Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2015.m1338

Vol. **8**, No. 3, pp. 313-335
August 2015

A Globally and Superlinearly Convergent Primal-dual Interior Point Method for General Constrained Optimization

Jianling Li[1], Jian Lv[2] and Jinbao Jian[3,*]

[1] *College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, 530004, China.*
[2] *School of Mathematics Science, Dalian University of Technology, Dalian, 116024, China.*
[3] *School of Mathematics and Information Science, Yulin Normal University; Guangxi Colleges and Universities Key Lab of Complex System Optimization and Big Data Processing, Yulin 537000, China.*

Received 14 December 2013; Accepted (in revised version) 28 June 2014

Abstract. In this paper, a primal-dual interior point method is proposed for general constrained optimization, which incorporated a penalty function and a kind of new identification technique of the active set. At each iteration, the proposed algorithm only needs to solve two or three reduced systems of linear equations with the same coefficient matrix. The size of systems of linear equations can be decreased due to the introduction of the working set, which is an estimate of the active set. The penalty parameter is automatically updated and the uniformly positive definiteness condition on the Hessian approximation of the Lagrangian is relaxed. The proposed algorithm possesses global and superlinear convergence under some mild conditions. Finally, some preliminary numerical results are reported.

AMS subject classifications: 65K05, 90C30, 90C31

Key words: general constrained optimization, primal-dual, active set, global convergence, superlinear convergence.

1. Introduction

Consider the following general constrained optimization problem

$$\text{(P)} \quad \begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \le 0, \quad i \in I = \{1, 2, \ldots, m\}, \\ & f_i(x) = 0, \quad i \in L = \{m+1, m+2, \ldots, m+l\}, \end{array}$$

*Corresponding author. *Email addresses:* jianlingli@126.com (J.L. Li), 710261120@qq.com (J. Lv), jianjb@gxu.edu.cn (J. B. Jian)

 ©2015 Global-Science Press

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

where functions $f_0, f_i\ (i \in I \cup L) : R^n \to R$ are all continuously differentiable.

It is well known that sequential quadratic programming (SQP) methods is one of efficient methods for constrained optimization. The interested readers are referred to Boggs et al. [1] and Gill et al. [5] for more information on SQP methods. However, SQP methods have to solve QP subproblems at each iteration, which is computationally expensive. Panier et al. in [10] presented a feasible QP-free algorithm (Algorithm PTH for short) for (P) only with inequality constraints. At each iteration, only three systems of linear equations (SLEs) need to be solved. Specifically, for the current iterative point x^k, the master search direction d^k is generated by solving two SLEs in (d, λ) with the following form

$$\begin{pmatrix} H_k & A_k \\ Z_k A_k^{\mathrm{T}} & G_k \end{pmatrix} \begin{pmatrix} d \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla f_0(x^k) \\ \eta \end{pmatrix}, \tag{1.1}$$

where H_k is an approximation of the Hessian of the Lagrangian associated with (P), the vector η is chosen in different ways, $A_k = (\nabla f_i(x^k), i \in I)$, $G_k = \mathrm{diag}(f_i(x^k), i \in I)$, $Z_k = \mathrm{diag}(z_i^k, i \in I)$ and $z_i^k\ (i \in I)$ the current estimate of the KKT multipliers. In order to avoid the Maratos effect, the search direction is modified by solving a least squares subproblem, which is equivalent to an SLE. The global convergence of Algorithm PTH requires a strong assumption, i.e., the number of stationary points is finite.

Algorithm PTH was later improved by Gao et al. in [6]. An extra SLE is solved and the algorithm globally converges to a KKT point under some conditions including a strong assumption that the multiplier sequence is bounded, which is impossible if the iteration matrix is ill-conditioned. Afterwards, based on the Fischer-Burmeister function and the KKT nonsmooth equation system, Qi and Qi in [12] also improved Algorithm PTH and proposed a new feasible QP-free algorithm. At each iteration, three SLEs, which differ from those in [10], are required to solve. It is shown that the coefficient matrix of the SLE is uniformly nonsingular and the multiplier approximation sequence is bounded, even if the strict complementarity is not satisfied. Without assuming the isolatedness of the stationary points, the global convergence is proved under some suitable conditions.

As we know, primal-dual interior point methods are also one of efficient solution methods for problem (P) (see, e.g., [2,11,13,15]). Especially, Bakhtiari and Tits in [2] proceeded along the lines of Algorithm PTH and proposed a simply feasible primal-dual interior-point method (Algorithm BT for short). For completeness, we first give a short introduction about feasible primal-dual interior point (PDIP) methods.

The KKT system of problem (P) only with inequality constraints is given as follows

$$\begin{cases} \nabla_x L(x, \lambda) = 0, \\ \lambda_i f_i(x) = 0,\ \lambda_i \geq 0,\ f_i(x) \leq 0,\ i \in I, \end{cases} \tag{1.2}$$

where the Lagrangian function is given by $L(x, \lambda) = f_0(x) + \sum_{i \in I} \lambda_i f_i(x)$. The key point of PDIP iteration is a perturbed variant of (1.2) in (x, λ) as follows

$$\begin{cases} \nabla_x L(x, \lambda) = 0, \\ \lambda_i f_i(x) = \mu_i,\ i \in I. \end{cases} \tag{1.3}$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

The vector $\mu \in R^m$ is the "barrier parameter" vector with component $\mu_i < 0,\ i \in I$. The main idea is then to attempt to solve (1.3) by means of Newton or quasi-Newton iterations, while driving μ to zero and enforcing primal and dual (strict) feasible at each iteration. Specifically, the following SLE in $(\triangle x, \triangle\lambda)$ is considered at pair (x^k, λ^k)

$$\begin{cases} H_k\triangle x + \sum\limits_{i\in I}\nabla f_i(x^k)\triangle\lambda_i + \nabla f_0(x^k) + \sum\limits_{i\in I}\nabla f_i(x^k)\lambda_i^k = 0, \\ \lambda_i^k\nabla f_i(x^k)^{\mathrm{T}}\triangle x + f_i(x^k)\triangle\lambda_i + \lambda_i^k f_i(x^k) = \mu_i,\ i \in I. \end{cases}$$

The system above can be equivalently rewritten as

$$\begin{pmatrix} H_k & A_k \\ \mathrm{diag}(\lambda^k)A_k^{\mathrm{T}} & G_k \end{pmatrix}\begin{pmatrix} \triangle x \\ \triangle\lambda + \lambda^k \end{pmatrix} = \begin{pmatrix} -\nabla f_0(x^k) \\ \mu \end{pmatrix}, \tag{1.4}$$

which is the same as (1.1) except that each component of μ is negative.

Now let us go back to [2]. The master search direction of Algorithm BT is yielded by solving only two SLEs similar to (1.4), and a higher correction direction is obtained by solving a least squares subproblem to avoid the Maratos effect. The barrier parameter vector μ is carefully constructed to guarantee that Algorithm BT possesses some good properties as follows: (1) the objective function value decreases at each iteration; (2) the positive definite restriction on the Hessian estimate is relaxed; (3) the initial point may lie on the boundary of the feasible set, not necessary in the strict interior of the feasible set; (4) Algorithm BT possesses global convergence under milder conditions.

The algorithms mentioned above share some common features: all the inequality constraints and their gradients are involved in the associated SLEs. Consequently, the size of the SLEs may be very large. Furthermore, the coefficient matrix of the third SLE differs from that of the first two SLEs, which leads to increasing computation cost. Algorithm PTH and Algorithm BT are only two-step superlinear convergent.

Tits et al. in [13] extended Algorithm PTH to solve problem (P) with equality and inequality constraints by incorporating a modified Mayne-Polak scheme (see [9]). In their algorithm (Algorithm TWBUL for short), the positive definiteness assumption on the Hessian estimate is relaxed and all the convergence results obtained in [10] still hold. Moreover, they developed a simpler penalty parameter update rule. Similar to Algorithm PTH and Algorithm BT, Algorithm TWBUL also has the features described above.

Recently, based on the idea of PDIP and a new identification technique of the active set, Jian et al. in [8] proposed a feasible QP-free algorithm for problem (P) only with inequality constraints. The size of SLEs solved at each iteration is smaller than that of (1.1) and (1.4) due to the use of the working set, which is an estimate of the active set. Moreover, the three SLEs solved at each iteration have the same coefficient matrix. The algorithm possesses global and superlinear convergence under suitable conditions, including the uniformly positive definiteness of the Hessian estimate and the strict complementarity.

In this paper, motivated from the idea of [8], we propose a new primal-dual method for problem (P) with equality and inequality constraints, which incorporates a penalty

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

function and a new identification technique of the active set. The main properties of the proposed algorithm are as follows:

- the penalty function is a semi-penalty function, i.e., only equality constraints are penalized and the penalty parameter is automatically updated at each iteration;
- the size of SLEs is smaller than that of (1.1) and (1.4) due to the introduction of the working set; and the coefficient matrices of the three SLEs solved at each iteration are the same. As a consequence, the computation cost is decreased;
- the uniformly positive definiteness assumption on the Hessian estimate is weakened, which is used in most of existing algorithms, such as [8, 10, 12] ;
- the algorithm possesses global convergence and one-step superlinear convergence under some mild conditions.

The remainder of the paper is organized as follows. The new algorithm is described in detail and its feasibility is discussed in Section 2. The global and superlinear convergence are analyzed in Section 3 and Section 4, respectively. Numerical experiments and results are stated in Section 5 and Finally, some concluding remarks are given in Section 6.

2. The algorithm

In this section, we will present our algorithm and then discuss its important properties. For the sake of simplicity, we use the following natations throughout the paper:

$$
\begin{aligned}
&X = \{x \in R^n :\ f_i(x) \le 0, i \in I;\ f_i(x) = 0, i \in L\},\\
&\bar{X} = \{x \in R^n :\ f_i(x) \le 0, i \in I\},\ \bar{X}_o = \{x \in R^n :\ f_i(x) < 0, i \in I\},\\
&I(x) = \{i \in I : f_i(x) = 0\},\ \ f_{_J}(x) = (f_i(x),\ i \in J \subseteq I \cup L),\\
&g_i(x) = \nabla f_i(x),\ i \in \{0\} \cup I \cup L,\ \ A_{_J}(x) = (g_i(x),\ i \in J \subseteq I).
\end{aligned}
$$

Before proceeding, we state a basic assumption throughout the paper.

(A1) Gradient vectors $\{g_i(x),\ i \in L\}$ are linearly independent for each $x \in \bar{X}_o$.

Remark 2.1. From our algorithm given in the sequel, all iteration points satisfy strict inequality constraints, so the assumption (A1) is weaker than the standard LICQ.

In order to estimate the active constraints, we define a vector-valued function by the way in [14]:

$$
\Phi(x, \lambda(x), \gamma(x)) = \begin{pmatrix} \nabla_x L(x, \lambda(x), \gamma(x)) \\ \min\{-f_{_I}(x), \lambda(x)\} \\ f_{_L}(x) \end{pmatrix},
$$

where $L(x, \lambda, \gamma)$ is the Lagrange function of problem (P), i.e.,

$$
L(x, \lambda, \gamma) = f_0(x) + \sum_{i \in I} \lambda_i f_i(x) + \sum_{i \in L} \gamma_i f_i(x).
$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Further, we define an optimal identification function for problem (P) as $\varphi:\ R^{n+m+l} \to R$:

$$\varphi(x, \lambda(x), \gamma(x)) = \|\Phi(x, \lambda(x), \gamma(x))\|^{\zeta}, \tag{2.1}$$

where $\zeta \in (0, 1)$, $\|\cdot\|$ denotes the Euclidean norm. It is clear that $(x^*, \lambda^*, \gamma^*)$ is a KKT pair of problem (P) if and only if $\varphi(x^*, \lambda^*, \gamma^*) = 0$. Moreover, we know from [14] that the index set $\{i \in I : f_i(x^k) + \varphi(x^k, \lambda(x^k), \gamma(x^k)) \geq 0\}$ is an exact identification set for the active set $I(x^*)$ if the pair $(x^k, \lambda(x^k), \gamma(x^k))$ is sufficiently close to a KKT pair $(x^*, \lambda^*, \gamma^*)$, and the Mangasarian-Fromovitz constraint qualification (MFCQ) and the second order sufficient conditions are satisfied at a KKT pair $(x^*, \lambda^*, \gamma^*)$.

In this paper, in order to reduce computation cost, inspired from the idea of [14], for the current iteration point $x^k \in \bar{X}_o$, the multiplier vectors $\lambda(x^k)$ and $\gamma(x^k)$ are chosen as follows:

$$\begin{aligned} &\lambda(x^0) = z^0, \quad \gamma(x^0) = 1; \\ &\lambda(x^k) = \bar{\lambda}^{k-1}, \ \ \gamma(x^k) = \bar{\gamma}^{k-1}, \ \ k \geq 1, \end{aligned}$$

where $z^0 > 0$, and $\bar{\lambda}^{k-1}$, $\bar{\gamma}^{k-1}$ are computed in the $(k-1)$-th iteration (see Algorithm 2.1 below). So we define the working set I_k by

$$I_k = \{i \in I :\ f_i(x^k) + \varphi(x^k,\ \lambda(x^k),\ \gamma(x^k)) \geq 0\}. \tag{2.2}$$

With the help of I_k, only two or three reduced SLEs with the same coefficient matrix are solved at each iteration in our algorithm, and the inactive constraints at a KKT point x^* will be neglected after finite number of iterations. In view of the equality constraints, the coefficient matrices of the SLEs in our algorithm are constructed with the form

$$M(x^k,\ z^k_{I_k},\ H_k) = \begin{pmatrix} H_k & A_k & B_k \\ Z_k A_k^{\mathrm{T}} & G_k & 0 \\ B_k^{\mathrm{T}} & 0 & 0 \end{pmatrix}, \tag{2.3}$$

where H_k is an approximation of the Hessian of the Lagrangian $L(x, \lambda(x), \gamma(x))$, and

$$\begin{aligned} &A_k = A_{I_k}(x^k) = (g_i(x^k),\ i \in I_k), \quad B_k = (g_i(x^k), i \in L),\ z^k_{I_k} \in R^{|I_k|}, \\ &Z_k = \mathrm{diag}(z^k_{I_k}), \quad G_k = G_{I_k}(x^k) = \mathrm{diag}(f_{I_k}(x^k)), \end{aligned}$$

where $|I_k|$ means the cardinality of I_k. The search directions of our algorithm are generated by solving two or three SLEs of the form

$$M(x^k,\ z^k_{I_k},\ H_k) \begin{pmatrix} d \\ \lambda_{I_k} \\ \gamma \end{pmatrix} = \begin{pmatrix} -g_0(x^k) \\ \mu^k \\ \varsigma^k \end{pmatrix}$$

with different vectors μ^k and ς^k.

In this paper, the merit function for arc search is a semi-penalty function defined by

$$\Psi(x, r) = f_0(x) + r \sum_{i \in L} |\ f_i(x)\ |, \tag{2.4}$$

where $r > 0$ is a penalty parameter.

Now, the proposed algorithm is described in detail.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Algorithm 2.1. *(Primal-dual Interior Point Method)*

Parameters: $\alpha \in (0, \frac{1}{2}),\ \sigma,\ \beta,\ \theta,\ \zeta \in (0,1),\ \tau \in (2,3),\ r_0 > 0, \nu > 2,\ \rho_1 > 0,\ \rho_2 > 0,\ \overline{M} > 0,\ p > 0$ *and* $z_{\max} > z_{\min} > 0$.

Data: $z^0 \in R^m$ *with* $z_j^0 \in (z_{\min}, z_{\max}],\ j \in I, H_0 \in R^{n\times n}$ *is a symmetric and positive definite matrix; a starting point* x^0 *satisfies* $x^0 \in \bar{X}_0$, *or* $x^0 \in \bar{X}$ *such that vectors* $\{g_i(x^0),\ i \in I(x^0) \cup L\}$ *are linearly independent. Set* $k = 0$.

Step 1. *Generate the working set* I_k *and matrix* H_k. *Compute* $\varphi(x^k, \lambda(x^k), \gamma(x^k))$ *by* (2.1), *then generate the working set* I_k *by* (2.2). *For* $k \geq 1$, *compute a symmetric matrix* H_k *such that it is an approximation of the Lagrange Hessian of the problem (P) such that matrix*

$$W_k \overset{\text{def}}{=} H_k - \sum_{i\in I_k} \frac{z_i^k}{f_i(x^k)} g_i(x^k) g_i(x^k)^{\mathrm{T}}$$

is positive definite.

Step 2. *Compute the search directions. Denote* $M_k = M(x^k,\ z_{I_k}^k,\ H_k)$.

(i) Compute $(\bar{d}^k, \bar{\lambda}_{I_k}^k, \bar{\gamma}^k)$ *by solving the first SLE in* (d, λ, γ):

$$M_k \begin{pmatrix} d \\ \lambda_{I_k} \\ \gamma \end{pmatrix} = \begin{pmatrix} -g_0(x^k) \\ 0 \\ -f_L(x^k) \end{pmatrix}. \tag{2.5}$$

Set $\bar{\lambda}^k = (\bar{\lambda}_{I_k}^k, 0_{I\setminus I_k})$. *If* $\bar{d}^k = 0$ *and* $\bar{\lambda}_{I_k}^k \geq 0$, *then* x^k *is a KKT point of (P), stop.*

(ii) Compute $(d^k, \lambda_{I_k}^k, \gamma^k)$ *by solving the second SLE in* (d, λ, γ):

$$M_k \begin{pmatrix} d \\ \lambda_{I_k} \\ \gamma \end{pmatrix} = \begin{pmatrix} -g_0(x^k) \\ \mu_{I_k}^k \\ -f_L(x^k) \end{pmatrix}, \tag{2.6a}$$

where

$$\mu_i^k = (1-\beta_k)\phi_i^k + \beta_k(-\|\bar{d}^k\|^\nu - |\psi_k|^\nu) z_i^k,\ \ i \in I_k; \tag{2.6b}$$

$$\psi_k = g_0(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in I_k} \frac{\bar{\lambda}_i^k}{z_i^k}\phi_i^k - \sum_{i\in L}\left(\bar{\gamma}_i^k f_i(x^k) + |f_i(x^k)|\right); \tag{2.6c}$$

$$\phi_i^k = \min\left\{0, -(\max\{-\bar{\lambda}_i^k, 0\})^p - \bar{M} f_i(x^k)\right\},\ \ i \in I_k; \tag{2.6d}$$

$$\beta_k = \begin{cases} 1, & \text{if } b_k \leq 0; \\ \min\{\frac{(1-\theta)|\psi_k|}{b_k}, 1\}, & \text{if } b_k > 0; \end{cases} \tag{2.6e}$$

$$b_k = \sum_{i\in I_k} \frac{\bar{\lambda}_i^k}{z_i^k}\left((\|\bar{d}^k\|^\nu + |\psi_k|^\nu) z_i^k + \phi_i^k\right). \tag{2.6f}$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Step 3. *Update the penalty parameter.* Denote $\eta_k = \max\limits_{i\in L}\{1+|2\bar{\gamma}_i^k-\gamma_i^k|, |\gamma_i^k|\}$, and set

$$r_k = \begin{cases} r_{k-1}, & \text{if } r_{k-1} \geq \eta_k + \rho_1; \\ \max\{r_{k-1}, \eta_k\} + \rho_1, & \text{otherwise.} \end{cases} \tag{2.7}$$

Step 4. *Trial of unit step-length.* If

$$\Psi(x^k + d^k,\ r_k) \leq \Psi(x^k,\ r_k) + \alpha\Gamma(x^k, d^k, r_k), \tag{2.8a}$$
$$f_i(x^k + d^k) < 0, \quad \forall i \in I, \tag{2.8b}$$

then set the step size $t_k \equiv 1$ and the higher correction direction $\widetilde{d}^k = 0$, enter Step 7. Here, the function $\Gamma(x^k, d^k, r_k)$ is defined by

$$\Gamma(x^k, d^k, r_k) = g_0(x^k)^{\mathrm{T}} d^k - r_k \sum_{i\in L} \mid f_i(x^k) \mid . \tag{2.9}$$

Step 5. *Generate correction direction.* Compute $(\widetilde{d}^k,\ \widetilde{\lambda}^k_{I_k}, \widetilde{\gamma})$ by solving the third SLE in (d, λ, γ):

$$M_k \begin{pmatrix} d \\ \lambda_{I_k} \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ \widetilde{\mu}^k_{I_k} \\ \widetilde{\varsigma}^k \end{pmatrix}, \tag{2.10a}$$

where

$$\widetilde{\mu}_i^k = \begin{cases} -\delta_k - z_i^k f_i(x^k + d^k) - \frac{(z_i^k)^2}{\lambda_i^k}\beta_k|\psi_k|^\nu, & \text{if } \lambda_i^k \neq 0, i \in I_k; \\ -\delta_k - z_i^k f_i(x^k + d^k), & \text{if } \lambda_i^k = 0, i \in I_k; \end{cases} \tag{2.10b}$$
$$\widetilde{\varsigma}_i^k = -\|d^k\|^\tau - f_i(x^k + d^k), \quad i \in L; \tag{2.10c}$$
$$\delta_k = \max\left\{\|d^k\|^\tau,\ \|d^k\|^2 \max_{i\in I_k, \lambda_i^k \neq 0}\{|1 - \frac{z_i^k}{\lambda_i^k}|^\sigma\}\right\}. \tag{2.10d}$$

If $\|\widetilde{d}^k\| > \|d^k\|$, set $\widetilde{d}^k = 0$.

Step 6. *Perform arc search.* Let t_k be the first number t of the sequence $\{1,\ \beta,\ \beta^2, \ldots\}$ satisfying the following inequalities:

$$\Psi(x^k + td^k + t^2\widetilde{d}^k,\ r_k) \leq \Psi(x^k, r_k) + \alpha t\Gamma(x^k, d^k, r_k), \tag{2.11a}$$
$$f_i(x^k + td^k + t^2\widetilde{d}^k) < 0, \quad i \in I. \tag{2.11b}$$

Step 7. *Update.* Set $x^{k+1} = x^k + t_k d^k + t_k^2\widetilde{d}^k$, $z_i^{k+1} = \min\{\max\{\|d^k\|^2 + z_{\min}, \lambda_i^k\}, z_{\max}\},\ i \in I$. Set $k := k+1$, go back to Step 1.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Remark 2.2. (1) ψ_k given in Step 2(ii) plays an important role in verifying that Algorithm 2.1 is well defined. We can prove that x^k is a KKT point when $\psi_k = 0$ (see Lemma 2.2 below). (2) $\mu^k_{I_k}$ defined in Step 2(ii) is used to generate a suitable search direction d^k, which can keep not only x^k staying in $\bar{X}_o$ but also $\Psi(x, r)$ decreasing.

Remark 2.3. The updating rule of z^k in Step 7 plays a key role in ensuring that the coefficient matrix M_k is uniformly nonsingular. With this aim, z^k can be updated in different ways, such as $z_i^{k+1} = \min\{\max\{z_{\min}, \lambda_i^k\}, z_{\max}\}$, $i \in I$.

Remark 2.4. Since d^k is the solution to the SLE (2.5), $\Gamma(x^k, d^k, r_k)$ defined in (2.9) is identical to the directional derivative of $\Psi(\cdot, r_k)$ at x^k along the direction d^k.

In what follows, we discuss some features and feasibility of Algorithm 2.1. It is not difficult to show that the following conclusion holds true.

Lemma 2.1. *Suppose that the assumption* $(A1)$ *holds and* W_k *is positive definite. Then the coefficient matrix* M_k *is nonsingular and the SLEs* (2.5), (2.6a) *and* (2.10a) *have a unique solution, respectively.*

The following results play a crucial role in analyzing the feasibility of Algorithm 2.1.

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

(i) $g_0(x^k)^{\mathrm{T}}\bar{d}^k - \sum_{i\in L} \bar{\gamma}_i^k f_i(x^k) = -(\bar{d}^k)^{\mathrm{T}} W_k(\bar{d}^k)$;

(ii) $g_0(x^k)^{\mathrm{T}} d^k \le \theta\psi_k + \sum_{i\in L}(1 + |2\bar{\lambda}_i^k - \lambda_i^k|)|f_i(x^k)|, \ \ \theta \in (0, 1)$;

(iii) $\psi_k \le 0$. *If* $\psi_k = 0$, *then* x^k *is a KKT point of the problem (P).*

Proof. (i) It follows from the SLE (2.5) that

$$H_k\bar{d}^k + \sum_{i\in I_k} \bar{\lambda}_i^k g_i(x^k) + \sum_{i\in L} \bar{\gamma}_i^k g_i(x^k) = -g_0(x^k), \tag{2.12a}$$

$$z_i^k g_i(x^k)^{\mathrm{T}}\bar{d}^k + \bar{\lambda}_i^k f_i(x^k) = 0, \ \ i \in I_k, \tag{2.12b}$$

$$g_i(x^k)^{\mathrm{T}}\bar{d}^k = -f_i(x^k), \ \ i \in L. \tag{2.12c}$$

Further, it follows from the above equalities that

$$g_0(x^k)^{\mathrm{T}}\bar{d}^k = -(\bar{d}^k)^{\mathrm{T}} H_k\bar{d}^k + \sum_{i\in I_k} \frac{z_i^k}{f_i(x^k)}(\bar{d}^k)^{\mathrm{T}} g_i(x^k) g_i(x^k)^{\mathrm{T}}\bar{d}^k + \sum_{i\in L} \bar{\gamma}_i^k f_i(x^k),$$

which together with the definition of W_k gives

$$g_0(x^k)^{\mathrm{T}}\bar{d}^k - \sum_{i\in L} \bar{\gamma}_i^k f_i(x^k) = -(\bar{d}^k)^{\mathrm{T}} W_k(\bar{d}^k).$$

That is, the result (i) holds.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

(ii) One has from the SLE (2.6a)

$$H_k d^k + \sum_{i\in I_k} \lambda_i^k g_i(x^k) + \sum_{i\in L} \gamma_i^k g_i(x^k) = -g_0(x^k), \tag{2.13}$$

$$z_i^k g_i(x^k)^{\mathrm{T}} d^k + \lambda_i^k f_i(x^k) = \mu_i^k, \quad i \in I_k; \quad g_i(x^k)^{\mathrm{T}} d^k = -f_i(x^k), \quad i \in L. \tag{2.14}$$

Thus it follows from (2.13) and (2.12a) that

$$\begin{aligned} & g_0(x^k)^{\mathrm{T}} d^k - g_0(x^k)^{\mathrm{T}} \bar{d}^k \\ = & \sum_{i\in I_k} \lambda_i^k g_i(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in I_k} \bar{\lambda}_i^k g_i(x^k) d^k + \sum_{i\in L} \gamma_i^k g_i(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in L} \bar{\gamma}_i^k g_i(x^k) d^k. \end{aligned}$$

Further, combining with (2.12b) and (2.14), it follows that

$$\begin{aligned} & g_0(x^k)^{\mathrm{T}} d^k \\ = \; & g_0(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in I_k} \frac{\bar{\lambda}_i^k}{z_i^k} \Big((1-\beta_k)\phi_i^k + \beta_k(-\|\bar{d}^k\|^\nu - |\psi_k|^\nu) z_i^k \Big) + \sum_{i\in L} (\bar{\gamma}_i^k - \gamma_i^k) f_i(x^k) \\ = \; & g_0(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in I_k} \frac{\bar{\lambda}_i^k}{z_i^k} \phi_i^k + \beta_k \sum_{i\in I_k} \frac{\bar{\lambda}_i^k}{z_i^k} \Big((\|\bar{d}^k\|^\nu + |\psi_k|^\nu) z_i^k + \phi_i^k \Big) + \sum_{i\in L} (\bar{\gamma}_i^k - \gamma_i^k) f_i(x^k) \\ \leq \; & \psi_k + \beta_k b_k + \sum_{i\in L} (1 + |2\bar{\gamma}_i^k - \gamma_i^k|) |f_i(x^k)|. \end{aligned} \tag{2.15}$$

On the other hand, it follows from the conclusion (i) and Step 7 that

$$g_0(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in L} \bar{\gamma}_i^k f_i(x^k) = -(\bar{d}^k)^{\mathrm{T}} W_k \bar{d}^k - \sum_{i\in L} f_i(x^k)^2 \leq -(\bar{d}^k)^{\mathrm{T}} W_k \bar{d}^k \leq 0,$$

and it follows from the definition (2.6d) of ϕ_i^k that $-\sum_{i\in I_k} \frac{\bar{\lambda}_i^k}{z_i^k} \phi_i^k \leq 0$. Therefore, from the definition (2.6c) of ψ_k, we can conclude that

$$\psi_k \leq 0. \tag{2.16}$$

Noting that $\beta_k = 1$ for $b_k \leq 0$ and $\theta \in (0,1)$, it follows from (2.15) that

$$g_0(x^k)^{\mathrm{T}} d^k \leq \theta \psi_k + \sum_{i\in L} (1 + |2\bar{\gamma}_i^k - \gamma_i^k|) |f_i(x^k)|.$$

Noting that $\beta_k = \min\{\frac{(1-\theta)|\psi_k|}{b_k}, 1\}$ for $b_k > 0$, it follows that $\beta_k b_k \leq (\theta - 1)\psi_k$, which together with (2.15) gives

$$g_0(x^k)^{\mathrm{T}} d^k \leq \theta \psi_k + \sum_{i\in L} (1 + |2\bar{\gamma}_i^k - \gamma_i^k|) |f_i(x^k)|.$$

Hence, the result (ii) holds true.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

(iii) The conclusion $\psi_k \leq 0$ is true from (2.16). If $\psi_k = 0$, then it implies from (2.16) and (2.6c) that

$$(\bar{d}^k)^{\mathrm{T}} W_k \bar{d}^k = 0, \quad \sum_{i \in I_k} \frac{\bar{\lambda}_i^k}{z_i^k} \phi_i^k = 0, \quad \sum_{i \in L} |f_i(x^k)| = 0. \tag{2.17}$$

So we obtain that $\bar{d}^k = 0$ from the positive definiteness of W_k and that $f_i(x^k) = 0$ $(i \in L)$. Further, together with (2.12b) and $\bar{d}^k = 0$, we have $\bar{\lambda}_i^k f_i(x^k) = 0$, $i \in I_k$. In view of the second equality of (2.17) and $\bar{\lambda}_i^k \phi_i^k \geq 0, i \in I_k$, we obtain

$$\bar{\lambda}_i^k \phi_i^k = 0, \ i \in I_k. \tag{2.18}$$

We can show that $\bar{\lambda}_i^k \geq 0$ for any $i \in I_k$. In fact, by contradiction, if there exists some $i_0 \in I_k$ such that $\bar{\lambda}_{i_0}^k < 0$, then $f_{i_0}(x^k) = 0$. Therefore, from (2.6d), we have $\phi_{i_0}^k < 0$. It implies that $\bar{\lambda}_{i_0}^k \phi_{i_0}^k > 0$. It is a contradiction to (2.18). Hence, from $\bar{d}^k = 0$, $\bar{\lambda}_{I_k}^k \geq 0$ and the SLE (2.5), we can conclude that x^k is a KKT point of (P). □

Based on the above discussion, we now verify the feasibility of Algorithm 2.1.

Lemma 2.3. *Suppose that the assumption* (A1) *holds. Then*

(i) *If x^k is not a KKT point of (P), then $\Gamma(x^k, d^k, r_k) < 0$;*
(ii) *Algorithm* 2.1 *is well defined.*

Proof. (i) From the definition (2.9) of $\Gamma(x^k, d^k, r_k)$ and Lemma 2.2(ii), we have

$$\begin{aligned}\Gamma(x^k, d^k, r_k) &\leq \theta\psi_k + \sum_{i \in L}(1 + |2\bar{\gamma}_i^k - \gamma_i^k|)|f_i(x^k)| - r_k \sum_{i \in L} |\ f_i(x)| \\ &\leq \theta\psi_k - \sum_{i \in L}(r_k - 1 - |2\bar{\gamma}_i^k - \gamma_i^k|)|f_i(x^k)|,\end{aligned}$$

which together with (2.7) and Lemma 2.2(iii) gives

$$\Gamma(x^k, d^k, r_k) \leq \theta\psi_k < 0. \tag{2.19}$$

(ii) In order to verify the result (ii), it is sufficient to show that the inequalities (2.11a) and (2.11b) hold for $t > 0$ sufficiently small.

Let's first consider the inequality (2.11a). From the definition (2.4) of Ψ, we have

$$\begin{aligned}&\Psi(x^k + td^k + t^2\widetilde{d}^k, r_k) - \Psi(x^k, r_k) \\ &= tg_0(x^k)^{\mathrm{T}} d^k + o(t) + r_k \sum_{i \in L}(|\ f_i(x^k) + tg_i(x^k)^{\mathrm{T}} d^k + o(\|td^k + t^2\widetilde{d}^k\|)\ | - |\ f_i(x^k)\ |).\end{aligned}$$

From the SLE (2.6a), we have $g_i(x^k)^{\mathrm{T}} d^k = -f_i(x^k)$, $\forall i \in L$. Substituting this formula into the above inequality, we obtain

$$\begin{aligned}&\Psi(x^k + td^k + t^2\widetilde{d}^k, r_k) - \Psi(x^k, r_k) \\ &= tg_0(x^k)^{\mathrm{T}} d^k + r_k \sum_{i \in L}(|\ (1-t)f_i(x^k)\ | - |\ f_i(x^k)\ |) + o(\|td^k + t^2\widetilde{d}^k\|).\end{aligned}$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Note that $1-t>0$ for t sufficiently small. It follows from the above inequality that

$$\begin{aligned}
&\Psi(x^k+td^k+t^2\widetilde{d}^k,r_k)-\Psi(x^k,r_k)\\
=\;& tg_0(x^k)^{\mathrm{T}}d^k-tr_k\sum_{i\in L}\mid f_i(x^k)\mid+o(\|td^k+t^2\widetilde{d}^k\|)\\
=\;& t\Gamma(x^k,d^k,r_k)+o(\|td^k+t^2\widetilde{d}^k\|)\\
\le\;& \alpha t\Gamma(x^k,d^k,r_k)+o(\|td^k+t^2\widetilde{d}^k\|),
\end{aligned}\tag{2.20}$$

the last inequality is duo to (2.19) and $\alpha\in(0,\frac{1}{2})$. Therefore, there exists a $\bar{t}_k>0$ such that (2.11a) holds for all $t\in(0,\bar{t}_k)$.

Next consider the inequality (2.11b). Since $f_i(x^k)<0$ for $i\in I$, the inequality (2.11b) follows for t sufficiently small due to the continuity of f_i.

By summarizing the above discussion, there exists a $\widehat{t}_k>0$ such that (2.11a) and (2.11b) hold for all $t\in(0,\widehat{t}_k)$. Consequently, Algorithm 2.1 is well defined. □

3. Global convergence

In this section, we will show that Algorithm 2.1 is globally convergent, i.e., any accumulation point of the generated sequence $\{x^k\}$ is a KKT point for (P). For this goal, we need the following assumptions.

(A2) (i) The sequence $\{x^k\}$ generated by Algorithm 2.1 is bounded.
(ii) For each accumulation point x^* of $\{x^k\}$, vectors $\{g_i(x^*),\ i\in I(x^*)\cup L\}$ are linearly independent.
(iii) There exist two positive constants a and b such that

$$\|H_k\|\le b,\quad d^{\mathrm{T}}W_kd\ge a\|d\|^2,\ \ \forall d\in R^n,\ \forall\ k.\tag{3.1}$$

In what follows, we first present several Lemmas which are used in the proof of global convergence.

Lemma 3.1. *Suppose that the assumptions* (A1) *and* (A2) *hold. Then*

(i) there exists a $\bar{c}>0$ such that $\|M_k^{-1}\|\le\bar{c}$;
(ii) the sequences $\{(\bar{d}^k,\bar{\lambda}^k,\bar{\gamma}^k)\}$, $\{(d^k,\lambda^k,\gamma^k)\}$ and $\{(\widetilde{d}^k,\widetilde{\lambda}^k,\widetilde{\gamma}^k)\}$ are all bounded.

The proof is similar to that of Lemma 3.1 in [8] and omitted here.

The following lemma can be shown from the updating formula (2.7).

Lemma 3.2. *Suppose that the assumptions* (A1) *and* (A2) *hold. Then there exists a positive integer* k_0 *such that* $r_k\equiv r_{k_0}\overset{\text{def}}{=}r$ *for all* $k\ge k_0$.

Based on Lemma 3.2, in the rest of this paper we suppose, without loss of generality, that $r_k\equiv r$.

By Lemma 3.1, (A2)(iii) and Lemma 2.2(iii), we can show that the following conclusion is true, which facilitates the proof of global convergence.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Lemma 3.3. *Suppose that the assumptions* (A1) *and* (A2) *hold. Let x^* be an accumulation point of the sequence $\{x^k\}$ generated by Algorithm* 2.1 *and suppose that subsequence $\{x^k\}_K \to x^*$. If $\{\psi_k\}_K \to 0$, then x^* is a KKT point of (P) and $\{(\bar{\lambda}^k, \bar{\gamma}^k)\}_K$ converges to the unique multiplier vector (λ^*, γ^*) corresponding to x^*.*

At the end of this section, based on the above results, we present a conclusion of global convergence of Algorithm 2.1.

Theorem 3.1. *Suppose that the assumptions* (A1) *and* (A2) *hold and x^* is an accumulation point of the sequence $\{x^k\}$ generated by Algorithm* 2.1. *Then x^* is a KKT point of (P), i.e., Algorithm* 2.1 *is globally convergent.*

Proof. Let sequence $\{x^k\}_K$ converge to x^*. Without loss of generality, we suppose that $\{\psi_k\}_K \to \psi^*$. It is clear from Lemma 2.2(iii) that $\psi^* \leq 0$. We suppose, by contradiction, that x^* is not a KKT point of (P). Then, by Lemma 3.3, one obtains $\psi^* < 0$. The remnant of the proof is divided into two steps as follows.

Step A. We show the inequalities (2.11a) and (2.11b) hold for $k \in K$ sufficiently large and $t > 0$ sufficiently small.

We first prove the inequality (2.11b), i.e., $f_i(x^k + td^k + t^2\widetilde{d}^k) < 0, i \in I$.

(i) For $i \in I \setminus I(x^*)$, i.e., $f_i(x^*) < 0$. The inequality (2.11b) follows from the continuity of f_i and the boundedness of $\{(d^k, \widetilde{d}^k)\}_K$.

(ii) For $i \in I(x^*)$. In view of Lemma 3.1(ii), without loss of generality, we can suppose that $\bar{\lambda}^{k-1} \to \bar{\lambda}^*$, $\bar{\gamma}^{k-1} \to \bar{\gamma}^*$, $k \in K$. Thus, $(x^*, \bar{\lambda}^*, \bar{\gamma}^*)$ is not a KKT pair of problem (P), further, we have $\varphi(x^*, \bar{\lambda}^*, \bar{\gamma}^*) > 0$. Since $I(x^*) \subseteq I_k$, from the Taylor expansion, (2.6a), (2.6b), Lemma 3.1(ii) and (2.6e), one gets for t sufficiently small

$$\begin{aligned} f_i(x^k + td^k + t^2\widetilde{d}^k) &= f_i(x^k) + t\frac{\mu_i^k - \lambda_i^k f_i(x^k)}{z_i^k} + o(t) \\ &= (1 - t\frac{\lambda_i^k}{z_i^k})f_i(x^k) + t\frac{1-\beta_k}{z_i^k}\phi_i^k - \beta_k t(\|\bar{d}^k\|^\nu + |\psi_k|^\nu) + o(t). \end{aligned} \tag{3.2}$$

The boundedness of $\{\lambda_i^k\}$ and $\{z_i^k\}$ implies $1 - t\frac{\lambda_i^k}{z_i^k} \geq 0$. It is clear that

$$\left(1 - t\frac{\lambda_i^k}{z_i^k}\right) f_i(x^k) \leq 0, \quad t\frac{1-\beta_k}{z_i^k}\phi_i^k \leq 0, \tag{3.3}$$

which shows the following inequality:

$$f_i(x^k + td^k + t^2\widetilde{d}^k) \leq -\beta_k t(\|\bar{d}^k\|^\nu + |\psi_k|^\nu) + o(t). \tag{3.4}$$

In view of (2.6e), (2.6f) and $\psi^* < 0$, there exists a scalar $\underline{\beta} > 0$ such that $\beta_k \geq \underline{\beta} > 0$. From Lemma 3.1 and Lemma 3.3, there exists a subsequence $K' \subseteq K$ and a scalar $\varpi > 0$ such that

$$\inf_{k \in K'}(\|\bar{d}^k\|^\nu + |\psi_k|^\nu) \geq \varpi > 0.$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

So, it follows, for k sufficiently large and $t > 0$ sufficiently small, that

$$f_i(x^k + td^k + t^2\widetilde{d}^k) \leq -\frac{1}{2}\underline{\beta}\varpi t + o(t) < 0. \tag{3.5}$$

Hence, the inequality (2.11b) holds true.

Next, we consider the inequality (2.11a). From (2.20) and the boundedness of $\{d^k, \widetilde{d}^k\}$, we have

$$\begin{aligned} & \Psi(x^k + td^k + t^2\widetilde{d}^k, r) - \Psi(x^k, r) - \alpha t\Gamma(x^k, d^k, r) \\ = \;& t\Gamma(x^k, d^k, r) - \alpha t\Gamma(x^k, d^k, r) + o(\|td^k + t^2\widetilde{d}^k\|) \\ \leq \;& (1-\alpha)t\theta\psi_k + o(t) \leq (1-\alpha)t\theta\widetilde{\alpha}\psi^* + o(t), \end{aligned} \tag{3.6}$$

where $\widetilde{\alpha} \in (0, 1)$. The above inequality implies that the inequality (2.11a) holds for $k \in K'$ sufficiently large and $t > 0$ sufficiently small.

By summarizing the above discussion, there exists a constant $\bar{t} > 0$ such that the step size $t_k \geq \bar{t}$ for all $k \in K'$.

Step B. Bring a contradiction based on $t_k \geq \bar{t} > 0$ $(k \in K')$. Noting that

$$\lim_{k \in K'} \Psi(x^k, r) = \Psi(x^*, r)$$

and the monotonic property of $\{\Psi(x^k, r)\}$, one obtains

$$\lim_{k \to \infty} \Psi(x^k, r) = \Psi(x^*, r).$$

In view of Step 6 in Algorithm 2.1 and (2.19), it follows for $k \in K'$ sufficiently large that

$$\Psi(x^{k+1}, r) - \Psi(x^k, r) \leq \alpha t_k \Gamma(x^k, d^k, r) \leq \alpha\theta\psi_k\bar{t}. \tag{3.7}$$

Passing to the limit $k \overset{K'}{\to} \infty$ in the inequality above, one gets $\alpha\theta\psi^*\bar{t} \geq 0$, which is a contradiction. Hence, x^* is a KKT point of (P). The whole proof is completed. □

4. Strong and superlinear convergence

In this section, we first analyze the strong convergence of Algorithm 2.1 and then analyze the acceptance of a full step of one for all k sufficiently large. Finally, we establish the superlinear convergence under some mild assumptions.

In the analysis of this section, the following assumptions are necessary.

- **(A3)** Functions f_0, f_i $(i \in I \cup L)$ are all twice continuously differentiable.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

- **(A4)** The sequence $\{x_k\}$ generated by Algorithm 2.1 possesses an accumulation point x^*, at which the second order sufficient conditions and the strict complementarity hold, i.e., the Lagrange multiplier vector $\lambda^* \in E^m$ at x^* satisfies $\lambda_i^* > 0$ $(\forall i \in I(x^*))$ and the Hessian matrix $\nabla^2_{xx} L(x^*, \lambda^*, \gamma^*)$ is positive definite on the subspace $\{y \in R^n \mid y \neq 0,\ g_i(x^*)^{\mathrm{T}} y = 0,\ \forall i \in I(x^*) \cup L\}$.

Lemma 4.1. *Suppose that the stated assumptions* (A1)-(A4) *hold. For any subset* K *of* $\{1, 2, \cdots\}$ *such that* $\lim\limits_{K} x^k = x^*$, *then*

$$\text{(i)}\ \lim_{K}(\bar{d}^k, \bar{\lambda}^k, \bar{\gamma}^k) = (0, \lambda^*, \gamma^*);$$
$$\text{(ii)}\ \lim_{K}(d^k, \lambda^k, \gamma^k) = (0, \lambda^*, \gamma^*).$$

Proof. (i) Since $\{(\bar{d}^k, \bar{\lambda}^k, \bar{\gamma}^k)\}$ and $\{Z_k\}$ are bounded, in view of (A4) and $I_k \subseteq I$ and the finiteness of I, without loss of generality, suppose that there exists an index set $\bar{I}$ such that $I_k \equiv \bar{I},\ \forall\, k \in K$, and there exists an infinite subset $K' \subseteq K$ such that

$$H_k \to H_*, \quad (\bar{\lambda}^k, \bar{\lambda}^{k-1}) \to (\bar{\lambda}, \bar{\lambda}^*), \quad (\bar{\gamma}^k, \bar{\gamma}^{k-1}) \to (\bar{\gamma}, \bar{\gamma}^*),$$
$$\bar{d}^k \to \bar{d}^*, \quad Z_{I_k} \to Z^*_{\bar{I}}, \quad G_{I_k} \to G_{\bar{I}} \overset{\text{def}}{=} \mathrm{diag}(f_{\bar{I}}(x^*)), \quad k \in K'.$$

If $(x^*, \bar{\lambda}^*, \bar{\gamma}^*)$ is a KKT pair of (P), then $\bar{\lambda}^* = \lambda^*$, $\bar{\gamma}^* = \gamma^*$, and it follows from Theorem 2.3 and Theorem 3.7 in [4] that $I_k \equiv I(x^*)$. If $(x^*, \bar{\lambda}^*, \bar{\gamma}^*)$ is not a KKT pair of (P), then $\varphi(x^*, \bar{\lambda}^*, \bar{\gamma}^*) > 0$. This implies that $I(x^*) \subseteq I_k,\ k \in K'$. Taking the limit as $k \to \infty,\ k \in K'$ in (2.5), we deduce that $(\bar{d}^*, \bar{\lambda}_{\bar{I}}, \bar{\gamma})$ solves the following SLE in (u, v_1, v_2)

$$\begin{pmatrix} H_* & A_{\bar{I}}(x^*) & B_{L}(x^*) \\ Z^*_{\bar{I}} A_{\bar{I}}(x^*)^{\mathrm{T}} & G_{\bar{I}}(x^*) & 0 \\ B_{L}(x^*)^{\mathrm{T}} & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -g_0(x^*) \\ 0 \\ -f_{L}(x^*) \end{pmatrix}. \tag{4.1}$$

Further, similar to the proof of Lemma 3.1 in [8], we can show that the coefficient matrix of (4.1) is nonsingular. So the solution of (4.1) is unique. On the other hand, noting that $(x^*, \lambda^*, \gamma^*)$ is a KKT pair of (P) from Theorem 3.4, we know that $(0, \lambda^*_{\bar{I}}, \gamma^*)$ is a solution of (4.1), too. So we conclude that $(\bar{d}^*, \bar{\lambda}_{\bar{I}}, \bar{\gamma}) = (0, \lambda^*_{\bar{I}}, \gamma^*)$ due to the uniqueness. Hence, $\lim_K (\bar{d}^k, \bar{\lambda}^k, \bar{\gamma}^k) = (0, \lambda^*, \gamma^*)$.

(ii) For any subset K such that $x^k \to x^*$ as $k \to \infty, k \in K$, in view of $\mu^k_{I_k} \to 0\ (k \in K)$, similar to the proof of the conclusion (i), we can get $\lim\limits_{K} (d^k, \lambda^k, \gamma^k) = (0, \lambda^*, \gamma^*)$. □

Next we present the strong convergence of Algorithm 2.1 and some associated conclusions.

Theorem 4.1. *Suppose that the stated assumptions* (A1)-(A4) *hold. Then the following conclusions hold true:*

- (i) *the entire sequence* $\{x^k\}$ *yielded by Algorithm* 2.1 *converges to* x^*, *i.e., Algorithm* 2.1 *is strongly convergent;*

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

- (ii) $(\bar{d}^k, \bar{\lambda}^k, \bar{\gamma}^k) \to (0, \lambda^*, \gamma^*)$ *as* $k \to \infty$;
- (iii) $\phi^k_{I_k} = 0,\ \mu^k_{I_k} = \beta_k(-\|\bar{d}^k\|^\nu - |\psi_k|^\nu) z^k_{I_k}$ *and* $I_k \equiv I(x^*)$ *for* k *sufficiently large;*
- (iv) $(d^k, \lambda^k, \gamma^k) \to (0, \lambda^*, \gamma^*)$ *as* $k \to \infty$;
- (v) $z^k \to \min\{\max\{z_{\min}\varpi, \lambda^*\}, z_{\max}\varpi\}$ *as* $k \to \infty$, *where* $\varpi = (1, 1, \cdots, 1)^{\mathrm{T}} \in R^m$.

The proof is similar to that of Theorem 4.1 in [8] and omitted here.

We now present some features of the solution of the SLE (2.10a), which will be used in the analysis of a full step of one in the sequel.

Lemma 4.2 *Suppose that the stated assumptions* (A1)-(A4) *hold. If* $\rho_2 \le \min\{\lambda^*_i, i \in I(x^*)\}$ *and* $z_{\max} \ge \max\{\lambda^*_i, i \in I(x^*)\}$, *then the solution of the SLE* (2.10a) *satisfies the following relations:*

$$\|\widetilde{d}^k\| = O\left(\max_{i\in I(x^*)}\{|\frac{z^k_i}{\lambda^k_i} - 1| \cdot \|d^k\|, \|d^k\|^2\}\right) = o(\|d^k\|); \tag{4.2}$$

$$\|\widetilde{\lambda}^k\| = O\left(\max_{i\in I(x^*)}\{|\frac{z^k_i}{\lambda^k_i} - 1| \cdot \|d^k\|, \|d^k\|^2\}\right); \tag{4.3}$$

$$\|\widetilde{d}^k\| \cdot \|d^k\| = o(\delta_k), \quad \|\widetilde{\lambda}^k\| \cdot \|d^k\| = o(\delta_k). \tag{4.4}$$

Proof. It follows from the SLEs (2.5) and (2.6a) that

$$z^k_i g_i(x^k)^{\mathrm{T}} \bar{d}^k = -\bar{\lambda}^k_i f_i(x^k), \quad i \in I_k, \tag{4.5a}$$

$$g_i(x^k)^{\mathrm{T}} d^k = -f_i(x^k), \quad i \in L. \tag{4.5b}$$

So it is clear that $\mid f_i(x^k) \mid = O(\|\bar{d}^k\|),\ i \in I_k;\ \mid f_i(x^k) \mid = O(\|d^k\|),\ i \in L$. In view of $\phi_{I_k} \equiv 0$ for k sufficient large, it follows from (2.6c) that

$$\psi_k = g_0(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in L} \bar{\gamma}^k_i f_i(x^k) - \sum_{i\in L} |f_i(x^k)|. \tag{4.6}$$

From Lemma 2.2(i) and (A1), we obtain

$$g_0(x^k)^{\mathrm{T}} \bar{d}^k - \sum_{i\in L} \bar{\gamma}^k_i f_i(x^k) = O(\|\bar{d}^k\|^\rho), \quad (\rho \in [1, 2]),$$

so (4.6) is equivalently rewritten as

$$\psi_k = O(\max\{\|d^k\|, \|\bar{d}^k\|^\rho\}). \tag{4.7}$$

Again, from the SLEs (2.6a) and (2.5), one obtains

$$M_k \begin{pmatrix} d^k - \bar{d}^k \\ \lambda_{I_k} - \bar{\lambda}_{I_k} \\ \gamma - \bar{\gamma} \end{pmatrix} = \begin{pmatrix} 0 \\ \mu^k_{I_k} \\ 0 \end{pmatrix}. \tag{4.8}$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Together with $\mu^k_{I_k} = \beta_k(-\|\bar{d}^k\|^\nu - |\psi_k|^\nu)z^k_{I_k}$, (4.7) and Lemma 3.1, the above equality implies that there exists a positive constant c such that

$$\begin{aligned}\|d^k - \bar{d}^k\| &\le c(\|\bar{d}^k\|^\nu + |\psi_k|^\nu)\\ &= c\|\bar{d}^k\|^\nu + O(\max\{\|d^k\|^\nu, \|\bar{d}^k\|^{\rho\nu}\}).\end{aligned} \tag{4.9}$$

Hence, one obtains $\|d^k\| \sim \|\bar{d}^k\|$.

In view of the strict complementarity, i.e., $\lambda^*_i > 0, i \in I(x^*) = I_k$, it follows from (2.10b) for any $i \in I(x^*)$, that

$$\widetilde{\mu}^k_i = -\delta_k - z^k_i\left[f_i(x^k) + g_i(x^k)^{\mathrm{T}}d^k\right] - \frac{(z^k_i)^2}{\lambda^k_i}\beta_k|\psi_k|^\nu + O(\|d^k\|^2). \tag{4.10}$$

Again, it follows from the SLE (2.6a) that

$$\begin{aligned}&z^k_i g_i(x^k)^{\mathrm{T}}d^k + f_i(x^k)\lambda^k_i\\ =\ &\mu^k_i = \beta_k(-\|\bar{d}^k\|^\nu - |\psi_k|^\nu)z^k_i = O(\|\bar{d}^k\|^2) - \beta_k|\psi_k|^\nu z^k_i,\end{aligned}$$

or equivalently,

$$f_i(x^k) = -\frac{z^k_i}{\lambda^k_i}g_i(x^k)^{\mathrm{T}}d^k - \frac{1}{\lambda^k_i}\beta_k|\psi_k|^\nu z^k_i + O(\|d^k\|^2).$$

Substituting this equality into (4.10), we obtain

$$\begin{aligned}\widetilde{\mu}^k_i =& -\delta_k - z^k_i(1 - \frac{z^k_i}{\lambda^k_i})g_i(x^k)^{\mathrm{T}}d^k + \frac{z^k_i}{\lambda^k_i}\beta_k|\psi_k|^\nu z^k_i - \frac{(z^k_i)^2}{\lambda^k_i}\beta_k|\psi_k|^\nu + O(\|d^k\|^2)\\ =& -\delta_k - z^k_i(1 - \frac{z^k_i}{\lambda^k_i})g_i(x^k)^{\mathrm{T}}d^k + O(\|d^k\|^2)\\ =& -\delta_k + O\left(\max\{|\frac{z^k_i}{\lambda^k_i} - 1| \cdot \|d^k\|, \|d^k\|^2, i \in I(x^*)\}\right) + O(\|d^k\|^2),\end{aligned}$$

which together with (2.10d) shows that

$$\|\widetilde{\mu}^k_{I_k}\| = O\left(\max\{|\frac{z^k_i}{\lambda^k_i} - 1| \cdot \|d^k\|, \|d^k\|^2, i \in I(x^*)\}\right) = o(\|d^k\|).$$

On the other hand, it follows from (2.10c) and the Taylor expansion that for any $i \in L$

$$\begin{aligned}\widetilde{\varsigma}^k_i =& -\|d^k\|^\tau - f_i(x^k) - g_i(x^k)^{\mathrm{T}}d^k + O(\|d^k\|^2)\\ =& -\|d^k\|^\tau + O(\|d^k\|^2) = O(\|d^k\|^2).\end{aligned}$$

The second equality and the third equality are due to (4.5a) and $\tau \in (2, 3)$, respectively. Thus, together with Lemma 3.1(i), the relations (4.2) and (4.3) are immediate. Further, the relation (4.4) follows from (2.10b) and (2.10d). The proof is completed. □

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

To ensure the step size $t_k \equiv 1$ for k sufficiently large, the following assumption on the Hessian approximation H_k is necessary.

(A5) Assume $\|P_k(\nabla^2_{xx}L(x^k,\lambda^k,\gamma^k) - H_k)d^k\| = o(\|d^k\|)$, where

$$P_k = E - (A_{_k}, B_{_k})[(A_{_k}, B_{_k})^{\mathrm{T}}(A_{_k}, B_{_k})]^{-1}(A_{_k}, B_{_k})^{\mathrm{T}}. \tag{4.11}$$

In what follows, we present an important result, i.e., a full step of one is accepted.

Theorem 4.2. *Suppose that the stated assumptions* (A1)-(A5) *hold and*

$$z_{\min} \le \min\{\lambda_i^*, i \in I(x^*)\}, \quad z_{\max} \ge \max\{\lambda_i^*, i \in I(x^*)\}.$$

Then, the arc search of Algorithm 2.1 *eventually accepts a full step of one, i.e.,* $t_k \equiv 1$ *for* k *sufficiently large.*

Proof. In view of the inequalities (2.11a) and (2.11b), it is sufficient to verify the following inequalities hold true.

$$\Psi(x^k + d^k + \widetilde{d}^k, r) \le \Psi(x^k, r) + \alpha\Gamma(x^k, d^k, r); \tag{4.12}$$

$$f_i(x^k + d^k + \widetilde{d}^k) < 0, \quad i \in I. \tag{4.13}$$

(1) We first prove the inequality (4.13) holds. For $i \in I\backslash I(x^*)$, i.e., $f_i(x^*) < 0$, using the continuity of f_i and $(x^k, d^k, \widetilde{d}^k) \to (x^*, 0, 0)$ as $k \to \infty$, the inequality (4.13) is immediate for k sufficiently large.

For $i \in I(x^*)$, i.e., $f_i(x^*) = 0$. We know from the strict complementarity that $\lambda_i^* > 0$. It follows from the SLE (2.10a) and (2.10b) that

$$z_i^k g_i(x^k)^{\mathrm{T}}\widetilde{d}^k + \widetilde{\lambda}_k^i f_i(x^k) = -\delta_k - z_i^k f_i(x^k + d^k) - \frac{(z_i^k)^2}{\lambda_i^k}\beta_k|\psi_k|^\nu,$$

which combining with $\lambda_i^k > 0, \beta_k > 0$ and $z_i^k > 0$ gives

$$f_i(x^k + d^k) + g_i(x^k)^{\mathrm{T}}\widetilde{d}^k \le -\frac{\delta_k}{z_i^k} - \frac{\widetilde{\lambda}_i^k}{z_i^k}f_i(x^k) = -\frac{\delta_k}{z_i^k} + o(\delta_k).$$

Therefore, by the Taylor expansion and the inequality above, one has

$$\begin{aligned} f_i(x^k + d^k + \widetilde{d}^k) &= f_i(x^k + d^k) + g_i(x^k)^{\mathrm{T}}\widetilde{d}^k + O(\|d^k\|\|\widetilde{d}^k\|) \\ &\le -\frac{\delta_k}{z_i^k} + o(\delta_k), \end{aligned}$$

which together with $z_i^k \ge z_{\min} > 0$ shows that the inequality (4.13) holds true .

(2) In what follows, we analyze the inequality (4.12). By the Taylor expansion and Lemma 4.2, we obtain

$$\begin{aligned} w_k &:= \Psi(x^k + d^k + \widetilde{d}^k, r) - \Psi(x^k, r) - \alpha\Gamma(x^k, d^k, r) \\ &= g_{_0}(x^k)^{\mathrm{T}}(d^k + \widetilde{d}^k) + \frac{1}{2}(d^k)^{\mathrm{T}}\nabla^2_{xx}f_{_0}(x^k)d^k + r\sum_{i\in L} \mid f_i(x^k) + g_i(x^k)^{\mathrm{T}}(d^k + \widetilde{d}^k) \\ &\quad + \frac{1}{2}(d^k)^{\mathrm{T}}\nabla^2_{xx}f_i(x^k)d^k \mid -r\sum_{i\in L} \mid f_i(x^k) \mid -\alpha\Gamma(x^k, d^k, r) + o(\|d^k\|^2). \end{aligned} \tag{4.14}$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

From (2.6a), (2.10a) and (2.10c), one has

$$g_i(x^k)^{\mathrm{T}}d^k = -f_i(x^k), \quad g_i(x^k)^{\mathrm{T}}\widetilde{d}^k = -\|d^k\|^\tau - f_i(x^k+d^k), \ \ i \in L,$$

further, by the Taylor expansion, one obtains

$$\begin{aligned} g_i(x^k)^{\mathrm{T}}\widetilde{d}^k \ = \ & -\|d^k\|^\tau - f_i(x^k) - g_i(x^k)^{\mathrm{T}}d^k \\ & -\frac{1}{2}(d^k)^{\mathrm{T}}\nabla^2_{xx}f_i(x^k)d^k + o(\|d^k\|^2), \ \ i \in L, \end{aligned}$$

equivalently,

$$f_i(x^k) + g_i(x^k)^{\mathrm{T}}(d^k+\widetilde{d}^k) + \frac{1}{2}(d^k)^{\mathrm{T}}\nabla^2_{xx}f_i(x^k)d^k = o(\|d^k\|^2), \quad i \in L. \tag{4.15}$$

On the other hand, it follows from (2.6a) and Theorem 4.1(iii) that

$$g_0(x^k) + H_k d^k + \sum_{i\in I(x^*)} \lambda_i^k g_i(x^k) + \sum_{i\in L} \gamma_i^k g_i(x^k) = 0,$$

which together with Lemma 4.2 gives

$$g_0(x^k)^{\mathrm{T}}d^k = -(d^k)^{\mathrm{T}}H_k d^k - \sum_{i\in I(x^*)} \lambda_i^k g_i(x^k)^{\mathrm{T}}d^k - \sum_{i\in L} \gamma_i^k g_i(x^k)^{\mathrm{T}}d^k, \tag{4.16a}$$

$$\begin{aligned} g_0(x^k)^{\mathrm{T}}(d^k+\widetilde{d}^k) = \ & -\sum_{i\in I(x^*)} \lambda_i^k g_i(x^k)^{\mathrm{T}}(d^k+\widetilde{d}^k) \\ & -\sum_{i\in L} \gamma_i^k g_i(x^k)^{\mathrm{T}}(d^k+\widetilde{d}^k) - (d^k)^{\mathrm{T}}H_k d^k + o(\|d^k\|^2). \end{aligned} \tag{4.16b}$$

Again, from the SLE (2.10a) and (2.10b), it follows that for $i \in I(x^*)$ and $\lambda_i^k \neq 0$

$$z_i^k g_i(x^k)^{\mathrm{T}}\widetilde{d}^k + \widetilde{\lambda}_i^k f_i(x^k) = -\delta_k - z_i^k f_i(x^k+d^k) - \frac{(z_i^k)^2}{\lambda_i^k}\beta_k|\psi_k|^v.$$

By Lemma 4.2, the above equality can be rewritten as

$$f_i(x^k+d^k) + g_i(x^k)^{\mathrm{T}}\widetilde{d}^k = -\frac{\widetilde{\lambda}_i^k}{z_i^k}f_i(x^k) - \frac{\delta_k}{z_i^k} - \frac{z_i^k}{\lambda_i^k}\beta_k|\psi_k|^\nu = o(\|d^k\|^2),$$

which together with the Taylor expansion of $f_i(x^k+d^k)$ shows that

$$\begin{aligned} & g_i(x^k)^{\mathrm{T}}(d^k+\widetilde{d}^k) \\ = \ & -f_i(x^k) - \frac{1}{2}(d^k)^{\mathrm{T}}\nabla^2_{xx}f_i(x^k)d^k + o(\|d^k\|^2), \ \ i \in I(x^*). \end{aligned} \tag{4.17}$$

Again, from the SLE (2.10a) and (2.10c), we easily obtain

$$g_i(x^k)^{\mathrm{T}}(d^k+\widetilde{d}^k) = -f_i(x^k) - \frac{1}{2}(d^k)^{\mathrm{T}}\nabla^2_{xx}f_i(x^k)d^k + o(\|d^k\|^2), \ i \in L. \tag{4.18}$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Thus, in view of (2.9), (4.15) and (4.17) $\sim$ (4.19), the equality (4.14) is rewritten as

$$w_k = \sum_{i\in I(x^*)} \lambda_i^k f_i(x^k) + \sum_{i\in L} \gamma_i^k f_i(x^k) + \frac{1}{2}(d^k)^{\mathrm{T}}(\nabla^2_{xx}L(x^k,\lambda^k,\gamma^k) - H_k)d^k$$
$$- \frac{1}{2}(d^k)^{\mathrm{T}}H_k d^k - \alpha g_0(x^k)^{\mathrm{T}}(d^k) + (\alpha-1)r\sum_{i\in L}|f_i(x^k)| + o(\|d^k\|^2).$$

It follows from the definition (4.11) of the projection matrix P_k that

$$d^k = P_k d^k + d_0^k,$$

where $d_0^k = (A_k, B_k)[(A_k, B_k)^{\mathrm{T}}(A_k, B_k)]^{-1}(A_k, B_k)^{\mathrm{T}}d^k$. Furthermore, it follows from the SLE (2.6a) and Theorem 4.1(iii) that

$$(A_k, B_k)^{\mathrm{T}}d^k = \begin{pmatrix} Z_k^{-1}\mu_{I_k}^k - Z_k^{-1}G_k\lambda_{I_k}^k \\ -f_L(x^k) \end{pmatrix},\ d_0^k = O(\|f_{I_k}(x^k)\|) + o(\|d^k\|^2).$$

Based on the above relations, it follows from (A5) and (4.16a) that

$$w_k = (\alpha - \frac{1}{2})(d^k)^{\mathrm{T}}W_k d^k + (\alpha - \frac{1}{2})\sum_{i\in I(x^*)}\frac{z_i^k}{f_i(x^k)})(d^k)^{\mathrm{T}}g_i(x^k)g_i(x^k)^{\mathrm{T}}d^k$$
$$+ \sum_{i\in I(x^*)}\lambda_i^k(1-\alpha\frac{\lambda_i^k}{z_i^k})f_i(x^k) + \sum_{i\in L}\gamma_i^k(1-\alpha)f_i(x^k)$$
$$+ (\alpha-1)r\sum_{i\in L}|f_i(x^k)| + o(\|f_{I_k}\|) + o(\|d^k\|^2),$$

further, it follows from A2(iii) and the SLE (2.6a) that

$$w_k \le \left(\alpha - \frac{1}{2}\right)a\|d^k\|^2 + \left(\alpha - \frac{1}{2}\right)\sum_{i\in I(x^*)}(\lambda_i^k)^2\frac{f_i(x^k)}{z_i^k}$$
$$+ \sum_{i\in I(x^*)}\lambda_i^k\left(1-\alpha\frac{\lambda_i^k}{z_i^k}\right)f_i(x^k) + \sum_{i\in L}\gamma_i^k(1-\alpha)f_i(x^k)$$
$$+(\alpha-1)r\sum_{i\in L}|f_i(x^k)| + o(\|f_{I_k}\|) + o(\|d^k\|^2). \tag{4.19}$$

Noting that $\alpha \in (0, \frac{1}{2})$ and $r \ge \eta_k > |\gamma_i^k|$, we obtain

$$\sum_{i\in L}\gamma_i^k(1-\alpha)f_i(x^k) + (\alpha-1)r\sum_{i\in L}|f_i(x^k)| < 0,$$

which together with (4.19) shows

$$w_k \le \left(\alpha - \frac{1}{2}\right)a\|d^k\|^2 + \sum_{i\in I(x^*)}\lambda_i^k\left(1 - \frac{1}{2}\frac{\lambda_i^k}{z_i^k}\right)f_i(x^k) + o(\|f_{I_k}\|) + o(\|d^k\|^2). \tag{4.20}$$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

It follows from Theorem 4.1(v) that

$$\lim_{k\to\infty} \lambda_i^k \left(1 - \frac{1}{2}\frac{\lambda_i^k}{z_i^k}\right) = \frac{1}{2}\lambda^* > 0, \quad \forall i \in I(x^*), \tag{4.21}$$

and noting that $f_i(x^k) < 0, \forall i \in I(x^*)$, so it follows from (4.20), (4.21) and $\alpha \in (0, \frac{1}{2})$ that for k sufficient large

$$w_k \leq (\alpha - \frac{1}{2})a\|d^k\|^2 + o(\|d^k\|^2) < 0,$$

which means that the inequality (4.12) holds for k sufficient large. The proof is finished. □

Based on Theorem 4.2, similar to Theorem 5.2 in [16], we can obtain the conclusion of the superlinear convergence of Algorithm 2.1.

Theorem 4.3 *Suppose that the stated assumptions* (A1)-(A5) *hold. Then*

$$\|x^{k+1} - x^*\| = o(\|x^k - x^*\|),$$

i.e., Algorithm 2.1 *is superlinearly convergent.*

5. Numerical experiments

In this section, we test Algorithm 2.1 on some problems which are chosen from [7]. Algorithm 2.1 was coded by MATLAB and ran on a computer with Windows 7 and 2.20 GHz CPU. The details about the implementation are described as follows.

(a) The Hessian estimate H_k is updated as follows:

$$H_k = h_k E_n + \nabla^2_{xx} L(x^k, \lambda^k, \gamma^k),$$

where

$$h_k = \begin{cases} 0, & \text{if } \varpi^k_{\min} > 10^{-5}; \\ -\varpi^k_{\min} + 10^{-5}, & \text{if } |\varpi^k_{\min}| \leq 10^{-5}; \\ 2|\varpi^k_{\min}|, & \text{otherwise,} \end{cases}$$

and $\varpi^k_{\min}$ is the smallest eigenvalue of the matrix

$$U_k \overset{\text{def}}{=} \nabla^2_{xx} L(x^k, \lambda^k, \gamma^k) - \sum_{i\in I_k} \frac{z_i^k}{f_i(x^k)} g_i(x^k) g_i(x^k)^{\mathrm{T}}.$$

It can be shown that the matrix W_k defined by Step 7 satisfies the assumption (A2)(iii).

(b) The parameters in Algorithm 2.1 are chosen as follows:
$\alpha = 0.45,\ \sigma = 0.01,\ \beta = 0.5,\ \theta = 0.99,\ \xi = 0.5,\ \tau = 2.5,\ \nu = 3,\ z_{\max} = 10^5,$

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Table 1: The numerical results.

Problem	n	(m,l)	Algorithm 2.1						Algorithm in [17]		
			Itr	Nf	Ng_i	$Aset$	$Time$	f_{final}	Itr	Nf	Ng
Svanberg-10	10	(30, 0)	15	47	489	6	0.21	1.5731517e+01	36	227	258
Svanberg-30	30	(90, 0)	26	98	2736	20	0.71	4.9142526e+01	101	777	864
Svanberg-50	50	(150, 0)	36	141	4976	43	1.38	8.2581912e+01	108	881	968
Svanberg-80	80	(240, 0)	48	239	12847	64	4.19	1.32749819e+02	190	1666	1835
Svanberg-100	100	(300, 0)	57	291	16948	73	6.43	1.66197171e+02	178	1628	1782
Svanberg-500	500	(1500, 0)	207	1603	474084	398	584.43	8.35186918e+02	402	4020	4407

$$\varepsilon_0 = 10^{-5},\ \ M = 1,\ p = 1,\ \rho_1 = 0.1,\ \rho_2 = 0.035.$$

(c) Termination criteria. Algorithm 2.1 stops if one of the following termination criteria is satisfied:

(i) $\|\Phi(x^k, \lambda(x^k))\| \le 10^{-5}$;

(ii) $\|\bar{d}^k\| \le 10^{-5}$ and $\max\{-\bar{\lambda}_i^k, i \in I\} \le 10^{-5}$.

In the numerical experiment we compared our algorithm, i.e., Algorithm 2.1, with Algorithm in [17] and Algorithm in [13], respectively. We use the initial points given in [7] if they are feasible. The numerical results are summarized in Table 1 and Table 2, in which the following notations are used:

n:	the number of variables;
m:	the number of inequality constraints;
l:	the number of equality constraints;
Itr:	the number of iterations;
N_f:	the number of function evaluations for f;
N_{g_i}:	the number of function evaluations for g_i;
$Aset$:	the number of indices in the final working set;
f_{final} :	the objective function value at the final iteration.

We can know from tables 1 and 2 that the cardinality of the final working set "Aset"is generally smaller than the number of constraints, which implies that subproblems of Algorithm 2.1 are generally smaller than that of the full dimensional methods. Our working set is very useful. Table 1 shows that the numbers of "Itr" and "Nf" in Algorithm 2.1 are much smaller than those in [17]. The numbers of "Ng_i" in Algorithm 2.1 are larger than those of "Ng" in [17] because we compute the number of g_i which is needed during the iterative procedure.

Remark 5.1. The numerical results of Algorithm TWBUL are obtained directly from [13].

The number of iterations for Algorithm 2.1 is less than that of Algorithm in [13] for half of the test problems. Especially, for problem HS66 and problem HS111, the numbers of iterations of Algorithm TWBUL exceed 1000, whereas the numbers of iterations of Algorithm 2.1 are only 10 and 298, respectively.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

Table 2: The numerical results.

Problem	n	(m,l)	Algorithm 2.1					Algorithm in [13]	
			Itr	N_f	N_{g_i}	$Aset$	f_{final}	Itr	f_{final}
HS2	2	(1, 0)	17	19	42	1	4.8321e+00	−−	− − − − −
HS3	2	(1, 0)	13	12	30	1	1.2702e−005	4	8.5023e−09
HS4	2	(1, 1)	5	5	14	1	2.6667e+00	4	2.6667e+00
HS8	2	(0, 2)	4	3	11	0	−1.0000e+00	14	−1.0000e+00
HS12	2	(1, 0)	6	6	18	1	−3.0000e+01	5	−3.0000e+01
HS24	2	(5, 0)	11	13	101	2	−1.0000e+00	14	−1.0000e+00
HS29	3	(1, 0)	12	13	33	1	−2.2627e+01	8	−2.2627e+01
HS33	3	(6, 0)	22	22	288	3	−4.5858e+00	29	−4.5867e+00
HS34	3	(8, 0)	23	25	339	4	−8.3403e−01	30	−8.3403e−01
HS35	3	(4, 0)	11	16	60	1	1.1111e−01	8	1.1111e−01
HS36	3	(7, 0)	10	42	140	3	−3.3000e+03	10	−3.3000e+03
HS37	3	(1, 7)	6	24	98	1	−3.4560e+03	7	−3.3860e+03
HS38	4	(8, 0)	17	89	198	0	4.0056e−14	37	3.1564e−24
HS43	4	(3, 0)	13	14	84	2	−4.4000e+01	9	−4.4000e+01
HS48	5	(0, 2)	8	9	19	0	5.2645e−016	6	0.0000e+00
HS49	5	(0, 2)	20	21	43	0	3.1927e−010	69	3.1836e−010
HS50	5	(0, 3)	7	5	5	0	4.8891e−020	11	4.0725e−017
HS66	3	(8, 0)	10	10	139	3	5.1816e−01	1000+	5.1816e−01
HS93	6	(8, 0)	19	21	427	7	1.3508e+02	12	1.3508e+02
HS100	7	(4, 0)	15	13	113	2	6.8063e+02	9	6.8063e+02
HS111	10	(20, 3)	298	301	5943	0	-4.77202e+01	1000+	-4.7720e+01
HS113	10	(8, 0)	18	16	249	4	2.4306e+01	10	2.4306e+01
HS117	15	(20, 0)	22	27	653	20	3.2349e+01	25	3.23487e+01

6. Concluding remarks

With the help of a penalty function and a new identification technique of active constraints, we proposed a new primal-dual interior point method for general constrained optimization. Because of the introduction of working set, the size of SLEs is smaller than that of SLEs in some existing algorithms. At each iteration, the master search direction and higher-correction direction are generated by only solving two or three reduced SLEs with the same coefficient matrix and the penalty parameter is automatically updated. The uniformly positive definiteness assumption on the Hessian estimate is weaken, which is used in most of existing algorithms. The global and superlinear convergence are shown under some mild conditions. The preliminary numerical results indicates the feasibility and efficiency of the proposed algorithm for small and medium size test problems.

Acknowledgments The authors thank the two anonymous referees and the Associate Editor, Xiaoqi Yang, for their insightful comments and suggestions on the original version. This work is supported by the National Natural Science Foundation No.11271086, the Natural Science Foundation of Guangxi Nos. 2012GXNSFAA053007

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press

and 2014GXNSFFA118001 and Science Foundation of Guangxi Education Department No. 201102ZD002 of China.

References

[1] P. T. Boggs and J. W. Tolle, *Sequential quadratic programming* , in: Acta No.4,(1995), pp. 1–51. Cambridge University Press, Cambridge, UK.
[2] S. Bakhtiari and A. L. Tits, *A simple primal-dual feasible interior-point method for nonlinear programming with monotone descent*, Comput. Optim. Appl., 25 (2003), pp. 17-38.
[3] A. Forsgren and P. E. Gill , *Primal-dual interior methods for nonconvex nonlinear programming*, SIAM J. Optim., 8 (1998), pp. 1132-1152.
[4] F. Facchinei, A. Fischer and C. Kanzow, *On the accurate identification of active constraints,* SIAM J. Optim., 9 (1998), pp. 14-32.
[5] P. E. Gill, W. Murray and M. A. Saunders, *SNOPT: An SQP algorithm for large-scale constrained optimization.* SIAM Review, 47 (2005), pp. 99-131.
[6] Z.Y. Gao, G.P. He and F. Wu, *Sequential systems of linear equations algorithm for nonlinear opti mization problems with general constraints,* J. Optim. Theory Appl., 95 (1997), pp. 371-397.
[7] W. Hock and K. Schittkowski, *Test examples for nonlinear programming codes. in: Lecture notes in Economics and Mathematical Systems,* vol.187 (1981), Springer-Verlag, Berlin Heidelberg, New York.
[8] J. B. Jian, R. Quan and W. X. Cheng, *A feasible QP-free algorithm combining the interior point method with active set for constrained optimization,* Comput. Math. Appl., 58 (2009), pp. 1520-1533.
[9] D. Q. Mayne and E. Polak, *Feasible direction algorithms for optimization problems with equality and inequality constraints,* Math. Program., 11 (1976), pp. 67-80.
[10] 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 (1988), pp. 788–811.
[11] J. Peng , C. Roos and T. Terlaky, *New complexity analysis of the primal-dual Newton method for linear optimization,* Ann. Oper. Res., 99 (2000), pp. 22–39.
[12] H.D. Qi, and L.Q. Qi, *A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization,* SIAM J. Optim., 11 (2000), pp. 113-132.
[13] A. L. Tits, A. Wachter, S. Bakhtiari, T.J. Urban and C. T. Lawrence, *A primal-dual interior-point method for nonlinear programming with strong global and local convergence properties,* SIAM J. Optim., 14 (2003), pp. 173-199.
[14] Y.L. Wang, L.F. Chen and G.P. He, *Sequential systems of linear equations method for general constrained optimization without strict complementarity,* J. Comput. Appl. Math., 182 (2005), pp. 447-471.
[15] H. Yamashita, H. Yabe and T. Tanabe, *A globally and superlinearly convergent primal dual interior point trust region method for large scale constrained optimization,* Math. Program., Ser.A, 102 (2005), pp. 111-151.
[16] F. Facchinei, S. Lucidi, *Quadratically and superlinearly convergent for the solution of inequality constrained optimization problem*, J. Optim. Theory Appl., 85 (1995), pp. 265-289.
[17] Y. F. Yang, D. H. Li and L. Q. Qi. *A feasible sequential linear equation method for inequality constrained optimization*, SIAM J. Optim., 13 (2003), pp. 1222-1244.

https://doi.org/10.4208/nmtma.2015.m1338 Published online by Cambridge University Press