Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

doi:10.3934/xx.xxxxxxx

pp. **X–XX**

GLOBAL CONVERGENCE OF AUGMENTED LAGRANGIAN METHOD APPLIED TO MATHEMATICAL PROGRAM WITH SWITCHING CONSTRAINTS

Lei Guo

School of Business, East China University of Science and Technology, Shanghai 200237, China

Gao-Xi Li*

School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China;
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

Xinmin Yang

School of Mathematics Science, Chongqing Normal University, Chongqing 401331, China;

(Communicated by the associate editor name)

Abstract. The mathematical program with switching constraints (MPSC) is a kind of problems with disjunctive constraints. The existing convergence results cannot directly be applied to this kind of problem since the required constraint qualifications for ensuring the convergence are very likely to fail. In this paper, we apply the augmented Lagrangian method (ALM) to solve the MPSC and an application of the MPSC (i.e., the either-or-constrained program). We show that, under the MPSC relaxed constant positive linear dependent condition recently proposed in the literature, the feasible accumulation points of the iterates generated by the ALM are guaranteed to be strongly stationary if the multiplier sequence is bounded. When the multiplier sequence is unbounded, the feasible accumulation points are weakly stationary if MPSC linear independence constraint qualification holds. Some numerical experiments are conducted and compared with the recently proposed relaxation method. The numerical results demonstrate the effectiveness of the ALM and show that ALM can find better solutions than the relaxation method.

1. **Introduction.**

The mathematical program with switching constraints (MPSC) considered in this paper is of the form

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

2020 *Mathematics Subject Classification.* 90C26; 90C30 .

Key words and phrases. mathematical program with switching constraints; either-or-constrained program; augmented Lagrangian method.

This first author was supported in part by NSFC Grant(No.72131007,11771287), the Natural Science Foundation of Shanghai, China(No.22ZR1415900),the Fundamental Research Funds for the Central Universities; This second author was supported in part by NSFC Grants (No.11901068); This third author was supported in part by Natural Science Foundation of Chongqing (No.cstc2019jcyj-zdxmX0016).

*Corresponding author: Gaoxi Li .

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 simplicity, 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 linear independence constraint qualification and Mangasarian-Fromovitz constraint qualification (MFCQ) fail [8, 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 Abadie constraint qualification fails at a degenerate feasible point.

The constraint structure of the MPSC is highly related with the mathematical program with complementarity constraints and mathematical program with vanishing constraints which have been widely studied in the literature [17, 15, 10, 12, 1, 6, 16, 5]. Recently, Mehlitz [8] proposed several stationarity concepts such as weak-stationarity, Mordukhovich (M-) stationarity, and strong (S-) stationarity for the MPSC (1), and for ensuring these stationary conditions to hold at optimum, then proposed some MPSC tailored constraint qualifications such as MPSC Mangasarian-Fromovitz constraint qualification, MPSC linear independence constraint qualification (MPSC LICQ), MPSC Abadie constraint qualification, and MPSC Guignard constraint qualification. Liang and Ye [13] consider the optimality conditions and exact penalty for this problem. Based on the theoretical findings in [8], a relaxation method for solving MPSC was proposed in [9] and second-order optimality conditions for MPSC were investigated in [14]. More recently, Li and Guo [11] further extended the weaker verifiable constraint qualifications for standard nonlinear programming to the MPSC (1) and discussed the relations among all the MPSC constraint qualifications.

This rest of the paper is organized as follows. In Section 2 we recall some basic concepts for the MPSC. In Section 3 we present the convergence results of the ALM for solving the MPSC. In Section 4, some numerical results are reported.

2. **Preliminaries.**

The used notation is standard as in the literature. We let $\|\cdot\|$ denote the Euclidean norm and $\|\cdot\|_\infty$ the infinite norm. Denote $[v]_i$ as the ith component of the vector v. If there is no possibility of confusion we may also use the notation v_i. We denote by $\circ$ the Hadamard product (i.e., the componentwise product of two vectors).

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

$$\begin{cases} \mathcal{I}_h := \{1,\dots,p\},\ \mathcal{I}_g^* := \{i \in \{1,\dots,m\} : g_i(x^*) = 0\}, \\ \mathcal{I}_G^* := \{t \in \{1,\dots,l\} : G_t(x^*) = 0 \wedge H_t(x^*) \neq 0\}, \\ \mathcal{I}_H^* := \{t \in \{1,\dots,l\} : G_t(x^*) \neq 0 \wedge H_t(x^*) = 0\}, \\ \mathcal{I}_{GH}^* := \{t \in \{1,\dots,l\} : G_t(x^*) = 0 \wedge H_t(x^*) = 0\}. \end{cases} \tag{2}$$

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

$$L_{MPSC}(x,\lambda_g,\lambda_h,\lambda_G,\lambda_H) := f(x) + g(x)^\top\lambda_g + h(x)^\top\lambda_h + G(x)^\top\lambda_G + H(x)^\top\lambda_H.$$

We recall some stationarity concepts of the MPSC (1) and some constraint qualifications [8, 11].

Definition 2.1. Let $x^* \in \mathcal{X}$. We say that x^* is W-stationary to problem (1) if there exist multipliers $(\lambda_g, \lambda_h, \lambda_G, \lambda_H) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ such that

$$\begin{cases} \nabla_x L_{MPSC}(x^*, \lambda_g, \lambda_h, \lambda_G, \lambda_H) = 0, \\ \lambda_g \geq 0,\ g(x^*)^\top \lambda_g = 0,\ [\lambda_G]_i = 0,\ i \in \mathcal{I}_H^*,\ [\lambda_H]_i = 0,\ i \in \mathcal{I}_G^*. \end{cases} \tag{3}$$

We say that x^* is M-stationary to problem (1) if there exists multipliers $(\lambda_g, \lambda_h, \lambda_G, \lambda_H) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ satisfying (3) and

$$[\lambda_G]_i [\lambda_H]_i = 0,\ i \in \mathcal{I}_{GH}^*. \tag{4}$$

We say that x^* is S-stationary to (1) if there exists multipliers $(\lambda_g, \lambda_h, \lambda_G, \lambda_H) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l$ satisfying (3) and

$$[\lambda_G]_i = 0 \text{ and } [\lambda_H]_i = 0,\ i \in \mathcal{I}_{GH}^*. \tag{5}$$

Definition 2.2. Let $x^* \in \mathcal{X}$. We say that MPSC LICQ holds at x^* if the following gradients are linearly independent:

$$\left\{ \nabla g_i(x^*), \nabla h_j(x^*), \nabla G_r(x^*), \nabla H_t(x^*) : \begin{array}{l} i \in \mathcal{I}_g^*, r \in \mathcal{I}_G^* \cup \mathcal{I}_{GH}^* \\ j \in \mathcal{I}_h, t \in \mathcal{I}_H^* \cup \mathcal{I}_{GH}^* \end{array} \right\}.$$

Definition 2.3. 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 $\{\lambda_g, \lambda_h, \lambda_G, \lambda_H\}$ with $[\lambda_g]_i \geq 0$ for each $i \in I_4$ and $[\lambda_G]_i [\lambda_H]_i = 0$ for each $i \in \mathcal{I}_{GH}^*$, not all zero, such that

$$\sum_{i \in I_4} [\lambda_g]_i \nabla g_i(x^*) + \sum_{j \in I_1} [\lambda_h]_j \nabla h_j(x^*) + \sum_{r \in I_2 \cup I_5} [\lambda_G]_r \nabla G_r(x^*) + \sum_{t \in I_3 \cup I_6} [\lambda_H]_t \nabla H_t(x^*) = 0, \tag{6}$$

  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

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

It is obvious that MPSC LICQ is stronger than MPSC RCPLD (see, e.g.,[11]).

3. Augmented Lagrangian method.

In this section, we first present the convergence of the ALM for solving the MPSC (1) and then apply the convergence results to either-or-constrained programs.

3.1. Augmented Lagrangian method for the MPSC.

In this section we apply the safeguarded ALM for nonlinear programming problems proposed in [2] to solve the MPSC (1) and then investigate its convergence. In general, constraint qualifications such as the RCPLD for ensuring the convergence of the safeguarded ALM imply that the tangent cone of the considered nonlinear programming problems are convex. Thus, the existing convergence results cannot be expected to be valid for the MPSC (1) due to that the tangent cone of the feasible region of the MPSC (1) is generally nonconvex. An independent analysis of the MPSC (1) is needed that takes into account the problem structure.

The usual Lagrangian function of the MPSC (1) is defined as

$$L(x, \mu_g, \mu_h, \mu_o) := f(x) + g(x)^\top \mu_g + h(x)^\top \mu_h + (G(x) \circ H(x))^\top \mu_o,$$

and the augmented Lagrangian function is defined as

$$L_\rho(x, \mu_g, \mu_h, \mu_o) := f(x) + \frac{1}{2\rho}(\|\mu_h + \rho h(x)\|^2 + \|\mu_o + \rho(G(x) \circ H(x))\|^2 + \|\max\{0, \mu_g + \rho g(x)\}\|^2).$$

In the following, we apply the safeguarded ALM for nonlinear programming problems proposed in [2] to solve the MPSC (1).

Algorithm 3.1. Choose parameters $\bar{\mu}_g^{\max} > 0$, $\bar{\mu}_h^{\min}$ and $\bar{\mu}_h^{\max}$ such that $\bar{\mu}_h^{\min} < \bar{\mu}_h^{\max}$, $\bar{\mu}_0^{\min}$ and $\bar{\mu}_0^{\max}$ such that $\bar{\mu}_0^{\min} < \bar{\mu}_0^{\max}$, $\rho_0 > 0$, $\theta \in [0, 1)$, and $\sigma > 1$. Set $k = 0$.

Step 1. Let $[\bar{\mu}_g]_j^k \in [0, \bar{\mu}_g^{\max}]$, $j = 1, ..., m$, $[\bar{\mu}_h]_j^k \in [\bar{\mu}_h^{\min}, \bar{\mu}_h^{\max}]$, $i = 1, ..., p$, and $[\bar{\mu}_0]_j^k \in [\bar{\mu}_0^{\min}, \bar{\mu}_0^{\max}]$, $i = 1, ..., l$. (For $k > 0$, the typical option is to take $(\bar{\mu}_g^k, \bar{\mu}_h^k, \bar{\mu}_0^k)$ as the Euclidean projection of $(\mu_g^k, \mu_h^k, \mu_0^k)$ onto the box $\bigotimes_{i=0}^m [0, \bar{\mu}_g^{\max}] \times \bigotimes_{j=1}^p [\bar{\mu}_h^{\min}, \bar{\mu}_h^{\max}] \times \bigotimes_{j=1}^l [\bar{\mu}_0^{\min}, \bar{\mu}_0^{\max}]$).

Compute x^{k+1} as a stationary point of the unconstrained optimization problem

$$\min_x L_{\rho_k}(x, \bar{\mu}_g^k, \bar{\mu}_h^k, \bar{\mu}_0^k). \tag{7}$$

Step 2. Set

$$\mu_h^{k+1} = \bar{\mu}_h^k + \rho^k h(x^{k+1}),\ \mu_o^{k+1} = \bar{\mu}_o^k + \rho^k G(x^{k+1}) \circ H(x^{k+1}), \tag{8}$$

$$\mu_g^{k+1} = \max\{0, \bar{\mu}_g^k + \rho^k g(x^{k+1})\}, \tau^{k+1} = \min\{\mu_g^{k+1}, -g(x^{k+1})\}, \tag{9}$$

and

$$\beta^{k+1} = \max\{\|h(x^{k+1})\|_\infty, \|G(x^{k+1}) \circ H(x^{k+1})\|_\infty, \|\tau^{k+1}\|_\infty\}.$$

Step 3. If $k = 0$ or $\beta^{k+1} \le \theta \beta^k$, select any $\rho_{k+1} = \rho_k$. Otherwise select $\rho_{k+1} = \sigma \rho_k$. Adjust k by 1, and go to Step 1.

Before giving the convergence results of Algorithm 3.1, we recall two Lemmas.

Lemma 3.1. [3, Lemma 1] *Let $0 \neq x = \sum_{i=1}^{m+p} \alpha_i v_i$, where $\{v_1, \ldots, v_m\}$ is linearly independent. 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.*

Lemma 3.2. [7, Lemma 3.1] *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ be differentiable in some neighborhood of $\bar{x} \in \mathbb{R}^n$, with their derivatives being continuous at $\bar{x}$. Let $\{x^k\} \subset \mathbb{R}^n$ be a sequence which converges to $\bar{x}$. Assume that $\nabla h(\bar{x})$ is full of row rank and the following*

$$\alpha^k \nabla f(x^k) - (\nabla h(x^k))^\top \lambda^k = w^k$$

holds for some $\alpha^k \in \mathbb{R}$, $\lambda^k \in \mathbb{R}^l$, and $\omega^k \in \mathbb{R}^n$ for all k.

If $\alpha_k \to \bar{\alpha}$ and $\{w^k\} \to 0$ as $k \to \infty$, then there exists a unique $\bar{\lambda} \in \mathbb{R}^s$ such that $\{\lambda^k\} \to \bar{\lambda}$ as $k \to \infty$ and

$$\bar{\alpha} \nabla f(\bar{x}) - (\nabla h(\bar{x}))^\top \bar{\lambda} = 0.$$

In particular, if $\bar{\alpha} = 0$, then $\{\lambda^k\} \to 0$.

We next give the convergence of the sequence generated by Algorithm 3.1.

Theorem 3.3. *Let $\{x^k\} \subset \mathbb{R}^n$ be a sequence generated by Algorithm 3.1 and let $x^* \in \mathbb{R}^n$ be an accumulation point. Suppose that the sequence $\{\mu_o^k\}$ generated by Algorithm 3.1 has a bounded subsequence such that the corresponding subsequence of $\{x^k\}$ converges to x^*. If x^* is feasible to the MPSC (1) and MPSC-RCPLD holds at x^*, then x^* is an S-stationary point of the MPSC (1).*

Proof. According to Step 1 and Step 2 in Algorithm 3.1, it follows that

$$\nabla f(x^k) + \sum_{i=1}^{m} [\mu_g^k]_i \nabla g_i(x^k) + \sum_{i=1}^{p} [\mu_h^k]_i \nabla h_i(x^k) + \sum_{i=1}^{l} [\mu_o^k]_i H_i(x^k) \nabla G_i(x^k) + \sum_{i=1}^{l} [\mu_o^k]_i G_i(x^k) \nabla H_i(x^k) = 0,\ \mu_g^k \geq 0. \tag{10}$$

Let

$$\lambda_G^k = \mu_o^k \circ H(x^k),\ \lambda_H^k = \mu_o^k \circ G(x^k),\ \lambda_g^k = \mu_g^k,\ \lambda_h^k = \mu_h^k. \tag{11}$$

Then (10) becomes

$$\nabla f(x^k) + \sum_{i=1}^{m} [\lambda_g^k]_i \nabla g_i(x^k) + \sum_{i=1}^{p} [\lambda_h^k]_i \nabla h_i(x^k) + \sum_{i=1}^{l} [\lambda_G^k]_i \nabla G_i(x^k) + \sum_{i=1}^{l} [\lambda_H^k]_i \nabla H_i(x^k) = 0,\ \mu_g^k \geq 0. \tag{12}$$

By a simple arrangement, it follows that

$$\nabla f(x^k) + w^k + \sum_{i \in \mathcal{I}_g^*} [\lambda_g^k]_i \nabla g_i(x^k) + \sum_{i=1}^{p} [\lambda_h^k]_i \nabla h_i(x^k) + \sum_{i \in \mathcal{I}_G^*} [\lambda_G^k]_i \nabla G_i(x^k) + \sum_{i \in \mathcal{I}_H^*} [\lambda_H^k]_i \nabla H_i(x^k) = 0. \tag{13}$$

where

$$w^k := \sum_{i \in \{1,\dots,m\} \backslash \mathcal{I}_*^g} [\lambda_g^k]_i \nabla g_i(x^k) + \sum_{i \in \mathcal{I}_H^* \cup \mathcal{I}_{GH}^*} [\lambda_G^k]_i \nabla G_i(x^k) + \sum_{i \in \mathcal{I}_G^* \cup \mathcal{I}_{GH}^*} [\lambda_H^k]_i \nabla H_i(x^k). \tag{14}$$

We now show that $\{w^k\}$ converges to 0 as $k \in K_1, k \to \infty$. By the assumptions, we assume that $K_1 \subset \{1, 2, \dots\dots\}$ is a subsequence such that $\{\mu_o^k | k \in K_1\}$ converges to $\bar{\mu}_o$ and $x^k \to x^*$ as $k \in K_1, k \to \infty$. It then follows from the first two conditions of (11) that

$$\{[\lambda_G^k]_{\mathcal{I}_H^* \cup \mathcal{I}_{GH}^*} | k \in K_1\} \to 0,\ \{[\lambda_H^k]_{\mathcal{I}_G^* \cup \mathcal{I}_{GH}^*} | k \in K_1\} \to 0,\ \text{as } k \to \infty. \tag{15}$$

Let us prove that $\{[\lambda_g^k]_{\{1,\dots,m\} \backslash \mathcal{I}_*^g} | k \in K_1\} \to 0$ as $k \to \infty$. We consider two cases: $\{\rho_k\}$ is bounded; and $\{\rho_k\}$ is unbounded. In the first case, from the implementation process of Algorithm 3.1, it follows that $\{\beta^k\}$ converges to 0 as $k \to \infty$. Thus, we readily have that $\tau^k \to 0$ as $k \to \infty$, which suggests the derived result immediately. In the second case, if $i \notin \mathcal{I}_*^g$, then $g_i(x^*) < 0$ and hence $g_i(x^k) < -\epsilon$ for some positive ϵ for all sufficiently large k. By the boundedness of $\{\bar{\mu}_g^k\}$ and the definition of $[\lambda_g]_i^k$, it is easy to verify that $[\lambda_g]_i^k = 0$ when $k \in K_1$ is sufficiently large.

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 $\mathcal{G}(x^*, \mathcal{I}_h, \mathcal{I}_G^*, \mathcal{I}_H^*)$. 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. By this fact and (13), there exists $\hat{\lambda}_h^k$, $\hat{\lambda}_G^k$, and $\hat{\lambda}_H^k$ such that

$$w^k + \nabla f(x^k) + \sum_{i \in \mathcal{I}_g^*} [\lambda_g^k]_i \nabla g_i(x^k) + \sum_{i \in I_1} [\hat{\lambda}_h^k]_i \nabla h_i(x^k) + \sum_{i \in I_2} [\hat{\lambda}_G^k]_i \nabla G_i(x^k) + \sum_{i \in I_3} [\hat{\lambda}_H^k]_i \nabla H_i(x^k) = 0. \tag{16}$$

When $w^k + \nabla f(x^k) = 0$ on some subsequence $K_2 \subseteq K_1$, it is easy to verify that $\nabla f(x^*) = 0$ which suggests that x^* is an S-stationary point.

From now on, we assume that $w^k + \nabla f(x^k) \neq 0$ all $k \in K_1$ sufficiently large. Then applying Lemma 3.1 to (16) yields that there exist $I_4^k \subseteq \mathcal{I}_g^*$ and $\bar{\lambda}_g^k \geq 0$, $\bar{\lambda}_h^k$, $\bar{\lambda}_G^k$, and $\bar{\lambda}_H^k$ such that

$$w^k + \nabla f(x^k) + \sum_{i \in I_4^k} [\bar{\lambda}_g^k]_i \nabla g_i(x^k) + \sum_{i \in I_1} [\bar{\lambda}_h^k]_i \nabla h_i(x^k) + \sum_{i \in I_2} [\bar{\lambda}_G^k]_i \nabla G_i(x^k) + \sum_{i \in I_3} [\bar{\lambda}_H^k]_i \nabla H_i(x^k) = 0, \tag{17}$$

and the vectors $\{\nabla g_i(x^k) : i \in I_4^k\}$, $\{\nabla h_j(x^k) : j \in I_1\}$, $\{\nabla G_r(x^k) : r \in I_2\}$, $\{\nabla H_t(x^k) : t \in I_3\}$ are linearly independent for all k sufficiently large. Without loss of generality, we assume that $I_4^k \equiv I_4$.

Applying Lemma 3.2 with $\alpha_k = 1$ to (17), there exists $\bar{\lambda}_g^*$, $\bar{\lambda}_h^*$, $\bar{\lambda}_G^*$ and $\bar{\lambda}_H^*$ such that

$$\nabla f(x^*) + \sum_{i \in I_4} [\bar{\lambda}_g^*]_i \nabla g_i(x^*) + \sum_{i \in I_1} [\bar{\lambda}_h^*]_i \nabla h_i(x^*) + \sum_{i \in I_2} [\bar{\lambda}_G^*]_i \nabla G_i(x^*) + \sum_{i \in I_3} [\bar{\lambda}_H^*]_i \nabla H_i(x^*) = 0,$$

which together with the facts that $I_4 \subseteq \mathcal{I}_g^*$, $I_2 \subseteq \mathcal{I}_G^*$, $I_3 \subseteq \mathcal{I}_H^*$ implies that x^* is an S-stationary point. □

Theorem 3.4. *Let $\{x^k\} \subset \mathbb{R}^n$ be a sequence generated by Algorithm 3.1 and let x^* be an accumulation point. Suppose that the sequence $\{\|\mu_o^k\|\}$ converges to ∞ as $k \to \infty$. If x^* is feasible to the MPSC (1) and MPSC-LICQ holds at x^*, then x^* is a W-stationary point of the MPSC (1).*

Proof. We assume without loss of generality that $\{x^k\}$ converges to x^* for some subsequence $K \subseteq \{1, 2, \ldots\ldots\}$ and $G(x^*) = 0$ (this can be achieved by moving the nonzero components of G to H and the corresponding zero components of H to G). It is easy to check from (2) that

$$\begin{aligned} \{i \in \{1, ..., l\} \mid H_i(x^*) = 0\} &= \mathcal{I}_{GH}^*, \\ \{i \in \{1, ..., l\} \mid H_i(x^*) \neq 0\} &= \mathcal{I}_G^*, \end{aligned} \tag{18}$$

and $\{\mathcal{I}_G^*, \mathcal{I}_{GH}^*\}$ is a disjoint partition of $\{1, \ldots, l\}$. According to the proof of Theorem 3.3, the equation (12) still holds true.

We first consider the case when $\{\rho_k\}$ is bounded. It then follows from (8) and (9) that $\{\mu_j^k | k \in K\}, \{\mu_0^k | k \in K\}, \{\mu_g^k | k \in K\}$ are bounded. Without loss of generality, we assume that they converges to $\{\mu_j^*\}, \{\mu_0^*\}, \{\mu_g^*\}$ respectively. Since $\{\mu_0^k\}$ is bounded, from the first two conditions of (11), the relation (15) holds. Moreover, from the implementation process of Algorithm 3.1, it follows that $\{\beta^k\}$ converges to 0 as $k \to \infty$. Thus, we readily have that $h(x^k) \to 0$, $G(x^k) \circ H(x^k) \to 0$ and $\tau^k \to 0$ as $k \to \infty$. These facts suggest that x^* is a feasible point and $\min\{\mu_g^*, -g(x^*)\} = 0$. Thus taking a limit in (12) implies that x^* is an S-stationary point.

From now on, we consider the case when $\rho^k \to \infty$. For convenience, we let

$$\mathcal{I}_\infty = \left\{ i \in \{1, \ldots, l\} \mid |[\mu_0^k]_i| \to \infty, \; as\; k \in K, \; k \to \infty \right\}.$$

According to the first condition of (9) and the third condition of (11), we readily have that for all $k \in K$ large enough,

$$[\lambda_g^k]_{\{1,\ldots,m\} \setminus \mathcal{I}_g^*} = 0$$

It then follows from the fact that $\mathcal{I}_G^* \cup \mathcal{I}_{GH}^* = \{1, \ldots, l\}$ and (12) that

$$\begin{aligned} 0 = \nabla f(x^k) + \sum_{i \in \mathcal{I}_g^*} [\lambda_g^k]_i \nabla g_i(x^k) + \sum_{i=1}^{p} [\lambda_h^k]_i \nabla h_i(x^k) + \sum_{i \in \mathcal{I}_{GH}^*} [\lambda_H^k]_i \nabla H_i(x^k) \\ + \sum_{i \in \mathcal{I}_{GH}^*} [\lambda_G^k]_i \nabla G_i(x^k) + \sum_{i \in \mathcal{I}_G^*} [\lambda_G^k]_i \nabla G_i(x^k) + \sum_{i \in \mathcal{I}_G^*} [\lambda_H^k]_i \nabla H_i(x^k). \end{aligned} \tag{19}$$

In order to apply Lemma 3.2, we need to show that $[\lambda_H^k]_i \to 0$ as $k \in K, k \to \infty$ for all $i \in \mathcal{I}_G^*$. First following from the second condition of (11), it is easy to verify that

$$[\lambda_H^k]_{\mathcal{I}_G^* \setminus \mathcal{I}_\infty} \to 0, \text{ as } k \in K, k \to \infty.$$

Let $\mathcal{I}_0 := \{i \in \mathcal{I}_G^* \cap \mathcal{I}_\infty | [\lambda_H^k]_i \to 0 \text{ as } k \in K, k \to \infty\}$ and $\mathcal{I}_0^c = \mathcal{I}_G^* \cap \mathcal{I}_\infty \setminus \mathcal{I}_0$. Let $\mathcal{I}_0^c \neq \emptyset$. We can find $\epsilon > 0$ and an infinite set $K_1 \subseteq K$ such that

$$|[\lambda_H^k]_i| \geq \epsilon, \quad i \in \mathcal{I}_0^c, k \in K_1.$$

Let $j \in \mathcal{I}_0^c$. According to the second condition of (8), it follows that

$$\frac{[\mu_0^k]_j}{\rho_{k-1}} = G_j(x^k) H_j(x^k) + \frac{[\bar{\mu}_0^{k-1}]_j}{\rho_{k-1}} \to 0, \text{ as } k \in K_1, k \to \infty \tag{20}$$

where the limit follows from the fact that $G_j(x^*) = 0$ and the boundedness of $\{\bar{\mu}_0^{k-1}\}$. It is not hard to see that for all $i \in \mathcal{I}_\infty$,

$$\frac{\rho^{k-1}}{[\mu_0^k]_i} G_i(x^k) H_i(x^k) = 1 - \frac{[\bar{\mu}_0^{k-1}]_i}{[\mu_0^k]_i} \to 1, \; as \; k \to \infty \tag{21}$$

where the equality follows from the second condition of (8). By (11) and (21), we have

$$\frac{\rho^{k-1} [\lambda_H^k]_j}{[\mu_0^k]_j^2} = \frac{\rho^{k-1}}{[\mu_0^k]_j} G_j(x^k) = (1 - \frac{[\bar{\mu}_0^{k-1}]_j}{[\mu_0^k]_j}) / H_j(x^k) \to \frac{1}{H_j(x^*)} \text{ as } k \to \infty. \tag{22}$$

This immediately implies that

$$\frac{(\rho^{k-1})^2 [\lambda_H^k]_j^2}{[\mu_o^k]_j^4} \to \frac{1}{H_j^2(x^*)}, \text{ as } k \to \infty. \tag{23}$$

Moreover, from (21) and the first two conditions of (11) it is easy to verify that

$$\frac{\rho^{k-1}}{[\mu_0^k]_j^3}[\lambda_H^k]_j[\lambda_G^k]_j \to 1 > 0 \text{ as } k \to \infty. \tag{24}$$

It then follows from (23) and (24) that

$$\frac{\mu_o^k}{\rho^{k-1}}\frac{[\lambda_G^k]_j}{[\lambda_H^k]_j} \to H_j^2(x^*) > 0 \text{ as } k \to \infty.$$

This together with (20) implies that

$$\frac{[\lambda_H^k]_j}{[\lambda_G^k]_j} \to 0 \text{ as } k \in K_1, k \to \infty.$$

Clearly, it follows that $\{[\lambda_G^k]_j | k \in K_1\} \to \infty$ as $k \to \infty$. Thus, we have

$$\frac{\|[\lambda_H^k]_{\mathcal{I}_0^c}\|}{\|([\lambda_g^k]_{\mathcal{I}_g^*}, [\lambda_h^k]_{\mathcal{I}_h}, [\lambda_H^k]_{\mathcal{I}_G^*}, \lambda_G^k, [\lambda_H^k]_{\mathcal{I}_{GH}^*})\|} \le \frac{\|[\lambda_H^k]_{\mathcal{I}_0^c}\|}{\|[\lambda_G^k]_{\mathcal{I}_0^c}\|} \to 0.$$

Dividing (19) by $\|([\lambda_g^k]_{\mathcal{I}_g^*}, [\lambda_h^k]_{\mathcal{I}_h}, [\lambda_g^k]_{\mathcal{I}_g^*}, \lambda_G^k, [\lambda_H^k]_{\mathcal{I}_{GH}^*})\|$ and applying Lemma 3.2 for $k \in K_1$ with

$$\alpha_k = \frac{1}{\|([\lambda_g^k]_{\mathcal{I}_g^*}, [\lambda_h^k]_{\mathcal{I}_h}, [\lambda_H^k]_{\mathcal{I}_G^*}, \lambda_G^k, [\lambda_H^k]_{\mathcal{I}_{GH}^*})\|}, \; w^k = \frac{\sum_{i \in \mathcal{I}_G^*}[\lambda_H^k]_i \nabla H_i(x^k)}{\|([\lambda_g^k]_{\mathcal{I}_g^*}, [\lambda_h^k]_{\mathcal{I}_h}, [\lambda_H^k]_{\mathcal{I}_G^*}, \lambda_G^k, [\lambda_H^k]_{\mathcal{I}_{GH}^*})\|}.$$

Since $\alpha_k \to 0$ and $w^k \to 0$, we obtain that

$$\frac{([\lambda_g^k]_{\mathcal{I}_g^*}, [\lambda_h^k]_{\mathcal{I}_h}, [\lambda_H^k]_{\mathcal{I}_G^*}, \lambda_G^k, [\lambda_H^k]_{\mathcal{I}_{GH}^*})}{\|([\lambda_g^k]_{\mathcal{I}_g^*}, [\lambda_h^k]_{\mathcal{I}_h}, [\lambda_H^k]_{\mathcal{I}_G^*}, \lambda_G^k, [\lambda_H^k]_{\mathcal{I}_{GH}^*})\|} \to 0 \text{ as } k \in K_1, k \to \infty,$$

which is evidently impossible. Thus, we have that $\mathcal{I}_0^c = \emptyset$. It then follows that

$$\sum_{i \in \mathcal{I}_G^*}[\lambda_H^k]_i \nabla H_i(x^k) = \sum_{i \in \mathcal{I}^0 \cup (\mathcal{I}_G^* \setminus \mathcal{I}_\infty)}[\lambda_H^k]_i \nabla H_i(x^k) \to 0 \text{ as } k \in K_1, k \to \infty.$$

Thus, applying Lemma 3.2 to (19) for $k \in K$ with

$$\alpha^k = 1, \quad w^k = \sum_{i \in \mathcal{I}_G^*}[\lambda_H^k]_i \nabla H_i(x^k)$$

suggests that there exist $[\lambda_g^*]_i, i \in \mathcal{I}_g^*$, λ_h^*, λ_G^*, and $[\lambda_H^*]_i, i \in \mathcal{I}_{GH}^*$ such that

$$\nabla f(x^*) + \sum_{i \in \mathcal{I}_g^*}[\lambda_g^*]_i \nabla g_i(x^*) + \sum_{i=1}^{p}[\lambda_h^*]_i \nabla h_i(x^*) \\ + \sum_{i \in \mathcal{I}_{GH}^*}[\lambda_H^*]_i \nabla H_i(x^*) + \sum_{i=1}^{l}[\lambda_G^*]_i \nabla G_i(x^*) = 0.$$

By the definition of the W-stationarity, we derive that x^* is a weakly stationary point. □

3.2. **Augmented Lagrangian method for the EOCP.** In portfolio selection problems, a minimum buy-in threshold is introduced to prevent the investors from holding a small position for assets [4]. In other words, the investor needs to make a portfolio selection that is either zero or above a positive weight. This class of problems can be modelled as the following either-or-constrained program (EOCP).

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0,\ i = 1, \dots, m, \\ & h_j(x) = 0,\ j = 1, \dots, p, \\ & c_k^1(x) \le 0 \ \vee\ c_k^2(x) \le 0,\ k = 1, \dots, l, \end{array} \tag{25}$$

where $\vee$ denotes the logical "or", and all functions $c_1^1, \dots, c_l^1, c_1^2, \dots, c_l^2 \colon \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable. For simplicity, we let $c^1 = (c_1^1, \dots, c_l^1)^\top$ and $c^2 = (c_1^2, \dots, c_l^2)^\top$. The constraints $c_k^1(x) \le 0 \vee c_k^2(x) \le 0$ $k = 1, \dots, l$ are called *either-or-constraints*. A binary variable can be introduced to transform the EOCP (25) into a mixed integer program but it makes the numerical treatment of this either-or-constraints challenging [8]. Another possible approach to deal with this either-or-constraints is to equivalently transform them into $\min\{c_k^1(x), c_k^2(x)\} \le 0$. In order to avoid the inherent nonsmoothness of "min" function, two variables $z_k^1, z_k^2 \in \mathbb{R}$ are introduced and the either-or-constraints are reformulated as the so-called *switching constraints*

$$(c_k^1(x) - z_k^1)(c_k^2(x) - z_k^2) = 0,\ z_k^1, z_k^2 \le 0.$$

The EOCP (25) is then reformulated as the following MPSC

$$\begin{array}{ll} \min_{x,z} & f(x) \\ \text{s.t.} & g_i(x) \le 0,\ i = 1, \dots, m, \\ & h_j(x) = 0,\ j = 1, \dots, p, \\ & z_k^1, z_k^2 \le 0,\ k = 1, \dots, l, \\ & (c_k^1(x) - z_k^1)(c_k^2(x) - z_k^2) = 0,\ k = 1, \dots, l. \end{array} \tag{26}$$

The equivalence between problem (25) and problem (26) is discussed in [8] from the viewpoint of global and local optimal solutions respectively.

In the following, we employ Algorithm 3.1 to solve problem (25). Fix a feasible point $x^* \in \mathbb{R}^n$ of the EOCP (25) and define the following index sets

$$\begin{aligned} I_{0-}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) = 0 \wedge c_k^2(x^*) < 0\}, \\ I_{-0}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) < 0 \wedge c_k^2(x^*) = 0\}, \\ I_{0+}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) = 0 \wedge c_k^2(x^*) > 0\}, \\ I_{+0}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) > 0 \wedge c_k^2(x^*) = 0\}, \\ I_{-+}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) < 0 \wedge c_k^2(x^*) > 0\}, \\ I_{+-}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) > 0 \wedge c_k^2(x^*) < 0\}, \\ I_{00}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) = 0 \wedge c_k^2(x^*) = 0\}, \\ I_{--}^* &:= \{k \in \{1, \dots, l\} | c_k^1(x^*) < 0 \wedge c_k^2(x^*) < 0\}. \end{aligned}$$

We recall some dual stationarity concepts of the EOCP(25) as given in [8].

Definition 3.5. Let $x^* \in \mathbb{R}^n$ be feasible to the EOCP (25). Then, x^* is called to be

(1) W-stationary, if there exist multipliers which solve the following system:

$$\left\{\begin{array}{ll} 0= & \nabla f(x^*)+\sum_{i\in\mathcal{I}_g^*}[\lambda_g]_i\nabla g_i(x^*)+\sum_{i\in\mathcal{I}_h}[\lambda_h]_i\nabla h_i(x^*) \\ & +\sum_{k\in I_{0+}^*\cup I_{00}^*}[\lambda_{c^1}]_k\nabla c_k^1(x^*)+\sum_{k\in I_{+0}^*\cup I_{00}^*}[\lambda_{c^2}]_k\nabla c^2(x^*), \\ & \forall i\in\mathcal{I}_g^*,\ [\lambda_g]_i\ge 0, \\ & k\in I_{0+}^*\cup I_{00}^*,\ [\lambda_{c^1}]_k\ge 0, \\ & k\in I_{+0}^*\cup I_{00}^*,\ [\lambda_{c^2}]_k\ge 0. \end{array}\right. \tag{27}$$

(2) S-stationary, if there exist multipliers which satisfy the conditions (27) and

$$[\lambda_{c^1}]_k=0\wedge[\lambda_{c^2}]_k=0,\ k\in I_{00}^*.$$

The following lemma follows from Section 7.2 of [8].

Lemma 3.6. *If (x^*,z^*) is a W-stationary (S-stationary) point of problem (26), then x^* is a W-stationary (S-stationary) point of the EOCP (25).*

With the help of Theorems 3.3 and 3.4, and Lemma 3.6, it is easy to obtain the convergence results of the ALM for solving the EOCP (25).

Corollary 1. *Let $\{(x^k,z^k)\}\subset\mathbb{R}^{n+l}$ be a sequence generated by Algorithm 3.1 applied to problem (26), and let $(x^*,z^*)\in\mathbb{R}^{n+l}$ be an accumulation point. Suppose that the point (x^*,z^*) is feasible to problem (26) and that MPSC RCPLD holds at (x^*,z^*). Moreover, we suppose that the sequence $\{\mu_o^k\}$ generated by Algorithm 3.1 has a bounded subsequence such that the corresponding subsequence of $\{(x^k,z^k)\}$ converges to (x^*,z^*). Then x^* is an S-stationary point of the EOCP (25).*

Corollary 2. *Let $\{(x^k,z^k)\}\subset\mathbb{R}^{n+l}$ be a sequence generated by Algorithm 3.1 applied to problem (26), and let $(x^*,z^*)\in\mathbb{R}^{n+l}$ be an accumulation point. Suppose that the point (x^*,z^*) is feasible to problem (26) and that MPSC-LICQ holds at (x^*,z^*). Moreover, we suppose that the sequence $\{\|\mu_o^k\|\}$ generated by Algorithm 3.1 tending to ∞ as $k\to\infty$. Then x^* is a W-stationary point of the EOCP (25).*

4. **Numerical experiments.** In this section we conduct numerical experiments to test the performance of Algorithm 3.1, which is coded in MATLAB. We apply it to solve some examples in the MPSC literature. All computations are performed on a computer with 2000 MHz Inter(R) Core i5 processor and all the augmented Lagrangian subproblems are solved by using the function "*fmincon*" in Optimization Toolbox with default tolerance in MATLAB R2016b. We use $\beta^{k+1}<10^{-3}$ as the termination condition of Algorithm 3.1 and we set $\rho_0=2$, $\theta=0.8$, $\sigma=10$, $\bar{\mu}_h^{\max}=\bar{\mu}_g^{\max}=\bar{\mu}_0^{\max}=10^5$, $\bar{\mu}_h^{\min}=\bar{\mu}_0^{\min}=-10^5$. The initial multiplier vector is chosen as an all-eights vector.

4.1. **Either-or-constrained examples.**

Example 4.1. [9, Subsection 6.2.1] Let us consider the problem

$$\begin{array}{ll} \min & (x_1-8)^2+(x_2+3)^2 \\ \text{s.t.} & x_1-2x_2+4\le 0\ \vee\ x_1-2\le 0, \\ & x_1^2-4x_2\le 0\ \vee\ (x_1-3)^2+(x_2-1)^2-10\le 0. \end{array} \tag{28}$$

It is easy to see that problem (28) possesses the unique global minimizer $(2,-2)$ and another local minimizer $(4,4)$. We can transform (28) into an MPSC by introducing additional variables:

$$\begin{array}{ll} \min & (x_1-8)^2+(x_2+3)^2 \\ \text{s.t.} & z_1, z_2, z_3, z_4 \le 0, \\ & (x_1-2x_2+4-z_1)(x_1-2-z_2)=0, \\ & (x_1^2-4x_2-z_3)((x_1-3)^2+(x_2-1)^2-10-z_4)=0. \end{array}$$

Example 4.2. [8, Example 7.1] Let us consider the problem

$$\begin{array}{ll} \min & (x_1-1)^2 \\ \text{s.t.} & x_1 \le 0 \ \vee \ x_2 \le 0, \end{array} \tag{29}$$

as well as its MPSC reformulation

$$\begin{array}{ll} \min & (x_1-1)^2 \\ \text{s.t.} & z_1, z_2 \le 0, \\ & (x_1-z_1)(x_2-z_2)=0. \end{array}$$

The set of global minimizers of problem (29) is given by $\{(1,t) | t \le 0\}$. There are no local minimizers of (29) which are different from its globally minimizers.

Example 4.3. [8, Example 7.2] Let us consider the problem

$$\begin{array}{ll} \min & (x_1+1)^2+x_2^2 \\ \text{s.t.} & -x_1+x_2 \le 0, \\ & -x_1 \le 0 \ \vee \ -x_2 \le 0, \end{array}$$

with a unique global minimizer $(0,0)$. Its MPSC reformulation is

$$\begin{array}{ll} \min & (x_1+1)^2+x_2^2 \\ \text{s.t.} & z_1, z_2 \le 0, -x_1+x_2 \le 0, \\ & (-x_1-z_1)(-x_2-z_2)=0. \end{array}$$

4.2. **MPSC examples.**

Example 4.4. [8, Example 4.1] Consider the problem

$$\begin{array}{ll} \min & (x_1-2)^2+(x_2-1)^2+(x_3-2)^2 \\ \text{s.t.} & x_1^2+x_2^2+x_3^2 \le 3, x_3 \le 1, \\ & (x_1-x_2^2)(x_2-x_1^2)=0. \end{array}$$

It is easy to see that $(1,1,1)$ is a globally optimal solution.

Example 4.5. [8, Example 5.1] Consider the problem

$$\begin{array}{ll} \min & (x_1-1)^2+x_2^2 \\ \text{s.t.} & x_1x_2=0, -x_2 \le 0, x_1 \le 0, \end{array}$$

whose unique global minimizer is $(0,0)$.

Example 4.6. [8, Example 5.2] Consider the problem

$$\begin{array}{ll} \min & x_1+x_2^2 \\ \text{s.t.} & -x_1+x_2 \le 0, x_1x_2=0, \end{array}$$

whose unique global minimizer is $(0,0)$.

Example 4.7. [8, Example 5.3] Consider the problem

$$\begin{array}{ll} \min & x_1+x_2 \\ \text{s.t.} & x_1^2-x_2 \le 0, x_1x_2=0, \end{array}$$

whose unique global minimizer is $(0,0)$.

TABLE 1. Numerical results of Algorithm 3.1 for solving Examples 4.1–4.7

	ExaSolut	AppSolut	β^{k+1}
Example 4.1	(2,-2)	(2.0,-2.0)	5.8e-4
Example 4.2	$\{(1,t)\|t \leq 0\}$	(1.0,-1.9)	3.2e-5
Example 4.3	(0,0)	(-0.0, 0.0)	5.1e-5
Example 4.4	(1,1,1)	(1.0,1.0, 1.0)	2.5e-6
Example 4.5	(0,0)	(0.0,0.0)	6.6e-6
Example 4.6	(0,0)	(-0.0,-0.0)	1.3e-7
Example 4.7	(0,0)	(-0.0,0.0)	9.5e-5

We summarize in Table 1 the numerical results of Algorithm 3.1 for solving Examples 4.1–4.7. The columns of the table list the exact solution labeled by **ExaSolut** and the approximate solution obtained by Algorithm 3.1 labeled by **AppSolut**. For either-or-constrained examples (Examples 4.1–4.3), we use Algorithm 3.1 to solve their MPSC reformulations and the exact and approximate solutions reported in Table 1 are those of their original problems. From Table 1, we see that the proposed method can find good approximate solutions.

4.3. **Portfolio optimization examples.** We present a concrete example of portfolio optimization based on the test examples in [4], which is described as

$$\begin{array}{ll} \min & x^\top Q x \\ \text{s.t.} & e^\top x = 1, \\ & \mu^\top x \geq \rho, \\ & x_i = 0 \ \vee \ x_i \in [l_i, u_i], \ i = 1, ..., q, \end{array} \tag{30}$$

with randomly generated $Q \in \mathbb{R}^{n \times n}$, $\mu, l, u \in \mathbb{R}^n$, and $\rho \in \mathbb{R}$. Here, $e \in \mathbb{R}^n$ represents the all-ones vector. More details can be found in [4]. The decision variable x need to fulfill the constraints

$$x_i = 0 \ \vee \ x_i \in [l_i, u_i], \ i = 1, ..., q.$$

That is, x is required to lie in some continuous interval, except for the outlier case when it is equal to zero. In the context, this variable is often called semi-continuous. Assuming that $u_i \geq 0$ holds for all $i = 1, ..., q$, we can simply treat the requirement $x_i \leq u_i$ as a standard inequality constraint which should be fulfilled at all times. Clearly, if $x_i = 0$ is valid, then the inequality $x_i \leq u_i$ holds automatically, so that we can rewrite problem (30) as the following EOCP

$$\begin{array}{ll} \min & x^\top Q x \\ \text{s.t.} & e^\top x = 1, \\ & \mu^\top x \geq \rho, u - x \geq 0, \\ & x_i = 0 \ \vee \ x_i \geq l_i \ i = 1, ..., q. \end{array} \tag{31}$$

Then it is easy to transform problem (31) into the following MPSC problem

$$\begin{array}{ll} \min & x^\top Q x \\ \text{s.t.} & e^\top x - 1 = 0, \\ & \rho - \mu^\top x \leq 0, \\ & x - u \leq 0, \ -y \leq 0, \\ & x_i(x_i - l_i - y_i) = 0, \ i = 1, ..., q. \end{array}$$

We apply Algorithm 3.1 to solve problem (31) with $\rho = 1$, $l_i = -10$, $u_i = 20$ for $i = 1, ..., q$, in which Q is randomly generated. The numerical results of Algorithm 3.1 for solving problem (31) are presented in Table 2. The problem size **n** is listed in the first column, the termination state β^{k+1} is listed in the second column, and the CPU time in second of the proposed method is listed in the third column labeled by **CPU**.

TABLE 2. Numerical results of Algorithm 3.1 for solving problem (31) with random instances

n	β^{k+1}	CPU
10	1.1e-5	5.7
50	9.5e-4	66.2
100	2.6e-4	110.9
150	1.6e-4	149.7

4.4. **Comparison with relaxation methods.** In this subsection, we give some examples given in [9] in which only W-stationary points are obtained by the recently proposed relaxation method but better S-stationary points are derived by the proposed method in this paper.

Example 4.8. [9, Example 3.3] Let us consider the MPSC problem

$$\begin{array}{ll} \min & \frac{1}{2}(x_1-1)^2 + \frac{1}{2}(x_2-1)^2 \\ \text{s.t.} & x_1 x_2 = 0. \end{array}$$

Its global minimizers are $(1, 0)$ and $(0, 1)$, which are both S-stationary. Furthermore, there exists a W-stationary point $(0, 0)$ which is not a local minimizer. In [9], they showed the Scholtes' relaxation scheme can only find the W-stationary point $(0, 0)$. Using Algorithm 3.1 to solve the problem, an S-stationary point can be found (see Table 3).

Example 4.9. [9, Example 5.1] Let us consider the MPSC problem

$$\begin{array}{ll} \min & x_1 x_2 - x_1 - x_2 \\ \text{s.t.} & x_1^2 + x_2^2 - 1 \leq 0, \\ & x_1 x_2 = 0. \end{array}$$

It is easy to verify that the global minimizers are $(1, 0)$ and $(0, 1)$, which are both S-stationary. Furthermore, there is a W-stationary point $(0, 0)$ which is not a local minimizer. The relaxation scheme of Steffensen and Ulbrich can only find W-stationary point $(0, 0)$ (see, e.g., [9]). An S-stationary point can be found by using Algorithm 3.1 to solve the problem (see Table 3).

We summarize in Table 3 the numerical results of Algorithm 3.1 for solving Examples 4.8 and 4.9. In this table, **S-Station** and **W-Station** respectively represent the S-stationary and W-stationary points of these examples. It is easy to see that the proposed method can find better solutions than the relaxation method.

Table 3. Numerical results of Algorithm 3.1 for solving Examples 4.8 and 4.9

	S-Station	W-Station	AppSolut	β^{k+1}
Example 4.8	(0,1), (1,0)	(0,0)	(0,1)	3.9e-5
Example 4.9	(0,1), (1,0)	(0,0)	(0,1)	2.6e-5

REFERENCES

[1] W. Achtziger , and C. Kanzow, Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications, *Math. Program.* **114**(2008),69–99.

[2] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt, On augmented Lagrangian methods with general lower-level constraints, *SIAM J. Optim.* **18**(2007), 1286–1309.

[3] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. Silva, A relaxed constant positive linear dependence constraint qualification and applications, *Math. Program.* **135**(2012), 255–273.

[4] A. Frangioni, C.Gentile, SDP diagonalizations and perspective cuts for a class of nonseparable MIQP, *Oper. Res. Lett.* **35**(2007), 181–185.

[5] X. Huang, X. Q. Yang and D. L. Zhu, A sequential smooth penalization approach to mathematical programs with complementarity constraints, *Numer. Funct. Anal. Optim.* **27**(2006), 71–78.

[6] A. F. Izmailov, and M. V. Solodov, Mathematical programs with vanishing constraints: optimality conditions, sensitivity, and a relaxation method, *J. Optim. Theory Appl.* **142**(2009), 501–532.

[7] A. F. Izmailov, M. V. Solodov, and E. I. Uskov, Global convergence of augmented lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints, *SIAM J. Optim.* **22**(2012), 1579–1606.

[8] P. Mehlitz, Stationary conditions and constraint qualification for mathematical programs with switching constraints, *Math. Program.* **181**(2020), 149-186.

[9] C. Kanzow, P. Mehlitz, and D. Steck, Relaxation schemes for mathematical programs with switching constraints, *Optim. Methods Softw.* DOI: 10.1080/10556788.2019.1663425.

[10] C. Kanzow, and A. Schwartz, Mathematical programs with equilibrium constraints: enhanced fritz john-conditions, new constraint qualifications, and improved exact penalty results, *SIAM J. Optim.* **20**(2010), 2730–2753.

[11] G. Li, and L. Guo, Mordukhovich stationarity for mathematical programs with switching constraints under weak constraint qualifications, *Optimization Online*, http://www.optimization-online.org/DB_HTML/2019/07/7288.html.

[12] G. H. Lin, and M. Fukushima, New reformulations for stochastic nonlinear complementarity problems, *Optim. Methods Softw.* **21**(2006), 551–564.

[13] Y. C. Liang and J. J. Ye, Optimality conditions and exact penalty for mathematical programs with switching constraints. *J Optim Theory Appl.* **190**(2021), 1–31.

[14] P. Mehlitz, On the linear independence constraint qualification in disjunctive programming, *Optimization*, **10**(2020), 2241–2277.

[15] Z.-Q. Luo, J.-S. Pang, and D. Ralph, Mathematical programs with equilibrium constraints, *Cambridge University Press*, 1996.

[16] Z. Wan and Y. Wang, Convergence of an inexact smoothing method for mathematical programs with equilibrium constraints, *Numer. Funct. Anal. Optim.* **27**(2006), 485–495.

[17] X. Q. Yang and X. X. Huang, Lower-order penalty methods for mathematical programs with complementarity constraints, *Optim. Methods Softw.* **19**(2004), 693–720.

Received xxxx 20xx; revised xxxx 20xx; early access xxxx 20xx.

E-mail address: ligaoxicn@126.com (G. Li)