mathematics

MDPI

Article

Optimality Conditions, Qualifications and Approximation Method for a Class of Non-Lipschitz Mathematical Programs with Switching Constraints

Jinman Lv [1,*], Zhenhua Peng [2,*] and Zhongping Wan [1]

[1] School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China; mathwanzhp@whu.edu.cn
[2] Department of Mathematics, School of Sciences, Nanchang University, Nanchang 330031, China
* Correspondence: lvjinman@whu.edu.cn (J.L.); zhenhuapeng@whu.edu.cn (Z.P.)

Abstract: In this paper, we consider a class of mathematical programs with switching constraints (MPSCs) where the objective involves a non-Lipschitz term. Due to the non-Lipschitz continuity of the objective function, the existing constraint qualifications for local Lipschitz MPSCs are invalid to ensure that necessary conditions hold at the local minimizer. Therefore, we propose some MPSC-tailored qualifications which are related to the constraints and the non-Lipschitz term to ensure that local minimizers satisfy the necessary optimality conditions. Moreover, we study the weak, Mordukhovich, Bouligand, strongly (W-, M-, B-, S-) stationay, analyze what qualifications making local minimizers satisfy the (M-, B-, S-) stationay, and discuss the relationship between the given MPSC-tailored qualifications. Finally, an approximation method for solving the non-Lipschitz MPSCs is given, and we show that the accumulation point of the sequence generated by the approximation method satisfies S-stationary under the second-order necessary condition and MPSC Mangasarian-Fromovitz (MF) qualification.

Keywords: switching constraint; stationary point; non-Lipschitz; optimality condition

check for updates

Citation: Lv, J.; Peng, Z.; Wan, Z. Optimality Conditions, Qualifications and Approximation Method for a Class of Non-Lipschitz Mathematical Programs with Switching Constraints. *Mathematics* **2021**, *9*, 2915. https://doi.org/10.3390/math9222915

Academic Editor: Sorana D. Bolboacă

Received: 18 October 2021
Accepted: 13 November 2021
Published: 16 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The mathematical program with switching constraints (MPSCs) is a class of optimization problems which consider products of two functions as their sum of equality constraints. "Switching" means that the product of two functions is equal to 0 if and only if at least one of two functions equals 0. The current applications for switching structures appear in optimal control frequently [1–3] and optimization problems with either–or constraints [4]. In real world applications, switching structures naturally appear via modelling the bacteria switch instantaneously between active and dormant states [5], and modelling the state at the two boundary points to control a finite string to the zero state in finite time [6]. From the perspective of problem structure, mathematical programs with switching constraints (MPSCs) are closely related to mathematical programs with complementarity constraints (MPCCs) [7,8], mathematical programs with vanishing constraints (MPVCs) [9,10], and mathematical programs with cardinality constraints [11,12]. By the structures of MPCC and MPSC, we know that any MPCC is a special form of MPSC. There are many studies on MPCCs; as is known to all, standard constraint qualifications such as linear independence constraint qualification (LICQ) and Mangasarian–Fromovitz constraint qualification (MFCQ) do not hold at the feasible points of MPCCs due to the complementarity constraints. Similarly, LICQ and MFCQ also do not hold at the feasible points of MPSCs due to the switching constraints. In order to resolve this problem, Mehlitz [13] studied the (W-, M-, S-) stationarity and some MPSC-constraint qualifications and their relationships are given. Thereafter, there are some studies on the optimality conditions of MPSCs [14,15]. The solution methods of MPSCs at present are the relaxation method [16] and penalty methods [15,17,18].

Sparse solutions of systems have been widely used to seek a sparse solution in regression, feature selection in machine learning and compressed sensing [19–22]. In fact, finding the sparsest solutions is to find the solution which has the smallest number of nonzero components. The number of nonzero components of a vector z denoted as $\|z\|_0$, the so-called l_0-norm, is a nonconvex discontinuous function, usually approximated by l_p-norm. The non-Lipschitz function is a better continuous approximation for l_0-norm. Moreover, some smoothing methods [23–26] are provided for solving such problems. Guo and Chen [27] derived associated qualifications, necessary optimality conditions and an approximation method for MPCCs with the objective function involving a non-Lipschitz term. However, we know MPCC is a special structure of MPSC. Existing theory for MPCCs with a non-Lipschitz term are not applicable to MPSCs due to the switching constrains. In order to replenish this shortcoming, it is necessary to introduce the concepts of stationarity points and compatible constraint qualifications for MPSCs with the objective function involving a non-Lipschitz term. This is the central purpose of this paper. Our main contributions are summarized as follows.

1. We prove that the basic qualification (BQ) fails at any feasible point of Equation (1) when D is full row rank.
2. We present the (W-, M-, B-, S-) stationarity conditions for Equation (1), propose 12 MPSC-tailored qualifications, and give the qualifications which can guarantee that M-stationarity and S-stationarity are necessary for local optimality of Equation (1), respectively. In addition, the diagram of the relations for 12 qualifications and (M-, S-) stationarity are given.
3. We propose an approximation method for solving problem Equation (1), in which a local Lipschitz function is used to approximate the non-Lipschitz term. We show that any accumulation point of the sequence generated by our method is W-stationary under MPSC-MF qualification, and is S-stationary under second-order necessary condition and MPSC-MF qualification.

The rest of this paper is organized as follows. In Section 2, the notations and preliminary definitions are given. In Section 3, we propose some qualifications which are related to the constraints and the non-Lipschitz term to ensure that local minimizers satisfy the necessary optimality conditions; and the relations among MPSC tailored qualifications are discussed. In Section 4, a relaxation scheme for solving the non-Lipschitz MPSCs is presented, and under the MPSC-LI qualification and second-order necessary condition, we can obtain its W-stationary and S-stationary. Conclusions are summarized in Section 5.

2. Notations and Preliminaries

In order to facilitate the reading of this article, some notations are given first:

$\mathbb{N}$	Set of natural numbers	$\widehat{\mathcal{N}}_A(x)$	The regular normal cone of A at x
$\mathbb{R}$	Set of real numbers	$\mathcal{N}_A(x)$	The limiting normal cone of A at x
$\mathbb{R}^n$	Set of real n-vectors	$T_A(x)$	The tangent cone of A at x
$\mathbb{R}^{m\times n}$	Set of real $m \times n$-matrices	$\widehat{\partial}\psi(x)$	The regular subdifferential of ψ at x
$\|\cdot\|$	A norm	$\nabla\psi(x)$	The gradient of ψ at x
$\|\cdot\|_p$	p norm	$\partial\psi(x)$	The limiting subdifferential of ψ at x
D^T	Transpose of matrix D	$\partial^\infty\psi(x)$	The horizon subdifferential of ψ at x
span$D$	The spanning set of set D	$\nabla^2\psi(x)$	The Hessian Matrix of ψ at x
$\langle\cdot,\cdot\rangle$	Inner product	$supp(A)$	Index set of non-zero elements of A

The following notation will be used throughout this paper. Given a nonempty closed set A and a point $x^* \in A$.

Definition 1. *(Regular normal cone) The regular normal cone of A at x^* is defined as*

$$\widehat{\mathcal{N}}_A(x^*) := \{d : \langle d, x - x^*\rangle \leq o(\|x - x^*\|),\ \forall x \in A\},$$

where $o(\cdot)$ means that $o(\alpha)/\alpha \to 0$ as $\alpha \downarrow 0$.

Definition 2. *(Limiting normal cone) The limiting normal cone of A at x^* is defined as*

$$\mathcal{N}_A(x^*) := \{d : \exists x^k \in A, x^k \to x^*, \exists d^k \in \widehat{N}_A(x^k),\ s.t.\ d^k \to d\}.$$

Definition 3. *(Tangent cone) The tangent cone of A at x^* is defined as*

$$T_A(x^*) := \{d : \exists x^k \in A, \exists t_k \searrow 0 : x^k \to x^*, and\ (z^* - z^k)/t_k \to d\}.$$

Given a continuous function $\psi : \mathbb{R}^s \to \mathbb{R}$ and a point $\bar{x} \in \mathbb{R}^s$, we recall the definitions of regular subdifferential, limiting subdifferential, and horizon subdifferential [28] (Definition 8.3).

Definition 4. *(Regular subdifferential) The regular subdifferential of ψ at $\bar{x}$ is defined as*

$$\widehat{\partial}\psi(\bar{x}) := \{v : \psi(x) \geq \psi(\bar{x}) + v^T(x - \bar{x}) + o(\{x - \bar{x}\|)\ \forall x \in \mathbb{R}^n\}.$$

Definition 5. *(Limiting subdifferential) The limiting subdifferential of ψ at $\bar{x}$ is defined as*

$$\partial\psi(\bar{x}) := \{v : \exists x^k \to \bar{x}, v^k \in \hat{\partial}\psi(x^k),\ s.t.\ v^k \to v\}.$$

Definition 6. *(Horizon subdifferential) The horizon subdifferential of ψ at $\bar{x}$ is defined as*

$$\partial^\infty\psi(\bar{x}) := \{v : \exists x^k \to \bar{x}, v^k \in \widehat{\partial}\psi(x^k)\ and\ t_k \downarrow 0,\ s.t.\ t_k v^k \to v\}.$$

It is worth noting that $\partial^\infty\psi(\bar{x}) = \{0\}$ if function ψ is strictly continuous at $\bar{x}$ [28] (Theorem 9.13); and $\partial\psi(\bar{x}) = \{\nabla\psi(\bar{x})\}$ if ψ is continuously differential at $\bar{x}$ [28] (Exercise 8.8).

3. The Non-Lipschitz MPSCs and Their Optimality Conditions

We mainly consider the MPSCs with the objective function involving a non-Lipschitz term $||\cdot||_p^p$ as below

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & F(x) = f(x) + \|Dx\|_p^p \\ s.t. & g_i(x) \leq 0,\ i \in \mathcal{I} = \{1, \cdots, m\}, \\ & h_j(x) = 0,\ j \in \mathcal{J} = \{1, \cdots, q\}, \\ & G_k(x)H_k(x) = 0,\ k \in \mathcal{K} = \{1, \cdots, l\}. \end{array} \tag{1}$$

where all functions $f, g_i, h_j, G_k, H_k : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable for all $i \in \mathcal{I}$, $j \in \mathcal{J}$, $k \in \mathcal{K}$, and D is a given real matrix in $\mathbb{R}^{r\times n}$, and $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$. We call $G_k(x)H_k(x) = 0$ the switching constraint since functions $G_k(x), H_k(x)$ are active at least one, that is to say, the last l constraints in problem (1) force $G_k(x) = 0$ or $H_k(x) = 0$ for all $k = 1, \cdots, l$ at any feasible point of (MPSC). Denote the feasible set of problem (1) by $X := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i \in \mathcal{I}, h_j(x) = 0, j \in \mathcal{J}, G_k(x)H_k(x) = 0, k \in \mathcal{K}\}$. Comparing with the classical MPSCs, the difficulty of problem (1) is to deal with the term $\|Dx\|_p^p$ which is non-convex and non-Lipschitz when $0 < p < 1$. The basic qualification (BQ) of problem (1) holds at $\bar{x} \in X$ if $-\partial^\infty F(\bar{x}) \cap \mathcal{N}_X(\bar{x}) = \{0\}$. The BQ ensures the validity of the sum rule of the subdifferentials of the objective function F and an indicator function, which plays an important role in deriving necessary optimality conditions for problem (1).

3.1. Stationary Conditions and Qualifications

We derive three different stationarity notions of problem (1) which are not stronger than the associated KKT conditions. We first introduce some index sets of a feasible point $x \in \Omega$ of problem (1), D_i is the transpose of the ith row of D:

$$I_{D0}(x) := \{i \in \{1, \cdots, r\} : D_i^T x = 0\},\ I_{D\neq}(x) := \{i \in \{1, \cdots, r\} : D_i^T x \neq 0\},$$
$$I_G(x) := \{k \in \mathcal{K} : G_i(x) = 0, H_k(x) \neq 0\},\ I_H(x) := \{k \in \mathcal{K} : G_i(x) \neq 0, H_k(x) = 0\},$$
$$I_{GH}(x) := \{k \in \mathcal{K} : G_i(x) = 0, H_k(x) = 0\},\ I_g(x) := \{i \in \mathcal{I} : g_i(x) = 0\}.$$

If there is no ambiguity, we abbreviate the above sets to $I_{D0}, I_{D\neq}, I_G, I_H, I_{GH}, I_g$.

Note that the index sets I_G, I_H, I_{GH} are a partition of $\mathcal{K}$, i.e., $I_G \cup I_H \cup I_{GH} = \mathcal{K}$, $I_G \cap I_H \cap I_{GH} = \emptyset$. In [13] (Lemma 4.1), we know that MFCQ and LICQ are violated at $\bar{x}$ when $I_{GH} \neq 0$.

Denote $\mathcal{C} := \{(a, b) \in \mathbb{R}^2 : ab = 0\}$, for any $(a, b) \in \mathcal{C}$, the following formulas are valid [13] (Lemma 3.2):

$$\widehat{N}_{\mathcal{C}}(a,b) = \begin{cases} \mathbb{R} \times \{0\} & \text{if } a = 0, b \neq 0, \\ \{0\} \times \mathbb{R} & \text{if } a \neq 0, b = 0, \\ \{(0,0)\} & \text{if } a = 0, b = 0, \end{cases} \quad N_{\mathcal{C}}(a,b) = \begin{cases} \mathbb{R} \times \{0\} & \text{if } a = 0, b \neq 0, \\ \{0\} \times \mathbb{R} & \text{if } a \neq 0, b = 0, \\ \mathcal{C} & \text{if } a = 0, b = 0. \end{cases} \tag{2}$$

Next, we would like to prove that the BQ fails for problem (1) at any nonzero feasible point, and the following lemma is crucial for the proof.

Lemma 1. *Let $x \in \Theta$. Then for any $(\mu, \nu, \alpha, \beta) \in \mathbb{R}^{m+p+2l}$ with $\mu_i = 0$, $i \in \mathcal{I} \setminus I_g$, $\alpha_k = 0$, $k \in I_H \cup I_{GH}$, and $\beta_k = 0$, $k \in I_G \cup I_{GH}$, we have that*

$$\mu^T \nabla g(x) + \nu^T \nabla h(x) + \alpha^T \nabla G(x) + \beta^T \nabla H(x) \in N_\Theta(x).$$

Proof. For the normal cone, define sets $\Omega := \{x : g_i(x) \leq 0, i \in \mathcal{I}, h_j(x) = 0, j \in \mathcal{J}\}$ and $\mathcal{C} := \{(a, b) \in \mathbb{R}^2 : ab = 0\}$, then $\Theta = \{x \in \Omega : (G(x), H(x)) \in \mathcal{C}\}$. From [28] (theorem 6.14) and the definition of normal cone, we have

$$\alpha^T \nabla G(x) + \beta^T \nabla H(x) + z \in \widehat{N}_\Theta(x) \subseteq N_\Theta(x),$$

where $(\alpha, \beta) \in \widehat{N}_{\mathcal{C}}(G(x), h(x))$, $z \in \widehat{N}_\Omega(x)$. According to $\widehat{N}_{\mathcal{C}}(a, b)$ in Equation (2), then

$$\widehat{N}_{\mathcal{C}}(G(x), H(x)) = \begin{cases} & \alpha_k \in \mathbb{R}, \beta_k = 0, \text{ if } k \in I_G, \\ (\alpha_k, \beta_k) : & \alpha_k = 0, \beta_k \in \mathbb{R}, \text{ if } k \in I_H, \\ & \alpha_k = 0, \beta_k = 0, \text{ if } k \in I_{GH}. \end{cases}$$

We let $z = \sum_{i \in I_g} \mu_i \nabla g_i(x) + \sum_{j \in \mathcal{J}} \nu_j \nabla h_j(x)$ with $\mu_i \geq 0, i \in I_g$. Then the result follows immediately form the above formulas. □

Proposition 1. *Let $\bar{x} \in \Theta$ and $I_{D0} \neq \emptyset$. Assume that D is full row rank. If there exists a vector $(\lambda, \mu, \nu, \alpha, \beta) \in \mathbb{R}^{|I_{D0}|+m+p+2l}$ with $\lambda \neq 0$ such that*

$$\begin{aligned} &\sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \nu_j \nabla h_j(\bar{x}) + \sum_{k=l}^{m} \alpha_k \nabla G_k(\bar{x}) + \sum_{k=1}^{l} \beta_k \nabla H_k(\bar{x}) + \sum_{i \in I_{D0}} \lambda_i D_i = 0, \\ &\mu_i = 0\ i \in \mathcal{I} \setminus I_g, \alpha_k = 0\ k \in I_H \cup I_{GH}, \beta_k = 0\ k \in I_G \cup I_{GH}, \end{aligned} \tag{3}$$

then the BQ fails at $\bar{x}$ for problem (1).

Proof. From Lemma 1 where for all $(\mu, \nu, \alpha, \beta)$ with $\mu_i = 0, i \in \mathcal{I} \setminus I_g$, $\alpha_k = 0, k \in I_H \cup I_{GH}$, and $\beta_k = 0, k \in I_G \cup I_{GH}$, there is

$$\sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \nu_j \nabla h_j(\bar{x}) + \sum_{k=l}^{m} \alpha_k \nabla G_k(\bar{x}) + \sum_{k=1}^{l} \beta_k \nabla H_k(\bar{x}) \in N_\Theta(\bar{x}).$$

Combine with the assumption Equation (3) with $\lambda \neq 0$ such that

$$-\sum_{i \in I_{D0}} \lambda_i D_i \in N_\Theta(\bar{x}). \tag{4}$$

It is clear that $\sum_{i \in I_{D0}} \lambda_i D_i \neq 0$ since D is full row rank. Now let $\phi(\cdot) = \|\cdot\|_p^p$ and $S(x) = Dx$ which imply $\|D\bar{x}\|_p^p = \phi(S(\bar{x}))$. By [28] (Exercise 10.7), we have

$$\partial^\infty(\|D\bar{x}\|_p^p) = \nabla S(x) * \partial^\infty \phi(S(\bar{x})),$$

where

$$\partial^\infty \phi(S(\bar{x})) = \{(a, b) \in \mathbb{R}^n : a_i = 0 \text{ if } S_i(\bar{x}) > 0, b = 0\},$$

by [27] (euqation (2.2)). Moreover, by [28] (Exercise 8.8(c)), it follows that

$$\partial^\infty F(\bar{x}) = \partial^\infty(\|D\bar{x}\|_p^p) = \left\{ \sum_{i \in I_{D0}} a_i D_i : a_i \in \mathbb{R}\ i \in I_{D0} \right\}. \tag{5}$$

According to Equations (4) and (5),

$$-\partial^\infty F(\bar{x}) \cap N_\Theta(\bar{x}) = -\sum_{i \in I_{D0}} \lambda_i D_i \neq 0.$$

Thus the BQ fails at point $\bar{x}$ and the proof is complete. □

3.1.1. The Stationary Condition

We give four stationary conditions for problem (1). However, first we introduce a tangent cone and a sign function. The MPSC-tailored linearized tangent cone of problem (1) is defined as

$$\mathcal{L}_{MPSC}(\bar{x}) = \left\{ d \;\middle|\; \begin{array}{ll} \nabla g_i(\bar{x})^T d \leq 0, & i \in I_g, \\ \nabla h_j(\bar{x})^T d = 0, & j \in \mathcal{J}, \\ \nabla G_k(\bar{x})^T d = 0, & k \in I_G, \\ \nabla H_k(\bar{x})^T d = 0, & k \in I_H, \\ (\nabla G_k(\bar{x})^T d)(\nabla H(\bar{x})^T d) = 0, & k \in I_{GH}. \end{array} \right\}, \tag{6}$$

and for any $\theta \in \mathbb{R}$, the sign function of θ is denoted as

$$sign\theta = \begin{cases} 1 & \text{if } \theta > 0, \\ [-1, 1] & \text{if } \theta = 0, \\ -1 & \text{if } \theta < 0. \end{cases} \tag{7}$$

Definition 7. *Let $\bar{x} \in \mathbb{R}^n$ be a feasible point of problem (1). Then, $\bar{x}$ is called*

- *weak stationary (W-stationary) point if there exist multipliers $(\lambda, \mu, \nu, \alpha, \beta)$ such that*

$$\begin{aligned} &\nabla f(\bar{x}) + \sum_{r=1}^{s} \lambda_r D_r + \sum_{i \in I_g} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x}) + \sum_{k \in \mathcal{K}} [\alpha_k \nabla G_k(\bar{x}) + \beta_k \nabla H_k(\bar{x})] = 0, \\ &\lambda_i = p|D_r^T \bar{x}|^{p-1} sign(D_r^T \bar{x}), r \in I_{D\neq}, \mu_i \geq 0, \forall i \in I_g, \alpha_k = 0, \forall k \in I_H, \beta_k = 0, \forall k \in I_G. \end{aligned} \tag{8}$$

- *Mordukhovich-stationary (M-stationary) point if there exist multipliers* $(\lambda, \mu, \nu, \alpha, \beta)$ *such that Equation* (8) *holds and satisfies the condition:*

$$\alpha_k \beta_k = 0, \forall k \in I_{GH}. \tag{9}$$

- *Bouligand-stationay (B-stationary) point if for* $\zeta \in \partial F(\bar{x})$*, the following condition holds:*

$$\zeta^T d \geq 0, \ \forall d \in \mathcal{L}_{MPSC}(\bar{x}). \tag{10}$$

- *strongly stationary (S-stationary) point if there exist multipliers* $(\lambda, \mu, \nu, \alpha, \beta)$ *such that Equation* (8) *holds and satisfies the condition:*

$$\alpha_k = 0, \beta_k = 0, \forall k \in I_{GH}. \tag{11}$$

Theorem 1. *Let* $\bar{x} \in \Theta$*. (i) If* $\bar{x}$ *is an S-stationary point of problem* (1)*, then* $\bar{x}$ *is a B-stationary point of problem* (1)*. (ii) If* $\bar{x}$ *is a B-stationary point of problem* (1)*, then* $\bar{x}$ *is an M-stationary point of problem* (1)*.*

Proof. (i) Let $\bar{x}$ be an S-stationary point of problem (1); there exist multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$ satisfying Equations (8) and (11). Then for any $d \in \mathbb{R}^n$, denote $\nabla F(\bar{x}) = \nabla f(\bar{x}) + \sum_{r=1}^{s} \lambda_r D_r$; we have

$$\nabla F(\bar{x})^T d + \sum_{i \in I_g} \mu_i \nabla g_i(\bar{x})^T d + \sum_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x})^T d + \sum_{k \in \mathcal{K}} [\alpha_k \nabla G_k(\bar{x})^T d + \beta_k \nabla H_k(\bar{x})^T d] = 0.$$

By contradiction, if $\bar{x}$ is not a B-stationary point of problem (1), which means that there exist $d^k \in \mathcal{L}_{MPSC}(\bar{x})$ such that $\nabla F(\bar{x})^T d^k < 0$. From Equations (6) and (11), we get

$$\nabla F(\bar{x})^T d^k + \sum_{i \in I_g} \mu_i \nabla g_i(\bar{x})^T d^k + \sum_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x})^T d^k + \sum_{k \in \mathcal{K}} [\alpha_k \nabla G_k(\bar{x})^T d^k + \beta_k \nabla H_k(\bar{x})^T d^k] < 0,$$

which is a contraction.

(ii) Let $\bar{x}$ be a B-stationary point of problem (1). For $d \in \mathcal{L}_{MPSC}(\bar{x})$, we have $\nabla F(\bar{x})^T d \geq 0$, which means that $d = 0$ is an optimal solution of

$$\min_{d \in \mathcal{L}_{MPSC}(\bar{x})} \nabla F(\bar{x})^T d. \tag{12}$$

In fact, problem (12) is an MPSC concerning which all constraints are affine. By [13] (Corollary 5.2, Theorem 5.1), MPSC-ACQ is valid for problem (12) at $d = 0$ which is an M-stationary point of problem (12). There exist multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$ such that

$$\nabla F(\bar{x}) + \sum_{i \in I_g} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x}) + \sum_{k \in \mathcal{K}} [\alpha_k \nabla G_k(\bar{x}) + \beta_k \nabla H_k(\bar{x})] = 0,$$
$$\mu_i \geq 0, \forall i \in I_g, \quad \alpha_k = 0, \forall k \in I_H, \quad \beta_k = 0, \forall k \in I_G, \quad \alpha_k \beta_k = 0, \forall k \in I_{GH},$$

implying that $\bar{x}$ is an M-stationary point of problem (1). □

By Theorem 1 and Definition 7, the following relations hold easily

$$\text{S-stationary} \implies \text{B-stationary} \implies \text{M-stationary} \implies \text{W-stationary}.$$

3.1.2. The Qualifications

Notice that $\|Dx\|_p^p$ is non-convex and non-Lipschitz when $0 < p < 1$, we need to consider the appropriate subdifferentiable and constraint qualification to ensure that a minimizer x^* of problem (1) is a stationary point.

Let $\mathcal{P}(I_{GH}(\bar{x}))$ be the set of all the disjoint bipartitions of $I_{GH}(\bar{x})$. For fixed $(\eta_1, \eta_2) \in \mathcal{P}(I_{GH}(\bar{x}))$, we reformulate problem (1) as the following

$$\begin{array}{ll} \min\limits_{x} & \varphi(x) = f(x) + \sum\limits_{r \in I_{D\neq}} |D_r^T x|^p \\ s.t. & D_r^T x = 0,\ r \in I_{D0}, \\ & g_i(x) \leq 0,\ i \in \mathcal{I}, \\ & h_j(x) = 0,\ j \in \mathcal{J}, \\ & G_k(x) = 0,\ k \in I_G \cup \eta_1, \\ & H_k(x) = 0,\ k \in I_H \cup \eta_2. \end{array} \tag{13}$$

Denote Ξ as the feasible set of problem (13).

Lemma 2. *The point $\bar{x} \in \Theta$ is a local minimizer of problem (1) if and only if there exists a partition $(\eta_1, \eta_2) \in \mathcal{P}(I_{GH}(\bar{x}))$ such that $\bar{x}$ is a local minimizer of problem (13).*

Proof. Assume that $\bar{x} \in \Theta$ is a local minimizer of problem (1). If the set $I_{GH}(\bar{x}) = \emptyset$, problem (13) is equivalent to problem (1), and the lemma is clearly true. Otherwise, there exists $(\eta_1, \eta_2) \in \mathcal{P}(I_{GH}(\bar{x}))$, but the point $\bar{x}$ is not a local minimizer of problem (13). Hence there exists $x^* \in \Xi$ such that $\varphi(x^*) \leq \varphi(\bar{x})$. By the feasibility of x^*, we know that $x^* \in \Theta$ and $D_r^T x^* = 0,\ r \in I_{D0}(x^*)$. Thus, $F(x^*) \leq F(\bar{x})$, which is a contradiction of $\bar{x}$, a local minimizer of problem (1).

Conversely, $\bar{x}$ is a local minimizer of problem (13). We assume that $\bar{x}$ is not a local minimizer of problem (1); there exists a point $x^* \in \Theta$ such that $F(x^*) \leq F(\bar{x})$. By the feasibility of x^*, we know $G_k(x^*)H_k(x^*) = 0, k \in \mathcal{K}$, which implies $G_k(x^*) = 0, k \in I_G(x^*) \cup I_{GH}(x^*)$ or $H_k(x^*) = 0, k \in I_H(x^*) \cup I_{GH}(x^*)$. Let $z^* = Dx^*$, and there exists a partition $(\eta_1, \eta_2) \in \mathcal{P}(I_{GH}(\bar{x}))$ such that x^* is a feasible point of problem (13) and $\varphi(x^*) \leq \varphi(\bar{x})$, which is a contradiction of $\bar{x}$, a local minimizer of problem (13). The proof is complete. □

Here, we give some qualifications of problem (1).

Definition 8. *Let $\bar{x} \in \Theta$. We say*

(a) MPSC linear independence (MPSC-LI) qualification holds at $\bar{x}$ for problem (1) if the following gradients are linearly indepedent:

$$\{D_r, \nabla g_i(\bar{x}), \nabla h_j(\bar{x}), \nabla G_k(\bar{x}), \nabla H_{k'}(\bar{x}) \mid r \in I_{D0}, i \in I_g, j \in \mathcal{J}, k \in I_G \cup I_{GH}, k' \in I_H \cup I_{GH}\}.$$

(b) MPSC linear constraint qualification (MPSC-LCQ) holds if all functions $\{g, h, G, H\}$ in the constraints are affine.

(c) MPSC Mangasarian-Fromovitz (MPSC-MF) qualification holds at $\bar{x}$ iff there are no nonzero multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$ such that

$$\begin{cases} \sum\limits_{r \in I_{D0}} \lambda_r D_r + \sum\limits_{i \in I_g} \mu_i \nabla g_i(\bar{x}) + \sum\limits_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x}) + \sum\limits_{k \in \mathcal{K}} (\alpha_k \nabla G_k(\bar{x}) + \beta_k \nabla H_k(\bar{x})) = 0, \\ \mu_i \geq 0, i \in I_g, \quad \mu^T g(\bar{x}) = 0, \quad \alpha_k = 0, k \in I_H, \quad \beta_k = 0, k \in I_G. \end{cases}$$

(d) MPSC no nonzero abnormal multiplier (MPSC-NNAM) qualification holds at $\bar{x}$ iff there are no nonzero multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$ such that

$$\begin{cases} \sum\limits_{r \in I_{D0}} \lambda_r D_r + \sum\limits_{i \in I_g} \mu_i \nabla g_i(\bar{x}) + \sum\limits_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x}) + \sum\limits_{k \in \mathcal{K}} (\alpha_k \nabla G_k(\bar{x}) + \beta_k \nabla H_k(\bar{x})) = 0, \\ \mu_i \geq 0, i \in I_g, \quad \mu^T g(\bar{x}) = 0, \quad \alpha_k = 0, k \in I_H, \quad \beta_k = 0, k \in I_G, \quad \alpha_k \beta_k = 0, i \in I_{GH}. \end{cases}$$

(e) *MPSC constant rank (MPSC-CR) qualification holds at $\bar{x}$ if there exists $\delta > 0$ such that, for any $I_1 \subseteq I_{D0}$, $I_2 \subseteq I_g$, $I_3 \subseteq \mathcal{J}$, $I_4 \subseteq I_G \cup I_{GH}$, and $I_5 \subseteq I_H \cup I_{GH}$, the gradients*

$$\{D_r, \nabla g_i(x), \nabla h_j(x), \nabla G_k(x), \nabla H_{k'}(x) : r \in I_1, i \in I_2, j \in I_3, k \in I_4, k' \in I_5\}$$

has the same rank for each $x \in \mathcal{B}_\delta(\bar{x}) := \{x : \|x - \bar{x}\| \leq \delta\}$.

(f) *MPSC relaxed constant rank (MPSC-RCR) qualification holds at $\bar{x}$ if there exists $\delta > 0$ such that, for any $I_1 \subseteq I_{D0}$, $I_2 \subseteq I_g$, $I_3, I_4 \subseteq I_{GH}$, the gradients*

$$\{D_r, \nabla g_i(x), \nabla h_j(x), \nabla G_k(x), \nabla H_{k'}(x) : r \in I_1, i \in I_2, j \in \mathcal{J}, k \in I_G \cup I_3, k' \in I_H \cup I_4\}$$

have the same rank for each $x \in \mathcal{B}_\delta(\bar{x})$.

(g) *MPSC constant positive linear dependent (MPSC-CPLD) holds at $\bar{x}$ if, for any $I_1 \subseteq I_{D0}$, $I_2 \subseteq I_g$, $I_3 \subseteq \mathcal{J}$, $I_4 \subseteq I_G \cup I_{GH}$, and $I_5 \subseteq I_H \cup I_{GH}$, whenever there exist multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$, not all zero, with $\mu_i \geq 0$ for each $i \in I_2$ and $\alpha_i \beta_i = 0$ for each $i \in I_{GH}$, such that*

$$\sum_{r \in I_1} \lambda_r D_r + \sum_{i \in I_2} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in I_3} \nu_j \nabla h_j(\bar{x}) + \sum_{k \in I_4} \alpha_k \nabla G_k(\bar{x}) + \sum_{k' \in I_5} \beta_{k'} \nabla H_{k'}(\bar{x}) = 0,$$

there exists $\delta > 0$ such that, for any $x \in B_\delta(\bar{x})$, the vectors

$$\{D_r, \nabla g_i(x), \nabla h_j(x), \nabla G_k(x), \nabla H_{k'}(x) : r \in I_1, i \in I_2, j \in I_3, k \in I_4, k' \in I_5\}$$

are linearly dependent.

(h) *Let $I_1 \subseteq I_{D0}$, $I_2 \subseteq \mathcal{J}$, $I_3 \subseteq I_G$, and $I_4 \subseteq I_H$, be such that $\mathcal{G}(\bar{x}, I_1, I_2, I_3, I_4)$ is a basis for span$\mathcal{G}(\bar{x}, I_{D0}, \mathcal{J}, I_G, I_H)$. MPSC relaxed constant positive linear dependence (MPSC-RCPLD) holds at $\bar{x}$ if there exists $\delta > 0$ such that*

- *$\mathcal{G}(x, \mathcal{J}, I_{D0}, I_G, I_H)$ has the same rank for $x \in B_\delta(\bar{x})$;*
- *for any $I_5 \subseteq I_g$, $I_6, I_7 \in I_{GH}$, if there exist multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$, not all zero, with $\mu_i \geq 0$ for each $i \in I_5$ and $\alpha_i \beta_i = 0$ for each $i \in I_{GH}$, such that*

$$\sum_{r \in I_1} \lambda_r D_r + \sum_{i \in I_5} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in I_2} \nu_j \nabla h_j(\bar{x}) + \sum_{k \in I_3 \cup I_6} \alpha_k \nabla G_k(\bar{x}) + \sum_{k' \in I_4 \cup I_7} \beta_{k'} \nabla H_{k'}(\bar{x}) = 0,$$

then for any $x \in B_\delta(\bar{x})$, the following vectors are linearly dependent:

$$\{D_r, \nabla g_i(x), \nabla h_j(x), \nabla G_k(x), \nabla H_{k'}(x) : r \in I_1, i \in I_5, j \in I_2, k \in I_3 \cup I_6, k' \in I_4 \cup I_7\},$$

where $\mathcal{G}(\bar{x}, I_1, I_2, I_3, I_4) := \{D_r, \nabla h_j(x), \nabla G_k(x), \nabla H_{k'}(x) : r \in I_1, j \in I_2, k \in I_3, k' \in I_4\}$.

(i) *MPSC pseudonormality holds at $\bar{x}$ if there are no nonzero multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$ such that*

- $\sum_{r \in I_{D0}} \lambda_r D_r + \sum_{i \in I_g} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x}) + \sum_{k \in \mathcal{K}} (\alpha_k \nabla G_k(\bar{x}) + \beta_k \nabla H_k(\bar{x})) = 0$;
- $\mu_i \geq 0, \mu^T g(\bar{x}) = 0, \alpha_k = 0, k \in I_H, \beta_k = 0, k \in I_G$, *and* $\alpha_k \beta_k = 0$ *for* $k \in I_{GH}$;
- *there exists a sequence $\{x^k\} \to \bar{x}$ such that, for each k*

$$\sum_{r=1}^{s} \lambda_r D_r + \sum_{i=1}^{m} \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \nu_j \nabla h_j(\bar{x}) + \sum_{k=1}^{l} \alpha_k \nabla G_k(\bar{x}) + \sum_{k'=1}^{l} \mu_{k'} \nabla H_{k'}(\bar{x}) > 0.$$

(j) *MPSC quasinormality holds at $\bar{x}$ if there are no nonzero multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$ such that*

- $\sum_{r \in I_{D0}} \lambda_r D_r + \sum_{i \in I_g} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{J}} \nu_j \nabla h_j(\bar{x}) + \sum_{k \in \mathcal{K}} (\alpha_k \nabla G_k(\bar{x}) + \beta_k \nabla H_k(\bar{x})) = 0$;
- $\mu_i \geq 0, \mu^T g(\bar{x}) = 0, \alpha_k = 0, k \in I_H, \beta_k = 0, k \in I_G$, *and* $\alpha_k \beta_k = 0$ *for* $k \in I_{GH}$;

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

$$\begin{aligned} &\lambda_r > 0 \Longrightarrow \lambda_r(D_r^T x^k) > 0,\ \mu_i > 0 \Longrightarrow \mu_i g_i(x^k) > 0,\ \nu_j > 0 \Longrightarrow \nu_j h_j(x^k) > 0, \\ &\alpha_k > 0 \Longrightarrow \alpha_k G_k(x^k) > 0,\ \beta_k > 0 \Longrightarrow \beta_k H_k(x^k) > 0. \end{aligned} \tag{14}$$

(k) MPSC Abadie qualification holds at $\bar{x}$ *if* $T_\Xi(\bar{x}) = \mathcal{L}_\Xi^{MPSC}(\bar{x})$, *where*

$$\mathcal{L}_\Xi^{MPSC}(\bar{x}) = \left\{ d \;\middle|\; \begin{array}{ll} \nabla g_i(\bar{x})^T d \le 0, & i \in I_g, \\ D_r^T d = 0, & r \in I_{D0}, \\ \nabla h_j(\bar{x})^T d = 0, & j \in \mathcal{J}, \\ \nabla G_k(\bar{x})^T d = 0, & k \in I_G, \\ \nabla H_k(\bar{x})^T d = 0, & k \in I_H, \\ (\nabla G_k(\bar{x})^T d)(\nabla H_(\bar{x})^T d) = 0, & k \in I_{GH}. \end{array} \right\} \tag{15}$$

(l) MPSC Guignard qualification holds at $\bar{x}$ *if* $T_\Xi(\bar{x})^\circ = \mathcal{L}_\Xi^{MPSC}(\bar{x})^\circ$.

In the following, we show that $\bar{x}$ is an S-stationary point of problem (1) when MPSC-LI qualification holds at $\bar{x}$, and is an M-stationary point of problem (1) when MPSC-RCLPD holds at $\bar{x}$.

Theorem 2. *Let* $\bar{x} \in \Theta$ *be a locally optimal solution of problem* (1). *If MPSC-LI qualification holds at* $\bar{x}$, *then* $\bar{x}$ *is an S-stationary point of problem* (1).

Proof. Assume $\bar{x} \in \Theta$ is a locally optimal solution of problem (1) and by Lemma 2, $\bar{x}$ is a local minimizer of problem (13). Note that $D_i \bar{x} \neq 0$ for all $i \in I_{D\neq}$, it is easy to see that the objective function of problem (13) is continuously differentiable at $\bar{x}$. By the assumption that MPSC-LI qualification holds at point $\bar{x}$, i.e.,

$$\{D_r, \nabla g_i(\bar{x}), \nabla h_j(\bar{x}), \nabla G_k(\bar{x}), \nabla H_{k'}(\bar{x}) : r \in I_{D0}, i \in I_g, j \in \mathcal{J}, k \in I_G \cup I_{GH}, k' \in I_H \cup I_{GH}\},$$

which implies that LICQ holds at point $\bar{x}$ for problem (13). Then for $(\eta_1, \eta_2) \in \{(I_{GH}, \emptyset)\}$, there exist multipliers $\{\lambda^1, \mu^1, \nu^1, \alpha^1, \beta^1\}$ satisfying

$$\begin{aligned} &\nabla f(\bar{x}) + \sum_{r \in I_{D\neq}} \rho_r D_r + \sum_{r \in I_{D0}} \lambda_r^1 D_r + \sum_{i \in I_g} \mu_i^1 \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{J}} \nu_j^1 \nabla h_j(\bar{x}) \\ &\quad + \sum_{k \in I_G \cup I_{GH}} \alpha_k^1 \nabla G_k(\bar{x}) + \sum_{k \in I_H} \beta_k^1 \nabla H_k(\bar{x}) = 0, \\ &\rho_r = p|D_r \bar{x}|^{p-1} sign(D_r \bar{x}), r \in I_{D\neq}, \quad \mu_i \ge 0, \forall i \in I_g, \end{aligned} \tag{16}$$

and for $(\eta_1, \eta_2) \in \{(\emptyset, I_{GH})\}$, there exist multipliers $\{\lambda^2, \mu^2, \nu^2, \alpha^2, \beta^2\}$ satisfying

$$\begin{aligned} &\nabla f(\bar{x}) + \sum_{r \in I_{D\neq}} \rho_r D_r + \sum_{r \in I_{D0}} \lambda_r^2 D_r + \sum_{i \in I_g} \mu_i^2 \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{J}} \nu_j^2 \nabla h_j(\bar{x}) \\ &\quad + \sum_{k \in I_G} \alpha_k^2 \nabla G_k(\bar{x}) + \sum_{k \in I_H \cup I_{GH}} \beta_k^2 \nabla H_k(\bar{x}) = 0, \\ &\rho_r = p|D_r \bar{x}|^{p-1} sign(D_r \bar{x}), r \in I_{D\neq}, \quad \mu_i \ge 0, \forall i \in I_g. \end{aligned} \tag{17}$$

Let Equations (16) and (17), we have

$$\begin{aligned} &\sum_{r \in I_{D0}} (\lambda_r^1 - \lambda_r^2) D_r + \sum_{i \in I_g} (\mu_i^1 - \mu_i^2) \nabla g_i(\bar{x}) + \sum_{j \in \mathcal{J}} (\nu_j^1 - \nu_j^2) \nabla h_j(\bar{x}) + \sum_{k \in I_G} (\alpha_k^1 - \alpha_k^2) \nabla G_k(\bar{x}) \\ &\quad + \sum_{k \in I_H} (\beta_k^1 - \beta_k^2) \nabla H_k(\bar{x}) + \sum_{k \in I_{GH}} \alpha_k^1 \nabla G_k(\bar{x}) - \sum_{k \in I_{GH}} \beta_k^2 \nabla H_k(\bar{x}) = 0. \end{aligned}$$

Due to the validity of MPSC-LI qualification of problem (13), $\lambda_r^1 = \lambda_r^2, r \in I_{D0}$, $\mu_i^1 = \mu_i^2, i \in I_g$, $\nu_j^1 = \nu_j^2, j \in \mathcal{J}$, $\alpha_k^1 = \alpha_k^2, k \in I_G$, $\beta_k^1 = \beta_k^2, k \in I_H$, and $\alpha_k^1 = \beta_k^2 = 0, k \in I_{GH}$ are obtained. Furthermore, $I_{D0} \cup I_{D\neq} = \{1, \cdots, s\}$ and $I_G \cup I_H \cup I_{GH} = \mathcal{K}$; hence, we can write Equation (16) or Equation (17) as Equations (8) and (11) and the proof is complete. □

Before we prove the following theorem, we give two lemmas in [28,29].

Lemma 3. [28] *A vector $d \in T_A(\bar{x})^\circ$ iff there exists a smooth function ϕ such that $-\nabla\phi(\bar{x}) = d$ and $\arg\min_{x\in A} \phi(x) = \{\bar{x}\}$.*

Lemma 4. [29] *Let $x = \sum_{i=1}^{m+p} \lambda_i v_i$ where $\{v_1, \cdots, v_m\}$ is linearly independent and $\lambda_i \neq 0$ for each $i = m+1, \cdots, m+p$. Then there exist $\mathcal{P} \subseteq \{m+1, \cdots, m+p\}$ and $\bar{\lambda}_i, i \in \{1, \cdots, m\} \cup \mathcal{P}$, such that $x = \sum_{i\in\{1,\cdots,m\}\cup\mathcal{P}} \bar{\lambda}_i v_i$ with $\lambda_i\bar{\lambda}_i > 0$ for each $i \in \mathcal{P}$ and $\{v_i\}_{i\in\{1,\cdots,m\}\cup\mathcal{P}}$ is linearly independent.*

Theorem 3. *Let $\bar{x} \in \Theta$ be a locally optimal solution of problem (1). If MPSC-RCLPD holds at $\bar{x}$, then $\bar{x}$ is an M-stationary of problem (1).*

Proof. Assume $\bar{x} \in \Theta$ is a locally optimal solution of problem (1) and by Lemma 2, $\bar{x}$ is a local minimizer of problem (13). Let $\bar{y} := G(\bar{x})$, $\bar{z} := H(\bar{x})$. For each $k \in \mathbb{N}$, we consider the following optimization problem

$$\begin{aligned} \min\ & F_k(x,y,z) := \varphi(x) + \frac{k}{2}\|\max\{0, g(x)\}\|^2 + \frac{k}{2}\|D_{I_{D0}}^T x\|^2 + \frac{k}{2}\|h(x)\|^2 \\ & + \frac{k}{2}\|G(x) - y\|^2 + \frac{k}{2}\|H(x) - z\|^2 + \frac{1}{2}\|(x,y,z) - (\bar{x},\bar{y},\bar{z})\|^2 \\ \text{s.t. } & (x,y,z) \in U \cap \Lambda, \end{aligned} \tag{18}$$

where $U := \mathbb{R}^n \times \mathcal{C}$ and $\Lambda := \{(x,y,z) : \|(x,y,z) - (\bar{x},\bar{y},\bar{z})\| \leq \varepsilon\}$. Since the set $U \cap \Lambda$ is compact and $F_k(x,y,z)$ is continuous, problem (18) has at least one optimal solution, denote as (x^k, y^k, z^k). Next, we show that $(x^k, y^k, z^k) \to (\bar{x}, \bar{y}, \bar{z})$ as $k \to \infty$. By the optimality at point (x^k, y^k, z^k), we have $F_k(x^k, y^k, z^k) \leq F_k(\bar{x}, \bar{y}, \bar{z}) = \varphi(\bar{x})$ for all $k \in \mathbb{N}$, that is

$$\begin{aligned} & \varphi(x^k) + \frac{k}{2}\|\max\{0, g(x^k)\}\|^2 + \frac{k}{2}\|D_{I_{D0}}^T x^k\|^2 + \frac{k}{2}\|h(x^k)\|^2 + \frac{k}{2}\|G(x^k) - y^k\|^2 \\ & + \frac{k}{2}\|H(x^k) - z^k\|^2 + \frac{1}{2}\|(x^k, y^k, z^k) - (\bar{x},\bar{y},\bar{z})\|^2 \leq \varphi(\bar{x}). \end{aligned} \tag{19}$$

By the compactness of $U \cap \Lambda$, $\varphi(\bar{x})$ is bounded, which implies that

$$\begin{aligned} & \lim_{k\to\infty} \max\{0, g(x^k)\} = 0, \quad \lim_{k\to\infty} D_{I_{D0}}^T x^k = 0, \quad \lim_{k\to\infty} h(x^k) = 0, \\ & \lim_{k\to\infty} G(x^k) - y^k = 0, \quad \lim_{k\to\infty} H(x^k) - z^k = 0. \end{aligned} \tag{20}$$

Assume (x^*, y^*, z^*) is an accumulation point of $\{(x^k, y^k, z^k)\}$. From Equations (19) and (20), we have

$$\begin{aligned} & \varphi(x^*) + \frac{1}{2}\|(x^*, y^*, z^*) - (\bar{x},\bar{y},\bar{z})\|^2 \leq \varphi(\bar{x}), \\ & g(x^*) \leq 0, D_{I_{D0}}^T x^* = 0, h(x^*) = 0, G(x^*) = y^*, H(x^*) = z^*. \end{aligned}$$

Note that $\varphi(\bar{x}) \leq \varphi(x^*)$ by the feasibility of x^* for problem (13) and local optimality of $\bar{x}$; thus the whole sequence $\{x^k, y^k, z^k\}$ converges to $(\bar{x}, \bar{y}, \bar{z})$.

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

$$-\nabla F_k(x^k, y^k, z^k) \in \widehat{N}_U(x^k, y^k, z^k),$$

where

$$\widehat{N}_U(x^k, y^k, z^k) = \begin{cases} a_i \in \mathbb{R}, b_i = 0 \text{ if } y_i = 0, z_i \neq 0, \\ (0, a, b) : a_i = 0, b_i \in \mathbb{R} \text{ if } y_i \neq 0, z_i = 0, \\ a_i = 0, b_i = 0 \text{ if } y_i = 0, z_i = 0, \end{cases} \tag{21}$$

by Equation (2). On the basis of Equation (21), we obtain that

$$\begin{aligned} &\nabla\varphi(x^k) + \sum_{r\in I_{D0}} \lambda_r^k D_r + \sum_{i=1}^{m} \mu_i^k \nabla g_i(x^k) + \sum_{j=1}^{p} \nu_j^k \nabla h_j(x^k) + (x^k - \bar{x}) \\ &+ \sum_{i\in I_G \cup I_{GH}} \alpha_i^k \nabla G_i(x^k) + \sum_{i\in I_H \cup I_{GH}} \beta_i^k \nabla H_i(x^k) = 0 \end{aligned} \tag{22}$$

where $\lambda_r^k = k(D_r x^k), r \in I_{D0}$, $\mu_i^k = k\max\{0, g_i(x^k)\}, i \in \mathcal{I}$, $\nu_j^k = kh_j(x^k), j \in \mathcal{J}$, $\alpha_t^k = k(G_t(x^k) - y_t^k), t \in I_G \cup I_{GH}$, $\beta_t^k = k(H_t(x^k) - z_i^k), t \in I_H \cup I_{GH}$, and

$$\begin{cases} y_i^k = 0, z_i^k \neq 0 \Rightarrow \beta_i^k = z_i^k - \bar{z}_i, \\ y_i^k \neq 0, z_i^k = 0 \Rightarrow \alpha_i^k = y_i^k - \bar{y}_i, \\ y_i^k = 0, z_i^k = 0 \Rightarrow \alpha_i^k = y_i^k - \bar{y}_i, \beta_i^k = z_i^k - \bar{z}_i. \end{cases} \tag{23}$$

From Equations (21) and (23), it is easy to obtain $(\alpha_i^k, \beta_i^k) \in N_{\mathcal{C}}(y_i^k, z_i^k)$. For convenience, we define the index sets as follows

$$\begin{aligned} &I_{D0}^k := \{i \in \{1, \cdots, s\} : D_i x^k = 0\}, \quad I_g^k := \{i \in \mathcal{I} : g_i(x^k) = 0\}, \\ &I_G^k := \{k \in \mathcal{K} : G_i(x^k) = 0, H_k(x^k) \neq 0\}, \quad I_H^k := \{k \in \mathcal{K} : G_i(x^k) \neq 0, H_k(x^k) = 0\}, \\ &I_{GH}^k := \{k \in \mathcal{K} : G_i(x^k) = 0, H_k(x^k) = 0\}, \end{aligned}$$

and it is obvious that $I_G \subseteq I_G^k$ and $I_H \subseteq I_H^k$ for each k sufficiently large. Define $supp(a) := \{i : a_i \neq 0\}$, then by Equations (22) and (23), we have

$$\begin{aligned} 0 = \Upsilon^k &+ \sum_{r\in I_{D0}} \lambda_r^k D_r + \sum_{i\in supp(\mu_k)} \mu_i^k \nabla g_i(x^k) + \sum_{j\in\mathcal{J}} \nu_j^k \nabla h_j(x^k) + \sum_{t\in I_G} a_t^k \nabla G_t(x^k) \\ &+ \sum_{t\in I_H} b_t^k \nabla H_t(x^k) + \sum_{t\in I_G^k \setminus I_G \cup I_{GH}^k \cap supp(a_t^k)} a_t^k \nabla G_t(x^k) + \sum_{t\in I_H^k \setminus I_H \cup I_{GH}^k \cap supp(b_t^k)} b_t^k \nabla H_t(x^k), \end{aligned}$$

where

$$\begin{aligned} &a_t^k = \alpha_t^k - (y_i^k - \bar{y}_i), b_t^k = \beta_t^k - (z_i^k - \bar{z}_i), \\ &\Upsilon^k = \nabla\varphi(x^k) + (x^k - \bar{x}) + \sum_{i\in I_G \cup I_{GH}} \alpha_i^k \nabla G_i(x^k) + \sum_{i\in I_H \cup I_{GH}} \beta_i^k \nabla H_i(x^k). \end{aligned}$$

Let $I_1 \subseteq I_{D0}$, $I_2 \subseteq \mathcal{J}$, $I_3 \subseteq I_G$, and $I_4 \subseteq I_H$ be index sets such that $\mathcal{G}(\bar{x}, I_1, I_2, I_3, I_4)$ is a basis for $\mathcal{G}(\bar{x}, I_{D0}, \mathcal{J}, I_G, I_H)$. Because $\bar{x}$ satisfies the MPSC-RCPLD, there exists a constant $\delta > 0$ such that the rank of $\mathcal{G}(\bar{x}, I_{D0}, \mathcal{J}, I_G, I_H)$ is constant for each $x \in B_\delta(\bar{x})$. Hence $\mathcal{G}(\bar{x}, I_1, I_2, I_3, I_4)$ is a basis for span$\mathcal{G}(\bar{x}, I_{D0}, \mathcal{J}, I_G, I_H)$ as k sufficiently large. From Lemma 4, there exist $I_5^k \subseteq supp(\mu^k)$, $I_6^k \subseteq I_G^k \setminus I_G \cup I_{GH}^k \cap supp(a_t^k)$, $I_7^k \subseteq I_H^k \setminus I_H \cup I_{GH}^k \cap supp(b_t^k)$ and $(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}^k, \bar{a}^k, \bar{b}^k)$ such that

$$\begin{aligned} 0 = &\Upsilon^k + \sum_{r\in I_1} \bar{\lambda}_r^k D_r + \sum_{j\in I_2} \bar{\nu}_j^k \nabla h_j(x^k) + \sum_{t\in I_3} \bar{a}_t^k \nabla G_t(x^k) + \sum_{t\in I_4} \bar{b}_t^k \nabla H_t(x^k) \\ &+ \sum_{i\in I_5^k} \bar{\mu}_i^k \nabla g_i(x^k) + \sum_{t\in I_6^k} \bar{a}_t^k \nabla G_t(x^k) + \sum_{t\in I_7^k} \bar{b}_t^k \nabla H_t(x^k), \end{aligned} \tag{24}$$

and the vectors $\{D_r, \nabla g_i(x), \nabla h_j(x), \nabla G_s(x), \nabla H_t(x) : r \in I_1, i \in I_5^k, j \in I_2, s \in I_3 \cup I_6^k, t \in I_4 \cup I_7^k\}$ are linearly independent for sufficiently large k. Set $\{\bar{\lambda}_r^k = \bar{\mu}_i^k = \bar{v}_j^k = \bar{a}_s^k = \bar{b}_t^k = 0 : r \notin I_1, i \notin I_5^k, j \notin I_2, s \notin I_3 \cup I_6^k, t \notin I_4 \cup I_7^k\}$, and by Lemma 4 follows that $\bar{\mu}_i^k \geq 0, i \in I_5^k$ and $(\bar{a}_i^k, \bar{b}_i^k) \in N_{\mathcal{C}}(y_i^k, z_i^k), i \in \mathcal{K}$. We assume that $I_5^k \equiv I_5$, $I_6^k \equiv I_6$, $I_7^k \equiv I_7$ for sufficiently large k; then the vectors

$$\{D_r, \nabla g_i(x), \nabla h_j(x), \nabla G_s(x), \nabla H_t(x) : r \in I_1, i \in I_5, j \in I_2, s \in I_3 \cup I_6, t \in I_4 \cup I_7\}$$

are linearly independent, and it is easy to obtain $I_5 \subseteq I_g$, $I_6, I_7 \subseteq I_{GH}$ with $I_G \cup I_H \cup I_{GH} = I_G^k \cup I_H^k \cup I_{GH}^k$. Denote

$$M_k := \max\{\bar{\lambda}_r^k, \bar{\mu}_i^k, \bar{v}_j^k, \bar{a}_s^k, \bar{b}_t^k : \ r \in I_1, i \in I_5, j \in I_2, s \in I_3 \cup I_6, t \in I_4 \cup I_7\}.$$

Now, we show that M_k is bounded. By contraction, assume $M_k \to \infty$; then there exists a subsequence such that for any $r \in I_1, i \in I_5, j \in I_2, s \in I_3 \cup I_6, t \in I_4 \cup I_7$,

$$\frac{(\bar{\lambda}_r^k, \bar{\mu}_i^k, \bar{v}_j^k, \bar{a}_s^k, \bar{b}_t^k)}{M_k} \to (\bar{\lambda}_r^*, \bar{\mu}_i^*, \bar{v}_j^*, \bar{a}_s^*, \bar{b}_t^*) \text{ as } k \to \infty.$$

Dividing Equation (24) by M_k, and taking the limit as $k \to \infty$, we have

$$\sum_{r \in I_1} \bar{\lambda}_r^* D_r + \sum_{j \in I_2} \bar{v}_j^* \nabla h_j(\bar{x}) + \sum_{t \in I_3} \bar{a}_t^* \nabla G_t(\bar{x}) + \sum_{t \in I_4} \bar{b}_t^* \nabla H_t(\bar{x}) + \sum_{i \in I_5} \bar{\mu}_i^* \nabla g_i(\bar{x}) + \sum_{t \in I_6} \bar{a}_t^* \nabla G_t(\bar{x}) + \sum_{t \in I_7} \bar{b}_t^* \nabla H_t(\bar{x}) = 0.$$

Furthermore, $\bar{\mu}_i^*, i \in I_5$, following that $(\bar{a}_i^*, \bar{b}_i^*) \in N_{\mathcal{C}}(\bar{y}_i, \bar{z}_i)$ by $(\bar{a}_i^k, \bar{b}_i^k) \in N_{\mathcal{C}}(y_i^k, z_i^k), i \in \mathcal{K}$ and the outer semi-continuity of limiting normal cone, which give a contradiction with MPSC-RCPLD. Therefore, M_k is bounded. There exists a subsequence such that for any $r \in I_1, i \in I_5, j \in I_2, s \in I_3 \cup I_6, t \in I_4 \cup I_7$,

$$\frac{(\bar{\lambda}_r^k, \bar{\mu}_i^k, \bar{v}_j^k, \bar{a}_s^k, \bar{b}_t^k)}{M_k} \to (\bar{\lambda}_r, \bar{\mu}_i, \bar{v}_j, \bar{a}_s, \bar{b}_t) \text{ as } k \to \infty,$$

and from Equation (24) it follows that

$$\nabla \varphi(\bar{x}) + \sum_{r \in I_1} \bar{\lambda}_r D_r + \sum_{j \in I_2} \bar{v}_j \nabla h_j(\bar{x}) + \sum_{t \in I_3} \bar{a}_t \nabla G_t(\bar{x}) + \sum_{t \in I_4} \bar{b}_t \nabla H_t(\bar{x}) + \sum_{i \in I_5} \bar{\mu}_i \nabla g_i(\bar{x}) + \sum_{t \in I_6} \bar{a}_t \nabla G_t(\bar{x}) + \sum_{t \in I_7} \bar{b}_t \nabla H_t(\bar{x}).$$

Note that $\varphi(x) = f(x) + \sum_{r \in I_{D_{\neq}}} |D_r^T x|^p$, and $(\bar{a}_i, \bar{b}_i) \in N_{\mathcal{C}}(\bar{y}_i, \bar{z}_i), i \in \mathcal{K}$ which implies that $\bar{a}_i \bar{b}_i = 0, i \in I_{GH}$. It is easy to verify that $\bar{x}$ satisfies the definition of M-stationary in definition 7 and the proof is complete. □

Theorem 4. *Let $\bar{x} \in \Theta$ be a locally optimal solution of problem (1) where MPSC Guignard qualification is valid. Then, $\bar{x}$ is a B-stationary of problem (1).*

Proof. Assume $\bar{x} \in \Theta$ is a locally optimal solution of problem (1) and by Lemma 2, $\bar{x}$ is a local minimizer of problem (13). It follows that

$$\nabla \varphi(\bar{x})^T d \geq 0, \ d \in T_{\Xi}(\bar{x}),$$

which implies that $-\nabla \varphi(\bar{x}) \in T_{\Xi}(\bar{x})^{\circ} = \mathcal{L}_{\Xi}^{MPSC}(\bar{x})^{\circ}$. From $-\nabla \varphi(\bar{x}) \in \mathcal{L}_{\Xi}^{MPSC}(\bar{x})^{\circ}$, we have

$$\nabla \varphi(\bar{x})^T d \geq 0, \ d \in \mathcal{L}_{\Xi}^{MPSC}(\bar{x}).$$

Note that $\nabla\varphi(\bar{x}) = \nabla f(\bar{x}) + \sum_{r\in I_{D\neq}} \lambda_r D_r$; denote

$$\zeta = \nabla f(x) + \sum_{r\in I_{D\neq}} \lambda_r D_r + \sum_{r\in I_{D0}} \lambda_r D_r \in \partial F(\bar{x}).$$

From $D_r^T d = 0, r \in I_{D0}$, for $d \in \mathcal{L}_{MPSC}(\bar{x})$, we have

$$\zeta^T d = (\nabla f(x) + \sum_{r\in I_{D\neq}} \lambda_r D_r + \sum_{r\in I_{D0}} \lambda_r D_r)^T d = \nabla\varphi(\bar{x})^T d + \sum_{r\in I_{D0}} \lambda_r D_r^T d \geq 0. \tag{25}$$

Therefore, $\bar{x}$ is a B-stationary of problem (1) and the proof is complete. □

3.1.3. The Relation between Various Qualifications

By Definition 8, the following relations are obtained immediately,

MPSC-LI qualification ⇒ MPSC-MF qualification ⇒ MPSC-NNAM qualification ⇒ MPSC-pseudonormality ⇒ MPSC-quasinormality;
MPSC-LI qualification ⇒ MPSC-CR qualification ⇒ MPSC-RCR qualification;
MPSC-LCQ ⇒ MPSC-CR qualification; MPSC-CPLD ⇒ MPSC-RCPLD;
MPSC-Abadie qualification ⇒ MPSC Guignard qualification.

Next, we show that MPSC-CR qualification(RCR qualification) ⇒ MPSC-CPLD (RCPLD), and MPSC-quasinormality ⇒ MPSC-Abadie qualification.

Theorem 5. *If the MPSC-CR qualification holds at $\bar{x} \in \Theta$, then MPSC-CPLD holds at $\bar{x}$. If the MPSC-RCR qualification holds at $\bar{x} \in \Theta$, then MPSC-RCPLD holds at $\bar{x}$.*

Proof. (i) Let the MPSC-CR qualification hold at $\bar{x} \in \Theta$. Let $I_1 \subseteq I_{D0}$, $I_2 \subseteq I_g$, $I_3 \subseteq \mathcal{J}$, $I_4 \subseteq I_G \cup I_{GH}$, and $I_5 \subseteq I_H \cup I_{GH}$. Assume that there exist multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$, not all zero, with $\mu_i \geq 0$ for each $i \in I_2$ and $\alpha_i\beta_i = 0$ for each $i \in I_{GH}$, such that

$$\sum_{r\in I_1} \lambda_r D_r + \sum_{i\in I_2} \mu_i \nabla g_i(\bar{x}) + \sum_{j\in I_3} \nu_j \nabla h_j(\bar{x}) + \sum_{k\in I_4} \alpha_k \nabla G_k(\bar{x}) + \sum_{k'\in I_5} \beta_{k'} \nabla H_{k'}(\bar{x}) = 0,$$

which means that the vectors

$$\{D_r, \nabla g_i(x), \nabla h_j(x), \nabla G_k(x), \nabla H_{k'}(x) \mid r \in I_1, i \in I_2, j \in I_3, k \in I_4, k' \in I_5\} \tag{26}$$

are linearly dependent at $\bar{x}$. By the MPSC-CR qualification, there exists $\delta > 0$, for any $x \in B_\delta(\bar{x})$, the vectors of Equation (26) have the same rank, implying that they are linearly dependent at $x \in B_\delta(\bar{x})$. Thus, MPSC-CLPD holds at $\bar{x}$.

(ii) Assume MPSC-RCR qualification holds at $\bar{x} \in \Theta$. Let $I_1 \subseteq I_{D0}$, $I_2 \subseteq \mathcal{J}$, $I_3 \subseteq I_G$ and $I_4 \subseteq I_H$ be such that $\mathcal{G}(\bar{x}, I_1, I_2, I_3, I_4)$ is a basis for span $\mathcal{G}(\bar{x}, I_{D0}, \mathcal{J}, I_G, I_H)$. Let $I_5 \subseteq I_g$, $I_6, I_7 \in I_{GH}$ be such that there exist multipliers $\{\lambda, \mu, \nu, \alpha, \beta\}$, not all zero, with $\mu_i \geq 0$ for each $i \in I_5$ and $\alpha_i\beta_i = 0$ for each $i \in I_{GH}$, such that

$$\sum_{r\in I_1} \lambda_r D_r + \sum_{i\in I_5} \mu_i \nabla g_i(\bar{x}) + \sum_{j\in I_2} \nu_j \nabla h_j(\bar{x}) + \sum_{k\in I_3\cup I_6} \alpha_k \nabla G_k(\bar{x}) + \sum_{k'\in I_4\cup I_7} \beta_{k'} \nabla H_{k'}(\bar{x}) = 0,$$

which means that

$$\{D_r, \nabla g_i(\bar{x}), \nabla h_j(\bar{x}), \nabla G_k(\bar{x}), \nabla H_{k'}(\bar{x}) \mid r \in I_1, i \in I_5, j \in I_2, k \in I_3 \cup I_6, k' \in I_4 \cup I_7\} \tag{27}$$

are linearly dependent at $\bar{x}$. Note that $\mathcal{G}(\bar{x}, I_1, I_2, I_3, I_4)$ and $\mathcal{G}(\bar{x}, I_{D0}, \mathcal{J}, I_G, I_H)$ have the same rank. Combine the MPSC-RCR qualification; there exists $\delta > 0$ such that the vectors of Equation (27) have the same rank at $x \in B_\delta(\bar{x})$, meaning that they are linearly dependent. Thus, MPSC-RCPLD holds at $\bar{x}$ and the proof is complete. □

Theorem 6. *If the MPSC-quasinormality holds at $\bar{x} \in \Theta$, the MPSC-Abadie qualification holds at $\bar{x}$.*

Proof. Assume the MPSC-quasinormality holds at $\bar{x} \in \Theta$. Note that the MPSC-quasinormality is the quasinormality condition for nonlinear optimization problem (13) and the MPSC-Abadie qualification is the ACQ for problem (13). By [30] (Theorem 10.5), the result is true and the proof is complete. □

Finally, we summarize the relations between the qualifications of non-Lipschitz MPSC in Figure 1.

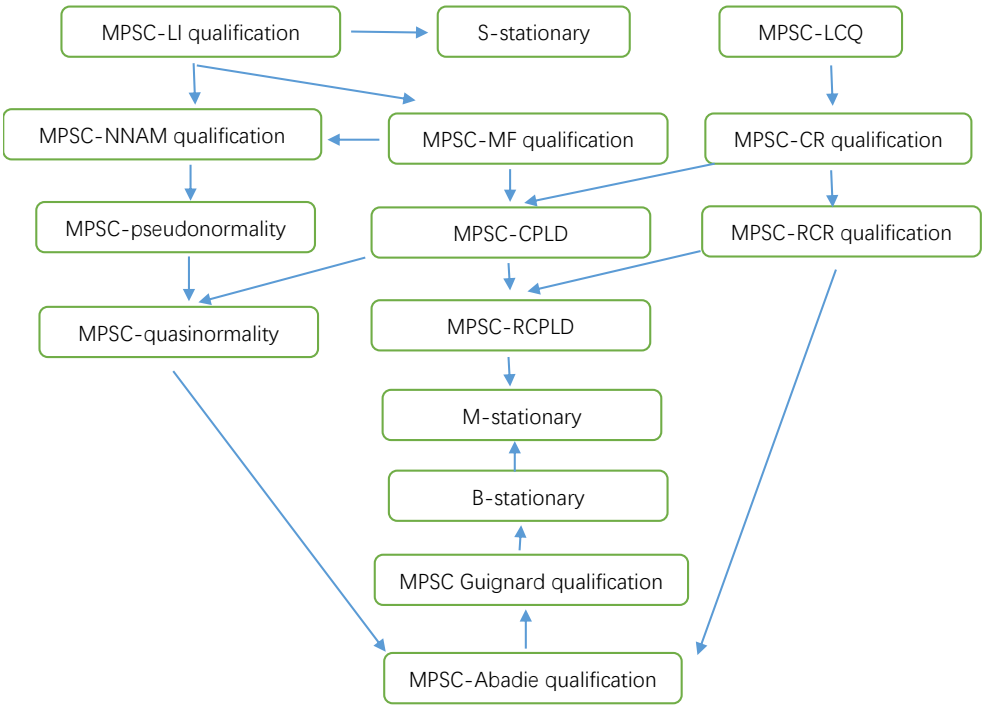

Figure 1. Relations between various qualifications.

4. The Approximation Method for the Non-Lipschitz MPSC Problem

The non-Lipschitz term and switching constraints make problem (1) difficult to solve. Here, we propose the relaxation method to problem (1) as follows

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & F(x) = f(x) + \varphi_\epsilon(x) \\ s.t. \quad & g_i(x) \le 0,\ i \in \mathcal{I}, \\ & h_j(x) = 0,\ j \in \mathcal{J}, \\ & t \le G_\ell(x)H_\ell(x) \le t,\ \ell \in \mathcal{K}. \end{aligned} \tag{28}$$

Here, the non-Lipschitz term $\|Dx\|_p^p$ is approximated by a local Lipschitz function

$$\varphi_\epsilon(x) := \sum_{i=1}^{s} (|D_i^T x| + \epsilon_i)^p,$$

where $\epsilon_i > 0, i = 1, \ldots, s$. The feasible region denotes

$$\Theta_t = \{g_i(x) \le 0, i \in \mathcal{I}, h_j(x) = 0, j \in \mathcal{J}, -t \le G_k(x)H_k(x) \le t, k \in \mathcal{K}\}.$$

This is the Scholtes global relaxation method, see [31], and is used in mathematical programs with switching constraints, see [16] (Section 3).

Theorem 7. *Let $\{t_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+$ be a sequence of positive relaxation parameters converging to zero. For each $k \in \mathbb{N}$, let $x^k \in \Theta_{t_k}$ be a KKT point of problem (28). Assume that the sequence $\{x_k\}_{k\in\mathbb{N}}$ converges to a point $\bar{x} \in \Theta$ where MPSC-MF qualification holds. Then $\bar{x}$ is a W-stationary point of problem (1).*

Proof. Noting that x^k is a KKT point of problem (28), we can find multipliers $\mu^k \in \mathbb{R}^m$, $\nu^k \in \mathbb{R}^p$, and $\lambda^k \in \mathbb{R}^l$ which satisfy the conditions as follows

$$\begin{aligned}
0 =& \nabla f(x^k) + p\sum_{i=1}^{s}(|D_i^T x| + \epsilon_i)^{p-1} sign(D_i^T x^k) D_i + \sum_{i\in I_g(x^k)} \mu_i^k \nabla g_i(x^k) + \sum_{j\in\mathcal{J}} \nu_j^k \nabla h_j(x^k) \\
&+ \sum_{\ell\in\mathcal{K}} \lambda_\ell^k [H_\ell(x^k)\nabla G_\ell(x^k) + G_\ell(x^k)\nabla H_\ell(x^k)], \\
&\forall i \in I_g(x^k) : \mu_i^k \geq 0, \ \forall i \in \mathcal{I} \setminus I_g(x^k) : \mu_i^k = 0, \\
&\forall \ell \in I_t^+(x^k) : \lambda_\ell^k \geq 0, \quad \forall \ell \in I_t^-(x^k) : \lambda_\ell^k \leq 0, \quad \forall \ell \in \mathcal{K} \setminus I_t^+(x^k) \cup I_t^-(x^k) : \lambda_\ell^k = 0,
\end{aligned}$$

where $I_t^+(x) := \{\ell \in \mathcal{K} : G_\ell(x)H_\ell(x) = t\}$, and $I_t^-(x) := \{\ell \in \mathcal{K} : G_\ell(x)H_\ell(x) = -t\}$. For each $k \in \mathbb{N}$ and $\ell \in \mathcal{K}$, we define multipliers $\alpha_\ell^k, \beta_\ell^k \in \mathbb{R}$ as stated below:

$$\alpha_\ell^k := \begin{cases} \lambda_\ell^k H_\ell(x^k) & \ell \in I_G \cup I_{GH}, \\ 0 & \ell \in I_H, \end{cases} \quad \beta_\ell^k := \begin{cases} \lambda_\ell^k G_\ell(x^k) & \ell \in I_H \cup I_{GH}, \\ 0 & \ell \in I_G, \end{cases} \tag{29}$$

and $\gamma_i^k = p(|D_i^T x^k| + \epsilon_i)^{p-1} sign(D_i^T x^k)$. Thus, we obtain

$$\begin{aligned}
0 = \nabla f(x^k) + \sum_{i=1}^{s} \gamma_i^k D_i + \sum_{i\in I_g(x^k)} \mu_i^k \nabla g_i(x^k) + \sum_{j\in\mathcal{J}} \nu_j^k \nabla h_j(x^k) + \sum_{\ell\in\mathcal{K}} [\alpha_\ell^k \nabla G_\ell(x^k) + \beta_\ell^k \nabla H_\ell(x^k)] \\
+ \sum_{\ell\in I_H} \lambda_\ell^k H_\ell(x^k)\nabla G_\ell(x^k) + \sum_{\ell\in I_G} \lambda_\ell^k G_\ell(x^k)\nabla H_\ell(x^k).
\end{aligned} \tag{30}$$

We next show that the sequence $\{(\gamma^k, \mu^k, \nu^k, \alpha^k, \beta^k, \lambda_{I_G\cup I_H}^k)\}_{k\in\mathbb{N}}$ is bounded. It is worth noting that $\gamma_i^k \in \mathbb{R}$ for $i \in I_{D0}(\bar{x})$, since $I_{D0}(x^k) \subset I_{D0}(\bar{x})$ and $sign(D_i^T x^k) = [-1, 1]$, and γ_i^k is bounded for $i \in I_{D\neq}$. Hence, we just show the boundness of $\{(\gamma_{I_{D0}}^k, \mu^k, \nu^k, \alpha^k, \beta^k, \lambda_{I_G\cup I_H}^k)\}_{k\in\mathbb{N}}$. In contrast, we assume that $\rho^k := \|(\gamma_{I_{D0}}^k, \mu^k, \nu^k, \alpha^k, \beta^k, \lambda_{I_G\cup I_H}^k)\| \to \infty$ when $k \to \infty$. Thus, $\frac{1}{\rho^k}(\gamma_{I_{D0}}^k, \mu^k, \nu^k, \alpha^k, \beta^k, \lambda_{I_G\cup I_H}^k) \to (\gamma_{I_{D0}}, \mu, \nu, \alpha, \beta, \lambda_{I_G\cup I_H})$ and $\|(\gamma_{I_{D0}}, \mu, \nu, \alpha, \beta, \lambda_{I_G\cup I_H})\| = 1$. Note that $I_g(x^k) \subset I_g(\bar{x})$ by the continuity of g for sufficiently large $k \in \mathbb{N}$. Dividing Equation (30) by ρ^k and taking the limit $k \to \infty$, we obtain

$$\begin{aligned}
&0 = \sum_{i\in I_{D0}} \gamma_i D_i + \sum_{i\in I_g} \mu_i \nabla g_i(\bar{x}) + \sum_{j\in\mathcal{J}} \nu_j \nabla h_j(\bar{x}) + \sum_{\ell\in\mathcal{K}} [\alpha_\ell \nabla G_\ell(\bar{x}) + \beta_\ell \nabla H_\ell(\bar{x})], \\
&\forall i \in I_g : \mu_i \geq 0, \quad \forall i \in \mathcal{I} \setminus I_g : \mu_i = 0, \quad \forall \ell \in I_G : \beta_\ell = 0, \quad \forall \ell \in I_H : \alpha_\ell = 0.
\end{aligned}$$

Therefore, we can yield $\gamma_{I_{D0}} = 0$, $\mu = 0$, $\nu = 0$, $\alpha = 0$, and $\beta = 0$ by the assumption that the MPSC-MF qualification holds at $\bar{x}$, which is defined as (a) of Definition 8. Hence, $\lambda_{\ell_0} \neq 0$ holds for at least one index $\ell_0 \in I_G \cup I_H$. Without loss of generality, we assume $\ell_0 \in I_G$, then we have $\alpha_{\ell_0}^k = \lambda_{\ell_0}^k H_{\ell_0}(x^k)$. Besides, we have

$$\alpha_{\ell_0} = \lim_{k\to\infty} \frac{\alpha_{\ell_0}^k}{\rho^k} = \lim_{k\to\infty} \frac{\lambda_{\ell_0}^k H_{\ell_0}(x^k)}{\rho^k} = \lambda_{\ell_0} H_{\ell_0}(\bar{x}) \neq 0,$$

which is a contradiction since $\alpha = 0$. As a consequence, the sequence $\{(\gamma^k, \mu^k, \nu^k, \alpha^k, \beta^k, \lambda_{I_G\cup I_H}^k)\}_{k\in\mathbb{N}}$ is bounded.

Thus, we may assume the sequence $\{(\gamma^k, \mu^k, \nu^k, \alpha^k, \beta^k, \lambda^k_{I_G \cup I_H})\}_{k\in\mathbb{N}}$ converges to $\{(\bar{\gamma}, \bar{\mu}, \bar{\nu}, \bar{\alpha}, \bar{\beta}, \bar{\lambda}_{I_G \cup I_H})\}$. We take the limit in Equation (30) and obtain

$$
\begin{aligned}
&0 = \nabla f(\bar{x}) + \sum_{i=1}^{s} \bar{\gamma}_i D_i + \sum_{i\in I_g} \bar{\mu}_i \nabla g_i(\bar{x}) + \sum_{j\in\mathcal{J}} \bar{\nu}_j \nabla h_j(\bar{x}) + \sum_{\ell\in\mathcal{K}} [\bar{\alpha}_\ell \nabla G_\ell(\bar{x}) + \bar{\beta}_\ell \nabla H_\ell(\bar{x})],\\
&\forall i \in I_{D\neq} : \bar{\gamma}_i = p(|D_i^T \bar{x}| + \epsilon_i)^{p-1} sign(D_i^T \bar{x}),\\
&\forall i \in I_g : \bar{\mu}_i \geq 0, \quad \forall i \in \mathcal{I} \setminus I_g : \bar{\mu}_i = 0, \quad \forall \ell \in I_G : \bar{\beta}_\ell = 0, \quad \forall \ell \in I_H : \bar{\alpha}_\ell = 0,
\end{aligned}
$$

which shows that $\bar{x}$ is a W-stationary point of problem (1). □

In some cases, even if the MPSC-LICQ is valid for the relaxation (28), a better solution may not be obtained than the W-stationary point. Study [16] (Example 3.3) has illustrated this point. Note that the second-order necessary condition is not valid at the stationary point $(\sqrt{t}, \sqrt{t})$. Consider its relaxation scheme similar to problem (28):

$$
\begin{aligned}
\min \quad & f(x) := \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}(x_2 - 1)^2\\
s.t. \quad & g_1(x) := x_1 x_2 - t \leq 0\\
& g_2(x) := -x_1 x_2 - t \leq 0
\end{aligned}
\tag{31}
$$

The first-order necessary condition is

$$
\nabla f(x) + \mu_1 \nabla g_1(x) + \mu_2 \nabla g_2(x) = 0,\ \mu_1,\ \mu_2 \geq 0,\ \mu_1 g_1(x) = 0,\ \mu_2 g_2(x) = 0,
$$

and the second-order necessary condition is

$$
\nabla^2 f(x) + \mu_1 \nabla^2 g_1(x) + \mu_2 \nabla^2 g_2(x) \geq 0,
$$

where

$$
\nabla^2 f(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \nabla^2 g_1(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
$$

Concerning problem (31), there are stationary points $(\frac{1}{2}, \frac{1}{2}) \pm \sqrt{\frac{1}{4} - t}(-1, 1)$ and $(\sqrt{t}, \sqrt{t})$. In fact, the constraint $g_1(x)$ is active and $g_2(x)$ is inactive at point $x^t := (\sqrt{t}, \sqrt{t})$. Hence, the corresponding multipliers $\mu_1 = \frac{1}{\sqrt{t}} - 1$ and $\mu_2 = 0$. Now, we verify that second-order necessary condition fails at x^t. For $d^t = (d_1^t, d_2^t)^T$, $d^t \in \mathcal{T}_d(x) := \{d \in \mathbb{R}^2 : \nabla g_1(x^t)^T d = 0\}$, we have $d_1^t = -d_2^t$. Then, the second-order necessary condition

$$
(d^t)^T [\nabla^2 f(x^t) + \mu_1 \nabla^2 g_1(x^t)] d^t = (d_1^t)^2 + (d_2^t)^2 + 2\mu_1 d_1^t d_2^t = 2(2 - \frac{1}{\sqrt{t}})(d_1^t)^2.
$$

Note that t is sufficiently small; then $2 - \frac{1}{\sqrt{t}} < 0$, which implies that the second-order necessary condition is not valid at x^t.

However, for points $(\frac{1}{2}, \frac{1}{2}) \pm \sqrt{\frac{1}{4} - t}(-1, 1)$, the corresponding multipliers $\mu_1 = 1$ and $\mu_2 = 0$. For all $d \in \mathcal{T}_d(x)$, the second-order necessary condition

$$
d^T [\nabla^2 f(x) + \mu_1 \nabla^2 g_1(x)] d = d_1^2 + d_2^2 + 2 d_1 d_2 = (d_1 + d_2)^2 \geq 0.
$$

Furthermore, we know that the limit points of $(\frac{1}{2}, \frac{1}{2}) \pm \sqrt{\frac{1}{4} - t}(-1, 1)$ are $(1, 0)$ and $(0, 1)$, which are S-stationary for the problem in [16] (Example 3.3).

Definition 9. *We say, for problem* (28)*, the second-order necessary condition holds at* x^k *if there exists* d *such that*

$$
d^T [\Phi(x^k) + M(x^k)] d \geq 0,\ \forall d \in \mathcal{T}_d(x^k),
$$

where

$$\Phi(x^k) := \nabla^2 f(x^k) + p(p-1)\sum_{r=1}^{s}(|D_r^T x^k| + \epsilon_r)^{p-2} D_r D_r^T,$$

$$M(x^k) := \sum_{i=1}^{m}\mu_i \nabla^2 g_i(x^k) + \sum_{j=1}^{p}\nu_j \nabla^2 h_j(x^k) + \sum_{\ell=1}^{l}\lambda_\ell [G_\ell(x^k)\nabla^2 H_\ell(x^k) + H_\ell(x^k)\nabla^2 G_\ell(x^k)]$$
$$+ \sum_{\ell=1}^{l}\lambda_\ell [\nabla G_\ell(x^k)\nabla H_\ell(x^k) + \nabla H_\ell(x^k)\nabla G_\ell(x^k)],$$

$$\mathcal{T}_d(x^k) := \{d \in \mathbb{R}^n : D_r^T d = 0\ r \in I_{D0}, \nabla g_i(x^k)^T d = 0\ i \in I_g, \nabla h_j(x^k)^T d = 0\ j \in \mathcal{J},$$
$$[G_\ell(x^k)\nabla H_\ell(x^k) + H_\ell(x^k)\nabla G_\ell(x^k)]^T d = 0\ \ell \in I_t^+(x^k) \cup I_t^-(\bar{x})\}.$$

Theorem 8. *Let $\{t_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+$ be a sequence of positive relaxation parameters converging to zero. For each $k \in \mathbb{N}$, let $x^k \in \Theta_{t_k}$ be a KKT point of problem (28), and the sequence $\{x_k\}_{k\in\mathbb{N}}$ converges to a point $\bar{x} \in \Theta$. Assume that the second-order necessary condition holds at x^k for all large k and the MPSC-MF qualification holds at $\bar{x}$. Then $\bar{x}$ is an S-stationary point of problem (1).*

Proof. In the proof of Theorem 7, we have shown that $\bar{x}$ is a W-stationary point under the MPSC-MF qualification condition. Assume to the contrary that $\bar{x}$ is not an S-stationary point. There must exist $\ell_0 \in I_{GH}$ such that $\alpha_{\ell_0}^k \to \bar{\alpha}_{\ell_0} \neq 0$ or $\beta_{\ell_0}^k \to \bar{\beta}_{\ell_0} \neq 0$. Without loss of generality, we assume $\bar{\alpha}_{\ell_0} \neq 0$. Let

$$A_k := \begin{bmatrix} D_i^T & i \in I_{D0} \\ \nabla g_i(x^k)^T & i \in I_g \\ \nabla h_i(x^k)^T & i \in \mathcal{J} \\ \nabla G_i(x^k)^T & i = \ell_0 \\ \nabla G_i(x^k)^T & i \in I_{GH} \setminus \{\ell_0\} \cup I_G(x^k) \\ \frac{G_i(x^k)}{H_i(x^k)}\nabla H_i(x^k)^T + \nabla G_i(x^k)^T & i \in I_G \setminus I_G(x^k) \\ \nabla H_i(x^k)^T & i = \ell_0 \\ \nabla H_i(x^k)^T & i \in I_{GH} \setminus \{\ell_0\} \cup I_H(x^k) \\ \frac{H_i(x^k)}{G_i(x^k)}\nabla G_i(x^k)^T + \nabla H_i(x^k)^T & i \in I_H \setminus I_H(x^k) \end{bmatrix}$$

and

$$z_k := \begin{bmatrix} 0 & i \in I_{D0} \\ 0 & i \in I_g \\ 0 & i \in \mathcal{J} \\ 1 & i = \ell_0 \\ 0 & i \in I_{GH} \setminus \{\ell_0\} \\ 0 & i \in I_G \\ -\frac{\lambda_{\ell_0} H_{\ell_0}^2(x^k)}{G_{\ell_0}(x^k)} & i = \ell_0 \\ 0 & i \in I_{GH} \setminus \{\ell_0\} \\ 0 & i \in I_H \end{bmatrix}$$

The index sets $I_G(x^k) \subseteq I_G$ and $I_H(x^k) \subseteq I_H$ for all k sufficiently large; then we have that $(I_{GH} \setminus \{\ell_0\} \cup I_G(x^k)) \cup (I_G \setminus I_G(x^k)) = I_{GH} \cup I_G \setminus \{\ell_0\}$ and $(I_{GH} \setminus \{\ell_0\} \cup I_H(x^k)) \cup (I_H \setminus I_H(x^k)) = I_{GH} \cup I_H \setminus \{\ell_0\}$. By the MPSC-LI qualification at $\bar{x}$, it follows that the limit of the matrix sequence $\{A_k\}$ has full row rank. Hence, A_k has full rank for all k sufficiently large. Let $d^k := A_k^T(A_k A_k^T)^{-1} z^k$, which satisfies $A_k d_k = z^k$. Furthermore, the sequence $\{z^k\}$ is convergent since $\frac{\lambda_{\ell_0} H_{\ell_0}^2(x^k)}{G_{\ell_0}(x^k)} \to \frac{\bar{\alpha}_\ell^2}{\bar{\beta}_\ell}$ with $k \to \infty$ by the definitions of α_ℓ and β_ℓ in the proof of Theorem 7, which means that $\{d^k\}$ is bounded. In the following, we show the second-order derivatives of the objective and constraints for problem (28).

Let $\psi_\ell(x) := G_\ell(x)H_\ell(x)$, for $\ell \in I_t^+(x^k) \cup I_t^-(x^k)$, we know $\ell \in I_G \setminus I_G(x^k) \cup I_H \setminus I_H(x^k)$; then we have

$$\begin{aligned}\nabla\psi_\ell(x^k)^T d^k &= G_\ell(x^k)\nabla H_\ell(x^k)^T d^k + H_\ell(x^k)\nabla G_\ell(x^k)^T d^k \\ &= G_\ell(x^k)\frac{-H_\ell(x^k)}{G_\ell(x^k)}\nabla G_\ell(x^k)^T d^k + H_\ell(x^k)\nabla G_\ell(x^k)^T d^k \\ &= 0.\end{aligned} \tag{32}$$

From $A^k d^k = z^k$ and Equation (32) we obtain that $d^k \in \mathcal{T}_d(x^k)$ for all sufficiently large k.

We next show that $\psi_k(x^k) := (d^k)^T(\Phi_k + M_k)d^k \to -\infty$ as $k \to \infty$, which contradicts the second-order necessary condition at x^k when k is sufficiently large. Note that $D_r^T d^k = 0$ for all $r \in I_{D0}$, we have that

$$(d_k)^T\Phi(x^k)d^k = (d_k)^T\nabla^2 f(x^k)d^k + p(p-1)\sum_{r\in I_{D\neq}}(|D_r^T x^k| + \epsilon_r)^{p-2}(D_r^T d^k)^2.$$

Since $\{d^k\}$, $\{x^k\}$ are bounded, and $|D_r^T x^k| \to |D_r^T \bar{x}| > 0$ for all $r \in I_{D\neq}$, it is easy to verify that the sequence $\{(d_k)^T\Phi_k d^k\}$ is bounded. For the constraints, we can obtain

$$\begin{aligned}(d_k)^T M(x^k)d^k = &\sum_{i=1}^m \mu_i^k (d_k)^T\nabla^2 g_i(x^k)d^k + \sum_{j=1}^p \nu_j^k (d_k)^T\nabla^2 h_j(x^k)d^k \\ &+ \sum_{\ell=1}^l \lambda_\ell^k[G_\ell(x^k)(d_k)^T\nabla^2 H_\ell(x^k)d^k + H_\ell(x^k)(d_k)^T\nabla^2 G_\ell(x^k)d^k] \\ &+ 2\sum_{\ell=1}^l \lambda_\ell^k \nabla G_\ell(x^k)^T d^k \nabla H_\ell(x^k)^T d^k.\end{aligned} \tag{33}$$

Note that $\lambda_\ell^k G_\ell(x^k) = \beta_\ell^k$ and $\lambda_\ell^k H_\ell(x^k) = \alpha_\ell^k$ according to Equation (29). In the proof of Theorem 7, we proved that $\{\mu^k\}$, $\{\nu^k\}$, $\{\alpha^k\}$ and $\{\beta^k\}$ are all bounded. Since $\{d^k\}$ is bounded, it is not hard to see that the top four items in Equation (33) are all bounded. At the same time, by $A^k d^k = z^k$ and the definitions of A^k and z^k, we have

$$\begin{aligned}&\sum_{\ell=1}^l \lambda_\ell^k \nabla G_\ell(x^k)^T d^k \nabla H_\ell(x^k)^T d^k \\ =&(\sum_{\ell\in I_G} + \sum_{\ell\in I_H} + \sum_{\ell\in I_{GH}\setminus\{\ell_0\}} + \sum_{\ell=\ell_0})(\lambda_\ell^k \nabla G_\ell(x^k)^T d^k \nabla H_\ell(x^k)^T d^k) \\ =&-\sum_{\ell\in I_G}\lambda_\ell^k\frac{G_\ell(x^k)}{H_\ell(x^k)}(\nabla H_\ell(x^k)^T d^k)^2 - \sum_{\ell\in I_H}\lambda_\ell^k\frac{H_\ell(x^k)}{G_\ell(x^k)}(\nabla G_\ell(x^k)^T d^k)^2 + \frac{-(\lambda_{\ell_0}^k H_{\ell_0}(x^k))^2}{G_{\ell_0}(x^k)}.\end{aligned}$$

Since $G_{\ell_0} \to 0$ and $\lambda_{\ell_0}^k H_{\ell_0}(x^k) = \alpha_{\ell_0}^k \to \bar{\alpha}_{\ell_0} \neq 0$, the term $\frac{-(\lambda_{\ell_0}^k H_{\ell_0}(x^k))^2}{G_{\ell_0}(x^k)} \to -\infty$ as $k \to \infty$. Thus, $(d_k)^T M_k d^k \to -\infty$ as $k \to \infty$. Together with the boundness of $\{(d_k)^T\Phi_k d^k\}$, which contradicts the second-order necessary condition. Thus, $\bar{\alpha}_{\ell_0} = 0$.

Similarly, we can show that $\bar{\beta}_{\ell_0} = 0$. Furthermore, the second-order necessary condition holding at point x^k implies that $\alpha_{\ell_0}^k$ and $\beta_{\ell_0}^k$ vanished at the same time. The proof is completed. □

5. Conclusions

In this paper, we study optimality conditions and qualifications for non-Lipschitz MPSCs. Our theoretical investigations enrich the landscape of available qualifications in the field of non-Lipschitz MPSCs. Moreover, we present an approximation method similar as Scholtes' relaxation scheme. Theoretically, the accumulation point of the sequences

produced by our method is a stationary point under certain conditions, however, there is no efficient computational solution method to solve it. Referring to the existing algorithms, some penalty methods for MPSCs do not apply for the non-Lipschitz term, and some smoothing methods for non-Lipschitz optimization do not take into account the switching constrains. Hence, combing the penalty method and the smoothing method to propose an efficient method to solve the non-Lipschitz MPSCs will be interesting and meaningful work in the future.

Author Contributions: Conceptualization, methodology, writing—original draft preparation and writing—review and editing, J.L.; methodology, writing—review and editing, Z.P.; writing—original draft preparation, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Natural Science Foundation of China (Grant No. 11871383, 12001260).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liberzon, D. *Switching in Systems and Control*; Birkhäuser: Boston, MA, USA, 2003.
2. Sager, S. Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control. *J. Process Control.* **2009**, *19*, 1238–1247. [CrossRef]
3. Wang, L.; Yan, Q. Time optimal controls of semilinear heat equation with switching control. *J. Optim. Theory Appl.* **2015**, *165*, 263–278. [CrossRef]
4. Sarker, R.A.; Newton, C.S. *Optimization Modelling*; CRC Press: Boca Raton,FL, USA, 2008.
5. Seidman, T.I. Optimal control of a diffusion/reaction/switching system. *Evol. Equ. Control Theory* **2013**, *2*, 723–731. [CrossRef]
6. Gugat, M. Optimal switching boundary control of a string to rest in finite time. *ZAMM J. Appl. Math. Mech.* **2008**, *88*, 283–305. [CrossRef]
7. Luo, Z.Q.; Pang, J.S.; Ralph, D. *Mathematical Programs with Equilibrium Constraints*; Cambridge University Press: Cambridge, UK, 1996.
8. Outrata, J.V.; Kocvara, M.; Zowe, J. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*; Kluwer Academic Publishers: Boston, MA, USA, 1998.
9. Achtziger, W.; Kanzow, C. Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications. *Math. Program.* **2008**, *114*, 69–99. [CrossRef]
10. Sadeghieh, A.; Kanzi, N.; Caristi, G. On stationarity for nonsmooth multiobjective problems with vanishing constraints. *J. Glob. Optim.* **2021**, 1–21. [CrossRef]
11. Burdakov, O.P.; Kanzow, C.; Schwartz, A. Mathematical programs with cardinality constraints: Reformulation by complementarity-type conditions and a regularization method. *SIAM J. Optim.* **2016**, *26*, 397–425. [CrossRef]
12. Červinka, M.; Kanzow, C.; Schwartz, A. Constraint qualifications and optimality conditions for optimization problemswith cardinality constraints. *Math. Program.* **2016**, *160*, 353–377. [CrossRef]
13. Mehlitz, P. Stationarity conditions and constraint qualifications for mathematical programs with switching constraints. *Math. Program.* **2020**, *181*, 149–186. [CrossRef]
14. Li, G.; Guo, L. Mordukhovich Stationarity for Mathematical Programs with Switching Constraints under Weak Constraint Qualifications. Available online: http://www.optimization-online.org/DB_FILE/2019/07/7288.pdf (accessed on 7 November 2021).
15. Liang, Y.-C.; Ye, J.J. Optimality conditions and exact penalty for mathematical programs with switching constraints. *J. Optim. Theory Appl.* **2021**, *190*, 1–31. [CrossRef]
16. Kanzow, C.; Mehlitz, P.; Steck, D. Relaxation schemes for mathematical programs with switching constraints. *Optim. Methods Softw.* **2019**, 1–36. [CrossRef]
17. Clason, C.; Rund, A.; Kunisch, K.; Barnard, R.C. A convex penalty for switching control of partial differential equations. *Syst. Control Lett.* **2016**, *89*, 66–73. [CrossRef]
18. Clason, C.; Rund, A.; Kunisch, K. Nonconvex penalization of switching control of partial differential equations. *Syst. Control Lett.* **2017**, *106*, 1–8. [CrossRef]
19. Brodie, J.; Daubechies, I.; De Mol, C.; Giannone, D.; Loris, I. Sparse and stable Markowitz portfolios. *Proc. Nat. Acad. Sci. USA* **2009**, *106*, 12267–12272. [CrossRef]
20. Bruckstein, A.M.; Donoho, D.L.; Elad, M. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Rev.* **2009**, *51*, 34–81. [CrossRef]

21. Bian, W.; Chen, X. Linearly constrained non-Lipschitz optimization for image restoration. *SIAM J. Imaging Sci.* **2015**, *8*, 2294–2322. [CrossRef]
22. Zeng, C.; Wu, C.; Jia, R. Non-Lipschitz Models for Image Restoration with Impulse Noise Removal. *SIAM J. Imaging Sci.* **2019**, *12*, 420–458. [CrossRef]
23. Chen, X.; Niu, L.; Yuan, Y. Optimality conditions and a smoothing trust region newton method for non-Lipschitz optimization. *SIAM J. Optim.* **2013**, *23*, 1528–1552. [CrossRef]
24. Lu, Z. Iterative reweighted minimization methods for l_p regularized unconstrained nonlinear programming. *Math. Program.* **2014**, *147*, 277–307. [CrossRef]
25. Huang, Y.; Liu, H. Smoothing projected Barzilai-Borwein method for constrained non-Lipschitz optimization. *Comput. Optim. Appl.* **2016**, *65*, 671–698. [CrossRef]
26. Liu, Y.; Ma, S.; Dai, Y.-H.; Zhang, S. A smoothing SQP framework for a class of composite l_p minimization over polyhedron. *Math. Program.* **2016**, *158*, 467–500. [CrossRef]
27. Guo, L.; Chen, X.J. Mathematical programs with complementarity constraints and a non-Lipschitz objective: Optimality and approximation. *Math. Program.* **2021**, *185*, 455–485. [CrossRef]
28. Rockafellar, R.T.; Wets, R.J.-B. *Variational Analysis*; Springer: Berlin, Germany, 1998.
29. Andreani, R.; Haeser, G.; Schuverdt, M.L.; Silva, P.J.S. A relaxed constant positive linear dependence constraint qualification and applications. *Math. Program.* **2012**, *135*, 255–273. [CrossRef]
30. Hestenes, M.R. *Optimization Theory: The Finite Dimensional Case*; Wiley: New York, NY, USA, 1975.
31. Scholtes, S. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM J. Optim.* **2001**, *11*, 918–936. [CrossRef]