# The Penalty Interior-Point Method Fails to Converge*

SVEN LEYFFER†

September 22, 2003; revised April 15, 2004

**Abstract**

Equilibrium equations in the form of complementarity conditions often appear as constraints in optimization problems. Problems of this type are commonly referred to as mathematical programs with complementarity constraints (MPCCs). A popular method for solving MPCCs is the penalty interior-point algorithm (PIPA). This paper presents a small example for which PIPA converges to a nonstationary point, providing a counterexample to the established theory. The reasons for this adverse behavior are discussed.

**Keywords:** Nonlinear programming, interior-point methods, PIPA, MPEC, MPCC, equilibrium constraints.

**AMS-MSC2000:** 90C30, 90C33, 90C51, 49M37, 65K10.

## 1 Introduction

Equilibrium equations in the form of complementarity conditions often appear as constraints in optimization problems. Problems of this type, commonly referred to as mathematical programs with complementarity constraints (MPCCs), arise in many engineering and economic applications; see the survey [7] and the monographs [16, 18] for further references. The growing collections of test problems [14, 5] indicate that this is an important area. MPCCs can be expressed in general as

$$\begin{array}{ll} \text{minimize} & f(x,y,w,z) \\ \text{subject to} & x \in X \\ & F(x,y,w,z) = 0 \\ & 0 \le y \ \perp \ w \ge 0, \end{array} \tag{1.1}$$

where $X \subset \mathbb{R}^n$ is a polyhedral set, $f$ and $F$ are twice continuously differentiable functions, and $y, w \in \mathbb{R}^m$. The complementarity constraint $y \perp w$ means that either a component of $y$ is zero or the corresponding component of $w$ is zero, which implies that $y^T w = 0$.

Many approaches exist for solving MPCCs. These include branch-and-bound methods [2], implicit nonsmooth approaches [18], piecewise SQP methods [16], perturbation and

*Preprint ANL/MCS-P1091-0903

†Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439 (`leyffer@mcs.anl.gov`).

penalization approaches [6] analyzed in [22], and the penalty interior-point algorithm (PIPA) [16], whose convergence properties are the subject of this note.

Recently, there has also been renewed interest in solving (1.1) with standard nonlinear programming (NLP) techniques by replacing the complementarity constraint by

$$y^T w \leq 0,\ y \geq 0,\ w \geq 0.$$

This approach had been regarded as numerically unsafe because (1.1) violates a constraint qualification at any feasible point [21]. However, numerical [8] and theoretical [1, 9] evidence suggests that sequential quadratic programming (SQP) solvers are able to solve large classes of MPCCs. A similar analysis is being extended to interior-point methods (IPMs) [3, 15, 20].

This note is organized as follows. The next section briefly describes PIPA and presents its convergence result. Section 3 presents a small counterexample that shows that PIPA can fail to converge to a stationary point for certain admissible choices of parameters. Section 4 discusses the example and suggests a remedy that avoids this problem.

## 2 The Penalty Interior-Point Algorithm (PIPA)

PIPA [16, Chapter 6.1] solves MPCCs of the form (1.1). It has been studied by Fukushima and Pang [11] and by Pieper [19] and has been applied to a number of real applications in finance [13], target classification [17], and electricity markets [12].

PIPA generates a sequence of iterates $(x^k, y^k, w^k, z^k)$ that are strictly interior in $y^k > 0$, $w^k > 0$. The iterates are computed by solving a quadratic direction-finding problem. In that sense, PIPA combines aspects of interior-point and SQP methods.

We use $d$ to denote the step or displacement computed by PIPA. Subscripts such as $d_x$ refer the part of $d$ corresponding to the $x$ variables; superscripts are used to denote iterates or evaluation of functions at a particular point, for example, $\nabla f^k = \nabla f(x^k, y^k, w^k, z^k)$. Diagonal matrices are denoted by $W = \mathrm{diag}(w)$ and $Y = \mathrm{diag}(y)$ and $e = (1, \ldots, 1)^T$.

The algorithm solves the following direction-finding problem at every iteration:

$$\begin{array}{ll} \text{minimize} & \nabla {f^k}^T d + \frac{1}{2} d_x^T Q^k d_x \\ \text{subject to} & x^k + d_x \in X \\ & F^k + \nabla {F^k}^T d = 0 \\ & Y^k d_w + W^k d_y = -Y^k w^k + \sigma \frac{{y^k}^T w^k}{m} e \\ & \|d_x\|_2^2 \leq c \left( \|F^k\| + {y^k}^T w^k \right), \end{array} \tag{2.2}$$

where $c > 0$ and $\sigma \in (0, 1)$ are parameters and $Q^k$ approximates second-order information in $f$ and $F$. Note that (2.2) has a quadratic trust-region type of constraint. If this constraint is replaced by the $\ell_\infty$ norm, then (2.2) becomes a quadratic program (QP). Note also that the linearization of the complementarity constraint $Yw = 0$ is relaxed by $\sigma \frac{{y^k}^T w^k}{m} e$, giving PIPA an interior-point flavor. Note that PIPA is not strictly an interior-point algorithm, as it remains interior only with respect to the complementary variables $y, w$ but not with respect to the other constraints. We will show that the final trust-region type constraint is the source of the problems affecting PIPA.

After determining a search direction, PIPA performs a backtracking line-search to ensure that $y^k, w^k$ remain positive and close to the central path. This is achieved by finding the zero $\overline{\tau}$ of

$$g_k(\tau) = (1-\rho)\sigma \frac{{y^k}^T w^k}{m} + \tau \left( \min_{1 \le i \le m} d^k_{y_i} d^k_{w_i} - \rho \frac{{d^k_y}^T d^k_w}{m} \right), \tag{2.3}$$

if it exists, or setting $\overline{\tau} = 1$ otherwise. It can be shown (see [16, Lemma 6.1.8]) that any step along $d$ with step length $\tau \in [0, \overline{\tau}]$ maintains positivity and centrality of the iterates. Global convergence is enforced by further reducing the step length $\overline{\tau}$ to yield sufficient reduction in a quadratic penalty function

$$P_\alpha(x, y, w, z) = f(x, y, w, z) + \alpha \left( \|F(x, y, w, z)\|^2 + y^T w \right), \tag{2.4}$$

where the penalty parameter $\alpha > 0$ is chosen as follows: Let $p \ge 1$ be the smallest integer such that the step ensures sufficient model decrease, that is,

$$\nabla {f^k}^T d - \alpha^p_{k-1} \left( 2\|F(x^k, y^k, w^k, z^k)\|^2 + (1-\sigma){y^k}^T w^k \right) < -\|F(x^k, y^k, w^k, z^k)\|^2 - {y^k}^T w^k, \tag{2.5}$$

and the penalty parameter is set to $\alpha = \alpha_k = \alpha^p_{k-1}$. Note that the penalty function $P_\alpha$ mixes a quadratic penalty for the nonlinear equation $F$ with an $\ell_1$ penalty for complementarity (since $y^k, w^k > 0$, it follows that $\|{y^k}^T w^k\|_1 = {y^k}^T w^k$). Sufficient reduction in the penalty function

$$\begin{aligned} &P_\alpha(x^k + \tau d_x, y^k + \tau d_y, w^k + \tau d_w, z^k + \tau d_z) - P_\alpha(x^k, y^k, w^k, z^k) \\ &\le \gamma\tau \left( \nabla {f^k}^T d - \alpha^p_{k-1} \left( 2\|F(x^k, y^k, w^k, z^k)\|^2 + (1-\sigma){y^k}^T w^k \right) \right) \end{aligned}$$

is achieved by performing a backtracking line-search, where $\gamma \in (0,1)$ and $\tau$ is the final step length. Further details of the algorithm can be found in [16]. PIPA can be summarized as follows.

**Penalty Interior-Point Algorithm (PIPA)**

1. *Initialization:*
   Choose parameters $c > 0$, and $0 < \sigma, \gamma, \rho < 1$, choose a starting point $(x^0, y^0, w^0, z^0)$ such that $y^0, w^0 > 0$ suitably centered, choose a penalty parameter $\alpha > 0$, and set $k = 0$.

**repeat**

2. *Direction-finding problem:* Solve problem (2.2) for a trial step $d = d^k$.
3. *Step size determination:* Find a step size $\tau = \tau_k$ to:
   3.1 Ensure centrality and positivity of $(y, w)$ by finding the root of $g_k(\tau)$ in (2.3) or setting $\tau = 1 - \epsilon$ if this root does not exist or is greater than 1.
   3.2 Ensure sufficient reduction in the quadratic penalty function $P_\alpha(x, y, w, z)$ in (2.4) by performing an Armijo-search on $P_\alpha$.
   Let $\tau_k$ be the step size determined in 3.1 and 3.2.
4. *Update:* $(x^{k+1}, y^{k+1}, w^{k+1}, z^{k+1}) = (x^k, y^k, w^k, z^k) + \tau_k(d^k_x, d^k_y, d^k_w, d^k_z)$, $k = k+1$.

**until** $\|d\| \le \epsilon$

This presentation of PIPA is less sophisticated than that in [16], which allows for instance $\sigma$ to vary with the iteration. Convergence of PIPA is established in [16] under the following two assumptions.

**[SC]** Strict complementarity of the solution, namely, $y^* + w^* > 0$

**[NS]** Nonsingularity of the matrix

$$\begin{bmatrix} \nabla_y F^* & \nabla_w F^* & \nabla_z F^* \\ W^* & Y^* & 0 \end{bmatrix}$$

A sufficient condition of **[NS]** is that the Jacobian of $F$ satisfies the mixed $P_0$ property (see [16, Definition 6.1.4 and Proposition 6.1.6]). The following theorem summarizes the convergence results in [16].

**Theorem 2.1 [16, Theorem 6.1.17]** Suppose that the Hessian matrices $W^k$ are bounded and that $\sigma > 0$. If the penalty parameter $\alpha$ is bounded, then every limit point of $(x^k, y^k, w^k, z^k)$ that satisfies **[SC]** and **[NS]** is a stationary point of (1.1).

In the next section, we present a small example for which PIPA converges to a non-stationary point, contradicting Theorem 2.1.

## 3 A Counterexample

In this section, we show that PIPA may fail to converge to a stationary point. This result contradicts the convergence results of [16]. This section examines the behavior of PIPA applied to the following small example:

$$\begin{array}{ll} \text{minimize} & x + w \\ \text{subject to} & -1 \le x \le 1 \\ & -1 + x + y = 0 \\ & 0 \le y \;\perp\; w \ge 0. \end{array} \tag{3.1}$$

The solution of (3.1) is $(x, y, w)^* = (-1, 2, 0)$, which satisfies the assumptions **[SC]** and **[NS]**, since $y^* + w^* = 2 > 0$ and

$$\begin{bmatrix} \nabla_y F^* & \nabla_w F^* \\ W^* & Y^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

is nonsingular. Next, we show that this example generates a limit point that is not stationary. Specifically, we perform a standard iteration of the algorithm and bound the limit away from stationarity. The following lemma summarizes our result.

**Lemma 3.1** PIPA applied to the Example (3.1) and started at $(x^0, y^0, w^0) = (0, 1, 0.02)$, with $c = 1, \sigma = 0.1$, $\gamma = 0.01, \rho = 0.9$, and $\alpha = 2$, generates a sequence of iterates satisfying

$$1 \leq y^k \leq y^{k+1} \tag{3.2}$$
$$w^{k+1} \leq \frac{1}{2} w^k \leq \frac{2}{100} \tag{3.3}$$
$$x^{k+1} \geq x^k - \sqrt{y^k w^k} > -1 \tag{3.4}$$
$$y^{k+1} w^{k+1} \leq \frac{1}{2} y^k w^k \tag{3.5}$$
$$\frac{5}{9} \leq \tau_k \leq 1. \tag{3.6}$$

**Proof.** The starting point satisfies the linear equation, which implies that the equality constraints are satisfied for all subsequent iterations. The theorem follows by induction. Clearly, the assertions hold for the starting point. Now assume that (3.2) to (3.6) hold for $k - 1$, and show that they also hold for $k$.

The direction-finding problem for this example can be simplified as follows:

$$\begin{array}{ll} \text{minimize} & d_x + d_w \\ \text{subject to} & -1 \leq x^k + d_x \leq 1 \\ & d_x + d_y = 0 \\ & w^k d_y + y^k d_w = (\sigma - 1) y^k w^k \\ & -\sqrt{y^k w^k} \leq d_x \leq \sqrt{y^k w^k}, \end{array} \tag{3.7}$$

where $Q^k = 0$ for simplicity. However, the example remains valid for certain positive semi-definite bounded $Q^k$ (as will be shown below).

First we show that the optimal basis $B$ to the LP (3.7) is given by

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & w^k & 0 \\ 0 & y^k & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 & -\frac{w^k}{y^k} \\ 0 & 0 & \frac{1}{y^k} \\ 1 & -1 & \frac{w^k}{y^k} \end{bmatrix}.$$

From this, multipliers can be computed as

$$\lambda^k = B^{-1} \nabla f = \begin{bmatrix} 0 & 1 & -\frac{w^k}{y^k} \\ 0 & 0 & \frac{1}{y^k} \\ 1 & -1 & \frac{w^k}{y^k} \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{w^k}{y^k} \\ \frac{1}{y^k} \\ 1 + \frac{w^k}{y^k} \end{pmatrix}.$$

Optimality follows, since $\lambda_3^k \geq 0$ for the only inequality constraint. The step $d$ can also be computed as

$$d^k = B^{-T} b^k = \begin{pmatrix} -\sqrt{y^k w^k} \\ \sqrt{y^k w^k} \\ -\frac{9}{10} w^k - \frac{w^k}{y^k} \sqrt{y^k w^k} \end{pmatrix},$$

where $b^k = (0, -\frac{9}{10}y^k w^k, -\sqrt{y^k w^k})^T$ is the right-hand side of the active constraints.

Note that for any positive semi-definite $Q^k$, this solution remains unchanged as long as $\lambda_3^k \geq 0$. The contribution of the Hessian to the multipliers can be estimated by adding the Hessian term to the gradient and re-solving for the multipliers,

$$\tilde{\lambda_3^k} = \lambda_3^k + B^{-1}Q^k d^k \geq 1 + \frac{w^k}{y^k} - \|B^{-1}Q^k\|\|d^k\|,$$

which remains positive for $\|Q^k\|$ or $\|d^k\|$ sufficiently small. Thus, the basis remains optimal and the same conclusions apply even if a positive definite $Q^k$ is used.

Next, we show that (3.4) follows by applying the step $d^k$, using $\tau^k \leq 1$ and induction:

$$\begin{aligned} x^{k+1} &= x^k + \tau^k d_x^k \geq x^k - \sqrt{y^k w^k} \geq x^0 - \sum_{l=0}^{k} \sqrt{y^l w^l} \geq x^0 - \sqrt{y^0 w^0} \sum_{l=0}^{k} \left(\frac{1}{\sqrt{2}}\right)^l \\ &\geq x^0 - \sqrt{y^0 w^0} \sum_{l=0}^{\infty} \left(\frac{1}{\sqrt{2}}\right)^l \geq 0 - \frac{\sqrt{2}}{10} \frac{\sqrt{2}}{1-\sqrt{2}} \simeq -0.4828 > -1. \end{aligned} \tag{3.8}$$

This implies that the "choice" of active set in the LP (3.7) was correct, and the lower bound on $x$ is never active during the iteration. Next we show that (3.2) follows by applying the step in the $y$ component,

$$y^{k+1} = y^k + \tau^k \sqrt{y^k w^k} \geq y^k \geq 1.$$

Next consider (3.6), and observe that $\tau^k \leq 1$ follows trivially. To obtain the lower bound, observe that the zero of $g_k(\tau)$ in (2.3) is given by the following expression, where the superscripts $k$ have been omitted for the sake of simplicity

$$\begin{aligned} \tau &= -\frac{1}{10}\frac{yw}{d_w d_y} = \frac{y^2}{\left(9y + 10\sqrt{yw}\right)\sqrt{yw}} \\ &= \frac{1}{9}\frac{\sqrt{y}}{\sqrt{w}} \frac{1}{1 + \frac{10}{9}\frac{\sqrt{w}}{\sqrt{y}}} = \frac{1}{9}\frac{\sqrt{y}}{\sqrt{w}} \sum_{l=0}^{\infty} \left(-\frac{10}{9}\frac{\sqrt{w}}{\sqrt{y}}\right)^l. \end{aligned}$$

Setting $q = (10/9)(\sqrt{w}/\sqrt{y})$, we can write this equation as

$$\begin{aligned} \tau &= \frac{1}{9}\frac{\sqrt{y}}{\sqrt{w}} \sum_{l=0}^{\infty} (-q)^l = \frac{1}{9}\frac{\sqrt{y}}{\sqrt{w}}(1-q)\left(1 + \sum_{l=0}^{\infty} q^{2l}\right) \\ &\geq \frac{1}{9}\frac{\sqrt{y}}{\sqrt{w}}(1-q) \geq \frac{10}{9\sqrt{2}}\left(1 - \frac{\sqrt{2}}{9}\right) > \frac{5}{9}. \end{aligned}$$

It remains to show that this lower bound on $\tau$ also passes the second line-search criterion. The penalty update rule (2.5) simplifies because $F(x^k, y^k, w^k, z^k) = 0$ for all $k$. Thus, we need to find the smallest integer $p \geq 1$ such that

$$\nabla f^{k^T} d - \alpha^p (1-\sigma) y^{k^T} w^k \;<\; -\alpha^p (1-\sigma) y^{k^T} w^k \;<\; -y^{k^T} w^k$$

holds. Since $\nabla f^{k^T} d < 0$ and $\alpha(1-\sigma) > 1$, this is always satisfied for $p = 1$, and the penalty parameter is never increased. Finally, this step size also satisfies the conditions for the Armijo search in Step 3.2. The actual reduction clearly satisfies

$$P_\alpha^{k+1} - P_\alpha^k = \tau d_x + \tau d_w + \alpha\left(y^{k+1}w^{k+1} - y^k w^k\right).$$

Since $y^{k+1}w^{k+1} \le \frac{1}{2}y^k w^k$, it follows that

$$P_\alpha^{k+1} - P_\alpha^k \le \tau d_x + \tau d_w - \alpha \frac{1}{2} y^k w^k,$$

which implies that the sufficient reduction condition,

$$P_\alpha^{k+1} - P_\alpha^k \le \gamma\tau d_x + \gamma\tau d_w - \alpha(1-\sigma)\gamma y^k w^k,$$

holds for any $\gamma \le 1$, since $(1-\sigma)\gamma \le \frac{1}{2}$.

Next consider (3.3). Since $\tau \ge \frac{5}{9}$, it follows that

$$w^{k+1} \;=\; w^k + \tau\left(-\frac{9}{10}w^k - \frac{w^k}{y^k}\sqrt{y^k w^k}\right) \;\le\; \frac{1}{2}w^k - \frac{5}{9}\frac{w^k}{y^k}\sqrt{y^k w^k}. \tag{3.9}$$

Note that $1/2 - (5/9)(\sqrt{w^k}/\sqrt{y^k}) \ge 1/2 - (5/9)(\sqrt{2}/10) > 0$. Thus, it is possible to bound the new complementarity violation after the step by

$$\begin{array}{rcl} y^{k+1}w^{k+1} & \le & \left(y^k + \sqrt{y^k w^k}\right)\left(\frac{1}{2}w^k - \frac{5}{9}\frac{w^k}{y^k}\sqrt{y^k w^k}\right) \\ & = & \frac{1}{2}y^k w^k - \frac{5}{9}w^k\sqrt{y^k w^k} + \frac{1}{2}w^k\sqrt{y^k w^k} - \frac{5}{9}(w^k)^2 \\ & \le & \frac{1}{2}y^k w^k \end{array}$$

obtaining (3.5). This concludes the proof. □

We have already shown in (3.8) that $x^{k+1} \ge -0.4828 > -1$. Similarly, an upper bound can be derived for the state iterates as

$$y^{k+1} = y^k + \tau^k d_y^k \le y^k + \sqrt{y^k w^k} \le y^0 + \sum_{l=0}^{\infty}\sqrt{y^l w^l} \le 1 + \frac{\sqrt{2}}{10}\frac{\sqrt{2}}{\sqrt{2}-1} \approx 1.4828; \tag{3.10}$$

see (3.8). These bounds can also be verified numerically. Example (3.1) and PIPA have been implemented in AMPL [10] by using the "looping extension" which allows the convenient implementation of algorithms in AMPL. The sequence generated is given in Table 1 and confirms the bounds given above. The columns headed "ared" and "pred" in Table 1 refer to the actual and predicted reduction. Their implications are discussed below.

Thus, it follows that all iterates remain in a compact set, namely,

$$\begin{array}{rcccccl} 0 & = & x^0 & \ge & x^k & \ge & -0.4828 \\ 1 & = & y^0 & \le & y^k & \le & 1.4828 \\ \frac{2}{100} & = & w^0 & \ge & w^k & \ge & 0, \end{array}$$

and the sequence $(x,y,w)^k$ has a limit point $(x,y,w)^\infty$. From (3.9) it follows that

$$w^{k+1} \;\le\; w^k\left(\frac{1}{2} - \frac{5}{9y^k}\sqrt{y^k w^k}\right) \;\le\; \frac{1}{2}w^k,$$

Table 1: Iterates, actual and predicted reduction of AMPL implementation of PIPA

| $k$ | $x^k$ | $y^k$ | $w^k$ | ared | pred |
|---|---|---|---|---|---|
| 1 | 0 | 1 | 0.02 | | |
| 2 | -0.096022613 | 1.0960226 | 0.0058578644 | -0.198 | -0.137 |
| 3 | -0.17606958 | 1.1760696 | 0.00016323495 | -0.0974 | -0.0982 |
| 4 | -0.18991126 | 1.1899113 | 1.4549224E-05 | -0.0143 | -0.0143 |
| 5 | -0.1940679 | 1.1940679 | 1.4171928E-06 | -0.00421 | -0.0042 |
| 6 | -0.19536745 | 1.1953675 | 1.4145236E-07 | -0.00131 | -0.0013 |
| 7 | -0.19577825 | 1.1957782 | 1.4223933E-08 | -0.000412 | -0.000411 |
| 8 | -0.19590853 | 1.1959085 | 1.433645E-09 | -0.00013 | -0.00013 |
| 9 | -0.1959499 | 1.1959499 | 1.4460519E-10 | -4.14e-05 | -4.14e-05 |
| 10 | -0.19596304 | 1.195963 | 1.4586827E-11 | -1.32e-05 | -1.31e-05 |

so that $w^\infty = 0$ and ${y^\infty}^T w^\infty = 0$. Therefore, PIPA converges to a limit point $(x, y, w)^\infty$ such that

$$\begin{array}{rcccl} 0 & \geq & x^\infty & \geq & -0.4828 \\ 1 & \leq & y^\infty & \leq & 1.4828 \\ & & w^\infty & = & 0. \end{array}$$

This limit point satisfies complementarity and the linear constraints for problem (3.1) (since $(x, y, w)^0$ satisfies the linear constraints). Moreover, this limit point also satisfies the assumptions of Theorem 2.1. Assumption [**SC**] is satisfied because $y^\infty + w^\infty \geq 1 > 0$, and Assumption [**NS**] holds because

$$\left[\begin{array}{cc} M & I \\ W^\infty & Y^\infty \end{array}\right] = \left[\begin{array}{cc} 1 & 0 \\ 0 & y^\infty \end{array}\right]$$

is nonsingular for all $1 \leq y^\infty \leq 1.4828$. However, the limit point is clearly not a stationary point, thereby contradicting Theorem 2.1.

## 4 Discussion and Remedy

In the preceding section, a small example has been presented for which PIPA converges to a feasible but nonstationary limit point. This example contradicts [16, Theorem 6.1.17], which establishes the convergence of PIPA. In particular, since $\tau \geq 5/9 > 0$, we observe that part (i) of the proof of [16, Theorem 6.1.17] is wrong.

The reason for the failure of PIPA is the trust-region constraint

$$\|d_x\|_2^2 \;\leq\; c\left(\|F^k\| + {y^k}^T w^k\right) \;=:\; \Delta_k \tag{4.11}$$

in the direction-finding problem (2.2). The trust-region radius $\Delta_k$ converges to zero as the iterates approach feasibility, thereby limiting the progress toward optimality in the controls $x$. This adverse behavior can be expected to occur whenever the iterates approach feasibility "faster" than optimality.

Linking the trust-region radius to the feasibility of the problem is counterintuitive. Normally, $\Delta_k$ is controlled by the algorithm, rather than by the iterates. In particular, it should reflect how well the model problem (2.2) approximates the original problem, measured, for instance, by the agreement between actual reduction,

$$\text{ared} \;=\; P_\alpha^k - P_\alpha^{k+1},$$

and predicted reduction,

$$\text{pred} \;=\; \nabla f^{k^T} d - \alpha^p (1-\sigma) y^{k^T} w^k.$$

Table 1 shows that, in this example, both the actual and the predicted reductions show near-perfect agreement toward the end. Thus, one would expect $\Delta_k$ to be increased rather than decreased. The adverse situation in which (4.11) limits progress toward optimality can be easily detected. The Lagrange multiplier of (4.11) indicates whether progress can be made by relaxing this constraint. In the present example, this multiplier converges to 1.

One possible way to remedy the adverse behavior of PIPA is to control the trust-region radius by using standard trust-region techniques. The direction-finding problem then becomes

$$\begin{array}{ll} \text{minimize} & \nabla f^{k^T} d + \frac{1}{2} d_x^T Q^k d_x \\ \text{subject to} & x^k + d_x \in X \\ & F^k + \nabla F^{k^T} d = 0 \\ & Y^k d_w + W^k d_y = -Y^k w^k + \sigma \frac{y^{k^T} w^k}{m} e \\ & \|d_x\| \le \Delta_k, \end{array} \tag{4.12}$$

where the trust-region is now controlled by the algorithm. A typical updating scheme for $\Delta_k$ is as follows. First compute the ratio of actual over predicted reduction

$$\rho_k := \frac{\text{ared}}{\text{pred}},$$

and then update $\Delta_k$ according to

$$\Delta_{k+1} \in \begin{cases} [\gamma_0 \Delta_k, \gamma_1 \Delta_k] & \text{if } \rho_k < \eta_1 \\ [\gamma_1 \Delta_k, \Delta_k] & \text{if } \rho_k \in [\eta_1, \eta_2) \\ [\Delta_k, \gamma_2 \Delta_k] & \text{if } \rho_k \ge \eta_2 \end{cases}$$

in Step 4, where $0 < \gamma_0 \le \gamma_1 < 1 \le \gamma_2$, and $0 < \eta_1 < \eta_2 < 1$. Typical values for these constants are $\gamma_0 = \gamma_1 = 0.5$, $\gamma_2 = 2$, $\eta_1 = 0.25$, and $\eta_2 = 0.75$.

This results in a mixed trust-region/line-search algorithm. Convergence can be established under the assumption that the Jacobian of $F$,

$$\begin{bmatrix} \nabla_y F & \nabla_w F & \nabla_z F \\ W & Y & 0 \end{bmatrix},$$

remains nonsingular and has a bounded inverse along the lines of the coupled trust-region approach of Dennis et al. [4], since the QP subproblem can be regarded as a problem in $d_x$ only after eliminating $(d_y, d_w, d_z)$ using the linearization of $F$, which is nonsingular.

## Acknowledgments

I am grateful to Roger Fletcher and Jorge Nocedal for many fruitful discussions on MPCCs and interior-point methods. I am also grateful to Jong-Shi Pang and two anonymous referee for useful comments on an earlier version of the paper. This work was 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.


## References


[1] M. Anitescu (2000). 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.

[2] J. F. Bard (1988). Convex two-level optimization. Mathematical Programming, 40(1):15–27.

[3] H. Benson, D. F. Shanno, and R. V. D. Vanderbei (2002). Interior-point methods for nonconvex nonlinear programming: Complementarity constraints. Technical Report ORFE 02-02, Princeton University, Operations Research and Financial Engineering.

[4] J. E. Dennis, M. Heinkenschloss, and L. N. Vicente (1998). Trust-region interior-point SQP algorithms for a class of nonlinear programming problems. SIAM J. Control & Optimization, 36(5):1750–1794.

[5] S. P. Dirkse (2001). MPEC world. Webpage, GAMS Development Corp., [www.gamsworld.org/mpec/].

[6] S. P. Dirkse, M. C. Ferris, and A. Meeraus (2002). Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constraint optimization. Numerical Analysis Group Research Report NA-02/11, Oxford University Computing Laboratory.

[7] M. C. Ferris and J. S. Pang (1997). Engineering and economic applications of complementarity problems. SIAM Review, 39(4):669–713.

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

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

[10] R. Fourer, D. M. Gay, and B. W. Kernighan (1993). AMPL: A Modeling Language for Mathematical Programming. Boyd & Fraser Publishing Company, Boston.

[11] M. Fukushima and J.-S. Pang (1998). Some feasibility issues in mathematical programs with equilibrium constraints. SIAM Journal on Optimization, 8(3):673–681.

[12] B. F. Hobbs, C. B. Metzler, and J.-S. Pang (2000). Strategic gaming analysis for electric power systems: An MPEC approach. IEEE Transactions on Power Systems, 15(2):638–645.

[13] J. Huang and J.-S. Pang (1999). A mathematical programming with equilibrium approach to the implied volatility surface of American options. Technical report, Dept. of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD.

[14] S. Leyffer (2000). MacMPEC: AMPL collection of MPECs. [www.mcs.anl.gov/~leyffer/MacMPEC/].

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

[16] Z.-Q. Luo, J.-S. Pang, and D. Ralph (1996). Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, UK.

[17] T. Olson, J.-S. Pang, and C. Priebe (2003). A likelihood-MPEC approach to target classification. Mathematical Programming, 96(1):1–31.

[18] J. Outrata, M. Kocvara, and J. Zowe (1998). Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer Academic Publishers, Dordrecht.

[19] H. Pieper (2001). Algorithms for Mathematical Programs with Equilibrium Constraints with Applications to Deregulated Electricity Markets. Ph.D. thesis, Department of Management Science and Engineering, Stanford University, Stanford, CA.

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

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

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