Journal of Computational Mathematics
Vol.41, No.1, 2023, 133–152.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2106-m2020-0059

# A SSLE-TYPE ALGORITHM OF QUASI-STRONGLY SUB-FEASIBLE DIRECTIONS FOR INEQUALITY CONSTRAINED MINIMAX PROBLEMS *

Jinbao Jian and Guodong Ma[1)]

*College of Mathematics and Physics, Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Center for Applied Mathematics and Artificial Intelligence, Guangxi University for Nationalities, Nanning 530006, China*

*Email: jianjb@gxu.edu.cn, mgd2006@163.com*

Yufeng Zhang

*School of Mathematics and Information Science, Guangxi University, Nanning 530004, China*

*Email: 1178452540@qq.com*

**Abstract**

In this paper, we discuss the nonlinear minimax problems with inequality constraints. Based on the stationary conditions of the discussed problems, we propose a sequential systems of linear equations (SSLE)-type algorithm of quasi-strongly sub-feasible directions with an arbitrary initial iteration point. By means of the new working set, we develop a new technique for constructing the sub-matrix in the lower right corner of the coefficient matrix of the system of linear equations (SLE). At each iteration, two systems of linear equations (SLEs) with the same uniformly nonsingular coefficient matrix are solved. Under mild conditions, the proposed algorithm possesses global and strong convergence. Finally, some preliminary numerical experiments are reported.

*Mathematics subject classification:* 90C30, 90C47, 65K05.

*Key words:* Inequality constraints, Minimax problems, Method of quasi-strongly sub-feasible directions, SSLE-type algorithm, Global and strong convergence.

## 1. Introduction

The minimax optimization may occur in engineering design [1], control system design [2], portfolio optimization [3], or as subproblems of algorithms for in semi-infinite minimax problems [4]. In this work, we discuss the nonlinear minimax problem with inequality constraints of the form

$$\begin{array}{ll} \min & F(x) \\ \text{s.t.} & f_i(x) \le 0, \quad i \in J = \{m+1, \cdots, m+l\}, \end{array} \tag{1.1}$$

where $F(x) = \max\{f_i(x), i \in I\}$ with $I = \{1, 2, \cdots, m\}$, and $f_i(x)(i \in I \cup J) : \mathbb{R}^n \to \mathbb{R}$. Since the objective function of this minimax problem is continuous but non-differentiable, the classical methods of smooth nonlinear programming cannot be used directly to solve the problem. Fortunately, by introducing an additional variable, a minimax problem can be equivalently reformulated as a smooth nonlinear programming, and many algorithms have been proposed,

* Received March 12, 2020 / Revised version received September 16, 2020 / Accepted June 9, 2021 / Published online April 1, 2022 /

1) Corresponding author

such as the sequential quadratic programming type methods [5–7], the interior point algorithm [8], the exponential smoothing algorithm [9], the trust-region method [10] and the sequential quadratically constrained quadratic programming algorithm [11].

It is known that the SSLE method (also call QP-free algorithm) is one of the effective methods for solving smooth nonlinear constrained optimization. In 1988, based on the KKT conditions, Panier, Tits and Herskovits [12] proposed a feasible QP-free algorithm for inequality constrained optimization, where two linear systems with a same coefficient matrix and a least squares subproblem need to be solved at each iteration. To overcome the calculation of the least squares subproblem, Yang, Li and Qi [13] proposed a feasible SSLE algorithm, where three reduced linear systems need to be solved at each iteration. Furthermore, to improve the convergence properties and numerical performance, many efforts have been made for research on SSLE-type (or QP-free) algorithms, in Refs. [14–18]. In fact, a feasible point is required to initialize the algorithm for the methods of feasible direction [6, 11–15], to overcome such kind of difficulty in a more general context, Jian and his collaborators proposed a method of strongly sub-feasible directions (MSSFD), see [18–20] and [21, Chapter 2]. The main features of the MSSFD can be described as follows: the initial point can be chosen arbitrarily without using any penalty parameters or penalty functions; the operations of initialization (Phase I) and optimization (Phase II) can be well unified automatically; the feasibility of a constraint is maintained through the iterations once it is reached, and therefore the number of feasible constraints is nondecreasing.

Recently, by improving the MSSFD, Jian et al. [22] presented the method of quasi-strongly sub-feasible directions (MQSSFD), the main characteristic is that the request $I^-(x^k) \subseteq I^-(x^{k+1})$ in MSSFD is relaxed by $|I^-(x^k)| \leq |I^-(x^{k+1})|$, where $|I^-(x^k)|$ means the number of functions satisfying the inequality constraints at $x^k$, thus the step size yielded by MQSSFD is larger than MSSFD. Furthermore, combining the idea of MQSSFD, Ma and Jian [16] presented a new QP-free algorithm for inequality constrained optimization with an arbitrary initial iteration point. At each iteration, this algorithm solves only two SLEs with a same uniformly nonsingular coefficient matrix to obtain the search direction. And the QP-free algorithm possesses nice theoretical properties and numerical results. In addition, the superiority of numerical performance for MQSSFD has also been verified by norm-relaxed sequential quadratic programming method in [23].

In this work, motivated by the idea of MQSSFD and QP-free algorithm for nonlinear inequality constrained optimization in Ref. [16], and based on the stationary point conditions of the nonlinear minimax problems with inequality constraints, we propose a SSLE-type algorithm of quasi-strongly sub-feasible directions for solving problem (1.1), in which the initial point is arbitrary. And our algorithm possesses the following features:

- A new technique for constructing the submatrix $\bar{F}_k$ in the lower right corner of the coefficient matrix is presented, thus the coefficient matrix possesses good sparsity;
- A new working set technique is introduced, then only functions indexed by working sets are considered, which can reduce the scale of the subproblems;
- At each iteration, two SLEs with a same uniformly nonsingular coefficient matrix need to be solved, which further reduce the computation cost;
- Under mild conditions, the proposed algorithm has global and strong convergence.

The paper is organized as follows. The next section describes the algorithm. Section 3 discusses the convergence analysis. Section 4 contains numerical experiments. Finally, a conclusion is given in Section 5.

## 2. Description of Algorithm

For convenience of discussion, we introduce the notations for the discussed problem (1.1) as follows.

$$\begin{array}{l}\varphi(x) := \max\{0,\ f_i(x), i\in J\},\ X := \{x:\ f_i(x)\le 0, i\in J\},\quad I(x) := \{i\in I: f_i(x) = F(x)\},\\ l_x := \min\{i:\ i\in I(x)\},\ J^+(x) := \{i\in J: f_i(x) > 0\},\quad J^-(x) := \{i\in J: f_i(x)\le 0\ \},\\ J(x) := \{i\in J: f_i(x) = 0 \text{ or } f_i(x) = \varphi(x)\},\quad L(x) := I(x)\cup J(x),\\ F_k := F(x^k),\quad \varphi_k := \varphi(x^k),\quad f_i^k := f_i(x^k),\ g_i^k := g_i(x^k) = \nabla f_i(x^k),\quad i\in I\cup J.\end{array}$$

A point $x\in\mathbb{R}^n$ is said to be a stationary point of the problem (1.1) with a multiplier vector $\lambda_{I\cup J}$ if the following relations hold

$$\left\{\begin{array}{l}\sum\limits_{i\in I\cup J}\lambda_i\nabla f_i(x) = 0,\quad \sum\limits_{i\in I}\lambda_i = 1;\quad \lambda_i\ge 0,\quad \lambda_i(F(x)-f_i(x)) = 0,\quad i\in I;\\ \lambda_i\ge 0,\quad \lambda_i f_i(x) = 0,\ i\in J;\quad \varphi(x) = 0.\end{array}\right. \tag{2.1}$$

To construct our search direction, the following basic assumption is necessary.

**Assumption 2.1.** (i) The functions $f_i$ $(i\in I\cup J)$ are all first order continuously differentiable over $\mathbb{R}^n$; (ii) For each $x\in\mathbb{R}^n$, the gradient vectors $\{\nabla f_i(x)-\nabla f_{l_x}(x),\ i\in I(x)\backslash\{l_x\};\nabla f_i(x),\ i\in J(x)\}$ are linearly independent.

**Remark 2.1.** As mentioned in Ref. [6], it is not difficult to show that Assumption 2.1 (ii) is really equivalent to: for each $t\in I(x)$, the vectors $\{\nabla f_i(x)-\nabla f_t(x),\ i\in I(x)\backslash\{t\};\nabla f_i(x), i\in J(x)\}$ are linearly independent.

For the current iteration point $x^k\in\mathbb{R}^n$, and a given suitable positive parameter $\varepsilon_k > 0$, then we generate the working set $L_k$,

$$J_k := \{i\in J^-(x^k):\ f_i^k\ge -\delta_k\}\cup\{i\in J^+(x^k): f_i^k-\varphi_k\ge -\delta_k\}, \tag{2.2}$$

$$I_k := \{i\in I: -\delta_k\le f_i^k - F_k\},\ l_k := l_{x^k},\quad I_k^0 := I_k\backslash\{l_k\},\ L_k := I_k^0\cup J_k, \tag{2.3}$$

where

$$\delta_0 = \varepsilon_0,\quad \delta_k = \min\{\varepsilon_k,\ \rho_{k-1}\},\quad k\ge 1, \tag{2.4}$$

and the optimal identification value $\rho_{k-1}$ is associated with the previous iterate $x^{k-1}$. The detailed computation of $\rho_{k-1}$ is shown in formula (2.17) below.

Based on the working set $L_k$, as well as $I(x^k)\subseteq I_k$ and $J(x^k)\subseteq J_k$, relations (2.1) are equivalent to the following stationary conditions with a multiplier vector $\lambda_{L_k}^k$

$$\left\{\begin{array}{l}\sum\limits_{i\in I_k^0}\lambda_i^k(g_i^k-g_{l_k}^k)+\sum\limits_{i\in J_k}\lambda_i^k g_i^k = -g_{l_k}^k,\quad \lambda_{l_k}^k := 1-\sum\limits_{j\in I_k^0}\lambda_i^k\ge 0;\\ \lambda_i^k\ge 0,\quad \lambda_i^k(F_k-f_i^k) = 0,\quad i\in I_k^0;\quad \lambda_i^k\ge 0,\quad \lambda_i^k f_i^k = 0,\quad i\in J_k;\ \varphi_k = 0.\end{array}\right. \tag{2.5}$$

In order to derive a suitable coefficient matrix in our SLEs, let us define

$$A_k := A_{L_k}(x^k) = (\bar g_i(x^k),\ i\in L_k),\quad \bar g_i^k := \bar g_i(x^k) := \left\{\begin{array}{ll} g_i^k-g_{l_k}^k, & i\in I_k^0;\\ g_i^k, & i\in J_k,\end{array}\right. \tag{2.6}$$

$$\bar F_k := \mathrm{diag}(\bar F_{ki},\ i\in L_k),\quad \bar F_{ki} = \left\{\begin{array}{ll} 0, & \text{if } \det(A_k^TA_k)\ge\varepsilon_k;\\ \bar f_i^k, & \text{if } \det(A_k^TA_k)<\varepsilon_k,\end{array}\right. \tag{2.7}$$

and

$$\bar{f}_i^k := \bar{f}_i(x^k) := \begin{cases} F_k - f_i^k, & i \in I_k^0; \\ \varphi_k - f_i^k, & i \in J^+(x^k) \cap J_k; \\ -f_i^k, & i \in J^-(x^k) \cap J_k. \end{cases} \tag{2.8}$$

Therefore, (2.5) implies, if $(x^k, \lambda_{L_k}^k)$ is a stationary pair, that $(0, \lambda_{L_k}^k)$ is a solution to the SLE

$$\begin{pmatrix} H_k & A_k \\ A_k^T & -\bar{F}_k \end{pmatrix} \begin{pmatrix} 0 \\ \lambda_{L_k}^k \end{pmatrix} = \begin{pmatrix} -g_{l_k}^k \\ 0 \end{pmatrix}, \tag{2.9}$$

where $H_k$ is an $n \times n$ matrix.

Based on the analysis above, we introduce the coefficient matrix in our SLEs as follows

$$V_k := V(x^k, H_k, L_k) := \begin{pmatrix} H_k & A_k \\ A_k^T & -\bar{F}_k \end{pmatrix}. \tag{2.10}$$

Next, in order to guarantee the non-singularity of the coefficient matrix $V_k$, the matrix $H_k$ needs to satisfy the general hypothesis, and it is given in the following Lemma 2.1 (or Assumption 3.1). Then the invertibility of the coefficient matrix $V_k$ constructed above is proved below.

**Lemma 2.1.** *Suppose that Assumption* 2.1 *holds, and the matrix $H_k$ is positive definite on the null space $NS_k := \{d \in \mathbb{R}^n : (\bar{g}_i^k)^T d = 0, i \in L(x^k) \backslash \{l_k\}\}$. Then the matrix $V_k$ yielded by* (2.10) *is invertible.*

*Proof.* It is sufficient to show that the system of equations $V_k(u^T, v^T)^T = 0$ has the unique solution zero. Taking into account (2.7), we consider the following two cases, i.e., $\det(A_k^T A_k) < \varepsilon_k$ and $\det(A_k^T A_k) \geq \varepsilon_k$.

(i) If $\det(A_k^T A_k) < \varepsilon_k$, then $\bar{F}_{ki} = \bar{f}_i^k$. We have from (2.10)

$$H_k u + A_k v = 0, \tag{2.11}$$

$$A_k^T u - \bar{F}_k v = 0, \quad \text{i.e.,} \quad (\bar{g}_i^k)^T u - \bar{f}_i^k v_i = 0, \quad i \in L_k. \tag{2.12}$$

Multiplying (2.11) by $u^T$ from left-hand side and combing with (2.12), we obtain

$$0 = u^T H_k u + u^T A_k v = u^T H_k u + v^T \bar{F}_k v.$$

For $i \in L(x^k) \backslash \{l_k\} \subseteq L_k$, it follows that $\bar{f}_i^k = 0$ from (2.8). Further, from (2.12), we get $(\bar{g}_i^k)^T u = 0$, which implies that $u \in NS_k$. Then $0 \leq u^T H_k u = -v^T \bar{F}_k v \leq 0$. So, taking into account the positive definiteness of $H_k$, it follows that $u = 0$. Again, (2.11) and (2.12) can be simplified as

$$A_k v = 0, \quad \bar{f}_i^k v_i = 0, \quad i \in L_k. \tag{2.13}$$

For $i \in L_k$ and $\bar{f}_i^k \neq 0$, by (2.13), one has $v_i = 0$. Then, based on the first equation of (2.13), one gets

$$\sum_{i \in L_k,\ \bar{f}_i^k = 0} v_i \bar{g}_i^k = 0.$$

This together with Assumption 2.1 shows that $v_i = 0$ for $i \in L_k$ and $\bar{f}_i^k = 0$. Therefore, $u = 0$ and $v = 0$, and $V_k$ is invertible.

(ii) If $\det(A_k^T A_k) \geq \varepsilon_k$, then matrix $A_k$ is full-column rank and $\bar{F}_k = 0$. According to the definition (2.10), we have

$$H_k u + A_k v = 0, \ A_k^T u = 0. \tag{2.14}$$

Multiplying the first equation of (2.14) by $u^T$ from left-hand side and combing with the second equation of (2.14), one knows $u^T H_k u = 0$. In addition, the second equation of (2.14) implies that $u \in NS_k$. Consequently, one knows $u = 0$ from the positive definiteness of $H_k$. Further, one has $A_k v = 0$. This, together with full-column rank of $A_k$, shows that $v = 0$.

According to the analysis above, the system of equations $V_k(u^T, v^T)^T = 0$ has the unique solution zero, so $V_k$ is invertible. The proof is completed. $\square$

Based on the analysis and preparation above, we introduce the main idea of our method. First, taking into account the stationary conditions (2.5), as well as (2.9) and Lemma 2.1, if we ignore the non-negativity request of the multiplier vector, then the form of the SLE that need to be solved in our algorithm are as follows:

$$V_k \begin{pmatrix} d \\ \lambda_{L_k} \end{pmatrix} = \begin{pmatrix} -g^k_{l_k} \\ 0 \end{pmatrix}. \tag{2.15}$$

Let $(d^{k0}, \lambda^{k0}_{L_k})$ be a solution to the SLE (2.15). In order to check whether or not the current iteration $x^k$ is a stationary point of the problem (1.1), we define

$$\omega_k := \sum_{i \in L_k} \max\{-\lambda^{k0}_i,\ \lambda^{k0}_i \bar{f}^k_i\}, \qquad \lambda^{k0}_{l_k} := 1 - \sum_{i \in I^0_k} \lambda^{k0}_i, \qquad \bar{\omega}_k := \max\{-\lambda^{k0}_{l_k},\ 0\}, \tag{2.16}$$

$$\rho_k := \frac{|(g^k_{l_k})^T d^{k0}| + \omega_k + \bar{\omega}^2_k + \varphi_k}{1 + |(e^k)^T \lambda^{k0}_{L_k}|}, \tag{2.17}$$

where $e^k = (1, 1, \cdots, 1)^T \in \mathbb{R}^{|L_k|}$.

Second, if $x^k$ is not a stationary point of the problem (1.1), this along with that the solution $d^{k0}$ of the SLE (2.15) may not be an improved direction since $(\bar{g}^k_i)^T d^{k0} = 0$ for $i \in (I(x^k) \backslash \{l_k\}) \cup J(x^k)$. Hence, in order to yield improved search direction with global and strong convergence, our algorithm will solve an additional SLE by perturbing the right-hand side vector of (2.15) with the same coefficient matrix as follows:

$$V_k \begin{pmatrix} d \\ \lambda_{L_k} \end{pmatrix} = \begin{pmatrix} 0 \\ \mu^k \end{pmatrix}, \tag{2.18}$$

where the vector $\mu^k = (\mu^k_i, i \in L_k)$ with elements

$$\mu^k_i(i \in I^0_k) := \begin{cases} \rho^\xi_k(-1 - \rho_k) - r\varphi_k + \rho^\xi_k \bar{\omega}_k, & \text{if } \lambda^{k0}_i < 0; \\ \rho^\xi_k(\bar{f}^k_i - \rho_k) - r\varphi_k + \rho^\xi_k \bar{\omega}_k, & \text{if } \lambda^{k0}_i \geq 0, \end{cases} \tag{2.19a}$$

$$\mu^k_i(i \in J_k) := \begin{cases} \rho^\xi_k(-1 - \rho_k) - r\varphi_k, & \text{if } \lambda^{k0}_i < 0; \\ \rho^\xi_k(\bar{f}^k_i - \rho_k) - r\varphi_k, & \text{if } \lambda^{k0}_i \geq 0, \end{cases} \tag{2.19b}$$

and the parameters $\xi,\ r > 0$. Let $(d^{k1}, \lambda^{k1}_{L_k})$ be a solution to SLE (2.18). In order to obtain nice theoretical properties of the main search direction and improve the numerical performance of our proposed algorithm, the main search direction $d^k$ is yielded by a convex combination of $\rho^\xi_k d^{k0}$ and $d^{k1}$:

$$d^k = (1 - \sigma)\rho^\xi_k d^{k0} + \sigma d^{k1}, \tag{2.20}$$

where the parameter $\sigma \in (0, \frac{1}{2})$. Based on the constructions above, it is not difficult to verify the following lemma.

**Lemma 2.2.** *Suppose that the conditions stated in Lemma* 2.1 *hold. Then*

(i) $(g_{l_k}^k)^T d^{k1} = -\sum_{i\in L_k} \lambda_i^{k0}\mu_i^k$;

(ii) $(g_{l_k}^k)^T d^{k0} \le -(d^{k0})^T H_k d^{k0} \le 0,\ d^{k0} \in NS_k$;

(iii) $x^k$ *is a stationary point of problem* (1.1) *if and only if* $\rho_k = 0$; *and*

(iv) *the following estimates of inequalities hold:*

$$(g_{l_k}^k)^T d^k \le -\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0}), \tag{2.21}$$

$$(g_i^k)^T d^k \le -\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0}) + r_{ki}\bar{F}_{ki} + \sigma\rho_k^\xi \bar{f}_i^k,\ \forall\ i \in I_k^0, \tag{2.22}$$

$$(g_i^k)^T d^k \le -\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0}),\ \forall\ i \in I(x^k), \tag{2.23}$$

$$(g_i^k)^T d^k \le -\sigma(\rho_k^{1+\xi} + r\varphi_k) + r_{ki}\bar{F}_{ki} + \sigma\rho_k^\xi \bar{f}_i^k,\ \forall\ i \in J_k, \tag{2.24}$$

$$(g_i^k)^T d^k \le -\sigma(\rho_k^{1+\xi} + r\varphi_k),\ \forall\ i \in J(x^k), \tag{2.25}$$

*where* $r_{ki} := (1-\sigma)\rho_k^\xi\lambda_i^{k0} + \sigma\lambda_i^{k1}$.

*Proof.* (i) Since $(d^{k0}, \lambda_{L_k}^{k0})$ and $(d^{k1}, \lambda_{L_k}^{k1})$ are solutions to SLE (2.15) and (2.18), respectively, we have

$$H_k d^{k0} + A_k\lambda_{L_k}^{k0} = -g_{l_k}^k,\quad A_k^T d^{k0} - \bar{F}_k\lambda_{L_k}^{k0} = 0. \tag{2.26}$$

$$H_k d^{k1} + A_k\lambda_{L_k}^{k1} = 0,\quad A_k^T d^{k1} - \bar{F}_k\lambda_{L_k}^{k1} = \mu^k. \tag{2.27}$$

Multiplying the first equations of (2.26) and (2.27) by $(d^{k1})^T$ and $(d^{k0})^T$ from left-hand side, respectively, one has

$$(d^{k1})^T H_k d^{k0} + (d^{k1})^T A_k\lambda_{L_k}^{k0} = -(d^{k1})^T g_{l_k}^k,\quad (d^{k0})^T H_k d^{k1} + (d^{k0})^T A_k\lambda_{L_k}^{k1} = 0.$$

Therefore, taking into account the symmetric matrix $H_k$ as well as the second equations of (2.26) and (2.27), we immediately get

$$\begin{aligned}(g_{l_k}^k)^T d^{k1} &= (d^{k0})^T A_k\lambda_{L_k}^{k1} - (d^{k1})^T A_k\lambda_{L_k}^{k0}\\ &= (A_k^T d^{k0})^T\lambda_{L_k}^{k1} - (A_k^T d^{k1})^T\lambda_{L_k}^{k0}\\ &= (\bar{F}_k\lambda_{L_k}^{k0})^T\lambda_{L_k}^{k1} - (\bar{F}_k\lambda_{L_k}^{k1} + \mu^k)^T\lambda_{L_k}^{k0}\\ &= (\lambda_{L_k}^{k0})^T\bar{F}_k^T\lambda_{L_k}^{k1} - (\lambda_{L_k}^{k1})^T\bar{F}_k^T\lambda_{L_k}^{k0} - (\mu^k)^T\lambda_{L_k}^{k0}\\ &= -(\mu^k)^T\lambda_{L_k}^{k0}\\ &= -\sum_{i\in L_k}\lambda_i^{k0}\mu_i^k.\end{aligned}$$

(ii) In view of (2.7) and (2.8), one has $\bar{F}_{ki} \ge 0$. This, together with (2.26), shows that

$$\begin{aligned}(g_{l_k}^k)^T d^{k0} &= -(d^{k0})^T H_k d^{k0} - \sum_{i\in L_k}\lambda_i^{k0}(\bar{g}_i^k)^T d^{k0}\\ &= -(d^{k0})^T H_k d^{k0} - \sum_{i\in L_k}(\lambda_i^{k0})^2\bar{F}_{ki}\\ &\le -(d^{k0})^T H_k d^{k0}.\end{aligned}$$

In addition, for $i \in (L(x^k)\backslash\{l_k\}) \subseteq L_k$, from (2.7)–(2.8), one has $\bar{F}_{ki} = 0$. Further, $(\bar{g}_i^k)^T d^{k0} = 0$ follows from (2.26), this implies $d^{k0} \in NS_k$. Hence, $(g_{l_k}^k)^T d^{k0} \le -(d^{k0})^T H_k d^{k0} \le 0$ holds due to the positive definiteness of $H_k$ on $NS_k$.

(iii) If $\rho_k = 0$, then from (2.17), one has

$$(g_{l_k}^k)^T d^{k0} = 0, \quad \omega_k = 0, \quad \bar{\omega}_k = 0, \quad \varphi_k = 0.$$

Further, in view of the positive definite property of $H_k$ and conclusion (ii), we have $d^{k0} = 0$. Again, from $\omega_k = 0$ and $\bar{\omega}_k = 0$ as well as (2.16), one gets $\lambda_{I_k \cup J_k}^{k0} \geq 0,\ \lambda_i^{k0} \bar{f}_i^k = 0,\ \forall\, i \in L_k$. Now, from the first equation of (2.26), we have

$$\begin{cases} \sum\limits_{i \in I_k^0} \lambda_i^{k0}(g_i^k - g_{l_k}^k) + \sum\limits_{i \in J_k} \lambda_i^{k0} g_i^k = -g_{l_k}^k, \quad \lambda_{l_k}^{k0} = 1 - \sum\limits_{i \in I_k^0} \lambda_i^{k0} \geq 0, \\ \lambda_i^{k0} \geq 0, \quad \lambda_i^{k0}(f_i^k - F_k) = 0, \quad i \in I_k^0; \quad \lambda_i^{k0} \geq 0, \quad \lambda_i^{k0} f_i^k = 0, \quad i \in J_k. \end{cases}$$

These, together with $\varphi_k = 0$ and (2.5), imply that $x^k$ is a stationary point of the problem (1.1).

Conversely, if $x^k$ is a stationary point of the problem (1.1) with multiplier vector $\hat{\lambda}^k \geq 0$, then it follows from (2.5) that

$$\varphi_k = 0,;\ A_k \hat{\lambda}_{L_k}^k = -g_{l_k}^k, \quad \hat{\lambda}_i^k \bar{f}_i^k = 0, \quad \forall\, i \in L_k, \quad \hat{\lambda}_{l_k}^k = 1 - \sum_{i \in I_k^0} \lambda_i^k.$$

Therefore,

$$\begin{pmatrix} H_k & A_k \\ A_k^T & -\bar{F}_k \end{pmatrix} \begin{pmatrix} 0 \\ \hat{\lambda}_{L_k}^k \end{pmatrix} = \begin{pmatrix} -g_{l_k}^k \\ 0 \end{pmatrix}.$$

This implies that $(0^T,\ (\hat{\lambda}_{L_k}^k)^T)^T$ is also a solution to SLE (2.15). Noting that the uniqueness of solution to the SLE (2.15), one has $d^{k0} = 0, \lambda_{L_k}^{k0} = \hat{\lambda}_{L_k}^k$, $\lambda_{l_k}^{k0} = \hat{\lambda}_{l_k}^k \geq 0$ and $\bar{\omega}_k = 0$. On the other hand, the conclusion $\omega_k = 0$ holds by (2.16). Therefore, it follows from (2.17) that $\rho_k = 0$.

(iv) For $i = l_k \in I(x^k)$, from (2.20) and conclusions (i) and (ii), we get

$$\begin{aligned} (g_{l_k}^k)^T d^k =& (1-\sigma)\rho_k{}^{\xi}(g_{l_k}^k)^T d^{k0} + \sigma (g_{l_k}^k)^T d^{k1} \\ =& (1-\sigma)\rho_k{}^{\xi}(g_{l_k}^k)^T d^{k0} - \sigma \sum_{i \in L_k} \lambda_i^{k0} \mu_i^k. \end{aligned} \tag{2.28}$$

In addition, taking into account

$$-\bar{\omega}_k \lambda_{l_k}^{k0} = \bar{\omega}_k^2 \quad \text{and} \quad \bar{\omega}_k \sum_{i \in I_k^0} \lambda_i^{k0} = \bar{\omega}_k + \bar{\omega}_k^2,$$

together with (2.16), (2.19a) and (2.19b), gives

$$\sum_{i \in L_k} \lambda_i^{k0} \mu_i^k = -\rho_k^{\xi} \left( \rho_k \sum_{i \in L_k} \lambda_i^{k0} - \omega_k - \bar{\omega}_k^2 - \bar{\omega}_k \right) - r\varphi_k \sum_{i \in L_k} \lambda_i^{k0}. \tag{2.29}$$

Furthermore, from (2.17) and conclusion (ii), we have

$$\rho_k \sum_{i \in L_k} \lambda_i^{k0} - \omega_k - \bar{\omega}_k^2 \leq |(g_{l_k}^k)^T d^{k0}| + \varphi_k - \rho_k = -(g_{l_k}^k)^T d^{k0} + \varphi_k - \rho_k. \tag{2.30}$$

Now, from (2.28)–(2.30) and $(g_{l_k}^k)^T d^{k0} \le 0$, one obtains

$$
\begin{aligned}
(g_{l_k}^k)^T d^k &= (1-\sigma){\rho_k}^{\xi}(g_{l_k}^k)^T d^{k0} + \sigma\{\rho_k^{\xi}(\rho_k \sum_{i\in L_k}\lambda_i^{k0} - \omega_k - \bar{\omega}_k^2 - \bar{\omega}_k) + r\varphi_k \sum_{i\in L_k}\lambda_i^{k0}\} \\
&\le (1-\sigma){\rho_k}^{\xi}(g_{l_k}^k)^T d^{k0} + \sigma\{\rho_k^{\xi}(-(g_{l_k}^k)^T d^{k0} - \rho_k + \varphi_k - \bar{\omega}_k) + r\varphi_k \sum_{i\in L_k}\lambda_i^{k0}\} \\
&= (1-2\sigma){\rho_k}^{\xi}(g_{l_k}^k)^T d^{k0} - \sigma{\rho_k}^{1+\xi} + \sigma\varphi_k(\rho_k^{\xi} + r\sum_{i\in L_k}\lambda_i^{k0}) - \sigma\rho_k^{\xi}\bar{\omega}_k \\
&\le -\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^{\xi} + r\sum_{i\in L_k}\lambda_i^{k0}) - \sigma\rho_k^{\xi}\bar{\omega}_k \\
&\le -\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^{\xi} + r\sum_{i\in L_k}\lambda_i^{k0}). \qquad (2.31)
\end{aligned}
$$

So, inequality (2.21) holds.

For $i \in I_k^0 = I_k \backslash \{l_k\}$ , from (2.20), the second equations of (2.26) and (2.27), one knows

$$
\begin{aligned}
(g_i^k - g_{l_k}^k)^T d^k &= (1-\sigma)\rho_k^{\xi}(g_i^k - g_{l_k}^k)^T d^{k0} + \sigma(g_i^k - g_{l_k}^k)^T d^{k1} \\
&= (1-\sigma)\rho_k^{\xi}\bar{F}_{ki}\lambda_i^{k0} + \sigma(\bar{F}_{ki}\lambda_i^{k1} + \mu_i^k) \\
&= ((1-\sigma)\rho_k^{\xi}\lambda_i^{k0} + \sigma\lambda_i^{k1})\bar{F}_{ki} + \sigma\mu_i^k, \\
&= r_{ki}\bar{F}_{ki} + \sigma\mu_i^k. \qquad (2.32)
\end{aligned}
$$

Therefore, by using (2.19a)–(2.19b) and the penultimate inequality of (2.31), we have

$$
\begin{aligned}
(g_i^k)^T d^k &= \sigma\mu_i^k + (g_{l_k}^k)^T d^k + r_{ki}\bar{F}_{ki} \\
&\le \sigma(\rho_k^{\xi}\bar{f}_i^k - \rho_k^{1+\xi} - r\varphi_k + \rho_k^{\xi}\bar{\omega}_k) + (g_{l_k}^k)^T d^k + r_{ki}\bar{F}_{ki} \\
&\le -\sigma\rho_k^{1+\xi} + \sigma\rho_k^{\xi}\bar{\omega}_k + (g_{l_k}^k)^T d^k + r_{ki}\bar{F}_{ki} + \sigma\rho_k^{\xi}\bar{f}_i^k \\
&\le -\sigma\rho_k^{1+\xi} + \sigma\rho_k^{\xi}\bar{\omega}_k - \sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^{\xi} + r\sum_{i\in L_k}\lambda_i^{k0}) - \sigma\rho_k^{\xi}\bar{\omega}_k + r_{ki}\bar{F}_{ki} + \sigma\rho_k^{\xi}\bar{f}_i^k \\
&\le -\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^{\xi} + r\sum_{i\in L_k}\lambda_i^{k0}) + r_{ki}\bar{F}_{ki} + \sigma\rho_k^{\xi}\bar{f}_i^k. \qquad (2.33)
\end{aligned}
$$

So the assertion (2.22) holds. In particular, for $i \in (I(x^k)\backslash\{l_k\}) \subseteq I_k$, from (2.7) and (2.8), one gets $\bar{F}_{ki} = \bar{f}_i^k = 0$, thus the conclusion (2.23) follows from (2.21) and (2.22).

For $i \in J_k$, relation $(g_i^k)^T d^k = r_{ki}\bar{F}_{ki} + \sigma\mu_i^k$ follows by a similar analysis to (2.32). So, based on (2.19b), we further have

$$
(g_i^k)^T d^k \le -\sigma(\rho_k^{1+\xi} + r\varphi_k) + r_{ki}\bar{F}_{ki} + \sigma\rho_k^{\xi}\bar{f}_i^k, \ \forall\, i \in J_k.
$$

and the assertion (2.24) holds. In particular, for $i \in J(x^k) \subseteq J_k$, from (2.7) and (2.8), it is easy to get $\bar{F}_{ki} = \bar{f}_i^k = 0$. This, along with (2.24), shows that conclusion (2.25) holds. And the proof is completed. □

Now, we give the steps of our algorithm for solving the problem (1.1) as follows.

**Algorithm 2.1. Parameters:** $\alpha, \beta \in (0,1),\ \sigma \in (0, \frac{1}{2}),\ \varepsilon_0, \xi, r > 0$.

*Data:* $x^0 \in \mathbb{R}^n$, an initial symmetric positive matrix $H_0 \in \mathbb{R}^{n\times n}$.

**Step 0 (Initialization).** Set $k := 0$.

**Step 1 (Generating the working set).** Generate the working set $L_k$ by (2.2)-(2.4).

**Step 2 (Computation of the search direction).**

**Substep 2.1.** Obtain $(d^{k0}, \lambda^{k0}_{L_k})$ by solving SLE (2.15). Yield $\lambda^{k0}_{l_k} := 1 - \sum_{i\in I^0_k} \lambda^{k0}_i \ge 0$, and let $\lambda^{k0} = (\lambda^{k0}_{L_k},\ \lambda^{k0}_{l_k},\ 0_{(I\cup J)\setminus(L_k\cup\{l_k\})})$, and go to Substep (2.2).

**Substep 2.2.** Yield $\omega_k,\ \bar{\omega}_k$ and $\rho_k$ by (2.16)–(2.17).If $\rho_k = 0$, then $x^k$ is a stationary point of the problem (1.1), stop; otherwise, generate $\mu^k = (\mu^k_i, i \in L_k)$ by (2.19a) and (2.19b).

**Substep 2.3.** Obtain $(d^{k1}, \lambda^{k1}_{L_k})$ by solving the SLE (2.18), compute $\lambda^{k1}_{l_k} := 1 - \sum_{i\in I^0_k} \lambda^{k1}_i$. Let $\lambda^{k1} = (\lambda^{k1}_{L_k},\ \lambda^{k1}_{l_k},\ 0_{(I\cup J)\setminus(L_k\cup\{l_k\})})$. Generate the main search direction $d^k$ by (2.20), and go to Step 3.

**Step 3 (Doing line search).** If $\varphi_k > 0$ and $f_i(x^k + d^k) \le 0$ hold for all $i \in J$ , then set $t_k = 1$, and go to Step 4. Otherwise, compute the step size $t_k$ which is the maximum $t$ in the sequence $\{1, \beta, \beta^2, \beta^3, \cdots\}$ that satisfying the following inequalities:

$$F(x^k + td^k) \le F_k + \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^{\xi} + r \sum_{i\in L_k} \lambda^{k0}_i)), \tag{2.34}$$

$$f_i(x^k + td^k) \le \max\{0,\ \varphi_k - \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k)\}, \qquad i \in J, \tag{2.35}$$

$$|J^-(x^k)| \le |J^-(x^k + td^k)|. \tag{2.36}$$

**Step 4 (Updating).** Let $x^{k+1} := x^k + t_k d^k$, update the parameter $\varepsilon_k$ by

$$\varepsilon_{k+1} := \begin{cases} \varepsilon_k, & \text{if } \det(A_k^T A_k) \ge \varepsilon_k; \\ \frac{1}{2}\varepsilon_k, & \text{if } \det(A_k^T A_k) < \varepsilon_k. \end{cases} \tag{2.37}$$

Generate a new symmetric positive definite matrix $H_{k+1}$, such that matrix $H_{k+1}$ is positive definite in the null space $NS_{k+1}$, set $k := k + 1$, and go back to Step 1.

**Remark 2.2.** From inequalities (2.34)–(2.36), it follows that one of the following two cases must take place:

*Case A.* There exists an integer $s$ such that $\varphi(x^s) = 0$, i.e., the iteration point $x^s$ gets into the feasible set $X$. Then one has

$$f_i^k \le 0, \quad \forall\ i \in J,\ \varphi_k = 0; \quad F_{k+1} \le F_k - \sigma\alpha t_k \rho_k^{1+\xi}, \quad \forall\ k \ge s; \tag{2.38}$$

*Case B.* For any $k = 0, 1, 2, \cdots$, $\varphi(x^k) > 0$. This case implies that

$$\varphi_k > 0, \quad \varphi_{k+1} \le \varphi_k - \alpha\sigma t_k(\rho_k^{1+\xi} + r\varphi_k) < \varphi_k, \quad k = 0, 1, \cdots. \tag{2.39}$$

The following lemma shows that Algorithm 2.1 is well-defined.

**Lemma 2.3.** *Suppose that the conditions stated in Lemma* 2.1 *hold. Then inequalities* (2.34)–(2.36) *hold for sufficiently small* $t > 0$, *so the line search in Step* 3 *can be carried out in a finite number of computations, i.e., Algorithm* 2.1 *is well-defined.*

*Proof.* In the whole discussion of this lemma, $t$ always means a positive and sufficiently small real number.

**A1.** Analyze the inequality (2.34): the proof can be divided into two cases, i.e., $i \in I(x^k)$ and $i \notin I(x^k)$.

A1.1. For $i \in I(x^k)$ , i.e., $f_i^k = F_k,$ using Taylor expansion and (2.23), we have

$$\begin{aligned} & f_i(x^k + td^k) - F_k - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0})) \\ = & f_i^k + t(g_i^k)^T d^k - F_k - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0})) + o(t) \\ \leq & t(-\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0})) - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0})) + o(t) \\ \leq & -(1-\alpha)\sigma t\rho_k^{1+\xi} + o(t) \leq 0. \end{aligned}$$

A1.2. For $i \in I\backslash I(x^k)$, that is $f_i^k < F_k,$ using Taylor expansion, we obtain

$$\begin{aligned} & f_i(x^k + td^k) - F_k - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0})) \\ = & f_i^k + t(g_i^k)^T d^k - F_k - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^\xi + r\sum_{i\in L_k}\lambda_i^{k0})) + o(t) \\ \leq & f_i^k - F_k + O(t) \leq 0. \end{aligned}$$

Summarizing the analysis above, we know that inequality (2.34) holds.

**A2.** Analyze inequalities (2.35) and (2.36). For convenience, denote

$$a_i^k(t) = f_i(x^k + td^k) - \max\{0, \varphi_k - \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k)\}. \tag{2.40}$$

Then it is sufficient to show $a_i^k(t) \leq 0$ for $t > 0$ sufficiently small and for all $i \in J$. Next, we consider the following four cases:

A2.1. For $i \in J^+(x^k) \cap J(x^k)$, i.e., $f_i^k = \varphi_k,$ by Taylor expansion, it follows from (2.25) that

$$\begin{aligned} a_i^k(t) & \leq f_i(x^k + td^k) - \varphi_k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) \\ & = f_i^k - \varphi_k + t(g_i^k)^T d^k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) + o(t) \\ & \leq -\sigma t(\rho_k^{1+\xi} + r\varphi_k) + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) + o(t) \\ & = -(1-\alpha)\sigma t(\rho_k^{1+\xi} + r\varphi_k) + o(t) \leq 0. \end{aligned}$$

A2.2. For $i \in J^+(x^k)\backslash J(x^k)$, that is $f_i^k < \varphi_k,$ from Taylor expansion, we have

$$\begin{aligned} a_i^k(t) & \leq f_i(x^k + td^k) - \varphi_k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) \\ & = f_i^k - \varphi_k + t(g_i^k)^T d^k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) + o(t) \\ & = f_i^k - \varphi_k + O(t) \leq 0. \end{aligned}$$

A2.3. For $i \in J^-(x^k) \cap J(x^k)$, that is $f_i^k = 0,$ expanding $f_i(x^k + td^k)$ at the current iteration point $x^k$ and combining relation (2.25), we get

$$\begin{aligned} a_i^k(t) \leq & f_i(x^k + td^k) = f_i^k + t(g_i^k)^T d^k + o(t) \\ \leq & -\sigma t\rho_k^{1+\xi} - \sigma tr\varphi_k + o(t) \leq -\sigma t\rho_k^{1+\xi} + o(t) \leq 0. \end{aligned}$$

A2.4. For $i \in J^-(x^k)\backslash J(x^k)$, i.e., $f_i^k < 0,$ we obtain

$$a_i^k(t) \leq f_i(x^k + td^k) = f_i^k + t(g_i^k)^T d^k + o(t) = f_i^k + O(t) \leq 0.$$

Therefore, for $i \in J^-(x^k)$, one knows that $f_i(x^k + td^k) \le 0$ holds from cases A2.3 and A2.4. So $|J^-(x^k)| \le |J^-(x^k + td^k)|$, i.e., (2.36) holds. Moreover, (2.35) holds for $i \in J^-(x^k)$. From cases A2.1 and A2.2, it shows that (2.35) holds for $i \in J^+(x^k)$. Therefore, (2.35) holds for all $i \in J$. And the lemma is proved. □

## 3. Convergence Analysis

If Algorithm 2.1 stops at the current iteration point $x^k$, from Substep 2.2 and Lemma 2.2 (iii), we know that $x^k$ is a stationary point of the problem (1.1). In this section, we assume that the algorithm yields an infinite iteration sequence $\{x^k\}$ of points, then discuss the global and strong convergence of Algorithm 2.1. For this purpose, the following assumption is necessary.

**Assumption 3.1.** The sequences both $\{x^k\}$ and $\{H_k\}$ are bounded, and there exists a positive constant $a$ such that $d^T H_k d \ge a\|d\|^2$ ($\forall\ d \in NS_k$) holds for $k$ large enough.

The assumption above, together with Lemma 2.2(ii), implies that (for $k$ large enough)

$$(g^k_{l_k})^T d^{k0} \le -(d^{k0})^T H_k d^{k0} \le -a\|d^{k0}\|^2. \tag{3.1}$$

**Lemma 3.1.** *Suppose that Assumptions* 2.1 *and* 3.1 *hold. Then*

(i) $\underline{\varepsilon} := \lim\limits_{k\to\infty} \varepsilon_k = \inf\{\varepsilon_k\} > 0$, *so* $\varepsilon_k \ge \underline{\varepsilon}$, $\forall\ k = 0, 1, 2, \cdots$ , *and furthermore, there exists an index* $\ell$ *such that* $\varepsilon_k = \varepsilon_\ell$, $\forall\ k \ge \ell$;

(ii) $\det(A_k^T A_k) \ge \varepsilon_k$ *and* $\bar{F}_k \equiv 0$ *for all* $k$ *large enough.*

*Proof.* (i) First, in view of that the whole sequence $\{\varepsilon_k\}$ being nonincreasing and bounded as well as positive, we know $\lim\limits_{k\to\infty} \varepsilon_k = \inf\{\varepsilon_k\} \ge 0$. Second, we show that $\underline{\varepsilon} > 0$. Suppose by contradiction that $\underline{\varepsilon} = 0$. Then there exists an infinite subset $\mathcal{K}$ such that $\varepsilon_{k+1} < \varepsilon_k$ for $k \in \mathcal{K}$. So, it follows, from the finite choice of $l_k$, $I_k$ and $J_k$ as well as the updating formula (2.37), that there is an infinite subset $\bar{\mathcal{K}} \subseteq \mathcal{K}$ satisfying

$$l_k \equiv l_*, \quad I_k \equiv \bar{I}, \ \ J_k \equiv \bar{J}, \ \ \bar{L} := \bar{I}\backslash\{l_*\} \cup \bar{J}, \ \det(A_k^T A_k) < \varepsilon_k, \ k \in \bar{\mathcal{K}}.$$

In view of the boundedness of the sequence $\{x^k\}$, without loss of generality, suppose that $x^k \xrightarrow{\bar{\mathcal{K}}} x^*$. Taking into account $\varepsilon_k \to 0$ and (2.3)-(2.3), it follows that $l_* \in J(x^*)$, $\bar{I} \subseteq I(x^*)$, $\bar{J} \subseteq J(x^*)$. Let matrix

$$A_* := (\nabla f_i(x^*) - \nabla f_{l_*}(x^*),\ i \in \bar{I}\backslash\{l_*\};\ \nabla f_i(x^*), i \in \bar{J}).$$

Then, passing to the limit in $\det(A_k^T A_k) < \varepsilon_k$, for $k \in \bar{\mathcal{K}}$, one has $\det(A_*^T A_*) = 0$, which contradicts with Assumption 2.1. So, $\underline{\varepsilon} > 0$ is at hand, and this together with the update (2.37) implies the second claim in (i) is also at hand.

(ii) From conclusion (i) and (2.37) as well as (2.7), it is easy to get that $\det(A_k^T A_k) \ge \varepsilon_k$ and $\bar{F}_k \equiv 0$ when $k$ is sufficiently large. So the proof is completed. □

**Lemma 3.2.** *Suppose that Assumptions* 2.1 *and* 3.1 *hold. Then*

(i) *the accumulation point* $V_*$ *of* $\{V_k\}$ *of matrices is nonsingular, and there exists a positive constant* $c$*, such that* $\|V_k^{-1}\| \le c$, $\forall\ k = 0, 1, 2, \cdots$*;*

(ii) *the sequences* $\{(d^{k0}, \lambda^{k0})\}$, $\{(d^{k1}, \lambda^{k1})\}$ *and* $\{(d^k, \rho_k,\ \mu^k)\}$ *are all bounded.*

*Proof.* (i) Assume that $V_*$ is a given accumulation point of the sequence $\{V_k\}$. Then exists an infinite index set $\hat{\mathcal{K}}$, such that $\lim\limits_{k\in\hat{\mathcal{K}}} V_k = V_*$. Further, in view of the finite choice for sets $J^+(x^k),\ J^-(x^k),\ I_k,\ J_k,\ L_k,\ l_k$ and the boundedness of $\{(x^k, H_k)\}_{\hat{\mathcal{K}}}$ , there exists an infinite index set $\mathcal{K} \subseteq \hat{\mathcal{K}}$, such that

$$\begin{array}{l} x^k \xrightarrow{\mathcal{K}} x^*, \quad H_k \xrightarrow{\mathcal{K}} H_*, \quad I_k \equiv I',\ J_k \equiv J', \quad J^+(x^k) \equiv J^+, \quad J^-(x^k) \equiv J^-, \\ l_k \equiv l', \quad I_k^0 \equiv I^{0'} := I' \backslash \{l'\}, \quad L_k \equiv L' := I^{0'} \cup J', \quad \forall\ k \in \mathcal{K}. \end{array} \tag{3.2}$$

Then, it follows from (2.10), (3.2) and $\bar{F}_k = 0$ (Lemma 3.1) that

$$V_k \to V_* := \begin{pmatrix} H_* & A_* \\ A_*^T & 0 \end{pmatrix},\ k \in \mathcal{K},$$

where

$$A_* = (\bar{g}_i^*,\ i \in L'), \quad \bar{g}_i^* = \begin{cases} g_i(x^*) - g_{l'}(x^*), & i \in I^{0'}; \\ g_i(x^*), & i \in J', \end{cases}$$

and $\det(A_*^T A_*) \geq \underline{\varepsilon} > 0$.

Next, we prove that $V_*$ is nonsingular. It is sufficient to show that the system of equations $V_*(y^T, z^T)^T = 0$ has a unique solution zero, and this is equivalent to system

$$H_* y + A_* z = 0,\ A_*^T y = 0 \tag{3.3}$$

has the unique solution zero.

Since $A_*$ is full-column rank, $A_*$ can be divided as $A_* = (C^T,\ D^T)^T$, where $C$ is nonsingular. Dividing the vector $y$ and the matrix $A_k$ in the same way, that is:

$$y = \begin{pmatrix} y_C \\ y_D \end{pmatrix},\ A_k = \begin{pmatrix} C_k \\ D_k \end{pmatrix},$$

Since $A_k \to A_*$, then $C_k \to C,\ D_k \to D,\ k \in \mathcal{K}$, and $C_k$ is nonsingular for $k \in \mathcal{K}$ large enough. Yield $y^k$ by

$$y_D^k = y_D, \quad y_c^k = -(C_k^{-1})^T D_k^T y_D^k \to -(C^{-1})^T D^T y_D, \quad k \in \mathcal{K}.$$

then $A_k^T y^k = 0,\ y^k \in NS_k$. Furthermore, since $A_*^T y = 0$, one knows that $y^k \to y,\ k \in \mathcal{K}$. Therefore, it follows that $(y^k)^T H_k y^k \geq a\|y^k\|^2$ from Assumption 3.1, and this further implies that $y^T H_* y \geq a\|y\|^2$.

On the other hand, multiplying the first equation of (3.3) by $y^T$ from left-hand side and combing with the second equation of (3.3), we have $0 = y^T H_* y \geq a\|y\|^2$. So, $y = 0$. Further, from the first equation of (3.3), we get $A_* z = 0$. And $A_*$ is full-column rank, one knows $z = 0$. Hence, the system of linear equation $V_*(y^T, z^T)^T = 0$ has a unique solution zero, and $V_*$ is nonsingular. The second claim in part (i) follows from the first one in part (i).

(ii) From conclusion (i), formulas (2.15)-(2.20) and Assumption 3.1, the result (ii) is at hand. The proof is completed. □

**Lemma 3.3.** *Suppose that Assumptions* 2.1 *and* 3.1 *hold. Then*

(i) $\lim\limits_{k\to\infty} t_k \rho_k^{1+\xi} = \lim\limits_{k\to\infty} t_k \rho_k^{\xi} = \lim\limits_{k\to\infty} t_k \varphi_k = 0$*; and*

(ii) $\lim\limits_{k\to\infty} \|x^{k+1} - x^k\| = 0.$

*Proof.* (i) If Case A in Remark 2.2 takes place, then $\varphi_k \equiv 0,\ \forall\ k \geq s$, and $\{F_k\}_{k\geq s}$ is decreasing and bounded. So the whole sequence $\{F_k\}_{k\geq s}$ is convergent. Consequently, it follows from (2.38) that $\lim\limits_{k\to\infty} t_k\rho_k^{1+\xi} = 0$. Furthermore,

$$\lim_{k\to\infty} t_k\rho_k^{\xi} = \lim_{k\to\infty} [t_k(t_k\rho_k^{1+\xi})^{\xi}]^{\frac{1}{1+\xi}} = 0.$$

If Case B in Remark 2.2 takes place, then $\{\varphi_k\}_{k\geq 0}$ is decreasing and bounded, and further is convergent. Therefore, from (2.39) we have

$$\lim_{k\to\infty} t_k\rho_k^{1+\xi} = \lim_{k\to\infty} t_k\varphi_k = 0.$$

(ii) From (2.19a), (2.19b) and conclusion (i), it follows that $\lim\limits_{k\to\infty} t_k\mu^k = 0$. In addition, since (2.18) and Lemma 3.2, one obtains $\lim\limits_{k\to\infty} t_k d^{k1} = 0$. So,

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = \lim_{k\to\infty} \|(1-\sigma)t_k\rho_k^{\xi} d^{k0} + \sigma t_k d^{k1}\| = 0.$$

This completes the proof of the lemma. □

**Theorem 3.1.** *Suppose that Assumptions* 2.1 *and* 3.1 *hold, and let* $x^*$ *be any accumulation point of the sequence* $\{x^k\}$ *generated by Algorithm* 2.1. *Then*

(i) *the accumulation point* $x^*$ *is a stationary point of the problem* (1.1)*, i.e., Algorithm* 2.1 *is globally convergent; and*

(ii) $\lim\limits_{k\to\infty} \varphi_k = \lim\limits_{k\to\infty} \varphi(x^k) = 0.$

*Proof.* (i) For the given accumulation point $x^*$, there exists an infinite index set $\mathcal{K}$ , such that $x^k \xrightarrow{\mathcal{K}} x^*$. And from Lemma 3.3 (ii), we have $x^{k-1} \xrightarrow{\mathcal{K}} x^*$. In view of the finite choice for related sets and the boundedness of $\{(H_k, d^{k0}, \lambda^{k0})\}$, one can assume, without loss of generality, that the index set $\mathcal{K}$ satisfies

$$\begin{aligned}&d^{k0} \xrightarrow{\mathcal{K}} d^{*0},\ H_k \xrightarrow{\mathcal{K}} H_*,\ \lambda^{k0} \xrightarrow{\mathcal{K}} \lambda^{*0},\ I_k \equiv I',\ J_k \equiv J',\ J^+(x^k) \equiv J^+,\ J^-(x^k) \equiv J^-,\\ &l_k \equiv l',\ I_k^0 \equiv {I^0}' := I'\backslash\{l'\},\ L_k \equiv L' := {I^0}' \cup J',\ L(x^k) \equiv \widetilde{L},\end{aligned} \tag{3.4}$$

$$\begin{aligned}&d^{(k-1)0} \xrightarrow{\mathcal{K}} \bar{d}^{*0},\ H_{k-1} \xrightarrow{\mathcal{K}} \bar{H}_*,\ \lambda^{(k-1)0} \xrightarrow{\mathcal{K}} \bar{\lambda}^{*0},\ I_{k-1} \equiv \bar{I}', J_{k-1} \equiv \bar{J}',\ J^+(x^{k-1}) \equiv \bar{J}^+,\\ &J^-(x^{k-1}) \equiv \bar{J}^-,\ l_{k-1} \equiv \bar{l}', I_{k-1}^0 \equiv {\bar{I}^0}{}' := \bar{I}'\backslash\{\bar{l}'\}, L_{k-1} \equiv \bar{L}' := {\bar{I}^0}{}' \cup \bar{J}',\ L(x^{k-1}) \equiv \widetilde{L}'.\end{aligned} \tag{3.5}$$

Further, we define

$$\bar{f}_i^* := \begin{cases} F(x^*) - f_i(x^*), & i \in {I^0}'; \\ \varphi(x^*) - f_i(x^*), & i \in J^+ \cap J'; \\ -f_i(x^*), & i \in J^- \cap J', \end{cases} \tag{3.6}$$

$$\omega_* = \sum_{i\in L'} \max\{-\lambda_i^{*0},\ \lambda_i^{*0}\bar{f}_i^*\},\quad \lambda_{l'}^{*0} = 1 - \sum_{i\in {I^0}'} \lambda_i^{*0},\quad \bar{\omega}_* = \max\{-\lambda_{l'}^{*0},\ 0\}, \tag{3.7}$$

$$\rho_* = \frac{|g_{l'}(x^*)^T d^{*0}| + \omega_* + \bar{\omega}_*^2 + \varphi(x^*)}{1 + |(e^*)^T\lambda_{L'}^{*0}|}, \tag{3.8}$$

where $e^* = (1, 1, \cdots, 1)^T \in \mathbb{R}^{|L'|}$. In a similar fashion to the definition of $\rho_*$, one can also define $\bar{\rho}_*$ at the limit $x^*$ of $\{x^{k-1}\}_{\mathcal{K}}$ corresponding to sets $\bar{I}',\ \bar{J}',\ \bar{L}'$ and $\bar{l}'$.

Subsequently, suppose by contradiction that $x^*$ isn't a stationary point of the problem (1.1). Then the proof is divided into three steps as follows.

**Step A.** Show that $\rho_* > 0$ and $\bar{\rho}_* > 0$. So, for $k \in \mathcal{K}$ large enough, it follows that

$$\rho_k \geq 0.5\rho_*, \quad \delta_k = \min\{\varepsilon_k,\ \rho_{k-1}\} \xrightarrow{\mathcal{K}} \delta_* := \min\{\underline{\varepsilon},\ \bar{\rho}_*\} > 0, \quad \delta_k \geq \frac{1}{2}\delta_* > 0.$$

If $\rho_* = 0$, then, from (3.8), (3.7), and (3.1), one has

$$\begin{cases} g_{l'}(x^*)^T d^{*0} = 0 \implies d^{*0} = 0, \\ \omega_* = 0 \implies \lambda_i^{*0} \geq 0,\ \lambda_i^{*0} \bar{f}_i^* = 0,\ \forall\, i \in L', \\ \bar{\omega}_* = 0, \implies \lambda_{l'}^{*0} = 1 - \sum\limits_{i \in I^{0\prime}} \lambda_i^{*0} \geq 0, \\ \varphi(x^*) = 0 \implies x^* \in X, \text{ i.e.,} x^* \text{ is a feasible point.} \end{cases}$$

Again, taking the limit in the first equality of (2.26) for $k \in \mathcal{K}$ and combining $d^{*0} = 0$, we have

$$\sum_{i \in I^{0\prime}} \lambda_i^{*0}(g_i(x^*) - g_{l'}(x^*)) + \sum_{i \in J'} \lambda_i^{*0} g_i(x^*) = -g_{l'}(x^*),$$

i.e.,

$$\sum_{i \in I'} \lambda_i^{*0} g_i(x^*) + \sum_{i \in J'} \lambda_i^{*0} g_i(x^*) = 0.$$

Hence, the two relations above show that $x^*$ is a stationary point of the problem (1.1), and this is a contradiction. So $\rho_* > 0$ is at hand. The analysis for $\bar{\rho}_* > 0$ is similar.

**Step B.** Prove $\underline{t} := \inf\{t_k,\ k \in \mathcal{K}\} > 0$.

It is sufficient to show that inequalities (2.34)-(2.36) are all satisfied for $k \in \mathcal{K}$ large enough and real number $t > 0$ sufficiently small (independent of $k$). In the remaining analysis, the statement of "$k \in \mathcal{K}$ large enough and real number $t > 0$ sufficiently small" isn't repeated.

**B1.** Analyze inequality (2.34). For $i \in I$ , the proof is further divided into two cases, i.e., $i \in I(x^*)$ and $i \notin I(x^*)$ .

**B1.1.** For $i \in I(x^*)$, one has $f_i^k - F_k \to f_i(x^*) - F(x^*) = 0 > -\frac{1}{2}\delta_* > -\delta_k,\ k \in \mathcal{K}$ large enough, then $i \in I_k$ by (2.2). Again, taking into account the differentiability of $f_i(x)$ and the boundedness of $\{d^k\}_{\mathcal{K}}$ , and using Taylor expansion, we have

$$\begin{aligned} w_{ki}(t) &:= f_i(x^k + td^k) - F_k - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^\xi + r \sum_{i \in L_k} \lambda_i^{k0})) \\ &= f_i^k + t(g_i^k)^T d^k + o(t) - F_k - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^\xi + r \sum_{i \in L_k} \lambda_i^{k0})). \end{aligned} \tag{3.9}$$

Noting that $\bar{F}_{ki} \equiv 0,\ \bar{f}_i^k = F_k - f_i^k \xrightarrow{\mathcal{K}} F(x^*) - f_i(x^*) = 0$ for $i \in I(x^*)\backslash\{l'\}$, we can obtain from (2.21)-(2.22)

$$(g_i^k)^T d^k \leq -\sigma\rho_k^{1+\xi} + \sigma\varphi_k(\rho_k^\xi + r \sum_{i \in L_k} \lambda_i^{k0}) + O(\bar{f}_i^k), \qquad i \in I(x^*). \tag{3.10}$$

Substituting (3.10) into (3.9), it follows that

$$\begin{aligned} w_{ki}(t) &\leq f_i^k - F_k - (1-\alpha)\sigma t\rho_k^{1+\xi} + tO(\bar{f}_i^k) + o(t) \\ &\leq -0.5^{1+\xi}(1-\alpha)\sigma t\rho_*^{1+\xi} + o(t) \leq 0. \end{aligned}$$

**B1.2.** For $i \in I\backslash I(x^*)$, one gets $f_i(x^*) < F(x^*)$, and then

$$\begin{aligned} w_{ki}(t) &= f_i^k + tg_i(x^k)^T d^k - F_k - \sigma t(-\alpha\rho_k^{1+\xi} + \varphi_k(\rho_k^{\xi} + r\sum_{i\in L_k}\lambda_i^{k0})) + o(t) \\ &= f_i^k - F_k + O(t) \le 0.5(f_i(x^*) - F(x^*)) + O(t) \le 0. \end{aligned}$$

**B2.** Analyze inequalities (2.35) and (2.36) via four cases as follows.

**B2.1.** Consider $i \in J^+(x^k)$ and $f_i(x^*) = \varphi(x^*)$. In view of $x^k \xrightarrow{\mathcal{K}} x^*$ and $\delta_k \ge \frac{1}{2}\delta_* > 0$, one gets $i \in J_k = J'$ by (2.3). In addition, $\bar{f}_i^k = \varphi_k - f_i^k \to \varphi(x^*) - f_i(x^*) = 0,\ k \in \mathcal{K}$. So, we have from $\bar{F}_{ki} = 0$ and (2.24)

$$g_i(x^k)^T d^k \le -\sigma(\rho_k^{1+\xi} + r\varphi_k) + O(\bar{f}_i^k).$$

Using Taylor expansion, one gets from (2.40)

$$\begin{aligned} a_i^k(t) &\le f_i(x^k + td^k) - \varphi_k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) \\ &= f_i^k - \varphi_k + t(g_i^k)^T d^k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) + o(t) \\ &\le f_i^k - \varphi_k - (1-\alpha)\sigma t\rho_k^{1+\xi} - (1-\alpha)\sigma t r\varphi_k + tO(\bar{f}_i^k) + o(t) \\ &\le -0.5^{1+\xi}(1-\alpha)\sigma t\rho_*^{1+\xi} + o(t) \le 0. \end{aligned}$$

**B2.2.** Consider $i \in J^+(x^k)$ and $f_i(x^*) < \varphi(x^*)$. In view of the boundedness of $\{(d^k,\ \rho_k,\ \varphi_k)\}$ , and using Taylor expansion, we have from (2.40)

$$\begin{aligned} a_i^k(t) &\le f_i(x^k + td^k) - \varphi_k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) \\ &= f_i^k - \varphi_k + t(g_i^k)^T d^k + \alpha\sigma t(\rho_k^{1+\xi} + r\varphi_k) + o(t) \\ &= f_i^k - \varphi_k + O(t) \le 0.5(f_i(x^*) - \varphi(x^*)) + O(t) \le 0. \end{aligned}$$

**B2.3.** Consider $i \in J^-(x^k)$ and $f_i(x^*) = 0$. Taking into account $\delta_k \ge \frac{1}{2}\delta_* > 0$, we have $i \in J_k = J'$ from (2.3), and it further follows that $\bar{f}_i^k = -f_i^k \to -f_i(x^*) = 0,\ k \in \mathcal{K}$. Therefore, from (2.24) and $\bar{F}_{ki} = 0$, one has

$$\begin{aligned} f_i(x^k + td^k) &= f_i^k + tg_i(x^k)^T d^k + o(t) \\ &\le f_i^k - \sigma t(\rho_k^{1+\xi} + r\varphi_k) + tO(\bar{f}_i^k) + o(t) \\ &\le f_i^k - \sigma t\rho_k^{1+\xi} + tO(\bar{f}_i^k) + o(t) \\ &\le -0.5^{1+\xi}\sigma t\rho_*^{1+\xi} + o(t) \le 0. \end{aligned}$$

**B2.4.** Consider $i \in J^-(x^k)$ and $f_i(x^*) < 0$. In view of the boundedness of $\{d^k\}$, we have

$$f_i(x^k + td^k) = f_i^k + tg_i(x^k)^T d^k + o(t) \le 0.5f_i(x^*) + O(t) \le 0.$$

Now, summarizing the analysis in cases B2.3 and B2.4 above, it follows that $f_i(x^k + td^k) \le 0,\ \forall\, i \in J^-(x^k)$, for $k \in \mathcal{K}$ large enough. So $J^-(x^k) \subseteq J^-(x^k + td^k)$ holds, and then $|J^-(x^k)| \le |J^-(x^k + td^k)|$, inequality (2.36) holds. On the other hand, from the analysis of the four cases above, it shows that inequality (2.35) holds.

**Step C.** Using $t_k \ge \underline{t} > 0$ ($\forall\ k \in \mathcal{K}$) to get a final contradiction. If Case A in Remark 2.2 takes place, then $\{F_k\}_{k\ge s}$ is decreasing and bounded. So it is convergent. On the other hand, it follows from (2.38) that

$$F_{k+1} \le F_k - \sigma\alpha t_k\rho_k^{1+\xi} \le F_k - 0.5^{1+\xi}\sigma\alpha\underline{t}\rho_*^{1+\xi}, \quad k \ge s,\ k \in \mathcal{K}.$$

Passing to the limit in the inequality above, we immediately get $0 \leq -0.5^{1+\xi}\alpha\sigma\underline{t}\rho_*^{1+\xi}$, this together with $\underline{t} > 0$ and $\rho_* > 0$ brings a contradiction.

If Case B in Remark 2.2 takes place, then $\{\varphi_k\}_{k\geq 0}$ is decreasing and bounded. Thus, it is convergent. Therefore, according to (2.39), we have

$$0 = \lim_{k\in\mathcal{K},k\to\infty}(\varphi_{k+1} - \varphi_k) \leq \lim_{k\in\mathcal{K},k\to\infty} -\alpha\sigma t_k(\rho_k^{1+\xi} + r\varphi_k) \leq -0.5^{1+\xi}\alpha\sigma\underline{t}\rho_*^{1+\xi}.$$

This also brings a contradiction. Consequently, we can conclude that $x^*$ is a stationary point of the problem (1.1).

(ii) Based on conclusion (i), as well as the monotone and bounded property of $\{\varphi(x^k)\}$, we have $\lim\limits_{k\to\infty}\varphi_k = \lim\limits_{k\to\infty}\varphi(x^k) = \varphi(x^*) = 0$, where $x^*$ is an accumulation point of the sequence $\{x^k\}$. And the theorem is proved. $\square$

Subsequently, we further analyze the strong convergence of our proposed algorithm. For this purpose, the following assumption which is used in [6, 7, 11] is necessary.

**Assumption 3.2.** (i) The functions $f_i(x)$ $(i \in I \cup J)$ are all twice continuously differentiable over $\mathbb{R}^n$; (ii) The stationary pair $(x^*, \lambda^*)$ satisfies the upper-level strict complementarity and the strong second–order sufficiency conditions (SSOSC).

The details for the Assumption 3.2 (ii) above can be seen in [11, Assumption 4.1 (iii)]. Next, we give a lemma to show that $x^*$ is an isolated accumulation point of the problem (1.1) under certain conditions.

**Lemma 3.4.** *Suppose that $x^*$ is the accumulation point of the problem* (1.1) *and the stated assumptions hold. Then $x^*$ is an isolated accumulation point of* (1.1).

The proof of this lemma is similar to the one of Theorem 1.4.2 in [21] or Lemma 4.1 in [6], thus it is omitted here. Now, we can present the strong convergence of Algorithm 2.1 as follow.

**Theorem 3.2.** *Suppose that Assumptions* 2.1, 3.1 *and* 3.2 *are all satisfied. Then $\lim\limits_{k\to\infty} x^k = x^*$, i.e., Algorithm* 2.1 *is strongly convergent.*

*Proof.* From Lemma 3.3 (ii), we have $\lim_{k\to\infty}\|x^{k+1} - x^k\| = 0$. This, together with the isolation of the accumulation point $x^*$, implies that $\lim_{k\to\infty} x^k = x^*$, see [21, Corollary 1.1.8]. $\square$

## 4. Numerical Experiments

In this section, in order to validate the practical effectiveness of our proposed algorithm, we test the middle-small-scale constrained minimax problems. The numerical experiments were implemented in MATLAB 9.3.0.713579 (R2017b) on a personal computer with Windows 10, Intel(R) Core(TM) i5-8250U CPU 1.80 GHz and 8 GB RAM. During the numerical experiments, we set the parameters $\alpha = \beta = 0.5$, $\sigma = 0.19$, $\xi = 0.01$, $r = 12$, $\varepsilon_0 = 10$, and the approximate Hessian matrix $H_k$ is updated by the BFGS formula with Powells modification (see [6] for details), where $H_0$ is the identity matrix. The unified terminated criterion is $\rho_k < 10^{-5}$ for all test problems.

The first test group consists of 4 problems (P1-P4) taken from [5], where problems P1-P4 correspond to the test problems 1, 2, 5 and 7 in [5], respectively. And the second group

Table 4.1: Numerical comparisons for Algorithms 2.1 and A.

| Problem | $n/m/l$ | $x^0$ | $\varphi(x^0)$ | Algorithm | Ni | $F(x^*)$ | Tcpu |
|---|---|---|---|---|---|---|---|
| P1 | 2/3/2 | $(3,3)^T$ | 42.5 | Algorithm 2.1 | $1+12$ | 1.952225 | 0.02 |
| | | | | Algorithm A | $3+7$ | 1.952224 | 0.10 |
| P2 | 2/6/2 | $(1,3)^T$ | 0 | Algorithm 2.1 | 8 | 0.616435 | 0.01 |
| | | | | Algorithm A | 7 | 0.616432 | 0.04 |
| P3 | 4/4/3 | $(2,2,2,2)^T$ | 10 | Algorithm 2.1 | $6+25$ | $-43.999996$ | 0.04 |
| | | | | Algorithm A | $1+22$ | $-44.000000$ | 0.13 |
| P4 | 7/5/4 | $(3,\cdots,3)^T$ | 188 | Algorithm 2.1 | $2+50$ | 680.630077 | 0.09 |
| | | | | Algorithm A | $4+60$ | 680.630057 | 0.36 |
| P5 | 50/3/49 | $(0.5,\cdots,0.5)^T$ | 0 | Algorithm 2.1 | 63 | 98.000671 | 1.13 |
| | | | | Algorithm A | 147 | 98.000001 | 10.28 |
| | | $(0,\cdots,0)^T$ | 1 | Algorithm 2.1 | $1+54$ | 98.000781 | 0.88 |
| | | | | Algorithm A | $1+72$ | 98.000001 | 5.31 |
| P6 | 50/2/48 | $(2,1,\cdots,2,1)^T$ | 0 | Algorithm 2.1 | 147 | $-36.436048$ | 1.58 |
| | | | | Algorithm A | 38 | $-36.436316$ | 3.18 |
| | | $(1,\cdots,1)^T$ | 0.5 | Algorithm 2.1 | $1+137$ | $-36.436061$ | 1.44 |
| | | | | Algorithm A | $48+18$ | $-36.436316$ | 4.64 |
| P7 | 50/2/49 | $(0,\cdots,0)^T$ | 0 | Algorithm 2.1 | 151 | $-56.580049$ | 1.85 |
| | | | | Algorithm A | 92 | $-56.580326$ | 6.29 |
| | | $(1,\cdots,1)^T$ | 2 | Algorithm 2.1 | $1+99$ | $-56.580079$ | 1.14 |
| | | | | Algorithm A | $1+48$ | $-56.580326$ | 4.01 |
| P8 | 50/3/49 | $(0.5,\cdots,0.5)^T$ | 0 | Algorithm 2.1 | 113 | 198.345882 | 1.26 |
| | | | | Algorithm A | 78 | 198.345372 | 5.70 |
| | | $(-1,\cdots,-1)^T$ | 2 | Algorithm 2.1 | $1+246$ | 198.345961 | 5.66 |
| | | | | Algorithm A | $1+298$ | 198.345369 | 20.05 |
| P9 | 50/2/48 | $(1.5,\cdots,1.5)^T$ | 0 | Algorithm 2.1 | 135 | 119.115732 | 1.47 |
| | | | | Algorithm A | 157 | 119.114299 | 10.17 |
| | | $(0.5,\cdots,0.5)^T$ | 0.875 | Algorithm 2.1 | $1+163$ | 119.115682 | 1.69 |
| | | | | Algorithm A | $36+98$ | 119.114299 | 8.74 |
| P10 | 50/2/49 | $(0.5,\cdots,0.5)^T$ | 0 | Algorithm 2.1 | 19 | 0.000006 | 0.06 |
| | | | | Algorithm A | 10 | 0 | 1.17 |
| | | $(2,\cdots,2)^T$ | 11 | Algorithm 2.1 | $1+20$ | 0.000010 | 0.06 |
| | | | | Algorithm A | $3+15$ | 0 | 1.37 |

consists of 7 problems (P5-P11), these problems are composed of the corresponding objective functions and constraint functions in [24]. In particular, $\text{P5} = 2.4 + 4.6(1)$ (which means the objective and constraints of the problem P5 are 2.4 and 4.6(1) in [24], respectively, and the same blew), $\text{P6} = 2.3 + 4.1(2)$, $\text{P7} = 2.3 + 4.6(2)$, $\text{P8} = 2.4 + 4.6(2)$, $\text{P9} = 2.9 + 4.7$, $\text{P10} = 2.9 + 4.6(2)$, $\text{P11} = 2.1 + 4.6(1)$.

The numerical results are listed in Tables 1 and 2. The following notations are used: "n": the dimensions of the variable $x$; "m": the number of all component objective functions; "l": the number of constraint functions; "$x^0$": the initial feasible point; "Ni": the number of iterations (when Ni is a sum of two numbers, the former and the latter indicate the number of iterations outside and inside the feasible set, respectively); "NF": the number of all component functions evaluations in the objective; "NC": the number of constraints evaluations; "$F(x^*)$": the final objective value; "$\rho_*$": the approximate identification value at the final iteration point; "Tcpu": computing time of CPU (in seconds); "Algorithm A": the algorithm of [7], and the corresponding data of Algorithms A in Table 4.1 are reported in [7].

Table 4.2: Numerical results of Algorithm 2.1.

| Problem | $x^0$ | $\varphi(x^0)$ | $n/m/l$ | Ni | NF | NC | $F(x^*)$ | $\rho_*$ | Tcpu |
|---|---|---|---|---|---|---|---|---|---|
| P5 | $(3.5,\cdots,3.5)^T$ | 23.75 | 100/3/99 | $25+124$ | 2557 | 90321 | 198.001766 | $9.8357e-06$ | 4.22 |
| | | | 150/3/149 | $25+79$ | 2411 | 146180 | 298.001952 | $7.1367e-06$ | 6.58 |
| | | | 200/3/199 | $25+61$ | 1667 | 126592 | 398.002627 | $8.8239e-06$ | 7.33 |
| | | | 300/3/299 | $25+105$ | 2178 | 249242 | 598.004368 | $8.8252e-06$ | 14.74 |
| P6 | $(1,\cdots,1)^T$ | 0.5 | 100/2/98 | $1+127$ | 2398 | 48204 | $-73.239238$ | $8.6121e-06$ | 2.98 |
| | | | 150/2/148 | $1+131$ | 2372 | 71970 | $-110.042615$ | $6.0118e-06$ | 4.92 |
| | | | 200/2/198 | $1+105$ | 1768 | 72140 | $-146.845472$ | $9.4005e-06$ | 5.87 |
| | | | 300/2/298 | $1+125$ | 2250 | 127198 | $-220.451924$ | $8.3525e-06$ | 10.21 |
| P7 | $(6,\cdots,6)^T$ | 107 | 100/2/99 | $50+137$ | 4025 | 125959 | $-114.314785$ | $9.7788e-06$ | 4.13 |
| | | | 150/2/149 | $62+170$ | 5069 | 240724 | $-172.049576$ | $9.2432e-06$ | 8.64 |
| | | | 200/2/199 | $62+178$ | 5235 | 323382 | $-229.784348$ | $9.1351e-06$ | 17.18 |
| | | | 300/2/299 | $62+232$ | 6213 | 539402 | $-345.253830$ | $9.3946e-06$ | 30.03 |
| P9 | $(0.5,\cdots,0.5)^T$ | 0.875 | 100/2/98 | $1+161$ | 2817 | 53370 | 248.406681 | $9.5585e-06$ | 3.54 |
| | | | 150/2/148 | $1+167$ | 2850 | 81340 | 377.697717 | $9.8484e-06$ | 5.85 |
| | | | 200/2/198 | $1+177$ | 2998 | 114242 | 506.988534 | $9.6540e-06$ | 9.27 |
| | | | 300/2/298 | $1+196$ | 3281 | 192362 | 765.573917 | $1.3291e-06$ | 20.94 |
| P10 | $(5,\cdots,5)^T$ | 74 | 100/2/99 | $12+23$ | 841 | 45051 | 0.000007 | $6.7653e-06$ | 0.15 |
| | | | 150/2/149 | $21+22$ | 1182 | 97607 | 0.000005 | $7.1377e-06$ | 0.25 |
| | | | 200/2/199 | $21+26$ | 1322 | 145891 | 0.000012 | $9.2633e-06$ | 0.48 |
| | | | 300/2/299 | $20+27$ | 1359 | 224573 | 0.000003 | $2.4150e-06$ | 0.89 |
| P11 | $(0.8,\cdots,0.8)^T$ | 0 | 100/100/99 | 95 | 62100 | 29674 | 0.111111 | $1.6572e-15$ | 3.79 |
| | | | 150/150/149 | 145 | 141750 | 67022 | 0.111111 | $2.2313e-15$ | 9.84 |
| | | | 200/200/199 | 181 | 205000 | 110180 | 0.111111 | $1.5667e-17$ | 18.98 |

In Table 4.1, we compare our algorithm with the SQP algorithm for minimax problem in [7], regardless of starting from the feasible points or not. The performance of these two algorithms is similar in terms of the approximate optimal objective value at the final iteration point, and two algorithms have their own advantages with respect to the total number of iterations, thus the comparisons show that our algorithm is promising. Furthermore, Algorithm 2.1 solves two SLEs with a same coefficient matrix at each iteration, while Algorithm A need to solve a norm-relaxed quadratic programming subproblem and a SLE at each iteration, so we point out that our algorithm can reduce the computing time of CPU in the experiments. In Table 4.2, we see that Algorithm 2.1 can efficiently solve all the problems with dimensions varying from 100 to 300, except for Problem 12 with the dimension $n = 300$, the reason is that the coefficient matrix of Problem 12 ($n = 300$) in the SLEs is close to singular or badly scaled during the implementation, so its numerical results are not given in Table 4.2.

In summary, through the numerical results in the two tables and the analysis above, we can get that our proposed algorithm is promising for the middle-small-scale constrained minimax problems.

## 5. Conclusion

In this work, based on the stationary conditions of the nonlinear minimax problems with inequality constraints, we propose a SSLE-type algorithm of quasi-strongly sub-feasible directions starting from an arbitrary initial iteration point. By means of a new working set, we develop a new technique for constructing the coefficient matrix of the SLE. At each iteration, two SLEs with the same uniformly nonsingular coefficient matrix are solved. Under mild conditions, the proposed algorithm possesses global and strong convergence.

In terms of further work, we think the ideas in this work can be extended to minimax problems with equality and inequality constraints or other optimization problems.

**Acknowledgments.** This work was supported by the Research Foundation of Guangxi University for Nationalities (No. 2021KJQD04), the Natural Science Foundation of Guangxi Province (No. 2018GXNSFAA281099) and NSFC (No. 11771383).

## References

[1] E. Polak, On the mathematical foundations of nondifferentiable optimization in engineering design, *SIAM Rev.*, **29**:1 (1987), 21–89.

[2] E. Polak, S.E. Salcudean, and D.Q. Mayne, Adaptive control of ARMA plants using worst case design by semi-infinite optimization, *IEEE T. Automat. Contr*, **32**:5 (1987), 388–396.

[3] X.Q. Cai, K.L. Teo, X.Q. Yang, and X.Y. Zhou, Portfolio optimization under a minimax rule, *Manage. Sci.*, **46**:7 (2000), 957–972.

[4] E.R. Panier and A.L. Tits, A globally convergent algorithm with adaptively refined discretization for semi-infinite optimization problems arising in engineering design, *IEEE T. Automat. Contr*, **34**:8 (1989), 903–908.

[5] B. Rustem and Q. Nguyen, An algorithm for the inequality-constrained discrete minimax problem, *SIAM J. Optim.*, **8**:1 (1998), 265–283.

[6] J.B. Jian, R. Quan, and X.L. Zhang, Feasible generalized monotone line search SQP algorithm for nonlinear minimax problems with inequality constraints, *J. Comput. Appl. Math.*, **205**:1 (2007), 406–429.

[7] J.B. Jian, Q.J. Hu, and C.M. Tang, Superlinearly convergent Norm-Relaxed SQP method based on active set identification and new line search for constrained Minimax problems, *J. Optimiz. Theory App.*, **163** (2014), 859–883.

[8] E. Obasanjo, G. Tzallas-Regas, and B. Rustem, An interior-point algorithm for nonlinear minimax problems, *J. Optimiz. Theory App.*, **144**:2 (2010), 291–318.

[9] E.Y. Pee and J.O. Royset, On solving large-scale finite Minimax problems using exponential smoothing, *J. Optimiz. Theory App.*, **148**:2 (2011), 390–421.

[10] F.S. Wang, C.L. Wang, and L. Wang, A new trust-region algorithm for finite minimax problem, *J. Comput. Math.*, **30**:3 (2012), 262–278.

[11] J.B. Jian, X.D. Mo, L.J. Qiu, S.M. Yang, and F.S. Wang, Simple sequential quadratically constrained quadratic programming feasible algorithm with active identification sets for constrained Minimax problems, *J. Optimiz. Theory App.*, **160**:1 (2014), 158–188.

[12] E.R. Panier, A.L. Tits, and J.N. Herskovits, A QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization, *SIAM J. Control. Optim.*, **26**:4 (1988), 788–811.

[13] Y.F. Yang, D.H. Li, and L.Q. Qi, A feasible sequnetial system of linear equations method for inequality constrained optimization, *SIAM J. Optim.*, **13**:4 (2003), 1222–1244.

[14] Z.B. Zhu, An interior point type QP-free algorithm with superlinear convergence for inequality constrained optimization, *Appl. Math. Model.*, **31**:6 (2007), 1201–1212.

[15] J.B. Jian, D.L. Han, and Q.J. Xu, A new sequential systems of linear equations algorithm of feasible descent for inequality constrained optimization, *Acta Math. Sin.*, **26**:12 (2010), 2399–2420.

[16] G.D. Ma and J.B. Jian, A QP-free algorithm of quasi-strongly sub-feasible directions for inequality constrained optimization, *J. Ind. Manag. Optim.*, **11**:1 (2015), 307–328.

[17] J.L. Li and Z.P. Yang, A QP-free algorithm without a penalty function or a filter for nonlinear general-constrained optimization, *Appl. Math. Comput.*, **316** (2018), 52–72.

[18] J.B. Jian, H.Q. Pan, C.M. Tang, and J.L. Li, A strongly sub-feasible primal-dual quasi interior-point algorithm for nonlinear inequality constrained optimization, *Appl. Math. Comput.*, **266** (2015), 560–578.

[19] J.B. Jian, H.Y. Zheng, C.M. Tang, and Q.J. Hu, A new superlinearly convergent norm-relaxed method of strongly sub-feasible direction for inequality constrained optimization, *Appl. Math. Comput.*, **182** (2006), 955–976.

[20] J.B. Jian, C.M. Tang, and H.Y. Zheng, Sequential quadratically constrained quadratic programming norm-relaxed algorithm of strongly sub-feasible directions, *Eur. J. Oper. Res.*, **200** (2010), 645–657.

[21] J.B. Jian, *Fast algorithms for smooth constrained optimization – theoretical analysis and numerical experiments*, Science Press, Beijing, 2010.

[22] J.B. Jian, Y.H. Chen, and C.H. Guo, A strongly convergent method of quasi-strongly sub-feasible directions for constrained optimization, *Pac. J. Optim.*, **7** (2011), 339–351.

[23] J.B. Jian and Y. Liu, A superlinearly convergent method of quasi-strongly sub-feasible directions with active set identifying for constrained optimization, *Nonlinear Anal-Real*, **12**:5 (2011), 2717–2729.

[24] N. Karmitsa, Test problems for large-scale nonsmooth minimization, Reports of the Department of Mathematical Information Technology, Series B, Scientific computing, No. B. 4 (2007).