Mathematical Programs with Complementarity Constraints: Convergence properties of a smoothing method

Gemayqzel Bouza
University of Havana, Department of Mathematics, Havana, Cuba.
email: gema@matcom.uh.cu

Georg Still (corresponding author)
University of Twente, Department of Mathematics, Enschede, the Netherlands
email: g.still@math.utwente.nl

Abstract. In the present paper, optimization problems P with complementarity constraints are considered. Characterizations for local minimizers $\overline{x}$ of P of order one and two are presented. We analyze a parametric smoothing approach for solving these programs in which P is replaced by a perturbed problem P_τ depending on a (small) parameter τ. We are interested in the convergence behavior of the feasible set $\mathcal{F}_\tau$ and the convergence of the solutions $\overline{x}_\tau$ of P_τ for $\tau \to 0$. In particular, it is shown, that under generic assumptions the solutions $\overline{x}_\tau$ are unique and converge to a solution $\overline{x}$ of P with a rate $\mathcal{O}(\sqrt{\tau})$. Moreover, the convergence for the Hausdorff distance $d(\mathcal{F}_\tau, \mathcal{F})$ between the feasible sets of P_τ and P is of order $\mathcal{O}(\sqrt{\tau})$.

Keywords: mathematical programs with complementarity constraints, smoothing method, rate of convergence, genericity.

MSC 2000 Subject Classification : Primary 90C30; Secondary 90C33, 65K05.

OR/MS Subject Classification : Programming, Complementarity.

1. Introduction

This paper deals with optimization problems of the form

$$\text{P:} \quad \min_x f(x) \quad \text{s.t.} \quad \begin{array}{rcll} g_j(x) & \geq & 0\,, & j \in J := \{1,\ldots,q\} \\ r_i(x)s_i(x) & = & 0\,, & i \in I := \{1,\ldots,m\} \\ r_i(x)\,,\ s_i(x) & \geq & 0\,, & i \in I\,. \end{array} \tag{1}$$

As usual, such a program will be called a MPEC problem. All functions $f, g_j, r_i, s_i : \mathbb{R}^n \to \mathbb{R}$ are assumed to be C^2-functions. The constraints $r_i(x)s_i(x) = 0,\ r_i(x), s_i(x) \geq 0$ are called *complementarity constraints*.

This class of MPEC problems is a topic of intensive recent research (see *e.g.*, [5, 6, 8, 10, 11, 12, 14, 15, 16] and the references in these contributions). Complementarity constraints arise in problems with equilibrium conditions (*cf.* Outrata,Kocvara and Zowe [13]) or as special cases in the so-called Kuhn Tucker approach for solving problems with a bilevel structure (see *e.g.* [18]).

We say that at a local solution $\overline{x}$ of P the *strict complementary slackness* is fulfilled if the relation

$$\text{(SC):} \qquad r_i(\overline{x}) + s_i(\overline{x}) > 0, \quad \forall i \in I \tag{2}$$

is satisfied. The problem in MPEC is that typically the condition SC is not satisfied at a solution $\overline{x}$ of P. It is also well-known that the *Mangasarian Fromovitz constraint qualification* (MFCQ) of standard finite programming (and thus the stronger *Linear Independency constraint qualification* (LICQ)) fails to hold at any feasible point of P (see *e.g.*, [2]). So, to solve these complementarity constrained programs numerically, we cannot use standard software of nonlinear programming since the standard algorithms always rely on LICQ.

To circumvent this problem the following *parametric smoothing* approach can be applied. Instead of P we consider the perturbed problem

$$\text{P}_\tau\text{:} \quad \min_x f(x) \quad \text{s.t.} \quad \begin{array}{rcll} g_j(x) & \geq & 0\,, & j \in J \\ r_i(x)s_i(x) & = & \tau\,, & i \in I \\ r_i(x), s_i(x) & \geq & 0\,, & i \in I\,, \end{array} \tag{3}$$

where $\tau > 0$ is a small perturbation parameter. In this paper we intend to analyze the convergence behavior of this approach.

Let in the following φ, φ_τ denote the marginal values, $\mathcal{F}, \mathcal{F}_\tau$ the feasible sets and $\mathcal{S}_\tau, \mathcal{S}$ the sets of minimizers of $P = P_0$, P_τ respectively. We expect, by letting $\tau \to 0$, that a solution $\overline{x}_\tau$ of P_τ converges to a solution $\overline{x}$ of P.

It will be shown that under natural (generic) assumptions the convergence rate for

$$\mathcal{F}_\tau \to \mathcal{F} \quad \text{and for} \quad \overline{x}_\tau \to \overline{x} \quad \text{is of order} \quad \mathcal{O}(\sqrt{\tau})\,.$$

The assumptions MPEC-LICQ , MPEC-SC and MPEC-SOC (*cf.*, (6),(15),(16)) will play a crucial role in the convergence analysis.

The paper is organized as follows. Section 2 illustrates the convergence behavior on some motivating examples and discusses natural regularity conditions. Section 3 reviews the genericity results in [16] and presents necessary and sufficient optimality conditions for a minimizer $\overline{x}$ of P of order one and two under natural assumptions. In Section 4 the convergence behavior of the perturbed feasible set $\mathcal{F}_\tau$ is analyzed from a local and global viewpoint. Finally, in the last section we prove the existence of local minimizers $\overline{x}_\tau$ of P_τ near a local minimizer $\overline{x}$ of P and their unicity. We show that generically the rate $\|\overline{x}_\tau - \overline{x}\| = O(\sqrt{\tau})$ takes place.

We introduce some notation. The distance between a point $\hat{x}$ and a set $\mathcal{F}$ is defined by $d(\hat{x}, \mathcal{F}) = \min\{\|x - \hat{x}\| \mid x \in \mathcal{F}\}$. We also use the notation $B_\varepsilon(\overline{x}) = \{x \mid \|x - \overline{x}\| < \varepsilon\}$ and denote its closure by $\overline{B}_\varepsilon(\overline{x})$. The norm $\|x\|$ will always be the Euclidean norm.

In the rest of this introduction we will discuss earlier results related to our investigations. The parametric approach (3) has been used for the first time by Luo et al. [12] in connection with equilibrium constrained problems. Here constraints $y_i w_i = 0$ had been perturbed to $y_i w_i = \sigma\mu$ (*cf.* [12, p.280]). For problems of the type (1) this smoothing method has been applied by Facchinei et *al.* [4], Fukushima and Pang [6] and Hu [9] (using NCP-functions). In these papers the convergence to a B-stationary point has been established (under appropriate regularity assumptions). In Stein and Still [18], such a convergence is obtained for a similar (interior point) approach for solving semi-infinite programming problems. A referee draw our attention to the (preprint) of Ralph and Wright [14]. Here a convergence $\|\overline{x}_\tau - \overline{x}\| \le O(\tau^{1/4})$ has been shown (see also Corollary 5.2). Under an additional MPEC-SC condition we will prove the convergence $\|\overline{x}_\tau - \overline{x}\| = O(\tau^{1/2})$ (*cf.* Theorem 5.1). With respect to this result the present contribution is complementary to the paper [14].

Other regularizations of MPEC problems have been considered in the literature such as:

$$\begin{array}{llrcll} \mathrm{P}_\tau^{\le}: & \min_x f(x) \quad \text{s.t.} & g_j(x) & \ge & 0\,, & j \in J \\ & & r_i(x)s_i(x) & \le & \tau\,, & i \in I \\ & & r_i(x), s_i(x) & \ge & 0\,, & i \in I\,. \end{array}$$

$$\begin{array}{llrcll} \hat{P}_\tau^{\le}: & \min_x f(x) \quad \text{s.t.} & g_j(x) & \ge & 0\,, & j \in J \\ & & r^T(x)s(x) & \le & \tau\,, & \\ & & r_i(x), s_i(x) & \ge & 0\,, & i \in I\,. \end{array}$$

Scholtes [17] answered the question under which assumptions a stationary point $x(\tau)$ of $\mathrm{P}_\tau^{\le}$, $\tau \downarrow 0$, converges to a B-stationary point of P. In [14] it is shown that (under natural conditions) the solution $x(\tau)$ of $\mathrm{P}_\tau^{\le}$ converge to a (nearby) solution $\overline{x}$ of MPEC with order $O(\tau)$. Similar results are stated for the problem $\hat{P}_\tau^{\le}$.

We emphasis that these regularizations $\mathrm{P}_\tau^{\le}$, $\hat{P}_\tau^{\le}$ structurally completely differ from the smoothing approach P_τ. For $\mathrm{P}_\tau^{\le}$, *e.g.*, the following is shown in [17, Th.3.1,Cor.3.2]: If $\overline{x}$ is a solution of P where MPEC-LICQ and MPEC-SC holds then for the (nearby) minimizers $\hat{x}_\tau$ of $\mathrm{P}_\tau^{\le}$ (for τ small enough) the complementarity constraints $r_i(x)\, s_i(x) \le \tau$, $i \in I_{rs}(\overline{x})$, are *not active* (*cf.* Section 2 for a definition of $I_{rs}(\overline{x})$). More precisely,

$$r_i(\hat{x}_\tau) = s_i(\hat{x}_\tau) = 0\,, \quad \forall i \in I_{rs}(\overline{x}),$$

is true. This fact can also be deduced from Corollary 3.1 (*cf.*, Section 3). In particular, in the case $I = I_{rs}(\overline{x})$ (for all small $\tau > 0$) the solution $\hat{x}_\tau$ of $\mathrm{P}_\tau^{\le}$ coincides with the solution $\overline{x}$ of P. In Hu and Ralph [8] the following parametric version of P has been studied.

$$\begin{array}{llrcll} \mathrm{P}(\tau): & \min_x f(x,\tau) \quad \text{s.t.} & g_j(x,\tau) & \ge & 0\,, & j \in J \\ & & r_i(x,\tau)\cdot s_i(x,\tau) & = & 0\,, & i \in I \\ & & r_i(x,\tau)\,,\ s_i(x,\tau) & \ge & 0\,, & i \in I \end{array}$$

under the assumption $f, g_j, r_i, s_i \in C^2$ (wrt. all variables). Let $\overline{x}$ be a local minimizer of $P(0)$ (*i.e.*, $\tau = 0$). In contrast to our perturbation P_τ in (3), under natural assumptions, the parametric program $P(\tau)$ can be analyzed using the (smooth) Implicit Function Theorem so that roughly speaking the perturbation $P(\tau)$ behaves *more smoothly* than the perturbation P_τ. (In fact, by using the result of Corollary 3.1, the problem $P(\tau)$ can be analyzed as the parametric version of the relaxed problem $P_R(\overline{x})$ (see Corollary 3.1), *i.e.*, it can be treated as a standard parametric optimization problem.) In particular under the assumption that MPEC-LICQ and MPEC-SOC holds at $\overline{x}$ the value function $\varphi(\tau)$ of $P(\tau)$ is differentiable at $\tau = 0$ implying

$$|\varphi(\tau) - \varphi| = O(\tau)$$

and a similar behavior for the minimizers. This contrasts with the *nonsmooth* behavior $|\varphi(\tau) - \varphi| = O(\sqrt{\tau})$ for the perturbation P_τ (see Example 2.1 and Corollary 5.1).

REMARK 1.1 For numerical purposes it is convenient to model the constraints $r_i(x)s_i(x) = \tau$ and $r_i(x), s_i(x) \ge 0$ equivalently by a unique constraint $\phi_\tau(r_i(x), s_i(x)) = 0$ where ϕ_τ is a so-called parameterized NCP-function (see *e.g.* [3] and [6]).

REMARK 1.2 We emphasize that all results in this paper remain valid for problems P containing additional equality constraints $c_l(x) = 0$ if we assume additional linear independence of the gradients $\nabla c_l(x)$. To keep the presentation as clear as possible we omit these equality constraints.

The smoothing approach P_τ is directly connected with the interior point method for solving finite optimization problems (FP). To solve a program

$$\text{FP:} \quad \min f(x) \quad \text{s.t.} \quad g_j(x) \ge 0\,, \quad j \in J$$

one tries to solve the perturbed KKT-system

$$E_\tau: \quad \begin{array}{rcl} \nabla f(x) - \nabla^T g(x)\mu &=& 0 \\ g_j(x)\mu_j &=& \tau, \quad \forall j \in J \end{array}$$

and $\mu_j, g_j(x) \geq 0$. This is a special case of a feasible set of a problem P_τ (including equality constraints). In Orban and Wright [20] the convergence behavior of solutions x_τ, μ_τ of E_τ has been analyzed (via properties of the log barrier function) also for the case that the strict complementarity condition (SC) is not satisfied at the solution $\overline{x}, \overline{\mu}$ of E_0. Here also a convergence rate $O(\sqrt{\tau})$ has been established (under the weaker MFCQ assumption). So the results of Section 5 can be seen as a generalization of (some of the) results in [20].

2. Motivating examples and regularity conditions

We begin with some illustrative examples and formulate regularity conditions to avoid some negative convergence behavior.

EXAMPLE 2.1 $\quad \min \ x_1 + x_2 \quad \text{s.t.} \quad x_1 \cdot x_2 = 0, \ \ x_1, x_2 \geq 0 \ .$

Here the set $\mathcal{F}_\tau$ *converges* to the set $\mathcal{F}$ and the solutions $\overline{x}_\tau = (\sqrt{\tau}, \sqrt{\tau})$ of P_τ converges to the solution $\overline{x} = 0$ of P with a rate $\|\overline{x}_\tau - \overline{x}\| = \sqrt{2} \cdot \sqrt{\tau}$ and $|\varphi_\tau - \varphi| = \sqrt{2} \cdot \sqrt{\tau}$.

EXAMPLE 2.2 $\quad \min \ (x_2 - 1)^2 \quad \text{s.t.} \quad x_2 \cdot e^{-x_1} = 0, \quad x_2, e^{-x_1} \geq 0 \ , g(x) := x_1 \geq 0.$

Here $\mathcal{F} = \{(x_1, 0) \mid x_1 \geq 0\}$ coincides with the set $\mathcal{S}$ of minimizers. The feasible set $\mathcal{F}_\tau = \{(x_1, \tau e^{x_1}) \mid x_1 \geq 0\}$ however does not *converge* to $\mathcal{F}$. The (unique) minimizer of P_τ is given by $\overline{x}_\tau = (-\ln \tau, 1)$, implying $d(\overline{x}_\tau, \mathcal{S}) \to \infty$. The problem here is that the feasible set is not compact.

In the next example (from a preliminary version of [16]) the perturbed feasible set $\mathcal{F}_\tau$ is *smaller* than $\mathcal{F}$.

EXAMPLE 2.3 $\quad \min \ (x_3 - 1)^2 + x_2^2 \quad \text{s.t.} \quad x_1 \cdot x_2 = 0, \ x_1 \cdot x_3 = 0, \quad x_1, x_2, x_3 \geq 0 \ .$

The minimizer is given by $\overline{x} = (0, 0, 1)$. The feasible set $\mathcal{F}_\tau$ is *smaller* than $\mathcal{F}$ and the (unique) minimizer $\overline{x}_\tau = (2\tau, 1/2, 1/2)$ does not converge to $\overline{x}$. The problem here is that the feasible set $\mathcal{F}$ does not satisfy MPEC-LICQ (at any point $(0, x_2, x_3) \in \mathcal{F}$, see (6)).

In the following example the feasible set $\mathcal{F}_\tau$ behaves well but the rate of convergence of $\|\overline{x}_\tau - \overline{x}\|$ is arbitrarily slow.

EXAMPLE 2.4 $\quad \min \ x_1^q + x_2 \quad \text{s.t.} \quad x_1 \cdot x_2 = 0, \quad x_1, x_2 \geq 0$

with $q > 0$. The minimizer $\overline{x} = (0, 0)$ of the problem and the solutions of P_τ, $\overline{x}_\tau = \left((\tau/q)^{1/(q+1)}, \ q^{1/(q+1)} \tau^{q/(q+1)}\right)$, show the convergence rate $\|\overline{x}_\tau - \overline{x}\| = O(\tau^{1/(q+1)})$.

In the sequel we are interested in the convergence behavior and the rate of convergence

$$\mathcal{F}_\tau \to \mathcal{F}, \quad \varphi_\tau \to \varphi \quad \text{and} \quad \overline{x}_\tau \to \overline{x} \quad \text{if} \quad \tau \to 0,$$

for the feasible sets, the value functions and the solutions of P and P_τ. To avoid the negative behavior in the Examples 2-4 we need some (natural) assumptions.

Firstly, motivated by Example 2.2, we assume throughout the paper that the feasible sets are compact. Note that in practice this does not mean a restriction since it is advisable to add (if necessary) to the constraints $g_j(x) \geq 0$, e.g. box constraints, $|x_\nu| \leq K, \ \nu = 1, \dots, n,$ for some large number $K > 0$. So, in the sequel we assume that for all $\tau \geq 0$

$$\mathcal{F}_\tau \subset X \qquad \text{where } X \subset \mathbb{R}^n \text{ is compact .} \tag{4}$$

Under this condition, in particular, global solutions of P and P_τ exist (unless the feasible set is empty). Moreover we assume throughout the paper that all functions f, g_j, r_i, s_i are from $C^2(X, \mathbb{R})$. Then, in particular, the functions are Lipschitz continuous on X, *i.e.*, there is some $L > 0$ such that

$$|f(\hat{x}) - f(x)| \leq L \cdot \|\hat{x} - x\| \qquad \forall \hat{x}, x \in X \ . \tag{5}$$

To avoid the bad behavior in Example 2.3 we have to assume a constraint qualification for the feasible set. To do so, for a point $x \in \mathcal{F}$ we define the active index sets $J(x) = \{j \in J \mid g_j(x) = 0\}$, $I_{rs}(x) = \{i \in I \mid r_i(x) = s_i(x) = 0\}$, $I_r(x) = \{i \in I \mid r_i(x) = 0, \ s_i(x) > 0\}$ and $I_s(x) = \{i \in I \mid r_i(x) > 0, s_i(x) = 0\}$. We say that at the feasible point $x \in \mathcal{F}$ the condition MPEC-LICQ holds, if the active gradients

$$\nabla g_j(x), \ j \in J(x), \ \ \nabla r_i(x), \ i \in I_{rs}(x) \cup I_r(x), \ \ \nabla s_i(x), \ i \in I_{rs}(x) \cup I_s(x) \tag{6}$$

are linearly independent.

As we shall see later on, this condition will imply that locally around x the set $\mathcal{F}_\tau$ converges to the set $\mathcal{F}$ with a rate $O(\sqrt{\tau})$. To assure the global convergence we have to assume that MPEC-LICQ holds globally, *i.e.*, that MPEC-LICQ is fulfilled at every point $x \in \mathcal{F}$. We emphasize that this assumption is generically fulfilled as will be shown in the next section. (*cf.* Theorem 3.1).

3. Optimality conditions for minimizer of P In this section we are interested in necessary and sufficient optimality conditions for local minimizers of P. New characterizations for minimizers of order one are given and known optimality conditions for solutions of order two (*cf., e.g.,* [12] and [16]) are extended. We also review the genericity results for problems P in [16] which will play an important role throughout the article.

Recall that $\bar{x} \in \mathcal{F}$ is said to be a local minimizer of P of order $\omega > 0$ if in a neighborhood $B_\varepsilon(\bar{x})$, $\varepsilon > 0$, of $\bar{x}$ with some $\kappa > 0$:

$$f(x) \geq f(\bar{x}) + \kappa \|x - \bar{x}\|^\omega \qquad \forall x \in \mathcal{F} \cap B_\varepsilon(\bar{x}) \ . \tag{7}$$

The point $\bar{x}$ is called a global minimizer of order ω if we can choose $\varepsilon = \infty$.

Perhaps the most natural way to obtain optimality conditions for P is to consider the MPEC problem as a problem which is piecewise built up by finitely many common finite programs. To this end let $\bar{x} \in \mathcal{F}$ be given. For any subset $I_0 \subset I_{rs}(\bar{x})$ we define $I_0^c = I_{rs}(\bar{x}) \setminus I_0$ and consider the common finite optimization problem

$$\begin{array}{llrcll} P_{I_0}(\bar{x}): & \min\limits_x f(x) \quad \text{s.t.} & g_j(x) & \geq & 0\,, & j \in J(\bar{x}) \\ & & r_i(x) = 0\,,\; s_i(x) & \geq & 0\,, & i \in I_0 \\ & & r_i(x) \geq 0\,,\; s_i(x) & = & 0\,, & i \in I_0^c \\ & & r_i(x) & = & 0\,, & i \in I_r(\bar{x}) \\ & & s_i(x) & = & 0\,, & i \in I_s(\bar{x}) \end{array} \tag{8}$$

With the feasible sets $\mathcal{F}_{I_0}(\bar{x})$ of $P_{I_0}(\bar{x})$, obviously, the following *piecewise (or disjunctive)* description holds (see also *e.g.* [12, Chapter 4], [15, p.6]).

LEMMA 3.1 *Let $\bar{x}$ be feasible for P. Then we have:*
(a) There exists a neighborhood $B_\varepsilon(\bar{x})$ ($\varepsilon > 0$) of $\bar{x}$ such that

$$\mathcal{F} \cap B_\varepsilon(\bar{x}) = \bigcup_{I_0 \subset I_{rs}(\bar{x})} \left(\mathcal{F}_{I_0}(\bar{x}) \cap B_\varepsilon(\bar{x}) \right) \ .$$

(b) The point $\bar{x} \in \mathcal{F}$ is a local minimizer of order ω of P if and only if $\bar{x}$ is a local minimizer of order ω of $P_{I_0}(\bar{x})$ for all $I_0 \subset I_{rs}(\bar{x})$.

By this lemma, all optimality conditions and genericity results for the common problems $P_{I_0}(\bar{x})$ directly lead to corresponding results for the complementarity constrained program P. To do so, let us recall some notation. $C_{I_0}(\bar{x})$ denotes the cone of *critical directions* for $P_{I_0}(\bar{x})$ at $\bar{x}$,

$$C_{I_0}(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \begin{array}{rclrcll} \nabla f(\bar{x})d & \leq & 0, & \nabla g_j(\bar{x})d & \geq & 0, & j \in J(\bar{x}) \\ \nabla r_i(\bar{x})d & = & 0, & \nabla s_i(\bar{x})d & \geq & 0, & i \in I_0 \\ \nabla r_i(\bar{x})d & \geq & 0, & \nabla s_i(\bar{x})d & = & 0, & i \in I_0^c \\ & & & \nabla r_i(\bar{x})d & = & 0, & i \in I_r(\bar{x}) \\ & & & \nabla s_i(\bar{x})d & = & 0, & i \in I_s(\bar{x}) \end{array} \right\} \tag{9}$$

The point $\bar{x} \in \mathcal{F}_{I_0}(\bar{x})$ is called a *Karush-Kuhn-Tucker point (KKT point)* for $P_{I_0}(\bar{x})$ if there exist multipliers γ, ρ, σ such that

$$\begin{aligned} \nabla_x L(\bar{x}, \gamma, \rho, \sigma) \quad := \quad & \nabla f(\bar{x}) - \sum_{j \in J(\bar{x})} \gamma_j \nabla g_j(\bar{x}) - \sum_{i \in I_{rs}(\bar{x})} \left[\rho_i \nabla r_i(\bar{x}) + \sigma_i \nabla s_i(\bar{x}) \right] \\ - \quad & \sum_{i \in I_r(\bar{x})} \rho_i \nabla r_i(\bar{x}) - \sum_{i \in I_s(\bar{x})} \sigma_i \nabla s_i(\bar{x}) \ = \ 0 \end{aligned} \tag{10}$$

$$\text{and} \qquad \gamma_j \geq 0,\ j \in J(\bar{x}),\ \rho_i \geq 0,\ i \in I_0^c, \sigma_i \geq 0,\ i \in I_0 \ , \tag{11}$$

where L denotes the *Lagrange function* as usual. The vector $(\bar{x}, \gamma, \rho, \sigma)$ is then called a *KKT solution* of $P_{I_0}(\bar{x})$ and the *strict complementary slackness* is said to hold if

$$(\text{SC}_{I_0}(\bar{x})): \qquad \gamma_j > 0,\ j \in J(\bar{x}), \quad \rho_i > 0,\ i \in I_0^c, \quad \sigma_i > 0,\ i \in I_0 \ ,$$

and the *second order condition* if

$$(\text{SOC}_{I_0}(\bar{x})): \qquad d^T \nabla_x^2 L(\bar{x}, \gamma, \rho, \sigma) d > 0 \qquad \forall d \in C_{I_0}(\bar{x}) \setminus \{0\} \ . \tag{12}$$

We now introduce some notation for P. We define

$$C_{\bar{x}} = \bigcup_{I_0 \subset I_{rs}(\bar{x})} C_{I_0}(\bar{x}) \tag{13}$$

and call $\bar{x} \in \mathcal{F}$ a MPEC-KKT *point* of P if $\bar{x}$ is a KKT point of $P_{I_0}(\bar{x})$ for all $I_0 \subset I_{rs}(\bar{x})$. A vector $(\bar{x}, \gamma, \rho, \sigma)$ is said to be a MPEC-KKT *solution* of P if it is a KKT solution of $P_{I_0}(\bar{x})$ for all $I_0 \subset I_{rs}(\bar{x})$. Note that for a MPEC-KKT solution of P from (11) it follows that

$$(10) \quad \text{holds with} \quad \gamma_j \geq 0,\ j \in J(\bar{x}), \quad \rho_i \geq 0\ ,\ \sigma_i \geq 0,\ i \in I_{rs}(\bar{x}) \ . \tag{14}$$

We say that such a MPEC-KKT solution satisfies the *strict complementary slackness* for MPEC if

$$(\text{MPEC-SC}\ ): \qquad \gamma_j > 0,\ j \in J(\bar{x}), \quad \rho_i > 0, \sigma_i > 0,\ i \in I_{rs}(\bar{x}) \ . \tag{15}$$

and the *second order condition* for MPEC if

$$(\text{MPEC-SOC}\ ): \qquad d^T \nabla_x^2 L(\bar{x}, \gamma, \rho, \sigma) d > 0 \qquad \forall d \in C_{\bar{x}} \setminus \{0\} \ . \tag{16}$$

Note that (wrt. the conditions for ρ_i, σ_i in (15)) the condition SC (*i.e.,* $I_{rs}(\bar{x}) = \emptyset$) is stronger than MPEC-SC . By definition, the condition MPEC-LICQ at $\bar{x}$ means that the common LICQ condition holds at $\bar{x}$ for all problems $P_{I_0}(\bar{x})$.

REMARK 3.1 In the context of MPEC problems there are different concepts of stationarity (or Fritz-John-, KKT-points) (see *e.g.* [15]). We emphasize that all these concepts coincide if the MPEC-LICQ assumption holds at $\overline{x}$ (even the weaker SMFCQ). In this case: $\overline{x}$ is a MPEC-KKT point $\Leftrightarrow$ $\overline{x}$ is a B-stationary point $\Leftrightarrow$ $\overline{x}$ is a strong stationary point (*cf.* [15, Th.4]). Therefore in this paper we will use the term MPEC-KKT point.

If at a MPEC-KKT point $\overline{x}$ the condition MPEC-LICQ holds then there is a unique corresponding MPEC-KKT solution $(\overline{x}, \gamma, \rho, \sigma)$ (same unique multipliers γ, ρ, σ for P and all $P_{I_0}(\overline{x})$). Moreover it is not difficult to see that in this case the set $C_{\overline{x}}$ simplifies (see (9) and (13)) to the cone:

$$C_{\overline{x}} = \left\{ d \;\middle|\; \begin{array}{ccccccl} \nabla g_j(\overline{x})d & \begin{array}{c} = \\ \geq \end{array} & 0 & \text{if} & \gamma_j & \begin{array}{c} > \\ = \end{array} & 0, \quad j \in J(\overline{x}) \\ \nabla r_i(\overline{x})d & \begin{array}{c} = \\ \geq \end{array} & 0 & \text{if} & \rho_i & \begin{array}{c} > \\ = \end{array} & 0, \quad i \in I_{rs}(\overline{x}) \\ \nabla s_i(\overline{x})d & \begin{array}{c} = \\ \geq \end{array} & 0 & \text{if} & \sigma_i & \begin{array}{c} > \\ = \end{array} & 0, \quad i \in I_{rs}(\overline{x}) \\ & & & & \nabla r_i(\overline{x})d & = & 0, \quad i \in I_r(\overline{x}) \\ & & & & \nabla s_i(\overline{x})d & = & 0, \quad i \in I_s(\overline{x}) \end{array} \right\} \qquad (17)$$

We now sketch some genericity results for problem P. Let in the sequel all functions f, g_j, s_i, r_i be in the space $C^2(\mathbb{R}^n, \mathbb{R})$ endowed with the C^2_s-topology (strong topology, cf. Guddat *et.al.* [7, p.23]). Then, for fixed n, m, q, the set of all problems P can be identified with the set $\mathcal{P} := \{(f, g, s, r)\} \equiv C^2(\mathbb{R}^n, \mathbb{R})^{q+2m+1}$. We say that a property holds generically for P if it holds for a (in the C^2_s-topology) dense and open subset $\mathcal{P}_0$ of $\mathcal{P}$. From the well-known genericity results for the problems $P_{I_0}(\overline{x})$ (see Guddat *et.al.* [7]) we directly obtain via the *piecewise formulation* in Lemma 3.1 the following genericity results (see also Scholtes and Stöhr [16]).

THEOREM 3.1 *There is a dense and open (generic) subset $\mathcal{P}_0$ of $\mathcal{P}$ such that for all MPEC problems $P \in \mathcal{P}_0$ the following holds. For any feasible point $x \in \mathcal{F}$ the condition* MPEC-LICQ *is satisfied and for any local minimizer $\overline{x}$ of P the conditions* MPEC-SC *and* MPEC-SOC *are fulfilled.*

REMARK 3.2 We shortly comment on the genericity concept. A generic subset $\mathcal{P}_0$ of $\mathcal{P}$ is an open and dense subset. Dense means that any MPEC problem from $\mathcal{P}$ can be approximated arbitrarily well by a problem in the (*nice*) generic set $\mathcal{P}_0$. The openness implies stability, *i.e.*, if we have given a problem P from the generic set $\mathcal{P}_0$ then all sufficiently small C^2_s-perturbations of P remain in the set $\mathcal{P}_0$. In other words when dealing with a MPEC problem theoretically or numerically we *can expect* (generically) that the problem has the structure of a problem in the (nice) generic set and a general purpose solver for MPEC should be designed in such a way that it is able to deal (at least) with all situations encountered by problems in the generic set $\mathcal{P}_0$. A problem which is not in the generic set can be seen as an exceptional case.

As an example of a typical genericity result, it can be shown that generically the Newton-method can be applied to solve nonlinear equations $F(x) = 0$ (see [7, Chapter 2]) in the following sense: For a generic set of functions $F \in C^2(\mathbb{R}^n, \mathbb{R}^n)$ the regularity conditions $\det(\nabla F(\overline{x})) \neq 0$ holds at all solutions $\overline{x}$ of the equation $F(\overline{x}) = 0$.

We now give some optimality conditions for MPEC problems P (see also [12], [15]). It is well-known (see [15, Th.2,Lem.2] that any minimizer of (1) which satisfies MPEC-LICQ (or the weaker SMFCQ) must necessarily be a MPEC-KKT point.

From the *piecewise description* of P we obtain the following characterizations for minimizers of order one. In the context of MPEC problems these results are new.

THEOREM 3.2 (primal conditions of order 1) *For a point $\overline{x}$ which is feasible for P:*

$$C_{\overline{x}} = \{0\} \quad \Rightarrow \quad \overline{x} \textit{ is a (isolated) local minimizer of order } \omega = 1 \textit{ of } P.$$

If MPEC-LICQ *holds at $\overline{x}$, also the converse is true.*

PROOF. It is well-known (see *e.g.* Still and Streng [19, Th.3.2, Th.3.6]) that $C_{I_0}(\overline{x}) = \{0\}$ implies that $\overline{x}$ is a (isolated) local minimizer of order 1 of $P_{I_0}(\overline{x})$ and under LICQ the converse holds. Recall that MPEC-LICQ coincides with the common LICQ condition for $P_{I_0}(\overline{x})$. With regard to the definition of $C_{\overline{x}}$ in (13) the result follows from Lemma 3.1. □

THEOREM 3.3 (dual conditions of order 1) *Let* MPEC-LICQ *hold at $\overline{x} \in \mathcal{F}$. Then $\overline{x}$ is a (isolated) local minimizer of order $\omega = 1$ of P if and only if one of the following equivalent conditions (a) or (b) is satisfied :*

(a) $\nabla f(\overline{x}) \in \operatorname{int} Q_{\overline{x}}$ *where*

$$Q_{\overline{x}} = \left\{ d = \sum_{j \in J(\overline{x})} \gamma_j \nabla g_j(\overline{x}) + \sum_{i \in I_{rs}(\overline{x})} \left[\rho_i \nabla r_i(\overline{x}) + \sigma_i \nabla s_i(\overline{x})\right] + \sum_{i \in I_r(\overline{x})} \rho_i \nabla r_i(\overline{x}) \right.$$
$$\left. + \sum_{i \in I_s(\overline{x})} \sigma_i \nabla s_i(\overline{x}) \quad , \quad \gamma_j \geq 0,\; j \in J(\overline{x}),\; \rho_i \geq 0, \sigma_i \geq 0,\; i \in I_{rs}(\overline{x}) \right\}.$$

(b) The vector $\overline{x}$ is a MPEC-KKT *point with (unique) multipliers γ, ρ, σ such that $|J(\overline{x})| + 2|I_{rs}(\overline{x})| + |I_r(\overline{x})| + |I_s(\overline{x})| = n$ and $\gamma_j > 0,\; j \in J(\overline{x}),\; \rho_i > 0, \sigma_i > 0,\; i \in I_{rs}(\overline{x})$,* i.e., MPEC-SC *holds.*

PROOF. It is well-known (*cf. e.g.* [19]) that the primal condition $C_{I_0}(\overline{x}) = \{0\}$ is equivalent with the condition $\nabla f(\overline{x}) \in \text{int}\, Q_{I_0}(\overline{x})$ where

$$\begin{aligned} Q_{I_0}(\overline{x}) \quad = \quad & \Big\{ d = \sum_{j\in J(\overline{x})} \gamma_j \nabla g_j(\overline{x}) + \sum_{i\in I_{rs}(\overline{x})} \big[\rho_i \nabla r_i(\overline{x}) + \sigma_i \nabla s_i(\overline{x})\big] + \sum_{i\in I_r(\overline{x})} \rho_i \nabla r_i(\overline{x}) \\ & + \sum_{i\in I_s(\overline{x})} \sigma_i \nabla s_i(\overline{x}) \quad , \quad \gamma_j \geq 0,\ j\in J(\overline{x}),\ \rho_i \geq 0,\ i\in I_0^c,\ \sigma_i\geq 0,\ i\in I_0 \Big\}. \end{aligned}$$

By Lemma 3.1 this yields (a).
(b) We now prove under MPEC-LICQ : (a) ⇔ (b). Note that the direction "⇐" is evident. To prove the converse let us assume that $\nabla f(\overline{x}) \in \text{int}\, Q_{\overline{x}}$ but $|J(\overline{x})| + 2|I_{rs}(\overline{x})| + |I_r(\overline{x})| + |I_s(\overline{x})| < n$. The latter means that there exists $d \in \mathbb{R}^n$ such that $d \notin S_0$,

$$S_0 := \text{span}\,\{\nabla g_j(\overline{x}),\ j\in J(\overline{x}),\ \nabla r_i(\overline{x}),\ i\in I_{rs}(\overline{x})\cup I_r(\overline{x}),\ \nabla s_i(\overline{x}),\ i\in I_{rs}(\overline{x})\cup I_s(\overline{x})\}.$$

Note that since in particular $\overline{x}$ is a MPEC-KKT point it follows $S_0 = \text{span}\,\{\{-\nabla f(\overline{x})\}\cup S_0\}$. Consequently for any $\varepsilon > 0$, $\varepsilon d \notin \text{span}\,\{\{-\nabla f(\overline{x})\}\cup S_0\}$ and thus $\nabla f(\overline{x}) + \varepsilon d \notin \text{span}\, S_0$ in contradiction to (a). Let us now assume that MPEC-SC does not hold, say $\gamma_1 = 0$. Then by MPEC-LICQ for any $\varepsilon > 0$ the vector $\nabla f(\overline{x}) - \varepsilon \nabla g_1(\overline{x})$ is not contained in $Q_{\overline{x}}$ a contradiction to (a). □

We now give a characterization of minimizers of order two. We refer to [15] for similar necessary and sufficient conditions (under weaker assumptions).

THEOREM 3.4 (dual conditions of order 2) *Let* MPEC-LICQ *hold at* $\overline{x}\in\mathcal{F}$ *and assume* $C_{\overline{x}} \neq \{0\}$ *(*i.e., *in view of Theorem 3.2,* $\overline{x}$ *is not a local minimizer of order 1). Then* $\overline{x}$ *is a (isolated) local minimizer of order* $\omega = 2$ *of* P *if and only if* $\overline{x}$ *is a* MPEC-KKT *point of* P *such that with (unique) multipliers* γ, ρ, σ *the condition* MPEC-SOC *holds.*
(Under this condition $\overline{x}$ *is locally the unique* MPEC-KKT *point of* P.*)*

PROOF. $C_{\overline{x}} \neq \{0\}$ implies $C_{I_0}(\overline{x}) \neq \{0\}$ for (at least) one set $I_0 \subset I_{rs}(\overline{x})$ so that $\overline{x}$ is not a local minimizer of order one of $P_{I_0}(\overline{x})$ (see the proof of Theorem 3.2) and thus not of P (*cf.* Lemma 3.1). By [19, Th.3.6] (under MPEC-LICQ) (for any $I_0 \subset I_{rs}(\overline{x})$) $\overline{x}$ is a (isolated) local minimizer of order 2 of $P_{I_0}(\overline{x})$ iff $\overline{x}$ is a KKT point for $P_{I_0}(\overline{x})$ satisfying (12). Under this condition, by [19, Th.3.23], $\overline{x}$ is locally the unique KKT point of $P_{I_0}(\overline{x})$. Again the result follows from Lemma 3.1. □

Note that in view of the genericity result in Theorem 3.1 we can state:

$$\text{generically each local minimizer of } P \text{ has either order } \omega = 1 \text{ or order } \omega = 2. \tag{18}$$

It is interesting to mention that with the common finite problem (a relaxation of P)

$$P_R(\overline{x}): \qquad \min_x\ f(x) \quad \text{s.t.} \quad \begin{array}{rcll} g_j(x) & \geq & 0\,, & j\in J(\overline{x}) \\ r_i(x)\geq 0\,,\ s_i(x) & \geq & 0\,, & i\in I_{rs}(\overline{x}) \\ r_i(x) & = & 0\,, & i\in I_r(\overline{x}) \\ s_i(x) & = & 0\,, & i\in I_s(\overline{x}) \end{array} \tag{19}$$

the following is true (*cf.* also [15]).

COROLLARY 3.1 *Let* MPEC-LICQ *hold at* $\overline{x}\in\mathcal{F}$. *Then* $\overline{x}$ *is a local minimizer of order* $\omega = 1$ *or* $\omega = 2$ *of* P *if and only if* $\overline{x}$ *is a local minimizer of order* $\omega = 1$ *or* $\omega = 2$ *of* $P_R(\overline{x})$.
(Recall that generically each local minimizer of P is either of order 1 or of order 2.)

PROOF. Under MPEC-LICQ any local minimizer $\overline{x}$ of P must be a MPEC-KKT point of P with unique multipliers γ, ρ, σ. Note that by (14) $(\overline{x}, \gamma, \rho, \sigma)$ is also a KKT solution of $P_R(\overline{x})$ with the same Lagrange function $L(x, \gamma, \rho, \sigma)$. Moreover the set of critical directions for $P_R(\overline{x})$ coincides with $C_{\overline{x}}$ (see (17)). So the first order optimality condition $C_{\overline{x}} = \{0\}$ (*cf.* Theorem 3.2) and the second order conditions (*cf.* Theorem 3.4) for P and $P_R(\overline{x})$ coincide. □

4. The convergence behavior of the feasible set $\mathcal{F}_\tau$ In this section we consider the convergence behavior of the feasible set $\mathcal{F}_\tau$ from a local and global viewpoint. The local convergence relies on a local MPEC-LICQ assumption and the global results are proven under a global assumption.

We begin with an auxiliary result.

LEMMA 4.1 *For* $x_\tau \in \mathcal{F}_\tau$ *and* $\tau \to 0$ *it follows* $d(x_\tau, \mathcal{F}) \to 0$ *uniformly: To any* $\varepsilon > 0$ *there exists* $\tau_0 > 0$ *such that for all* $0 < \tau \leq \tau_0$ *and for all* $x_\tau \in \mathcal{F}_\tau$ *the bound* $d(x_\tau, \mathcal{F}) < \varepsilon$ *holds.*

PROOF. Assuming that the statement is not true, there must exist $\gamma > 0$ and a sequence $x_\tau \in \mathcal{F}_\tau$ such that for $\tau \to 0$, $d(x_\tau, \mathcal{F}) \geq \gamma$.

Due to the compactness assumption (4) we can choose a convergent subsequence $x_{\tau_\nu} \to \overline{x} \in X$. The condition $r_i(x_{\tau_\nu})s_i(x_{\tau_\nu}) = \tau_\nu$, $g_j(x_{\tau_\nu}) \geq 0$ together with the continuity of the functions r_i, s_i, g_j leads for $\tau_\nu \to 0$ to $r_i(\overline{x})s_i(\overline{x}) = 0$ and $g_j(\overline{x}) \geq 0$, *i.e.*, $\overline{x}\in\mathcal{F}$, a contradiction. □

To prove our main results on the behavior of $\mathcal{F}_\tau$ we make use of a local (local) diffeomorphism. The idea is to transform the problem into an equivalent problem with simpler structure so that the proofs of the results become technically much simpler. However this approach relies on the MPEC-LICQ assumption. Such a transformation has been mentioned in [16] to illustrate the local behavior of $\mathcal{F}_\tau$. Here we present a complete global analysis.

Consider a point $\overline{x} \in \mathcal{F}$ satisfying MPEC-LICQ with $|J(\overline{x})| = q_0,\ |I_{rs}(\overline{x})| = p$ where $p \le m, q_0 \le q$ and $m + p + q_0 \le n$. Wlog. we can assume

$$J(\overline{x}) = \{1, \dots, q_0\},\ I_{rs}(\overline{x}) = \{1, \dots, p\},\ I_r(\overline{x}) = \{p+1, \dots, m\}\ ,\ I_s(\overline{x}) = \emptyset\ .$$

By MPEC-LICQ the gradients $\nabla g_j(\overline{x}),\ j \in J(\overline{x}),\ \nabla r_i(\overline{x}),\ i = 1, \dots, m,\ \nabla s_i(\overline{x}),\ i = 1, \dots, p$, are linearly independent and we can complete these vectors to a basis of $\mathbb{R}^n$ by adding vectors $v_i,\ i = m+p+q_0+1, \dots, n$. Now we define the transformation $y = T(x)$ by

$$\begin{array}{rcll} y_i &=& r_i(x), & i = 1, \dots, m, \\ y_{i+m} &=& s_i(x), & i = 1, \dots, p, \\ y_{m+p+i} &=& g_i(x), & i = 1, \dots, q_0, \\ y_i &=& v_i^T(x - \overline{x}), & i = m+p+q_0+1, \dots, n\ . \end{array} \tag{20}$$

By construction, the Jacobian $\nabla T(\overline{x})$ is regular and T defines locally a diffeomorphism. This means that there exists $\varepsilon = \varepsilon(\overline{x}) > 0$ and neighborhoods $B_\varepsilon(\overline{x})$ of $\overline{x}$ and $U_\varepsilon(\overline{y}) := T(B_\varepsilon(\overline{x}))$ of $\overline{y} = 0$ such that $T : B_\varepsilon(\overline{x}) \to U_\varepsilon(\overline{y})$ is a bijective mapping with $T, T^{-1} \in C^1$, $T(\overline{x}) = \overline{y}$ and for $y = T(x)$ it follows:

$$x \in \mathcal{F}_\tau \cap B_\varepsilon(\overline{x}) \quad \Leftrightarrow \quad \begin{array}{rcll} y_{m+p+j} &\ge& 0 & j = 1, \dots, q_0 \\ y_i \cdot y_{m+i} &=& \tau & i = 1, \dots, p \\ y_i \cdot \widetilde{s}_i(y) &=& \tau & i = p+1, \dots, m \\ y_i &\ge& 0 & i = 1, \dots, m \\ y_{m+i} &\ge& 0 & i = 1, \dots, p \end{array} \quad , \quad y \in U_\varepsilon(\overline{y})\ ,$$

where $\widetilde{s}_i(\overline{y}) := s_i(T^{-1}(\overline{y})) = s_i(\overline{x}) > 0,\ i = p+1, \dots, m$ and $\widetilde{g}_j(\overline{y}) := g_j(T^{-1}(\overline{y})) = g_j(\overline{x}) > 0,\ j = q_0+1, \dots, q$.

In particular, since T is a diffeomorphism, the distance between two points remains equivalent in the sense that with constants $0 < \kappa_- < \kappa_+$:

$$\kappa_- \|y_1 - y_2\| \le \|x_1 - x_2\| \le \kappa_+ \|y_1 - y_2\| \quad \forall x_1, x_2 \in B_\varepsilon(\overline{x}),\ \ y_1 = T(x_1),\ y_2 = T(x_2)\ .$$

So (after applying a diffeomorphism T) we may assume $\overline{x} = 0$,

$$\begin{array}{rcll} g_j(x) &=& x_{m+p+j} & j = 1, \dots, q_0 \\ r_i(x) &=& x_i & i = 1, \dots, m \\ s_i(x) &=& x_{m+i} & i = 1, \dots, p\ , \end{array} \tag{21}$$

$$c_i^s := s_i(\overline{x}) > 0\ , \quad i = p+1, \dots, m \tag{22}$$

and that there is some $\varepsilon > 0$ such that

$$x \in \mathcal{F}_\tau \cap B_\varepsilon(\overline{x}) \quad \Leftrightarrow \quad \begin{array}{rcll} g_j(x) = x_{m+p+j} &\ge& 0 & j = 1, \dots, q_0 \\ h_i(x) = x_i \cdot x_{m+i} &=& \tau & i = 1, \dots, p \\ h_i(x) = x_i \cdot s_i(x) &=& \tau & i = p+1, \dots, m \\ x_i &\ge& 0 & i = 1, \dots, m \\ x_{m+i} &\ge& 0 & i = 1, \dots, p\ . \end{array} \tag{23}$$

By choosing ε small enough we also can assume

$$s_i(x) \ge \frac{c_i^s}{2}\ , \quad i = p+1, \dots, m, \quad \forall x \in B_\varepsilon(\overline{x})\ . \tag{24}$$

By making use of the previously described transformation we are now able to prove the local convergence result for $\mathcal{F}_\tau$.

LEMMA 4.2 *Let* MPEC-LICQ *hold at* $\overline{x} \in \mathcal{F}$.
(a) Then there exist $\varepsilon, \tau_0, \alpha, \beta > 0$ *such that for all* $0 < \tau \le \tau_0$ *the following holds: There exist* $\overline{x}_\tau \in \mathcal{F}_\tau$ *with*

$$\|\overline{x}_\tau - \overline{x}\| \le \alpha\sqrt{\tau} \tag{25}$$

and for any $\overline{x}_\tau \in \mathcal{F}_\tau \cap B_\varepsilon(\overline{x})$ *there exists a point* $\hat{x}_\tau \in \mathcal{F} \cap B_\varepsilon(\overline{x})$ *satisfying*

$$\|\hat{x}_\tau - \overline{x}_\tau\| \le \beta\sqrt{\tau}\ . \tag{26}$$

Moreover, if SC holds at $\overline{x}$ *the statements are true with* $\sqrt{\tau}$ *replaced by* τ.
(b) If the condition SC is not *fulfilled at* $\overline{x}$ *then the convergence rate* $O(\sqrt{\tau})$ *in (25) is optimal. More precisely, there is some* $\gamma > 0$ *such that for all* $\overline{x}_\tau \in \mathcal{F}_\tau$ *the relation* $\|\overline{x}_\tau - \overline{x}\| \ge \gamma\sqrt{\tau}$ *holds for all small* τ.

PROOF. (a) Let MPEC-LICQ hold at $\overline{x} \in \mathcal{F}$. As discussed before (after applying a diffeomorphism) we can assume that $\overline{x} = 0$ and that in a neighborhood $B_\varepsilon(\overline{x})$ of $\overline{x}$ the set $B_\varepsilon(\overline{x}) \cap \mathcal{F}_\tau$ is described by (23). To construct a suitable element $x^\tau \in \mathcal{F}_\tau$ we fix the components $x_i^\tau = x_{m+i}^\tau = \sqrt{\tau},\ i = 1, \dots, p$ and $x_i^\tau = 0,\ i = m+p+1, \dots, n$. From (23) we then find

$$\begin{array}{rcll} g_j(x^\tau) &=& 0 & j = 1, \dots, q_0 \\ h_i(x^\tau) &=& \tau & i = 1, \dots, p \\ h_i(x^\tau) = x_i^\tau \cdot s_i(x^\tau) &=& \tau & i = p+1, \dots, m\ , \end{array}$$

where the first two relations are already satisfied. So, we only need to consider the remaining equations

$$h_i(\widetilde{x}) := x_i^\tau \cdot s_i(x^\tau) = \tau, \quad i = p+1, \dots, m \tag{27}$$

which (for fixed τ) only depend on the remaining variables $\widetilde{x} = (x_{p+1}^\tau, \dots, x_m^\tau)$. For $\widetilde{x} = 0$ the gradients $\nabla h_i(0) = e_i s_i(0) = e_i c_i^s$, $i = p+1, \dots, m$ (*cf.* (22)), are linearly independent. As usual e_i denote the unit vectors. So, the function $h : \mathbb{R}^{m-p} \to \mathbb{R}^{m-p}$, $h = (h_{p+1}, \dots, h_m)$, $h(0) = 0$ has locally near $\widetilde{x} = 0$ a C^1-inverse such that (for small τ) the vector $\widetilde{x}^\tau := h^{-1}(e\,\tau)$ (with $(e = (1, \dots, 1) \in \mathbb{R}^{m-p})$ defines a solution of (27). Because of $h^{-1}(0) = 0$ it follows $\|\widetilde{x}^\tau\| = O(\tau)$.

Altogether, with the other fixed components x_i^τ, this vector $\widetilde{x}^\tau$ defines a feasible point $x^\tau \in \mathcal{F}_\tau$ which satisfies

$$\|x^\tau - \overline{x}\| \le O(\sqrt{\tau}) \ .$$

We now prove (26). As shown above (*cf.* (23)) for some $\varepsilon > 0$ the point $\overline{x}_\tau \in B_\varepsilon(\overline{x})$ is in $\mathcal{F}_\tau$ if and only if $x := \overline{x}_\tau$ satisfies the relations

$$\begin{array}{rcll} g_j(x) = x_{m+p+j} & \ge & 0 & j = 1, \dots, q_0 \\ h_i(x) = x_i \cdot x_{m+i} & = & \tau & i = 1, \dots, p \\ h_i(x) = x_i \cdot s_i(x) & = & \tau & i = p+1, \dots, m \ . \end{array}$$

Obviously, $\min\{x_i, x_{m+i}\} \le \sqrt{\tau}$, $i = 1, \dots, p$, so that wlog. $x_i \le \sqrt{\tau}$, $i = 1, \dots, p$. By (24) for $x = \overline{x}_\tau \in B_\varepsilon(\overline{x})$ it follows

$$x_i = \frac{\tau}{s_i(x)} \le \frac{\tau}{c_i^s/2} \ , \qquad i = p+1, \dots, m \ . \tag{28}$$

Given this element $x = \overline{x}_\tau \in \mathcal{F}_\tau$ we now choose the point $\hat{x}_\tau$ of the form $\hat{x}_\tau = (0, \dots, 0, x_{m+1}, \dots, x_n)$ which is contained in $\mathcal{F}$. By using (28) and $x_i \le \sqrt{\tau}$, $i = 1, \dots, p$, and by putting $c_s = \min\{c_i^s/2, \ i = p+1, \dots, m\}$ we find $(x = \overline{x}_\tau)$

$$\|\hat{x}_\tau - \overline{x}_\tau\| \le \sqrt{p\,\tau + (m-p)\,\frac{\tau^2}{c_s^2}} \le O(\sqrt{\tau}) \ .$$

Let now SC be satisfied at $\overline{x} \in \mathcal{F}$ (see (2)). Then locally in $B_\varepsilon(\overline{x})$ the set $\mathcal{F}_\tau$ is defined by $(\overline{x} = 0)$

$$\begin{array}{rcll} g_j(x) = x_{m+j} & \ge & 0 & j = 1, \dots, q_0 \\ x_i \cdot s_i(x) & = & \tau & i = 1, \dots, m \ , \end{array} \tag{29}$$

where $s_i(x) \ge c_i^s/2$ for all $x \in B_\varepsilon(\overline{x})$. As in the first part of the proof we can fix the coefficients of x^τ by $x_i^\tau = \overline{x}_i \ (= 0)$, $i = m+1, \dots, n$, and find a solution $x = x^\tau \in \mathcal{F}_\tau$ by applying the Inverse Function Theorem to the remaining m equations

$$h_i(\widetilde{x}) := x_i s_i(x) = \tau \ , \quad i = 1, \dots, m \ ,$$

only depending on the remaining variables $\widetilde{x} := (x_1, \dots, x_m)$. This provides us with a solution x^τ of (29) satisfying

$$\|x^\tau - \overline{x}\| = O(\tau) \ .$$

On the other hand for any solution $x := \overline{x}_\tau$ of (29) in $B_\varepsilon(\overline{x})$ the point $\hat{x}_\tau = (0, \dots, 0, \ x_{m+1}, \dots, x_n)$ is an element in $\mathcal{F}$ with $\|\hat{x}_\tau - \overline{x}_\tau\| = O(\tau)$.

(b) Suppose now that SC is not fulfilled at $\overline{x}$, *i.e.*, for some $i_0 \in \{1, \dots, m\}$ (see (a))

$$h_i(\overline{x}) = \overline{x}_{i_0} \cdot \overline{x}_{m+i_0} = 0 \quad \text{with} \quad \overline{x}_{i_0} = \overline{x}_{m+i_0} = 0 \ .$$

Then near $\overline{x}$ any point $x^\tau \in \mathcal{F}_\tau$ must satisfy $x_{i_0}^\tau \cdot x_{m+i_0}^\tau = \tau$ which implies $(\overline{x} = 0)$

$$\|x^\tau - \overline{x}\| \ge \max\{x_{i_0}^\tau \ , \ x_{m+i_0}^\tau\} \ge \sqrt{\tau} \ .$$

Recall that (because of the diffeomorphism applied) this inequality only holds up to a constant $\gamma > 0$. □

Lemma 4.2 yields the local convergence of $\mathcal{F}_\tau$ near a point $\overline{x} \in \mathcal{F}$. We now are interested in the global convergence behavior (on the whole compact set X, *cf.* (4)).

LEMMA 4.3 *Let* MPEC-LICQ *hold at each point* $\overline{x} \in \mathcal{F}$. *Then there are* $\tau_0, \alpha, \beta > 0$ *such that for all* $0 < \tau \le \tau_0$ *the following holds: For each* $\overline{x} \in \mathcal{F}$ *there exists* $\overline{x}_\tau \in \mathcal{F}_\tau$ *with*

$$\|\overline{x}_\tau - \overline{x}\| \le \alpha\sqrt{\tau} \tag{30}$$

and for any $\overline{x}_\tau \in \mathcal{F}_\tau$ *there exists a point* $\hat{x}_\tau \in \mathcal{F}$ *satisfying*

$$\|\hat{x}_\tau - \overline{x}_\tau\| \le \beta\sqrt{\tau} \ . \tag{31}$$

Moreover, if SC holds at all $\overline{x} \in \mathcal{F}$ *the statements are true with* $\sqrt{\tau}$ *replaced by* τ.

PROOF. We firstly prove (31). To extend the analysis from the local to a global statement we have to apply a compactness argument. Recall the local transformation constructed above near any point $\overline{x} \in \mathcal{F}$ (see 23)). The union $\cup_{\overline{x} \in \mathcal{F}} B_{\varepsilon(\overline{x})}(\overline{x})$ forms an open cover of the compact feasible set $\mathcal{F} \subset X$. Consequently, by definition of compactness, we can choose a finite cover, *i.e.*, points $x_\nu \in \mathcal{F}$, $\nu = 1, \dots, N$, such that with $\varepsilon_\nu = \varepsilon(x_\nu)$ the set $\cup_{\nu=1,\dots,N} B_{\varepsilon_\nu}(x_\nu)$ provides an open cover of $\mathcal{F}$ and with $\beta_\nu > 0$

the corresponding condition (26) holds. By defining $B_\varepsilon(\mathcal{F}) = \{x \in X \mid d(x, \mathcal{F}) < \varepsilon\}$ we can choose some $\varepsilon_0 > 0$ (small) such that

$$B_{\varepsilon_0}(\mathcal{F}) \subset \bigcup_{\nu=1,\dots,N} B_{\varepsilon_\nu}(x_\nu)$$

By choosing $\varepsilon = \varepsilon_0$ and τ_0 in Lemma 4.1 we find for all $0 \le \tau \le \tau_0$:

$$\mathcal{F}_\tau \subset B_{\varepsilon_0}(\mathcal{F}) \subset \bigcup_{\nu=1,\dots,N} B_{\varepsilon_\nu}(x_\nu)\ .$$

The second convergence result (31) now directly follows by combining the finite cover argument with the local convergence and by noticing that we can choose as convergence constant the number $\beta = \min\{\beta_\nu\ ,\ \nu = 1, \dots, N\}$.

To prove (30) we have to show that the following sharpening of the local bound (25) holds: For $\overline{x} \in \mathcal{F}$ there exist $\tau_0 > 0$, $\varepsilon > 0$ such that for any $x \in \mathcal{F} \cap B_\varepsilon(\overline{x})$ and for any $0 \le \tau \le \tau_0$ there is a point $x_\tau \in \mathcal{F}_\tau$ with

$$\|x_\tau - x\| \le \alpha\sqrt{\tau}\ . \tag{32}$$

Then a finite cover argument as above yields the global relation (30). We only sketch the proof of (32). Let $\overline{x} \in \mathcal{F}$ be fixed. In the proof of Lemma 4.2(a) we made use of a local diffeomorphism $T_{\overline{x}}(x)$ leading to relation (25). This transformation $T_{\overline{x}}$ is constructed depending on the active index set $I_a(\overline{x}) := I_{rs}(\overline{x}) \cup I_r(\overline{x}) \cup I_s(\overline{x}) \cup J(\overline{x})$ (see (20)). For any x near $\overline{x}$ we have $I_a(x) \subset I_a(\overline{x})$ and there are only finitely many choices $I_\mu,\ \mu = 1, \dots, R$, for $I_\mu = I_a(x)$. So if we fix $I_\mu,\ I_\mu \subset I_a(\overline{x})$ any point $\hat{x}$ near $\overline{x}$ yields a local diffeomorphism $T_{\hat{x}}$ which depends smoothly on $\hat{x}$ (see the construction (20)). So we find a common bound: There exist $\alpha_\mu, \varepsilon_\mu > 0$ such that for any $x \in \mathcal{F} \cap B_{\varepsilon_\mu}(\overline{x})$ with $I_a(x) = I_\mu$ there is a point $x_\tau \in \mathcal{F}_\tau$ such that (for all τ small)

$$\|x_\tau - x\| \le \alpha_\mu\sqrt{\tau}\ .$$

Then by choosing $\varepsilon = \min\{\varepsilon_\mu \mid \mu = 1, \dots, R\}$ and $\alpha = \min\{\alpha_\mu \mid \mu = 1, \dots, R\}$ we have shown the relation (32). □

Note that Lemma 4.3 proves that the convergence in the Hausdorff distance

$$d(\mathcal{F}_\tau, \mathcal{F}) := \max\{\max_{x_\tau \in \mathcal{F}_\tau} d(x_\tau, \mathcal{F})\ ,\ \max_{x \in \mathcal{F}} d(x, \mathcal{F}_\tau)\}$$

between $\mathcal{F}_\tau$ and $\mathcal{F}$ satisfies $d(\mathcal{F}_\tau, \mathcal{F}) = O(\sqrt{\tau})$.

5. Convergence results for the value function and for the solutions of P_τ Let in this section $\overline{x} \in \mathcal{F}$ denote a global or local minimizer of P and $\overline{x}_\tau$ a nearby local solution of P_τ. Recall that by our compactness assumptions (4) a global minimizer of P_τ always exists (assuming $\mathcal{F}_\tau \neq \emptyset$).

In the present section we are interested in the convergence behavior and the convergence rate

$$\varphi_\tau \to \varphi \quad \text{and} \quad \overline{x}_\tau \to \overline{x} \qquad \text{if } \tau \to 0$$

for the value functions and the solutions of P and P_τ. From a viewpoint of parametric optimization to assure convergence the following assumptions are needed.

A$_1$. There exists a (global) solution $\overline{x}$ of P and a continuous function $\alpha : [0, \infty) \to [0, \infty)$, $\alpha(0) = 0$ such that for any $\tau > 0$ (small enough) we can find a point $x_\tau \in \mathcal{F}_\tau$ satisfying

$$\|x_\tau - \overline{x}\| \le \alpha(\tau)\ .$$

A$_2$. There exists a continuous function $\beta : [0, \infty) \to [0, \infty)$, $\beta(0) = 0$ such that for any $\tau > 0$ (small enough) the following holds: We can find a (global) solution $\overline{x}_\tau$ of P_τ and a corresponding point $\hat{x}_\tau \in \mathcal{F}$ such that

$$\|\hat{x}_\tau - \overline{x}_\tau\| \le \beta(\tau)\ .$$

It now appears that A$_1$ is connected to the upper semicontinuity of φ_τ (and $\mathcal{F}_\tau$) (see Lemma 5.1) and A$_2$ to the lower semicontinuity (see Lemma 5.2). To show this we have to use that by Lemma 4.2 the condition A$_1$ is satisfied with $\alpha(\tau) = O(\sqrt{\tau})$ if MPEC-LICQ holds at a (at least one) solution $\overline{x} \in \mathcal{S}$ and (see Lemma 5.2) A$_2$ holds with $\beta(\tau) = O(\sqrt{\tau})$ if MPEC-LICQ is satisfied at all $\overline{x} \in \mathcal{S}$.

LEMMA 5.1 *Let* MPEC-LICQ *hold at a point $\overline{x} \in \mathcal{S}$ (at least one). Then:*

(a) There exist constants $L > 0$, $\alpha > 0$ such that for all τ small enough the relation

$$\varphi_\tau - \varphi \le L\alpha(\tau)$$

is true with $\alpha(\tau) = \alpha\sqrt{\tau}$. If moreover SC is satisfied at $\overline{x}$ the inequality holds with $\alpha(\tau) = \alpha\tau$.

(b) To any $\varepsilon_1 > 0$ there is some τ_1 such that

$$d(\overline{x}_\tau, \mathcal{S}) < \varepsilon_1 \quad \textit{for all } \overline{x}_\tau \in \mathcal{S}_\tau \textit{ and for all } 0 \le \tau \le \tau_1\ .$$

PROOF. (a) By Lemma 4.2 the relation A_1 holds with the given function $\alpha(\tau)$. So with the solution $\overline{x}$ of P and the points x_τ in A_1 by using the Lipschitz continuity (5) we find

$$\varphi_\tau - \varphi \leq f(x_\tau) - f(\overline{x}) \leq L\|x_\tau - \overline{x}\| \leq L\,\alpha(\tau)\;.$$

(b) Suppose to the contrary that there is some $\varepsilon > 0$ and some sequence $\tau_\nu \to 0$ with corresponding $\overline{x}_{\tau_\nu} \in \mathcal{S}_{\tau_\nu}$ satisfying

$$d(\overline{x}_{\tau_\nu}, \mathcal{S}) \geq \varepsilon\;. \tag{33}$$

By compactness assumption without restriction we can assume $\overline{x}_{\tau_\nu} \to \hat{x}$. In view of Lemma 4.1 it follows $\hat{x} \in \mathcal{F}$ and from (a) we find

$$f(\overline{x}_{\tau_\nu}) = \varphi_{\tau_\nu} \leq \varphi + L\alpha(\tau_\nu) \to \varphi$$

and thus $f(\hat{x}) = \varphi$ implying $\hat{x} \in \mathcal{S}$ in contradiction to (33). □

LEMMA 5.2 *Let* MPEC-LICQ *hold at every point* $\overline{x} \in \mathcal{S}$. *Then the condition* A_2 *is satisfied with* $\beta(\tau) = \beta\sqrt{\tau}$ *for some* $\beta > 0$ *(and with* $\beta(\tau) = \beta\tau$ *in case that SC holds at all* $\overline{x} \in \mathcal{S}$*) and there exists* $L > 0$ *such that for all* τ *small enough:*

$$-L\beta(\tau) \leq \varphi_\tau - \varphi\;.$$

PROOF. The set $\mathcal{S} \subset X$ is compact and arguing as in the proof of Lemma 4.3 it follows that $\cup_{\overline{x}\in\mathcal{S}} B_{\varepsilon(\overline{x})}(\overline{x})$ forms an open cover of the set $\mathcal{S}$. So we can choose a finite cover, $\mathcal{S} \subset \cup_{\nu=1,\ldots,K} B_{\varepsilon_\nu}(\overline{x}_\nu)$, $\overline{x}_\nu \in \mathcal{S}$, and define $B_\varepsilon(\mathcal{S}) := \{x \mid d(x, S) < \varepsilon\}$. Now we choose $\varepsilon_1 > 0$ such that

$$B_{\varepsilon_1}(\mathcal{S}) \subset \bigcup_{\nu=1,\ldots,K} B_{\varepsilon_\nu}(\overline{x}_\nu)\;.$$

By Lemma 5.1(b) there exists some $\tau_1 > 0$ such that

$$\overline{x}_\tau \in B_{\varepsilon_1}(\mathcal{S}) \qquad \text{for all } \overline{x}_\tau \in \mathcal{S}_\tau \text{ and for all } 0 \leq \tau \leq \tau_1\;.$$

By construction, for $0 \leq \tau \leq \tau_1$, any point $\overline{x}_\tau \in \mathcal{S}_\tau$ is contained in (at least) one of the balls $B_{\varepsilon_\nu}(\overline{x}_\nu)$, $\nu \in \{1, \ldots, K\}$ ($\nu = \nu_\tau$) and in view of Lemma 4.2(a) we can choose a point $\hat{x}_\tau \in \mathcal{F}$ such that

$$\|\hat{x}_\tau - \overline{x}_\tau\| < \beta_\nu\sqrt{\tau} \qquad (\text{ resp. } \quad < \beta_\nu\tau)\;,$$

(β_ν corresponding to $\overline{x}_\nu$). By defining $\beta = \min\{\beta_\nu \mid \nu = 1, \ldots, K\}$ we have proven A_2 and with these points $\overline{x}_\tau, \hat{x}_\tau$ by using (5) again we find

$$\varphi_\tau - \varphi \geq f(\overline{x}_\tau) - f(\hat{x}_\tau) \geq -L\,\beta(\tau)\;.$$

□

To obtain qualitative results on the rate of convergence for the solutions $\overline{x}_\tau$ of P_τ we have to assume some growth condition at the solution $\overline{x}$ of P. We will assume that $\overline{x}$ is a minimizer of order $\omega \geq 1$ (see (7)). Sufficient and necessary conditions for these assumptions are given in Section 3. Note that in this case $\mathcal{S} = \{\overline{x}\}$. For minimizers of order $\omega = 2$ the next result, *i.e.*, a convergence $O(\tau^{1/4})$, is also proven in [14]. (However with a different technique.)

COROLLARY 5.1 *Let* $\overline{x}$ *be a global minimizer of* P *of order* $\omega \geq 1$ *and let* MPEC-LICQ *hold at* $\overline{x}$. *Then* $|\varphi_\tau - \varphi| \leq O(\sqrt{\tau})$ *and there is some* $c > 0$ *such that for any global minimizer* $\overline{x}_\tau$ *of* P_τ *it follows*

$$\|\overline{x}_\tau - \overline{x}\| \leq c \cdot \sqrt{\tau}^{1/\omega}\;. \tag{34}$$

If SC holds at $\overline{x}$ $\sqrt{\tau}$ *can be replaced by* τ.

PROOF. By Lemma 5.1 and 5.2 the convergence for the value function φ_τ is immediate. Moreover the assumptions A_1 and A_2 hold with functions $\alpha(\tau) = \alpha\sqrt{\tau}$ etc. Then with the points $\overline{x}, \hat{x}_\tau \in \mathcal{F}$, $\overline{x}_\tau, x_\tau \in \mathcal{F}_\tau$ in A_1 and A_2 we obtain

$$f(\overline{x}) \leq f(\hat{x}_\tau) \leq f(\overline{x}_\tau) + L\,\beta(\tau) \leq f(x_\tau) + L\,\beta(\tau) \leq f(\overline{x}) + L\,\alpha(\tau) + L\,\beta(\tau)$$

and thus

$$0 \leq f(\hat{x}_\tau) - f(\overline{x}) \leq L\,\alpha(\tau) + L\,\beta(\tau).$$

Again by taking the point $\hat{x}_\tau \in \mathcal{F}$ in A_2 in view of (7) this inequality yields

$$\begin{aligned} \|\overline{x}_\tau - \overline{x}\| &\leq \|\overline{x}_\tau - \hat{x}_\tau\| + \|\hat{x}_\tau - \overline{x}\| \leq \beta(\tau) + \left(\frac{f(\hat{x}_\tau) - f(\overline{x})}{\kappa}\right)^{1/\omega} \\ &\leq \beta(\tau) + \frac{1}{\kappa^{1/\omega}}\,(L\alpha(\tau) + L\beta(\tau))^{1/\omega} \end{aligned}$$

which in view of $\omega \geq 1$, proves the statement. □

The preceeding corollary presents a result on the global minimizers which always exist. Recall that $\mathcal{F}_\tau, \mathcal{F}$ are compact (see (4)). In the next corollary also the existence of local minimizers for P_τ is established.

COROLLARY 5.2 *Let $\bar{x} \in \mathcal{F}$ be a local minimizer of order $\omega \geq 1$ of P such that* MPEC-LICQ *holds at $\bar{x}$. Then for any $\tau > 0$ small enough there exist (nearby) local minimizers $\bar{x}_\tau$ of P_τ and (for each of these minimizers) it follows:*

$$\|\bar{x}_\tau - \bar{x}\| \leq O(\sqrt{\tau}^{1/\omega}) \ .$$

If SC holds at $\bar{x}$ $\sqrt{\tau}$ can be replaced by τ.

PROOF. Let $\bar{x}$ be a local minimizer of P satisfying MPEC-LICQ . Then with some $\delta > 0$ (small enough) $\bar{x}$ is a global solution of the problem restricted to $\mathcal{F}_\tau \cap \overline{B}_\delta(\bar{x})$. Note that we have chosen a closed ball $\overline{B}_\delta(\bar{x})$ to assure the existence of a minimizer $\bar{x}_\tau$. By Corollary 5.1 the statements follow for the problem restricted to $\mathcal{F}_\tau \cap \overline{B}_\delta(\bar{x})$. But since $\bar{x}_\tau \to \bar{x}$ for $\tau \to 0$ the points $\bar{x}_\tau$ are also elements of the open set $B_\delta(\bar{x})$, *i.e.*, $\bar{x}_\tau$ are local minimizers of the problems P_τ. □

We emphasize that in general (without SC), for the minimizer $\bar{x}_\tau$ we cannot expect a faster convergence rate than $O(\sqrt{\tau})$. More precisely from Lemma 4.2(b) we deduce that at a minimizer $\bar{x}$ of P where SC does *not* hold the following is true with some $c_2 > 0$:

$$\|\bar{x}_\tau - \bar{x}\| \geq c_2\sqrt{\tau} \ . \tag{35}$$

If $\bar{x}$ is a local minimizer of order $\omega = 1$ the optimal convergence rate $\|\bar{x}_\tau - \bar{x}\| \leq O(\sqrt{\tau})$ occurs (*cf.* Corollary 5.2) (optimal, unless SC holds). Recall that generically all local minimizers of P are either of order $\omega = 1$ or $\omega = 2$ (see (18)). We give a counterexample for the remaining case $\omega = 2$.

EXAMPLE 5.1 $\quad \min\ x_1^2 + x_2, \quad \text{s.t.} \quad x_1 \cdot x_2 = 0, \ \ x_1, x_2 \geq 0 \ ,$

i.e., $r(x) = x_1, \ s(x) = x_2$. The minimizer $\bar{x} = (0,0)$ is of order $\omega = 2$ and it is a MPEC-KKT point satisfying the KKT condition $\nabla f(\bar{x}) = 0 \cdot \nabla r(\bar{x}) + 1 \cdot \nabla s(\bar{x})$. So, the MPEC-SC condition is not fulfilled. Here, the minimizers of P_τ read: $\bar{x}_\tau = \big((\frac{\tau}{2})^{\frac{1}{3}}, (2\tau^2)^{\frac{1}{3}}\big)$.

The preceding example (see also [14]) shows that at a local minimizer $\bar{x}$ of order two even under MPEC-LICQ the convergence rate for $\|\bar{x}_\tau - \bar{x}\|$ can be slower than $O(\sqrt{\tau})$. Note however that this example is not a generic one since the MPEC-SC condition does not hold. We will now show that in the generic case this bad behavior can be excluded. More precisely under the condition MPEC-LICQ , MPEC-SC and MPEC-SOC at $\bar{x}$ we prove that the minimizers $\bar{x}_\tau$ of P_τ are (locally) unique and the (*optimal*) convergence rate $\|\bar{x}_\tau - \bar{x}\| = O(\sqrt{\tau})$ takes place. The proof again makes use of the local transformation of the problem into an equivalent simpler one (*cf.*, Section 4).

THEOREM 5.1 *Let $\bar{x}$ be a local minimizer of P such that* MPEC-LICQ , MPEC-SC *and* MPEC-SOC *holds. Then for all $\tau > 0$ (small enough) the local minimizers $\bar{x}_\tau$ of P_τ (near $\bar{x}$) are uniquely determined and satisfy $\|\bar{x}_\tau - \bar{x}\| = O(\sqrt{\tau})$.*
The same statement holds for the global minimizers $\bar{x}$ and $\bar{x}_\tau$ of P and P_τ, respectively.

PROOF. To prove this statement we again consider the problem P_τ in standard form (see Section 4, (23)),

$$P_\tau : \quad \min\ f(x) \quad \text{s.t.} \quad \begin{array}{rcll} h_i(x) = x_i \cdot x_{m+i} & = & \tau & i = 1,\ldots,p \\ h_i(x) = x_i \cdot s_{p+i}(x) & = & \tau & i = 1,\ldots,m-p \\ g_i(x) = x_{m+p+i} & \geq & 0 & i = 1,\ldots,q_0 \\ x_i, x_{m+i} & \geq & 0 & i = 1,\ldots,p \\ x_i, s_i(x) & \geq & 0 & i = p+1,\ldots,m \end{array} \tag{36}$$

where $\bar{x} = 0$ is the local solution of P_0 with $s_{p+i}(0) = c_i^s > 0, \ i = 1,\ldots,m-p$. Under MPEC-LICQ , the KKT condition for $\bar{x}$ reads

$$\nabla f(\bar{x}) - \sum_{i=1}^{p}\big(\gamma_i^1 e_i + \gamma_{i+m}^2 e_{i+m}\big) - \sum_{i=1}^{m-p}\gamma_i^3 e_{p+i} - \sum_{i=i}^{q_0}\gamma_j^4 e_{m+p+i} \ = \ 0 \tag{37}$$

with multiplier vector $(\gamma^1,\gamma^2,\gamma^4) > 0$, by MPEC-SC . So in (36) the function $f(x)$ has the form

$$f(\bar{x}) = \sum_{i=1}^{p}\big(\gamma_i^1 x_i + \gamma_{i+m}^2 x_{i+m}\big) + \sum_{i=1}^{m-p}\gamma_i^3 x_{p+i} + \sum_{i=i}^{q_0}\gamma_i^4 x_{m+p+i} + q(x) \tag{38}$$

where $|q(x)| = O(\|x\|^2)$. For convenience we now introduce the abbreviation $x^1 = (x_1,\ldots,x_p)$, $x^2 = (x_{m+1},\ldots,x_{m+p})$, $x^3 = (x_{p+1},\ldots,x_m)$, $x^4 = (x_{m+p+1},\ldots,x_{m+p+q_0})$ and $x^5 = (x_{m+p+q_0+1},\ldots,x_n)$ and write $x = (x^1,\ldots,x^5)$. In this setting the tangent space at $\bar{x}$ becomes $T_{\bar{x}} = \text{span}\,\{e_i\ ,\ i = m+p+q_0+1,\ldots,n)$ ($T_{\bar{x}} = C_{\bar{x}}$ *cf.* (17)), MPEC-SOC takes the form

$$\nabla_x^2 f(x) \quad \text{is positive definite on } T_{\bar{x}} \text{ or } \nabla_{x^5}^2 f(x) \text{ is positive definite} \tag{39}$$

and (36) reads:

$$P_\tau : \qquad \min\ (\gamma^1)^T x^1 + (\gamma^2)^T x^2 + (\gamma^3)^T x^3 + (\gamma^4)^T x^4 + q(x) \quad \text{s.t.}$$
$$\begin{array}{rcll} x_i^1 \cdot x_i^2 & = & \tau & i = 1,\ldots,p \\ x_i^3 \cdot s_{p+i}(x) & = & \tau & i = 1,\ldots,m-p \\ x_i^4 & = & 0 & i = 1,\ldots,q_0 \ . \end{array} \tag{40}$$

Note that by the condition $\gamma^4 > 0$, near $\bar{x}$, all inequalities $x_i^4 \geq 0$ must be active.

The minimizers $\overline{x}_\tau$ of P_τ are solutions of the following KKT system of (40) in the variables (x, λ, μ, ν) (we omit the variable x),

$$\begin{pmatrix} \gamma^1 + \nabla_{x^1} q \\ \\ \gamma^2 + \nabla_{x^2} q \\ \\ \gamma^3 + \nabla_{x^3} q \\ \gamma^4 + \nabla_{x^4} q \\ \nabla_{x^5} q \end{pmatrix} - \begin{pmatrix} x_1^2 & & \\ & \ddots & \\ & & x_p^2 \\ x_1^1 & & \\ & \ddots & \\ & & x_p^1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda - \begin{pmatrix} x_1^3 \nabla_{x^1} s_{p+1} & .. & x_{m-p}^3 \nabla_{x^1} s_m \\ \\ x_1^3 \nabla_{x^2} s_{p+1} & .. & x_{m-p}^3 \nabla_{x^2} s_m \\ \\ x_1^3 \nabla_{x^3} s_{p+1} + s_{p+1} e_1 & .. & x_{m-p}^3 \nabla_{x^3} s_m + s_m e_{m-p} \\ x_1^3 \nabla_{x^4} s_{p+1} & .. & x_{m-p}^3 \nabla_{x^4} s_m \\ x_1^3 \nabla_{x^5} s_{p+1} & .. & x_{m-p}^3 \nabla_{x^5} s_m \end{pmatrix} \mu - \begin{pmatrix} 0 \\ \\ 0 \\ \\ 0 \\ \nu \\ 0 \end{pmatrix} = 0$$

together with the constraints in (40). In this system the vectors $e_1, \ldots, e_{m-p}$ are unit vectors in $\mathbb{R}^{m-p}$.

Now the trick is to eliminate the unknown λ and to simplify (regularize) the equations $x_i^1 \cdot x_i^2 = \tau$ as follows. We define

$$\begin{aligned} \hat{\gamma}^1 &:= \gamma^1 + \nabla_{x^1} q - [x_1^3 \nabla_{x^1} s_{p+1} \,..\, x_{m-p}^3 \nabla_{x^1} s_m]\mu \ , \\ \hat{\gamma}^2 &:= \gamma^2 + \nabla_{x^2} q - [x_1^3 \nabla_{x^2} s_{p+1} \,..\, x_{m-p}^3 \nabla_{x^2} s_m]\mu \end{aligned}$$

and note that due to $\gamma^1, \gamma^2 > 0$ and $|q(x)| = O(\|x\|^2)$, near $\overline{x} = 0$, the vectors satisfy $\hat{\gamma}^1, \hat{\gamma}^2 > 0$. So near $\overline{x} = 0$ the functions

$$\sqrt{\hat{\gamma}^1/\hat{\gamma}^2} \quad \text{and} \quad \sqrt{\hat{\gamma}^2/\hat{\gamma}^1}$$

are C^1-functions of x. From the system we deduce $\hat{\gamma}_i^1 = x_i^2 \lambda_i$, $\hat{\gamma}_i^2 = x_i^1 \lambda_i$ and $\hat{\gamma}_i^1 \hat{\gamma}_i^2 = \tau(\lambda_i)^2$ or $\lambda_i = \sqrt{\frac{\hat{\gamma}_i^1 \hat{\gamma}_i^2}{\tau}}$ and finally

$$x_i^1 = \sqrt{\hat{\gamma}_i^2/\hat{\gamma}_i^1} \cdot \sqrt{\tau} \quad , \quad x_i^2 = \sqrt{\hat{\gamma}_i^1/\hat{\gamma}_i^2} \cdot \sqrt{\tau} \ .$$

So the system above can be subdivided into the two equations:

$$\begin{array}{ccccc} x_i^1 & - & \sqrt{\hat{\gamma}_i^2/\hat{\gamma}_i^1} \cdot \sqrt{\tau} & = 0 \\ x_i^2 & - & \sqrt{\hat{\gamma}_i^1/\hat{\gamma}_i^2} \cdot \sqrt{\tau} & = 0 \\ (x_1^3 \nabla_{x^3} s_{p+1} + s_{p+1} e_1 \ldots x_{m-p}^3 \nabla_{x^3} s_m + s_{p+1} e_{m-p})\mu & - & \gamma^3 - \nabla_{x^3} q & = 0 \\ (x_1^3 \nabla_{x^5} s_{p+1} \ldots x_{m-p}^3 \nabla_{x^5} s_m)\mu & - & \nabla_{x^5} q & = 0 \\ x_i^3 s_{p+i} & - & \tau & = 0 \\ x_i^4 & & & = 0 \end{array} \tag{41}$$

and the system corresponding to the multiplier ν:

$$-(x_1^3 \nabla_{x^5} s_{p+1} \ldots x_{m-p}^3 \nabla_{x^5} s_m)\mu \quad + \quad \gamma^4 + \nabla_{x^4} q \quad = \nu \ . \tag{42}$$

The relation (41) represents a system $F(x, \mu, \tau) = 0$ of $n + m - p$ equations in $n + m - p + 1$ variables (x, μ, τ). The point $(\overline{x}, \overline{\mu}, \overline{\tau})$ with $\overline{x} = 0$, $\overline{\tau} = 0$ and $\overline{\mu} = (\gamma_1^3/s_{p+1}(\overline{x}), \ \ldots, \gamma_{m-p}^3/s_m(\overline{x}))$ (recall $s_i(\overline{x}) > 0$) solves (41). The Jacobian with respect to (x, μ) at this point $(\overline{x}, \overline{\mu}, \overline{\tau})$ has the form

$$\begin{array}{cccccc} x^1 & x^2 & x^3 & x^4 & x^5 & \mu \\ \hline I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ X & X & X & X & X & \begin{bmatrix} s_{p+1} & & \\ & \ddots & \\ & & s_m \end{bmatrix} \\ 0 & 0 & X & 0 & \nabla_{x^5}^2 q & 0 \\ 0 & 0 & \begin{bmatrix} s_{p+1} & & \\ & \ddots & \\ & & s_m \end{bmatrix} & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \end{array}$$

(X is some matrix of appropriate dimension; recall $\nabla_{x^i} q(\overline{x}) = 0$). Since $\nabla_{x^5}^2 q(\overline{x})$ is positive definite (*cf.* (39)) and $s_i(\overline{x}) > 0$, $i = p+1, \ldots, m$, we see that this matrix is regular. So we can apply the Implicit Function Theorem to the equation $F = 0$ which near $\overline{\tau} = 0$ yields a unique solution $x(\tau), \mu(\tau)$ differentiable in the parameter $\sqrt{\tau}$. This implies $x(\tau) = \overline{x} + O(\sqrt{\tau})$ $\mu(\tau) = \overline{\mu} + O(\sqrt{\tau})$. Substituting this solution $x(\tau), \mu(\tau)$ into the equation (42) determines the variable $\nu(\tau)$. Since the (local) minimizers $\overline{x}_\tau$ of P_τ must solve the systems (41), (42) clearly $\overline{x}_\tau = x(\tau)$ is uniquely determined. The unique multipliers wrt. P_τ are $\nu_i(\tau)$ corresponding to $x_i^4 = 0$, $\mu_i(\tau)$ corresponding to $x_i^3 s_{p+i}(x) = \tau$ and $\hat{\gamma}_i^1, \hat{\gamma}_i^2$ belonging to x_i^1, x_i^2. This proves the statement for the local minimizers.

If $\overline{x}$ is a global minimizer we can argue as in the second part of the proof of Corollary 5.2. Firstly by restricting the minimization to a neighborhood $\overline{B}_\delta(\overline{x})$ the result follows as above. The compactness assumption for F_τ and the fact that $\overline{x}$ is a global minimizer (of order $\omega = 2$) exclude global minimizers $\overline{x}_\tau$ of P_τ outside $\overline{B}_\delta(\overline{x})$. □

In the next remark we indicate that the result of Theorem 5.1 is also true for C-stationary points.

REMARK 5.1 Let $\overline{x}$ be a feasible point of the complementarity constrained problem P. It is called C-stationary point if the condition (10) holds with some multiplier (γ, ρ, σ), satisfying $\gamma_j \geq 0,\ j \in J(\overline{x})$ and $\rho_i \cdot \sigma_i \geq 0,\ i \in I_{rs}(\overline{x})$ (see *e.g.* [15]). If MPEC-LICQ holds at $\overline{x}$ the multiplier is uniquely determined. In this case we define

$$(\text{MPEC-SC}') : \qquad \gamma_j > 0,\ j \in J(\overline{x}), \quad \rho_i \cdot \sigma_i > 0,\ i \in I_{rs}(\overline{x}) .$$

$$(\text{MPEC-SOC}\ ') : \qquad d^T \nabla_x^2 L(\overline{x}, \gamma, \rho, \sigma) d \neq 0 \qquad \forall d \in C_{\overline{x}} \backslash \{0\} .$$

The genericity result in Theorem 3.1 then also holds for C-stationary points:

Generically in C_s^2 at all C-stationary points of a problem P the conditions MPEC-LICQ, MPEC-SC' *and* MPEC-SOC' *hold.*

By modifying the proof of Theorem 5.1 in an obvious way (use $\gamma_i^1 \cdot \gamma_i^2 > 0$ instead of $\gamma_i^1, \gamma_i^2 > 0$ etc.) the statement of Theorem 5.1 is also true for C-stationary points:

Let $\overline{x}$ be a C-stationary point of P such that MPEC-LICQ, MPEC-SC' *and* MPEC-SOC' *holds. Then for all $\tau > 0$ (small enough) there exist (locally) unique stationary points $\overline{x}_\tau$ of P_τ and $\|\overline{x}_\tau - \overline{x}\| = O(\sqrt{\tau})$.*
Note that C-stationarity is a weaker concept than the concept of local minimizers. As shown (*e.g.*, in [15]) under a certain MFCQ assumption at $\overline{x}$ (which is weaker than MPEC-LICQ) any local minimizer of P is necessarily a C-stationary point. Moreover the limit points of of a sequence of minimizers $\overline{x}_\tau$ of P_τ (for $\tau \to 0$) are typically C-stationary points of P.

We end up with some further observations.

REMARK 5.2 Let us note that from the results of this paper we also can deduce the convergence results of [14] for the relaxation $P_\tau^{\leq}$ of Section 1 (under the stronger MPEC-LICQ condition).

Suppose we have given a local solution $\overline{x}$ of P such that MPEC-LICQ holds and with a corresponding KKT-solution MPEC-SC , MPEC-SOC is satisfied (*i.e.*, by Theorem 3.4 $\overline{x}$ is a minimizer of order $\omega = 2$). In view of Corollary 3.1 it is also a solution of the relaxed problem $P_R(\overline{x})$ in (19) and by using MPEC-SC it follows that for the solutions $\hat{x}_\tau$ of $P_\tau^{\leq}$ (near $\overline{x}$) (see Section 1) the conditions $r_i(x)\ s_i(x) \leq \tau$, $i \in I_{rs}(\overline{x})$, are not active but that for all $\tau > 0$ small enough

$$r_i(\hat{x}_\tau) = s_i(\hat{x}_\tau) = 0\ , \quad \forall i \in I_{rs}(\overline{x}), \tag{43}$$

holds. So to analyze the behavior of the solution $\hat{x}_\tau$ the whole analysis can be done under the condition (43), *i.e.*, we are in the situation as for the case that the strong SC-condition holds. So instead of the convergence $O(\sqrt{\tau})$ (*cf., e.g.* Lemma 4.2) we obtain a rate $O(\tau)$ and in the same way the analysis in Section 5 simplifies resulting in a convergence behavior $\|\hat{x}_\tau - \overline{x}\| = O(\tau)$.

REMARK 5.3 We wish to emphasize that the convergence results of this paper can be generalized in a straightforward way to problems P_μ containing constraints of the product form

$$r_1(x) r_2(x) \cdots r_\mu(x) = 0, \quad r_1(x), r_2(x), \ldots, r_\mu(x) \geq 0\ .$$

Here at a solution $\overline{x}$ of P_μ where all constraints r_i are active, *i.e.*,

$$r_1(\overline{x}) = r_2(\overline{x}) = \ldots = r_\mu(\overline{x}) = 0\ ,$$

a perturbation $r_1(x) r_2(x) \cdots r_\mu(x) = \tau$ will lead to a convergence rate

$$\|\overline{x}_\tau - \overline{x}\| \approx O(\tau^{1/\mu})$$

for the solutions $\overline{x}_\tau$ of the perturbed problem. Also all other results in the present paper can be extended in a straightforward way to this generalization.

Acknowledgements. The authors are very indebted to the referees for their many valuable comments.

References

[1] Bank G., Guddat J., Klatte D., Kummer B., Tammer D., *Non-linear Parametric Optimization*, Birkhäuser Verlag, Basel, (1983).

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

[3] Chen C., Mangasarian O.L., *Smoothing methods for convex inequalities and linear complementarity problems*, Mathematical Programming 71, 51-69, (1995).

[4] Facchinei F., Jiang H., Qi L., *A smoothing method for mathematical programs with equilibrium problems*, Mathematical Programming 95, 107-134, (1999).

[5] Fukushima M., Lin G-H., , *Smoothing methods for mathematical programs with equilibrium constraints*, Preprint.

[6] Fukushima M., Pang J.-S., *Convergence of a smoothing continuation method for mathematical programs with complementarity constraints*, In Ill-posed variational problems and regularization techniques, Lecture Notes in Econom. and Math. Systems, 477, Springer, 99-110, (1999).

[7] Guddat J., Guerra F. and Jongen H. Th., *Parametric optimization: singularities, pathfollowing and jumps*, Teubner and John Wiley, Chichester (1990).

[8] Hu, X., Ralph, D., *A note on sensitivity of value functions of mathematical programs with complementarity constraints*, Mathematical Programming, Ser. A 93, 265-279, (2002).

[9] Hu, X., *On the convergence of general regularization and smoothing schemes for mathematical programs with complementarity constraints*, Optimization 53, 39-50, (2004).

[10] Leyffer S. *Mathematical programs with complementarity constraints*, Argonne National Lab. Preprint ANA/MCSP1026-0203 (2003).

[11] Lin G.H., Fukushima M., *New relaxation method for mathematical programs with complementarity constraints*, Journal of Optimization Theory and Applications 118, no. 1, 81-116, (2003).

[12] Luo, Z. Q., J. S. Pang, D. Ralph., *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, U.K. (1996).

[13] Outrata J., Kocvara M., Zowe J., *Nonsmooth approach to optimization problems with equilibrium constraints.* Nonconvex Optimization and its Applications, 28, Kluwer Academic Publishers, Dordrecht, (1998).

[14] Ralph D., Wright S.J., *Some properties of regularization and penalization schemes for MPECS*, Optimization Methods and Software 19, 527-556, (2004).

[15] Scheel H., Scholtes S., *Mathematical Programs with Complementarity Constraints: Stationarity, Optimality and Sensitivity*, Mathematics of Operations Research 25, No. 1, 1-21, (2000).

[16] Scholtes S., Stöhr M., *How Stringent is the Linear Independence Assumption for Mathematical Programs with Stationary Constraints?*, Mathematics of Operations Research 26, No. 4, 851-863, (2001).

[17] Scholtes S., *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM J. Optim. 11, No. 4, 918-936, (2001).

[18] Stein O., Still G., *Solving semi-infinite optimization problems with interior point techniques,* SIAM J. Control Optim. 42, no. 3, 769–788, (2003).

[19] Still G., Streng M., *Survey Paper: Optimality conditions in smooth nonlinear optimization*, Journal of Optimization Theory and Applications 90 (3), 483-516, (1996).

[20] Wright, S.J.; Orban, D., *Properties of the log-barrier function on degenerate nonlinear programs*, Mathematics of Operations Research 27, no. 3, 585–613, (2002).