MATHEMATICS OF OPERATIONS RESEARCH
Vol. 25, No. 1, February 2000
Printed in U.S.A.

MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS: STATIONARITY, OPTIMALITY, AND SENSITIVITY

HOLGER SCHEEL AND STEFAN SCHOLTES

We study mathematical programs with complementarity constraints. Several stationarity concepts, based on a piecewise smooth formulation, are presented and compared. The concepts are related to stationarity conditions for certain smooth programs as well as to stationarity concepts for a nonsmooth exact penalty function. Further, we present Fiacco-McCormick type second order optimality conditions and an extension of the stability results of Robinson and Kojima to mathematical programs with complementarity constraints.

1. Introduction. Mathematical programs with equilibrium constraints (MPECs) have recently found considerable attention alongside a growing interest in the areas of optimal design of economical or mechanical equilibrium systems and inverse equilibrium problems. In the former applications certain parameters of the system have to be set in such a way that some performance measure is optimized in an equilibrium state. In the latter applications equilibrium states are observed and the observations are used to estimate unknown data in a corresponding equilibrium model. We refer to the monographs of Luo et al. (1996) and Outrata et al. (1998) for detailed applications. Adequate modeling of the equilibrium conditions is eminently important for the computational tractability of such optimization problems. In the simplest case, the equilibrium conditions can be formulated as a well-posed system of smooth equations and the MPEC turns into an ordinary nonlinear program. A more complicated situation arises if the equilibrium laws are amenable to a complementarity formulation, either directly or by employing suitable stationarity conditions of underlying constrained optimization problems or variational inequalities (cf. Cottle et al. 1992 and Ferris and Pang 1997). In this case the MPEC typically will comprise constraints of the form

$$F(z) \geq 0, \qquad G(z) \geq 0, \qquad F(z)^{\top} G(z) = 0$$

for suitable smooth functions $F, G : \mathbb{R}^n \to \mathbb{R}^m$. From a nonlinear programming point of view, such complementarity constraints are problematic, irrespective of the data F, G, since there is no feasible solution which satisfies all inequalities strictly. Therefore the Mangasarian-Fromovitz constraint qualification, which implies the existence of a strictly feasible arc, is violated at every feasible point and consequently any constraint system that comprises complementarity constraints of the above form is inherently unstable (cf. Robinson 1976 and Jongen and Weber 1991). Related to this observation is the obvious drawback that a linearization of the above constraints, and indeed of any smooth formulation of complementarity constraints, leads to a polyhedral convex cone which cannot reflect the structural nonconvexity, due to the combinatorial nature of the complementarity conditions, in the vicinity of points where strict complementarity is violated, i.e.,

Received June 24, 1998; revised March 30, 1999 and August 12, 1999.
AMS 1991 subject classification. Primary: 90C33, 90C30.
OR/MS subject classification. Primary: Programming complementarity, Nonlinear.
Key words. Bouligand stationarity, Clarke stationarity, vertical complementarity constraints, equilibrium constraints, exact penalty function, sensitivity, stability, strong stationarity, weak stationarity.

0364-765X/00/2501/0001/$05.00
1526-5471 electronic ISSN, © 2000, INFORMS

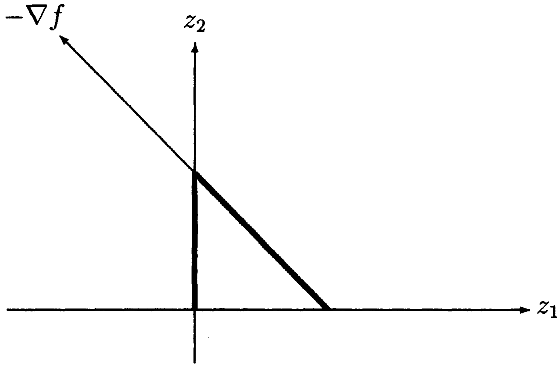

FIGURE 1. $\min\{z_1 - z_2 \mid \min\{z_1, 1 - z_1 - z_2\} = 0,\ z_2 \geq 0\}$.

$F(z) + G(z) \not> 0$; see Figure 1. In this paper we shall therefore focus attention on an equivalent nonsmooth formulation of the complementarity constraints as

$$\min\{F_i(z), G_i(z)\} = 0, \quad i = 1, \dots, m.$$

A first order approximation of these equations captures the nonconvexity of the constraint set adequately. It turns out that many results that can be obtained for programs with standard nonlinear complementarity constraints hold, mutatis mutandis, also for the more general case of vertical complementarity constraints as introduced by Cottle and Dantzig (1970). Our objects of study are therefore mathematical programs with complementarity constraints (MPCC) of the form

$$\begin{aligned} \text{MPCC} \quad \min \quad & f(z) \\ \text{s.t.} \quad & \min\{F_{k1}(z), \dots, F_{kl}(z)\} = 0, \quad k = 1, \dots, m, \\ & g(z) \leq 0, \\ & h(z) = 0, \end{aligned}$$

where $z \in \mathbb{R}^n$, $f(z) \in \mathbb{R}$, $g(z) \in \mathbb{R}^p$, $h(z) \in \mathbb{R}^q$, and $F(z) \in \mathbb{R}^{m \times l}$ with

$$F(z) = \begin{bmatrix} F_{11}(z) & \cdots & F_{1l}(z) \\ \vdots & \ddots & \vdots \\ F_{m1}(z) & \cdots & F_{ml}(z) \end{bmatrix}.$$

We assume throughout that the functions f, F, g, h are smooth and that $m \geq 1$ and $l \geq 2$. Figure 1 gives a graphical illustration of a simple but typical MPCC.

If $z = (x, y) \in \mathbb{R}^m \times \mathbb{R}^{n-m}$ and $F_{k1}(x, y) = x_k$ then $\min\{F_{k1}(z), F_{k2}(z)\} = 0$, $k = 1, \dots, m$, is equivalent to the standard parametric nonlinear complementarity constraint

$$x \geq 0, \qquad F_2(x, y) \geq 0, \qquad x^\top F_2(x, y) = 0.$$

Karush-Kuhn-Tucker conditions of parametric nonlinear programs or of variational inequalities can be similarly reformulated and incorporated into the constraints of MPCC.

In the following section we introduce several stationarity concepts for MPCCs and study their relations. Section 3 contains second order optimality conditions, while the final section presents some stability results for MPCCs.

In the remainder of this introductory section we set some notation and explain some terminology that we will use throughout. We associate with an MPCC the *Lagrangian function*

$$L(\alpha, z, \Gamma, \lambda, \mu) = \alpha f(z) - F(z)\Gamma + g(z)\lambda + h(z)\mu, \tag{1}$$

where $F(z)\Gamma = \sum_i \sum_j F_{ij}(z)\Gamma_{ij}$ is the inner product of the two $m \times l$-matrices $F(z)$ and Γ.

Given a feasible point z, we furthermore associate with an MPCC two ordinary nonlinear programs, the *tightened NLP*

$$\begin{array}{lll} \text{TNLP} & \min & f(x) \\ & \text{s.t.} & F_{ki}(x) = 0 \quad \text{if } F_{ki}(z) = 0, \\ & & F_{kj}(x) \geq 0 \quad \text{if } F_{kj}(z) > 0, \\ & & g(x) \leq 0, \\ & & h(x) = 0, \end{array}$$

and the *relaxed NLP*

$$\begin{array}{lll} \text{RNLP} & \min & f(x) \\ & \text{s.t.} & F_{ki}(x) = \begin{cases} = 0 & \text{if } F_{ki}(z) = 0 \text{ and } F_{kj}(z) > 0 \text{ for } j \neq i, \\ \geq 0 & \text{otherwise}, \end{cases} \\ & & g(x) \leq 0, \\ & & h(x) = 0. \end{array}$$

A simple and elegant way of dealing with an MPCC is the local decomposition approach proposed in Scheel (1995) and Luo et al. (1996). Given a feasible point z of an MPCC let

$$\mathcal{I}(z) = \{I \mid I \subseteq \{(k,i) \mid F_{ki}(z) = 0\} \text{ and } \forall k \ \exists i : (k,i) \in I\} \tag{2}$$

and associate with an index set $I \in \mathcal{I}(z)$ the ordinary nonlinear program

$$\begin{array}{lll} \text{NLP}_I & \min & f(x) \\ & \text{s.t.} & F_I(x) = 0, \\ & & F_{I^c}(x) \geq 0, \\ & & g(x) \leq 0, \\ & & h(x) = 0, \end{array}$$

where F_I denotes the vector valued function with components $(k,i) \in I$ and I^c denotes the complement of I. Notice that the Lagrangian of each program NLP_I coincides with the Lagrangian (1) of MPCC. Following Luo et al. (1996), we denote by $\mathcal{F}_I$, $\mathcal{F}_{\text{MPCC}}$, $\mathcal{F}_{\text{TNLP}}$, and $\mathcal{F}_{\text{RNLP}}$ the feasible sets of the programs NLP_I, MPCC, TNLP, and RNLP, respectively, and obtain the relations

$$\mathcal{F}_{\text{TNLP}} = \bigcap_{I \in \mathcal{I}(z)} \mathcal{F}_I \subseteq \mathcal{F}_I \subseteq \mathcal{F}_{\text{MPCC}} = \bigcup_{I \in \mathcal{I}(z)} \mathcal{F}_I \subseteq \mathcal{F}_{\text{RNLP}} \tag{3}$$

locally around the point z. In particular, z is a local minimizer of an MPCC if and only if it is a local minimizer of program NLP_I for every $I \in \mathcal{I}(z)$. If z is a local minimizer of RNLP then it is a local minimizer of MPCC and if z is a local minimizer of MPCC then it is a local minimizer of TNLP. The reverse implications hold in general only if *strict complementarity* holds at z, i.e., if there is only one vanishing component in each row of the matrix F. In this case $\mathcal{I}(z)$ consists of a single set and equality holds throughout (3). In the complementarity literature, strict complementarity is often referred to as nondegeneracy. We avoid this terminology here since nondegeneracy has a different meaning in the context of optimization problems. The simple example in Figure 1 shows that we cannot expect strict complementarity to hold at a minimizer. In fact, strict complementarity is not only violated at the minimizer $(z_1, z_2) = (0, 1)$, but will remain violated at the locally unique local minimizer of a perturbed problem if the perturbations added to the objective and constraint functions are smooth and have sufficiently small function values and partial derivatives in a neighbourhood of $(0, 1)$

We shall use three constraint qualifications for MPCCs in the sequel which are defined in terms of well-known constraint qualifications for the tightened program TNLP. We say that an MPCC satisfies the Mangasarian-Fromovitz constraint qualification (MFCQ), the strict Mangasarian-Fromovitz constraint qualification (SMFCQ) or the linear independence constraint qualification (LICQ) at a feasible point z if TNLP satisfies the respective condition at the point z. Recall that LICQ is satisfied at the feasible point z of TNLP if the gradients

$$\begin{aligned} &\nabla F_{ki}(z), && (k,i)\colon F_{ki}(z) = 0, \\ &\nabla g_r(z), && r\colon g_r(z) = 0, \\ &\nabla h_s(z), && s = 1, \dots q \end{aligned}$$

are linearly independent, and that this condition implies MFCQ which assumes that the gradients

$$\nabla F_{ki}(z), \quad (k,i)\colon F_{ki}(z) = 0, \qquad \nabla h_s(z), \qquad s = 1, \dots, q$$

are linearly independent and that there exists a vector v orthogonal to these gradients such that

$$\nabla g_r(z) v < 0, \qquad r : g_r(z) = 0.$$

A feasible point z of TNLP satisfies SMFCQ if there exist Lagrange multipliers Γ, λ, μ such that the gradients

$$\begin{aligned} &\nabla F_{ki}(z), && (k,i)\colon F_{ki}(z) = 0, \\ &\nabla g_j(z), && j : \lambda_j > 0, \\ &\nabla h_s(z), && s = 1, \dots, q \end{aligned}$$

are linearly independent and there exists a vector v such that

$$\begin{aligned} \nabla F_{ki}(z) v &= 0, && (k,i) : F_{ki}(z) = 0, \\ \nabla h(z) v &= 0, && \\ \nabla g_j(z) v &= 0, && j : \lambda_j > 0, \\ \nabla g_r(z) v &< 0, && r : g_r(z) = \lambda_r = 0. \end{aligned} \tag{4}$$

It is known that MFCQ holds at a Karush-Kuhn-Tucker point of a nonlinear program if and only if there exists a compact set of Lagrange multipliers; cf. Gauvin (1977), while SMFCQ is equivalent to the existence of a unique Lagrange multiplier (cf. Kyparisis 1985). Notice, if MPCC satisfies MFCQ or LICQ, then the respective constraint qualification is satisfied for each of the smooth programs TNLP, RNLP, and NLP_I, $I \in \mathcal{I}(z)$.

At first glance it seems that LICQ is a rather restrictive assumption. However, as pointed out in Jongen et al. (1986, Remark 7.3.4), LICQ is satisfied for a generic program

$$
\begin{aligned}
\min \quad & f(z) \\
\text{s.t.} \quad & F(z) \geq 0, \\
& g(z) \leq 0, \\
& h(z) = 0,
\end{aligned}
\tag{5}
$$

at every feasible point z. Since the feasible set of MPCC is a subset of the feasible set of (5) and LICQ holds for MPCC, provided it holds for (5), LICQ is indeed a generic constraint qualification for MPCCs in the sense of Jongen et al. (1986).

2. Stationarity conditions.

2.1. Bouligand stationarity. Stationarity concepts in nonlinear programming either rely on the geometry of the feasible set and are formulated in terms of suitable tangent cones or they rely on the representation of the feasible set and are formulated with the aid of first order approximations of the data functions. Constraint qualifications are used to unify both approaches. Here, we follow the latter avenue and refer to Luo et al. (1996) for a thorough study of tangent cones for MPCCs and related problems. We call a feasible point z of an MPCC *Bouligand stationary* (B-stationary) if it is a local minimizer of the *linearized* MPCC which is obtained by linearizing all data functions at the point z. In other words, z is B-stationary if $\nabla f(z)d \geq 0$ for every d satisfying

$$
\begin{aligned}
\min\{\nabla F_{ki}(z)d \mid i\colon F_{ki}(z)=0\} &= 0 \qquad k=1,\dots,m, \\
\nabla g_r(z)d &\leq 0 \qquad r: g_r(z)=0, \\
\nabla h(z)d &= 0.
\end{aligned}
\tag{6}
$$

One can equivalently characterize a B-stationary point as a feasible point of an MPCC which is a stationary point of each smooth program NLP_I with $I \in \mathscr{I}(z)$ (cf. (2)). Thus, in general, the B-stationarity concept has the drawback that its verification involves the consistency of a possibly large number of inequality systems (cf. Theorem 3.3.4 in Luo et al. 1996). Bouligand stationarity is a necessary optimality condition for MPCC under appropriate constraint qualifications, the obvious one being that each of the smooth programs NLP_I with $I \in \mathscr{I}(z)$ satisfies a constraint qualification at z. A stronger condition, which may be easier to verify, is the generalized Mangasarian Fromovitz constraint qualification for piecewise smooth programs suggested by Kuntz and Scholtes (1994).

2.2. Dual stationarity conditions. In standard nonlinear programming B-stationarity is often rephrased by means of linear programming duality. This results in the Karush–Kuhn–Tucker conditions and allows the verification of stationarity through the provision of suitable Lagrange multipliers. For MPCCs meaningful multipliers are available not only at B-stationary points but indeed at every stationary point of the corresponding tightened nonlinear program TNLP. We call such points weakly stationary, i.e., a feasible point z of the MPCC is called *weakly stationary point* if there exist multipliers Γ, λ, μ such that

$$
\begin{aligned}
\nabla_z L(1,z,\Gamma,\lambda,\mu) &= 0,\\
\lambda &\geq 0,\\
g(z)\lambda &= 0,\\
F(z) * \Gamma &= 0,
\end{aligned}
\tag{7}
$$

where $A * B$ denotes the Hadamard, i.e., componentwise, product of two matrices of the same size. While B-stationarity requires that z is a Karush-Kuhn-Tucker point of all programs NLP_I with $I \in \mathscr{I}(z)$; cf. (2), weak stationarity holds already if z is a Karush-Kuhn-Tucker point of one such program. Indeed, a weakly stationary point need not even be a Karush-Kuhn-Tucker point of any adjacent program NLPI — in the problem of Figure 1 the point $(0,1)$ is weakly stationary for any objective function. Nevertheless, weakly stationary points play an important role since, under appropriate assumptions, the multiplier information can be used to identify programs NLP_I for which z is feasible but not stationary. This observation lies at the heart of piecewise SQP methods for MPCCs (cf. e.g., Clark and Westerberg 1990, Luo et al. 1998, Scholtes and Stöhr 1999, and Stöhr 1999).

As a preparation for the next steps on the ladder of dual stationarity concepts we give the following Fritz John type condition which is an immediate consequence of Clarke's (1976) stationarity condition for programs with locally Lipschitzian functions (cf. also Facchinei et al. 1999).

LEMMA 1. *If z is a local minimizer of MPCC then there exist nonvanishing multipliers $(\alpha, \Gamma, \lambda, \mu)$ such that*

$$
\begin{aligned}
\nabla_z L(\alpha,z,\Gamma,\lambda,\mu) &= 0,\\
\alpha &\geq 0,\\
F(z) * \Gamma &= 0,\\
\Gamma_{ki}\Gamma_{kj} &\geq 0, \qquad (i,j): F_{ki}(z) = F_{ki}(z) = 0,\\
\lambda &\geq 0,\\
g(z)\lambda &= 0.
\end{aligned}
\tag{8}
$$

PROOF. By Theorem 1 of Clarke (1976) there exist nonvanishing multipliers $(\alpha, u, \lambda, \mu)$ with $\alpha \geq 0$, $\lambda \geq 0$, $g(z)\lambda = 0$, and

$$0 = \alpha \nabla f(z) + \sum_{k=1}^{m} u_k x_k + \nabla g(z)\lambda + \nabla h(z)\mu$$

with $x_k \in \partial \min\{F_{k1}(z), \ldots, F_{kl}(z)\}$, where ∂ denotes Clarke's subdifferential operator

$$\partial \min\{F_{k1}(z), \ldots, F_{kl}(z)\} = \operatorname{conv}\{\nabla F_{ki}(z) \mid i : F_{ki}(z) = 0\}$$

(cf. Clarke 1983). Hence

$$x_k = \sum_{i=1}^{l} \beta_{ki} \nabla F_{ki}(z)$$

with $\beta_{ki} \geq 0$, $\beta_{ki} F_{ki}(z) = 0$, and $\sum_{i=1}^{l} \beta_{ki} = 1$ and thus the numbers $\Gamma_{ki} = u_k \beta_{ki}$ have all the same sign as u_k which proves the assertion. □

Recall that if z is a local minimizer of MPCC, then it is a local minimizer of TNLP and the reverse statement holds if strict complementarity holds at z. However, if strict complementarity is violated, then the Fritz John condition corresponding to TNLP is weaker than condition (8) since the former condition does not require $\Gamma_{ki}\Gamma_{kj} \geq 0$ for those pairs (i,j) with $F_{ki}(z) = F_{kj}(z) = 0$.

In terms of its applicability the above result has two clear drawbacks. On the one hand, as in standard nonlinear programming, stationarity may result from an ill-posed representation of the feasible set, indicated by a vanishing α. On the other hand, in contrast to standard nonlinear programming, checking the conditions (8) for a given feasible point z is a combinatorial problem if strict complementarity fails. Indeed, the conditions $\Gamma_{ki}\Gamma_{kj} \geq 0$ for all (i,j) with $F_{ki}(z) = F_{kj}(z) = 0$ translate into the alternatives $\Gamma_{ki} \geq 0$ for all i with $F_{ki}(z) = 0$ *or* $\Gamma_{ki} \leq 0$ for all i with $F_{ki}(z) = 0$. The following theorem addresses both concerns separately. The second part of the theorem is reminiscent of Propositions 4.3.5 and 4.3.7 in Luo et al. (1996).

THEOREM 2. *Let z be a local minimizer of the MPCC.*
(1) *If MFCQ holds at z, then there exist multipliers Γ, λ, μ such that*

$$\begin{aligned} \nabla_z L(1, z, \Gamma, \lambda, \mu) &= 0, \\ F(z) * \Gamma &= 0, \\ \Gamma_{ki}\Gamma_{kj} &\geq 0, \qquad (i,j) : F_{ki}(z) = F_{kj}(z) = 0, \\ \lambda &\geq 0, \\ g(z)\lambda &= 0. \end{aligned} \tag{9}$$

(2) *If SMFCQ holds at z, then there exist unique multipliers Γ, λ, μ with*

$$\begin{aligned} \nabla_z L(1, z, \Gamma, \lambda, \mu) &= 0, \\ F(z) * \Gamma &= 0, \\ \Gamma_{ki} &\geq 0, \qquad \text{if } \exists j \neq i : F_{ki}(z) = F_{kj}(z) = 0, \\ \lambda &\geq 0, \\ g(z)\lambda &= 0. \end{aligned} \tag{10}$$

PROOF. (1) If z is a local minimizer then Lemma 1 implies the existence of a nonvanishing vector $(\alpha, \Gamma, \lambda, \mu)$ which satisfies (8). Suppose $\alpha = 0$. Then

$$\sum_{r : g_r(z) = 0} \lambda_r \nabla g_r(z) = - \sum_{(k,i) : F_{ki}(z) = 0} \nabla F_{ki}(z) \Gamma_{ki} - \sum_{s=1}^{q} \nabla h_s(z) \mu_s. \tag{11}$$

Recall that MFCQ holds at z if the gradients

$$\begin{aligned} &\nabla F_{ki}(z), \qquad (k,i) : F_{ki}(z) = 0, \\ &\nabla h_s(z), \qquad s = 1, \ldots, q, \end{aligned} \tag{12}$$

are linearly independent and there exists a vector v in the orthogonal complement of these gradients such that $\nabla g_r(z)v < 0$ for every $r : g_r(z) = 0$. Hence

$$\sum_{r:g_r(z)=0} \lambda_r \nabla g_r(z)v = 0$$

and thus $\lambda = 0$ since $\lambda \geq 0$, $g(z)\lambda = 0$, and $\nabla g_r(z)v < 0$ for every r with $g_r(z) = 0$. However, if $\lambda = 0$, then Γ and μ vanish in view of (11) and the assumed linear independence of the gradients (12). This contradicts the assumption that $(\alpha, \Gamma, \lambda, \mu)$ is nonvanishing and thus proves that $\alpha > 0$. Scaling yields $\alpha = 1$.

(2). Let Γ, λ, μ be multipliers corresponding to the stationary point z of TNLP and suppose $\Gamma_{ki} < 0$ for some i with $F_{ki} = F_{kj} = 0$ and $j \neq i$. We may assume without loss of generality that $i = k = 1$ and $j = 2$. Recall that SMFCQ holds at z if the gradients $\nabla F_{ki}(z)$, $(k,i) : F_{ki}(z) = 0$, $\nabla g_j(z)$, $j : \lambda_j > 0$, $\nabla h_s(z)$, $s = 1, \ldots, q$, are linearly independent and there exists a vector v satisfying (4). Hence there exists a vector d such that

$$\begin{aligned}
\nabla F_{11}(z)d &= 1, & & \\
\nabla F_{ki}(z)d &= 0, & & (k,i) : F_{ki}(z) = 0,\ (k,i) \neq (1,1), \\
\nabla g_j(z)d &= 0, & & j : \lambda_j > 0, \\
\nabla h(z)d &= 0, & &
\end{aligned}$$

and, passing to a vector $d + \sigma v$ for sufficiently large $\sigma > 0$ if necessary, $\nabla g_r(z)d < 0$ for every r such that $g_r(z) = \lambda_r = 0$. Since conditions (8) hold with $\alpha = 1$, we conclude that $\nabla f(z)d = \Gamma_{11} < 0$. Hence z is not a stationary point of the smooth program NLP_I with $I = \{(k,i) \mid F_{ki}(z) = 0,\ (k,i) \neq (1,1)\}$. Notice that SMFCQ implies MFCQ for program NLP_I; hence z is not a local minimizer of NLP_I. It follows from (3) that z is not a local minimizer of MPCC, which, however, contradicts the assumption. We conclude that $\Gamma_{ki} \geq 0$, provided there exists an index $j \neq i$ such that $F_{ki}(z) = F_{kj}(z) = 0$. $\square$

Notice that the proof of part (ii) not only shows that the point z is not optimal if the multiplier Γ_{11} is negative, but also provides a decomposition piece NLP_I along which a descent is possible. This fact has been exploited algorithmically in Scholtes and Stöhr (1999).

We call a feasible point z of the MPCC *Clarke stationary* (C-stationary) if there exist multipliers Γ, λ, μ such that (9) holds. If there exist multipliers such that (10) holds, then z is called *strongly stationary*. Note that the conditions (10) coincide with the Karush-Kuhn-Tucker conditions for the relaxed program RNLP at z which, in general, has locally a larger feasible set than the MPCC. The following example shows that

- one cannot replace SMFCQ in the second statement of the foregoing theorem by MFCQ and
- even in the case of linear data functions an additional constraint qualification is necessary to ensure that a local minimizer is strongly stationary.

EXAMPLE 3. The origin is the unique minimizer of the program

$$\begin{aligned}
\min \quad & z_1 + z_2 - z_3, \\
\text{s.t} \quad & \min\{z_1, z_2\} = 0, \\
& -4z_1 + z_3 \leq 0, \\
& -4z_2 + z_3 \leq 0,
\end{aligned}$$

while the relaxed program RNLP has no minimizer since $z_1 = \rho,\ z_2 = \rho,\ z_3 = 3\rho$ is feasible for every $\rho \geq 0$ and yields an objective value of $-\rho$.

Whether and how the SMFCQ assumption of the second part of Theorem 2 can be substantially weakened seems to be an interesting and apparently difficult question.

The concepts of weak stationarity, C-stationarity, B-stationarity, and strong stationarity are all equivalent if strict complementarity holds at z. However, in general B-stationarity is stronger than C-stationarity which, in turn, is stronger than weak stationarity. In fact, if we choose the objective function $f(z_1, z_2) = 2z_1 + z_2$ in the example of Figure 1, then the point $(z_1, z_2) = (0, 1)$ is weakly stationary but neither C-stationary, nor B-stationary. If we choose $f(z_1, z_2) = z_1 - 2z_2$ as the objective function, then the same point is C-stationary but not B-stationary. The next result clarifies the relation between strongly stationary points and B-stationary points.

THEOREM 4. *If z is a strongly stationary point of the MPCC, then z is a B-stationary point of the MPCC. If SMFCQ holds at z, then the reverse implication holds as well.*

PROOF. Recall that z is a strongly stationary point of the MPCC if and only if z is a KKT point of the relaxed program RNLP. Since a KKT multiplier for the RNLP is a KKT multiplier for every program NLP_I, $I \in \mathcal{I}(z)$, z is a B-stationary point of the MPCC. To see the converse, suppose z is B-stationary. Recall that under SMFCQ the KKT multiplier (Γ, λ, μ) of TNLP is unique. Since the multipliers of every program NLP_I are multipliers of TNLP as well, (Γ, λ, μ) is the unique multiplier for every program NLP_I. Hence $\Gamma_{ki} \geq 0$ if there is an index $j \neq i$ with $F_{ki}(z) = F_{kj}(z) = 0$ which implies that z is a strongly stationary point of the MPCC. □

Note that the reverse implication of the statement of the above theorem does not hold if SMFCQ is replaced by MFCQ as is illustrated by Example 3. The following table summarizes the relations between the stationarity concepts that we have discussed.

$$
\begin{array}{ccc}
\text{B-stationary point} & \Rightarrow & \text{weakly stationary point} \\
\Uparrow & & \Uparrow \\
\text{strongly stationary point} & \Rightarrow & \text{C-stationary point.}
\end{array}
$$

Some implications can be added if additional assumptions are imposed. We have already mentioned that all concepts coincide under strict complementarity. In view of Theorem 2, B-stationarity implies C-stationarity if MFCQ holds, while B-stationarity implies strong stationarity if SMFCQ or the stronger LICQ holds. We emphasize again that weak and strong stationarity, are, in principle, easy to verify since they are optimality conditions of the standard nonlinear programs TNLP and RNLP, respectively. The B-stationarity concept, however, is more natural than the former since it takes the inherent combinatorial structure of the feasible region of the MPCC into account. The combinatorial aspects of B- and C-stationarity make these conditions difficult to verify in general. It is a very fortunate circumstance that strong stationarity and B-stationarity are equivalent under the generic constraint qualification LICQ and that therefore B-stationarity is easy to verify for many MPCCs.

We refer the reader to the recent paper by Pang and Fukushima (1998) for a more extensive study of the B-stationarity condition and to Outrata (1999) for a study of a further stationarity concept, based on Mordukhovich's generalized differential calculus, which bridges the gap between C-stationarity and strong stationarity but retains a combinatorial aspect which complicates its verification if strict complementarity fails.

2.3. Stationary points of exact penalty functions. We next relate the stationarity concepts for an MPCC to nonsmooth stationary point concepts for a corresponding exact penalty function. The starting point is the reformulation of the MPCC as

$$\begin{array}{ll} \min & f(z) \\ \text{s.t.} & \phi(z)=0, \end{array} \tag{13}$$

where

$$\phi(z)=\max_{k,r,s}\{|\min\{F_{k1}(z),\ldots,F_{kl}(z)\}|,\ g_r(z),|h_s(z)|\}. \tag{14}$$

The function ϕ is a real-valued nonnegative piecewise smooth function. We say that a continuous nonnegative function ϕ satisfies the *exact penalty constraint qualification* (EPCQ) at a root z if for every smooth function f the point z is a local minimizer of $f+\pi\phi$ for every sufficiently large π, provided it was a local minimizer of (13); EPCQ is said to be satisfied at a feasible point z of an MPCC if it holds for the function ϕ given in (14).

It follows from the results of Robinson (1981) that (EPCQ) holds for piecewise affine functions and thus for problems with linear equations and inequalities and suitably parametrized linear complementarity constraints. Conditions which guarantee EPCQ to hold for piecewise differentiable functions ϕ are given in Scholtes and Stöhr (1999). It follows, e.g., from the results in the latter reference that function (14) satisfies EPCQ at a root z if MFCQ holds at the feasible point z of the MPCC.

Condition EPCQ allows the derivation of optimality conditions for constrained programs from unconstrained optimality conditions for nonsmooth functions (cf., e.g., Burke 1991). Our first result relates the weak stationary point concept for MPCCs to Clarke's (1983) optimality condition for the penalty function.

PROPOSITION 5. *Let ϕ be given by* (14). *A feasible point z of an MPCC is weakly stationary if and only if $0\in\partial(f+\pi\phi)(z)$ for every sufficiently large π.*

The proof of this proposition is straightforward and therefore omitted. The result implies that a local minimizer of an MPCC is a weak stationary point if EPCQ holds. The proposition highlights the weakness of stationarity derived via Clarke's calculus from exact penalty functions in the presence of nonsmooth equality constraints (cf. Polak et al. 1983). Indeed, even in the presence of LICQ, a Clarke stationary point of an exact penalty function need not be a C-stationary point of the MPCC. Consider, e.g., the program $\min z_1-z_2$ s.t. $\min\{z_1,z_2\}=0$. The origin is not a C-stationary point of the MPCC but it is a Clarke stationary point of $z_1-z_2+\pi\phi(z_1,z_2)$ for every sufficiently large penalty parameter $\pi\geq 1$ and, indeed, the latter statement remains true for any smooth objective function.

The next result is the pendant to the former result for B-stationarity. In contrast to the stationarity concept for penalty functions based on Clarke's derivative, the stationarity concept based on the Bouligand derivative does indeed capture the B-stationarity concept for the MPCC. Recall that a function g is called Bouligand differentiable at z if the directional derivative

$$g'(z;d)=\lim_{\substack{\Delta\to 0\\ \Delta>0}}\frac{g(z+\Delta d)-g(z)}{\Delta}$$

exists for every direction d and the approximation is uniform, i.e.,

$$\lim_{d\to 0}\frac{g(z+d)-g(z)-g'(z;d)}{\|d\|}=0$$

(cf. Robinson 1987). Notice that the exact penalty function $g(z) = f(z) + \pi\phi(z)$ is Bouligand differentiable (cf., e.g., Scholtes 1994). Having minimization problems in mind, we call a point of z a *B-stationarity point of* g if $g'(z; d) \geq 0$.

PROPOSITION 6. *Let ϕ be given by* (14). *A feasible point z of an MPCC is B-stationary if and only if it is a B-stationary point of the function $f + \pi\phi$ for every sufficiently large π.*

PROOF. The calculus rules for B-derivatives imply that (6) holds if and only if $\phi'(z; d) = 0$ (cf., e.g., Scholtes 1994). Hence $\nabla f(z)d \geq 0$ for every d satisfying (6) if and only if the origin is a minimizer of the program

$$\begin{array}{ll} \min & \nabla f(z)d \\ \text{s.t.} & \phi'(z; d) = 0. \end{array} \tag{15}$$

Since the B-derivative of a piecewise smooth function is piecewise linear (cf. Scholtes 1994), the result of Robinson (1981) implies the existence of a local error bound for the constraints of (15) and thus, in view of the principle of exact penalization as outlined, e.g., in Clarke (1983), Luo et al. (1996), and Scholtes and Stöhr (1999), the origin is a local minimizer of (15) if and only if it is a local minimizer of the function $\nabla f(z)d + \pi\phi'(z; d)$ for every sufficiently large π (cf. Scholtes and Stöhr 1999, Corollary 2.4). $\square$

Notice that this result implies that a local minimizer of an MPCC is a B-stationary point, provided EPCQ holds.

3. Second order conditions. Using the decomposition approach based on (3), the classical second order optimality conditions of Fiacco and McCormick (1968) can be readily extended to MPCCs. To illustrate this we provide a short proof of the following theorem which is similar to results obtained in Chapter 5 of Luo et al. (1996). Here, a direction d is called a *critical direction* for an MPCC at a weakly stationary point z if (6) holds and $\nabla f(z)d = 0$. Notice that d is a critical direction for an MPCC at z if and only if it is a critical direction for one of the smooth programs NLP_I, $I \in \mathcal{I}(z)$ at z.

THEOREM 7 (SECOND ORDER OPTIMALITY CONDITIONS).

(1) *Let z be a local minimizer of MPCC and suppose SMFCQ holds at z. Then z is a strongly stationary point with unique multipliers (Γ, λ, μ) satisfying* (10) *and the inequality*

$$d^\top \nabla^2_{zz} L(1, z, \Gamma, \lambda, \mu) d \geq 0 \tag{16}$$

holds for every critical direction d.

(2) *Let z be a strongly stationary point of an MPCC. If for every nonvanishing critical direction d there exist multipliers (Γ, λ, μ) satisfying the conditions* (10) *and*

$$d^\top \nabla^2_{zz} L(1, z, \Gamma, \lambda, \mu) d > 0, \tag{17}$$

then z is a strict local minimizer of the MPCC.

PROOF.

(1) The existence and uniqueness of the multipliers follow from Theorem 2. Notice that, in view of SMFCQ, the unique KKT multipliers for TNLP are also the unique KKT multipliers of every program NLP_I, i.e., SMFCQ and, a fortiori, MFCQ holds at z for every program NLP_I. Since for any critical direction d, there is an index set $I \in \mathcal{I}(z)$ such that d is critical for NLP_I, we can apply the second order necessary condition for this smooth program (cf., e.g., Burke 1991), to obtain the first assertion.

(2) To see that the second statement holds, notice that every multiplier (Γ, λ, μ) satisfying the conditions (10) is also a KKT multiplier of every program NLP_I, $I \in \mathscr{I}(z)$. Since every critical direction of one of the latter programs is a critical direction of the MPCC, the assumptions imply that the second order sufficient condition holds for every smooth program NLP_I, $I \in \mathscr{I}(z)$. Hence z is a strict local minimizer of every program NLP_I and thus, in view of (3), a strict local minimizer of the MPCC. □

In view of Theorem 4 the sufficient condition of the latter theorem is, in the presence of SMFCQ or LICQ, also sufficient for a B-stationary point to be a strict local minimizer. Recall that the strong stationarity condition for an MPCC coincides with the KKT condition for the relaxed program RNLP. This equivalence does not hold for the above second order sufficient condition. To see this, consider the program

$$\begin{aligned} \min \quad & (z_1+1)^2 + (z_2+1)^2 - 3z_1z_2 \\ \text{s.t.} \quad & \min\{z_1, z_2\} = 0. \end{aligned}$$

The origin is a strongly stationary point and thus a KKT point of the relaxed program RNLP:

$$\begin{aligned} \min \quad & (z_1+1)^2 + (z_2+1)^2 - 3z_1z_2 \\ \text{s.t.} \quad & z_1 \geq 0,\ z_2 \geq 0. \end{aligned}$$

However, although the above second order sufficient condition is satisfied for the original program, $d = (1, 1)$ is a feasible second order descent direction for the relaxed problem.

We shall next see that the second order condition for MPCCs is equivalent to the second order sufficiency condition of the relaxed problem with a suitably augmented objective function. To this end we recall that one can express the constraint $\min\{F_{k1}(z), \ldots, F_{kl}(z)\} = 0$ equivalently as $F_{ki}(z) \geq 0$, $i = 1, \ldots, l$, and $\prod_{i=1}^{l} F_{ki}(z) = 0$.

THEOREM 8. *Let z be a feasible point of an MPCC and consider the nonlinear program*

$$\begin{aligned} \min \quad & f(x) + \pi\psi_K(x) \\ s.t. \quad & F_I(x) = 0, \\ & F_{I^c}(x) \geq 0, \\ & g(x) \leq 0, \\ & h(x) = 0, \end{aligned} \tag{18}$$

where $I = \{(k,i) \mid F_{kj}(z) > 0 \text{ for all } j \neq i\}$,

$$K = \{k \mid \text{there exist } i \neq j\colon\ F_{ki}(z) = F_{kj}(z) = 0\},$$

$$\pi \geq 0,\ \psi_K(x) = \sum_{k \in K} \prod_{i=1}^{l} F_{ki}(x).$$

(1) *The point z is a strongly stationary point of the MPCC if and only if it is a KKT point of* (18). *Moreover, the set of multipliers for both problems coincide.*

(2) *If a strongly stationary point z of the MPCC satisfies the second order sufficient condition for problem* (18), *then it also satisfies the second order sufficient conditions*

for the MPCC. Moreover, if the index set

$$S = \{k \mid \text{there exist distinct } r,s,t \text{ with } F_{kr}(z) = F_{ks}(z) = F_{kt}(z) = 0\}$$

is empty then the reverse statement holds for every sufficiently large π.

PROOF. Direct calculation shows that

$$\nabla \psi_K(x) = \sum_{k \in K} \sum_{i=1}^{l} \nabla F_{ki}(x) \prod_{\substack{j=1 \\ j \neq i}}^{l} F_{kj}(x)$$

and, in view of the definition of K and the feasibility of z, we obtain $\nabla \psi_K(z) = 0$, which implies the first statement.

To see the second statement note that the Hessian with respect to z of the Lagrangian corresponding to (18) is of the form

$$\nabla_{zz}^2 L(1, z, \Gamma, \lambda, \mu) + \pi \nabla^2 \psi_K(z),$$

where L is the Lagrangian (1) corresponding to the MPCC. Hence the strongly stationary point z satisfies the second order sufficient condition for (18) if and only if

$$d^\top \nabla_{zz}^2 L(1, z, \Gamma, \lambda, \mu) d + \pi d^\top \nabla^2 \psi_K(z) d > 0$$

for every nonvanishing d such that

$$\begin{aligned} \nabla F_I(z) d &= 0, \\ \nabla F_J(z) d &\geq 0, \\ \nabla g_R(z) d &\leq 0, \\ \nabla h(z) d &= 0, \\ \nabla f(z) d &= 0, \end{aligned} \tag{19}$$

where $R = \{r \mid g_r(z) = 0\}$ and $J = \{(k,i) \in I^c \mid F_{ki}(z) = 0\}$, while the second order sufficient condition for the MPCC requires that

$$d^\top \nabla_{zz}^2 L(1, z, \Gamma, \lambda, \mu) d > 0$$

for every nonvanishing d such that

$$\begin{aligned} \min\{\nabla F_{ki}(z) d \mid i\colon\ F_{ki}(z) = 0\} &= 0, \qquad k = 1, \ldots, m, \\ \nabla g_R(z) d &\leq 0, \\ \nabla h(z) d &= 0, \\ \nabla f(z) d &= 0. \end{aligned} \tag{20}$$

Direct calculation yields

$$\begin{aligned} \nabla^2 \psi_K(x) = &\sum_{k \in K} \sum_{i=1}^{l} \sum_{\substack{j=1 \\ j \neq i}}^{l} \left(\prod_{\substack{s=1 \\ s \neq i,j}}^{l} F_{ks}(x) \right) [\nabla F_{ki}(x)^\top \nabla F_{kj}(x)] \\ &+ \sum_{k \in K} \sum_{i=1}^{l} \left(\prod_{\substack{j=1 \\ j \neq i}}^{l} F_{kj}(x) \right) \nabla^2 F_{ki}(x), \end{aligned}$$

where the product over an empty set of indices, should it appear in these formulae, is understood to equal unity. In view of the definition of K and the feasibility of z we obtain

$$\nabla^2\psi_K(z)=\sum_{k\in K}\sum_{i=1}^{l}\sum_{\substack{j=1\\ j\neq i}}^{l}\beta_{kij}[\nabla F_{ki}(z)^\top\nabla F_{kj}(z)]$$

with

$$\beta_{kij}=\prod_{\substack{s=1\\ s\neq i,j}}^{l}F_{ks}(z)\geq 0.$$

Now suppose the second order sufficient condition for (18) holds. Since every d satisfying (20) also satisfies (19) we obtain

$$d^\top\nabla^2_{zz}L(1,z,\Gamma,\lambda,\mu)+\pi d^\top\nabla^2\psi_K(z)d>0$$

for every d satisfying (20). Since $\beta_{kij}>0$ if and only if $F_{ks}(z)>0$ for every $s\neq i,j$ and in this case the first constraint of (20) implies that at least one of the terms $\nabla F_{ki}(z)d$ and $\nabla F_{kj}(z)d$ vanishes, we further conclude that

$$d^\top\nabla^2\psi_K(z)d=\sum_{k\in K}\sum_{i=1}^{l}\sum_{\substack{j=1\\ j\neq i}}^{l}\beta_{kij}d^\top\nabla F_{ki}(z)^\top\nabla F_{kj}(z)d=0$$

for every d satisfying (20); hence the second order sufficient condition for MPCCs holds.

Now suppose the second order sufficient condition holds for the MPCC and S is empty. Then there exists a number $\varepsilon>0$ such that

$$d^\top\nabla^2_{zz}L(1,z,\Gamma,\lambda,\mu)d>0$$

for every d with $\|d\|=1$ and

$$\begin{aligned}\nabla F_I(z)d &= 0,\\ \nabla F_J(z)d &\geq 0,\\ \min\{\nabla F_{ki}(z)d\mid i:F_{ki}(z)=0\} &< \varepsilon, \qquad k\in K,\\ \nabla g_R(z)d &\leq 0,\\ \nabla h(z)d &= 0,\\ \nabla f(z)d &= 0.\end{aligned}\tag{21}$$

In fact, if no such ε exists then one could construct a sequence d_n converging to a direction d of unit norm which is feasible for (20) and satisfies $d^\top\nabla^2_{zz}L(1,z,\Gamma,\lambda,\mu)d\leq 0$. This, however, contradicts the assumption that the second order sufficient condition holds for the MPCC. Now let

$$\kappa=\min\{d^\top\nabla^2_{zz}L(1,z,\Gamma,\lambda,\mu)d\mid (19)\text{ holds and }\|d\|=1\},$$

and let π be such that

$$\pi\sum_{i=1}^{l}\sum_{\substack{j=1\\ j\neq i}}^{l}\beta_{kij}>\frac{-\kappa}{\varepsilon^2}\tag{22}$$

for every $k \in K$. Since S is empty, such a number π exists since for every $k \in K$ there exist indices $i \neq j$ such that $\beta_{kij} > 0$ and thus $\sum_{i=1}^{l} \sum_{j=1; j \neq i}^{l} \beta_{kij} > 0$. Now let d be a direction satisfying (19) with $\|d\| = 1$. If, on the one hand, $\min\{\nabla F_{ki}(z)d \mid i : F_{ki}(z) = 0\} < \varepsilon$ for every $k \in K$ then

$$d^{\top} \nabla_{zz}^2 L(1, z, \Gamma, \lambda, \mu) d + \pi d^{\top} \nabla^2 \psi_K(z) d > 0.$$

Indeed, the positivity of the first term was established above, while the nonnegativity of the second term is a consequence of the formula for $\nabla^2 \psi_K(z)$ and the validity of the inequalities $\nabla F_{ki}(z) d \geq 0$ for every $(k, i) \in I \cup J$. If, on the other hand, there exists an index $k \in K$ such that $\nabla F_{ki}(z) d \geq \varepsilon$ for every $i : F_{ki}(z) = 0$ then

$$d^{\top} \nabla_{zz}^2 L(z, \Gamma, \lambda, \mu) d + \pi d^{\top} \nabla^2 \psi_K(z) d \geq \kappa + \pi \sum_{i=1}^{l} \sum_{\substack{j=1 \\ j \neq i}}^{l} \beta_{kij} \varepsilon^2 > 0$$

in view of (22). $\square$

Note that the set S in part (ii) of the above theorem is always empty if $l = 2$, i.e., if the constraints of the MPCC are standard nonlinear complementarity constraints.

4. Stability of stationary points. In this final section we extend the stability results of Robinson (1980) and Kojima (1980) to MPCCs. For simplicity we confine ourselves to parametric perturbations; Kojima's results for nonparametric perturbations can be extended in a similar way. Our analysis is based on the following version of the results of Robinson and Kojima (cf. Scholtes 1994).

THEOREM 9. *Consider the parametric program*

$$\begin{array}{lll} NLP_I(t) & \min\limits_{z} & f(z, t) \\ & \text{s.t.} & F_I(z, t) = 0, \\ & & F_{I^c}(z, t) \geq 0, \\ & & g(z, t) \leq 0, \\ & & h(z, t) = 0, \end{array}$$

where $t \in \mathbb{R}^s$ is a parameter, $I \subseteq \{(k, i) \mid k = 1, \ldots, m, \ i = 1, \ldots, l\}$, and I^c denotes the complement of I. Suppose that all functions involved in the formulation of $NLP_I(t)$ are twice continuously differentiable. Let $\bar{z}$ be a Karush-Kuhn-Tucker point of $NLP_I(\bar{t})$ and let $\bar{\Gamma}, \bar{\lambda}, \bar{\mu}$ be corresponding multipliers associated with the Lagrangian

$$L(1, z, \Gamma, \lambda, \mu, t) = f(z, t) - F(z, t)\Gamma + g(z, t)\lambda + h(z, t)\mu.$$

Suppose for all index sets K and R with

$$I \cup \{(k, i) \mid \bar{\Gamma}_{ki} > 0\} \subseteq K \subseteq \{(k, i) \mid F_{ki}(\bar{z}, \bar{t}) = 0\},$$

$$\{r \mid \bar{\lambda}_r > 0\} \subseteq R \subseteq \{r \mid g_r(\bar{z}, \bar{t}) = 0\},$$

the matrices

$$\begin{pmatrix} \nabla^2_{zz}L(1,\bar z,\bar\Gamma,\bar\lambda,\bar\mu,\bar t) & -\nabla_z F_K(\bar z,\bar t)^\top & \nabla_z g_R(\bar z,\bar t)^\top & \nabla_z h(\bar z,\bar t)^\top \\ \nabla_z F_K(\bar z,\bar t) & 0 & 0 & 0 \\ -\nabla_z g_R(\bar z,\bar t) & 0 & 0 & 0 \\ \nabla_z h(\bar z,\bar t) & 0 & 0 & 0 \end{pmatrix} \tag{23}$$

have the same nonvanishing determinantal sign.

(1) *There exist open neighbourhoods Z of $\bar z$ and T of $\bar t$, and piecewise smooth functions $z(\cdot),\Gamma(\cdot),\lambda(\cdot),\mu(\cdot)$ defined on T such that for every $t\in T$ the point $z(t)$ is the unique Karush-Kuhn-Tucker point of $NLP_I(t)$ in Z and $\Gamma(t),\lambda(t),\mu(t)$ are the unique corresponding Lagrange multipliers.*

(2) *If, in addition, $d^\top\nabla^2_{zz}L(1,\bar z,\bar\Gamma,\bar\lambda,\bar\mu,\bar t)d>0$ for every nonvanishing vector d with*

$$\begin{aligned} \nabla_z F_{ki}(\bar z,\bar t)d &= 0, \qquad (k,i)\in I \text{ or } \bar\Gamma_{ki}>0,\\ \nabla_z g_r(\bar z,\bar t)d &= 0, \qquad r:\bar\lambda_r>0,\\ \nabla_z h(\bar z,\bar t)d &= 0, \end{aligned}$$

then $z(t)$ is a strict local minimizer of $NLP_I(t)$ for every t sufficiently close to $\bar t$.

Note that the nonvanishing determinantal sign property implies that LICQ holds at $\bar z$ and thus that the multipliers $\bar\Gamma,\bar\lambda,\bar\mu$ are unique.

We assume the MPCC to be embedded in a parametric family of problems

$$\begin{aligned} \text{MPCC}(t)\quad \min_z\quad & f(z,t)\\ \text{s.t.}\quad & \min\{F_{k1}(z,t),\dots,F_{kl}(z,t)\}=0, \qquad k=1,\dots,m,\\ & g(z,t)\le 0,\\ & h(z,t)=0, \end{aligned}$$

where $t\in\mathbb{R}^s$ is a parameter, and all functions involved are C^2-functions. The weak stationary point conditions for the parametric problem MPCC(t) are

$$\begin{aligned} \nabla_z L(1,z,\Gamma,\lambda,\mu,t) &=0,\\ \min\{F_{k1}(z,t),\dots,F_{kl}(z,t)\} &=0, \qquad k=1,\dots,m,\\ h(z,t) &=0,\\ g(z,t) &\le 0,\\ \lambda &\ge 0,\\ g_r(z,t)\lambda_r &=0, \qquad r=1,\dots,p,\\ F_{ki}(z,t)\Gamma_{ki} &=0, \qquad k=1,\dots,m,\ i=1,\dots,l, \end{aligned} \tag{24}$$

where

$$L(\alpha,z,\Gamma,\lambda,\mu,t)=\alpha f(z,t)-F(z,t)\Gamma+g(z,t)\lambda+h(z,t)\mu$$

is the Lagrangian corresponding to MPCC(t). Recall that C-stationarity requires in addition that $\Gamma_{ki}\Gamma_{kj}\ge 0$ if $F_{ki}(z,t)=F_{kj}(z,t)=0$. Moreover, by Theorem 4, a weakly stationary point is B-stationary, provided $\Gamma_{ki}\ge 0$ for every (k,i) such that $F_{ki}=F_{kj}(z,t)=0$ for some

$j \neq i$ and the reverse implication holds in the presence of LICQ. In the sequel we shall study the local behaviour of stationary points as well as local minimizers as a function of the parameter t. We begin with two elementary examples.

EXAMPLE 10.

(1) Consider the parametric program

$$\min_z \quad (z_1 - t)^2 + (z_2 - t)^2$$
$$\text{s.t.} \quad \min\{z_1, z_2\} = 0,$$

with parameter $t \in \mathbb{R}$. The solution is the projection of the point $t(1,1)$ onto the boundary of the positive orthant in $\mathbb{R}^2$. Thus the problem has the origin as the unique minimizer if $t \leq 0$, while it has the two minimizers $(t, 0)$ and $(0, t)$ if $t > 0$. Moreover, the origin is a C-stationary point for every t, while it is B-stationary only if $t \leq 0$.

(2) Consider the parametric program

$$\min_z \quad -(z_1 - 1)^2 - (z_2 - t)^2$$
$$\text{s.t.} \quad \min\{z_1, z_2\} = 0,$$

where $t \in \mathbb{R}$ is the parameter. Here, the origin is a local minimizer for $t > 0$, a B-stationary point for $t \geq 0$, and only a weakly stationary point for $t < 0$.

In order to rule out the branching behaviour of the solution set that we observe in the first example at $t = 0$ we employ the following condition.

(ULSC). A solution $(z, \Gamma, \lambda, \mu)$ of (7) is said to satisfy the *upper level strict complementarity condition* (ULSC) if $\Gamma_{ki} \neq 0$ for every (k, i) such that $F_{ki}(z) = F_{kj}(z) = 0$ for some $j \neq i$.

This condition is satisfied if strict complementarity holds at the lower level, i.e., for every index k the equation $F_{ki}(z) = 0$ holds for exactly one index i. However, upper level strict complementarity is less stringent than lower level strict complementarity. In Example 10, lower level strict complementarity is violated at the origin for every t while (ULSC) is violated only for $t = 0$. Recall that z is a strongly stationary point of an MPCC if it is a KKT-point of the relaxed program RNLP. If strict complementarity holds for a multiplier vector (Γ, λ, μ) corresponding to the relaxed program RNLP, then (ULSC) holds for this multiplier vector. The reverse statement, however, is not true since (ULSC) makes no assumptions on the multipliers λ_r corresponding to the inequalities $g_r(z) \leq 0$. The following theorem is a generalization of the stability results of Robinson (1980) and Kojima (1980) to MPCCs.

THEOREM 11. *Let $\bar{z}$ be a weakly stationary point of the program MPCC($\bar{t}$) and let $\bar{\Gamma}, \bar{\lambda}, \bar{\mu}$ be corresponding multipliers satisfying* (24). *Suppose (ULSC) holds at $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ and all matrices*

$$\begin{pmatrix} \nabla^2_{zz} L(1, \bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}, \bar{t}) & -\nabla_z F_I(\bar{z}, \bar{t})^\top & \nabla_z g_R(\bar{z}, \bar{t})^\top & \nabla_z h(\bar{z}, \bar{t})^\top \\ \nabla_z F_I(\bar{z}, \bar{t}) & 0 & 0 & 0 \\ -\nabla_z g_R(\bar{z}, \bar{t}) & 0 & 0 & 0 \\ \nabla_z h(\bar{z}, \bar{t}) & 0 & 0 & 0 \end{pmatrix} \tag{25}$$

with $I = \{(k, i) \mid F_{ki}(\bar{z}, \bar{t}) = 0\}$ and $\{r \mid \bar{\lambda}_r > 0\} \subseteq R \subseteq \{r \mid g_r(\bar{z}, \bar{t}) = 0\}$ have the same non-vanishing determinantal sign.

(1) *There exist open neighbourhoods Z of $\bar{z}$ and T of $\bar{t}$, and piecewise smooth functions $z(\cdot), \Gamma(\cdot), \lambda(\cdot), \mu(\cdot)$, defined on T such that for every $t \in T$ the point $z(t)$ is the unique weakly stationary point of MPCC(t) in Z and $\Gamma(t), \lambda(t), \mu(t)$ are the unique corresponding multipliers.*

(2) *If $\bar{z}$ is a C-stationary (B-stationary) point, then $z(t)$ is a C-stationary (B-stationary) point for every t sufficiently close to $\bar{t}$.*

(3) *If $\bar{z}$ is a B-stationary point and $d^\top \nabla^2_{zz} L(1, \bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}, \bar{t}) d > 0$ for every nonvanishing vector d with*

$$\begin{aligned} \nabla_z F_{ki}(\bar{z}, \bar{t}) d &= 0, \qquad (k,i): F_{ki}(\bar{z}, \bar{t}) = 0, \\ \nabla_z g_r(\bar{z}, \bar{t}) d &= 0, \qquad r: \bar{\lambda}_r > 0, \\ \nabla_z h(\bar{z}, \bar{t}) d &= 0, \end{aligned}$$

then $z(t)$ is a strict local minimizer of MPCC(t) for every t sufficiently close to $\bar{t}$.

Proof. If (ULSC) holds at a weakly stationary point $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu})$ of problem MPCC(t) corresponding to the parameter value $\bar{t}$, then the conditions (24) hold for some $(z, \Gamma, \lambda, \mu, t)$ in a neighbourhood of $(\bar{z}, \bar{\Gamma}, \bar{\lambda}, \bar{\mu}, \bar{t})$ if and only if

$$\begin{aligned} \nabla_z L(1, z, \Gamma, \lambda, \mu, t) &= 0, \\ F_{ki}(z,t) &= 0, \qquad (k,i): F_{ki}(\bar{z}, \bar{t}) = 0, \\ h(z,t) &= 0, \\ g(z,t) &\le 0, \\ \lambda &\ge 0, \\ g_r(z,t)\lambda_r &= 0, \qquad r = 1, \dots, p, \\ \Gamma_{ki} &= 0, \qquad (k,i): F_{ki}(\bar{z}, \bar{t}) > 0. \end{aligned} \tag{26}$$

These conditions are equivalent to the Karush-Kuhn-Tucker conditions for the smooth parametric program $\mathrm{NLP}_I(t)$ with $I = \{(k,i) \mid F_{ki}(\bar{z}, \bar{t}) = 0\}$. Theorem 9 thus implies the first statement.

The second statement is an immediate consequence of the continuity of the functions $z(t)$ and $\Gamma(t)$ and (ULSC). Note the nonvanishing determinantal sign condition implies that LICQ holds at $\bar{z}$ and hence by Theorem 4, B-stationarity is equivalent to the condition (10).

To see that the third statement holds, note that, in view of (ULSC), the conditions of part (ii) of Theorem 9 are satisfied for the index set

$$I = \{(k,i) \mid F_{kj}(\bar{z}, \bar{t}) > 0 \text{ for every } j \neq i\}.$$

Thus $z(t)$ is a local minimizer of the corresponding program $\mathrm{NLP}_I(t)$. Since $z(t)$ is feasible for MPCC(t), it follows that $z(t)$ is a local minimizer for the latter program as well. □

In view of the results of Ralph and Dempe (1995) the B-derivative of the functions $z(.), \Gamma(.), \lambda(.), \mu(.)$ in the direction τ can be calculated by solving a quadratic program. In fact, if the assumptions of Theorem 11 hold then for each fixed τ the quadratic program

$$\begin{aligned}
\min \quad & \tfrac{1}{2}\zeta^\top \nabla^2_{zz} L(1,\bar z,\bar\Gamma,\bar\lambda,\bar\mu,\bar t)\zeta + \zeta^\top \nabla^2_{zt} L(1,\bar z,\bar\Gamma,\bar\lambda,\bar\mu,\bar t)\tau \\
\text{s.t.} \quad & -\nabla_z F_I(\bar z,\bar t)\zeta = \nabla_t F_I(\bar z,\bar t)\tau, \\
& \nabla_z g_r(\bar z,\bar t)\zeta \le -\nabla_t g_r(\bar z,\bar t)\tau, \qquad r: g_r(\bar z,\bar t) = \bar\lambda_r = 0, \\
& \nabla_z g_r(\bar z,\bar t)\zeta = -\nabla_t g_r(\bar z,\bar t)\tau, \qquad r: \bar\lambda_r > 0, \\
& \nabla_z h(\bar z,\bar t)\zeta = -\nabla_t h(\bar z,\bar t)\tau,
\end{aligned}$$

has a unique stationary point $\tilde\zeta(\tau)$ and unique corresponding multipliers $\tilde\Gamma_I(\tau), \tilde\lambda_R(\tau), \tilde\mu(\tau)$ and the relations

$$\begin{aligned}
z'(\bar t;\tau) &= \tilde\zeta(\tau), \\
\Gamma'_{ki}(\bar t;\tau) &= \begin{cases} \tilde\Gamma_{ki}(\tau) & \text{if } (k,i)\in I, \\ 0 & \text{otherwise,} \end{cases} \\
\lambda'_r(\bar t;\tau) &= \begin{cases} \tilde\lambda_r(\tau) & \text{if } g_r(\bar z,\bar t) = 0, \\ 0 & \text{otherwise,} \end{cases} \\
\mu'(\bar t;\tau) &= \tilde\mu(\tau),
\end{aligned}$$

hold (cf. Ralph and Dempe 1995, Ralph and Scholtes 1997, or Theorem 5.2.1 and Remark 5.2.4 of Scholtes 1994).

The programs of Example 10 satisfy all conditions of the theorem at $\bar t = 0$ except for (ULSC). This shows that an additional condition like (ULSC) has to be imposed in order to guarantee the local uniqueness of the stationary point as t varies. Note that the second problem of Example 10 shows that B-stationarity is not necessarily preserved under perturbations if (ULSC) is replaced by the weaker condition that at least one of the multipliers Γ_{ki} is positive for those k with $F_{ki}(z) = F_{kj}(z) = 0$ for some $i \ne j$.

If the assumptions of Theorem 11 except for (ULSC) hold, then the parametric program $\mathrm{NLP}_I(t)$ with $I = \{(k,i) \mid F_{ki}(\bar z,\bar t) = 0\}$ still has a locally unique stationary point function $z(t)$, and $z(t)$ will be a weakly stationary point of MPCC(t) for the parameter value t; hence the existence of weak stationary points for varying parameter t can still be guaranteed. If $\bar t = 0$ in the programs of Example 10, then this function is $z(t) = 0$ for every t. However, as one can see from the first problem of the example, branching can occur if (ULSC) is violated. The following theorem refines this observation.

THEOREM 12. *Let $\bar z$ be a weakly stationary point of the program MPCC($\bar t$) and let $\bar\Gamma, \bar\lambda, \bar\mu$ be multipliers satisfying* (24). *If all matrices*

$$(27) \qquad \begin{pmatrix} \nabla^2_{zz} L(1,\bar z,\bar\Gamma,\bar\lambda,\bar\mu,\bar t) & -\nabla_z F_K(\bar z,\bar t)^\top & \nabla_z g_R(\bar z,\bar t)^\top & \nabla_z h(\bar z,\bar t)^\top \\ \nabla_z F_K(\bar z,\bar t) & 0 & 0 & 0 \\ -\nabla_z g_R(\bar z,\bar t) & 0 & 0 & 0 \\ \nabla_z h(\bar z,\bar t) & 0 & 0 & 0 \end{pmatrix}$$

with

$$\{(k,i)\,|\,\bar{\Gamma}_{ki}\neq 0\}\subseteq K\subseteq\{(k,i)\,|\,F_{ki}(\bar{z},\bar{t})=0\}\text{ and }\forall k=1,\ldots,m\ \exists i\in\{1,\ldots,l\}\colon\ (k,i)\in K$$

$$\{r\,|\,\bar{\lambda}_r>0\}\subseteq R\subseteq\{r\,|\,g_r(\bar{z},\bar{t})=0\},$$

have the same nonvanishing determinantal sign, then there exist open neighbourhoods Z of $\bar{z}$ and T of $\bar{t}$ and a nonempty Hausdorff continuous multifunction $S: T\to Z$ such that for fixed $t\in T$ the point z is a weakly stationary point of MPCC(t) if and only if $z\in S(t)$. Moreover, the graph of $S(t)$ is the union of the graphs of

$$\prod_{k\in K_{\neq}} 2^{l_k}\prod_{k\in K_0}(2^{l_k}-1)$$

single valued piecewise smooth functions, where $K_{\neq}$ is the set of all indices k such that there exists an index i with $\bar{\Gamma}_{ki}\neq 0$, K_0 is the complement of $K_{\neq}$ in $\{1,\ldots,m\}$, and l_k is the number of indices j such that $\bar{\Gamma}_{kj}=F_{kj}(\bar{z},\bar{t})=0$.

Proof. Note first that the nonvanishing determinantal sign property implies that LICQ holds at $\bar{z}$; hence the multipliers $\bar{\Gamma},\bar{\lambda},\bar{\mu}$ are unique. Define

$$I_F(z,t)=\{(k,i)\,|\,F_{ki}(z,t)=0\},\qquad I_{\neq}=\{(k,i)\,|\,\bar{\Gamma}_{ki}\neq 0\}.$$

We first show that there exist neighbourhoods T of $\bar{t}$ and Z of $\bar{z}$ such that $z\in Z$ is a weakly stationary point of MPCC(t) for a parameter vector $t\in T$ if and only if z is a Karush-Kuhn-Tucker point of a problem of type NLP$_I(t)$ with

$$(28)\qquad I_{\neq}\subseteq I\subseteq I_F(\bar{z},\bar{t})\text{ and }\forall k=1,\ldots,m,\ \exists i\in\{1,\ldots,l\}\colon\ (k,i)\in I.$$

There exist neighbourhoods Z of $\bar{z}$ and T of $\bar{t}$ such that for every $z\in Z$ and $t\in T$ the inclusion $I_F(z,t)\subseteq I_F(\bar{z},\bar{t})$ holds. If, on the one hand, z is a weakly stationary point of MPCC(t) for the parameter vector t, then it is a Karush-Kuhn-Tucker point of problem NLP$_I(t)$, with $I=I_F(z,t)$. If z is sufficiently close to $\bar{z}$ and t is sufficiently close to $\bar{t}$, then, in view of LICQ, the corresponding multipliers Γ_{ki} are close to $\bar{\Gamma}_{ki}$. Hence, reducing Z and T if necessary, we may assume that $I_{\neq}\subseteq I_F(z,t)$ for every $z\in Z$ and $t\in T$. Moreover, for every k there is an index i such that $F_{ki}(z,t)=0$ since z is feasible for MPCC(t). Hence we conclude that z is a Karush-Kuhn-Tucker point of problem NLP$_I(t)$ for some index sets I satisfying (28). To see the converse, suppose z^I is a Karush-Kuhn-Tucker point of problem NLP$_I(t)$. If I satisfies (28) then z is feasible for MPCC(t) since for every k there exists at least one index i with $F_{ki}(z,t)=0$. Moreover, the Karush-Kuhn-Tucker conditions for the smooth program NLP$_I(t)$ imply the weak stationary point conditions of MPCC(t).

In view of Theorem 9, the assumptions imply that program NLP$_I(t)$ admits a unique piecewise smooth stationary point function $z^I(t)$ in a neighbourhood of $(\bar{z},\bar{t})$ and unique corresponding multipliers $\Gamma^I(t),\lambda^I(t),\mu^I(t)$ for every I satisfying (28). Hence if $S(t)$ is the union of the vectors $z^I(t)$ over all index sets I satisfying (28) then there exists a neighbourhood Z of $\bar{z}$ such that for every t sufficiently close to $\bar{t}$, a point $z\in Z$ is a weakly stationary point of MPCC(t) if and only if $z\in S(t)$. The mapping $S(.)$ is the union of finitely many piecewise smooth functions and thus Hausdorff continuous.

Finally, the number of stationary point functions z^I with I satisfying (28) coincides with the number of index sets I satisfying (28) and there are

$$\prod_{k\in K_{\neq}} 2^{l_k}\prod_{k\in K_0}(2^{l_k}-1)$$

such index sets. In fact, all such index sets can be generated in the following way. If $k\in K_{\neq}$ then every subset of the set of all pairs (k,j) with $\bar{\Gamma}_{kj}=F_{kj}(\bar{z},\bar{t})=0$ can be contained in I, giving 2^{l_k} possibilities. If $k \in K_0$ then, in addition, at least one of the index pairs (k,j) has to be contained in I, i.e., the empty set has to be excluded, thus resulting in $2^{l_k} - 1$ possibilities. This can be done independently for every index k. □

5. Conclusion. We have studied mathematical programs with vertical complementarity constraints (MPCC) and derived several stationary point conditions. The relationships between these conditions were discussed and second order optimality conditions were given. The final part of the paper was concerned with the stability of stationary points of MPCCs in the sense of Robinson (1980) and Kojima (1980). Most results extend directly to equality constraints of the form $\min\{F_{k1}(z), \ldots, F_{kl_k}(z)\} = 0$, where the number l_k of selection functions changes with the index k. This observation is worth mentioning since a mere replication of selection functions to achieve the same numbers l_k may violate constraint qualifications. Moreover, some of the results do not make explicit use of the minimum operation and therefore remain valid for more general piecewise smooth equations. In order not to overload the exposition with technicalities, we have confined ourselves to a unified presentation of the results in the framework of mathematical programs with complementarity constraints.

Acknowledgments. This paper is a revised version of the authors' working paper "Mathematical Programs with Equilibrium Constraints: Stationarity, Optimality, and Sensitivity," Judge Institute of Management Studies, University of Cambridge, 1997. Many colleagues have given us advise on this earlier version. We are particularly grateful for Daniel Ralph's detailed comments on a first draft of the manuscript. Thanks are also due to the anonymous referees for valuable suggestions.

References

Burke, J. V. 1991. An exact penalization viewpoint of constrained optimization. *SIAM J. Control and Optim.* **29** 1340–1359.

Clarke, F. H. 1976. A new approach to Lagrange multipliers. *Math. of Oper. Res.* **1** 165–174.

——. 1983. *Optimization and Nonsmooth Analysis.* J. Wiley & Sons, Chichester, U.K.

Clark, P. A., A. W. Westerberg. 1990. Bilevel programming for steady-state chemical process design—I. Fundamentals and algorithms. *Comput. & Chemical Engrg.* **14** 87–97.

Cottle R. W., G. B. Dantzig. 1970. A generalization of the linear complementarity problem. *J. Combinatorial Theory* **8** 79–90.

——, J. S. Pang, R. E. Stone. 1992. *The Linear Complementarity Problem.* Academic Press, Boston, MA.

Facchinei, F., H. Jiang, L. Qi. 1999. A smoothing method for mathematical programs with equilibrium constraints. *Math. Programming*, Forthcoming.

Ferris, M. C., J. S. Pang. 1997. Engineering and economic applications of complementarity problems. *SIAM Rev.* **39** 669–713.

Fiacco, A., G. McCormick. 1968. *Nonlinear programming: Sequential unconstrained minimization techniques.* John Wiley, New York.

Gauvin J. 1977. A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming. *Math. Programming* **12** 136–138.

Jongen H. Th., P. Jonker, F. Twilt. 1986. *Nonlinear Optimization in* $\mathbb{R}^n$ II: *Transversality, Flows, Parametric Aspects.* Peter Lang Verlag, Frankfurt, Germany.

——, G. W. Weber. 1991. Nonlinear optimization: Characterization of structural stability. *J. Global Optim.* **1** 47–64.

Kojima, M. 1980. Strongly stable stationary solutions in nonlinear programming. S. M. Robinson (Ed.) *Analysis and Computation of Fixed Points.* Academic Press, New York, 93–138.

Kuntz, L., S. Scholtes. 1994. A nonsmooth variant of the Mangasarian-Fromovitz constraint qualification. *J. Optim. Theory and Appl.* **82** 59–75.

Kyparisis, J. 1985. On the uniqueness of Kuhn-Tucker multipliers in nonlinear programming. *Math. Programming* **32** 203–213.

Luo Z. Q., J. S. Pang, D. Ralph. 1996. *Mathematical programs with equilibrium constraints.* Cambridge University Press, Cambridge, U.K.

——, ——, ——. 1998. Piecewise sequential quadratic programming for mathematical programs with nonlinear complementarity constraints. A. Migdalas et al. (Eds.), *Multilevel Optimization: Algorithms, Complexity and Applications*. Forthcoming, Kluwer Academic Publishers, Boston, MA.

Outrata, J. V. 1999. Optimality conditions for a class of mathematical programs with equilibrium constraints. *Math. Oper. Res.*, Forthcoming.

——, M. Kocvara, J. Zowe. 1998. *Nonsmooth approach to optimization problems with equilibrium constraints: Theory, applications and numerical results*. Kluwer Academic Publishers, Boston, MA.

Pang, J. S., M. Fukushima. 1998. Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. Working paper, Department of Mathematical Sciences, Whiting School of Engineering, The Johns Hopkins University, Baltimore, MD. 21218-2682.

Polak, E., D. Q. Mayne, Y. Wardi. 1983. On the extension of constrained optimization algorithms from differentiable to nondifferentiable problems. *SIAM J. Control and Optim.* **21** 179–203.

Ralph, D., S. Dempe. 1995. Directional derivatives of solutions of a parametric nonlinear program. *Math. Programming* **70** 159–172.

——, S. Scholtes. 1997. Sensitivity analysis of composite piecewise smooth equations. *Math. Programming* **76** 593–612.

Robinson, S. M. 1976. Stability theory for systems of inequalities, part II: Differentiable nonlinear systems. *SIAM J. Numer. Anal.* **13** 497–513.

——. 1980. Strongly regular generalized equations. *Math. Oper. Res.* **5** 43–62.

——. 1981. Some continuity properties of polyhedral multifunctions. *Math. Programming Stud.* **14** 206–214.

——. 1987. Local structure of feasible sets in nonlinear programming, part III: Stability and sensitivity. *Math. Programming Stud.* **30** 45–66.

Scheel, H. 1995. *Ein Straffunktionsansatz für Optimierungsprobleme mit Gleichgewichtsrestriktionen*. Diploma Thesis, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, 76128 Karlsruhe, Germany.

Scholtes, S. 1994. *Introduction to piecewise differentiable equations*. Habilitation Thesis, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, 76128 Karlsruhe, Germany.

——, M. Stöhr. 1999. Exact penalization of mathematical programs with equilibrium constraints. *SIAM J. Control and Optim.* **37** 617–652.

M. Stöhr, 1999. *Nonsmooth trust region methods and their applications to mathematical programs with equilibrium constraints*. PhD Thesis, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, 76128 Karlsruhe, Germany.

H. Scheel: Universität Dortmund, Operations Research and Wirtschaftsinformatik, 44221 Dortmund, Germany

S. Scholtes: University of Cambridge, Judge Institute of Management Studies and Department of Engineering, Cambridge CB2 1AG, England