Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT

ELSEVIER

Journal of Computational and Applied Mathematics 151 (2003) 171–199

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Global convergence enhancement of classical linesearch interior point methods for MCPs☆

Stefania Bellavia, Benedetta Morini*

Dipartimento di Energetica "S. Stecco", University of Florence, via C. Lombroso 6/17, 50134 Florence, Italy

Received 23 November 2001; received in revised form 8 July 2002

Abstract

Recent works have shown that a wide class of globally convergent interior point methods may manifest a weakness of convergence. Failures can be ascribed to the procedure of linesearch along the Newton step. In this paper, we introduce a globally convergent interior point method which performs backtracking along a piecewise linear path. Theoretical and computational results show the effectiveness of our proposal.

MSC: 65K05; 65H10; 65H20; 90C33

Keywords: Complementarity problems; Interior point methods; Backtracking strategy

1. Introduction

The problem we address in this paper is to find a solution of a mixed complementarity problem (MCP), i.e., we seek for a vector $x=(v,s,z)\in\mathbb{R}^{m+2n}$ with $s,z\in\mathbb{R}^n_+$, that satisfies

$$H(x)=\begin{pmatrix} F(v,s,z) \\ SZe \end{pmatrix}=0, \tag{1.1}$$

where $F:\mathbb{R}^{m+2n}\mapsto\mathbb{R}^{m+n}$, $S=\mathrm{diag}(s)$, $Z=\mathrm{diag}(z)$, $e=(1,1,\ldots,1)^{\mathrm{T}}\in\mathbb{R}^n$.

MCPs are constrained nonlinear systems of equations which arise frequently in practice. In fact, many economics and engineering applications can be modeled by MCPs, see e.g., [12,13]. Further,

☆ Work supported by MIUR, Rome, Italy, through "Cofinanziamenti Programmi di Ricerca Scientifica di Interesse Nazionale".
* Corresponding author.
E-mail addresses: bellavia@ciro.de.unifi.it (S. Bellavia), morini@ciro.de.unifi.it (B. Morini).

PII: S0377-0427(02)00745-8

problem (1.1) is a generalization of nonlinear complementarity problems (NCPs) and the Karush–Khun–Tucker (KKT) equations for nonlinear programming and variational inequalities are particular cases of (1.1). Note that when $n=0$ an MCP reduces to a nonlinear system of equations.

MCPs problems can be effectively solved by interior point methods and here we consider the problem of enhancing the convergence properties of a class of widely used infeasible interior point methods. Infeasible interior point methods for MCPs start from a point (v_0, s_0, z_0) such that $s_0, z_0 \in \mathbb{R}_{++}$ and generate sequences of points which remain in the region $\mathbb{R}^m \times \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$.

In fact, the early study of interior point was motivated from the desire to find algorithms for linear programming problems with better theoretical properties than the simplex method. A great deal of work on this topic has given rise to procedures which display great efficiency in solving linear programming problems. Also, algorithms and software for linear programming have become quite sophisticated [32].

Extensions to more general classes of problems such as complementarity problems and nonlinear programming problems have been studied too, and now we give a brief account of some major works. For a detailed review of interior point algorithms we refer the reader to the recent survey [21].

Considerable research effort has been devoted to the study of infeasible interior point methods for monotone linear complementarity problems; in [35] Zhang described an algorithm with polynomial complexity, in later works Potra [19], Potra and Sheng [23], Wright [30,31] described fast convergent interior Point methods. In particular, the procedure given in [23] shows the best complexity bound. Also, the problem of solving nonmonotone linear complementarity problems by infeasible methods was addressed, e.g., see [20,25].

Infeasible interior point methods have been extended also to NCPs and nonlinear programming problems (NLPs). In the field of this latter class of problems, an important contribution to methods for convex NLPs is due to El Bakry et al. [11], while interior point methods for nonconvex NLPs were studied in [4,15,27,34]. The problem of solving nonlinear complementarity problems was addressed in the early paper [22] where several results for linear and nonlinear optimization are generalized to NCPs. Moreover, infeasible interior point methods for NCPs and MCPs were proposed by Kojima et al. [16], Wright and Ralph [33], Tseng [26].

Infeasible interior point methods for MCPs and NLPs are based upon a common scheme and differ for the choice of the merit function, the strategy of updating the barrier parameter and the globalization strategy used (linesearch [1,9,11,15,16,22,26,27,33] or trust region [4,34]).

In this paper, we focus on convergence enhancement of classical linesearch interior point methods. Given an initial guess $x_0 = (v_0, s_0, z_0)$, with s_0, z_0 positive component-wise vectors, at the kth iteration these methods compute the Newton step p_k^N by solving the linear system

$$H'(x_k)p_k^N = -H(x_k) + \mu_k e_0, \tag{1.2}$$

where H' is the Jacobian of H, μ_k is a positive scalar and $e_0 = (0, \ldots, 0, e^T)^T \in \mathbb{R}^{m+2n}$. Then, they apply a backtracking scheme along p_k^N in order to maintain s_{k+1} and z_{k+1} positive and to decrease a suitable merit function ψ.

Classical linesearch interior point methods are relevant because of their simplicity and their efficiency. However, several examples illustrate that they can fail to converge, see [5,17,24,28]. These papers show that the generated sequence $\{x_k\}$ can be attracted to a point $\tilde{x}$ that is neither a solution of problem (1.1) nor a stationary point for ψ. The point $\tilde{x}$ is called a singular nonstationary point with respect to the merit function ψ and it is such that $H(\tilde{x}) \neq 0$, $H'(\tilde{x})$ is singular and $\nabla\psi(\tilde{x}) \neq 0$, [5].

Convergence failures can occur with or without the involvement of the bounds. Concerning NCPs and the KKT conditions for nonlinear programming problems several examples of failures are provided in [5,24,28]. In some cases the bounds play a key role in blocking progress towards feasibility. On the other hand, similar phenomenon can take place regardless the bounds. This latter occurrence is closely related to failures that can be observed in the solution of nonlinear systems of equations by linesearch Newton methods. In this context, Powell [5,17] provided an example where a linesearch Newton method sticks a singular nonstationary point.

In [5,17] it was pointed out that the motivation for such convergence failures lies on an intrinsic flaw of the Newton direction. Therefore, in such situations, the Newton direction must be dropped and modifications of the basic linesearch approach must be devised. Regarding the proposal of procedures designed to overcome difficulties of classical linesearch interior point methods, we are only aware of the papers [3,17,2].

In [3] Benson et al. derived and discussed three possible solutions to such failures. One of these can be applied when the current iterate is near a bound. In particular, if a slack variable is small and at the same time the temptative increment is much greater than it, the variable is shifted. The current version of LOQO [3] emploies this technique and does not fail in solving the hard test problem proposed by Wachter and Biegler in [28].

In [17] Marazzi and Nocedal discussed interior point methods that generate steps employing trust region techniques. This way steps different from the Newton one can be selected. They proposed this approach as a resolution of the convergence difficulties of the classical linesearch strategy.

In [2] the authors proposed a method for MCPs which turns out to be a modification of the interior point method studied in [1,25]. The resulting method was denoted Piecewise Linear Interior Point (PLIP) method and the used globalization technique was designed to leave the Newton direction when it reveals to be unsatisfactory, i.e., when too many backtracks are required to maintain positive the bounded variables or to decrease the value of ψ. The key feature of [2] is the definition of the new piecewise linear path exploited by the backtracking strategy. The new path has a simple and inexpensive formulation. Moreover, the theoretical analysis of its properties yields to a strategy that allows for an automatic transition from the Newton direction to alternative directions. The conducted numerical experiments on hard test problems given in [2] showed that the PLIP method is a promising procedure for solving MCPs. However, no attempt was made to study it from a theoretical point of view.

In this paper, the convergence properties of the PLIP method are investigated. The global convergence properties of the method are studied taking into account the features of the piecewise linear path. First, assuming the invertibility of H' in a neighborhood of the bounded sequence $\{x_k\}$ it is proved that the PLIP method converges to a solution of the MCP problem and fast local convergence can be retained. Further, the problem of whether the PLIP method can be attracted to a singular nonstationary point is investigated. It is shown that under suitable hypotheses the backtracking procedure along the piecewise linear path prevents the iterates from sticking singular nonstationary points that belong to the interior $\mathbb{R}^m \times \mathbb{R}^n_+ \times \mathbb{R}^n_+$. It is important to point out that the hypotheses used to prove this result can actually be satisfied by problems that classical methods fail to solve. Finally, we report numerical results obtained with the PLIP method. The used set of test problems is constituted by hard problems considered in [5,24,28]. The obtained results indicate that our global strategy copes with the flaw of Newton step direction, shows fast local rate of convergence and low computational cost.

1.1. Notations

Through the paper, we will use (v,s,z) as shorthand for the vector $(v^T,s^T,z^T)^T$, and the vectors x and Δx for $x=(v,s,z)$, $\Delta x=(\Delta v,\Delta s,\Delta z)$.

For any generic vector y the subscript i will be used to indicate the ith component y_i of y. The corresponding capital letter Y denotes the diagonal matrix whose (i,i)th entry is given by y_i. In addition, $y>0$ $(y\geqslant 0)$ means that all the components of y are positive (nonnegative). Moreover, $\min(y)$ stands for $\min_i(y_i)$.

For any vector, $\|\cdot\|$ is the standard Euclidean norm and $\|\cdot\|_1$ is the 1-norm. Further, $N_\varepsilon(u)$ denotes the closed ball $N_\varepsilon(u)=\{y\in\mathbb{R}^n \| y-u\|\leqslant\varepsilon\}$.

If $H(y)$, $y\in\mathbb{R}^m$, is a given smooth vector function, the Jacobian matrix is denoted by $H'(y)$, while the gradient vector of a given smooth real function $h(y)$, $y\in\mathbb{R}^m$, is denoted by $\nabla h(y)$ and the Hessian matrix is denoted by $\nabla^2 h(y)$. For $F=F(v,s,z)$, $F'_v\in\mathbb{R}^{(m+n)\times m}$, $F'_s\in\mathbb{R}^{(m+n)\times n}$, $F'_z\in\mathbb{R}^{(m+n)\times n}$, are the jacobian matrices of F when F is considered as a function of v,s,z, respectively. When clear from the context, the argument of a mapping is omitted and, for any function H, the notation H_k is used to denote $H(x_k)$.

Finally, we recall that, if $\psi:\mathbb{R}^n\to\mathbb{R}$ is differentiable, then a descent direction p for ψ at x_k satisfies $\nabla\psi_k^T p<0$ [18] and this means that there exists a $\lambda_0>0$ such that $\psi(x_k+\lambda p)<\psi(x_k)$ for all $\lambda<\lambda_0$.

2. Backgrounds

In this section, first we briefly summarize the main features of classical linesearch interior point methods for MCPs. Then, we focus on a specific method of this class that we denote CLIP method. This method was studied in [1] and [25] and it is based on the classical interior point framework given in [11].

At each iteration of an interior point method, the numerical solution of the linear system (1.2) is required. Due to the structure of H'_k

$$H'_k=\begin{pmatrix} F'_{v,k} & F'_{s,k} & F'_{z,k}\\ 0 & Z_k & S_k\end{pmatrix}, \tag{2.1}$$

it is easy to characterize the affine space $P_{a,k}$ which contains the vectors p satisfying the last block of n equations of (1.2). In fact, any vector $p=(p_v,p_s,p_z)$ with $p_v\in\mathbb{R}^m$, p_s, $p_z\in\mathbb{R}^n$ and such that it satisfies exactly the last block of n equations of (1.2), has the form

$$p_z=-D_k p_s+\tilde q_k,$$

where

$$D_k=S_k^{-1}Z_k,\quad \tilde q_k=-Z_k e+\mu_k S_k^{-1}e. \tag{2.2}$$

Thus, letting

$$W_k = \begin{pmatrix} I_m & 0 \\ 0 & I_n \\ 0 & -D_k \end{pmatrix} \in \mathbb{R}^{(m+2n)\times(m+n)}, \qquad q_k = \begin{pmatrix} 0 \\ 0 \\ \tilde{q}_k \end{pmatrix}, \tag{2.3}$$

$P_{a,k}$ is defined by

$$P_{a,k} = \{ p \in \mathbb{R}^{m+2n} | p = W_k y + q_k \text{ for some } y \in \mathbb{R}^{m+n} \}.$$

Clearly, the Newton step p_k^N belongs to $P_{a,k}$ as it solves the linear system (1.2). Moreover, the form (2.1) of H_k' can be exploited to compute p_k^N. Specifically, letting $p_k^N = (p_{k,v}^N,\ p_{k,s}^N,\ p_{k,z}^N)$, where $p_{k,v}^N \in \mathbb{R}^m$ and $p_{k,s}^N, p_{k,z}^N \in \mathbb{R}^n$, $p_{k,z}^N$ is given by

$$p_{k,z}^N = -D_k\ p_{k,s}^N + \tilde{q}_k \tag{2.4}$$

with D_k and $\tilde{q}_k$ defined in (2.2). Consequently, the vector $(p_{k,v}^N, p_{k,s}^N)$ is the solution of the reduced linear system

$$J_k \begin{pmatrix} p_{k,v}^N \\ p_{k,s}^N \end{pmatrix} = -F_k - F_{z,k}' \tilde{q}_k, \tag{2.5}$$

where the matrix $J(x) \in \mathbb{R}^{(m+n)\times(m+n)}$ has the form

$$J(x) = (\, F_v'(x) \quad F_s'(x) - F_z'(x) D(x)\,). \tag{2.6}$$

Then, if the merit function ψ used is such that p_k^N is a descent direction for ψ at x_k, the classical linesearch technique considers trial iterates of the form

$$x_{k+1} = x_k + \Delta x_k, \tag{2.7}$$

where $\Delta x_k = \lambda p_k^N$ and $\lambda \in (0, 1]$. A vector $x_{k+1} = (v_{k+1}, s_{k+1}, z_{k+1}) \in \mathbb{R}^{m+2n}$ will be accepted as the new iterate if it sufficiently decreases the function ψ and satisfies $s_{k+1} > 0, z_{k+1} > 0$.

Now, for sake of completeness let us briefly describe the CLIP method. The used merit function is

$$\psi(x) = \|H(x)\|^2 \tag{2.8}$$

and in (2.8) μ_k is

$$\mu_k = \sigma_k s_k^{\mathrm{T}} z_k / n, \quad \sigma_k \in (0, 1). \tag{2.9}$$

Consequently, the vector p_k^N is a descent direction for ψ at x_k as

$$-\nabla\psi_k^{\mathrm{T}} p_k^N = -2H_k^{\mathrm{T}} H_k' p_k^N = 2\left(H_k^{\mathrm{T}} H_k - \sigma_k \frac{s_k^{\mathrm{T}} z_k}{n} H_k^{\mathrm{T}} e_0 \right) \geqslant 2(1-\sigma_k)\|H_k\|^2. \tag{2.10}$$

The trial iterate of the CLIP method has form (2.7) with $\Delta x_k = \lambda_k p_k^N$ and is *acceptable* if the following conditions hold when $\lambda = \lambda_k$:

$$\psi(x_{k+1}) \leqslant \psi(x_k) + \alpha \nabla\psi(x_k)^{\mathrm{T}} \Delta x_k, \tag{2.11}$$

$$f^1(\lambda) := \min(S_{k,\lambda}(z_k + \Delta z_k)) - \tau_1\, \gamma_k (s_k + \Delta s_k)^{\mathrm{T}} (z_k + \Delta z_k)/n \geqslant 0, \tag{2.12}$$

$$f^2(\lambda) : = (s_k + \Delta s_k)^{\mathrm{T}} (z_k + \Delta z_k) - \tau_2 \gamma_k \|F_{k+1}\| \geqslant 0. \tag{2.13}$$

Here $S_{k,\lambda} = \mathrm{diag}(s_k + \Delta s_k)$, $\alpha \in (0, 1/2)$, $\gamma_k \in [\hat{\gamma}, 1)$, $\hat{\gamma} > 0$ and

$$\tau_1 = \frac{\min(S_0 z_0)}{s_0^{\mathrm{T}} z_0/n}, \quad \tau_2 = \frac{s_0^{\mathrm{T}} z_0}{\|F_0\|}.$$

Condition (2.11) is the classical Armijo condition [18]. It enforces a sufficient decrease on $\|H\|$ and makes the progress to a solution of the MCP likely.

Conditions (2.12) and (2.13) are two widely used centering conditions. The first centering condition ensures that every iterate stays in the set defined by

$$\{s, z \in \mathbb{R}^n, s > 0, z > 0 \,|\, \min(SZe) \geqslant \tau_1 \hat{\gamma} s^{\mathrm{T}} z/n\}.$$

This way, whenever the sequences $\{s_k\}$ and $\{z_k\}$ are bounded, the vectors s_k and z_k cannot approach the boundary of the positive orthant of $\mathbb{R}^{2n}$ prematurely. The second centering condition (2.13) prevents improvement in the complementarity gap $s_{k+1}^{\mathrm{T}} z_{k+1}/n$ from outpacing improvement in the infeasibility measured by $\|F_{k+1}\|$.

The steplength λ_k is usually computed applying a backtracking strategy. At this regard, it should be noted that it is easy to compute the scalar $\lambda_{k,1}$ such that $\Delta_k = \lambda p_k^N$ satisfies condition (2.12) for all $\lambda \in (0, \lambda_{k,1}]$. In fact, this can be done by solving n scalar quadratic equations since $f^1(\lambda)$ is a componentwise quadratic function. Therefore, in order to avoid backtracking along p_k^N until the point $\lambda_{k,1} p_k^N$ is met, the accepted step has usually the form $\Delta x_k = \lambda_k p_k^N$ where $\lambda_k = \chi^i \lambda_{k,1}$, $\chi \in (0, 1)$ and i is the smallest integer such that (2.11) and (2.13) are satisfied.

We end this section considering the occurrence where $\{x_k\}$ is attracted to a singular nonstationary point $\tilde{x}$. Since $\tilde{x}$ is a point of singularity of H', when x_k approaches $\tilde{x}$ the norm of p_k^N becomes very large and the selected steplength λ_k eventually becomes tiny. Thus, the algorithm becomes stuck. More precisely two different types of failure can be detected:

(i) The sequence $\{x_k\}$ approaches the boundary prematurely and sticks the bounds. Thus, the steplengths $\lambda_{k,1}$ and λ_k tend to zero to maintain feasibility.
(ii) The direction p_k^N becomes increasingly perpendicular to $\nabla \psi_k$ and very small scalars λ_k must be taken to impose the Armijo condition while the sequence $\{\lambda_{k,1}\}$ remains bounded away from zero.

3. The PLIP method

The interior point method we study in this paper is a simple modification of the CLIP method. It adopts many features of the CLIP method: the merit function (2.8), the barrier parameter (2.9), the acceptance conditions (2.11)–(2.13) while it differs in the employed globalization strategy. In particular, at the kth iteration the PLIP method attempts to overcome the convergence failures of the classical linesearch interior point methods by applying a backtracking strategy along the piecewise path $\zeta_k(\lambda)$ proposed in [2].

Since the design of the path $\zeta_k(\lambda)$ and its properties provide the basis of the PLIP method, we begin summarizing results obtained in [2, Section 2].

In order to construct $\zeta_k(\lambda)$, first we searched for a descent direction for ψ at x_k that can be easily computed and belongs to the affine space $P_{a,k}$. This way the last block of n equations in (1.2) is satisfied and the direction points towards the central path.

The reference direction d_k we introduced is defined as follows:

$$d_k = \underset{\substack{p\in P_{a,k},\\ \|p\|^2 \leqslant \|p_k^N\|^2}}{\operatorname{argmin}} \nabla\psi_k^{\mathrm{T}} p. \tag{3.1}$$

In [2, Theorem 2.1], it is shown that if $W_k^{\mathrm{T}}\nabla\psi_k \neq 0$, the solution of problem (3.1) has the form

$$d_k = -\hat{W}_k(q_k + \alpha_k\nabla\psi_k) + q_k, \tag{3.2}$$

where

$$\hat{W}_k = W_k(W_k^{\mathrm{T}}W_k)^{-1}W_k^{\mathrm{T}} \quad \text{and} \quad \alpha_k = \sqrt{\frac{q_k^{\mathrm{T}}\hat{W}_k q_k - \|q_k\|^2 + \|p_k^N\|^2}{\nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k}} \tag{3.3}$$

and it is such that $\|d_k\| = \|p_k^N\|$. Therefore, d_k is the steepest descent direction for ψ at x_k among the vectors that are restricted to belong to $P_{a,k}$ and to have 2-norm equal to $\|p_k^N\|$.

Note that d_k can be computed at a low computational cost since forming the symmetric and semidefinite positive matrix $\hat{W}_k$

$$\hat{W}_k = \begin{pmatrix} I_m & 0 & 0 \\ 0 & (I_n + D_k^2)^{-1} & -D_k(I_n + D_k^2)^{-1} \\ 0 & -D_k(I_n + D_k^2)^{-1} & D_k^2(I_n + D_k^2)^{-1} \end{pmatrix}, \tag{3.4}$$

is simple and inexpensive.

On the other hand, if $W_k^{\mathrm{T}}\nabla\psi_k = 0$, for all $p \in P_{a,k}$ we have $\nabla\psi_k^{\mathrm{T}} p = \nabla\psi_k^{\mathrm{T}} p_k^N = \nabla\psi_k^{\mathrm{T}} q_k$. In this case, we select d_k as the vector of minimum norm out of the vectors of $P_{a,k}$, i.e.,

$$d_k = W_k\bar{y} + q_k \quad \text{where} \quad \bar{y} = \underset{y\in\mathbb{R}^{m+n}}{\operatorname{argmin}} \|W_k y + q_k\|^2.$$

Due to the structure of W_k it trivially follows:

$$\bar{y} = \begin{pmatrix} 0 \\ (I + D_k^2)^{-1}D_k\tilde{q}_k \end{pmatrix} \quad \text{and} \quad d_k = \begin{pmatrix} 0 \\ (I + D_k^2)^{-1}D_k\tilde{q}_k \\ (I + D_k^2)^{-1}\tilde{q}_k \end{pmatrix}. \tag{3.5}$$

This way, d_k is the steepest descent direction for ψ at x_k among the vectors that are restricted to belong to $P_{a,k}$ and to have 2-norm less than or equal to $\|p_k^N\|$.

In order to make some comments on the properties of d_k we begin introducing the function

$$f(x) = \|F(x)\|^2,$$

the angle θ_k between $-\nabla\psi_k$ and d_k, the angle v_k between $-\nabla\psi_k$ and p_k^N, i.e.,

$$\cos\theta_k = \frac{-\nabla\psi_k^{\mathrm{T}} d_k}{\|\nabla\psi_k\|\|d_k\|}, \quad \cos v_k = \frac{-\nabla\psi_k^{\mathrm{T}} p_k^N}{\|\nabla\psi_k\|\|p_k^N\|}. \tag{3.6}$$

Also, we let $m_k(p)=\|H_k' p+H_k\|^2$ be the standard quadratic model for ψ at x_k and $l_k(p)=\|F_k' p+F_k\|^2$ be the standard quadratic model for f at x_k. Then, assuming that H_k' is invertible, [2, Lemma 2.1] proved three relevant properties of d_k.

First, the angle between d_k and $-\nabla\psi_k$ is smaller than the angle between p_k^N and $-\nabla\psi_k$, i.e.,

$$\cos\theta_k \geqslant \cos v_k. \tag{3.7}$$

Second, if $f_k \neq 0$ we have

$$\nabla f_k^{\mathrm{T}} d_k \leqslant \nabla f_k^{\mathrm{T}} p_k^N \leqslant -2\|F_k\|^2. \tag{3.8}$$

Hence, d_k is a descent direction for the function f at x_k and it guarantees a sufficient progress toward feasibility.

Third, there exists $\beta^* \in [0,1)$ given by

$$\beta^* = \frac{-(\nabla\psi_k^{\mathrm{T}} p_k^N - \nabla\psi_k^{\mathrm{T}} d_k)}{(p_k^N)^{\mathrm{T}} H_k'^{\mathrm{T}} H_k' p_k^N - d_k^{\mathrm{T}} H_k'^{\mathrm{T}} H_k' d_k} \tag{3.9}$$

such that

$$m_k(\rho p_k^N) \leqslant m_k(\rho d_k) \quad \text{and} \quad l_k(\rho p_k^N) \leqslant l_k(\rho d_k) \quad \forall \rho \in [\beta^*, 1],$$

$$m_k(\rho p_k^N) \geqslant m_k(\rho d_k) \quad \text{and} \quad l_k(\rho p_k^N) \geqslant l_k(\rho d_k) \quad \forall \rho \in [0, \beta^*].$$

In other words, both the quadratic models $m_k(p)$ and $l_k(p)$ take lower values along p_k^N than along d_k until the point $\beta^* p_k^N$ is met. This result suggests that the strategy for leaving the Newton step will depend critically on the step $\beta^* p_k^N$. In particular, it is appropriate to backtrack along p_k^N until the point $\beta^* p_k^N$ is reached. Then, the next stage is to leave p_k^N and move towards a reference point belonging to d_k, say $t^* d_k$. Such point should be selected not farther from x_k than $\beta^* p_k^N$ and should be such that

$$m_k(t^* d_k) \leqslant m_k(\beta^* p_k^N).$$

Thus, $t^* d_k$ is chosen as the minimizer of m_k along the direction d_k subject to $0 \leqslant t^* \leqslant \beta^* \|p_k^N\|/\|d_k\|$. Namely, t^* solves

$$t^* = \operatorname*{argmin}_{0 \leqslant t \leqslant \beta^* \frac{\|p_k^N\|}{\|d_k\|}} m_k(t d_k)$$

and it takes the form

$$t^* = \min\left\{ -\frac{\nabla\psi_k^{\mathrm{T}} d_k}{2\|H_k' d_k\|^2},\ \beta^* \frac{\|p_k^N\|}{\|d_k\|} \right\}. \tag{3.10}$$

On the basis of the above properties, in [2] we defined the piecewise linear path $\zeta_k(\lambda)$ for $\lambda \in (0,1]$ that has four nodes: the point zero, $t^* d_k$, $\beta^* p_k^N$ and p_k^N. Now, letting

$$l_{1,k} = (1-\beta^*)\|p_k^N\|, \quad l_{2,k} = \|\beta^* p_k^N - t^* d_k\|, \quad l_{3,k} = t^* \|d_k\|, \tag{3.11}$$

$$l_k = l_{1,k} + l_{2,k} + l_{3,k} \tag{3.12}$$

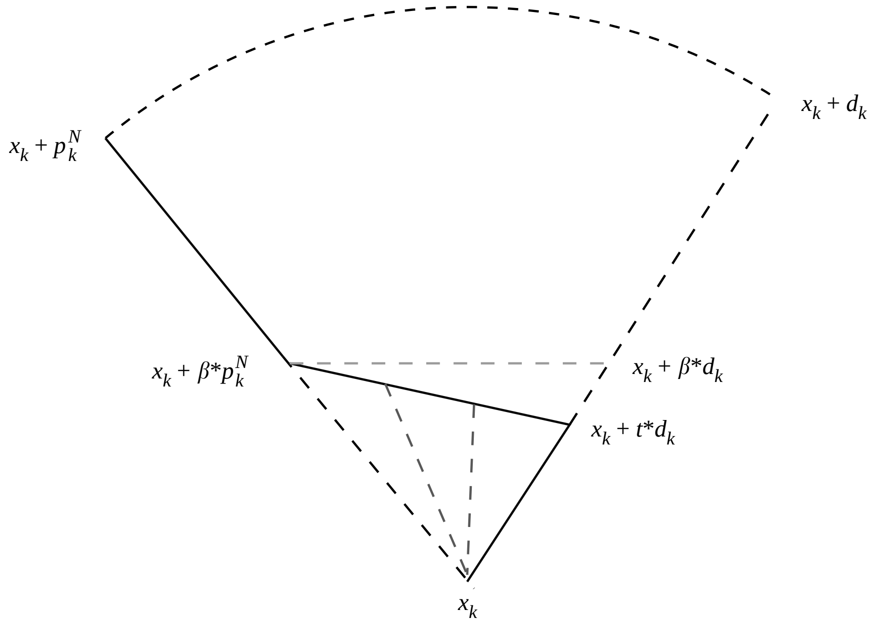

Fig. 1. The piecewise linear path $x_k + \zeta_k(\lambda)$, $\lambda \in [0, 1]$.

and

$$I_{1,k} = \left(\frac{l_{2,k} + l_{3,k}}{l_k}, 1 \right], \tag{3.13}$$

$$I_{2,k} = \left(\frac{l_{3,k}}{l_k}, \frac{l_{2,k} + l_{3,k}}{l_k} \right], \tag{3.14}$$

$$I_{3,k} = \left[0, \frac{l_{3,k}}{l_k} \right], \tag{3.15}$$

$\zeta_k(\lambda)$ can be parametrized as follows:

$$\zeta_k(\lambda) = \begin{cases} \dfrac{\lambda l_k - l_{2,k} - l_{3,k} + \beta^* \| p_k^N \|}{\| p_k^N \|} \, p_k^N, & \text{if } \lambda \in I_{1,k}, \\[2ex] \dfrac{\lambda l_k - l_{3,k}}{l_{2,k}} \, \beta^* p_k^N + \left(1 - \dfrac{\lambda l_k - l_{3,k}}{l_{2,k}} \right) t^* d_k & \text{if } \lambda \in I_{2,k}, \\[2ex] \dfrac{\lambda l_k}{\| d_k \|} d_k & \text{if } \lambda \in I_{3,k}. \end{cases} \tag{3.16}$$

Fig. 1 illustrates the path $x_k + \zeta_k(\lambda)$ which can be viewed as a double dogleg curve [8,18]. Clearly, since p_k^N and d_k are descent directions for ψ at x_k, each vector $\Delta x_k = \zeta_k(\lambda)$ for $\lambda \in (0, 1]$ is a descent direction for ψ at x_k.

Besides (3.6), the path $\zeta_k(\lambda)$ has a further important property that was proved in [2, Lemma 2.2]. If H_k' is invertible, the angle η_k between $-\nabla\psi_k$ and $\Delta x_k = \zeta_k(\lambda)$, is bounded above from the angle v_k between p_k^N and $-\nabla\psi_k$ for any $\lambda \in I_{2,k} \cup I_{3,k}$, i.e.,

$$\cos\eta_k \geqslant \cos v_k, \tag{3.17}$$

where

$$\cos\eta_k = \frac{-\nabla\psi_k^{\mathrm{T}}\Delta x_k}{\|\nabla\psi_k\|\|\Delta x_k\|}. \tag{3.18}$$

Summarizing, $\zeta_k(\lambda)$ is angled away from $-\nabla\psi_k^{\mathrm{T}}$ with respect to p_k^N for $\lambda \in I_{2,k} \cup I_{3,k}$.

Now with $\zeta_k(\lambda)$ at hand we can continue the description of the PLIP method. Its trial points have the form (2.7) with

$$\Delta x_k = \zeta_k(\lambda) \text{ for some } \lambda \in (0,1]$$

and are tested using conditions (2.11)–(2.13). Note that the existence of a $\lambda \in I_{3,k}$ such that $\Delta x_k = \zeta_k(\lambda)$ satisfies (2.11) is ensured because d_k is a descent direction for ψ at x_k [18, Lemma 3.1].

The general algorithmic description of the kth iteration of PLIP method is now sketched.

Algorithm (kth iteration of PLIP method). Let $x_k = (v_k, s_k, z_k)$, $\chi \in (0,1), \sigma_k \in (0,1), \tau_1, \tau_2, \hat{\gamma} > 0$, $\gamma_k \in [\hat{\gamma}, 1), \hat{\beta} > 0, \hat{\beta}_k \in [\hat{\beta}, 1]$, $\alpha \in (0, 1/2), \hat{v} > 0$ be given.

1. Let $D_k = S_k^{-1}Z_k$, $\mu_k = \sigma_k s_k^{\mathrm{T}} z_k/n$, $\tilde{q}_k = -Z_k e + \mu_k S_k^{-1} e$, $\bar{\lambda}_k = 1$.
2. Solve the linear system
   $$H_k' p_k^N = -H_k + \mu_k e_0.$$
3. Form W_k, q_k, $\cos v_k$, by using (2.3) and (3.6).
4. If $W_k^{\mathrm{T}}\nabla\psi_k \neq 0$ then
   compute d_k using (3.2)
   Else
   compute d_k using (3.5).
5. Compute β^* given in (3.9). Set $\beta^* = \max(\beta^*, \hat{\beta}_k)$.
6. Compute t^* given in (3.10).
7. Compute $l_{1,k}, l_{2,k}, l_{3,k}$ and l_k by (3.11) and (3.12).
8. If $\cos v_k < \hat{v}$, let $\bar{\lambda}_k = l_{3,k}/l_k$.
9. If $\bar{\lambda}k = 1$ then
   Compute $\lambda_{k,1}$ s.t. $\forall\lambda \in (0, \lambda_{k,1}]$ the step λp_k^N satisfies (2.12).
   Set $\lambda_k = \max(\lambda_{k,1}, \beta^*)$.

   Else

   Set $\lambda_k = \bar{\lambda}k$.
10. Compute $\zeta_k(\lambda_k)$ by (3.16) and set $\Delta x_k = \zeta_k(\lambda_k)$.
11. While $(\psi(x_k + \Delta x_k) > \psi_k + \alpha\nabla\psi_k^{\mathrm{T}}\Delta x_k)$ or $(f^1(\lambda_k) < 0)$ or $(f^2(\lambda_k) < 0)$
    11.1 Set $\lambda_k = \chi\lambda_k$.
    11.2 Compute $\zeta_k(\lambda_k)$ by (3.16) and set $\Delta x_k = \zeta_k(\lambda_k)$.
12. Let $x_{k+1} = x_k + \Delta x_k$.
13. Choose $\hat{\beta}_{k+1} \in [\hat{\beta}, 1]$, $\sigma_{k+1} > 0$.

We remark that the PLIP algorithm is specifically designed for leaving the Newton direction either when it is nearly orthogonal to the gradient of ψ or when too many backtracks are required in order to satisfy conditions (2.11)–(2.13). At this regard, Steps 5 and 8 are essential aspects of our algorithm. They concern the construction of $\zeta_k(\lambda)$, after p_k^N and d_k were evaluated.

First, let us make some comments on Step 8. Assuming $\hat{v}$ be a small given scalar, if $\cos v_k < \hat{v}$ we drop p_k^N and consider only the segment of $\zeta_k(\lambda)$ that belongs to d_k. The motivation for this issue lies on (3.7), i.e., on the fact that d_k is angled away from $\nabla\psi_k$ with respect to p_k^N.

In the case $\cos v_k \geqslant \hat{v}$, the point $\beta^* p_k^N$ is necessary in order to form $\zeta_k(\lambda)$. Then, we turn our attention to Step 5 where if β^* given in (3.9) is less than the threshold $\hat{\beta}_k$, we set $\beta^*=\hat{\beta}_k$. Specifically, the actually emploied β^* is ensured to be uniformly bounded away from zero in order to overcome both the occurrences where β^* in (3.9) is equal to zero or tiny. In fact, the former case occurs if $W_k^{\mathrm{T}}\nabla\psi_k=0$ and implies that $\zeta_k(\lambda)$ reduces to the line segment p_k^N, while in the latter case too many backtracks along p_k^N would have to be performed before reaching $\beta^* p_k^N$.

We conclude this section considering the case where the PLIP algorithm breaks down, i.e., it is precluded from determining the iterate x_{k+1}. The PLIP algorithm breaks down if H_k' is singular or if $\|\nabla\psi_k\|=0$. In the former case the steps p_k^N and d_k cannot be computed while the latter case occurs if $H_k=0$ or if H_k' is singular and x_k is a stationary point of ψ.

Moreover, it is worth noting that d_k vanishes if only if $H_k=0$. In fact due to (3.1), $d_k=0$ implies $\|\nabla\psi_k^{\mathrm{T}} p_k^N\|=0$ and from (2.10) we have $H_k=0$. If the PLIP method does not break down, it can select the direction d_k since $l_{3,k}/l_k$ in (3.15) is not null. Namely, after a finite number of reductions in Step 11.1, a value of λ_k in $I_{3,k}$ is obtained.

4. Convergence results

In this section, we discuss the theoretical properties of the PLIP method. First, we let $\Omega(\varepsilon)$ be the set

$$\begin{aligned}\Omega(\varepsilon) = \{x = (v,s,z)\in\mathbb{R}^{m+2n} \mid \varepsilon \leqslant \|H(x)\| \leqslant \|H(x_0)\|,\\ \min(ZSe) \geqslant (\tau_1\hat{\gamma})s^{\mathrm{T}}z/n,\ s^{\mathrm{T}}z \geqslant (\tau_2\hat{\gamma})\|F(x)\|\}\end{aligned}$$

with ε a given nonnegative scalar. The sequence $\{x_k\}$ generated by the proposed method belongs to $\Omega(0)$.

In our theoretical analysis, we will assume that the PLIP method does not break down. Hence, we can introduce a neighborhood L of the entire sequence $\{x_k\}$ of the form

$$L = \cup_{k=0}^{\infty}\{x\in\mathbb{R}^{m+2n} \mid \|x - x_k\| \leqslant r\},$$

where $r > 0$ is a fixed constant.

Under the assumptions:

(A1) H is continuously differentiable in $\Omega(0)$;
(A2) F' is Lipschitz continuous with constant L_Ω in $\Omega(\varepsilon), \varepsilon > 0$;
(A3) $\|H'\|$ is bounded above in $L\cap\Omega(0)$,

we will prove that if the sequence $\{x_k\}$ is bounded and $H'(x)$ is invertible in $L \cap \Omega(\varepsilon)$, $\varepsilon > 0$, then $\|H_k\| \to 0$. Further, if there exists a limit point x^* of $\{x_k\}$ such that $H'(x^*)$ is invertible then $x_k \to x^*$.

Next, we will discuss what happens if we drop the assumptions that $H'(x)$ is invertible in $L \cap \Omega(\varepsilon)$, $\varepsilon > 0$, and study whether the PLIP method can be attracted to a singular nonstationary point.

Note that from $z_k^{\mathrm{T}} s_k \geqslant \|Z_k S_k e\|$ and $x_k \in \Omega(0)$ it follows that

$$\begin{aligned} z_k^{\mathrm{T}} s_k &\geqslant \sqrt{(\|Z_k S_k e\|^2 + (\tau_2 \hat{\gamma} \|F_k\|)^2)/2} \\ &\geqslant \sqrt{2} \min(1, \tau_2 \hat{\gamma}) \|H_k\|/2. \end{aligned} \tag{4.1}$$

Therefore, we have $z_k^{\mathrm{T}} s_k$ bounded away from zero in $\Omega(\varepsilon)$, $\varepsilon > 0$. Further, since $\|Z_k S_k e\| \leqslant \|H_k\| \leqslant \|H_0\|$ in $\Omega(0)$, we can conclude that $\{z_k\}$ must be bounded if $\liminf_{k\to\infty} \|s_k\| \nrightarrow 0$ and vice versa.

In the sequel we will use the following technical result.

Lemma 4.1. *Assume that* (A1) *and* (A3) *are satisfied. If* $\|\nabla \psi_k\| \nrightarrow 0$ *there exist two constants* $\varepsilon_1 > 0$ *and* $\varepsilon > 0$ *such that* $\|\nabla \psi_k\| > \varepsilon_1$, $\|H_k\| \geqslant \varepsilon$, *for any* $k > 0$.

Proof. We proceed by contradiction. Assume that there exists a subsequence $\{x_{k_j}\}$ such that $\|\nabla \psi_{k_j}\| \to 0$. This implies $\|H_{k_j}\| \to 0$, because $\nabla \psi_k = H'^{\mathrm{T}}_k H_k$ and $\|H'_k\|$ is bounded by hypothesis. Since the sequence $\{\|H_k\|\}$ is monotone decreasing and bounded it is convergent, consequently $\|H_k\| \to 0$ and this yields $\|\nabla \psi_k\| \to 0$, that is a contradiction. Analogously, assume $\|H_k\| \to 0$. This implies $\|\nabla \psi_k\| \to 0$ and we have again a contradiction. □

The next stage is to show some relevant features of the path $\zeta_k(\lambda)$. In the next two lemmas we will give conditions under which the angle θ_k between d_k and $-\nabla \psi_k$ defined in (3.6) is bounded away from $\pi/2$ whenever $\|\nabla \psi_k\| \nrightarrow 0$.

First, assume that $\{\|p_k^N\|\}$ is bounded above.

Lemma 4.2. *Assume that* (A1) *and* (A3) *are satisfied and* σ_k *is bounded above from one. Let* θ_k *be the angle defined in* (3.6). *If there exists a constant* $\Gamma > 0$ *such that* $\|p_k^N\| \leqslant \Gamma$ *and* $\|\nabla \psi_k\| \nrightarrow 0$ *then there exists* $\delta > 0$ *such that*

$$\cos \theta_k > \delta. \tag{4.2}$$

Proof. Let $K = \sup_{x \in L \cap \Omega(0)} \|H'(x)\|$. Since σ_k is bounded above from one, $\sigma_k \in (0, \bar{\sigma}]$, for some constant $\bar{\sigma} > 0$. Further, condition (2.10) and Lemma 4.1 yields

$$-\nabla \psi_k^{\mathrm{T}} p_k^N \geqslant 2(1 - \bar{\sigma})\varepsilon^2,$$

where $\varepsilon > 0$ is the constant such that $\|H_k\| \geqslant \varepsilon$. Also, recalling (3.7) and that $\|p_k^N\| \leqslant \Gamma$ by hypothesis, we get

$$\cos \theta_k \geqslant \frac{2(1-\bar{\sigma})\varepsilon^2}{\|\nabla \psi_k\| \Gamma} \geqslant \frac{2(1-\bar{\sigma})\varepsilon^2}{K \|H_0\| \Gamma}$$

and the thesis follows with $\delta = (2(1-\bar{\sigma})\varepsilon^2)/(K\|H_0\|\Gamma)$. □

If we drop the assumption on the boundness of $\{\|p_k^N\|\}$, we obtain the following result.

Lemma 4.3. *Assume that* (A1) *and* (A3) *are satisfied and* σ_k *is bounded above from one. If* $\|\nabla\psi_k\| \nrightarrow 0$, $\liminf_{k\to\infty}\|s_k\| \neq 0$ *and* $\liminf_{k\to\infty}\|\nabla\psi_k^{\mathrm{T}} W_k\| \neq 0$, *then there exists* $\delta > 0$ *such that* (4.2) *holds for k sufficiently large.*

Proof. Let us consider the set of indices $\tilde{K}$ such that $\{\|p_k^N\|\}_{k\in\tilde{K}} \to \infty$. The previous lemma yields the existence of $\delta > 0$ such that (4.2) holds for all $k \notin \tilde{K}$. Hence, assume $k \in \tilde{K}$ and k sufficiently large that $\|\nabla\psi_k^{\mathrm{T}} W_k\| > 0$.

From form (3.2) of d_k we get

$$\cos\theta_k = \frac{\nabla\psi_k^{\mathrm{T}}\hat{W}_k q_k - \nabla\psi_k^{\mathrm{T}} q_k + \alpha_k \nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k}{\|\nabla\psi_k\|\|d_k\|}, \tag{4.3}$$

where $\hat{W}_k$ and α_k are defined in (3.3). In order to prove that the angle θ_k is bounded away from $\pi/2$ we need to analyze and bound some quantities.

First, we turn our attention to $\nabla\psi_k^{\mathrm{T}}\hat{W}_k q_k - \nabla\psi_k^{\mathrm{T}} q_k$. By using the following partition of the vector $\nabla\psi_k$, $\nabla\psi_k = (\nabla\psi_{k,1}, \nabla\psi_{k,2}, \nabla\psi_{k,3})$, where $\nabla\psi_{k,1} \in \mathbb{R}^m$ and $\nabla\psi_{k,2}, \nabla\psi_{k,3} \in \mathbb{R}^n$, and the structure of $\hat{W}_k$ (see (3.4)), we have

$$\begin{aligned}\nabla\psi_k^{\mathrm{T}}\hat{W}_k q_k - \nabla\psi_k^{\mathrm{T}} q_k = &-\sum_{i=1}^{n}(\nabla\psi_{k,2})_i\left(-(z_k)_i + \mu_k\frac{1}{(s_k)_i}\right)\left(\frac{(z_k)_i/(s_k)_i}{1+((z_k)_i/(s_k)_i)^2}\right)\\ &-\sum_{i=1}^{n}(\nabla\psi_{k,3})_i\left(-(z_k)_i + \mu_k\frac{1}{(s_k)_i}\right)\left(\frac{1}{1+((z_k)_i/(s_k)_i)^2}\right).\end{aligned}$$

Now, by using Assumption (A3), the inequalities $|(s_k)_i(z_k)_i| \leqslant \|H_0\|$, $(s_k)_i(z_k)_i \geqslant \tau_1\hat{\gamma}s_k^{\mathrm{T}}z_k/n$ and (4.1), it is easy to verify that $\nabla\psi_k^{\mathrm{T}}\hat{W}_k q_k - \nabla\psi_k^{\mathrm{T}} q_k$ is bounded above in $\Omega(\varepsilon), \varepsilon > 0$. Further, since $\|\nabla\psi_k\| \nrightarrow 0$, Lemma (4.1) implies $\|\nabla\psi_k\| \geqslant \varepsilon_1, \varepsilon_1 > 0$. Therefore, the limit $\{\|d_k\|\}_{k\in\tilde{K}} \to \infty$ yields

$$\liminf_{\substack{k\in\tilde{K}\\ k\to\infty}} \cos\theta_k = \liminf_{\substack{k\in\tilde{K}\\ k\to\infty}} \frac{\alpha_k\nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k}{\|\nabla\psi_k\|\|d_k\|} \tag{4.4}$$

and using definition (3.3) of α_k we get

$$\begin{aligned}\frac{\alpha_k\nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k}{\|\nabla\psi_k\|\|d_k\|} &= \frac{\sqrt{\nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k}\sqrt{q_k^{\mathrm{T}}\hat{W}_k q_k - q_k^{\mathrm{T}} q_k + \|p_k^N\|^2}}{\|\nabla\psi_k\|\|p_k^N\|}\\ &= \frac{\sqrt{\nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k}\sqrt{(q_k^{\mathrm{T}}\hat{W}_k q_k - q_k^{\mathrm{T}} q_k)/\|p_k^N\|^2 + 1}}{\|\nabla\psi_k\|}.\end{aligned} \tag{4.5}$$

Now, we consider the quantity $q_k^{\mathrm{T}}\hat{W}_k q_k - q_k^{\mathrm{T}} q_k$ in (4.5). Since

$$q_k^{\mathrm{T}}\hat{W}_k q_k - q_k^{\mathrm{T}} q_k = \sum_{i=1}^{n}\left(-(z_k)_i + \mu_k\frac{1}{(s_k)_i}\right)^2\left(-\frac{1}{1+((z_k)_i/(s_k)_i)^2}\right),$$

it is easy to see that there exists a constant C such that $|q_k^{\mathrm{T}}\hat{W}_k q_k - q_k^{\mathrm{T}} q_k| \leqslant C$ for k sufficiently large. Then, from (4.5) it follows

$$\liminf_{\substack{k\in\tilde{K}\\ k\to\infty}} \cos\theta_k = \liminf_{\substack{k\in\tilde{K}\\ k\to\infty}} \frac{\sqrt{\nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k}}{\|\nabla\psi_k\|}.$$

Recalling definition (3.3) of $\hat{W}_k$, we can write

$$\hat{W}_k = W_k\tilde{W}_k^2 W_k^{\mathrm{T}},$$

where

$$\tilde{W}_k = \begin{pmatrix} I_m & 0 \\ 0 & (I_n + D_k^2)^{-1/2} \end{pmatrix}.$$

Then

$$\nabla\psi_k^{\mathrm{T}}\hat{W}_k\nabla\psi_k = \|\tilde{W}_k W_k^{\mathrm{T}}\nabla\psi_k\|^2 \geqslant \left(\min\left(1, \min_{1\leqslant i\leqslant n}\frac{1}{\sqrt{1+((z_k)_i/(s_k)_i)^2}}\right)\right)^2 \|W_k^{\mathrm{T}}\nabla\psi_k\|^2.$$

Therefore, since $\{s_k\}$ is bounded away from zero, the inequality $\|Z_k S_k e\| \leqslant \|H_0\|$ implies that $(z_k)_i$ are bounded above for $i = 1,\ldots,n$. Then, $\liminf_{k\to\infty}\|\nabla\psi_k^{\mathrm{T}} W_k\| \neq 0$ and Assumption (A3) yield the thesis. □

Next two lemmas show that the last segment of the curve ζ_k is uniformly bounded from zero.

Lemma 4.4. *Assume that H_k' is invertible for each k, σ_k is bounded above from one and* (A1) *and* (A3) *are satisfied. If $\|\nabla\psi_k\| \nrightarrow 0$, then there exists a constant C such that*

$$t^*\|d_k\| \geqslant C \tag{4.6}$$

for all the indeces k such that (4.2) *holds.*

Proof. Let $K = \sup_{x\in L\cap\Omega(0)}\|H'(x)\|$ and $\bar{\sigma} > 0$ such that $\sigma_k \in (0,\bar{\sigma}]$. Since $\|\nabla\psi_k\| \nrightarrow 0$, Lemma 4.1 ensures that there exist $\varepsilon_1 > 0$, $\varepsilon > 0$ such that $\|\nabla\psi_k\| \geqslant \varepsilon_1$ and $\|H_k\| \geqslant \varepsilon$.

To prove (4.6), we recall form (3.10) of t^*. If $t^* = |\nabla\psi_k^{\mathrm{T}} d_k|/(2\|H_k' d_k\|^2)$, using the bound (4.2) we get

$$t^*\|d_k\| = \frac{|\nabla\psi_k^{\mathrm{T}} d_k|}{\|\nabla\psi_k\|\|d_k\|}\,\frac{\|\nabla\psi_k\|\|d_k\|^2}{2\|H_k' d_k\|^2} \geqslant \delta\frac{\|\nabla\psi_k\|}{2\|H_k'\|^2} \geqslant \delta\frac{\varepsilon_1}{2K^2}.$$

On the contrary, if $t^* = \beta^*\|p_k^N\|/\|d_k\|$, we have

$$t^*\|d_k\| = \beta^*\|p_k^N\|. \tag{4.7}$$

Let us examine $\|p_k^N\|$; since p_k^N is the solution of (1.2) it satisfies the following inequality:

$$\|p_k^N\| \geqslant \|-S_k Z_k e + \mu_k e\|/\|H_k'\|. \tag{4.8}$$

Let $\hat{\imath}$ be the index such that $(s_k)_{\hat{\imath}}(z_k)_{\hat{\imath}} = \max_i (s_k)_i (z_k)_i$. Taking into account that

$$\begin{aligned} \| -S_k Z_k e + \mu_k e\| &\geqslant \| -S_k Z_k e + \mu_k e\|_1/\sqrt{n} \\ &\geqslant |-(s_k)_{\hat{\imath}}(z_k)_{\hat{\imath}} + \mu_k|/\sqrt{n} \\ &= ((s_k)_{\hat{\imath}}(z_k)_{\hat{\imath}} - \mu_k)/\sqrt{n} \\ &\geqslant (1-\sigma_k) s_k^{\mathrm{T}} z_k/(n\sqrt{n}) \end{aligned} \tag{4.9}$$

and using inequality (4.1) we get

$$\| -S_k Z_k e + \mu_k e\| \geqslant (1-\bar{\sigma})\sqrt{2}\min(1, \tau_2\hat{\gamma})\|H_k\|/(2n\sqrt{n}).$$

Thus,

$$\|p_k^N\| \geqslant (1-\bar{\sigma})\sqrt{2}\min(1, \tau_2\hat{\gamma})\|H_k\|/(2Kn\sqrt{n}). \tag{4.10}$$

This latter inequality along with (4.7) and (4.8) yields

$$t^*\|d_k\| \geqslant \hat{\beta}(1-\bar{\sigma})\sqrt{2}\min(1, \tau_2\hat{\gamma})\varepsilon/(2Kn\sqrt{n}).$$

Hence, letting

$$C = \min\left\{\delta\frac{\varepsilon_1}{2K^2}, \hat{\beta}(1-\bar{\sigma})\frac{\varepsilon\sqrt{2}\min(1, \tau_2\hat{\gamma})}{2Kn\sqrt{n}}\right\},$$

the thesis follows. □

Lemma 4.5. *Assume that H_k' is invertible for each k and* (A1) *and* (A3) *are satisfied. Let $\hat{\nu} > 0$ be a given constant, $l_{3,k}$ and l_k be the quantities defined in* (3.11) *and* (3.12). *If $\|\nabla\psi_k\| \nrightarrow 0$ then there exists a constant C_1 such that*

$$\frac{l_{3,k}}{l_k} > C_1 \tag{4.11}$$

for all the indeces k such that (4.2) *holds and* $\cos\nu_k > \hat{\nu}$.

Proof. Since $\|\nabla\psi_k\| \nrightarrow 0$ there exists $\varepsilon_1 > 0$ such that $\|\nabla\psi_k\| \geqslant \varepsilon_1$ (see Lemma 4.1). Then, noting that (3.10) yields $t^* \leqslant \beta^*\|p_k^N\|/\|d_k\|$ and by using (3.11), (3.12) we have

$$\begin{aligned} \frac{l_{3,k}}{l_k} &= \frac{t^*\|d_k\|}{t^*\|d_k\| + (1-\beta^*)\|p_k^N\| + \|t^* d_k - \beta^* p_k^N\|} \\ &\geqslant \frac{t^*\|d_k\|}{2t^*\|d_k\| + \|p_k^N\|} \geqslant \frac{t^*\|d_k\|}{3\|p_k^N\|}. \end{aligned}$$

Further, from Lemma 4.4, we get

$$\frac{l_{3,k}}{l_k} \geqslant \frac{C}{3\|p_k^N\|}. \tag{4.12}$$

Let us consider $\|p_k^N\|$. Since $\cos v_k > \hat{v}$ holds by hypothesis, we have

$$\begin{aligned}\hat{v} &< \frac{-\nabla\psi_k^{\mathrm{T}} p_k^N}{\|\nabla\psi_k\|\|p_k^N\|}\\ &= 2\frac{\|H_k\|^2 - \sigma_k(s_k^{\mathrm{T}} z_k)^2/n}{\|\nabla\psi_k\|\|p_k^N\|}\\ &\leqslant \frac{2\|H_k\|^2}{\|\nabla\psi_k\|\|p_k^N\|}.\end{aligned}$$

Hence, we obtain

$$\|p_k^N\| < \frac{2\|H_0\|^2}{\varepsilon_1\hat{v}} \tag{4.13}$$

and (4.12) and (4.13) yield the thesis. □

Next we turn our attention to the centering conditions. We will prove that (2.12) and (2.13) can be satisfied along d_k with a steplength bounded away from zero whenever $\{x_k\}$ belongs to $\Omega(\varepsilon), \varepsilon > 0$, $\{p_k^N\}$ is bounded above and the sequence $\{x_k\}$ is bounded. To this end, let $\{\lambda_k^{\mathrm{c}}\}$ denote the sequence of steplengths such that the centering conditions (2.12), (2.13) are satisfied. Namely, at the kth iteration λ_k^{c} is given by

$$\lambda_k^{\mathrm{c}} = \min(\lambda_{k,1}, \lambda_{k,2}), \tag{4.14}$$

where $\lambda_{k,1}$ and $\lambda_{k,2}$ are such that

$$f^1(\lambda) \geqslant 0 \quad \forall\lambda\in(0,\lambda_{k,1}], \qquad f^2(\lambda) \geqslant 0 \quad \forall\lambda\in(0,\lambda_{k,2}].$$

Theorem 4.1. *Assume that Assumptions* (A1)–(A3) *are satisfied and* $\{\sigma_k\}$ *is bounded away from zero and one, i.e.* $\sigma_k\in[\underline{\sigma}, \bar{\sigma}]$, *where* $0 < \underline{\sigma} < \bar{\sigma} < 1$. *If* $\|\nabla\psi_k\| \nrightarrow 0$, $\|p_k^N\| \nrightarrow \infty$ *and the sequence* $\{x_k\}$ *is bounded, then* $\{\lambda_k^{\mathrm{c}}\}$ *is bounded away from zero.*

Proof. Let $K = \sup_{x\in L\cap\Omega(0)}\|H'(x)\|$. Since $\|\nabla\psi_k\| \nrightarrow 0$, from Lemma 4.1 it follows that there exist $\varepsilon_1 > 0$ and $\varepsilon > 0$ such that $\|H_k\| \geqslant \varepsilon$ and $\|\nabla\psi_k\| \geqslant \varepsilon_1$. Hence, $x_k\in\Omega(\varepsilon)$. Further, from our assumptions there exists a constant Γ such that $\|d_k\| \leqslant \|p_k^N\| \leqslant \Gamma$.

First, we stress that under our assumptions $l_{3,k}/l_k$ is bounded away from zero and consequently after a finite number of backtracks the PLIP method switches to the direction d_k. Hence, to prove the thesis, we need to show that the centering conditions can be satisfied with a bounded λ_k along the last segment of the curve. Namely, we focus on steps Δx_k of the form $\Delta x_k = \lambda l_k d_k/\|d_k\|$, $0 < \lambda \leqslant l_{3,k}/l_k$.

In order to simplify the notation, the iteration index k is omitted in the following analysis and ρ is used as a shorthand for $\lambda l/\|d\|$, i.e., we will consider steps $\Delta x = \rho d$, $0 < \rho \leqslant t^*$. Further, we will use the following partition of d: $d = (d_v, d_s, d_z)$ where $d_v\in\mathbb{R}^m$, $d_s, d_z\in\mathbb{R}^n$ and we will use $s(\rho)$ and $z(\rho)$ as shorthand for $s + \rho d_s$ and $z + \rho d_z$.

We first show that λ_1 is bounded away from zero. Since $\|d\| \leqslant \Gamma$ there exists a positive constant M_1 such that

$$|(d_z)_i(d_s)_i - (\tau_1\gamma/n)d_z^{\mathrm{T}}d_s| \leqslant M_1$$

holds for $i = 1,\dots,n$. Recalling that $d \in P_a$, i.e., it satisfies the last block of n equations in (1.2), we obtain

$$\begin{aligned} s_i(\rho)z_i(\rho) - (\tau_1\gamma/n)s(\rho)^{\mathrm{T}}z(\rho) &= (1-\rho)(z_is_i - (\tau_1\gamma/n)s^{\mathrm{T}}z) + (1-\tau_1\gamma)\sigma\rho\frac{s^{\mathrm{T}}z}{n} \\ &\quad + \rho^2((d_z)_i(d_s)_i - (\tau_1\gamma/n)d_z^{\mathrm{T}}d_s) \\ &\geqslant \rho(1-\tau_1\gamma)\sigma\frac{s^{\mathrm{T}}z}{n} - \rho^2M_1. \end{aligned}$$

Consequently, $f^1(\lambda) \geqslant 0$ for $\lambda \in (0, \lambda_1]$ where λ_1 is given by

$$\lambda_1 = \frac{\bar\rho\|d\|}{l} \quad \text{and} \quad \bar\rho \geqslant \min\left\{\frac{s^{\mathrm{T}}z(1-\tau_1\gamma)\sigma}{nM_1}, t^*\right\}.$$

From (4.1) we know that $s^{\mathrm{T}}z$ is bounded away from zero in $\Omega(\varepsilon)$ and Lemma 4.4 ensures that $t^*\|d\|$ is bounded away from zero. Moreover l is bounded from above since $\|p^N\|$ is bounded by hypothesis. Hence to prove the boundness of λ_1 we need to show that there exists $C > 0$ such that

$$\|d\| \geqslant C\varepsilon. \tag{4.15}$$

To this end, note that if $W^{\mathrm{T}}\nabla\psi \neq 0$, then by construction $\|d\| = \|p^N\|$ and from (4.15) we get that there exists $C > 0$ such that (4.15) holds. On the other hand, when d is given by (3.5), letting $\hat\imath$ be the index such that $s_{\hat\imath}z_{\hat\imath} = \max_i s_iz_i$ we can proceed as follows:

$$\|d\| \geqslant \frac{\|d\|_1}{\sqrt{n}} \geqslant \frac{s_{\hat\imath}z_{\hat\imath} - \sigma s^{\mathrm{T}}z/n}{\sqrt{n}(s_{\hat\imath}^2 + z_{\hat\imath}^2)}s_{\hat\imath} \geqslant \frac{(1-\bar\sigma)s^{\mathrm{T}}z/n}{\sqrt{n}(s_{\hat\imath}^2 + z_{\hat\imath}^2)}s_{\hat\imath}.$$

Now, note that from (4.1) it follows that $s^{\mathrm{T}}z$ is bounded away from zero. Then, since $s_{\hat\imath}z_{\hat\imath} \geqslant s^{\mathrm{T}}z/n$ and the sequence $\{x_k\}$ is bounded by hypothesis, we can conclude that (4.15) holds and consequently λ_1 is bounded away from zero.

Next, we show that λ_2 is bounded away from zero, too. Since $\|d\|$ is bounded above there exists a positive constant M_2 such that

$$2\sqrt{n}\|H_0\||d_s^{\mathrm{T}}d_z| + \frac{(\tau_2\gamma)^2}{2}\tilde L_\Omega\|d\|^2 \leqslant M_2.$$

Since $F'(x)$ is Lipschitz continuous in $\Omega(\varepsilon)$ with constant L_Ω, $\nabla f(x)$ is Lipschitz continuous in $\Omega(\varepsilon)$ with constant $\tilde L_\Omega = (L_\Omega L_R + L_R^2)$ where $L_R = \max(\sup_{x\in\Omega(\varepsilon)} f(x), \sup_{x\in\Omega(\varepsilon)}\|F'(x)\|)$ (see [18]). By using the mean-value theorem, and inequality (3.8) we obtain

$$\begin{aligned} f(x+\rho d) &= f(x) + \rho\nabla f(x)^{\mathrm{T}}d + \rho\int_0^1((\nabla f(x+u\rho d) - \nabla f(x))^{\mathrm{T}}d)\,\mathrm{d}u \\ &\leqslant (1-2\rho)\|F(x)\|^2 + \rho\int_0^1((\nabla f(x+u\rho d) - \nabla f(x))^{\mathrm{T}}d)\,\mathrm{d}u \end{aligned}$$

and the Lipschitz continuity of $\nabla f(x)$ yields

$$f(x+\rho d) \leqslant (1-2\rho)f(x) + \frac{\rho^2}{2}\tilde{L}_\Omega\|d\|^2.$$

Then, assuming $\rho \leqslant \frac{1}{2}$ and recalling that x_k satisfies (2.12)–(2.13) we get

$$\begin{aligned}(s(\rho)^{\mathrm{T}}z(\rho))^2 - (\tau_2\gamma)^2 f(x+\rho d) &\geqslant (z^{\mathrm{T}}s)^2(1-\rho+\rho\sigma)^2 + (\rho)^4(d_s^{\mathrm{T}}d_z)^2\\ &\quad + 2\rho^2(1-\rho+\rho\sigma)(z^{\mathrm{T}}s)d_s^{\mathrm{T}}d_z\\ &\quad - (\tau_2\gamma)^2((1-2\rho)f(x) + (\rho^2/2)\tilde{L}_\Omega\|d\|^2)\\ &\geqslant (z^{\mathrm{T}}s)^2(1-\rho+\rho\sigma)^2 - 2\rho^2(z^{\mathrm{T}}s)|d_s^{\mathrm{T}}d_z|\\ &\quad - (1-2\rho)(z^{\mathrm{T}}s)^2 - (\tau_2\gamma)^2(\rho^2/2)\tilde{L}_\Omega\|d\|^2.\end{aligned}$$

Moreover, noting that $z^{\mathrm{T}}s = \|SZe\|_1 \leqslant \sqrt{n}\|SZe\| \leqslant \|H_0\|$ we have

$$\begin{aligned}(s(\rho)^{\mathrm{T}}z(\rho))^2 - (\tau_2\gamma)^2 f(x+\rho d) &\geqslant (z^{\mathrm{T}}s)^2((1-\rho+\rho\sigma)^2 - (1-2\rho))\\ &\quad - \rho^2(2\sqrt{n}\|H_0\||d_s^{\mathrm{T}}d_z| + ((\tau_2\gamma)^2/2)\tilde{L}_\Omega\|d\|^2)\\ &\geqslant (z^{\mathrm{T}}s)^2(1+\rho^2(\sigma-1)^2 + 2(\sigma-1)\rho - 1 + 2\rho) - \rho^2 M_2\\ &\geqslant 2(z^{\mathrm{T}}s)^2\sigma\rho - \rho^2 M_2.\end{aligned}$$

Hence, $f^2(\lambda) \geqslant 0$ for $\lambda \in (0, \lambda_2]$ where λ_2 is given by

$$\lambda_2 = \frac{\bar{\rho}\|d\|}{l} \quad \text{and} \quad \bar{\rho} \geqslant \min\left\{2\underline{\sigma}(s^{\mathrm{T}}z)^2/M_2, t^*, \frac{1}{2}\right\}$$

and λ_2 is bounded away from zero because of (4.1), (4.15) and Lemma 4.4. Hence the thesis follows. □

Now we state convergence results for the PLIP method when H' is invertible in $L \cap \Omega(\varepsilon)$ and the sequence $\{x_k\}$ is bounded.

Theorem 4.2. *Let $\{x_k\}$ be generated by the PLIP method. Assume that* (A1)–(A3) *are satisfied and σ_k is bounded above from one. Further, assume that σ_k is bounded away from zero whenever $\|H_k\| \nrightarrow 0$. If $H'(x)$ is invertible in $L \cap \Omega(\varepsilon)$ for any $\varepsilon > 0$ and the sequence $\{x_k\}$ is bounded, then $\|H_k\| \to 0$. Further, if there exists an accumulation point x^** of *$\{x_k\}$ such that $H'(x^*)$ is invertible then $x_k \to x^*$.*

Proof. Note that the sequence $\{\|H_k\|\}$ is decreasing and bounded; hence it is convergent. Suppose that the limit is $\tilde{\gamma} > 0$ and let $\bar{\sigma}$ be such that $\sigma_k \leqslant \bar{\sigma}$.

Let x^* be a limit point of $\{x_k\}$ and $\{x_{k_j}\}$ be a subsequence such that $x_{k_j} \to x^*$. Since $\|H(x^*)\| = \tilde{\gamma}$, it follows that $x^* \in \Omega(\tilde{\gamma})$ and therefore $H'(x^*)$ is invertible. This implies $\|\nabla\psi_k\| \nrightarrow 0$ and $\|p_{k_j}^N\|$ bounded. Hence, from Lemma 4.2 we have that (4.2) holds for $k = k_j$ and from Theorem 4.1 we get that the sequence $\{\lambda_{k_j}^c\}$ is bounded away from zero.

The backtracking linesearch based on the Armijo condition (2.11) produces

$$\frac{\nabla\psi_k^{\mathrm{T}}\Delta x_k}{\|\Delta x_k\|} \to 0$$

(see [18, Theorem 3.2]). From (2.10) and (3.17) we get

$$\frac{\nabla\psi_k^{\mathrm{T}}\Delta x_k}{\|\Delta x_k\|} \leqslant \frac{\nabla\psi_k^{\mathrm{T}} p_k^N}{\|p_k^N\|} \leqslant 0$$

and since $\|p_{k_j}^N\|$ is bounded, we conclude that $\nabla\psi_{k_j}^{\mathrm{T}} p_{k_j}^N \to 0$. This contradicts our assumption that $\tilde{\gamma} > 0$ because from (2.10) we have

$$\nabla\psi_{k_j}^{\mathrm{T}} p_{k_j}^N \leqslant -2(1-\sigma_{k_j})\|H_{k_j}\|^2 \leqslant -2(1-\bar{\sigma})\|H_{k_j}\|^2 < 0.$$

Hence $\|H_k\|$ must converge to zero.

To prove the second part of the theorem we need to show that our search directions Δx_k satisfy

$$\|H_k'\Delta x_k + H_k\| \leqslant \|H_k\|.$$

In fact, the Newton direction is such that

$$\|H_k' p_k^N + H_k\| \leqslant \sigma_k\|H_k\| \tag{4.16}$$

and from the definition of t^* it follows that $\|H_k' t^* d_k + H_k\| \leqslant \|H_k\|$. Further, it can be easily verified that the following inequalities:

$$\|\rho H_k' p_k^N + H_k\| \leqslant \|H_k\|, \qquad \|\rho H_k' t^* d_k + H_k\| \leqslant \|H_k\|$$

hold for any $\rho \in [0,1]$. Regarding the directions $\Delta x_k = (1-\xi)\beta^* p_k^N + \xi t^* d_k$ we have

$$\begin{aligned}\|H_k'\Delta x_k + H_k\| &= \|(1-\xi)\beta^* H_k' p_k^N + (1-\xi)H_k + \xi t^* H_k' d_k + \xi H_k\| \\ &\leqslant (1-\xi)\|\beta^* H_k' p_k^N + H_k\| + \xi\|t^* H_k' d_k + H_k\| \leqslant \|H_k\|.\end{aligned}$$

Hence our search directions are inexact Newton directions for the problem $H(x) = 0$ and by using Theorem 3.3 of [10] we get that $x_k \to x^*$. □

Now we investigate the asymptotic rate of convergence to a point x^* such that $H'(x^*)$ is invertible. The next result shows that under a suitable choice of the centering parameter σ_k, the step $\Delta x_k = p_k^N$ is eventually chosen and the ultimate rate of convergence is superlinear.

Theorem 4.3. *Assume that* (A1)–(A3) *are satisfied. Let* $\{x_k\}$ *be generated by the PLIP method and suppose that* $\{x_k\} \to x^*$ *such that* $H(x^*) = 0$ *and* $H'(x^*)$ *is nonsingular. If* $\sigma_k = O(\|H_k\|^p)$, $0 < p < 1$, *and there is an open neighborhood* D *of* x^* *such that* H *is twice differentiable, with* $\|\nabla^2 H(x)_i\|$, $i = 1, \ldots, n$, *bounded for* $x \in D$, *then there exists an index* $k_0 > 0$ *such that* $\lambda_k = 1$ *for* $k \geqslant k_0$. *Furthermore* $x_k \to x^*$ *superlinearly.*

Proof. From [1, Theorem 4.1] it follows that, if $\sigma_k = O(\|H_k\|^p)$, $0 < p < 1$ is chosen, then the centering conditions are satisfied with $\lambda_k = 1$ for sufficiently large k.

In order to show that $\lambda_k = 1$ will eventually satisfy the decrease condition (2.11) too, we show that the so-called Dennis–Morè condition holds, i.e.,

$$\lim_{k\to\infty} \frac{\|\nabla\psi_k + \nabla^2\psi_k p_k^N\|}{\|p_k^N\|} = 0. \tag{4.17}$$

To this end, note that $\nabla^2\psi_k = 2(H_k'^{\mathrm{T}}H_k' + \sum_{i=1}^n (H_k)_i\nabla^2(H_k)_i)$ and therefore

$$\begin{aligned}\frac{\|\nabla\psi_k + \nabla^2\psi_k p_k^N\|}{\|p_k^N\|} &\leqslant 2\,\frac{\|H_k'^{\mathrm{T}}H_k + H_k'^{\mathrm{T}}H_k' p_k^N\|}{\|p_k^N\|} + 2\left\|\sum_{i=1}^n (H_k)_i\nabla^2(H_k)_i\right\| \\ &\leqslant 2\,\frac{\|H_k'\|\mu_k\sqrt{n}}{\|p_k^N\|} + 2\left\|\sum_{i=1}^n (H_k)_i\nabla^2(H_k)_i\right\|.\end{aligned}$$

This latter inequality along with (4.8) and (4.9) yields

$$\frac{\|\nabla\psi_k + \nabla^2\psi_k p_k^N\|}{\|p_k^N\|} \leqslant 2\frac{\sigma_k n\|H_k'\|^2}{1-\sigma_k} + 2\left\|\sum_{i=1}^n (H_k)_i\nabla^2(H_k)_i\right\|.$$

Therefore, since $\sigma_k \to 0$ and $\|\sum_{i=1}^n (H_k)_i\nabla^2(H_k)_i\| \to 0$, we can conclude that condition (4.17) holds. Then, from [7, Theorem 6.4] it follows that $\lambda_k = 1$ will eventually satisfy (2.11).

Hence, the method reduces to an inexact Newton method with forcing term σ_k (see (4.16)), and $\sigma_k \to 0$ implies the superlinear convergence rate [6, Theorem 3.3]. □

So far, we have studied the convergence behavior of the PLIP method assuming the invertibility of the Jacobian of H in $L \cap \Omega(\varepsilon)$, $\varepsilon > 0$. We end this section by discussing the occurrence where H' is not ensured to be invertible everywhere in $\Omega(\varepsilon)$, i.e., singular nonstationary points might be present.

At the end of Section 2 we pointed out that when a classical interior point method approaches a singular nonstationary point $\tilde{x}$, the length of p_k^N tends to infinity and the generated sequence $\{x_k\}$ might get trapped around $\tilde{x}$ for two different reasons. In the first type of failure, the scalars λ_k^c (see (4.14)) and λ_k tends to zero to enforce the bounds. In this case, we cannot prove that the convergence properties of our method are superior to those of the basic linesearch methods. In fact, when $\nabla\psi_k^{\mathrm{T}} W_k \neq 0$ our alternative direction d_k has the same norm of p_k^N and it is not ensured to point inwards.

On the contrary, the second type of failure is not related to the presence of bounds: the Newton direction tends to become orthogonal to the gradient of the merit function and classical methods become stuck near $\tilde{x}$ in order to satisfy a sufficient decrease condition on ψ. The PLIP method can prevent this failure by using d_k since the angle between d_k and $-\nabla\psi_k$ is bounded away from $\pi/2$ when $\liminf_{k\to\infty}\|\nabla\psi_k^{\mathrm{T}} W_k\| \neq 0$, see Lemma 4.3. Next theorem proves this fact: if the centering conditions are satisfied with bounded λ_k^c, then the sequence $\{x_k\}$ cannot tend to a singular point $\tilde{x} = (\tilde{v}, \tilde{s}, \tilde{z})$ with $\tilde{s} > 0$, except in the special case when $F(\tilde{x}) \in \mathrm{Ker}(J^{\mathrm{T}}(\tilde{x}))$.

Theorem 4.4. *Let* $\{x_k\}$ *be generated by the PLIP algorithm. Assume that* (A1)–(A3) *are satisfied and* σ_k *is bounded away from one. Let* $\tilde{x} = (\tilde{v}, \tilde{s}, \tilde{z})$ *be an accumulation point of the sequence such that* $\tilde{s} > 0$ *an* $J(\tilde{x})$ *be the matrix* (2.6) *at* $\tilde{x}$. *If the sequence* $\{\lambda_k^c\}$, *defined in* (4.14) *is bounded away from zero, then* $\nabla\psi(\tilde{x})^T W(\tilde{x}) = 0$, *i.e., one of the following situations occurs*:

(a) $\nabla\psi(\tilde{x}) = 0$,
(b) $\nabla\psi(\tilde{x}) \neq 0$ *and* $F(\tilde{x}) \in \mathrm{Ker}(J^T(\tilde{x}))$.

Proof. Let $\{x_{k_j}\}$ be a subsequence such that $x_{k_j} \to \tilde{x}$.

Suppose that $\liminf_{k\to\infty} \|\nabla\psi_{k_j}^T W_{k_j}\| \neq 0$. Taking into account that $\tilde{s} > 0$ implies that $\|W_{k_j}\|$ is bounded above, we have that $\|\nabla\psi_{k_j}\| \nrightarrow 0$. Since the sequence $\{\lambda_{k_j}^c\}$ is bounded away from zero, the backtracking linesearch used in our algorithm, produces

$$\frac{\nabla\psi_k^T \Delta x_k}{\|\Delta x_k\|} \to 0$$

(see [18, Theorem 3.2]). From Lemma 4.3 there exists $\delta > 0$ such that

$$\frac{\nabla\psi_{k_j}^T \Delta x_{k_j}}{\|\Delta x_{k_j}\|} > \delta \|\nabla\psi_{k_j}\|$$

for k sufficiently large. Consequently $\|\nabla\psi_{k_j}\| \to 0$. This is a contradiction.

Thus, it must be $\|\nabla\psi_{k_j}^T W_{k_j}\| \to 0$ and $\nabla\psi(\tilde{x})^T W(\tilde{x}) = 0$. Since

$$\nabla\psi(x)^T W(x) = F(x)^T J(x),$$

condition $\nabla\psi(\tilde{x})^T W(\tilde{x}) = 0$ occurs if $\nabla\psi(\tilde{x}) = 0$ or $\nabla\psi(\tilde{x}) \neq 0$ and $F(\tilde{x}) \in \mathrm{Ker}(J^T(\tilde{x}))$. □

5. Numerical results

In this section, we are mainly interested in how the PLIP method compares to the basic linesearch interior point methods. Hence, we apply it in the solution of the hard tests given in [5,14,24,28,29], that cannot be solved by classical linesearch interior point methods.

We implemented the PLIP and CLIP methods as MATLAB programs and run them under MATLAB version 5.3 with machine precision about 10^{-16}. The numerical experiments were done on an HP 9000 C200 workstation.

In (2.9) the value of σ_k needed to form μ_k was set equal to

$$\sigma_k = \min\{0.5, \|H_k\|^{1/2}\}.$$

This choice ensures superlinear asymptotic rate of convergence of the PLIP method (see Theorem 4.3) and of the CLIP method, see [1]. Then, $p_{k,v}^N$ and $p_{k,s}^N$ were computed solving the linear system (2.5) by Gaussian elimination and $p_{k,z}^N$ was computed by (2.4).

The parameter γ_k used in the centering conditions (2.12) and (2.13) was taken constant and set to 10^{-6}, the parameter α in the Armijo condition (2.11) was set to 10^{-4}.

Regarding the backtracking process, in the PLIP and CLIP methods we used the same strategy to shrink the step. First, $\lambda_{k,1}$ was computed by solving n quadratic equations. Then in Step 11 of the

PLIP algorithm λ_k was updated using the constant $\chi = 0.1$. Analogously, in the CLIP method, the selected step had the form $\lambda_k\, p_k^N$ where $\lambda_k = \chi^i \lambda_{k,1}$, $\chi = 0.1$ and i is the smallest integer such that (2.11) and (2.13) are satisfied.

In the PLIP method the thresholds $\hat{v}$ and $\hat{\beta}$ were set to 10^{-8} and 10^{-2}, respectively. Further, in Step 13 the threshold $\hat{\beta}_{k+1}$ was computed using the following rule:

if $\|H_{k+1}\| \geqslant 0.9\|H_k\|$ then
  $\hat{\beta}_{k+1} = \min(1, 2\hat{\beta}_k)$
else if $\|H_{k+1}\| \leqslant 0.6\|H_k\|$ then
  $\hat{\beta}_{k+1} = \max(\hat{\beta}, 0.5\,\hat{\beta}_k)$
else
  $\hat{\beta}_{k+1} = \hat{\beta}_k$.

Clearly, we update the value of the threshold $\hat{\beta}_k$ depending on the actual reduction of $\|H_{k+1}\|$ with respect to $\|H_k\|$. The threshold is halved if a great reduction in the value of $\|H\|$ occurs. On the contrary, if a poor reduction in the value of norm of H is detected, $\hat{\beta}_{k+1}$ is the double of $\hat{\beta}_k$. Finally, i.e., for $0.6 < \|H_{k+1}\|/\|H_k\| < 0.9$, the threshold is kept the same for the next step.

For both methods we terminated the iteration when

$$\|H_k\| \leqslant 10^{-6}\sqrt{m+2n}.$$

Failure was declared when 50 backtracks were not enough to satisfy the centering conditions and the Armijo condition, or if within 300 iterations the stopping criterion was not met. Also, failure of the CLIP method was declared if $\lambda_{k,1} < 10^{-12}$ was detected.

Through the iterations we monitored the following quantities:

$\lambda_{k,1}$:	the steplength taken in order to satisfy condition (2.12) in the CLIP method;
λ_k:	the steplength taken in order to satisfy conditions (2.11)–(2.13);
$\cos(v_k)$, $\cos(\eta_k)$:	the angles defined in (3.6) and (3.18).

Concerning the PLIP method it is important to monitor the steps used, too. In the following tables for a given iterate k we indicate the line segment I of $x_k + \zeta_k(\lambda)$ on which the accepted point x_{k+1} lies. Specifically, since the accepted step Δx_k has the form $\zeta_k(\lambda)$ where λ belongs to one of the intervals $I = I_{1,k}, I_{2,k}, I_{3,k}$ given in (3.13),(3.14), (3.15) we drop the index k for brevity and let $I = I_1, I_2, I_3$, respectively.

In the sequel we report some results obtained with meaningful test problems from Wachter and Biegler [28], Byrd et al. [5] and Simantiraky and Shanno [24].

Example 1 (Wachter and Biegler [28]). Consider the problem:

$$\begin{aligned} \min \quad & w_1 \\ \text{s.t.} \quad & w_1^2 - w_2 - 1 = 0, \\ & w_1 - w_3 - 2 = 0, \\ & w_2 \geqslant 0 \;\; w_3 \geqslant 0, \end{aligned}$$

Table 1
Example 1. CLIP method with starting point in $\mathscr{D}$

k	$\|H_k\|$	$\lambda_{k,1}$	λ_k	$\|\cos(v_k)\|$
1	5.2e0	2.0e-1	2.0e-1	2.0e-1
⋮				
10	4.8e0	5.2e-3	5.2e-4	8.2e-4
⋮				
50	4.8e0	6.7e-4	6.7e-6	1.6e-6
⋮				
100	4.8e0	4.3e-4	4.3e-6	3.9e-7
⋮				
200	4.8e0	1.7e-4	1.7e-6	2.5e-8
⋮				
300	4.8e0	7.9e-5	7.9e-8	2.3e-9

which is a special case of the example presented in [28]. It has only one stationary point which is the global minimizer, too.

In [28] the authors showed that starting from an initial point belonging to the set

$$\mathscr{D} = \{w_1 < -\sqrt{w_2 + 1},\ w_2 > 0,\ w_3 > 0\},$$

any classical linesearch interior point method generates a sequence that is confined in a region where $w_1 - w_3 - 2$ is bounded away from zero. Therefore, classical methods cannot generate a sequence converging to the solution of the problem.

We solved the MCP problem given by the KKT condition. This way we obtain an MCP problem of form (1.1) with $v = (w, y)$ where $y \in \mathbb{R}^2$ is the vector of Lagrange multipliers of the equality constraints, $s \in \mathbb{R}^2$ are the slack variables and z is the vector of Lagrange multipliers of the inequality constraints.

First, we focus on the behavior of the PLIP method and of the CLIP method when the starting point $v^0 = (-2, 1, 1, 1, 1)^T$, $s^0 = z^0 = (1, 1)^T$ is used. This initial guess belongs to $\mathscr{D}$.

In Table 1 we report the results obtained using the CLIP method. Note that the method is not able to converge within 300 iterations and both $\lambda_{k,1}$ and λ_k tend to zero. Further, the Newton step p_k^N tends to become orthogonal to $\nabla\psi_k$. We remark also that the matrix $H'(x_{300})$ is numerically singular and $\nabla\psi(x_{300}) = 8.4e2$.

Table 2 shows that the PLIP method succeeds in solving the problem. The iteration history highlights that the use of the alternative direction d_k (3.1) is crucial to obtain convergence. In fact, this direction is selected in the first five iterations and at the third iteration, x_3 leaves the area in which the CLIP method gets trapped. Also, we point out that the angle η_k remains bounded away from $\pi/2$ and in the last six iterates the full Newton step is taken.

Now, we consider an "easy" starting point i.e., a point that does not belong to $\mathscr{D}$: $v^0 = (20, 1, 1, 1, 1)^T$, $s^0 = z^0 = (1, 1)^T$. Starting from this point both methods succeeded. In Table 3 for both methods we report the number *Nit* of performed iterations, the number *Nbt* of performed backtracks and the

Table 2
Example 1. PLIP method with starting point in $\mathscr{D}$

k	$\|H_k\|$	λ_k	$\|\cos(\eta_k)\|$	I
1	5.3e0	5.3e-2	9.9e-1	I_3
⋮				
5	3.4e0	1.0e-1	6.0e-1	I_3
⋮				
15	1.1e0	1.0e-1	4.1e-1	I_2
⋮				
32	9.5e-8	1.0e0	1.0e-1	I_1

Table 3
Example 1. Performance of the CLIP and PLIP method with starting point $\notin \mathscr{D}$

	Nit	Final $\|H\|$	*Nbt*	*NH*
CLIP	46	1.e-8	23	70
PLIP	45	1.e-7	40	86

number *NH* of performed H-evaluation. Results in Table 3 are typical. Our approach and the classical back tracking interior point method have similar cost, i.e., the use of the alternative path does not effect the overall performance of the basic interior point method.

Example 2 (Byrd et al. [5]). Consider the problem:

$$\begin{aligned} \min \quad & (w_1)^2 + (w_2)^2 + (w_3)^2 \\ \text{s.t.} \quad & \tfrac{1}{2}(w_1 + w_2 + \sqrt{2}w_3 + (w_2 - w_1)^2) = 0, \\ & \tfrac{\sqrt{2}}{2}(w_1 + w_2 + \sqrt{2}w_3 - 2)(w_2 - w_1) = 0, \\ & w_3 \geqslant -1. \end{aligned}$$

In [5] it was enlighted that, for a range of infeasible initial points any interior point method that performs backtracking along p_k^N will fail. More precisely, starting from such set of initial points, any linesearch algorithm whose search direction satisfies the linearization of the equality constraints will never achieve feasibility. In this case, the failure is not related to the presence of bounds, actually the iterates approach a singular nonstationary point and the Newton direction becomes increasing orthogonal to the gradient of the merit function. Hence, the Armijo condition forces the steps to be truncated. This happens regardless of the choice of the merit function and of the step selection strategy.

We considered the MCP given by KKT conditions for this problem. Hence, we have $v = (w, y)$, where $y \in \mathbb{R}^2$ is the vector of Lagrange multipliers of the equality constraints, $s \in \mathbb{R}$ is the slack variable and $z \in \mathbb{R}$ is the Lagrange multiplier of the inequality constraint.

Table 4
Example 2. CLIP method with the starting point (5.1)

k	$\|H_k\|$	$\lambda_{k,1}$	λ_k	$\|\cos(v_k)\|$
1	3.2e0	7.7e-1	7.7e-1	4.4e-1
⋮				
10	3.2e0	1.0e0	1.0e-3	4.6e-3
⋮				
50	3.1e0	1.0e0	1.0e-4	3.2e-4
⋮				
100	3.1e0	1.0e0	1.0e-4	6.1e-5
⋮				
200	3.1e0	1.0e0	1.7e-5	2.1e-5
⋮				
300	3.1e0	1.0e0	1.0e-5	1.2e-5

Table 5
Example 2. PLIP method with the starting point (5.1)

k	$\|H_k\|$	λ_k	$\|\cos(\eta_k)\|$	I
1	3.1e0	7.7e-1	4.4e-1	I_1
2	2.0e0	1.0e-2	9.7e-1	I_3
3	3.8e-1	1.0e0	7.3e-1	I_1
⋮				
8	1.5e-7	1.0e0	7.6e-1	I_1

We used the following starting guess:

$$v^0 = (-\sqrt{2}/2, \sqrt{2}/2, \sqrt{2}, 1, 1)^{\mathrm{T}}, \quad s_0 = z_0 = 1, \tag{5.1}$$

which is one of the difficult initial guesses.

In Table 4 we report the iteration history of the CLIP method. Note that the Newton direction tends to become orthogonal to $\nabla\psi_k$ and therefore the steplength λ_k tends to zero. However, it should be noted that the steplength $\lambda_{k,1}$ does not tend to zero and eventually it is equal to one. In fact, Δs_k^N and Δz_k^N are bounded while $\|\Delta v_k^N\| \to \infty$. Therefore, this is an example of the second type of failure: the sequence generated by the CLIP method approaches a singular nonstationary point $\tilde{x} = (\tilde{v}, \tilde{s}, \tilde{z})^{\mathrm{T}}$ such that $\tilde{s}$ and $\tilde{z}$ are strictly positive and the steplength $\lambda_{k,1}$ is bounded away from zero. Moreover $F(\tilde{x}) \notin \mathrm{Ker}(J^{\mathrm{T}}(\tilde{x}))$.

Theorem 4.4 ensures that the sequence generated by the PLIP method cannot approach such a singular nonstationary point. Actually, from Table 5 we see that the PLIP method succeeded. In particular, at the second iteration, the direction d_k is used and for this reason the iterate x_2 leaves the area in which the CLIP method gets stuck. In the last six iterates the full Newton step is taken and the asymptotical behavior of the classical interior point method is recovered. Finally, note that the angle η_k remains bounded away from $\pi/2$.

Example 3 (Simantiraky and Shanno [24]). In [24] the problem of computing equilibria of oligopolistic pricing models was formulated as a nonlinear complementarity problem with defining function f as follows:

$$f(y,w,r) = \begin{pmatrix} \sum_{j=1}^{n} r_{ij} - 1 & i = 1,\dots,n \\ w_i - \sum_{j=1}^{n} \pi_{ij} r_{ij} - \delta \sum_{j=1}^{n} y_j r_{ij} & i = 1,\dots,n \\ y_i - \delta w_j - \pi_{ji} & i,j = 1,\dots,n \end{pmatrix},$$

where $\delta \in \mathbb{R}$ and $\pi \in \mathbb{R}^n \times \mathbb{R}^n$ are given. In particular, in [24] the authors considered an example where $n = 7$, $\delta = 0.9$ and

$$\pi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5/2 & 5 & 5 & 5 & 5 & 5 \\ 0 & 0 & 4 & 8 & 8 & 8 & 8 \\ 0 & 0 & 0 & 9/2 & 9 & 9 & 9 \\ 0 & 0 & 0 & 0 & 4 & 8 & 8 \\ 0 & 0 & 0 & 0 & 0 & 5/2 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Letting $s_i = y_i$ for $i = 1,\dots,7$, $s_i = w_{i-7}$ for $i = 8,\dots,14$ and $s_{i+14+7(j-1)} = r_{ij}$ for $i,j = 1,\dots,7$, and $z = f(s)$, the following 126-variables problem is obtained:

$$H(s,z) = \begin{pmatrix} f(s) - z \\ SZe \end{pmatrix} = 0.$$

Although the problem is known to have numerous solution, in [24] it was shown that the CLIP method failed to converge when started from certain starting points. Essentially, the same type of failure that occurs in the Example 1 was observed: the sequence approaches a singular nonstationary point, the Newton step becomes very large and the step $\lambda_{k,1}$ is forced to zero.

We performed several experiments applying both the CLIP method and the PLIP method and both methods resulted to be very sensitive to the selection of the parameters γ_k and σ_k, (see also [24]).

With the choices of the parameters indicated at the beginning of the section, first we considered the initial guess $s_i = z_i = 1$, $i = 1,\dots,63$. The CLIP method failed in solving the problem because at the 44th iteration $\lambda_{k,1}$ was less than 10^{-12}. On the contrary, the PLIP method succeeded in solving the problem, even if the convergence was very slow. In fact, it needed 237 iterations to satisfy the stopping criterion and only in the last six iterations the full Newton step was taken. Convergence to the following equilibrium occurred:

$$\hat{R}_1^* = \begin{pmatrix} \alpha_1 & 0 & 0 & 0 & \alpha_2 & \alpha_3 & \alpha_4 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad \text{with } \alpha_i > 0 \;\; i = 1, \dots, 4 \sum_{i=1}^{4} \alpha_i = 1,$$

$$v = (22.88, 25, 27.5, 28.25, 31.27, 31.27, 31.27),$$

$$w = (25.42, 25, 22.5, 24.75, 25.42, 25.42, 25.42).$$

As reported in [24], the model has at least two symmetric Markov perfect equilibria, one of this (the kinked demand curve) is

$$R_1^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \beta(\delta) & 0 & 1-\beta(\delta) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad \text{with } \beta(\delta) = (5+\delta)/(5\delta + 9\delta^2),$$

$$v = (40.5, 40.5, 43, 45, 49.5, 49.5, 49.5),$$

$$w = (40.5, 42.22, 36.45, 45, 40.5, 40.5, 40.5).$$

Following [24], we perturbed the equilibrium R_1^* by adding an $\varepsilon > 0$ to all its zero elements, and we used this new point as a starting point. This way we investigated the behavior of the CLIP and PLIP method near this equilibrium. In Table 6 we report for different values of ε the performance of both methods and the equilibrium to which they converged. The symbol '*' means that the method failed because 300 iterations were not enough to satisfy the stopping criterion. It should be noted that for $\varepsilon \leqslant 10^{-3}$, no backtracks were performed and therefore the PLIP method reduces to the CLIP method. By an $\varepsilon > 10^{-3}$ the PLIP method no longer converges to R_1^*, while the CLIP method converged to this equilibrium also for $\varepsilon = 10^{-3}$. However, with $\varepsilon \geqslant 10^{-8}$ the CLIP method failed in solving this problem, while the PLIP method succeeded and converged to $\hat{R}_1^*$.

Table 6
Example 3. Behavior of the CLIP method and PLIP method near the point $\hat{R}_1^*$

	CLIP method		PLIP method	
ε	*Nit*	Equilibrium	*Nit*	Equilibrium
10^{-6}	1	R_1^*	1	R_1^*
10^{-3}	5	R_1^*	5	R_1^*
10^{-2}	8	R_1^*	8	R_1^*
10^{-1}	19	R_1^*	14	$\hat{R}_1^*$
$3 * 10^{-1}$	21	$\hat{R}_1^*$	34	$\hat{R}_1^*$
$5 * 10^{-1}$	21	$\hat{R}_1^*$	39	$\hat{R}_1^*$
$8 * 10^{-1}$	*	*	42	$\hat{R}_1^*$
1	*	*	39	$\hat{R}_1^*$

Acknowledgements

The authors thank Prof. Maria Macconi for the helpful discussions that helped them to prepare this paper.

References

[1] S. Bellavia, An inexact interior point method, J. Optim. Theory Appl. 96 (1998) 109–121.
[2] S. Bellavia, B. Morini, A globalization strategy for interior point methods for mixed complementarity problems, Proceedings of the Conference on High Performance Algorithms and Software for Nonlinear Optimization, Erice, 2001, to appear.
[3] H.Y. Benson, D.F. Shanno, R.J. Vanderbei, Interior point methods for nonconvex nonlinear programming: jamming and comparative numerical testing, Technical report ORFE-00-02, Operation Research and Financial Engineering, Princeton University, 2000.
[4] R.H. Byrd, M.E. Hribar, J. Nocedal, An interior point algorithm for large-scale nonlinear programming, SIAM J. Optim. 9 (1999) 877–900.
[5] R.H. Byrd, M. Marazzi, J. Nocedal, On the convergence of Newton iterations to non-stationary points, Report OTC 2001/01, Optimization Technology Center, 2001.
[6] R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal. 19 (1982) 400–408.
[7] J.E. Dennis, J.J. Moré, Quasi-Newton methods, motivation and theory, SIAM Rev. 19 (1977) 46–89.
[8] J.E. Dennis, R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Englewood Cliffs, NJ, 1983.
[9] C. Durazzi, On the Newton interior-point method for nonlinear programming problems, J. Optimization Theory Appl. 104 (2000) 73–90.
[10] S.C. Eisenstat, H.F. Walker, Globally convergent inexact Newton methods, SIAM J. Optim. 4 (1994) 393–422.
[11] A.S. El-Bakry, R.A. Tapia, T. Tsuchiya, Y. Zhang, On the formulation and theory of the Newton interior-point method for nonlinear programming, J. Optim. Theory Appl. 89 (1996) 507–541.
[12] M.C. Ferris, C. Kanzow, Complementarity and related problems: a survey, in: P.M. Pardalos, M.G.C. Resende (Eds.), Handbook of Applied Optimization, Oxford University Press, Oxford, 2002, pp. 514–530.
[13] M.C. Ferris, J.S. Pang, Engineering and economic applications of complementarity problems, SIAM Rev. 39 (1997) 669–713.
[14] C.A. Floudas, et al., Handbook of Test Problems in Local and Global Optimization, in: Nonconvex Optimization and its Applications, Vol. 33, Kluwer Academic Publishers, Dordrecht, 1999.

[15] D.M. Gay, M.L. Overton, M.H. Wright, A primal-dual interior method for nonconvex nonlinear programming, in: Y. Yuan (Ed.), Advances in Nonlinear Programming, Kluwer Academic Publishers, Dordrecht, 1998, pp. 31–56.
[16] M. Kojima, T. Noma, A. Yoshise, Global convergence in Infeasible-Interior-Point algorithms, Math. Programming 65 (1994) 43–72.
[17] M. Marazzi, J. Nocedal, Feasibility control in nonlinear optimization, Foundations of Computational Mathematics, Conference, Oxford, 1999, Vol. 284, London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, MA, 2001, pp. 125–154.
[18] J. Nocedal, S.J. Wright, Numerical Optimization, Springer Series in Operations Research, Springer, Berlin, 1999.
[19] F.A. Potra, An $O(nL)$ infeasible interior point algorithm for LCP with quadratic convergence, Ann. Oper. Res. 62 (1996) 81–102.
[20] F.A. Potra, R. Sheng, A large step infeasible interior point method for the P_*-matrix LCP, SIAM J. Optim. 7 (1997) 318–335.
[21] F.A. Potra, S.J. Wright, Interior-point methods, J. Comput. Appl. Math. 124 (2000) 281–302.
[22] F.A. Potra, Y. Ye, Interior-point methods for nonlinear complementarity problems, J. Optim. Theory Appl. 88 (1996) 617–647.
[23] R. Sheng, F.A. Potra, A quadratically convergent infeasible interior point algorithm for LCP with polynomial complexity, SIAM J. Optim. 7 (1997) 304–317.
[24] E.M. Simantiraky, D.F. Shanno, Computing equilibria of oligopolistic pricing models, Rutcor Research Report RRR (1995) 41–95.
[25] E.M. Simantiraky, D.F. Shanno, An infeasible-interior-point method for linear complementarity problems, SIAM J. Optim. 7 (1997) 620–640.
[26] P. Tseng, An infeasible path following method for monotone complementarity problems, SIAM J. Optim. 7 (1997) 386–402.
[27] R.J. Vanderbei, D.F. Shanno, An interior point algorithm for nonconvex nonlinear programming, Comput. Optim. Appl. 13 (1999) 231–252.
[28] A. Wachter, L.T. Biegler, Failure of global convergence for a class of interior point methods for nonlinear programming, Math. Programming Series A 88 (2000) 565–574.
[29] L.T. Watson, Solving the nonlinear complementarity problem by a homotopy method, SIAM J. Control Optim. 17 (1979) 36–46.
[30] S.J. Wright, A path-following infeasible-interior-point algorithm for linear complementarity problems, Optim. Meth. Software 2 (1993) 79–106.
[31] S.J. Wright, An infeasible-interior-point algorithm for linear complementarity problems, Math. Programming Ser. A 67 (1994) 29–52.
[32] S.J. Wright, Primal-Dual Interior-Point Methods, SIAM, Philadelphia, PA, 1997.
[33] S.J. Wright, D. Ralph, A superlinear infeasible-interior point algorithm for monotone complementarity problems, Math. Oper. Res. 21 (1996) 815–838.
[34] H. Yamashita, H. Yabe, T. Tanabe, A globally and superlinearly convergent primal-dual interior point trust region method for large scale constrained optimization, Technical Report, Mathematical System Inc., 1997.
[35] Y. Zhang, On the convergence of a class of infeasible interior-point algorithms for the horizontal linear complementarity problems, SIAM J. Optim. 4 (1994) 208–227.