SOME PROPERTIES OF REGULARIZATION AND PENALIZATION SCHEMES FOR MPECS*

DANIEL RALPH† AND STEPHEN J. WRIGHT‡

Abstract. Some properties of regularized and penalized nonlinear programming formulations of mathematical programs with equilibrium constraints (MPECs) are described. The focus is on the properties of these formulations near a local solution of the MPEC at which strong stationarity and a second-order sufficient condition are satisfied. In the regularized formulations, the complementarity condition is replaced by a constraint involving a positive parameter that can be decreased to zero. In the penalized formulation, the complementarity constraint appears as a penalty term in the objective. Existence and uniqueness of solutions for these formulations are investigated, and estimates are obtained for the distance of these solutions to the MPEC solution under various assumptions.

Key words. Nonlinear Programming, Equilibrium Constraints, Complementarity Constraints

1. Introduction. We consider mathematical programs with equilibrium constraints in the form of complementarity constraints:

$$
\begin{array}{lll}
\min_x & f(x) & \text{subject to} \\
& g(x) \geq 0, & h(x) = 0, \\
& 0 \leq G(x) \perp H(x) \geq 0, &
\end{array}
\tag{1.1}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $h : \mathbb{R}^n \to \mathbb{R}^q$, $G : \mathbb{R}^n \to \mathbb{R}^m$, and $H : \mathbb{R}^n \to \mathbb{R}^m$ are all twice continuously differentiable functions, and the notation $G(x) \perp H(x)$ signifies that $G(x)^T H(x) = 0$. These problems have been the subject of much recent investigation because of both their importance in applications and their theoretical interest, which arises from the fact that their most natural nonlinear programming formulations (for example, replacing $G(x) \perp H(x)$ by $G(x)^T H(x) = 0$) do not satisfy constraint qualifications [4, 29] at any feasible point.

In this paper, we study a regularization scheme analyzed by Scholtes [28] in which (1.1) is approximated by the following nonlinear program, which is parameterized by the nonnegative scalar t:

$$
\begin{array}{llll}
\text{Reg}(t): & \min_x & f(x) & \text{subject to} \\
& & g(x) \geq 0, \quad h(x) = 0, & \\
& & G(x) \geq 0, \quad H(x) \geq 0, \quad G_i(x)H_i(x) \leq t, & i = 1, 2, \ldots, m.
\end{array}
\tag{1.2}
$$

We denote the solution of this problem by $x(t)$. Since Reg(0) is equivalent to (1.1), the regularization scheme can be put to use by applying a nonlinear programming algorithm to Reg(t) for a sequence of problems where t is positive and decreasing to 0, deriving a starting point for each minimization from approximate minimizers for previous problems in the sequence.

Scholtes [28, Theorem 4.1], restated later as Theorem 3.1, shows that in the neighborhood of a solution x^* of (1.1) satisfying certain conditions, there is a unique stationary point $x(t)$ for Reg(t) for all positive t sufficiently small. Moreover, this

*The final version of this paper appeared in *Optimization Methods and Software*, 19 (2004), pp. 527–556.

†Judge Institute of Management, University of Cambridge, Trumpington St, Cambridge, CB2 1AG, U.K. danny.ralph@jims.cam.ac.uk

‡Computer Sciences Department, 1210 W. Dayton Street, University of Wisconsin, Madison, WI 53706, U.S.A. swright@cs.wisc.edu

local solution mapping is piecewise smooth in t, and thus satisfies $\|x(t)-x^*\| = O(t)$. One of our main results (Theorem 3.7 in Section 3.3) shows that the same conclusion holds in the absence of one of the less natural assumptions—a strict complementarity condition—made in [28, Theorem 4.1]. Both results rely on a strong-second order condition, termed RNLP-SSOSC and defined below.

In Section 3.1, we investigate existence of solutions to Reg(t) near x^*, under weaker second-order and strict complementarity conditions. Theorem 3.2 replaces RNLP-SSOSC with a weaker second-order sufficient condition (MPEC-SOSC, also defined below), and drops the strict complementarity assumptions. This result shows that Reg(t) has a (possibly nonunique) local solution within a distance $O(t^{1/2})$ of x^*. Under RNLP-SOSC, a condition that is intermediate between MPEC-SOSC and RNLP-SSOSC, Theorem 3.3 gives an improved $O(t)$ bound, still without requiring the strict complementarity assumptions. Corollary 3.4 shows that a partial strict complementarity condition, in conjunction with MPEC-SOSC, leads to the $O(t)$ estimate again.

In Section 3.2, we show that Lagrange multipliers for solutions of Reg(t) satisfying the $O(t)$ estimate are bounded. Section 3.3 contains Theorem 3.7 mentioned above, which gives sufficient conditions for $x(t)$ to be piecewise smooth and locally unique for small $t > 0$.

Section 4 studies properties of solutions of some alternative regularized formulations. Scholtes [28, Section 5.1] also considers the following regularization scheme, in which the approximate complementarity condition is gathered into a single constraint:

$$\text{(1.3)} \qquad \begin{array}{lll} \text{RegComp}(t): & \min_x & f(x) \quad \text{subject to} \\ & & g(x) \geq 0, \quad h(x) = 0, \\ & & G(x) \geq 0, \quad H(x) \geq 0, \quad G(x)^T H(x) \leq t. \end{array}$$

Section 4.1 points out that analogs of Theorems 3.2 and 3.3 hold for RegComp(t), but that local uniqueness results like those of Section 3.3 do not hold. In another plausible regularization, the inequalities of the regularization terms in Reg(t) are replaced by equalities:

$$\text{(1.4)} \qquad \begin{array}{lll} \text{RegEq}(t): & \min_x & f(x) \quad \text{subject to} \\ & & g(x) \geq 0, \;\; h(x) = 0, \\ & & G(x) > 0, \quad H(x) > 0, \quad G_i(x) H_i(x) = t, \;\; i = 1, 2, \ldots, m. \end{array}$$

Section 4.2 shows that an existence result similar to Theorem 3.2 holds for this formulation, but with the $O(t^{1/2})$ estimate replaced by $O(t^{1/4})$. (The proof technique is quite different; unlike the proofs in Section 3.1, it does not rely on the results of Bonnans and Shapiro [3].)

Finally, in Section 5, we discuss a nonlinear programming reformulation based on the exact ℓ_1 penalty function. For a given nonnegative parameter ρ, this reformulation is as follows:

$$\text{(1.5)} \qquad \begin{array}{lll} \text{PF}(\rho): & \min_x & f(x) + \rho G(x)^T H(x) \quad \text{subject to} \\ & & g(x) \geq 0, \quad h(x) = 0, \quad G(x) \geq 0, \quad H(x) \geq 0, \end{array}$$

We show that this formulation has the appealing property that under standard assumptions, the MPEC solution x^* is a local solution of PF(ρ), for all ρ sufficiently large, and that regularity conditions for the MPEC imply regularity of PF(ρ).

While this paper focuses on certain regularization and penalization schemes, there are several other nonlinear programming approached to (1.1) with similar motivations, starting with Fukushima and Pang's analysis [8] of the smoothing scheme of Facchinei et al. [6], and including the penalty approaches analyzed by Hu and Ralph [12] and Huang, Yang, and Zhu [13]. Lin and Fukushima [18] have studied the issue of identifying active constraints in smoothing, regularization, and penalty methods. More recently, Anitescu [1] has studied the "elastic mode" for nonlinear programming, in conjunction with a sequential quadratic programming (SQP) algorithm, and focuses particularly on MPECs. Anitescu's formulation is similar to (1.5), but it introduces an extra variable into the formulation to represent the maximum of $G(x)^T H(x)$ and the violation of the other constraints.

On a slightly different tack, decomposition methods which recognize the disjunctive nature of MPEC constraints are well studied. We mention the globally convergent methods for MPECs with linear constraint functions proposed or analyzed by Jiang and Ralph [15] (see [20, Chapter 6] and [16] for local convergence analysis); Tseng and Fukushima [9], who use an ϵ-active set method; and Zhang and Liu [30], who use an extreme-ray descent method. SQP-based methods for MPECs can be found in Liu et al. [19] and Fletcher et al. [7]. Interior-point methods have been proposed by de Miguel, Friedlander, Nogales and Scholtes [5] and Raghunathan and Biegler [23], while Benson, Shanno, and Vanderbei [2] have performed a computational study involving the LOQO interior-point code and the MacMPEC test set (Leyffer [17]).

An anonymous referee has alerted us to a forthcoming paper by Izmailov [14]. We do not have access to an English translation of this paper, but believe that it includes analysis similar to some of that which appears in our proofs below (in particular, the proof of Theorem 3.2). See the acknowledgments at the end of this paper for further details.

In the remainder of the paper we use $\|\cdot\|$ to denote the Euclidean norm $\|\cdot\|_2$, unless otherwise specified. We write $b = O(a)$ for nonnegative scalars a and b if there is a constant C such that $b \le Ca$ for all a sufficiently small, or all a sufficiently large, depending on the context. We write $b = o(a)$ if for some sequence of nonnegative values a_k and corresponding b_k with either $a_k \to \infty$ or $a_k \to 0$, we have that $b_k/a_k \to 0$.

2. Assumptions and Background.

We now summarize some known results concerning constraint qualifications and optimality conditions, for use in subsequent sections. We discuss first-order conditions and constraint qualifications in Section 2.1 and second-order conditions in Section 2.2, concluding with a result concerning local quadratic increase of the objective in a feasible neighborhood of x^* in Section 2.3.

2.1. First-Order Conditions and Constraint Qualifications.

We start by defining the following active sets at the point x^*, feasible for (1.1):

$$I_g \stackrel{\text{def}}{=} \{i = 1, 2, \ldots, p \,|\, g_i(x^*) = 0\}, \tag{2.1a}$$

$$I_G \stackrel{\text{def}}{=} \{i = 1, 2, \ldots, m \,|\, G_i(x^*) = 0\}, \tag{2.1b}$$

$$I_H \stackrel{\text{def}}{=} \{i = 1, 2, \ldots, m \,|\, H_i(x^*) = 0\}. \tag{2.1c}$$

Because x^* is feasible, we have $I_G \cup I_H = \{1, 2, \ldots, m\}$. The set $I_G \cap I_H$ is called the *biactive set*.

Our first definition of stationarity is as follows.

DEFINITION 2.1. *A point x^* that is feasible for (1.1) is* Bouligand- *or* B-stationary *if $d = 0$ solves the following linear program with equilibrium constraints (LPEC):*

$$
\begin{array}{ll}
\min_d & \nabla f(x^*)^T d \quad \text{subject to} \\
& g(x^*) + \nabla g(x^*)^T d \geq 0, \quad h(x^*) + \nabla h(x^*)^T d = 0, \\
& 0 \leq G(x^*) + \nabla G(x^*)^T d \perp H(x^*) + \nabla H(x^*)^T d \geq 0.
\end{array}
\tag{2.2}
$$

Checking B-stationarity is difficult in general, as it may require the solution of $2^{\bar{m}}$ linear programs, where $\bar{m}$ is the cardinality of the biactive set $I_G \cap I_H$. However, B-stationarity is implied by the following condition, which is more restrictive but much easier to check.

DEFINITION 2.2. *A point x^* that is feasible for (1.1) is* strongly stationary *if $d = 0$ solves the following linear program:*

$$
\begin{array}{ll}
\min_d & \nabla f(x^*)^T d \quad \text{subject to} \\
& g(x^*) + \nabla g(x^*)^T d \geq 0, \quad h(x^*) + \nabla h(x^*)^T d = 0, \\
& \nabla G_i(x^*)^T d = 0, \quad i \in I_G \setminus I_H, \\
& \nabla H_i(x^*)^T d = 0, \quad i \in I_H \setminus I_G, \\
& \nabla G_i(x^*)^T d \geq 0, \quad \nabla H_i(x^*)^T d \geq 0, \quad i \in I_G \cap I_H.
\end{array}
\tag{2.3}
$$

Note that (2.3) is the linearized approximation to the following nonlinear program, which is referred to as the *relaxed nonlinear program* (RNLP) for (1.1):

$$
\begin{array}{ll}
\min_x & f(x) \quad \text{subject to} \\
& g(x) \geq 0, \quad h(x) = 0, \\
& G_i(x) = 0, \quad i \in I_G \setminus I_H, \qquad\qquad \text{(RNLP)} \\
& H_i(x) = 0, \quad i \in I_H \setminus I_G, \\
& G_i(x) \geq 0, \quad H_i(x) \geq 0, \quad i \in I_G \cap I_H.
\end{array}
\tag{2.4}
$$

We also mention an interesting and useful observation of Anitescu [1, Theorem 2.2] that x^* is strongly stationary if and only if it is stationary for Reg(0), that is, there are Lagrange multipliers such that the KKT conditions are satisfied for this problem. A similar result by Fletcher et al. [7, Proposition 4.1] gives equivalence between strongly stationary points and stationary points of RegComp(0).

By introducing Lagrange multipliers, we can combine the optimality conditions for (2.3) with the feasibility conditions for x^* as follows:

$$
0 = \nabla f(x^*) - \sum_{i \in I_g} \lambda_i^* \nabla g_i(x^*) - \sum_{i=1}^{q} \mu_i^* \nabla h_i(x^*) - \sum_{i \in I_G} \tau_i^* \nabla G_i(x^*) - \sum_{i \in I_H} \nu_i^* \nabla H_i(x^*),
\tag{2.5a}
$$

$$0 = h_i(x^*), \quad i = 1, 2, \ldots, q, \tag{2.5b}$$

$$0 = g_i(x^*), \quad i \in I_g, \tag{2.5c}$$

$$0 < g_i(x^*), \quad i \in \{1, 2, \ldots, q\} \setminus I_g, \tag{2.5d}$$

$$0 \leq \lambda_i^*, \quad i \in I_g, \tag{2.5e}$$

$$0 = G_i(x^*), \quad i \in I_G, \tag{2.5f}$$

$$0 < G_i(x^*), \quad i \in \{1, 2, \ldots, m\} \setminus I_G, \tag{2.5g}$$

$$0 = H_i(x^*), \quad i \in I_H, \tag{2.5h}$$

$$0 < H_i(x^*), \quad i \in \{1, 2, \ldots, m\} \setminus I_H, \tag{2.5i}$$

$$(2.5j)\qquad 0 \le \tau_i^*, \quad i \in I_G \cap I_H,$$

$$(2.5k)\qquad 0 \le \nu_i^*, \quad i \in I_G \cap I_H.$$

Clearly, the Lagrange multipliers in (2.5) suffice for all $2^{\bar{m}}$ of the LPECs in (2.2). For a strongly stationary point x^*, we can now define the following sets:

$$(2.6a)\qquad I_g^+ \stackrel{\text{def}}{=} \{i \in I_g \,|\, \lambda_i^* > 0 \text{ for some } (\lambda^*, \mu^*, \tau^*, \nu^*) \text{ satisfying (2.5)}\},$$

$$(2.6b)\qquad I_g^0 \stackrel{\text{def}}{=} I_g \setminus I_g^+,$$

$$(2.6c)\qquad J_G^+ \stackrel{\text{def}}{=} \{i \in I_G \cap I_H \,|\, \tau_i^* > 0 \text{ for some } (\lambda^*, \mu^*, \tau^*, \nu^*) \text{ satisfying (2.5)}\},$$

$$(2.6d)\qquad J_G^0 \stackrel{\text{def}}{=} (I_G \cap I_H) \setminus J_G^+,$$

$$(2.6e)\qquad J_H^+ \stackrel{\text{def}}{=} \{i \in I_G \cap I_H \,|\, \nu_i^* > 0 \text{ for some } (\lambda^*, \mu^*, \tau^*, \nu^*) \text{ satisfying (2.5)}\},$$

$$(2.6f)\qquad J_H^0 \stackrel{\text{def}}{=} (I_G \cap I_H) \setminus J_H^+.$$

It is easy to show that there exists a multiplier $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) such that

$$(2.7a)\qquad i \in I_g^+ \;\Rightarrow\; \lambda_i^* > 0, \qquad i \in I_g^0 \;\Rightarrow\; \lambda_i^* = 0,$$

$$(2.7b)\qquad i \in J_G^+ \;\Rightarrow\; \tau_i^* > 0, \qquad i \in J_G^0 \;\Rightarrow\; \tau_i^* = 0,$$

$$(2.7c)\qquad i \in J_H^+ \;\Rightarrow\; \nu_i^* > 0, \qquad i \in J_H^0 \;\Rightarrow\; \nu_i^* = 0.$$

(The set of optimal multipliers is convex, so we can simply take an average of the multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ that satisfy (2.6a), (2.6c), (2.6e) individually.)

If the MPEC-LICQ (defined next) is satisfied, then the Lagrange multipliers for (2.3) are in fact unique, and in this case strong stationarity and B-stationarity are equivalent.

DEFINITION 2.3. *The MPEC-LICQ is satisfied at the point x^* if the following set of vectors is linearly independent:*

$$(2.8)\qquad \begin{array}{c} \{\nabla g_i(x^*) \,|\, i \in I_g\} \cup \{\nabla h_i(x^*) \,|\, i = 1, 2, \ldots, q\} \cup \\ \{\nabla G_i(x^*) \,|\, i \in I_G\} \cup \{\nabla H_i(x^*) \,|\, i \in I_H\}. \end{array}$$

In other words, the linear independence constraint qualification (LICQ) is satisfied for the RNLP (2.4).

We have the following result concerning first-order necessary conditions dating back to Luo, Pang and Ralph [21] but stated in the form of Scheel and Scholtes [27, Theorem 2].

THEOREM 2.4. *Suppose that x^* is a local minimizer of (1.1). Then if the MPEC-LICQ condition holds at x^*, then x^* is strongly stationary; and the multiplier vector $(\lambda^*, \mu^*, \tau^*, \nu^*)$ that satisfies the conditions (2.5) is unique.*

A number of our results use the following weaker Mangasarian-Fromovitz constraint qualification (MFCQ).

DEFINITION 2.5. *The MPEC-MFCQ is satisfied at x^* if the MFCQ is satisfied for the RNLP (2.4); that is, if there is a nonzero vector $d \in \mathbb{R}^n$ such that*

$$\begin{gathered}
\nabla G_i(x^*)^T d = 0, \;\; i \in I_G \setminus I_H, \;\; \nabla H_i(x^*)^T d = 0, \;\; i \in I_H \setminus I_G, \\
\nabla h_i(x^*)^T d = 0, \;\; i = 1, 2, \ldots, q, \;\; \nabla g_i(x^*)^T d > 0, \;\; i \in I_g, \\
\nabla G_i(x^*)^T d > 0 \text{ and } \nabla H_i(x^*)^T d > 0, \;\; i \in I_G \cap I_H; \text{ and} \\
\nabla G_i(x^*), \;\; i \in I_G \setminus I_H, \;\; \nabla H_i(x^*), \;\; i \in I_H \setminus I_G, \\
\nabla h_i(x^*), \;\; i = 1, 2, \ldots, q \text{ are all linearly independent.}
\end{gathered}$$

(It is easy to show, by using an argument like that of Gauvin [10] for nonlinear programming, that MPEC-MFCQ holds if and only if the set of multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) is bounded.)

We now define three varieties of strict complementarity at a strongly stationary point. To our knowledge, the second of these has only appeared before in the conditions for superlinear convergence of the elastic-mode penalty approach to MPCC analyzed in [1, Section 4].

Definition 2.6. *Let x^* be a strongly stationary point at which MPEC-LICQ is satisfied.*

(a) The upper-level strict complementarity *(USC) condition holds if $J_G^+ = J_H^+ = I_G \cap I_H$.*

(b) The partial strict complementarity *(PSC) condition holds if $J_G^+ \cup J_H^+ = I_G \cap I_H$.*

(c) Lower-level strict complementarity *(LSC) holds if $I_G \cap I_H = \emptyset$.*

It is obvious that LSC $\Rightarrow$ USC $\Rightarrow$ PSC. Strong stationarity and B-stationarity are equivalent when lower-level strict complementarity holds, since in this case the LPEC (2.2) reduces to the LP (2.3).

2.2. Second-Order Conditions. The set $\bar{S}$ of *normalized critical directions* for the RNLP (2.4) is defined as follows:

$$
\begin{aligned}
\bar{S} \stackrel{\text{def}}{=} & \{s \mid \|s\|_2 = 1\} \cap \\
& \{s \mid \nabla h(x^*)^T s = 0\} \cap \\
& \{s \mid \nabla g_i(x^*)^T s = 0 \ \text{ for all } i \in I_g^+\} \cap \\
& \{s \mid \nabla g_i(x^*)^T s \geq 0 \ \text{ for all } i \in I_g^0\} \cap \\
& \{s \mid \nabla G_i(x^*)^T s = 0 \ \text{ for all } i \in I_G \setminus I_H\} \cap \\
& \{s \mid \nabla G_i(x^*)^T s \geq 0 \ \text{ for all } i \in J_G^0\} \cap \\
& \{s \mid \nabla G_i(x^*)^T s = 0 \ \text{ for all } i \in J_G^+\} \cap \\
& \{s \mid \nabla H_i(x^*)^T s = 0 \ \text{ for all } i \in I_H \setminus I_G\} \cap \\
& \{s \mid \nabla H_i(x^*)^T s \geq 0 \ \text{ for all } i \in J_H^0\} \cap \\
& \{s \mid \nabla H_i(x^*)^T s = 0 \ \text{ for all } i \in J_H^+\}.
\end{aligned} \tag{2.9}
$$

By enforcing the additional condition that either $\nabla H_i(x^*)^T s = 0$ or $\nabla G_i(x^*)^T s = 0$, for all $i \in J_G^0 \cap J_H^0$, we obtain the set of normalized critical directions S^* for the MPEC (1.1) (see Scheel and Scholtes [27, eq. (6) and Section 3]); that is,

$$
S^* \stackrel{\text{def}}{=} \bar{S} \cap \{s \mid \min(\nabla H_i(x^*)^T s, \nabla G_i(x^*)^T s) = 0 \ \text{ for all } i \in J_G^0 \cap J_H^0\}. \tag{2.10}
$$

The difference between S^* and $\bar{S}$ vanishes if $J_G^0 \cap J_H^0 = \emptyset$, that is, if USC, LSC or PSC is satisfied.

We also define the MPEC Lagrangian as in Scholtes [28, Sec. 4]:

$$
L(x, \lambda, \mu, \tau, \nu) = f(x) - \lambda^T g(x) - \mu^T h(x) - \tau^T G(x) - \nu^T H(x). \tag{2.11}
$$

(Note that the expression in (2.5a) is the partial derivative of L with respect to x at the point $(x^*, \lambda^*, \mu^*, \tau^*, \nu^*)$, omitting the terms corresponding to inactive constraints.)

We are now ready to define second-order sufficient conditions.

DEFINITION 2.7. *Let x^* be a strongly stationary point. The MPEC-SOSC holds at x^* if there is $\sigma > 0$ such that for every $s \in S^*$, there are multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) such that*

$$s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) s \geq \sigma. \tag{2.12}$$

The RNLP-SOSC holds at x^* if for every $s \in \bar{S}$, there are multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) such that (2.12) holds.

Likewise, we define strong second-order sufficient conditions for the MPEC and RNLP. For the latter, the normalized critical direction set at x^* is as follows:

$$\begin{aligned}\bar{T} \stackrel{\text{def}}{=} &\{s \,|\, \|s\|_2 = 1\} \cap \\ &\{s \,|\, \nabla h(x^*)^T s = 0\} \cap \\ &\{s \,|\, \nabla g_i(x^*)^T s = 0 \text{ for all } i \in I_g^+\} \cap \\ &\{s \,|\, \nabla G_i(x^*)^T s = 0 \text{ for all } i \text{ with } \tau_i^* \neq 0\} \cap \\ &\{s \,|\, \nabla H_i(x^*)^T s = 0 \text{ for all } i \text{ with } \nu_i^* \neq 0\}.\end{aligned}$$

For the MPECs, the critical directions may be different for every "branch" of the feasible set containing x^*; see [20] for various "piecewise" optimality conditions using this motivation. For any partition $I \cup J$ of $J_G^0 \cap J_H^0$, let

$$T^*(I,J) \stackrel{\text{def}}{=} \bar{T} \cap \{s \,|\, \nabla G_i(x^*)^T s = 0 \text{ for all } i \in I\} \cap \{s \,|\, \nabla H_i(x^*)^T s = 0 \text{ for all } i \in J\}.$$

DEFINITION 2.8. *Let x^* be a strongly stationary point. The MPEC-SSOSC holds at x^* if there is $\sigma > 0$ such that for every partition $I \cup J$ of $J_G^0 \cap J_H^0$ and each $s \in T^*(I,J)$, there are multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) such that (2.12) holds. The RNLP-SSOSC holds at x^* if for every $s \in \bar{T}$, there are multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) such that (2.12) holds.*

When $J_G^0 \cap J_H^0$ is empty (that is, PSC holds), the index sets I and J in the definition of MPEC-SSOSC are also empty, so that $T^*(I,J) = \bar{T}$ and the strong second-order sufficient conditions of Definition 2.8 coincide. In general, we have $\bar{T} \supset \bar{S} \supset S^*$, so that RNLP-SSOSC $\Longrightarrow$ RNLP-SOSC $\Longrightarrow$ MPEC-SOSC. Similarly, we have MPEC-SSOSC $\Longrightarrow$ MPEC-SOSC. The following example, which will be referred to again later, shows how the direction sets above are defined and demonstrates that MPEC-SOSC is strictly weaker than RNLP-SOSC, and that MPEC-SSOSC is strictly weaker than RNLP-SSOSC. (A similar example appears in Scheel and Scholtes [27, p. 12].)

EXAMPLE 1. *Let $x = (x_1, x_2) \in \mathbb{R}^2$ and*

$$Q = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

The MPEC

$$\min\, x^T Q x \quad \text{subject to} \quad 0 \leq x_1 \perp x_2 \geq 0$$

has the origin $x^* = (0,0)$ as a global minimizer, and no other local minimizers or stationary points. The MPEC-LICQ holds at x^* and, taking $G(x) = x_1$ and $H(x) = x_2$, and the corresponding multipliers are $\tau^* = 0$ and $\nu^* = 0$. Hence, we have

$$I_G = I_H = I_G \cap I_H = J_G^0 = J_H^0 = \{1\}, \qquad J_G^+ = J_H^+ = \emptyset.$$

The Hessian of the MPEC-Lagrangian (2.11) is Q, and we have

$$S^* = \{(1,0),(0,1)\}, \qquad \bar{S} = \{(s_1,s_2) \geq 0 \,|\, s_1^2 + s_2^2 = 1\},$$

and

$$T^*(\{1\},\emptyset) = \{(0,1),(0,-1)\}, \qquad T^*(\emptyset,\{1\}) = \{(1,0),(-1,0)\},$$
$$\bar{T} = \{(s_1,s_2) \,|\, s_1^2 + s_2^2 = 1\}.$$

It is easy to check that MPEC-SSOSC, hence MPEC-SOSC, holds. However, RNLP-SOSC does not hold, and neither does RNLP-SSOSC, as there exists a direction of zero curvature in $\bar{S}$, namely $s = (1/\sqrt{2}, 1/\sqrt{2})$. We mention for later reference that the solution set of Reg(t) can easily be seen to be a continuum $\{(x_1,x_2) : 0 \leq x_1 = x_2 \leq \sqrt{t}\}$ for $t > 0$.

2.3. Local Quadratic Increase. We have the following result concerning quadratic growth of the objective function in a feasible neighborhood of a strongly stationary x^* at which MPEC-SOSC is satisfied.

THEOREM 2.9. *Suppose that x^* is a strongly stationary point of (1.1) at which MPEC-SOSC is satisfied. Then x^* is a strict local minimizer of (1.1) and in fact for any $\hat{\sigma} \in (0,\sigma)$ (where σ is from (2.12)), there is $r_0 > 0$ such that*

$$f(x) - f(x^*) \geq \hat{\sigma}\|x - x^*\|_2^2, \quad \text{for all } x \text{ feasible in (1.1) with } \|x - x^*\| \leq r_0. \tag{2.13}$$

Proof. This result follows from Scheel and Scholtes [27, Theorem 7(2)] and basic theory concerning quadratic growth for standard nonlinear programming; see for example Maurer and Zowe [22] and Robinson [26, Theorem 2.2]. □

We can still prove quadratic increase if we drop the strong stationarity assumption, and assume instead B-stationarity of x^* along with an SOSC for all nonlinear programs of the form

$$\begin{array}{ll} \min_x & f(x) \quad \text{subject to} \\ & g(x) \geq 0, \quad h(x) = 0, \\ & G_i(x) = 0, \quad \text{for all } i \in I'_G, \\ & G_i(x) \geq 0, \quad \text{for all } i \notin I'_G, \\ & H_i(x) = 0, \quad \text{for all } i \in I'_H, \\ & H_i(x) \geq 0, \quad \text{for all } i \notin I'_H. \end{array}$$

where I'_G and I'_H form a partition of $\{1,2,\ldots,m\}$ such that $I'_G \subset I_G$ and $I'_H \subset I_H$. (We do not give a formal statement or proof of this result, since it is not needed for subsequent sections of this paper.)

Note that if we assume RNLP-SOSC rather than the less stringent MPEC-SOSC, the quadratic increase result becomes a trivial consequence of standard nonlinear programming theory; see again Robinson [26].

3. Properties of Solutions of Reg(t). In this section, we investigate the minimizers of Reg(t) for small values of t. Our starting point is a result of Scholtes [28, Theorem 4.1], which we state in a slightly modified form below. This result requires the RNLP-SSOSC as well as an additional (and somewhat artificial) complementarity assumption involving the multipliers τ_i^*, $i \in I_G \setminus I_H$ and ν_i^*, $i \in I_H \setminus I_G$.

THEOREM 3.1. *Suppose that x^* is a strongly stationary point of (1.1) at which MPEC-LICQ, RNLP-SSOSC, and USC are satisfied. Assume in addition that $\tau_i^* \neq 0$*

for all $i \in I_G$ and $\nu_i^* \neq 0$ for all $i \in I_H$. Then for all $t > 0$ sufficiently small, the problem (1.2) has a unique stationary point $x(t)$ in a neighborhood of x^* that satisfies second-order sufficient conditions for (1.2) and hence is a strict local solution. Moreover, we have that $\|x(t) - x^*\| = O(t)$.

The original result also notes that $x(t)$ is a piecewise smooth function of t for small nonnegative t.

Our results in this section are of two main types: existence results and uniqueness results for solutions of Reg(t). We prove the existence results in Section 3.1. In Theorem 3.2, we weaken the assumptions in the theorem above by replacing RNLP-SSOSC by MPEC-SOSC and dropping the complementarity condition. The result is correspondingly weaker; we do not prove uniqueness of the solution of Reg(t) in the neighborhood of x^*, and show only that the distance from $x(t)$ to x^* satisfies an $O(t^{1/2})$ estimate. In Theorem 3.3, we recover the $O(t)$ estimate at the expense of using the RNLP-SOSC instead of MPEC-SOSC.

Section 3.2 demonstrates boundedness of the Lagrange multipliers for Reg(t) at solutions $x(t)$ for which $\|x(t)-x^*\| = O(t)$. In Section 3.3, we discuss local uniqueness of these solutions, and piecewise smoothness of the solution mapping $x(t)$, making use of the SSOSC of Definition 2.8.

3.1. Estimating Distance Between Solutions of Reg(t) and the MPEC Optimum. We now prove our first result concerning existence of a solution to Reg(t) near the solution x^* of (1.1) and its distance to x^*. This result is obtained by applying Bonnans and Shapiro [3, Theorem 5.57] to the problem Reg(0), which is

$$\begin{array}{lll} \text{Reg}(0): & \min_x & f(x) \quad \text{subject to} \\ & & g(x) \geq 0, \quad h(x) = 0, \\ & & G(x) \geq 0, \quad H(x) \geq 0, \quad G_i(x)H_i(x) \leq 0, \ \ i = 1, 2, \ldots, m. \end{array} \tag{3.1}$$

THEOREM 3.2. *Suppose that x^* is a strongly stationary point of (1.1) at which MPEC-MFCQ and MPEC-SOSC are satisfied. Then there are positive constants $\hat{r}_0$, $\bar{t}_2$, and M_2 such that for all $t \in (0, \bar{t}_2]$, the global solution $x(t)$ of the localized problem* Reg(t) *with the additional ball constraint $\|x - x^*\| \leq \hat{r}_0$ that lies closest to x^* satisfies $\|x(t) - x^*\| \leq M_2 t^{1/2}$.*

Proof. We prove the result by verifying that the conditions of [3, Theorem 5.57] are satisfied. First, because x^* is a strict local solution of (1.1) (and hence of (3.1)), we can choose $\hat{r}_0$ and impose the additional condition $\|x - x^*\|_2 \leq \hat{r}_0$ in (3.1). With this additional constraint, x^* is the unique global solution of the problem, so the first condition of [3, Theorem 5.57] holds. Moreover, since the feasible set for Reg(t) contains the feasible set for Reg(0), we have by applying the additional condition $\|x - x^*\|_2 \leq \hat{r}_0$ to (1.2) that the feasible set for the latter problem is nonempty and uniformly bounded, thereby ensuring that the fifth condition of [3, Theorem 5.57] is also satisfied.

The second condition in [3, Theorem 5.57] is Gollan's condition [3, (5.111)]. This condition reduces for our problem to the existence of a nonzero vector $\tilde{d} \in \mathbb{R}^n$ such that

$$\begin{gathered} \nabla h_i(x^*), \ \ i = 1, 2, \ldots, q \ \text{ are linearly independent;} \\ \nabla g_i(x^*)^T \tilde{d} > 0, \ \text{ for all } i \in I_g; \\ \nabla G_i(x^*)^T \tilde{d} > 0, \ \text{ for all } i \in I_G; \quad \nabla H_i(x^*)^T \tilde{d} > 0, \ \text{ for all } i \in I_H; \\ G_i(x^*) \nabla H_i(x^*)^T \tilde{d} + H_i(x^*) \nabla G_i(x^*)^T \tilde{d} < 1, \ \ i = 1, 2, \ldots, m. \end{gathered} \tag{3.2}$$

The linear independence condition in Definition 2.5 implies that we can choose $s \in \mathbb{R}^n$ such that

$$\nabla h_i(x^*)^T s = 0, \;\; i \in I_g,$$
$$\nabla G_i(x^*)^T s = 1, \;\; i \in I_G \setminus I_H, \;\; \nabla H_i(x^*)^T s = 1, \;\; i \in I_H \setminus I_G.$$

By setting $\tilde{d} = d + \alpha s$, where d is from Definition 2.5 and $\alpha > 0$ is sufficiently small, we can ensure that all conditions but the final one in (3.2) are satisfied. By scaling $\tilde{d}$ by an appropriate factor we can ensure that this condition is satisfied too.

The third condition of [3, Theorem 5.57]—existence of Lagrange multipliers for (3.1) at x^*—follows from (2.5) in a similar fashion to the proof of [7, Proposition 4.1]; see also the recent result of Anitescu [1, Theorem 2.2]. We seek Lagrange multipliers $\bar{\lambda}$, $\bar{\mu}$, $\bar{\tau}$, $\bar{\nu}$, and $\rho \in \mathbb{R}^m$ (the last one for the constraints $G_i(x)H_i(x) \le 0$) such that

$$\begin{aligned}
&(3.3\text{a}) \;\; 0 = \nabla f(x^*) - \sum_{i \in I_g} \bar{\lambda}_i \nabla g_i(x^*) - \sum_{i=1}^{q} \bar{\mu}_i \nabla h_i(x^*) \\
&\qquad\qquad - \sum_{i \in I_G} (\bar{\tau}_i - \rho_i H_i(x^*)) \nabla G_i(x^*) - \sum_{i \in I_H} (\bar{\nu}_i - \rho_i G_i(x^*)) \nabla H_i(x^*), \\
&(3.3\text{b}) \;\; 0 = h_i(x^*), \;\; i = 1, 2, \ldots, q, \\
&(3.3\text{c}) \;\; 0 = g_i(x^*), \;\; i \in I_g, \\
&(3.3\text{d}) \;\; 0 > g_i(x^*), \;\; i \in \{1, 2, \ldots, q\} \setminus I_g, \\
&(3.3\text{e}) \;\; 0 \le \bar{\mu}_i, \;\; i \in I_g, \\
&(3.3\text{f}) \;\; 0 = G_i(x^*), \;\; i \in I_G, \\
&(3.3\text{g}) \;\; 0 < G_i(x^*), \;\; i \in \{1, 2, \ldots, m\} \setminus I_G, \\
&(3.3\text{h}) \;\; 0 = H_i(x^*), \;\; i \in I_H, \\
&(3.3\text{i}) \;\; 0 < H_i(x^*), \;\; i \in \{1, 2, \ldots, m\} \setminus I_H, \\
&(3.3\text{j}) \;\; 0 \le \bar{\tau}_i, \;\; i \in I_G, \\
&(3.3\text{k}) \;\; 0 \le \bar{\nu}_i, \;\; i \in I_H, \\
&(3.3\text{l}) \;\; 0 \le \rho_i, \;\; i = 1, 2, \ldots, m.
\end{aligned}$$

Note that, in contrast to (2.5j) and (2.5k), nonnegativity is required of all $\bar{\tau}_i$, $i \in I_G$ and all $\bar{\nu}_i$, $i \in I_H$, not just the components in the biactive set $I_G \cap I_H$. Given any set of multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) and (2.7), we can set

$$\begin{aligned}
&(3.4\text{a}) && \bar{\lambda}_i = \lambda_i^*, \quad i \in I_g, \\
&(3.4\text{b}) && \bar{\mu}_i = \mu_i^*, \quad i = 1, 2, \ldots, q, \\
&(3.4\text{c}) && \bar{\tau}_i = \tau_i^* + \rho_i H_i(x^*), \quad i \in I_G, \\
&(3.4\text{d}) && \bar{\nu}_i = \nu_i^* + \rho_i G_i(x^*), \quad i \in I_H,
\end{aligned}$$

where the multipliers ρ_i, $i = 1, 2, \ldots, m$ satisfy

$$\begin{aligned}
&(3.5\text{a}) && \rho_i \ge \bar{\rho}_i \stackrel{\text{def}}{=} \max\left(0, \frac{-\tau_i^*}{H_i(x^*)}\right), \;\; i \in I_G \setminus I_H; \\
&(3.5\text{b}) && \rho_i \ge \bar{\rho}_i \stackrel{\text{def}}{=} \max\left(0, \frac{-\nu_i^*}{G_i(x^*)}\right), \;\; i \in I_H \setminus I_G; \\
&(3.5\text{c}) && \rho_i \ge \bar{\rho}_i \stackrel{\text{def}}{=} 0, \;\; i \in I_G \cap I_H.
\end{aligned}$$

It is easy to check that the resulting multipliers satisfy (3.3). Note in particular that

$$\tau_i^* = \bar{\tau}_i, \;\; \nu_i^* = \bar{\nu}_i, \quad i \in I_G \cap I_H. \tag{3.6}$$

The fourth condition in [3, Theorem 5.57] requires second-order sufficient conditions for (3.1) to hold. Because of (3.6), the critical direction set for this problem is $\bar{S}$—the same as for the RNLP (2.4). Defining $\bar{L}$ to be the Lagrangian for (3.1), it is easy to see from the relations (3.4) that

$$\begin{aligned} &\nabla^2_{xx}\bar{L}(x^*, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho) = \nabla^2_{xx}L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) \\ &+ \sum_{i=1}^m \rho_i \left(\nabla G_i(x^*) \nabla H_i(x^*)^T + \nabla H_i(x^*) \nabla G_i(x^*)^T \right). \end{aligned} \tag{3.7}$$

By using Definition 2.7 and the definition (2.10) of S^*, we can find an $\epsilon > 0$ such that for each

$$s \in \bar{S} \cap \{ s \,|\, \min(\nabla H_i(x^*)^T s, \nabla G_i(x^*)^T s) \le \epsilon \;\text{ for all } i \in J_G^0 \cap J_H^0 \}, \tag{3.8}$$

there exists a tuple of MPEC multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ (satisfying (2.5)), hence a corresponding tuple of multipliers $(\bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho)$ satisfying (3.4) and (3.5), such that

$$s^T \nabla^2_{xx}\bar{L}(x, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho) s \ge s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) s \ge \sigma/2$$

where σ is from Definition 2.7. For all $s \in \bar{S}$ but not in the set (3.8), we have $\nabla H_i(x^*)^T s > \epsilon$ and $\nabla G_i(x^*)^T s > \epsilon$ for at least one $i \in J_G^0 \cap J_H^0$, so that

$$\begin{aligned} &\sup_{(\bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho)} s^T \nabla^2_{xx}\bar{L}(x, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho) s \\ &\qquad \ge \sup_{(\lambda^*, \mu^*, \tau^*, \nu^*)} s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) s + 2 \min_{i \in J_G^0 \cap J_H^0} \rho_i \epsilon^2 \end{aligned}$$

where, here and below, the supremum at left (right) is taken over the multipliers for (3.1) (MPEC multipliers, respectively). In addition to (3.5), we now require that $\rho_i \ge \hat{\rho}$ for all $i \in J_G^0 \cap J_H^0$, where $\hat{\rho}$ is large enough that the following condition holds:

$$\inf_{s \in \bar{S}} \sup_{(\lambda^*, \mu^*, \tau^*, \nu^*)} s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) s + 2\hat{\rho}\epsilon^2 \ge \sigma/2$$

Under these additional conditions on ρ, we have that

$$\sup_{(\bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho)} s^T \nabla^2_{xx}\bar{L}(x, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho) s \ge \sigma/2, \qquad \text{for all } s \in \bar{S}.$$

Hence, second-order sufficient conditions for (3.1) are satisfied at x^*, so the fourth condition of [3, Theorem 5.57] is also satisfied.

The result now follows immediately from [3, Theorem 5.57]. $\square$

When the RNLP-SOSC replaces MPEC-SOSC and MPEC-LICQ replaces MPEC-MFCQ, we can strengthen the bound to $\|x(t) - x^*\| = O(t)$.

THEOREM 3.3. *Suppose that x^* is a strongly stationary point of (1.1) at which MPEC-LICQ and RNLP-SOSC are satisfied, and let $\hat{r}_0$ be the positive constant defined in Theorem 3.2. Then there is a value $\bar{t}_3 > 0$ and a constant M_3 such that for all $t \in (0, \bar{t}_3]$, the global solution $x(t)$ of the localized problem* Reg(t) *with the additional ball constraint $\|x - x^*\| \le \hat{r}_0/2$ that lies closest to x^* satisfies $\|x(t) - x^*\| \le M_3 t$.*

Proof. We prove the result by invoking [3, Theorem 4.55]. Our task is to show that the three conditions of this theorem are satisfied by the limiting problem (3.1). We discuss these three conditions in the order (i), (iii), (ii).

To make the connections with the notation in [3], we write Reg(t) in the following general form:

$$\min f(x) \quad \text{subject to} \quad C(x,t) \in K, \tag{3.9}$$

where K in our case is a polyhedral convex cone (a Cartesian product of zeros and half-lines), and t appears in the constraints $C(x,t)$ as the linear term tv, where v is a vector consisting of zeros, except for -1 in the locations corresponding to the constraints $G_i(x)H_i(x) - t \le 0$.

Condition (i) of the cited theorem requires the Lagrange multiplier set for (3.1) to be nonempty and a "directional regularity" condition to be satisfied. We verified existence of Lagrange multipliers already in the proof of Theorem 3.2, while the directional regularity condition reduces for this problem to Gollan's condition, which has also been verified in our earlier proof.

Condition (iii) is automatic for our problem since K above is polyhedral and convex; see [3, Remark 4.59].

We turn now to condition (ii), which is a second-order sufficient condition [3, (4.139)]. Note first that the σ term in [3, (4.139)] can be ignored because of the polyhedral convex nature of our set K in (3.9). We start by expanding on results in the proof of Theorem 3.2, and then discuss the set of optimal multipliers for Reg(0) and define linearized dual problem for Reg(t) in terms of this set.

Let us introduce the Lagrangian $\tilde{L}$ for Reg(t), where

$$\begin{aligned} &\tilde{L}(x,t,\lambda,\mu,\tau,\nu,\rho) \\ &\quad = f(x) - \lambda^T g(x) - \mu^T h(x) - \tau^T G(x) - \nu^T H(x) + \sum_{i=1}^m \rho_i (G_i(x)H_i(x) - t). \end{aligned} \tag{3.10}$$

Note that when $t = 0$, we have

$$\tilde{L}(x,0,\lambda,\mu,\tau,\nu,\rho) = \bar{L}(x,\lambda,\mu,\tau,\nu,\rho), \tag{3.11}$$

for $\bar{L}$ defined in the proof of Theorem 3.2. As shown there, the set of optimal multipliers for Reg(0) can be defined by taking the union, over all MPEC multipliers $(\lambda^*,\mu^*,\tau^*,\nu^*)$ (satisfying (2.5)), of the corresponding multipliers $(\bar{\lambda},\bar{\mu},\bar{\tau},\bar{\nu},\rho)$ defined in (3.4), (3.5), where $\rho \ge \bar{\rho}$ and the components of $\bar{\rho}$ are defined in (3.5). Since we assume MPEC-LICQ, the MPEC multiplier $(\lambda^*,\mu^*,\tau^*,\nu^*)$ is in fact unique, so the multipliers $(\bar{\lambda},\bar{\mu},\bar{\tau},\bar{\nu},\rho)$ depend only on ρ, a dependence we indicate explicitly by writing $(\bar{\lambda}(\rho),\bar{\mu}(\rho),\bar{\tau}(\rho),\bar{\nu}(\rho),\rho)$. The linearized dual problem for Reg(0), following the general definition in [3, (4.46)], is as follows:

$$\max_{(\bar{\lambda}(\rho),\bar{\mu}(\rho),\bar{\tau}(\rho),\bar{\nu}(\rho),\rho);\rho\ge\bar{\rho}} \quad D_t\tilde{L}(x^*,0,\bar{\lambda}(\rho),\bar{\mu}(\rho),\bar{\tau}(\rho),\bar{\nu}(\rho),\rho), \tag{3.12}$$

From the definition (3.10), this problem reduces to

$$\min_{(\bar{\lambda}(\rho),\bar{\mu}(\rho),\bar{\tau}(\rho),\bar{\nu}(\rho),\rho);\rho\ge\bar{\rho}} \quad \sum_{i=1}^m \rho_i,$$

whose (unique) solution is obviously $(\bar{\lambda}(\bar{\rho}),\bar{\mu}(\bar{\rho}),\bar{\tau}(\bar{\rho}),\bar{\nu}(\bar{\rho}),\bar{\rho})$.

The condition [3, (4.139)] now reduces to the following:

$$s^T \nabla^2_{xx} \tilde{L}(x^*, 0, \bar{\lambda}(\bar{\rho}), \bar{\mu}(\bar{\rho}), \bar{\tau}(\bar{\rho}), \bar{\nu}(\bar{\rho}), \bar{\rho}) s \; > \; 0, \;\; \text{for all } s \in \bar{S}, \tag{3.13}$$

since, as we mentioned in the proof of Theorem 3.2, the critical direction set for (3.1) is the same as the critical direction set $\bar{S}$ (2.9) for the RNLP (2.4). From (3.7) and (3.11), we have that

$$\begin{aligned} &\nabla^2_{xx} \tilde{L}(x^*, 0, \bar{\lambda}(\bar{\rho}), \bar{\mu}(\bar{\rho}), \bar{\tau}(\bar{\rho}), \bar{\nu}(\bar{\rho}), \bar{\rho}) \\ &= \nabla^2_{xx} L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) + \sum_{i=1}^m \bar{\rho}_i \left(\nabla G_i(x^*) \nabla H_i(x^*)^T + \nabla H_i(x^*) \nabla G_i(x^*)^T \right), \end{aligned}$$

so that

$$\begin{aligned} &s^T \nabla^2_{xx} \tilde{L}(x^*, 0, \bar{\lambda}(\bar{\rho}), \bar{\mu}(\bar{\rho}), \bar{\tau}(\bar{\rho}), \bar{\nu}(\bar{\rho}), \bar{\rho}) s \\ &= s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) s + 2 \sum_{i=1}^m \bar{\rho}_i \left(\nabla G_i(x^*)^T s \right) \left(\nabla H_i(x^*)^T s \right). \end{aligned}$$

Because $s \in \bar{S}$, and because $(\lambda^*, \mu^*, \tau^*, \nu^*)$ is the unique multiplier satisfying (2.5), we have by RNLP-SOSC (Definition 2.7) that the first term on the right-hand side of this equation is at least $\sigma > 0$. Moreover, since $\bar{\rho} \geq 0$, $\nabla G(x^*)^T s \geq 0$, and $\nabla H(x^*)^T s \geq 0$, the summation in the final term is nonnegative. We conclude that (3.13), and hence condition (ii) of [3, Theorem 4.55] is satisfied.

We conclude that the three conditions of [3, Theorem 4.55] are satisfied, so our result follows directly from the cited theorem. □

The next result follows immediately from Theorem 3.3 when we note that the MPEC-SOSC and RNLP-SOSC conditions are identical when PSC holds.

COROLLARY 3.4. *Suppose that x^* is a strongly stationary point of (1.1) at which MPEC-LICQ and MPEC-SOSC are satisfied, and that the partial strict complementarity (PSC) condition holds. Then there is a value $\bar{t}_3 > 0$ and a constant M_3 such that for all $t \in (0, \bar{t}_3]$, the global solution $x(t)$ of the localized problem* Reg(t) *with the additional ball constraint $\|x - x^*\| \leq \hat{r}_0/2$ that lies closest to x^* satisfies $\|x(t) - x^*\| \leq M_3 t$, where $\hat{r}_0$ is as defined in Theorem 3.2.*

We conclude this subsection by illustrating the difference between Theorems 3.2 and 3.3 using Example 1. There we can take $x(t) = (\sqrt{t}, \sqrt{t})$, hence $\|x(t) - x^*\| = O(t^{1/2})$. The $O(t)$ estimate of Theorem 3.3 does not hold because RNLP-SOSC is not satisfied.

3.2. Boundedness of Lagrange Multipliers in Reg(t). We now establish a companion result for Theorem 3.3 and Corollary 3.4 concerning boundedness of the Lagrange multipliers at the solutions of Reg(t) described in those results. The main result, Proposition 3.6, is proved after the following simple technical preliminary.

LEMMA 3.5. *Consider any $i \in I_G \cap I_H$ and suppose that $\nabla G_i(x^*)$ and $\nabla H_i(x^*)$ are nonzero vectors. Then there exist a neighborhood U_i of x^* and positive constant c_i such that for any $x \in U_i$ and $t \geq 0$ with $G_i(x) H_i(x) = t$, we have $\|x - x^*\| \geq c_i \sqrt{t}$.*

Proof. Suppose for contradiction that there is a sequence $t_k \downarrow 0$, and corresponding x^k with $G_i(x^k) H_i(x^k) = t_k$, $G_i(x^k) > 0$, $H_i(x^k) > 0$, such that

$$\sqrt{t_k} / \|x^k - x^*\| \to \infty. \tag{3.14}$$

By taking a subsequence if necessary, we have that either $G_i(x^k) \geq \sqrt{t_k}$ for all k, or a similar bound on $H_i(x^k)$. In the former case, for all k sufficiently large, we have from $\nabla G_i(x^*) \neq 0$ that

$$\begin{aligned}\sqrt{t_k} &\leq G_i(x^k) = G_i(x^k) - G_i(x^*) = \nabla G_i(x^*)(x^k - x^*) + o(\|x^k - x^*\|) \\ &\leq 2\|\nabla G_i(x^*)\|\|x^k - x^*\|,\end{aligned}$$

which contradicts (3.14). A similar contradiction occurs in the latter case. ∎

PROPOSITION 3.6. *Let x^* be a strongly stationary point of (1.1) at which the MPEC-LICQ holds. If the regularized solution $x(t)$ satisfies $\|x(t) - x^*\| = O(t)$ for small positive t, then*

(i) $G_i(x(t))H_i(x(t)) < t$ for each small positive t and each $i \in I_G \cap I_H$; and

(ii) the Lagrange multipliers corresponding to $x(t)$ are bounded as $0 < t \to 0$.

Proof. Because of MPEC-LICQ, we have that each biactive pair $\nabla G_i(x^*)$ and $\nabla H_i(x^*)$ are linearly independent. Apply Lemma 3.5 to each $i \in I_G \cap I_H$ and combine the results to obtain a neighborhood U of x^* and positive constants $\hat{c}$ and $\hat{t}$ with the following property: If $0 \leq t \leq \hat{t}$, $x \in U$ and $G_i(x)H_i(x) = t$ for some biactive index $i \in I_G \cap I_H$, then $\|x - x^*\| \geq c\sqrt{t}$. Since $\|x(t) - x^*\| = O(t)$, we have for small $t > 0$ that $x(t) \in U$, $0 \leq t \leq \hat{t}$, and $\|x(t) - x^*\| < c\sqrt{t}$. Hence the constraint $G_i(x(t))H_i(x(t)) \leq t$ must be inactive, proving (i).

It follows from (i) that $\delta_i(t) = 0$ for all $i \in I_G \cap I_H$ and all t sufficiently small. From Scholtes [28, Theorem 3.1], we have the following convergence result for multipliers of Reg(t):

$$\lambda_i(t) \to \lambda_i^*, \quad \text{for all } i \in I_g, \tag{3.15a}$$

$$\mu(t) \to \mu^*, \tag{3.15b}$$

$$\tau_i(t) - \delta_i(t)H_i(x(t)) \to \tau_i^*, \quad \text{for all } i \in I_G, \tag{3.15c}$$

$$\nu_i(t) - \delta_i(t)G_i(x(t)) \to \nu_i^*, \quad \text{for all } i \in I_H. \tag{3.15d}$$

Since $\delta_i(t) = 0$ for $i \in I_G \cap I_H$, it follows from (3.15c) and (3.15d) that $\tau_i(t) \to \tau_i^*$ and $\nu_i(t) \to \nu_i^*$ for these indices. For $i \in I_G \setminus I_H$, we cannot have both $G_i(x(t)) = 0$ and $G_i(x(t))H_i(x(t)) = t$, so either or both of $\tau_i(t)$ and $\delta_i(t)$ must be zero. Checking (3.15c) in each case shows that the resulting multipliers $\tau_i(t)$ and $\delta_i(t)$ must be bounded. Boundedness of $\nu_i(t)$ and $\delta_i(t)$ likewise follows from (3.15d), for $i \in I_H \setminus I_G$. This completes the proof of (ii). ∎

3.3. Local Uniqueness of Solutions to Reg(t). In this subsection, we present a further refinement of Scholtes' result [28, Theorem 4.1] that has been mentioned several times above. The main difference between Theorem 3.7, below, and the existence results of Section 3.1 is that, in addition to an $O(t)$ bound on $\|x^* - x(t)\|$, it provides local uniqueness of $x(t)$ under RNLP-SSOSC. While a strong second-order sufficient condition is to be expected as a sufficient condition for uniqueness, one might hope to use the weaker MPEC-SSOSC. However, Example 1 dispels this hope: the MPEC-LICQ and MPEC-SSOSC hold at the strongly stationary point $x^* = (0, 0)$, but the solution of Reg(t) is not unique for positive t.

We present two main results below. In the first, Theorem 3.7, we weaken the assumptions of [28, Theorem 4.1] by dropping LSC altogether, while retaining similar conclusions. The second result, Corollary 3.9, assumes MPEC-SSOSC instead of RNLP-SSOSC and replaces LSC by the weaker PSC condition.

THEOREM 3.7. *Suppose that x^* is a strongly stationary point of (1.1) at which MPEC-LICQ and RNLP-SSOSC are satisfied. Then there exist a neighborhood U of*

x^*, a scalar $\bar t > 0$, and a piecewise smooth function $z : (-\bar t, \bar t) \to U$ such that $x(t) = z(t)$ is the unique stationary point of $\mathrm{Reg}(t)$ *in U for every $t \in [0, \bar t)$. A consequence of piecewise smoothness is that, for $s, t \in [0, \bar t)$, we have $\|x(s) - x(t)\| = O(|s - t|)$; in particular $\|x(t) - x^*\| = O(t)$.*

The proof of the theorem relies first on showing that $x(t)$ is one of finitely many local solution mappings of strongly stable NLPs and, second, on a somewhat involved argument to establish uniqueness of $x(t)$ within a neighborhood of x^*. By contrast, under LSC, the good behavior including uniqueness of $x(t)$ follows immediately by observing that it is the solution of a single, strongly stable nonlinear program whose constraints are identified by the sign of the multipliers of the active constraints of G and H; see [28].

A key step toward the proof of Theorem 3.7 is the following technical result.

LEMMA 3.8. *Let f, g, h, γ, and ϕ be functions from $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ to $\mathbb{R}$, $\mathbb{R}^l$, $\mathbb{R}^m$, $\mathbb{R}$, and $\mathbb{R}$ respectively. Suppose that each of the following parametric problems is strongly stable about $(x^*, 0)$, meaning that there is a neighborhood of x^* such that for small perturbations of t about zero the parametric problem has a unique solution (stationary point or local minimizer) in that neighborhood:*

$$\begin{array}{ll} \min\limits_x & f(x,t) \quad \text{subject to} \\ & g(x,t) \ge 0, \quad h(x,t) = 0; \end{array} \tag{3.16a}$$

$$\begin{array}{ll} \min\limits_x & f(x,t) \quad \text{subject to} \\ & g(x,t) \ge 0, \quad h(x,t) = 0, \quad \gamma(x,t) \ge 0; \end{array} \tag{3.16b}$$

$$\begin{array}{ll} \min\limits_x & f(x,t) \quad \text{subject to} \\ & g(x,t) \ge 0, \quad h(x,t) = 0, \quad \phi(x,t) \ge 0. \end{array} \tag{3.16c}$$

Suppose further, for each x near x^* with $g(x, t) \ge 0$ and $h(x, t) = 0$, that $\gamma(x, t) \le 0$ implies $\phi(x, t) \ge 0$; and $\phi(x, t) \le 0$ implies $\gamma(x, t) \ge 0$. Then the problem

$$\begin{array}{ll} \min\limits_x & f(x,t) \quad \text{subject to} \\ & g(x,t) \ge 0, \quad h(x,t) = 0, \quad \gamma(x,t) \ge 0, \quad \phi(x,t) \ge 0 \end{array} \tag{3.17}$$

is also strongly stable at $(x^*, 0)$, and the local solution mapping $x(t)$ for this problem is a selection of the local solution mappings for the previous problems.

Proof. Let $x^1(t)$, $x^2(t)$, and $x^3(t)$ denote the local solution mappings of (3.16a), (3.16b) and (3.16c) respectively. We discuss existence and uniqueness of the solution $x(t)$ of (3.17) in turn.

a) Existence. If any one of $x^1(t)$, $x^2(t)$, and $x^3(t)$ is feasible for (3.17) then it is a solution of this problem because the feasible set of (3.17) is contained in the feasible set of each of the other problems. Suppose $x^1(t)$ is not feasible for (3.17), for example $\gamma(x^1(t)) < 0$ and, of course, $x^1(t) \neq x^2(t)$. Then $\gamma(x^2(t)) = 0$, otherwise $\gamma(x^2(t)) > 0$, in which case $x^2(t)$ is a local minimizer of both (3.16a) and (3.16b), which implies $x^1(t) = x^2(t)$ by uniqueness, a contradiction. Now use the relationship between γ and ϕ, which requires that $\phi(x^2(t)) \ge 0$, i.e. $x^2(t)$ is a solution of (3.17). A similar argument exchanging the roles of γ and ϕ shows that $x^3(t)$ is a solution of (3.17) if $\phi(x^1(t)) < 0$.

b) Uniqueness. Let $x^4(t)$ be a solution of (3.17) near x^*, for t near 0. If $\gamma(x^4(t))$ and $\phi(x^4(t))$ are both positive then $x^4(t)$ is also a solution of (3.16a), hence coincides with $x^1(t)$ by uniqueness of the latter. Similarly, $x^4(t) = x^2(t)$ if $\gamma(x^4(t)) = 0 <$

$\phi(x^4(t))$ and $x^4(t) = x^3(t)$ if $\gamma(x^4(t)) > 0 = \phi(x^4(t))$. That is, $x^4(t)$ is a selection of $\{x^1(t), x^2(t), x^3(t)\}$.

If $x^1(t)$ and $x^2(t)$ are both solutions of (3.17) then obviously the former is also a solution of (3.16b), and they coincide by uniqueness of the latter. Likewise if $x^1(t)$ and $x^3(t)$ are both solutions of (3.17) then they coincide.

Finally, let $x^2(t)$ and $x^3(t)$ be solutions of (3.17). We show by contradiction that $x^1(t)$ must be feasible for this problem, hence $x^4(t) = x^1(t) = x^2(t) = x^3(t)$. Assume $x^1(t)$ is infeasible for (3.17), say $\gamma(x^1(t)) < 0$. The relationship between γ and ϕ requires that $\phi(x^1(t)) \geq 0$, i.e. $x^1(t)$ is feasible for (3.16c) and therefore coincides with $x^3(t)$. But $x^3(t)$ is feasible for (3.17), a contradiction. A similar argument shows a contradiction if we assume $\phi(x^1(t)) < 0$. $\square$

Proof of Theorem 3.7. To unburden notation we assume without loss of generality, by exchanging G_i with H_i if necessary, that $I_G = \{1, \ldots, m\}$. Define $I^0 = \{i \in I_G \setminus I_H : \tau_i^* = 0\}$; note that the corresponding set $\{i \in I_H \setminus I_G : \nu_i^* = 0\}$ is empty.

Define "minimal core" constraints as follows:

$$\begin{array}{ll} g(x) \geq 0, & h(x) = 0, \\ G_i(x) \geq 0, & \text{if } i \in I_G \cap I_H \text{ or } \tau_i^* > 0 \\ H_i(x) \geq 0, & \text{if } i \in I_G \cap I_H \text{ or } \nu_i^* > 0 \\ F_i(x) \leq t, & \text{if } \tau_i^* + \nu_i^* < 0. \end{array}$$

Define "core" constraints as any set composed of the minimal core as well as, for each $i \in I_G \setminus I_H$ with $\tau_i^* = 0$, at most one of $G_i(x) \geq 0$ and $F_i(x) \leq t$.

Choose any set of core constraints and consider the corresponding "core NLP" which is parametric in t,

$$\min_x f(x) \quad \text{subject to} \quad x \text{ satisfies the chosen core constraints.}$$

When $t = 0$, because of MPEC-LICQ, x^* is a solution of this core NLP at which the LICQ and SSOSC hold; hence classical perturbation theory [3, 26] says that the core NLP is strongly stable at $(x^*, 0)$ and the local solution mapping is piecewise smooth in t. Call this problem NLP(1). Take $i \in I^0$ such that neither $G_i(x,t) \geq 0$ nor $F_i(x,t) \leq t$ is in the core. Define NLP(2) by adding the constraint $G_i(x,t) \geq 0$ to NLP(1); and NLP(3) by adding the constraint $F_i(x,t) \leq t$ to NLP(1). Then each of NLP(1)-(3) is a core NLP (using a different set of core constraints), hence is strongly stable at $(x^*, 0)$. It is easy to see that Lemma 3.8 can be applied by taking (3.16a), (3.16b), (3.16c) as NLP(1), NLP(2), NLP(3) respectively, yielding strong stability of the new problem (corresponding to (3.17)): $\min_x f(x)$ subject to the constraints of NLP(1) and also constraints

$$G_i(x,t) \geq 0 \quad \text{and} \quad F_i(x,t) \leq t. \tag{3.18}$$

The lemma also says that the local solution mapping for the fourth problem (call it $x^{(4)}(t)$) is a selection of the local solution mappings of NLP(1)-(3), therefore $x^{(4)}(t)$ is also piecewise smooth. Thus we have fulfilled the following induction hypothesis for $k = 1$.

Induction hypothesis[1] k: Choose any distinct $i_1, \ldots, i_k \in I^0$ and any set of core constraints that includes neither $G_i(x,t) \geq 0$ nor $F_i(x,t) \leq t$ for $i = i_1, \ldots, i_k$. Then

[1] The assumption $I_G = \{1, \ldots, m\}$ means we need not also consider pairs of constraints $H_i(x,t) \geq 0$ or $F_i(x,t) \leq t$.

the NLP with constraints given by the chosen core and (3.18) for all $i = i_1, \ldots, i_k$ is strongly stable at $(x^*, 0)$, and the associated local solution mapping is piecewise smooth in t.

Let k be at least one and less than the cardinality of I^0. We now show that the induction hypothesis holds for $k+1$. Choose any distinct $i_1, \ldots, i_{k+1} \in I^0$ and any set of core constraints that includes neither $G_i(x,t) \geq 0$ nor $F_i(x,t) \leq t$ for $i = i_1, \ldots, i_{k+1}$. Consider three NLPs, each with the objective function f. The first problem, NLP(i), has constraints given the by chosen core with the additional constraints (3.18) for $i = i_1, \ldots, i_k$. The second (third resp.) problem NLP(ii) (NLP(iii) resp.) is derived from NLP(i) by adding the constraint $G_{i_{k+1}}(x) \geq 0$ ($F_{i_{k+1}}(x) \leq t$ resp.). The constraints of each of NLP(i)-(iii) can be written as the union of a core set together with (3.18) for $i = i_1, \ldots, i_k$, i.e. in the form of the NLP specified in Induction Hypothesis k. This is obvious for NLP(i). For NLP(ii), take the core to be the chosen core as well as $G_{i_{k+1}}(x) \geq 0$; and for NLP(iii), the chosen core as well as $F_{i_{k+1}}(x) \leq t$. Hence each NLP(i)-(iii) is strongly stable at $(x^*, 0)$. Lemma 3.8 says that the NLP with objective f and constraints consisting of the chosen core and the pair (3.18) for all $i = i_1, \ldots, i_{k+1}$ is also strongly stable at $(x^*, 0)$, and that its local solution mapping, denoted $x^{(iv)}(t)$, is the selection of the local solution mappings of NLP(i)-(iii); so $x^{(iv)}(t)$ is also piecewise smooth. □

The last result here follows from the above theorem simply because, under PSC, MPEC-SSOSC implies (is equivalent to) RNLP-SSOSC.

Corollary 3.9. *The conclusions of Theorem 3.7 hold if x^* is a strongly stationary point of (1.1) at which MPEC-LICQ, MPEC-SSOSC, and PSC hold.*

4. Alternative Regularized Formulations. We now consider the alternative regularized formulations RegComp(t) and RegEq(t), and discuss the possibility of results like Theorems 3.2 and 3.3 holding for these formulations.

4.1. Properties of Solutions of RegComp(t). For RegComp(t), in which the individual constraints $G_i(x)H_i(x) \leq t$ are replaced by a single "approximate complementarity" constraint $G(x)^T H(x) \leq t$, the feasible region contains that of the original problem (1.1) and is a subset of the feasible region for Reg(t). Analogs of Theorems 3.2 and 3.3 hold, with RegComp(t) replacing Reg(t), and the proofs are quite similar. (We omit the details.) However, local uniqueness of the solution of RegComp(t) is difficult to ensure. Scholtes [28] mentions a private communication of Hu which shows that [28, Theorem 4.1] does not extend to RegComp(t). In Hu's counterexample, which is presented in [11, Example 2.3.2], all conditions of [28, Theorem 4.1], hence of Theorem 3.7 above, are shown to hold but the solutions of the RegComp(t) are not unique.

4.2. Properties of Solutions of RegEq(t). A result like Theorem 3.2 holds for the RegEq(t) formulation (1.4) as well, but only if the $O(t^{1/2})$ estimate is replaced by a weaker $O(t^{1/4})$ estimate. The following result differs from Theorem 3.2 also in that MPEC-LICQ is assumed in place of MPEC-MFCQ. The result [3, Theorem 5.57] cannot be applied here, as Gollan's directional regularity condition (constraint qualification) does not hold for this formulation. Our proof is based on more elementary results.

Theorem 4.1. *Suppose that x^* is a strongly stationary point of (1.1) at which MPEC-LICQ and MPEC-SOSC are satisfied. Then there are positive constants $\hat{r}_2$, $\bar{t}_4$, and M_7 such that for all $t \in (0, \bar{t}_4]$, the global solution $x(t)$ of the localized problem* RegEq(t) *with the additional ball constraint $\|x - x^*\| \leq \hat{r}_2$ that lies closest to x^**

satisfies $\|x(t) - x^*\| \le M_7 t^{1/4}$.

Proof. Our strategy is to define two balls about x^* with the following properties:

- The inner ball has radius $O(t^{1/4})$, while the outer ball has a constant radius;
- There is at least one feasible point $z(t)$ for RegEq(t) in the inner ball;
- All feasible points for RegEq(t) in the annulus between the two balls have a larger function value than $f(z(t))$.

It follows from these facts that the minimizer $x(t)$ described in the proof of the theorem lies inside the inner ball, so the $O(t^{1/4})$ estimate is satisfied.

Consider first the following projection problem, a nonlinear program parametrized by t:

$$\begin{gathered}\min_x \tfrac{1}{2}\|x - x^*\|_2^2 \quad \text{subject to} \\ g(x) \le 0, \quad h(x) = 0, \\ G_i(x) = t^{1/2}, \quad H_i(x) = t^{1/2} \quad (i \in I_G \cap I_H), \\ G_i(x)H_i(x) = t \quad (i \notin I_G \cap I_H).\end{gathered} \tag{4.1}$$

When $t = 0$, the solution is x^* and the gradients of the active constraints are linearly independent, by the MPEC-LICQ assumption (Definition 2.3). Since the objective is strongly convex, standard perturbation theory shows that the solution $z(t)$ of this problem satisfies

$$\|z(t) - x^*\| \le M_6 t^{1/2}, \tag{4.2}$$

for some constant $M_6 > 0$ and all t sufficiently small.

We now choose $\hat{r}_2$ such that the following properties hold:

$$\hat{r}_2 \le r_0, \tag{4.3a}$$
$$\hat{r}_2 \le r_1, \tag{4.3b}$$
$$\|\nabla f(x)\| \le 2\|\nabla f(x^*)\| \quad \text{for all } x \text{ with } \|x - x^*\| \le \hat{r}_2, \tag{4.3c}$$

where r_0 is defined in Theorem 2.9 and r_1 is defined in Lemma A.2. We now define a constant M_7 large enough that the following are true:

$$M_7 \ge 2M_1, \tag{4.4a}$$
$$\hat{\sigma} M_7^2 \ge 16 M_1 \|\nabla f(x^*)\|, \tag{4.4b}$$
$$\hat{\sigma} M_7^2 \ge 32 M_6 \|\nabla f(x^*)\|, \tag{4.4c}$$

where M_1 is defined in Lemma A.2 and $\hat{\sigma}$ is defined in Theorem 2.9. We further define $\bar{t}_4$ small enough that the following conditions hold:

$$\bar{t}_4 \le 1, \tag{4.5a}$$
$$M_6 \bar{t}_4^{1/2} \le \hat{r}_2/2, \tag{4.5b}$$
$$M_7 \bar{t}_4^{1/4} < \hat{r}_2/2, \tag{4.5c}$$
$$M_1 \bar{t}_4^{1/2} \le \hat{r}_2/2. \tag{4.5d}$$

From (4.2) and (4.5b) and (4.3c), we have

$$f(z(t)) \le f(x^*) + 2M_6 t^{1/2} \|\nabla f(x^*)\|, \quad \text{for all } t \in (0, \bar{t}_4]. \tag{4.6}$$

For a given $t \le \bar{t}_4$, we define the radius of the inner ball to be $M_7 t^{1/4}$ and of the outer ball to be $\hat{r}_2/2$. (Because of (4.5c), the inner ball is truly contained in the outer ball.) Now let x be any point in the annulus between the two balls that is feasible for RegEq(t). Since $\|x - x^*\| \le \hat{r}_2/2 < r_1$, we have from Lemma A.2 that there is a z feasible for (1.1) such that

$$\|z - x\| \le M_1 t^{1/2}. \tag{4.7}$$

Since from (4.5d) we have

$$\|z - x^*\| \le \|z - x\| + \|x - x^*\| \le M_1 t^{1/2} + \hat{r}_2/2 \le \hat{r}_2/2 + \hat{r}_2/2 = \hat{r}_2,$$

we have using (4.3c) again that

$$|f(z) - f(x)| \le 2\|\nabla f(x^*)\| \, \|z - x\| \le 2\|\nabla f(x^*)\| M_1 t^{1/2}. \tag{4.8}$$

Moreover, we have from (4.7) and the definition of x that

$$\|z - x^*\| \ge \|x - x^*\| - \|z - x\| \ge M_7 t^{1/4} - M_1 t^{1/2} > 0,$$

where the final inequality follows from (4.5a) and (4.4a). Hence, from Theorem 2.9 and (4.8), we have

$$\begin{aligned} f(x) - f(x^*) &\ge f(z) - f(x^*) - |f(z) - f(x)| \\ &\ge \hat{\sigma}\|z - x^*\|^2 - 2\|\nabla f(x^*)\| M_1 t^{1/2} \\ &\ge \hat{\sigma}[M_7 t^{1/4} - M_1 t^{1/2}]^2 - 2\|\nabla f(x^*)\| M_1 t^{1/2}. \end{aligned} \tag{4.9}$$

Because of (4.5a) and (4.4a), we have $M_1 t^{1/2} \le (1/2) M_7 t^{1/4}$, so

$$\hat{\sigma}[M_7 t^{1/4} - M_1 t^{1/2}]^2 \ge (1/4)\hat{\sigma} M_7^2 t^{1/2}.$$

By substituting into (4.9) and using (4.4b), we have

$$f(x) - f(x^*) \ge (1/4)\hat{\sigma} M_7^2 t^{1/2} - 2\|\nabla f(x^*)\| M_1 t^{1/2} \ge (1/8)\hat{\sigma} M_7^2 t^{1/2}. \tag{4.10}$$

By comparing with (4.6), and using (4.4c), we have

$$\begin{aligned} &f(x) \ge f(x^*) + (1/8)\hat{\sigma} M_7^2 t^{1/2} \\ &\ge f(z(t)) - 2M_6 t^{1/2}\|\nabla f(x^*)\| + (1/8)\hat{\sigma} M_7^2 t^{1/2} \ge f(z(t)) + (1/16)\hat{\sigma} M_7^2 t^{1/2}, \end{aligned}$$

thereby confirming that any feasible point for RegEq(t) in the space between the two balls has a higher function value than the point $z(t)$ defined by (4.1), which is feasible for RegEq(t) and which lies inside the inner ball. This observation establishes the result.

∎

The stronger $O(t^{1/2})$ estimate of Theorem 3.2 cannot apply, at least not under the assumptions of Theorem 4.1.

Example 2. *The simple MPEC*

$$\min \; x_1 + \tfrac{1}{2}x_2^2 \quad \text{subject to} \quad 0 \le x_1 \perp x_2 \ge 0$$

has a strongly stationary point $x^* = (0,0)$ at which MPEC-LICQ and RNLP-SSOSC (hence MPEC-SOSC) hold, with MPEC-multipliers $\tau^* = 1$ and $\nu^* = 0$. RegEq(t) is

$$\min \; x_1 + \tfrac{1}{2}x_2^2 \quad \text{subject to} \quad x_1 x_2 = t, \; x_1, x_2 > 0.$$

It can be easily shown, starting from KKT conditions, that RegEq(t) has a unique solution $x(t) = (t^{2/3}, t^{1/3})$, so that $\|x^* - x(t)\| = O(t^{1/3}) \neq O(t^{1/2})$.

We demonstrate this fact more generally for the case in which $I_G \cap I_H \neq \emptyset$ and exactly one of τ_i^* and ν_i^* is nonzero for each $i \in I_G \cap I_H$.

Logic like that of Scholtes [28, Theorem 3.1] can be used to show that first-order sufficient conditions hold at the solution of RegEq(t), so we have

$$\begin{aligned}\nabla f(x(t)) = \sum_{i \in I_g} \lambda_i(t) \nabla g_i(x(t)) + \sum_{i=1}^{q} \mu_i(t) \nabla h_i(x(t)) \\ - \sum_{i=1}^{m} \delta_i(t) \left[G_i(x(t)) \nabla H_i(x(t)) + H_i(x(t)) \nabla G_i(x(t)) \right],\end{aligned} \tag{4.11}$$

for t sufficiently small. Consider now the limit of this expression as $t \downarrow 0$ (and therefore $x(t) \to x^*$). Suppose that for some $i \in I_H \setminus I_G$, we have $\delta_i(t) H_i(x(t)) \not\to 0$. Then since $i \in I_H$ and hence $H_i(x^*) = 0$, it follows that $|\delta_i(t)| \to \infty$, and since $G_i(x^*) > 0$, we have $|\delta_i(t) G_i(x(t))| \to \infty$ and that eventually $G_i(x(t)) \gg H_i(x(t))$. Likewise, if there is $i \in I_G \setminus I_H$ for which $\delta_i(t) G_i(x(t)) \not\to 0$, we have that $|\delta_i(t) H_i(x(t))| \to \infty$ and eventually $H_i(x(t)) \gg G_i(x(t))$. If there exist indices i that fall into one of these two categories, we therefore have that

$$\max \left(\max_{i \in I_g} \lambda_i(t),\ \max_{i=1,2,\dots,q} |\mu_i(t)|,\ \max_{i \in I_G} |\delta_i(t) H_i(x(t))|,\ \max_{i \in I_H} |\delta_i(t) G_i(x(t))| \right) \to \infty,$$

so by dividing both sides of (4.11) by this largest absolute multiplier, and taking limits as $t \downarrow 0$, we have that there is a vector $(\hat\lambda, \hat\mu, \hat\tau, \hat\nu)$ with $\|(\hat\lambda, \hat\mu, \hat\tau, \hat\nu)\|_\infty = 1$ such that

$$0 = \sum_{i \in I_g} \hat\lambda_i \nabla g_i(x^*) + \sum_{i=1}^{q} \hat\mu_i \nabla h_i(x^*) - \sum_{i \in I_G} \hat\tau_i \nabla G_i(x^*) - \sum_{i \in I_H} \hat\nu_i \nabla H_i(x^*).$$

However, the MPEC-LICQ condition now implies that $(\hat\lambda, \hat\mu, \hat\tau, \hat\nu) = 0$, a contradiction. We conclude therefore that $\delta_i(t) H_i(x(t)) \to 0$ for all $i \in I_H \setminus I_G$ and $\delta_i(t) G_i(x(t)) \to 0$ for all $i \in I_G \setminus I_H$. A comparison of (4.11) with (2.5a) then yields that

$$\lambda_i(t) \to \lambda_i^*, \quad \text{for all } i \in I_g, \tag{4.12a}$$

$$\mu(t) \to \mu^*, \tag{4.12b}$$

$$-\delta_i(t) H_i(x(t)) \to \tau_i^*, \quad \text{for all } i \in I_G, \tag{4.12c}$$

$$-\delta_i(t) G_i(x(t)) \to \nu_i^*, \quad \text{for all } i \in I_H. \tag{4.12d}$$

These limits suggest that the multipliers $(\lambda(t), \mu(t), \delta(t))$ are bounded if $I_G \cap I_H = \emptyset$; that is, if LSC holds. Otherwise, we see from (4.12c) and (4.12d) that the multiplier $\delta_i(t) \to \infty$ for all $i \in I_G \cap I_H$ such that either $\tau_i^* > 0$ or $\nu_i^* > 0$. Moreover, if for any index $i \in I_G \cap I_H$ we have that exactly one of $\tau_i^* > 0$ and $\nu_i^* > 0$ is zero, then the solution $x(t)$ cannot satisfy the $O(t^{1/2})$ error estimate. For contradiction, let i be such an index, and assume without loss of generality that $\tau_i^* = 0$ and $\nu_i^* > 0$. By multiplying (4.12c) and (4.12d) and using $G_i(x(t)) H_i(x(t)) = t$, we obtain that

$$\delta_i(t)^2 G_i(x(t)) H_i(x(t)) = \delta_i(t)^2 t \to \tau_i^* \nu_i^* = 0,$$

so that $|\delta_i(t)| = o(t^{-1/2})$. Then from (4.12d) we have that $G_i(x(t)) \to \nu_i^* / |\delta_i(t)| = \nu_i^* / o(t^{-1/2})$, which is incompatible with $G_i(x(t)) = O(\|x(t) - x^*\|) = O(t^{1/2})$.

5. Properties of Solutions of PF(ρ). In this section, we prove results concerning exactness of the penalty formulation (1.5). We show first that the penalty function formulation (1.5) is exact, in the sense that a strongly stationary point of (1.1) is a local minimizer of (1.5) under certain assumptions. Note in particular that no strict complementarity condition is required.

THEOREM 5.1. *Suppose that x^* is a strongly stationary point of (1.1). Then for all ρ sufficiently large, the following claims are true.*

(a) x^* is a stationary point of PF(ρ).

(b) If MPEC-LICQ holds at x^*, then LICQ holds for PF(ρ) at x^*. If MPEC-MFCQ holds at x^*, then MFCQ holds for PF(ρ) at x^*.

(c) If MPEC-SOSC (Definition 2.7) is satisfied at x^*, then there is $\hat{\rho} > 0$ such that SOSC is satisfied for PF(ρ) at x^* for all $\rho \geq \hat{\rho}$.

Proof. We start by proving (a). The KKT conditions for (1.5) will be satisfied at x^* if we can find Lagrange multipliers $\bar{\lambda}$, $\bar{\mu}$, $\bar{\tau}$, and $\bar{\nu}$ such that the following conditions hold:

$$\begin{aligned}
&(5.1a) \quad 0 = \nabla f(x^*) - \sum_{i \in I_g} \bar{\lambda}_i \nabla g_i(x^*) - \sum_{i=1}^{q} \bar{\mu}_i \nabla h_i(x^*) \\
&\qquad\qquad - \sum_{i \in I_G} (\bar{\tau}_i - \rho H_i(x^*)) \nabla G_i(x^*) - \sum_{i \in I_H} (\bar{\nu}_i - \rho G_i(x^*)) \nabla H_i(x^*), \\
&(5.1b) \quad 0 \leq \bar{\mu}_i, \;\; i \in I_g, \\
&(5.1c) \quad 0 \leq \bar{\tau}_i, \;\; i \in I_G, \\
&(5.1d) \quad 0 \leq \bar{\nu}_i, \;\; i \in I_H,
\end{aligned}$$

where I_g, I_G, and I_H are the index sets defined in (2.1). Given $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) and (2.7), we set

$$\begin{aligned}
&(5.2a) \qquad \bar{\lambda}_i = \lambda_i^*, \quad i \in I_g, \\
&(5.2b) \qquad \bar{\mu}_i = \mu_i^*, \quad i = 1, 2, \ldots, q, \\
&(5.2c) \qquad \bar{\tau}_i = \tau_i^* + \rho H_i(x^*), \quad i \in I_G, \\
&(5.2d) \qquad \bar{\nu}_i = \nu_i^* + \rho G_i(x^*), \quad i \in I_H,
\end{aligned}$$

and choose ρ to satisfy

$$(5.3) \qquad \rho \geq \hat{\rho}_1 \stackrel{\text{def}}{=} 1 + \max\left(0, \max_{i \in I_G \setminus I_H : \tau_i^* < 0} \frac{-\tau_i^*}{H_i(x^*)}, \max_{i \in I_H \setminus I_G : \nu_i^* < 0} \frac{-\nu_i^*}{G_i(x^*)}\right),$$

it is easy to verify by comparison with (2.5) that the conditions (5.1) are satisfied.

For (b), we note first that the LICQ condition for (1.5) follows immediately from MPEC-LICQ (Definition 2.3). Since we established in the proof of Theorem 3.2 that MPEC-MFCQ implies existence of a vector $\tilde{d}$ satisfying (3.2), then MPEC-MFCQ implies MFCQ for PF(ρ) at x^* as well, since MFCQ for PF(ρ) consists of all conditions in (3.2) except the final one.

We now prove (c). First, it is easy to show that for ρ and the multipliers $(\bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu})$ defined as in (5.2), and denoting the Lagrangian for (1.5) by $\hat{L}(x, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}; \rho)$, we have

$$\begin{aligned}
&\nabla_{xx}^2 \hat{L}(x^*, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}; \rho) = \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*) \\
(5.4) \qquad &+ \rho \sum_{i=1}^{m} \left(\nabla G_i(x^*) \nabla H_i(x^*)^T + \nabla H_i(x^*) \nabla G_i(x^*)^T\right).
\end{aligned}$$

In a similar fashion to the proof of Theorem 3.2, we can show that the critical direction set for PF(ρ) at x^* is again $\bar{S}$, the critical direction set for the RNLP (2.4). We can now apply an almost identical argument to the one in the proof of Theorem 3.2 to deduce that there is a threshold $\hat{\rho}$ such that for all $\rho \geq \hat{\rho}$ we have

$$s^T \nabla^2_{xx} \hat{L}(x^*, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}; \rho)s \geq \sigma/2, \qquad \text{for all } s \in \bar{S},$$

where σ is the quantity from Definition 2.7. Hence, the SOSC for PF(ρ) is satisfied, as claimed. ∎

Part (c) of the theorem above is quite similar to the recent result [1, Corollary 3.2], the essential difference being that we show that MPEC-SOSC implies SOSC for PF(ρ), while [1] proves a similar implication for the quadratic increase condition.

We now prove a partial converse of Theorem 5.1, showing that local solutions of (1.5) for which $G(x)^T H(x) = 0$ are local solutions of (1.1).

THEOREM 5.2. *Suppose that x^* is a stationary point for PF(ρ) (1.5) and that $G(x^*)^T H(x^*) = 0$. Then x^* is strongly stationary for (1.1).*

If in addition x^* satisfies LICQ for PF(ρ) at x^*, then (1.1) satisfies MPEC-LICQ.

Suppose in addition to stationarity that SOSC is satisfied for PF(ρ) at x^*. Then MPEC-SOSC is satisfied at x^*, so that x^* is a strict local minimizer of (1.1).

Proof. Stationarity of x^* for PF(ρ) means that the conditions (5.1) are satisfied for some $\bar{\lambda}$, $\bar{\mu}$, $\bar{\tau}$, and $\bar{\nu}$, for I_g, I_G, and I_H defined by (2.1). Since $G(x^*)^T H(x^*) = 0$, we have $I_G \cup I_H = \{1, 2, \ldots, m\}$. If we define λ^*, μ^*, τ^*, and ν^* by the relations (5.2), we see that (2.5) are satisfied, so that x^* is strongly stationary for (1.1), as claimed.

The LICQ condition for PF(ρ) is exactly the condition (2.8), so that MPEC-LICQ is satisfied at this point also.

For the final statement of the theorem, we need to show that there is some $\sigma > 0$ such that for any $s \in S^*$, there are multipliers $(\lambda^*, \mu^*, \tau^*, \nu^*)$ satisfying (2.5) such that (2.12) holds. Since $S^* \subset \bar{S}$, and since $\bar{S}$ is the critical direction set for PF(ρ), we have from the SOSC for x^* in (1.5) that there exist multipliers $(\bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu})$ for (1.5) such that

$$s^T \nabla^2_{xx} \bar{L}(x^*, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho)s \geq \hat{\sigma},$$

where $\hat{\sigma} > 0$ is a constant. Defining $(\lambda^*, \mu^*, \tau^*, \nu^*)$ from $(\bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu})$ using (5.2), and using (5.4), we have that

$$s^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*, \tau^*, \nu^*)s = s^T \nabla^2_{xx} \hat{L}(x^*, \bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu}, \rho)s \geq \hat{\sigma},$$

since a close examination of the conditions defining S^* show that the final term in (5.4) contributes nothing to the product. Hence, we have shown that MPEC-SOSC conditions are satisfied at x^* for $\sigma = \hat{\sigma}$. ∎

The condition $G(x)^T H(x) = 0$ appears to be essential. As observed by Hu and Ralph [12], there is nothing to stop an approach based on (1.5) to get stuck at a minimizer of the "limiting" problem

$$\begin{array}{ll} \min_x & G(x)^T H(x) \quad \text{subject to} \\ & g(x) \leq 0, \quad h(x) = 0, \quad G(x) \geq 0, \quad H(x) \geq 0, \end{array} \tag{5.5}$$

Under the PSC assumption, we are able to prove an extension of the results of Theorems 5.1 and 5.2, in which all stationary points for PF(ρ) in the neighborhood of a strongly stationary point x^* for (1.1) are also strongly stationary for (1.1).

Proposition 5.3. *Suppose that x^* is a strongly stationary point of (1.1) and that MPEC-LICQ and PSC hold at x^*. Then there is a neighborhood U of x^* and a scalar $\rho^* > 0$ such that for all $\rho \geq \rho^*$, every stationary point for PF(ρ) in U is a strongly stationary point for (1.1).*

Proof. Given the point x^* that is strongly stationary for (1.1), we note that, because of MPEC-LICQ, the multiplier vector $(\lambda^*, \mu^*, \tau^*, \nu^*)$ is uniquely defined. We now define the neighborhood U and the threshold value ρ^* for the penalty parameter ρ. We first define the constant $\epsilon > 0$ as follows:

$$\epsilon = (1/2) \min \left(\min_{i \notin I_H} H_i(x^*), \ \min_{i \notin I_G} G_i(x^*), \ \min_{i \,|\, \tau_i^* > 0} \tau_i^*, \ \min_{i \,|\, \nu_i^* > 0} \nu_i^* \right).$$

Now choose U small enough that

$$x \in U \quad \Rightarrow \quad H_i(x) \geq \epsilon \ \text{ for all } i \notin I_H \ \text{ and } \ G_i(x) \geq \epsilon \ \text{ for all } i \notin I_G.$$

By using the MPEC-LICQ property, we can shrink U further if necessary such that if x is any point in U for which there exist multipliers $(\lambda, \mu, \tau, \nu)$ satisfying $\nabla_x L(x, \lambda, \mu, \tau, \nu) = 0$, then we have

$$\|(\lambda, \mu, \tau, \nu) - (\lambda^*, \mu^*, \tau^*, \nu^*)\|_\infty \leq \epsilon. \tag{5.6}$$

Choose ρ^* large enough that

$$\rho^* \epsilon + \tau_i^* - \epsilon > 0 \ \text{ and } \ \rho^* \epsilon + \nu_i^* - \epsilon > 0 \ \text{ for all } i = 1, 2, \ldots, m.$$

Let $\bar{x} \in U$ be stationary for PF(ρ) for some $\rho \geq \rho^*$. We aim to show that $\bar{x}$ is feasible for MPEC, in particular, $G(\bar{x})^T H(\bar{x}) = 0$. Once we have established this fact, the proof follows immediately from Theorem 5.2.

By choice of U, $I_G(\bar{x}) \subset I_G$ and $I_H(\bar{x}) \subset I_H$, where $I_G(\bar{x})$ contains the indices for which $G_i(\bar{x}) = 0$, and similarly for $I_H(\bar{x})$. To show MPEC feasibility of $\bar{x}$, we first select some index $i \in I_G \setminus I_H$. As shown in (5.1a), stationarity of $\bar{x}$ for PF(ρ) implies that there are multipliers $(\bar{\lambda}, \bar{\mu}, \bar{\tau}, \bar{\nu})$ such that $\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{\tau} - \rho H(\bar{x}), \bar{\nu} - \rho G(\bar{x})) = 0$. Hence, from (5.6),

$$\|\bar{\tau} - \rho H(\bar{x}) - \tau^*\|_\infty \leq \epsilon, \qquad \|\bar{\nu} - \rho G(\bar{x}) - \nu^*\|_\infty \leq \epsilon.$$

Hence, for the given $i \in I_G \setminus I_H$, we have that $\bar{\tau}_i \geq \rho H_i(\bar{x}) + \tau_i^* - \epsilon > 0$. From complementarity of $\bar{\tau}_i$ with $G_i(\bar{x})$ in the KKT conditions for $\bar{x}$ in PF(ρ), we deduce that $G_i(\bar{x}) = 0$. Similar logic yields that $H_i(\bar{x}) = 0$ for all $i \in I_H \setminus I_G$.

Next, select $i \in I_G(x^*) \cap I_H(x^*)$. Because of the PSC assumption, either τ_i^* or ν_i^* is positive. Suppose $\tau_i^* > 0$. From (5.6) we have $\bar{\tau}_i - \rho H_i(\bar{x}) \geq \tau_i^* - \epsilon$, where the quantity at right is positive by definition of ϵ. Since $\rho H_i(\bar{x}) \geq 0$, we see that $\bar{\tau}_i > 0$. Hence, complementarity of $\bar{\tau}_i$ with $G_i(\bar{x})$ yields that $G_i(\bar{x}) = 0$. Likewise, if $\nu_i^* > 0$ then $H_i(\bar{x}) = 0$.

We have shown that for all i, at least one of $G_i(\bar{x})$ and $H_i(\bar{x})$ is zero, so that $\bar{x}$ is not only stationary for PF(ρ) but also feasible for the MPEC (1.1). The result now follows from Theorem 5.2. □

Let x^* be strongly stationary point that satisfies PSC and U and ρ^* be given by Proposition 5.3. An immediate corollary of this result is that U is a "neighborhood of finite termination" for the penalty method: If $\rho_k \to \infty$ and, for each k, x^k is a stationary point of PF(ρ_k), then for any iterate with $\rho^k \geq \rho^*$ and $x^k \in U$ we have strong stationarity of x^k, hence termination of the penalty approach.

6. Conclusions. We have examined several properties of the solutions to the regularized formulation Reg(t) (1.2) to the MPEC (1.1)—distance between solutions of (1.2) and (1.1), boundedness of Lagrange multipliers, local uniqueness, and smoothness of the solution mapping—under various assumptions on (1.1) at a local solution x^*. We have obtained similar results for the alternative regularized formulations (1.3) and (1.4). We have also looked at the penalty formulation PF(ρ) (1.5), deriving relationships between solutions of this problem and solutions of the original MPEC.

Further work is needed on making use of the observations above in algorithms based on Reg(t). It may be possible to devise a method with an overall superlinear convergence rate (and desirable global convergence properties) by applying an SQP-like method to approximately solve Reg(t) for a decreasing sequence of t values. Near x^*, it may be possible to decrease t at a "superlinear" rate while taking only one SQP step for each t. For the penalty formulation, an SQP strategy in conjunction with a technique to find an appropriately large value of ρ is needed. For both regularization and penalization techniques, we are also interested in algorithms that converge when LICQ conditions are replaced by corresponding MFCQ conditions.

Acknowledgments. We thank Mihai Anitescu for several discussions about these results in the summer of 2003. We are most grateful to an anonymous referee for pointing out that forthcoming paper [14] contains a proof that MPEC-SOSC implies that SOSC holds for Reg(0), a fact that is used in the proof of Theorem 3.2. (We had proved a similar property in connection with Theorem 5.1 but did not recognize its relevance to the earlier results.) The referee's observation allowed us to replace our earlier proofs of Theorems 3.2 and 3.3 with proofs that relied on results of Bonnans and Shapiro [3]. The earlier proofs, which are more self-contained but considerably longer, can be found in the technical report [24].

The second author acknowledges the support of the National Science Foundation under Grants #ACI-0082065, #ACI-0113051, #ATM-0086559, and #EIA-0127857.

Appendix A. Perturbation Results for Regularized Formulations.

We start by stating a special case of Robinson [25, Theorem 1].

LEMMA A.1. *Suppose that x^* is a feasible point for the system*

$$c(x) = 0, \quad d(x) \le 0, \tag{A.1}$$

where $c : \mathbb{R}^n \to \mathbb{R}^{n_c}$ and $d : \mathbb{R}^n \to \mathbb{R}^{n_d}$ are continuously differentiable at x^*. Suppose that MFCQ is satisfied at x^*, that is, the vectors $\nabla c_i(x^*)$, $i = 1, 2, \dots, n_c$ are linearly independent and there exists a vector $v \neq 0$ such that

$$\begin{aligned} &\nabla c_i(x^*)^T v = 0, \ \text{ for all } i = 1, 2, \dots n_c, \\ &\nabla d_i(x^*)^T v < 0, \ \text{ for all } i = 1, 2, \dots, n_d \text{ such that } d_i(x^*) = 0. \end{aligned}$$

Then there exists a radius $r > 0$ and a constant $M > 0$ such that for all x with $\|x - x^*\| < r$, there is a vector $z \in \mathbb{R}^n$ satisfying (A.1) such that

$$\|z - x\| \le M \left\| \begin{pmatrix} c(x) \\ \max(d(x), 0) \end{pmatrix} \right\|.$$

Next, we show that in a neighborhood of x^*, any point x that is feasible for (1.2) is at most a distance $O(t^{1/2})$ from a point that is feasible for (1.1).

LEMMA A.2. *Let x^* be a solution of (1.1) at which strong stationarity and MPEC-LICQ are satisfied. Then there exist a radius $r_1 > 0$ and a constant $M_1 > 0$*

such that the following property holds. When x is a feasible point for $\text{Reg}(t)$ *defined in (1.2) for* $t \in [0,1]$, *and in addition* $\|x - x^*\| \le r_1$, *then there is a point z feasible for (1.1) such that* $\|z - x\| \le M_1 t^{1/2}$. *If LSC holds, this estimate can be improved to* $\|z - x\| \le M_1 t$.

Proof. Consider any subset $I_P \subset I_G \cap I_H$, and define the following system of inequalities

$$\begin{aligned}
G_i(z) &= 0, \quad i \in I_G \setminus I_H,\\
G_i(z) &= 0, \quad i \in I_P,\\
G_i(z) &\ge 0, \quad i \in I_H \setminus I_P,\\
H_i(z) &= 0, \quad i \in I_H \setminus I_G,\\
H_i(z) &= 0, \quad i \in I_P^c,\\
H_i(z) &\ge 0, \quad i \in I_G \setminus I_P^c,\\
g_i(z) &\ge 0, \quad i = 1, 2, \ldots, p,\\
h_i(z) &= 0, \quad i = 1, 2, \ldots, q,
\end{aligned} \tag{A.2}$$

where I_P^c denotes $(I_G \cap I_H) \setminus I_P$. Note first that any z satisfying (A.2) is certainly feasible for (1.1). Note too that x^* is feasible for this system, for all choices of I_P, and that the active constraint gradients at x^* are simply the vectors in (2.8), which are linearly independent by assumption. Hence, the MFCQ condition of Lemma A.1 is satisfied by (A.2) at x^*.

Next, define $\bar{r}$ such that the following properties hold for all x with $\|x - x^*\| \le \bar{r}$:

$$G_i(x) \ge \tfrac{1}{2} G_i(x^*), \quad i \in I_H \setminus I_G, \tag{A.3a}$$

$$H_i(x) \ge \tfrac{1}{2} H_i(x^*), \quad i \in I_G \setminus I_H, \tag{A.3b}$$

$$g_i(x) \ge \tfrac{1}{2} g_i(x^*), \quad i \notin I_g. \tag{A.3c}$$

We now apply Lemma A.1 to (A.2) at x^*. By this result, we can choose $\bar{r}(I_P) \in (0, \bar{r}]$ and $\bar{M}(I_P) > 0$ such that for any x with $\|x - x^*\| \le \bar{r}(I_P)$, there is a solution z of (A.2) such that the following condition is satisfied:

$$\begin{aligned}
\|z - x\| \le \bar{M}(I_P) \Bigg\{ & \sum_{i \in I_P \cup (I_G \setminus I_H)} |G_i(x)| + \sum_{i \in I_P^c} \max(-G_i(x), 0) \\
& + \sum_{i \in I_P^c \cup (I_H \setminus I_G)} |H_i(x)| + \sum_{i \in I_P} \max(-H_i(x), 0) \\
& + \sum_{i \in I_g} \max(-g_i(x), 0) + \sum_{i=1}^{q} |h_i(x)| \Bigg\}.
\end{aligned} \tag{A.4}$$

(We have used the equivalence of the 1 and 2$-$norms and have omitted certain terms from the right-hand side of the bound because of the inactivities (A.3).) Let us now define r_1 and $\hat{M}_1$ as follows:

$$r_1 \stackrel{\text{def}}{=} \min_{I_P \subset I_G \cap I_H} \bar{r}(I_P), \qquad \hat{M}_1 \stackrel{\text{def}}{=} \max_{I_P \subset I_G \cap I_H} \bar{M}(I_P).$$

Consider any x feasible for $\text{Reg}(t)$ (1.2) that also satisfies $\|x - x^*\| \le r_1$. For this x, we define I_P as follows:

$$I_P = \{ i \in I_G \cap I_H \mid G_i(x) \le H_i(x) \}.$$

For this x, we have from the constraints in (1.2) that

$$G_i(x) \geq 0,\ H_i(x) \geq 0,\ G_i(x)H_i(x) \leq t \ \ \Rightarrow \ \ 0 \leq G_i(x) \leq t^{1/2}, \quad \text{for all } i \in I_P.$$

Similarly, we have $0 \leq H_i(x) \leq t^{1/2}$ for all $i \in I_P^c$. We have from $r_1 \leq \bar{r}$ and (A.3) that

$$\begin{aligned} 0 \leq G_i(x) \leq t/H_i(x) \leq 2t/H_i(x^*), \quad i \in I_G \setminus I_H, \\ 0 \leq H_i(x) \leq t/G_i(x) \leq 2t/G_i(x^*), \quad i \in I_H \setminus I_G. \end{aligned}$$

We also have for all x feasible in (1.2) that $G(x) \geq 0$, $H(x) \geq 0$, $g(x) \geq 0$ and $h(x) = 0$. Hence, by applying (A.4) for this choice of I_P, we find that there is a z satisfying (A.2) (and hence feasible for (1.1)) such that

$$\begin{aligned} \|z - x\| &\leq \hat{M}_1 \left\{ \sum_{i \notin I_H} 2t/H_i(x^*) + \sum_{i \in I_P} t^{1/2} + \sum_{i \notin I_G} 2t/G_i(x^*) \sum_{i \in I_P^c} t^{1/2} \right\} \\ &= 2\hat{M}_1 t \left\{ \sum_{i \notin I_H} 1/H_i(x^*) + \sum_{i \notin I_G} 1/G_i(x^*) \right\} + \hat{M}_1 |I_G \cap I_H| t^{1/2} \\ &\leq M_1 t^{1/2} \end{aligned}$$

for all $t \in [0, 1]$ and for an obvious definition of M_1. For the final statement of the theorem, we have $I_G \cap I_H = \emptyset$, so that the final bound can be strengthened to $M_1 t$ (for a different value of M_1). $\square$

REFERENCES

[1] M. Anitescu, On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints, Preprint ANL/MCS-P864-1200, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, USA, December 2000. Revised December, 2003.

[2] H. Y. Benson, D. F. Shanno, and R. J. Vanderbei, Interior-point methods for nonconvex nonlinear programming: Complementarity constraints, ORFE 02-02, Department of Operations Research and Financial Engineering, Princeton University, July 2002.

[3] J. F. Bonnans and A. Shapiro (2000). *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer.

[4] Y. Chen and F. Florian (1995). The nonlinear bilevel programming problem: formulations, regularity and optimality conditions, *Optimization*, **32**, 193–209.

[5] A. de Miguel, M. Friedlander, F. Nogales, and S. Scholtes, An interior-point method for MPECs, LBS working paper, London Business School, 2003.

[6] F. Facchinei, H. Jiang, and L. Qi (1999). A smoothing method for mathematical programs with equilibrium constraints, *Mathematical Programming*, **85**, 81–106.

[7] 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, May 2002.

[8] M. Fukushima and J. Pang, Convergence of a smoothing continuation method for mathematical programs with complementarity constraints, in M. Théra and R. Tichatschke (Eds.). *Ill-posed Variational Problems and Regularization Techniques* vol. 447 of Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin/Heidelberg, 99–110, 1999.

[9] M. Fukushima and P. Tseng (2002). An implementable active-set algorithm for computing a b-stationary point of a mathematical program with linear complementarity constraints, *SIAM Journal on Optimization* **12**, 724–739.

[10] J. Gauvin (1977). A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming, *Mathematical Programming*, **12**, 136–138.

[11] X. Hu, Mathematical programs with equilibrium constraints and game theory models in electricity markets, PhD thesis, Department of Mathematics and Statistics, The University of Melbourne, 2003.

[12] X. Hu and D. Ralph, Convergence of a penalty method for mathematical programming with complementarity constraints, Manuscript, Judge Institute of Management, University of Cambridge, revised 2003. To appear in *Journal of Optimization Theory and Applications*.

[13] X. X. Huang, X. Q. Yang, and D. L. Zhu, A sequential smooth penalization approach to mathematical programs with complementarity constraints, Manuscript, Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China, 2001.

[14] A. F. Izmailov, Optimization problems with complementarity constraints: Regularity, optimality conditions, and sensitivity, 2004. To appear in *Computational Mathematics and Mathematical Physics*.

[15] H. Jiang and D. Ralph (1999). QPECgen, a MATLAB generator for mathematical programs with quadratic objectives and affine variational inequality constaints, *Computational Optimization and Applications*, **13**, 25–29.

[16] H. Jiang and D. Ralph (2002). Extension of quasi-newton methods to mathematical programs with complementarity constraints, *Computational Optimization and Applications*, **25**, 123–150.

[17] S. Leyffer, MacMPEC AMPL collection of mathematical programs with equilibrium constraints, 2000.

[18] G. Lin and M. Fukushima, Hybrid algorithms with active set identification for mathematical programs with complementarity constraints, Preprint, Department of Applied Mathematics and Physics, Kyoto University, November 2002.

[19] X. Liu, G. Perakis, and J. Sun, A robust SQP methods for mathematical programs with complementarity constraints, Research Report, Singapore-MIT alliance, National University of Singapore, 2003.

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

[21] Z. Luo, J.-S. Pang, and D. Ralph, Piecewise sequential quadratic programming for mathematical programs with nonlinear complementarity constraints, In: A. Migdalas, P. Pardalos, and P. Värbrand, (Eds.), *Multilevel Optimization: Algorithms, Complexity and Applications*, Kluwer Academic Publishers, pp. 209–229, 1998.

[22] H. Maurer and J. Zowe (1979). First- and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, *Mathematical Programming*, **16**, 98–110.

[23] A. U. Raghunathan and L. T. Biegler, Interior point methods for mathematical programs with complementarity constraints, CAPD technical report, Department of Chemical Engineering, Carnegie Mellon University, June 2003.

[24] D. Ralph and S. J. Wright, Some properties of regularization schemes for MPECs, Technical Report, Computer Sciences Department, University of Wisconsin-Madison, December 2003.

[25] S. M. Robinson (1976). Stability theory for systems of inequalities, Part II: Differentiable nonlinear systems, *SIAM Journal on Numerical Analysis*, **13**, 497–513.

[26] S. M. Robinson (1982). Generalized equations and their solutions. Part II: Applications to nonlinear programming, *Mathematical Programming Study*, **19**, 200–221.

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

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

[29] J. Ye, D. L. Zhu, and Q. J. Zhu (1997). Exact penalization and necessary optimality conditions for generalized bilevel programming problems, *SIAM Journal on Optimization*, **7**, 481–507.

[30] J. Zhang and G. Liu (2001). A new extreme point algorithm and its application in psqp algorithms for solving mathematical programs with linear complementarity constraints, *Journal of Global Optimization*, **19**, 335–361.