21, 22]. Since problem (2.1) is smooth, we can safely apply standard nonlinear programming algorithms, such as interior methods, to solve it. The barrier problem associated to (2.1) is

$$\begin{array}{ll} \text{minimize} & f(x) + \pi x_1^T x_2 - \mu \sum_{i \in \mathcal{I}} \log s_i - \mu \sum_{i=1}^{p} \log x_{1i} - \mu \sum_{i=1}^{p} \log x_{2i} \\ \text{subject to} & \begin{array}{rcll} c_i(x) & = & 0, & i \in \mathcal{E}, \\ c_i(x) - s_i & = & 0, & i \in \mathcal{I}, \end{array} \end{array} \tag{2.2}$$

where $\mu > 0$ is the barrier parameter and $s_i > 0$, $i \in \mathcal{I}$, are slack variables. The Lagrangian of this barrier problem is given by

$$\begin{aligned} \mathcal{L}_{\mu,\pi}(x, s, \lambda) \;=\; & f(x) + \pi x_1^T x_2 - \mu \sum_{i \in \mathcal{I}} \log s_i - \mu \sum_{i=1}^{p} \log x_{1i} - \mu \sum_{i=1}^{p} \log x_{2i} \\ & - \sum_{i \in \mathcal{E}} \lambda_i c_i(x) - \sum_{i \in \mathcal{I}} \lambda_i (c_i(x) - s_i), \end{aligned} \tag{2.3}$$

and the first-order Karush-Kuhn-Tucker (KKT) conditions of (2.2) can be written as

$$\begin{array}{rcll} \nabla_x \mathcal{L}_{\mu,\pi}(x, s, \lambda) = \nabla f(x) - \nabla c_{\mathcal{E}}(x)^T \lambda_{\mathcal{E}} - \nabla c_{\mathcal{I}}(x)^T \lambda_{\mathcal{I}} - \begin{pmatrix} 0 \\ \mu X_1^{-1} e - \pi x_2 \\ \mu X_2^{-1} e - \pi x_1 \end{pmatrix} & = & 0, & \\ s_i \lambda_i - \mu & = & 0 & i \in \mathcal{I}, \\ c_i(x) & = & 0 & i \in \mathcal{E}, \\ c_i(x) - s_i & = & 0 & i \in \mathcal{I}, \end{array} \tag{2.4}$$

where we have grouped the components $c_i(x), i \in \mathcal{E}$ into the vector $c_{\mathcal{E}}(x)$, and similarly for $c_{\mathcal{I}}(x), \lambda_{\mathcal{E}}, \lambda_{\mathcal{I}}$. We also define $\lambda = (\lambda_{\mathcal{E}}, \lambda_{\mathcal{I}})$. X_1 denotes the diagonal matrix containing the elements of x_1 on the diagonal (the same convention is used for X_2 and S), and e is a vector of ones of appropriate dimension.

The KKT conditions (2.4) can be expressed more compactly as

$$\nabla_x \mathcal{L}_{\mu,\pi}(x, s, \lambda) = 0, \tag{2.5a}$$

$$S \lambda_{\mathcal{I}} - \mu e = 0, \tag{2.5b}$$

$$c(x, s) = 0, \tag{2.5c}$$

where we define

$$c(x, s) = \begin{pmatrix} c_{\mathcal{E}}(x) \\ c_{\mathcal{I}}(x) - s \end{pmatrix}. \tag{2.6}$$

In Figure 1, we describe an interior method for MPCCs based on the ℓ_1-penalty formulation. In addition to requiring that the optimality conditions (2.7) of the barrier problem are satisfied approximately, we impose a reduction in the complementarity term by means of (2.8). For now, the only requirement on the sequence of barrier parameters $\{\mu^k\}$ and the stopping tolerances $\{\epsilon_{pen}^k\}, \{\epsilon_{comp}^k\}$ is that they all converge to 0 as $k \to \infty$. Later, in the local analysis of Section 4, we will impose further conditions on the relative speed of convergence of these sequences.

We use $\| \min\{x_1^k, x_2^k\} \|$ in (2.8) as a measure of complementarity, rather than $x_1^{kT} x_2^k$, because it is less sensitive to the scaling of the problem and is independent of the number of variables. Moreover, this measure is accurate even when both x_{1i}^k and x_{2i}^k converge to zero.

Algorithm I: Interior-Penalty Method for MPCCs

Initialization: Let x^0, s^0, λ^0 be the initial primal and dual variables. Set $k = 1$.

repeat

1. Choose a barrier parameter μ^k, stopping tolerances ϵ^k_{pen} and ϵ^k_{comp}, and a penalty parameter π^k

2. Apply an algorithm to approximately solve the barrier problem (2.2), updating the penalty parameter π^k if necessary, until the following conditions are satisfied for some (x^k, s^k, λ^k), with $x_1^k > 0, x_2^k > 0, s^k > 0, \lambda^k_{\mathcal{I}} > 0$:

$$\begin{aligned} \|\nabla_x \mathcal{L}_{\mu^k,\pi^k}(x^k, s^k, \lambda^k)\| &\leq \epsilon^k_{pen}, &\text{(2.7a)}\\ \|S^k \lambda^k_{\mathcal{I}} - \mu^k e\| &\leq \epsilon^k_{pen}, &\text{(2.7b)}\\ \|c(x^k, s^k)\| &\leq \epsilon^k_{pen}, &\text{(2.7c)} \end{aligned}$$

   and

$$\|\min\{x_1^k, x_2^k\}\| \leq \epsilon^k_{comp} \tag{2.8}$$

3. Let $k \leftarrow k + 1$

until a stopping test for the MPCC is satisfied

Figure 1: An interior-penalty method for MPCCs.

Our formulation of Algorithm I is sufficiently general to permit various updating strategies for the penalty parameter in Step 2. One option is to keep π^k fixed during the solution of the barrier problem governed by μ^k. If condition (2.8) holds, then we proceed to Step 3. Otherwise, we increase π^k and solve (2.2) again using the same barrier parameter μ^k. The process is repeated, if necessary, until (2.8) is satisfied. We show in Section 5 that Algorithm I with this penalty update strategy is much more robust and efficient than the direct application of an interior method to (1.3). Nevertheless, there are some flaws in a strategy that holds the penalty parameter fixed throughout the minimization of a barrier problem, as illustrated by the following examples.

The results reported next were obtained with an implementation of Algorithm I that uses the penalty update strategy described in the previous paragraph. The initial parameters are $\pi^1 = 1, \mu^1 = 0.1$, and we set $\epsilon^k_{comp} = (\mu^k)^{0.4}$ for all k. When the penalty parameter is increased, it is multiplied by 10. The other details of the implementation are discussed in Section 5 and are not relevant to the discussion that follows.

Example 1 (`ralph2`). Consider the MPCC

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 - 4x_1x_2 \\ \text{subject to} \quad & 0 \leq x_1 \perp x_2 \geq 0, \end{aligned} \tag{2.9}$$

whose solution is $(0,0)$. The associated penalty problem is

$$\begin{array}{ll} \text{minimize} & (x_1 - x_2)^2 + (\pi - 2)x_1 x_2 \\ \text{subject to} & x_1 \geq 0,\ x_2 \geq 0, \end{array} \tag{2.10}$$

which is unbounded for any $\pi < 2$. Starting with $\pi^1 = 1$, the first barrier problem is never solved. The iterates increase monotonically because, by doing so, the objective function is reduced and feasibility is maintained for problem (2.10). Eventually, the iterates diverge. Table 1 shows the values of $x_1 x_2$ during the first eight iterations of the inner algorithm in Step 2.

Table 1: Complementarity values for problem `ralph2`.

Iterate	1	2	3	4	5	6	7	8
Complementarity	0.0264	0.0916	0.1480	51.70	63.90	79.00	97.50	120.0

The upward trend in complementarity should be taken as a warning sign that the penalty parameter is not large enough, since no progress is made toward satisfaction of (2.8). This suggests that we should be prepared to increase the penalty parameter *during* the solution of the barrier problem. How to do so, in a robust manner, is not a simple question because complementarity can oscillate or increase slowly. We return to this issue in Section 5, where we describe a dynamic strategy for updating the penalty parameter. □

Example 2 (`scale1`). Even if the penalty problem is bounded, there are cases where efficiency can be improved with a more flexible strategy for updating π^k. For example, consider the MPCC

$$\begin{array}{ll} \text{minimize} & (100x_1 - 1)^2 + (x_2 - 1)^2 \\ \text{subject to} & 0 \leq x_1 \perp x_2 \geq 0, \end{array} \tag{2.11}$$

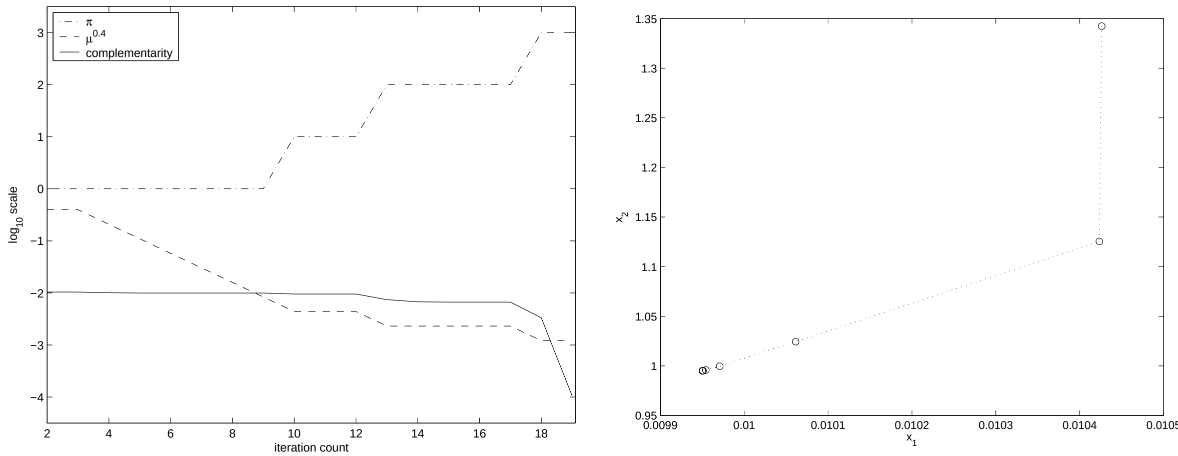
which has two local solutions: $(0.01, 0)$ and $(0, 1)$. Table 2 shows the first seven values of x^k satisfying (2.7) and (2.8), and the corresponding values of μ^k.

Table 2: Solutions of 7 consecutive barrier-penalty problems for `scale1`.

k	1	2	3	4	5	6	7
μ^k	0.1	0.02	0.004	0.0008	0.00016	0.000032	0.0000064
x_1^k	0.010423	0.010061	0.009971	0.009954	0.009951	0.009950	0.009950
x_2^k	1.125463	1.024466	0.999634	0.995841	0.995186	0.995057	0.995031
ϵ_{comp}^k	0.398107	0.209128	0.109856	0.057708	0.030314	0.015924	0.008365

We observe that complementarity, as measured by $\min\{x_1^k, x_2^k\}$, stagnates. This result is not surprising because the minimum penalty parameter required to recover the solution $(0,1)$ is $\pi^* = 200$ and we have used the value $\pi^1 = 1$. In fact, for any $\pi < 200$, there is a saddle point close to $(0,1)$, and the iterates approach that saddle point. Seven barrier problems must be solved before the test (2.8) is violated for the first time, triggering the first update of π^k.

The behavior of the algorithm is illustrated in Figure 2(a), which plots three quantities as a function of the inner iterations. Complementarity (continuous line) stalls at a nonzero value during the first ten iterations, while μ^k (dashed line) decreases monotonically. The penalty parameter (dashed-dotted line) is increased for the first time at iteration 9. It must be increased three times to surpass the threshold value $\pi^* = 200$, which finally forces complementarity down to zero.



(a) Complementarity gap. (b) Solution path.

Figure 2: A numerical solution of problem `scale1`.

Figure 2(b) shows the path of the iterates up to the solution of the seventh barrier problem. There is a clear pattern of convergence to the stationary point where none of the variables is zero. If this convergence pattern can be identified early, the penalty parameter can be increased sooner, saving some iterations in the solution of the MPCC. □

Example 3. One could ask whether the penalty parameter needs to be updated at all. Suppose that we choose a very large value of π and hold it fixed during the execution of Algorithm I. Could this prove to be a more effective strategy? In Section 5 we show that excessively large penalty parameters can result in substantial loss of efficiency. More important, this strategy does not constitute a reliable approach because, for any given choice of π, one can find problems and initial points for which a penalty method can generate steps that lead away from the solution. For example, the MPCC

$$\begin{array}{rl} \text{minimize} & x_1^2 + x_2^2 - x_1^2 x_2^2 \\ \text{subject to} & 0 \leq x_1 \perp x_2 \geq 0 \end{array} \tag{2.12}$$

has a unique minimizer at the origin. The associated penalty problem is

$$\begin{array}{rl} \text{minimize} & x_1^2 + x_2^2 - x_1^2 x_2^2 + \pi x_1 x_2 \\ \text{subject to} & x_1 \geq 0, \; x_2 \geq 0. \end{array} \tag{2.13}$$

For any fixed penalty parameter π, choose $x_1^0 = x_2^0 > \sqrt{1 + \pi/2}$, and consider the direction $d = (1,1)^T$. A simple computation shows that d is a feasible descent direction at x^0 for (2.13), which clearly leads away from the solution. If π is increased fast enough, however, it eventually handles complementarity appropriately, and the iterates will not diverge. □

In Section 5, we describe a dynamic strategy for updating the penalty parameter that permits changes during the solution of the barrier problem. We show that it is able to promptly identify the undesirable behavior described in these examples and to react accordingly.

3 Global Convergence Analysis

In this section, we present the global convergence analysis of an interior-penalty method. We start by reviewing an MPCC constraint qualification that suffices to derive first-order optimality conditions for MPCCs. We then review stationarity concepts.

Definition 3.1 *We say that the* MPCC linear independence constraint qualification *(MPCC-LICQ) holds at a feasible point x for the MPCC (1.1) if and only if the standard LICQ holds at x for the set of constraints*

$$\begin{aligned} c_i(x) = 0, &\quad i \in \mathcal{E}, \\ c_i(x) \geq 0, &\quad i \in \mathcal{I}, \\ x_1 \geq 0, &\quad x_2 \geq 0. \end{aligned} \tag{3.1}$$

We denote indices of the active constraints at a feasible point x by

$$\begin{aligned} \mathcal{A}_c(x) &= \{i \in \mathcal{I} : c_i(x) = 0\}, \\ \mathcal{A}_1(x) &= \{i \in \{1, \ldots, p\} : x_{1i} = 0\}, \\ \mathcal{A}_2(x) &= \{i \in \{1, \ldots, p\} : x_{2i} = 0\}. \end{aligned} \tag{3.2}$$

For ease of notation, we use $i \notin \mathcal{A}_1(x)$ as shorthand for $i \in \{1, \ldots, p\} \setminus \mathcal{A}_1(x)$ (likewise for $\mathcal{A}_2, \mathcal{A}_c$). We sometimes refer to variables satisfying $x_{1i} + x_{2i} > 0$ as *branch variables*; those for which $x_{1i} + x_{2i} = 0$, that is, variables indexed by $\mathcal{A}_1(x) \cap \mathcal{A}_2(x)$, are called *corner variables.*

The next theorem establishes the existence of multipliers for minimizers that satisfy MPCC-LICQ. It can be viewed as a counterpart for MPCCs of the first-order KKT theorem for NLPs.

Theorem 3.2 *Let x^* be a minimizer of the MPCC (1.1), and suppose MPCC-LICQ holds at x^*. Then, there exist multipliers $\lambda^*, \sigma_1^*, \sigma_2^*$ that, together with x^*, satisfy the system*

$$\nabla f(x) - \nabla c_{\mathcal{E}}(x)^T \lambda_{\mathcal{E}} - \nabla c_{\mathcal{I}}(x)^T \lambda_{\mathcal{I}} - \begin{pmatrix} 0 \\ \sigma_1 \\ \sigma_2 \end{pmatrix} = 0 \tag{3.3a}$$

$$c_i(x) = 0, \;\; i \in \mathcal{E} \tag{3.3b}$$

$$c_i(x) \geq 0, \;\; i \in \mathcal{I} \tag{3.3c}$$

$$x_1 \geq 0, x_2 \geq 0 \tag{3.3d}$$

$$x_{1i} = 0 \;\; or \;\; x_{2i} = 0, \;\; i = 1, \ldots, p \tag{3.3e}$$

$$c_i(x)\lambda_i = 0, \;\; i \in \mathcal{I} \tag{3.3f}$$

$$\lambda_i \geq 0, \;\; i \in \mathcal{I} \tag{3.3g}$$

$$x_{1i}\sigma_{1i} = 0 \;\; and \;\; x_{2i}\sigma_{2i} = 0, \;\; i = 1, \ldots, p \tag{3.3h}$$

$$\sigma_{1i} \geq 0, \sigma_{2i} \geq 0, \;\; i \in \mathcal{A}_1(x) \cap \mathcal{A}_2(x). \tag{3.3i}$$

For a proof of this theorem, see [22] or an alternative proof in [18].

We note, that the multipliers σ_1, σ_2 are required to be nonnegative only for corner variables. This requirement reflects the geometry of the feasible set: If $x_{1i} > 0$, then $x_{2i} = 0$ acts like an equality constraint, and the corresponding multiplier can be positive or negative. Theorem 3.2 motivates the following definition.

Definition 3.3 *(a) A point x^* is called a* strongly stationary point *of the MPCC (1.1) if there exist multipliers $\lambda^*, \sigma_1^*, \sigma_2^*$ such that (3.3) is satisfied. (b) A point x^* is called a* C-stationary point *of the MPCC (1.1) if there exist multipliers $\lambda^*, \sigma_1^*, \sigma_2^*$ such that conditions (3.3a)–(3.3h) hold and*

$$\sigma_{1i}^* \sigma_{2i}^* \geq 0, \;\; i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*). \tag{3.4}$$

Strong stationarity implies the absence of first-order feasible descent directions. These are the points that the algorithms should aim for. Although C-stationarity does not characterize the solutions of an MPCC, since it allows descent directions if $\sigma_{1i} < 0$ or $\sigma_{2i} < 0$, we consider C-stationary points because they are attractors of iterates generated by Algorithm I. One can find examples in which a sequence of stationary points of the penalty problem converge to a C-stationary point where descent directions exist, and this phenomenon can actually be observed in practice (see case 1 of Example 3.1 in [15] and the comments on problem `scale4` in Section 5). The reader further interested in stationarity for MPCCs is referred to [22].

3.1 Global Convergence of the Interior-Penalty Algorithm

We now study the global convergence properties of Algorithm I. Many algorithms have been proposed to solve the barrier problem in Step 2; see, for example, [6, 12] and the references therein. As is well known, these inner algorithms may fail to satisfy (2.7), and therefore Algorithm I can fail to complete Step 2. The analysis of the inner algorithm is beyond the scope of this paper, and we concentrate only on the analysis of the outer iterations in Algorithm I. We will assume that the inner algorithm is always successful and that Algorithm I generates an infinite sequence of iterates $\{x^k, s^k, \lambda^k\}$ satisfying conditions (2.7) and (2.8). We now study the properties of limit points x^* of the sequence $\{x^k\}$.

We present the following result in the slightly more general setting in which a vector of penalties $\pi = (\pi_1, \ldots, \pi_p)$ is used, with the objective function as

$$f(x) + \pi^T X_1 x_2, \tag{3.5}$$

and with minor changes in the Lagrangian of the problem and the stopping test (2.8). This will allow us to extend the global convergence result to the relaxation approach in the next subsection. For the implementation, however, we use of a uniform (i.e., scalar-valued) penalty.

Theorem 3.4 *Suppose that Algorithm I generates an infinite sequence of iterates $\{x^k, s^k, \lambda^k\}$ and parameters $\{\pi^k, \mu^k\}$, satisfying conditions (2.7) and (2.8) for sequences $\{\epsilon^k_{pen}\}, \{\epsilon^k_{comp}\}, \{\mu^k\}$ converging to zero. If x^* is a limit point of the sequence $\{x^k\}$ and f and c are continuously differentiable in an open neighborhood $\mathcal{N}(x^*)$ of x^*, then x^* is feasible for the MPCC (1.1). If, in addition, MPCC-LICQ holds at x^*, then x^* is a C-stationary point of (1.1). Moreover, if $\pi^k_i x^k_{ji} \to 0$ for $j = 1, 2$ and $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$, then x^* is a strongly stationary point of (1.1).*

Proof. Let x^* be a limit point of the sequence $\{x^k\}$ generated by Algorithm I, and let $\mathcal{K}$ be an infinite index set such that $\{x^k\}_{k\in\mathcal{K}} \to x^*$. Then, $x^k \in \mathcal{N}(x^*)$ for all k sufficiently large; from the assumption of continuous differentiability on $\mathcal{N}(x^*)$, and $\{x^k\}_{k\in\mathcal{K}} \to x^*$, we conclude that the sequences $\{f(x^k)\}, \{c(x^k)\}, \{\nabla f(x^k)\}, \{\nabla c_{\mathcal{E}}(x^k)\}, \{\nabla c_{\mathcal{I}}(x^k)\}, k \in \mathcal{K}$ have limit points and are therefore bounded.

Since the inner algorithm used in Step 2 enforces positivity of the slacks s^k, by continuity of c and the condition $\epsilon^k_{pen} \to 0$ we have

$$\begin{aligned} c_i(x^*) = \; & 0 \qquad i \in \mathcal{E} \\ c_i(x^*) = s^*_i \geq \; & 0 \qquad i \in \mathcal{I}, \end{aligned}$$

where $s^*_i = \lim_{k\in\mathcal{K}} s^k_i$. Therefore x^* satisfies (3.3b) and (3.3c), and it also satisfies (3.3d) because the inner algorithm enforces the positivity of x^k. The complementarity condition (3.3e) follows directly from (2.8) and $\epsilon^k_{comp} \to 0$. Therefore, x^* is feasible for the MPCC (1.1).

Existence of Multipliers. Let us define

$$\sigma_{1i}^k = \frac{\mu^k}{x_{1i}^k} - \pi_i^k x_{2i}^k, \qquad \sigma_{2i}^k = \frac{\mu^k}{x_{2i}^k} - \pi_i^k x_{1i}^k \tag{3.6}$$

and

$$\alpha^k = \|(\lambda^k, \sigma_1^k, \sigma_2^k)\|_\infty. \tag{3.7}$$

We first show that $\{\alpha^k\}_{k\in\mathcal{K}}$ is bounded, a result that implies that the sequence of multipliers $(\lambda^k, \sigma_1^k, \sigma_2^k)$ has a limit point. Then we show that any limit point satisfies C-stationarity at x^*.

We can assume without loss of generality, that $\alpha_k \geq \tau > 0$ for all $k \in \mathcal{K}$. Indeed, if there were a further subsequence $\{\alpha_k\}_{k\in\mathcal{K}'}$ converging to 0, this subsequence would be trivially bounded, and we would apply the analysis below to $\{\alpha_k\}_{k\in\mathcal{K}\setminus\mathcal{K}'}$, which is bounded away from 0, to prove the boundedness of the entire sequence $\{\alpha_k\}_{k\in\mathcal{K}}$.

Let us define the "normalized multipliers"

$$\hat{\lambda}^k = \frac{\lambda^k}{\alpha^k}, \quad \hat{\sigma}_1^k = \frac{\sigma_1^k}{\alpha^k}, \quad \hat{\sigma}_2^k = \frac{\sigma_2^k}{\alpha^k}. \tag{3.8}$$

We now show that the normalized multipliers corresponding to inactive constraints converge to 0 for $k \in \mathcal{K}$. Consider an index $i \notin \mathcal{A}_c(x^*)$, where $\mathcal{A}_c$ is defined by (3.2). Since $s_i^k \to c_i(x^*) > 0$ and $s_i^k \lambda_i^k \to 0$ by (2.7b), we have that λ_i^k converges to 0, and so does $\hat{\lambda}_i^k$.

Next consider an index $i \notin \mathcal{A}_1(x^*)$. We want to show that $\hat{\sigma}_{1i}^k \to 0$. If $i \notin \mathcal{A}_1(x^*)$, then $x_{1i}^k \to x_{1i}^* > 0$, which implies that $x_{2i}^k \to 0$, by (2.8) and $\epsilon_{comp}^k \to 0$. We also have, from (3.6), that for any $k \in \mathcal{K}$,

$$\sigma_{1i}^k \neq 0 \quad \Rightarrow \quad \frac{\mu^k}{x_{1i}^k} - \pi_i^k x_{2i}^k \neq 0 \quad \Rightarrow \quad \frac{\mu^k}{x_{2i}^k} - \pi_i^k x_{1i}^k \neq 0 \quad \Rightarrow \quad \sigma_{2i}^k \neq 0. \tag{3.9}$$

Using this and the fact that $|\sigma_{2i}^k| \leq \alpha^k$, we have that, if there is any subsequence of indices k for which $\sigma_{1i}^k \neq 0$, then

$$\begin{aligned} |\hat{\sigma}_{1i}^k| = \frac{|\sigma_{1i}^k|}{\alpha^k} \leq \frac{|\sigma_{1i}^k|}{|\sigma_{2i}^k|} &= \frac{\left|\frac{\mu^k}{x_{1i}^k} - \pi_i^k x_{2i}^k\right|}{\left|\frac{\mu^k}{x_{2i}^k} - \pi_i^k x_{1i}^k\right|} \\ &= \frac{\left|\frac{\mu^k - \pi_i^k x_{1i}^k x_{2i}^k}{x_{1i}^k}\right|}{\left|\frac{\mu^k - \pi_i^k x_{1i}^k x_{2i}^k}{x_{2i}^k}\right|} = \frac{x_{2i}^k}{x_{1i}^k} \to 0. \end{aligned}$$

Since clearly $\hat{\sigma}_{1i}^k \to 0$ for those indices with $\sigma_{1i}^k = 0$, we have that the whole sequence $\hat{\sigma}_{1i}^k$ converges to zero for $i \notin \mathcal{A}_1(x^*)$. The same argument can be applied to show that $\hat{\sigma}_{2i}^k \to 0$ for $i \notin \mathcal{A}_2(x^*)$. Therefore we have shown that the normalized multipliers (3.8) corresponding to the inactive constraints converge to zero for $k \in \mathcal{K}$.

To prove that $\{\alpha^k\}_{k\in\mathcal{K}}$ is bounded, we proceed by contradiction and assume that there exists $\mathcal{K}' \subseteq \mathcal{K}$ such that $\{\alpha^k\}_{k\in\mathcal{K}'} \to \infty$. By definition, the sequences of normalized multipliers (3.8) are bounded, so we restrict $\mathcal{K}'$ further, if necessary, so that the sequences of normalized multipliers are convergent within $\mathcal{K}'$. Given that $\mathcal{K}' \subseteq \mathcal{K}$, all the sequences of gradients $\{\nabla f(x^k)\}$, $\{\nabla c_{\mathcal{E}}(x^k)\}$, $\{\nabla c_{\mathcal{I}}(x^k)\}$, $k \in \mathcal{K}'$ are convergent. We can then divide both sides of (2.7a) by α^k and take limits to get

$$\lim_{k\to\infty, k\in\mathcal{K}'} \|\frac{1}{\alpha^k} \nabla_x \mathcal{L}_{\mu^k,\pi^k}(x^k, s^k, \lambda^k)\| \quad \leq \quad \lim_{k\to\infty, k\in\mathcal{K}'} \frac{\epsilon_{pen}^k}{\alpha^k} = 0$$

or

$$\lim_{k\to\infty, k\in\mathcal{K}'} \left[\frac{1}{\alpha^k}\nabla f^k - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \hat\lambda_i^k \nabla c_i(x^k) - \sum_{i=1}^{p} \hat\sigma_{1i}^k \begin{pmatrix} 0 \\ e_i \\ 0 \end{pmatrix} - \sum_{i=1}^{p} \hat\sigma_{2i}^k \begin{pmatrix} 0 \\ 0 \\ e_i \end{pmatrix} \right] = 0. \tag{3.10}$$

It is immediate that the first term of (3.10) converges to 0. We showed that the coefficients (the normalized multipliers (3.8)) of the inactive constraints also converge to zero. Since the corresponding sequences of gradients have limits (hence are bounded), all the terms corresponding to inactive constraints get cancelled in the limit, and we have

$$\sum_{i\in\mathcal{E}\cup\mathcal{A}_c(x^*)} \hat\lambda_i^* \nabla c_i(x^*) + \sum_{i\in\mathcal{A}_1(x^*)} \hat\sigma_{1i}^* \begin{pmatrix} 0 \\ e_i \\ 0 \end{pmatrix} + \sum_{i\in\mathcal{A}_2(x^*)} \hat\sigma_{2i}^* \begin{pmatrix} 0 \\ 0 \\ e_i \end{pmatrix} = 0.$$

If the limit point x^* satisfies MPCC-LICQ, then the constraint gradients involved in this expression are linearly independent, and we get

$$\hat\lambda^* = 0, \;\; \hat\sigma_1^* = 0, \;\; \hat\sigma_2^* = 0.$$

This result, however, contradicts the fact that $\|(\hat\lambda^k, \hat\sigma_1^k, \hat\sigma_2^k)\|_\infty = 1$ for all $k \in \mathcal{K}'$, which follows from (3.7), (3.8) and the assumption that $\lim_{k\to\infty,k\in\mathcal{K}'} \alpha^k \to \infty$. Therefore, we conclude that no such unbounded subsequence exists, and hence all the sequences $\{\lambda^k\}, \{\sigma_1^k\}, \{\sigma_2^k\}$, with $k \in \mathcal{K}$, are bounded and have limit points.

C-Stationarity. Choose any such limit point $(\lambda^*, \sigma_1^*, \sigma_2^*)$, and restrict $\mathcal{K}$, if necessary, so that

$$(x^k, s^k, \lambda^k, \sigma_1^k, \sigma_2^k) \to (x^*, s^*, \lambda^*, \sigma_1^*, \sigma_2^*).$$

By (2.7a) and (2.4) and by continuity of f and c, we have that

$$\nabla f(x^*) - \nabla c_{\mathcal{E}}(x^*)^T \lambda_{\mathcal{E}}^* - \nabla c_{\mathcal{I}}(x^*)^T \lambda_{\mathcal{I}}^* - \begin{pmatrix} 0 \\ \sigma_1^* \\ \sigma_2^* \end{pmatrix} = 0,$$

which proves (3.3a). We have already shown that the limit point x^* satisfies conditions (3.3b) through (3.3e). The nonnegativity of λ^*, condition (3.3g), follows from the fact that the inner algorithm maintains $\lambda^k > 0$. Condition (3.3f) holds because, if $c_i(x^*) = s_i^* > 0$, then since $s_i^k \lambda_i^k \to 0$, we must have $\lambda_i^* = 0$.

We now establish that conditions (3.3h) hold at the limit point $(x^*, s^*, \lambda^*, \sigma_1^*, \sigma_2^*)$. They are clearly satisfied when $i \in \mathcal{A}_1(x^*)$ and $i \in \mathcal{A}_2(x^*)$. Consider an index $i \notin \mathcal{A}_1(x^*)$. If there is any infinite subset $\mathcal{K}'' \subseteq \mathcal{K}$ with $\sigma_{1i}^k \neq 0$ for all $k \in \mathcal{K}''$, then, as argued in (3.9), $\sigma_{1i}^k \neq 0 \Rightarrow \sigma_{2i}^k \neq 0$ for all $k \in \mathcal{K}''$ and

$$\lim_{k\to\infty,k\in\mathcal{K}''} \frac{|\sigma_{1i}^k|}{|\sigma_{2i}^k|} = \lim_{k\to\infty,k\in\mathcal{K}''} \frac{\left| \frac{\mu^k}{\pi_i^k x_{1i}^k} - x_{2i}^k \right|}{\left| \frac{\mu^k}{\pi_i^k x_{2i}^k} - x_{1i}^k \right|} = \lim_{k\to\infty,k\in\mathcal{K}''} \frac{x_{2i}^k}{x_{1i}^k} = 0, \tag{3.11}$$

where the limit follows from the fact that $x_{1i}^* > 0$, which implies that $x_{2i}^k \to 0$. $\{\sigma_{2i}^k\}$ has a limit and is therefore bounded. Hence, (3.11) can hold only if $\lim_{k\to\infty,k\in\mathcal{K}''} \sigma_{1i}^k = 0$ and, by definition, $\sigma_{1i}^k = 0$ for all $k \in \mathcal{K} \setminus \mathcal{K}''$. We conclude that $\sigma_{1i}^* = 0$ for $i \notin \mathcal{A}_1(x^*)$. A similar argument can be used to get $\sigma_{2i}^* = 0$ if $i \notin \mathcal{A}_2(x^*)$.

To prove (3.4), we consider an index $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$. If $\sigma_{1i}^* = 0$, we immediately have $\sigma_{1i}^* \sigma_{2i}^* = 0$. If $\sigma_{1i}^* > 0$, then for all $k \in \mathcal{K}$ large enough, $\sigma_{1i}^k > 0$. Then

$$\frac{\mu^k}{x_{1i}^k} > \pi_i^k x_{2i}^k, \quad \Rightarrow \quad \frac{\mu^k}{x_{2i}^k} > \pi_i^k x_{1i}^k,$$

or $\sigma_{2i}^k > 0$. Hence, $\sigma_{1i}^* \sigma_{2i}^* \geq 0$, as desired. The same argument can be used to show that, if $\sigma_{1i}^* < 0$, then $\sigma_{2i}^* < 0$, and hence $\sigma_{1i}^* \sigma_{2i}^* \geq 0$. Therefore, condition (3.4) holds, and x^* is a C-stationary point of the MPCC.

Strong Stationarity. Let $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$. If $\pi_i^k x_{2i}^k \to 0$, then

$$\sigma_{1i}^* = \lim_{k \in \mathcal{K}} \sigma_{1i}^k = \lim_{k \in \mathcal{K}} \left(\frac{\mu^k}{x_{1i}^k} - \pi_i^k x_{2i}^k \right) = \lim_{k \in \mathcal{K}} \frac{\mu^k}{x_{1i}^k} \geq 0. \tag{3.12}$$

A similar argument shows that $\sigma_{2i}^* \geq 0$. Therefore, condition (3.3i) holds, and x^* is a strongly stationary point for the MPCC (1.1). □

The proof builds on a similar proof in [15], where a similar result is derived for *exact* subproblem solves. Our result is related to the analysis in [2] (derived independently), except that we explicitly work within an interior-method framework and we do not analyze the convergence of the inner algorithm. In [2], stronger assumptions are required (e.g., that the lower-level problem satisfies a mixed-P property) to guarantee that the inner iteration always terminates.

For strong-stationarity, we required a condition on the behavior of the penalty parameter, relative to the sequences converging to the corners. This is the same condition that Scholtes requires for strong-stationarity in [23]. A simpler, though stronger, assumption on the penalties is a boundedness condition, which we use for the following corollary that corresponds to the particular case of our implementations.

Corollary 3.5 *Suppose Algorithm I is applied with a uniform (i.e., scalar-valued) penalty parameter, and let the assumptions of Theorem 3.4 hold. Then, if the sequence of penalty parameters* $\{\pi^k\}$ *is bounded,* x^* *is a strongly stationary point for* (1.1). □

In our algorithmic framework, the sequence of penalty parameters does not have to be monotone, although practical algorithms usually generate nondecreasing sequences. Monotonicity is required neither in the description of the algorithm nor in the proof. This flexibility could be exploited to correct unnecessarily large penalty parameters in practice. For theoretical purposes, on the other hand, this nonmonotonicity property will be important for deriving Theorem 3.6 in the next subsection.

3.2 Relationship to Interior-Relaxation Methods

An alternative to exact penalization for regularizing the complementarity constraints of an MPCC is to relax the complementarity constraints. This approach has been combined with interior methods in [17, 20]; we refer to it as the "interior-relaxation" method. The objective of this subsection is to show that there is a correspondence between interior-penalty and interior-relaxation approaches and that this correspondence can be exploited to give an alternative global convergence proof for an interior-relaxation method as a corollary of Theorem 3.4.

Interior-relaxation methods solve a sequence of the barrier problems associated to (1.3) with one modification, namely, the complementarity constraints (1.3e) are relaxed by introducing a parameter $\theta^k > 0$ that goes to 0 as the barrier parameter μ^k approaches 0. Effectively, a sequence of problems

$$
\begin{array}{lrcll}
\text{minimize} & \multicolumn{4}{l}{f(x) - \mu^k \sum_{i\in\mathcal{I}} \log s_i - \mu^k \sum_{i=1}^{p} \log s_{ci} - \mu^k \sum_{i=1}^{p} \log x_{1i} - \mu^k \sum_{i=1}^{p} \log x_{2i}} \\
 & c_i(x) & = & 0, & i \in \mathcal{E}, \\
\text{subject to} & c_i(x) - s_i & = & 0, & i \in \mathcal{I}, \\
 & \theta^k - x_{1i}x_{2i} - s_{ci} & = & 0, & i = 1, \dots, p,
\end{array}
\tag{3.13}
$$

has to be solved, where s_c are the slacks for the relaxed complementarity constraints, the multipliers of which are denoted by ξ. $\mathcal{L}_{\mu^k,\theta^k}$ denotes the Lagrangian of (3.13).

An approximate solution of (3.13), for some μ^k and θ^k, is given by variables $x^k, s^k, s_c^k, \lambda^k, \xi^k$, with $x_1^k > 0, x_2^k > 0, s^k > 0, s_c^k > 0, \lambda_{\mathcal{I}}^k > 0, \xi^k > 0$, satisfying the following inexact KKT system, where $\epsilon_{rel}^k > 0$ is some tolerance

$$
\begin{aligned}
\|\nabla_x \mathcal{L}_{\mu^k,\theta^k}(x^k, s^k, \lambda^k, \xi^k)\| &\leq \epsilon_{rel}^k, && \text{(3.14a)} \\
\|S^k \lambda_{\mathcal{I}}^k - \mu^k e\| &\leq \epsilon_{rel}^k, && \text{(3.14b)} \\
\|S_c^k \xi^k - \mu^k e\| &\leq \epsilon_{rel}^k, && \text{(3.14c)} \\
\|c(x^k, s^k)\| &\leq \epsilon_{rel}^k, && \text{(3.14d)} \\
\|\theta^k e - X_1^k x_2^k - s_c^k\| &\leq \epsilon_{rel}^k. && \text{(3.14e)}
\end{aligned}
$$

Theorem 3.6 *Suppose an interior-relaxation method generates an infinite sequence of solutions* $\{x^k, s^k, s_c^k, \lambda^k, \xi^k\}$ *and parameters* $\{\mu^k, \theta^k\}$ *satisfying conditions (3.14) for sequences* $\{\mu^k\}, \{\theta^k\}$ *and* $\{\epsilon_{rel}^k\}$*, all converging to 0. If* x^* *is a limit point of the sequence* $\{x^k\}$ *and* f *and* c *are continuously differentiable in an open neighborhood* $\mathcal{N}(x^*)$ *of* x^**, then* x^* *is feasible for the MPCC (1.1). If, in addition, MPCC-LICQ holds at* x^**, then* x^* *is a C-stationary point of (1.1). Moreover, if* $\xi_i^k x_{ji}^k \to 0$ *for* $j = 1, 2$ *and* $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$*, then* x^* *is a strongly stationary point of (1.1).*

Proof. We provide an indirect proof. Given sequences of variables $\{x^k, s^k, s_c^k, \lambda^k, \xi^k\}$, parameters $\{\mu^k, \theta^k\}$, and tolerances $\{\epsilon_{rel}^k\}$ satisfying the assumptions, we define sequences of parameters $\{\mu^k, \pi^k := \xi^k\}$ and tolerances $\{\epsilon_{pen}^k := \epsilon_{rel}^k, \epsilon_{comp}^k := (\theta^k + \epsilon_{rel}^k)^{1/2}\}$; for the variables, we will keep $\{x^k, s^k, \lambda^k\}$ only. Note that we have not changed the sequence of decision variables $\{x^k\}$, so the limit points are unchanged. We will show that the sequences that we just defined satisfy the assumptions of Theorem 3.4. Observe that there is no reason why the sequence of multipliers $\{\xi^k\}$ should be monotone. This is not a problem, however, because there is no monotonicity requirement for the sequence $\{\pi^k\}$ in Theorem 3.4, as noted earlier.

First, it is clear that $\{\mu^k\}, \{\epsilon_{pen}^k\}, \{\epsilon_{comp}^k\}$ all converge to 0. Next, it is easy to see that

$$
\nabla_x \mathcal{L}_{\mu^k,\pi^k}(x^k, s^k, \lambda^k) = \nabla_x \mathcal{L}_{\mu^k,\theta^k}(x^k, s^k, \lambda^k, \xi^k).
$$

This, together with conditions (3.14a), (3.14b), and (3.14d), yields (2.7).

We assume, without loss of generality, that the infinity norm is used for (2.8). Combining (3.14e) with $\min\{x_1^k, x_2^k\} \leq x_1^k$ and $\min\{x_1^k, x_2^k\} \leq x_2^k$, we get

$$
\begin{aligned}
0 \leq \min\{x_{1i}^k, x_{2i}^k\} &\leq (x_{1i}^k x_{2i}^k)^{1/2} \\
\leq (x_{1i}^k x_{2i}^k + s_{ci}^k)^{1/2} &\leq (\theta^k + \epsilon_{rel}^k)^{1/2} = \epsilon_{comp}^k.
\end{aligned}
$$

Therefore, the sequence $\{x^k, s^k, \lambda^k\}$, with corresponding parameters $\{\mu^k, \pi^k\}$, satisfies conditions (2.7) and (2.8) for all k. The conclusions follow from a direct application of Theorem 3.4. □

A similar global result is proved directly in [17], under somewhat different assumptions. The key for the proof presented here is that there exists a one-to-one correspondence between KKT points of problems (2.2) and (3.13), which is easily seen by comparing the corresponding first-order conditions. In fact, this one-to-one relation between KKT points of relaxation and penalization schemes can be extended to general NLPs. Such an extension is useful because some convergence results can be derived directly for one approach only and then extended to the alternative regularization scheme in a simple way.

4 Local Convergence Analysis

In this section, we show that, if the iterates generated by Algorithm I approach a solution x^* of the MPCC that satisfies certain regularity conditions and if the penalty parameter is sufficiently large, then this parameter is never updated, and the iterates converge to x^* at a superlinear rate.

We start by defining a second-order sufficient condition for MPCCs (see [22]). For this purpose, we define the Lagrangian

$$\mathcal{L}(x, \lambda, \sigma_1, \sigma_2) = f(x) - \lambda_{\mathcal{E}}^T c_{\mathcal{E}}(x) - \lambda_{\mathcal{I}}^T c_{\mathcal{I}}(x) - \sigma_1^T x_1 - \sigma_2^T x_2. \tag{4.1}$$

Definition 4.1 *We say that the* MPCC second-order sufficient condition *(MPCC-SOSC) holds at x^* if x^* is a strongly stationary point of (1.1) with multipliers $\lambda^*, \sigma_1^*, \sigma_2^*$ and*

$$d^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*, \sigma_1^*, \sigma_2^*)\, d > 0 \tag{4.2}$$

for all critical direction $d \in T_{mpcc}(x^*)$, with $\|d\| = 1$, satisfying

$$\begin{aligned}
\nabla f(x)^T d &= 0, && (4.3a)\\
\nabla c_i(x)^T d &= 0 \ \ \textit{for all} \ \ i \in \mathcal{E}, && (4.3b)\\
\nabla c_i(x)^T d &\geq 0 \ \ \textit{for all} \ \ i \in \mathcal{A}(x), && (4.3c)\\
\min_{\{j : x_{ji} = 0\}} \{d_{ji}\} &= 0 \ \ \textit{for all} \ \ i = 1, \ldots, p. && (4.3d)
\end{aligned}$$

Notice that (4.3d) is a convenient way to summarize the following conditions, which characterize the set of feasible directions with respect to the complementarity constraints: If $x_{1i} = 0, x_{2i} > 0$, then $d_{1i} = 0$ and d_{2i} is free; if $x_{2i} = 0, x_{1i} > 0$, then $d_{2i} = 0$ and d_{1i} is free; and if $x_{1i} = x_{2i} = 0$, then $0 \leq d_{1i} \perp d_{2i} \geq 0$.

For the local analysis, we make the following assumptions.

Assumptions 4.2 *There exists a strongly stationary point x^* of the MPCC (1.1), with multipliers $\lambda^*, \sigma_1^*, \sigma_2^*$, satisfying the following conditions:*

1. *f and c are twice Lipschitz continuously differentiable in an open neighborhood of x^*.*

2. *MPCC-LICQ holds at x^*.*

3. *The following primal-dual strict complementarity holds at x^*: $\lambda_i^* \neq 0$ for all $i \in \mathcal{E} \cup \mathcal{A}_c(x^*)$, and $\sigma_{ji}^* > 0$ for all $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$, for $j = 1, 2$.*

4. *MPCC-SOSC holds at* x^*.

The following lemma shows that the penalty formulation inherits the desirable properties of the MPCC for a sufficiently large penalty parameter. The multipliers for the bound constraints $x_1 \geq 0, x_2 \geq 0$ of the penalty problem (2.1) will be denoted by $\nu_1 \geq 0, \nu_2 \geq 0$, respectively.

Lemma 4.3 *If Assumptions 4.2 hold at* x^* *and* $\pi > \pi^*$, *where*

$$\pi^* = \pi^*(x^*, \sigma_1^*, \sigma_2^*) = \max\left\{0, \max_{\{i:x_{1i}^*>0\}} \frac{-\sigma_{2i}^*}{x_{1i}^*}, \max_{\{i:x_{2i}^*>0\}} \frac{-\sigma_{1i}^*}{x_{2i}^*}\right\}, \tag{4.4}$$

then it follows that

1. *LICQ holds at* x^* *for* (2.1).
2. x^* *is a KKT point of* (2.1).
3. *Primal-dual strict complementarity holds at* x^* *for* (2.1); *that is,* $\lambda_i^* \neq 0$ *for all* $i \in \mathcal{E} \cup \mathcal{A}_c(x^*)$ *and* $\nu_{ji}^* > 0$ *for all* $i \in \mathcal{A}_j(x^*)$, *for* $j = 1, 2$.
4. *The second-order sufficiency condition holds at* x^* *for* (2.1).

Proof. LICQ at x^* for (2.1) follows from the definition of MPCC-LICQ.

The proof of Part 2 is similar to the proof of Proposition 4.1 in [11]. The key for the proof is the relationship between the multipliers σ_1^*, σ_2^* of (1.1) and $\nu_1^* \geq 0, \nu_2^* \geq 0$ of (2.1), given by

$$\nu_1^* = \sigma_1^* + \pi x_2^* \quad \text{and} \quad \nu_2^* = \sigma_2^* + \pi x_1^*. \tag{4.5}$$

The result is evident when the strongly stationarity conditions (3.3) and the first-order KKT conditions of (2.1) are compared, except for the nonnegativity of ν_1^* and ν_2^*. To see that $\nu_1^*, \nu_2^* \geq 0$, suppose first that $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$. In that case, from (4.5), we have $\nu_{ji} = \sigma_{ji}, j = 1, 2$, and the nonnegativity follows directly from (3.3i). If, on the other hand, $i \notin \mathcal{A}_2(x^*)$, then (4.5) and $\pi > \pi^*$ imply

$$\nu_{1i}^* = \sigma_{1i}^* + \pi x_{2i}^* > \sigma_{1i}^* + \frac{-\sigma_{1i}^*}{x_{2i}^*} x_{2i}^* = 0. \tag{4.6}$$

The same argument applies for $i \notin \mathcal{A}_1(x^*)$, which completes the proof of Part 2.

Note that $\pi \geq \pi^*$ suffices for the nonnegativity of ν_1, ν_2. The strict inequality $\pi > \pi^*$ is required for Part 3; that is, we need it for primal-dual strict complementarity at x^* for (2.1). In fact, (4.6) yields primal-dual strict complementarity for $i \notin \mathcal{A}_2(x^*)$ (and a similar argument works for $i \notin \mathcal{A}_1(x^*)$). For $i \in \mathcal{E} \cup \mathcal{A}_c(x^*)$, strict complementarity comes directly from the assumptions. For $i \in \mathcal{A}_2(x^*) \cap \mathcal{A}_1(x^*)$, relation (4.5) shows that $\nu_{ji}^* = \sigma_{ji}^*, j = 1, 2$, which is positive by Assumption 4.2 (3).

For Part 4, Assumption 4.2 (3) implies that the multipliers of the complementarity variables satisfy $\nu_{1i}^* + \nu_{2i}^* > 0$ for all $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$, which, together with $\pi > \pi^*$, constitutes a sufficient condition for SOSC of the penalty problem (2.1); see [18] for details. Therefore, SOSC hold at x^* for (2.1). $\square$

We note that Assumption 4.2 (3) can be weakened and still get SOSC for the penalized problem (2.1). In [18], two alternative sufficient conditions for SOSC of (2.1) are given. The first involves $\nu_{1i}^* + \nu_{2i}^* > 0$ for all $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$ (which is called partial strict complementarity in [21]) and $\pi > \pi^*$. The second condition involves a possibly larger penalty parameter and shows how the curvature term of the complementarity constraint $x_1^T x_2$ can be exploited to ensure the penalized problem satisfies a second-order condition. We state the result here for completeness.

Lemma 4.4 *Let MPCC-SOSC hold at x^*, and assume that one of the following conditions holds:*

1. $\pi > \pi^*$ *and* $\nu_{1i}^* + \nu_{2i}^* > 0$ *for all* $i \in \mathcal{A}_1(x^*) \cap \mathcal{A}_2(x^*)$;
2. $\pi > \max\{\pi^*, \pi_{SO}\}$, *with* π_{SO} *as in (4.7) below,*

then SOSC holds at x^* for (2.1).

Proof. We only sketch the proof here (a complete derivation can be found in [18]). In Case 1, it is not hard to see that the tangent cone of the penalty problem $T_\pi(x^*)$ and the tangent cone of the MPCC $T_{mpcc}(x^*)$ coincide, from which the result easily follows. In Case 2, following Scheel and Scholtes [22] or Ralph and Wright [21], we define the following partition of $T_\pi(x^*)$, for any $\epsilon > 0$:

$$\begin{aligned} T_{\pi,\epsilon}^+(x) &:= T_\pi(x) \cap \Big\{ d : \min\{d_{1i}, d_{2i}\} < \epsilon, \ \forall i \in \mathcal{A}_1(x) \cap \mathcal{A}_2(x) \Big\}, \\ T_{\pi,\epsilon}^-(x) &:= T_\pi(x) \setminus T_{\pi,\epsilon}^+(x). \end{aligned}$$

If we let

$$\beta = \min\Big\{ d^T \nabla_{xx}^2 \mathcal{L}(x,\lambda)\, d : d \in T_{\pi,\epsilon}^-(x) \Big\}$$

and

$$\pi_{SO} = \max\left\{ 0, -\frac{\beta}{2\epsilon^2} \right\}, \tag{4.7}$$

then $\pi > \max\{\pi^*, \pi_{SO}\}$ ensures that Hessian has positive curvature over $T_\pi(x)$. □

We now show that an adequate penalty parameter stabilizes near a regular solution and superlinear convergence takes place.

We group primal and dual variables in a single vector $z = (x, s, \lambda)$. Given a strongly stationary point x^* with multipliers $\lambda^*, \sigma_1^*, \sigma_2^*$, we associate to it the triplet $z^* = (x^*, s^*, \lambda^*)$, where $s^* = c_{\mathcal{I}}(x^*)$. We also group the left-hand side of (2.5) in the function

$$F_\mu(z;\pi) = \begin{pmatrix} \nabla_x \mathcal{L}_{\mu,\pi}(x,\lambda) \\ S\lambda_{\mathcal{I}} - \mu e \\ c(x,s) \end{pmatrix}. \tag{4.8}$$

At every inner iteration in Step 2 of Algorithm I, a step d is computed by solving a system of the form

$$\nabla F_\mu(z;\pi) d = -F_\mu(z;\pi). \tag{4.9}$$

Note (2.7) is equivalent to $\|F_\mu(z;\pi)\| \le \epsilon_{pen}$.

The following theorem shows that there are practical implementations of Algorithm I that, near a regular solution x^* of the MPCC and for a sufficiently large penalty parameter, satisfy the stopping tests (2.7) and (2.8) at every iteration, with no backtracking and no updating of the penalty parameter. Using this fact one can easily show that the iterates converge to x^* superlinearly. To state this result, we introduce the following notation. Let z be an iterate satisfying $\|F_\mu(z;\pi)\| \le \epsilon_{pen}$ and $\|\min\{x_1, x_2\}\| \le \epsilon_{comp}$. We define z^+ to be the new iterate computed using a barrier parameter $\mu^+ < \mu$, namely,

$$z^+ = z + d, \qquad \text{with} \quad F_{\mu^+}(z;\pi) d = -F_{\mu^+}(z;\pi). \tag{4.10}$$

Theorem 4.5 *Suppose that Assumptions 4.2 hold at a strongly stationary point x^*. Assume that $\pi > \pi^*$, with π^* given by (4.4) and that the tolerances $\epsilon_{pen}, \epsilon_{comp}$ in Algorithm I are functions of μ that converge to 0 as $\mu \to 0$. Furthermore, assume that the barrier parameter and these tolerances are updated so that the following limits hold as $\mu \to 0$:*

$$\frac{(\epsilon_{pen}+\mu)^2}{\epsilon_{pen}^+} \;\to\; 0, \tag{4.11a}$$

$$\frac{(\epsilon_{pen}+\mu)^2}{\mu^+} \;\to\; 0, \tag{4.11b}$$

$$\frac{\mu^+}{\epsilon_{comp}^+} \;\to\; 0. \tag{4.11c}$$

Assume also that

$$\frac{\mu^+}{\|F_0(z;\pi)\|} \to 0, \quad as \quad \|F_0(z;\pi)\| \to 0. \tag{4.12}$$

Then, if μ is sufficiently small and z is sufficiently close to z^*, the following conditions hold:

1. *The stopping criteria (2.7) and (2.8), with parameters $\mu^+, \epsilon_{pen}^+, \epsilon_{comp}^+$ and π, are satisfied at z^+.*

2. *$\|z^+ - z^*\| = o(\|z - z^*\|)$.*

Proof. By the implicit function theorem, Assumptions 4.2, the condition $\pi > \pi^*$, and Lemma 4.3, it follows that, for all sufficiently small μ, there exists a solution $z^*(\mu)$ of problem (2.2); see, for example, [12]. If, in addition, z is close to z^*, then

$$L_1\mu \;\le\; \|z^* - z^*(\mu)\| \;\le\; U_1\mu, \tag{4.13}$$

$$L_2\|F_\mu(z;\pi)\| \;\le\; \|z - z^*(\mu)\| \;\le\; U_2\|F_\mu(z;\pi)\|. \tag{4.14}$$

(Condition (4.13) is Corollary 3.14 in [12], and (4.14) is Lemma 2.4 in [5].) Here and in the rest of the proof L_i and U_i denote positive constants, and we assume without loss of generality that $\|\cdot\|$ denotes the infinity norm. By standard Newton analysis (see, e.g., Theorem 2.3 in [5]) we have that

$$\|z^+ - z^*(\mu^+)\| \le\; U_3\|z - z^*(\mu^+)\|^2. \tag{4.15}$$

We will also use the inequality

$$\|z^+ - z^*(\mu^+)\| \le U_4(\epsilon_{pen}+\mu)^2, \tag{4.16}$$

which is proved as follows:

$$\begin{aligned}
\|z^+ - z^*(\mu^+)\| &\le U_3\|z - z^*(\mu^+)\|^2 \quad \text{(from (4.15))}\\
&\le U_3\left(\|z - z^*(\mu)\| + \|z^*(\mu) - z^*\| + \|z^* - z^*(\mu^+)\|\right)^2\\
&\le U_3\left(U_2\|F_\mu(z;\pi)\| + U_1\mu + U_1\mu^+\right)^2 \quad \text{(from (4.14) and (4.13))}\\
&\le U_4\left(\epsilon_{pen}+\mu\right)^2,
\end{aligned}$$

where the last inequality holds because z satisfies (2.7) with μ, ϵ_{pen}, π and because $\mu^+ < \mu$.

We now show that (2.7) holds at z^+, with parameters $\mu^+, \epsilon_{pen}^+, \pi$, as follows:

$$\begin{aligned}\|F_{\mu^+}(z^+;\pi)\| &\leq L_2^{-1}\|z^+ - z^*(\mu^+)\| \quad \text{(from (4.14))}\\ &\leq L_2^{-1}U_4(\epsilon_{pen}+\mu)^2 \quad \text{(from (4.16))}\\ &= L_2^{-1}U_4\frac{(\epsilon_{pen}+\mu)^2}{\epsilon_{pen}^+}\epsilon_{pen}^+\\ &\leq \epsilon_{pen}^+ \quad \text{(from (4.11a))}.\end{aligned}$$

To see that $x_1^+ > 0$, we can apply (4.16) componentwise to get

$$|x_{1i}^+ - x_{1i}^*(\mu^+)| \leq U_4(\epsilon_{pen}+\mu)^2,$$

from which we have that

$$x_{1i}^+ \geq x_{1i}^*(\mu^+) - U_4(\epsilon_{pen}+\mu)^2. \tag{4.17}$$

If $x_{1i}^* = 0$, we have by (4.13) and the positivity of $x_{1i}^*(\mu^+)$ that $x_{1i}^*(\mu^+) \geq L_1\mu^+$. Therefore

$$\begin{aligned}x_{1i}^+ &\geq L_1\mu^+ - U_4\frac{(\epsilon_{pen}+\mu)^2}{\mu^+}\mu^+ \quad \text{(from (4.13))}\\ &\geq L_5\mu^+ \quad \text{(from (4.11b))}.\end{aligned}$$

If, on the other hand, $x_{1i}^* > 0$, then from (4.13) and (4.17), we get

$$\begin{aligned}x_{1i}^+ &\geq x_{1i}^* - U_1\mu^+ - U_4(\epsilon_{pen}+\mu)^2\\ &= x_{1i}^* - U_1\mu^+ - U_4\frac{(\epsilon_{pen}+\mu)^2}{\mu^+}\mu^+\\ &> 0 \quad \text{(from (4.11b))}.\end{aligned}$$

Similar arguments can be applied to get $x_2^+ > 0, s^+ > 0, \lambda_{\mathcal{I}}^+ > 0$.

To prove (2.8), we first observe that

$$\begin{aligned}\|z^+ - z^*\| &\leq \|z^+ - z^*(\mu^+)\| + \|z^*(\mu^+) - z^*\|\\ &\leq U_4(\epsilon_{pen}+\mu)^2 + U_1\mu^+ \quad \text{(from (4.16) and (4.13))}\\ &= U_4\frac{(\epsilon_{pen}+\mu)^2}{\mu^+}\mu^+ + U_1\mu^+\\ &\leq U_5\mu^+ \quad \text{(from (4.11b))}.\end{aligned} \tag{4.18}$$

Let $i \in \{1,\ldots,p\}$, and assume, without loss of generality, that $x_{1i}^* = 0$. Then,

$$\begin{aligned}|\min\{x_{1i}^+, x_{2i}^+\}| &= \min\{x_{1i}^+, x_{2i}^+\} \quad \text{(because } x_1^+ > 0, x_2^+ > 0)\\ &\leq x_{1i}^+ = |x_{1i}^+ - x_{1i}^*|\\ &\leq U_5\mu^+ \quad \text{(from (4.18))}\\ &= U_5\frac{\mu^+}{\epsilon_{comp}^+}\epsilon_{comp}^+ \leq \epsilon_{comp}^+,\end{aligned}$$

where the last inequality follows from (4.11c). Since this argument applies to all $i \in \{1,\ldots,p\}$, we have that (2.8) is satisfied. This concludes the proof of Part 1 of the theorem.

For Part 2, we have that

$$\begin{aligned}
\|z^+ - z^*\| &\leq \|z^+ - z^*(\mu^+)\| + \|z^*(\mu^+) - z^*\| \\
&\leq U_3\|z - z^*(\mu^+)\|^2 + U_1\mu^+ \quad \text{(from (4.15) and (4.13))} \\
&\leq U_3\left(\|z - z^*\| + \|z^* - z^*(\mu^+)\|\right)^2 + U_1\mu^+ \\
&\leq U_3\left(2\|z - z^*\|^2 + 2\|z^* - z^*(\mu^+)\|^2\right) + U_1\mu^+ \\
&\leq 2U_3\|z - z^*\|^2 + 2U_3(U_1\mu^+)^2 + U_1\mu^+ \quad \text{(from (4.15))} \\
&\leq U_6\left(\|z - z^*\|^2 + \mu^+\right).
\end{aligned}$$

This implies that

$$\frac{\|z^+ - z^*\|}{\|z - z^*\|} \leq U_6\left(\|z - z^*\| + \frac{\mu^+}{\|z - z^*\|}\right).$$

We apply the left inequality in (4.14), evaluated at z and with barrier parameter 0, to get

$$\frac{\|z^+ - z^*\|}{\|z - z^*\|} \leq U_6\left(\|z - z^*\| + \frac{1}{L_2}\frac{\mu^+}{\|F_0(z;\pi)\|}\right). \tag{4.19}$$

Note that, from (4.14), if $\|z - z^*\|$ is sufficiently small, so is $\|F_0(z;\pi)\|$, which in turn, by (4.12), implies that the second term in the right-hand side is also close to 0. Hence, if $\|z - z^*\|$ is sufficiently small, it follows that the new iterate z^+ will be even closer to z^*. Moreover, by applying (4.19) recursively, we conclude that the iterates will converge to z^*. From the same relation, it is clear that this convergence happens at a superlinear rate, which concludes the proof. □

Many practical updating rules for μ and ϵ_{pen} satisfy conditions (4.11a)–(4.12). For example, we can define $\epsilon_{pen} = \theta\mu$ with $\theta \in [0, \sqrt{|\mathcal{I}|})$. In this case, it is not difficult to show [5] that (4.11a), (4.11b), (4.12) are satisfied if we update μ by the rule

$$\mu^+ = \mu^{1+\delta}, \quad 0 < \delta < 1.$$

The same is true for the rule

$$\mu^+ = \|F_\mu(z;\pi)\|^{1+\delta}, \quad 0 < \delta < 1.$$

A simple choice for ϵ_{comp} that ensures (4.11c) is μ^γ, with $0 < \gamma < 1$.

5 Implementation and Numerical Results

We begin by describing two practical implementations of Algorithm I that use different strategies for updating the penalty parameter. The first algorithm, *Classic*, is described in Figure 3; it updates the penalty parameter only after the barrier problem is solved, and provided the complementarity value has decreased sufficiently as stipulated in Step 3. We index by k the major iterates that satisfy (2.7) and (2.8); this notation is consistent with that of Section 2. We use j to index the sequence of all minor iterates generated by the algorithm Classic. Since $\gamma \in (0,1)$, the tolerance ϵ^k_{comp} defined in Step 1 converges to 0 slower than does $\{\mu^k\}$; this is condition (4.11c) in Theorem 4.5.

In the numerical experiments, we use $\gamma = 0.4$ for the following reason: The distance between iterates x^k and the solution x^* is proportional to $\sqrt{\mu^k}$, if primal-dual strict complementarity does not hold at x^*. By choosing the complementarity tolerance to be $\epsilon^k_{comp} = (\mu^k)^{0.4}$, we ensure that

Algorithm Classic: A Practical Interior-Penalty Method for MPCCs

Initialization: Let $z^0 = (x^0, s^0, \lambda^0)$ be the initial primal and dual variables. Choose an initial penalty π^0 and a parameter $\gamma \in (0,1)$. Set $j = 0, k = 1$.

repeat (barrier loop)

1. Choose a barrier parameter μ^k, a stopping tolerance ϵ^k_{pen}, let $\epsilon^k_{comp} = (\mu^k)^\gamma$ and let $\pi^k = \pi^{k-1}$.

2. **repeat** (inner iteration)

   (a) Let $j \leftarrow j+1$ and let the current point be $z^c = z^{j-1}$.

   (b) Using a globally convergent method, compute a primal-dual step d^j based on the KKT system (2.4), with $\mu = \mu^k, \pi = \pi^k$ and $z = z^c$.

   (c) Let $z^j = z^c + d^j$.

   until conditions (2.7) are satisfied for ϵ^k_{pen}.

3. **If** $\|\min\{x_1^j, x_2^j\}\| \le \epsilon^k_{comp}$, let $z^k = z^j$, set $k \leftarrow k+1$;
   else set $\pi^k \leftarrow 10\pi^k$ and go to Step 2.

until a stopping test for the MPCC is satisfied.

Figure 3: Description of the Algorithm Classic.

the test (2.8) can be satisfied in this case. All other details of the interior method will be described below.

The second algorithm we implemented, *Dynamic*, is described in Figure 4. It is more flexible than Classic in that it allows changes in the penalty parameter at every iteration of the inner algorithm. The strategy of Step 2(c) is based on the following considerations: If the complementarity pair is relatively small according to the preset tolerance ϵ^k_{comp}, then there is no need to increase π. Otherwise, we check whether the current complementarity value, $x_1^{jT} x_2^j$, is less than a fraction of the maximum value attained in the m previous iterations (in our tests, we use $m = 3$ and $\eta = 0.9$). If not, we increase the penalty parameter. We believe that it is appropriate to look back at several previous steps, and not require decrease at every iteration, because the sequence $\{x_1^{jT} x_2^j\}$ is frequently nonmonotone, especially for problems in which primal-dual strict complementarity is violated (see, e.g., Figure 6(a)). Note that the algorithms Classic and Dynamic are both special cases of Algorithm I of Section 2.

We implemented these two algorithms as an extension of our MATLAB solver IPM-D. This solver is based on the interior algorithm for nonlinear programming described in [25], with one change: IPM-D handles negative curvature by adding a multiple of the identity to the Hessian of the Lagrangian, as in [24], instead of switching to conjugate-gradient iterations. We chose to work with IPM-D because it is a simple interior solver that does not employ the regularizations, scalings, and other heuristics used in production packages that alter the MPCC, making it harder to assess the impact of the approach proposed in this paper.

In our implementation, all details of the interior-point iteration, such as the update of the barrier parameter, the step selection, and the choice of merit function, are handled by IMP-D. The

Algorithm Dynamic: A Practical Interior-Penalty Method for MPCCs

Initialization: Let $z^0 = (x^0, s^0, \lambda^0)$ be the initial primal and dual variables. Choose an initial penalty π^0, parameters $\gamma, \eta \in (0,1)$, and an integer $m \geq 1$. Set $j = 0,\ k = 1$.

repeat (barrier loop)

1. Choose a barrier parameter μ^k, a stopping tolerance ϵ^k_{pen} and let $\epsilon^k_{comp} = (\mu^k)^\gamma$.
2. **repeat** (inner iteration)

   (a) Set $j \leftarrow j+1$, let the current point be $z^c = z^{j-1}$, and let $\pi^j = \pi^{j-1}$.

   (b) Using a globally convergent method, compute a primal-dual step d^j based on the KKT system (2.4), with $\mu = \mu^k, \pi = \pi^k$ and $z = z^c$.

   (c) If $\|\min\{x_1^j, x_2^j\}\| > \epsilon^k_{comp}$ and

   $$x_1^{jT} x_2^j > \eta \max\left\{x_1^{jT} x_2^j, \ldots, x_1^{(j-m+1)T} x_2^{(j-m+1)}\right\}, \tag{5.1}$$

   then set $\pi^j \leftarrow 10\pi^j$, adjust λ^j and go to Step 2.

   until conditions (2.7) are satisfied for ϵ^k_{pen}.
3. **If** $\|\min\{x_1^j, x_2^j\}\| \leq \epsilon^k_{comp}$, let $z^k = z^j$ and $k = k+1$
   else set $\pi^k \leftarrow 10\pi^k$ and go to Step 2

until a stopping test for the MPCC is satisfied

Figure 4: Description of the Algorithm Dynamic.

main point of this section is to demonstrate how to adapt an existing interior-point method to solve MPCCs efficiently and reliably.

We tested the algorithms on a collection of 74 problems, listed in Table 3, where we report the number of variables n (excluding slacks), the number of constraints m (excluding complementarity constraints), and the number of complementarity constraints p. Most of these problems are taken from the MacMPEC collection [16], but we added a few problems to test the sensitivity of our implementations to bad scalings in the MPCC. All the methods tested were implemented in IPM-D, and since this MATLAB program is not suitable for very large problems, we restricted our test set to a sample of problems with fewer than 1,000 variables. We report results for four methods, which are labeled in the figures as follows:

NLP is the direct application of the interior code IPM-D to the nonlinear programming formulation (1.3) of the MPCC.

Fixed is a penalty method in which IPM-D is applied to (2.1) with a fixed penalty of 10^4. The penalty parameter is not changed.

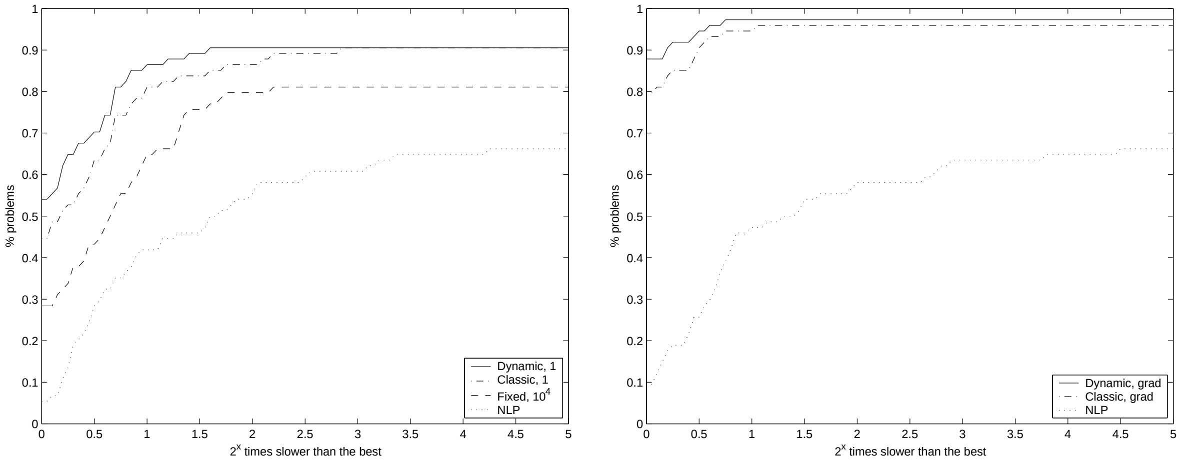
Classic is the algorithm given in Figure 3, implemented in the IPM-D solver.

Dynamic is the algorithm given in Figure 4 implemented in the IPM-D solver.

In Figure 5 we report results for these four methods in terms of total number of iterations (indexed by j). The figures use the logarithmic performance profiles described in [8]. An important

Table 3: Test Problem Characteristics.

Name	n	m	p	Name	n	m	p
bar-truss-3	29	22	6	bard1	5	1	3
bard3	6	3	2	bilevel1	10	9	6
bilevel3	12	7	4	bilin	8	1	6
dempe	4	2	2	design-cent-1	12	9	3
design-cent-4	22	9	12	desilva	6	2	2
df1	2	2	1	ex9.1.1	13	12	5
ex9.1.3	23	21	6	ex9.1.5	13	12	5
ex9.1.6	14	13	6	ex9.1.7	17	15	6
ex9.1.8	14	12	5	ex9.1.9	12	11	5
ex9.1.10	14	12	5	ex9.2.1	10	9	4
ex9.2.2	10	11	4	ex9.2.4	8	7	2
ex9.2.5	8	7	3	ex9.2.6	16	12	6
ex9.2.7	10	9	4	ex9.2.8	6	5	2
ex9.2.9	9	8	3	flp2	4	2	2
flp4-1	80	60	30	gauvin	3	0	2
gnash10	13	4	8	gnash11	13	4	8
gnash12	13	4	8	gnash13	13	4	8
gnash14	13	4	8	gnash15	13	4	8
gnash16	13	4	8	gnash17	13	4	8
gnash18	13	4	8	gnash19	13	4	8
hakonsen	9	8	4	hs044-i	20	14	10
incid-set1-16	485	491	225	incid-set2c-16	485	506	225
kth1	2	0	1	kth2	2	0	1
kth3	2	0	1	liswet1-050	152	103	50
outrata31	5	0	4	outrata32	5	0	4
outrata33	5	0	4	outrata34	5	0	4
pack-comp1-16	332	151	315	pack-comp2c-16	332	166	315
pack-rig1c-16	209	148	192	pack-rig2-16	209	99	192
pack-rig3-16	209	99	192	portfl-i-2	87	25	12
portfl-i-6	87	25	12	qpec-100-1	105	102	100
ralph1	2	0	1	ralph2	2	0	1
ralphmod	104	0	100	scale1	2	0	2
scale2	2	0	2	scale3	2	0	2
scale4	2	0	2	scale5	2	0	2
scholtes1	3	1	1	scholtes2	3	1	1
scholtes3	2	0	2	scholtes4	3	2	2
scholtes5	3	2	2	tap-09	86	68	32



(a) $\pi = 1$, initially. (b) Initial π based on gradient norm.

Figure 5: Performance of 4 methods for solving MPCCs.

choice in the algorithms Classic and Dynamic is the initial value of π. In Figure 5(a) we show results for $\pi^0 = 1$, and in Figure 5(b) for $\pi^0 = \|\nabla f(x^0)\|$ (the latter rule is also used, for example, in the elastic phase of SNOPT [19]).

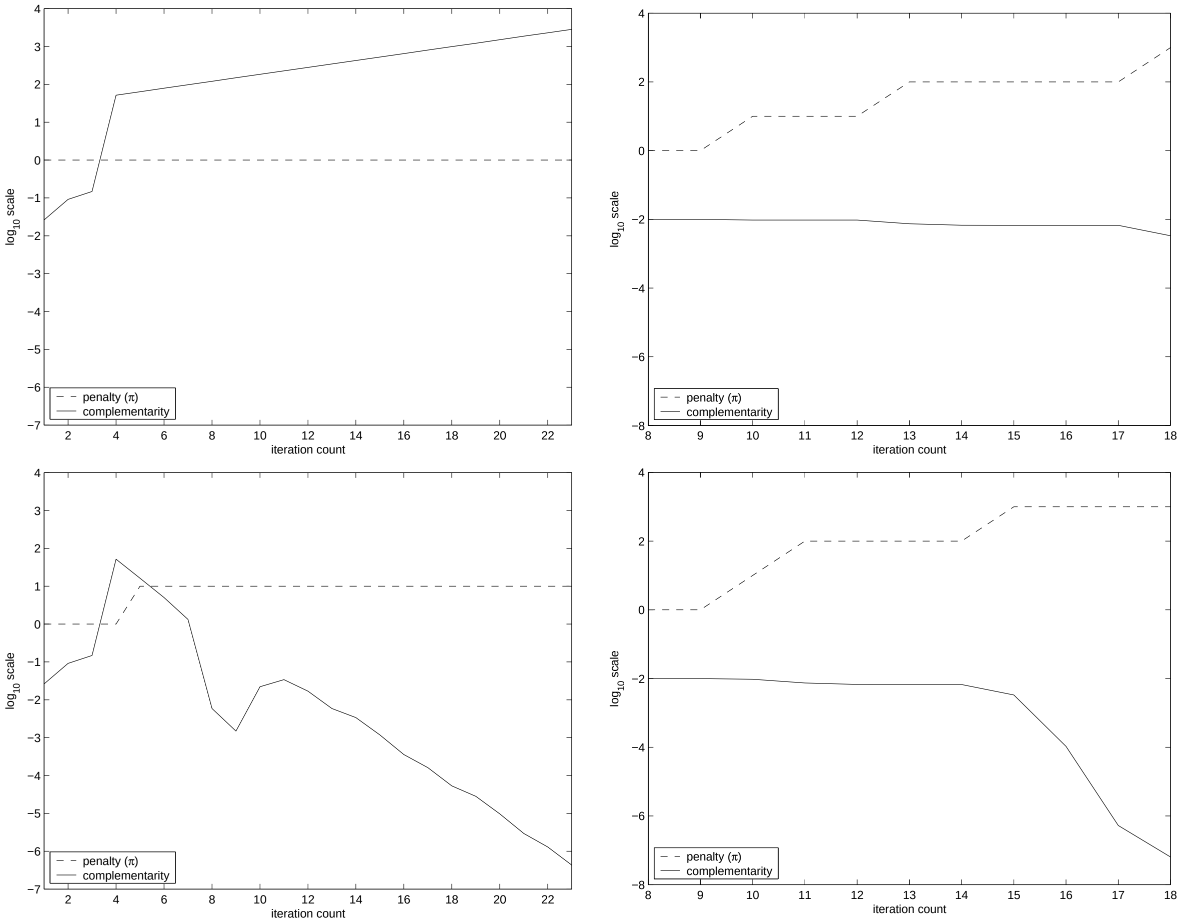
Comparing the results in Figure 5, we note that the direct application of the interior method (option NLP) gives the poorest results. Option Fixed (dashed curve in Figure 5(a)) is significantly more robust and efficient, but it is clearly surpassed by the Classic and Dynamic methods. Apart from the fact that option FIXED fails more often, it requires, in general, more iterations to solve each barrier problem. In extreme cases, like `bar-truss-3`, Dynamic (with $\pi^0 = 1$) solves the first barrier problem in 15 iterations, whereas Fixed needs 43 iterations. Moreover, we frequently find that, near a solution, the algorithms Classic and Dynamic take one iteration per barrier problem, as expected, whereas Fixed keeps taking several steps to find a solution every time μ is updated.

Classic and Dynamic perform remarkably well with the initial value $\pi^0 = 1$ (Figure 5(a)), which is too small for most problems. Both algorithms, however, adjust π efficiently, especially Dynamic. The choice $\pi^0 = \|\nabla f(x^0)\|$ attempts to estimate the norm of the multipliers and can certainly be unreliable. Nonetheless, it performed very well on this test set. We note from Figure 5(b) that the performance of both algorithms improves for $\pi^0 = \|\nabla f(x^0)\|$.

The MacMPEC collection is, however, composed almost exclusively of well-scaled problems, and `ralph2` is the only problem that becomes unbounded for the initial penalty (with either initialization of π). As a result, Dynamic does not differ significantly from Classic on this test set. We therefore take a closer look at the performance of these methods on problems `ralph2` and `scale1` discussed in Section 2. We believe that the results for these examples support the choice of Dynamic over Classic for practical implementations.

Example 1 (`ralph2`), revisited. Figure 6(a) plots the complementarity measure $(x_1^{jT} x_2^j)$ (continuous line) and the value of the penalty parameter π^j (dashed line) for problem (2.9) (using a $\log_{10}$ scale). The top figure corresponds to Classic and the bottom figure to Dynamic; both used an initial penalty parameter of 1. Recall that $\pi = 1$ gives rise to an unbounded penalty problem. The two algorithms perform identically up to iteration 4. At that point, the Dynamic algorithm increases π, whereas the Classic algorithm never changes π, because it never solves the first barrier problem. Classic fails on this problem, and complementarity grows without bound. $\square$

Example 2 (`scale1`), revisited. Problem (2.11) requires $\pi \geq 200$ so that the penalty problem re-



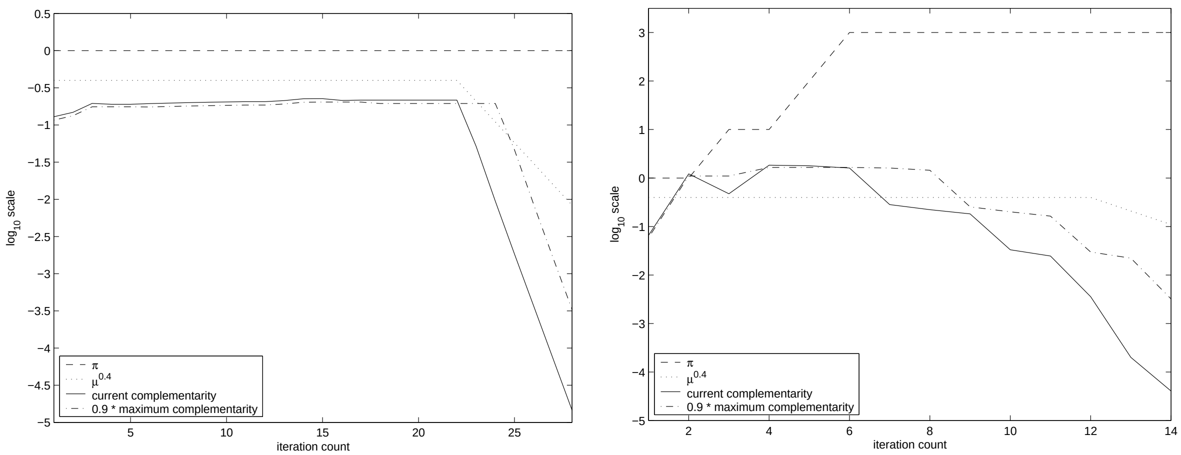
(a) `ralph2`: Classic (top) vs. Dynamic (bottom). (b) `scale1`: Classic (top) vs. Dynamic (bottom).

Figure 6: Evolution of penalty and complementarity values ($\log_{10}$ scale).

covers the solution of the MPCC. We again initialize Dynamic and Classic with $\pi = 1$. Figure 6(b) plots the complementarity measure and the penalty parameter values for both implementations. The two algorithms increase π three times (from 1 to 10, to 100, to 1000). While the Classic implementation (top figure) is performing the third update of π, the Dynamic implementation (bottom figure) has converged to the solution. The Dynamic algorithm detects earlier that complementarity has stagnated (and is not sufficiently small) and takes corrective action by increasing π. Not all plateaus mean that π needs to be changed, however, as we discuss next. □

To study in more detail the algorithm Dynamic, we consider two other problems, **`bard3`** and **`bilin`**, from the MacMPEC collection (we initialize the penalty parameter to 1, as before).

Example 4 (`bard3`). Figure 7(a) shows the results for problem **`bard3`**. The continuous line plots $x_1^{jT} x_2^j$, and the dashed-dotted line plots the $\eta = 0.9$ curve of the maximum value of $x_1^{iT} x_2^i$, over the last three iterations. Note that the complementarity measure increases at the beginning and does not decrease during the first 20 iterations. However, Dynamic does not increase the value of π (dashed line) because the value of complementarity is small enough, compared to the threshold $\mu^{0.4}$, the dotted line plots $(\mu^j)^{0.4}$. This is the correct action because, if the algorithm increased π simply because the maximum value of complementarity over the last three iterations is not decreasing, π would take on large values that would slow the iteration and could even cause failure. □



(a) Problem `bard3`. (b) Problem `bilin`.

Figure 7: Illustration of the Dynamic updating strategy.

Example 5 (`bilin`). A different behavior is observed for problem `bilin`; see Figure 7(b). The value of complementarity (continuous line) not only lies above the line that plots the $\eta = 0.9$ curve of the maximum complementarity over the last three iterations (dashed-dotted line) but is also above the line plotting $(\mu^j)^{0.4}$. Thus the penalty parameter is increased quickly (dashed line). The sufficient reduction condition is satisfied at iteration 3 but is then again violated, so π is increased again, until complementarity finally starts converging to zero. □

These results suggest that Dynamic constitutes an effective technique for handling the penalty parameter in interior-penalty methods for MPCCs.

We conclude this section by commenting on some of the failures of our algorithms. All implementations converge to a C-stationary point for problem `scale4` (which is a rescaling of problem `scholtes3`). We find it interesting that convergence to C-stationary points is possible in practice and is not simply allowed by the theory. We note that convergence to C-stationary points cannot be ruled out for SQP methods, and in this sense interior-point methods are no less robust than SQP methods applied to MPCCs. Another failure, discussed already, is problem `ralph2` for the algorithm Classic.

The rest of the failures can be attributed to various forms of problem deficiencies beyond the MPCC structure. All implementations have difficulties solving problems for which the minimizer is not a strongly stationary point, that is, problems for which there are no multipliers at the solution. This is the case in `ex9.2.2`, where our algorithms obtain good approximations of the solution but the penalty parameter diverges, and for `ralphmod`, where our algorithms fail to find a stationary point. These difficulties are not surprising because the algorithms strongly rely upon the existence of multipliers at the solution. SQP methods also fail to find strongly stationary solutions to these problems, and generate a sequence of multipliers that diverge to infinity.

Test problems in the groups `incid-set`$*$, `pack-rig`$*$ and `pack-comp`$*$ include degenerate constraints other than those defining complementarity. Our implementations are able to solve most of these problems, but the number of iterations is high, and the performance is very sensitive to changes in the implementation. In some of these problems our algorithms have difficulty making progress near the solution. Problem `tap-09` has a rank-deficient constraint Jacobian that causes difficulties for our algorithms. All of these point to the need for more general regularization schemes for interior methods that can cope with both MPCCs and with other forms of degeneracy. This topic is the subject of current investigation [13, 14].

6 Conclusions

Interior methods can be an efficient and robust tool for solving MPCCs, when appropriately combined with a regularization scheme. In this article, we have studied an interior-penalty approach and have carefully addressed issues related to efficiency and robustness. We have provided global and local convergence analysis to support the interior-penalty methods proposed here. We have also shown how to extend our global convergence results to interior methods based on the relaxation approach described by [17, 20].

We have presented two practical implementations. The first algorithm, Classic, is more flexible than the approach studied in [2, 15], which solves the penalty problem (2.1) with a fixed penalty parameter and then updates π if necessary. The approach in [2, 15] has the advantage that it can be used in combination with any off-the-shelf nonlinear programming solver; the disadvantage is that it can be very wasteful in terms of iterations if the initial penalty parameter is not appropriate. The second algorithm, Dynamic, improves on Classic by providing a more adaptive penalty update strategy. This can be particularly important in dealing with unbounded penalty problems and also yields an improvement in efficiency when the scaling of the problem complicates the detection of complementarity violation. The numerical results presented in this paper are highly encouraging. We plan to implement the penalty method for MPCCs in the KNITRO package, which will allow us to solve large-scale MPCCs.

The penalty methods considered here are designed specifically for MPCCs. However, lack of regularity other than that caused by complementarity constraints occurs in practice, and a more general class of interior-penalty methods for degenerate NLPs is the subject of current research [13, 14]. Some of the techniques proposed here may be useful in that more general context.

Acknowledgments

This work was in part supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract W-31-109-ENG-38. Support was also provided by National Science Foundation grant CCR-0219438 and Department of Energy grant DE-FG02-87ER25047-A004.

References

[1] M. Anitescu. On solving mathematical programs with complementarity constraints as nonlinear programs. Preprint ANL/MCS-P864-1200, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, 2000.

[2] Anitescu, M. Global convergence of an elastic mode approach for a class of mathematical programs with complementarity constraints. Preprint ANL/MCS-P1143-0404, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, 2004.

[3] H. Benson, A. Sen, D. F. Shanno, and R. V. D. Vanderbei. Interior-point algorithms, penalty methods and equilibrium problems. Technical Report ORFE-03-02, Princeton University, Operations Research and Financial Engineering, October 2003.

[4] R. H. Byrd, M. E. Hribar, and J. Nocedal. An interior point algorithm for large scale nonlinear programming. *SIAM J. Optimization*, 9(4):877–900, 1999.

[5] R. H. Byrd, G. Liu, and J. Nocedal. On the local behavior of an interior point method for nonlinear programming. In D. F. Griffiths, D. J. Higham, and G. A. Watson, editors, *Numerical Analysis 1997*, pages 37–56, Harlow, UK, 1997. Addison Wesley Longman.

[6] A. R. Conn, N. I. M. Gould, and Ph. Toint. *Trust-region methods.* MPS-SIAM Series on Optimization. SIAM publications, Philadelphia, PA, 2000.

[7] De Miguel, A.V., M.P. Friedlander, F.J. Nogales, and S. Scholtes. An interior-point method for MPECs based on strictly feasible relaxations. Technical report, London Business School, London NW1 4SA, UK, 2004.

[8] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming, Series A*, 91:201–213, 2002.

[9] R. Fletcher and S. Leyffer. Nonlinear programming without a penalty function. *Mathematical Programming*, 91:239–270, 2002.

[10] R. Fletcher and S. Leyffer. Solving mathematical program with complementarity constraints as nonlinear programs. *Optimization Methods and Software*, 19(1):15–40, 2004.

[11] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes. Local convergence of SQP methods for mathematical programs with equilibrium constraints. Numerical Analysis report NA/209, Department of Mathematics, University of Dundee, Dundee, UK, May 2002.

[12] A. Forsgren, P. E. Gill, and M. H. Wright. Interior methods for nonlinear optimization. *SIAM Review*, 4(4):525–597, 2002.

[13] D. Goldfarb and L. Chen. Interior-point ℓ_2 penalty methods for nonlinear programming with strong global convergence properties. Technical report, IEOR Dept, Columbia University, New York, 2004.

[14] N. I. M. Gould, D. Orban, and Ph. Toint. An interior-point l1-penalty method for nonlinear optimization. Technical report RAL-TR-2003-022, Rutherford Appleton Laboratory Chilton, Oxfordshire, UK, 2003.

[15] X. Hu and D. Ralph. Convergence of a penalty method for mathematical programming with complementarity constraints. Technical report, Judge Institute of Management Science, England, UK, 2002. To appear in *Journal of Optimization Theory and Applications.*

[16] S. Leyffer. MacMPEC: AMPL collection of MPECs. Web page, `www.mcs.anl.gov/~leyffer/MacMPEC/`, 2000.

[17] X. Liu and J. Sun. Generalized stationary points and an interior point method for mathematical programs with equilibrium constraints. Preprint, Singapore, School of Business, 2002.

[18] G. López-Calva. *Algorithms for Nonlinear Optimization: Problems with Complementarity Constraints and Derivative-Free Methods.* Ph.D. thesis, Industrial Engineering & Management Sciences, Northwestern University, Evanston, IL, in progress, 2004.

[19] P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: An SQP algorithm for large–scale constrained optimization. Report NA 97-2, Dept. of Mathematics, University of California, SanDiego, November 1997.

[20] A. Raghunathan and L. T. Biegler. Barrier methods for mathematical programs with complementarity constraints (MPCCs). Technical report, Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA, December 2002.

[21] D. Ralph, and S. J. Wright. Some properties of regularization and penalization schemes for MPECs. Technical Report 03-04, Computer Science Department, University of Wisconsin, December 2003. Revised April 2004, to appear in *Computational Optimization and Applications*.

[22] H. Scheel and S. Scholtes. Mathematical program with complementarity constraints: Stationarity, optimality and sensitivity. *Mathematics of Operations Research*, 25:1–22, 2000.

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

[24] R. J. Vanderbei and D. F. Shanno. An interior point algorithm for nonconvex nonlinear programming. *COAP*, 13:231–252, 1999.

[25] R. A. Waltz, J. L. Morales, J. Nocedal, and D. Orban. KNITRO-Direct: A hybrid interior algorithm for nonlinear optimization. Technical report 2003-10, Optimization Technology Center, Northwestern University, Evanston, IL, August 2003.