Preprint ANL/MCS-P864-1200

ON USING THE ELASTIC MODE IN NONLINEAR PROGRAMMING APPROACHES TO MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

MIHAI ANITESCU*

Abstract. We investigate the possibility of solving mathematical programs with complementarity constraints (MPCCs) using algorithms and procedures of smooth nonlinear programming. Although MPCCs do not satisfy a constraint qualification, we establish sufficient conditions for their Lagrange multiplier set to be nonempty. MPCCs that have nonempty Lagrange multiplier sets and that satisfy the quadratic growth condition can be approached by the elastic mode with a bounded penalty parameter. In this context, the elastic mode transforms MPCC into a nonlinear program with additional variables that has an isolated stationary point and local minimum at the solution of the original problem, which in turn makes it approachable by sequential quadratic programming algorithms. One such algorithm is shown to achieve local linear convergence once the problem is relaxed. Under stronger conditions, we also prove superlinear convergence to the solution of an MPCC using an adaptive elastic mode approach for a sequential quadratic programming algorithm recently analyzed in an MPCC context by Fletcher and al. [16]. Our assumptions are more general since we do not use a critical assumption from that reference. In addition, we show that the elastic parameter update rule will not interfere locally with the super linear convergence once the penalty parameter is appropriately chosen.

Key words. Nonlinear Programming, Elastic Mode, SQP, MPEC, MPCC, complementarity constraints.

AMS subject classifications. 90C30, 90C33, 90C55,49M37, 65K10

1. Introduction. Complementarity constraints can be used to model numerous economics or engineering applications [27, 33]. Solving optimization problems with complementarity constraints may prove difficult for classical nonlinear optimization, however, given that, at a solution x^*, such problems cannot satisfy a constraint qualification [27, Chapter 3]. As a result, algorithms based on the linearization of the feasible set, such as sequential quadratic programming (SQP) algorithms, may fail because feasibility of the linearization can no longer be guaranteed in a neighborhood of the solution [27].

Several methods have been recently proposed to accommodate such problems. For example, a nondifferentiable penalty term in the objective function can be used to replace the complementarity constraints [28], while maintaining the same solution set. Although the new problem may now satisfy the constraint qualification the nondifferentiability of the objective function is an obstacle to the efficient computation of an optimal point, at least from the practical perspective of the range of available software. Another method is the disjunctive nonlinear programming (disjunctive NLP) approach [27], though this may lead to a large number of subcases to account for the alternatives involving degenerate complementarity constraints. If all constraint functions, with the exception of the complementarity constraints, are linear, then efficient active set approaches can be defined, if the linear independence constraint

* Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Ave., Argonne, Illinois, 60439, (anitescu@mcs.anl.gov)

qualification holds [20]. Still other approaches have been defined for problems whose complementarity constraints originate in equilibrium conditions [27].

A nonsmooth approach has been proposed in [33] for MPCCs in which the underlying complementarity constraints originate in a variational inequality with strong regularity properties. A bundle trust-region algorithm is defined in which each element of the bundle is generated from the generalized gradient of the reduced objective function. The key step is to produce an element of the generalized gradient [33, Equations (7.24), (7.25)], which may be quite costly for general cases at points where there are a substantial number of degenerate complementarity constraints.

In this work we investigate the possibility of solving MPCCs by applying certain SQP algorithms to their nonlinear programming formulation. This endeavor is important because it allows one to extend the considerable body of analytical and computational expertise of smooth nonlinear programming to this new class of problems. The advantage of such an approach over disjunctive programming, for example, is that it considers simultaneously all the alternatives involving degenerate complementarity constraints. The disadvantage is that the description of the constraint set is considerably less well behaved.

Recognizing that the potential infeasibility of the subproblems with linearized constraints may prevent normal termination of SQP algorithms, we discuss their use in conjunction with the elastic mode [22]. The elastic mode is a standard technique of approaching infeasible subproblems by relaxing the constraints and introducing a differentiable penalty term in the objective function. To show that such an approach can accommodate a large class of MPCCs, we use the framework from [36] to determine sufficient conditions for MPCCs to have nonempty Lagrange multiplier sets.

As in [36], the first- and second-order optimality properties of an MPCC are compared with the similar properties of two nonlinear programs that involve no complementarity constraints and may thus satisfy a constraint qualification. Here, however, we consider the optimality properties of an MPCC formulated as a nonlinear program with differentiable data. In [36] the MPCC is equivalently described with the complementarity constraints replaced by an equality involving the nondifferentiable function $\min\{x_1, x_2\}$. The two formulations will ultimately have similar properties, but the smooth description is important in anticipation of the use of a standard nonlinear programming algorithm to solve MPCCs.

The elastic mode approach we present here is different from other nonlinear programming approaches for MPCC in the following important respect. Virtually all smooth nonlinear programming approaches currently described in the literature for finding a solution x^* of MPCC consist of transforming it into another nonlinear program depending on a parameter p, MPCC(p) and then finding the solution x^p of the modified problem [23, 27, 37]. The problem MPCC(p) will have enough constraint regularity for x^p to be found reasonably efficiently. The solution x^* is then obtained in the limit as $p \to 0$, and $x^p \neq x^*$ for any p. The program MPCC(0) is undefined, or does not satisfy a constraint qualification (if the parameter is a penalty parameter c, the same observation is valid by choosing $p = \frac{1}{c}$).

For the elastic mode, under conditions to be specified in the body of this work, MPCC is transformed into a problem MPCC(c) that satisfies a constraint qualification and has x^* as a local solution for all c sufficiently large but finite. So MPCC is transformed by a finite procedure in a nonlinear program with the same solution that satisfies a constraint qualification, which does not happen for the other approaches. To our knowledge, at the time of the initial issue of this technical report in December

2000, the developments presented here are the first systematic approach of this type that is valid for a generic instance of mathematical programs with complementarity constrains.

We note that SQP was applied before in connection with a smoothing method for MPCC [19]. However, that approach is different from the one used here, in that the smoothing parameter had to be taken to a limit point, at which the nonlinear programming constraint qualification did not hold. In the context described here, that approach is closer to the MPCC(p) approach presented in the preceding paragraphs, though its appeal is that for a given smoothing parameter, the resulting problem was not solved until convergence, as opposed to other MPCC(p) approaches.

The paper is structured as follows. In the remainder of Section 1 we review the relevant nonlinear programming concepts. In Section 2 we discuss sufficient conditions for MPCC to have a nonempty Lagrange multiplier set, in spite of not satisfying a constraint qualification at any point. This allows us to argue in Section 3 that the elastic mode applied to an instance of the MPCC class will retrieve a local solution of the problem for a finite value of the penalty parameter, a point which is supported by several numerical examples. In section 4 we prove that an adaptive elastic mode approach built around an algorithm recently analyzed in Fletcher and al. [16] in the MPCC context will result in super linear convergence near the solution of an MPCC under assumptions weaker than in [16]. Specifically, here we do not assume that the iterates are either feasible or satisfy the complementarity constraints for the unrelaxed problem. In addition, we show that the elastic parameter update rule will not affect locally the super linear convergence once the penalty parameter is appropriately chosen.

1.1. Optimality Conditions for General Nonlinear Programming. We review the optimality conditions for a general nonlinear program

$$\min_x \tilde{f}(x) \quad \text{subject to } \tilde{g}(x) \leq 0,\ \tilde{h}(x) = 0. \tag{1.1}$$

Here $\tilde{g} : \mathcal{R}^n \to \mathcal{R}^m$, $\tilde{h} : \mathcal{R}^n \to \mathcal{R}^r$. We assume that $\tilde{f}$, $\tilde{g}$, and $\tilde{h}$ are twice continuously differentiable.

We call x a stationary point of (1.1) if the Fritz-John condition holds: There exist multipliers $0 \neq \tilde{\lambda} = (\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_{m+r}) \in \mathcal{R}^{m+r+1}$, such that

$$\nabla_x \mathcal{L}(x, \tilde{\lambda}) = 0,\ \ \tilde{h}(x) = 0;\ \ \tilde{\lambda}_i \geq 0,\ \ \tilde{g}_i(x) \leq 0, \quad \text{for } i = 1, 2, \ldots, m;\ \ \sum_{i=1}^m \tilde{\lambda}_i \tilde{g}_i(x) = 0. \tag{1.2}$$

Here $\mathcal{L}$ is the Lagrangian function

$$\mathcal{L}(x, \tilde{\lambda}) = \tilde{\lambda}_0 \tilde{f}(x) + \sum_{i=1}^m \tilde{\lambda}_i \tilde{g}_i(x) + \sum_{j=1}^r \tilde{\lambda}_{m+j} \tilde{h}_j(x). \tag{1.3}$$

A local solution x^* of (1.1) is a stationary point [34]. We introduce the sets of generalized Lagrange multipliers

$$\Lambda^g(x) = \left\{ 0 \neq \tilde{\lambda} \in \mathcal{R}^{m+r+1} \mid \ \tilde{\lambda} \text{ satisfies (1.2) at } x \right\}, \tag{1.4}$$

$$\Lambda_1^g(x) = \left\{ \tilde{\lambda} \in \Lambda^g(x) \mid \tilde{\lambda}_0 = 1 \right\}. \tag{1.5}$$

The set of active inequality constraints at a stationary point x is

$$\tilde{\mathcal{A}}(x) = \{i \in \{1,2,\ldots,m\} \mid \tilde{g}_i(x) = 0\}\,. \tag{1.6}$$

The set of inactive inequality constraints at x is the complement of $\tilde{\mathcal{A}}(x)$:

$$\tilde{\mathcal{A}}^c(x) = \{1,2,\ldots,m\} - \tilde{\mathcal{A}}(x). \tag{1.7}$$

With this notation, the complementarity condition from (1.2), $\sum_{i=1}^m \tilde{\lambda}_i g_i(x) = 0$, becomes $\tilde{\lambda}_{\tilde{\mathcal{A}}^c(x)} = 0$.

If certain regularity conditions hold at a stationary point x (discussed below), there exist $\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2, \ldots, \tilde{\mu}_{m+r}) \in \mathcal{R}^{m+r}$ that satisfy the Karush-Kuhn-Tucker (KKT) conditions [3, 4, 14]:

$$\begin{array}{c} \nabla_x \tilde{f}(x) + \sum_{i=1}^m \tilde{\mu}_i \nabla_x \tilde{g}_i(x) + \sum_{j=1}^r \tilde{\mu}_{m+j} \nabla_x \tilde{h}_j(x) = 0, \;\; \tilde{h}(x) = 0; \\ \tilde{\mu}_i \geq 0, \;\; \tilde{g}_i(x) \leq 0, \;\; \tilde{\mu}_i \tilde{g}_i(x) = 0, \quad \text{for } i = 1,2,\ldots,m. \end{array} \tag{1.8}$$

In this case, $\tilde{\mu}$ are referred to as the Lagrange multipliers, and x is called a Karush-Kuhn-Tucker (KKT) point. We denote the set of Lagrange multipliers by

$$\Lambda(x) = \left\{ \tilde{\mu} \in \mathcal{R}^{m+r} \mid \; \tilde{\mu} \text{ satisfies (1.8) at } x \;\right\}. \tag{1.9}$$

A simple inspection of the definitions of $\Lambda(x)$ and $\Lambda_1^g(x)$ reveals that:

$$\tilde{\mu} \in \Lambda(x) \Leftrightarrow (1, \tilde{\mu}) \in \Lambda_1^g(x).$$

Also, because of the first-order homogeneity of the conditions (1.2), and from (1.8), it immediately follows that

$$\Lambda(x) \neq \emptyset \Leftrightarrow \Lambda_1^g(x) \neq \emptyset \Leftrightarrow \exists \tilde{\lambda} \in \Lambda^g(x), \text{ such that } \tilde{\lambda}_0 \neq 0. \tag{1.10}$$

The regularity condition, or constraint qualification, ensures that a linear approximation of the feasible set in the neighborhood of a stationary point x captures the geometry of the feasible set. The regularity condition that we will use at times at a stationary point x is the Mangasarian-Fromovitz constraint qualification (MFCQ) [30, 29]:

$$\text{(MFCQ)} \qquad \begin{array}{l} 1.\; \nabla_x \tilde{h}_j(x), \; j = 1,2,\ldots,r, \text{ are linearly independent and} \\ 2.\; \exists p \neq 0 \text{ such that } \nabla_x \tilde{h}_j(x)^T p = 0, \;\; j = 1,2,\ldots,r \\ \text{and } \nabla_x \tilde{g}_i(x)^T p < 0, \;\; i \in \tilde{\mathcal{A}}(x). \end{array}$$

It is well known [21] that (MFCQ) is equivalent to the fact that the set $\Lambda(x)$ of Lagrange multipliers of (1.1) is not empty and bounded at a stationary point x of (1.1). Note that $\Lambda(x)$ is certainly polyhedral in any case.

Another condition that we will use on occasion is the strict Mangasarian-Fromovitz constraint qualification (SMFCQ). We say that this condition is satisfied by (1.1) at a KKT point x if

$$\text{(SMFCQ)} \qquad \begin{array}{l} 1)\text{ (MFCQ) is satisfied at } x \text{ and} \\ 2)\text{ the Lagrange multiplier set } \Lambda(x) \text{ contains exactly one element.} \end{array}$$

The critical cone at a stationary point x is [13, 38]

$$
\begin{aligned}
\mathcal{C}(x) \quad = \quad & \Big\{ u \in \mathcal{R}^n \mid \nabla_x \tilde{h}_j(x)^T u = 0,\ j = 1, 2, \ldots, r, \\
& \nabla_x \tilde{g}_i(x)^T u \leq 0,\ i \in \tilde{\mathcal{A}}(x);\ \nabla_x \tilde{f}(x)^T u \leq 0 \Big\}.
\end{aligned} \tag{1.11}
$$

We now review the conditions for a point x^* to be a solution of (1.1). The second-order necessary conditions for x^* to be a local minimum are that $\Lambda^g(x^*) \neq \emptyset$ and [24]

$$
\forall u \in \mathcal{C}(x^*),\ \ \exists \tilde{\lambda}^* \in \Lambda^g(x^*),\ \text{such that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u \geq 0. \tag{1.12}
$$

The second-order sufficient conditions for x^* to be a local minimum are that $\Lambda^g(x^*) \neq \emptyset$ and [24]

$$
\forall u \in \mathcal{C}(x^*),\ \ u \neq 0,\ \ \exists \tilde{\lambda}^* \in \Lambda^g(x^*),\ \text{such that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u > 0. \tag{1.13}
$$

Stronger second-order conditions are Robinson's conditions. These conditions are that, at a solution x^*, the following condition holds:

$$
\forall u \in \mathcal{C}(x^*),\ \ u \neq 0,\ \ \forall \tilde{\lambda}^* \in \Lambda^g(x^*),\ \text{we have that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u > 0.
$$

In a fact we will invoke Robinson's conditions for the case where $\Lambda^g_1(x^*) \neq \emptyset$. In the latter situation, Robinson's conditions are equivalent to:

$$
\text{(RSOSC)} \quad \forall u \in \mathcal{C}(x^*),\ \ u \neq 0,\ \ \forall \tilde{\lambda}^* \in \Lambda^g_1(x^*),\ \text{we have that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u > 0.
$$

1.2. Notation. For a mapping $q : \mathcal{R}^n \to \mathcal{R}^l$, we define:

$$
q^+(x) = \begin{pmatrix} \max\{q_1(x), 0\} \\ \max\{q_2(x), 0\} \\ \vdots \\ \max\{q_l(x), 0\} \end{pmatrix} \ \text{and}\ q^-(x) = \begin{pmatrix} \max\{-q_1(x), 0\} \\ \max\{-q_2(x), 0\} \\ \vdots \\ \max\{-q_l(x), 0\} \end{pmatrix}.
$$

With this definition, it immediately follows that $q(x) = q^+(x) - q^-(x)$ and that $|q_i(x)| = q_i^+(x) + q_i^-(x)$, $i = 1, 2, \ldots, l$.

We denote the L_∞ nondifferentiable penalty function by

$$
\tilde{P}_\infty(x) = \max\left\{ \tilde{g}_1(x), \tilde{g}_2(x), ..., \tilde{g}_m(x), \left|\tilde{h}_1(x)\right|, \left|\tilde{h}_2(x)\right|, \ldots, \left|\tilde{h}_r(x)\right|, 0 \right\}. \tag{1.14}
$$

We also define the L_1 penalty function as

$$
\tilde{P}_1(x) = \sum_{i=1}^{m} \tilde{g}_i^+(x) + \sum_{j=1}^{r} \left|\tilde{h}_j(x)\right|. \tag{1.15}
$$

It is immediate that:

$$
0 \leq \tilde{P}_\infty(x) \leq \tilde{P}_1(x) \leq (m + r)\tilde{P}_\infty(x).
$$

An obvious consequence of (1.15) and (1.14) is that x is a feasible point of (1.1) if and only if $\tilde{P}_1(x) = \tilde{P}_\infty(x) = 0$.

We say that the nonlinear program (1.1) satisfies the quadratic growth condition with a parameter $\tilde{\sigma}$ at x^* if

$$\max\left\{\tilde{f}(x)-\tilde{f}(x^*),\tilde{P}_\infty(x)\right\}\geq\tilde{\sigma}\,||x-x^*||^2 \tag{1.16}$$

holds for some $\tilde{\sigma}>0$ and all x in a neighborhood of x^*. The quadratic growth condition is equivalent to the second-order sufficient conditions (1.13), [6, 7, 24, 25, 38] and it is the weakest possible second-order condition.

For the case in which (MFCQ) holds at a solution x^* of (1.1), the quadratic growth condition at x^* is equivalent to [6]

$$\tilde{f}(x)-\tilde{f}(x^*)\geq\tilde{\sigma}_{\tilde{f}}\,||x-x^*||^2 \tag{1.17}$$

for some $\tilde{\sigma}_{\tilde{f}}>0$ and all x feasible in a neighborhood of x^*.

We will also use the Landau notation. We say that a is of order b which we denote by $a=O(b)$, if there exists c such that $a\leq cb$ for all a and b sufficiently small. We denote by $a=\Omega(b)$ quantities a and b that satisfy $a=O(b)$ and $b=O(a)$.

We will use certain symbols twice, to denote related data of different programs. However, to avoid confusion, we will use a $\tilde{}$ sign for the data of the general nonlinear programming problem (1.1), whereas the same objects associated with the (MPCC) problem (to be defined later) are denoted it without the $\tilde{}$ sign. For instance, $\tilde{f}$, $\tilde{g}$, $\tilde{h}$, denote, respectively, the objective, the inequality constraint, and the equality constraints of the general nonlinear programming problem, whereas f,g,h denote, respectively, the objective, the inequality constraints, and the equality constraints of the MPCC problem. The MPCC problem, however, has, in addition, complementarity constraints.

In the literature problems of the type we treat here are also called mathematical programs with equilibrium constraints (MPEC), an acronym that we may use when we invoke optimality conditions from the respective references.

1.3. Exact Penalty Conditions for Degenerate Nonlinear Programming. We now assume that at a solution x^* of the nonlinear program (1.1) the following conditions hold:

1. The Lagrange multiplier set at x^*, $\Lambda(x^*)$, is not empty.
2. The quadratic growth condition (1.16) is satisfied.

Then there exists a neighborhood $\mathcal{V}(x^*)$, some penalty parameters $\tilde{c}_1\geq 0$, $\tilde{c}_\infty\geq 0$ and some growth parameters $\sigma_1>0$ and $\sigma_\infty>0$ such that [7, Theorem 3.113]

$$\begin{aligned}\forall x\in\mathcal{V}(x^*),\psi_1(x)=\tilde{f}(x)+\tilde{c}_1\tilde{P}_1(x)&\geq\tilde{f}(x^*)+\sigma_1\,||x-x^*||^2\\&=\psi_1(x^*)+\sigma_1\,||x-x^*||^2,\end{aligned} \tag{1.18}$$

$$\begin{aligned}\forall x\in\mathcal{V}(x^*),\psi_\infty(x)=\tilde{f}(x)+\tilde{c}_\infty\tilde{P}_\infty(x)&\geq\tilde{f}(x^*)+\sigma_\infty\,||x-x^*||^2\\&=\psi_\infty(x^*)+\sigma_\infty\,||x-x^*||^2.\end{aligned} \tag{1.19}$$

Therefore, x^* becomes an unconstrained strict local minimum for the nondifferentiable functions $\psi_1(x)$ and $\psi_\infty(x)$. Such functions are called nondifferentiable exact merit functions for the nonlinear program (1.1) [3, 4, 14]. If (1.18) and (1.19) are satisfied then we say that the functions $\psi_1(x)$ and $\psi_\infty(x)$ satisfy a quadratic growth condition near x^*.

1.4. Formulation of Mathematical Programs with Complementarity Constraints. We use notation similar to the one in [36] to define a mathematical program with complementarity constraints (MPCC).

$$
\begin{array}{llll}
\text{(MPCC)} \quad \min_x & f(x) & & \\
\text{subject to} & g_i(x) & \leq 0, & i = 1, 2, \ldots, n_i \\
& h_j(x) & = 0, & j = 1, 2, \ldots, n_e \\
& F_{k,1}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\
& F_{k,2}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\
& F_{k,1}(x)F_{k,2}(x) & \leq 0, & k = 1, 2, \ldots, n_c.
\end{array}
$$

In this work we assume that the data of (MPCC) ($f(x), h(x), g(x)$ and $F_{k,i}(x)$, for $k = 1, 2, \ldots, n_c$, and $i = 1, 2$) are twice continuously differentiable.

For a given k, the constraints $F_{k,1}(x) \leq 0$, $F_{k,2}(x) \leq 0$ imply that $F_{k,1}(x)F_{k,2}(x) \leq 0$ is equivalent to $F_{k,1}(x)F_{k,2}(x) = 0$. The constraints $F_{k,1}(x)F_{k,2}(x) \leq 0$ are therefore called complementarity constraints and are active at any feasible point of (MPCC).

Since we cannot have $F_{k,1}(x) < 0$, $F_{k,2}(x) < 0$, and $F_{k,1}(x)F_{k,2}(x) < 0$ simultaneously, it follows that (MFCQ) cannot hold at any feasible point x [27, 36].

1.5. MPCC Notation. In this section, which previews our general convergence results, we use the same notation from [36] to denote certain index sets, because at some point we invoke a theorem from that reference. Later, in our super linear convergence results we will use notation from [16] to denote similar index sets, because we will use results from the latter reference.

If i is one of $1, 2$ we define $\bar{i} = 2 - i + 1$. Therefore $i = 1 \Rightarrow \bar{i} = 2$, and $i = 2 \Rightarrow \bar{i} = 1$. The complementarity constraints can thus be written as $F_{k,i}(x)F_{k,\bar{i}}(x) \leq 0$, $k = 1, 2, \ldots, n_c$. We use the notation

$$F(x) = \left(F_{11}(x), F_{12}(x), F_{21}(x), F_{22}(x), \ldots, F_{n_c1}(x), F_{n_c2}(x)\right)^T. \tag{1.20}$$

The active set of the inequality constraints $g_i(x) \leq 0,\ 1 \leq i \leq m$, at a feasible point x is

$$\mathcal{A}(x) = \{i \in \{1, 2, \ldots, n_i\} \mid g_i(x) = 0\}. \tag{1.21}$$

We use the following notation:

$$\mathcal{I}(x) = \left\{(k, i) \in \{1, 2, \ldots, n_c\} \times \{1, 2\} \mid F_{k,\bar{i}}(x) < 0\right\}, \tag{1.22}$$

$$\overline{\mathcal{I}}(x) = \{(k, i) \in \{1, 2, \ldots, n_c\} \times \{1, 2\} \mid F_{k,i}(x) < 0\}, \tag{1.23}$$

$$\overline{\mathcal{D}}(x) = \left\{(k, i) \in \{1, 2, \ldots, n_c\} \times \{1, 2\} \mid F_{k,i}(x) = F_{k,\bar{i}}(x) = 0\right\}, \tag{1.24}$$

$$\mathcal{I}^c(x) = \{1, 2, \ldots, n_c\} \times \{1, 2\} - \mathcal{I}(x), \tag{1.25}$$

$$\mathcal{K}(x) = \{k \in \{1, 2, \ldots, n_c\} \mid (k, 1) \in \mathcal{I}(x) \ \text{ or } \ (k, 2) \in \mathcal{I}(x)\}, \tag{1.26}$$

$$\overline{\mathcal{K}}(x) = \{k \in \{1, 2, \ldots, n_c\} \mid F_{k,1}(x) = F_{k,2}(x) = 0\} = \{1, 2, \ldots, n_c\} - \mathcal{K}(x). \tag{1.27}$$

There are two cases for the constraints involved in the complementarity constraints at a feasible point x.

1. $F_{k,1}(x) + F_{k,2}(x) < 0$. In this case there is an $i(k) \in \{1, 2\}$ such that $F_{k,i(k)} = 0$ and $F_{k,\bar{i}(k)} < 0$. Therefore, with our notation $k \in \mathcal{K}(x)$, $(k, i(k)) \in \mathcal{I}(x)$

and $(k, \bar{i}(k)) \in \overline{\mathcal{I}}(x)$. We call $F_{k,1}(x), F_{k,2}(x)$ a nondegenerate (or strictly complementary) pair. In the rest of the paper $i(k)$ and $\bar{i}(k)$ will have the meaning defined in this paragraph, whenever $k \in \mathcal{K}$.

2. $F_{k,1}(x) + F_{k,2}(x) = 0$, or $F_{k,1}(x) = F_{k,2}(x) = 0$. In this case $k \in \overline{\mathcal{K}}(x)$, $(k,1) \in \overline{\mathcal{D}}(x)$ and $(k,2) \in \overline{\mathcal{D}}(x)$. We call $F_{k,1}(x), F_{k,2}(x)$ a degenerate pair.

Therefore $\mathcal{I}(x)$ and $\overline{\mathcal{I}}(x)$ contain the indices of the active constraints at which strict complementarity occurs, whereas $\overline{\mathcal{D}}(x)$ contains the indices of the constraints that are degenerate at x from the point of view of complementarity. The set $\mathcal{K}(x)$ represents the indices k at which strict complementarity occurs and $\overline{\mathcal{K}}(x)$ the indices k at which complementarity degeneracy occurs.

Since we are interested in the behavior of (MPCC) at a solution point x^*, we may avoid the dependence of these index sets on x. Therefore we denote $\mathcal{I} = \mathcal{I}(x^*)$, $\overline{\mathcal{D}} = \overline{\mathcal{D}}(x^*)$, $\mathcal{K} = \mathcal{K}(x^*)$, and $\mathcal{A} = \mathcal{A}(x^*)$.

For a set of pairs $\mathcal{J} \subset \{1, 2, \ldots, n_c\} \times \{1, 2\}$ we denote by $F_{\mathcal{J}}$ a map whose components are $F_{k,i}$ with $(k,i) \in \mathcal{J}$.

1.6. Associated Nonlinear Programs at x^*. In this section we associate two nonlinear programs to (MPCC). This will help with characterizing the stationarity conditions for (MPCC). The notation is from [36].

At x^* we associate the relaxed nonlinear program (RNLP) to (MPCC).

$$
\text{(RNLP)} \qquad \begin{array}{rlll} \min_x & f(x) & & \\ \text{subject to} & g_i(x) & \leq 0, & i = 1, 2, \ldots, n_i \\ & h_j(x) & = 0, & j = 1, 2, \ldots, n_e \\ & F_{\overline{\mathcal{D}}}(x) & \leq 0, & \\ & F_{\mathcal{I}}(x) & = 0. & \end{array}
$$

As it can be seen, (RNLP) is obtained from (MPCC) by dropping the elements from $F(x)$ that are inactive at x^*, as well as the complementarity constraints, but enforcing the complements of inactive constraints as equality constraints.

We also associate at x^* the tightened nonlinear program (TNLP), in which all the complementarity constraints in (MPCC) are dropped and all active constraints at x^* connected to complementarity constraints are replaced by equality constraints.

$$
\text{(TNLP)} \qquad \begin{array}{rlll} \min_x & f(x) & & \\ \text{subject to} & g_i(x) & \leq 0, & i = 1, 2, \ldots, n_i \\ & h_j(x) & = 0, & j = 1, 2, \ldots, n_e \\ & F_{\overline{\mathcal{D}}}(x) & = 0, & \\ & F_{\mathcal{I}}(x) & = 0. & \end{array}
$$

We immediately see that, near x^*, (TNLP) is a more constrained problem than (MPCC), which in turn is more constrained than (RNLP), and all three programs have the same objective function. As a result, if x^* is a local solution of (RNLP), then it must be a local solution of (MPCC). Also, if x^* is a local solution of (MPCC), then it will be a local solution of (TNLP). None of the reverse implications hold in general for either local solutions or stationary points.

However, if (TNLP) satisfies (SMFCQ) at a solution x^* of (MPCC), then x^* is a Karush-Kuhn-Tucker point of (TNLP) and (RNLP) [36].

2. The Lagrange Multiplier Set of (MPCC). In this section we analyze the relationship between the relevant mathematical objects of (MPCC) and (RNLP) at a solution x^*. The (RNLP) formulation does not immediately violate (MFCQ), the way

(MPCC) does. By establishing a correspondence between the Lagrange multiplier sets of (RNLP) and (MPCC) we ensure that, under certain conditions, (MPCC) has a nonempty Lagrange multiplier set, although it does not satisfy a constraint qualification.

2.1. Critical Cones. In this section we compare the critical cones of (MPCC) and (RNLP). The active sets play a structural part in the definition of the critical cones. We have that:

$$\nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = F_{k,1}(x^*)\nabla_x F_{k,2}(x^*) + F_{k,2}(x^*)\nabla_x F_{k,1}(x^*).$$

We distinguish two cases.

1. If $k \in \overline{\mathcal{K}}$, we have that $F_{k,1}(x^*) = F_{k,2}(x^*) = 0$, and, as a result,

$$k \in \overline{\mathcal{K}} \Rightarrow \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = 0. \tag{2.1}$$

   Therefore, if $k \in \overline{\mathcal{K}}$, the constraint $F_{k,1}(x)F_{k,2}(x) \leq 0$, which is active at x^*, has no bearing on the definition (1.11) of the critical cone (it would just add the constraint $0 \leq 0$).
2. If $k \in \mathcal{K}$, then there exist an $i(k)$ such that $(k, i(k)) \in \mathcal{I}$ and $(k, \bar{i}(k)) \in \overline{\mathcal{I}}$. The constraints $F_{k,i(k)}(x) \leq 0$ and $F_{k,i(k)}(x)F_{k,\bar{i}(k)}(x) \leq 0$ are active at x^*, whereas $F_{k,\bar{i}(k)}(x^*) < 0$ and the corresponding constraint is inactive at x^*. Therefore we have that

$$\nabla_x \left(F_{k,i(k)}F_{k,\bar{i}(k)}\right)(x^*) = F_{k,\bar{i}(k)}(x^*)\nabla_x F_{k,i(k)}(x^*), \tag{2.2}$$

   and thus the constraints connected to k that enter the definition of the critical cone (1.11) are:

$$\nabla_x F_{k,i(k)}(x^*)u \leq 0, \quad F_{k,\bar{i}(k)}(x^*)\left(\nabla_x F_{k,i(k)}(x^*)\right)u \leq 0$$

   for u an element of the critical cone.

Using the definition (1.11) we get that the critical cone of (MPCC) is

$$\begin{array}{llcll}
\mathcal{C}_{\mathsf{MPCC}} = \{u \in R^n \mid & \nabla_x f(x^*)u & \leq & 0, & \\
& \nabla_x g_i(x^*)u & \leq & 0, & i \in \mathcal{A} \\
& \nabla_x h_j(x^*)u & = & 0, & j \in 1, 2, \ldots, n_e \\
& \nabla_x F_{k,1}(x^*)u & \leq & 0, & k \in \overline{\mathcal{K}} \\
& \nabla_x F_{k,2}(x^*)u & \leq & 0, & k \in \overline{\mathcal{K}} \\
& \nabla_x F_{k,i(k)}(x^*)u & \leq & 0, & (k, i(k)) \in \mathcal{I} \\
& F_{k,\bar{i}(k)}(x^*)\nabla_x F_{k,i(k)}(x^*)u & \leq & 0, & (k, i(k)) \in \mathcal{I}\}\,.
\end{array} \tag{2.3}$$

Note that the definition that we use here of the critical cone corresponds to the nonlinear programming interpretation of (MPCC). There exists another, combinatorial, definition of the critical cone, that is used in connection to disjunctive approaches [27, Equation 5.3.(2)].

We use (1.11) again to determine the critical cone of the relaxed nonlinear program. It is immediate from the definition of the index sets $\mathcal{I}, \mathcal{K}$, and $\overline{\mathcal{D}}$ that all constraints involving components of $F(x)$ are active at x^* for (RNLP). It thus follows that the critical cone of (RNLP) is

$$
\mathcal{C}_{\mathsf{RNLP}} = \{u \in R^n \mid \begin{array}{lcll} \nabla_x f(x^*)u & \leq & 0, & \\ \nabla_x g_i(x^*)u & \leq & 0, & i \in \mathcal{A} \\ \nabla_x h_j(x^*)u & = & 0, & j \in 1,2,\ldots,n_e \\ \nabla_x F_{k,1}(x^*)u & \leq & 0, & k \in \overline{\mathcal{K}} \\ \nabla_x F_{k,2}(x^*)u & \leq & 0, & k \in \overline{\mathcal{K}} \\ \nabla_x F_{k,i(k)}(x^*)u & = & 0, & (k,i(k)) \in \mathcal{I}\}\,. \end{array} \tag{2.4}
$$

LEMMA 2.1. $\mathcal{C}_{\mathsf{MPCC}} = \mathcal{C}_{\mathsf{RNLP}}$.

Proof The conclusion is immediate, by noting that all the constraints involving the critical cones are the same with the exception of the ones involving indices k for which $(k,i(k)) \in \mathcal{I}$. For these k, from the definition (1.22) of the index sets it follows that $F_{k,\bar{i}(k)}(x^*) < 0$. We therefore have that:

$$
\begin{array}{lclcrcl} \nabla_x F_{k,i(k)}(x^*)u & \leq & 0 & \text{and} & F_{k,\bar{i}(k)}(x^*)\nabla_x F_{k,i(k)}(x^*)u & \leq & 0 \Leftrightarrow \\ \nabla_x F_{k,i(k)}(x^*)u & \leq & 0 & \text{and} & \nabla_x F_{k,i(k)}(x^*)u & \geq & 0 \Leftrightarrow \\ & & & & \nabla_x F_{k,i(k)}(x^*)u & = & 0. \end{array}
$$

Since the remaining constraints of (RNLP) and (MPCC) are the same this equivalence proves the claim. ◇

2.2. Generalized Lagrange Multipliers. The set of generalized Lagrange multipliers of (MPCC) at x^* is a set of multiples

$$
0 \neq (\alpha,\nu,\pi,\mu,\eta) \in \mathcal{R} \times \mathcal{R}^{n_i} \times \mathcal{R}^{n_e} \times \mathcal{R}^{2n_c} \times \mathcal{R}^{n_c}
$$

that satisfies the Fritz-John conditions (1.2). Since μ are the multipliers corresponding to the components of $F(x)$, we will index them by elements in $(1,2,\ldots,n_c) \times (1,2)$. The Fritz-John conditions for (MPCC) at x^* are that x^* is feasible for (MPCC) and that

$$
\alpha \nabla_x f(x^*) + \sum_{i=1}^{n_i} \nu_i \nabla_x g_i(x^*) + \sum_{j=1}^{n_e} \pi_j \nabla_x h_j(x^*) + \tag{2.5}
$$

$$
\sum_{k=1}^{n_c} \left[\mu_{k,1}\nabla_x F_{k,1}(x^*) + \mu_{k,2}\nabla_x F_{k,2}(x^*) + \eta_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*)\right] = 0
$$

$$
\begin{array}{rrrl} F_{k,i}(x^*) \leq 0, & \mu_{k,i} \geq 0, & \mu_{k,i}F_{k,i}(x^*) = 0, & k = 1,2,\ldots,n_c, \\ & & & i = 1,2 \\ g_i(x^*) \leq 0, & \nu_i \geq 0, & \nu_i g_i(x^*) = 0, & i = 1,2,\ldots,n_i \\ F_{k,1}(x^*)F_{k,2}(x^*) \leq 0, & \eta_k \geq 0, & \eta_k F_{k,1}(x^*)F_{k,2}(x^*) = 0, & k = 1,2,\ldots,n_c. \end{array} \tag{2.6}
$$

From our definition of the index sets it follows that $F_{\overline{\mathcal{I}}}(x^*) < 0$ and $g_{\mathcal{A}^c}(x^*) < 0$. Therefore, from the complementarity conditions (2.6), it follows that $\mu_{\overline{\mathcal{I}}} = 0$ and $\nu_{\mathcal{A}^c} = 0$.

We can also determine the relations satisfied by the generalized Lagrange multipliers of (RNLP). As discussed above, the index sets that define (RNLP) have been chosen such that all constraints involving components of $F(x)$ are active. Therefore the generalized Lagrange multipliers are:

$$
0 \neq (\tilde{\alpha},\tilde{\nu},\tilde{\pi},\tilde{\mu},\tilde{\eta}) \in \mathcal{R} \times \mathcal{R}^{n_i} \times \mathcal{R}^{n_e} \times \mathcal{R}^{n_{\overline{\mathcal{D}}}} \times \mathcal{R}^{n_{\mathcal{I}}}
$$

that satisfy the Fritz-John conditions:

$$\tilde{\alpha}\nabla_x f(x^*) + \sum_{i=1}^{n_i} \tilde{\nu}_i \nabla_x g_i(x^*) + \sum_{j=1}^{n_e} \tilde{\pi}_j \nabla_x h_j(x^*) + \sum_{k\in\overline{\mathcal{K}}} \left[\tilde{\mu}_{k,1}\nabla_x F_{k,1}(x^*) + \tilde{\mu}_{k,2}\nabla_x F_{k,2}(x^*)\right] + \sum_{k\in\mathcal{K}} \tilde{\eta}_{k,i(k)} \nabla_x F_{k,i(k)}(x^*) = 0 \tag{2.7}$$

$$\begin{array}{llll} g_i(x^*) \leq 0, & \tilde{\nu}_i \geq 0, & \tilde{\nu}_i g_i(x^*) = 0, & i = 1, 2, \ldots, n_i \\ & \tilde{\mu}_{k,1} \geq 0, & \tilde{\mu}_{k,2} \geq 0, & k \in \overline{\mathcal{K}}. \end{array} \tag{2.8}$$

Here $\tilde{\mu}$ is a vector that is indexed by elements of $\overline{\mathcal{D}}$, and $\tilde{\eta}$ is indexed by elements of $\mathcal{I}$.

2.3. Relations between the generalized Lagrange Multiplier Sets of (MPCC) and (RNLP). Take $\tilde{\lambda} = (\tilde{\alpha}, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta}) \in \Lambda^g_{\mathsf{RNLP}}$. We construct from the generalized multiplier $\tilde{\lambda}$ of (RNLP) a generalized multiplier λ° of (MPCC). We define the following types of components of λ°.

1. Components that correspond to the objective function or the inequality constraints $g_i(x) \leq 0$ and equality constraints $h_j(x) = 0$

$$\alpha^\circ = \tilde{\alpha}; \quad \nu^\circ = \tilde{\nu}; \quad \pi^\circ = \tilde{\pi}. \tag{2.9}$$

2. Components connected to the pairwise degenerate constraints. For these we have $k \in \overline{\mathcal{K}}$ and $(k, 1), (k, 2) \in \overline{\mathcal{D}}$ or $F_{k,1}(x^*) = F_{k,2}(x^*) = 0$. We define

$$\mu^\circ_{k,i} = \tilde{\mu}_{k,i}, \ (k, i) \in \overline{\mathcal{D}}; \quad \eta^\circ_k = 0, \ k \in \overline{\mathcal{K}}. \tag{2.10}$$

Similar to the equation (2.1) we have that:

$$\nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = 0,$$

and therefore

$$\tilde{\mu}_{k,1}\nabla_x F_{k,1}(x^*) + \tilde{\mu}_{k,2}\nabla_x F_{k,2}(x^*) = \mu^\circ_{k,1}\nabla_x F_{k,1}(x^*) + \mu^\circ_{k,2}\nabla_x F_{k,2}(x^*) + \eta^\circ_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*). \tag{2.11}$$

3. Components connected to pairwise strictly complementary constraints. In this case we have $k \in \mathcal{K}$, $(k, i(k)) \in \mathcal{I}$, and $(k, \bar{i}(k)) \in \overline{\mathcal{I}}$. Therefore $F_{k,\bar{i}(k)}(x^*) < 0$, $F_{k,i(k)}(x^*) = 0$, and we thus define the multipliers

$$\begin{array}{rcll} \mu^\circ_{k,i(k)} & = & \max\left\{\tilde{\eta}_{k,i(k)}, 0\right\}, & (k, i(k)) \in \mathcal{I} \\ \mu^\circ_{k,\bar{i}(k)} & = & 0, & (k, \bar{i}(k)) \in \overline{\mathcal{I}} \\ \eta^\circ_k & = & \frac{1}{F_{k,\bar{i}(k)}(x^*)} \min\left\{\tilde{\eta}_{k,i(k)}, 0\right\}, & k \in \mathcal{K}. \end{array} \tag{2.12}$$

It is immediate from these definitions that $\mu^\circ_{k,i(k)} \geq 0$ and $\eta^\circ_k \geq 0$. Since, for fixed k, $\tilde{\eta}_{k,i(k)}$ is the only multiplier of (RNLP) involved in definition (2.12), we obtain using (2.2) that

$$\begin{aligned} \tilde{\eta}_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) &= \left[\max\left\{\tilde{\eta}_{k,i(k)}, 0\right\} + \min\left\{\tilde{\eta}_{k,i(k)}, 0\right\}\right] \nabla_x F_{k,i(k)}(x^*) \\ &= \mu^\circ_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) + \eta^\circ_k F_{k,\bar{i}(k)}(x^*)\nabla_x F_{k,i(k)}(x^*) = \\ &\mu^\circ_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) + \mu^\circ_{k,\bar{i}(k)}\nabla_x F_{k,\bar{i}(k)}(x^*) + \eta^\circ_k \nabla_x \left(F_{k,i(k)}F_{k,\bar{i}(k)}\right)(x^*). \end{aligned} \tag{2.13}$$

After we compare the terms that, following (2.11) and (2.13), are equal in (2.7) and (2.5), we get that $\lambda^\circ = (\alpha^\circ, \nu^\circ, \pi^\circ, \mu^\circ, \eta^\circ)$ satisfies (2.5) as well as (2.6). By tracing the definition of λ° we also have that $\tilde{\lambda} \neq 0 \Rightarrow \lambda^\circ \neq 0$. Therefore λ° is a generalized Lagrange multiplier of (MPCC) or

$$\lambda^\circ = (\alpha^\circ, \nu^\circ, \pi^\circ, \mu^\circ, \eta^\circ) \in \Lambda^g_{\mathsf{MPCC}}, \tag{2.14}$$

where $\alpha^\circ = \tilde{\alpha}$ from (2.9).

THEOREM 2.2. *If the set of Lagrange multipliers of (RNLP) is not empty, then the set of Lagrange multipliers of (MPCC) is not empty.*

Proof Since the Lagrange multiplier set of (RNLP) is not empty, we can choose $\tilde{\lambda} = (1, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta}) \in \Lambda^g_{1,\ \mathsf{RNLP}}$. From (2.14) it follows that $\lambda^\circ = (1, \nu^\circ, \pi^\circ, \mu^\circ, \eta^\circ) \in \Lambda^g_{1,\mathsf{MPCC}}$ is a generalized multiplier of (MPCC). From (1.10) it follows that the Lagrange multiplier set of (MPCC) is not empty. $\diamond$

COROLLARY 2.3. *Assume that (TNLP) satisfies (SMFCQ) at a solution x^* of (MPCC), i.e.*

1. $\nabla_x F_{\overline{\mathcal{D}}}(x^*)$, $\nabla_x F_{\mathcal{I}}(x^*)$, *and* $\nabla_x h(x^*)$ *are linearly independent.*
2. *There exists* $p \neq 0$ *such that* $\nabla_x F_{\overline{\mathcal{D}}}(x^*)p = 0$, $\nabla_x F_{\mathcal{I}}(x^*)p = 0$, $\nabla_x h(x^*)p = 0$, $\nabla_x g_i(x^*)p < 0$, *for* $i \in \mathcal{A}(x^*)$.
3. *The Lagrange multiplier set of (TNLP) at* x^* *has a unique element.*

Then the Lagrange multiplier set of (MPCC) is not empty.

Proof From [36, Theorem 2], since (TNLP) satisfies (SMFCQ) at x^*, the Lagrange multiplier set of (RNLP) is not empty. Following Theorem 2.2, we obtain that the Lagrange multiplier set of (MPCC) is not empty, which proves the claim. $\diamond$

Unfortunately, the reverse statement of Theorem 2.2 does not hold in the absence of (SMFCQ), as is shown in [36]. Indeed, consider the following example:

$$\begin{array}{lll} \min_{y,x} & y - x & \\ & y & \leq 0 \\ & y + x & \leq 0 \\ & y(y+x) & \leq 0 \\ & x & \leq 0. \end{array} \tag{2.15}$$

The unique minimum of this problem is $(0,0)$. However, if we construct the associated (RNLP) formulation, we obtain

$$\begin{array}{lll} \min_{y,x} & y - x & \\ & y & \leq 0 \\ & y + x & \leq 0 \\ & x & \leq 0. \end{array} \tag{2.16}$$

The point $(y, 0)$ is feasible for $y < 0$ for the now-linear program (2.16). Thus (2.16) is unbounded and cannot have $(0,0)$ as a stationary point. Therefore Theorem 2.2 cannot be applied, since the Lagrange multiplier set of (2.16) is empty. In this situation (TNLP) associated to (2.15) of (2.15) does not satisfy either (MFCQ) or (SMFCQ).

2.4. An alternative formulation. We also investigate the following equivalent formulation of (MPCC), where the complementarity constraints have been replaced by one constraint:

$$
(2.17)\qquad
\begin{array}{rrll}
\min_x & f(x) & & \\
\text{subject to} & g_i(x) & \leq 0, & i = 1, 2, \ldots, n_i \\
& h_j(x) & = 0, & j = 1, 2, \ldots, n_e \\
& F_{k,1}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\
& F_{k,2}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\
& \sum_{k=1}^{n_c} F_{k,1}(x) F_{k,2}(x) & \leq 0. &
\end{array}
$$

At a feasible point of the above program, we must have that $\sum_{k=1}^{n_c} F_{k,1}(x)F_{k,2}(x) = 0$ and the equivalence between (2.17) and (MPCC) follows immediately. This formulation is of interest in computations because it has less constraints than (MPCC).

Lemma 2.4. *If the Lagrange multiplier set of (MPCC) is not empty, there exists a generalized Lagrange multiplier* $(1, \nu, \pi, \mu, \eta) \in \Lambda^g_{\mathsf{MPCC}}$ *such that* $\eta_k = \eta_1$, $k = 2, 3, \ldots, n_c$. **Proof** Let $\lambda^\circ = (1, \nu^\circ, \pi^\circ, \mu^\circ, \eta^\circ) \in \Lambda^g_{\mathsf{MPCC}}$ be a generalized Lagrange multiplier of (MPCC). Now let $d \in \mathcal{R}^{n_c}$ such that $d \geq 0$.

If k corresponds to degenerate complementarity constraints, $k \in \overline{K}$, we have, as argued above, that $F_{k,1}(x^*) = F_{k,2}(x^*) = 0$, and thus

$$d_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = 0.$$

For this case, define

$$\eta^*_k = \eta^\circ_k + d_k, \quad \mu^*_{k,1} = \mu^\circ_{k,1}, \quad \mu^*_{k,2} = \mu^\circ_{k,2},$$

which results in

$$
(2.18)\qquad
\begin{aligned}
&\mu^*_{k,1}\nabla_x F_{k,1}(x^*) + \mu^*_{k,2}\nabla_x F_{k,2}(x^*) + \eta^*_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = \\
&\quad \mu^\circ_{k,1}\nabla_x F_{k,1}(x^*) + \mu^\circ_{k,2}\nabla_x F_{k,2}(x^*) + \eta^\circ_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*).
\end{aligned}
$$

If k corresponds to strict complementarity constraints, $k \in \mathcal{K}$, we have that $F_{k,i(k)}(x^*) = 0$, $F_{k,\bar{i}(k)}(x^*) < 0$ and thus $\mu_{k,\bar{i}(k)} = 0$. Define

$$\eta^*_k = \eta^\circ_k + d_k, \quad \mu^*_{k,i(k)} = \mu^\circ_{k,i(k)} - d_k F_{k,\bar{i}(k)}(x^*) \geq 0, \quad \mu^*_{k,\bar{i}(k)} = \mu^\circ_{k,\bar{i}(k)} = 0.$$

Since $F_{k,i(k)}(x^*) = 0$ we have that

$$\nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = F_{k,\bar{i}(k)}(x^*)\nabla_x F_{k,i(k)}(x^*).$$

In turn, the last equation implies that

$$
\begin{aligned}
&\mu^*_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) + \mu^*_{k,\bar{i}(k)}\nabla_x F_{k,\bar{i}(k)}(x^*) + \eta^*_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = \\
&\quad \left(\mu^\circ_{k,i(k)} - d_k F_{k,\bar{i}}(x^*)\right)\nabla_x F_{k,i(k)}(x^*) + \mu^\circ_{k,\bar{i}(k)}\nabla_x F_{k,\bar{i}(k)}(x^*) + \\
&\quad \left(\eta^\circ_k + d_k\right)\nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = \mu^\circ_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) + \\
&\quad \mu^\circ_{k,\bar{i}(k)}\nabla_x F_{k,\bar{i}(k)}(x^*) + \eta^\circ_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*).
\end{aligned}
$$

Since λ° satisfies (2.5) and (2.6), it follows from the preceding equation and (2.18), in a manner similar to the proof of the Theorem 2.2, that $\lambda^* = (1, \nu^\circ, \pi^\circ, \mu^*, \eta^*)$ satisfies (2.5) and (2.6) and thus $\lambda^* \in \Lambda^g_{\mathsf{MPCC}}$ for any $0 \leq d \in \mathcal{R}^{n_c}$ where $\eta^* = \eta^\circ + d$. The conclusion is immediate, since we can always choose a vector $d \geq 0$ such that $\eta^*_k = \eta^*_1$, $k = 1, 2, \ldots, n_c$. One such choice, for example, is $d = ||\eta^\circ||_\infty (1, 1, \ldots, 1)^T - \eta^\circ$. ⋄

We now describe the generalized Lagrange multiplier set of the alternative formulation (2.17). We denote mathematical objects connected to (2.17) by the subscript ${}_{MPCC1}$. We write the Fritz-John conditions (1.2) for (2.17) at the point x^*, and we obtain

$$\begin{aligned} &\alpha^\diamond \nabla_x f(x^*) + \sum_{i=1}^{n_i} \nu_i^\diamond \nabla_x g_i(x^*) + \sum_{j=1}^{n_e} \pi_j^\diamond \nabla_x h_j(x^*) + \\ &\sum_{k=1}^{n_c} \sum_{i=1}^{2} \mu_{k,i}^\diamond \nabla_x F_{k,i}(x^*) + \eta_1^\diamond \sum_{k=1}^{n_c} \nabla_x \left(F_{k,1} F_{k,2}\right)(x^*) = 0 \end{aligned} \tag{2.19}$$

$$\begin{array}{llll} F_{k,i}(x^*) \le 0, & \mu_{k,i}^\diamond \ge 0, & \mu_{k,i}^\diamond F_{k,i}(x^*) = 0, & k = 1, 2, \ldots, n_c, \\ & & & i = 1, 2 \\ g_i(x^*) \le 0, & \nu_i^\diamond \ge 0, & \nu_i^\diamond g_i(x^*) = 0, & i = 1, 2, \ldots, n_i \end{array} \tag{2.20}$$

and $\eta_1^\diamond \ge 0 \in \mathcal{R}$. A generalized multiplier of (2.17) is thus

$$\lambda^\diamond = (\alpha^\diamond, \nu^\diamond, \pi^\diamond, \mu^\diamond, \eta_1^\diamond) \in \Lambda^g_{MPCC1} \subset \mathcal{R} \times \mathcal{R}^{n_i} \times \mathcal{R}^{n_e} \times \mathcal{R}^{2n_c} \times \mathcal{R},$$

where $\lambda^\diamond$ satisfies the Fritz-John conditions (2.19), (2.20).

THEOREM 2.5.

- **i)** *The formulation (2.17) has a not empty Lagrange multiplier set if and only if (MPCC) has a not empty Lagrange multiplier set.*
- **ii)** *If (MPCC) has a not empty Lagrange multipliers set and it satisfies the quadratic growth condition at x^*, then (2.17) has a not empty Lagrange multipliers set and satisfies the quadratic growth condition at x^*.*

Proof If the Lagrange multiplier set of (2.17) is not empty, then there exists $\lambda^\diamond = (1, \nu^\diamond, \pi^\diamond, \mu^\diamond, \eta_1^\diamond) \in \mathcal{R} \times \mathcal{R}^{n_i} \times \mathcal{R}^{n_e} \times \mathcal{R}^{2n_c} \times \mathcal{R}$ that satisfies (2.19–2.20). Define $\eta^* = (\eta_1^\diamond, \eta_1^\diamond, \ldots, \eta_1^\diamond)^T \in \mathcal{R}^{n_c}$ and $\lambda^* = (1, \nu^\diamond, \pi^\diamond, \mu^\diamond, \eta^*)$. It follows by inspection that λ^* satisfies (2.5), (2.6) at x^*. Therefore λ^* is a generalized Lagrange multiplier of (MPCC), which means that $(\nu^\diamond, \pi^\diamond, \mu^\diamond, \eta^*)$ is a Lagrange multiplier of (MPCC). Thus the Lagrange multiplier set of (MPCC) is not empty. Conversely, applying Lemma 2.4, if the Lagrange multiplier set of (MPCC) is not empty, there exists the generalized Lagrange multiplier $\lambda = (1, \nu, \pi, \mu, \eta)$ of (MPCC) that satisfies $\eta_k = \eta_1$, for $k = 1, 2, \ldots, n_c$. It immediately follows that, since λ satisfies (2.5) and (2.6), $(1, \nu, \pi, \mu, \eta_1)$ satisfies (2.19) and (2.20) and is thus a generalized Lagrange multiplier of (2.17). Therefore (ν, π, μ, η_1) is a Lagrange multiplier of (2.17) at x^*. The proof of part **i)** is complete.

For the second part, if (MPCC) has a not empty Lagrange multiplier set, then, from part **i)**, the formulation (2.17) has a not empty Lagrange multiplier set. Since (MPCC) satisfies the quadratic growth condition, we have that [7, Theorem 3.113] there exists $c_1 > 0$ and σ_1 such that, for any x in a neighborhood of x^*, we have that:

$$\begin{aligned} f(x) &+ c_1 \left|\left|g^+(x)\right|\right|_1 + c_1 \left|\left|h(x)\right|\right|_1 + c_1 \left|\left|F_1^+(x)\right|\right|_1 + c_1 \left|\left|F_2^+(x)\right|\right|_1 \\ &+ c_1 \textstyle\sum_{k=1}^{n_c} \left(F_{k1}(x) F_{k2}(x)\right)^+ \ge f(x^*) + \sigma_1 \left|\left|x - x^*\right|\right|^2. \end{aligned} \tag{2.21}$$

Here we denote by $F_1(x) = (F_{11}(x), F_{21}(x), \ldots, F_{n_c 1}(x))$ and by $F_2(x) = (F_{12}(x), F_{22}(x), \ldots, F_{n_c 2}(x))$. We now analyze in detail one term of the rightmost sum. We have that:

$$F_{k1}(x) F_{k2}(x) = F_{k1}^+(x) F_{k2}^+(x) + F_{k1}^-(x) F_{k2}^-(x) - F_{k1}^+(x) F_{k2}^-(x) - F_{k1}^-(x) F_{k2}^+(x).$$

Since only one term in the preceding sum can be a nonzero, we obtain that:

$$\begin{aligned}(F_{k1}(x)F_{k2}(x))^+ &= F_{k1}^+(x)F_{k2}^+(x) + F_{k1}^-(x)F_{k2}^-(x),\\ (F_{k1}(x)F_{k2}(x))^- &= F_{k1}^+(x)F_{k2}^-(x) + F_{k1}^-(x)F_{k2}^+(x).\end{aligned}$$

This implies that there exists a $c_{FF} > 0$ such that, for any x in a neighborhood of x^*, and for any $k = 1, 2, \ldots, n_c$, we will have that:

$$(F_{k1}(x)F_{k2}(x))^- \leq c_{FF}\left(F_{k1}^+(x) + F_{k2}^+(x)\right).$$

In turn, this implies that for any x in a neighborhood of x^*, we will have that:

$$\begin{array}{rcl}\left(\sum_{k=1}^{n_c} F_{k1}(x)F_{k2}(x)\right)^+ &=& \left(\sum_{k=1}^{n_c}(F_{k1}(x)F_{k2}(x))^+ - \sum_{k=1}^{n_c}(F_{k1}(x)F_{k2}(x))^-\right)^+\\ &\geq& \sum_{k=1}^{n_c}(F_{k1}(x)F_{k2}(x))^+ - \sum_{k=1}^{n_c}(F_{k1}(x)F_{k2}(x))^-\\ &\geq& \sum_{k=1}^{n_c}(F_{k1}(x)F_{k2}(x))^+\\ &-& c_{FF}\left(\left|\left|F_1^+(x)\right|\right|_1 + \left|\left|F_2^+(x)\right|\right|_1\right).\end{array} \tag{2.22}$$

Using the last inequality we obtain that:

$$\begin{array}{rl} f(x) & +\ (c_1 + c_1c_{FF})\left|\left|g^+(x)\right|\right|_1 + (c_1 + c_1c_{FF})\left|\left|h(x)\right|\right|_1 + (c_1 + c_1c_{FF})\left|\left|F_1^+(x)\right|\right|_1\\ & +\ (c_1 + c_1c_{FF})\left|\left|F_2^+(x)\right|\right|_1 + (c_1 + c_1c_{FF})\left(\displaystyle\sum_{k=1}^{n_c} F_{k1}(x)F_{k2}(x)\right)^+ \geq\\ f(x) & +\ (c_1 + c_1c_{FF})\left|\left|g^+(x)\right|\right|_1 + (c_1 + c_1c_{FF})\left|\left|h(x)\right|\right|_1 + (c_1 + c_1c_{FF})\left|\left|F_1^+(x)\right|\right|_1\\ & +\ (c_1 + c_1c_{FF})\left|\left|F_2^+(x)\right|\right|_1 + c_1\left(\displaystyle\sum_{k=1}^{n_c} F_{k1}(x)F_{k2}(x)\right)^+ \overset{(2.22)}{\geq}\\ f(x) & +\ (c_1 + c_1c_{FF})\left|\left|g^+(x)\right|\right|_1 + (c_1 + c_1c_{FF})\left|\left|h(x)\right|\right|_1 + c_1\left|\left|F_1^+(x)\right|\right|_1\\ & +\ c_1\left|\left|F_2^+(x)\right|\right|_1 + c_1\displaystyle\sum_{k=1}^{n_c}(F_{k1}(x)F_{k2}(x))^+\\ & \overset{(2.21)}{\geq}\ f(x^*) + \sigma_1\left|\left|x - x^*\right|\right|^2. \end{array}$$

From the last inequality it follows that (2.17) satisfies the quadratic growth condition (1.16) with parameter $\min\left\{\frac{1}{2}, \frac{\sigma_1}{2(n_i+n_e+3n_c+1)(c_1+c_1c_{FF})}\right\}$ The proof of part **ii)** is complete. ◇

Theorems 2.2 and 2.5 give sufficient conditions for (MPCC) and (2.17) to have a nonempty Lagrange multiplier set in spite of the fact that neither satisfy a constraint qualification at any point in the usual sense of nonlinear programming. In Section 3 these conditions will imply that a relaxed version of either (MPCC) or (2.17) will have the same solution as (MPCC) and will satisfy (MFCQ), which makes either approachable by SQP algorithms.

3. The Elastic Mode. An important class of techniques for solving nonlinear programs (1.1) is sequential quadratic programming. The main step in an algorithm of this type is solving the quadratic program

$$\begin{array}{rll} \min_d & \nabla_x\tilde{f}(x)^Td + d^T\widetilde{W}d, & \\ \text{subject to} & \tilde{g}_i(x) + \nabla_x\tilde{g}_i(x)^Td \leq 0, & i = 1, 2, \ldots, m\\ & \tilde{h}_j(x) + \nabla_x\tilde{h}_j(x)^Td = 0, & j = 1, 2, \ldots, r. \end{array} \tag{3.1}$$

The matrix $\widetilde{W}$ can be the Hessian of the Lagrangian (1.1) at (x, λ) [15], where λ is a suitable approximation of a Lagrange multiplier or a positive definite matrix that approximates the Hessian of the Lagrangian on a certain subspace [14, 22, 31]. A trust-region type constraint may be added to (3.1) to enforce good global convergence properties [15]. The solution $\overline{d}$ of (3.1) is then used in conjunction with a merit function and/or line search to determine a new iterate. We give here only a brief description of SQP algorithms, since our interest is solely in showing how the difficulties regarding the potential infeasibility of (3.1) when applied to (MPCC) can be circumvented. For more details about SQP methods see [14, 15, 22, 31].

If a nonlinear program satisfies (MFCQ) at x^* then the quadratic program will be feasible in a neighborhood of x^*. If (MFCQ) does not hold at x^*, however, the possibility exists that (3.1) is infeasible, no matter how close to x^* [27, 33, 36]. This is an issue in the context of this paper because (MPCC) does not satisfy the (MFCQ) at a solution x^*.

In the case of infeasible subproblems some of the SQP algorithms initiate the *elastic mode* [22]. This consists of modifying the nonlinear program (1.1) by relaxing the constraints and adding a penalty term to the objective function.

When the elastic mode is implemented, only the nonlinear constraints are relaxed [22]. To represent this situation in our approach, we assume that $\tilde{g}_i(x)$ for $i = 1, 2, \ldots, l_i$, and $\tilde{h}_j(x)$, for $j = 1, 2, \ldots, l_e$, are linear.

For these constraints, we assume that

[B1] The set $\mathcal{F}_l$ is feasible, where

$$\mathcal{F}_l = \left\{ x \,\middle|\, \tilde{g}_i(x) \leq 0,\ i = 1, 2, \ldots, l_i, \quad \tilde{h}_j(x) = 0, j = 1, 2, \ldots, l_e \right\},$$

[B2] The preceding representation of $\mathcal{F}_l$ is minimal: $\nabla_x \tilde{h}_j(x)$ are linearly independent, $j = 1, 2, \ldots, l_e$, and $\exists d$ such that $\nabla_x \tilde{h}_j(x) d = 0$, $\nabla_x \tilde{g}_i(x) d < 0$.

None of these assumptions induces any loss of generality. Indeed, if $\mathcal{F}_l = \emptyset$, then the original nonlinear program (1.1) is infeasible. Most software for nonlinear programming starts with an analysis of the linear constraints and the infeasibility of the problem, which is the correct outcome when $\mathcal{F}_l = \emptyset$, is immediately detected. Clearly, the interesting situation is when $\mathcal{F}_l$ is feasible, which is our assumption **[B1]**.

If the set $\mathcal{F}_l$ is polyhedral and nonempty, it must have a minimal representation [39, 11]. In addition, this representation can be computed by solving only one linear program [18]. The methods we use in this work are of the sequential quadratic programming type, where all constraints of the nonlinear program are linearized, and the nonlinear constraints are perhaps relaxed. Since the set $\mathcal{F}_l$ is invariant under linearization, any of its representations will produce the same quadratic program subproblems. As long as we do not involve the Lagrange Multipliers of the constraints defining $\mathcal{F}_l$, that are not invariant to a change in representation, assumption **[B2]** does not result in any loss of generality as long as assumption **[B1]** holds.

Depending on the type of the relaxation, we can have either an L_1 or an L_∞ approach. Our results are related to the situation in which the merit functions $\psi_1(x)$ and $\psi_\infty(x)$ are exact for the nonlinear program (1.1). Since both merit functions are widely used we will state our results for both cases.

First we consider the case in which the added penalty term is of the L_∞ [3] type:

Define $\tilde{c}_\infty > 0$, x^0, $k = 0$, $\sigma \in (0, \frac{1}{2})$, $\tau \in (0,1)$, $s > 0$.

QP Find the solution $d = d^k$ of the quadratic program.

$$\begin{array}{lrcl} \text{minimize}_d & \frac{1}{2}d^T d + \nabla_x \tilde{f}(x^k)^T d & & \\ \text{subject to} & \tilde{h}(x^k) + \nabla_x \tilde{h}(x^k)^T d & = & 0 \\ & \tilde{g}(x^k) + \nabla_x \tilde{g}(x^k)^T d & \leq & 0 \end{array}$$

Find the smallest integer $m = m^k$ that satisfies

$$\psi_\infty(x^k + \tau^m s d^k) - \psi_\infty(x^k) \geq \sigma \tau^m s {d^k}^T d^k.$$

Define $x^{k+1} = x^k + \tau^{m_k} s d^k$, and $k = k+1$.

Go to **QP**.

TABLE 3.1
The model algorithm

$$\begin{array}{rrcl} \min_{x,\zeta} & \tilde{f}(x) + c_\infty \zeta & & \\ \text{subject to} & \tilde{g}_i(x) & \leq & 0,\ i = 1,2,\ldots,l_i, \\ & \tilde{g}_i(x) & \leq & \zeta,\ i = l_i+1,\ldots,m, \\ & \tilde{h}_j(x) & = & 0,\ j = 1,2,\ldots,l_e \\ & -\zeta \leq \tilde{h}_j(x) & \leq & \zeta,\ j = l_e+1,\ldots,r \\ & \zeta & \geq & 0. \end{array} \tag{3.2}$$

An alternative elastic mode strategy consists of using an L_1 approach [14, 22]. The modified nonlinear program becomes

$$\begin{array}{lrcl} \min_{x,u,v,w} & \tilde{f}(x) + c_1\left(e_{m-l_i}^T u + e_{r-l_e}^T (v+w)\right) & & \\ \text{subject to} & \tilde{g}_i(x) & \leq & 0,\ i = 1,2,\ldots,l_i, \\ & \tilde{g}_i(x) & \leq & u_i,\ i = l_i+1,\ldots,m, \\ & \tilde{h}_j(x) & = & 0,\ j = 1,2,\ldots,l_e \\ & -v_j \leq \tilde{h}_j(x) & \leq & w_j,\ j = l_e+1,\ldots,r \\ & u,v,w & \geq & 0, \end{array} \tag{3.3}$$

where e_{m-l_i} and e_{r-l_e} are vectors whose entries are all ones, of dimension $m - l_i$ and $r - l_e$, respectively. We call c_∞ and c_1 the penalty parameters.

All the constraints are now inequality constraints. A quadratic program analogous to (3.1) is constructed for (3.2) or (3.3), which, since **[B1]** and **[B2]** hold, now satisfies (MFCQ) at any feasible point. A feasible point of (3.2) or (3.3), respectively, can be immediately obtained by choosing ζ and u, v, w, respectively, to be sufficiently large.

We will make specific claims about one algorithm, presented in Table 3.1. The algorithm is not necessarily practical, but it serves to show that rates of convergence results can be obtained under very general assumptions. We now define the algorithm for the general nonlinear program (1.1), though we later applied it to (3.2) and (3.3).

For fixed penalty parameter c_1, the problem (3.3) can be approached by the above SQP algorithms without resulting in an infeasible QP, since the linearization of the problem (3.3) is always feasible if **[B1]** and **[B2]** hold. If for a solution of (3.3) we have that u, v, w are all equal to zero then the x component of the solution of (3.3) is also a solution of the original, unrelaxed nonlinear program (1.1).

The possibility exists, however, that c_1 may have to be increased indefinitely before a solution of (1.1) is obtained. In the following theorem we discuss sufficient conditions that ensure that the elastic mode relaxations (3.2) and (3.3) have x^* as a component of the solution for sufficiently large but finite penalty parameter.

THEOREM 3.1. *Assume that, at a solution x^* of (1.1), we have that*
- *the Lagrange multiplier set of (1.1) is not empty,*
- *the quadratic growth condition (1.16) is satisfied at x^*, and*
- *the data of (1.1) are twice continuously differentiable.*

Then,
1. *For sufficiently large but finite values of the penalty parameter c_∞ and, respectively, c_1, we have that the points $(x^*, 0)$ and, respectively, $(x^*, 0_{m-l_i}, 0_{r-l_e}, 0_{r-l_e})$, are local minima of (3.2) and (3.3) at which both (MFCQ) and the quadratic growth condition (1.16) are satisfied.*
2. *For the same values c_∞ and, respectively, c_1 we have that the points $(x^*, 0)$ and $(x^*, 0_{m-l_i}, 0_{n-l_e}, 0_{n-l_e})$ are isolated stationary points of (3.2) and (3.3).*
3. *For the same values c_∞ and, respectively, c_1, if the algorithm in Table 3.1 is applied to (3.2) and, respectively, (3.3), and is initialized sufficiently close to $(x^*, 0)$ and, respectively, $(x^*, 0_{m-l_i}, 0_{r-l_e}, 0_{r-l_e})$, with sufficiently large penalty parameter $\tilde{c}_\infty$, then the sequence x^k of iterates converges R-linearly.*

Proof We will prove part 1 of the Theorem only for the L_∞ case, the L_1 case following in the same manner. We define the fully relaxed nonlinear program

$$
\begin{array}{rrcl}
\min_{x,\zeta} & \tilde{f}(x) + c_\infty \zeta & & \\
\text{subject to} & \tilde{g}_i(x) & \leq & \zeta, \ i = 1, 2, \ldots, m, \\
& -\zeta \leq \tilde{h}_j(x) & \leq & \zeta, \ j = 1, 2, \ldots, r \\
& \zeta & \geq & 0.
\end{array}
\tag{3.4}
$$

If (x, ζ) is a feasible point of (3.4), it immediately follows from the definition (1.14) of the L_∞ penalty function, $\tilde{P}_\infty(x)$, that $\zeta \geq \tilde{P}_\infty(x)$. From (1.19), under the assumptions stated in this Theorem, we have that there exists $\tilde{c}_\infty > 0$ such that the penalty function $\psi_\infty(x)$ satisfies a quadratic growth condition at x^*. Choose now

$$c_\infty = \tilde{c}_\infty + 1.$$

Using (1.19), we obtain that, in a sufficiently small neighborhood of x^*, we must have that:

$$\tilde{f}(x) + \tilde{c}_\infty \zeta \geq \tilde{f}(x) + \tilde{c}_\infty \tilde{P}_\infty(x) \geq \sigma_1 \left|\left|x - x^*\right|\right|^2.$$

Whenever $\zeta \leq \frac{1}{\sigma_1}$, we will have that $\sigma_1 \zeta^2 \leq \zeta$. Therefore, in a sufficiently small neighborhood of $(x^*, 0)$, for all (x, ζ) feasible, we will have that:

$$\tilde{f}(x) + c_\infty \zeta = \tilde{f}(x) + \tilde{c}_\infty \zeta + \zeta \geq \sigma_1 \left(\left|\left|x - x^*\right|\right|^2 + \zeta^2 \right).$$

Therefore, for our choice of c_∞ we have that (3.4) satisfies the quadratic growth condition for feasible points (x, ζ). Since any feasible point of (3.2) is feasible for (3.4), if follows that (3.2) also satisfies quadratic growth at $(x^*, 0)$ for every feasible point (x, ζ). Since (3.2) clearly satisfies (MFCQ) everywhere from assumption **[B2]**, this is equivalent to the quadratic growth condition (1.16) holding for all (x, ζ) in a neighborhood of $(x^*, 0)$ [6, 7]. The proof of part 1 of the theorem is complete.

From the conclusion of part 1 we have that, since (MFCQ) and the quadratic growth condition holds for (3.2) and, respectively, (3.3), at $(x^*, 0)$ and, respectively,

$(x^*, 0_{m-l_i}, 0_{n-l_e}, 0_{n-l_e})$, these points must be isolated stationary points of the respective nonlinear programs [1]. This concludes the proof of part 2.

Part 3 immediately follows from [1] since (3.2) and, respectively, (3.3) satisfy the quadratic growth condition and (MFCQ) at $(x^*, 0)$ and, respectively, $(x^*, 0_{m-l_i}, 0_{n-l_e}, 0_{n-l_e})$. Note that $\tilde{c}_\infty$ that enters the definition of ψ_∞ used in the algorithm is not the same as the one in the proof of part 1. It can be shown that now we need $\tilde{c}_\infty \geq c_\infty$, once the latter is chosen. $\diamond$

Discussion

- Note that the conditions used in the Theorem are fairly weak. The quadratic growth is the weakest possible second-order sufficient condition. Relaxing our Lagrange multiplier requirement would result in a problem with an empty Lagrange multiplier set, for which few regularity results are known that could be algorithmically useful.
- Determining that a solution point is an isolated stationary point is an important issue in nonlinear programming [35, 13, 7]. In practical terms, it means that a nonlinear programming algorithm with global convergence safeguards that does not leave a neighborhood of the solution point x^* will in fact converge to x^*. Example of such algorithms are provided in the references [8, 9, 10, 12, 26, 15].
- A difficulty with the definition of the algorithm is that the successful completion of the algorithm depends on the choice of the parameters c_∞, $\tilde{c}_\infty$, c_1, that need to be sufficiently large but finite. Note that here c_∞ and $\tilde{c}_\infty$ have different purposes: c_∞ is needed to enforce $\zeta = 0$ at a solution of (3.2), whereas $\tilde{c}_\infty$ is the parameter of the merit function $\tilde{\psi}_\infty(x)$ when the preceding algorithm is applied to (3.2) or (3.3). The two are related, but we need at least $\tilde{c}_\infty > c_\infty$ if we follow the proof of linear convergence from [1].
  If (MFCQ) does not hold, the usual update [2, 4], that depends on the Lagrange multipliers, for the value of the penalty parameter cannot be used to adapt c_∞, if the original value is insufficient to result in $\zeta = 0$. An adaptive elastic mode can be implemented where if ζ is not sufficiently small, then the penalty parameters are increased [22]. In Section 4 we show that such an update works under certain conditions.

We now apply Theorem 3.1 for the case of interest in this work, MPCC. The following corollary is a simple restatement of Theorem 3.1 for (MPCC).

COROLLARY 3.2. *Assume that (MPCC) satisfies the following conditions, at a solution x^*:*

- *The Lagrange multiplier set of (MPCC) not empty. From Theorem 2.2, (SMFCQ) holding for (TNLP) is a sufficient condition for this assumption to hold.*
- *The quadratic growth condition (1.16) is satisfied at x^*.*
- *The data of (MPCC) are twice continuously differentiable.*

Then the conclusions of Theorem 3.1 hold for (MPCC) and for (2.17).

Note The part of this result that leads to of Theorem 3.1 **i)** holding for (MPCC) under the stated conditions is related to [36, Theorem 8]. Note, however, that the cited result is rather an analysis tool since it involves a nonlinear program whose setup requires the knowledge of the various active sets at the solution. Such an information is not available in an algorithmic framework, unless we are very near convergence. In the approach that we present in this work, the elastic mode is applied directly to the original formulation either in (MPCC) or (2.17), and no further knowledge of the

Table 3.2
Results obtained with MINOS

Problem	Var-Con-CC	Value	Status	Feval	Infeas
gnash14	21-13-1	-0.17904	Optimal	80	0.0
gnash15	21-13-1	-354.699	Infeasible	236	7.1E0
gnash16	21-13-1	-241.441	Infeasible	272	1.0E1
gnash17	21-13-1	-90.7491	Infeasible	439	5.3E0
gne	16-17-10	0	Infeasible	259	2.6E1
pack-rig1-8	89-76-1	0.721818	Optimal	220	0.0E0
pack-rig1-16	401-326-1	0.742102	Optimal	1460	0.0E0
pack-rig1-32	1697-1354-1	N/A	Interrupted	N/A	N/A

Table 3.3
Results obtained with SNOPT

Problem	Var-Con-CC	Value	Status	Feval	Elastic,Why
gnash14	21-13-1	-0.17904	Optimal	27	Yes, Inf
gnash15	21-13-1	-354.699	Optimal	12	None
gnash16	21-13-1	-241.441	Optimal	7	None
gnash17	21-13-1	-90.7491	Optimal	9	None
gne	16-17-10	0	Optimal	10	Yes, Inf
pack-rig1-8	89-76-1	0.721818	Optimal	15	None
pack-rig1-16	401-326-1	0.742102	Optimal	21	None
pack-rig1-32	1697-1354-1	0.751564	Optimal	19	None

problem is required.

Proof From Theorem 3.1 the conclusion immediately applies for (MPCC). For (2.17) we apply Theorem 2.5 followed by Theorem 3.1 to obtain the conclusion. $\diamond$

Consequently, when started sufficiently close to a solution and with a sufficiently large penalty parameter, the algorithm will converge to that solution of (MPCC) or (2.17) with a sufficiently large but finite c_∞ and $\tilde{c}_\infty$ as soon as (MPCC) satisfies the quadratic growth condition and has a nonempty Lagrange multiplier set at a solution x^*. Since (SMFCQ) is a generic condition for (MPCC) and holds with probability 1 for instances of problems in the MPCC class [36] and the quadratic growth condition is the weakest second-order sufficient condition, this convergence property is expected to hold with probability 1.

3.1. Numerical Experiments. We conducted some numerical experiments on MPCCs from the collection MacMPEC of Sven Leyffer. To validate the conclusions of this work, we used two widely employed nonlinear solvers MINOS [31] and SNOPT [22]. SNOPT implements an adaptive L_1 elastic mode approach.

We considered three types of problem, all of which appear in [33]

1. Stackelberg games [33, Section 12.1], which characterize market complementarity problems in which one of the players has a temporal advantage over the others. In our table these are the `gnash` problems.
2. Generalized Nash complementarity points [33, Section 12.2]. In our table this is the `gne` problem, an instance of the problem 12.34 in [33]. This is a restricted market complementarity problem.
3. Optimum packaging problem. The problem involves designing the support

of a membrane such that the area of contact between the membrane and a specified rigid obstacle is minimized, subject to the constraint that a certain region must be in contact [33, Chapter 10]. The underlying variational inequality is defined by a two-dimensional elliptic operator, which is discretized on a grid of 8×8, 16×16, and 32×32 elements, which are the problems `pack-rig` followed by the discretization index in our table.

With the exception of `gne`, all the problems have the complementarity constraints lumped together as one inequality, as in the formulation (2.17).

In the tables showing the results for MINOS and SNOPT, we indicate the number of variables, constraints, and complementarity constraints ("Var-Con-CC" in the first column), the final value of the objective function, the number of function evaluations and the final status of the run. The runs for both MINOS and SNOPT were done on the NEOS server [32] at Argonne National Laboratory. For SNOPT, we also indicate if the elastic mode was started and in what conditions. SNOPT can initialize the elastic mode in two circumstances: if the subproblem is infeasible or if the Lagrange multipliers exceed some user-specified quantity. In both examples for which the elastic mode was started, this happened because of an infeasible subproblem, a fact that we indicated in the *Elastic, Why* column of Table 3.3 by an "Inf" symbol. All the runs except one completed: the exception was `pack-rigid-32`, in MINOS, which we were forced to interrupt after it had been running on the World Wide Web interface of NEOS for about 8 hours.

The fact that (MPCC) does not satisfy (MFCQ) does not immediately result in the algorithm's running into an infeasible QP and failure. But it suggests a significant expectation that this would occur. Indeed, it can be seen that MINOS fails in more than half of the instances of MPCCs with an "infeasible" message and a large value of the measure of infeasibility. SNOPT, by contrast, solves all the problems presented in a reasonable number of iterations, needing to initiate the elastic mode for two problems after encountering elastic subproblems as shown in Table 3.3.

MINOS is a widely-used augmented Lagrangian approach whereas SNOPT is an SQP algorithm that implements an adaptive elastic mode approach. We do not use this comparison to demonstrate that the elastic mode approach is what makes MPCC solvable with an SQP approach. Rather, we have used MINOS to show that otherwise robust NLP solvers can have unpredictable failures on problems from the MPCC class and that the difficulty of MPCC is not trivial. Also, the fact that the elastic mode was initiated for SNOPT shows that the use of the elastic mode considerably increases the robustness of sequential quadratic programs, since otherwise SNOPT would have failed with an *Infeasible* diagnostic. The elastic mode approach is guaranteed to succeed for a finite penalty parameter under the conditions discussed in this paper.

4. A superlinearily convergent algorithm for MPCC. In the following, we present a superlinear convergence result for a special but widely-encountered type of algorithm that uses exact second derivatives. Here we extend the work in [16], in that we relax one assumption that was critical to the convergence proof: that the iterates satisfy either the complementarity constraints or the feasibility conditions for the unrelaxed quadratic program. We will show that in certain distinguished cases, an adaptive elastic mode approach can be used to induce either feasibility or complementarity for all iterates from which superlinear convergence follows from [16].

We assume that the complementary variables are the last components of the unknown vector x. This is not a restrictive assumption: any MPCC can be recast in such a form by using slack variables [16].

Consider the MPCC:

$$
\text{(4.1)} \qquad \text{(MPCCS)} \qquad
\begin{array}{llcll}
\text{minimize} & f(x) & & & \\
\text{subject to} & g_i(x) & \leq & 0 & i = 1,2,\ldots,n_i \\
 & h_j(x) & = & 0 & j = 1,2,\ldots,n_j \\
 & x_{k1} & \leq & 0 & k = 1,2,\ldots,n_c \\
 & x_{k2} & \leq & 0 & k = 1,2,\ldots,n_c \\
 & \sum_{k=1}^{n_c} x_{1k}x_{2k} & \leq & 0. &
\end{array}
$$

To simplify the subsequent notation, we assume that we are interested in local convergence to a solution x^* that satisfies:

$$
x^*_{k1} = 0, k = 1,2,\ldots,n_c, \quad x^*_{k2} = 0, k = 1,2,\ldots,n_d, \quad x^*_{k2} < 0, k = n_{d+1},\ldots,n_c. \tag{4.2}
$$

We also denote by $x_1 = (x_{11}, x_{21}, \ldots, x_{n_c1})$ and by $x_2 = (x_{12}, x_{22}, \ldots, x_{n_c2})$. With this notation, the complementarity constraint becomes $x_1^T x_2 = 0$, or, equivalently over the feasible set of (MPCCS), $x_1^T x_2 \leq 0$.

Note that here we use a more specialized notation convention for the various index sets, same as in Fletcher and al. [16], that is related to the objects defined in (1.22)–(1.26), for which the notation is the same as in [36]. This is because we will invoke at times results from [16]. The connection between the two notations, however, can be immediately determined. We have the following relations that tie in the two notations.

$$
\begin{aligned}
&\mathcal{I}(x^*) = \{(n_d+1,1),(n_d+2,1),\ldots,(n_c,1)\}\,, \\
&\bar{\mathcal{I}}(x^*) = \{(n_d+1,2),(n_d+2,2),\ldots,(n_c,2)\}\,, \\
&\overline{\mathcal{D}}(x^*) = \{(1,1),(1,2),(2,1),(2,2),\ldots,(n_d,1),(n_d,2)\}\,, \\
&\mathcal{K}(x^*) = \{n_d+1, n_d+2, \ldots, n_c\}\,, \\
&\bar{\mathcal{K}}(x^*) = \{1,2,\ldots,n_d\}\,.
\end{aligned}
$$

To prove our convergence results, we will invoke stronger conditions than in our preceding sections. One such condition is MPEC-LICQ. We say that (MPCCS) satisfies MPEC-LICQ at x^*, if the associated relaxed nonlinear program satisfies the linear independence constraint qualification (LICQ) at x^*. Specifically, the condition is expressed as:

$$
\text{MPEC-LICQ:} \quad \begin{array}{l} \nabla_x g_i(x^*)|_{i\in\mathcal{A}(x^*)}\,,\ \nabla_x h_j(x^*)|_{j=1,2,\ldots,n_e}\,,\ e_{k1}|_{k=1,2,\ldots,n_c}\,, \\ e_{k2}|_{k=1,2,\ldots,n_d}\,,\ \text{ are linearly independent.} \end{array}
$$

As opposed to the preceding sections, where a symbol e was used to denote the vector of all ones, in this section we denote by $e_* \in \mathcal{R}^n$ a vector that has zeroes everywhere, except in the $*$ position, where it has a 1. We also denote by $\mathcal{A}(x^*)$ the set of inequality constraints that is active:

$$
\mathcal{A}(x^*) = \{i\,|i = 1,2,\ldots,n_i,\ g_i(x^*) = 0\}\,.
$$

For an arbitrary point x, we denote by:

$$
\mathcal{A}(x) = \{i\,|i = 1,2,\ldots,n_i,\ g_i(x) \geq 0\}\,.
$$

The associated (RNLP) is:

$$
(4.3) \qquad \text{(RNLPS)} \qquad
\begin{array}{rrcll}
\text{minimize} & f(x) & & & \\
\text{subject to} & g_i(x) & \le & 0 & i = 1, 2, \ldots, n_i \\
& h_j(x) & = & 0 & j = 1, 2, \ldots, n_j \\
& x_{k1} & \le & 0 & k = 1, 2, \ldots, n_d \\
& x_{k2} & \le & 0 & k = 1, 2, \ldots, n_d \\
& x_{k1} & = & 0 & k = n_{d+1}, \ldots, n_c.
\end{array}
$$

and, if x^* is a local solution of (MPCCS) and MPEC-LICQ holds at x^*, then x^* is a Karush-Kuhn-Tucker point of (MPCCS) [36] as well as a local solution of (RNLPS). If MPEC-LICQ holds at x^*, we have that LICQ holds for (RNLPS) at x^*, and (RNLPS) has a unique Lagrange multiplier at the solution: $(\tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})$, that satisfies

$$\tilde{\nu} \ge 0,\ \tilde{\mu}_{k1} \ge 0,\ \tilde{\mu}_{k2} \ge 0,\ k = 1, 2, \ldots, n_d$$

$$
\begin{array}{rcl}
-\nabla_x f(x^*) & = & \sum_{i \in \mathcal{A}(x^*)} \tilde{\nu}_i \nabla_x g_i(x^*) + \sum_{i=1}^{n_e} \nabla_x h_j(x^*) \tilde{\pi}_j \\
& + & \sum_{k=1}^{n_d} (\tilde{\mu}_{k1} e_{k1} + \tilde{\mu}_{k2} e_{k2}) + \sum_{k=n_d+1}^{n_c} \tilde{\mu}_{k1} e_{k1}.
\end{array}
$$

Here we use the natural implicit convention that $\tilde{\nu}_i = 0$ for $i \notin \mathcal{A}(x^*)$, which we will use at other points since this way we can treat the multiplier as a vector that satisfies the complementarity constraints corresponding to the optimality conditions. However, at times we will refer to $\tilde{\nu}$ or similar quantities as to a vector with n_i components, especially in a solution stability context.

Associated to the MPEC-LICQ assumption, we can define strong second-order sufficient conditions.

$(MPEC - SOSC)$ LICQ holds at x^* and $s^T \nabla^2_{xx} \mathcal{L}^* s > 0$, $\forall s \in \mathcal{C}_{RNLPS}$ where $\nabla^2_{xx} \mathcal{L}^*$ is the Hessian of the Lagrangian evaluated at $(x^*, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})$ and $\mathcal{C}_{RNLPS}$ is the critical cone of (RNLPS) at x^*.

In the rest of this work we invoke the following assumptions,

- **[A1]** f, g, h are twice continuously differentiable.
- **[A2]** (MPCCS) satisfies MPEC-LICQ at the solution x^*.
- **[A3]** (MPCCS) satisfies MPEC-SOSC at the solution x^*.
- **[A4]** $\tilde{\nu}_i > 0$, $i \in \mathcal{A}(x^*)$, $\pi_j \neq 0$, $j = 1, 2, \ldots, n_e$, and either $\tilde{\mu}_{k1} > 0$ or $\tilde{\mu}_{k2} > 0$ for $k = 1, 2, \ldots, n_d$.
- **[A5]** When a QP is solved, the QP solver picks a linearly independent basis.

In relation to assumption **[A4]**, if $\tilde{\mu}_{k1} > 0$, for k one of $1, 2, \ldots, n_d$, we say that the constraint $x_{k1} \le 0$ is a constraint in a degenerate pair that is strongly active (at x^*). Similarly, if $\tilde{\mu}_{k2} > 0$, for k one of $1, 2, \ldots, n_d$, we say that the constraint $x_{k2} \le 0$ is a constraint in a degenerate pair that is strongly active. In these terms, the portion of assumption **[A4]** referring to degenerate pairs can be stated as: at least one constraint from each degenerative pair is strongly active.

The assumptions **[A1]**–**[A5]** are identical to the assumptions with the same name used in [16]. Note, however, that at no point in this work we will invoke either Assumption **[A6]** (that the current point satisfies the complementarity constraint $x_1^T x_2 = 0$) or Assumption **[A7]** (that all subproblems are feasible) from the same reference. As we later show, for suitable values of the penalty parameter, the elastic mode approach that we present here will induce one of the situations covered by these two assumptions, without the need to assume it from the outset.

We now explore in more details the relationship between the Lagrange multipliers and the second-order conditions of (RNLPS), (MPCCS), and of an elastic mode relaxation of (MPCCS).

In the following analysis, an important role will be played by the fact that we will assume that the elastic mode implementation enforces the linear constraints exactly. To that end, we assume that:

$$g_i(x),\ i=1,2,\ldots,l_i, \quad h_j(x),\ j=1,2,\ldots,l_e \text{ are linear functions.}$$

This is not a restrictive assumption: it is only a bookkeeping convention, since we allow l_i and l_e to be 0. Also note, following from our assumption [**A2**], that MPEC-LICQ holds at x^*, that we have that the assumptions [**B1**] and [**B2**] that were used for the elastic mode defined in Section 3, will automatically hold here.

We now present an L_∞ elastic mode approach. An L_1 approach with similar properties can be defined, exactly as used in SNOPT [22]. We use the L_∞ approach because of notational convenience: only one extra variable is needed. In this case, the relaxed MPCC becomes:

$$(MPCC(c)) \qquad \begin{array}{rcccl} \text{minimize} & f(x)+c_\infty\zeta & & & \\ \text{subject to} & g_i(x) & \leq & 0 & i=1,2,\ldots,l_i \\ & g_i(x) & \leq & \zeta & i=l_i+1,\ldots,n_i \\ & h_j(x) & = & 0 & j=1,2,\ldots,l_e \\ -\zeta \leq & h_j(x) & \leq & \zeta & j=l_e+1,\ldots,n_e \\ & x_{k1} & \leq & 0 & k=1,2,\ldots,n_c \\ & x_{k2} & \leq & 0 & k=1,2,\ldots,n_c \\ & \sum_{k=1}^{n_c} x_{k1}x_{k2} & \leq & \zeta & \\ & \zeta & \geq & 0. & \end{array}$$

When MPEC-LICQ holds, we can say more about the relationship between the multipliers of (RNLPS), (MPCCS) and (MPCC(c)), than we did in Section 2. The Lagrange multiplier set of (MPCCS) is not empty at x^*, following [36]. Therefore, there must exist ν_i, $i \in \mathcal{A}(x^*)$; π_j, $j=1,2,\ldots,n_e$; μ_{k1}, for $k=1,2,\ldots,n_c$; and μ_{k2} for $k=1,2,\ldots,n_d$; and $\eta \in \mathcal{R}$, that satisfy the KKT conditions for (MPCCS):

$$\begin{array}{rcl} -\nabla_x f(x^*) & = & \sum_{i\in\mathcal{A}(x^*)} \nu_i \nabla_x g_i(x^*) + \sum_{j=1}^{n_e} \nabla_x h_j(x^*)\pi_j + \\ & + & \sum_{k=1}^{n_d} (\mu_{k1}e_{k1}+\mu_{k2}e_{k2}) + \sum_{k=n_d+1}^{n_c} \mu_{k1}e_{k1} + \eta \sum_{k=n_d+1}^{n_c} x^*_{k2}e_{k1}, \end{array}$$

as well as the inequality constraints $\nu_i \geq 0$, $i \in \mathcal{A}(x^*)$, $\mu_{k1} \geq 0$, $k=1,2,\ldots,n_c$, $\mu_{k2} \geq 0$, $k=1,2,\ldots,n_d$ and $\eta \geq 0$. Here η is the Lagrange multiplier of the complementarity constraint of (MPCCS) $x_1^T x_2 \leq 0$.

One immediate consequence is that (MPCCS) has a Lagrange multiplier that is minimal (in terms of the 1 norm) [16]. We call that multiplier the *fundamental multiplier*. Comparing the algebraic expression of its components in terms of the components of the multiplier of (RNLPS) we obtain:

$$(FMC) \quad \begin{array}{rcll} \nu^* & = & \tilde{\nu} & \\ \pi^* & = & \tilde{\pi} & \\ \mu^*_{k1} & = & \tilde{\mu}_{k1} & \geq 0, \quad k=1,2,\ldots,n_d \\ \mu^*_{k2} & = & \tilde{\mu}_{k2} & \geq 0, \quad k=1,2,\ldots,n_d \\ \mu^*_{k1} & = & \tilde{\mu}_{k1} - \eta^* x^*_{k2} & \geq 0, \quad k=n_d+1,\ldots,n_c \\ \eta^* & = & \max\left\{0, \max_{k=n_d+1,\ldots,n_c}\left\{\frac{\tilde{\mu}_{k1}}{x^*_{k2}}\right\}\right\} & \geq 0. \end{array}$$

Since LICQ holds, it is immediate that (MPCCS) has a degeneracy of order 1 and thus its Lagrange multiplier set will have dimension at most 1. Therefore any multiplier of (MPCCS) must satisfy:

$$
\begin{array}{rcll}
\nu &=& \nu^* & \\
\pi &=& \pi^* & \\
\mu_{k1} &=& \tilde{\mu}^*_{k1} & \geq 0, \quad k = 1, 2, \ldots, n_d \\
\mu_{k2} &=& \tilde{\mu}^*_{k2} & \geq 0, \quad k = 1, 2, \ldots, n_d \\
\mu_{k1} &=& \tilde{\mu}^*_{k1} - a x_{k2} & \geq 0, \quad k = n_d + 1, \ldots, n_c \\
\eta &=& \eta^* + a & \geq 0,
\end{array}
\tag{4.4}
$$

where $a \geq 0$.

We now write the KKT conditions for the relaxed problem (MPCC(c)) at $(x^*, 0)$, assuming that $(x^*, 0)$ is a stationary point of (MPCC(c)), an assumption for which we will later determine sufficient conditions for it to hold. We have that, for a Lagrange multiplier of MPCC(c) with components $\hat{\nu}_i$, $i \in \mathcal{A}(x^*)$; $\hat{\pi}_j$, for $j = 1, 2, \ldots, l_e$; $\hat{\pi}^+_j$, associated with the inequality $h_j(x) \leq \zeta$ and $\hat{\pi}^-_j$, associated with the inequality $-h_j(x) \leq \zeta$, $j = l_e + 1, \ldots, n_e$; $\hat{\mu}_{k1}$ for $k = 1, 2, \ldots, n_c$; $\hat{\mu}_{k2}$ for $k = 1, 2, \ldots, n_d$; and $\hat{\eta}$, associated with the complementarity constraint $x_1^T x_2 \leq \zeta$; and $\hat{\theta} \geq 0$, associated with the inequality $\zeta \geq 0$, we can write the KKT conditions:

$$
\begin{array}{rcl}
-\nabla_x f(x^*) &=& \sum_{i \in \mathcal{A}(x^*)} \hat{\nu}_i \nabla_x g_i(x^*) + \sum_{j=1}^{l_e} \nabla_x h_j(x^*) \hat{\pi}_j \\
&+& \sum_{j=l_e+1}^{n_e} \nabla_x h_j(x^*) \left(\hat{\pi}^+_j - \hat{\pi}^-_j \right) + \sum_{k=1}^{n_d} \left(\hat{\mu}_{k1} e_{k1} + \hat{\mu}_{k2} e_{k2} \right) \\
&+& \sum_{k=n_d+1}^{n_c} \hat{\mu}_{k1} e_{k1} + \hat{\eta} \sum_{k=n_d+1}^{n_c} \left(x^*_{k2} e_{k1} \right) \\
c_\infty &=& \sum_{i=l_i+1, \ldots, n_i,\ i \in \mathcal{A}(x^*)} \hat{\nu}_i + \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}^+_j + \hat{\pi}^-_j \right) + \hat{\eta} + \hat{\theta}.
\end{array}
$$

If MPEC-LICQ holds, and $(x^*, 0)$ is a stationary point of (MPCC(c)), then, using (4.4), we obtained that any Lagrange multiplier of (MPCC(c)) must satisfy, in terms of the components of the fundamental multiplier, the following relations:

$$
\begin{array}{rclll}
\hat{\nu}_i &=& \nu^*_i & \geq 0 & i \in \mathcal{A}(x^*) \\
\hat{\pi}_j &=& \pi^*_j & & j = 1, 2, \ldots, l_e, \\
\hat{\pi}^+_j &=& \max\{\pi^*_j, 0\} + f_j & \geq 0 & j = l_e + 1, \ldots, n_e \\
\hat{\pi}^-_j &=& \max\{-\pi^*_j, 0\} + f_j & \geq 0 & j = l_e + 1, \ldots, n_e \\
\hat{\mu}_{k1} &=& \mu^*_{k1} & \geq 0 & k = 1, 2, \ldots, n_d \\
\hat{\mu}_{k2} &=& \mu^*_{k2} & \geq 0 & k = 1, 2, \ldots, n_d \\
\hat{\mu}_{k1} &=& \mu^*_{k1} - a x_{k2} & \geq 0 & k = n_d + 1, \ldots, n_c \\
\hat{\eta} &=& \eta^* + a & \geq 0 & \\
\hat{\theta} &=& c_\infty - \sum_{i=l_{i+1}, i \in \mathcal{A}(x^*)}^{n_i} \hat{\nu}^*_i & & \\
&-& \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}^+_j + \hat{\pi}^-_j \right) - \hat{\eta} & \geq 0, &
\end{array}
\tag{4.5}
$$

where $a \geq 0$, and $f_j \geq 0$, for $j = l_e + 1, \ldots, n_e$. Here $\hat{\theta}$ is the Lagrange multiplier of the constraint $\zeta \geq 0$ and the requirement $\hat{\theta} \geq 0$ results in a condition on the value that c_∞ must take for $(x^*, 0)$ to be a stationary point of (MPCC(c)). We do not add the complementarity constraints that appear when applying the first order optimality condition, since these are automatically satisfied by our choice of active set and multiplier components.

Rewriting $\hat{\theta}$ in terms of the fundamental multiplier (FMC), we obtain that:

$$
\hat{\theta} = c_\infty - \sum_{i=l_{i+1}, i \in \mathcal{A}(x^*)}^{n_i} \nu^*_i - \sum_{j=l_e+1}^{n_e} \left(|\pi^*_j| + 2 f_j \right) - \eta^* - a \geq 0.
\tag{4.6}
$$

We denote by ν_0 the following quantity, that is also defined in terms of the components of the fundamental multiplier (FMC):

$$\nu_0 = \sum_{i=l_{i+1}, i\in\mathcal{A}(x^*)}^{n_i} \nu_i^* + \sum_{j=l_e+1}^{n_e} |\pi_j^*| + \eta^*. \tag{4.7}$$

Since the values of the fundamental multiplier are fixed, and since $a \geq 0$ and $f_j \geq 0$, $j = l_e+1, \ldots, n_e$, the condition for $(x^*, 0)$ to be a stationary point of (MPCC(c)) is:

$$c_\infty \geq \sum_{i=l_{i+1}, i\in\mathcal{A}(x^*)}^{n_i} \nu_i^* + \sum_{j=l_e+1}^{n_e} \left|\pi_j^*\right| + \eta^* = \nu_0.$$

The set of Lagrange multipliers of (MPCC(c)) is defined by $f_j \geq 0$, $j = l_e+1, \ldots, n_e$ and $a \geq 0$ that satisfy the inequality $\hat{\theta} \geq 0$, that is,

$$c_\infty - \nu_0 \geq \sum_{j=l_e+1}^{n_e} (2f_j) + a. \tag{4.8}$$

An immediate consequence is that, if $c_\infty = \nu_0$, then the program (MPCC(c)) has a unique multiplier! This is formally stated in the next result.

4.1. Second-order conditions. LEMMA 4.1. *Assume that (MPCCS) satisfies MPEC-LICQ and MPEC-SOSC. Assume that c_∞ satisfies $c_\infty \geq \nu_0$. Then (MPCC(c)) satisfies (MFCQ) and (RSOSC) at $(x^*, 0)$. In addition, if $c_\infty = \nu_0$, then the Lagrange multiplier set of (MPCC(c)) at $(x^*, 0)$ has a unique element.*

Proof Consider one multiplier of (MPCC(c)), whose components are: $\hat{\nu_i} \geq 0$ for $i \in \mathcal{A}(x^*)$; $\hat{\pi}_j$ for $j = 1, 2, \ldots, l_e$; $\hat{\pi}_j^+ \geq 0$ and $\hat{\pi}_j^- \geq 0$ for $j = l_e+1, \ldots, n_e$; $\hat{\mu}_{k1} \geq 0$ and $\hat{\mu}_{k2} \geq 0$ for $k = 1, 2, \ldots, n_d$; $\hat{\mu}_{k1} \geq 0$ for $k = 1, 2, \ldots, n_d$; $\hat{\eta} \geq 0$ and $\hat{\theta} \geq 0$.

The Hessian of the Lagrangian at x^* corresponding to this Lagrange Multiplier, accounting for only the nonlinear terms, and using (FMC) and (4.5), is the following:

$$\begin{array}{rcl}
\nabla^2_{xx}\mathcal{L}_{MPCC(c)}(x^*, 0) & = & \nabla^2_{xx} f(x^*) + \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}_j^+ \nabla^2_{xx} h_j(x^*) - \hat{\pi}_j^- \nabla^2_{xx} h_j(x^*)\right) \\
& + & \sum_{i=l_i+1, i\in\mathcal{A}(x^*)}^{n_i} \hat{\nu}_i \nabla^2_{xx} g(x^*) + \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k_1}^T e_{k2} + e_{k2}^T e_{k1}\right) \\
& = & \nabla^2_{xx} f(x^*) + \sum_{j=l_e+1}^{n_e} \pi_j^* \nabla^2_{xx} h_j(x^*) \\
& + & \sum_{i=l_i+1, i\in\mathcal{A}(x^*)}^{n_i} \nu_i^* \nabla^2_{xx} g(x^*) + \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k_1}^T e_{k2} + e_{k2}^T e_{k1}\right) \\
& = & \nabla^2_{xx}\mathcal{L}_{RNLPS}(x^*) + \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k_1}^T e_{k2} + e_{k2}^T e_{k1}\right).
\end{array}$$

The Hessian of the Lagrangian of (MPCC(c)) should also be computed with respect to ζ, but, since the contribution of ζ is linear both in the constraints and in the objective function of (MPCC(c)), it follows that the Hessian of the Lagrangian is:

$$\nabla^2_{(x,\zeta)(x,\zeta)}\mathcal{L}_{MPCC(c)}(x^*, 0) = \begin{pmatrix} \nabla^2_{xx}\mathcal{L}_{MPCC(c)}(x^*, 0) & 0_{n\times 1} \\ 0_{1\times n} & 0 \end{pmatrix}.$$

Note that, from Lemma 2.1, we have that $\mathcal{C}_{MPCCS} = \mathcal{C}_{RNLPS}$ at x^*, where by $\mathcal{C}$ we denote the critical cone of the respective nonlinear program. When $c_\infty \geq \nu_0$, since x^* is a stationary point for (MPCC) and $(x^*, 0)$ is a stationary point for (MPCC(c)) we will have that $\mathcal{C}_{MPCC(c)} = \mathcal{C}_{MPCCS} \bigoplus \{0\}$. Let now $(u, 0) \in \mathcal{C}_{MPCC(c)}$ and thus $u \in \mathcal{C}_{RNLPS}$. We have that:

$$\begin{array}{rcl}
(u, 0)^T \nabla^2_{(x,\zeta)(x,\zeta)}\mathcal{L}_{MPCC(c)}(u, 0) & = & u^T \mathcal{L}_{MPCC(c)} u = u^T \mathcal{L}_{RNLPS} u \\
& + & 2\hat{\eta} \sum_{k=1}^{n_c} (e_{k1}^T u)(e_{k2}^T u) \geq u^T \mathcal{L}_{RNLPS} u > 0,
\end{array}$$

where the last two inequalities follow, respectively, from the fact that, on the critical cone of (RNLPS) we have $(e_{k1}^T u)(e_{k2}^T u) \geq 0$, for $k = 1, 2, \ldots, n_d$, and $(e_{k1}^T u)(e_{k2}^T u) = 0$, for $k = n_d + 1, \ldots, n_c$, and $\hat{\eta} \geq 0$, and, respectively, from the MPEC-SOSC assumption. Since we obtained, for any multiplier of MPCC(c), that:

$$\forall \hat{u} \in \mathcal{C}_{MPCC(c)} \quad \Rightarrow \quad \hat{u}^T \nabla^2_{(x,\zeta)(x,\zeta)} \mathcal{L}_{MPCC(c)}(x^*, 0)\hat{u} > 0,$$

this means that Robinson's condition (RSOSC) holds for (MPCC(c)) at $(x^*, 0)$. Since (MFCQ) clearly holds at $(x^*, 0)$, the conclusion of the lemma follows. The uniqueness of the multiplier for the case $c_\infty = \nu_0$ follows from (4.8). $\diamond$

4.2. The algorithm. We now consider the effect of applying SQP to MPCC in either the original or the relaxed form. As matrix W of the QP we use the Hessian of the Lagrangian, with the Lagrange multipliers computed at the previous step.

We define the following matrix $\widehat{W}$ at the point x:

$$\widehat{W} = \begin{cases} \nabla^2_{xx} f(x) & + \sum_{i=l_i+1}^{n_i} \hat{\nu}_i \nabla^2_{xx} g_i(x) \\ & + \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}_j^+ - \hat{\pi}_j^-\right) \nabla^2_{xx} h_j(x) \\ & + \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k1}^T e_{k2} + e_{k2}^T e_{k1}\right). \\ \text{or, alternatively,} & \\ \nabla^2_{xx} f(x) & + \sum_{i=l_i+1}^{n_i} \nu_i \nabla^2_{xx} g_i(x) \\ & + \sum_{j=l_e+1}^{n_e} \pi_j \nabla^2_{xx} h_j(x) \\ & + \eta \sum_{k=1}^{n_c} \left(e_{k1}^T e_{k2} + e_{k2}^T e_{k1}\right), \end{cases}$$

where the first branch is followed in the quadratic program associated with the relaxation (QPMC(c)) was solved at the preceding iteration, and the second branch is followed if the quadratic program associated with the unrelaxed problem ($QP(x)$, to be defined later) was solved at the preceding iteration.

Consider now the following quadratic program associated to (MPCC(c)) at point (x, ζ).

$$(QPMC(c)) \begin{cases} \min_{d,d_\zeta} & \frac{1}{2} d^T \widehat{W} d + \nabla_x f(x) d + c_\infty(\zeta + d_\zeta) & & & \\ & g_i(x) + \nabla_x g_i(x) d & \leq & 0 & i = 1, 2, \ldots, l_i \\ & g_i(x) + \nabla_x g_i(x) d & \leq & \zeta + d_\zeta & i = l_i + 1, \ldots, n_i \\ & h_j(x) + \nabla_x h_j(x) d & = & \zeta + d_\zeta & j = 1, 2, \ldots, l_e \\ & h_j(x) + \nabla_x h_j(x_k) d & \leq & \zeta + d_\zeta & j = l_{e+1}, \ldots, n_e \\ \text{sbj. to} & -h_j(x) - \nabla_x h_j(x_k) d & \leq & \zeta + d_\zeta & j = l_{e+1}, \ldots, n_e \\ & x_{k1} + d_{k1} & \leq & 0 & k = 1, 2, \ldots, n_c \\ & x_{k2} + d_{k2} & \leq & 0 & k = 1, 2, \ldots, n_c \\ & x_1^T d_2 + x_2^T d_1 + x_1^T x_2 & \leq & \zeta + d_\zeta & \\ & \zeta + d_\zeta & \geq & 0. & \end{cases}$$

To obtain results for the application of the algorithm to (MPCCS), we state and prove the following stability result for general nonlinear programming.

LEMMA 4.2. *Consider the nonlinear program:*

$$\begin{array}{lrcl} \mathit{min}_x & \tilde{f}(x) & & \\ \mathit{subject\ to} & \tilde{h}_j(x) & = & 0, j = 1, 2, \ldots, r \\ & \tilde{g}_i(x) & \leq & 0, i = 1, 2, \ldots, m. \end{array}$$

Assume that it satisfies (MFCQ) and (RSOSC) at a solution x^*. We denote the compact Lagrange multiplier set of this nonlinear program by $\Lambda(x^*)$.

Consider the Quadratic Program:

$$\begin{array}{rlcl} \mathit{min}_\delta & \nabla_x \tilde{f}(x)d + \frac{1}{2}d^T\tilde{W}d & & \\ \mathit{subject\ to} & \tilde{g}_i(x) + \nabla_x\tilde{g}_i(x)d & \leq & 0,\ i=1,2,\ldots,m \\ & \tilde{h}_j(x) + \nabla_x\tilde{h}_j(x)d & = & 0,\ j=1,2,\ldots,r, \end{array}$$

where $\tilde{W} = \nabla^2_{xx}\tilde{f} + \sum_{j=1}^{n_e}\pi_j\nabla^2_{xx}\tilde{h}_j(x) + \sum_{i=1}^{n_i}\nu_i\nabla^2_{xx}\tilde{g}_i(x)$. Let d be a solution of this quadratic program, and ν^+ and π^+ its Lagrange multipliers.

Then, there exist $\epsilon > 0$, $c_1 > 0$, $c_2 > 0$ such that $\operatorname{dist}((\nu,\pi);\Lambda(x^*)) \leq \epsilon$ *and* $\operatorname{dist}(x;x^*) \leq \epsilon$ *imply that*

- **(i)** $\frac{1}{c_1}\,||x-x^*|| \leq ||d|| \leq c_2\,||x-x^*||$,
- **(ii)** $\operatorname{dist}((\nu^+,\pi^+);\Lambda(x^*)) \leq c_2\,||x-x^*||$.

Here $\operatorname{dist}(\cdot,\cdot)$ *denotes the distance between two sets.*

Proof The rightmost inequalities in both part **(i)** and **(ii)** are a consequence of [40, Theorem A1]. In that reference, the inequality-only case is treated, but the conclusion can be immediately extended for the case where there are also equality constraints, that are linearly independent and that, together with the inequality constraints, satisfy (MFCQ). Since (MFCQ) holds, we have that:

$$\begin{array}{rcl} \tilde{g}(x+d) & \leq & O(||d||^2) \leq O(||x-x^*||^2) \\ \tilde{h}(x+d) & \leq & O(||d||^2) \leq O(||x-x^*||^2). \end{array}$$

This implies that the size of the infeasibility, $\mathcal{I}(x)$, where

$$\mathcal{I}(x) = \max\left\{\max_{i=1,2,\ldots,n_i}\{\tilde{g}_i^+(x)\},\ \max_{j=1,2,\ldots,n_e}\{|\tilde{h}_j(x)|\}\right\}$$

satisfies $\mathcal{I}(x) = O(||d||)$. Using the result [40, Theorem A1], as well as $\nu^+ \geq 0$ and the optimality conditions for the QP, we obtain that, since $Wd = -\nabla_x\tilde{f}^T(x) - \nabla_x\tilde{g}^T(x)\nu^+ - \nabla_x\tilde{h}^T(x)\pi^+$, we have that:

$$\begin{array}{l} \Omega\left(||x-x^*|| + \operatorname{dist}((\nu^+,\pi^+);\Lambda(x^*))\right) = \\ \max\left\{\left|\left|\nabla_x\tilde{f}(x) + \nabla_x\tilde{g}(x)\nu^+ + \nabla_x\tilde{h}(x)\pi^+\right|\right|, \mathcal{I}(x)\right\} = O(||d||). \end{array}$$

In turn, this implies that there exists a $c_1 > 0$ such that $||x-x^*|| \leq c_1\,||d||$, which completes the proof. $\diamond$

THEOREM 4.3. *Consider the quadratic program (QPMC(c)), whose solution is (d, d_ζ). Assume that $c_\infty - \nu_0 = \gamma_0 > 0$. There exists $\epsilon > 0$ and $c_0 > 0$, such that, if*

$$\operatorname{dist}\left(\left(\hat{\nu},\hat{\pi},\hat{\pi^+},\hat{\pi^-},\hat{\mu}_1,\hat{\mu}_2,\hat{\eta},\hat{\theta}\right);\Lambda_{MPCC(c)}\right) < \epsilon \text{ and } \operatorname{dist}((x,\zeta);(x^*,0)) \leq \epsilon,$$

then

- **i)** $||d|| + ||d_\zeta|| \leq c_0(||x-x^*|| + \zeta)$.
- **ii)** $\operatorname{dist}\left(\left(\hat{\nu}^+,\hat{\pi}^+,\hat{\pi}^{+^+},\hat{\pi}^{-^+},\hat{\mu}_1^+,\hat{\mu}_2^+,\hat{\eta}^+,\hat{\theta}^+\right);\Lambda_{MPCC(c)}\right) \leq c_0(||x-x^*||+\zeta)$. *Here the first vector is a Lagrange multiplier solution of (QPMC(c)) at (x,ζ).*
- **iii)** $\left|\hat{\pi}_j^{+^+} - \hat{\pi}_j^{-^+} - \pi_j^*\right| \leq c_0(||x-x^*|| + \zeta)$, *for $j = l_e+1,\ldots,n_e$. Here π_j^* is the corresponding component of the fundamental multiplier (FMC).*

- **iv)**
  - *Either* $\zeta + d_\zeta = 0$,
  - *Or the point* $x + d$ *satisfies the complementarity constraints, and the strongly active variables that are in degenerate pairs at* x^* *are active.*

Note Here the components of the multipliers are as follows: $\hat{\nu}$ is considered to have n_i components, though only the ones corresponding $\mathcal{A}(x^*)$ may be nonzero close to the solution $(x^*, 0)$, $\hat{\pi}$ has l_e elements, $\hat{\pi}^+$ and $\hat{\pi}^-$ have $n_e - l_e$ elements, $\hat{\mu}_1$ and $\hat{\mu}_2$ have n_c elements. The last $n_c - n_d$ components of $\hat{\mu}_2$ are equal to 0, by virtue of the optimality conditions for (MPCCS), (MPCC(c)) and (4.2).

Proof Parts **i)** and **ii)** are an immediate consequence of Lemma 4.1 and Lemma 4.2. Part **iii)** follows from Part **ii)**, (FMC) and (4.4), using the relations:

$$\begin{aligned}\left|\hat{\pi}_j^{+^+} - \hat{\pi}_j^{-^+} - \pi_j^*\right| &= \left|\hat{\pi}_j^{+^+} - \hat{\pi}_j^{-^+} - \left(\max\{\pi_j^*, 0\} + f_j - \max\{-\pi_j^*, 0\} - f_j\right)\right| \\ &\leq \left|\hat{\pi}_j^{+^+} - \left(\max\{\pi_j^*, 0\} + f_j\right)\right| + \left|\hat{\pi}_j^{-^+} - \left(\max\{-\pi_j^*, 0\} + f_j\right)\right|\end{aligned}$$

for $j = l_e + 1, \ldots, n_e$.

For Part **iv)**, since $\zeta + d_\zeta \geq 0$, we need to analyze only the case where $\zeta + d_\zeta > 0$. In the latter case, we obtain that, for any j, $j = l_i + 1, \ldots, n_i$, only one of the constraints $-\zeta - d_\zeta \leq h_j(x) + \nabla_x h_j(x) d$ and $h_j(x) + \nabla_x h_j(x) d \leq \zeta + d_\zeta$ can be active and therefore only one of the multipliers $\hat{\pi}_j^{+^+} \geq 0$ and $\hat{\pi}_j^{-^+} \geq 0$ can be positive, that is,

$$\hat{\pi}_j^{+^+} \hat{\pi}_j^{-^+} = 0. \tag{4.9}$$

Using Part **ii)**, we get that there exists

$$\left(\hat{\nu}^*, \hat{\pi}^*, \hat{\pi}^{+^*}, \hat{\pi}^{-^*}, \hat{\mu}_1^*, \hat{\mu}_2^*, \hat{\eta}^*, \hat{\theta}^*\right) \in \Lambda_{MPCC(c)}$$

such that:

$$\begin{aligned} & ||\hat{\nu}^+ - \hat{\nu}^*||_1 + ||\hat{\pi}^* - \hat{\pi}||_1 + \left|\left|\hat{\pi}^{+^*} - \hat{\pi}^{+^+}\right|\right|_1 + \left|\left|\hat{\pi}^{-^*} - \hat{\pi}^{-^+}\right|\right|_1 + \left|\left|\hat{\mu}_1^* - \hat{\mu}_1^+\right|\right|_1 \\ + & \left|\left|\hat{\mu}_2^* - \hat{\mu}_2^+\right|\right|_1 + ||\hat{\eta}^+ - \hat{\eta}^*||_1 + \left|\left|\hat{\theta}^+ - \hat{\theta}^*\right|\right|_1 \leq c_0 \left(||x - x^*|| + \zeta\right). \end{aligned} \tag{4.10}$$

In particular, for $j = l_e + 1, \ldots, n_e$, we obtain from (4.10) that:

$$\left|\left|\hat{\pi}_j^{+^+} - \hat{\pi}_j^{+^*}\right|\right| + \left|\left|\hat{\pi}_j^{-^+} - \hat{\pi}_j^{-^*}\right|\right| \leq c_0 \left(||x - x^*|| + \zeta\right). \tag{4.11}$$

Note that, from (4.5) we have, for $j = l_e + 1, \ldots, n_e$, that

$$\hat{\pi}_j^{+^*} = \max\{\pi_j^*, 0\} + f_j, \quad \hat{\pi}_j^{-^*} = \max\{-\pi_j^*, 0\} + f_j,$$

for some $f_j \geq 0$. We have two cases, for each $j = l_e + 1, \ldots, n_e$. One of them can be discarded by using the first part of the assumption **[A4]**. However, to maintain the assumptions we use here at a minimum we will not invoke it at this point.

1. $\pi_j^* \neq 0$, where π_j^* is the corresponding component of the fundamental multiplier. Assume that $\pi_j^* > 0$. Then (FMC) and (4.5), together with part **ii)** implied that for ϵ sufficiently small we must have that $\hat{\pi}_j^{+^+} > 0$, and thus, using the equation (4.9) we obtain that $\hat{\pi}_j^{-^+} = 0$. Since from the last displayed equation we must have that $\hat{\pi}_j^{-^*} = f_j$, we use (4.11) to obtain that:

$$f_j \leq c_0 \left(||x - x^*|| + \zeta\right). \tag{4.12}$$

The same inequality holds for the case where $\hat{\pi}_j^* < 0$.

2. Assume now that $\pi_j^* = 0$. In this case we have that $\hat{\pi}_j^{+^*} = \hat{\pi}_j^{-^*} = f_j$. After we apply again (4.9) followed by (4.11), we obtain that (4.12) holds.

Therefore (4.12) holds in any case.

Since we were under the assumption that $\zeta + \delta\zeta > 0$, the optimality conditions of (QPMC(c)) imply that $\hat{\theta}^+ = 0$. In turn, this and (4.10) imply that $\hat{\theta}^* \leq c_0 \left(||x - x^*|| + \zeta\right)$. Using the last inequality as well as (4.6) and (4.12) we obtain that:

$$(4.13)\; a = c_\infty - \hat{\theta}^* - \nu_0 - 2 \sum_{j=l_e+1}^{n_e} f_j \geq \gamma_0 - c_0 \left(2(n_e - l_e) + 1\right) \left(||x - x^*|| + \zeta\right).$$

Using now (4.10), and the relationship between the multipliers of (MPCCS) and the fundamental multipliers (4.5), we obtain that:

$$\left|\hat{\mu}_{k1}^+ - \hat{\mu}_{k1}^*\right| = \left|\hat{\mu}_{k1}^+ - (\mu_{k1}^* - a x_{k2}^*)\right| \leq c_0 \left(||x - x^*|| + \zeta\right), \quad k = n_d + 1, \ldots, n_c.$$

This implies, using (4.13), and that, from (FMC), $\mu_{k1}^* \geq 0$ for $k = n_d + 1, \ldots, n_c$, that:

$$\begin{aligned}\hat{\mu}_{k1}^+ &\geq (\mu_{k1}^* - a x_{k2}^*) - c_0 \left(||x - x^*|| + \zeta\right) \\ &\geq -\gamma_0 x_{k2}^* - c_0 \left(||x - x^*|| + \zeta\right) \left((2(n_e - l_e) + 1)\, x_{k2}^* + 1\right).\end{aligned}$$

We clearly obtain that, if we choose $\epsilon > 0$, and thus $||x - x^*|| + \zeta$, sufficiently small then we must have $\hat{\mu}_{k1}^+ > 0$, $k = n_{d+1}, \ldots, n_c$. This means that, in (QPMC(c)), the constraints $x_{k1} + d_{k1} \leq 0$ must be active, that is, $x_{k1} + d_{k1} = 0$, for $k = n_{d+1}, \ldots, n_c$. In addition,using part **ii)** of the proof, (FMC), (4.4), Assumption **[A4]** and the complementarity relation in the optimality conditions of (QPMC(c)), we obtain that at least one of the inequalities

$$x_{k1} + d_{k1},\; x_{k2} + d_{k2}.$$

must be active for any $k = 1, 2, \ldots, n_d$, the one that corresponds to the positive fundamental multiplier. This implies that when we define $x^+ = x + d$, the complementarity constraints will be satisfied, and all constraints that are in a degenerate pair with positive multipliers will be active. This completes the proof. ◇

Note that we have used only the last part of the assumption **[A4]**, and we have not assumed anything about the strict complementarity of the multipliers of the nonlinear inequality and equality constraints of (MPCCS). In effect, in this work we need the assumption **[A4]** as stated only at the points where we invoke results from [16].

THEOREM 4.4. *Consider the quadratic program (QPMC(c)). Assume that $c_\infty - \nu_0 = \gamma_0 > 0$. There exists $\epsilon > 0$ and such that, if*

1. dist $\left((x, \zeta); (x^*, 0)\right) \leq \epsilon$,

2. dist $\left(\left(\hat{\nu}, \hat{\pi}, \hat{\pi^+}, \hat{\pi^-}, \hat{\mu}_1, \hat{\mu}_2, \hat{\eta}, \hat{\theta}\right); \Lambda_{MPCC(c)}\right) < \epsilon$, *and*

3. x satisfies the complementarity constraints, and the strongly active constraints from the degenerate pairs exactly.

Then

i) $\zeta + d_\zeta = 0$,

ii) *the strongly active constraints from degenerate pairs are active for (QPMC(c)) at x,*

iii) *$x+d$ satisfies the complementarity constraints.*

Proof

Let

$$\begin{array}{rcl} W & = & \nabla^2_{xx} f(x) + \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}_j^+ - \hat{\pi}_j^-\right) \nabla^2_{xx} h_j(x) + \sum_{i=l_i+1}^{n_i} \hat{\nu}_i \nabla^2_{xx} g_i(x) \\ & + & \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k1}^T e_{k2} + e_{k2}^T e_{k1}\right). \end{array}$$

Let

$$\overline{W} = \nabla^2_{xx} f(x) + \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}_j^+ - \hat{\pi}_j^-\right) \nabla^2_{xx} h_j(x) + \sum_{i=l_i+1}^{n_i} \hat{\nu}_i \nabla^2_{xx} g_i(x).$$

Note that, from (FMC) and (4.5), we obtain, similarly to the conclusion of Theorem 4.3 **iii)** that:

$$(4.14) \quad |\hat{\pi}_j^+ - \hat{\pi}_j^- - \pi_j^*| \le \epsilon, \;\; j = l_e+1, \ldots, n_e \quad \text{and} \quad |\hat{\nu}_i - \nu_i^*| \le \epsilon, \;\; i = l_i+1, \ldots, n_i.$$

Let

$$\mathcal{Z}_1(x) = \{k \in 1, 2, \ldots, n_c | x_{k1} = 0\}, \mathcal{Z}_2(x) = \{k \in 1, 2, \ldots, n_c | x_{k2} = 0\}.$$

Also, denote by

$$\mathcal{Z}_1(x)^\perp = \{1, 2, \ldots, n_c\} - \mathcal{Z}_1(x), \quad \mathcal{Z}_2(x)^\perp = \{1, 2, \ldots, n_c\} - \mathcal{Z}_2(x).$$

Recall, we assume that x satisfies the complementarity constraints which means that $\mathcal{Z}_1(x) \cup \mathcal{Z}_2(x) = \{1, 2, \ldots, n_c\}$. The degenerate indices are the indices k at which both complementarity variables are 0, that is $x_{k1} = x_{k2} = 0$. We denote by $\mathcal{D}(x)$ their set and we have that $\mathcal{D}(x) = \mathcal{Z}_1(x) \cap \mathcal{Z}_2(x)$. In the definition of these index sets, we use the same notation as in [16] since we will invoke key results from that reference. These index sets are similar to the ones defined in (1.20–1.26), although we will use them at points other than x^*.

Due to our assumption that the complementarity constraints hold at x, we must have that $\mathcal{Z}_1^\perp(x), \mathcal{Z}_2^\perp(x), \mathcal{D}(x)$ form a partition of $\{1, 2, \ldots, p\}$. We now examine the difference between $\mathcal{Z}_1^\perp(x^*)$ and $\mathcal{Z}_1^\perp(x)$. The following discussion is similar to the one following proposition 5.1 in [16]. We clearly must have $\mathcal{Z}_1^\perp(x^*) \subset \mathcal{Z}_1^\perp(x)$, since $\mathcal{Z}_1^\perp(x^*)$ contains only indices that are inactive at x^* and, therefore, in an entire neighborhood of x^*. Let $k \in \mathcal{Z}_1^\perp(x) \backslash \mathcal{Z}_1^\perp(x^*)$, which means that $x_{k1} < 0$ and $x_{k1}^* = 0$. Since x satisfies the complementarity constraints, we must have that $k \notin \mathcal{Z}_2^\perp(x)$ and therefore $k \notin \mathcal{Z}_2^\perp(x^*)$. This implies that k must be a degenerate index at x^*, and thus $k \in \mathcal{D}(x^*)$. We therefore have that:

$$(4.15) \; \mathcal{Z}_1(x^*) \subset \mathcal{Z}_1(x) \subset \mathcal{Z}_1(x) \cup \mathcal{D}(x^*) \text{ and } \mathcal{Z}_2(x^*) \subset \mathcal{Z}_2(x) \subset \mathcal{Z}_2(x) \cup \mathcal{D}(x^*).$$

An important consequence of this relationship is, using (4.2), that we must have $\{n_d + 1, \ldots, n_c\} \subset \mathcal{Z}_1(x)$ for all x in a neighborhood of x^* that satisfy assumption 3 from the statement of this Theorem and that $\{1, 2, \ldots, n_d\} = \mathcal{D}(x^*)$.

Consider the following unrelaxed quadratic program, which is (QPMC(c)), with the additional requirement that $\zeta + d_\zeta = 0$.

$$(QP(x)) \left\{ \begin{array}{lrcll} \min_d & \frac{1}{2} d^T W d + \nabla_x f(x) d & & & \\ & g_i(x) + \nabla_x g_i(x) d & \le & 0 & i = 1, 2, \ldots, n_i \\ & h_j(x) + \nabla_x h_j(x) d & = & 0 & j = 1, 2, \ldots, n_e \\ \text{sbj. to} & x_{k1} + d_{k1} & \le & 0 & k = 1, 2, \ldots, n_c \\ & x_{k2} + d_{k2} & \le & 0 & k = 1, 2, \ldots, n_c \\ & x_1^T d_2 + x_2^T d_1 + x_1^T x_2 & \le & 0. & \end{array} \right.$$

This quadratic program corresponds to a direct application of the SQP algorithm to MPEC (without relaxation). Also, consider the following quadratic program at x, corresponding to a $\mathcal{Z}(\hat{x}) = \mathcal{Z} = \{\mathcal{Z}_1(\hat{x}), \mathcal{Z}_2(\hat{x})\}$ that defines a partition $\{1, 2, \ldots, n_c\} = \mathcal{Z}_1^{\perp}(\hat{x}) \cup \mathcal{Z}_2^{\perp}(\hat{x}) \cup \mathcal{D}(\hat{x})$, as introduced in the preceding paragraphs. Here we consider the situation where $\mathcal{Z} = \mathcal{Z}(\hat{x})$ is associated to a point $\hat{x}$ in a neighborhood of x^* that satisfies assumption 3 of this theorem.

$$
(QP_R(x, \mathcal{Z})) \left\{ \begin{array}{llcll} \min_d & \frac{1}{2} d^T \overline{W} d + \nabla_x f(x)^T d & & & \\ & g_i(x) + \nabla_x g_i(x)^T d & \leq & 0 & i = 1, 2, \ldots, n_i \\ & h_j(x) + \nabla_x h_j(x)^T d & = & 0 & j = 1, 2, \ldots, n_e \\ \text{sbj. to} & d_{k1} & = & 0 & k \in \mathcal{Z}_2^{\perp} \\ & d_{k2} & = & 0 & k \in \mathcal{Z}_1^{\perp} \\ & x_{k1} + d_{k1} & \leq & 0 & k \in \mathcal{Z}_2 \\ & x_{k2} + d_{k2} & \leq & 0 & k \in \mathcal{Z}_1. \end{array} \right.
$$

From assumptions 1-3 of this Theorem and from (4.14), we obtain that $(QP_R(x, \mathcal{Z}(x)))$ and $(QP(x))$ have the same feasible set and their objective functions has the same values on the feasible set [16, Lemma 5.1, Lemma 5.2] and thus have the same solution. The solution d of $(QP_R(x, \mathcal{Z}(x)))$ is the only strict local minimum in a neighborhood of $d = 0$ and the corresponding multipliers $\nu, \pi, \mu_{k1}, \mu_{k2}$ are unique, due to the MPEC-LICQ assumption **[A2]**. Moreover, d is also the only strict local minimizer of $(QP(x))$ in a neighborhood of $d = 0$ [16]. However, since we need a more specialized version of the results in [16], we will provide our own proof, at least for the part of $(QP_R(x, \mathcal{Z}(x)))$ having the same stationary point as $(QP(x))$.

Stability of the primal-dual solution of $(QP_R(x, \mathcal{Z}))$. We note that if $\mathcal{Z} = \mathcal{Z}(\hat{x}) = \{\mathcal{Z}_1(\hat{x}), \mathcal{Z}_2(\hat{x})\}$, where $\hat{x}$ is a point in the neighborhood of x^* that has the properties stated in assumption 3 of this Theorem, then we will have that:

$$
\{n_d + 1, \ldots, n_c\} \subset \mathcal{Z}_2^{\perp}(\hat{x}) \cap \mathcal{Z}_1(\hat{x}).
$$

Therefore the constraints of $(QP_R(x, \mathcal{Z}(\hat{x})))$ that correspond to $k \in \{n_d + 1, \ldots, n_c\}$, will be the following:

$$
d_{k1} = 0, \quad x_{k2} + d_{k2} \leq 0. \tag{4.16}
$$

If we take $x = x^*$ in $(QP_R(x, \mathcal{Z}(\hat{x})))$, we obtain that $d = 0$ is a solution at which $x^*_{k2} + d_{k2} < 0$, for $k = n_d + 1, \ldots, n_c$. . Therefore all active constraints of $QP_R(x^*, \mathcal{Z}(\hat{x}))$ are active constraints of (RNLPS). In addition, from (4.16) we get that the active inequality constraints of $QP_R(x^*, \mathcal{Z}(\hat{x}))$ are a subset of the ones in (RNLPS).

Therefore, we obtain that

α) Since both (RNLPS) and $QP_R(x^*, \mathcal{Z}(\hat{x}))$ satisfy LICQ, then $\tilde{\nu}_i$ for $i \in \mathcal{A}(x^*)$; $\tilde{\pi}_j$ for $j = 1, 2, \ldots, n_e$; $\tilde{\mu}_{k1} \geq 0$ and and $\tilde{\mu}_{k2}$ for $k = 1, 2, \ldots, n_d$; and $\tilde{\mu}_{k1}$ for $k = n_d + 1, \ldots, n_c$, the components of the unique Lagrange multiplier of (RNLPS) are also the components of the Lagrange multiplier of $QP_R(x^*, \mathcal{Z}(\hat{x}))$.

β) From part α), and our observation concerning active sets, we also obtain the following relationship between the critical cones $\mathcal{C}_{QP_R(x, \mathcal{Z}(\hat{x}))} \subset \mathcal{C}_{RNLPS}$.

From assumptions **[A2]** and **[A3]** we obtain that, for $x = x^*$ and at $d = 0$, $QP_R(x, \mathcal{Z}(\hat{x}))$ satisfies (LICQ) and (RSOSC). Therefore, using the stability results from [35], we obtain that for any $\hat{x}$ that satisfies assumption 3 in a neighborhood of x^*, we will have that there exists a $c_{\mathcal{Z}}$, that truly depends only on $\mathcal{Z}(\hat{x})$, that satisfies:

$$
||d(x, \mathcal{Z}(\hat{x}))|| \leq c_{\mathcal{Z}} \, ||x - x^*|| \,, \quad ||(\nu, \pi, \mu, \eta) - (\tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})|| \leq c_{\mathcal{Z}} \, ||x - x^*|| \,,
$$

where $d(x, \mathcal{Z}(\hat{x}))$ and, respectively, (ν, π, μ, η), are the solution and, respectively, the Lagrange multiplier of $(QP_R(x, \mathcal{Z}(\hat{x})))$. Since the number of partitions that lead to a particular $\mathcal{Z}$ is finite, it follows that there exists $\hat{c}$ such that for any $\hat{x}$ that satisfies assumption 3 a neighborhood of x^*, and for any x in a neighborhood of x^*, we have that:

$$(STAB) \quad ||d(x, \mathcal{Z}(\hat{x}))|| \leq \hat{c}\,||x - x^*||\,, \quad ||(\nu, \pi, \mu, \eta) - (\tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})|| \leq \hat{c}\,||x - x^*||\,.$$

In the rest of the proof, we will use this result for $x = \hat{x}$, but this choice does not affect (STAB).

Constructing a Lagrange multiplier of $(QP(x))$ **starting from a Lagrange multiplier of** $(QP_R(x, \mathcal{Z}(x)))$**.** Consider the Lagrange multiplier $\nu, \pi, \mu_{k1}, \mu_{k2}$ of $(QP_R(x, cZ(x)))$. We must have that:

$$0 = \nabla_x f(x)^T + \overline{W}d + \nabla_x g(x)^T \nu + \nabla_x h(x)^T \pi + \sum_{k=1}^{n_c} \mu_{k1} e_{k1}^T + \mu_{k2} e_{k2}^T. \tag{4.17}$$

From the optimality conditions of $(QP_R(x, \mathcal{Z}(x)))$, we have that $\mu_{k1} \geq 0$, $k \in \mathcal{Z}_2(x)$, and $\mu_{k2} \geq 0$ for $k \in \mathcal{Z}_1(x)$.

We want to show that there exists a choice of $\eta \geq 0$ and $\mu_{k1}^\sharp$, and $\mu_{k2}^\sharp$, $k = 1, 2, \ldots, n_c$, such that:

$$\begin{array}{rcl} 0 & = & \nabla_x f^T(x) + Wd + \nabla_x g^T(x)\nu + \nabla_x h^T(x)\pi \\ & + & \sum_{k=1}^{n_c} \left(\mu_{k1}^\sharp e_{k1}^T + \mu_{k2}^\sharp e_{k2}^T + \eta \left(x_{k2} e_{k1}^T + x_{k1} e_{k2}^T \right) \right), \end{array} \tag{4.18}$$

where $\mu_{k1}^\sharp \geq 0$, $\mu_{k2}^\sharp \geq 0$, for $k = 1, 2, \ldots, n_c$, and $x_{k1} + d_{k1} < 0 \Rightarrow \mu_{k1}^\sharp = 0$, and $x_{k2} + d_{k2} < 0 \Rightarrow \mu_{k2}^\sharp = 0$, for $k = 1, 2, \ldots, n_c$. This would result in d being a stationary point of (QP(x)).

Since $Wd = \overline{W}d + \hat{\eta} \sum_{k=1}^{n_c} (e_{k2} d_{k1} + e_{k1} d_{k2})$, we obtain from (4.17), for any η, that:

$$\begin{array}{rcl} 0 & = & \nabla_x f(x)^T + Wd + \nabla_x g(x)^T \nu + \nabla_x h(x)^T \pi + \eta \sum_{k=1}^{n_c} \left(x_{k2} e_{k1}^T + x_{k1} e_{k2}^T \right) \\ & + & \sum_{k=1}^{n_c} \left(\left(\mu_{k1} - \eta x_{k2} - \hat{\eta} d_{k2} \right) e_{k1}^T + \left(\mu_{k2} - \eta x_{k1} - \hat{\eta} d_{k1} \right) e_{k2}^T \right). \end{array} \tag{4.19}$$

We want to show, based on the previous equality, that the choice

$$\mu_{ki}^\sharp = \mu_{ki} - \eta x_{k\bar{i}} - \hat{\eta} d_{k\bar{i}}, \quad \text{where } \bar{i} = 2 - i + 1,\ i = 1, 2 \tag{4.20}$$

satisfies the multiplier conditions, for the appropriate choice of $\eta \geq 0$.

Recall, we assume that x satisfies Assumption 3 from the statement of this Theorem. We have the following cases.

- $k \in \mathcal{Z}_1(x) \cap \mathcal{Z}_2(x) = \mathcal{D}(x)$. From Assumption **[A4]**, if x is sufficiently close to x^* we must have that $k \in \mathcal{Z}_1(x^*) \cap \mathcal{Z}_2(x^*)$ and that at least one of $\tilde{\mu}_{k1}$ and $\tilde{\mu}_{k2}$ is positive. Using (STAB) we get that at least one of μ_{k1} and μ_{k2} is positive. We assume without loss of generality that $\mu_{k1} > 0$ which means that we must also have $x_{k1} + d_{k1} = 0$. This implies that, since $k \in \mathcal{Z}_1(x)$, we must have that $x_{k1} = 0$, and thus $d_{k1} = 0$. Since $k \in \mathcal{Z}_2(x)$, we also have that $x_{k2} = 0$ and thus, from the inequality $x_{k2} + d_{k2} \leq 0$, we obtain that $d_{k2} \leq 0$. Since $k \in \mathcal{Z}_1(x) \cap \mathcal{Z}_2(x)$, from the optimality conditions for $(QP_R(x, \mathcal{Z}(x)))$, we obtain that $\mu_{k1} \geq 0$ and $\mu_{k2} \geq 0$.

In this situation, the choice (4.20) results in the following relations:

$$\mu^\sharp_{k1} = \mu_{k1} - \eta d_{k2} \geq 0,\ \mu^\sharp_{k2} = \mu_{k2} \geq 0,\ \forall \eta \geq 0.$$

In addition, if $x_{k2} + d_{k2} < 0$, then $\mu_{k2} = 0$ and thus $\mu^\sharp_{k2} = 0$.

- $k \in \mathcal{Z}_1^\perp(x)$. Then $x_{k1} < 0$, $x_{k2} = 0$, $d_{k2} = 0$, and $\mu_{k1} \geq 0$. In this situation, the choice (4.20) results in the following relations:

$$\mu^\sharp_{k1} = \mu_{k1} - \eta x_{k2} - \hat{\eta} d_{k2} = \mu_{k1} \geq 0, \forall \eta \geq 0.$$

In addition, $x_{k1} + d_{k1} < 0$ implies that $\mu_{k1} = 0$, from the optimality conditions of $(QP_R(x, \mathcal{Z}(x)))$ which in turn implies that $\mu^\sharp_{k1} = 0$. For the second multiplier, the choice (4.20) results in the following relation:

$$\mu^\sharp_{k2} = \mu_{k2} - \eta x_{k1} - \hat{\eta} d_{k1}$$

which is nonnegative as soon as

$$\eta \geq \frac{-\mu_{k2} + \hat{\eta} d_{k1}}{-x_{k1}}. \tag{4.21}$$

- $k \in \mathcal{Z}_2^\perp(x)$. Then $x_{k2} < 0$, $x_{k1} = 0$, $d_{k1} = 0$, and the choice (4.20) results in the following relation:

$$\mu^\sharp_{k2} = \mu_{k2} - \eta x_{k1} - \hat{\eta} d_{k1} = \mu_{k2} \geq 0, \forall \eta \geq 0.$$

In addition, $x_{k2} + d_{k2} < 0$ implies that $\mu_{k2} = 0$, from the optimality conditions of $(QP_R(x, \mathcal{Z}(x)))$ which in turn implies that $\mu^\sharp_{k2} = 0$. For the second multiplier, the choice (4.20) results in the following relation:

$$\mu^\sharp_{k1} = \mu_{k1} - \eta x_{k2} - \hat{\eta} d_{k2},$$

which is nonnegative as soon as

$$\eta \geq \frac{-\mu_{k1} + \hat{\eta} d_{k2}}{-x_{k2}}. \tag{4.22}$$

From equation (4.21) and (4.22) and using the choice (4.20) to define $\mu^\sharp_{k1}$ and $\mu^\sharp_{k2}$ we have that (4.18) holds together with,

$$\mu^\sharp_{k1} \geq 0,\ \mu^\sharp_{k2} \geq 0, \quad k = 1, 2, \ldots, n_c$$

as well as

$$x_{k1} + d_{k1} < 0 \Rightarrow \mu^\sharp_{k1} = 0,\ k = 1, 2, \ldots, n_c;\ x_{k2} + d_{k2} < 0 \Rightarrow \mu^\sharp_{k2} = 0,\ k = 1, 2, \ldots, n_c,$$

provided that the following holds:

$$\eta = \max\left\{0, \max_{k \in Z_1^\perp(x)} \left\{\frac{-\mu_{k2} + \hat{\eta} d_{k1}}{-x_{k1}}\right\}, \max_{k \in Z_2^\perp(x)} \left\{\frac{-\mu_{k1} + \hat{\eta} d_{k2}}{-x_{k2}}\right\}\right\}. \tag{4.23}$$

This shows that $\nu, \pi, \mu^\sharp_1, \mu^\sharp_2, \eta$, is a Lagrange multiplier of $(QP(x))$.

Constructing a Lagrange multiplier of (QPMC(c)). We want to now show that we can construct from this multiplier a Lagrange multiplier for $(QPMC(c))$. This would mean that the solution d of $(QP(x))$ can be extended to the solution $(d, 0)$ of (QPMC(c)) and would prove the claim **i)**.

To that end, we claim that the multiplier we chose above approaches the fundamental multiplier. Tracking back the expression (4.20) we obtained for $\mu^\sharp_{k1}$ and $\mu^\sharp_{k2}$, and since $d \to 0$ as $x \to x^*$, from (STAB), this claim follows if we can show that η approaches η^*, where the latter is the corresponding component of the fundamental multiplier.

Comparing the definition of η^*, with the definition of η, we see that they are different in two respects: η includes components of d, the solution of $(QP_R(x, \mathcal{Z}(x)))$ and the maximums defining η are done over larger sets.

We now show that sufficiently close to the solution we have that:

$$\eta = \max\left\{0, \max_{k \in \mathcal{Z}_1^\perp(x^*)}\left\{\frac{-\mu_{k2} + \hat{\eta} d_{k1}}{-x_{k1}}\right\}, \max_{k \in \mathcal{Z}_2^\perp(x^*)}\left\{\frac{-\mu_{k1} + \hat{\eta} d_{k2}}{-x_{k2}}\right\}\right\}, \tag{4.24}$$

and that in effect, for x sufficiently close to x^*, η can be defined as if the sets $\mathcal{Z}_1(x)$ and $\mathcal{Z}_2(x)$ are computed at $x = x^*$. In turn, using (4.2) and (STAB), this will result in the following relation:

$$\eta = \max\left\{0, \max_{k \in \{n_d+1, \dots, n_c\}}\left\{\frac{-\mu_{k1} + \hat{\eta} d_{k2}}{-x_{k2}}\right\}\right\} \overset{d \to 0}{\to} \eta^*. \tag{4.25}$$

Using (STAB) we obtain that:

$$\tilde{\mu}_{ki} > 0 \quad \Rightarrow \mu_{ki} > 0, \quad \text{for } k = 1, 2, \dots, n_d \text{ and } i = 1, 2, \tag{4.26}$$

where $\tilde{\mu}_{ki}$, for $k = 1, 2, \dots, n_d$ and $i = 1, 2$, denote multipliers of (RNLPS) and μ_{ki}, for $k = 1, 2, \dots, n_d$ and $i = 1, 2$ denote multipliers of $(QP(x, \mathcal{Z}(x)))$.

We now show that any index k in $\mathcal{Z}_2^\perp(x) \backslash \mathcal{Z}_2^\perp(x^*)$ cannot win the comparison test that defines η, (4.24). Using (4.15), we obtain that $k \in \mathcal{D}(x^*)$. From Assumption **[A4]** and (4.26) we determine that at least one of $\tilde{\mu}_{k1}$ and $\tilde{\mu}_{k2}$ must be positive. If both are positive, from our assumption 3 we would have that $x_{k1} = x_{k2} = 0$ and thus $k \in \mathcal{D}(x)$ and $k \notin \mathcal{Z}_2^\perp(x)$, which would be a contradiction. The only way in which k can be in $\mathcal{Z}_2^\perp(x)$ is when $\tilde{\mu}_{k1} > 0$ but $\tilde{\mu}_{k2} = 0$, and, using (4.26), we also have that $\mu_{k1} > 0$. In that case we have, since $k \in Z_2^\perp(x)$, that $x_{k1} = 0$, $x_{k2} < 0$, $d_{k1} = 0$. The entry corresponding to k in the definition of η is

$$l_{k2} = \frac{-\mu_{k1} + \hat{\eta} d_{k2}}{-x_{k2}}.$$

Using (STAB) in conjunction with $\hat{\eta} \le c_\infty$ results in $l_{k2} < 0$ whenever ϵ is sufficiently small, and thus k cannot win the comparison defining η (4.24). Similarly an index in $\mathcal{Z}_1^\perp(x) \backslash \mathcal{Z}_1^\perp(x^*)$ cannot win the comparison defining η, which means that, for x sufficiently close to x^*, we have that η is in fact defined by (4.24), and, as a result of (4.2), by (4.25).

Since $d \to 0$ as $\epsilon \to 0$ from (STAB), where here d is the solution of $(QP(x, \mathcal{Z}(x)))$, we obtain that η approaches η^* from (FMC) and (STAB) thus, from (FMC), (4.20) and (STAB) the Lagrange multiplier $(\nu, \pi, \mu_1^\sharp, \mu_2^\sharp, \eta)$ of $(QP(x, \mathcal{Z}(x)))$ approaches the fundamental multiplier, whose components are: ν^*, π^*, μ_1^*, μ_2^*, η^*, as $\epsilon \to 0$. Define now

$$\theta = c_\infty - \sum_{i=l_i+1}^{n_i} \nu_i - \sum_{j=l_e+1}^{n_e} |\pi_j| - \eta.$$

As $\epsilon \to 0$, from the definition of ν_0 (4.7), we obtain that θ will approach $c_\infty - \nu_0 = \gamma_0 > 0$. Therefore comparing (QP(x)) with (QPMC(c)), we see that $\nu, \pi, \mu_1^\sharp, \mu_2^\sharp, \eta, \theta$, is a Lagrange multiplier of (QPMC(c)) for the stationary point $(d, -\zeta)$. Since the solution is locally unique, due to the Robinson conditions holding at $(x^*, 0)$ for (MPCC(c)) for ϵ sufficiently small from Lemma 4.1 coupled with (4.14), we obtain that $\zeta + d_\zeta = 0$ at the solution of (QPMC(c)). This proves the claim of part **i)**.

Following (4.26), we obtain that the strongly active constraints from degenerate pairs are active at the solution of (QPMC(c)). This proves part **ii)** of the claim.

What is now left to prove is part **iii)**, that the point $x + d$, where d is the solution of (QPMC(c)), also satisfies the complementarity conditions.

If the index $k \notin D(x)$, we have that exactly one of x_{k1}, x_{k2} is negative, since the point x satisfies the complementarity constrains. We denote that entry by $x_{ki(k)}$. We thus have that $x_{ki(k)} < 0$ and that $x_{k\bar{i}(k)} = 0$, where $\bar{i}(k) = 3 - i(k)$.

The linear constraint from (QPMC(c)) derived from the complementarity constraints of (MPCCS) becomes:

$$\sum_{k \in \{1,2,\dots n_c\} \setminus D(x)} x_{ki(k)} \left(x_{k\bar{i}(k)} + d_{k\bar{i}(k)} \right) \le \zeta + d_\zeta = 0,$$

where the last part of the equation follows part **i)** of this Theorem. Since $x_{k\bar{i}(k)} + d_{k\bar{i}(k)} \le 0$, from the constraints of (MPCC(c)) and since $x_{ki(k)} < 0$ for $k \notin D(x)$, it follows that:

$$k \notin D(x) \;\Rightarrow\; x_{k\bar{i}(k)} + d_{k\bar{i}(k)} = 0.$$

If, on the other hand, we have that $k \in D(x)$, then we must have that $k \in D(x^*)$, and using Assumption [**A4**], as well as (4.26) together with our conclusion that the multiplier of (QPMC(c)) that we constructed approaches the fundamental multiplier (FMC), we obtain that the least one of the constraints $x_{k1} + d_{k1} \le 0$, $x_{k2} + d_{k2} \le 0$ is active for the solution of (QPMC(c)). We therefore obtain that the point $x + d$ satisfies the complementarity constraints, which completes the proof. $\diamond$

4.3. A superlinearily convergent algorithm. We now state our algorithm for general nonlinear programming, and then show that, under certain conditions, it will converge superlinearily when applied to (MPCCS). The problem to be solved is the following:

$$\begin{array}{rlcll} \min_x & \tilde{f}(x) & & & \\ \text{sbj. to} & \tilde{g}_i(x) & \le & 0, & i = 1, 2, \dots, m \\ & \tilde{h}_j(x) & = & 0, & j = 1, 2, \dots, r. \end{array} \tag{4.27}$$

We assume that $\tilde{g}_i(x)$, $i = 1, 2, \dots, l_i$ and $\tilde{h}_j(x)$, $j = 1, 2, \dots, l_e$, are linear constraints that do not get relaxed in an elastic mode approach.

$x^0 = x$, $c_\infty = c_0$, $k = 0$.
NLP: Solve (QP).
If $\sum_{i=l_i+1}^{m} \nu_i^k + \sum_{j=l_e+1}^{r} |\pi_j^k| \le c_\mu$ and (QP) is feasible
    $x^{k+1} = x^k + d^k$, $k = k+1$, return to **NLP**.
Else
    NLPC: solve (QPC).
    $x^{k+1} = x^k + d^k$, $\zeta^{k+1} = \zeta^k + \delta_\zeta$, $k = k+1$.
    If $\sqrt{||d^k|| + \left\|\delta_\zeta^k\right\|} \le \zeta^k$,
        $c_\infty = c_\infty c_\gamma$, $k = k+1$ return to **NLPC**.
    End If
End If

Table 4.1
The elastic mode algorithm

In the algorithm that we describe we use one of the two following quadratic programs:

$$
(QP)\quad \begin{array}{rlcl} \min_d & \nabla_x \tilde{f}(x)d + \frac{1}{2}d^T\tilde{W}d & & \\ \text{sbj. to} & \tilde{g}_i(x)d + \nabla_x\tilde{g}_i(x)^T d & \le & 0, \quad i = 1,2,\ldots,m \\ & \tilde{h}_j(x)d + \nabla_x\tilde{h}_j(x)^T d & = & 0, \quad j = 1,2,\ldots,r \end{array}
$$

$$
(QPC)\quad \begin{array}{rrcll} \min_{d,d_\zeta} & \nabla_x \tilde{f}(x)d + \frac{1}{2}d^T\tilde{W}d & + & c_\infty(\zeta + d_\zeta) & \\ \text{sbj. to} & \tilde{g}_i(x)d + \nabla_x\tilde{g}_i(x)d & \le & 0, & i = 1,2,\ldots,l_i \\ & \tilde{g}_i(x)d + \nabla_x\tilde{g}_i(x)d & \le & \zeta + d_\zeta, & i = l_i+1,\ldots,m \\ & \tilde{h}_j(x)^T d + \nabla_x\tilde{h}_j(x)d & = & 0, & j = 1,2,\ldots,l_e \\ -\zeta - d_\zeta \le & \tilde{h}_j(x)^T d + \nabla_x\tilde{h}_j(x)d & \le & \zeta + d_\zeta, & j = l_e+1,\ldots,r \\ & \zeta + d_\zeta & \ge & 0. & \end{array}
$$

When (QP) is solved, we obtain a direction d and Lagrange multipliers ν_i, $i = 1,2,\ldots,n_i$ and π_j, for $j = 1,2,\ldots,n_e$. When (QPC) is solved, we obtain Lagrange multipliers ν_i, $i = 1,2,\ldots,n_i$; π_j, $j = 1,2,\ldots,l_i$; π_j^+ and π_j^-, $j = l_e+1,\ldots,n_j$; and θ. We define the matrix $\tilde{W}$ to be used to in the next quadratic program, which is either (QP) or (QPC) as follows.

$$
\tilde{W} = \begin{cases} \nabla_{xx}^2 \tilde{f}(x) + \sum_{i=l_i+1}^{r} \nu_i \nabla_{xx}^2 \tilde{g}_i(x) + \sum_{j=l_e+1}^{m} \pi_j \nabla_{xx}^2 \tilde{h}_j(x), \\ \qquad \text{if (QP) was last solved} \\ \nabla_{xx}^2 \tilde{f}(x) + \sum_{i=l_i+1}^{r} \nu_i \nabla_{xx}^2 \tilde{g}_i(x) + \sum_{j=l_e+1}^{m} \left(\pi_j^+ - \pi_j^-\right) \nabla_{xx}^2 \tilde{h}_j(x), \\ \qquad \text{if (QPC) was last solved.} \end{cases}
$$

We now define our algorithm in Table 4.1. The algorithm depends upon the parameters c_μ, $c_\gamma > 1$, c_0.

The (QP) subproblem of this algorithm is the same subproblem as in the algorithm by [17], minus the trust-region constraint that is imposed for globalization. The (QPC) subproblem is the natural extension of (QP) when using the elastic mode. The elastic mode strategy used here is identical to the one used in SNOPT [22], except that we use here the L_∞ function.

Note that once the subproblem (QPC) is solved the algorithm never solves the problem (QP) again. So we either solve (QP) till convergence or (QPC) till convergence.

We now analyze the effect of applying this algorithm to (MPCC). For this case, the (QP) subproblem becomes (QP(x)), whereas the (QPC) subproblem becomes (QPMC(c)).

THEOREM 4.5. *Assume that Assumptions* **[A1]**–**[A5]** *hold near a solution* x^* *of (MPCCS). Assume that the point* x^{k_0} *is sufficiently close to* x^* *and either*

i) The elastic mode is never invoked, and the algorithm uses at k_0*, for the purpose of constructing the matrix* $\tilde{W}$*, an estimate of the Lagrange multiplier that is sufficiently close to a multiplier of (MPCCS).*

ii) The elastic mode is invoked at k_0 *with* $c_\infty \geq \nu_0$ *and at all subsequent iterates, and the algorithm uses at* k_0*, for the purpose of constructing the matrix* $\tilde{W}$*, an estimate of the Lagrange multiplier that is sufficiently close to a multiplier of (MPCC(c)), and a* ζ^{k_0} *that is sufficiently close to* 0.

Then x^k *converges to* x^* *superlinearily in case* **i)** *and* (x^k, ζ^k) *converges to* $(x^*, 0)$ *superlinearily in case* **ii)**.

Proof For case **i)** to hold, we must have that the test involving the 1 norm of the nonlinear multipliers is always satisfied. Therefore, for all iteration indices k, we have that the following inequality holds:

$$\sum_{i=l_i+1}^{n_i} \nu_i^k + \sum_{j=l_e+1}^{n_e} |\pi_j^k| + \eta^k \leq c_\mu,$$

where the left hand side is composed of multipliers of $(\mathrm{QP}(x^k))$. Since we are sufficiently close to x^*, this must imply that $c_\mu \geq \nu_0$. In the latter case, we will have that a solution of $(\mathrm{QP}(x^k))$ can be completed to a solution of (QPMC(c)) with $\zeta = 0$ and $\zeta + \delta_\zeta = 0$, and $c_\infty = c_\mu \geq \nu_0$. We get therefore treat the case **i)** as a special case of case **ii)** to which we now confine our attention. We have the following cases.

- Case 1 $c_\infty = \nu_0$. In this case, following Lemma 4.1, (MPCC(c)) has a unique multiplier, and the result from [5] applies to give superlinear convergence of (x^k, ζ^k) to $(x^*, 0)$.
- Case 2 $c_\infty > \nu_0$ and $\zeta^k + \delta\zeta^k = 0$ for all $k \geq k_0$. Then the solution obtained of (QPMC(c)) is a also a solution of $(\mathrm{QP}(x^k))$ which is feasible for all $x = x^k$. The result claimed follows from [16].
- Case 3 $c_\infty > \nu_0$ and $\zeta^{k_0} + \delta\zeta^{k_0} > 0$. Then using Theorem 4.3 we obtain that the point $x^{k_0+1} = x^{k_0} + d^{k_0}$, where d^{k_0} are the components of the solution of the quadratic program (QPMC(c)) corresponding to x at (x^k, ζ^k), satisfies the complementarity constraints and that the strongly active variables from degenerate pairs are active at the solution of the quadratic program (QPMC(c)). Using now Theorem 4.4 for all $k \geq k_0 + 1$, we obtain that x^k satisfies the complementarity constraints for $k \geq k_0 + 1$ and that $\zeta^k = 0$ for $k \geq k_0 + 2$. The superlinear convergence result again follows from [16].

The proof is complete. ◇

We have the following observations:

1. We do not prove here the more desirable result that if x^k is sufficiently close to x^* we obtain superlinear convergence to x^* or of (x^k, ζ^k) to $(x^*, 0)$. The difficulty is that, unless we have an estimate for ν_0, the subsequent iterates may find themselves far away from x^* once the elastic mode is entered. In effect, if the penalty parameter c_∞ is too small, the iterates may even be

unbounded, even if the objective function is bounded on the feasible set. Such an adverse outcome can be prevented only by a global convergence result that will be the result of future research.

2. From the Theorem 4.3 **i)**, we have that $||\zeta^k|| = O(||d^k|| + |d^k_\zeta|)$, so the c_∞ update rule from Table 4.1 will not be triggered, for x^k sufficiently close to x^* and for $c_\infty > \nu_0$. **Therefore, the update rule does not interfere with super linear convergence**. On the other hand, it is also clear that, if $c_\infty < \nu_0$, then $\zeta^k > 0$ and the rule will eventually be triggered assuming that the iterates approach a stationary point. The test we use here is important because we do not spend an infinite amount of steps with an inappropriate c_∞. The complete study of an appropriate rule should also involve global convergence issues, since one possible outcome of the penalty parameter adaptation rule is to obtain unbounded iterates.
3. The appropriate value of c_μ is ν_0 (or a slightly larger value) which of course cannot be determined unless we specifically use the MPCC structure. A general NLP approach cannot be guaranteed to succeed in determining the appropriate value for c_∞ through c_μ by looking at the multipliers of (QPMC(c)) alone, since the Lagrange multiplier set of (MPCCS) is infinite, whereas such rules are based on the assumption of finiteness of the Lagrange multiplier set at the solution x^* [2]. Whether c_μ can be adaptively defined is a matter for future research. For NLP implementations, a user-defined value for c_μ is considered to be an acceptable approach [22].
4. Note that the only meaningful difference between the proofs outlined here and the ones in [16], once the Theorems 4.3 and 4.4 have been established, is the one involve in the case where $c_\infty = \nu_0$, for which the well-known result from [5] has been invoked. Our contribution to this class of super linear convergent results has been essentially to show that, for the case $c_\infty > \nu_0$, the elastic mode approach, as presented in Table 4.1, will force the algorithm to choose points that are either always feasible or always satisfy the complementarity constraints and in so doing we do not have to apriori assume that this holds, as it was done in [16].
5. Note that we did not use Assumption [**A5**] and the first part of Assumption [**A4**] in the proof of any of the results in Section 4, except when we needed to invoke results from [16]. The proof of our main helping results Theorem 4.3 and Theorem 4.4 did not need the Assumption [**A5**] and the first part of Assumption [**A4**].

Acknowledgments. Thanks to Danny Ralph and Stefan Scholtes for the comments made on the presentation of partial results of this work at ISMP 2000 in Atlanta. Thanks to Alexander Shapiro for countless discussions concerning degenerate nonlinear programming and particularly for pointing out reference [5]. I am particularly grateful to my colleague Sven Leyffer for providing the AMPL examples and guidance for the numerical results and to Steve Wright for pointing out the result [40, Theorem A1]. Finally, I acknowledge the detailed comments from the referees and the editor, Danny Ralph, that have prompted the development of Section 4, and have markedly improved the quality of the results. This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing, U.S. Department of Energy, under Contract W-31-109-Eng-38. This work was also supported by award DMS-9973071 of the National Science Foundation.

REFERENCES

[1] M. Anitescu, *Degenerate nonlinear programming with a quadratic growth condition*, SIAM Journal of Optimization, 10(4) (2000), pp. 1116–1135.

[2] ———, *On the rate of convergence of sequential quadratic programming with nondifferentiable exact penalty function in the presence of constraint degeneracy*, Mathematical Programming, 92 (2002), pp. 359–386.

[3] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.

[4] ———, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, 1995.

[5] F. Bonnans, *Local analysis of newton-type methods for variational inequalities and nonlinear programming*, Applied Mathematics and Optimization, 29 (1994), pp. 161–186.

[6] F. Bonnans and A. Ioffe, *Second-order sufficiency and quadratic growth for nonisolated minima*, Mathematics of Operations Research, 20(4) (1995), pp. 801–819.

[7] F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, 2000.

[8] J. V. Burke, *A sequential quadratic programming method for potentially infeasible mathematical programs*, Journal of Mathematical Analysis and Applications, 139 (1989), pp. 319 – 351.

[9] ———, *A robust trust region method for constraint nonlinear programming problems*, SIAM Journal on Optimization, 2 (1992), pp. 325–347.

[10] J. V. Burke and S.-P. Han, *A robust sequential quadratic programming method*, Mathematical Programming, 43 (1989), pp. 277 – 303.

[11] R. J. Caron, H. J. Greenberg, and A. G. Holder, *Analytic centers and repelling inequalities*, European Journal of Operations Research, 143(2) (2002), pp. 268–290.

[12] F. Facchinei, *a robust recursive quadratic programming algorithm with global and super linear convergence properties*, Journal of optimization theory and applications, 92 (1997), pp. 543–579.

[13] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983.

[14] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, Chichester, 1987.

[15] R. Fletcher and S. Leyffer, *Nonlinear programming without a penalty function*, Mathematical Programming, (to appear).

[16] R. Fletcher, S. Leyffer, S. Sholtes, and D. Ralph, *Local convergence of sqp methods for mathematical programs with equilibrium constraints*, Tech. Rep. NA 210, University of Dundee, Dundee, UK, 2002.

[17] R. Fletcher, S. Leyffer, and P. L. Toint, *On the global convergence of a filter-sqp type algorithm*, Numerical Analysis Report NA/183, Department of Mathematics, University of Dundee, UK, 1998.

[18] R. Freund, R. Roundy, and M. Todd, *Identifying the set of always-active constraints in a sytem of linear inequalities*, working paper, Sloan School of Management, MIT, Cambridge, MA, 1985.

[19] M. Fukushima, Z.-Q. Luo, and J.-S. Pang, *A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints*, Computational Optimization and Applications, 10 (1998), pp. 5–34.

[20] M. Fukushima and P. Tseng, *An implementable active-set algorithm for computing a b-stationary point of a mathematical program with linear complementarity constraints*, SIAM J. of Optimization, (To appear).

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

[22] P. E. Gill, W. Murray, and M. A. Saunders, *User's guide for snopt 5.3: A fortran package for large-scale nonlinear programming*, Report NA 97-5, Department of Mathematics, University of California, San Diego, 1997.

[23] X. Hu and D. Ralph, *Convergence of a penalty method for mathematical programming with complementarity constraints*, Journal of Optimization Theory and Applications, (to appear).

[24] A. Ioffe, *Necessary and sufficient conditions for a local minimum.3: Second order conditions and augmented duality*, SIAM Journal of Control and Optimization, 17(2) (1979), pp. 266–288.

[25] ———, *On sensitivity analysis of nonlinear programs in banach spaces: the approach via composite unconstrained optimization*, SIAM Journal of Optimization, 4(1) (1994), pp. 1–43.

[26] S. Lucidi, *Results on a continuously differentiable exact penalty function*, SIAM Journal on

optimization, 2 (1992), pp. 558–574.
[27] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Complementarity Constraints*, Cambridge University Press, Cambridge, UK, 1996.
[28] Z.-Q. Luo, J.-S. Pang, D. Ralph, and S.-Q. Wu, *Exact penalization and stationarity conditions of mathematical programs with complementarity constraints*, Mathematical Programming, 34(2) (1986), pp. 142–162.
[29] O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969.
[30] O. L. Mangasarian and S. Fromovitz, *The fritz john necessary optimality conditions in the presence of equality constraints*, Journal of Mathematical Analysis and Applications, 17 (1967), pp. 34–47.
[31] B. Murtagh and M. A. Saunders, *Minos 5.0 user's guide*, Technical Report SOL 83-20, Stanford University, Stanford, 1983.
[32] *The neos guide*. Available online at *http://www.mcs.anl.gov/otc/Guide*.
[33] J. Outrata, M. Kocvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, 1998.
[34] E. Polak, *Optimization*, Springer, New York, 1997.
[35] S. M. Robinson, *Generalized equations and their solutions, part ii: Applications to nonlinear programming*, Mathematical Programming Study, 19 (1980), pp. 200–221.
[36] H. Scheel and S. Scholtes, *Mathematical programs with equilibrium constraints: Stationarity, optimality and sensitivity*, Mathematics of Operations Research, 25(1) (2000), pp. 1–25.
[37] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM Journal of Optimization, 11(4) (2001), pp. 918–936.
[38] A. Shapiro, *Sensitivity analysis of nonlinear programs and differentiability properties of metric projections*, SIAM Journal of Control and Optimization, 26(3) (1988), pp. 628–645.
[39] J. Telgen, *Minimal representation of convex polyhedral sets*, Journal of Optimization Theory and Applications, 38 (1982), pp. 1–24.
[40] S. J. Wright, *Modifying sqp for degenerate problems*, Preprint ANL/MCS-P699-1097, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill, 1997.