manuscript No.
(will be inserted by the editor)

Mordukhovich Stationarity for Mathematical Programs with Switching Constraints under Weak Constraint Qualifications

Gaoxi Li · Lei Guo

Received: date / Accepted: date

Abstract The mathematical program with switching constraints (MPSC), which is recently introduced, is a difficult class of optimization problems since standard constraint qualifications are very likely to fail at local minimizers. MPSC arises from the discretization of optimal control problems with switching constraints which appears frequently in the field of control. Due to the failure of standard constraint qualifications, it is reasonable to propose some constraint qualifications for local minimizers to satisfy some stationarity conditions that are generally weaker than Karush-Kuhn-Tucker stationarity such as Mordukhovich (M-) stationarity. First we propose the weakest constraint qualification for M-stationarity of MPSC to hold at local minimizers. Then we extend some weak verifiable constraint qualifications for nonlinear programming to allow the existence of switching constraints, which are all strictly weaker than MPSC linear independence constraint qualification and/or MPSC Mangasarian-Fromovitz constraint qualification used in the literature. We show that these newly introduced constraint qualifications are suffi-

Gaoxi Li
School of Mathematics and Statistics, Chongqing Technology and Business University
Chongqing, 400067, China.

Lei Guo, Corresponding author
School of Business, East China University of Science and Technology
Shanghai, 200237, China.
lguo@ecust.edu.cn.

cient for local minimizers to be M-stationary. Finally, the relations among MPSC tailored constraint qualifications are discussed.

Keywords Mathematical program with switching constraints · Constraint qualification · Stationarity

Mathematics Subject Classification (2010) 90C26 · 90C30 · 90C46

1 Introduction

Very recently, the mathematical program with switching constraints (MPSC) is introduced to investigate the discretization of optimal control problems with switching constraints (Clason et al., 2016; Gugat, 2010; Hante & Sager, 2013; Wang & Yan, 2016) in Mehlitz (2019b). Moreover, MPSC can be used to reformulate either-or-constrained programming problems (Dempe & Schreier, 2007). The constraint structure of MPSC is highly related with the mathematical program with complementarity constraints (MPEC) that has been widely studied (Lin & Fukushima, 2005; Luo et al., 1996a,b; Outrata, 1999; Shim et al., 2013) and the mathematical program with vanishing constraints (MPVC) (Achtziger & Kanzow, 2008; Izmailov & Solodov, 2009; Hoheisel et al., 2010; Hoheisel, 2009; Mishra et al., 2016). In Mehlitz (2019b), it was shown that the standard constraint qualifications are very likely to fail at local minimizers. This means that local minimizers of MPSC may not be KKT points. Some alternative stationarities such as Mordukhovich (M-) and strong (S-) stationarities are proposed in Mehlitz (2019b). Moreover, the standard linear independence constraint qualification (LICQ), Mangasarian-Fromovitz constraint qualification (MFCQ), Abadie constraint qualification (ACQ), and Guignard constraint qualification (GCQ) are extended to allow the existence of switching constraints. More recently, based on the theoretical findings in Mehlitz (2019b), a relaxation method for solving MPSC was proposed in Kanzow et al. (2018) and second-order optimality conditions for MPSC were investigated in Mehlitz (2019a).

In this paper, we first investigate sufficient and necessary conditions for M-stationarity of MPSC to hold. We then extend some weak verifiable constraint qualifications for nonlinear programming to MPSC such as constant rank constraint qualification (CRC-

Q)(Janin, 1984), relaxed constant rank constraint qualification (RCRCQ) (Minchenko & Stakhovski, 2011), constant positive linear dependent condition (CPLD)(Qi & Wei, 2000), relaxed constant positive linear dependent condition (RCPLD) (Andreani et al., 2012), and quasi-normality and pseudo-normality (Bertsekas & Ozdaglar, 2002). These newly introduced constraint qualifications are called MPSC tailored constraint qualifications, which are all strictly weaker than MPSC LICQ and MPSC MFCQ used in Mehlitz (2019b). We show that all these introduced MPSC tailored constraint qualifications are sufficient for local minimizers of MPSC to be M-stationary. Finally, we discuss the relations among MPSC tailored constraint qualifications.

The paper is organized as follows. Section 2 contains some background materials. In Section 3 we give necessary and sufficient conditions for Bouligand (B-) stationarity and M-stationarity of MPSC, and show that B-stationarity is strictly stronger than M-stationarity. In Section 4, we introduce some new MPSC tailored constraint qualifications and discuss the relations among them.

2 Preliminaries

The notation used in the paper is standard as in the literature. We denote by $\|\cdot\|$ the Euclidean norm and denote by $\mathcal{B}_\delta(x) := \{z \in \mathbb{R}^n : \|z - x\| < \delta\}$ the open ball centered at x with radius $\delta > 0$. For a differentiable function $\phi : \mathbb{R}^n \to \mathbb{R}^m$ and a vector $x \in \mathbb{R}^n$, $\nabla\phi(x)$ denotes the transposed Jacobian of ϕ at x. Moreover, we let $x^k \to_\Omega x^*$ stand for $x^k \in \Omega$ for each k and $x^k \to x^*$ as $k \to \infty$ and let $\mathrm{cl}\Omega$ denote the closure of a set Ω.

In what follows, we review some basic concepts and results in variational analysis, which will be used later on.

Definition 2.1 (Rockafellar & Wets, 1998) The polar cone of a cone K is a closed and convex cone defined by $K^o := \{d : d^\top x \leq 0 \text{ for each } x \in K\}$. The tangent cone of a set Ω at $x^* \in \mathrm{cl}\Omega$ is a closed cone defined by

$$\mathcal{T}_\Omega(x^*) := \{d : d = \lim_{k\to\infty} t_k(x^k - x^*) \text{ with } t_k \geq 0 \text{ and } x^k \to_\Omega x^*\}.$$

The regular normal cone of a set Ω at $x^* \in \mathrm{cl}\Omega$ is a closed and convex cone defined by $\widehat{\mathcal{N}}_\Omega(x^*) := \mathcal{T}_\Omega(x^*)^o$. The limiting normal cone of a set Ω at $x^* \in \mathrm{cl}\Omega$ is a closed cone defined by

$$\mathcal{N}_\Omega(x^*) := \{d : d = \lim_{k\to\infty} d^k \text{ with } d^k \in \widehat{\mathcal{N}}_\Omega(x^k) \text{ and } x^k \to_\Omega x^*\}.$$

The MPSC considered in this paper is of this form

$$\begin{aligned} &\min\ f(x)\\ &\text{s.t. } g_i(x) \le 0,\ \ i = 1, ..., m,\\ &\qquad h_j(x) = 0,\ \ j = 1, ..., p,\\ &\qquad G_t(x)H_t(x) = 0,\ \ t = 1, ..., l, \end{aligned} \tag{1}$$

where all functions $f, g_1, \dots, g_m, h_1, \dots, h_p, G_1, \dots, G_l, H_1, \dots, H_l$: $\mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable. For brevity, we let $g = (g_1, \dots, g_m)^\top$, $h = (h_1, \dots, h_p)^\top$, $G = (G_1, \dots, G_l)^\top$, and $H = (H_1, \dots, H_l)^\top$. Let $\mathcal{X}$ denote the feasible region of problem (1). The last l constraints in $\mathcal{X}$ force either $G_t(x)$ or $H_t(x)$ to be zero. We call $x \in \mathcal{X}$ a degenerate feasible point if there exists t_0 such that $G_{t_0}(x) = H_{t_0}(x) = 0$. For such degenerate feasible points, the standard LICQ and MFCQ fail (Mehlitz, 2019b, Lemma 4.1). Moreover, the tangent cone to $\mathcal{X}$ at a degenerate feasible point is generally nonconvex but the linearized tangent cone is convex. Thus, the standard ACQ also fails at a degenerate feasible point; see details in Section 3. These facts make problem (1) different from standard nonlinear programming problems.

In order to facilitate the notation, we define some index sets which depend on a feasible point $x^* \in \mathcal{X}$:

$$\begin{cases} \mathcal{I}^h := \{1, \dots, p\},\ \mathcal{I}^g_* := \{i \in \{1, \dots, m\} : g_i(x^*) = 0\},\\ \mathcal{I}^G_* := \{t \in \{1, \dots, l\} : G_t(x^*) = 0 \wedge H_t(x^*) \neq 0\},\\ \mathcal{I}^H_* := \{t \in \{1, \dots, l\} : G_t(x^*) \neq 0 \wedge H_t(x^*) = 0\},\\ \mathcal{I}^{GH}_* := \{t \in \{1, \dots, l\} : G_t(x^*) = 0 \wedge H_t(x^*) = 0\}. \end{cases}$$

Note that $\{\mathcal{I}^G_*, \mathcal{I}^H_*, \mathcal{I}^{GH}_*\}$ is a disjoint partition of $\{1, \dots, l\}$. The MPSC Lagrangian function of problem (1) is defined by

$$L_{MPSC}(x, \lambda, \rho, \mu, \nu) := f(x) + g(x)^\top\lambda + h(x)^\top\rho + G(x)^\top\mu + H(x)^\top\nu.$$

We note that problem (1) can be rewritten as an optimization problem with a geometric constraint:

$$\min f(x) \quad \text{s.t.} \quad F(x) \in \Lambda, \tag{2}$$

where

$$F(x) := (g(x), h(x), \psi(x))^\top, \quad \Lambda := (-\infty, 0]^m \times \{0\}^p \times \mathcal{S}^l,$$

and

$$\psi(x) := (G_1(x), H_1(x), \dots, G_l(x), H_l(x))^\top, \quad \mathcal{S} := \{(a,b) \in \mathbb{R}^2 : ab = 0\}.$$

We call the nonconvex cone $\mathcal{S}$ the switching cone. By direct calculation, we have the following results (Mehlitz, 2019b).

Proposition 2.1 *For any $(a,b) \in \mathcal{S}$, we have that*

$$\mathcal{T}_{\mathcal{S}}(a,b) := \left\{ \begin{array}{ll} \{0\} \times \mathbb{R} & \text{if } a = 0, b \neq 0 \\ \mathbb{R} \times \{0\} & \text{if } a \neq 0, b = 0 \\ \mathcal{S} & \text{if } a = 0, b = 0 \end{array} \right\}, \quad \widehat{\mathcal{N}}_{\mathcal{S}}(a,b) := \left\{ \begin{array}{ll} \mathbb{R} \times \{0\} & \text{if } a = 0, b \neq 0 \\ \{0\} \times \mathbb{R} & \text{if } a \neq 0, b = 0 \\ \{(0,0)\} & \text{if } a = 0, b = 0 \end{array} \right\},$$

$$\mathcal{N}_{\mathcal{S}}(a,b) := \left\{ \begin{array}{ll} \mathbb{R} \times \{0\} & \text{if } a = 0, b \neq 0 \\ \{0\} \times \mathbb{R} & \text{if } a \neq 0, b = 0 \\ \mathcal{S} & \text{if } a = 0, b = 0 \end{array} \right\}. \tag{3}$$

The linearized cone of $\mathcal{X}$ at x^* is defined by

$$\mathcal{L}_{MPSC}(x^*) := \{d : \nabla F(x^*)^\top d \in \mathcal{T}_\Lambda(F(x^*))\}.$$

By Proposition 2.1, the linearized cone $\mathcal{L}_{MPSC}(x^*)$ can be directly calculated as an explicit form

$$\mathcal{L}_{MPSC}(x^*) = \left\{ d : \begin{array}{ll} \nabla g_i(x^*)^\top d \leq 0 & i \in \mathcal{I}_*^g \\ \nabla h_i(x^*)^\top d = 0 & i \in \mathcal{I}^h \\ \nabla G_i(x^*)^\top d = 0 & i \in \mathcal{I}_*^G \\ \nabla H_i(x^*)^\top d = 0 & i \in \mathcal{I}_*^H \\ \nabla G_i(x^*)^\top d \cdot \nabla H_i(x^*)^\top d = 0 & i \in \mathcal{I}_*^{GH} \end{array} \right\}.$$

Let $\mathcal{P}(\mathcal{I}_*^{GH})$ be the set of all disjoint bipartitions of $\mathcal{I}_*^{GH}$, i.e.,

$$\mathcal{P}(\mathcal{I}_*^{GH}) := \{(\beta_1, \beta_2) : \beta_1 \cup \beta_2 = \mathcal{I}_*^{GH},\ \beta_1 \cap \beta_2 = \emptyset\}.$$

For any given $(\beta_1, \beta_2) \in \mathcal{P}(\mathcal{I}_*^{GH})$, we define a standard nonlinear programming problem as follows

$$
\text{NLP}(\beta_1,\beta_2) \qquad \begin{aligned} \min\ & f(x) \\ \text{s.t.}\ & g_i(x) \le 0, \ \ i = 1, ..., m, \\ & h_j(x) = 0, \ \ j = 1, ..., p, \\ & G_k(x) = 0, \ \ k \in \mathcal{I}_*^G \cup \beta_1, \\ & H_k(x) = 0, \ \ k \in \mathcal{I}_*^H \cup \beta_2. \end{aligned} \tag{4}
$$

We let $\mathcal{X}_{(\beta_1,\beta_2)}$ denote the feasible region of the above problem. We say that $\text{NLP}(\beta_1, \beta_2)$ is a branch of problem (1) since it is easy to verify that locally around x^*, the union of $\mathcal{X}_{(\beta_1,\beta_2)}$ over $(\beta_1, \beta_2) \in \mathcal{P}(\mathcal{I}_*^{GH})$ is exactly the feasible set $\mathcal{X}$. Then one easily has the following results(Mehlitz, 2019b, Lemma 5.1).

Lemma 2.1 *Let $x^* \in \mathcal{X}$. Then the following formulas hold true:*

$$
\mathcal{T}_{\mathcal{X}}(x^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\mathcal{I}_*^{GH})} \mathcal{T}_{\mathcal{X}(\beta_1,\beta_2)}(x^*),
$$

$$
\mathcal{L}_{MPSC}(x^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\mathcal{I}_*^{GH})} \mathcal{L}_{\mathcal{X}(\beta_1,\beta_2)}(x^*),
$$

where $\mathcal{L}_{\mathcal{X}(\beta_1,\beta_2)}(x^*)$ is the linearized cone of problem (4) *defined by*

$$
\mathcal{L}_{\mathcal{X}(\beta_1,\beta_2)}(x^*) := \left\{ d : \begin{array}{l} \nabla g_i(x^*)^\top d \le 0 \ \ i \in \mathcal{I}_*^g \\ \nabla h_i(x^*)^\top d = 0 \ \ i \in \mathcal{I}^h \\ \nabla G_i(x^*)^\top d = 0 \ i \in \mathcal{I}_*^G \cup \beta_1 \\ \nabla H_i(x^*)^\top d = 0 \ i \in \mathcal{I}_*^H \cup \beta_2 \end{array} \right\}.
$$

3 Stationarities

In this section, we first define the prime and dual stationarities of problem (1). We then investigate sufficient and necessary conditions for stationarities of problem (1).

Definition 3.1 Let $x^* \in \mathcal{X}$. We say that x^* is a Bouligand (B-) stationary point of problem (1) iff

$$
\nabla f(x^*)^\top d \ge 0, \ \ \forall d \in \mathcal{L}_{MPSC}(x^*).
$$

We next give some dual stationarities of problem (1) as given in Mehlitz (2019b).

Definition 3.2 Let $x^* \in \mathcal{X}$. We say that x^* is weakly stationary (W-stationary) to problem (1) iff there exists $(\lambda, \rho, \mu, \nu) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ such that

$$\begin{cases} \nabla_x L_{MPSC}(x^*, \lambda, \rho, \mu, \nu) = 0, \\ \lambda \geq 0,\ g(x^*)^\top \lambda = 0,\ \mu_i = 0,\ i \in \mathcal{I}_*^H,\ \nu_i = 0,\ i \in \mathcal{I}_*^G. \end{cases} \tag{5}$$

We say that x^* is M-stationary to problem (1) iff there exists $(\lambda, \rho, \mu, \nu) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ satisfying (5) and

$$\mu_i \nu_i = 0,\ i \in \mathcal{I}_*^{GH}. \tag{6}$$

We say that x^* is S-stationary to (1) iff there exists $(\lambda, \rho, \mu, \nu) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ satisfying (5) and

$$\mu_i = 0 \text{ and } \nu_i = 0,\ i \in \mathcal{I}_*^{GH}. \tag{7}$$

By Proposition 2.1, we can restate the B-, M-, and S-stationarities in the following compact forms.

Proposition 3.1 *Let $x^* \in \mathcal{X}$. We have the following results.*

(i) The B-stationarity is equivalent to $0 \in \nabla f(x^*) + \mathcal{L}_{MPSC}(x^*)^o$.

(ii) The M-stationarity is equivalent to $0 \in \nabla f(x^*) + \nabla F(x^*) \mathcal{N}_\Lambda(F(x^*))$.

(iii) The S-stationarity is equivalent to $0 \in \nabla f(x^*) + \nabla F(x^*) \widehat{\mathcal{N}}_\Lambda(F(x^*))$.

3.1 Necessary and sufficient conditions for B-/M-stationarity

In this subsection, we investigate necessary and sufficient conditions for B-stationarity and M-stationarity, respectively. We point out that although Theorems 3.1 and 3.2 in Guo & Lin (2013) are obtained for MPEC, their proofs are not dependent of the characterization of $\mathcal{X}$. Thus, Theorems 3.1 and 3.2 in Guo & Lin (2013) are valid for any set $\mathcal{X}$. Then, we immediately have the following two results.

Proposition 3.2 *If $x^* \in \mathcal{X}$ is a local minimizer of $\min_{x \in \mathcal{X}} \theta(x)$, where θ is a smooth function, and the following MPSC GCQ holds:*

$$\mathcal{T}_\mathcal{X}(x^*)^o = \mathcal{L}_{MPSC}(x^*)^o,$$

then x^* must be a B-stationary point. Conversely, if $x^* \in \mathcal{X}$ is B-stationary to $\min_{x\in\mathcal{X}} \theta(x)$ for any smooth function θ with x^* being a locally optimal solution, then MPSC GCQ holds at x^*.

Proposition 3.2 implies that MPSC GCQ is the weakest constraint qualification for B-stationarity of MPSC to hold at local minimizers.

Proposition 3.3 *Suppose that $x^* \in \mathcal{X}$ is a local minimizer for $\min_{x\in\mathcal{X}} \theta(x)$, where θ is a smooth function, and*

$$\mathcal{T}_{\mathcal{X}}(x^*)^o \subseteq \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*)). \tag{8}$$

Then x^* must be an M-stationary point. Conversely, if $x^* \in \mathcal{X}$ is M-stationary to $\min_{x\in\mathcal{X}} \theta(x)$ for any smooth function θ with x^* being a locally optimal solution, then (8) holds.

From Proposition 3.3, it is easy to see that (8) is the weakest constraint qualification for M-stationarity of MPSC to hold at local minimizers. For simplicity, we call condition (8) MPSC MCQ.

We next investigate the relation between B-stationarity and M-stationarity, i.e., the relation between MPSC GCQ and MPSC MCQ.

Theorem 3.1 *B-stationarity (resp. MPSC GCQ) implies M-stationarity (resp. MPSC MCQ).*

Proof For any $d \in \mathcal{T}_{\mathcal{X}}(x^*)^o$, by Rockafellar & Wets (1998, Theorem 6.11), there exists a smooth function φ such that $-\nabla\varphi(x^*) = d$ and $\arg\min_{x\in\mathcal{X}} \varphi(x) = \{x^*\}$. Since MPSC GCQ is valid at x^*, we have $-\nabla\varphi(x^*) \in \mathcal{L}_{MPSC}(x^*)^o$, i.e.,

$$-\nabla\varphi(x^*)^\top p \leq 0, \ \ \forall p \in \mathcal{L}_{MPSC}(x^*). \tag{9}$$

We next show that $d \in \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*))$. For any partition $(\beta_1, \beta_2) \in \mathcal{P}(\mathcal{I}_*^{GH})$, from Lemma 2.1, we have $\mathcal{L}_{\mathcal{X}(\beta_1,\beta_2)}(x^*) \subseteq \mathcal{L}_{MPSC}(x^*)$. Combining this with (9) implies that

$$\nabla\varphi(x^*)^\top p \geq 0, \ \ \forall p \in \mathcal{L}_{\mathcal{X}(\beta_1,\beta_2)}(x^*). \tag{10}$$

This means that $p = 0$ is a minimizer of the linear programming problem

$$\begin{aligned} \min\nolimits_p \ & \nabla\varphi(x^*)^\top p \\ s.t. \ & \nabla g_i(x^*)^\top p \leq 0, \ \ i \in \mathcal{I}_*^g, \\ & \nabla h_i(x^*)^\top p = 0, \ \ i \in \mathcal{I}^h, \\ & \nabla G_i(x^*)^\top p = 0, \ \ i \in \mathcal{I}_*^G \cup \beta_1, \\ & \nabla H_i(x^*)^\top p = 0, \ \ i \in \mathcal{I}_*^H \cup \beta_2. \end{aligned} \tag{11}$$

Then there exist $\lambda_i \geq 0$ $(i \in \mathcal{I}_*^g)$, ρ_i $(i \in \mathcal{I}^h)$, η_i $(i \in \mathcal{I}_*^G \cup \beta_1)$, θ_i $(i \in \mathcal{I}_*^H \cup \beta_2)$ such that

$$\nabla\varphi(x^*) + \sum_{i\in\mathcal{I}_*^g} \lambda_i \nabla g_i(x^*) + \sum_{i\in\mathcal{I}^h} \rho_i \nabla h_i(x^*) + \sum_{i\in\mathcal{I}_*^G\cup\beta_1} \eta_i \nabla G_i(x^*) + \sum_{i\in\mathcal{I}_*^H\cup\beta_2} \theta_i \nabla H_i(x^*) = 0. \tag{12}$$

By letting

$$\mu_i := \begin{cases} \eta_i & i \in \mathcal{I}_*^G \cup \beta_1, \\ 0 & i \in \beta_2, \end{cases} \qquad \nu_i := \begin{cases} 0 & i \in \beta_1, \\ \theta_i & i \in \mathcal{I}_*^H \cup \beta_2, \end{cases}$$

we can rewrite (12) as

$$-\nabla\varphi(x^*) = \sum_{i\in\mathcal{I}_*^g} \lambda_i \nabla g_i(x^*) + \sum_{i\in\mathcal{I}^h} \rho_i \nabla h_i(x^*) + \sum_{i\in\mathcal{I}_*^G\cup\mathcal{I}_*^{GH}} \mu_i \nabla G_i(x^*) + \sum_{i\in\mathcal{I}_*^H\cup\mathcal{I}_*^{GH}} \nu_i \nabla H_i(x^*), \tag{13}$$

satisfying $\mu_i\nu_i = 0$ for each $i \in \mathcal{I}_*^{GH}$. By the explicit expression of $\mathcal{N}_\Lambda(F(x^*))$, we have that $d = -\nabla\varphi(x^*) \in \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*))$. $\square$

For standard nonlinear programs with equality and inequality constraints, both B-stationarity and M-stationarity are KKT stationarity. The following example illustrates that B-stationarity is strictly stronger than M-stationarity.

Example 3.1 Consider an MPSC problem

$$\min\ x_1 + x_2 \ \text{ s.t. }\ x_2 - x_1^2 = 0,\ \ x_1x_2 = 0.$$

It is easy to see that the unique feasible point $x^* = (0, 0)$ is the unique optimal solution. By direct calculation, we have that

$$\mathcal{L}_{MPSC}(x^*) = \left\{ (d_1, d_2)^\top : \begin{array}{c} d_2 = 0 \\ d_1 \in \mathbb{R} \end{array} \right\} \text{ and } \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*)) = \mathbb{R}^2.$$

Thus, x^* is an M-stationary point but it is not a B-stationary point since

$$(1, 1) \cdot (0, -1)^\top < 0, \text{ and } (0, -1)^\top \in \mathcal{L}_{MPSC}(x^*).$$

4 Constraint Qualifications

We first review the existing MPSC tailored constraint qualifications for MPSC which are all given in Mehlitz (2019b). Then we extend weak verifiable constraint qualifications for nonlinear programming to MPSC.

Definition 4.1 Let $x^* \in \mathcal{X}$. (a) We say that MPSC LICQ holds at x^* iff the family of gradients

$$\left\{ \nabla g_i(x^*), \nabla h_j(x^*), \nabla G_r(x^*), \nabla H_t(x^*) : \begin{array}{c} i \in \mathcal{I}^g_*, r \in \mathcal{I}^G_* \cup \mathcal{I}^{GH}_* \\ j \in \mathcal{I}^h, t \in \mathcal{I}^H_* \cup \mathcal{I}^{GH}_* \end{array} \right\}$$

is linearly independent.

(b) We say that MPSC MFCQ holds at x^* iff there is no nonzero $\{\lambda, \rho, \mu, \nu\}$ such that

$$\begin{cases} \nabla_x L_{MPSC}(x^*, \lambda, \rho, \mu, \nu) = 0, \\ \lambda \geq 0, g(x^*)^\top \lambda = 0, \\ \mu_i = 0, i \in \mathcal{I}^H_*, \ \ \nu_i = 0, i \in \mathcal{I}^G_*. \end{cases}$$

(c) We say that MPSC NNAMCQ holds at x^* iff there is no nonzero $\{\lambda, \rho, \mu, \nu\}$ such that

$$\begin{cases} \nabla_x L_{MPSC}(x^*, \lambda, \rho, \mu, \nu) = 0, \\ \lambda \geq 0, g(x^*)^\top \lambda = 0, \\ \mu_i = 0, i \in \mathcal{I}^H_*, \ \ \nu_i = 0, i \in \mathcal{I}^G_*, \\ \mu_i \nu_i = 0, i \in I^{GH}_*. \end{cases}$$

(d) We say that MPSC ACQ holds at x^* iff $\mathcal{T}_{\mathcal{X}}(x^*) = \mathcal{L}_{MPSC}(x^*)$.

It is not hard to derive the following relationships:

$$\begin{aligned} \text{MPSC LICQ} \Rightarrow \text{MPSC MFCQ} \Rightarrow \text{MPSC NNAMCQ} \\ \Rightarrow \text{MPSC ACQ} \Rightarrow \text{MPSC GCQ}. \end{aligned}$$

Moreover, in Mehlitz (2019b), it has been shown that all these constraint qualifications are sufficient for M-stationarity of MPSC to hold at local minimizers. From our result Theorem 3.2, one can easily have that all these constraint qualifications are even sufficient for B-stationarity (better than M-stationarity) of MPSC to hold at local minimizers.

4.1 New verifiable constraint qualifications for M-stationarity

For standard nonlinear programming, the verifiable classical constraint qualifications in the literature are LICQ, Slater's CQ, and MFCQ. Moreover, it is well known that if all constraint functions are affine, then the KKT condition holds at local minimizers without requiring constraint qualifications. Quite a few weaker verifiable constraint qualifications have been introduced in the literature. Some relaxed versions of LICQ have been introduced such as CRCQ and RCRCQ. Some relaxed versions of MFCQ have also proposed such as CPLD, RCPLD, pseudo-normality, and quasi-normality. All these mentioned constraint qualifications have been extended to allow the existence of complementarity constraints, ensuring M-stationarity of MPEC holds at local minimizers (Ye & Zhang, 2014; Kanzow & Schwartz, 2010; Kanzow et al., 2018; Guo & Lin, 2013; Guo et al., 2013a). In the following, we extend the weak verifiable constraint qualifications for nonlinear programming to allow the existence of switching constraints.

Definition 4.2 Let $x^* \in \mathcal{X}$. (a) We say that MPSC CRCQ holds at x^* iff there exists $\delta > 0$ such that for any $I_1 \subseteq \mathcal{I}_*^g$, $I_2 \subseteq \mathcal{I}^h$, $I_3 \subseteq \mathcal{I}_*^G \cup I_*^{GH}$, and $I_4 \subseteq I_*^H \cup \mathcal{I}_*^{GH}$, the family of gradients

$$\{\nabla g_i(x), \nabla h_j(x), \nabla G_r(x), \nabla H_t(x) : i \in I_1, j \in I_2, r \in I_3, t \in I_4\}$$

has the same rank for each $x \in \mathcal{B}_\delta(x^*)$.

(b) We say that MPSC RCRCQ holds at x^* iff there exists $\delta > 0$ such that for any $I_1 \subseteq \mathcal{I}_*^g$ and $I_3, I_4 \subseteq \mathcal{I}_*^{GH}$, the family of gradients

$$\{\nabla g_i(x), \nabla h_j(x), \nabla G_r(x), \nabla H_t(x) : i \in I_1, j \in \mathcal{I}^h, r \in I_3 \cup \mathcal{I}_*^G, t \in I_4 \cup \mathcal{I}_*^H\}$$

has the same rank for each $x \in \mathcal{B}_\delta(x^*)$.

(c) We say that MPSC CPLD holds at x^* iff for any $I_1 \subseteq \mathcal{I}_*^g$, $I_2 \subseteq \mathcal{I}^h$, $I_3 \subseteq \mathcal{I}_*^G \cup \mathcal{I}_*^{GH}$, and $I_4 \subseteq \mathcal{I}_*^H \cup \mathcal{I}_*^{GH}$, whenever there exist $\{\lambda, \rho, \mu, \nu\}$ not all zero, with $\lambda_i \geq 0$ for each $i \in I_1$ and $\mu_i \nu_i = 0$ for each $i \in \mathcal{I}_*^{GH}$, such that

$$\sum_{i \in I_1} \lambda_i \nabla g_i(x^*) + \sum_{j \in I_2} \rho_j \nabla h_j(x^*) + \sum_{r \in I_3} \mu_r \nabla G_r(x^*) + \sum_{t \in I_4} \nu_t \nabla H_t(x^*) = 0, \tag{14}$$

there exists $\delta > 0$ such that, for any $x \in \mathcal{B}_\delta(x^*)$, the vectors $\{\nabla g_i(x) : i \in I_1\}$, $\{\nabla h_j(x) : j \in I_2\}$, $\{\nabla G_r(x) : r \in I_3\}$, $\{\nabla H_t(x) : t \in I_4\}$ are linearly dependent.

(d) Let $x^* \in \mathcal{X}$ and $I_1 \subseteq \mathcal{I}^h$, $I_2 \subseteq \mathcal{I}_*^G$, $I_3 \subseteq \mathcal{I}_*^H$ be index sets such that $\mathcal{G}(x^*; I_1, I_2, I_3)$ is a basis for span $\mathcal{G}(x^*; \mathcal{I}^h, \mathcal{I}_*^G, \mathcal{I}_*^H)$. We say that MPSC RCPLD holds at x^* iff there exists $\delta > 0$ such that

– $\mathcal{G}(x; \mathcal{I}^h, \mathcal{I}_*^G, \mathcal{I}_*^H)$ has the same rank for each $x \in \mathcal{B}_\delta(x^*)$;
– for each $I_4 \subseteq \mathcal{I}_*^g$, $I_5, I_6 \subseteq \mathcal{I}_*^{GH}$, if there exist $\{\lambda, \rho, \mu, \nu\}$ not all zero, with $\lambda_i \geq 0$ for each $i \in I_4$ and $\mu_i \nu_i = 0$ for each $i \in \mathcal{I}_*^{GH}$, such that

$$\sum_{i \in I_4} \lambda_i \nabla g_i(x^*) + \sum_{j \in I_1} \rho_j \nabla h_j(x^*) + \sum_{r \in I_2 \cup I_5} \mu_r \nabla G_r(x^*) + \sum_{t \in I_3 \cup I_6} \nu_t \nabla H_t(x^*) = 0, \tag{15}$$

then for any $x \in \mathcal{B}_\delta(x^*)$, the vectors $\{\nabla g_i(x) : i \in I_4\}$, $\{\nabla h_j(x) : j \in I_1\}$, $\{\nabla G_r(x) : r \in I_2 \cup I_5\}$, $\{\nabla H_t(x) : t \in I_3 \cup I_6\}$ are linearly dependent.

Here $\mathcal{G}(x; I_1, I_2, I_3)$ is a set of gradients defined by

$$\mathcal{G}(x; I_1, I_2, I_3) := \{\nabla h_j(x), \nabla G_r(x), \nabla H_t(x) : j \in I_1, r \in I_2, t \in I_3\}.$$

(e) We say that MPSC pseudo-normality hold at x^* iff there is no nonzero $\{\lambda, \rho, \mu, \nu\}$ such that

– $\sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \rho_i \nabla h_i(x^*) + \sum_{i=1}^l \mu_i \nabla G_i(x^*) + \sum_{i=1}^l \nu_i \nabla H_i(x^*) = 0$;
– $\lambda_i \geq 0$ for $i \in \mathcal{I}_*^g$, $\lambda_i = 0$ for $i \notin \mathcal{I}_*^g$, $\mu_i = 0$ for $i \in \mathcal{I}_*^H$, $\nu_i = 0$ for $i \in \mathcal{I}_*^G$, and $\mu_i \nu_i = 0$ for $i \in \mathcal{I}_*^{GH}$;
– there exists a sequence $\{x^k\} \to x^*$ such that for each k,

$$\sum_{i=1}^m \lambda_i g_i(x^k) + \sum_{i=1}^p \rho_i h_i(x^k) + \sum_{i=1}^l \mu_i G_i(x^k) + \sum_{i=1}^l \nu_i H_i(x^k) > 0.$$

(f) We say that MPEC quasi-normality holds at x^* iff there is no nonzero $\{\lambda, \rho, \mu, \nu\}$ such that

– $\sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \rho_i \nabla h_i(x^*) + \sum_{i=1}^l \mu_i \nabla G_i(x^*) + \sum_{i=1}^l \nu_i \nabla H_i(x^*) = 0$;
– $\lambda_i \geq 0$ for $i \in \mathcal{I}_*^g$, $\lambda_i = 0$ for $i \notin \mathcal{I}_*^g$, $\mu_i = 0$ for $i \in \mathcal{I}_*^H$, $\nu_i = 0$ for $i \in \mathcal{I}_*^G$, and $\mu_i \nu_i = 0$ for $i \in \mathcal{I}_*^{GH}$;

– there exists a sequence $\{x^k\} \to x^*$ such that for each k,

$$\lambda_i > 0 \Rightarrow \lambda_i g_i(x^k) > 0, \quad \mu_i \neq 0 \Rightarrow \mu_i G_i(x^k) > 0,$$
$$\rho_i \neq 0 \Rightarrow \rho_i h_i(x^k) > 0, \quad \nu_i \neq 0 \Rightarrow \nu_i H_i(x^k) > 0.$$

Following from Definitions 4.1 and 4.2, it is easy to see that MPSC NNAMCQ implies MPSC CPLD and MPSC pseudo-normality, MPSC CRCQ implies MPSC RCRCQ, MPSC CPLD implies MPSC RCPLD, and MPSC pseudo-normality implies MPSC quasi-normality.

In what follows, we discuss the relations among these MPSC tailored constraint qualifications mentioned in Section 4.

Theorem 4.1 *Suppose that MPSC RCRCQ holds at $x^* \in \mathcal{X}$. Then MPSC ACQ is also valid at x^*.*

Proof For any given partition $(\beta_1, \beta_2) \in \mathcal{P}(\mathcal{I}_*^{GH})$, consider $\mathrm{NLP}(\beta_1, \beta_2)$ as given in (4). Since MPSC RCRCQ holds at x^*, it is easy to see that RCRCQ holds at x^* for $\mathrm{NLP}(\beta_1, \beta_2)$. It then follows from Minchenko & Stakhovski (2011, Lemma 6) that

$$\mathcal{T}_{\mathcal{X}(\beta_1,\beta_2)}(x^*) = \mathcal{L}_{\mathcal{X}(\beta_1,\beta_2)}(x^*).$$

Then by the arbitrariness of $(\beta_1, \beta_2) \in \mathcal{P}(\mathcal{I}_*^{GH})$ and Lemma 2.1, one can easily obtain that $\mathcal{T}_{\mathcal{X}}(x^*) = \mathcal{L}_{MPSC}(x^*)$, i.e., MPSC Abadie CQ holds at x^*. □

Theorem 4.2 *MPSC CRCQ (resp. MPSC RCRCQ) implies MPSC CPLD (resp. MPSC RC-PLD).*

Proof (i) Let MPSC CRCQ hold at x^*. We will show that MPSC CPLD holds at x^*. Let $I_1 \subseteq \mathcal{I}_*^g$, $I_2 \subseteq \mathcal{I}^h$, $I_3 \subseteq \mathcal{I}_*^G \cup \mathcal{I}_*^{GH}$, and $I_4 \subseteq \mathcal{I}_*^H \cup \mathcal{I}_*^{GH}$. Assume that there exist $\{\lambda, \rho, \mu, \nu\}$ not all zero, with $\lambda_i \geq 0$ for each $i \in I_1$ and $\mu_i \nu_i = 0$ for each $i \in \mathcal{I}_*^{GH}$, such that

$$\sum_{i \in I_1} \lambda_i \nabla g_i(x^*) + \sum_{j \in I_2} \rho_j \nabla h_j(x^*) + \sum_{r \in I_3} \mu_r \nabla G_r(x^*) + \sum_{t \in I_4} \nu_t \nabla H_t(x^*) = 0.$$

This means that the vectors $\{\nabla g_i(x^*), \nabla h_j(x^*), \nabla G_r(x^*), \nabla H_t(x^*) : i \in I_1, j \in I_2, r \in I_3, t \in I_4\}$ are linearly dependent. Since MPSC CRCQ holds at $x^* \in \mathcal{X}$, it is easy to verify that there exists $\delta > 0$ such that the vectors

$$\{\nabla g_i(x), \nabla h_j(x), \nabla G_r(x), \nabla H_t(x) : i \in I_1, j \in I_2, r \in I_3, t \in I_4\}$$

are linearly dependent for each $x \in \mathcal{B}_\delta(x^*)$. Otherwise we can find a sequence $x^k \to x^*$ such that the vectors $\{\nabla g_i(x), \nabla h_j(x), \nabla G_r(x), \nabla H_t(x) : i \in I_1, j \in I_2, r \in I_3, t \in I_4\}$ has bigger rank at x^k than at x^*, contracting MPSC CRCQ. Thus, MPSC CPLD holds at x^*.

(ii) Let MPSC RCRCQ holds at x^*. We will show that MPSC RCPLD holds at x^*. The first claim in MPSC RCPLD follows by setting $I_1 = I_3 = I_4 = \emptyset$ in the definition of MPSC RCRCQ.

Let $I_1 \subseteq \mathcal{I}^h$, $I_2 \subseteq \mathcal{I}_*^G$, $I_3 \subseteq \mathcal{I}_*^H$ be index sets such that $\mathcal{G}(x^*, I_1, I_2, I_3)$ is a basis for span $\mathcal{G}(x^*, \mathcal{I}^h, \mathcal{I}_*^G, \mathcal{I}_*^H)$. Let $I_4 \subseteq \mathcal{I}_*^g$, $I_5, I_6 \subseteq \mathcal{I}_*^{GH}$ be such that there exist $\{\lambda, \rho, \mu, \nu\}$ not all zero, with $\lambda_i \geq 0$ for each $i \in I_4$ and $\mu_i \nu_i = 0$ for each $i \in \mathcal{I}_*^{GH}$, such that

$$\sum_{i \in I_4} \lambda_i \nabla g_i(x^*) + \sum_{j \in I_1} \rho_j \nabla h_j(x^*) + \sum_{r \in I_2 \cup I_5} \mu_r \nabla G_r(x^*) + \sum_{t \in I_3 \cup I_6} \nu_t \nabla H_t(x^*) = 0.$$

This means that the vectors

$$\{\nabla g_i(x), \nabla h_j(x), \nabla G_r(x), \nabla H_t(x) : i \in I_4, j \in I_1, r \in I_5 \cup I_2, t \in I_6 \cup I_3\} \tag{16}$$

are linearly dependent at x^*. It is easy to see that the set $\mathcal{G}(x^*; I_1, I_2, I_3)$ can generate the set $\mathcal{G}(x^*; \mathcal{I}^h, \mathcal{I}_*^G, \mathcal{I}_*^H)$ in a neighborhood of x^*. Thus, using MPSC RCRCQ at x^*, we can have that the set of gradients in (16) has the same rank in some neighborhood $\mathcal{B}_\delta(x^*)$ of x^*. Then this set must be linearly dependent for any $x \in \mathcal{B}_\delta(x^*)$. The proof is complete.□

We next consider the relation between MPSC quasi-normality and MPSC MCQ. To this end, we give a Fritz John type stationarity for problem (1) whose proof scheme follows from Bertsekas & Ozdaglar (2002, Proposition 1), Kanzow & Schwartz (2010, Theorem 3.1), or Guo et al. (2013b, Theorem 3.1).

Lemma 4.1 *Let x^* be a local minimizer of problem* (1). *Then there exist vectors $\{\alpha, \lambda, \rho, \mu, \nu\}$ such that*

(i) $\alpha \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \rho_i \nabla h_i(x^*) + \sum_{i=1}^{l} \mu_i \nabla G_i(x^*) + \sum_{i=1}^{l} \nu_i \nabla H_i(x^*) = 0;$

(ii) $\alpha \geq 0$, $\lambda_i \geq 0$ *for* $i \in \mathcal{I}_*^g$, $\lambda_i = 0$ *for* $i \notin \mathcal{I}_*^g$, $\mu_i = 0$ *for* $i \in \mathcal{I}_*^H$, $\nu_i = 0$ *for* $i \in \mathcal{I}_*^G$, *and* $\mu_i \nu_i = 0$ *for* $i \in \mathcal{I}_*^{GH}$*;*

(iii) $\alpha, \lambda, \rho, \mu, \nu$ *are not all equal to zero;*

(iv) if λ, ρ, μ, ν are not all equal to zero, then there is a sequence $\{x^k\} \to x^*$ such that for all $k \in \mathbb{N}$,

$$\lambda_i > 0 \Rightarrow g_i(x^k) > 0, \qquad \mu_i \neq 0 \Rightarrow \mu_i G_i(x^k) > 0,$$
$$\rho_i \neq 0 \Rightarrow \rho_i h_i(x^k) > 0, \quad \nu_i \neq 0 \Rightarrow \nu_i H_i(x^k) > 0.$$

Proof Let x^* be a local minimizer of problem (1), $y^* := G(x^*)$, and $v^* := H(x^*)$. For each $k \in \mathbb{N}$, we consider the following optimization problem

$$\begin{aligned}\min F_k(x,y,z) &:= f(x) + \frac{k}{2}\|\max\{0,g(x)\}\|^2 + \frac{k}{2}\|h(x)\|^2 + \frac{k}{2}\|G(x)-y\|^2 \\ &+ \frac{k}{2}\|H(x)-z\|^2 + \frac{1}{2}\|(x,y,z)-(x^*,y^*,z^*)\|^2 \quad \text{s.t. } (x,y,z) \in \Omega \cap C.\end{aligned}$$

where $C := \mathbb{R}^n \times \mathcal{S}^l$ and $\Omega := \{(x,y,z) : \|(x,y,z)-(x^*,y^*,z^*)\|^2 \leq \varepsilon\}$. Here $\varepsilon > 0$ is such that x^* is the global minimizer of minimizing $f(x)$ over $\mathcal{X} \cap \mathcal{B}_\varepsilon(x^*)$.

Since the set $\Omega \cap C$ is compact and $F_k(x,y,z)$ is continuous, the problem above has at least one optimal solution, say (x^k,y^k,z^k). We next show that $(x^k,y^k,z^k) \to (x^*,y^*,z^*)$ as $k \to \infty$. Due to the optimality of (x^k,y^k,z^k), we have that $F_k(x^k,y^k,z^k) \leq F_k(x^*,y^*,z^*) = f(x^*)$ for all $k \in \mathbb{N}$, i.e.,

$$\begin{aligned}&f(x^k) + \frac{k}{2}\|\max\{0,g(x^k)\}\|^2 + \frac{k}{2}\|h(x^k)\|^2 + \frac{k}{2}\|G(x^k)-y^k\|^2 \\ &+ \frac{k}{2}\|H(x^k)-z^k\|^2 + \frac{1}{2}\|(x^k,y^k,z^k)-(x^*,y^*,z^*)\|^2 \leq f(x^*), k \in \mathbb{N}.\end{aligned} \tag{17}$$

Since $\{f(x^k)\}$ is bounded due to the compactness of $\Omega \cap C$, it follows from the above inequality that

$$\lim_{k\to\infty} \max\{0,g(x^k)\} = 0, \qquad \lim_{k\to\infty} h(x^k) = 0,$$
$$\lim_{k\to\infty} G(x^k) - y^k = 0, \qquad \lim_{k\to\infty} H(x^k) - z^k = 0.$$

Let $(\bar{x},\bar{y},\bar{z})$ be an arbitrary accumulation point of $\{(x^k,y^k,z^k)\}$. Then from (17), it follows that

$$f(\bar{x}) + \frac{1}{2}\|(\bar{x},\bar{y},\bar{z})-(x^*,y^*,z^*)\|^2 \leq f(x^*).$$

Note that $f(x^*) \leq f(\bar{x})$ by the local optimality of x^*. This and the above inequality imply that $(\bar{x},\bar{y},\bar{z}) = (x^*,y^*,z^*)$. Thus the whole sequence $\{(x^k,y^k,z^k)\}$ converges to (x^*,y^*,z^*) as $k \to \infty$.

Without loss of generality, we may assume that (x^k, y^k, z^k) is an interior point of Ω for all $k \in \mathbb{N}$. By Fermat's rule, it follows that

$$-\nabla F_k(x^k, y^k, z^k) \in \widehat{\mathcal{N}}_C(x^k, y^k, z^k), \quad k \in \mathbb{N}. \tag{18}$$

By Proposition 2.1, the regular normal cone $\widehat{\mathcal{N}}_C(x^k, y^k, z^k)$ can be written as

$$\widehat{\mathcal{N}}_C(x^k, y^k, z^k) := \left\{ \begin{pmatrix} 0 \\ \xi \\ \zeta \end{pmatrix} : \begin{array}{l} \xi_i \in \mathbb{R}, \zeta_i = 0 \text{ if } z_i^k \neq 0 \\ \xi_i = 0, \zeta_i \in \mathbb{R} \text{ if } y_i^k \neq 0 \\ \xi_i = 0, \zeta_i = 0 \text{ if } y_i^k = 0, z_i^k = 0 \end{array} \right\}. \tag{19}$$

Combining (18) with (19), it is easy to derive that for $k \in \mathbb{N}$,

$$\begin{aligned} \nabla f(x^k) &+ \sum_{i=1}^m k \max\{0, g_i(x^k)\} \nabla g_i(x^k) + \sum_{i=1}^p k h_i(x^k) \nabla h_i(x^k) + (x^k - x^*) \\ &+ \sum_{i=1}^l k(G_i(x^k) - y_i^k) \nabla G_i(x^k) + \sum_{i=1}^l k(H_i(x^k) - z_i^k) \nabla H_i(x^k) = 0, \end{aligned} \tag{20}$$

and

$$\begin{aligned} k(G_i(x^k) - y_i^k) &= y_i^k - y_i^* \quad \text{if } y_i^k \neq 0, \\ k(H_i(x^k) - z_i^k) &= z_i^k - z_i^* \quad \text{if } z_i^k \neq 0, \\ k(G_i(x^k) - y_i^k) &= y_i^k - y_i^* \quad \text{if } y_i^k = z_i^k = 0, \\ k(H_i(x^k) - z_i^k) &= z_i^k - z_i^* \quad \text{if } y_i^k = z_i^k = 0. \end{aligned} \tag{21}$$

Let

$$\begin{aligned} \delta^k &:= \Big(1 + \|k \max\{0, g(x^k)\}\|^2 + \|kh(x^k)\|^2 + \|k(G(x^k) - y^k)\|^2 \\ &\qquad + \|(H(x^k) - z^k)\|^2\Big)^{1/2}, \\ \alpha^k &:= \frac{1}{\delta^k}, \quad \lambda^k := \frac{k \max\{0, g(x^k)\}}{\delta^k}, \quad \rho^k := \frac{kh(x^k)}{\delta^k}, \\ \mu^k &:= \frac{k(G(x^k) - y^k)}{\delta^k}, \quad \nu^k := \frac{k(H(x^k) - z^k)}{\delta^k}. \end{aligned}$$

It is easy to see that $\|(\alpha^k, \lambda^k, \rho^k, \mu^k, \nu^k)\| = 1$ for all $k \in \mathbb{N}$. We may assume that $(\alpha, \lambda, \rho, \mu, \nu) \neq 0$ is an accumulation point of $(\alpha^k, \lambda^k, \rho^k, \mu^k, \nu^k)$. Dividing (20) by σ^k and taking limits on both sides of (20) yield that

$$\alpha \nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \rho_i \nabla h_i(x^*) + \sum_{i=1}^l \mu_i \nabla G_i(x^*) + \sum_{i=1}^l \nu_i \nabla H_i(x^*) = 0.$$

It is easy to see that $\alpha \geq 0$ and $\lambda \geq 0$. Moreover, by the definition of λ^k and the continuity of the function g, we have that $\lambda_i = 0$ for each $i \notin \mathcal{I}_*^g$. When $i \in \mathcal{I}_*^H$, then $y_i^* = G_i(x^*) \neq 0$ and thus $y_i^k \neq 0$ for all k sufficient large. Then by (21) and the fact that $y^k \to y^*$ as $k \to \infty$, we have that

$$\mu_i = \lim_{k\to\infty} \frac{k(G_i(x^k) - y_i^k)}{\delta^k} = \lim_{k\to\infty} \frac{y_i^k - y_i^*}{\delta^k} = 0,\ i \in \mathcal{I}_*^H.$$

In the same way, we can show that $\nu_i = 0$ for all $i \in \mathcal{I}_*^G$. For $i \in \mathcal{I}_*^{GH}$, at least one of the following three cases occurs: (a) When there exist infinitely many k such that $y_i^k \neq 0,\ z_i^k = 0$: in this case, we can similarly show that $\mu_i = 0$. (b) When there exist infinitely many k such that $y_i^k = 0,\ z_i^k \neq 0$: we can similarly show that $\nu_i = 0$. (c) When there exist infinitely many k such that $y_i^k = 0,\ z_i^k = 0$: it follows from (21) that

$$\mu_i = \lim_{k\to\infty} \frac{k(G_i(x^k) - y_i^k)}{\delta^k} = \lim_{k\to\infty} \frac{y_i^k - y_i^*}{\delta^k} = 0,$$

$$\nu_i = \lim_{k\to\infty} \frac{k(H_i(x^k) - z_i^k)}{\delta^k} = \lim_{k\to\infty} \frac{v_i^k - z_i^*}{\delta^k} = 0.$$

In sum, we have shown that $\mu_i \nu_i = 0$ for any $i \in \mathcal{I}_*^{GH}$.

We now show condition (iv). Assume that $(\lambda, \rho, \mu, \nu) \neq 0$. By the definition of λ^k and ρ^k, it is easy to see that for all k sufficiently large,

$$\lambda_i > 0 \Rightarrow \lambda_i^k > 0 \Rightarrow g_i(x^k) > 0,$$
$$\rho_i \neq 0 \Rightarrow \rho_i \rho_i^k > 0 \Rightarrow \rho_i h_i(x^k) > 0.$$

For any $\mu_i \neq 0$, it follows that $\mu_i \mu_i^k > 0$ or equivalently $\mu_i(G_i(x^k) - y_i^k) > 0$ for all k sufficiently large. When $\mu_i \neq 0$, we have that $y_i^k = 0$ for all k sufficiently large. Otherwise if there exists a subsequence of $\{y_i^k\}$ not equal to 0, then $\mu_i = 0$ follows from (20) immediately. Thus, $\mu_i G_i(x^k) > 0$ for all k sufficiently large when $\mu_i \neq 0$. By symmetry, we can also show that $\nu_i H_i(x^k) > 0$ for all k sufficiently large when $\nu_i \neq 0$. The proof is complete. □

The following theorem follows from Lemma 4.1 immediately.

Theorem 4.3 *If MPSC quasi-normality holds at $x^* \in \mathcal{X}$, then MPSC MCQ holds at x^* as well.*

Proof For any $d \in \mathcal{T}_{\mathcal{X}}(x^*)^o$, it suffices to show that $d \in \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*))$. Since $d \in \mathcal{T}_{\mathcal{X}}(x^*)^o$, it follows from Rockafellar & Wets (1998, Theorem 6.11) that there exists a s-mooth function φ that achieves a minimizer relative to $\mathcal{X}$ at x^* with $-\nabla\varphi(x^*) = d$. Then by Lemma 4.1, there exist vectors $\{\alpha, \lambda, \rho, \mu, \nu\}$ satisfying (i)-(iv) in Lemma 4.1 with function φ in place of function f. Since MPSC quasi-normality holds at x^*, it is easy to see that $\alpha \neq 0$. This implies that x^* is an M-stationary point of minimizing φ over $\mathcal{X}$. By Proposition 3.1, we have that $0 \in \nabla\varphi(x^*)+\nabla F(x^*)\mathcal{N}_\Lambda(F(x^*))$. Thus, $d \in \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*))$. The proof is complete. □

In the rest of this section, we investigate the relation between MPSC RCPLD and MP-SC MCQ. To this end, we recall a lemma which can be seen as a corollary of Carathéodory's lemma (Andreani et al., 2012).

Lemma 4.2 *Let* $0 \neq x = \sum_{i=1}^{m+p} \alpha_i v_i$, *where* $\{v_1, \ldots, v_m\}$ *is linearly independent and* $\alpha_i \neq 0$ *for all* $i = m+1, \ldots, m+p$. *Then there exist* $\mathcal{J} \subset \{m+1, \ldots, m+p\}$ *and* $\bar{\alpha}_i$, $i \in \{1, \ldots, m\} \cup \mathcal{J}$, *such that* $x = \sum_{i\in\{1,\ldots,m\}\cup\mathcal{J}} \bar{\alpha}_i v_i$ *with* $\alpha_i\bar{\alpha}_i > 0$ *for every* $i \in \mathcal{J}$ *and* $\{v_i : i \in \{1, \ldots, m\} \cup \mathcal{J}\}$ *is linearly independent.*

Theorem 4.4 *If MPSC RCPLD holds at* $x^* \in \mathcal{X}$, *then MPSC MCQ holds at* x^*.

Proof For any $d \in \mathcal{T}_{\mathcal{X}}(x^*)^o$, by Rockafellar & Wets (1998, Theorem 6.11), it follows that there exists a smooth function φ such that $-\nabla\varphi(x^*) = d$ and $\arg\min_{x\in\mathcal{X}} \varphi(x) = \{x^*\}$. In the same way as the proof of Lemma 4.1, we have that there exists a sequence (x^k, y^k, z^k) converging to (x^*, y^*, z^*) such that (20) and (21) hold with function φ in place of function f. By a simple arrangement, we have

$$\xi^k + \sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla h_i(x^k) + \sum_{i=1}^{l} \nu_i \nabla G_i(x^k) + \sum_{i=1}^{l} \nu_i \nabla H_i(x^k) = 0, \tag{22}$$

where

$$\begin{aligned}
&\lambda_i^k := k \max\{0, g_i(x^k)\},\ \rho_i^k := k h_i(x^k),\\
&\mu_i^k := k(G_i(x^k) - y_i^k) - (y_i^k - y_i^*),\ \nu_i^k := k(H_i(x^k) - z_i^k) - (z_i^k - z_i^*),\\
&\xi^k := \nabla\varphi(x^k) + (x^k - x^*) + \sum_{i=1}^{l}(y_i^k - y_i^*)\nabla G_i(x^k) + \sum_{i=1}^{l}(z_i^k - z_i^*)\nabla H_i(x^k).
\end{aligned}$$

By (21), one can easily have $(\mu_i^k, \nu_i^k) \in \mathcal{N}_{\mathcal{S}}(y_i^k, z_i^k)$ for all $i = 1, \ldots, l$. For the sake of convenience, we denote

$$\begin{cases} \mathcal{I}_k^g := \{i \in \{1, ..., m\} : g_i(x^k) = 0\}, \\ \mathcal{I}_k^G := \{i \in \{1, ..., l\} : G_i(x^k) = 0 \wedge H_i(x^k) \neq 0\}, \\ \mathcal{I}_k^H := \{i \in \{1, ..., l\} : G_i(x^k) \neq 0 \wedge H_i(x^k) = 0\}, \\ \mathcal{I}_k^{GH} := \{i \in \{1, ..., l\} : G_i(x^k) = 0 \wedge H_i(x^k) = 0\}. \end{cases}$$

It is obvious that $\mathcal{I}_*^G \subseteq \mathcal{I}_k^G$ and $\mathcal{I}_*^H \subseteq \mathcal{I}_k^H$ for each k sufficiently large. Denote $\text{supp}\,(a) := \{i : a_i \neq 0\}$. It then follows from (22) that

$$\begin{aligned} 0 = \xi^k + \sum_{i \in \text{supp}\,(\lambda^k)} \lambda_i^k \nabla g_i(x^k) + \sum_{j \in \mathcal{I}^h} \rho_j^k \nabla h_j(x^k) + \sum_{r \in \mathcal{I}_*^G} \mu_r^k \nabla G_r(x^k) \\ + \sum_{t \in \mathcal{I}_*^H} \nu_t^k \nabla H_t(x^k) + \sum_{r \in \mathcal{I}_k^G \setminus \mathcal{I}_*^G \cup \mathcal{I}_k^{GH} \cap \text{supp}\,(\mu^k)} \mu_r^k \nabla G_r(x^k) \\ + \sum_{t \in \mathcal{I}_k^H \setminus \mathcal{I}_*^H \cup \mathcal{I}_k^{GH} \cap \text{supp}\,(\nu^k)} \nu_t^k \nabla H_t(x^k). \end{aligned} \tag{23}$$

If there exists a subsequence such that $\xi^k = 0$. Then $d = \lim_{k \to \infty} \nabla \varphi(x^k) = 0$ and hence $d \in \nabla F(x^*) \mathcal{N}_\Lambda(F(x^*))$. Thus, without loss of generality, we assume that $\xi^k \neq 0$ for all k. Let $I_1 \subseteq \mathcal{I}^h$, $I_2 \subset \mathcal{I}_*^G$, and $I_3 \subset \mathcal{I}_*^H$ be index sets such that $\mathcal{G}(x^*, I_1, I_2, I_3)$ is a basis for span $\mathcal{G}(x^*, \mathcal{I}^h, \mathcal{I}_*^G, \mathcal{I}_*^H)$. Then $\mathcal{G}(x^k, I_1, I_2, I_3)$ is linearly independent for k sufficiently large. Since MPSC RCPLD holds at x^*, by the definition, there is a constant $\delta > 0$ such that the rank of $\mathcal{G}(x, \mathcal{I}^h, \mathcal{I}_*^G, \mathcal{I}_*^H)$ is constant for each $x \in \mathcal{B}_\delta(x^*)$. Thus $\mathcal{G}(x^k, I_1, I_2, I_3)$ is a basis for span $\mathcal{G}(x^k, \mathcal{I}^h, \mathcal{I}_*^G, \mathcal{I}_*^H)$ for all k sufficiently large. Then it follows from Lemma 4.2 that there exist $I_4^k \subseteq \text{supp}\,(\lambda^k)$, $I_5^k \subseteq \mathcal{I}_k^G \setminus \mathcal{I}_*^G \cup \mathcal{I}_k^{GH} \cap \text{supp}\,(\mu^k)$, $I_6^k \subseteq \mathcal{I}_k^H \setminus \mathcal{I}_*^H \cup \mathcal{I}_k^{GH} \cap \text{supp}\,(\nu^k)$, and $\{\bar{\lambda}^k, \bar{\rho}^k, \bar{\mu}^k, \bar{\nu}^k\}$ such that

$$\begin{aligned} 0 = \xi^k + \sum_{i \in I_4^k} \bar{\lambda}_i^k \nabla g_i(x^k) + \sum_{j \in I_1} \bar{\rho}_j^k \nabla h_j(x^k) + \sum_{r \in I_2} \bar{\mu}_r^k \nabla G_r(x^k) \\ + \sum_{t \in I_3} \bar{\nu}_t^k \nabla H_t(x^k) + \sum_{r \in I_5^k} \bar{\mu}_r^k \nabla G_r(x^k) + \sum_{t \in I_6^k} \bar{\nu}_t^k \nabla H_t(x^k) \end{aligned} \tag{24}$$

and the vectors $\{\nabla g_i(x^k) : i \in I_4^k\}$, $\{\nabla h_j(x^k) : j \in I_1\}$, $\{\nabla G_r(x^k) : r \in I_2 \cup I_5^k\}$, $\{\nabla H_t(x^k) : t \in I_3 \cup I_6^k\}$ are linearly independent for all k sufficiently large. From the implementation process and Lemma 4.2, we also have $\bar{\lambda}_i^k \geq 0$ for all $i \in \mathcal{I}_4^k$ and $(\bar{\mu}_i^k, \bar{\nu}_i^k) \in$

$\mathcal{N}_{\mathcal{S}}(y_i^k, z_i^k)$ for all $i = 1, \dots, l$. Without any loss of generality, we assume that $I_4^k \equiv I_4$, $I_5^k \equiv I_5$, and $I_6^k \equiv I_6$. Then the vectors

$$\begin{aligned}\{\nabla g_i(x^k) : i \in I_4\}, \{\nabla h_j(x^k) : j \in I_1\},\ \{\nabla G_r(x^k) : r \in I_2 \cup I_5\},\\ \{\nabla H_t(x^k) : t \in I_3 \cup I_6\} \text{ are linearly independent.}\end{aligned} \tag{25}$$

It is not hord to get that $I_4 \subseteq \mathcal{I}_*^g$, $I_1 \subseteq \mathcal{I}^h$, and $I_5, I_6 \subseteq \mathcal{I}_*^{GH}$ by $\mathcal{I}_*^G \cup \mathcal{I}_*^H \cup \mathcal{I}_*^{GH} = \mathcal{I}_k^G \cup \mathcal{I}_k^H \cup \mathcal{I}_k^{GH}$. Define

$$M_k := \max\{\bar{\lambda}_i^k, \bar{\rho}_j^k, \bar{\mu}_r^k, \bar{\nu}_t^k : i \in I_4, j \in I_1, r \in I_2 \cup I_5, t \in I_3 \cup I_6\}.$$

If there exists a subsequence such that $M_k \to \infty$, then we may take a subsequence such that for any $i \in I_4$, $j \in I_1$, $r \in I_2 \cup I_5$, $t \in I_3 \cup I_6$,

$$\frac{(\bar{\lambda}_i^k, \bar{\rho}_j^k, \bar{\mu}_r^k, \bar{\nu}_t^k)}{M_k} \to (\bar{\lambda}_i^*, \bar{\rho}_j^*, \bar{\mu}_r^*, \bar{\nu}_t^*) \quad \text{as } k \to \infty.$$

Dividing (24) by M_k and taking limits yield

$$\begin{aligned}\sum_{i\in I_4} \bar{\lambda}_i^* \nabla g_i(x^*) + \sum_{j\in I_1} \bar{\rho}_j^* \nabla h_j(x^*) + \sum_{r\in I_2} \bar{\mu}_r^* \nabla G_r(x^*) + \sum_{t\in I_3} \bar{\nu}_t^* \nabla H_t(x^*)\\ + \sum_{r\in I_5} \bar{\mu}_r^* \nabla G_r(x^*) + \sum_{t\in I_6} \bar{\nu}_t^* \nabla H_t(x^*) = 0.\end{aligned}$$

Moreover, we also have $\bar{\lambda}_i^* \geq 0$ for all $i \in I_4$ and the fact that $(\bar{\mu}_i^k, \bar{\nu}_i^k) \in \mathcal{N}_{\mathcal{S}}(y_i^k, z_i^k)$ implies $(\bar{\mu}_i^*, \bar{\nu}_i^*) \in \mathcal{N}_{\mathcal{S}}(y_i^*, z_i^*)$ from the outer semi-continuity of limiting normal cones. The last two facts and (25) give a contradiction with MPSC RCPLD. Hence $\{M_k\}$ is bounded. Taking limits for a suitable subsequence such that for any $i \in I_4$, $j \in I_1$, $r \in I_2 \cup I_5$, $t \in I_3 \cup I_6$,

$$(\bar{\lambda}_i^k, \bar{\rho}_j^k, \bar{\mu}_r^k, \bar{\nu}_t^k) \to (\bar{\lambda}_i^*, \bar{\rho}_j^*, \bar{\mu}_r^*, \bar{\nu}_t^*) \quad \text{as } k \to \infty,$$

it follows from (24) that

$$\begin{aligned}-\nabla\varphi(x^*) = \sum_{i\in I_4} \bar{\lambda}_i^* \nabla g_i(x^*) + \sum_{j\in I_1} \bar{\rho}_j^* \nabla h_j(x^*) + \sum_{r\in I_2} \bar{\mu}_r^* \nabla G_r(x^*)\\ + \sum_{t\in I_3} \bar{\nu}_t^* \nabla H_t(x^*) + \sum_{r\in I_5} \bar{\mu}_r^* \nabla G_r(x^*) + \sum_{t\in I_6} \bar{\nu}_t^* \nabla H_t(x^*).\end{aligned}$$

This prove that x^* is an M-stationary point of minimizing φ over $\mathcal{X}$. By Proposition 3.1, it follows that $0 \in \nabla\varphi(x^*) + \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*))$. Thus,

$$d \in \nabla F(x^*)\mathcal{N}_\Lambda(F(x^*)).$$

The proof is complete. □

Finally, we summarize the relations among the constraint qualifications for MPSC in Fig. 1.

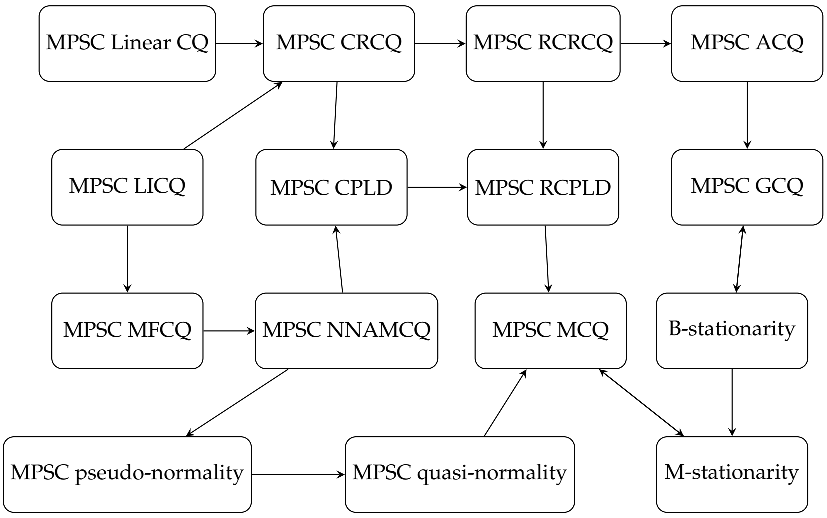

Fig.1 Relations among various CQs

Acknowledgements This work was supported in part by NSFC Grants (No.11771287, No.11401379), the Fundamental Research Funds for the Central Universities, the High Level Introduction of Talent Research Start-up Fund (No.1856009), and the Scientific and Technological Research Program of Chongqing Municipal Education Commission (No.KJQN201800810).

References

Achtziger, W., & Kanzow, C. (2008). Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications. *Mathematical Programming*, *114*, 69–99.

Andreani, R., Haeser, G., Schuverdt, M. L., & Silva, P. J. (2012). A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, *135*, 255–273.

Bertsekas, D. P., & Ozdaglar, A. E. (2002). Pseudonormality and a lagrange multiplier theory for constrained optimization. *Journal of Optimization Theory and Applications*, *114*, 287–343.

Clason, C., Rund, A., & Kunisch, K. (2016). Nonconvex penalization of switching control of partial differential equations. *Systems and Control Letters*, *106*, 1–8.

Dempe, S., & Schreier, H. (2007). *Operations research: deterministische Modelle und Methoden*. Springer-Verlag.

Gugat, M. (2010). Optimal switching boundary control of a string to rest in finite time. *Journal of Applied Mathematics and Mechanics*, *88*, 283–305.

Guo, L., & Lin, G. H. (2013). Notes on some constraint qualifications for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, *156*, 600–616.

Guo, L., Lin, G. H., & Ye, J. J. (2013a). Second-order optimality conditions for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, *158*, 33–64.

Guo, L., Ye, J. J., & Zhang, J. (2013b). Mathematical programs with geometric constraints in banach spaces: enhanced optimality, exact penalty, and sensitivity. *SIAM Journal on Optimization*, *23*, 2295–2319.

Hante, F. M., & Sager, S. (2013). Relaxation methods for mixed-integer optimal control of partial differential equations. *Computational Optimization and Applications*, *55*, 197–225.

Hoheisel, T. (2009). Mathematical programs with vanishing constraints. *Ph.D. thesis, Julius-Maximilians-Universitat Wärzburg*, .

Hoheisel, T., Kanzow, C., & Outrata, J. (2010). Exact penalty results for mathematical programs with vanishing constraints. *Nonlinear Analysis: Theory, Methods and Applications*, *72*, 2514–2526.

Izmailov, A. F., & Solodov, M. V. (2009). Mathematical programs with vanishing constraints: Optimality conditions, sensitivity, and a relaxation method. *Journal of Opti-*

mization Theory and Applications, *142*, 501–532.

Janin, R. (1984). Directional derivative of the marginal function in nonlinear programming. In *Sensitivity, Stability and Parametric Analysis* (pp. 110–126). Springer.

Kanzow, C., Mehlitz, P., & Steck, D. (2018). Relaxation schemes for mathematical programs with switching constraints. *Optimization Online*, .

Kanzow, C., & Schwartz, A. (2010). Mathematical programs with equilibrium constraints: enhanced fritz john-conditions, new constraint qualifications, and improved exact penalty results. *SIAM Journal on Optimization*, *20*, 2730–2753.

Lin, G.-H., & Fukushima, M. (2005). A modified relaxation scheme for mathematical programs with complementarity constraints. *Annals of Operations Research*, *133*, 63–84.

Luo, Z.-Q., Pang, J.-S., & Ralph, D. (1996a). *Mathematical programs with equilibrium constraints*. Cambridge University Press.

Luo, Z. Q., Pang, J. S., Ralph, D., & Wu, S. Q. (1996b). Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints. *Mathematical Programming*, *75*, 19–76.

Mehlitz, P. (2019a). On the linear independence constraint qualification in disjunctive programming. *Optimization Online*, .

Mehlitz, P. (2019b). Stationary conditions and constraint qualification for mathematical programs with switching constraints. *Mathematical Programming, https://doi.org/10.1007/s10107-019-01380-5*, .

Minchenko, L., & Stakhovski, S. (2011). Parametric nonlinear programming problems under the relaxed constant rank condition. *SIAM Journal on Optimization*, *21*, 314–332.

Mishra, S. K., Singh, V., & Laha, V. (2016). On duality for mathematical programs with vanishing constraints. *Annals of Operations Research*, *243*, 249–272.

Outrata, J. V. (1999). Optimality conditions for a class of mathematical programs with equilibrium constraints. *Mathematics of Operations Research*, *24*, 627–644.

Qi, L., & Wei, Z. (2000). On the constant positive linear dependence condition and its application to sqp methods. *SIAM Journal on Optimization*, *10*, 963–981.

Rockafellar, R. T., & Wets, R. J. B. (1998). Variational analysis. *Springer,Berlin*, .

Shim, Y., Fodstad, M., Gabriel, S. A., & Tomasgard, A. (2013). A branch-and-bound method for discretely-constrained mathematical programs with equilibrium constraints. *Annals of Operations Research*, *210*, 5–31.

Wang, L., & Yan, Q. (2016). Bang-bang property of time optimal controls for some semilinear heat equation. *SIAM Journal on Control and Optimization*, *54*, 477–499.

Ye, J. J., & Zhang, J. (2014). Enhanced Karush-Kuhn-Tucker conditions for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, *163*, 777–794.